Pre-tactical advice using machine learning for Air Traffic Flow Management delay estimation

Sergi Mas-Pujol
Luis Delgado
Paolino De Falco

11-13 May 2022

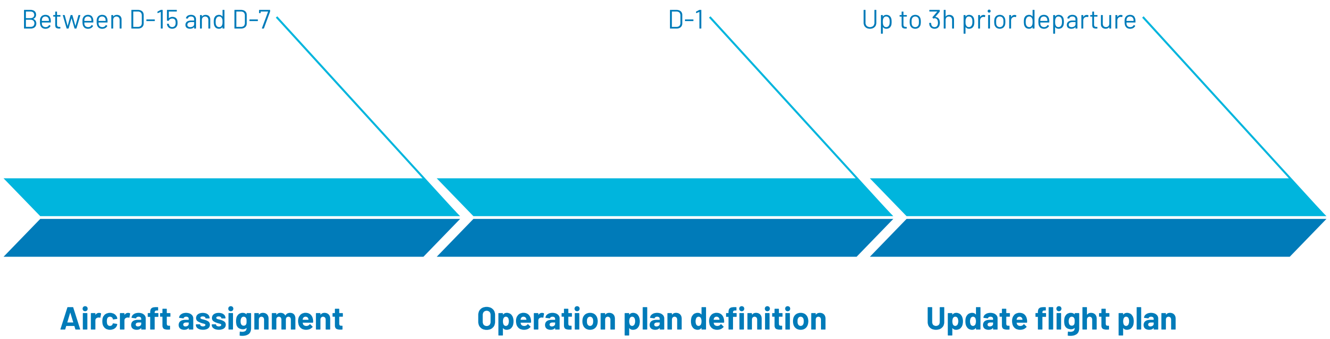
Grant: 886461

Outline

- Flight plans lifecycle
- Dispatcher3
- ATFM regulations
  - Analysis
  - Outcome of the system
- Data availability
- Methodology
- Evaluation
- Results
- Case study
- Conclusions
- Future development

Flight plans lifecycle

- Between D-15 and D-7: Aircraft assignment
- D-1 : Operation plan definition
- Up to 3 hours prior departure: Flight plans will be updated, and pre-tactical actions implementation



Dispatcher3

- Dispatcher3, a CleanSky 2 innovation action, uses machine learning techniques to support pre-departure processes

- Dispatcher3 is composed of three layers:
  - Data infrastructure
  - Predictive capabilities
  - Advice capabilities

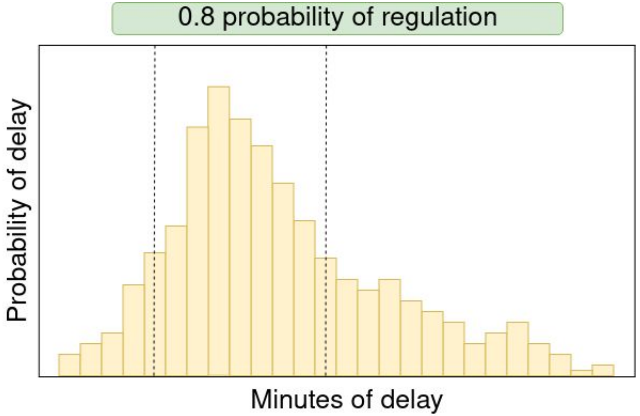
https://dispatcher3.eu/

- Flights might experience discrepancies between their plan and execution due to many factors
  - In particular, demand-capacity imbalances leading to ATFM regulations.

- Euro-centric approach
- We will focus on flights from Vueling

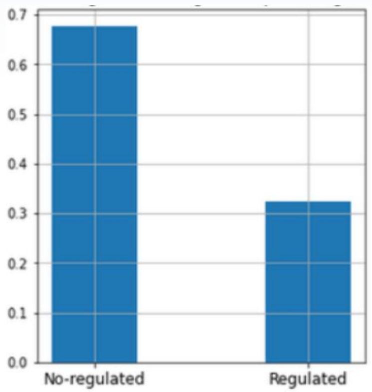
ATFM regulations

- Current work predicting ATFM regulations usually focuses on the network, or on specific OD pairs
- Early indication of potential disruptions at the flight level is important to plan and implement pre-tactical actions to minimise the potential propagation of these disruptions.

