

A Generalisable Machine Learning Framework for Taxi Time Prediction at A-CDM Airports

Angelika Swatowska, Lucie Gabagnou
SESAR Innovation Days 2025

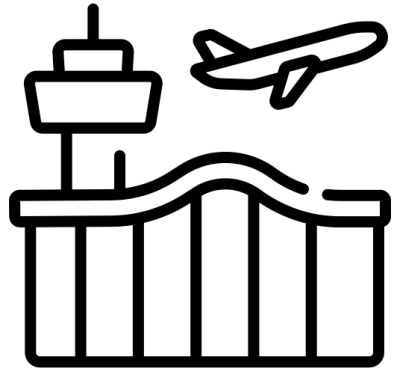
Better taxi time estimations through ML

OBJECTIVE: Deliver realistic, operationally-driven taxi-time predictions

Introduction

- Objective and research GAP

Model inputs (FEATURES)



Airport

Source:
APDF/Airport
Corner

Aircraft

Source: BADA

Flight

Source:
Airport Operator
Data Flow
(APDF)

METAR
METEOROLOGICAL AERODROME REPORT

**METAR OMAA 091000Z 32015
KT 9999 FEW030 SCT040
28/16 Q1012**

OMAA
ICAO CODE FOR
ABU DHABI INTL AIRPORT

FEW030 SCT040
FEW CLOUDS AT 3,000 FT,
SCATTERED AT 4,000 FT

091000Z
9TH DAY,
10:00 UTC

28/16
TEMPERATURE 28°C,
DEW POINT 16°C

Q1012
PRESSURE (QNH)
1012 HPA

9999
VISIBILITY 10 KM
OR MORE

NOSIG
NO SIGNIFICANT
WEATHER CHANGE
EXPECTED

COMMON ABBREVIATIONS

CODE	MEANING
FEW	Few clouds (1-2 oktas)
SCT	Scattered (3-4 oktas)
BKN	Broken (5-7 oktas)
OVC	Overcast (8 oktas)
CAVOK	Ceiling And Visibility OK
TSR	Vertical Visibility

Weather & Calendar

Source:
Meteorological
Aerodrome Reports
(SADIS)

SESAR INNOVATION DAYS 2025

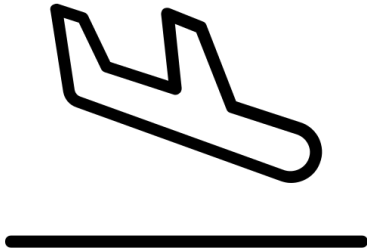
METHODOLOGY

Problem statement



ARRIVALS (TAXI IN)

$$AXIT = AIBT - ALDT$$



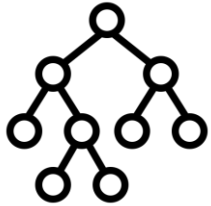
DEPARTURES (TAXI OUT)

$$AXOT = ATOT - AOBT$$

Methodology Overview



Training dataset (features)



ML Model (GBDT CatBoost)



Target (AXOT/AXIT predictions)



