Acta Polytechnica CTU Proceedings **52**:17–23, 2025
https://doi.org/10.14311/APP.2025.52.0017

CALCULATION OF QUEUE FORMATION AND QUEUE LENGTH AT AIRPORT CHECK-IN COUNTERS

Petr Had*, Jiří Volt, Slobodan Stojić

Czech Technical University in Prague, Faculty of Transportation Sciences, Department of Air Transport, Konviktská 20, 110 00 Praha, Czech Republic

* corresponding author: hadpetr1@fd.cvut.cz

Abstract. This paper focuses on calculation of airport passenger queue formation and its length. The main goal of the paper is to evaluate the appropriateness of the combination of machine learning and determistic algorithm for queue length determination. The study focuses on the departing passenger flows to the check-in counters. In this study the Random Forest algorithm was applied for passenger number prediction, and followed by application of the deterministic algorithm for passenger queue length estimation. The study was performed using the real data from the Václav Havel Airport Prague operations. Presented results indicate that the combined approach of machine learning and a deterministic algorithm offers useful insights into passenger queue formation. This approach provides valuable insights into queue dynamics at check-in counters, enabling improvements in planning and management of airport operations during operational peak.

Keywords: Queue prediction, machine learning, passenger flows, airport processor.

1. Introduction

Airports around the world are facing increasing pressure to improve operational efficiency due to rising passenger numbers. One of the key challenges that airports have to face is queuing at individual processors such as check-in counters, security checkpoints and border control. Inefficient queue management leads to longer waiting times, which has a negative impact not only on passenger satisfaction but also on the operating costs of both airport operators and airlines.

Traditional approaches to determining queue lengths rely primarily on historical data and static capacity models, which often fail to respond flexibly to dynamic changes in demand or to specific airport conditions at a given time. With increasing demands for more accurate planning and optimisation of airport processes, it is necessary to look for innovative ways to predict and manage flows of the passengers.

The objective of this study is to determine the queue lengths at airport terminal processors through application of the machine learning algorithms for prediction of the number of passengers arriving on individual flights, followed by calculation of the queue length using a deterministic algorithm. This approach enables the prediction of queue lengths in short time intervals for specific check-in counters, allowing better planning and more flexible management of the airport processors.

2. Research background

The management of queues at the airport terminals has become an increasingly critical issue, as passenger volumes continue to grow. Effective queue management is essential for reducing waiting times, improving passenger experience, and maintaining airport operational efficiency. Various approaches have been developed to model, analyse, and optimise airport queuing systems, ranging from classical queuing theory to more advanced techniques using machine learning algorithms.

2.1. Queuing theory and simulation models

Queuing theory represents a strong foundation for airport processes performance analysis. Thiagaraj and Seshaiah [1] present how queuing models are utilised in delay prediction and assessment of airport capacity under varying demand conditions. Queuing models provide insights into how to manage the allocation of service counters in order to minimise waiting times during busy time periods. Sharma et al. [2] demonstrate the use of simulation models to predict and mitigate queue lengths at airport security checkpoints, reducing the likelihood of queuing during operational peaks.

Another application of queuing theory is presented by Xue et al. [3], who discusses the synergy between queuing models and machine learning for scheduling tasks in high-traffic areas. The combination of these approaches ensures that service level objectives (SLOs) are met while maximising resource utilisation. This principle can be also applied to airport queues in order to improve efficiency of passenger flows management.

Al-Sultan [4] explores the optimisation of airport check-in processes through a combination of integer programming (IP) and simulation models. While this approach focuses on the optimisation of resource allocation, it emphasises the importance of simulation for improving passenger throughput and reducing congestion. Simulation techniques are often used together with queuing models to provide more realistic and detailed predictions of airport capacity and passenger flow.

2.2. Machine learning

Recent advances in the area of machine learning have enabled more dynamic and accurate predictions of queue lengths. Wang et al. [5] use machine learning algorithms, such as decision trees, random forests, and support vector machines (SVM), to predict traffic flow patterns. The study highlighted the importance of data-driven approaches for predicting non-linear traffic conditions, which is a concept also applicable to passenger flow at the airports. By incorporating real-time data, machine learning models are able to adjust resource allocation in order to minimise waiting times and prevent development of bottlenecks.

