Supporting
European
Aviation

EUROCONTROL

Using machine learning to predict the evolution and propagation of delays
Main results and lessons learnt

Engage workshop – Machine learning, AI and automation in ATM

Ramon Dalmau-Codina
EUROCONTROL

3rd of September 2021

FOUNDING MEMBER SESAR JOINT UNDERTAKING

NETWORK MANAGER

Outline

1. Evolution of air traffic flow management (ATFM) delay
    1. Introduction
    2. Model
    3. Features
    4. Data
    5.
2. … rfews
    1.
    2. Mo…
    3. Features
    4. Data
    5. Results
3. Discussion

Other members of the team:
- Brice Genestier
- Camille Anoraud
- Peter Choroba
- Darren Smith

Other members of the team:
- Giuseppe Murgese
- Yves de Wandeler
- Ricardo Correira
- Alan Marsden

Join us in the USA/Europe ATM R&D Seminar 2021

http://www.atmseminar.org/

Outline

1. **Evolution of air traffic flow management (ATFM) delay**
    1. Introduction
    2. Model
    3. Features
    4. Data
    5. Results
2. Delay propagation and night curfews
    1. Introduction
    2. Model
    3. Features
    4. Data
    5. Results
3. Discussion

Other members of the team:

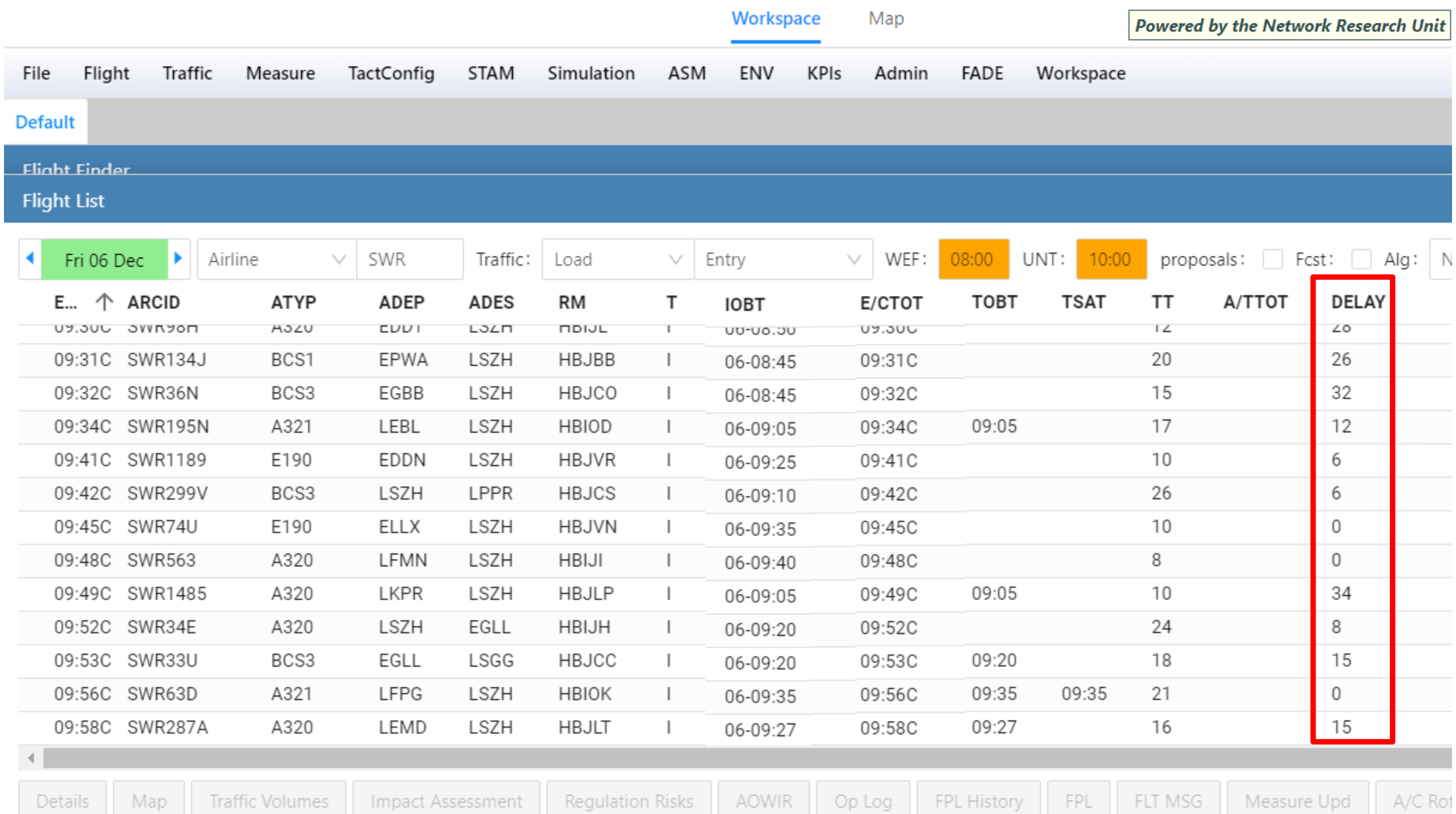
- **Brice Genestier**
- **Camille Anoraud**
- **Peter Choroba**
- **Darren Smith**

Other members of the team:

- Giuseppe Murgese
- Yves de Wandeler
- Ricardo Correira
- Alan Marsden

Introduction

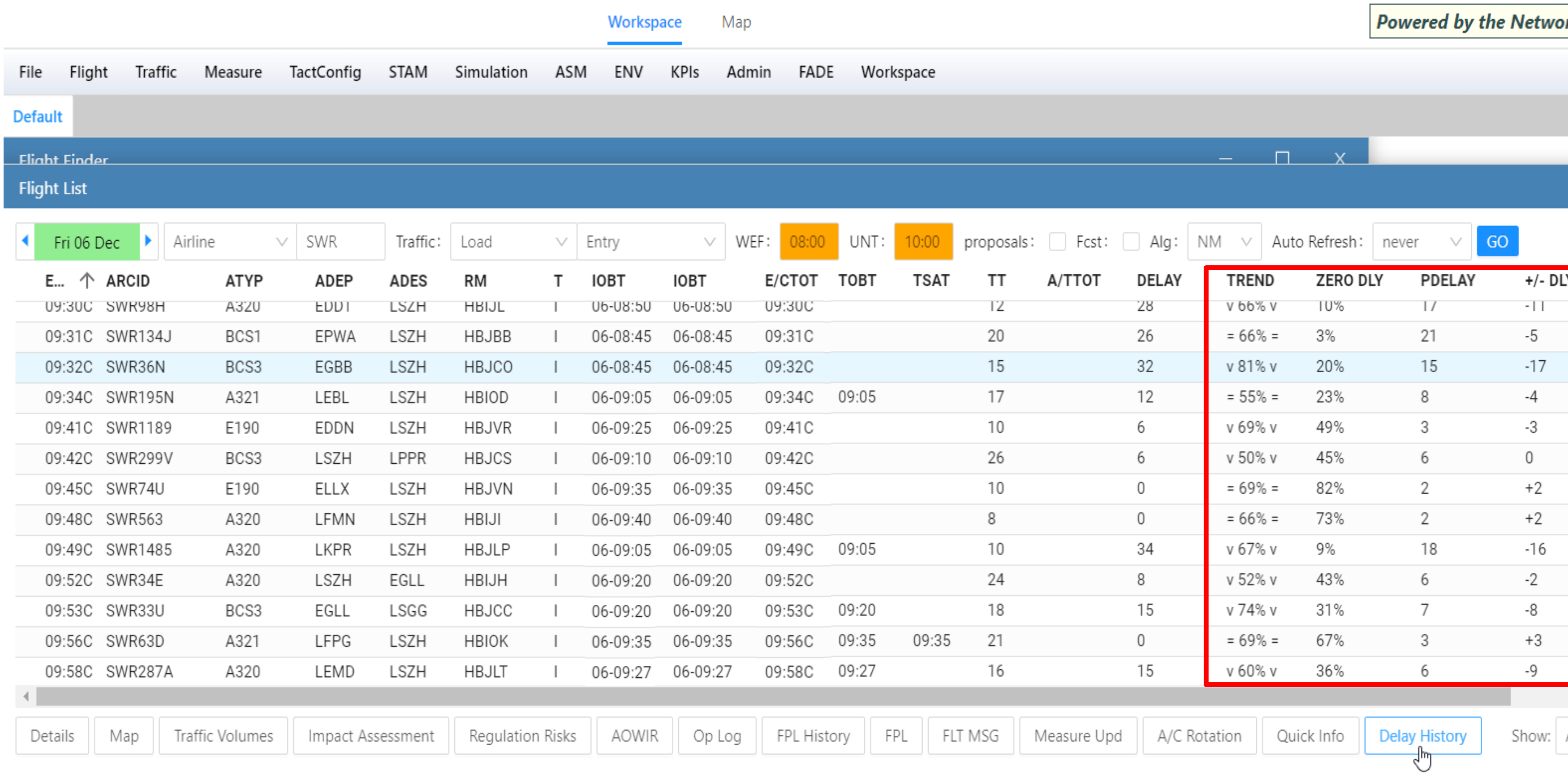
- At present, when a flight is caught by an air traffic flow management **(ATFM) regulation**, the airline has very **limited information about the evolution of the ATFM delay**



E...	ARCID	ATYP	ADEP	ADES	RM	T	IOBT	E/CTOT	TOBT	TSAT	TT	A/TTOT	DELAY
09:30C	SWR98H	A320	EDDT	LSZH	HBIJL	I	06-08:50	09:30C			12		28
09:31C	SWR134J	BCS1	EPWA	LSZH	HBJBB	I	06-08:45	09:31C			20		26
09:32C	SWR36N	BCS3	EGBB	LSZH	HBJCO	I	06-08:45	09:32C			15		32
09:34C	SWR195N	A321	LEBL	LSZH	HBIOD	I	06-09:05	09:34C	09:05		17		12
09:41C	SWR1189	E190	EDDN	LSZH	HBJVR	I	06-09:25	09:41C			10		6
09:42C	SWR299V	BCS3	LSZH	LPPR	HBJCS	I	06-09:10	09:42C			26		6
09:45C	SWR74U	E190	ELLX	LSZH	HBJVN	I	06-09:35	09:45C			10		0
09:48C	SWR563	A320	LFMN	LSZH	HBIJI	I	06-09:40	09:48C			8		0
09:49C	SWR1485	A320	LKPR	LSZH	HBJLP	I	06-09:05	09:49C	09:05		10		34
09:52C	SWR34E	A320	LSZH	EGLL	HBIJH	I	06-09:20	09:52C			24		8
09:53C	SWR33U	BCS3	EGLL	LSGG	HBJCC	I	06-09:20	09:53C	09:20		18		15
09:56C	SWR63D	A321	LFPG	LSZH	HBIOK	I	06-09:35	09:56C	09:35	09:35	21		0
09:58C	SWR287A	A320	LEMD	LSZH	HBJLT	I	06-09:27	09:58C	09:27		16		15

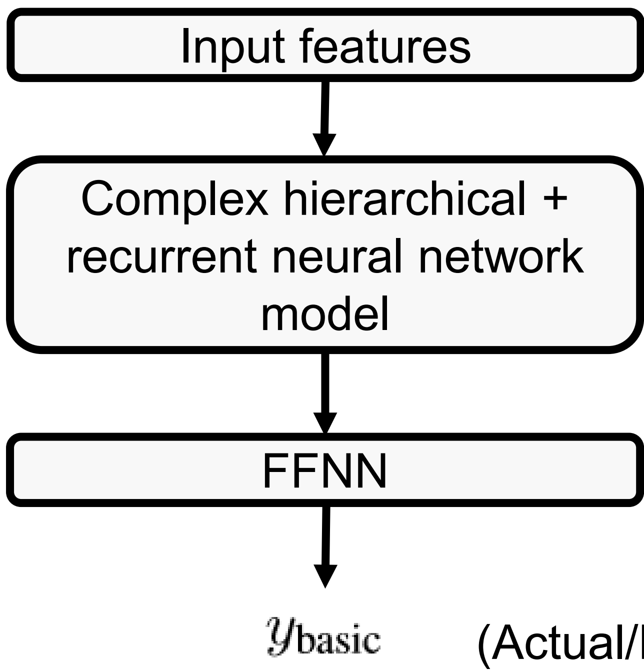
Introduction

- The objective of this project was to develop a **machine learning model** that, trained on **historical data** (the past), can predict the **evolution** (the future) **of the current ATFM delay** for a regulated flight



