CHALLENGES IN EUROPEAN AEROSPACE
5TH CEAS AIR & SPACE CONFERENCE

Developing Generic Flight Schedules for Airport Clusters

Peter Bießlich
Institute of Air Transportation Systems, Hamburg University of Technology (TUHH)
Research associate and PhD student
Blohmstr. 18, 21079, Hamburg, Germany
peter.biesslich@tu-harburg.de

Stefan Reitmann (Dresden University of Technology, TUD), Volker Gollnick (Hamburg University of Technology, TUHH), Klaus Lütjens (German Aerospace Center, DLR), Karl Nachtigall, (Dresden University of Technology, TUD), Steffen Marx (Dresden University of Technology, TUD)

NOMENCLATURE

AC	Aircraft
ADI	Sabre Airport Data Intelligence
ARR	Arrival Flight
BogS	Builder of generic Schedules
Cat	Category
DEP	Departure Flight
DIST	Distance
FoAM	Forecast of Aircraft Movements
IFR	Instrument Flight Rules
PAX	Passenger
O&D	Origin & Destination
SLF	Seat Load Factor
RWY	Runway

ABSTRACT

Flight schedules represent a crucial element of airline as well as airport operations. Information which airline fly on which time, to which destination, using which aircraft type are listed in a typical flight schedule and can be extended by any further information. But how a flight schedule of a cluster of similar but not identical airports looks like?
The paper's main objective is to implement a methodology of developing generic flight schedules, one for each pre-defined airport cluster, out of a set of real flight schedules and segments of the cluster representatives. Statistical and probability distributions are used to determine a suitable weekly distribution of arrival and departure flights, which is filled up with flight-specific data of such an airport cluster. Every airport-specific data like origin/destination information, specific aircraft type, passenger volume etc. are transferred into generic categories.
The methodology is implemented in a model called "Builder of generic Schedules", short BogS. It is shown that a generic flight schedule for a present situation can be provided and is suitable for airport cluster representatives. Furthermore a future scenario can be outlined, issuing the attributes future development, especially the future frequency of aircraft movements.

1 INTRODUCTION

1.1 Motivation

Airports have their operational focus on the short term. Nevertheless, a long term view is crucial for planning and developing their infrastructure and business. Software-assisted airport models are proper for the airport's expansion and business assessment but also for investigations in the field of airport research, whereas such models often focus on single airport elements only. This scope is not sufficient to address the mechanisms of the air transportation system as a whole.

From an holistic, air transport system point of view a generic airport model is useful, that covers operational aspects, e.g. passenger and aircraft movement demand, infrastructural aspects as well as resulting economic changes (revenues). The granularity should be as rough as possible, reducing the overall complexity and computing time, but accurately enough to model crucial intra-airport relationships. Therefore a generic airport model shall not represent one single airport in detail, but shall be suitable for an airport cluster. In order to assess intra-airport relationships, a network of generic airports is necessary, where the BogS model will be used to determine sound flight schedules. Our hypothesis is, that similar airports within a cluster have quite comparable flight operations over a week and, hence, only one generic flight schedule is representing the cluster. Using the outcome of this paper, a possible, scenario-based evolution of a certain network shall be shown in the long term, as well as the impact of introducing emerging technologies.

1.2 Literature review

Literature illustrates some approaches for grouping airports, usually distinguishing them by defined threshold values. Airport Council International (ACI), an association of airport operators, classifies four groups of airports simply using their yearly passenger volumes [1]. Azzam [2] introduces a new airport taxonomy based on flight plan data from 1979 to 2007 using network performance figures. By this, specific statements can be made about the airports evolution and function within the air transportation network at a defined point in time, putting them into a geographical context. Azzam uses a hierarchical cluster analysis, defining 12 different airport classes based on six network parameters. Oettel et al. [3] use clustering techniques to develop an application-oriented airport classification for air traffic simulation purposes. In particular, the single linkage algorithm is applied to identify outliers and similarities among a set of airports. Oettel et al. indicate that it makes sense to limit the set of parameters according to the application. The more parameters are considered for classification, the bigger the application field becomes, but also the smaller the sample and representativity of results.

Several activities are outlined in the literature concerning the development of airline schedules. But this issue cannot be seen on its own. Jacobs et al. [4] characterize it as an integrated, intermeshed process with strong overlapping to the airline marketing and distribution. Hence, the airline scheduling determines not only where and when the airline will fly (also called flight leg or segment from A to B), furthermore it focuses on the passenger origin and destination (O&D) market. The schedule is built to satisfy the passenger demand and to maximize airline profitability. The bigger and more competitive the airline business became over the past decades, the more sophisticated the development of flight schedules became as well. Therefore methods of operations research and mathematical modeling are applied in order to solve the industry problems (for details see [4], [5]).

In contrast to that, an airport flight schedule is only a mixture of different airline schedules and the airport has only marginal influence on that. Nevertheless typical attributes and influencing circumstances must keep in mind. Objectives are a most optimal utilization of the infrastructure and the airport's own profit maximization. The operating airline is subsidiary. Nevertheless similar activities to the paper in hand, resulting in a generic flight schedule suitable for different airports, could not be found.

2 APPROACH FUNDAMENTALS

Two beforehand accomplished analyses are incorporated in this approach of developing generic flight schedules. The fundamental and main input builds an airport cluster analysis. Furthermore the Forecast of Aircraft Movements (FoAM) model is used to determine the future arrival and/or departure frequency of a certain airport cluster. Both analysis will be introduced briefly in the following.

2.1 Clustering of Airports

Clustering is an algorithm for forming functional groups, whereby the objects of one group feature maximum similarity and likewise minimal similarity to objects in other groups [6]. Numerous methods of clustering can be found in literature (amongst others [6], [7]), but there is no single approach that is applicable for all kinds of analysis. Choosing an appropriate clustering algorithm largely depends on the investigated objects. A hierarchical agglomerative clustering method was applied to determine airport similarities. Representing a common linkage rule, the minimum-variance-linkage/Ward-linkage has been selected. Advantages of an hierarchical approach are the flexibility regarding the number of clusters and a non-specification of initial conditions. The algorithm merges the cluster pair that causes the smallest increase of the sum of squared errors, thus ensuring a maximum homogeneity per time step. Having calculated the distances between all pairs of objects, the two objects merge being most similar. Furthermore the Euclidean distance is chosen as measure of similarity/proximity. Seven clusters result out of this analysis and characterized by attributes out of different categories. Table 1 and Table 2 show the attributes used for the cluster analysis (blue) and the means of attributes for every cluster. Highlighted are the smallest (red) and biggest (green) value per attribute.

