

# RELIABILITY OF SHELL STRUCTURES THROUGH SENSITIVITY ANALYSIS

M. Oberguggenberger, J. King and B. Schmelzer

University of Innsbruck, Unit for Engineering Mathematics and Department of Mathematics

Technikerstraße 13, A – 6020 Innsbruck, Austria

## OVERVIEW

The reliability of shell structures is a major theme in aerospace engineering. This article argues that the reliability issue can and should be addressed through sensitivity analysis, that is, the assessment of the impact of individual input parameters (loads, material constants, geometry) or sets of input parameters on the failure of the structure. A major challenge in aerospace applications lies in the fact that high computational costs have to be faced. Methods have to be developed that admit assertions about the sensitivity of the output with as few computations as possible. This article explores and compares four probabilistic and semi-probabilistic techniques – interval estimates, Monte Carlo simulation, random set methods, fuzzy sets – and reports on our experience with the suitability of each method.

## 1. INTRODUCTION

This article addresses results of a research project in aerospace engineering, in collaboration of academia and industry. The project name is ICONA (Innovative Concept for Nonlinear Analysis of Lightweight Structures) and is carried out jointly by Intales GmbH Engineering Solutions, three departments of the University of Innsbruck (Mathematics, Civil Engineering, Computer Science), supported by EADS Astrium ST and TransIT Innsbruck and involves a collaboration with Delft University of Technology (Aerospace Engineering). The project is concerned with the determination of the buckling load of the frontskirt of the ARIANE 5 launcher under various loading and flight scenarios. It involves development of new shell elements, databanks for the flight scenarios, new numerical substructure techniques, and the development of concepts of sensitivity analysis in order to more accurately assess the reliability and the design of the structure or eventually suggest guidelines for redesign.

The sensitivity analysis question will be addressed in this article. Assessing the reliability of a structure is one of the core tasks in engineering. We believe that sensitivity analysis supplies more conclusive information in this respect, rather than safety margins or failure probabilities. It is probably less common in use due to the computational effort involved. Facing the computational costs, the current challenge is to conceive and develop practicable tools. The paper explores various probabilistic and semi-probabilistic techniques that promise a computationally feasible ap-

proach to sensitivity analysis.

The frontskirt is a reinforced light weight shell structure. The computation of the decisive parameter indicating failure, the load proportionality factor (LPF), is based on a large finite element model. The load proportionality factor is defined as the limiting value in an incremental procedure, in which the dynamic loads during a flight scenario are increased stepwise until breakdown of the structure is reached. The goal is to determine the most influential input parameters (loads, material constants, geometry) on the load proportionality factor in a sensitivity analysis. The calculation of the output variable LPF - under a given single set of input parameters - takes about 32 hours on a high performance computer. In addition to the extremely high computational cost, the LPF may depend in a non-differentiable manner on some of the input parameters, especially variations in the geometry. A classical sensitivity analysis of the complete structure is currently out of reach.

For this reason, we introduce probabilistic methods and employ tools from the rather new field of imprecise probability, in order to attain a conclusive representation of the respective effects of parameter changes (the term imprecise probability theory subsumes random sets, fuzzy sets, interval arithmetic and interval probability, among others). Engineering information on the variability of the input parameters usually consists of a central value and a coefficient or range of variation. The basic strategy for arriving at a sensitivity assessment will be to successively freeze the input parameters and study the effect on the variability of the output. We wish to do this without artificial parametric assumptions and with as few calls of the finite element program as possible.

In this exploratory study, we model the input variability by the following set of methods:

- (a) interval bounds estimated by sampling from a Cauchy distribution;
- (b) partial rank correlation coefficients obtained by Monte-Carlo simulation, together with resampling (to estimate accuracy while keeping the number of calls to the program low).
- (c) random sets constructed from Tchebycheff's inequality;
- (d) fuzzy sets and Hartley-like measures.

We test and compare the efficiency of the respective

methods with the aid of a simplified finite element model simulating part of a space craft launcher (Figure 1). The computational cost for the simplified model is one hour per call of the program. In the sensitivity analysis, up to 17 input parameters are taken into account. A tentative description of the meaning of the parameters as well as their nominal values can be read off from Table 1. It turned out that the partial rank correlation coefficient method gives the best results, requiring about 100 calls to the program. We are working on further reducing the computational cost by implementing the sensitivity analysis on a lower level in the iterative solution procedure, thus making application to the ARIANE 5 frontskirt feasible.

A detailed description of the respective methods will follow in four sections, with a final section devoted to a comparison of the methods. The question of modelling correlations between the input variables will be addressed in the appropriate sections. Tables and figures are collected at the end of the paper.

For background material on sensitivity analysis we refer to the Special Issue [7], in particular the survey article [8] and to [6], for random sets, to [16, 17], for random and fuzzy sets, to [4, 12], for probability boxes, to [2], for a review on probabilistic treatment of uncertainty in structural engineering as well as information on variability of typical input parameters, to [23], for a survey on the mathematical methods employed, to [18].

## 2. INTERVAL BOUNDS

We begin by showing how an interval estimate on the desired output variable LPF can be obtained, given knowledge about intervals between which the input parameters lie. The computation of the LPF can be viewed as an input-output map that assigns to the input variables  $(x_1, \dots, x_d)$  the output value  $y = g(x_1, \dots, x_d)$ . In our application, the  $x_i$  will be taken as some or all of the parameters listed in Table 1, and  $y$  will be LPF. To describe the method, suppose that the variability of each input parameter  $x_i$  is described by an interval  $[\mu_i - \Delta_i, \mu_i + \Delta_i]$  of spread  $\Delta_i$  around a central value  $\mu_i$ . It has been argued in [14], that an estimate of the output interval can be obtained by Monte Carlo simulation using the Cauchy distribution. Monte Carlo simulation enters only as a computational tool; the model of the uncertainty is non-parametric.