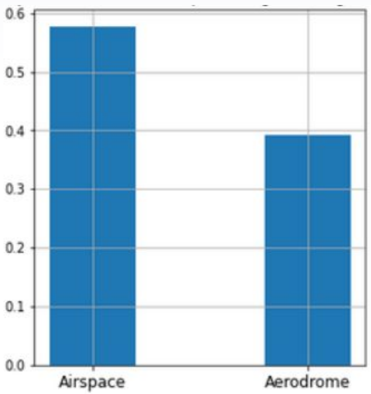


ATFM regulations - Analysis

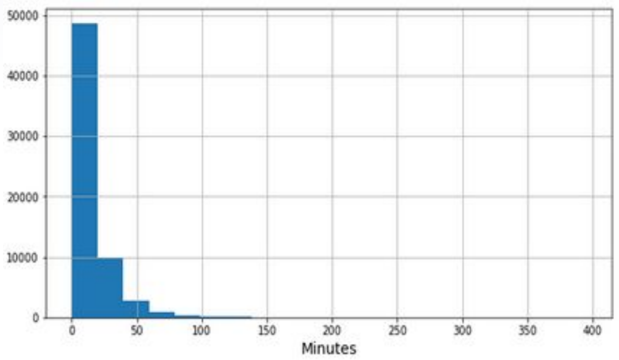
Non-regulated Vs Regulated



Type of delay

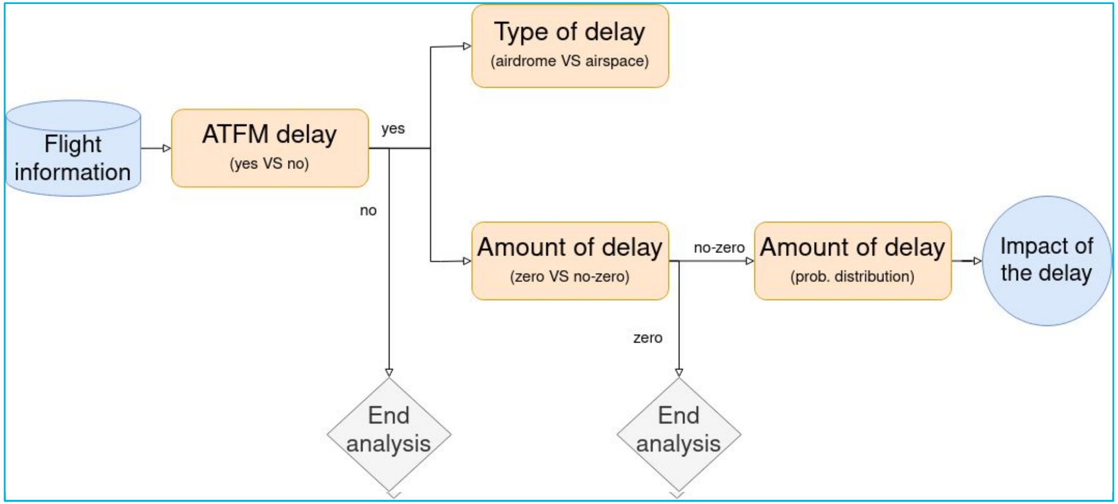


Distribution minutes of delay



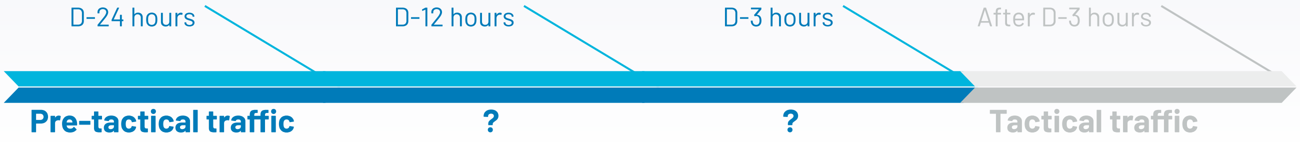
Outcome of the system

- Combination of machine learning models to create higher level interpretable predictions for D-1
  - Different levels of granularity
  - Take into account different scenarios (flexibility for D-1)



Data availability

- Data challenges: Available traffic
  - Ideal pre-tactical traffic:



- Assumption:
  - Airline has access to pre-tactical traffic
  - Static pre-tactical traffic

- Datasets used:

Data source	Description
Eurocontrol DDR (ALLFT+)	'Extension' of R&D data containing more detailed information
METAR	Forecasted historical weather information at the airports