Table 1: Attribute means of clusters (1)

Cluster	N	Name	Passengers [PAX/Year]	Movements [MOV/Year]	Cargo [Tons/Year]	Transfer PAX [%]
1	38	Small Regional Airports	5 001 000	65 000	59 000	8.20
2	20	Medium Regional Airports	8 477 000	126 000	176 000	12.79
3	23	International Airports	17 236 000	176 000	199 000	15.05
4	29	Secondary Hub Airports	25 460 000	282 000	350 000	16.48
5	24	International Hub Airports	45 843 000	398 000	1 039 000	38.84
6	4	High Frequency Hubs	66 372 000	780 000	600 000	43.34
7	8	Cargo Hubs	25 098 000	230 000	2 838 000	20.46

Table 2: Attribute means of clusters (2)

Cluster	Nb. of RWYs	Traffic Mix [%]				Revenue/PAX [US$]	Revenue [US$]	Distance to City Center [km]
		S	H	M	L			
1	2	0.00	8.29	90.46	1.25	18.18	93 664 000	18.97
2	4	0.00	2.29	97.26	0.45	12.30	105 840 000	17.87
3	2	0.00	15.72	83.64	0.65	22.48	384 324 000	24.01
4	3	0.00	17.24	81.97	0.79	16.67	408 486 000	19.87
5	3	0.33	29.69	69.65	0.33	19.15	907 114 000	23.72
6	6	0.00	10.39	89.12	0.49	8.86	553 736 000	31.03
7	3	0.06	38.88	60.14	0.91	26.05	828 540 000	36.90

Distinctive clusters are International Hub Airports, High Frequency Hubs (graphical visualized in Figure 1) and Cargo Hubs. International Hub Airports represent typical hubs of major airlines, e.g. Frankfurt Airport (FRA) or New York´s John F. Kennedy Airport (JFK). A high proportion of transfer passengers (38 %) and share of heavy aircraft (30 %) are characteristically. In addition, about 60 % of the airports of this class are approached by the currently largest airliner, the Airbus A380.
High Frequency Hubs represent typical major U.S. airports and include 4 airports only, e.g. Hartsfield-Jackson Atlanta International Airport (ATL). Providing six runways or more, these airports serve in average 780.000 aircraft movements per year, by far the largest value compared to the other clusters. This frequency allows for a volume of around 66 million passengers annually and also a variety of transfer options for passengers, reflecting in the highest transfer passenger share of 43 %. Despite these figures, only a revenue of US$ 8.86 per passenger is generated, representing the lowest value of all classes. Airports characterized by an annual cargo volume of 2.8 million tons are defined as Cargo Hubs. Due to the large volume of cargo, a considerable total revenue of US$ 828 million is generated. At the same time only 25 million passengers per anno use these airports, resulting in the best revenue per passenger value with US$ 26. Furthermore the highest share of heavy aircraft (around 39 %) is achieved by this airport cluster. For instance, this cluster features Memphis International Airport (MEM) and Hong Kong Chek Lap Kok International Airport (HKG). A dendrogram, showing all representatives of the clusters can be found at the appendix and further detailed information about the airport clustering in [8].

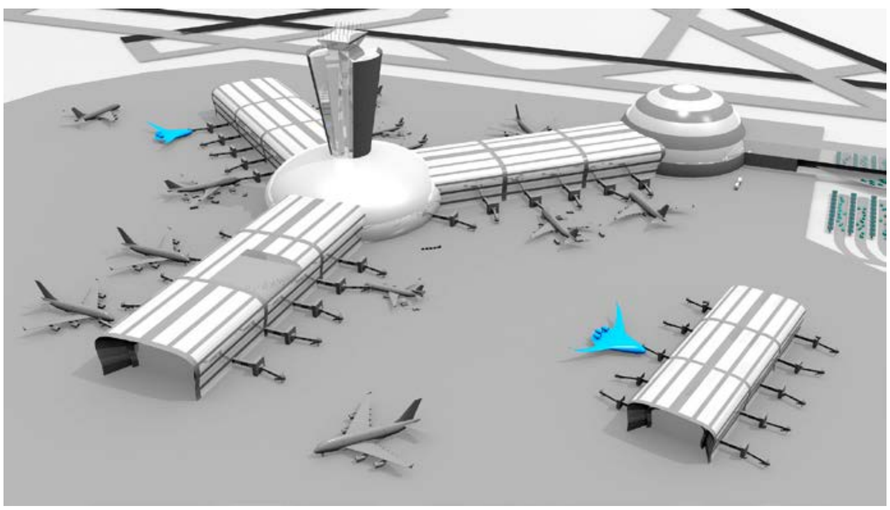

Figure 1: Visualization of High Frequency Hub

Depending on the number of objects (here airports) and choice of clustering attributes the result of the clustering changes, which reveal the advantages as well as disadvantages of a clustering methodology. Most positive effect is the volatility of the clustering approach due to an adaptability to the scope of research. A focused analysis on geographical regions or a certain characterizing category (e.g. effectiveness of an airport) can be executed easily. Furthermore functional groups on statistical data are derived, in contrast of using rigid threshold values. On the other hand this volatility is also disadvantageous. Changes in the number of investigated objects, the clustering parameters or the clustering method can cause a different result, even different clusters. Therefore, it is important to have a basic understanding of the desired results and previous knowledge of the research topic. However, the method of determining generic flight schedules presented in this paper can fully adapt to any further clustering.

2.2 Forecast of Aircraft Movements (FoAM) Model

FoAM represents a forecast model that has been developed to determine the behavior of airlines when it comes to a passenger growth on certain flight legs. The airlines are able to deal with this situation by either increasing the frequency or the aircraft size, which may entail different numbers of aircraft movements. Purpose of FoAM is to forecast a typical fleet mix and the growth of aircraft movements on flight segments worldwide based on an assumed passenger growth.
The model´s basic assumption is that the size of assigned aircraft and the flight frequency depends on two essential determinants: First, the passenger volume and, second, the segment distance. This approach was chosen because it supports the commonly accepted claim that the average aircraft size increases with growing distance and passenger volumes. Initially, each flight leg worldwide is assigned to a distance, passenger number and aircraft category. Each type of category comprises around seven to eight subcategories itself. It means, that a flight segment can be assigned to one of 448 different combinations of categories. Figure 2 illustrates this categorization, which can be adjusted in the categories number and boundaries for further investigations. [9]

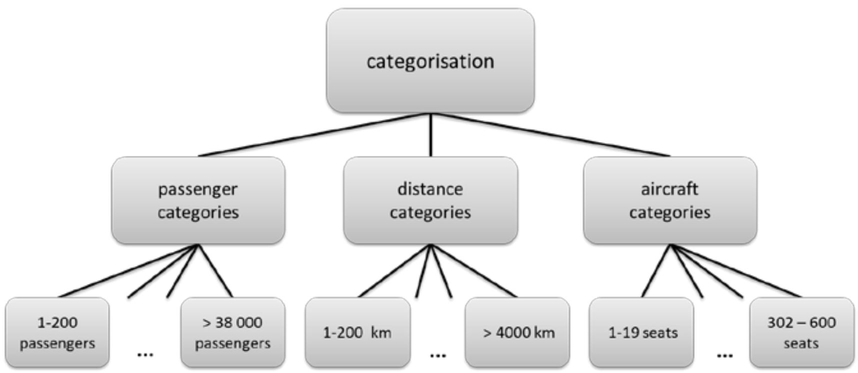

Figure 2: Pattern of categorization within FoAM

A typical mix of aircraft size is derived empirically for all categories, exemplarily shown at Figure 3 for the distance category 401 - 800 km.

