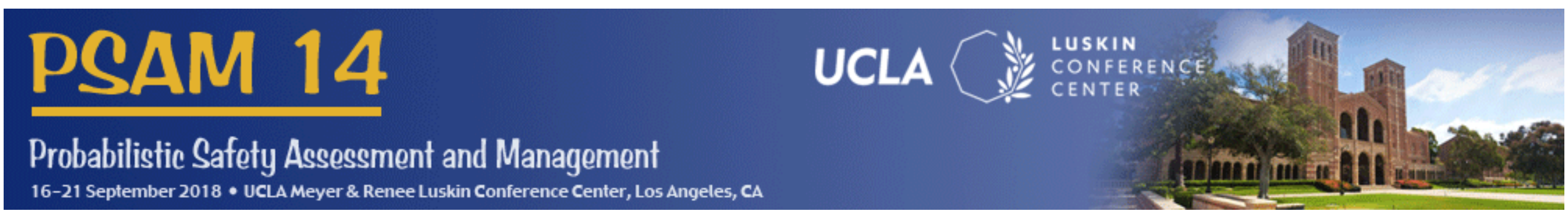
# Identification of Safety Critical Scenarios for Airlines using Machine Learning in Filter Trees

Lukas Höhndorf, Florian Holzapfel



# Agenda

- **Introduction**
- Filter and Filter Trees
- Outlier Detection
- Example
- Summary

# Introduction

1

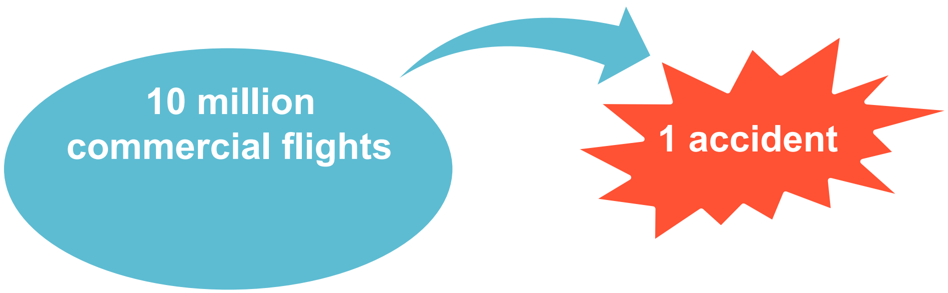
ICAO DOC 9859

- Airlines are required to implement a safety management system (SMS)
- SMS requires operators also to define their own **Acceptable Level of Safety (ALoS).**

2

Flightpath 2050

- **Europe** aims at less than one accident per ten million flights (i.e. **accident probability of $10^{-7}$ per flight**).



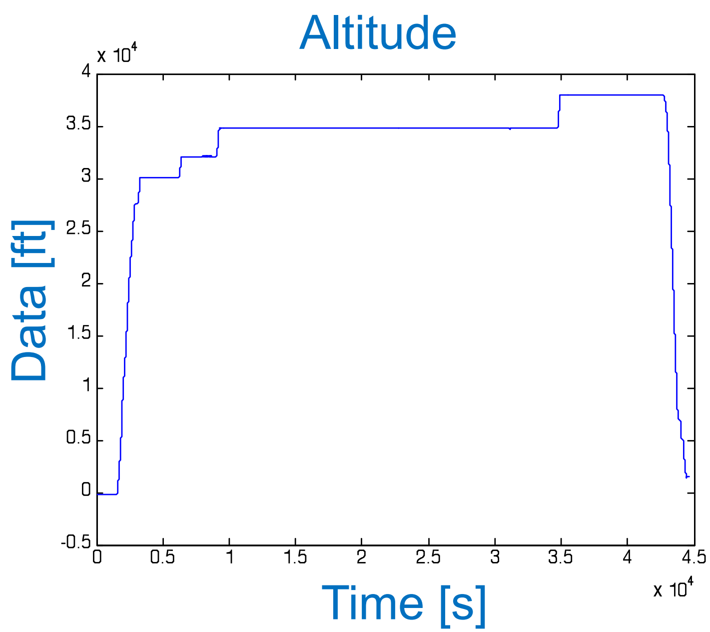

**Central question: What is the current safety level of an airline?**

This question is aimed to be answered using **data analysis**

# Introduction

- Data is obtained from the "Quick Access Recorder"
- Up to 2400 parameters are recorded.
  Depending on the aircraft and the airline.
- Frequency usually between ¼ Hertz and 16 Hertz
  Depending on the parameter.
- Number of the recorded parameters increased significantly in the last years.
- Current Flight Data Monitoring (FDM) algorithms often do not use the full potential of the data.

