Co-financed by the Connecting Europe
Facility of the European Union

EUROCONTROL

Applying Machine Learning Modeling to Enhance Runway Throughput at a Big European Airport

10th EASN International Conference

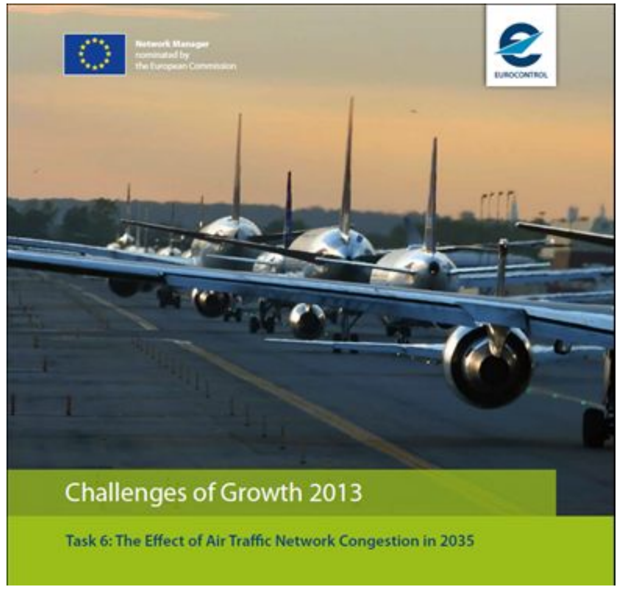
Guillaume Stempfel, Victor Brossard, Antoine Bonnefoy
Eura Nova
Ivan De Visscher
WaPT
Mohamed Ellejmi, Vincent Treve
EUROCONTROL

September 2nd 2020

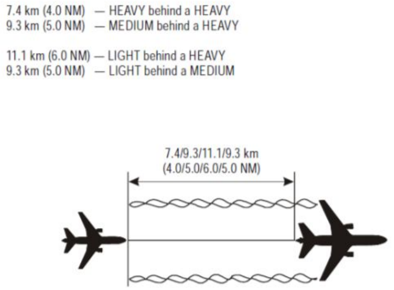
WaPT

EURA NOVA

Context



- Traffic growth puts pressure on the European Airport's capacity
- For peak operations, arrival traffic separation and throughput are constrained by:



Wake turbulence separation

Surveillance minima

Runway Occupancy Time (ROT)

- SESAR2020 Project 02 EnhAnced Runway THroughput (EARTH) aims to optimize runway performance with existing infrastructure
  - Through reduction of wake separation (e.g RECATpairwise) or surveillance separation

  ⇒ The most limiting constraint may then become the leader aircraft ROT
- ROT depends on many factors hard to anticipate for the ATC
  - Aircraft type, runway configuration, runway status, weather conditions, landing profile….
  - Significant buffers have to be taken in account to cover all possible cases

