



SUPPORTING
EUROPEAN
AVIATION

Flight Load Factor Predictions Based On Analysis Of Ticket Prices And Other Factors

SESAR Innovation Days 2025

EURV
NOVA





SUPPORTING
EUROPEAN
AVIATION

Miroslav Spak, Eurocontrol
Lorenzo Frigerio, Euranova
Lenka Hanakova, CTU in Prague
Vladimir Socha, CTU in Prague
Rocio Barragan Montes, Eurocontrol
Vincent Treve, Eurocontrol

SESAR Innovation Days 2025

EUR
NOVA



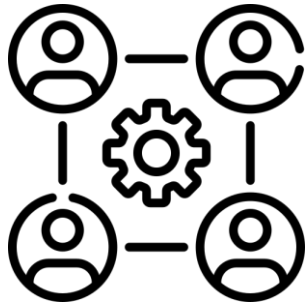
A-G



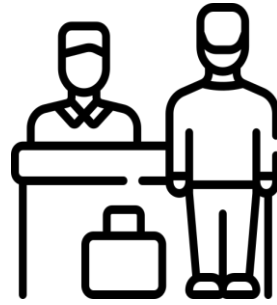
The ability to forecast traffic and size operations accordingly is a determining factor for airports



Passenger Demand to support Airport Operations Planning



Improved **resource
management**

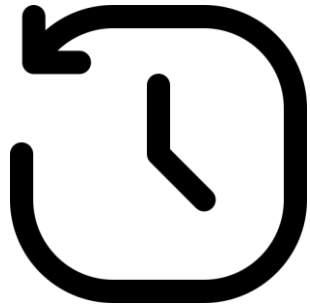


Enhanced **passenger
experience**



Operational
resilience

How can the demand be forecasted ?



Historical data



Scheduled **traffic**
evolution



External information such
as **ticket prices**

Improve the ability to forecast future passenger demand

Evaluate the effectiveness of a passenger demand predictive machine learning model based on ticket price information

We introduced a novel synthetic variable known as the "Price Ratio" which tracks the demand evolution

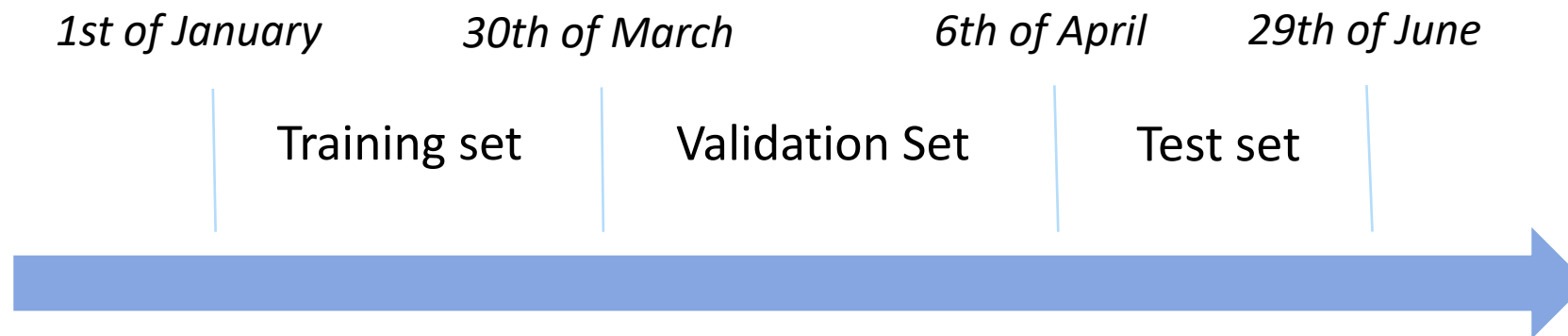
Dataset

Collect the data

Combine ticket prices and historical load factors

Data was collected for **72 818** flights between the 1st of January 2025 and the 29th of June 2025.