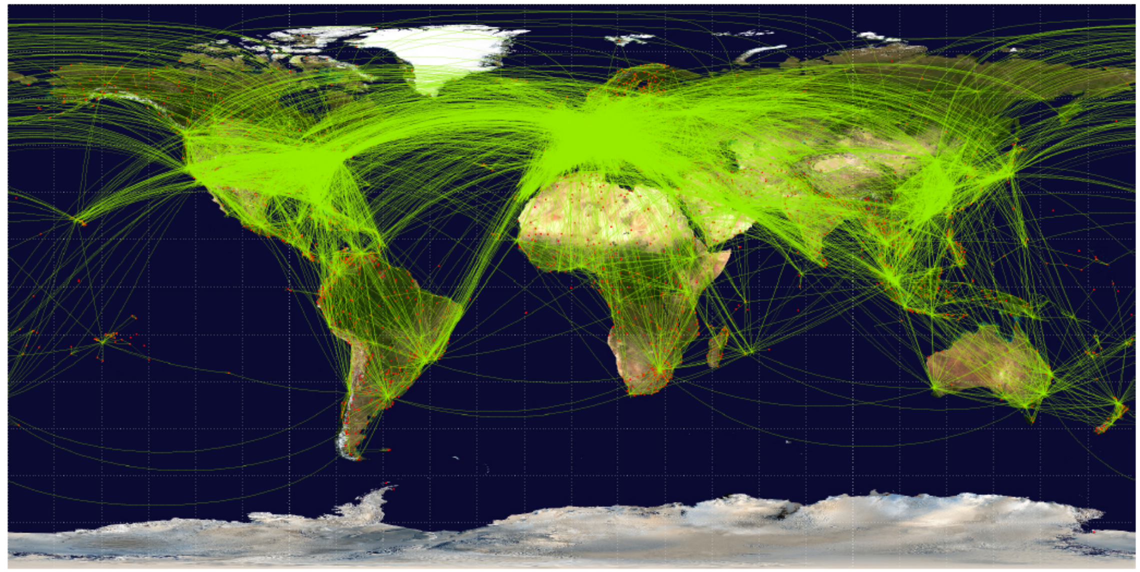

*Source: http://www.cwc-ae.com/sites/default/files/Enhanced_MultiPurpose_Flight_Recorder_EMPFR.jpg*



# Filter and Filter Trees

There are various different characteristics of an airline operation:

- Aircraft types
- Airports
- Runways
- Weather situation
- Operational criteria
  (e.g. flap setting at landing)
- …