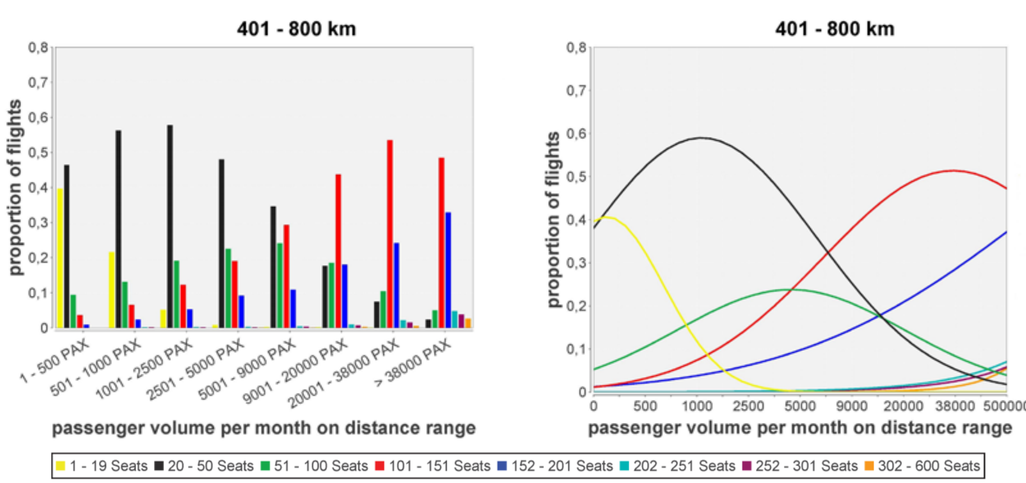

Figure 3: Discrete (left) and continuous (right) distribution of aircraft sizes, distance category 401 - 800 km

It indicates that the share of any specific aircraft sizes reaches a maximum at a certain passenger volume. As the passenger volume shifts away from this value, the share of the specific aircraft size decreases. The findings for other distance categories look quite similar. Once the mix of aircraft sizes α is identified, the frequency on one specific segment can be retrieved with means of the passenger volume pax, the seat load factor SLF and the average equipment capacity in a passenger category $\text{Ø}\,seats$. Therefore, the average passengers per flight are calculated from the weighted sum over all aircraft categories of the products of the capacity of all aircraft categories and the seat load factor. The segment frequency can then be modeled by dividing the total passenger number by the average passengers per flight.

$$\text{segment frequency} = \frac{\text{pax}}{\text{Øpax per flight}} = \frac{\text{pax}}{\sum_{\text{AC}} \alpha_{\text{AC}} * \text{SLF}_{\text{Dist}} * \text{Ø seats}_{\text{AC}}} \tag{1}$$

The global number of air traffic movements is the sum of the frequencies of the segments worldwide. Using the given formula and the passenger volumes of the base year allows modeling the global frequency for the base year. Eventually, applying an assumed passenger growth (e.g. 4.7 % p.a.) to all segments, a forecast of frequencies per segment is calculated (e.g. 2030). The global growth of air traffic movements per anno can thus be estimated by the sum of all frequencies. Further detailed information about FoAM are presented in [9].

3 FLIGHT SCHEDULE DETERMINATION

The determination of generic flight schedules for the aforementioned clusters is outlined in the following.

The fundamental methodology is divided into four phases. Basis of this study are real flight plan data from ADI [10] of 144 airports considered in the clustering (section 2.1). Appropriate flight schedules are merged and filtered according to predefined requirements (phase A). In a second step specific attributes are determined, amongst others the weekly flight distribution, using common distribution functions (phase B).

Phase C comprise the final modeling of the present flight schedule (here 2012), which can also be described as generic re-modeling of the initial, real flight schedules. The algorithm concludes with a forecast flight schedule (here 2030), primarily based on the findings of phase B.

The methodology is implemented in JAVA and is adaptable to further user-defined requirements. Figure 4 gives a graphical overview of the four phases.

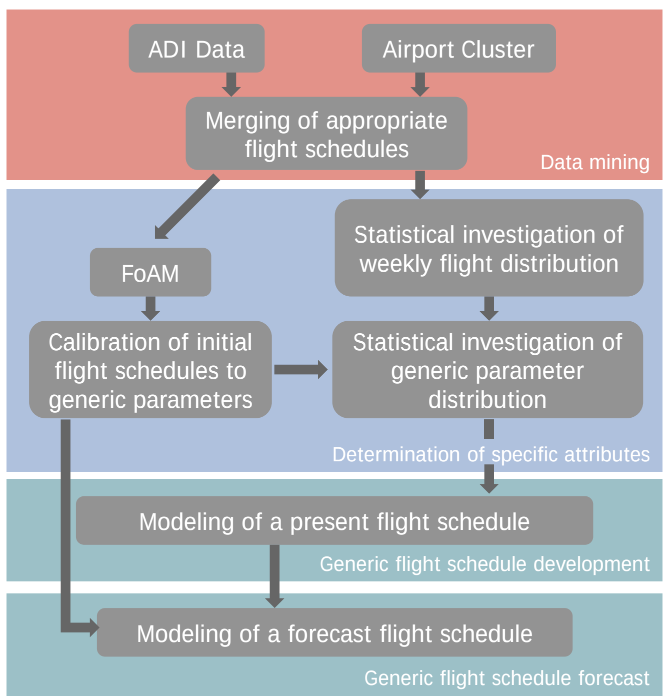

Figure 4: Methodology of determination a generic flight schedule (4 phases approach)

3.1 Phase A: Data mining

Phase A can be entitled as a simplified data mining process, the application of statistical methods to big data amounts in order to extract relevant data or to identify new cross connections.
First step within phase A is a data filtration of the real flight schedules and segments. In order to get a vital consistency, requirements were defined which derive from the character of flight schedules. The algorithm distinguishes between real flight schedules out of a cluster, meaning a certain number of airports, of a time, meaning summer (1st Apr – 31st Oct) or winter (1st Nov – 31st Mar) and of a flight type, which means if departure or arrival flight are considered. Furthermore night curfews and regional differences are taken into account.
After filtration the single attributes are transformed into generic attributes. Considered are three classifications: an aircraft category, a distance category and a passenger category, which are adopted from FoAM (see Figure 2). Finally, the outcome of phase A is a sorted schedule, representing the fundamental for further process steps. Figure 5 summarize the steps of phase A.