Historical data was collected for **5 airports** while an external data broker provided ticket information



Price Ratio

**Price Ratio
tracks the
evolution of the
price on a route**

The Price Ratio compares the current price with the historical one, it is computed as:

$$PR_i = \frac{TP_i}{RP_i}$$

PR_i : price ratio

TP_i : ticket price of the lowest available fare

RP_i : median price calculated over the last 56 days for a given flight i

Train a machine learning model based on Price Ratio and other information to predict the load factor

Gradient boosting regressor based on the following features:

- **Price Ratio:** price indicator computed at D-10;
- **Price:** ticket price of the lowest available fare at D-10;
- **Departure/Arrival airport:** one-hot encoding of the departure/arrival airport;
- **Airline:** one-hot encoding of the airline;
- **Seats:** Total number of seats on the aircraft;
- **Time of day:** morning, afternoon, evening or night;
- **Frequency:** number of flights per day on a given route;
- **Low cost:** true if the airline is a low cost airline;
- **Weekday:** day of the week expressed in number.

Convert load factor to predicted passengers

Predicted load factor can be then converted into number of passengers:

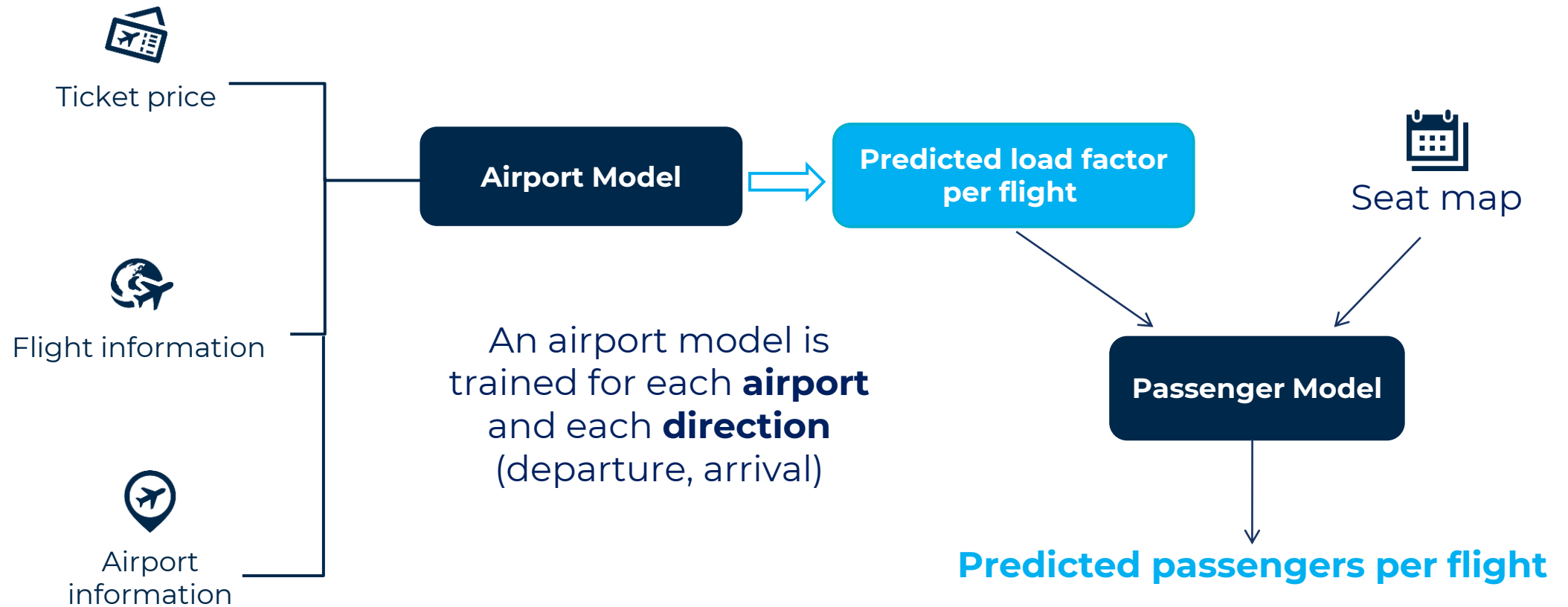
$$PAX_i = LF_i \cdot Nseats_i$$

PAX_i : predicted number of passengers

LF_i : predicted load factor

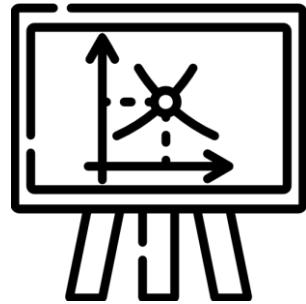
$Nseats_i$: number of available seats on flight i

Airport and Passenger Models



Validation

To **assess** the performance of the models, we designed a set of **validation experiments**