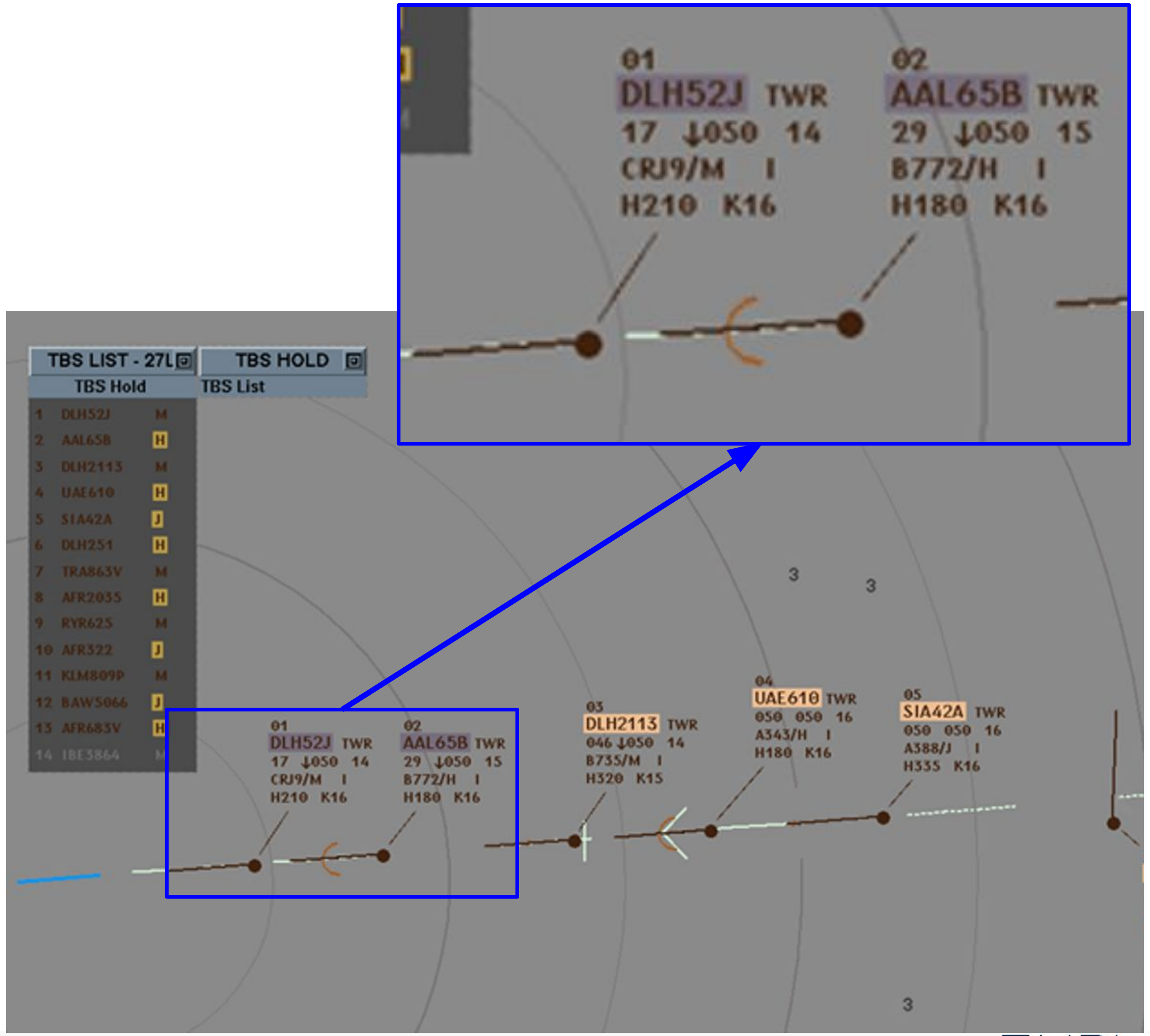
ROT definition

- Difference between the time an aircraft is observed to overfly the runway threshold and the time at which it has vacated the runway

- An aircraft is considered as having vacated the runway if
  - Either it is on the runway exit
  - Or it is still on the runway but sufficiently far from the threshold (> 2400m)

Our strategy

- Helping the ATC with a Machine Learning model able to predict the ROT depending on all impact parameters
- This model can directly be used to provide the required ROT spacing using a simple separation delivery tool
- 3 years of operations at LSZH Zürich airport
  - Covers a period from 2015 to 2017
  - 3 datasets:
    - Surveillance radar tracks of all landing flights
    - METAR data (wind speed, wind gust, visibility, pressure, temperature, meteorological events...)
    - Surface wind speed and direction
- Focus on peak operation periods
- 233.000 exploitable flights after filtering



Zurich airport topology

- 3 arrival runways:
  - 14 (main arrival runway)
  - 28
  - 34
- Runway 14 is significantly different
  - First exit is located at more than 2200m from the threshold
  - Other exits are located after the 2400m limit
  - All are high speed exits
- We expect the ROT to be driven by:
  - The aircraft final TAS on runway 14
  - The runway exit chosen by the aircraft on 28 and 34