Moreover, Xue et al. [3] introduce a hybrid model combining neural networks and queuing theory for prediction of the traffic intensity. This hybrid approach can be adapted for airport operations, where accurate predictions of passenger arrivals and queue lengths are crucial for optimisation of the terminal resources allocation. Machine learning models enhance the queuing model's ability to predict variations in passenger flow, providing more responsive and adaptive management of the queues at the airport processors.

Loureiro et al. [6] developed a machine learning approach for prediction of waiting times across different types of queues, including those at the airports. By using supervised learning algorithms, such as decision trees and neural networks, this study shows that machine learning can offer improvements in the accuracy of queue time predictions. Such models are particularly valuable in dynamically changing environments, where they can enhance real-time decision-making for resource allocation and improve the overall passenger experience.

2.3. Hybrid approaches to queue management

The integration of queuing theory with machine learning has shown promise in optimising queue management. Chocron et al. [7] demonstrate how combining queuing models with machine learning improves delay prediction in multi class service environments, such as airports. This hybrid approach allows for better precision in resource allocation by accounting for the different needs of various passenger classes, leading to reduced waiting times and improved overall efficiency.

Additionally, Loureiro et al. [6] highlight the potential of machine learning to outperform traditional queuing theory in environments with high variability, such as airport terminals. By continuously learning from real-time data, machine learning models can predict queue lengths more accurately, allowing airports to dynamically adjust staffing and service levels to meet fluctuating passenger demand.

Based on the limitations of the existing experiments, the presented research focuses attention on the integration of machine learning algorithm with deterministic queue length calculations, specifically applied for passengers arrivals and queueing at the airport terminal check-in counters on a per-flight basis. Limiting the approach to the specific process brings the set of the data and available data sources required for the estimations. The selected method combines the machine learning algorithm for passenger arrival predictions with a deterministic algorithm for queue length estimation in 15 minute intervals, providing a flexible and accurate tool for operational planning at airport terminals. This dual approach leverages the strengths of machine learning for forecasting number of arriving passengers, while using deterministic methods to model queue dynamics based on known processing rates and counter availability. It must be noted that proposed estimation approach is a first phase of the wider research intention to evaluate a possibility of more advanced concept of active passengers flow management. Therefore, the research question in this phase of the research is targeting the appropriateness of the passengers flow estimation using the combined machine learning and traditional queueing techniques.

3. Methodology

The approach used in this study to predict queue lengths combines machine learning algorithm for prediction of number of passengers and deterministic algorithm for queue length calculation. The study is limited only to the predictions of the passenger queue length formed at the airport check-in counters. This involves only the departing passenger flows arriving to the airport for the flight. For the purposes of this study the real data from the Václav Havel Airport Prague Terminal 2 were used.

3.1. Available data

To develop the prediction model and calculate the length of the queues, historical operational data and passenger arrival information were utilised. The dataset included a general arrival curve for departing passengers, the proportion of passengers using check-in counters, and details on all departing flights from Terminal 2 (serving flights to and from the Schengen area) during the period from June to August 2024. To maintain confidentiality of the provided information, the presented data have been anonymized. In total, the dataset covered 10 898 flights. The following parameters were available for each flight:

- Flight number;
- Airline;
- Destination;
- Scheduled time of departure (STD);

- Designation of all check-in counters;
- Scheduled opening and closing times of all the check-in counters;
- Number of passengers;
- Number of checked-in bags.

All timestamps used in this research were in the local time zone of Prague airport, i.e. Europe/Prague.

3.2. Machine learning for predicting the number of arriving passengers

Given the available data, the following predictors were chosen to build the prediction models:

- Airline;
- Destination;
- Number of checked-in bags;
- Low-cost indicator;
- Weekday;
- Month;
- Part of the day (night, morning, afternoon);
- Max. number of available check-in counters;
- Handling company;
- Aircraft type.