The underlying theory from [14] is as follows. Suppose we wish to estimate the difference

$$\Delta y = g(x_1, \dots, x_d) - g(\mu_1, \dots, \mu_d)$$

where  $|\Delta x_i| = |x_i - \mu_i| \leq \Delta_i$ . Linearization around the mean value gives

$$|\Delta y| \leq \Delta = \sum_{i=1}^d |c_i| \Delta_i, \quad c_i = \frac{\partial g}{\partial x_i}(\mu_1, \dots, \mu_d).$$

If the  $X_i$  are independent random variables following a Cauchy distribution with scale parameter  $\Delta_i$ , then  $Y = c_1 X_1 + \dots + c_d X_d$  obeys a Cauchy distribution with scale parameter  $\Delta$ . This offers the possibility of computing the bound  $\Delta$  on the output spread by Monte Carlo simulation.

The algorithm runs along the following lines. To produce a single realization, a  $d$ -dimensional sample  $(z_1, \dots, z_d)$  of Cauchy distributed variables with scale parameters 1 is taken. Setting  $K = \max_{1 \leq i \leq d} |z_i|$ , one has that  $\delta_i = \Delta_i z_i / K$  has a Cauchy distribution with scale parameter  $\Delta_i / K$ . Putting  $x_i = \mu_i + \delta_i$  it follows that

$$Z = K(g(x_1, \dots, x_d) - g(\mu_1, \dots, \mu_d))$$

is a realization of a Cauchy distributed variable with desired scale parameter  $\Delta$  (this is true exactly when  $g$  is linear and otherwise approximately). An  $n$ -fold repetition yields the Monte Carlo sample of size  $n$  of the variable  $Z$ . Fitting a Cauchy distribution – e.g. by the maximum likelihood method – produces an estimate of the spread  $\Delta$  of the output interval  $[g(\mu_1, \dots, \mu_d) - \Delta, g(\mu_1, \dots, \mu_d) + \Delta]$ . The computational effort for this estimate is  $n$  calls of the finite element program and thus independent of the dimension  $d$ . This offers the possibility to include a larger number of input variables in the analysis.

**Example 1.** In this calculation, 17 input parameters were included with nominal values displayed in Table 1. The spreads  $\Delta_i$  were taken as 0.15-times the nominal values  $\mu_i$ . We used a direct Monte Carlo method to produce a sample of size  $n = 100$ . The value of the load proportionality factor LPF was obtained as  $\mu = g(\mu_1, \dots, \mu_d) = 3.5443$ . The simulation resulted in an estimate for its spread of  $\hat{\Delta} = 0.2924$ .

This method allows to produce a confidence estimate for the desired spread  $\Delta$  as well – without further essential computational effort, namely by means of resampling, thus approximating the empirical distribution of  $\Delta$ . We employed 10000 random subsamples of size 100 (with repetition), following the suggestions in [22]. This resulted in a 95%-confidence interval for  $\Delta$  of  $CI_{0.95}(\hat{\Delta}) = [0.2281, 0.3685]$ . Since the resampling procedure is based on the original  $n = 100$  calls of the finite element program, the additional computational effort is negligible.

It should be noted here that the accuracy of the Cauchy method is in question in this context. In fact, the justification of the Cauchy method relies on the fact that the output function  $g$  is near linear. This is not the case in our situation where the output function  $g$  is a nonlinear finite element computation resulting in the LPF. It turned out that the simulations of the auxiliary variable  $Z$  actually failed the Kolmogorov-Smirnov-test for being Cauchy distributed. In fact, a direct Monte Carlo simulation of size  $n = 100$  of the output variable LPF, with uni-

formly distributed input variables, produced an output range of [3.45, 3.65]. This indicates that the range was largely overestimated by the Cauchy method. This could possibly be overcome by the suggestion of [14] of repeated bisection of the input interval, though at an increase in computational cost.

### 3. MONTE CARLO SIMULATION

In this section, we turn to direct Monte Carlo simulation in sensitivity analysis. Given a certain statistical distribution of the input variables, Monte Carlo simulation is a numerical procedure for calculating the distribution of the output. By means of a random number generator,  $n$  independent realizations of the input variables are produced. The finite element program yields a histogram of  $n$  realizations of the output variable, from which its statistical properties can be estimated. The method is parametric in as much as assumptions on the type of input distributions have to be made. We always assumed uniformly distributed input variables.

Further, we employed Latin hypercube sampling, an efficient stratified sampling strategy. To obtain a sample of size  $n$ , the Latin hypercube sampling plan divides the range of each variable  $X_i$  into  $n$  disjoint subintervals of equal probability. First,  $n$  values of each variable  $X_i$ ,  $i = 1, \dots, d$ , belonging to the respective subintervals are randomly selected. Then the  $n$  values for  $X_1$  are randomly paired without replacement with the  $n$  values for  $X_2$ . The resulting pairs are then randomly combined with the  $n$  values of  $X_3$  and so on, until a set of  $n$   $d$ -tuples is obtained. This set forms the Latin hypercube sample. The advantage of Latin hypercube sampling is that sampled points are evenly distributed through design space, thereby covering regions possibly important for the input-output map which might be missed by direct Monte Carlo simulation. It can be shown that the variance of an estimator based on Latin hypercube sampling is asymptotically smaller than the variance of the direct Monte Carlo estimator, and possibly markedly smaller when the input-output map is partially monotonic [6, 15, 25].

For additional accuracy in view of the rather small sample size we subjected the simulated variables to correlation control (see [10, 11]). This procedure consists in a rearrangement of the originally simulated values such that the resulting empirical rank correlation matrix is close to diagonal. Based on such a direct Monte Carlo simulation, a considerable number of methods is available for sensitivity analysis.