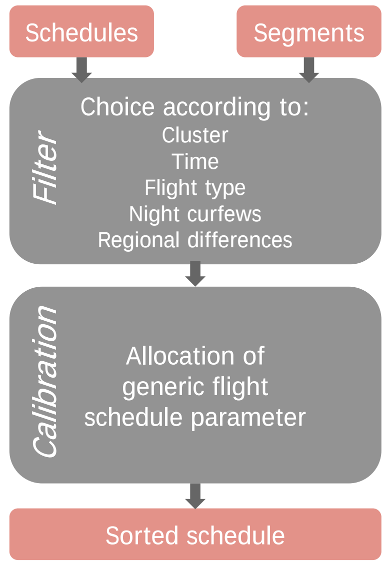

Figure 5: Data mining

3.2 Phase B: Determination of specific attributes

Initially, phase B includes the determination of a weekly flight distribution, i.e. the number of flights performed within a distinct time slot (for example three flights within slot 8:01 – 8:30) over a weekly time period (Monday - Sunday). This discrete curve is encompassed by an upper and lower confidence interval. Figure 6 illustrates this weekly flight frequency distribution.

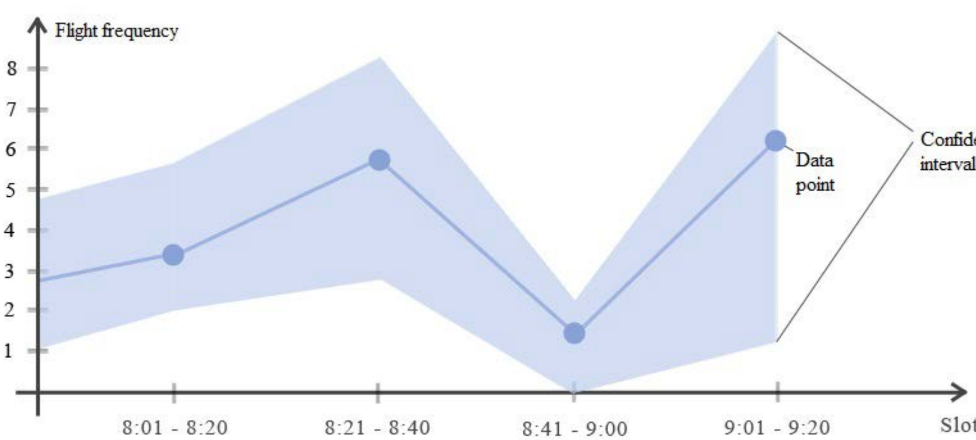

Figure 6: Exemplary weekly flight distribution

To determine a most precisely weekly flight distribution for every cluster the algorithm tries to describe every single time slot by a common probability mass function and verify this choice by a chi-square test. Assuming the airports within a cluster have similar flight operations, the frequencies should accumulate around the mean. Therefore, and most likely, a modeling by using a Gaussian distribution is verified firstly. Due to the symmetry of the Gaussian distribution the confidence intervals are calculated according to (2) and (3), with the expectation μ and the standard deviation σ.

$$I_{upper} = \mu + 1.96 * \sigma \quad (2)$$

$$I_{lower} = \mu - 1.96 * \sigma \quad (3)$$

If the null hypothesis is rejected by the chi-squared test, i.e. the distribution of the sample is not a Gaussian distribution, a possible skewness of the distribution is checked. This characteristic can appear at clusters with small airports or clusters with a large number of representatives (sample size). We use a Weibull distribution for distributions skewed to the left. Hence, the confidence intervals are calculated according to (4) and (5) for a Weibull-distributed sample (approximated by a F-distribution) [11], whereby n represents the sample size, $F_{(m_1, m_2,\ s)}$ is the F-distribution value and i = 1. Distributions skewed to the right are modeled by using exponential distributions. Confidence intervals are calculated according to (6) and (7), whereby $\bar{X}$ represents the mean of X and $\chi^2_{2n;\frac{\alpha}{2}}$ is the chi-squared value. If the three aforementioned hypothesis are rejected by the chi-squared test a uniform distribution is chosen. This would be the worst case solution and does not speak for the clustering, because no significant peak of flight frequencies per slot can be detected.

$$I_{i,upper} = 1 - \frac{1}{1 + \frac{i}{n-i+1} F_{2i,2(n-i+1),\alpha/2}} \quad (4)$$

$$I_{i,lower} = \frac{1}{\frac{n-i+1}{i} F_{2(n-i+1),2i,\alpha/2} + 1} \quad (5)$$

$$I_{upper} = \frac{2n\bar{X}}{\chi^2_{2n;\frac{\alpha}{2}}} \quad (6)$$

$$I_{lower} = \frac{2n\bar{X}}{\chi^2_{2n;1-\frac{\alpha}{2}}} \quad (7)$$

3.3 Phase C: Generic flight schedule development

Phase C encompasses the generation and allocation of appropriate flights out of the sorted schedule of phase A to the calculated frequencies of phase B. The crux of the matter is a determination of generic flight schedules for every single run of the algorithm using a random experiment, but getting no distorted result according to the real sample. The objective is to find a possible result, not an optimal one. The algorithm of phase C bases on the urn model, which is commonly used in probability theory.
The algorithm starts with the generation of the aircraft category according to their occurrence in the real schedules. For every category selection (aircraft, distance, passenger) a pool of data, an urn respectively, is generated. The probability of choice $P(A)$ is calculated by the number of results, the event takes place $|A|$ divided by the total number of possible results or sample space $|\Omega|$ (Laplace formula). We assume, that the size of aircraft is linked to particular slots and, thus, the choice of an appropriate distance and passenger category is limited automatically, which are the second and third step within the determination of random flights. The selection of a distance and passenger category takes place randomly from data out of the whole week. The following example and Figure 7 shall illustrate this process.