Based on the chosen predictors, three machine learning algorithms were selected to predict the expected number of passengers arriving at the check-in counters. These models were selected for their performance with structured data and their ability to manage complex relationships between the input features and the target variable effectively.

- **Random Forest (RF)**
  This model was selected due to its ability to handle large datasets with mixed data types (categorical and numerical) and its resistance to overfitting by averaging the results of multiple decision trees, resulting in more reliable predictions. Random Forest also provides feature importance scores, allowing to understand influences of the factors to the predictions [8].
- **XGBoost (Extreme Gradient Boosting)**
  XGBoost is a widely-used implementation of GBM (Gradient Boosting Machines). GBM is ensemble learning technique that builds models sequentially, where each new model corrects the errors made by the previous ones. This model was selected due to its high predictive power and efficiency in handling structured data [9].
- **Support Vector Machines (SVM)**
  SVM was selected due to its effectiveness in modeling complex relationships in data with relatively small to medium-sized datasets. SVM constructs a hyperplane that best fits the data points in high-dimensional space. This model can also effectively handle overfitting through the use of regularization and kernel functions [10].

3.2.1. Model training and testing

All models were trained using above mentioned historical data. The dataset was split into training and testing sets, with first 80 % of the data from each month used for training and 20 % for testing. Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and R-squared (R^2) were used to evaluate the performance of each model for prediction of number of passengers. Results are provided in Table 1.

	MAE [PAX]	RMSE [PAX]	R^2
RF	11.05	15.72	0.85
GBM	11.18	15.73	0.84
SVM	11.89	16.98	0.82

Table 1. Performance metrics for selected models for PAX prediction.

Based on the results provided in Table 1, Random Forest (RF) emerges as the best-performing model for predicting the set target value. The importance of each feature is shown in the Figure 1 and ordered from the most impactful feature on the outcome target variable. As shown in the Figure, the aircraft type has the highest importance in creating the prediction with an importance score corresponding to 32 %, followed by the total number of baggage with the importance score of 25 %. The remaining parameters considered had the importance levels of less than 10 % each.

As the Random Forest model performs the best in terms of overall error and variance it will be used for creating the input values for the length calculations of the check-in queues.

3.3. Queue length calculation using deterministic algorithm

This section describes the steps involved in prediction of queue formation and length at the airport terminal check-in counters. The methodology combines deterministic and stochastic elements, including the use of triangular distributions for processing times and deterministic calculations for queue lengths.

To prepare the queuing model, the general passenger arrival curve at the airport was used as one of the key inputs. This curve represents the percentage of passengers arriving at the airport in defined 15-minute intervals leading up to the scheduled time of departure of the flight. Additionally, the model incorporates the passenger processing times at check-in counters, the opening times of the counters, and the proportion of passengers using the check-in counters.

Passenger data, including flight details and airline-specific behaviour (in particular, differences in passenger check-in patterns according to different airlines), were integrated to adjust the number of passengers who really use the airport check-in counters. Each