E...	ARCID	ATYP	ADEP	ADES	RM	T	IOBT	IOBT	E/CTOT	TOBT	TSAT	TT	A/TTOT	DELAY	TREND	ZERO DLY	PDELAY	+/- DLY
09:30C	SWR98H	A320	EDDT	LSZH	HBIJL	I	06-08:50	06-08:50	09:30C			12		28	v 66% v	10%	17	-11
09:31C	SWR134J	BCS1	EPWA	LSZH	HBJBB	I	06-08:45	06-08:45	09:31C			20		26	= 66% =	3%	21	-5
09:32C	SWR36N	BCS3	EGBB	LSZH	HBJCO	I	06-08:45	06-08:45	09:32C			15		32	v 81% v	20%	15	-17
09:34C	SWR195N	A321	LEBL	LSZH	HBIOD	I	06-09:05	06-09:05	09:34C	09:05		17		12	= 55% =	23%	8	-4
09:41C	SWR1189	E190	EDDN	LSZH	HBJVR	I	06-09:25	06-09:25	09:41C			10		6	v 69% v	49%	3	-3
09:42C	SWR299V	BCS3	LSZH	LPPR	HBJCS	I	06-09:10	06-09:10	09:42C			26		6	v 50% v	45%	6	0
09:45C	SWR74U	E190	ELLX	LSZH	HBJVN	I	06-09:35	06-09:35	09:45C			10		0	= 69% =	82%	2	+2
09:48C	SWR563	A320	LFMN	LSZH	HBIJI	I	06-09:40	06-09:40	09:48C			8		0	= 66% =	73%	2	+2
09:49C	SWR1485	A320	LKPR	LSZH	HBJLP	I	06-09:05	06-09:05	09:49C	09:05		10		34	v 67% v	9%	18	-16
09:52C	SWR34E	A320	LSZH	EGLL	HBIJH	I	06-09:20	06-09:20	09:52C			24		8	v 52% v	43%	6	-2
09:53C	SWR33U	BCS3	EGLL	LSGG	HBJCC	I	06-09:20	06-09:20	09:53C	09:20		18		15	v 74% v	31%	7	-8
09:56C	SWR63D	A321	LFPG	LSZH	HBIOK	I	06-09:35	06-09:35	09:56C	09:35	09:35	21		0	= 69% =	67%	3	+3
09:58C	SWR287A	A320	LEMD	LSZH	HBJLT	I	06-09:27	06-09:27	09:58C	09:27		16		15	v 60% v	36%	6	-9

Model – Basic Regression

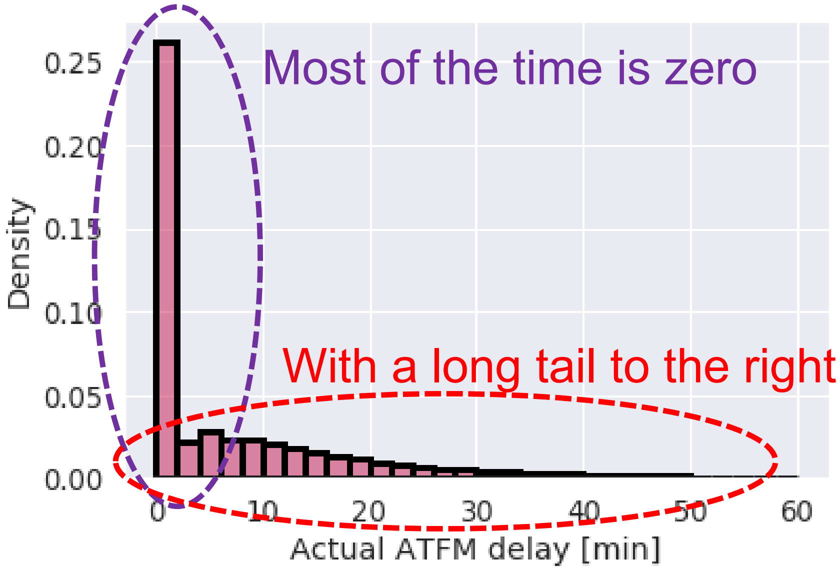


(Actual/Last ATFM delay, in minutes)

Works well, but does not provide a measure of uncertainty …

Model – Poisson Regression

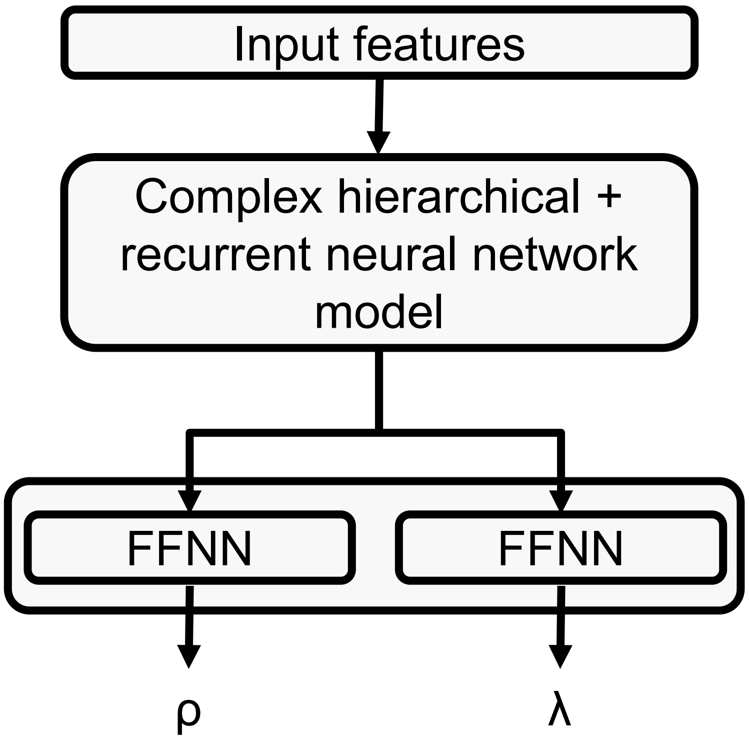
The last ATFM delay distribution fits a Zero-Inflated Poisson (**ZIP**)



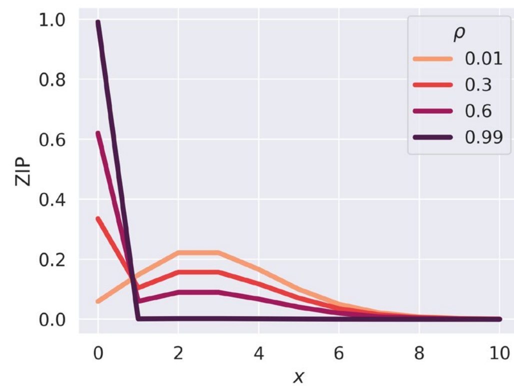
λ : Poisson rate

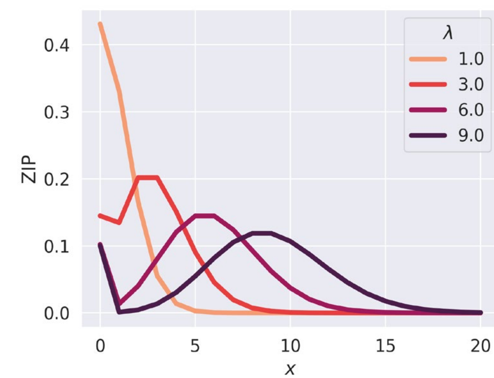
ρ : probability of extra zeros

Model – Poisson Regression



Parameters of the last ATFM delay distribution (under ZIP assumption)





