Identification of Safety Critical Scenarios for Airlines using Machine Learning in Filter Trees

Lukas Höhndorf[a*] and Florian Holzapfel[a]
[a]Institute of Flight System Dynamics, Technical University of Munich, Garching, Germany

Abstract: During flight, civil aircraft record data that are analyzed within Flight Data Monitoring (FDM) activities of an airline. Based on the characteristics of the operation and network, the flights can be filtered according to aircraft types, departure and arrival airports, and runways. The considered criteria that describe a set of flights are called a filter. When several filters are considered, they often can be arranged according to their hierarchical structure in a filter tree. Subsequently, every filter in this tree can be assigned certain values that describe the performance, e.g. in terms of safety, of the underlying flights. For example, the mean and the standard deviation of the landing masses can be assigned to any available flights arriving at Munich airport on runway 08L. Finally, outstanding filters can be identified using machine learning algorithms for a (potentially very big) filter tree.

Keywords: Flight Data Monitoring (FDM), machine learning, filter trees, safety critical scenarios

1. INTRODUCTION

Civil aircraft record various parameters describing the flight state, aircraft characteristics, engine properties, and many further details. Besides the Flight Data Recorder (FDR) that is widely known as "black box" and used for aircraft accident investigations, the so-called Quick Access Recorder (QAR) collects data for routine analyses. This data is stored and analyzed as part of the airline Flight Data Monitoring (FDM) program, which is occasionally also referred as Flight Operational Quality Assurance (FOQA).

The characteristics of an airline operation in terms of route network, airports, runways, aircraft types and more can generate many different scenarios. Using the FDM data, in which the necessary parameters for this filtering can also be found, the flights of a specific scenario can be easily filtered from the entire flight list. One selected scenario, for example all the flights with a specific aircraft type to a specific airport and runway, is called a filter. When several filters are considered simultaneously, they often can be organized in a filter tree based on their hierarchical structure. For example, a certain filter can describe all flights to Munich airport EDDM. On the next filter level, the flights could be further categorized by the arrival runway in EDDM, i.e. 08L, 08R, 26L, or 26R. These filter trees are useful for the application of machine learning algorithms such as FLAME, which is applied within this paper to detect outliers.

Specific parameters of the FDM data describe the operation and performance of the flight, such as speed, altitude, or mass. To represent the safety performance of the flight, safety critical measures such as the distance from the runway threshold to the touchdown point for deep landing analysis or the vertical load at touchdown for hard landing analysis can be derived for any individual flight. To examine the safety performance of various flights, e.g. the ones given by a specific filter in a filter tree, statistical properties of the safety critical measures as well as for other variables can be calculated for the filtered flights. For deep landing analysis, the mean and standard deviation of the distances from the runway threshold

[*] Corresponding author: lukas.hoehndorf@tum.de

to the touchdown points can be derived. Furthermore, they can be combined in a vector with factors that contribute to this incident, such as target speed deviation during approach or flight path deviations. Doing this for any filter of a filter tree results in a matrix that is describing the safety performance of a complex flight operation, see Table 1.

In this setting, the goal of this paper is to apply outlier detection algorithms that are common in machine learning to identify scenarios unusual from a safety performance perspective. These outstanding scenarios might then be relevant for airline safety officers and can be further discussed in the safety management of an airline.

In chapter 2 the available FDM data is briefly described. The notion of filters and filter trees are presented in chapter 3. In chapter 4, an overview of machine learning and the FLAME algorithm that is used for outlier detection within this paper is given. Subsequently, chapter 5 describes how the information of the FDM data is attached to the considered filter tree. Finally, all concepts are combined and outstanding scenarios are identified in chapter 6. Chapter 7 concludes the paper.

2. RECORDED FLIGHT DATA CHARACTERISTICS

Civil aircraft record data throughout the flight. The frequency of the data recording depends on the parameter, the aircraft type and furthermore, the airline can adjust the recording characteristics based on the airline specific requirements of the available parameters. In Figure 1, an exemplary recording of the barometric altitude is given. At this long haul flight, the characteristic step climbs of the aircraft can be identified.

