Assessment of local aircraft crash risk

Application of a cluster analysis as a statistical method for detecting similar airports

Christoph Thiel
Institute of Logistics and Aviation
Technische Universität Dresden
Dresden, Germany
thiel.christoph@googlemail.com

Hartmut Fricke
Institute of Logistics and Aviation
Technische Universität Dresden
Dresden, Germany

The assessment of local aircraft crash risks in the vicinity of airports is of primary importance in numerous safety studies relating to the determination of Third Party Risk due to aircraft accidents. This paper presents an approach of determining local aircraft crash rates by means of a cluster analysis. This statistical method detects similarities between airports in consideration of safety relevant parameters.

Safety, aircraft crash risk, accident ratio, External Risk, similarity analysis, cluster analysis

I. Background

A. Introduction

In the context of planning approval procedures at many major European airports, safety relevant issues today gain increasing importance over e.g. noise or environmental issues. When analyzing the safety in air traffic, one essential model for airport related safety studies is the determination of External or Third Party Risk, that is the risk of death due to an aircraft accident for people who do not participate in the air transport system, generally people living or working in the vicinity of an airport.

Despite the lack of regulations relating to External Risk in most European countries (with the exception of The Netherlands and Great Britain), determining External Risk becomes a central instrument for evaluating risks due to aircraft accidents for people living around airports.

The External Risk model consists of three sub-models: the accident ratio (AR), the accident location (AL) and lastly the accident consequence (AC) sub-model. The scope of this paper is the accident ratio sub-model, for which a statistical method of determining a local accident ratio for a specific airport by means of cluster analysis is presented.

In order to further clarify the topic the following section gives an introduction into the External Risk model itself.

B. The External Risk Model

The External Risk (ER) expresses the statistical potential of a human being receiving fatal injuries as the result of a severe aircraft accident or its potential consequences in form of secondary effects on the ground (damages to an industrial plant, for example). This potential is important around airports, because the operational accident risk for aircraft is highest during takeoff and landing and so ER calculations generally refer to an airport. The term *external* refers to the fact that the risk is calculated for those people who are not formally participating in the air transport system during a given time period. Typically, this is the population residing in the area around that airport, or people who work there (employees). More precisely, these are people located at least temporarily within a selected investigation area around the airport.

The External Risk effectively consists of two types of measurable risk figures, the local or Individual Risk and the group or Societal Risk:

- The Individual Risk is the probability of an imaginary person being killed in a particular location within the investigation area as a result of an aircraft accident during a period of one year. It is therefore not important to know whether a person is actually present or not.
- When calculating risk, it might additionally be important to consider the population that is actually present and the distribution of this population around the airport rather than an imaginary person. Calculations are made in order to determine the size of the risk of one or more simultaneous casualties within this population. This probability of a disaster of a certain size is known as the Societal Risk.

The External Risk model furthermore consists of three sub-models:

- an accident probability model, providing a local Accident Ratio (AR) for fatal aircraft accidents according to the definitions given in ICAO Annex 13 [1],
- an accident location model, providing an Accident Location (AL) distribution probability function referred to the Air Traffic Route System and linked to a runway and/or threshold, and
- an accident consequence model, providing the local Accident Consequence (AC) Area with regard to local terrain and industrial site details.

ISBN: 978-0-615-20720-9

With these three sub-models and all required airport related parameters (e.g. movements per departure/ arrival route, traffic mix, industrial sites) Individual Risk will be calculated and shown in individual risk contours around the airport area. Following Figure 1 demonstrate such individual risk contours as an example at an imaginary airport:

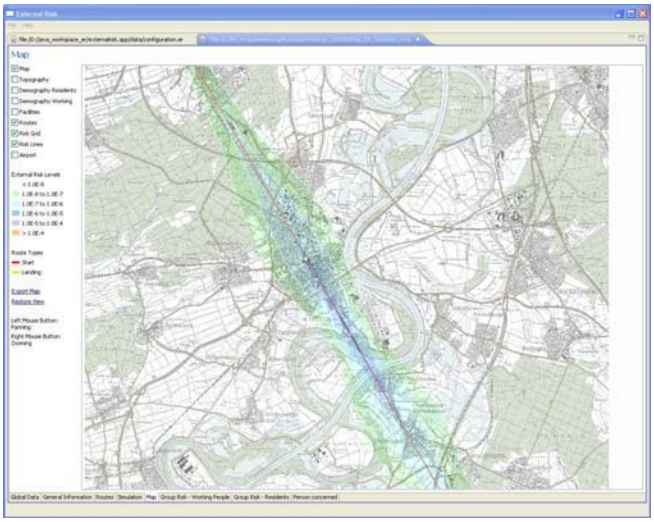

Figure 1. Individual Risk

With information about population density (residents/ employees) within the investigation area (usually an area of 40 km x 40 km centred in the airport reference point) societal risk will be calculated as so called F/n- curves.

As an essential element of External Risk calculation, the accident ratio sub-model and current approaches for calculation are described more precisely in the following chapter II. Furthermore in chapter III is presented an alternative approach of assessing a local accident ratio at a certain airport by means of statistical analysis.

II. Current Methods Of Local AR Determination

A. *Definition of local Accident Ratio*

Generally, aircraft crash risk may be assessed by using theoretical models which would use the measured probabilities of all possible causal factors to predict the probability of a crash. However, such a theoretical approach is very problematic, since aircraft accidents are usually the result of a combination of many separate causal factors with unknown probabilities and complex interrelationships. An alternative method is to use empirical data on accidents and on aircraft movements to calculate aircraft crash risks. This non-causal approach assumes that the historical accident ratio will continue into the future, which, if there are future safety improvements, as they are currently aimed by the SESAR Consortium [2], may lead to an overestimation of accident ratio in future years.

The accident ratio itself is defined as the probability of an aircraft accident per movement and is calculated by dividing the number of accidents in a certain time period by the number of relevant movements within this period:

$$AR = \frac{Number\ Of\ \ Accidents}{Number\ Of\ \ Movements} \tag{1}$$

Because crash probabilities differ considerably between airports, a selection of data from the available worldwide accident data is required to make the accident ratios suitable for a specific airport.

Many sources provide global AR values, calculated by a simple division of worldwide accidents by worldwide movement data. For example the “Statistical Summary of Commercial Jet Airplane Accidents” annually published by Boeing [3] provides a good overview about the global AR per year in worldwide commercial operations, as well as a trend in global AR.

But this global AR may not be valid for a specific airport; in fact, a local AR differs from the global AR, depending on the airport, because there are numerous factors which influence the safety within a certain investigation area (certain airport) and lead to a local accident crash probability.

Calculating a local accident crash probability suitable for a specific airport may be done in many different ways. The current essential (more or less standardized) methods, which are used by several national organizations to determine local aircraft crash risks (e.g. for calculating External Risk), are presented in the following sections.

B. *NATS -Method*

The National Air Traffic Services (NATS) of the United Kingdom uses an AR Model specific for generic aircraft groups within their model of Public Safety Zone (PSZ) calculation [4]. On an empiric basis, crash rates per classified aircraft group will be calculated.

The NATS defines a classification of aircrafts according to their type of engine (jets, turboprops and piston-engine), the region of their manufacturer (eastern, western) and their date of first delivery. The essential breakdown of aircraft classification is therefore:

- Class I western airliner jets (e.g. Boeing B707, Comet)
- Class II – IV western airliner jets (Boeing B727, B747, Airbus A310, etc.)
- Other jets (Eastern Jets / Executive Jets)
- Turboprops (western airliner Turboprops before 1970/ after 1970, unclassified Turboprops)
- Piston-engine

By means of historical aircraft accident data and statistics, NATS calculates an AR for every group of aircraft and with knowledge about the share of every aircraft group in the traffic mix at a certain airport, the overall AR for this certain airport can be calculated.

For most of the major, worldwide airports the current traffic-rate of modern western airliner jets (Class II – IV) is more than 95% and the share of e.g. eastern jets or older turboprops is marginal. So, according to the NATS-method this

ISBN: 978-0-615-20720-9

approach would calculate nearly the same overall AR for every major airport.

