DENSITY PLOTS FOR AIRPORT CAPACITY ASSESSMENT

S. Kellner, RWTH Aachen, 52074 Aachen, Germany

Abstract

This paper introduces genomic data visualization techniques to aircraft flows at airports. The density plots are enhanced by envelopes and centroids to allow the assessment of capacity figures. Results of four European airports are presented and support the applicability of the approach.

1. INTRODUCTION

In the past years, air traffic increased constantly. Current forecasts propose further growth. Reference [1] estimates that 11% of demand cannot be served by the airports in the European area (ESRA) in 2030. To increase available airport capacity, processes and infrastructure need to be addressed. One promising approach for improving the former is to compare similar airports. This paper proposes a visual method for assessing derived capacity values of airports by combining the Gilbo approach [2] with density plots. This is motivated by needs for easy to follow communication of capacity figures.

2. LITERATURE REVIEW AND THEORY

This chapter introduces the two pillars of theory relevant for this paper: deriving figures for airport capacity and visualization of data.

2.1. Airport Capacity

Capacity is defined as the "processing capability of a service facility over some period" [3]. With respect to airports, in most cases the capacity of the runway system is limiting total system capacity. Capacity is always connected to the level of service (LOS), which is often measured by delays. Given an agreed LOS, three categories of deriving capacity figures can be distinguished:

1. Analytical approaches,

2. Simulation studies and

3. Empirical analyses.

In the following, the categories are briefly summarized. For a detailed overview please refer to [4] or, more recently [5].

2.1.1. Analytical approaches

They apply queuing theory and stick to separation minima with respect to wake turbulence classes. A good example is the Blumstein model [6], focusing on estimating the capacity of a single runway with arrivals only. Analytical models are limited to simple runway configurations and cannot handle complex dependencies with the terminal maneuvering area (TMA) or apron of an airport. This is where simulation studies can be used.

2.1.2. Simulation studies

They require a detailed model of an airport system, which consists of infrastructure, e.g. apron, taxiways, runways, TMA; procedures, e.g. arrival separation, runway crossings, runway usage, taxiing concept; and the schedule, which covers the traffic mix and demand pattern. The effort even for simulation studies covering only parts of the airport is high. Often only the airport operator has sufficient data to build up a simulation model. This leads to empirical analyses of traffic flows.

2.1.3. Empirical analyses

They take the flow of an airport, eliminate outliners and derive a figure as sustainable capacity value. Common methods are the k-%-hour or the 30th busy day. This paper contributes to the third category by applying the approach described by Gilbo [2]: Envelopes are constructed for all arrival-departure pairs occurring at least once, twice and so on. These envelopes help eliminating outliners. The capacity value point is found by the intersection of the bisecting line and the respective envelope. The truncated sum of the arrival and departure values of the capacity value point is the resulting capacity figure. It is a measure for the sum of arrivals and departures only. The relation of the capacity figure to the handled movements of the airport can be found by comparing its frequency with the frequency of other arrival-departure pairs. In this paper a visual approach for this assessment is suggested.

2.2. Visualization of data

In general, two applications of appropriate visualization of data can be named: data pre-processing and presentation of results. The former is part of every empirical work and contributes to the reliability of the resulting analyses. Depending on the type of data, visualizations help to eliminate outliners, point to scaling problems and provide a first idea of potential patterns. With respect to high-dimensional data, plotting while keeping relations between the values becomes a complex task. This is addressed by scaling techniques, which are out of scope of this paper (see for example [7], with an application to aviation and weather data). The power of charts in presenting results

can be simply summarized by the proverb "a picture is worth a thousand words". In the following common chart types and typical examples in the context of airport capacity are briefly summarized.

2.2.1. Bar and line charts

Bar charts are composed of bars with lengths proportional to the value they represent. They are used to compare values plotted next to each other. Line charts can be derived from bar charts by connecting the peaks of the bars. They draw the reader's attention on trends in the data, like in time series. Applied to aviation, one could order hourly flight movements by magnitude, starting with the highest value. Summing up the area below the line chart, the ordinate, which denotes the number of movements, and an orthographic line, a percentage of movements can be defined. The related value on the abscissa is then referred to as the k-%-hour of an airport. Common values for the percentage are 1 or 5.

2.2.2. Scatter plots

Plotting two variables scatter plots can be generated. Often a functional relationship between the two variables is assumed. In this case the independent variable is denoted on the ordinate and the dependent variable on the abscissa. These plots are mainly used to detect patterns in data, which can be analytically found by for example clustering algorithms. Depicting the capabilities of an airport runway system, arrivals and departures within a defined time interval are plotted. As arrivals are commonly prioritized over departures, they are denoted as the independent variable. The Gilbo approach employs these plots.