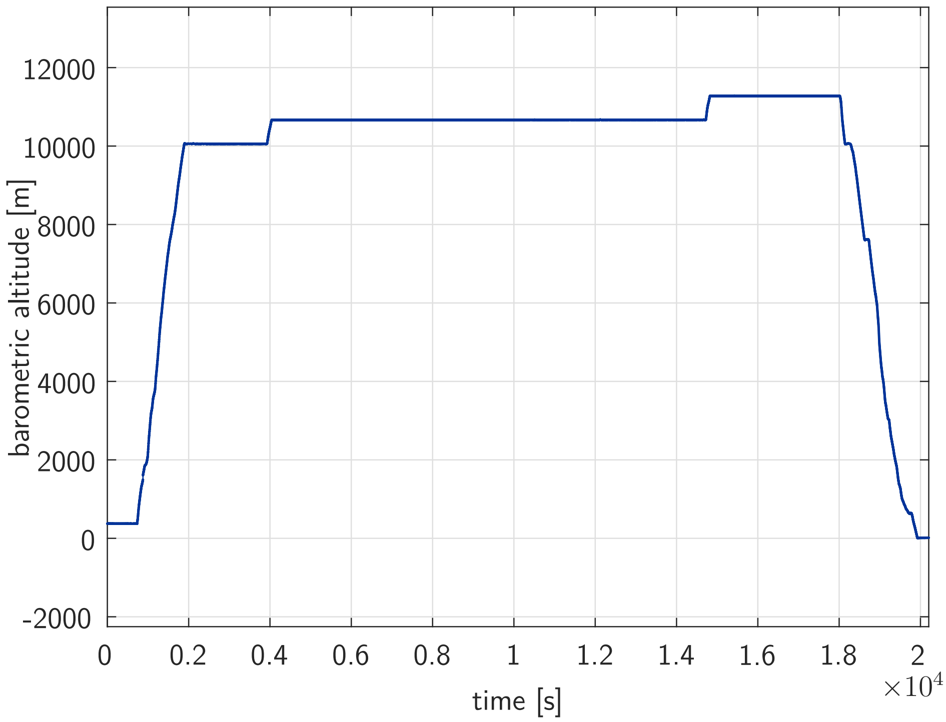


Figure 1: Time series recording of barometric altitude

Based on the recorded time series flight data, specific time points and so called measurements can be derived. One very important time point is the touchdown of the aircraft [1]. Subsequently, so called measurements such as ground speed at touchdown can be calculated based on the time points and the recorded time series. In addition, physical motivated smoothing can assist to reduce the errors and uncertainties in the recorded flight data [2, 3, 4] and eventually also in the calculated time points and measurements.

For aircraft exceeding a certain Maximum Certified Take Off Mass (MCTOM), this recorded data needs to be analysed. "*A flight data monitoring programme for those aeroplanes in excess of 27 000 kg MCTOM. Flight Data Monitoring (FDM) is the pro-active use of digital flight data from routine operations to improve aviation safety.*" [5]

Several commercial FDM software packages are available to assist the airline departments in storing, handling, and analyzing the massive amount of data.

3. FILTER AND FILTER TREES

The network of an airline can get highly complex. In Figure 2, an overview of the global route network is illustrated and gives an impression of its complexity[1].

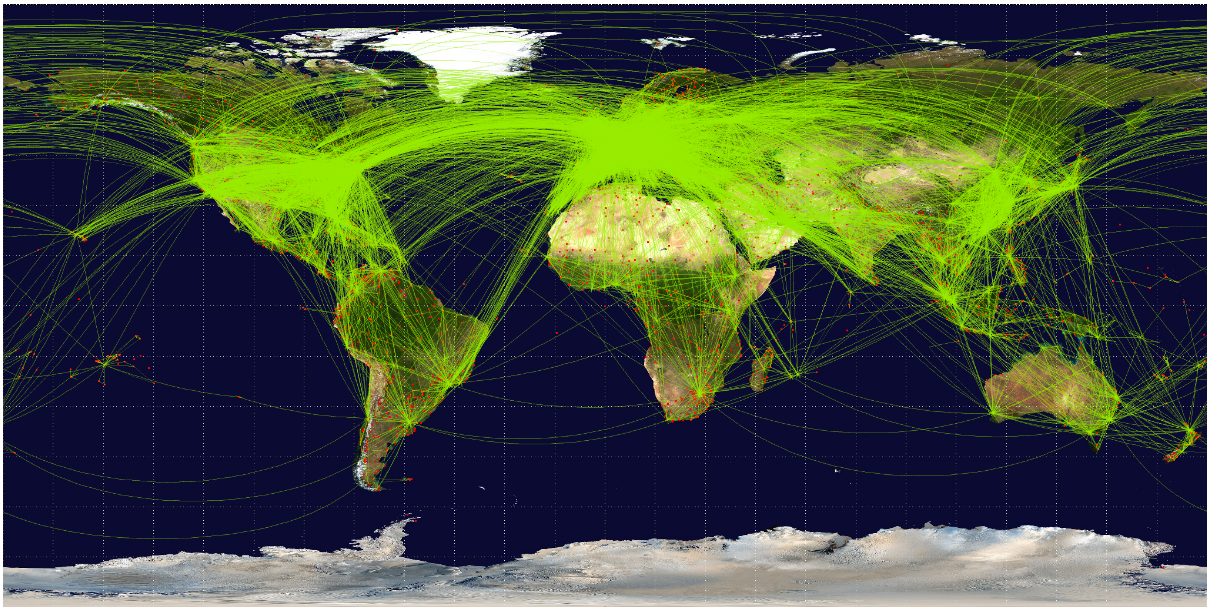


Figure 2: Global aviation routes, Source: Open Flights [1]

In Chapter 2 it was mentioned that the recorded data of any flight is automatically analyzed as part of the airline's FDM program. Calculating time points and measurements is conducted for any considered flight individually.