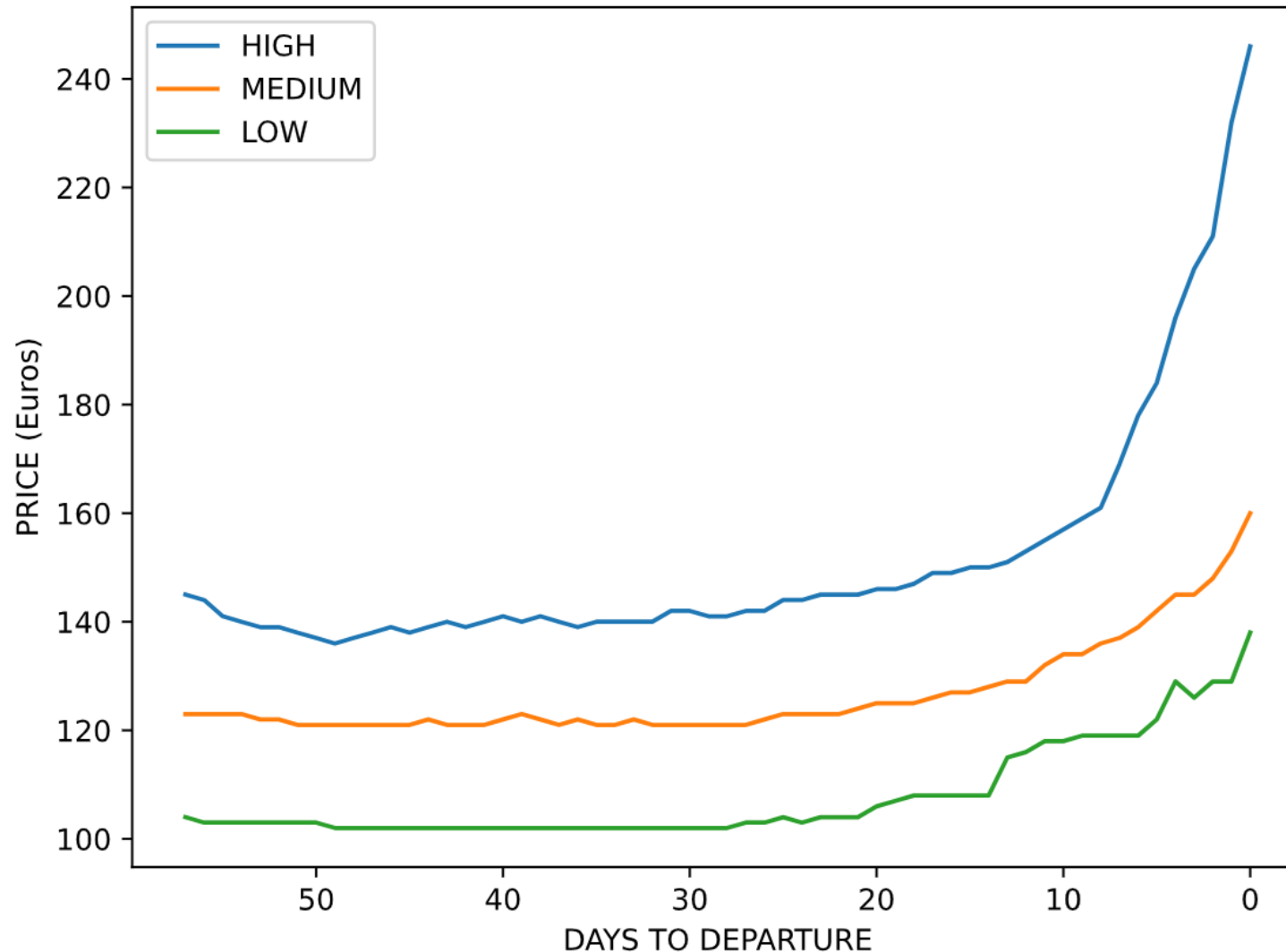


Explore the **relation** between
prices and **passenger demand**



Quantitatively **evaluate** the
performance of the **airport**