First layer: An urn contains 47 flights including 3 AC categories. Probabilities are 10/47 (Cat 2), 15/47 (Cat 4) und 22/47 (Cat 5). The choice of AC Cat 5 is most likely, here exemplarily AC Cat 4 is chosen. Second layer: AC Cat 4 flies over a week 300 times DIST Cat 3 and 500 times DIST Cat 4 out of 800 flights in total. Exemplarily DIST Cat 4 is chosen. Layer 3: Over a week a passenger volume is recognized, which fits 250 times to PAX Cat 5, 200 times to PAX Cat 6 and 100 times to PAX Cat 7. Exemplarily PAX Cat 6 is chosen for this flight.

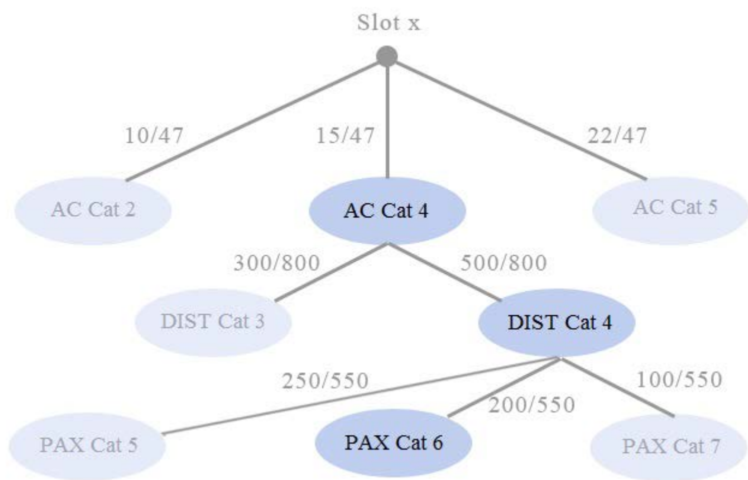

Figure 7: Exemplary decision tree of determining a flight

In parallel, currently a non-generic airline is selected following the same scheme of choosing a distance category. Hence, varying fleets are taken into account due to the adjustment of every aircraft category used by an airline. Finally, a flight consists of a data quadruplet of airline, distance category, aircraft category and passenger category.

After the determination of random flights we have to assign them to the frequencies of phase B in a random manner. As there exist no real segments (we waive the precise O&D information), a segment builds a container, defined by its distance and passenger category combination. The first flight allocated randomly sets the container fundament and it obtains a certain passenger volume interval depending on the passenger category. Consequently this container has to be filled with flights such that the boundaries of the FoAM passenger category are matched. The passenger increase per flight is calculated according to (8) by the product of the SLF, the upper/lower seat capacity bound of the aircraft category and 4.37, which is the average number of a weekday per month (valid for both summer and winter).

$$passenger\ increase_{\ min/max} = SLF *\ Seat\ capacity_{AC\ Cat,\ min/max} * 4.37 \tag{8}$$

> *Example: First flight allocated to a container is AC Cat 4, DIST Cat 5 and PAX Cat 6. Thus, the passenger volume of this container has to be between 9 001 and 20 000 passengers. According to (8) the minimum passenger increase amounts to 296 passengers and the maximum increase to 443 passengers, assuming a SLF of 65 %. The container can comprise a slot maximum of 960 [min per week] divided by the chosen slot interval [min].*

The segment filling process is implemented by using a array (n x m), exemplarily shown in Figure 8. In order to account for a regularity of flights over the week, the algorithm primarily generates flights with an identical aircraft category, using different slots (indicated red at Figure 8) and adds the flights to the days of appearance. Not used aircraft categories are added at the bottom of the matrix (indicated as slot number with *). This attachment of new rows ensures that no flights get lost and, only in a worst case, a multiple occurrence of flights of a segment within one time slot. A segment is considered to be full for the first time, if the passenger volume exceeds the lower category boundary. If no new segment container can be initialized the remaining flights are allocated to the already implemented containers, maintaining the upper boundary of the current passenger category.

Slot\ Week day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
19	4	3	2	4	4	4	4
.	.	.	.	.	.	.	.
41	2	4	3	3	4	2	2
2	3	2	4	3	2	4	2
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
19*							
41*							
2*							
960/slot interval	.	.	.	.	.	.	.

Figure 8: Array of AC categories per flight slot and weekday

3.4 Phase D: Generic flight schedule forecast

Last step (phase D) performs a forecast of the present flight schedule into the future, here 2030. The forecast bases on the flight frequencies per slot and the data pool of flight schedule parameters of phase B, extended by appropriate FoAM scale factors. FoAM provides future values of the flight frequency and the generic flight plan parameters per distance and passenger category combination, as well as the seat load factor per distance category. The flight frequency is adapted according to:

$$flight\ frequency_{new}(slot\ x) = flight\ freuquency_{old}(slot\ x) * \frac{total\ flight\ frequency_{2030}}{total\ flight\ frequency_{2012}} \tag{9}$$

This adjustment automatically results in a capacity overload of the airport infrastructure. The capacity is constrained according to the typical runway layout of the cluster, using default figures for IFR traffic by Horonjeff [12]. If the capacity limit is exceeded, the frequencies are shifted to the next free slot (Figure 9).

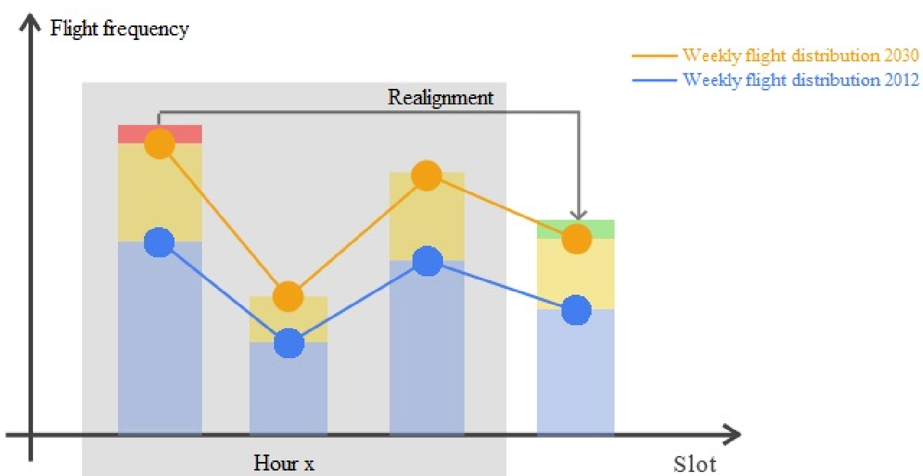

Figure 9: Shift of frequencies exceeding the airport capacity limit

$$SLF_{year,DIST\ CAT} = SLF_{base\ year,DIST\ Cat} + (year - base\ year) * (0.0077 - 0.0001 * (year - base\ year)) \tag{10}$$

Furthermore the SLF evolves according to (10) depending on the distance category, but never exceeding a SLF of 90 %. The SLF increases 0.77 % per year, which equals to the empirical calculated increase of the SLF worldwide within the last decade, and a damping of 0.1 % per anno. After the adjustments, phase C is performed once again resulting in a 2030 flight schedule.

