Prediction of reactionary delay and cost using machine learning

Dr. Paolino De Falco
Dr. Luis Delgado

UNIVERSITY OF WESTMINSTER

PILOT3

DISPATCHER3

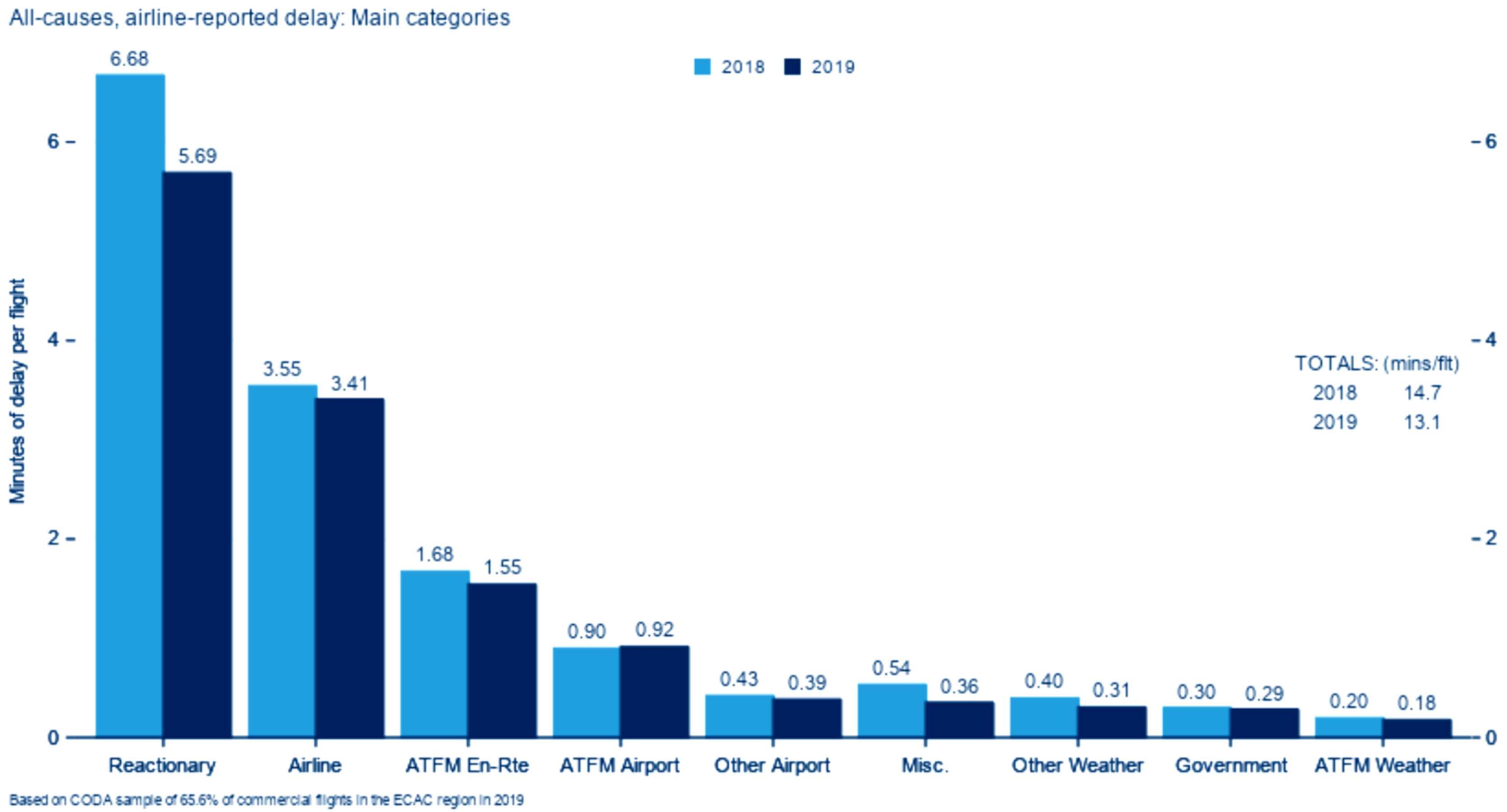
Outline

UNIVERSITY OF WESTMINSTER

- An overview on air traffic and delay
- Reactionary and rotational delay
- An approach to model reactionary delay using machine learning
- Dispatcher3 and Pilot3: two projects benefitting from this approach
- Future developments of the approach

Traffic and Average Delay per Flight Overview

UNIVERSITY OF WESTMINSTER



https://www.eurocontrol.int/sites/default/files/2020-04/eurocontrol-coda-digest-annual-report-2019.pdf

Reactionary delay

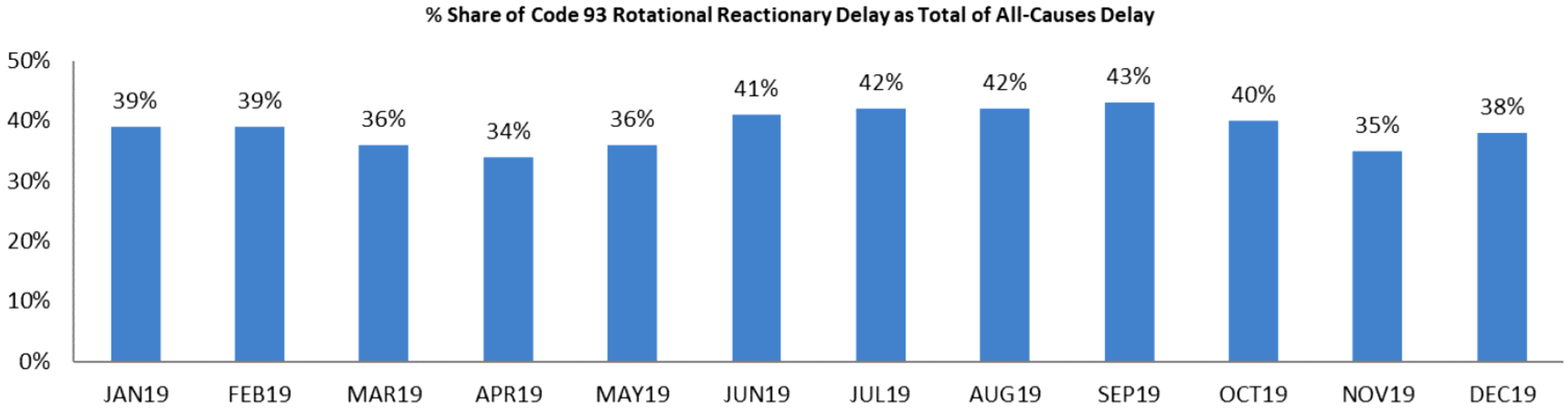
UNIVERSITY OF WESTMINSTER

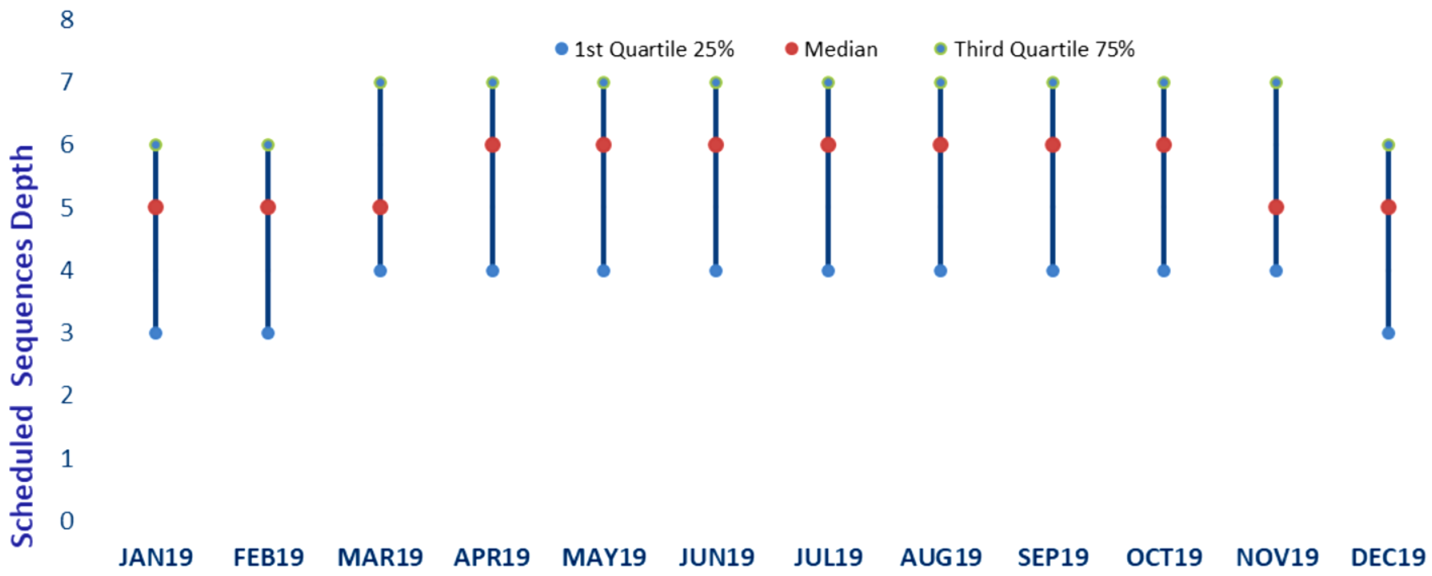
- **91 (RL)**: Passenger or Load Connection, awaiting load or passengers from another flight. Protection of stranded passengers onto a new flight.
- **92 (RT)**: Through Check-in error, passenger and baggage
- **93 (RA)**: Aircraft rotation, late arrival of aircraft from another flight or previous sector
- **94 (RS)**: Cabin crew rotation
- **95 (RC)**: Crew rotation, awaiting crew from another flight (flight deck or entire crew)
- **96 (RO)**: Operations control, rerouting, diversion, consolidation, aircraft change for reasons other than technical

https://www.eurocontrol.int/sites/default/files/2020-04/eurocontrol-coda-digest-annual-report-2019.pdf

Rotational reactionary delay

UNIVERSITY OF WESTMINSTER



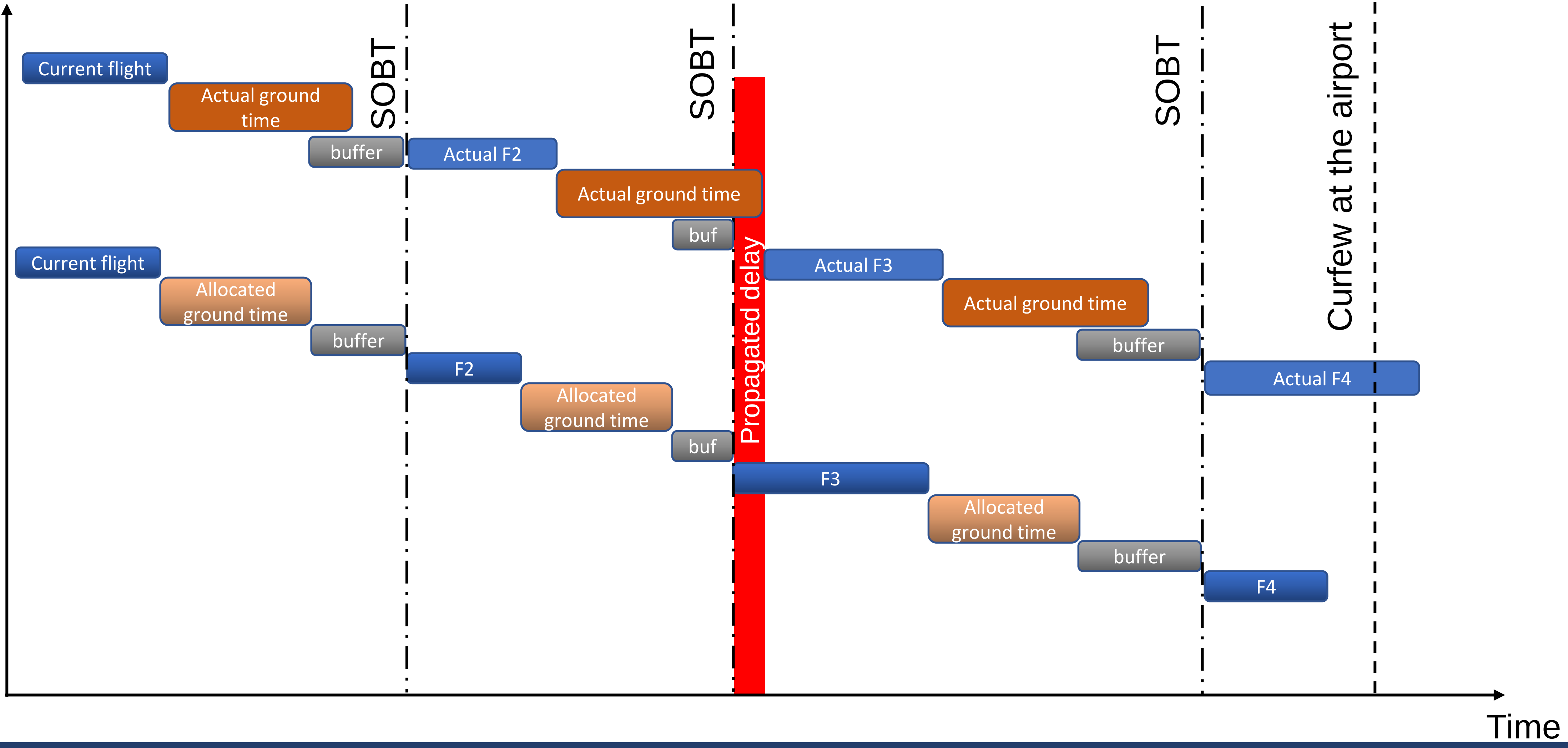


https://www.eurocontrol.int/sites/default/files/2020-04/eurocontrol-coda-digest-annual-report-2019.pdf