and **passenger models**

Validation - Ticket Price Evolution



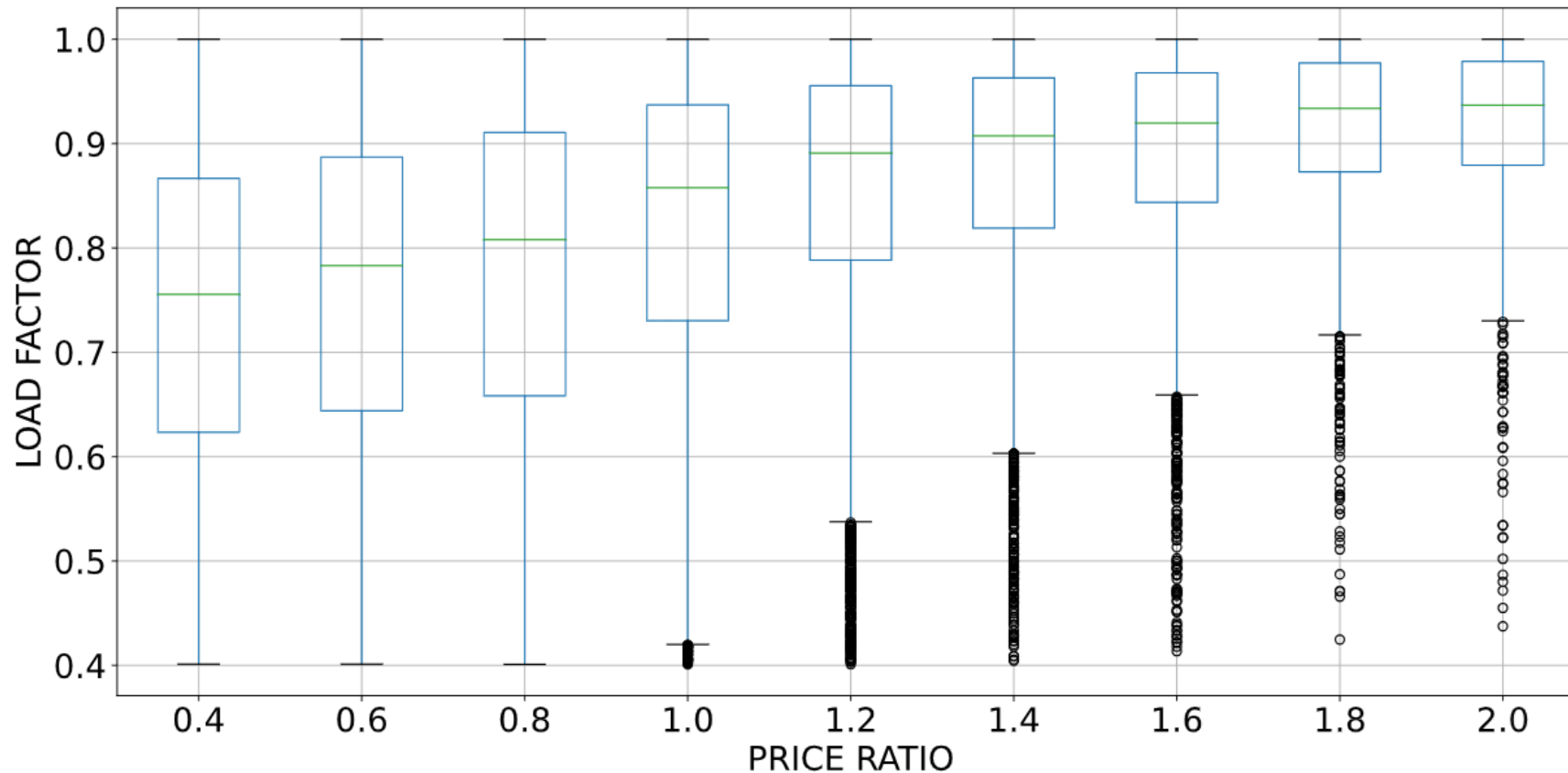
**Price evolution over time
for flights with different
demand**