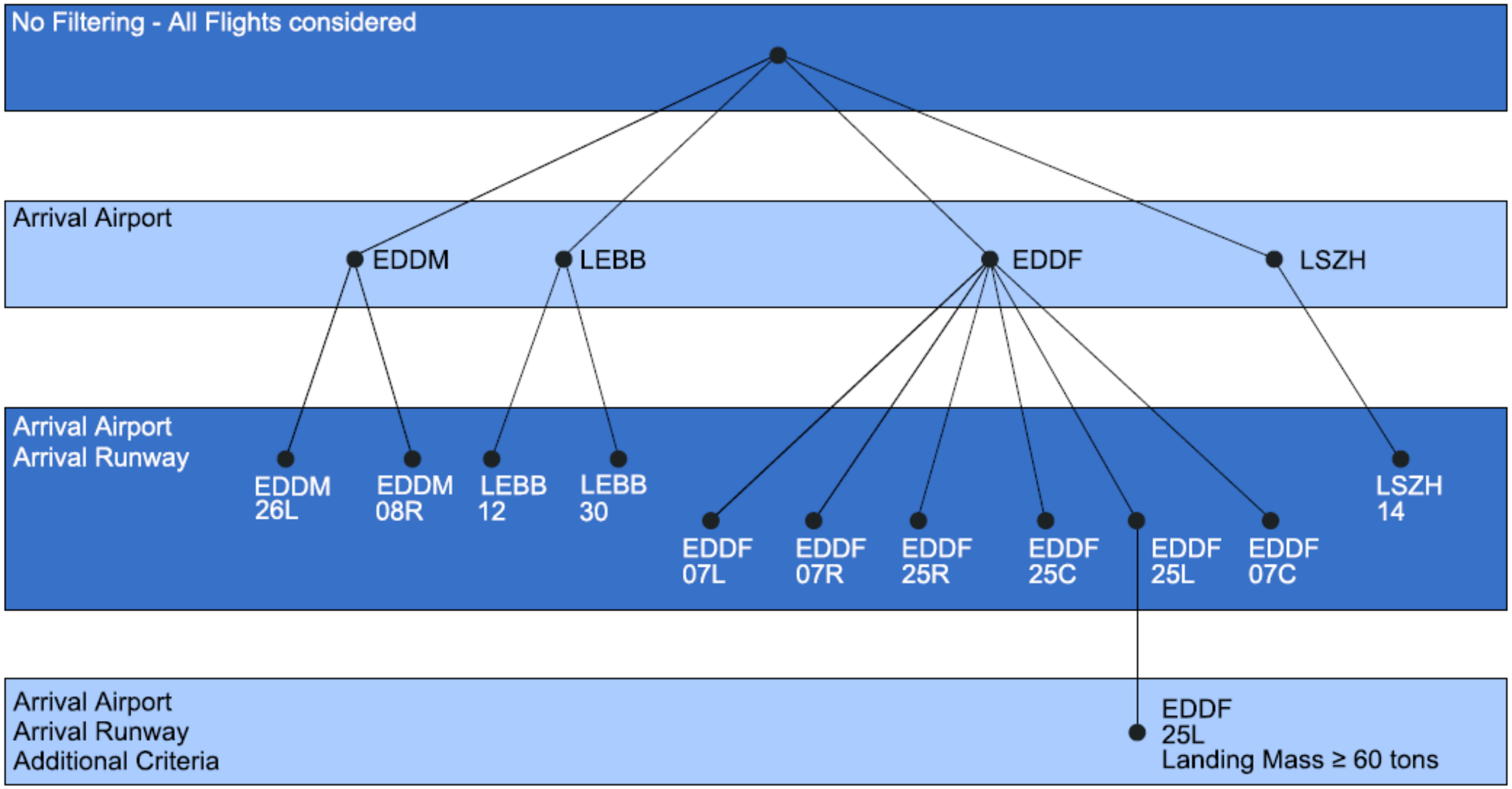


Different FDM algorithms require different level of flight filtering:

- Runway overrun probability estimation (depends on aircraft type, runway, weather, operational criteria,…)
- Analysis of remaining fuel at landing