**Example 2.** To estimate the influence of each of the 17 input parameters from Table 1 on the output LPF, we performed a Monte Carlo simulation of size  $n = 100$  with uniformly distributed input variables (on the intervals as in Example 1). A first scan of the influence of the individual input parameters on the output can

be done by means of scatterplots (input  $X_i$  against output  $Y$ ), see Figure 2.

It appears that parameter  $X_{13}$  (booster load) is most significantly correlated with the output LPF. To reveal more subtle dependencies, a detailed calculation is required, as e.g. computing the weighted contribution of each input variable to the variance of the output (this can be done by successively freezing variables at their mean value and observing the effect on the total output variance). However, hidden interactions may have a significant effect on the decomposition of the variance (as well as on the scatterplot). We therefore turn to a method which intends to remove the influence of co-variates on the correlation between a given input variable  $X_i$  and the output variable  $Y$ . This method is based on the partial rank correlation coefficient (PRCC).

We recall that partial correlation between two random variables  $X_i$  and  $Y$  given a set of co-variates  $X_{\setminus i} = \{X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_d\}$  is defined as the correlation between the two residuals  $e_{X_i \cdot X_{\setminus i}}$  and  $e_{Y \cdot X_{\setminus i}}$  obtained by regressing  $X_i$  on  $X_{\setminus i}$  and  $Y$  on  $X_{\setminus i}$ , respectively. More precisely, one first constructs the two regression models

$$\hat{X}_i = \alpha_0 + \sum_{j \neq i} \alpha_j X_j, \quad \hat{Y} = \beta_0 + \sum_{j \neq i} \beta_j X_j,$$

obtaining the residuals

$$e_{X_i \cdot X_{\setminus i}} = X_i - \hat{X}_i, \quad e_{Y \cdot X_{\setminus i}} = Y - \hat{Y}.$$

Since  $e_{X_i \cdot X_{\setminus i}}$  and  $e_{Y \cdot X_{\setminus i}}$  are those parts of  $X_i$  and  $Y$  that remain after subtraction of the best linear estimates in terms of  $X_{\setminus i}$ , the partial correlation coefficient

$$\rho_{X_i, Y \cdot X_{\setminus i}} = \rho(e_{X_i \cdot X_{\setminus i}}, e_{Y \cdot X_{\setminus i}})$$

quantifies the linear relationship between  $X_i$  and  $Y$  after removal of any part of the variation due to the linear influence of  $X_{\setminus i}$ . Applying a rank transformation to the variables  $X_i$  and  $Y$  leads to the partial rank correlation coefficient (PRCC). For further background on PCCs and PRCCs, see [5, 9, 21].

**Example 3.** The same Monte Carlo sample as in Example 2 was used to calculate the PRCCs. The result can be seen in Figure 3. For further statistical confirmation, we performed a resampling procedure as in Example 1, producing bootstrap confidence intervals for the partial rank correlation coefficients as displayed in Figure 4. Accordingly, only the PRCCs of the parameters  $X_1$ ,  $X_3$ ,  $X_9$ ,  $X_{13}$  and  $X_{14}$  test to be nonzero (initial temperature, temperature cylinder 2, hydrostatic pressure sphere 2, booster load node 1 in  $z$ -direction, booster load node 2 in  $z$ -direction).

The outcome shows that among the parameters  $X_3$ ,  $X_{13}$  and  $X_{14}$ , the one with the biggest influence is  $X_{13}$ , followed by  $X_3$  and  $X_{14}$ .

We also ran various tests with correlated input which confirmed the observed sensitivities. However, each

test required a new Monte Carlo simulation with sample size  $n = 100$ . In addition, we computed Sobol indices [24] for groups of variables; this, however, again requires additional Monte Carlo simulations.

#### 4. RANDOM SET METHODS

In this and the following section we present two completely non-parametric methods for a sensitivity assessment. Both methods produce good visualizations of the influence of each individual parameter, but they turn out to be computationally expensive. For this reason, we restrict their application to the three input parameters  $X_3, X_{13}, X_{14}$  (temperature cylinder 2, booster load node 1 in  $z$ -direction, booster load node 2 in  $z$ -direction) which turned out to be the most important parameters in the previous analysis.

As has been demonstrated in [19, 20], random intervals constructed by Tchebycheff's inequality can serve as a non-parametric model of the variability of a parameter, given its mean value and variance as sole information. We first explain the univariate case of a real-valued random variable  $X$ . Let  $\mu = E(X)$  be its expectation and  $\sigma^2 = V(X)$  be its variance. Tchebycheff's inequality asserts that

$$(1) \quad P(|X - \mu| > d_\alpha) \leq \alpha, \quad d_\alpha = \sigma/\sqrt{\alpha}.$$

Equipping the unit interval  $(0, 1]$  with the uniform probability distribution, the non-parametric confidence intervals

$$(2) \quad I_\alpha = [\mu - d_\alpha, \mu + d_\alpha], \quad \alpha \in (0, 1]$$

define a random set, i.e. a set-valued random variable. By construction, the following formulas for the belief in the set  $I_\alpha$  and the plausibility of its complement  $I_\alpha^c$  hold:

$$\begin{aligned} \underline{P}(I_\alpha) &= \int_{\{\beta \in (0, 1] : I_\beta \subset I_\alpha\}} d\beta = 1 - \alpha \leq P(I_\alpha), \\ \overline{P}(I_\alpha^c) &= \int_{\{\beta \in (0, 1] : I_\beta \cap I_\alpha^c \neq \emptyset\}} d\beta = \alpha \geq P(I_\alpha^c). \end{aligned}$$

This shows that the random set description provides a conservative assessment of the variability  $X$ . In applications, the range of the parameter  $X$  may be confined to a compact interval  $[x_{\min}, x_{\max}]$ . In this case, the random set will be truncated to

$$I_\alpha = [(\mu - d_\alpha) \vee x_{\min}, (\mu + d_\alpha) \wedge x_{\max}].$$

In the multivariate case  $X = (X_1, \dots, X_d)$  where each parameter  $X_i$  is modelled as a random set as in (2), we form the joint random set (assuming random set independence)

$$\alpha = (\alpha_1, \dots, \alpha_d) \rightarrow A_\alpha = I_{\alpha_1}^1 \times \dots \times I_{\alpha_d}^d$$

again with the uniform distribution on the probability space  $(0, 1]^d$ .