Evaluation & interpretation

Machine Learning Model



CatBoost

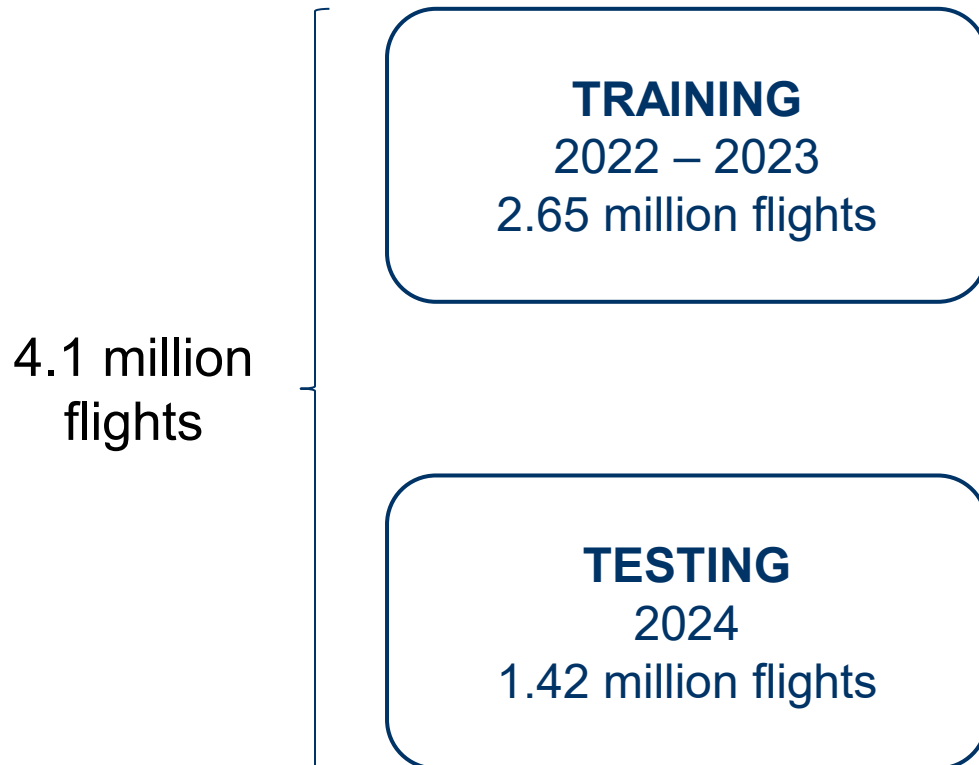
- **Supervised Learning Regression Problem**

continuous value in *min*

- **Gradient-Boosting Decision Trees**
 - *Handling of categorical variables*
 - *Ordered boosting*
 - *Interpretability for operations*

Other tested models: linear regression, Ridge, Random Forest

Evaluation setup



- **Temporal split** → OPERATIONAL USAGE simulation
- Top and bottom 0.1% **extreme values removed** per airport/phase
- **K-Fold Cross-Validation** (CV) → 5 folds
- Stratification per airport/phase: 12 strata

Error metric

Main loss metric: Mean Absolute Error

$$MAE = \frac{1}{n} \sum |y - \hat{y}|$$

Divide by the total number of data points

Predicted output value

Actual output value

Sum of

The absolute value of the residual

$$MAPE = \frac{100\%}{n} \sum \left| \frac{y - \hat{y}}{y} \right|$$

Multiplying by 100% converts to percentage

The residual

Each residual is scaled against the actual value

Evaluated against BASELINE

- Standard Taxi Times (STTs) from airports
- Customizes : mean/median historical values
- Sometimes stratified by A/C type

L	M	N	O	P
	AIRPORT	RWY	STAND	DURATION
	LSZH	10	101	
	LSZH	10	102	
	LSZH	10	103	
	LSZH	10	104	
	LSZH	10	105	
	LSZH	10	106	
	LSZH	10	107	

example of STT (LSZH airport) data

Model selection

- **CatBoost demonstrated best performance, even if difference is marginal**
 - MAE error in (min)

Model	MAE (min)	RMSE	Training time (s)
Linear Regression	2.92	4.21	2.84
Ridge	2.92	4.21	2.86
Random Forest	2.92	4.34	11.53
CatBoost	2.73	4.05	7.09

CatBoost hyperparameters

Parameter	Value	Description
Iterations	300	Number of boosting trees
Learning Rate	0.10	Step size at each iteration
Depth	8	Maximum tree depth
L2 leaf reg.	6.6	L2 regularization strength
Random strength	6.0	Noise added for regularization
Loss function	MAE	Objective minimized

Hyperparameters optimized using the Optuna framework to minimize validation MAE

SESAR INNOVATION DAYS 2025

RESULTS

RESULTS – AXIT (ARRIVALS)

Model vs. BASELINE

- With **MAE reductions ranging from 2.8% to 44%**, the model consistently outperforms Standard Taxi Times Baseline.
- Largest gains occur at Prague, Brussels, Zurich, and Vienna Airports.