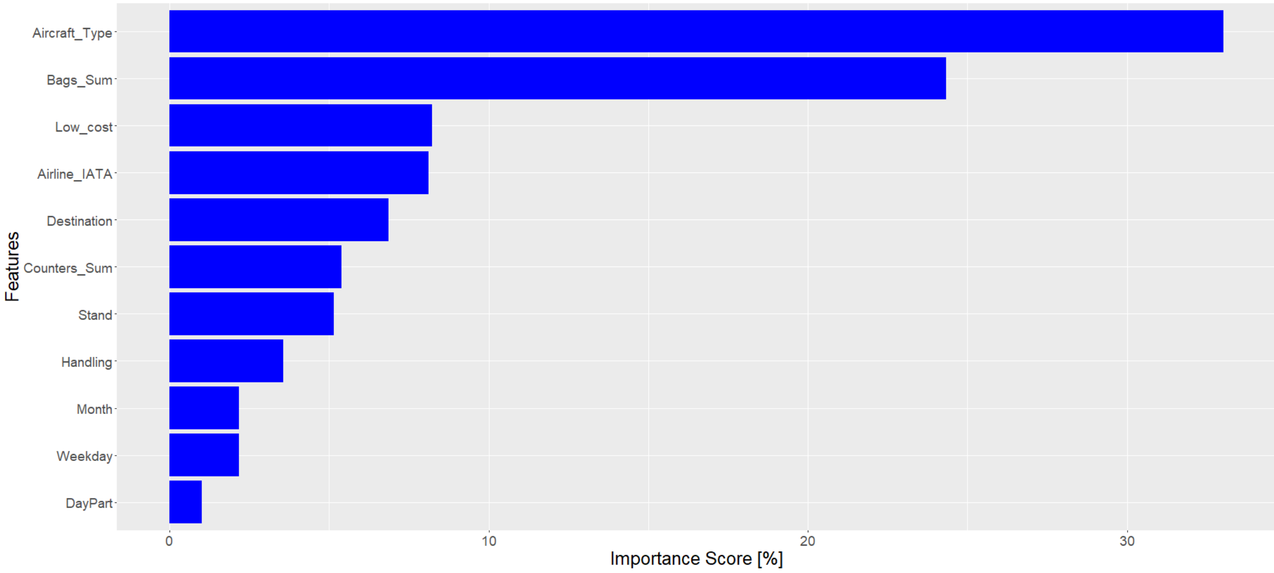

Figure 1. Feature Importance of Random Forest model.

flight's predicted number of passengers was therefore adjusted based on the proportion of passengers using self and online services. This adjustment is important to accurately estimate the number of passengers arriving at the check-in counters.

The number of available check-in counters varies over time. The developed algorithm checks for each time window whether counters are open based on the scheduled opening and closing times and calculate the number of open counters in each time interval. This solution enables the model to dynamically adjust processing capacity based on counter availability. The maximum number of passengers that can be processed in time window j is given by Eq. (1):

$$M_j = C_j \times \frac{T_{\text{window}}}{\overline{t_{\text{process}}}}, \tag{1}$$

where:

M_j is the maximum number of passengers processed in window j,

C_j is the number of open counters in window j,

T_{window} is the length of the time window (e.g., 15 minutes),

$\overline{t_{\text{process}}}$ is the average processing time per passenger.

A triangular distribution is used to account for variability in processing time of each passenger. The processing time for each passenger is sampled from this distribution, defined by a minimum, maximum, and mode.

The queue length for each time window is calculated based on the number of arriving passengers, the number of passengers already in the queue from previous windows, and the processing capacity of the open counters. The processing capacity is determined by the time window length, the processing times, and the number of open counters. This ensures that the number of passengers processed does not exceed the capacity of the available counters in each interval, and any unprocessed passengers remain in the queue.

The formula for queue length calculation is in Eq. (2). The queue length at the end of time window j is the number of passengers left unprocessed from the previous time windows plus the new arrivals, minus the passengers processed during window j:

$$Q_j = \max\left(Q_{j-1} + P_j - M_j, 0\right), \tag{2}$$

where:

Q_j is the queue length at the end of time window j,

Q_{j-1} is the queue length from the end of previous time window,

P_j is the number of passengers arriving in time window j,

M_j is the maximum number of passengers processed in time window j.

The queue length is updated by adding the passengers from the current window, subtracting the passengers processed, and ensuring the queue cannot be negative.

4. Results

The dataset analyzed for calculating queue lengths consisted of flights from the same data used to test the machine learning algorithm. This was because both the predicted and actual passenger numbers were included in the dataset. These covered the last 5 days in June and the last 6 days in both July and August. The analysis included a total of 2010 flights.

The result of the calculation for each analyzed flight was a dataset containing the flight name, scheduled flight date, current arrival curve interval, current