given in minutes of flight time (directly comparable for different aircraft type, airports, runways…)

# Filter and Filter Trees

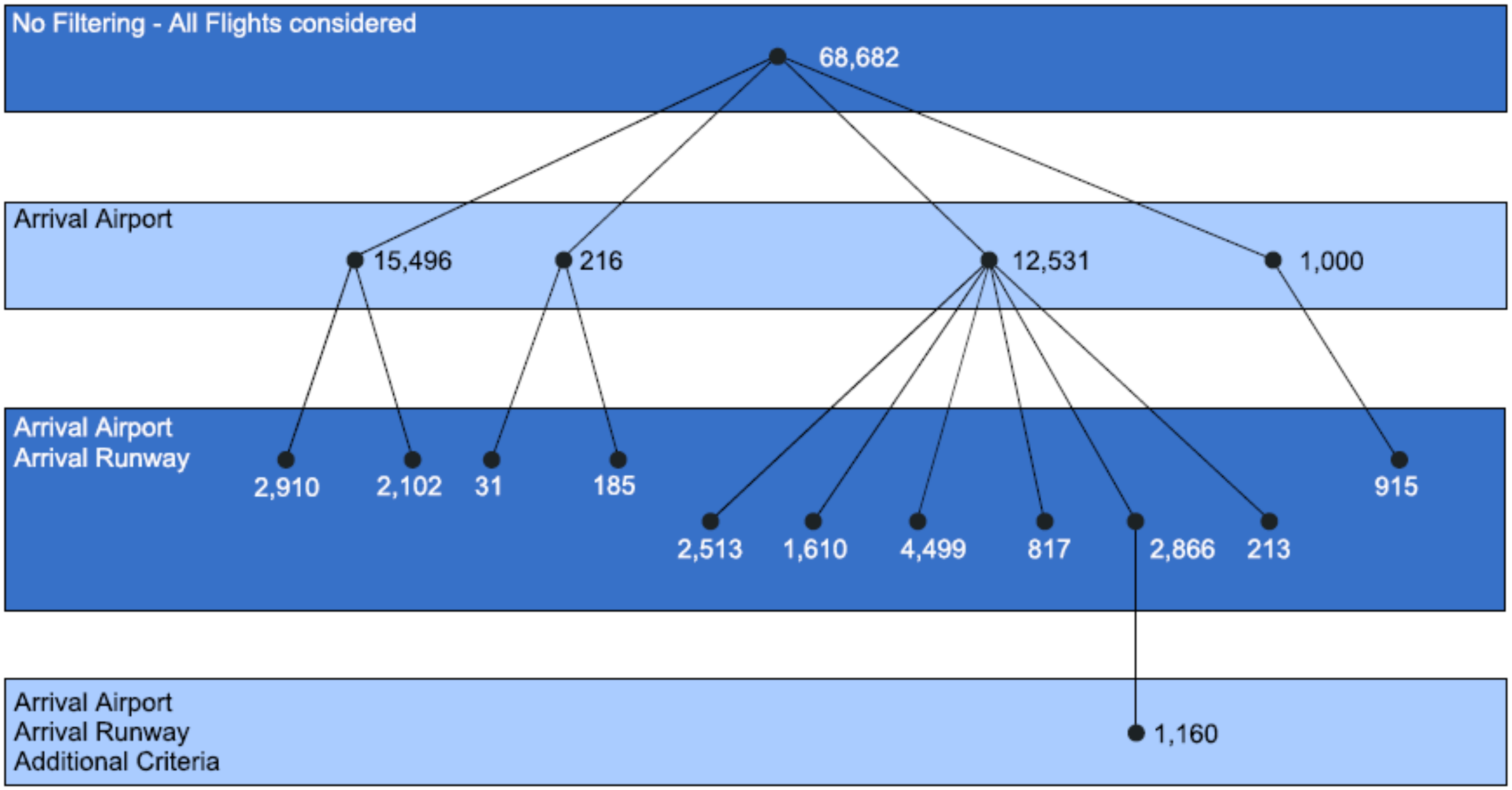
These different characteristics can often be hierarchically arranged in a filter tree:




This example is small, however, “full” trees in practice can get immense.

# Filter and Filter Trees

The number of flights of different filters (i.e. filter tree nodes) vary significantly:

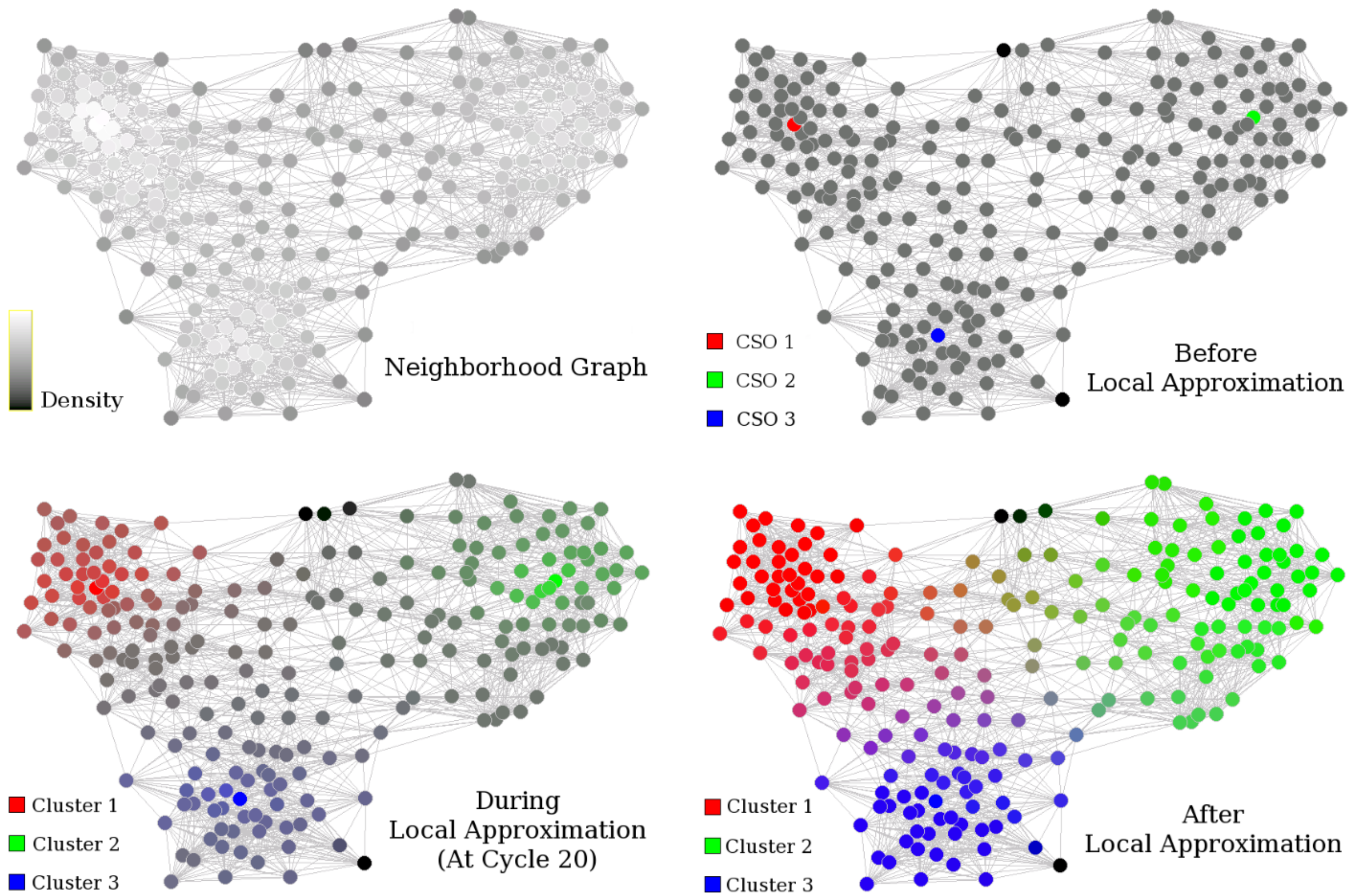



Any algorithms utilizing these filter trees need to take those different numbers into account.

# Outlier Detection

- The goal is to detect filters outstanding from a safety perspective.
- For the outlier detection, the machine learning algorithm FLAME is used:
  - Suppose $x_1, \dots, x_n \in \mathbb{R}^d$. For $x_i$, the $k$-nearest neighbors are $x_{i_{knn,1}}, \dots, x_{i_{knn,k}}$, the distance to $x_i$ are denoted by $d_{i_{knn,1}}, \dots, d_{i_{knn,k}}$.
  - For any $x_i$, these distances can be averaged $\overline{d_i} = \frac{1}{k} \sum_{j=1}^{k} d_{i_{knn,j}}$
  - And transformed into a density $\rho_i = \frac{\max_{j \in \{1,\dots,n\}} \overline{d_j}}{\overline{d_i}}$
- Within this paper, the scenarios with the lowest densities are interpreted (i.e. the clusters determined by the FLAME algorithm are not further considered within this paper).
- Since distances between data points are utilized, a normalization of the data is necessary (otherwise variables with high ranges outweigh variables with lower ranges). For our purposes, this is conducted linearly onto $[-1,1]$ with the minimal value to $-1$ and the maximal to $1$.

# Outlier Detection




Source: https://en.wikipedia.org/w/index.php?curid=14449059

# Example

The chosen example considers the pilot's braking behavior:

- 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 $\delta$ 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). Low $\delta$ is critical, high $\delta$ uncritical.

$$\delta = \frac{LDA - LDR}{LDA}$$

The upper/lower tail dependence are metrics to describe the dependence specifically at the boundary domains. Please see the paper for details.