This approach is therefore not very useful for calculating an AR suitable for a specific airport, because a causal differentiation primarily for major airports is not given. Furthermore a differentiated analysis of e.g. the airports Air Traffic Control (ATC) infrastructure or operational performance which may influence a local AR is not taken into consideration.

C. DOE- Standard

The U.S. Department of Energy (DOE) Standard DOE-STD-3014-2006 [5] describes a method of risk analysis for hazardous facilities due to aircraft crashes. The chosen approach is very similar to the method of External Risk calculation, as here also the probability of an aircraft crash independent from the crash location is one part of the risk calculation process.

For near-airport facilities, the aircraft crash rate is empirically calculated for accidents during take off or landing at a specific airport (close to a hazardous facility) differentiated by type of traffic (general aviation, commercial aviation and military). Furthermore, there are some defined sub-categories: for general aviation by type and number of engines and for commercial operations the two sub-categories: air carrier and air taxi.

This classification by type of traffic implies the same problem as the classification by aircraft type as favoured by NATS: a differentiation for many major airports is not given, because one type of traffic is omnipresent (here commercial air carrier) and therefore this method would also calculate nearly the same overall AR for most major airports.

D. NLR- Method

In 1993 the National Aerospace Laboratory of the Netherlands (NLR) published the first documentation about External Risk Calculation [6]. This document gives, amongst others, a detailed description how to determine a local accident ratio. Despite of numerous model updates [7] this accident ratio model remains more or less unchanged until today.

The NLR does not evaluate local AR by means of classification by aircraft or type of traffic only; in fact it uses a method of detecting a certain set of airports similar to the airport under investigation by means of expert justice. The similarity analysis may differ depending on the airport under investigation, but it takes following essential criteria into consideration:

- ATC Infrastructure (precision/ non-precision approaches, terminal approach radar, etc.)
- Size of airport (e.g. number of runways)
- Operational performance (annual movements)
- Local geographical peculiarities (e.g. mountainous area)

So, the NLR- approach does not include the disadvantages of the NATS or DOE method, as it uses more than one parameter for the AR assessment. However this expert based analysis is a very subjective and irreproducible method, as there are no clearly defined selection criteria (this differs depending on the airport) and the selection process itself is not traceable.

E. Conclusion

Each of the presented methods for assessing a local aircraft crash risk includes certain disadvantages: the NATS and DOE methods do not apply a real similarity analysis, as they only take one parameter into consideration and the NLR method by means of expert justice is irreproducible. To avoid these disadvantages, an alternative method of detecting similar airports is presented here. This method takes into consideration the operational performance of the airports as well as reproducibility, since it is calculated by means of statistical analysis.

III. Similarity Between Airports, A Statistical Approach

A. Introduction

1) AR determination process overview

A selection of a certain number of airports empirically similar to the airport under investigation and the total count of accidents at these similar airports has to be quantified to determine a local accident ratio.

So, the first step of determining a local AR is the selection of a certain number of similar airports. Here, similarity is understood as similar operational performance of the airports in terms of yearly passenger traffic, quantity of handled cargo and air traffic movements. By means of correlation analyses, it could be shown, that the chosen parameters (flight operations, cargo, passengers) provide the highest correlation between traffic load and traffic safety: a high correlation between the available Air Traffic Control Infrastructure, type of traffic at the airport (IFR/VFR) and the size of airport, for example, could be identified.

Once a certain set of similar airports is defined, all relevant accidents at these airports have to be investigated. The local AR for the airport than is calculated as the division of the sum of all accidents by sum of all movements. The whole process of determining a local AR at a certain airport is presented in following figure 2:

ISBN: 978-0-615-20720-9

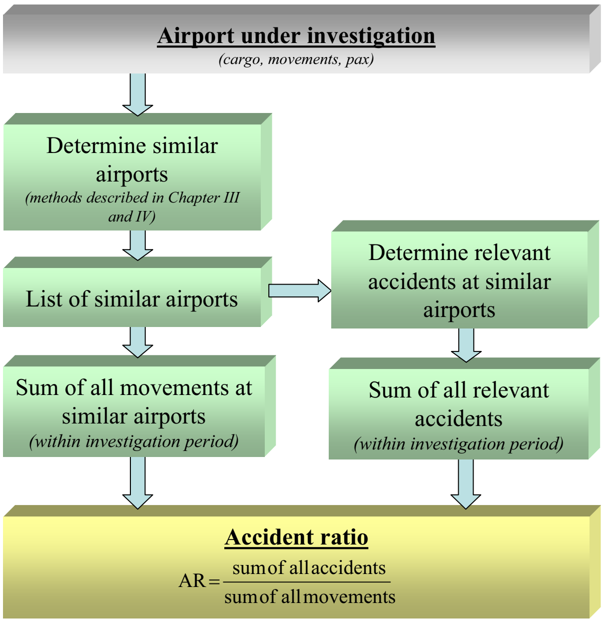

Figure 2. AR determination process

In order to estimate crash rates from historical data, it is necessary to have complete data on airport related crashes and the corresponding numbers of movements. The following sections give an overview how to investigate this information.

2) Investigation of traffic data

The first step to determine a local AR, suitable for the airport under investigation, is the selection of a certain number of similar airports. As described above, operational performance of the airports in terms of yearly passenger traffic, quantity of handled cargo and air traffic movements gives a good indication for similarity analysis. Therefore these three parameters for a huge set of airports must be investigated.

The database of the Airports Council International (ACI) [8] is a very comprehensive source of worldwide airport operations data. It provides operational data for more than 1500 airports worldwide from 1991 until today and is updated monthly. As mentioned above, there are three variables:

- Number of air traffic movements per year
- Number of passengers per year and
- Amount of handled cargo per year.

These three variables are used to filter similar airports based on the mean value per variable from 1991 until today (17 years) for each airport. This period was selected to provide an empirically stable reference data set while taking into consideration the poor data quality of the 80ies decade. Furthermore for validation of the analysis results, the data used should be applicable to current aviation which implies that only recent data should be taken into account.

The similarity selection process itself will be performed by means of statistical methods and is described later in section B of this chapter.

3) Investigation of relevant accidents

In order to determine relevant accidents at the identified similar airports there are various international accident databases available. A very comprehensive and easy accessible database for worldwide aircraft accidents is the Aviation Safety Net (ASN) [9]. It covers nearly every aircraft accident having occurred during the last 50 years and provides a huge amount of additional information (e.g. investigation agency, link to full accident report). Another comprehensive and free online accessible database is the NTSB- Database [10] which today includes more than 65,000 accidents having occurred since 1982.

Accidents taken into consideration should be at least consistent with following selection criteria:

- Occurrence with at least one fatal injured person on ground or on board the accident aircraft (according to definition of fatal accident in ICAO Annex 13),
- Occurrence during take off or landing phase and within a certain area around the airport (as mentioned above, usually 40 km * 40 km),
- Occurrence within the investigation period (as mentioned above, usually about 15 years)
- Occurrences not involving sabotage, hijacking or military action

Depending on the airport under investigation, additional selection criteria may be defined, e.g. no occurrences with aircraft below 5.7 to MTOM, if the airport under investigation does not include this kind of traffic or its traffic count is marginal.

B. *Clusteranalysis in AR- Determination*

1) Introduction

Cluster analysis is a multivariate procedure for detecting natural groupings (clusters) in data. Cluster analysis classification is based upon the placing of objects into more or less homogeneous groups, in a manner such that the relationship between groups is revealed. This means, the formed clusters should be internally homogenous (members are similar to each other) and externally heterogeneous (members are different to members of other clusters). Figure 3 below shows such a clustered dataset by means of a schematic diagram (note, that the given figure does not represent the true spread of airports operational performance and grouping, as it is only used for purpose of clarification):