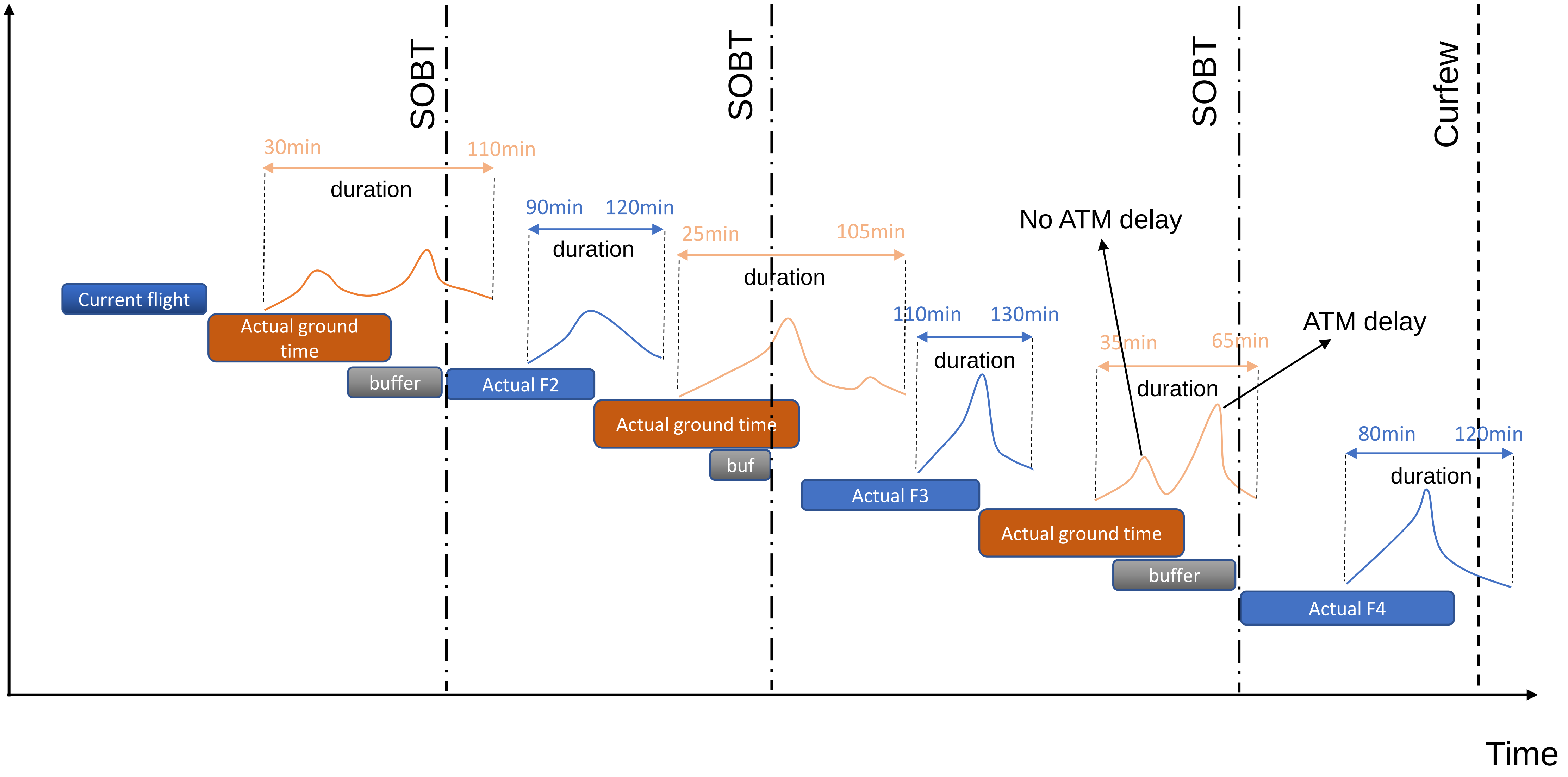
Aircrafts operations

UNIVERSITY OF WESTMINSTER

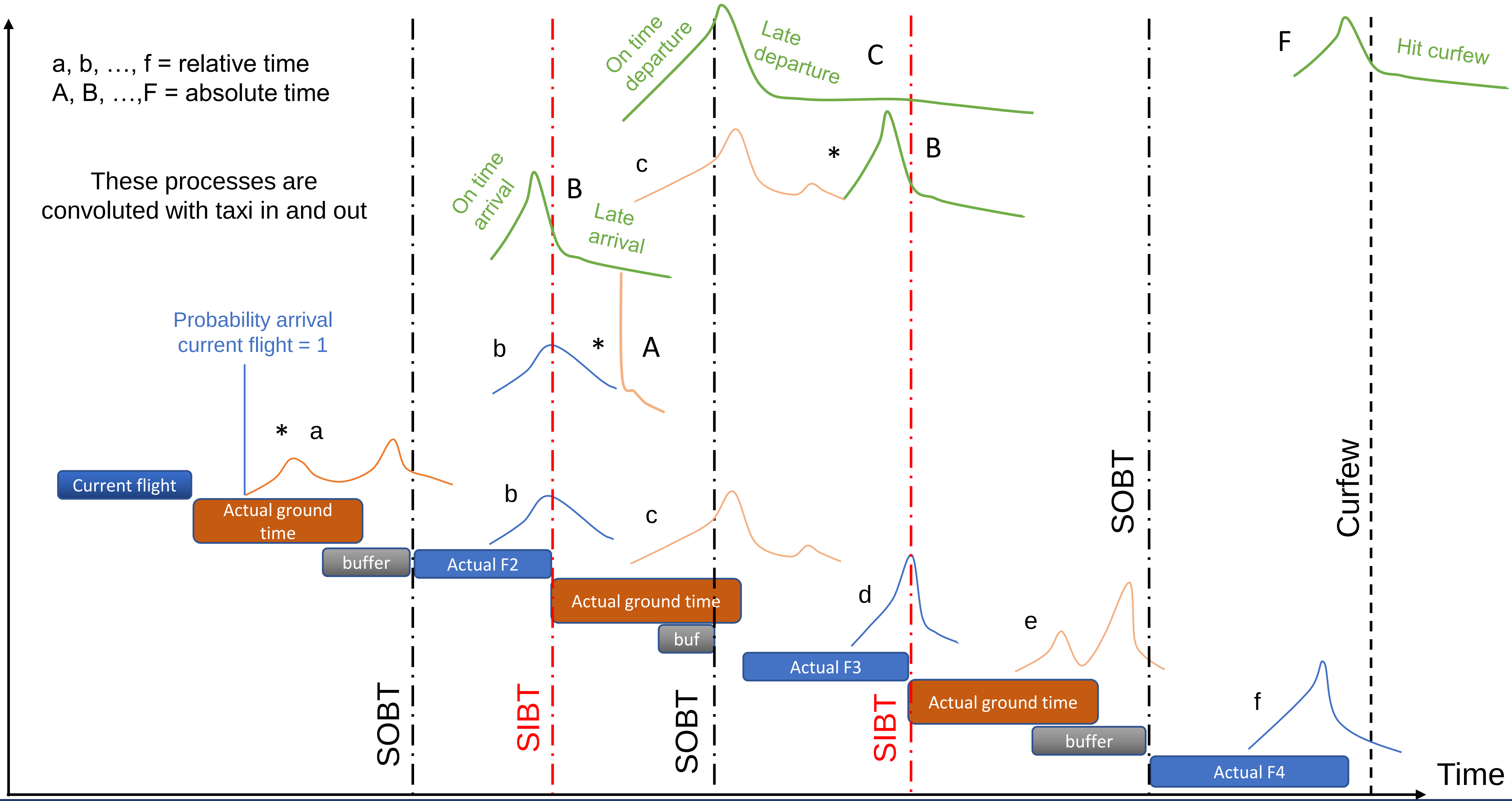
Aircraft leaves the parking position (off-block time) → Taxi out → Flight → Landing → Taxi in → Ground time (turnaround + buffer)



Aircrafts operations



Aircrafts operations



Modelling approach to predict reactionary delay and related costs

UNIVERSITY OF WESTMINSTER

Flight duration (gate-to-gate) → Time on the ground (at the gate) → Flight duration (gate-to-gate) →

4 time horizons

15h before departure — 6h before departure — 3h before departure — Departure time

Outcome

- Reactionary delay
- Costs associated to reactionary delay

Training of ML models using data from September 2018

Selection of features for ML models

UNIVERSITY OF WESTMINSTER

Static features	Dynamic features
Time_departure: morning, afternoon, evening	ATMAP score at departure airport
Airline_type: REG, LCC,FSC, Other	ATMAP score at arrival airport
Size airport departure: small, medium, big	Wind speed at departure airport
Size aircraft: low, medium, high, jumbo	Wind speed at arrival airport
Congestion at departure during the day of operations	Temperature at departure airport
Congestion at arrival during the day of operations	Temperature at arrival airport
Hub (yes/no)	Landing direction
Regulations (yes/no)	Average wind along trajectory
Great circle distance	Congestion_arrival: how many planes are landing in the hour of operations
Direction of flight (e.g., North-West, etc…)	Congestion_departure: how many planes are departing in the hour of operations
Size airport arrival: small, medium, big	

Features used in both ground and flight models

Features used in flight model only

Features used in ground model only

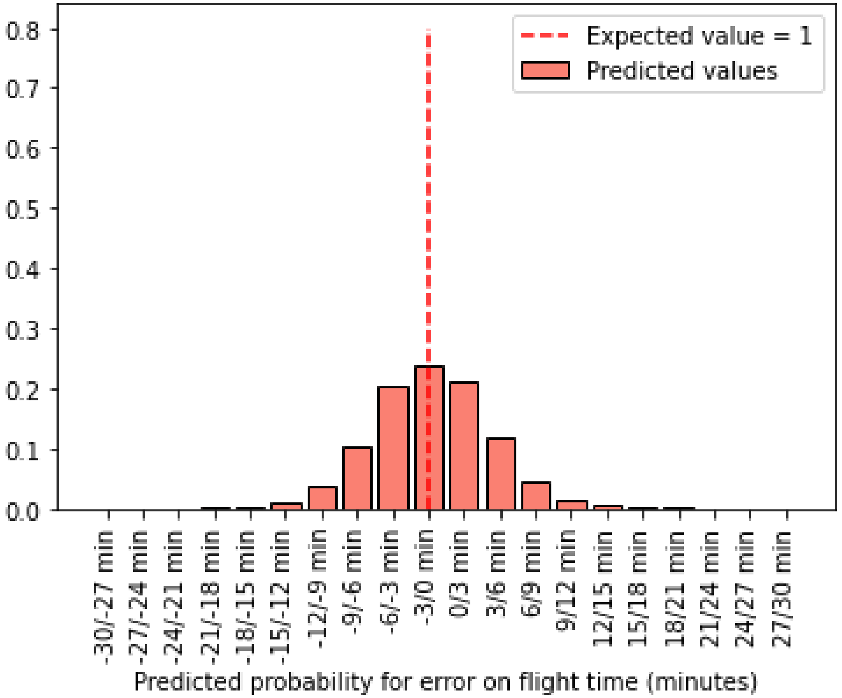
‘Used’ and ‘potential’ data

UNIVERSITY OF WESTMINSTER

Data Source	Description
ADS-B (Automatic Dependent Surveillance–Broadcast)	Technology allowing to determine aircraft position via satellite navigation or other sensors. Data is periodically broadcasted enabling to track aircrafts
ERA5 (ECMWF (European Centre for Medium-Range Weather Forecasts) Reanalysis 5th Generation)	ERA5 provides hourly estimates of a large number of atmospheric, land and oceanic climate variables
Eurocontrol R&D	The archive contains a list of flights with key data for each flight; the airspace structure that applied at the time; filed and actual flight trajectory
Eurocontrol DDR (demand data repository)	‘Extension’ of R&D data containing more detailed information
METAR (Meteorological Aerodrome Report)	Weather information (current and historical) at the airports
SIGMET (Significant Meteorological Information)	A SIGMET provides information issued by a Meteorological Watch Office (MWO) concerning the occurrence or expected occurrence of en-route weather that may affect the safety of aircraft operations.
TAF (Terminal Aerodrome Forecasts)	Forecasts weather at the airports
FDM (Flight Data Monitoring)	FDM contains analyses on data generated by an aircraft in order to improve flight safety and increase overall operational efficiency

Modelling approach to predict flight duration

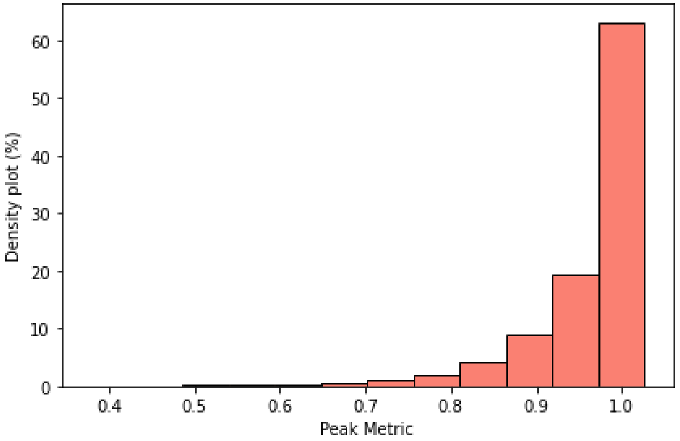
1. Extract the average flight time (take-off to landing) from data
2. Calculate the 'error' as (average flight time – actual flight time)
3. Predict this error as a discrete distribution with a classifier (problem of binning)

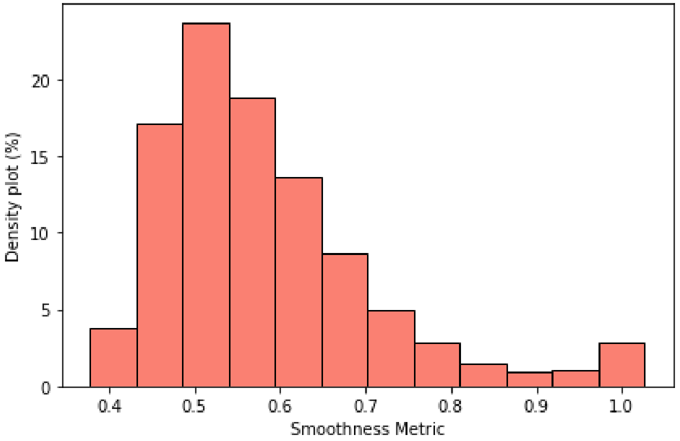


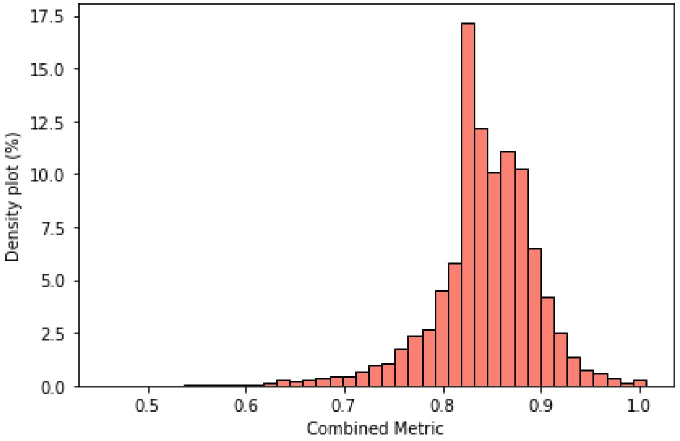
Classification model to predict the error as a difference with respect to the average flight duration

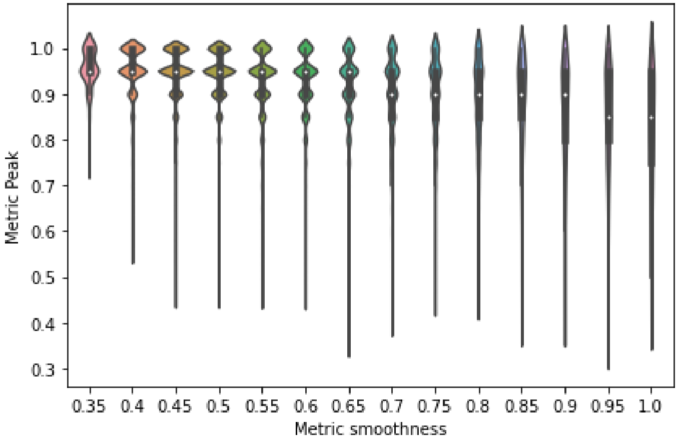
Modelling approach to predict flight duration

Classification model to predict the error as a difference with respect to the average flight duration









Peak metric

$$\left(1 - \frac{\text{distance in bins between predicted and expected peak values}}{\text{total number of bins}}\right)$$

Smoothness metric

$$\frac{\text{number of bins } (\neq 0)}{\text{total number of bins}}$$

Combined metric

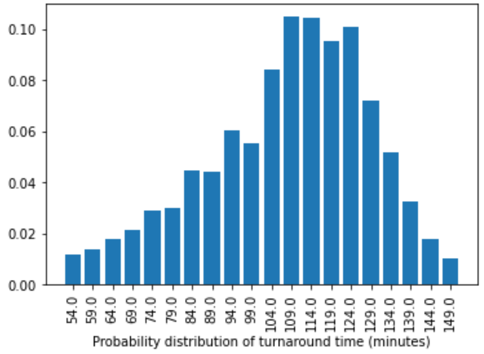
$$(\text{Peak metric} * \text{w1}) + (\text{Smoothness metric} * \text{w2})$$

w1 = 0.75 if Peak metric > 0.5 otherwise w1 = 0.25

w2 = 0.25 if Peak metric < 0.5 otherwise w2 = 0.75

Modelling approach to predict ground time

UNIVERSITY OF WESTMINSTER



$\times\ p_t$

ML classifier for probability of ATM delay

$p_t = 0,7$

ML classifier for turnaround time (using data from **ATM delayed flights**)