2 Problems

1) The ATCo requires an information on the applicable separation/spacing at interception, therefore ROT spacing must be available several (typically 10) minutes before landing

⇒ The model must only use data available at least 10 minutes before the landing of the target flight

2) Is it possible to correct the prediction at shorter term ?
   - Using refreshed data (or data unavailable in the 10-minutes term)
   - Goal: anticipate a potential sub separation and order a go-around for the follower aircraft

⇒ We define 3 additional set-ups at shorter-term:
   - At 8 NM from the threshold
   - At 2 NM from the threshold
   - At the threshold

10-minutes set-up

- Features:
  - Aircraft type
  - Wake turbulence RECAT category
  - Landing runway
  - Landing week, weekday and hour
  - Airline
  - The departure airport and its country
  - Surface wind data (aggregated over 3 and 15 minutes)
  - Pressure
  - Visibility
  - Temperature
  - Weather events: rain, fog, snow, haze, convective clouds…
  - Event intensity
  - Ceiling clouds altitude

10-minutes set-up

- Modelisation as a regression problem where the target is the ROT value

- Dataset split into two subsets
  - ⅔ for the learning
  - ⅓ for the evaluation

- Selection of the model among a set of non-linear regressors through a 5-fold cross-validation
  - XGBoost with a maximum of 5000 estimators of depth 8

- Improvement brought by the model is assessed through a baseline confrontation

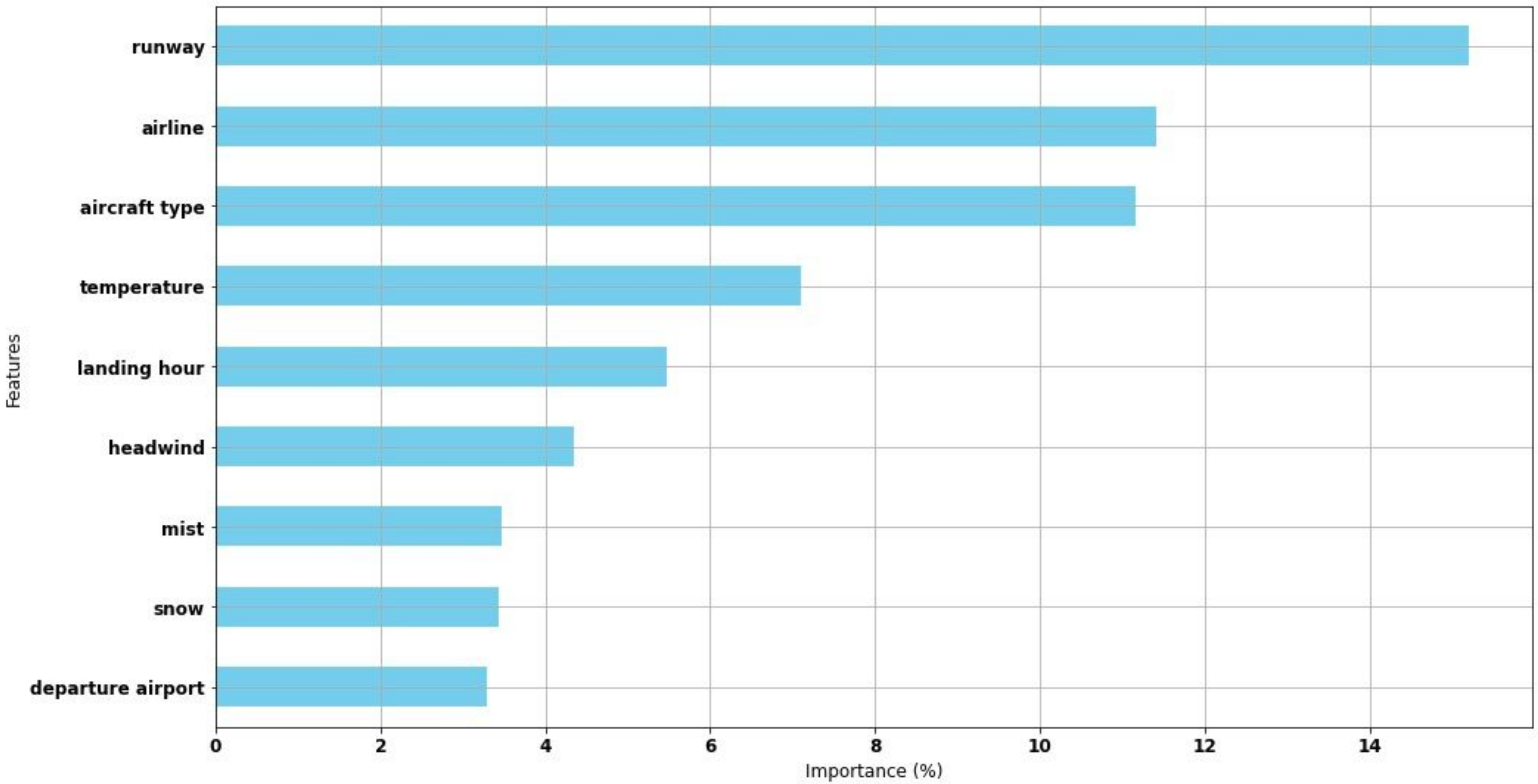
- The baseline decision is defined as the average ROT observed per aircraft type and per runway