ISBN: 978-0-615-20720-9

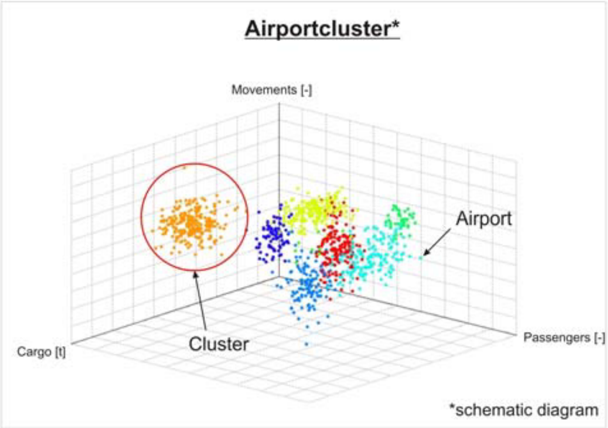

Figure 3. schematic cluster diagram

The selection of airports similar to a certain airport is derived here by a statistical similarity analysis – the cluster analysis. There are some more statistical methods for such similarity analyses, but irrespective of the method every similarity analysis underlies the same diametrical effect: The higher the number of similar elements (airports) is, the lower is the similarity to the reference element (airport under investigation). This means, the number of similar airports has to be high enough to achieve a statistically stable data basis for AR calculation, as well as the similarity to the airport under investigation has to be given for every considered airport. Otherwise, a set of e.g. 5 to 10 similar airports may induct a high similarity, but it's not enough to ensure a statistically stable data base. Therefore, finding a balanced proportion between the number of airports and similarity measure is of primary importance.

2) Standardization of variables

Standardization of variables has to be executed to enable the comparison of variables to minimize the bias in weighting which may result from differing measurement scales and ranges.

The variables used for clustering - the yearly passenger traffic, quantity of handled cargo and air traffic movements, may show very high differences , e.g. depending on the airport the number of yearly passengers is about 100-times higher than the number of yearly movements (generally speaking, average number of passengers per movement). For comparative purposes and in order not to overestimate the passenger variable, every variable has to be normalized before starting the cluster algorithm. The best method of normalization for the chosen approach is the so called *Z- Transformation*, which results in a mean value of zero and a variance of one for each variable:

$$z_i = \frac{x_i - x_m}{S_i} \tag{2}$$

With x_m as mean value and S_i as standard deviation:

$$x_m = \frac{1}{N_x} \sum_{i=1}^{N_x} x_i \tag{3}$$

$$S_i = \sqrt{\frac{1}{N_x - 1} \sum_{i=1}^{N_x} (x_i - x_m)^2} \tag{4}$$

3) Clustering with Ward's linkage

Generally, hierarchical cluster analyses are comprised of agglomerative and divisive methods. The divisive methods start with all of the observations in one cluster and then proceed to split (partition) them into smaller clusters. The agglomerative methods initially consider each observation as a separate cluster and then proceeds to combine them until all observations belong to one cluster. The here applied cluster algorithm is of agglomerative nature.

The two key steps within cluster analysis are in the first place the measurement of distances between objects and secondly to group the objects based upon the resultant distances (linkages). The distances are a measure of similarity between objects and may be measured in a variety of ways, e.g. as Euclidean distance. The criteria used to link (group) the variables may also be undertaken in a variety of manners, as a result significant variation in results may be seen. Linkages are based upon how the association between groups is measured. Four of the better-known algorithms for hierarchical clustering are average linkage, complete linkage, single linkage and Ward's linkage. Ward's is a popular default linkage which produces compact groups of well distributed size. The final algorithm according to Ward's procedure is applied here: it uses an analysis of variance approach to evaluate the distances between clusters. This method is regarded as very efficient for the existing approach, as it forms clusters of approximately the same size.

Ward's linkage is a specific weighting method applied in the hierarchical cluster analysis. The linkage function specifying the distance between two clusters is computed as the increase in the error sum of squares (ESS) after fusing two clusters into a single one. Ward's Method seeks to choose the successive clustering steps so as to minimize the increase in ESS at each step.

The ESS of a set x of N_x values is the sum of squares of the deviation from the mean value x_m. For a set x the ESS is therefore described by the following expression:

$$ESS(x) = \sum_{i=1}^{N_x} (x_i - x_m)^2 \tag{5}$$

The distance or linkage function for the distance between the two Clusters X and Y is described as:

$$D(X,Y) = ESS(XY) - [ESS(X) + ESS(Y)] \tag{6}$$

with XY being the combined cluster resulting from fusion clusters X and Y. So, the distance function describes the increase in ESS by fusing Cluster X and Cluster Y into one combined cluster XY. This distance function has to be calculated for every combination of clusters within the dataset and put into the so called distance matrix. The minimum increase in ESS between two certain clusters marks the next

 ISBN: 978-0-615-20720-9

agglomeration step. The two clusters with minimum increase in ESS will be combined into one cluster.