Let  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  be a continuous function. If the input variables  $X = (X_1, \dots, X_d)$  are modelled as a random

set  $A_\alpha$ ,  $\alpha \in (0, 1]^d$  (equipped with the uniform probability distribution), the output variable is given by the random set  $g(A_\alpha)$ ,  $\alpha \in (0, 1]^d$ . A visualization of the output can be obtained by means of the upper and lower distribution functions (or *probability box*, [2])

$$(3) \quad \begin{aligned} \overline{F}(x) &= P(\alpha : g(A_\alpha) \cap (-\infty, x] \neq \emptyset) \\ \underline{F}(x) &= P(\alpha : g(A_\alpha) \subset (-\infty, x]). \end{aligned}$$

In the numerical evaluation, the joint random set is approximated by a finite random set with focal elements

$$I_{\alpha_1}^1 \times \dots \times I_{\alpha_d}^d, \quad \alpha_j \in \{\frac{1}{n}, \frac{2}{n}, \dots, 1\},$$

each with probability weight  $n^{-d}$ . The input-output function is evaluated as follows: First, an interval  $Q \subset \mathbb{R}^d$  is determined that bounds the relevant range of the input variables  $X$ . Next, the values of the function  $g$  are computed at the  $m^d$  nodes of a uniform grid on  $Q$ . The output  $g(Q)$  is approximated by a response surface  $\hat{g}(Q)$  obtained by multilinear splines. More precisely, to compute the image of one of the sets  $A_\alpha$ ,  $\hat{g}(Q)$  is evaluated at all grid points inside  $A_\alpha$  and all points on its edges intersecting one of the grid lines. The interval  $g(A_\alpha)$  is approximated by the minimum and maximum value thus obtained. Finally, the probability box (3) is calculated by adding the weights when appropriate. The essential computational effort thus amounts to  $m^d$  calls of the finite element program.

Figure 5 shows the result of the calculation of the load proportionality factor (LPF) where the three input parameters  $X_3, X_{13}, X_{14}$  were kept variable. The variance  $\sigma$  for the Tchebycheff model was adjusted such that the base intervals  $[x_{\min}, x_{\max}]$  for each of the parameters was symmetric around the corresponding mean  $\mu$  with spread  $\pm 0.15\mu$ . In this case,  $d = 3$  and we chose  $m = 5$  so that 125 calls to the FE-program were required.

**Example 4.** To assess the sensitivity of the load proportionality factor LPF with respect to the parameters  $X_3, X_{13}, X_{14}$  we again use the Tchebycheff model for each of the parameters with spread 0.15 times their mean values. Then we successively set one of the resulting  $\sigma_3, \sigma_{13}, \sigma_{14}$  equal to zero (while keeping the others at their given value), go through the calculation indicated above and plot the resulting probability box (solid lines – the thin lines indicate the unperturbed result from Figure 5). This is displayed in Figure 6 and shows that setting  $\sigma_{13} = 0$  produces the biggest reduction of the width of the probability box, while setting  $\sigma_{14} = 0$  has little effect. This confirms that the parameter  $X_{14}$  has the least influence on the variability of the response, while  $X_{13}$  exerts the biggest influence.

The strategy of successively freezing variables, together with probability boxes is further explicated in [3] and applied in [20]. Questions of dependence or interactivity of the input variables are postponed to

the next section. Dependence could be modelled by copulas on the underlying probability space  $(0, 1]^d$  or by restrictions on the set of probability measures on  $\mathbb{R}^d$  defined by the random set.

## 5. FUZZY SETS

Another non-parametric, in fact, non-probabilistic method for sensitivity analysis can be based on fuzzy sets. In this section, one-dimensional input variables will be modelled as *normalized fuzzy numbers*, that is as fuzzy subsets  $B$  of the real line with upper semi-continuous membership function  $\pi_B(x)$  that attains the value 1. The  $\alpha$ -level set of  $B$  is the set

$$B_\alpha = \{x \in \mathbb{R} : \pi_B(x) \geq \alpha\}, \quad \alpha \in (0, 1].$$

In the multivariate case, the non-interactive joint fuzzy set is defined as follows. Given  $d$  univariate fuzzy sets  $B^1, \dots, B^d$ , the joint fuzzy set has the  $\alpha$ -level sets

$$B_\alpha = B_\alpha^1 \times \dots \times B_\alpha^d, \quad \alpha \in (0, 1].$$

Correlation – or rather interactivity in the language of fuzzy logic – can be modelled by certain parametric restrictions on the  $\alpha$ -level sets. To avoid combinatorial complications, we shall treat interactivity of at most two out of the  $d$  variables. Since an  $\alpha$ -level set of the form  $B_\alpha^i \times B_\alpha^j$  is an affine image of the unit square, it suffices to give the definitions for  $B_\alpha^1 = B_\alpha^2 = [0, 1]$ . Following [26], interactivity will be modelled by replacing the unit square by a diamond-shaped region, symmetric around one of the diagonals. Let  $0 \leq \rho \leq 1$  and define the points  $P_1, \dots, P_4$  by

$$\begin{aligned} P_1 &= (\rho/2, \rho/2), & P_2 &= (1 - \rho/2, \rho/2), \\ P_3 &= (1 - \rho/2, 1 - \rho/2), & P_4 &= (\rho/2, 1 - \rho/2). \end{aligned}$$

Interactivity of *positive degree*  $\rho$  is modelled by taking the rhombus with corners  $\{(0, 0), P_2, (1, 1), P_4\}$  as joint level set, while interactivity of *negative degree*  $-\rho$  is modelled by the rhombus with corners  $\{(0, 1), P_1, (1, 0), P_3\}$  as joint level set (Figure 7).