The next step is to analyze several flights together. Which flights are considered is given by certain characteristics. For example, all available flights landing in Munich airport EDDM on runway 26L can be considered. The set of characteristics that lead to a certain selection of flights is called "filter" within this paper. Depending on the properties of the specific FDM analyses, a certain level of filtering of the flights is required. For example, the analyses of the remaining fuel in the aircraft tanks at touchdown given in minutes of flight time can be considered as independent of any filter and can be combined for

[1] Downloaded from https://openflights.org/demo/openflights-routedb-2048.png on February 21st 2018 made available here under the Open Database License (ODbL) https://opendatacommons.org/licenses/odbl/1.0/.

any aircraft type and any airport. On the other hand, an analyses of the runway overrun probability depends on aircraft, runway, and further characteristics, see Section 9 of [6] as well as [7, 8].

Taking one further step leads to the setting considered in this paper. Thereby, not only a single filter is considered but a set of several filters, see Figure 3. As it can be seen, a hierarchical structure is generated which motivates using the term “filter tree”. Based on these filter trees, calculations can be conducted that compare different filters, i.e. different collections of flights, with each other. The central goal of this paper is to detect filters with an outstanding behavior using outlier detection algorithms. For these filters, the underlying flights show a different behavior than the ones for the remaining filters which might be relevant in terms of safety management.

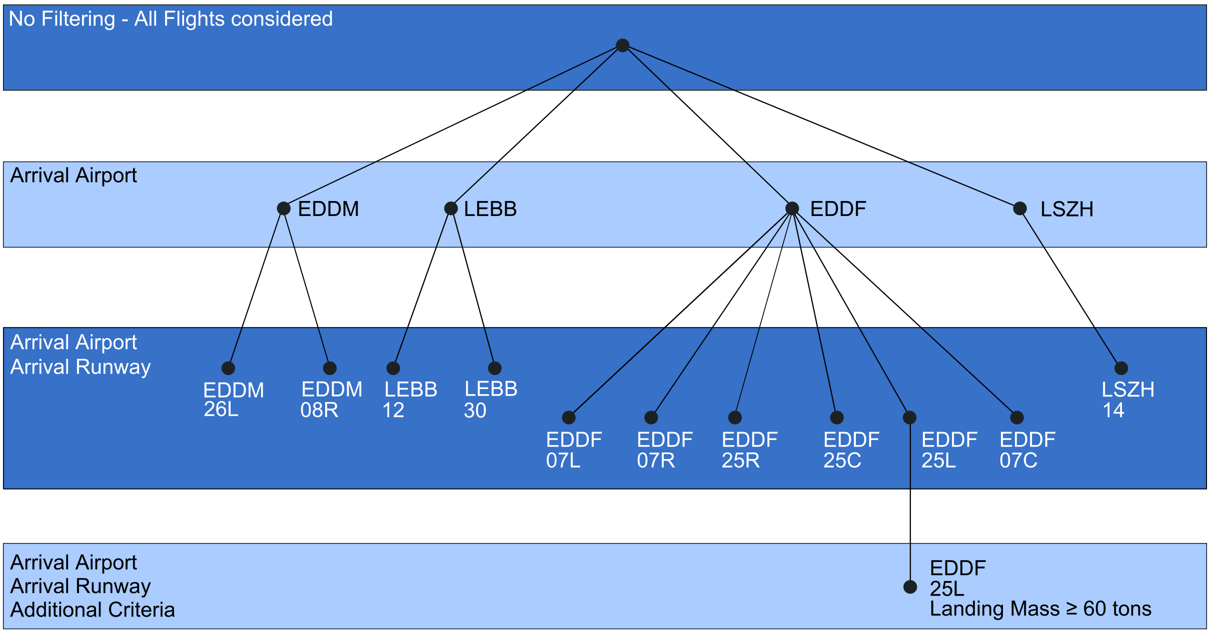


Figure 3: Filter tree

Figure 3 shows a filter tree with four different levels. For the first level, no filtering takes place and all available flights, in this case of Airbus A320 aircraft, are considered. On the second level, the flights are filtered according to their arrival airport. In particular, the four airports Munich EDDM, Bilbao LEBB, Frankfurt EDDF, and Zurich LSZH are considered. On the next level, the flights are further filtered according to their arrival runways. The last level consists of additional filter criteria, in this case according to the landing mass of the aircraft. The maximal landing mass of the considered A320 is 66 tons [9], so the filter is reasonable. In Figure 4, the numbers of flights fulfilling each filter of Figure 3 are given.

It is pointed out that the characteristics of the specific FDM software used for the analysis influence the decision whether or not these filter trees can be generated and which kind of properties can be used for filtering. The calculations within this paper have been conducted with the IT System developed by the Flight Safety working group at the Institute of Flight System Dynamics of the Technical University of Munich. This software allows very flexible filtering with respect to any available flight characteristic or measurement. Note that the first three filter levels of Figure 3 use general information about the flight. The fourth filter level uses the landing mass, which is considered as a measurement, see chapter 2, and is a more detailed description of the flight.