The essential process of agglomerative clustering with Ward's method is presented in following figure 4:

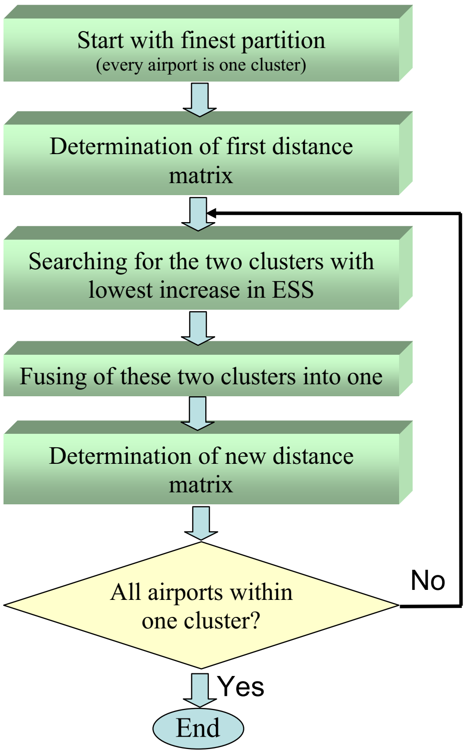

Figure 4. Process of agglomerative clustering

The results of the applied cluster analysis may be presented as a so called dendogram. This presents the fusion of clusters per cluster step as shown in the following Figure 5, at an imaginary example:

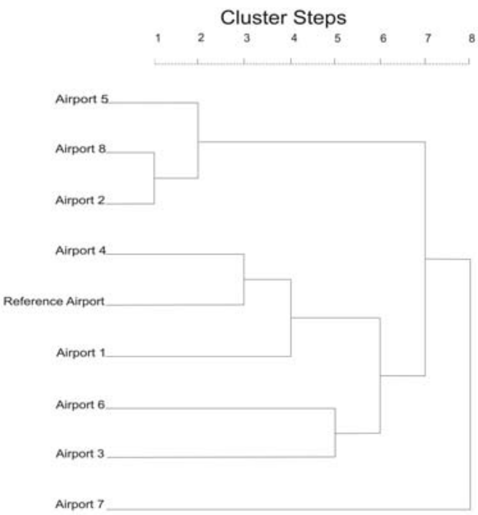

Figure 5. Dendogram – an imaginary example

As seen, within the first step of cluster analysis, *airport 8* and *airport 2* will be combined to one single cluster and step three indicates the first clustering of the *reference airport* (combined with *airport 4*). Within step 6 the master cluster contains 5 airports and within step 8 all airports belonging to one cluster, which marks the end of the cluster agglomeration.

This imaginary example also shows the difficulty of setting the best point of interrupting the cluster algorithm. Theoretically every cluster step reaching the minimum count of similar airports may be chosen as final cluster step.

4) *Interrupting the Analysis*

Once the similarity threshold through a minimum reachable data variance is reached, that cluster will be set to the master cluster, which contains the reference airport. All other airports belonging to the master cluster form the set of data that will be used in calculating the AR.

Detecting this similarity threshold (finding the best point of interrupting the cluster agglomeration) may be unfrugal. As described earlier there must be a balanced proportion between the number of similarity airports and similarity measure. The higher the number of included airports in the master cluster is, the lower is the similarity to the reference airport. A statistical method for finding this point may be the so called *F-test,* which provides an indication about the homogeneity of a certain group of airports. Therefore for each step of the cluster agglomeration the F-value for each variable of the reference cluster has to be calculated as the quotient of variance of reference cluster and variance of the entire data set:

$$F = \frac{V(J,G)}{V(J)} \tag{7}$$

With:

V(J,G) = variance of variable J in group G

V(J) = variance of variable J in the entirety

With variance defined as:

$$V(x) = \frac{1}{N_x - 1} \sum_{i=1}^{N_x} (x_i - x_m)^2 \tag{8}$$

As a result, it has to be mentioned that the smaller the F-value is, the higher is the homogeneity of the reference cluster. A cluster has to be considered as homogeneous if the F-value for each variable is not bigger than one.