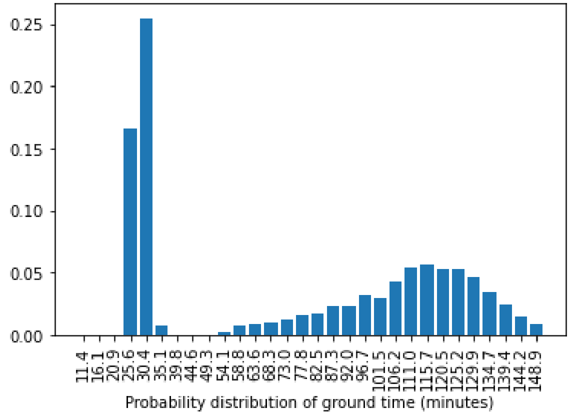
$p_m = 0,3$

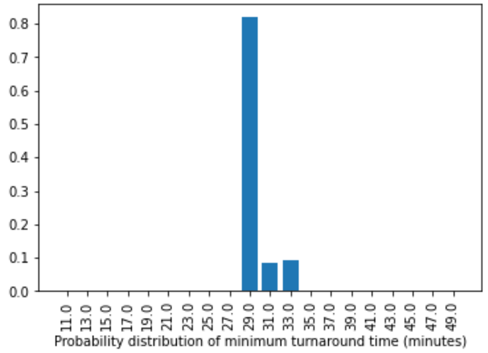
ML classifier for minimum turnaround time (using data from non affected by ATM delayed flights)

sum

sum

Final time on the ground

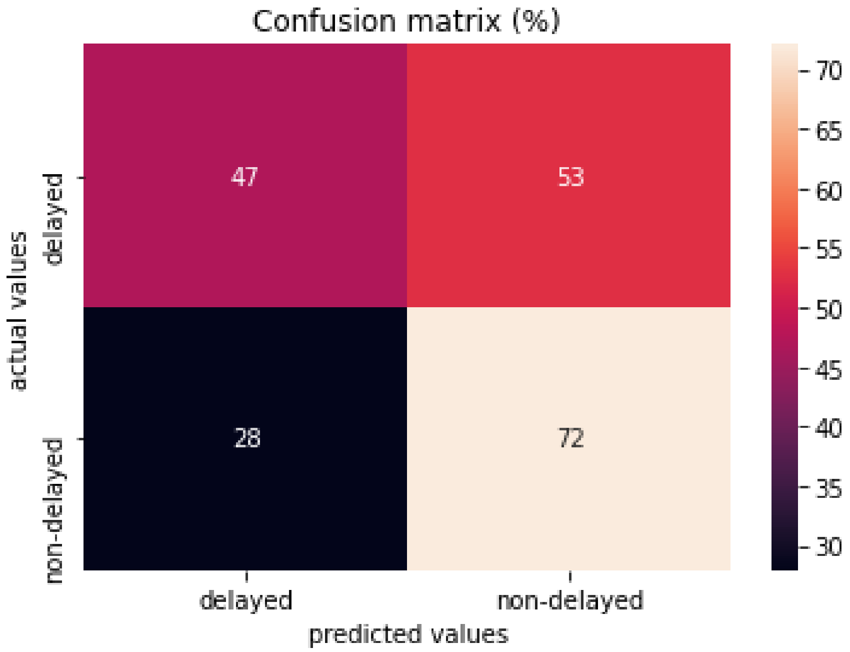




$\times\ p_m$

Modelling approach to predict probability of ATM delay

ML classifier for probability of having ATM delay



Balanced dataset

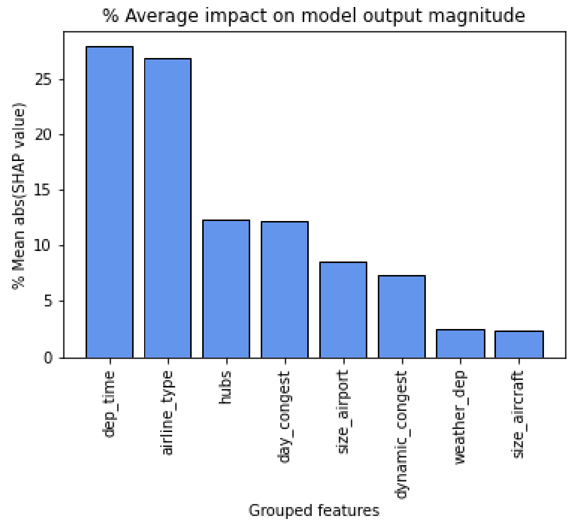
Example of output

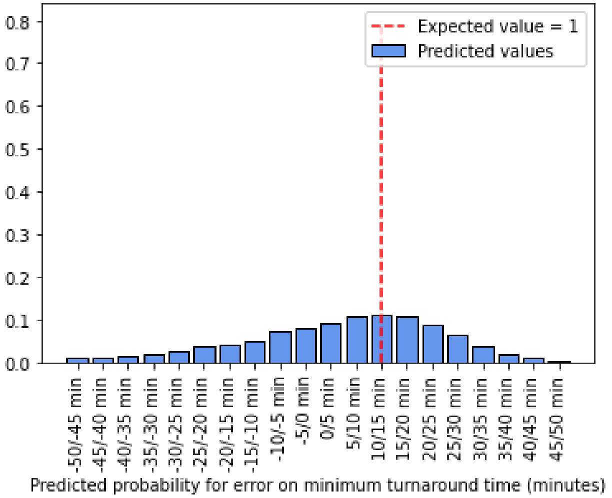
Probability ATM delay	Probability Non-delayed
0,6	0,4

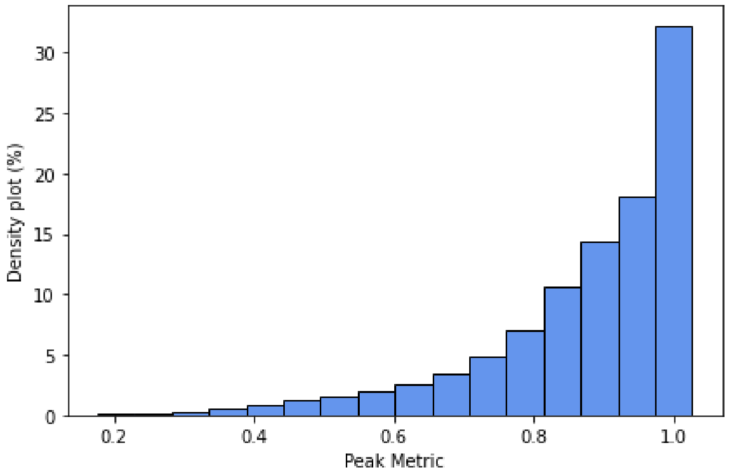
Modelling approach to predict turnaround time (ATM delay)

UNIVERSITY OF WESTMINSTER

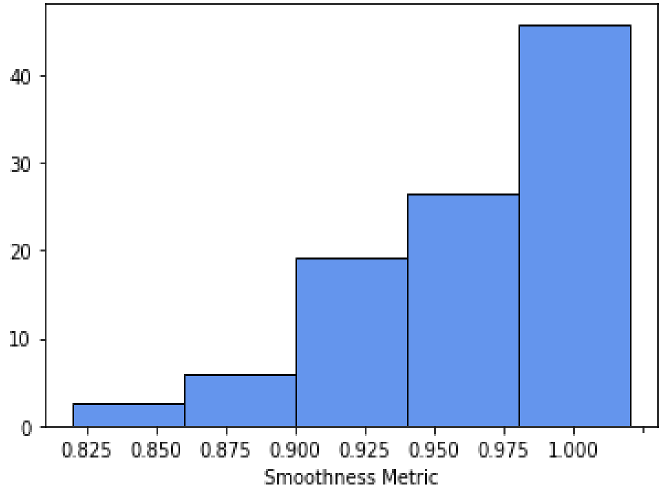
Regression + classification models to characterise each prediction as a probability distribution

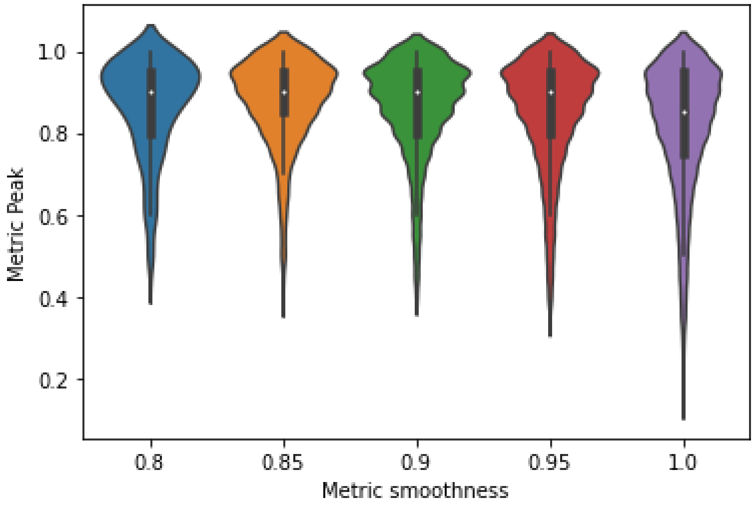


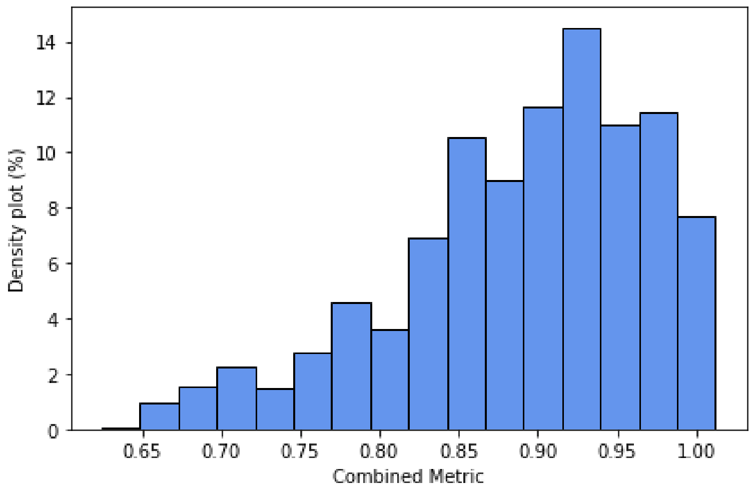




The performance of the 4 models in terms of ‘combined metric’ is **0,88**







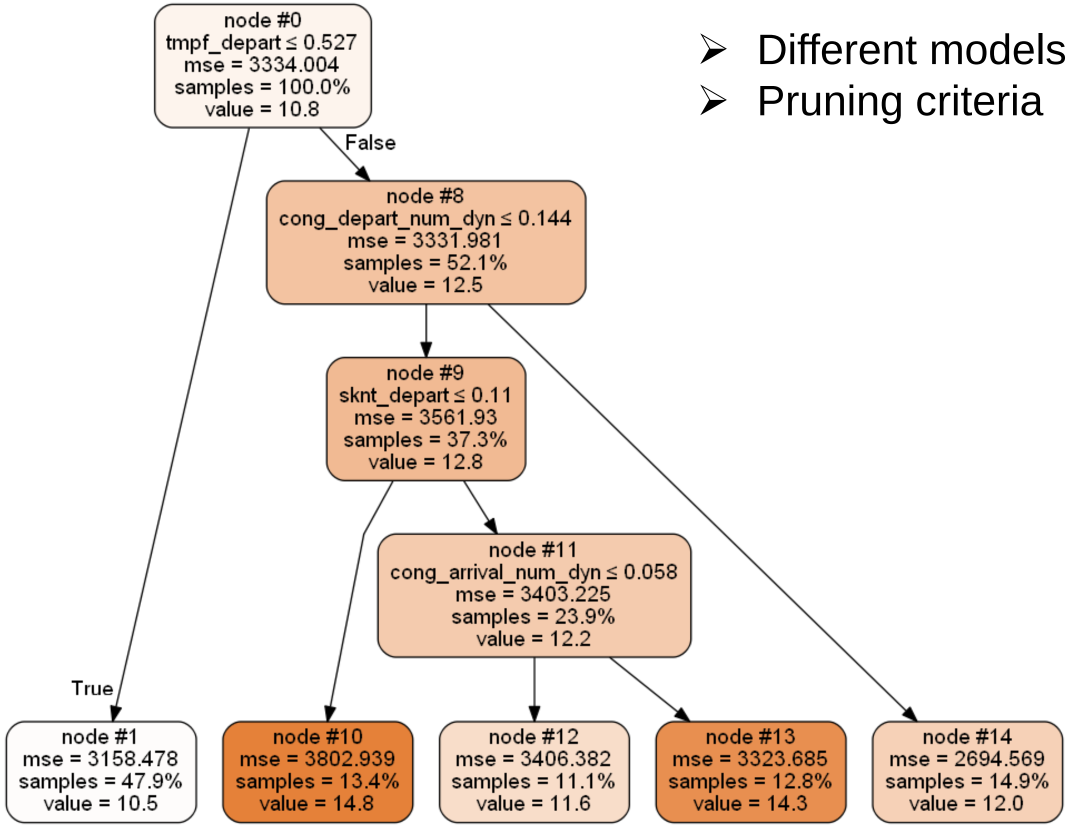
Modelling approach to predict minimum turnaround time

UNIVERSITY OF WESTMINSTER

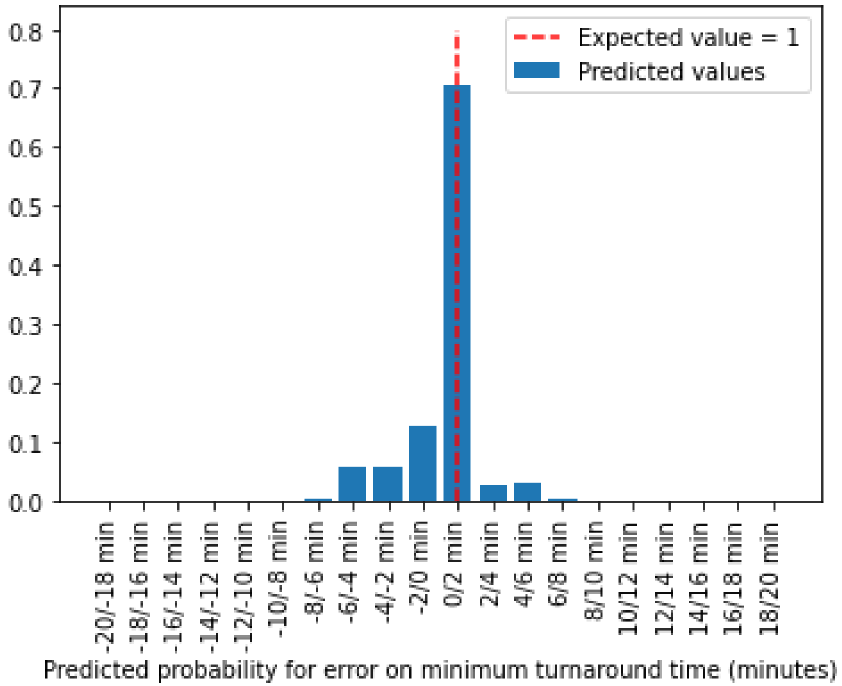
Regression decision tree for **labelling** the output as minimum turnaround time from turnaround time