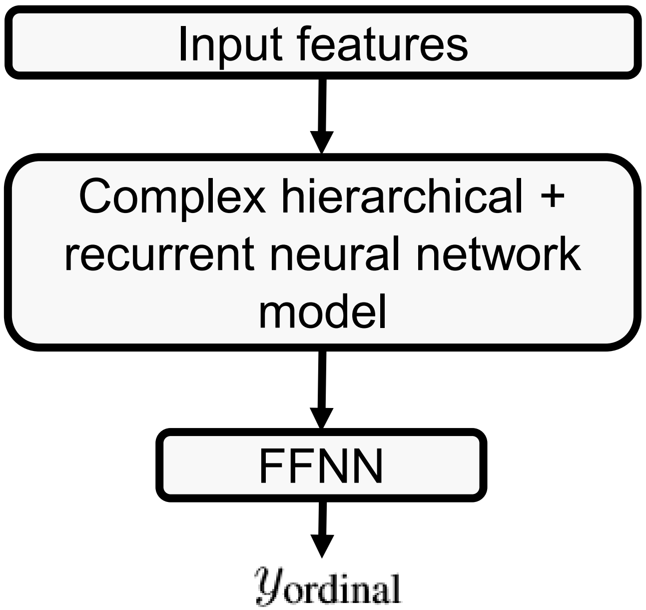
Expected last ATFM delay = **(1- ρ) λ**

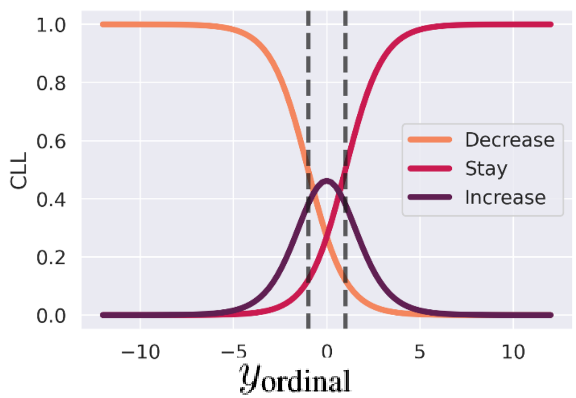
Example: ρ=0.9 λ=10
Expected last ATFM delay = 1

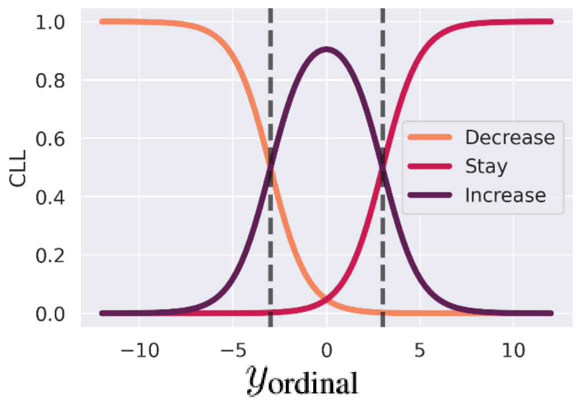
Model – Ordinal Regression

In the current implementation, the **trend** of the ATFM delay is classified as:

- **Increase:** The last ATFM delay is higher than the current + 5 min
- **Decrease:** The last ATFM delay is lower than the current - 5 min
- **Stay stable:** Otherwise







The *cutpoints* are additional parameters to be learned from the data

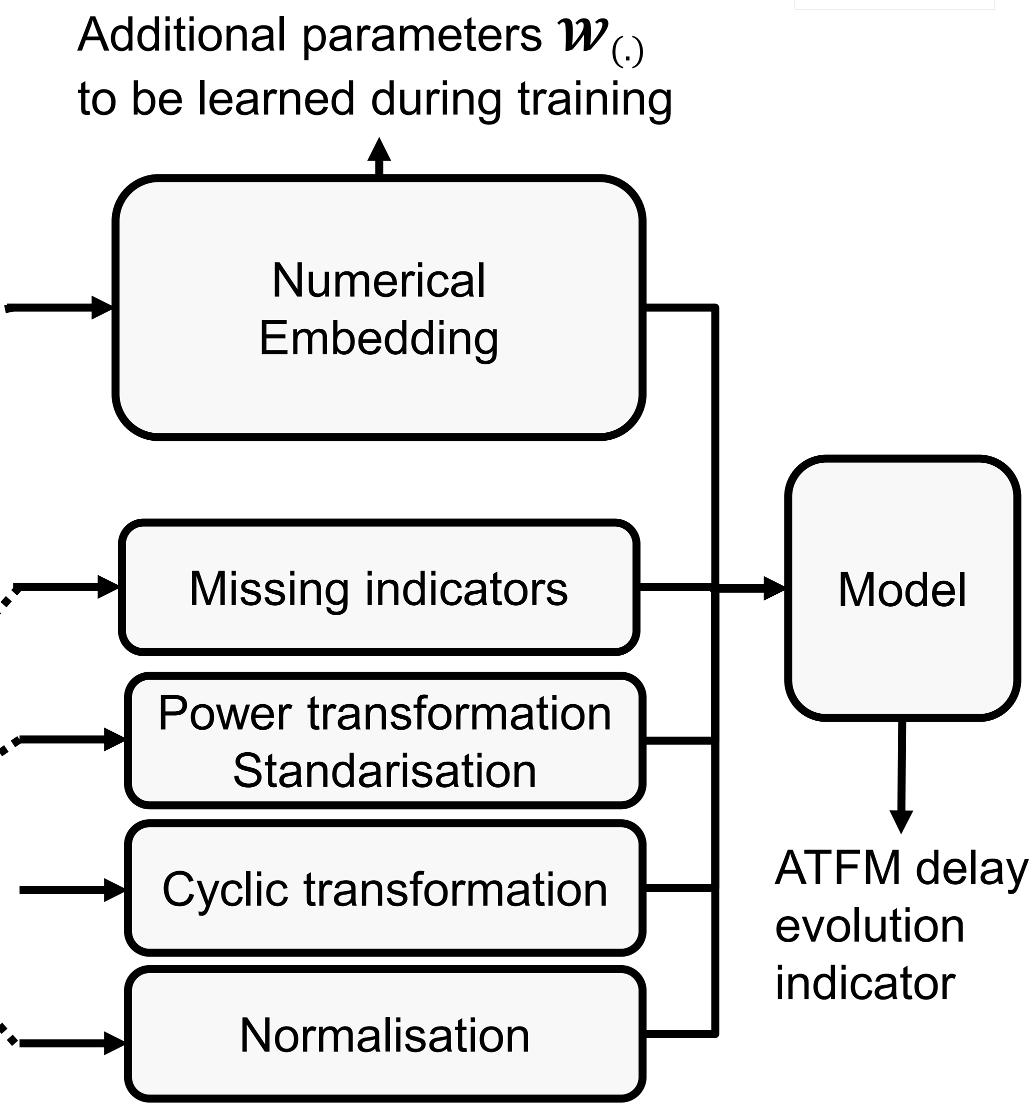
Features

Categorical features:

- Departure airport
- Destination airport
- ATFM flight state
- Traffic volume
- Geographical entity
- Type of geographical entity
- Reason of the regulation
- Airline …

Numerical features:

- Time to start of the regulation
- Current ATFM delay
- Hour of the day
- Number of regulations …



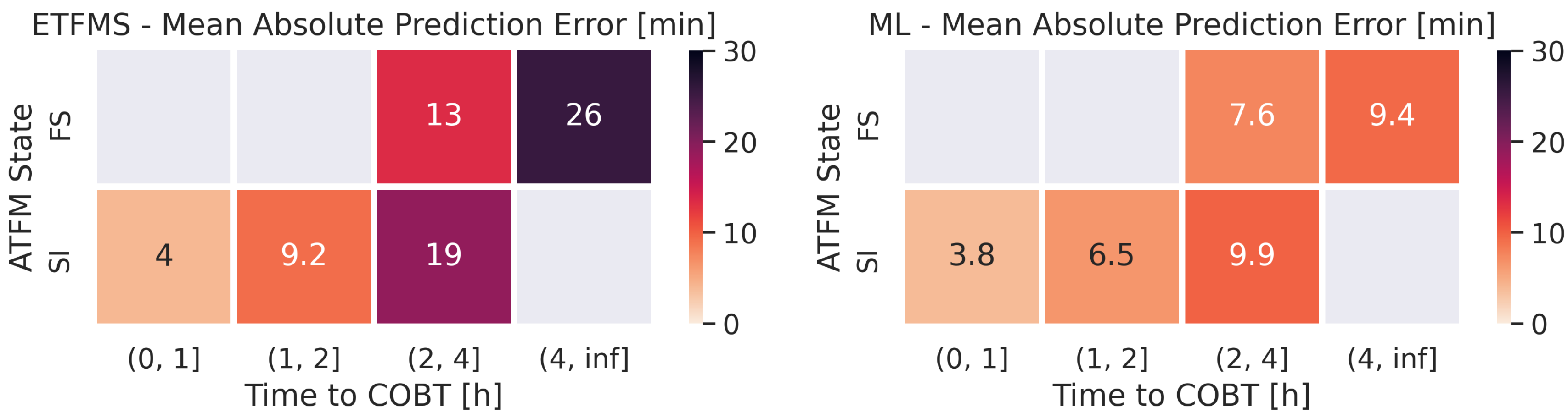
Data

- Enhanced tactical flow management system (**ETFMS**) flight data (EFD) messages of all flights that were regulated during **2019**
  - Departure and destination airports
  - Airline
  - Estimated off-block time (EOBT)
  - ATFM delay
  - Taxi time …
- Evolution of the corresponding **ATFM regulations**
  - Time of activation
  - Regulation rate
  - Reason of the regulation
  - Start and end time …

Example

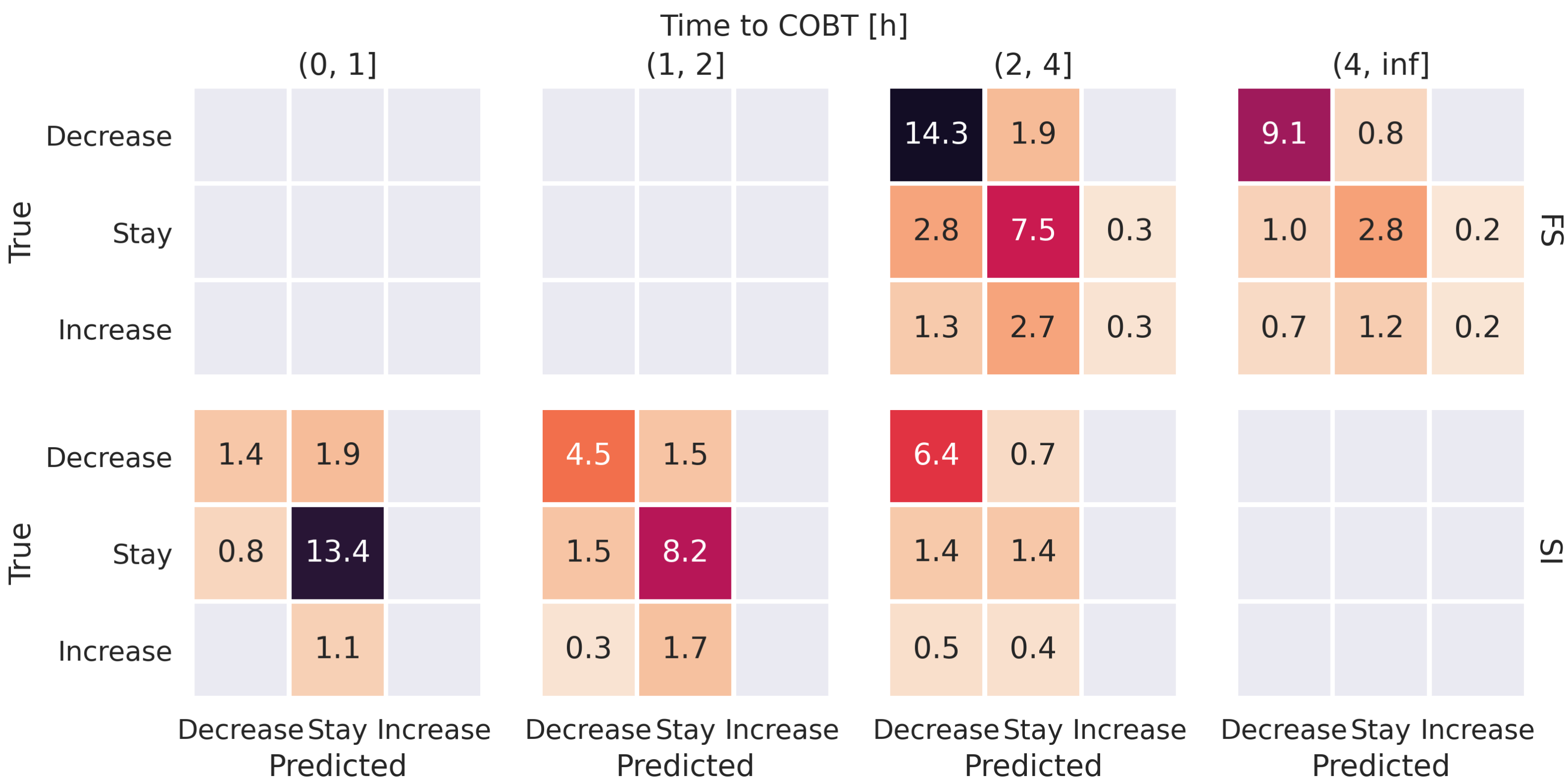
Time to EOBT	State	Delay	y_{basic}	λ	ρ	p_d	p_s	p_i
03:43:34	FS	79	12	23	0.25	0.98	0.02	0
03:31:37	FS	64	10	23	0.21	0.97	0.03	0
03:30:54	FS	79	11	26	0.19	0.98	0.02	0
02:58:47	FS	76	11	26	0.17	0.98	0.02	0
02:39:55	FS	46	09	21	0.20	0.93	0.07	0
02:28:42	FS	76	11	25	0.17	0.98	0.02	0
02:07:22	FS	76	12	26	0.16	0.98	0.02	0
02:02:54	SI	76	12	26	0.16	0.98	0.02	0
02:00:47	SI	76	12	25	0.17	0.98	0.02	0
00:42:48	SI	76	21	30	0.11	0.97	0.03	0
00:39:45	SI	76	21	31	0.12	0.96	0.04	0
00:39:38	SI	76	23	31	0.12	0.96	0.04	0
00:35:07	SI	76	31	39	0.08	0.96	0.04	0
00:23:46	SI	34	22	23	0.05	0.80	0.20	0
00:15:46	SI	19	15	14	0.08	0.61	0.37	0
00:09:27	SI	64	34	40	0.03	0.93	0.07	0
00:07:31	SI	34	23	24	0.03	0.74	0.26	0
00:01:52	SI	19	15	15	0.06	0.54	0.43	0
00:01:51	SI	19	16	16	0.03	0.41	0.55	0