The scope of this paper is to present the idea of these calculations and to show an illustrative example.

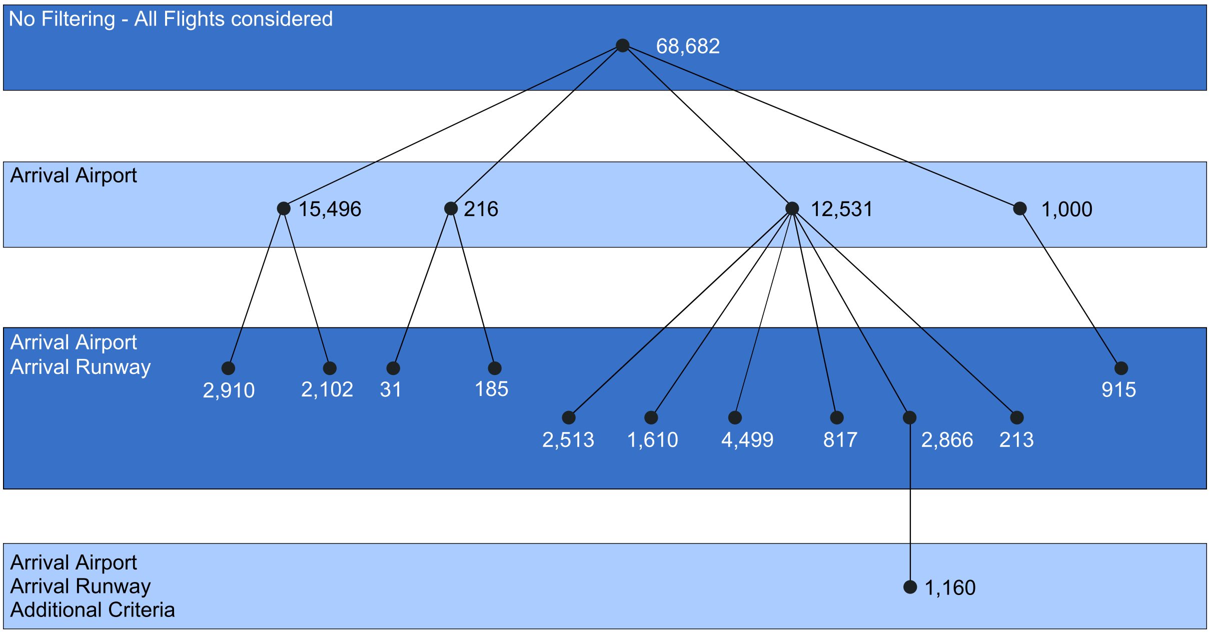


Figure 4: Number of flights fulfilling the filters

The mentioned software developed at the Flight Safety working group is also capable to generate full filter trees automatically. This means that for example any available arrival airport and arrival runway can be considered and arbitrary many filter levels can be added. Referring to Figure 2, the filter trees can then obviously get highly complex. This justifies the application of powerful machine learning algorithms to detect outstanding filters, which are discussed in the next chapter.

4. OUTLIER DETECTION AND MACHINE LEARNING

Starting with the available regression analysis at the beginning of the 19th century, data points "far" away from the regression model could be identified [10]. Nowadays, outlier detection can be considered as part of Machine Learning gaining more and more attention in data analytics. "*Another use of machine learning is outlier detection, which is finding instances that do not obey the general rule and are exceptions. The idea is that typical instances share characteristics that can be simply stated and instances that do not have those characteristics are atypical. In such a case, we are interested in finding a rule that is as simple as possible and covers as large a proportion of our typical instances as possible. Any instance that falls outside is an exception, which may be an anomaly requiring attention such as fraud; or it may be a novel, previously unseen but valid case, and hence the other name novelty detection.*", see page 9 of [11].

The outlier detection method that is used within this paper is Fuzzy clustering by Local Approximation of MEmberships (FLAME) [12]. At the Flight Safety working group at TUM, this method was first implemented in [13] [2]. In the following, the FLAME algorithm is briefly described.

Suppose that a set of data points $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^d$ is given. First, the k-nearest neighbors algorithm [14] is used to identify the neighbors of every data point. For any data point $\boldsymbol{x}_i$ with $i = 1, \ldots, n$ we denote

[2] Within [13], an illustrative video has been created by Max Schwenzer and uploaded to vimeo: https://vimeo.com/78348227. Availability of the link verified by the author of this paper on 13.03.2018.

the k-nearest neighbors by $\boldsymbol{x}_{i_{knn_1}}, \ldots, \boldsymbol{x}_{i_{knn_k}}$. Furthermore, we describe the euclidean distance of $\boldsymbol{x}_i$ to its neighbors by $d_{i_{knn_1}}, \ldots, d_{i_{knn_k}}$. These distances can be averaged for any data point

$$\bar{d}_i := \frac{1}{k} \cdot \sum_{j=1}^{k} d_{i_{knn_j}} \tag{1}$$

and furthermore transfered into a density

$$\rho_i := \frac{\max\limits_{j \in \{1,\ldots,n\}} \bar{d}_j}{\bar{d}_i}, \tag{2}$$

see page 19 of [13].