2.2.3. Adding more variables

A third variable may be visualized by so called 3D-plots. In mathematical terms these plots are two-dimensional projections of three-dimensional objects. A vertical axis is added which denotes the third variable. The projection tends to result in difficulties of reading the plot if elements are hidden due to the angle of vision. Interactive media like computers can adjust the projection to the needs of the user and thus circumvent this problem. A possible solution for printed publications is to express the additional variable by colors or symbols. The use of colors to depict density is proposed by [8], originating in the field of visualizing genomic data. In [9] a combination of symbols and colors is applied to illustrate the frequency of arrival-departure pairs.

3. METHODOLOGY

Three elements need to be arranged:

• 	scatter plots of arrival-departure pairs,

• 	envelopes derived according to [2] and

• 	the frequency of the pairs.

The scatter plots can be constructed by simply plotting the arrival-departure pairs in a two-dimensional chart. For deriving an hourly capacity value it is suggested applying the rolling hour criteria. It counts the number of movements within each 60-minute interval. Opposed to counting the number of movements within hours according to the clock scheme, rolling hours are defined as follows: Define a shift interval of a time units. Define a starting point s. The number of movements in the rolling hour r_i is then given by the number of movements in the interval $[s + i \cdot a; s + i \cdot a + 60)$. This is done for arrivals and departures separately.

The envelopes are added to the plot and the bisecting line is used to find the capacity value. The selection of the appropriate curve depends on the number of points which should be excluded. In this paper the envelope of pairs occurring at least three times is used. It can be justified by comparing the percentage of movements excluded. The bisecting line could be replaced by a line starting from the origin with the slope of the relation of arrivals to departures. To keep the plot simple this adjustment has been left out in this paper.

The frequency of the pairs is derived by counting the number of equal arrival-departure pairs. Then, each value of frequency is assigned to a color value. It is suggested using a color palette based on one color and black as well as correlating the darkness of the color to the frequency. Treating the data as empirical and for smoothing the plot, we propose using a 2D kernel density estimate as described in [10]. It is implemented in the recommended R package 'KernSmooth' [11]. For the generation of the color palette the 'smoothScatter' function of the 'geneplotter' R package is used [12].

The resulting plot allows comparing the location of the capacity figure c with the centroid g of operations, i.e. the darkest area and the envelopes. The capacity figure is not limited to empirically derived capacity figures and may include declared capacities from Europe, which are used for defining available airport slots according to [13] or capacities depending on weather classes, as proposed by [9]. Assuming an airport which, at least on a few days, operates close to saturation, three cases of location of centroid and capacity figure can be distinguished:

• 	Capacity figure below centroid: conservative figure below the capabilities of the airport

• 	Capacity figure close to centroid: represents predominant number of movements

• 	Capacity figure above centroid: figure exceeds usual number of movements

If the capacity value is beyond the envelopes, the meaning of this value is questionable. Capacity figures close to saturation capacity should be located in proximity of the envelopes.

The capacity figure should represent total movements, i.e. the sum of arrivals and departures and be located on the bisecting line. The terms "below", "close to" and "above"

then refer to the position of the terminal point of the orthogonal projection of $\vec{g}$ onto the vector $\vec{c}$ and the terminal point of the vector $\vec{c}$.

The centroid is calculated as the average of all arrival-departure pairs.

4. RESULTS

The methodology is tested with data collected by the Central Flow Management Unit (CFMU) covering the period of 2008-01-01 to 2008-12-31 of 20 European airports. The dataset includes the actual time of departure (ATOT) and the actual landing time (ALDT) of flights operated under Instrument Flight Rules (IFR), only. An internal comparison of airport movement data with CFMU data revealed a divergence of 1% to 2.5%, with airport movement data being the higher value. The divergence is negligible in the context of this paper. The time frame for calculating rolling hours is 06:00-22:00 local time. The time interval applied is 5 minutes. Capacity figures derived by the intersection of bisecting line and envelopes are not listed here. The declared capacity value represents an average declared hourly capacity.