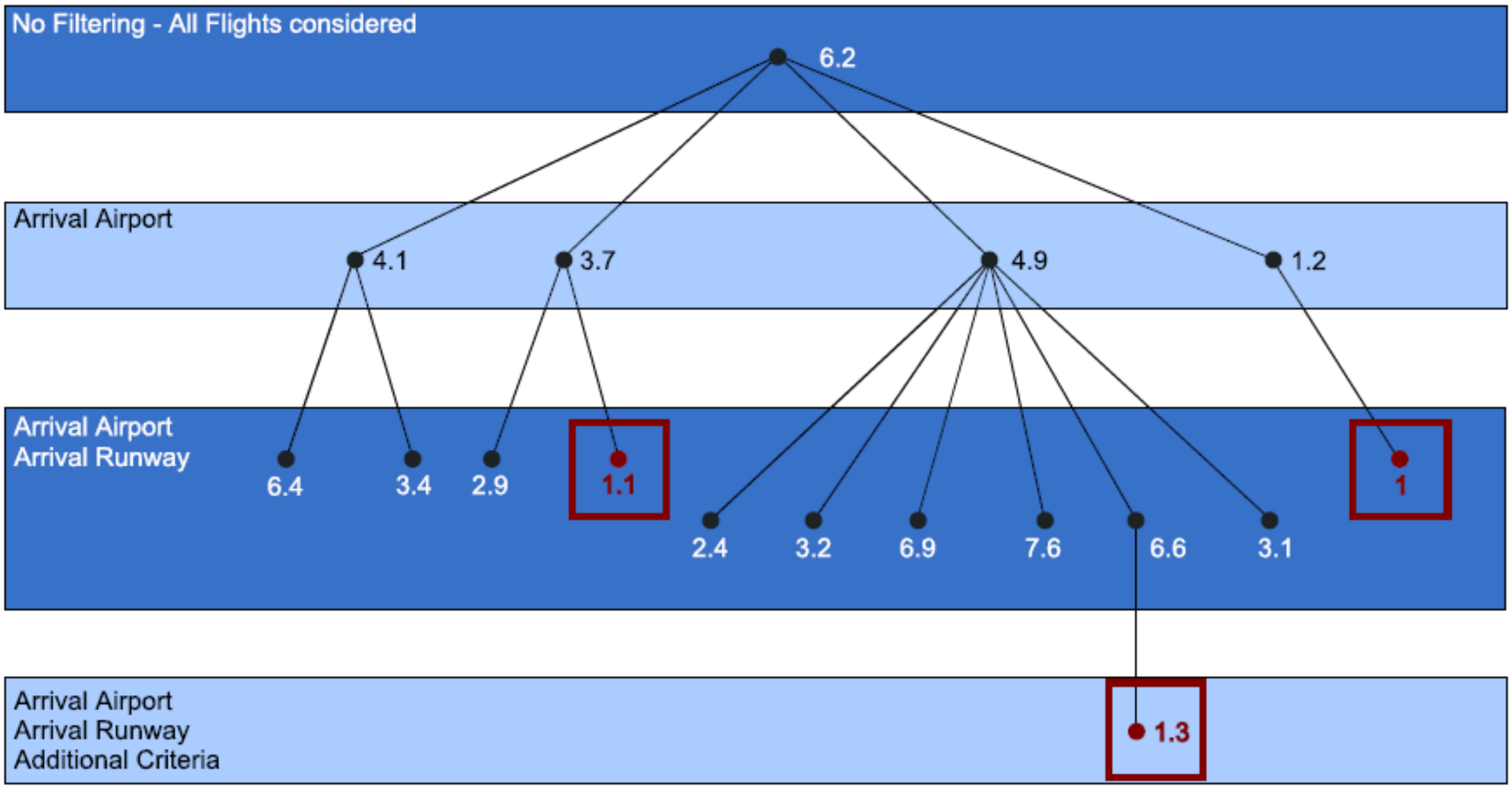
# Example

These values can be calculated for any node in the filter tree:

| 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 |

# Example

The densities determined by the FLAME algorithm are:



# Example

The three filters with the lowest FLAME density on the 3rd and 4th level of the filter tree are:

- 1. LSZH, Zurich Airport, Runway 14
- 2. LEBB, Bilbao Airport, Runway 30
- 3. EDDF, Frankfurt Airport, Runway 25L,
  Landing Mass ≥ 60 tons


Airport Layout: LSZH, Zurich Airport, Runway 14

# Example

$2^{nd}$ lowest density:

- Runway 30 of LEBB in Bilbao.
- Landings in Bilbao are famous to be challenging due to common significant wind situations.
- Runway length is 2600m which is together with a displaced threshold of 460m rather short.

$3^{rd}$ lowest density:

- Frankfurt EDDF, runway 25L with 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 the data table and are directly influencing the FLAME algorithm.

# Summary

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

Filter trees allow to handle complex networks of airline operations

Powerful outlier detection algorithms are available in machine learning

**Identification of Safety Critical Scenarios**

Huge amount of flights can be categorized in filter trees

Scenarios outstanding from a safety performance can be identified


**Lukas Höhndorf**, lukas.hoehndorf@tum.de
**Florian Holzapfel**, florian.holzapfel@tum.de

**Javensius Sembiring**, javensius.sembiring@tum.de
**Chong Wang**, chong.wang@tum.de
**Phillip Koppitz**, phillip.koppitz@tum.de
**Xiaolong Wang**, xiaolong.wang@tum.de
**Serçin Höhndorf**, sercin.hoehndorf@tum.de
**Florian Schwaiger**, f.schwaiger@tum.de

# Thank you for your attention