Improvement w.r.t baseline

Runway\Model	R^2-Score XG-Boost	R^2-Score Baseline	# Data
All runways	**0.516**	0.352	70211
14	**0.518**	0.300	54504
28	**0.222**	0.146	11161
34	**0.340**	0.207	4546

- Improvement of around 50% of the R2-score w.r.t to the baseline
- Strong improvement on the runway 14
- As expected, the prediction task seems to be more complicated on runways 28 and 34

Features importance



Shorter terms set-ups

- Additional features:
  - Aircraft ground speed (for 2NM and threshold set-ups)
  - Presence of an aircraft on the landing runway (for 2NM set-up)
  - Time elapsed since the last landing on the runway (for 2NM set-up)

- Updated weather features
  - Last recorded wind speed and direction at decision time
  - If available, updated METAR data

- Same algorithm and parametrization than for 10-minutes set-up

Shorter terms set-ups

Setup	R^2 Score
10 Minutes	0.524
8 Nautical Miles	0.528
2 Nautical Miles	0.554
Threshold	0.582

- The gain in 8 NM set-up is negligible. The refreshment of weather data does not seem to bring much information
- The gain at 2NM and at threshold is substantial. The aircraft ground speed and the leader landing data seems to be informative for ROT prediction
- Shorter-term ROT predictions may be then used to correct the initial one

Application: Definition of ROT-based MRS

- ICAO PANS ATM doc 4444 allows for reduced MRS from 3.0 NM to 2.5 NM between succeeding aircraft which are established on the same final approach track within 10 NM of the runway threshold provided that the average ROT of landing aircraft is proven not to exceed 50 seconds.

- Following the same logic as for MRS=2.5 NM, reduced MRS below 3 NM can be applied behind all aircraft types for which the average ROT <60 s and following

  MRS [NM]= ROT [s] x 180 kts/3600

- Indeed: 2.5 NM traveled in 50 s corresponds to a speed of 180 kts.