> -29% ↓

Airport	MAE MODEL (min)	MAE BASELINE (min)	MAPE MODEL (%)
LKPR (Prague)	1m02s	1m51s	20.5%
EBBR (Brussels)	1m05s	1m31s	18.4%
LSZH (Zurich)	0m54s	1m37s	16.5%
LOWW (Vienna)	0m52s	1m19s	16.8%
LFPO (Paris Orly)	1m20s	1m24s	19.7%
LTFM (Istanbul)	2m21s	2m25s	17.4%

RESULTS – AXOT (DEPARTURES)

Model vs. BASELINE

- With **MAE reductions ranging from 1% to 52%**, the model consistently outperforms Standard Taxi Times Baseline.
- Largest gains occur at Vienna, Zurich, Prague and Vienna Airports.

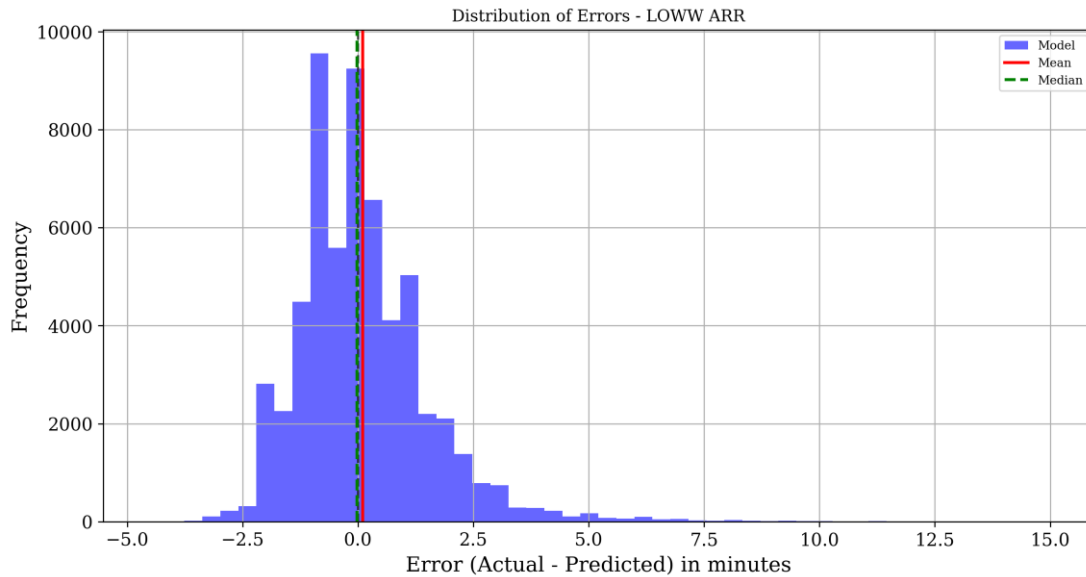
> -28% ↓

Airport	MAE MODEL (min)	MAE BASELINE (min)	MAPE MODEL (%)
LOWW (Vienna)	1m42s	3m31s	19.0%
LSZH (Zurich)	2m18s	3m34s	19.2%
LKPR (Prague)	1m55s	2m48s	17.8%
EBBR (Brussels)	1m56s	2m41s	17.2%
LFPO (Paris Orly)	2m02s	2m03s	19.5%
LTFM (Istanbul)	3m32s	3m45s	20.1%