To detect outliers, [12] proposes to choose a density threshold. Data points with a density less than the threshold are considered as outlier. In [12], the threshold is given by the mean of all densities minus two times the standard deviation of them, however, this choice is not unique. In addition, simply the data points with the lowest density can be considered representing the most unusual characteristics in the given data set.

For the calculations performed within this paper, the value of k is set to $k = 3$. In addition to the outlier detection that is used in this paper, the FLAME algorithm can be used for clustering, i.e. assigning the filters to different categories. Thereby, so called Cluster Support Objects (CSO) are identified and the data points are iteratively assigned to the different categories using fuzzy membership. Since this clustering is not used within this paper, it is not described here and the reader is referred to [12].

5. DATA ASSIGNMENT AND NORMALIZATION

In the previous chapters, the available data, filter trees, and machine learning algorithms used for outlier detection are described. The goal of this chapter is to define the components of the considered data vectors assigned to any filter in the filter tree.

Considering the hierarchical structure of filter trees, the numbers of flights fulfilling a particular filter vary significantly for the different tree nodes, see Figure 4. It is important that these different numbers of flights do not falsify the calculations. For example, the sum of differences from the maximal landing mass and the actual landing mass is not suitable since it heavily depends on the number of considered flights. Alternatively, the average of all these differences can be considered without problems.

For the purpose of this paper, the characteristics mean value and standard deviation of measurements are chosen. In addition, so-called tail dependence coefficients related to the copula theory [15] are considered to represent dependencies between measurements. They particularly describe the behavior of measurements in the boundary areas, i.e. if they get excessively small or large. For example, the lower tail dependence is defined for a random vector $X = (X_1, X_2)$ with marginal Cumulative Distribution Functions (CDF) F_1 and F_2 as

$$\lambda_L = \lim_{u \searrow 0} \mathbb{P}(X_2 \leq F_2^{-1}(u) | X_1 \leq F_1^{-1}(u)). \tag{3}$$

For any details of the tail dependence coefficients and the copula theory, the reader is referred to [15].

The example chosen within this paper describes the braking behavior of pilots based on the filter tree of Figure 3. The considered components are

- Mean value of *Time Touchdown to Start Manual Braking*
- Standard deviation of *Time Touchdown to Start Manual Braking*
- Mean value of *Landing Mass*
- Standard deviation of *Landing Mass*
- Lower tail dependence between *Landing Buffer* and *Distance Threshold to Touchdown*
- Upper tail dependence between *Landing Buffer* and *Distance Threshold to Touchdown*

Thereby, the *Landing Buffer* δ is a value that describes the criticality of an aircraft landing that is determined on board before the landing. It is given by a relation between the Landing Distance Available (LDA) and the Landing Distance Required (LDR), see Equation (4).

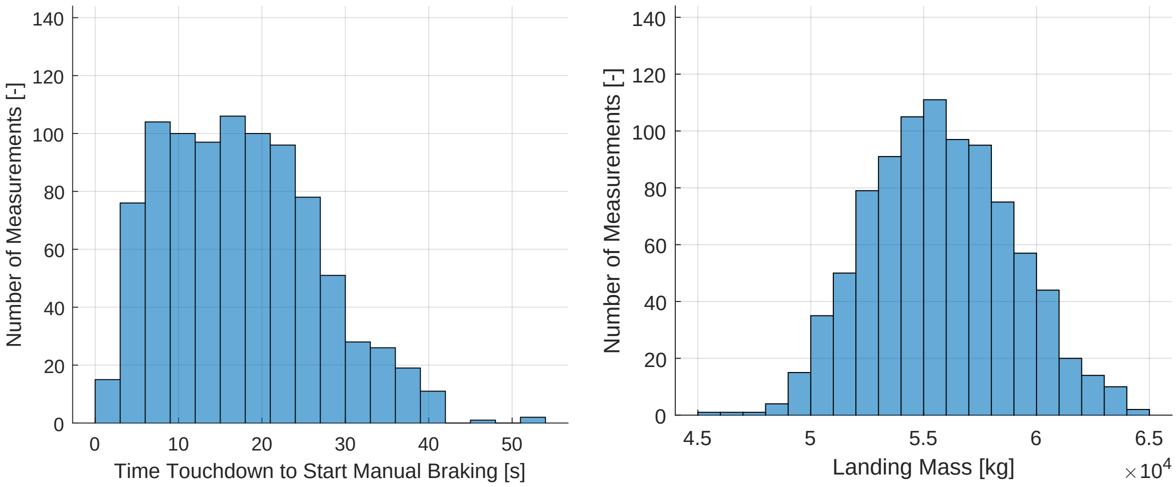
$$\delta = \frac{\text{LDA} - \text{LDR}}{\text{LDA}} \tag{4}$$

Further details can be found on page 205 of [7].