Generally, every cluster step can be chosen as final master cluster, whose F-value is lower than one and whose reaches a minimum sample size (from experience more than at least 20 airports). Practically, this cluster step should be used which marks the similarity threshold: where the F-value for each variable is still slightly below one and within the next cluster step one of the variables would exceed an F-value of one.

C. *Summary*

Once this similarity threshold is defined, all airports belonging to the master cluster within the derived cluster step can be set as similar airports. Afterwards all relevant accidents at these airports in the given time period have to be investigated as well as the total count of movements at all

ISBN: 978-0-615-20720-9

airports in the given time period to determine a local accident ratio, suitable for the airport under investigation.

Note, that the given mathematical descriptions in this paper do not represent the full Ward's cluster algorithm, as the determination process of such multivariate clustering process is much more complex. The full mathematical process is documented in a comprehensible way in various statistical volumes (e.g. [11], [12]). Presenting the full cluster algorithm, would go beyond the scope of this paper, nonetheless it could be shown that the application of a cluster analysis is a good instrument for detecting similar airports.

The following chapter IV gives an overview of the results of a hierarchical cluster analysis conducted according to the method described above, in order to identify prevailing trends.

IV. RESULTS

Based on the described method of airport clustering a cluster analysis was performed by using traffic data from the ACI- airports data base [8]. From this database airports with more than 30,000 movements per year (mean value from 1991 to 2006) and with traffic data from at least six out of 17 years per variable were selected. Within these selection criteria a dataset of 398 worldwide airports was used as input value for the cluster analysis.

The cluster analysis was performed without a reference airport, as the results should not be used for determination of a single AR for a specific airport, but should be used for identifying prevailing trends and correlations between traffic load and AR.

As a result, several clusters that reach a minimum sample size of at least 20 airports and with an F- value lower than one could be found. For each of the 398 airports within these clusters all relevant accidents were examined by using various international accident databases (e.g. [9], [10]). Following Table I present four clusters as an example:

TABLE I. SAMPLE OF AIRPORT CLUSTER

Sample of airport cluster				
Cluster- No.	***1***	***7***	***9***	***11***
No of airports	54	111	35	21
movements[a]	132,252	102,266	215,236	430,603
cargo[a]	156,436	27,016	331,266	338,934
passengers[a]	8,684,462	2,600,680	17,040,752	24,767,204
No of accidents (last 17 years)	23	95	8	9
AR [per movement]	**1.89E-07**	**4.92E-07**	**6.25E-08**	**5.85E-08**

a. mean value per airport and per year (last 17 years)

Based on these "characteristic" clusters, a correlation between traffic load and accident ratio was to be verified. Therefore, a large amount of random airports with typical passenger/ movement ratios and cargo/ movements ratios were produced and based on the determined clusters a local AR was calculated for each of these airports.

Finally, it could be found that the higher the number of movements (respectively passengers and cargo volume) is, the lower is the number of accidents per movement. Generally speaking, a decreasing AR by increasing traffic volume could be detected.

Following Figure 6 gives an example of this general trend, for imaginary airports with a passenger/ movement ratio of 60 (60 passengers per movement) and a cargo/ movement ratio of 1.25 (1.25 t cargo per movement):

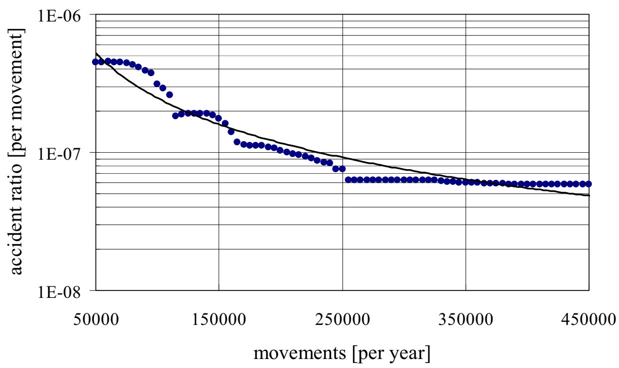

Figure 6. Decreasing accident ratio with increasing traffic volume

As seen in Figure 6 the AR values ranges between $5*10^{-7}$ per movement and $5*10^{-8}$ per movement, with a significant decreasing trend.