Regression decision tree for 'clustering'/**'regrouping**'

Classification ML model to characterize the error when predicting the output



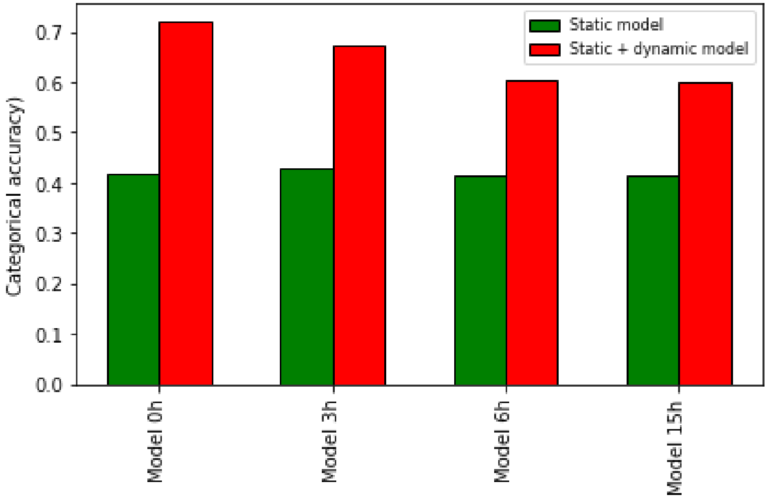
- ➢ Different models for each wake size
- ➢ Pruning criteria

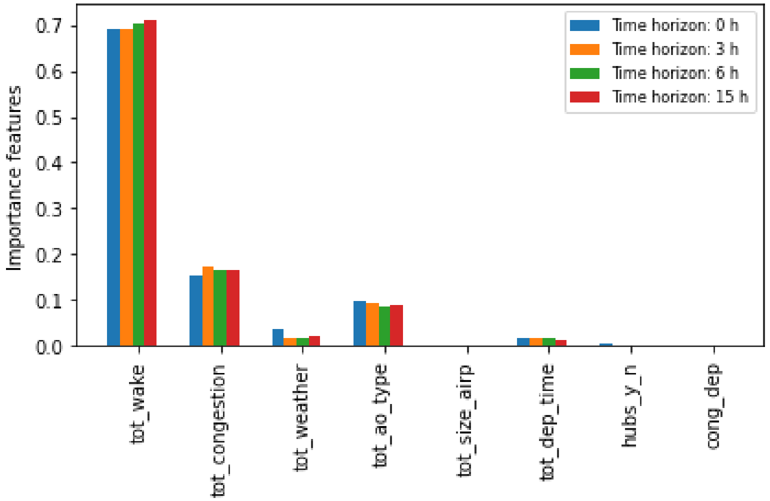


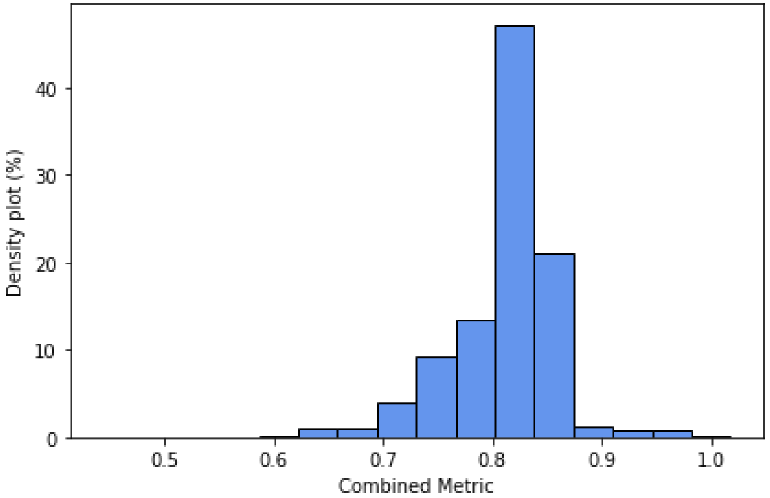
NB. Minimum turnaround time is calculated as 2% percentile of turnaround distributions

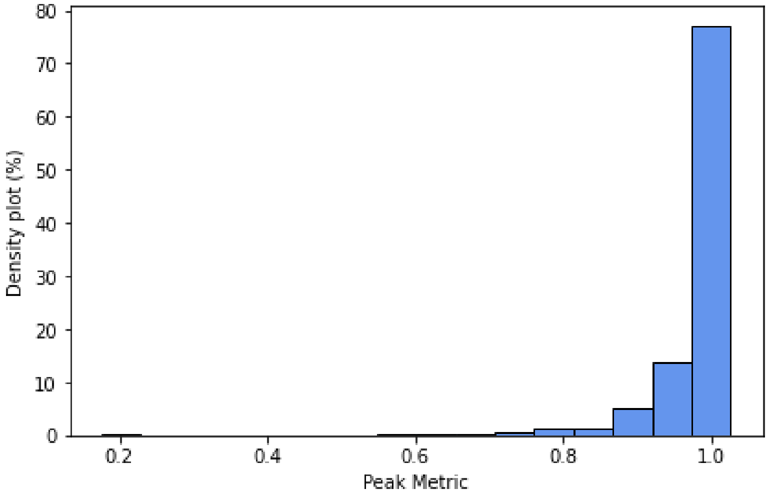
Modelling approach to predict minimum turnaround time

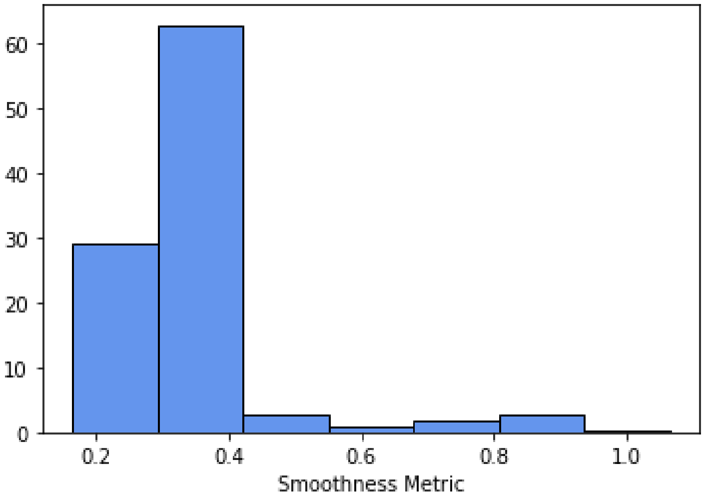
UNIVERSITY OF WESTMINSTER

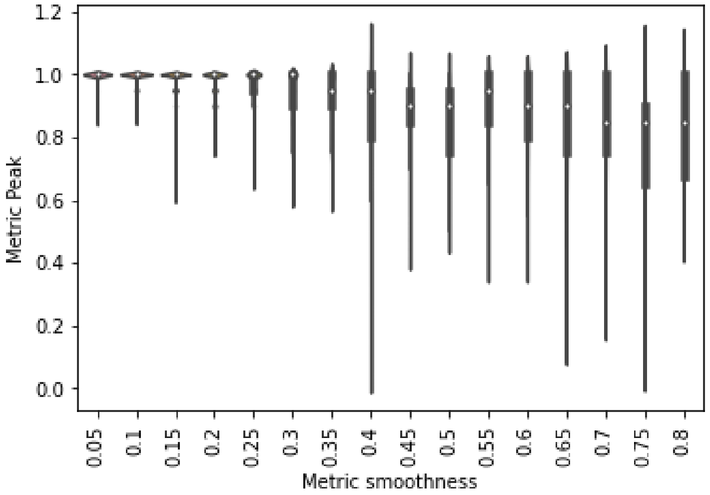








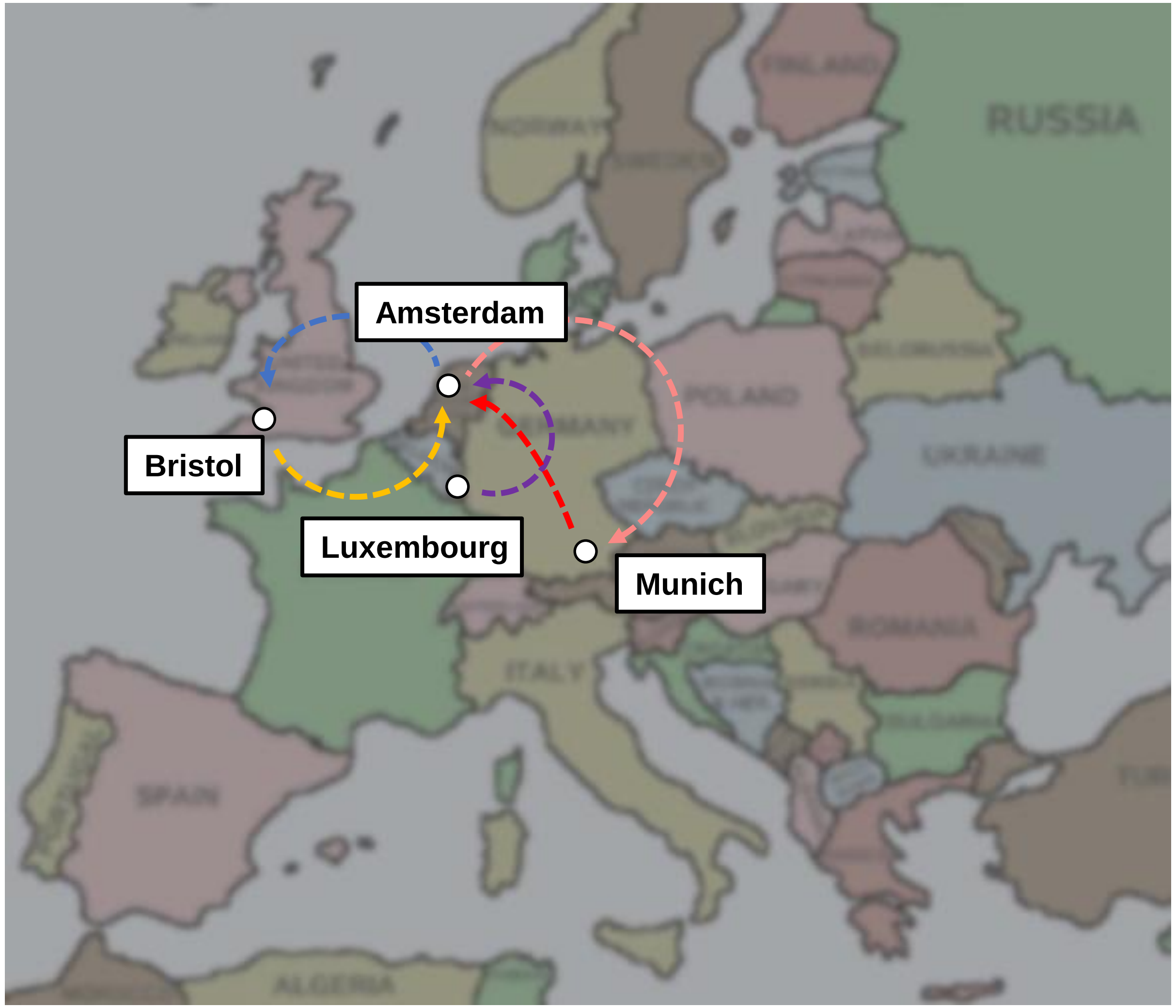




The performance of the 4 models in terms of ‘combined metric’ is **0,79**

An experiment

UNIVERSITY OF WESTMINSTER

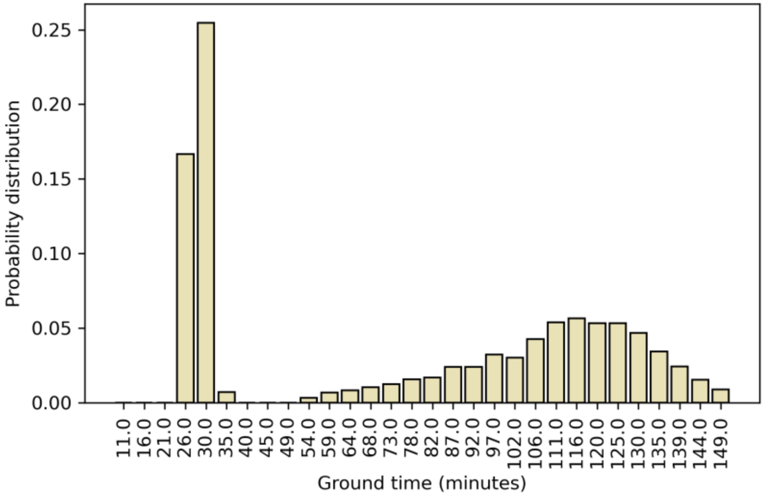


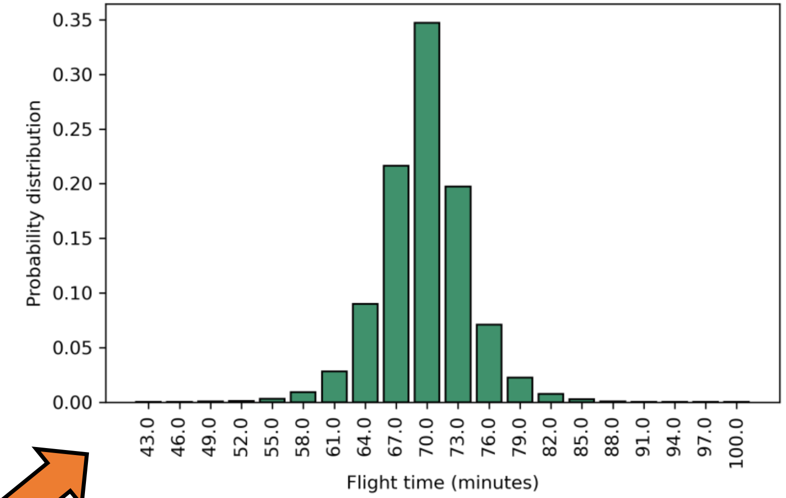
Origin	Destination	SOBT	SIBT
1. ELLX	EHAM	08:55am	10:00am
2. EHAM	EGGD	10:40am	11:50am
3. EGGD	EHAM	12:25pm	13:45pm
4. EHAM	EDDM	16:00pm	17:25pm
5. EDDM	EHAM	18:05pm	19:40pm