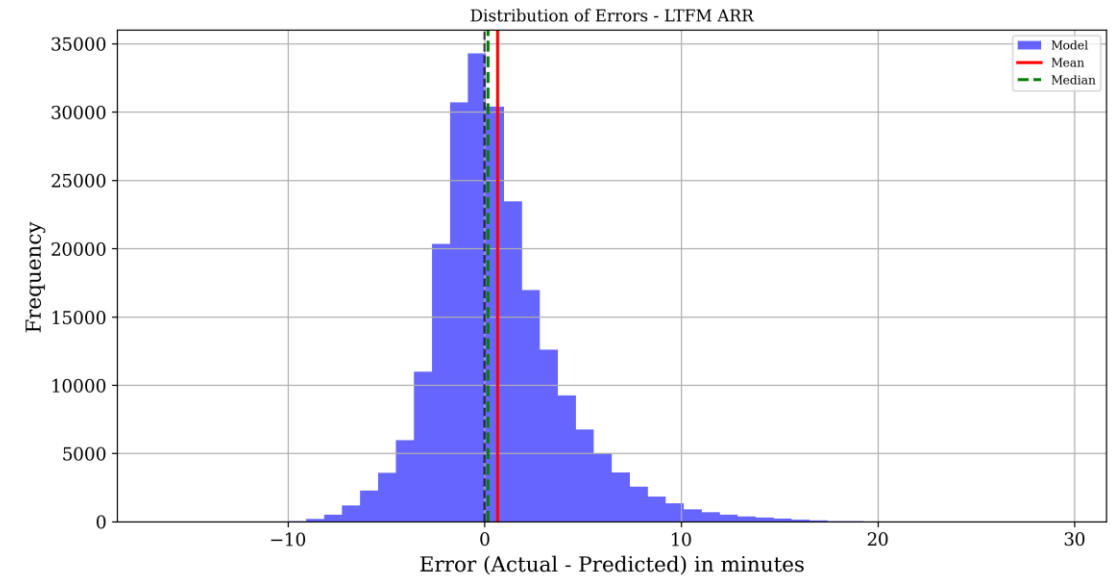
RESULTS – AXIT (ARRIVALS)

ERROR DISTRIBUTION

Taxi-in (AXIT) error distribution
Vienna (LOWW) Airport



Taxi-in (AXIT) error distribution
Istanbul (LTFM) Airport

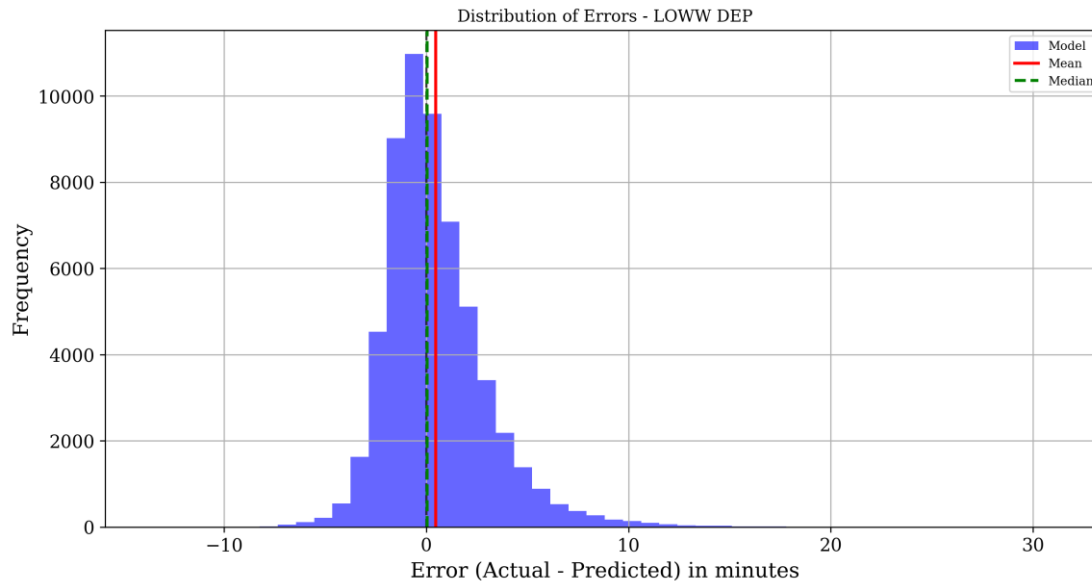


- **Sources of uncertainty:** *runway-to-gate routing complexity, runway-crossing requirements, congestion near apron entry points.*