Results – Basic regression



Similar results for the Poisson regression

Results – Ordinal regression



Outline

1. Evolution of air traffic flow management (ATFM) delay
    1. Introduction
    2. Model
    3. Features
    4. Data
    5. Results
2. **Delay propagation and night curfews**
    1. Introduction
    2. Model
    3. Features
    4. Data
    5. Results
3. Discussion

Other members of the team:

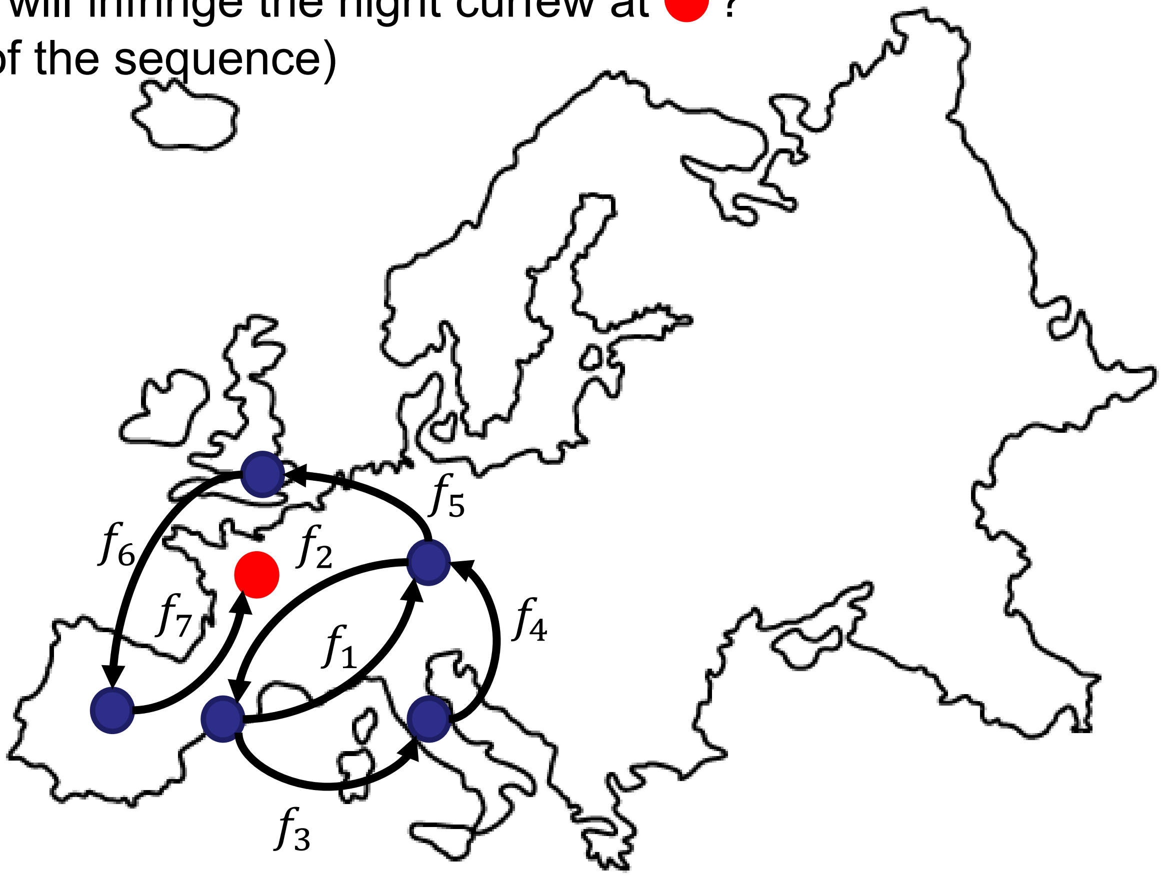
- Brice Genestier
- Camille Anoraud
- Peter Choroba
- Darren Smith

Other members of the team:

- **Giuseppe Murgese**
- **Yves de Wandeler**
- **Ricardo Correira**
- **Alan Marsden**

Introduction

Is it possible to predict the likelihood (probability) that the last flight in the sequence (f_7) will infringe the night curfew at ● ? (knowing the current state of the sequence)