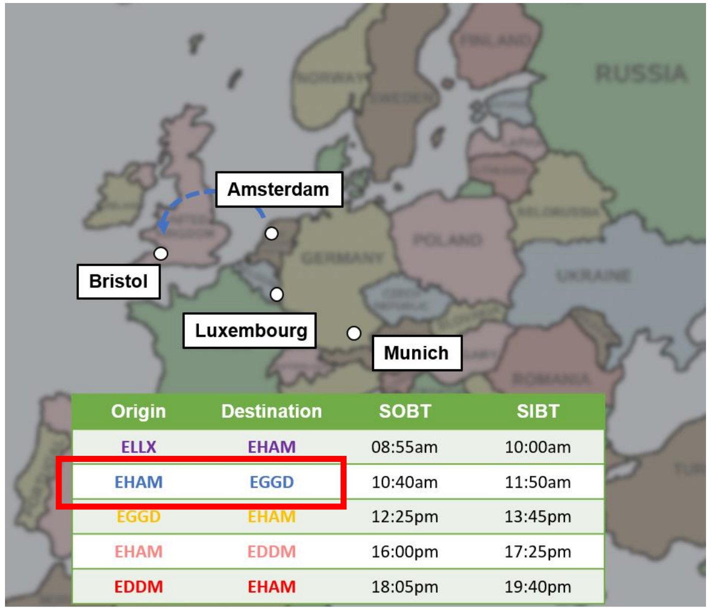
An experiment

UNIVERSITY OF WESTMINSTER

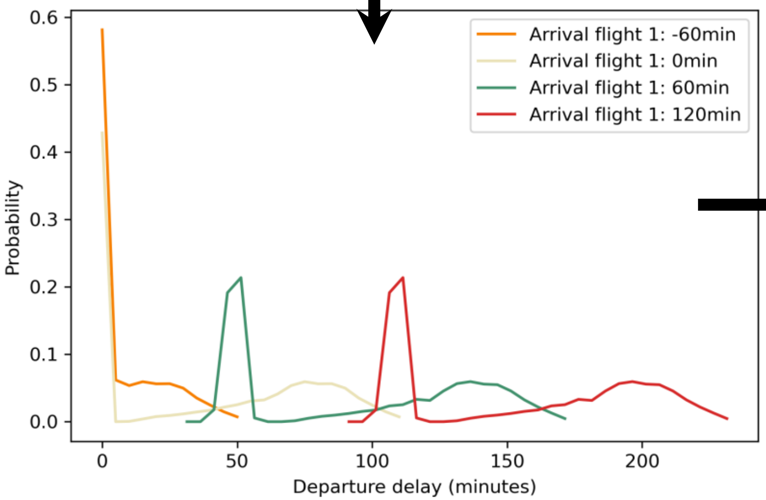
Flight 2: Amsterdam - Bristol

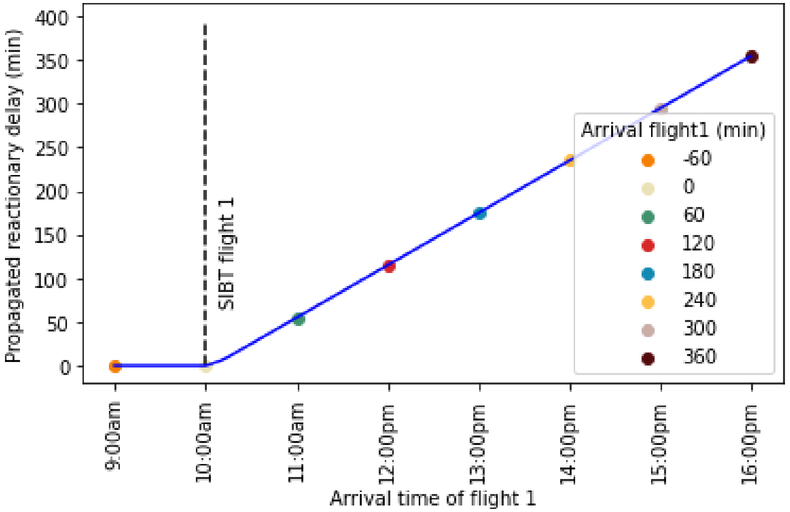


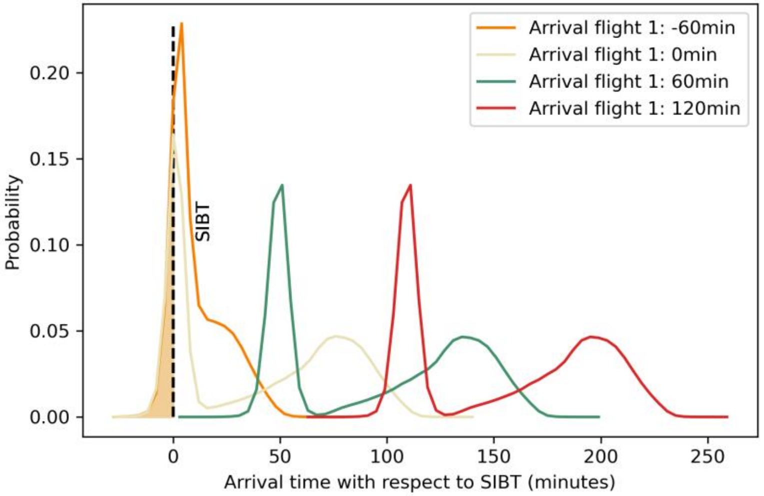




Origin	Destination	SOBT	SIBT
ELLX	EHAM	08:55am	10:00am
EHAM	EGGD	10:40am	11:50am
EGGD	EHAM	12:25pm	13:45pm
EHAM	EDDM	16:00pm	17:25pm
EDDM	EHAM	18:05pm	19:40pm



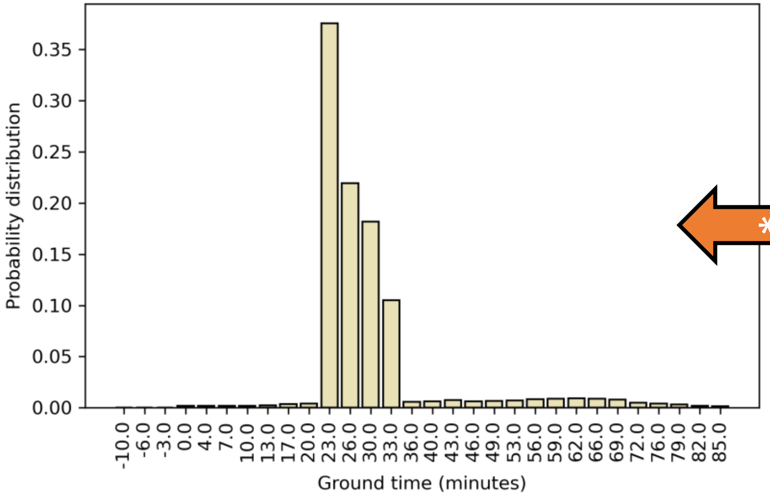


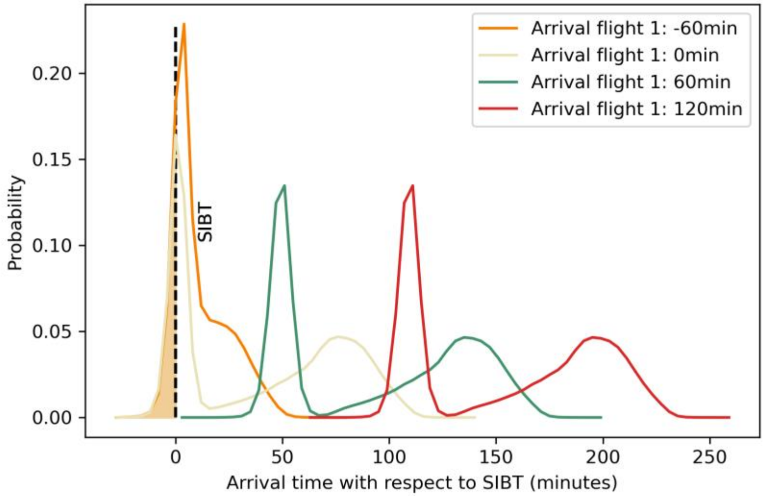


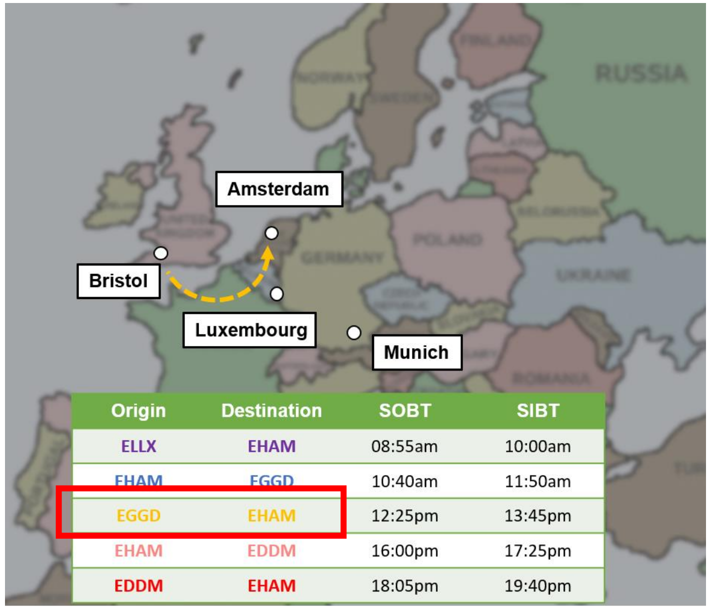
An experiment

UNIVERSITY OF WESTMINSTER

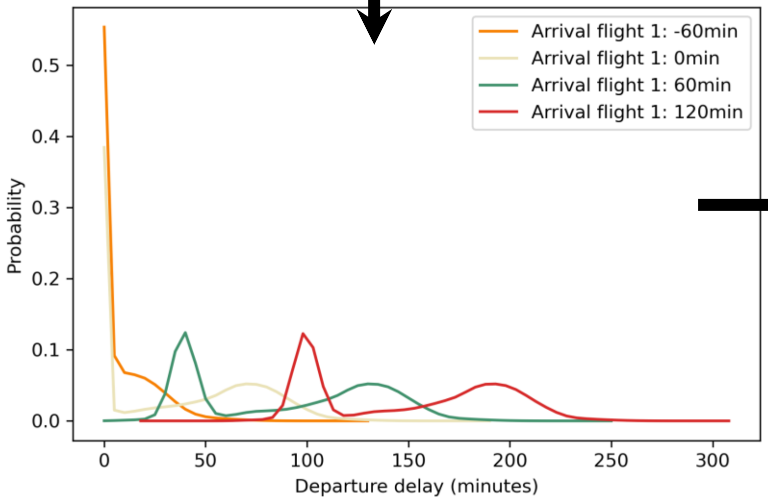
Flight 3: Bristol – Amsterdam

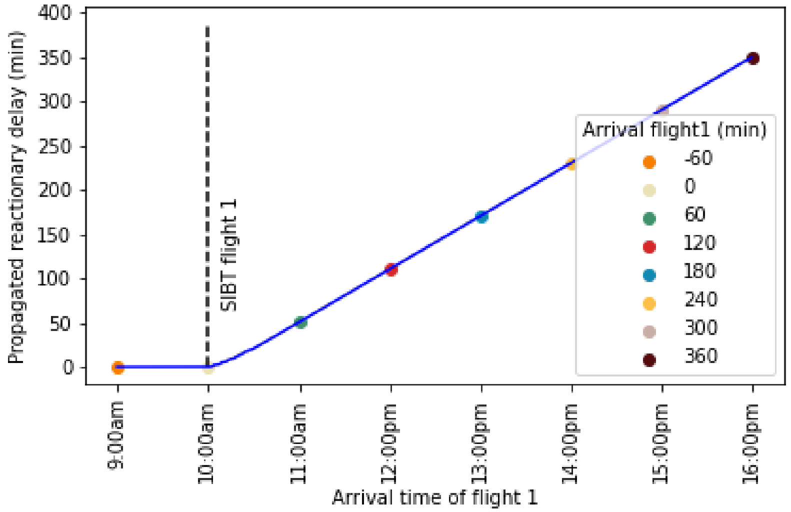






Origin	Destination	SOBT	SIBT
ELLX	EHAM	08:55am	10:00am
EHAM	EGGD	10:40am	11:50am
EGGD	EHAM	12:25pm	13:45pm
EHAM	EDDM	16:00pm	17:25pm
EDDM	EHAM	18:05pm	19:40pm