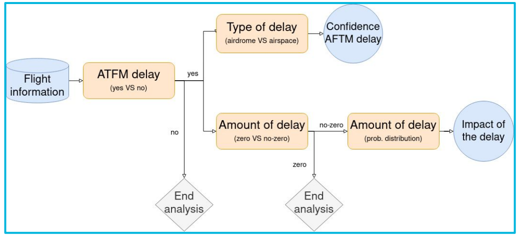
Methodology - Input features

Static features	Dynamic features
Time of departure (hourly discretization)	ATMAP score at departure/arrival airport (numerical)
Size departure airport (small, big, medium)	Temperature at departure/arrival airport (numerical)
Size arrival airport (small, big, medium)	Wind speed at departure/arrival airport (numerical)
	Visibility at departure/arrival airport (numerical)
	'Normalized' congestion at departure/arrival (in the day of operations)
	'Normalized' congestion at departure/arrival(within the hour of departure/arrival)
	Highest 'normalized' Occupancy Count (OC) within crossed sector
	Highest 'normalized' Entry Count (EC) within crossed sector

Methodology - Individual models

- Algorithms used for the individual models:

Models / Algorithms	
Probability ATFM delay (yes VS no)	Random Forest Classifier
Type of delay (airdrome VS airspace)	Decision Tree Classifier
Amount of delay (zero VS non-zero delay)	Decision Tree Classifier



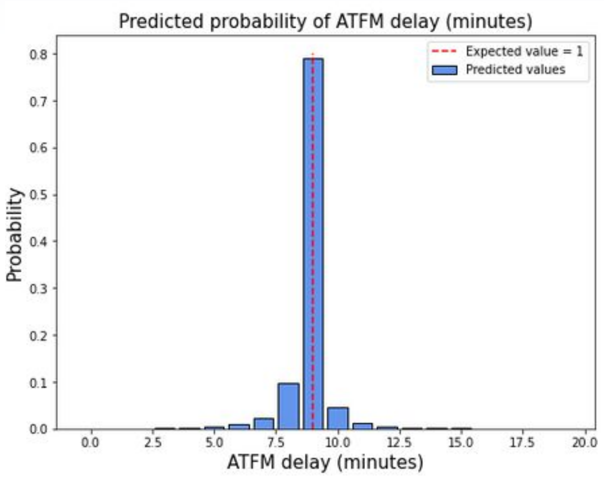
Methodology - Confidence metric

- Visual higher level interpretable information easier to be processed by the duty-manager
- Predictions inside percentile(90) of TN (or TP) -> Model sure about the prediction
- Example:
  - Prediction prob. ATFM delay = 0.87 -> Model very sure about the need of a regulation
  - Prediction type of delay = 0.59 -> Uncertain prediction for aerodrome regulation
  - Prediction amount of delay = 0.17 -> Model sure about the delay is going to be zero

Prob. ATFM
Airdrome
Zero delay

Methodology - Delay distribution

- Machine learning models produce probabilistic outputs
- Distribution of delay to better assess the impact/severity of the expected delay
  - Regression: Estimate severity (exact minutes)
  - Classification: Estimate impact (uncertainty/spread possible delay)



Evaluation - Individual models

Accuracy

Fraction of predictions right

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

Recall

Proportion of positive samples correctly identified

$$Recall = \frac{TP}{TP+FN}$$

Precision

Proportion of positive identification correctly

$$Precision = \frac{TP}{TP+FP}$$

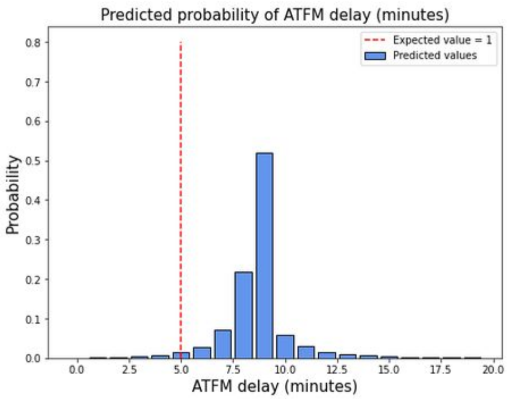
F1-score

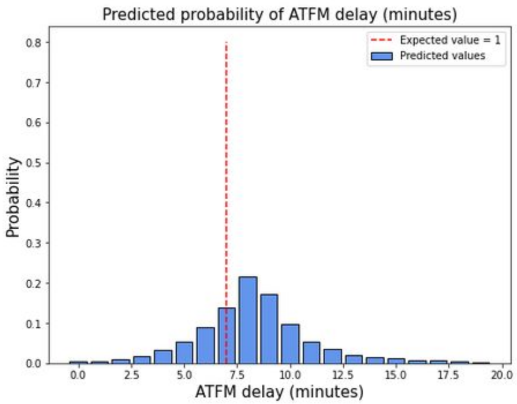
Harmonic mean of the precision and recall