● Last arrival airport

Model

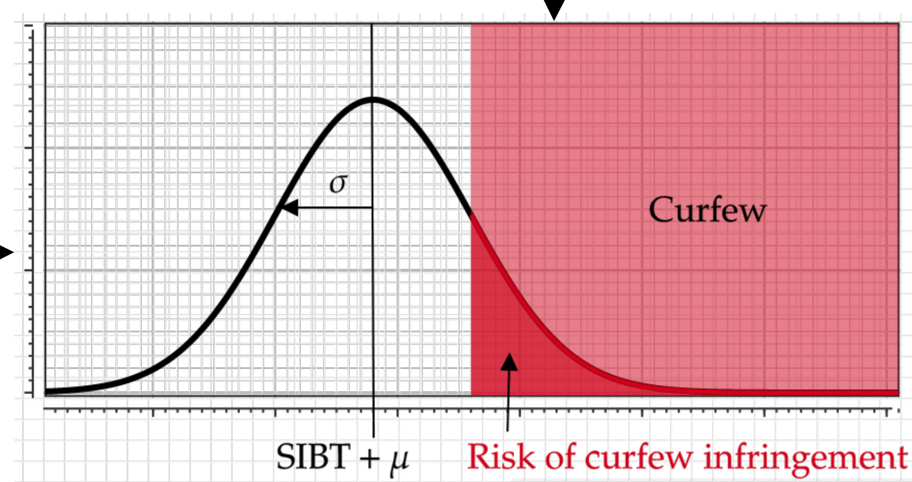
	Departure Destination …
f_1 f_2 … f_T	**Latest** flight data for the **sequence** of **flights** that a given **aircraft** is going to operate during the day

Features Generation

$\boldsymbol{x}$

Model based on Recurrent neural networks (**RNN**) and trained with **historical data**

σ_T, μ_T



Starting time of the night curfew (given) at the **destination airport of the last flight** in the sequence (i.e., f_T)

SIBT: Scheduled in-block time (SIBT)

Dynamic features ($\boldsymbol{x}_t$) – Per flight

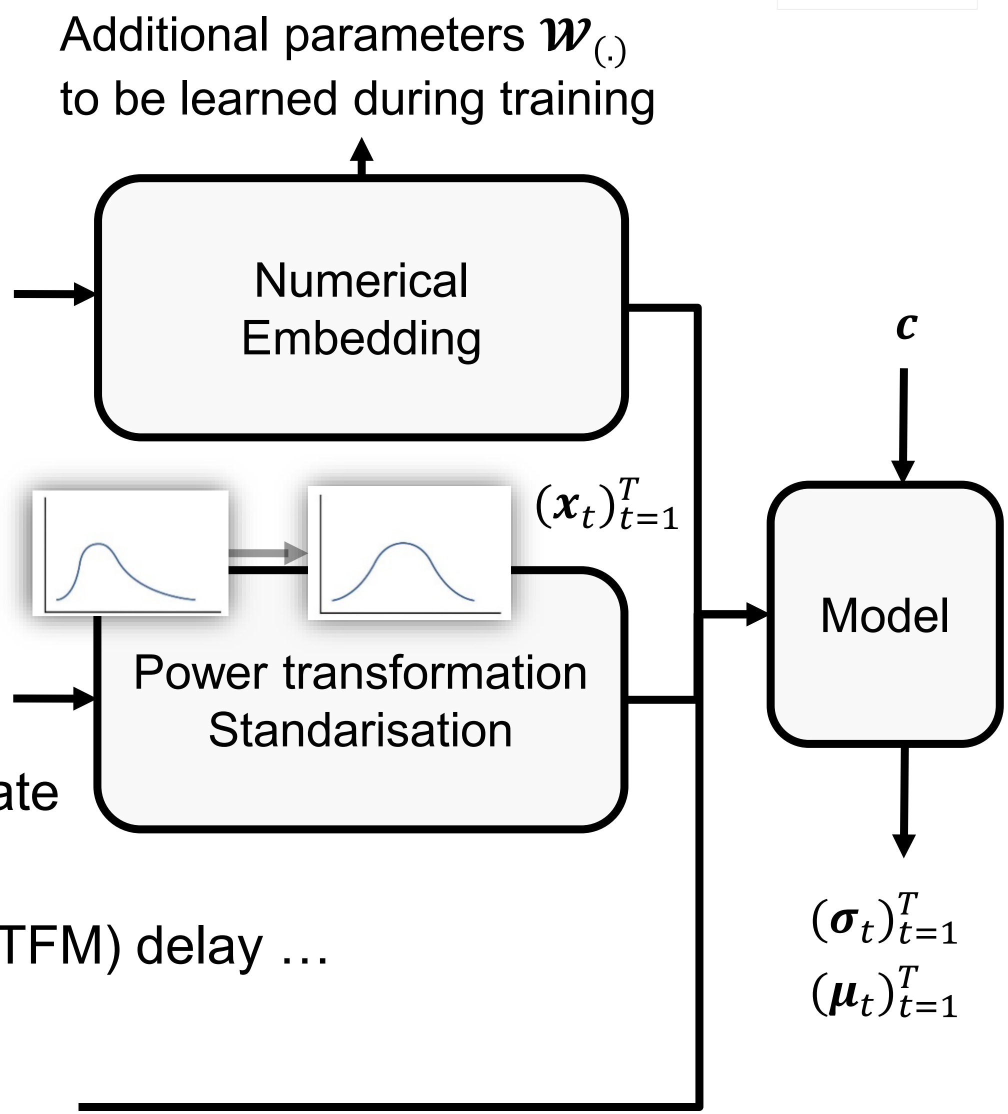
Categorical features:

- Departure airport
- Destination airport
- Flight state (e.g., terminated)
- Airline …

Numerical features:

- Turn around-time
- Departure delay
- Taxi time
- Time since last flight data update
- Time to SIBT
- Air traffic flow management (ATFM) delay …