- Requires ROT prediction

Naïve vs M/L approach
Definition of ROT distance spacing

EUROCONTROL

- ## Naive approach:
  - ROT defined per aircraft type and per runway QFU as the mean of the observed values
  - MRS_{naive} [NM]= ROT_{naive} [s] x180 kts /3600
  - Obtained from the learning dataset

	Mean ROT [s]	ROT-based MRS [NM]
A321	46	2.3
A319	48	2.4
B738	50	2.5
A318	52	2.6
GLF5	58	2.9

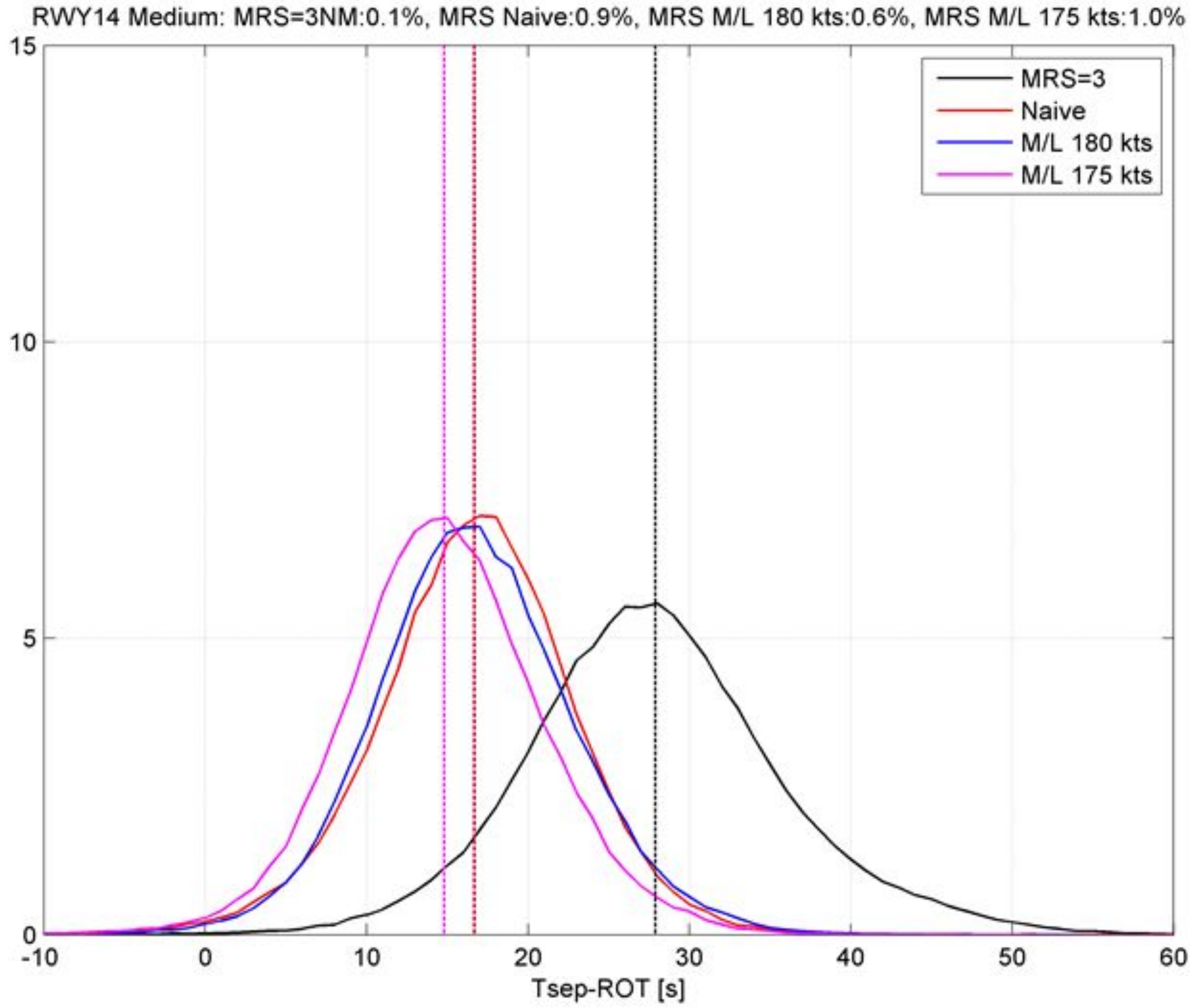
- ## M/L approach:
  - ROT defined using model with many features (aircraft type, QFU, airline, met data, etc.) calibrated through M/L
  - Obtained from the learning dataset
  - 2 spacing definitions:
    - $MRS_{M/L,\ 180kts}$ [NM]= $ROT_{M/L}$ [s] x180 kts /3600
    - $MRS_{M/L,\ 175kts}$ [NM]= $ROT_{M/L}$ [s] x**175** kts /3600