$$F1\ score = 2\frac{Precision*Recall}{Precision+Recall}$$

Evaluation - Delay distribution

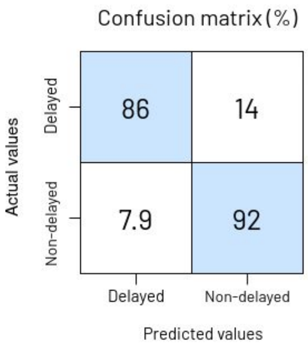
- How close is the prediction to the actual ATFM delay?
  - Compute the difference of minutes between ground-truth and the expected value from distribution
- How sure is the model about the expected delay?
  - Compute dispersion of predicted values
  - The more sure the model's prediction, the fewer bars will be present on the chart



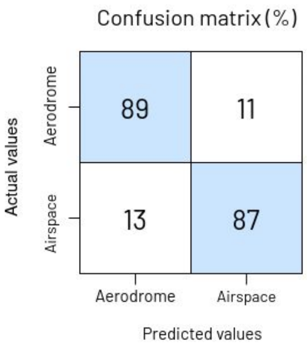


Results - Individual models

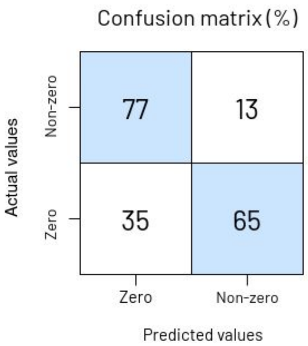
Prob. ATFM delay (yes vs no)	
Accuracy	0.88
F1-score	0.87



Type of delay (Airdrome vs Airspace)	
Accuracy	0.87
F1-score	0.86



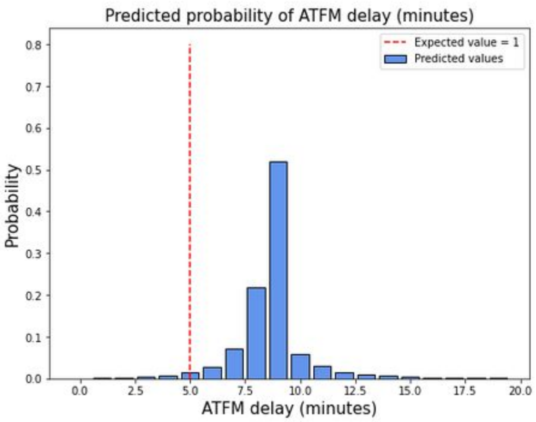
Amount of delay (zero vs non-zero delay)	
Accuracy	0.71
F1-score	0.66