Let  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  be a continuous function. If the input variables  $X = (X_1, \dots, X_d)$  are modelled as a non-interactive or interactive fuzzy set with  $\alpha$ -level sets  $B_\alpha$  as above, Zadeh's extension principle yields the output variable as the fuzzy number with level sets  $g(B_\alpha)$ ,  $\alpha \in (0, 1]$ .

While a fuzzy set can be interpreted as a random set (cf. e.g. [4]) and the procedure appears similar to the one of Section 4, there is a fundamental difference in the multivariate case: in fuzzy set theory, only  $\alpha$ -level sets of the same level are combined to produce the joint fuzzy set, while for random sets, the focal elements are obtained as products with respect to any combination and thus are indexed by the product space  $(0, 1]^d$ .

**Example 5.** In the assessment of the sensitivity of the load proportionality factor LPF with respect to

the input parameters  $X_3, X_{13}, X_{14}$ , these parameters were modelled as symmetric triangular fuzzy numbers, with central values  $\mu_i$  from Table 1 and spread  $\pm 0.15\mu_i$  as before. The numerical calculation is based on the response surface method explained in Section 4. The images of the  $\alpha$ -level sets are again computed by piecewise multilinear combination. To handle possible lack of monotonicity of the function  $g$ , we start with level  $\alpha = 1$  and go the way down to  $\alpha = 0$ , insuring at each step that the approximations satisfy  $g(A_\beta) \subset g(A_\alpha)$  for  $\alpha < \beta$ .

In the non-interactive case, the procedure for determining the sensitivity of the output with respect to the input variables is the same as in Example 4. The initial calculation is performed with proportional spreads  $\pm 0.15\mu_i$ . Then we successively replace one of the triangular fuzzy numbers by its crisp central value  $\mu_i$ , and compute the output as a fuzzy number. The result gives a good visual representation of the change of variability. This can be quantified using e.g. the Hartley-like measure

$$\text{HL}(B) = \int_0^1 \log(1 + \lambda(B_\alpha)) \, d\alpha$$

of fuzzy sets  $B$  as proposed by [12] (see also [1] for further implementation of this idea in sensitivity analysis).

The result is depicted in Figure 8, where the outer contour is the membership function of the fuzzy LPF with all input parameters fuzzy, while the shaded region is bounded by the membership function of the fuzzy LPF with successively frozen input parameters. It confirms the observations obtained by the random set method:  $X_{13}$  is the most influential parameter, followed by  $X_3$  and then  $X_{14}$ . This can be explained by the model setup:  $X_{13}$  refers to a large booster load on one side of the frontskirt, while  $X_{14}$  signifies a much smaller booster load on the opposite side. The Hartley-like measures displayed in Table 2, though, show that some, albeit small, influence of parameter  $X_{14}$  is detectable.

**Example 6.** This example serves to show how the effect of possible correlations between two of the input parameters on the sensitivity can be assessed. Correlation will be interpreted here as degree of interactivity as described above. In this example, we assume a degree of interactivity  $\rho = 0.98$  between parameters  $X_{13}$  and  $X_{14}$ . The remaining parameters are treated as non-interactive. The  $\alpha$ -level sets are of cylindrical shape with a rhombic base  $R_\alpha$ , say. Their images are again computed by piecewise multilinear combination. Otherwise, the procedure of successively freezing variables is similar: For example, when  $X_{13}$  is frozen at its central value  $\mu_{13}$ , the interactivity restricts  $X_{14}$  to vary along the intersection of  $R_\alpha$  with the line through  $\mu_{13}$  parallel to the  $x_{14}$ -axis, while  $X_3$  varies in its original  $\alpha$ -level interval.

The result is shown in Figure 9; the meaning of the

contour and the shaded region is the same as in Figure 8. The outcome confirms the prominence of parameter  $X_{13}$ ; as a consequence of the correlation, parameter  $X_{14}$  is seen to exert a comparable influence. The result also demonstrates that the correlation changes the sensitivity of the output with respect to parameter  $X_3$ . Table 3 shows the Hartley-like measures of the fuzzy output under successive freezing of input variables. One may note that the study of the influence of correlations can be implemented in the fuzzy approach with ease.

As in Example 5, the computational effort using the response surface consisted in 125 calls of the finite element program. The vertical jumps of the membership function in Figure 9 indicate that the output does not depend monotonically on the input variables. Closer inspection (done by producing an array of two-dimensional plots of the partial maps  $X_i \rightarrow \text{LPF}$ ) showed that this is indeed the case. Therefore, the accuracy of the method using just 125 grid values is in question. A number of additional explicit evaluations showed that the accuracy of the boundaries of the  $\alpha$ -level sets for the LPF is in the range of  $\pm 0.02$  in absolute value.

## 6. SUMMARY AND CONCLUSIONS

Starting from a research project in aerospace engineering one of whose goals was to determine the sensitivity of the buckling load of the frontskirt of the ARIANE 5 launcher with respect to certain input parameters, we explored various methods from probability and imprecise probability theory. In view of the excessive computational costs of a single run of the finite element program, the major challenge was to develop methods with as few calls of the program as possible. We used a simplified model of the launcher for the numerical tests of the methods.

The methods under scrutiny were random sets and Tchebycheff's inequality, fuzzy sets and Hartley-like measures, intervals and sampling from a Cauchy distribution, standard Monte-Carlo simulation and resampling. Criteria for the evaluation are: computational effort; applicability to large scale problems; accuracy; avoidance of tacit assumptions; reliability and clarity of interpretation; possibility of analyzing correlated input.