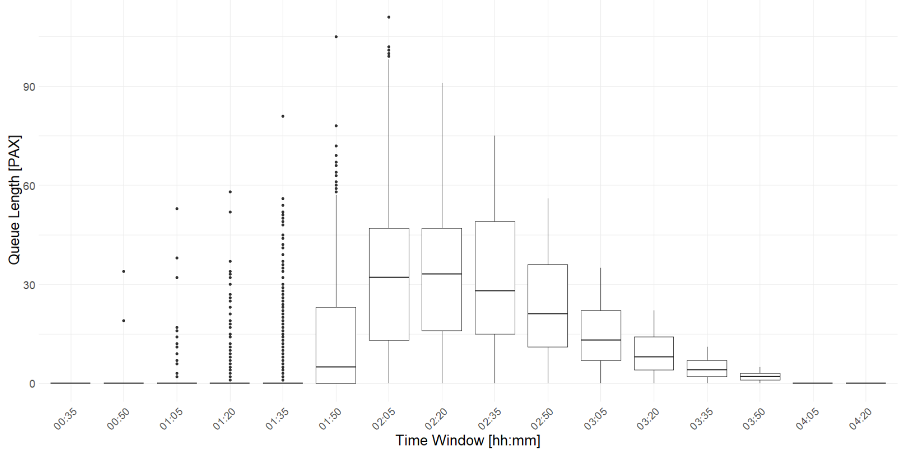

FIGURE 2. Queue length distribution by time window.

queue length in that interval, number of open check-in counters, and number of passengers processed in the given time window. The overall number of records for each flight matched the total number of arrival curve intervals.

Table 2 provides an overview of queue metrics used across different airline types, highlighting the differences in average, median and standard deviation of queue lengths for Legacy and Low Cost airlines. Cumulative metrics for both types of airlines combined are also presented.

Airline Type	Average [PAX]	Median [PAX]	Standard Deviation [PAX]
Legacy	13	2	19
Low Cost	5	0	11
Cumulative	10	0	17

TABLE 2. Descriptive statistics of queue lengths by airline type.

The average, median and standard deviation of queue lengths at various time windows leading up to scheduled time of departure times are presented in Table 3. Each row in the table corresponds to 15-minute time interval before STD in hh:mm format according to the defined arrival curve.

The Figure 2 presents distribution of queue lengths over different time windows before STD. Each of the boxes represents the spread of queue lengths within the time window. This figure provides insights into typical, minimum, maximum queue lengths and shows the overall variability in the queues.

Based on this initial analysis and purpose of this research, the following two metrics were defined and calculated.

4.1. Queue lengths for specific flights across time intervals

For each flight, queue length data were collected at defined intervals of the arrival curve, from 4 hours and 20 minutes to 35 minutes before the scheduled time of departure. The results show how queue lengths fluctuated during these intervals, reflecting the variability in passenger arrival patterns and check-in counter operations.

	Average [PAX]	Median [PAX]	Standard Deviation [PAX]
00:35	0	0	0
00:50	0	0	1
01:05	0	0	2
01:20	0	0	3
01:35	2	0	7
01:50	12	4	15
02:05	32	33	23
02:20	32	33	21
02:35	31	28	22
02:50	23	21	16
03:05	15	13	10
03:20	9	8	7
03:35	5	4	3
03:50	2	2	1
04:05	0	0	0
04:20	0	0	0

TABLE 3. Descriptive statistics of queue lengths by each time window [hh:mm] before STD for both airline types.

For example, for flight ABC123 on July 26th, the queue length peaked at 72 passengers, 2 hours and 5 minutes before the scheduled time of departure. Afterward, the queue gradually decreased as the check-in counters opened, and all passengers were processed 1 hour and 20 minutes before departure.

The progression of the queue for this flight, along with five other similar flights over the following days, is shown in Figure 3.

4.2. Cumulative queue formation during day

A cumulative analysis was also performed to evaluate queue formation during the period of one day. A day