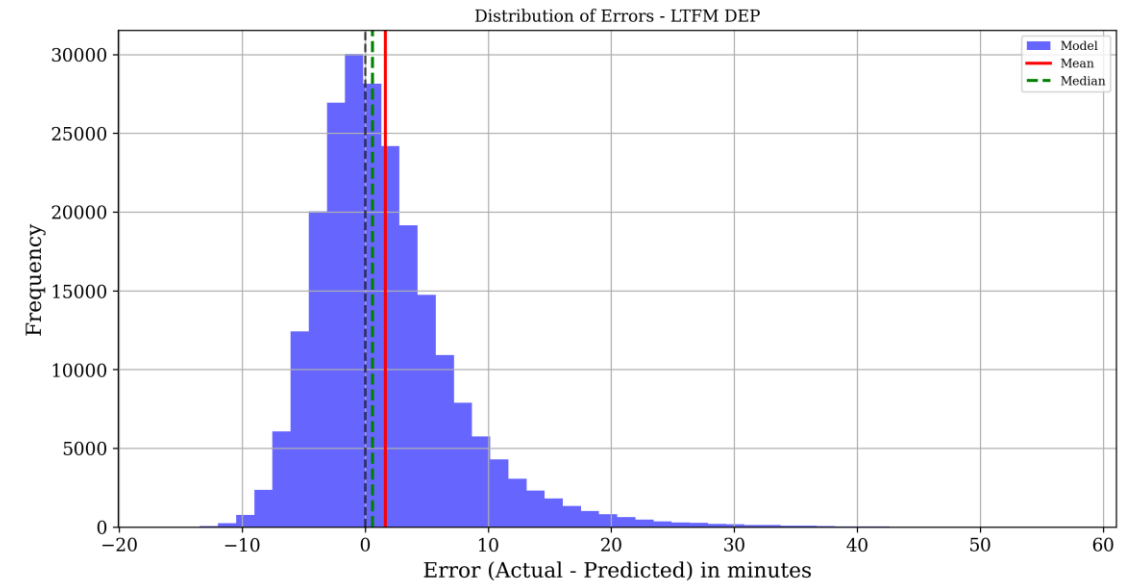
RESULTS – AXOT (DEPARTURES)