4 RESULTS

The aforementioned methodology is implemented in a JAVA environment, which enables an easy adaption of user-specific requirements. These can be adjusted directly at the graphic user interface of BogS. Despite the cluster, the user can choose between a time period (summer or winter), the type of operation (arrival or departure flight) and the slot duration (10, 15, 20 or 30 min). The functionality of the algorithm was tested by the following scenario, which comprises the International Hub Airports arrival flights in the winter and a slot interval of 15 minutes. No night curfews are taken into account.

Scenario I: *International Hub Airports*
Winter, ARR, slot interval: 15 min, no night curfew

The 2012 list of flights contains a total number of 8 152 flights, whereof 162 are allocated incorrectly (methodological error of 1.99 %). This error results from an insufficient amount of flights within a container, allocatable to the current slot.

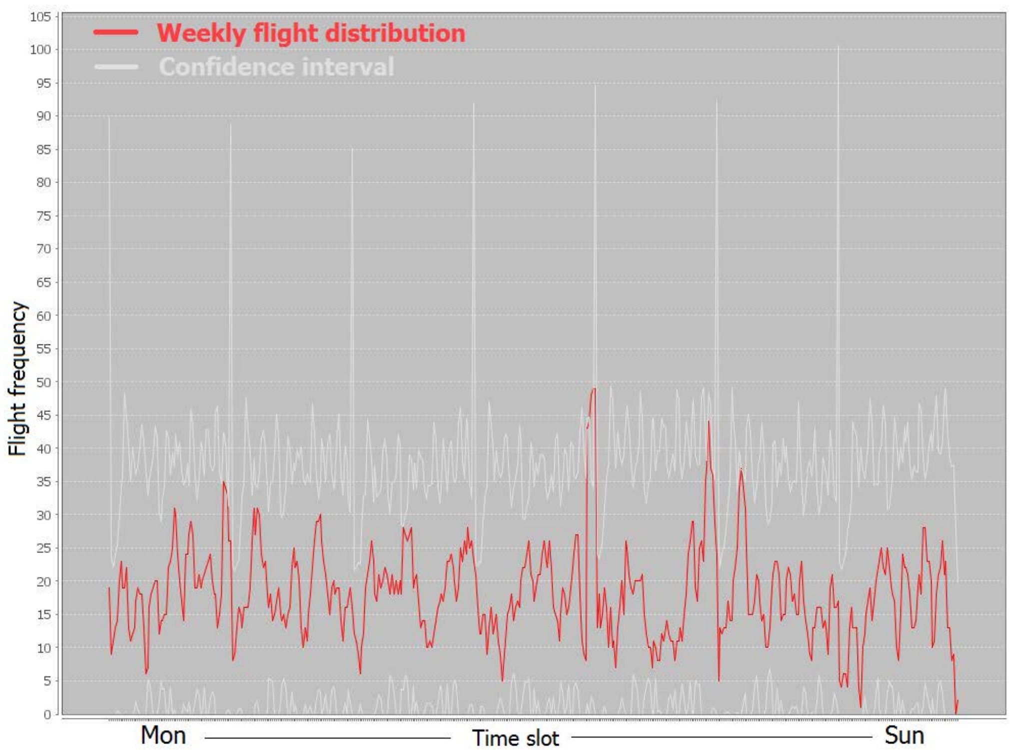

Figure 10: Scenario I – Weekly flight distribution

Figure 10 shows the weekly flight distribution of this scenario. Visible are distinctive peaks in the morning and a daily wave structure, typical for hub airports. In particular, aircraft category five occurs most frequent in the aircraft distribution, followed by category two and eight (see Figure 11) . This represents the outcome of the clustering and also real International Hub Airports very well. Aircraft category five, a 152 – 201 seater as well as smaller regional jets (aircraft category two) act as feeder traffic for long haul flights operated with larger aircraft (category eight). Furthermore aircraft category eight has a higher share than category six and seven (also in comparison with other scenarios not depicted here), which fit to the clustering outcome, that this airports are primarily destinations of the Airbus A380.