Generally speaking, the Monte Carlo simulation methods are computationally least expensive. For our sensitivity study, a sample size of  $n = 100$  appeared sufficient. In addition – as is well known – the sample size can be chosen independently of the number of input variables, so that we could include all 17 variables in our study. These methods are clearly applicable to large scale problems. Disadvantages are that parametric assumptions on the input variables have to be made and that freezing of variables requires repetition of the  $n = 100$  simulations. Thus computing PRCCs

plus resampling is possible irrespective of the problem scale, but variance decomposition by freezing variables is not. The same applies to analyzing sensitivity with respect to input correlations, which requires repetition of the simulation as well. The numerical accuracy of the Monte Carlo simulation is well known to be of order  $1/\sqrt{n}$  times the standard deviation of the simulated variable. In view of the coefficients of variation which were in the range of 10% this appeared sufficient for the sensitivity study.

We emphasize that the results of a Monte Carlo simulation are amenable to resampling, which introduces little additional computational effort (no further evaluations of the costly input-output map are needed). In this way, bootstrap confidence intervals can be obtained that may serve as statistical estimates of the accuracy of the results. For example, we estimated the bias of each partial rank correlation coefficient, that is, the absolute value of the difference of the mean of the resampled data and the initial estimate. The estimated bias resulted to be less than 2% of the initial estimate. Further, the significance of the resulting ranking of the influence of the respective input parameters can be assessed by comparing the bootstrap confidence intervals.

The Cauchy method is a simulation method for estimating the spread of the output interval. The resulting estimate is non-parametric in as much as only the spreads of the input variables enter. As a subcase of Monte Carlo simulation, everything that has been said above applies here as well. A problematic point is that the method is derived under the assumption that the output function is approximately linear. In our case, the output function is substantially nonlinear. By means of repeated simulations we observed a quite substantial lack of accuracy of the estimate of the output spread, as described in Section 2.

Both in the fuzzy set and random set methods, the output  $\alpha$ -level sets and focal sets, respectively, are computed by searching for the maximum and minimum of the corresponding output range. Sufficient accuracy can only be obtained by a larger number of calls of the output function, evaluated on a grid of input data. In addition, the grid size increases exponentially with the number of input variables. These methods appear feasible only in the case of medium size problems and a small number of input variables. Monotonicity or partial monotonicity of the output function increases accuracy and helps reducing the number of computations required.

Test runs with finer grids showed that the numerical error of the interpolation (i.e. replacing the true output function by a piecewise bilinear response surface) was less than 1%, thus definitely satisfactory. However, the optimization error introduced when calculating the boundaries of the output level sets turned out to be about  $\pm 0.02$  in absolute value, which is around

10 - 20% of the spread of the base level (see end of Section 5).

The numerical error in the boundaries of the output level sets appears less influential in the random set method. This is due to a certain averaging effect. Indeed, in the fuzzy model the computation of  $\ell$  output level sets corresponds to  $\ell$  input level sets, whereas in the random set model – at least when using random set independence – a combination of  $\ell^d$  input focal sets enters ( $d$  the number of variables).

Both methods are essentially non-parametric. The random set model we used is generated by Tchebycheff's inequality and hence non-parametric by definition. In the fuzzy set model, we used triangular fuzzy numbers as input. These can be seen as a collection of intervals of linearly changing length. The  $\alpha$ -level sets resulting from the computation determine the output range when the input varies over  $d$ -dimensional intervals of length proportional to  $1 - \alpha$ .

The fuzzy model in combination with the response surface technique has an additional advantage: it allows the a-posteriori introduction of interactivity between the input variables without the need for new calls of the output function. The effect of interactive input can simply be evaluated by interpolation in the response surface.

We finally comment on the practicality of upscaling to the full problem. This remains a major challenge. The computational structure of the given problem consists in a nonlinear, incremental procedure. The LPF is obtained as the ultimate load value beyond which the computed solution cannot be prolonged. This may be either due to a bifurcation point or to a breakdown of the structure. We currently pursue two strategies. One strategy is a perturbation method that replaces the full model by a quadratic approximation when a bifurcation point is reached. This is based on Koiter's asymptotic analysis of post-buckling of shells, see e.g. [13]. The sensitivity analysis would be done with the asymptotic model in place of the full model. The second strategy is to start the sensitivity analysis at a later stage of the iterative procedure. Both methods require to access the finite element code at a deeper level. A certain difficulty which we expect to encounter stems from the fact that the incremental procedure is path dependent. Thus varying the input parameters late in the process could be misleading, as initial variations might result in a quite different path to breakdown. Further directions of research will focus on the sensitivity of the eigenmodes computed by means of domain decomposition with respect to the boundary conditions connecting the domains. This will be particularly relevant in a situation where computations of different parts of a shell structure are done separately and should clarify what additional information is needed to arrive at a robust assessment.

## 7. ACKNOWLEDGEMENTS

We are grateful to Herbert Haller for providing the finite element models and advice on engineering questions. Thanks are also due to Christof Neuhauser and Alexander Ostermann for discussions on numerical problems, to Hermann Starmans for the project management and to Robert Winkler for help with the finite element codes. We thank Eelco Jansen for advice on Koiter's theory and Johann Arbocz for discussions on the buckling of cylindrical shells.