    ⇒ taking advantage of higher ROT prediction accuracy

Naïve vs M/L approach Assessment

- For all ICAO Medium landplane aircraft flights
- Computation of the failure rate (= separation delivery below actual ROT) when aircraft are spaced at 3NM (current MRS), MRS_{naive}, $MRS_{M/L,180kts}$, or $MRS_{M/L,175kts}$
- Using the time-to-fly distribution of all aircraft landing within +/- 5 min (assumed to be possible follower) considering all landplane followers (except Lights)
- Obtained using the test dataset (independent from the learning dataset)
- Distinction between follower category

Naïve vs M/L approach Assessment



0.9% of the pairs spaced below observed ROT if all separated at MRS_{naive}

0.6% of the pairs spaced below observed ROT if all separated at $MRS_{M/L}$ computed with 180 kts

1.0% of the pairs spaced below observed ROT if all separated at $MRS_{M/L}$ computed with 175 kts

Naïve vs M/L approach
Assessment results

Runway	mean Tsep-rot MRS 3 NM	mean Tsep-rot MRS_{naive}	mean Tsep-rot $MRS_{M/L,180}$	mean Tsep-rot $MRS_{M/L,175}$	Percentage below ROT MRS_{naive}	Percentage below ROT $MRS_{M/L, 180}$	Percentage below ROT $MRS_{M/L,175}$
14	27.8 s	16.7 s	16.7 s	14.8 s	0.9%	0.6%	1.0%
28	23.9 s	14.3 s	14.1 s	12.3 s	1.1%	0.6%	0.9%
34	27.9 s	16.7 s	16.9 s	15.1 s	0.9%	0.6%	1.0%