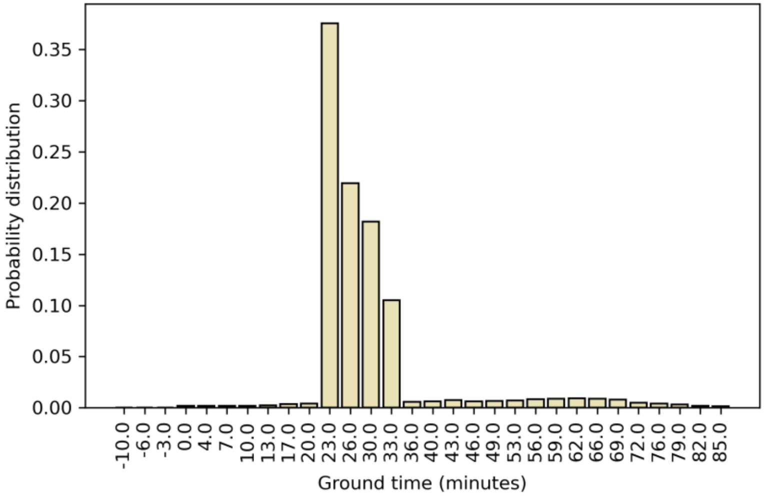


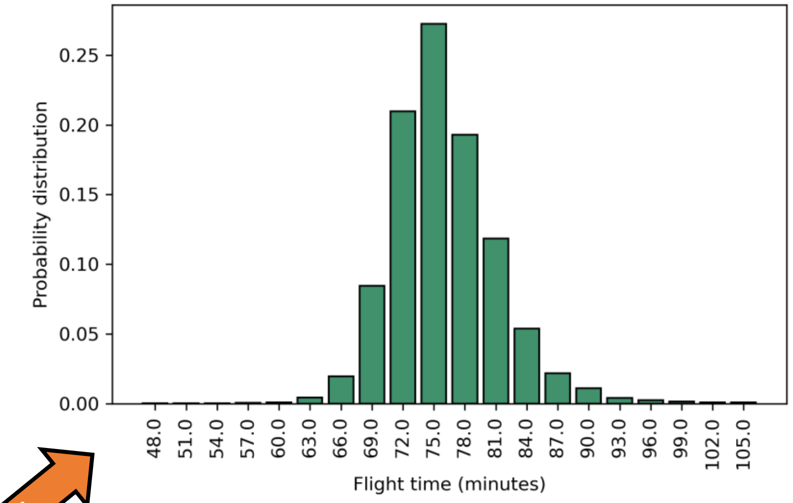


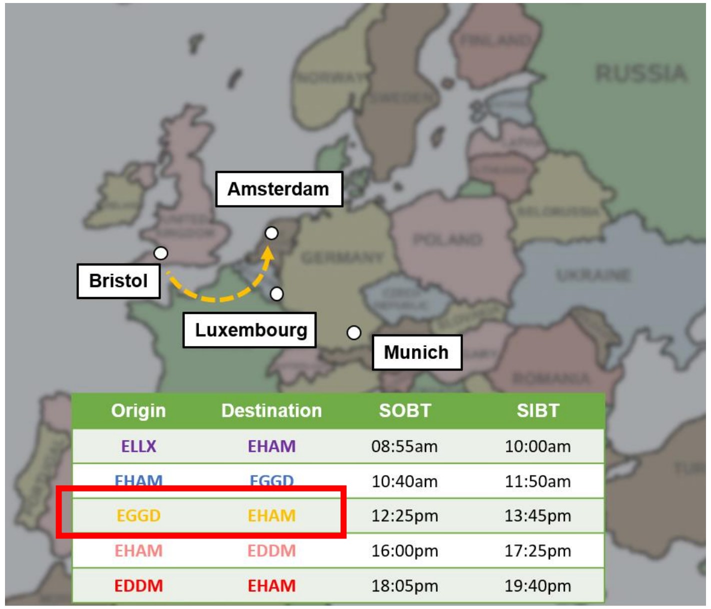
An experiment

UNIVERSITY OF WESTMINSTER

Flight 3: Bristol – Amsterdam

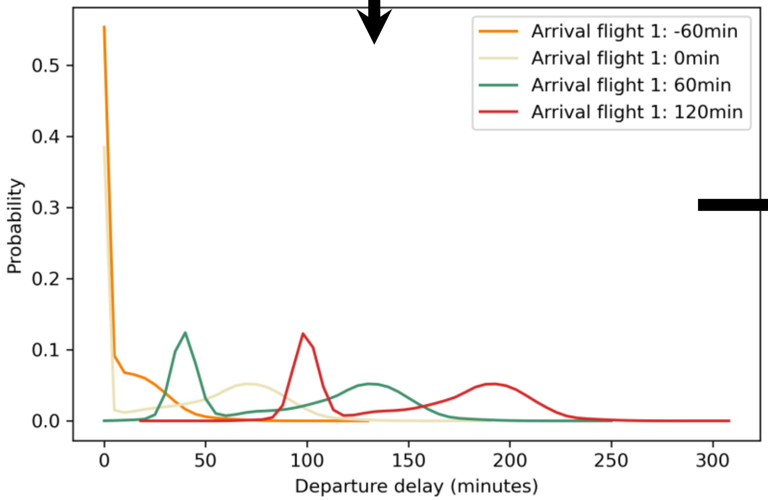


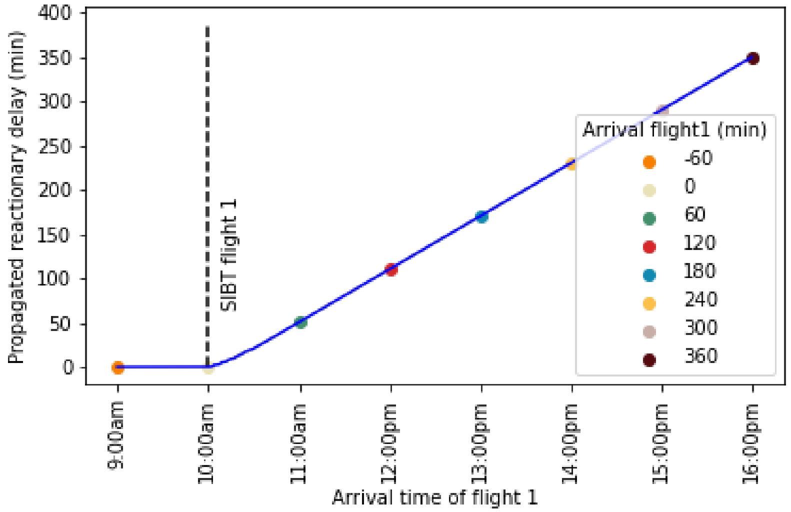


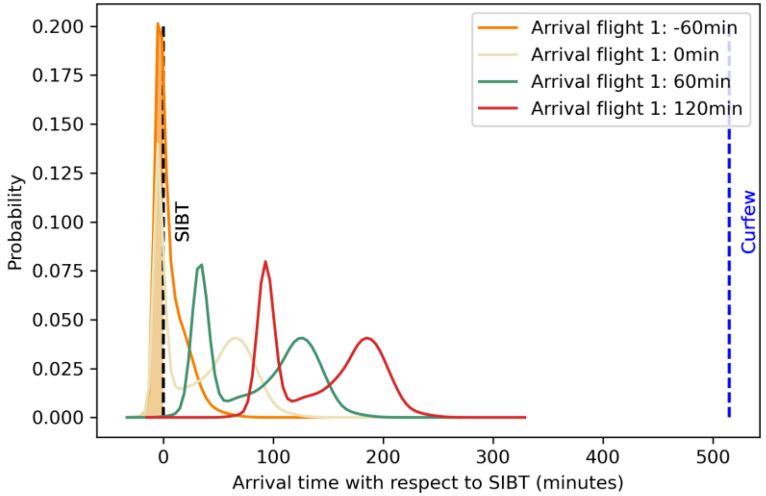


Origin	Destination	SOBT	SIBT
ELLX	EHAM	08:55am	10:00am
EHAM	EGGD	10:40am	11:50am
EGGD	EHAM	12:25pm	13:45pm
EHAM	EDDM	16:00pm	17:25pm
EDDM	EHAM	18:05pm	19:40pm

*

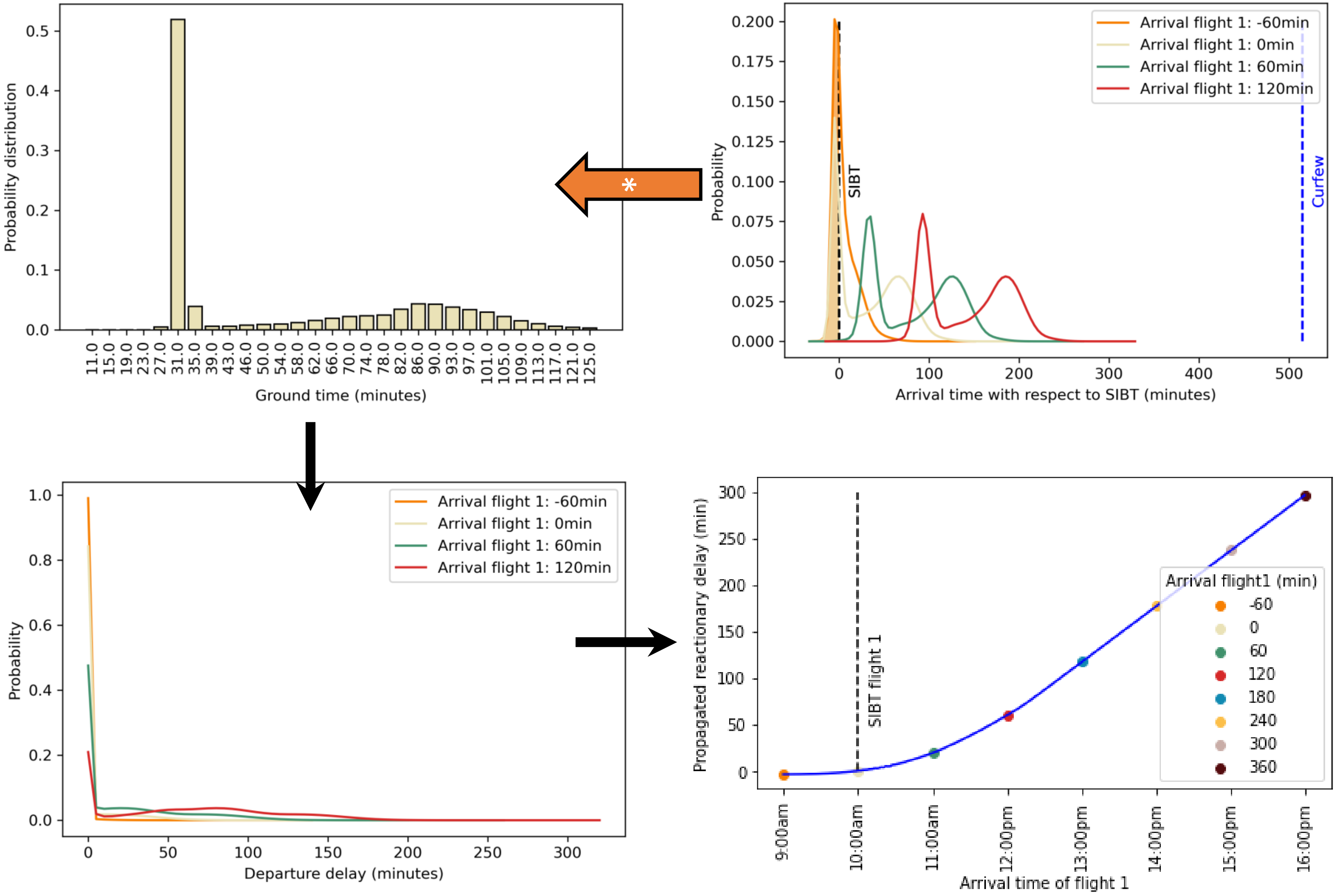


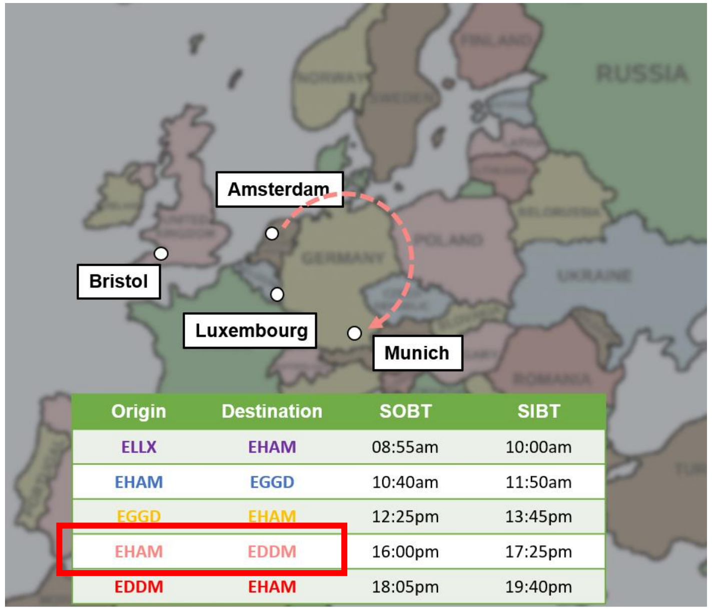




An experiment

Flight 4: Amsterdam - Munich



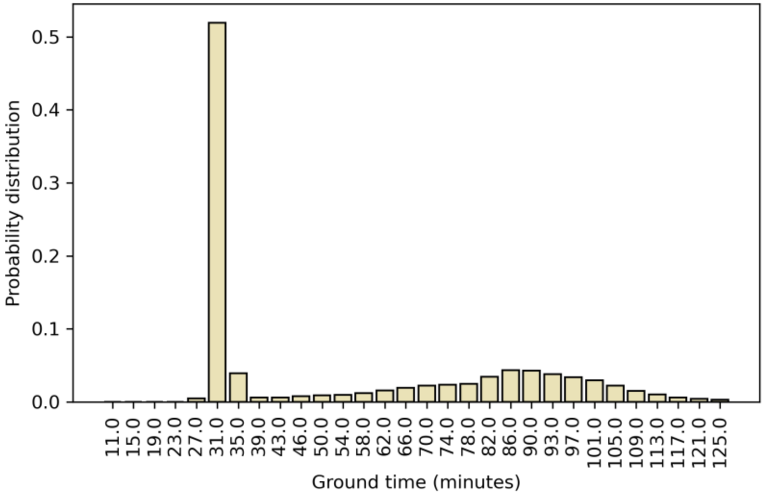


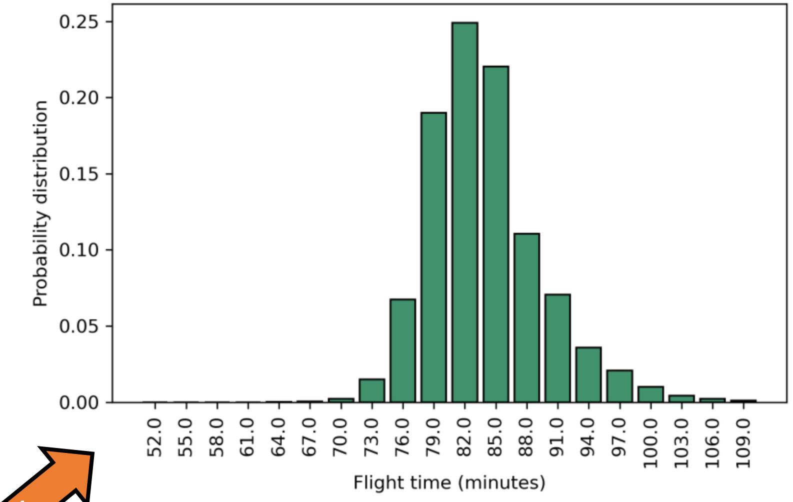
Origin	Destination	SOBT	SIBT
ELLX	EHAM	08:55am	10:00am
EHAM	EGGD	10:40am	11:50am
EGGD	EHAM	12:25pm	13:45pm
EHAM	EDDM	16:00pm	17:25pm
EDDM	EHAM	18:05pm	19:40pm

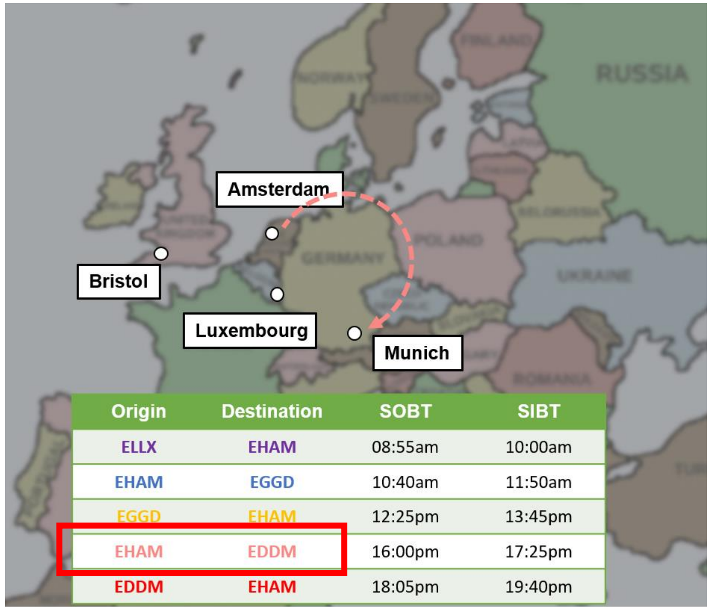
An experiment

UNIVERSITY OF WESTMINSTER

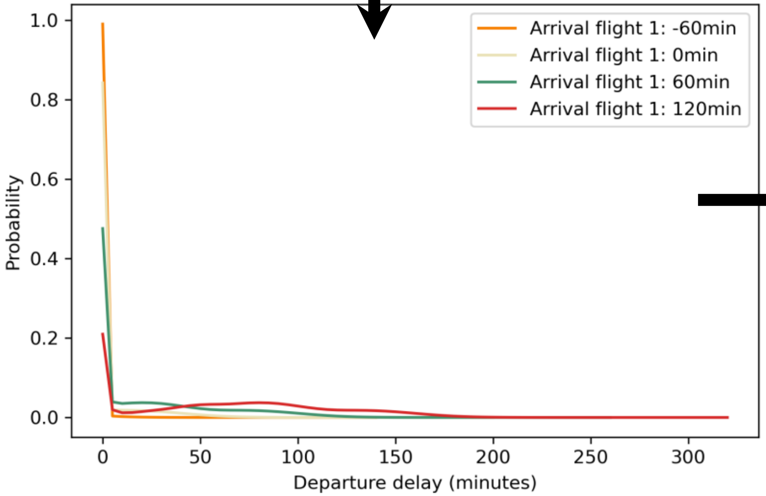
Flight 4: Amsterdam - Munich

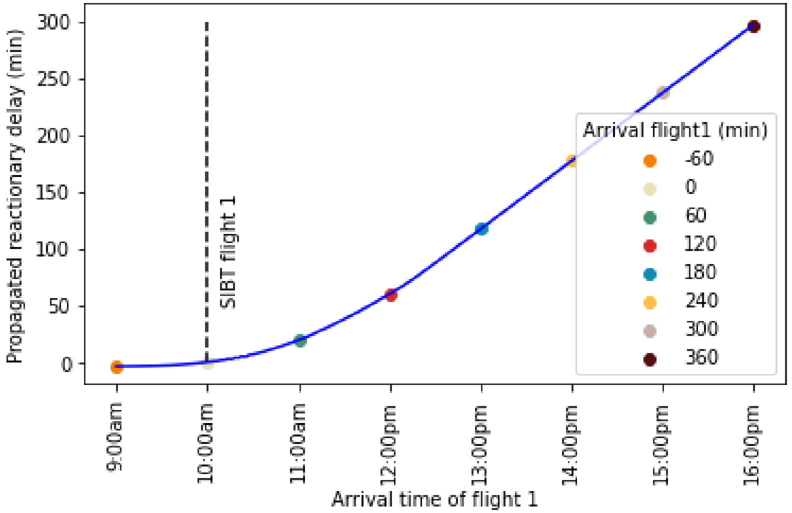


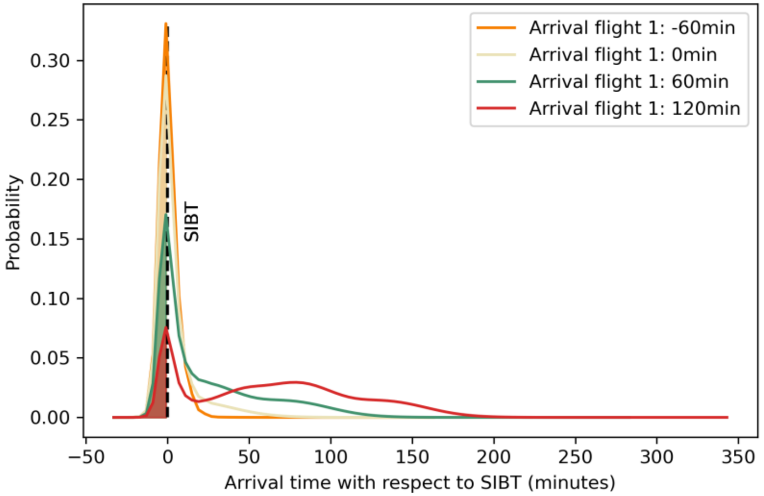




Origin	Destination	SOBT	SIBT
ELLX	EHAM	08:55am	10:00am
EHAM	EGGD	10:40am	11:50am
EGGD	EHAM	12:25pm	13:45pm
EHAM	EDDM	16:00pm	17:25pm
EDDM	EHAM	18:05pm	19:40pm