## 8. BIBLIOGRAPHY

- [1] D. A. Alvarez. Nonspecificity for infinite random sets. Preprint, Unit for Engineering Mathematics, University of Innsbruck, 2006.
- [2] S. Ferson, V. Kreinovich, L. Ginzburg, D. S. Myers and K. Sentz. Constructing probability boxes and Dempster-Shafer structures. *SANDIA Report SAND2002-4015*, Sandia National Laboratories, Albuquerque, 2003.
- [3] S. Ferson and W. T. Tucker. Sensitivity analysis using probability bounding. *Reliability Engineering & System Safety*, 91:1435-1442, 2006.
- [4] I. R. Goodman, H. T. Nguyen. Fuzziness and randomness. In: C. Bertoluzza, M. Á. Gil, D. A. Ralescu, editors. *Statistical Modeling Analysis and Management of Fuzzy Data*. Physica-Verlag, Heidelberg, 2002.
- [5] J. C. Helton, F. J. Davis. Sampling-based methods for uncertainty and sensitivity analysis. *SANDIA Report SAND99-2240*, Sandia National Laboratories, Albuquerque, 2000.
- [6] J. C. Helton, F. J. Davis. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *SANDIA Report SAND2001-0417*, Sandia National Laboratories, Albuquerque, 2002.
- [7] J. C. Helton, R. M. Cooke, M. D. McKay, A. Saltelli, editors. *The Fourth International Conference on Sensitivity Analysis of Model Output (SAMO 2004)*, Santa Fe, New Mexico. Special Issue, *Reliability Engineering & System Safety*, 91(10-11):1105-1474, 2006.
- [8] J. C. Helton, J. D. Johnson, C. J. Sallaberry, C. B. Storlie. Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering & System Safety*, 91:1175-1209, 2006.
- [9] R. L. Iman, W. J. Conover. The use of the rank transformation in regression. *Technometrics*, 21:499-509, 1979.

- [10] R. L. Iman, W. J. Conover. A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics - Simulation and Computation*, 11:311–334, 1982.
- [11] R. L. Iman, M. J. Shortencarier, J. D. Johnson. A FORTRAN 77 program and user’s guide for the calculation of partial correlation and standardized regression coefficients. *SANDIA Report SAND85-0044*, Sandia National Laboratories, Albuquerque, 1985.
- [12] G. Klir, M. J. Wiermann. *Uncertainty-Based Information. Elements of Generalized Information Theory*. Physica-Verlag, Heidelberg, 1998.
- [13] W. T. Koiter. Current trends in the theory of buckling. In: B. Budiansky, editor. *Buckling of Structures. Proceedings IUTAM Symposium, Cambridge 1974*. Springer-Verlag, Berlin 1976.
- [14] V. Kreinovich, S. A. Ferson. A new Cauchy-based black-box technique for uncertainty in risk analysis. *Reliability Engineering & System Safety*, 85:267–279, 2004.
- [15] M. D. McKay, R. J. Beckman, W. J. Conover. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21:239-245, 1979.
- [16] I. Molchanov. *Theory of Random Sets*. Springer-Verlag, Berlin, 2005.
- [17] H. T. Nguyen. *An Introduction to Random Sets*. Chapman & Hall, Boca Raton, 2006.
- [18] M. Oberguggenberger. The mathematics of uncertainty: models, methods and interpretations. In: W. Fellin, H. Lessmann, M. Oberguggenberger, R. Vieider, editors. *Analyzing Uncertainty in Civil Engineering*. Springer-Verlag, Berlin, 2005.
- [19] M. Oberguggenberger, W. Fellin. Assessing the sensitivity of failure probabilities: a random set approach. In: G. Augusti, G. I. Schuëller, M. Ciampoli, editors. *Safety and Reliability of Engineering Systems and Structures. ICOSSAR 2005 - Rome*. Millpress, Rotterdam 2005, 1755–1760.
- [20] M. Oberguggenberger, W. Fellin. Reliability bounds through random sets: nonparametric methods and geotechnical applications. *Computers & Structures*, to appear.
- [21] A. Saltelli, I. M. Sobol. About the use of rank transformation in sensitivity analysis of model output. *Reliability Engineering & System Safety*, 50:225-239, 1995.
- [22] J. Shao, D.-S. Tu. *The Jackknife and Bootstrap*. Springer Verlag, New York, 1995.
- [23] G. I. Schuëller. On the treatment of uncertainties in structural mechanics and analysis. *Computers & Structures*, 85:235–243, 2007.
- [24] I. M. Sobol. Sensitivity analysis for nonlinear mathematical models. *Mathematical Modeling & Computational Experiment*, 1:407-414, 1993.
- [25] M. Stein. Large sample properties of simulations using Latin hypercube sampling. *Technometrics*, 29:143-151, 1987.
- [26] F. Tonon, C. L. Pettit. Toward a definition and understanding of correlation for variables constrained by random relations. In: G. Augusti, G. I. Schuëller, M. Ciampoli, editors. *Safety and Reliability of Engineering Systems and Structures. ICOSSAR 2005 - Rome*. Millpress, Rotterdam 2005, 1741–1745.

## 9. TABLES

$i$	Parameter $X_i$	Mean $\mu_i$
1	Initial temperature	293 K
2	Step1 thermal loading cylinder1	450 K
3	Step1 thermal loading cylinder2	350 K
4	Step1 thermal loading cylinder3	150 K
5	Step1 thermal loading sphere1	150 K
6	Step1 thermal loading sphere2	110 K
7	Step2 hydrostatic pressure cylinder3	0.4 MPa
8	Step2 hydrostatic pressure sphere1	0.4 MPa
9	Step2 hydrostatic pressure sphere2	0.4 MPa
10	Step3 aerodynamic pressure	−0.05 MPa
11	Step4 booster loads y-direction node1	40000 N
12	Step4 booster loads y-direction node2	20000 N
13	Step4 booster loads z-direction node1	3.e6 N
14	Step4 booster loads z-direction node2	1.e6 N
15	Step4 mechanical loads x-direction	100 N
16	Step4 mechanical loads y-direction	50 N
17	Step4 mechanical loads z-direction	300 N

TABLE 1: Description of input parameters no. 1 – 17.

Fuzzy set	HL-measure
no fixing	0.1481
$X_{13}$ fixed	0.0398
$X_{14}$ fixed	0.1430
$X_3$ fixed	0.1268

TABLE 2: Hartley-like measures of outputs, non-interactive input.

Fuzzy set	HL-measure
no fixing	0.1357
$X_{13}$ fixed	0.0287
$X_{14}$ fixed	0.0329
$X_3$ fixed	0.1011

TABLE 3: Hartley-like measures of outputs, interactive input.