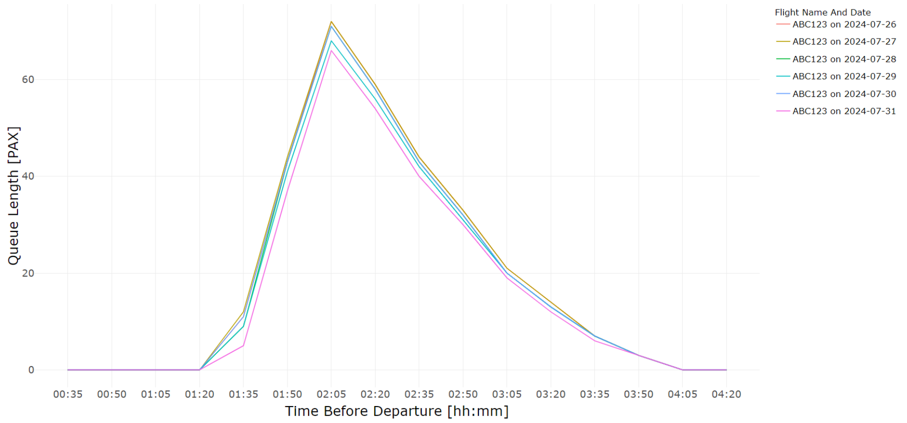

FIGURE 3. Queue lengths for flight ABC123 in last 6 days of July 2024.

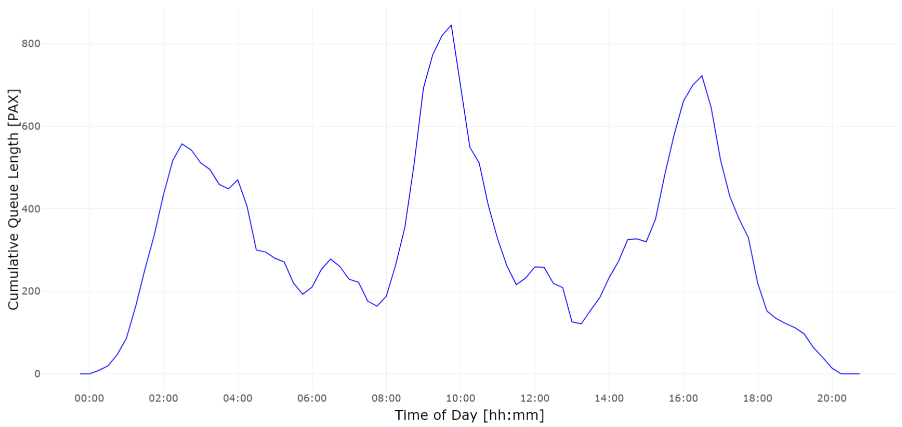

FIGURE 4. Cumulative queue formation on June 28th across all departing flights from Terminal 2.

with the highest number of passengers departing from Terminal 2 was selected. During this day, queue formation was analyzed across all flights operating within the day, aggregating queue lengths from individual time windows of the passenger arrival curve to provide an overall picture of passenger flow. The result can be seen in Figure 4.

Three operational peaks are visible in Figure 4. The busiest period occurred at 9:45 a.m., when the cumulative queue length reached its highest point at 841 passengers across all flights. Two smaller peaks were observed during the day, one around 2:30 a.m. with 556 passengers in the queues and another at 04:30 p.m. in the afternoon with 716 passengers waiting to be processed.

5. Discussion

The combination of machine learning algorithm with a deterministic algorithm for queue length estimation has demonstrated its potential in predicting queue formation at airport check-in counters. Reached results replicate the realistic behaviour of the passenger flow at the chosen processor. These were evaluated through the operational peaks of the executed flight schedule. Queue lengths for the particular flight within 6 days interval showed the same pattern, reaching the maximum values around two hours before the flight. This indicates similar conditions and absence of any kind of perturbations of the schedule and check-in counter allocation.

In terms of cumulative queuing for the chosen day, seasonality presented in three operational waves was correctly indicated. Summer charters peak in the night wave, mid-day and evening waves, clearly show the maximum queue capacity demand during the set time frame.

Based on the results it can be concluded that integration of these two techniques creates the method, which allows for the predictions of passenger numbers and the translation of these predictions into queue dynamics based on specific operational conditions, such as check-in counter throughput and passenger arrival patterns.

Main limitations of the presented approach are related to its accuracy and potential for further enhancements in this characteristic. The use of dynamic or more granular arrival curves with smaller time intervals or differentiated by flight type (e.g., low-cost, charter, or legacy carriers) could improve the precision of queue length predictions. This can allow the model to capture more details regarding passenger arrival patterns and to better reflect the variability in different types of flights.