-10 to 11 s

-10 to 11 s

Lower error rate

-12 to 13 s

Same error rate but …
Better coverage of extreme cases

Conclusions and next steps

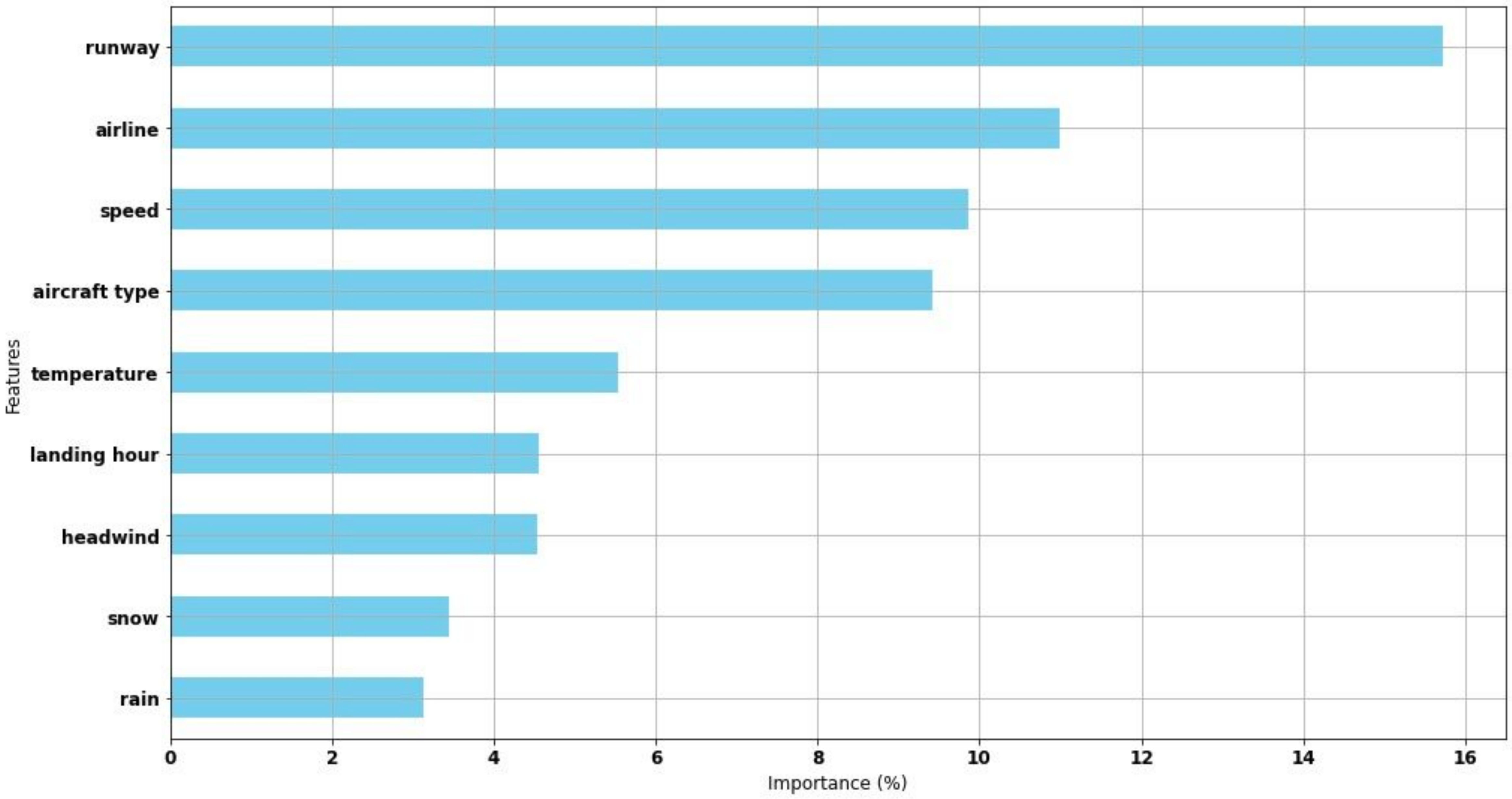
- Machine Learning can support the development of accurate models for ROT
  - the XGBoost model outperforms an average-based baseline
  - the prediction may be refined at shorter-term (especially using the aircraft ground speed)
- Based on Zurich case, compared to a naive approach, a model allows
  - reducing ROT-spacing by 3% leading to improved runway performance
  - slight improvement of safety by better predicting extreme large ROT
- ROT ML predictions can also be used as input in a more advanced ATC separation Delivery tool such as the TBS-ORD tool concept developed in SESAR
  - 10-minutes predictor for initial indicator
  - 8 NM predictor for indicator update if needed
  - 2 NM and threshold predictors for alerts
- Perspectives
  - Possible extension to a multi-airport context using runway topologies
  - On-going work in two projects
    - VLD3 (Heathrow and Zürich)
    - SESAR W2 AART (PJ02.14-10 ROCAT)