This decreasing trend may be due to the increasing navigation infrastructure (ILS Equipment, RNP-RNAV procedures, etc.) and increasing professionalism of all stakeholders with increasing traffic volume. On the other hand the results lead to the assumption that the complexity of the airport layout (e.g. numbers of runways) does not have a negative influence on the local accident ratio. However, more research is needed to identify all underlying causes for this decreasing trend.

V. CONCLUSION

The aim of the applied hierarchical cluster analysis is to achieve a reproducible, statistically based algorithm of detecting airports which are empirically similar to a certain airport, in order to assess a local accident ratio for this airport.

The cluster analysis is a well known statistical grouping method and applied for the given problem it gives a very good indication which set of airports can be assumed as similar to a specific airport under investigation. In terms of a legalization of External Risk calculations, a fully expert based analysis of similar airports (as favoured by the NLR), which is irreproducible and subjective method, may not lead to a legalization process, as it is not a specific and standardized calculation method. The applied cluster analysis avoids this disadvantage by means of a statistical and reproducible algorithm.

Finally a cluster analysis was performed, that uses traffic and accident data of nearly 400 worldwide airports. It could be

 ISBN: 978-0-615-20720-9

found, that the higher the number of traffic volume at a specific airport the lower is the number of accidents per movement.

ACKNOWLEDGMENT

The author thanks Mr. Norbert Gronak and Mr. Daniel Fiedler from the Gesellschaft für Luftverkehrsforschung (Organization of Air Traffic Research) for supporting him with their expertise in External Risk calculations, and for granting access to their traffic and accident databases. A special thanks goes to Ms. Angela Mantay for her enormous support with the set-up of the cluster algorithm and with the time consuming accident researches.

REFERENCES

[1] ICAO, "Aircraft accident and incident investigation (Annex 13)", 9th Edition, July 2001

[2] SESAR Consortium, "D2 - The performance target", SESAR Definition Phase – Delivarable 2, 22.12.2006

[3] Boeing Commercial Airplane Group, "Statistical summary of commercial jet airplane accident, worldwide operations 1959 – 2006", Seattle (USA), Juni 2007

[4] A. W. Evans, P. B. Foot, S. M. Mason, I. G. Parker and K. Slater, "Third party risk near airports and public safety zone policy", National Air Traffic Services Ltd (NATS), R&D Report 9636, London, June 1997

[5] U.S. Department of Energy (DOE), "DOE Standard, Accident analysis for aircraft crash into hazardous facilities", DOE-STD-3014-2006, Washington DC, May 2006

[6] M. A. Piers, M. P. Loog et al, "The development of a method for the analysis of societal and individual risk due to aircraft accidents in the vicinity of airports" , CR 93372 L, National Aerosapce Laboratory (NLR), Netherlands, November 1993

[7] A. J. Pikaar, M. A. Piers and B. Ale, "External risk around airports. A model update", NLR-TP-2000-400, presented at the 5th international conference on probabilistic safety assessment and management, Osaka, Japan, 2000

[8] Airports Council International (ACI), Top 500 - Annual Airport Traffic Statistics, Internet: www.airports.org

[9] Aviation Safety Network (ASN), Aircraft accident database, on permanent update, Internet: http://www.aviation-safety.net

[10] National Transportation Safety Board (NTSB), US Aircraft accident database, on permanent update, Internet: http://www.ntsb.gov/aviation

[11] W. Härdle and L. Simar, "Applied Multivariate Statistical Analysis", Springer Verlag, New York, 2003

[12] C. Romesburg, "Cluster analysis for researchers", LuLu Press, North Carolina, 2004

ISBN: 978-0-615-20720-9