The combination of these components shall characterize the braking behavior. The central component is the *Time Touchdown to Start Manual Braking*. Furthermore, the *Landing Mass* also plays a role in the pilot's braking behavior. The last two components describe the dependence between the *Landing Buffer* and the *Distance Threshold to Touchdown* with the upper and lower tail dependence coefficients. As it was identified on page 205 of [7], the criticality of the specific approach and landing influences the pilot's landing behavior. This shall be represented by the two tail dependence coefficients to take filter specific properties into account.

Figure 5 shows the histograms for the variables *Time Touchdown to Start Manual Braking* (Figure 5a) and *Landing Mass* (Figure 5b) exemplary for the filter described by the arrival airport LSZH with arrival runway 14. In addition, the copula contour plot for the variables *Landing Buffer* and *Distance Threshold to Touchdown* for the arrival airport LEBB and arrival runway 30 is given in Figure 6a. The copula contour plot and also the tail dependence coefficients are calculated with the R package "VineCopula" [16]. In case of two independent parameters, the indicated contour lines would be concentric circles. Since this is not the case in Figure 6a, considerable dependence structures exist. Furthermore, Figure 6b highlights the areas with outstanding dependence using a heat plot. It compares the dependence structure of the chosen copula (180 degree rotated Tawn copula) with the dependence structure of a bi-variate normal distribution. The lighter an area in Figure 6b, the more the prevailing dependence differs from the bi-variate normal distribution. It can be seen that especially for small landing buffers, there are remarkable light areas that eventually lead to the tail dependence coefficient of 0.10, see Equation (3). Observe the different scales that are used for the copula contour plot (Figure 6a) and the heat plot (Figure 6b). For detailed information regarding the interpretation of copulas and the copula contour plot, the reader is again referred to [15].

Equation (1) shows that distances between data points are essential for the identification of outliers. Since the ranges of different measurements are in general completely different, a normalization step has to be performed. For example, the landing mass of an Airbus A320 might be 60,000 kg and the approach



(a) *Time Touchdown to Start Manual Braking*

(b) *Landing Mass*

Figure 5: Histograms for LSZH, runway 14

speed 70 m/s. Furthermore, the value of the standard deviation will be much higher for the landing mass such that eventually the landing mass will outweigh the approach speed in terms of distances between data points. Therefore, it was chosen that any data point considered for the outlier detection is linearly mapped to the interval $[-1, 1]$ in each component. This mapping is designed such that the minimal value of a component is mapped to -1 and the maximal value to 1.

6. IDENTIFICATION OF SAFETY CRITICAL SCENARIOS

The presented variables are calculated for any filter of the filter tree given in Figure 3. Flights fulfilling a particular filter carry the specific characteristics of that filter, e.g. of the airport or runway. The utilized data for the example of this paper is illustrated in Table 1 (some abbreviations are necessary, "M" corresponds to mean, "S" to standard deviation, "Ma" to mass, "L" to lower, and "U" to upper). Based on this data, the FLAME algorithm is applied and identifies the filters with outstanding characteristics, in this example with respect to the braking behavior.

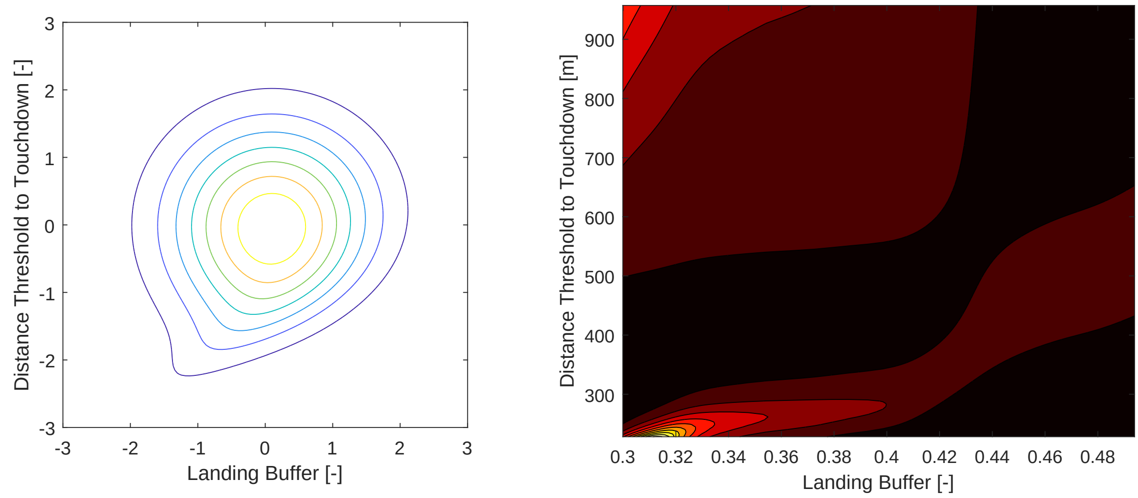
In the following, the three filters with the lowest density, see Equation (2) are presented starting with the lowest. Thereby, scenarios that were filtered at least until the runway level are considered. The overall result can be seen in Figure 7.