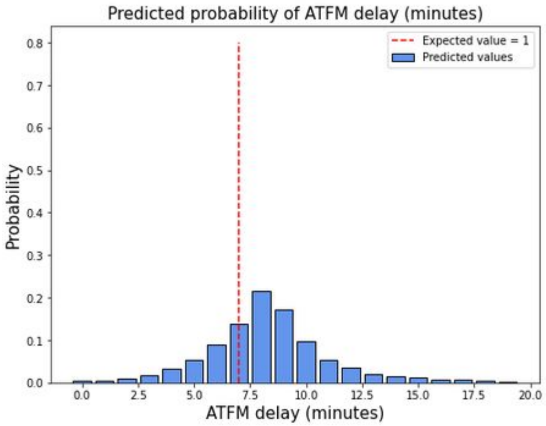


Results - Delay distribution

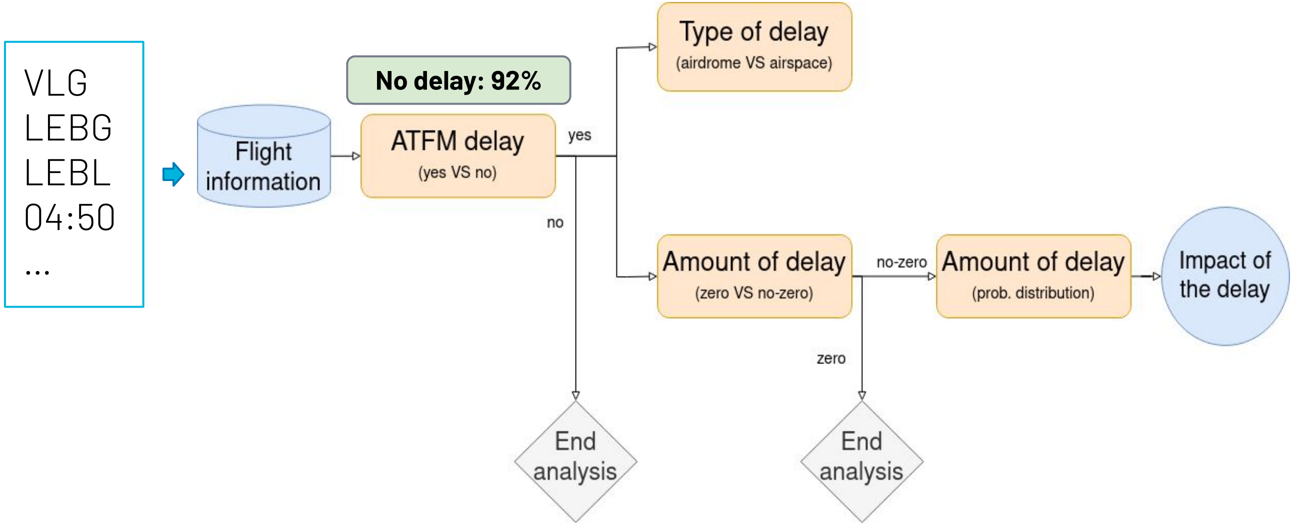
- Actual delay VS Predicted delay: 9,14 minutes
  - Mean difference between actual delay and expected value from the distribution