The results of four airports are presented in detail. These are Brussels (EBBR), London Heathrow (EGLL), Helsinki-Vantaa (EGHK) and Lisbon Portela (LPPT).
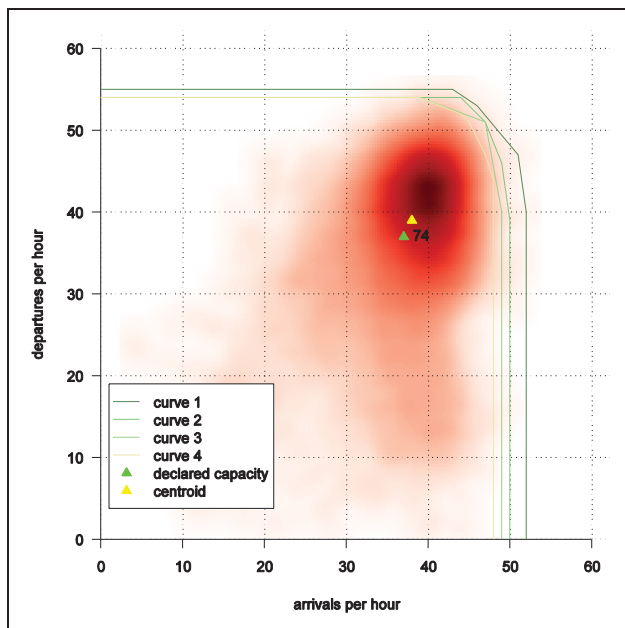

FIGURE 1. Density plot for London Heathrow (EGLL)

Starting with EGLL (see Figure 1), limited variations in movements become clear. All operations are concentrated on a small area, with little deviations. Centroid and capacity value are close and below the envelopes. Nevertheless, a certain portion of arrival-departure pairs is above the declared capacity.
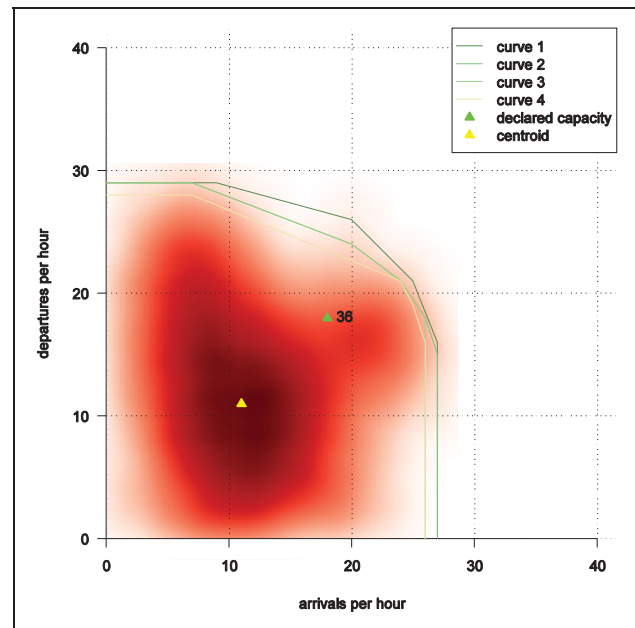

FIGURE 2. Density plot for Lisbon Portela (LPPT)

Although the magnitude of total movements is lower, LPPT (see Figure 2) can be compared to EGLL with regard to the location of centroid and capacity value. Differences are obvious, that LPPT uses the whole range of possible arrival-departure pairs.
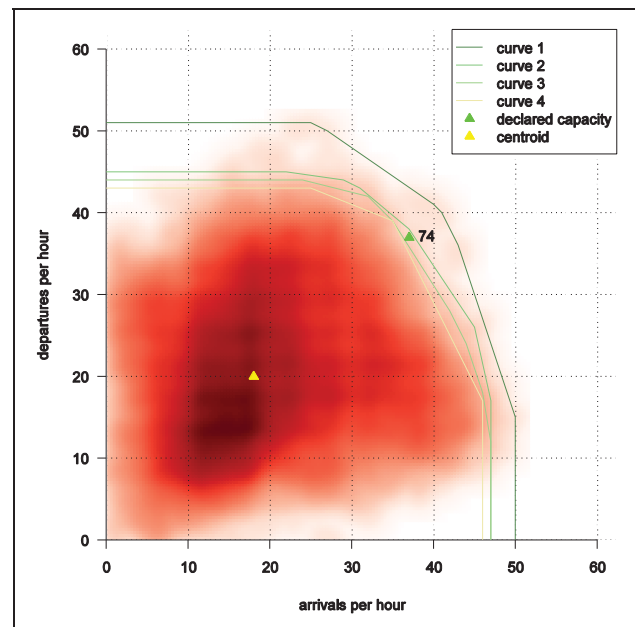

FIGURE 3. Density plot for Brussels (EBBR)