1. LSZH, Zurich Airport, Runway 14

2. LEBB, Bilbao Airport, Runway 30

3. EDDF, Frankfurt Airport, Runway 25L, Landing Mass $\geq$ 60 tons

For runway 14 of LSZH in Zurich, the reason of the low density is given by the airport layout, see Figure 8. Runway 14 is highlighted by the blue arrow and the relevant terminal by a blue circle. The runway exits are constructed at the runway end which leads to a special and late braking behavior and eventually contributes to a low density calculated by the FLAME algorithm. In addition, according to Table 1, it is the only filter with an upper tail dependence coefficient not equal to 0 and the average *Landing Mass* is lower than for the other filters.



(a) Copula contour plot

(b) Heat plot

Figure 6: Dependence *Landing Buffer* and *Distance Threshold to Touchdown* for LEBB, runway 30

Table 1: Data utilized by the FLAME algorithm

Filter	M Ma [kg]	S Ma [kg]	M Time [s]	S Time [s]	L Tail [-]	U Tail [-]
All	58,013	3,492	11.6	9.9	0	0
EDDM	57,925	3,369	9.0	8.3	0	0
LEBB	60,137	2,630	9.8	4.7	0	0
EDDF	58,557	3,296	13.2	10.2	0	0
LSZH	55,748	3,205	16.4	9.2	0.08	0
EDDM, 26L	58,665	3,186	8.0	5.4	0	0
EDDM, 08R	58,271	3,258	17.4	15.7	$2 \cdot 10^{-6}$	0
LEBB, 12	60,699	2,325	8.8	5.3	0	0
LEBB, 30	60,042	2,672	9.9	4.6	0.10	0
EDDF, 07L	58,520	3,237	18.6	9.9	0	0
EDDF, 07R	58,356	3,366	20.0	14.8	$8 \cdot 10^{-9}$	0
EDDF, 25R	58,621	3,258	9.4	5.4	$1 \cdot 10^{-4}$	0
EDDF, 25C	58,708	3,361	8.6	5.7	0	0
EDDF, 25L	58,541	3,360	12.1	9.8	$2 \cdot 10^{-5}$	0
EDDF, 07C	58,844	3,110	15.2	14.5	0	0
LSZH, 14	55,756	3,153	16.9	9.2	0.005	0.005
EDDF, 25L, $\geq$ 60 t	61,745	1,066	11.6	9.1	0	0

The filter with the second lowest density is runway 30 of LEBB in Bilbao. Landings in Bilbao are famous to be challenging due to common significant wind situations [17]. In addition, the runway length is 2600 m which is together with a displaced threshold of 460 m rather short [18]. Table 1 indicates a high average *Landing Mass*, a low average *Time Touchdown to Start Manual Braking* as well as the highest tail dependence of all filters.

The filter with the third lowest FLAME density is Frankfurt EDDF, runway 25L with the additional filter *Landing Mass* greater or equal 60 tons. The outstanding behavior of the mean value and standard deviation of the *Landing Mass* can be also seen in Table 1 and are directly influencing the FLAME algorithm. Therefore, this additional criteria is directly leading to this low density.