- Average dispersion of the prediction: 22,35 minutes



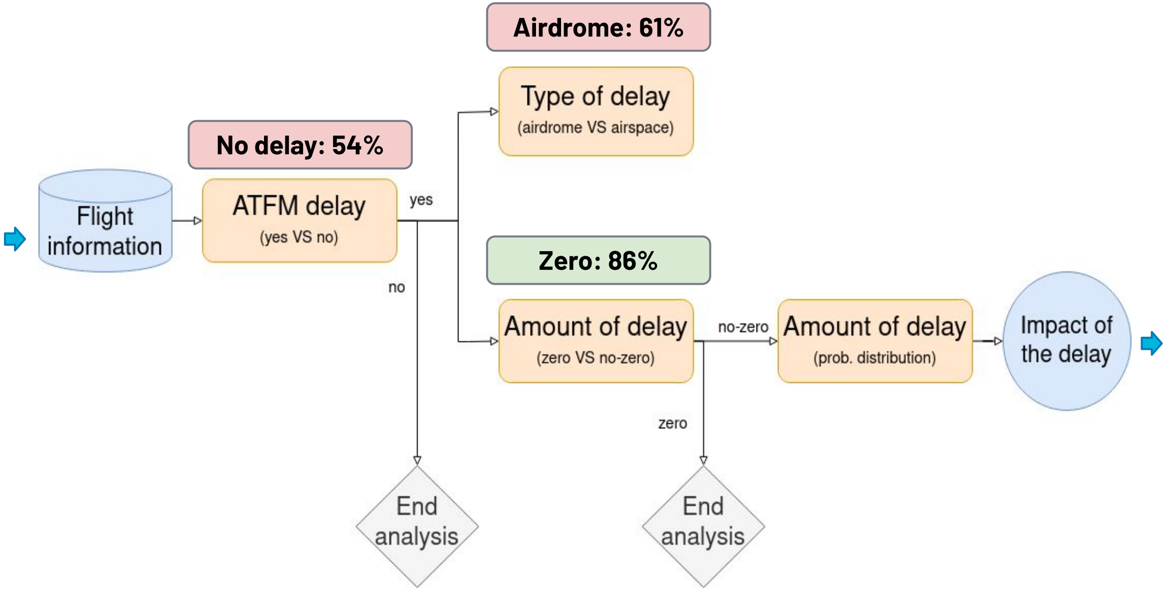


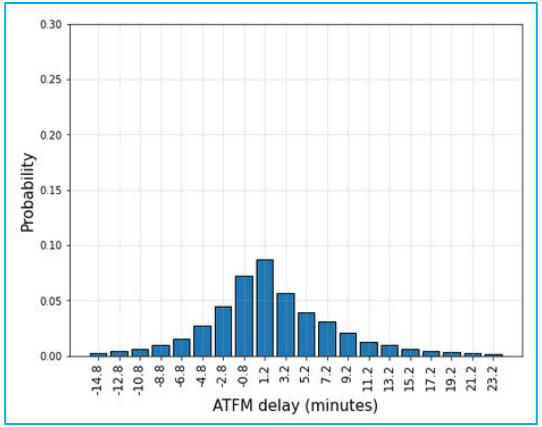
Case study - No ATFM delay



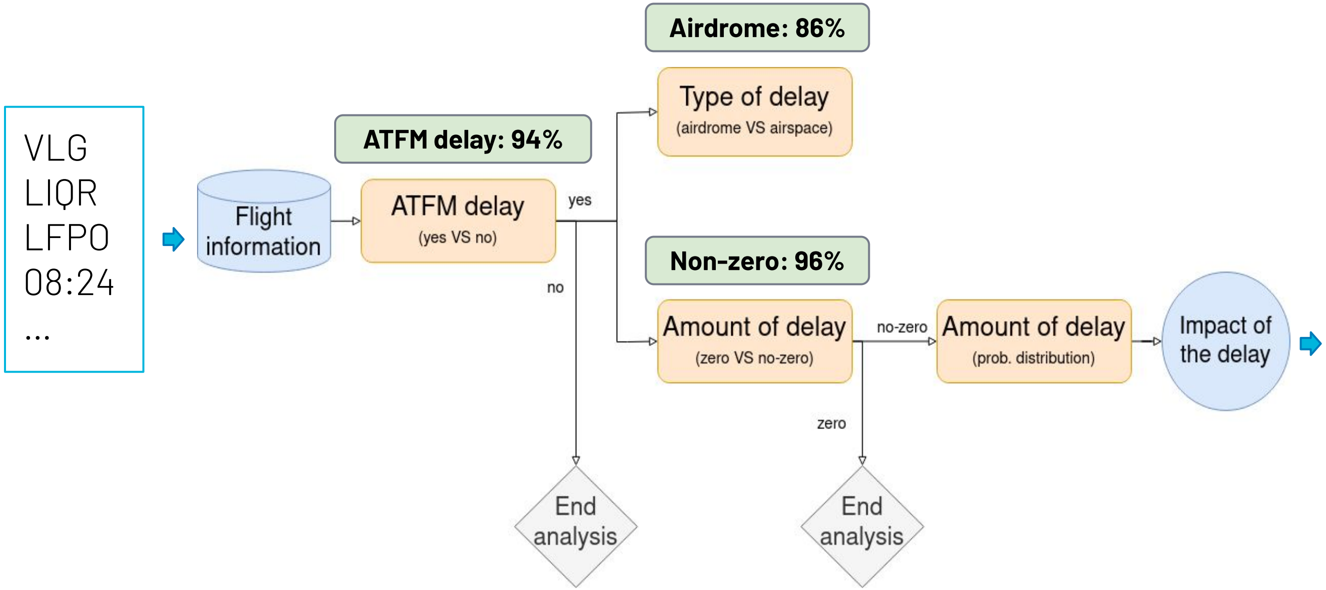
Case study - No ATFM delay

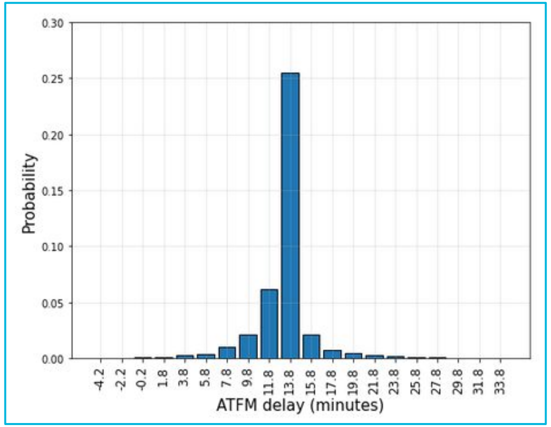
VLG
LEBL
LFPO
05:45
...