Limitations are also related to the machine learning algorithm. The predictors, such as information regarding significant events at the flight destinations, could be added to improve prediction accuracy. Limited time period data are also considered a limiting factor.

This study creates a good foundation for the next phases of the research, which will focus on active management of the distribution of arriving passengers. The primary objective is to create a tool, which will actively assign passengers to specific time windows, optimising the flow of passengers through the airport terminal processors and therefore reducing excessive queues on subsequent terminal processors. The queuing curves calculated for the individual flights represents a milestone and foundation for the further development in the research.

6. Conclusions

This study explored the use of a combined approach to predict the length of passenger queues at airport check-in counters. By utilizing real-world operational data from Václav Havel Airport Prague, the Random Forest model was trained and predicted passenger numbers for individual departing flights from Terminal 2. The deterministic model then used these predictions and calculated queue length estimates based on counter availability and passenger processing rates of the check-in counters.

These findings provided valuable insights into passenger behavior upon arrival at the first airport processor, as well as the dynamics of queue formation during the check-in process for individual flights. This information enables further research in the field of active management of departing passengers' arrivals at the airport.

While the methodology offers promising results, further research is needed to refine the accuracy of predictions. The findings also provide a foundation for future work, focusing on active passenger distribution management and coordination with other airport processors to reduce excessive queues and improve overall terminal efficiency.

Acknowledgements

The outcomes of the close cooperation between the Faculty of Transportation Sciences, CTU in Prague, and Prague Airport enabled the development of the presented model. All input data was received according to agreement with the Prague Airport. This work was supported by the internal grant No. SGS24/107/OHK2/2T/16 "Development of an optimisation tool to reduce queues in the airport terminal using managed passenger arrivals" of the Czech Technical University in Prague.

References

[1] H. B. Thiagaraj, C. V. Seshaiah. A queueing model for airport capacity and delay analysis. *Applied Mathematical Sciences* **8**(72):3561–3575, 2014. https://doi.org/10.12988/ams.2014.43179

[2] D. Sharma, D. S. Tege, D. Mehta. Simulation of queuing process of airport passenger at security check-in by using Simulink – MATLAB. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology* **9**(6):15–20, 2023. https://doi.org/10.32628/cseit2390537

[3] J. Xue, F. Yan, A. Riska, E. Smirni. Scheduling data analytics work with performance guarantees: Queuing and machine learning models in synergy. *Cluster Computing* **19**(1):849–864, 2016. https://doi.org/10.1007/s10586-016-0563-z

[4] A. T. Al-Sultan. Airport check-in optimization by IP and simulation in combination. *World Academy of Science, Engineering and Technology, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering* **9**(1):403–406, 2015. https://doi.org/10.5281/zenodo.1100262

[5] H. Wang, X. Wei, J. Yao, Y. Zhang. Traffic flow prediction using machine learning methods. In *2021 3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, pp. 30–35. 2021. https://doi.org/10.1109/MLBDBI54094.2021.00014

[6] C. Loureiro, P. Pereira, P. Cortez, et al. Predicting multiple domain queue waiting time via machine learning. In *Computational Science and Its Applications — ICCSA 2023*, pp. 404–421. 2023. https://doi.org/10.1007/978-3-031-36805-9_27

[7] E. Chocron, I. Cohen, P. Feigin. Delay prediction for managing multiclass service systems: An investigation of queueing theory and machine learning approaches. *IEEE Transactions on Engineering Management* **71**:4469–4479, 2024. https://doi.org/10.1109/TEM.2022.3222094

[8] M. Schonlau, R. Y. Zou. The random forest algorithm for statistical learning. *The Stata Journal* **20**(1):3–29, 2020. https://doi.org/10.1177/1536867X20909688

[9] Y. Zhang, A. Haghani. A gradient boosting method to improve travel time prediction. *Transportation Research Part C: Emerging Technologies* **58**:308–324, 2015. https://doi.org/10.1016/j.trc.2015.02.019

[10] C. D. Manning, P. Raghavan, H. Schütze. *Support vector machines and machine learning on documents*, p. 293–320. Cambridge University Press, 2008. https://doi.org/10.1017/CBO9780511809071.016