## 10. FIGURES

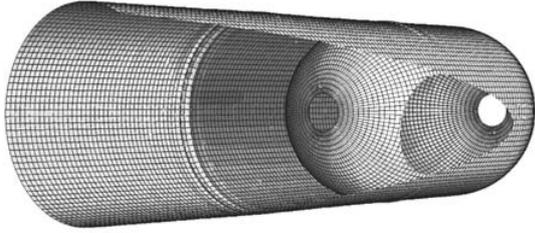


FIG. 1: Simplified finite element model.

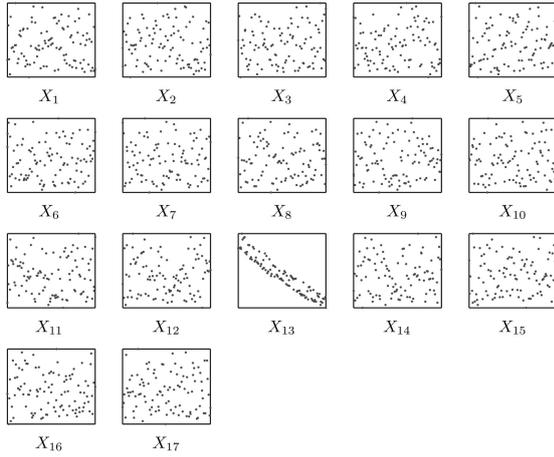


FIG. 2: Scatterplot input – output.

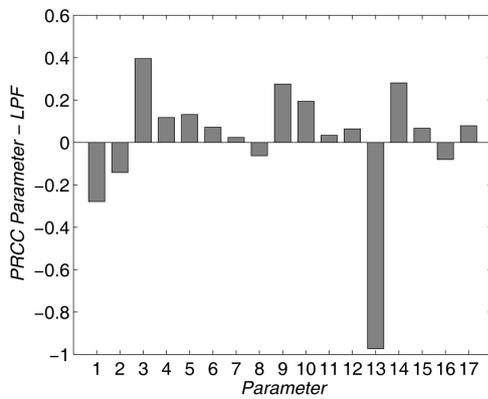


FIG. 3: Partial rank correlation coefficients.

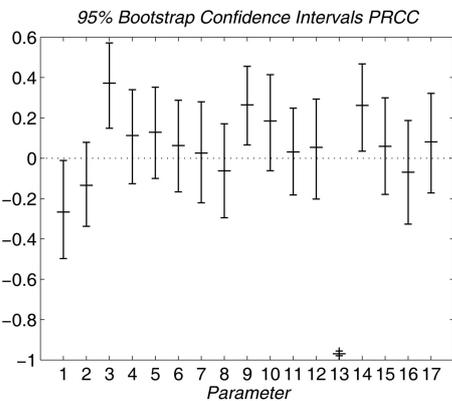


FIG. 4: Partial rank correlation coefficients, confidence intervals.

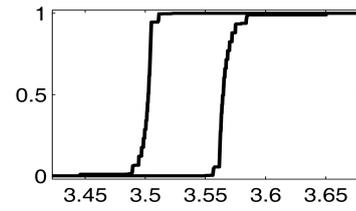


FIG. 5: Probability box: LPF, 3 input variables.

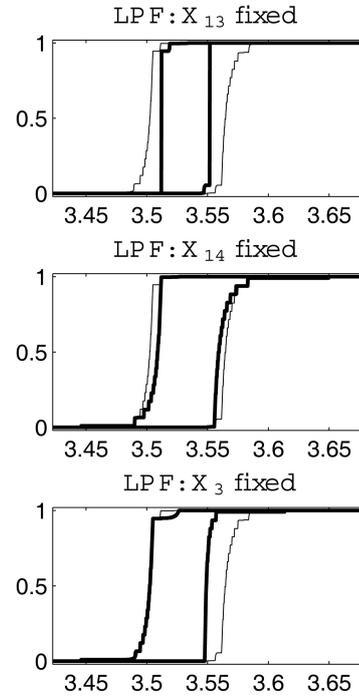


FIG. 6: Probability box: LPF, frozen variables.

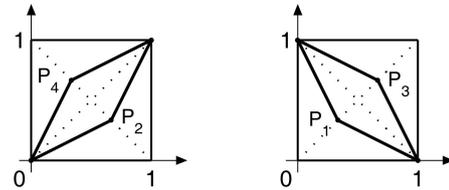


FIG. 7: Positive/negative interactivity.

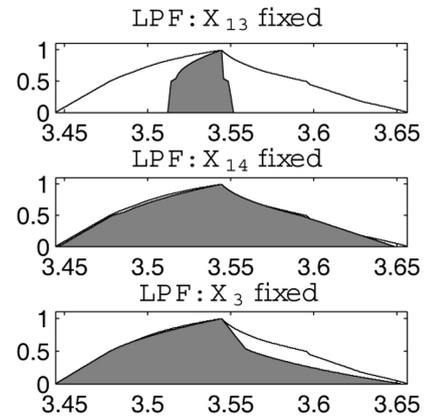


FIG. 8: Fuzzy sets: LPF, frozen variables, noninteractive case.

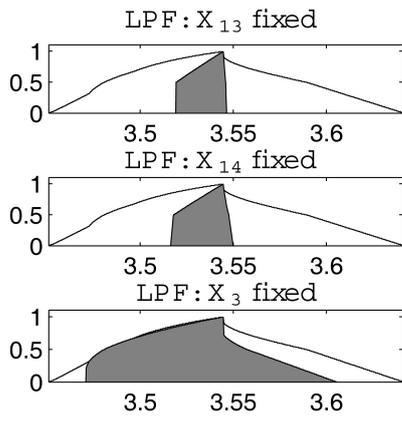


FIG. 9: Fuzzy sets: LPF, frozen variables, interactive case.