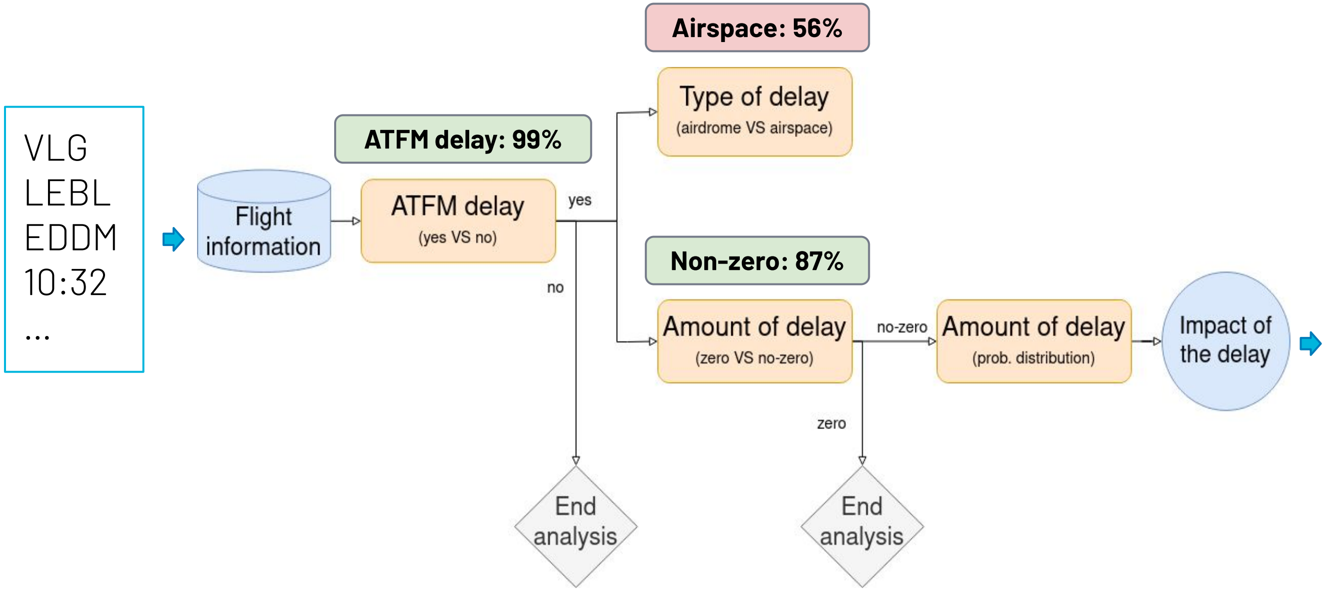


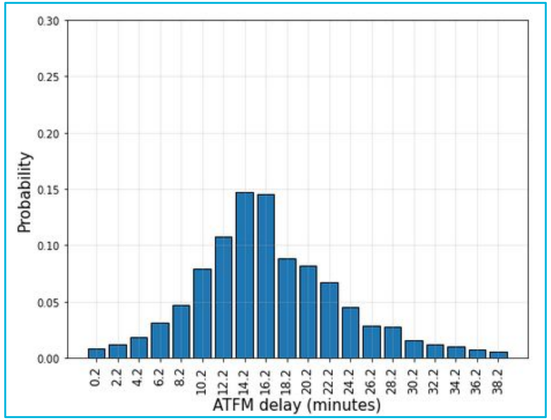
Case study - ATFM delay





Case study - ATFM delay





Conclusions

Benefits:

- Models can be used to identify ATFM regulation pre-tactically
- Individual models between 70% and 90% accuracy
- Impact/severity can be assessed with distribution of possible delay (mean error of 9 minutes with dispersion of 22 minutes)
- Models can be improved even further

Drawbacks:

- Assumed airlines have access to network information (M1 traffic)
- Assumed a static pre-tactical flight plan has been defined for each flight
- The less accurate individual model is the zero VS non-zero delay

Future development

- Feature selection analysis (e.g PCA, SHAP values)
- Fine-tune less accurate models
- Release predictions according to specific time horizons
- Integrate other data sources (e.g. network weather information)
- Provide additional information about the network status
- Validate the proposed representation of the predicted information with experts in the field

THANKS!

You can find me at:

- sergi.mas.pujol@upc.edu