Appendix: Example of difficult cases

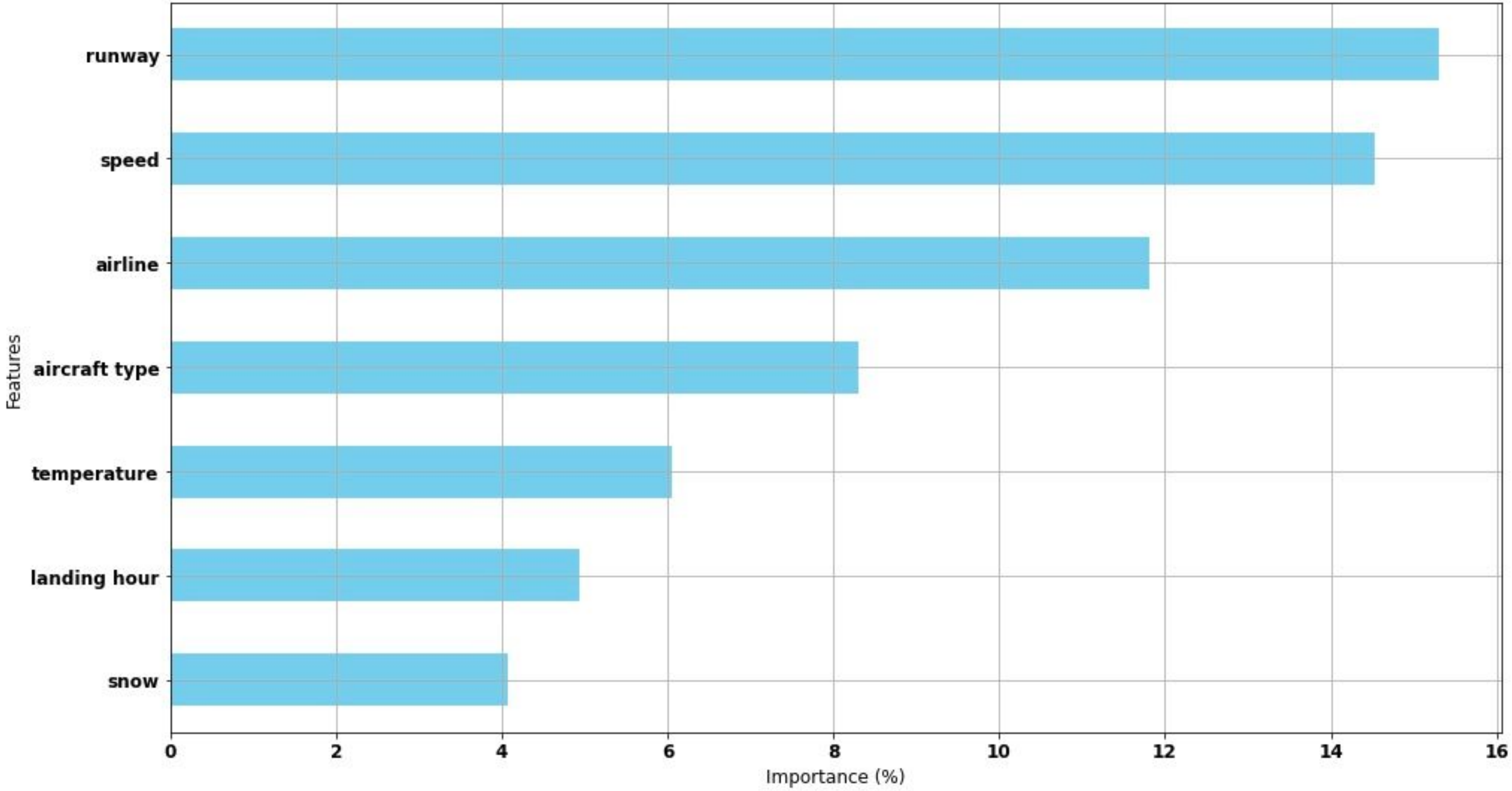
Category\Model	XG-Boost	Baseline	# Data
Super (S)	**0.452**	0.022	149
Upper Heavy (UH)	**0.525**	0.207	4946
Lower Heavy (LH)	**0.374**	0.221	794
Upper Medium (UM)	**0.513**	0.206	27542
Lower Medium (LM)	**0.367**	0.148	18262
Light (L)	**0.281**	0.251	2958

Low-cost	R^2 Score	Mean Absolute Error (s)
False	0.526	3.33
True	0.381	3.80

Features importance (2 NM)



Features importance (threshold)



Influence of additional features

	R^2 Score	Mean Absolute Error (s)
2NM with speed and leader information	0.564	3.16
2NM with speed & without leader information	0.561	3.17
Threshold with speed	0.583	3.07
Threshold without speed	0.521	3.36

- The leader position and the time elapsed since leader landing does not seem to help to improve the prediction
- On the other hand, the speed allows to increase the R^2-score by around 10%