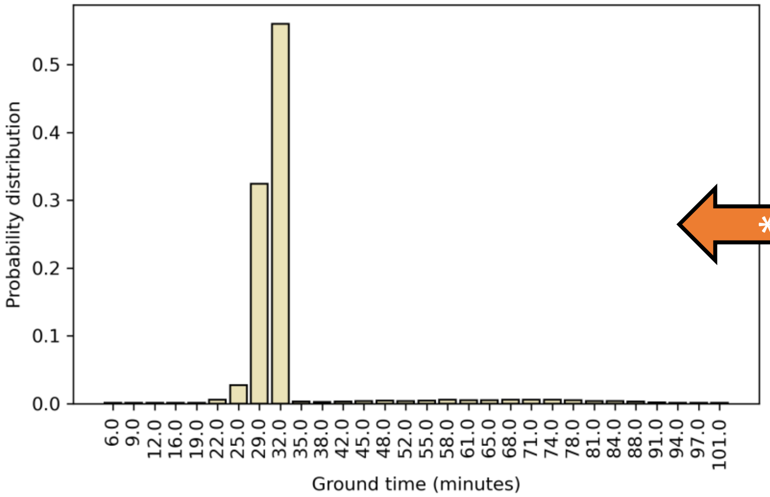


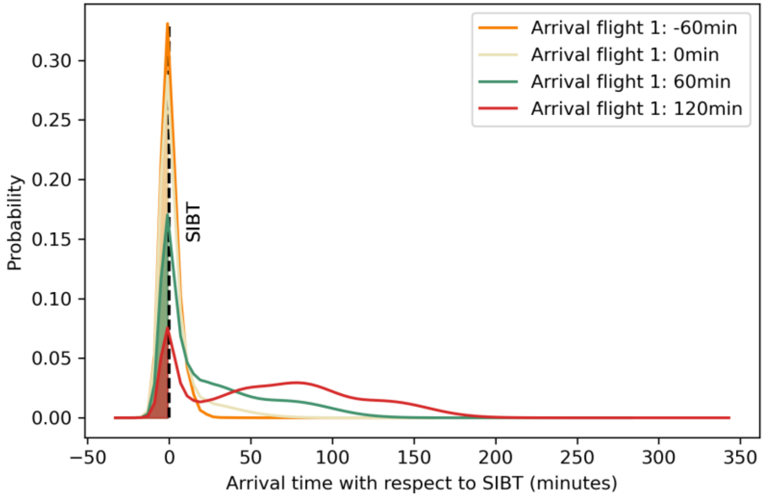


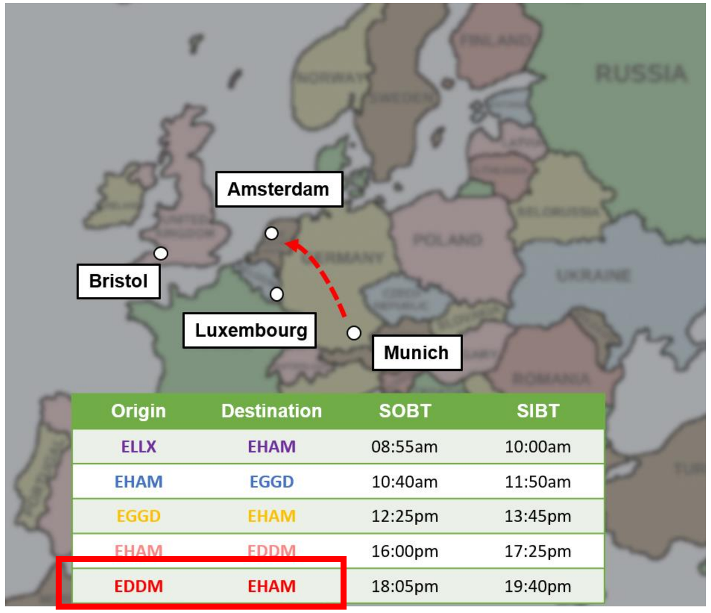


An experiment

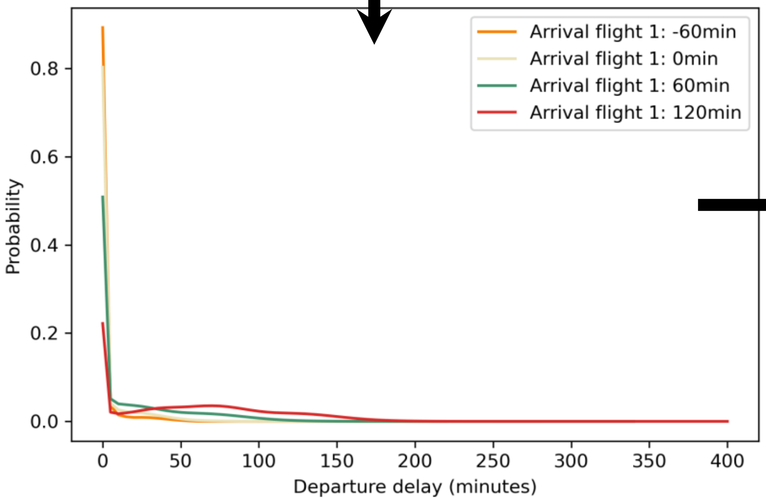
Flight 5: Munich - Amsterdam

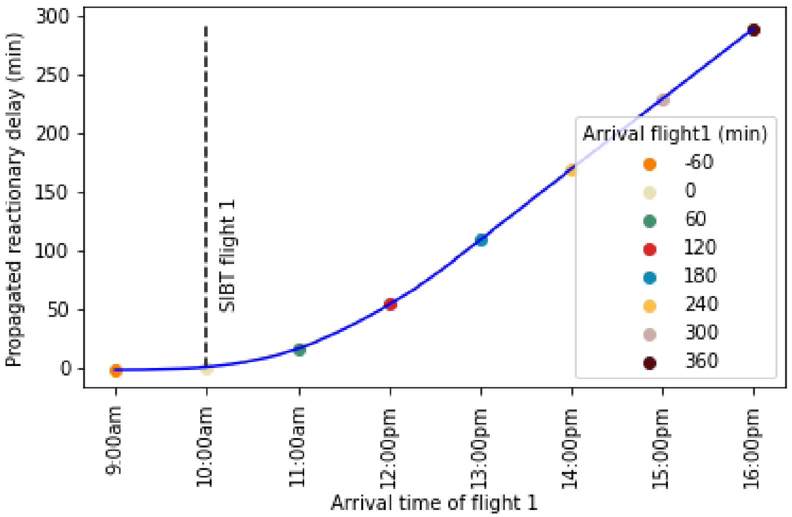






Origin	Destination	SOBT	SIBT
ELLX	EHAM	08:55am	10:00am
EHAM	EGGD	10:40am	11:50am
EGGD	EHAM	12:25pm	13:45pm
EHAM	EDDM	16:00pm	17:25pm
EDDM	EHAM	18:05pm	19:40pm

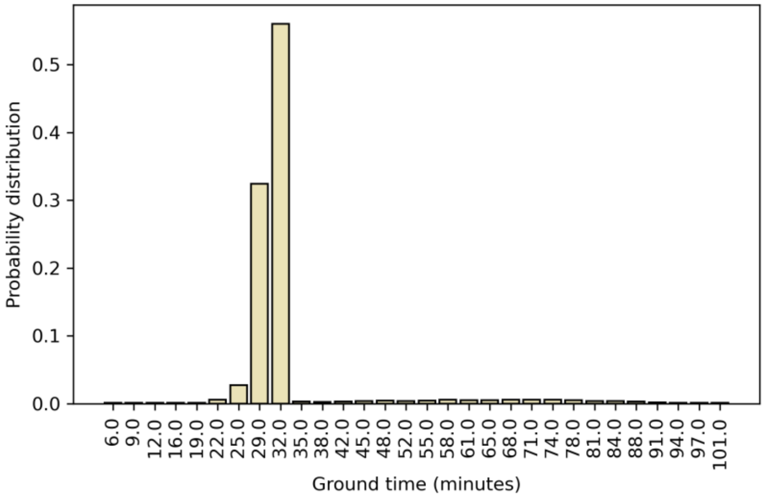


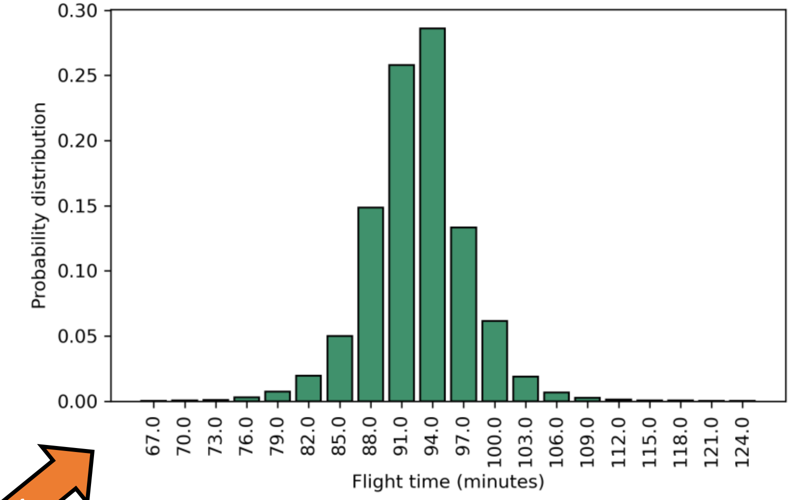


An experiment

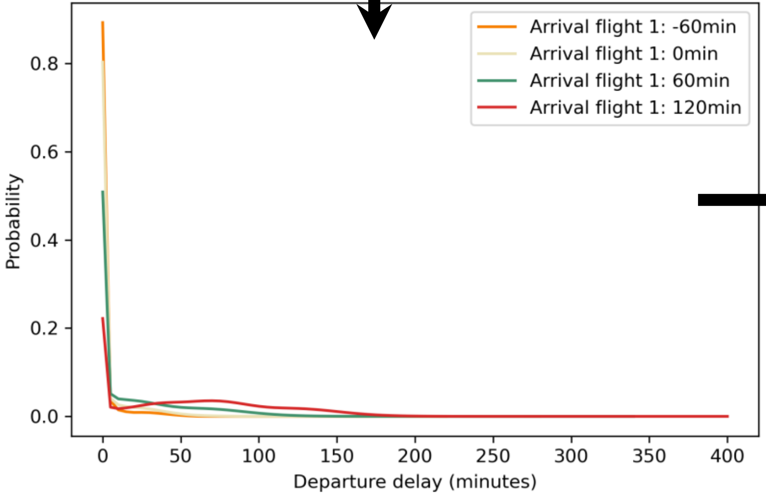
UNIVERSITY OF WESTMINSTER

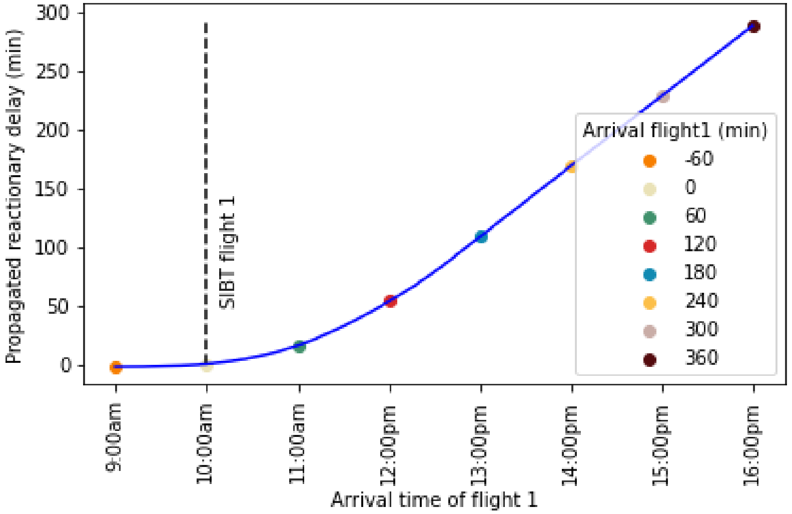
Flight 5: Munich - Amsterdam

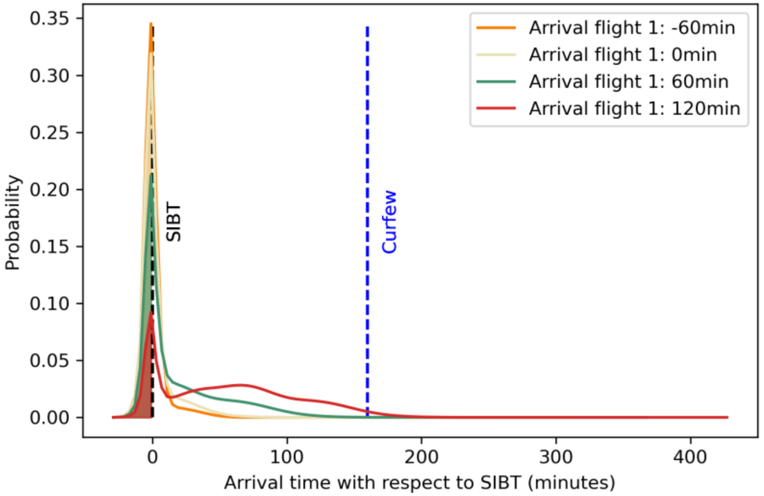




Origin	Destination	SOBT	SIBT
ELLX	EHAM	08:55am	10:00am
EHAM	EGGD	10:40am	11:50am
EGGD	EHAM	12:25pm	13:45pm
EHAM	EDDM	16:00pm	17:25pm
EDDM	EHAM	18:05pm	19:40pm