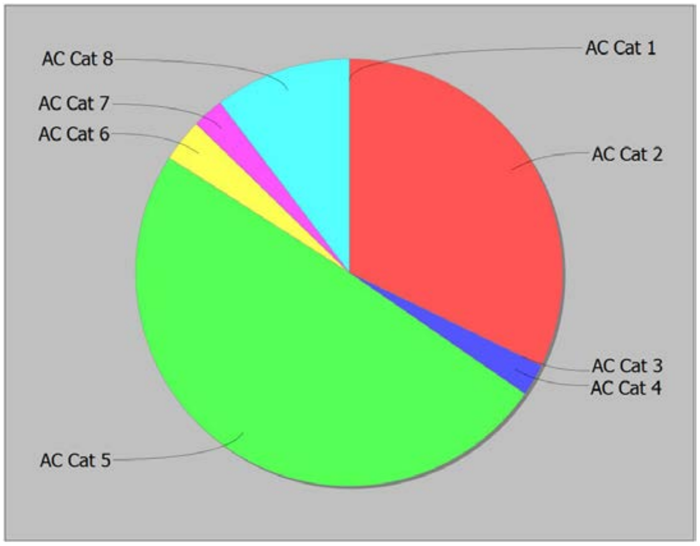

Figure 11: Scenario I - AC Cat distribution

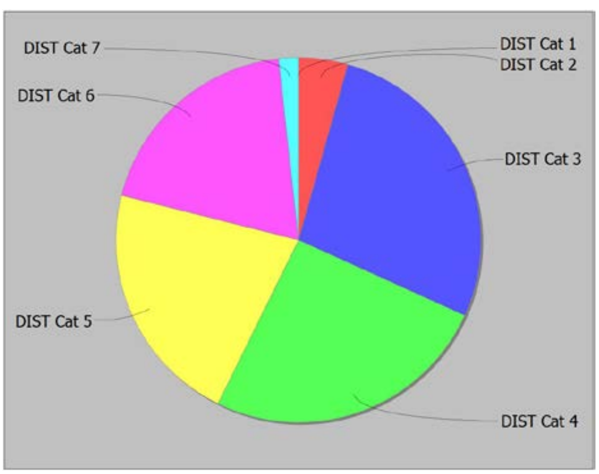

Figure 12: Scenario I - DIST Cat distribution

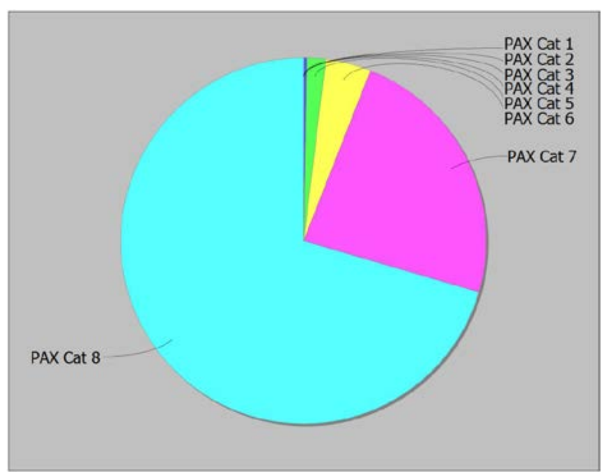

Figure 13: Scenario I - PAX Cat distribution

Beside ultra-long and ultra-short distances the other distance categories are equally spread at International Hub airports, visualized in Figure 12. In contrast to that, passenger category eight dominates the passenger distribution (see Figure 13). This means, more than 38 000 passengers per month fly on segments typical for this airport category. Due to a dominant market position not unusual for International Hub Airports.

Scenario I: Forecast

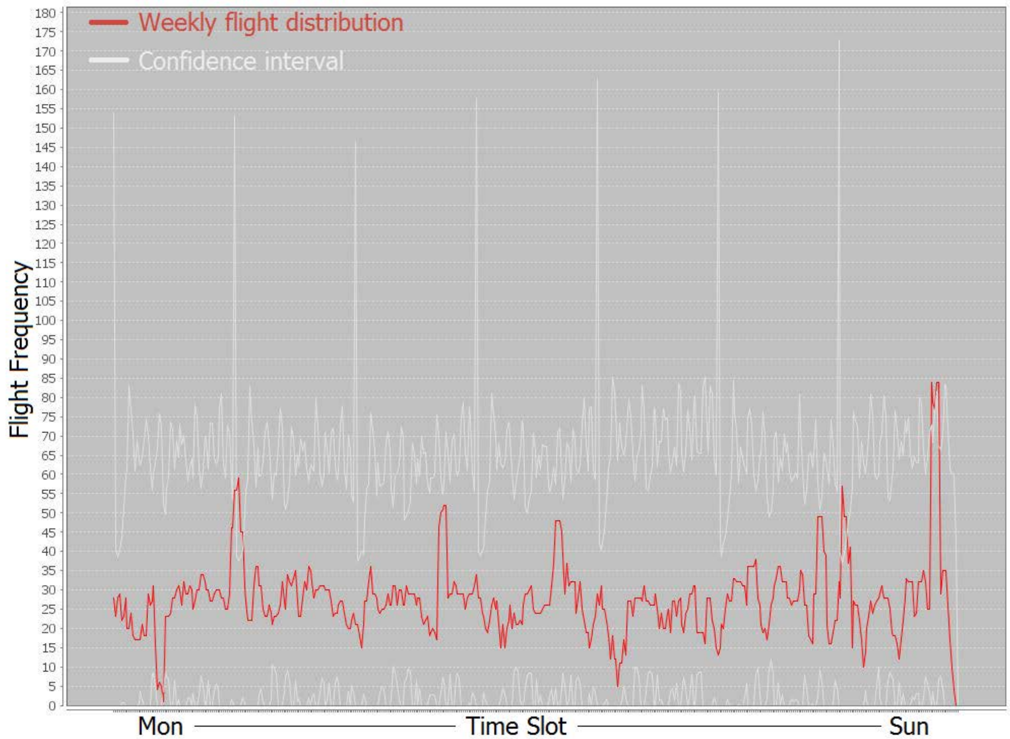

Figure 14: Scenario I – Weekly flight distribution in 2030

Due to a constant growth of the passenger demand, the number of flights is forecasted to 10 425 per week in 2030. Without infrastructural changes, this development naturally leads to capacity shortages. The methodological approach shifts flights exceeding the runway capacity to the next free slot, hence, the peaks level out and an equalization or homogenization of the weekly flight distribution is recognized, visible in Figure 14.

Further scenarios were executed, but cannot depicted here. Result is a slightly shift of relative shares per category due to the nonexistent restrictive allocation of the passenger and distance category. Thus, it can be observed, that the algorithm choses mainly categories with a large relative share within the empirical data, increasing the dominant position of this category. Moreover, a small sample size results in a large methodological error (up to 40 %), emerging at scenarios of the High Frequency Hubs. Due to the small sample size, outliers have a wide influence to the generation of the parameters used for the distribution function. If an outliner defines the frequency data point, only very few data points (or may only this one point) can be provided to describe the combinatorial possibilities of determining a flight (aircraft, distance, passenger category). Hence, the algorithm finds only a few (or may only one flight) that fit to the slot. The validity of these results shall be further verified.

5 CONCLUSION AND FUTURE WORK

The paper in hand introduces a methodology for determining generic flight schedules for airport clusters out of real flown schedules using common probability functions and further statistical methods. It is shown that the algorithm generates appropriate generic flight schedules highly matching real ones for both present and future points in time. These flight schedules include for every single airport slot an aircraft, distance, passenger and airline information only. Our hypotheses, similar airports have quite comparable flight operations, can be confirmed. Some limitations of the algorithm have to be taken into account, if the sample size (e.g. High Frequency Airports) is low. Further investigations and analyses will be done on that issue.

Although the methodology produces valid results, there are some possible future enhancements. In order to increase the accuracy of modeling the weekly flight distribution, further steady probability functions shall be added, e.g. the t-distribution. Particularly for large sample clusters with a inhomogeneous flight operation, a differentiation would be valuable.

Furthermore the implementation of generic airlines, the only non-generic parameter in the method at the moment, is in progress. Similar to the airport clustering generic airline types will be implemented in the model. Finally, the upcoming of new segments shall be taken into account in further methodology evolution steps.

The algorithm as well as the models mentioned in chapter 2 will be used to evaluate intra-airport relationships in a generically modeled air transportation network. Possible, scenario-based evolutions in the long term of a certain network and impacts on airports shall be shown, in a most macroscopic manner. Such scenarios could be the introduction of revolutionary aircraft configurations (e.g. a blended wing body), the airports evolution and function within the air transportation network at a defined point in time, as well as resulting network impacts.