In the case of EBBR (see Figure 3) the distance between centroid and capacity value is larger. But the density plot reveals that arrival-departure pairs at EBBR are more equally distributed. The declared capacity value is located above envelope three. This suggests that declared capacities in EBBR are relatively high.
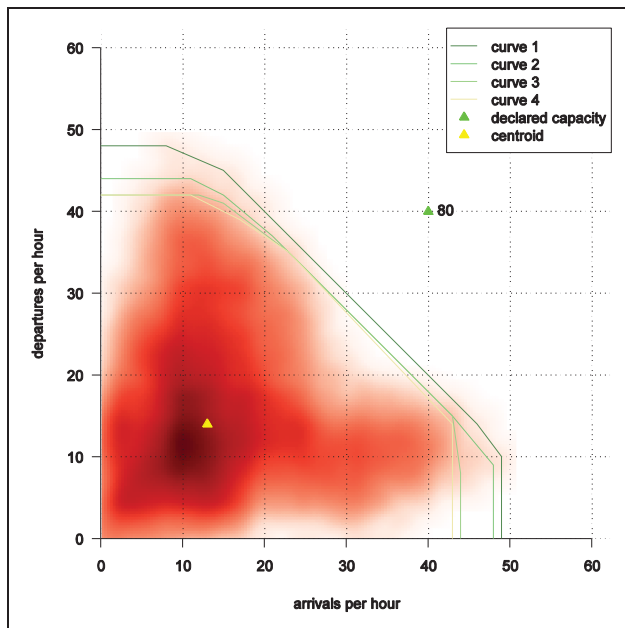
FIGURE 4. Density plot for Helsinki-Vantaa (EGHK)

An example where the method of defining a reasonable centroid seems to fail is found for EFHK (see Figure 4). The pattern is unbalanced and shows phases of departure and arrival priority. In addition the declared capacity value is far beyond the envelopes. This suggests that the airport does not operate close to saturation.

5. CONCLUSIONS AND OUTLOOK

The preliminary results seem promising in understanding capacity figures of an airport. The density plots provide easy to read charts which offer a detailed picture of the operations. Envelopes further sharpen the image. Derived capacity values can be put into context and judged by comparison with envelopes and centroids.

As the results suggest, more detailed investigations with appropriate data are possible. In this paper, no differentiation with respect to weekday/weekend, aircraft mix, runway configuration or weather has been applied, but could offer insight into the change of capacity under variation of these parameters. In addition, LOS criteria like delays should be included. The method for deriving the reasonable centroid of the density plot is subject to further studies. Once the exact place of the centroid is found, the distance between the terminal point of the orthogonal projection of $\vec{g}$ onto the vector $\vec{c}$ and the terminal point of the vector $\vec{c}$ can be applied to further increase the level of detail in comparisons.

6. ACKNOWLEDGMENT

7. REFERENCES

[1] EUROCONTROL, "Challenges of growth", Brussels, 2008.

[2] E. P. Gilbo, "Airport capacity: representation, estimation, optimization", IEEE Transactions on Control Systems Technology, vol. 1, no. 3, September 1993.

[3] R. Horonjeff, F. X. McKelvey, Planning and Design of Airports. Boston, MA: McGraw-Hill, 1994.

[4] International Center for Air Transportation, "Existing and required modeling capabilities for evaluating ATM systems and concepts", Massachusetts Institute of Technology, Cambridge, MA, March 1997.

[5] C. Barnhart, P. Belobaba, A. R. Odoni, "Applications of operations research in the air transport industry", Transportation Science, vol. 37, no. 4, pp. 368–391, November 2003.

[6] A. Blumstein, "The landing capacity of a runway", Operations Research, vol. 7, no. 6, pp. 752-763, 1959.

[7] F. Rehm, F. Klawonn, G. Ruß, R. Kruse, "Modern data visualization for air traffic management", Proceedings of the Annual Meeting of the North American Fuzzy Information Processing Society, pp. 19-24, 2007.

[8] R. Gentleman, F. Hahne, W. Huber, "Visualizing genomic data", Bioconductor Project Working Papers, bepress, 2006.

[9] Federal Aviation Authoriy, "Airport capacity benchmark report 2004", Washington, D.C., 2004.

[10] M. P. Wand, M. C. Jones, Kernel Smoothing. CRC Press, 1995.

[11] Documentation available online: http://rss.acs.unt.edu/Rdoc/library/KernSmooth/html/00Index.html

[12] Documentation available online: http://rss.acs.unt.edu/Rdoc/library/geneplotter/html/00Index.html

[13] Council Regulation (EEC) No 95/93: "Common rules for the allocation of slots at Community airports" 1993-01-13.