- Relative flight in the sequence



Static features ($\boldsymbol{c}$) – Per aircraft

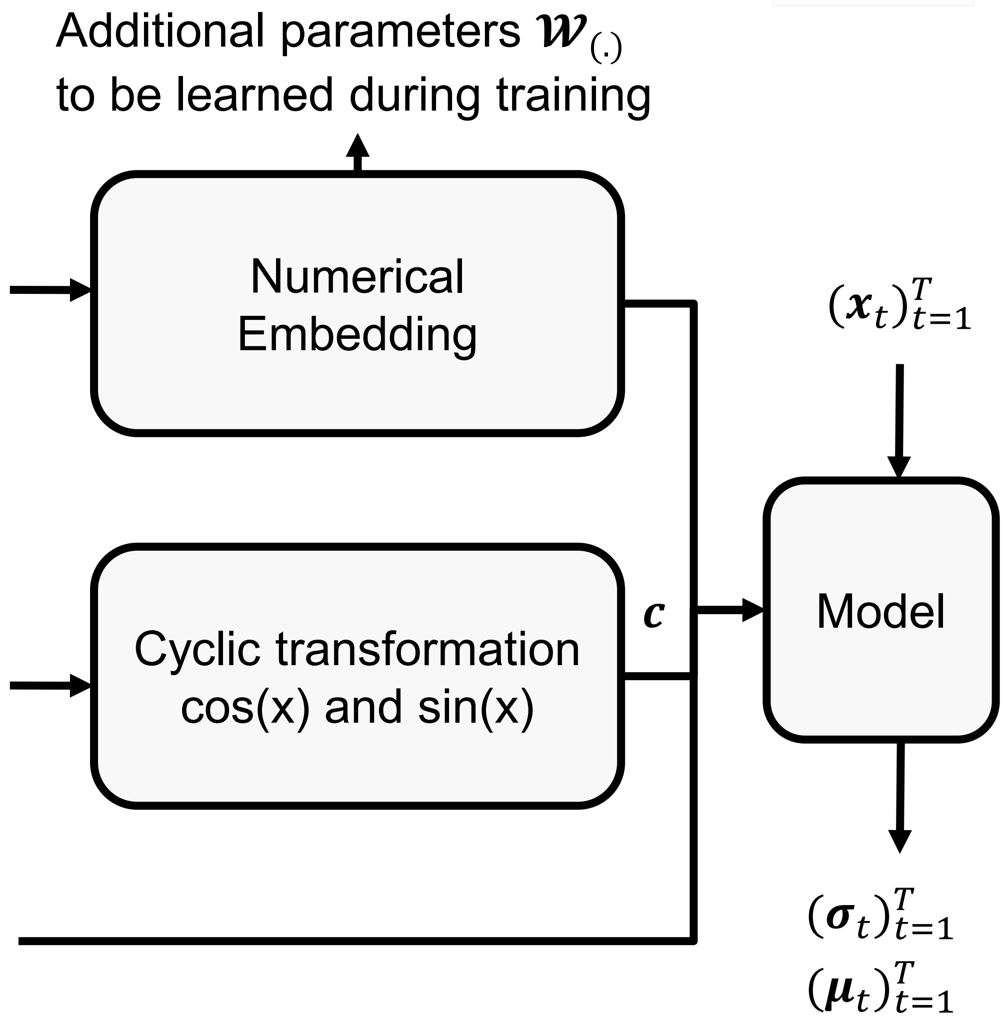
Categorical features:

- Aircraft type

Numerical features:

- Hour of the day
- Day of the week
- Month of the year

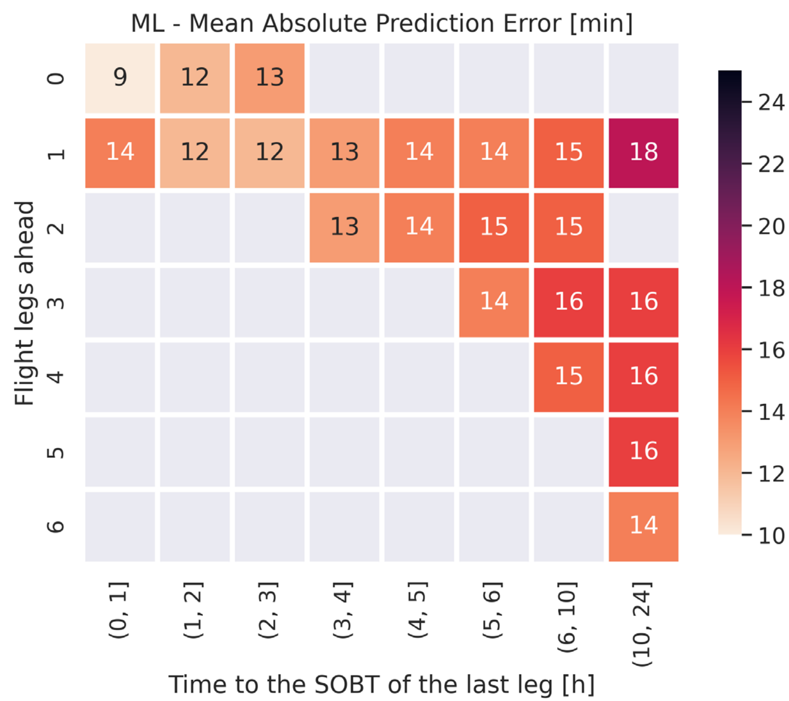
- Length of the sequence

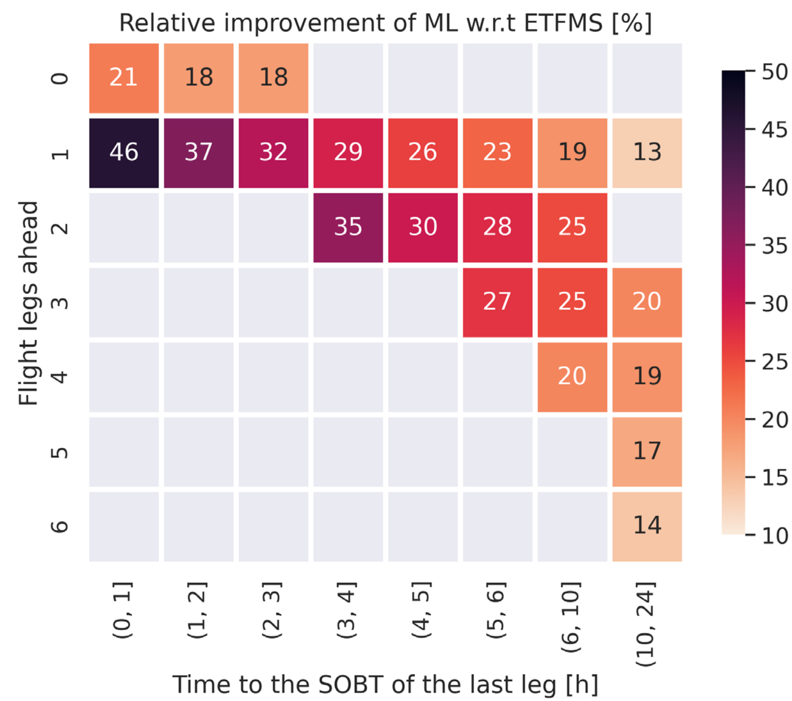


Data

- History of enhanced tactical flow management system (**ETFMS**) flight data (EFD) messages of all flights that were operated during **2019**
  - Departure and destination airports
  - Airline
  - Estimated off-block time (EOBT)
  - ATFM delay
  - Taxi time …
- Corresponding **airline schedules**
  - Scheduled in-block time (SIBT)
  - Scheduled off-block time (SOBT)

Results – January-May 2019





Discussion

- So far, lack of explainability. Methods based on gradients (e.g., integrated gradients) can help …
- Model may learn human *re*actions. Is this behaviour actually desired? How to identify human actions? …
- A lot of other external factors are not (explicitly) considered by the model. However, they may have an important impact on the delay evolution and propagation. All in all obtaining a perfect prediction in such noisy and uncertain environment is not feasible …
- A flight (or an aircraft) is not alone in the network. Interactions between entities are not considered by the models. Could graph neural networks capture such interactions? Community detection algorithms could also help to *reduce* the size of the problem …