Appendix

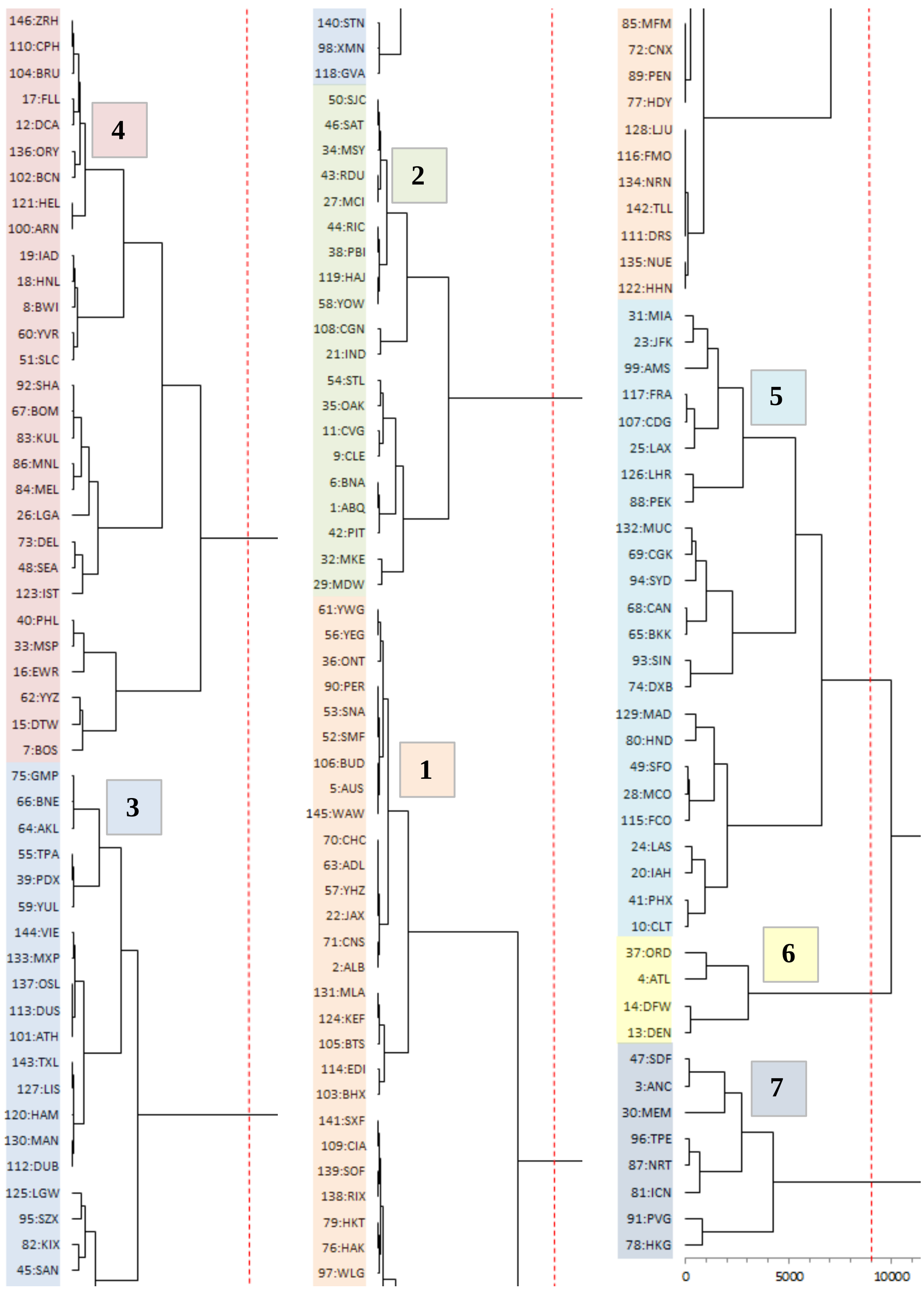

Figure 15: Dendrogram of airport cluster analysis

Literature

1. SCHULZ, A., BAUMANN, S. und WIEDENMANN, S. *Flughafen Management.* München: Oldenbourg, 2010. Lehr- und Handbücher zu Tourismus, Verkehr und Freizeit. ISBN 3486591797.
2. AZZAM, M. *Eine Analyse der Evolution und Effizienz des Luftverkehrsnetzwerks.* Dissertation. Fachbereich Maschinenbau, TU Darmstadt. Darmstadt, 2012.
3. ÖTTL, G. und BÖCK, P. An approach to airport categorization - category definition and quantification for air traffic related simulation purposes [online]. *Journal of Aerospace Operations,* 2011, 1(1-2), 155-170. doi:10.3233/AOP-2011-0009.
4. JACOBS, T.L., GARROW, L.A., LOHATEPANONT, M., KOPPELMAN, F.S., COLDREN, G.M. und PURNOMO, H. Airline Planning and Schedule Development. In: C. BARNHART und B. SMITH, Hg. *Quantitative Problem Solving Methods in the Airline Industry.* Boston, MA: Springer US, 2012, S. 35-99. ISBN 978-1-4614-1607-4.
5. LOHATEPANONT, M. und BARNHART, C. Airline Schedule Planning: Integrated Models and Algorithms for Schedule Design and Fleet Assignment [online]. *Transportation Science,* 2004, 38(1), 19-32. ISSN 0041-1655. doi:10.1287/trsc.1030.0026.
6. LAROSE, D.T. *Discovering Knowledge in Data. An Introduction to Data Mining.* Hoboken, N.J.: Wiley-Interscience, 2005. ISBN 0-471-66657-2.
7. BACHER, J. *Clusteranalyse. Anwendungsorientierte Einführung.* 3. Aufl. München: Oldenbourg, R, 2008. ISBN 978-3-486-58457-8.
8. BIEßLICH, P. und LIEBHARDT, B. *Parametric Reference Airports.* 17th World Conference of the Air-Transport-Research-Society. Bergamo, Italy, July 2013.
9. KÖLKER, K., BIEßLICH, P. und LÜTJENS, K. *FoAM - From Passenger Growth to Aircraft Movements.* 18th World Conference of the Air-Transport-Research-Society. Bordeaux, France, July 2014.
10. SABRE AIRLINE SOLUTIONS. *Airport Data Intelligence (ADI)* [online]. *[online data base].* http://www.sabreairlinesolutions.com/home/software_solutions/airports
11. SACHS, L. und HEDDERICH, J. *Angewandte Statistik. Methodensammlung mit R ; mit ... 180 Tabellen.* 12., vollst. neu bearb. Aufl. Berlin: Springer, 2006. ISBN 3-540-32160-8.
12. HORONJEFF, R. und MCKELVEY, F.X. *Planning & design of airports.* 4th edition. New York: McGraw-Hill, 1994. ISBN 0-07-113352-6.