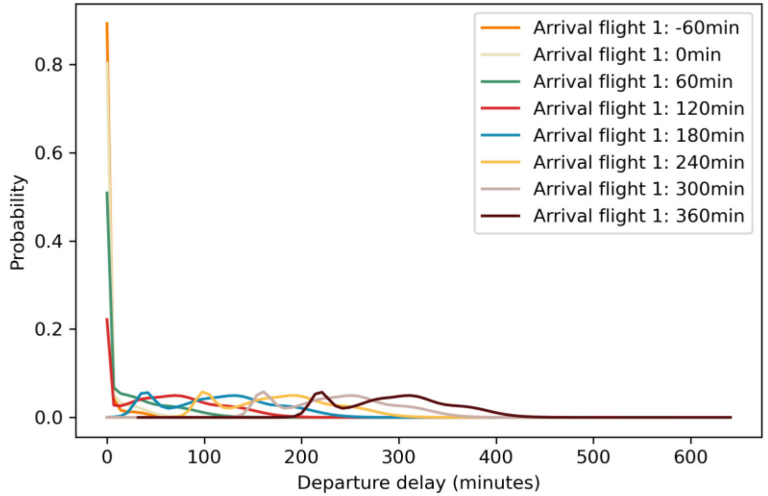


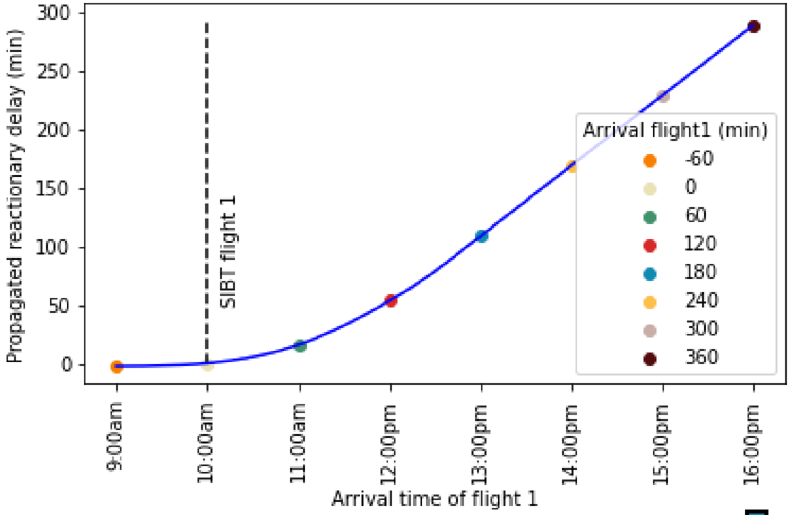


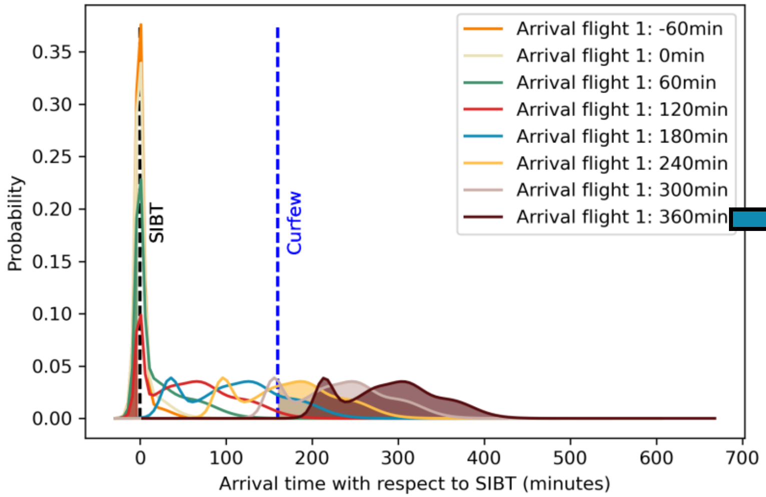
An experiment

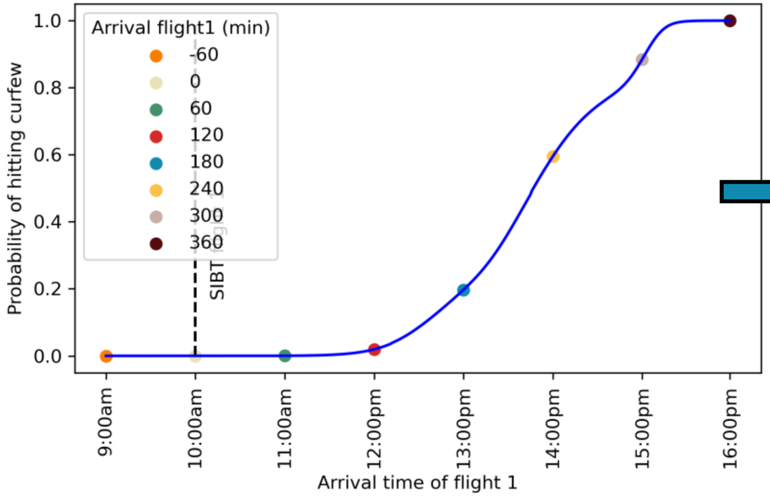
Flight 5: Munich - Amsterdam

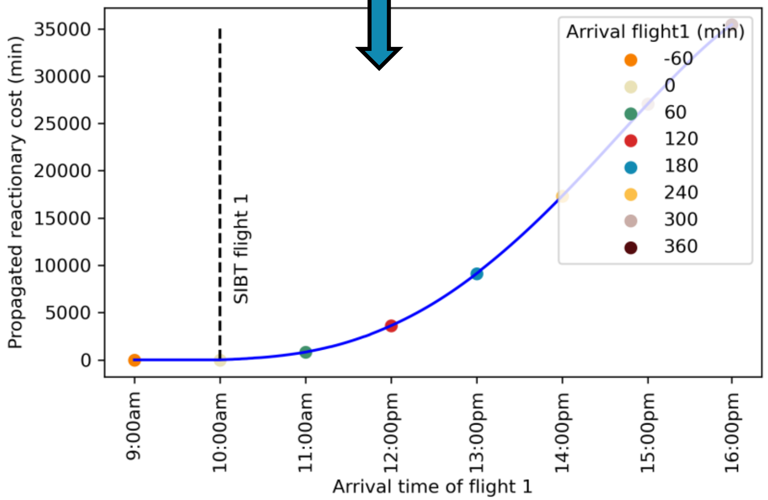
Probability of hitting the curfew





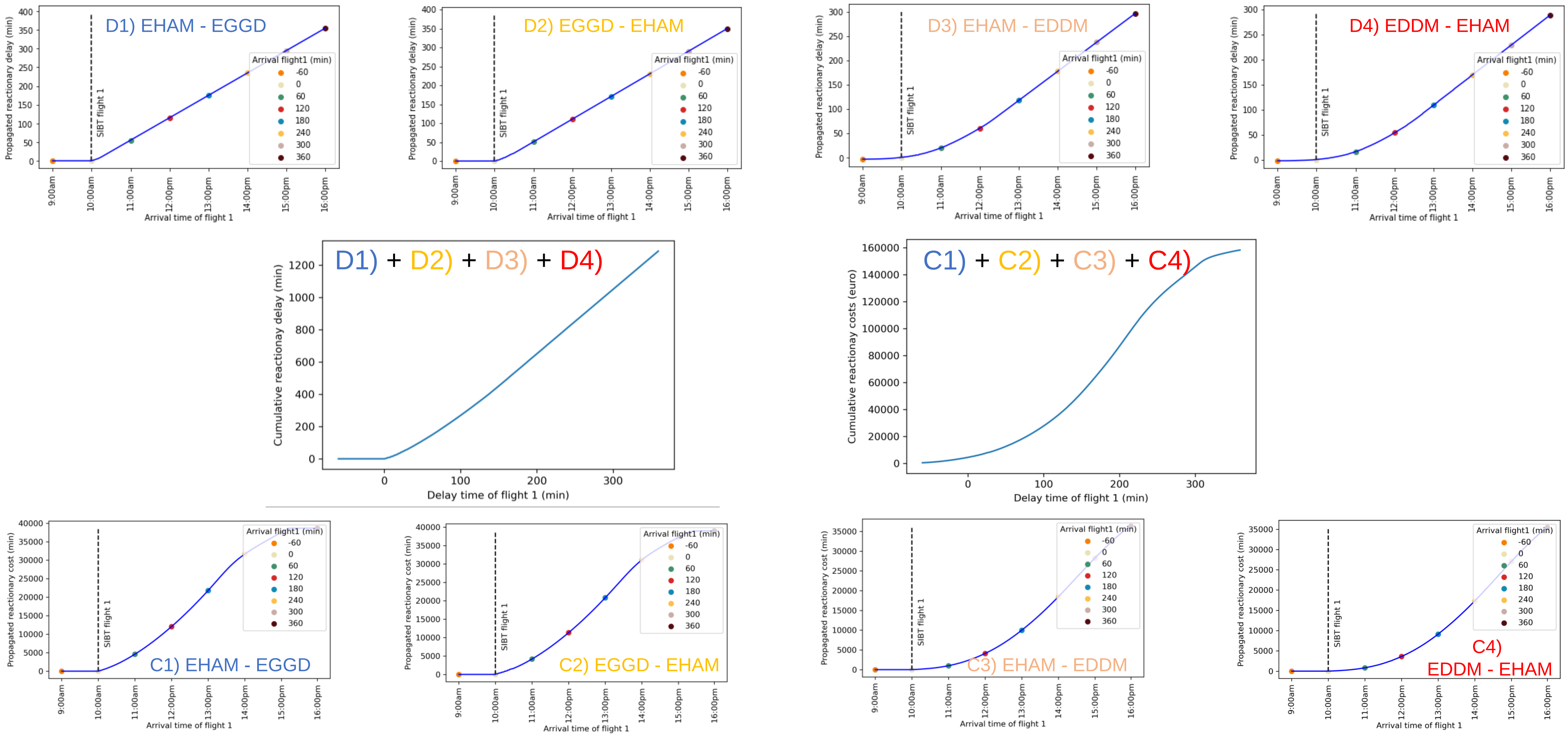






Cook, Andrew J., and Graham Tanner. "European airline delay cost reference values." (2011)

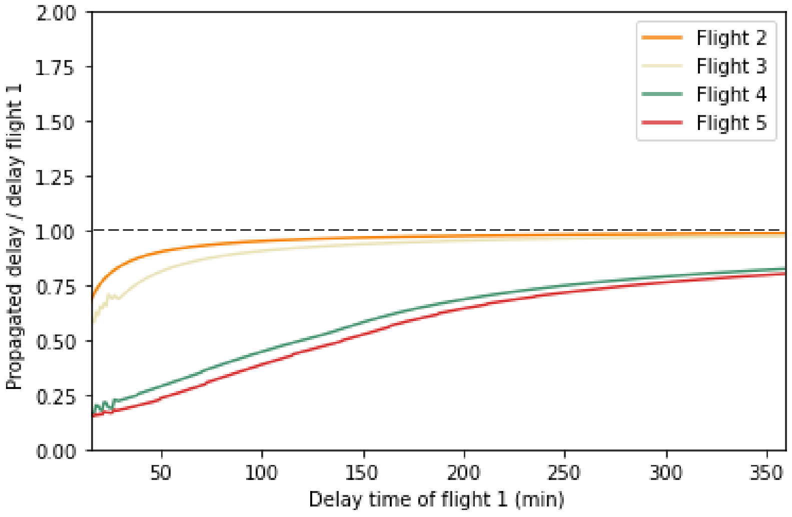
An overview on reactionary delay and costs

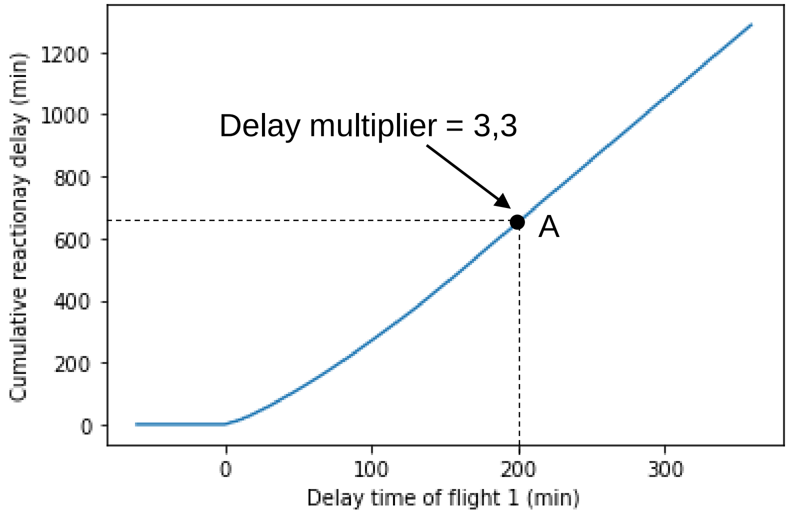


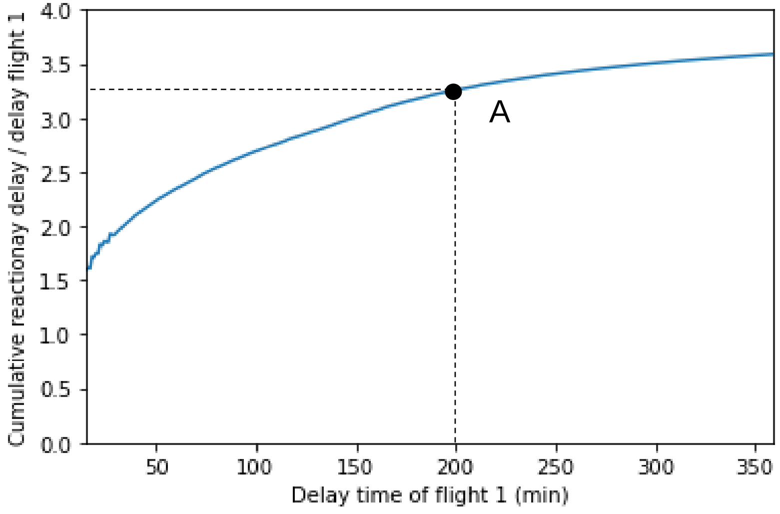
Cook, Andrew J., and Graham Tanner. "European airline delay cost reference values." (2011)

An overview on reactionary delay and costs

UNIVERSITY OF WESTMINSTER







Pilot3

- Pilot3 will develop a software engine model for supporting crew decisions for civil aircraft.

- Pilot3 will integrate airlines flight policies and overall performance targets to select and rank the alternatives.

- The system does not only consider the flight but the whole network operations of the airline (for example, it will estimate the cost of reactionary delay).

- Pilot3 will allow the airline to select how to predict fields of interest: using airborne information, ground information, with analysis of data and heuristics or with machine learning predictors.

https://pilot3.eu/

Dispatcher3

- Dispatcher3 will develop a software prototype for the acquisition and preparation of historical flight data in order to give support to the optimisation of future flights providing predictive capabilities and advice to flight managers.

- Dispatcher3 focuses on activities prior to departure: dispatching and pilot advice on how to operate the flight.

- Dispatcher3 is composed of three layers: data infrastructure, predictive capabilities and advice capabilities.

- The predictive capabilities will be provided by the development of predictive models using machine learning algorithms for targeted airlines' KPIs.

UNIVERSITY OF WESTMINSTER · innaxis · UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH · PACE a TXT company · vueling · skeyes · THALES · Clean Sky

SWISS · norwegian · airBaltic · EUROCONTROL · A3 Aviation Consulting · ACK · Valentín Lago Independant consultant · DISPATCHER3

https://dispatcher3.eu/

Future developments

UNIVERSITY OF WESTMINSTER

- Progressively increase the amount of data for the training/validation steps
- Integrate with other data sources (e.g. TAF data) and increase the amount of features of the ML models
- Interpolation within the time horizons
- Create models for specific ‘o/d pair’ airports
- Extend the approach to the strategic level to improve the buffer assignment