HIGH demand: $LF \geq 90\%$

MEDIUM demand: $60\% < LF < 90\%$

LOW demand: $LF \leq 60\%$

Validation - Price and Demand



There is a clear relation between Price Ratio and Load Factor

Flights with an **high price ratio** tend to have an **high load factor**

Metrics to evaluate the Airport and Passenger Models

Error per flight

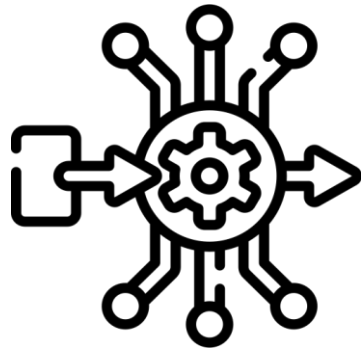
$$\textit{FlightBasedLevelError} = \frac{1}{n} \sum_{i=1}^n \textit{PredictedLoadFactor}_i - \textit{RealLoadFactor}_i$$

Error per day

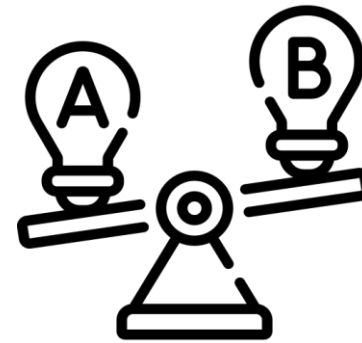
$$\textit{DailyLevelError} = \frac{1}{N} \sum_{j=1}^N \frac{\textit{PredictedDailyPassengers}_j - \textit{RealDailyPassengers}_j}{\textit{RealDailyPassengers}_j}$$

Validation - Baseline Model

To assess the performance of the models, we set up a comparative **baseline**:

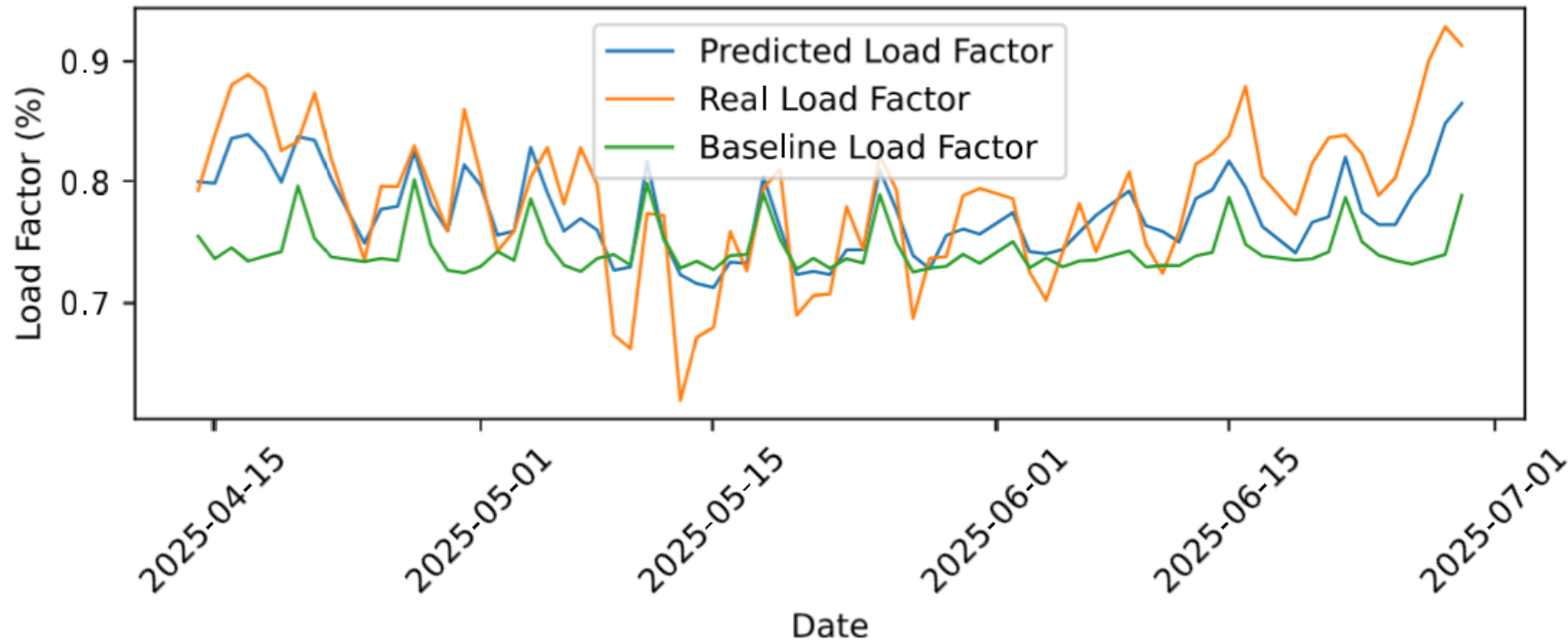


Airport and Passenger
models



Baseline model without
access to ticket information

Validation - Airport Model vs Baseline



The Airport Model captures demand peaks over the Easter period and the beginning of the summer

On the contrary, the baseline cannot reproduce the demand fluctuations over the test period and it remains flat

FlightBasedLevelError

Type	MAE-7days	MAE-10days	MAE-14days
Airport Model	10.8 pp	11.2 pp	11.3 pp
Baseline Model	13.7 pp	13.7 pp	13.7 pp

DailyLevelError

Passenger Model	5.75 %
Baseline Model	8.5 %

**Airport and
Passenger Models
perform
significantly better
than the baseline**

Improvement is significant
both at flight based level
and at daily level

Conclusion

The inclusion of ticket prices in the airport model provides a **significant improvement** to performance

Price Ratio proved to be an **effective metric** to track the demand evolution

FlightBasedLevelError around 10 pp per flight constitutes **credible input** for airport staff when preparing the **operational roster**.

The model was validated for **several diversified airports** in different European regions

Performance was maintained during a **challenging test period** which compassed Easter holidays and the different May and June bank holidays.

Pricing policies are constantly **changing**, and this causes deterioration of the predictions on certain conditions



SUPPORTING
EUROPEAN
AVIATION

Thank you!

lorenzo.frigerio.ext@eurocontrol.int

www.eurocontrol.int



NETWORK
MANAGER