ERROR DISTRIBUTION

Taxi-out (AXOT) error distribution
Vienna (LOWW) Airport

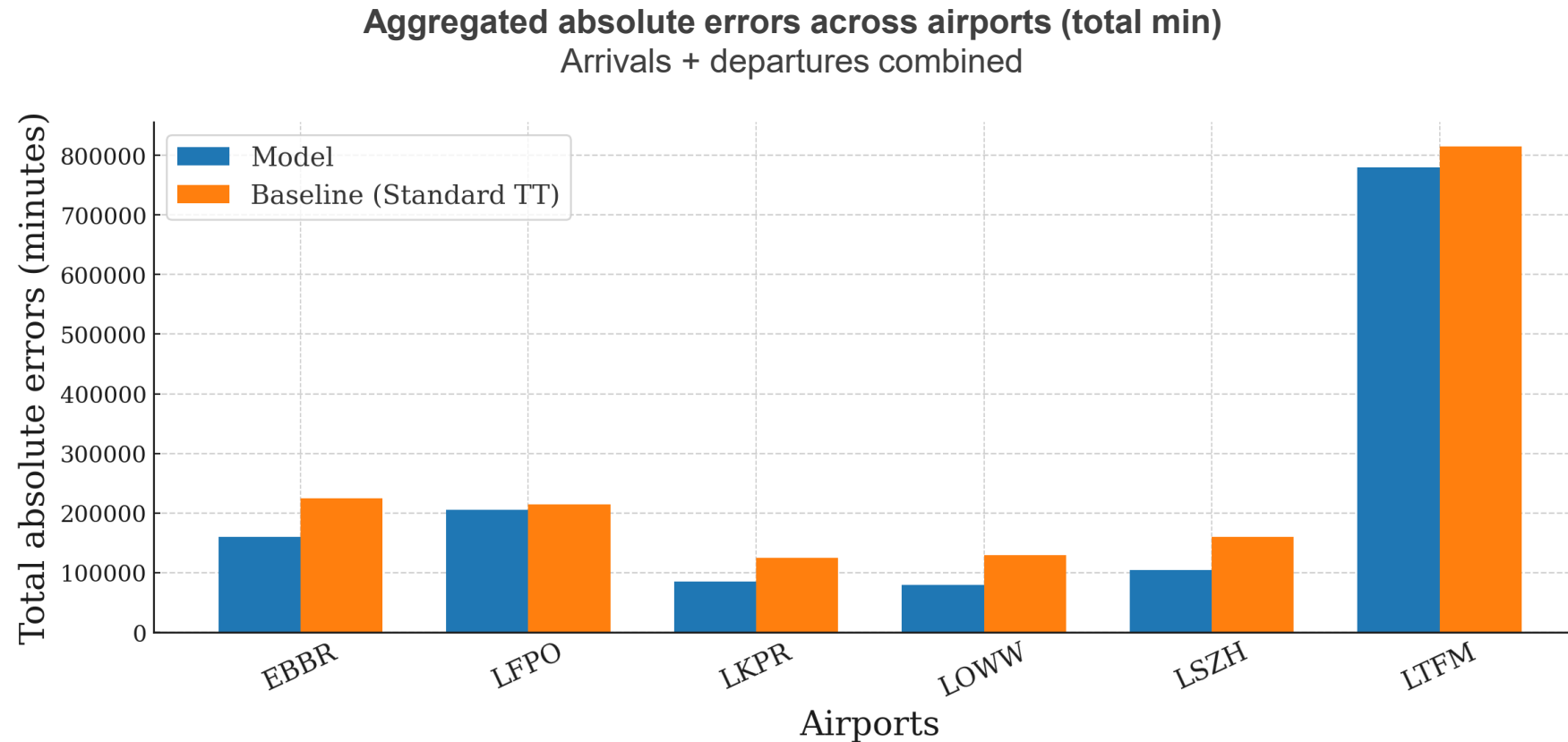


Taxi-out (AXOT) error distribution
Istanbul (LTFM) Airport



- **Sources of uncertainty:** *pushback sequencing, queueing at runway, runway configurations, de/anti-icing.*

RESULTS – AGGREGATED ERRORS



- **Consistent operational time savings across all airports:** 30k minutes saved at Istanbul airport, 50-65k min saved at Brussels, Vienna, Prague, and Zurich.

RESULTS – INFLUENTIAL FACTORS

- Operational and airport-surface factors dominate model predictions,** accounting for 71% of total importance for arrivals and 62% for departures.
- Non-surface layout factors play larger role in AXOT (departures) operations.

Feature	Share (%) AXIT	Share (%) AXOT
Phase	23.39%	29.07
Runway	31.60%	19.12
Stand	16.30%	13.50
Airport	8.31%	9.32
De-/Anti-Icing	4.94% 84%	7.37 78%
Airline	4.70%	5.44
Aircraft MTOW	2.03%	3.00
Hour of Day	1.97%	2.90
Aircraft Type	0.98%	1.79

>61%

KEY TAKEAWAYS

DISCUSSION

- **CatBoost model generalizes well across A-CDM airports with standard inputs.**
- **Consistent gains over STTs, with stronger relative improvements at regional airports.**
- **Runway, phase, and stand dominate the predictive signal**
- **Model captures real-time variability**

POTENTIAL IMPROVEMENTS

- **Add surface-congestion indicators (queues, occupancy, A-SMGCS).**
- **Richer weather features for extreme conditions.**
- **Airline operational data (pushback sequencing, turnaround state).**
- **Probabilistic outputs for DPI/flow-management use.**
- **Adaptive retraining for infrastructure.**