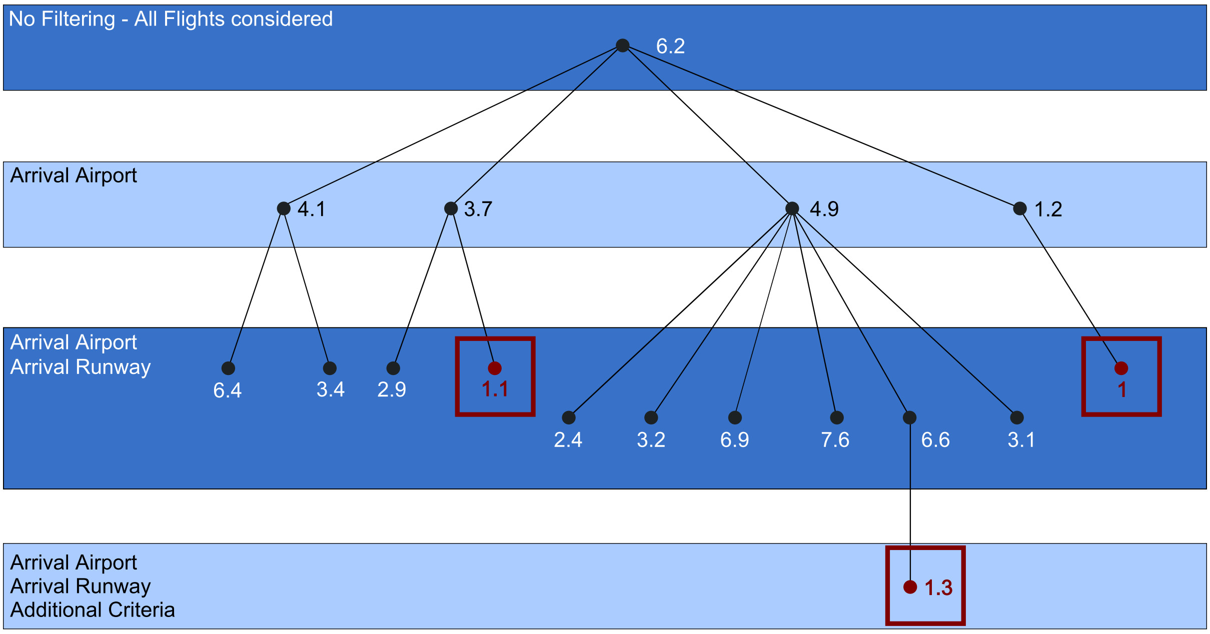


Figure 7: Densities calculated by the FLAME algorithm

Figure 8: LSZH runway 14 layout, source: Google Earth, Image Landsat / Copernicus

7. CONCLUSIONS AND OUTLOOK

Within this paper, the concept of filter trees is introduced to Flight Data Monitoring (FDM) to be able to identify scenarios outstanding in terms of safety relevant performance. Based on complex route networks and fleet structures of an airline, the filter trees considered in real life airline operations can become immense. To be able to find outstanding filters in such complex filter trees, concepts of the machine learning algorithm FLAME are used. Within this paper, the proposed methodology is shown on a simplistic example. The IT environment developed at the Institute of Flight System Dynamics allows automatic generation of full filter trees to provide more complex and realistic applications in the future.

Acknowledgements

This paper was supported by the Deutsche Forschungsgemeinschaft (DFG) within the project "Copula based dependence analysis of functional data for validation and calibration of dynamic aircraft models" with the project identifier HO 4190/10-1.

References

[1] P. Koppitz, J. Siegel, N. Romanow, L. Höhndorf, and F. Holzapfel, "Touchdown point detection for operational flight data using quality measures and a model based approach," in *2018 AIAA Atmospheric Flight Mechanics Conference*, AIAA SciTech Forum, AIAA, Inc, 2018.

[2] J. Siegel, *Reconstruction and Safety Assessment of Aircraft Landings Based on Operational Quick Access Recorder Data*. Master's thesis, Technische Universität München, 25.04.2017.

[3] J. Siegel, *Touchdown Point Reconstruction Based on Operational Quick Access Recorder Data*. Semester's thesis, Technische Universität München, 18.11.2015.

[4] L. Höhndorf, J. Siegel, J. Sembiring, P. Koppitz, and F. Holzapfel, "Reconstruction of aircraft states during landing based on quick access recorder data," *Journal of Guidance, Control, and Dynamics*, vol. 40, no. 9, pp. 2393–2398, 2017.

[5] European Aviation Safety Agency, "Position paper on the compliance of easa system and eu-ops with icao annex 6 safety management systems (sms) standards and recommended practices for air operators.," 20.12.2007.

[6] International Air Transport Association, *Safety report 2013: Issued April 2014*. Montréal, Québec: International Air Transport Association (IATA), 50th edition ed., 2014.

[7] L. Drees, *Predictive Analysis: Quantifying Operational Airline Risks*. München: Dr. Hut, 2017.

[8] C. Wang, *Quantifizierung von Unfallwahrscheinlichkeiten für Runway Overrun über Subset-Simulation*. Master's thesis, Technische Universität München, 29.10.2013.

[9] Airbus, "Airbus family figures: March 2016 edition," 2016.

[10] P. J. Rousseeuw and A. M. Leroy, *Robust regression and outlier detection*. Wiley series in probability and statistics, Hoboken, NJ: Wiley-Interscience, 2003.

[11] E. Alpaydin, *Introduction to Machine Learning*. Adaptive Computation and Machine Learning series / Ethem Alpaydin, Cambridge: MIT Press, 3 ed., 2014.

[12] L. Fu and E. Medico, "Flame, a novel fuzzy clustering method for the analysis of dna microarray data," *BMC bioinformatics*, vol. 8, p. 3, 2007.

[13] M. Schwenzer, *Advanced Flight Data Analysis: a study on Machine Learning algorithms*. Bachelor's thesis, Technische Universität München, 29.11.2013.

[14] B. W. Silverman and M. C. Jones, "E. fix and j.l. hodges (1951): An important contribution to nonparametric discriminant analysis and density estimation: Commentary on fix and hodges (1951)," *International Statistical Review / Revue Internationale de Statistique*, vol. 57, no. 3, p. 233, 1989.

[15] H. Joe, *Dependence modelling with copulas*, vol. 134 of *Monographs on statistics and applied probability*. Boca Raton: CRC Press, op. 2015.

[16] U. Schepsmeier, J. Stoeber, E. C. Brechmann, B. Graeler, T. Nagler, T. Erhardt, C. Almeida, A. Min, C. Czado, M. Hofmann, M. Killiches, H. Joe, and T. Vatter, "Package 'vinecopula': Manual," 16.08.2017.

[17] "Landing at bilbao airport not for the faint-hearted," *The Telegraph*, 14.12.2012.

[18] "Anexo vii: aip. aeropuerto de bilbao (2011)," 2011.