https://doi.org/10.1590/jatm.v17.1382
THEMATIC SECTION | AIR TRANSPORTATION SYSTEMS: INFRASTRUCTURE, OPERATIONS, AND ENVIRONMENTAL INTEGRATION

Setting Airport Boarding Strategies Based on Passengers' Operational Data through Machine Learning Techniques

Marco Aurelio Gehlen[1,*], Giovanna Miceli Ronzani Ronzani[1]

1. Departamento de Ciência e Tecnologia Aerospacial – Instituto Tecnológico de Aeronáutica – Divisão de Engenharia Civil – São José dos Campos/SP – Brazil.

*Corresponding author: marco.gehlen@gmail.com

ABSTRACT

Boarding is crucial to turnaround time and can cause significant delays, with the Federal Aviation Administration (FAA) estimating $30 billion in pre-pandemic losses. Previous studies on airport boarding focus on pre-defined strategies that often overlook passenger behavior. This has led to a lack of consensus on the best way to reduce boarding time and improve the level of service (LoS) in different contexts. To address this, this study proposes modeling boarding time using passenger behavior variables across different strategies by combining different techniques. A simulation of three boarding strategies is conducted using screening design of experiments (DOE) with 24 runs each, resulting in 72 samples for A320 boarding time estimation. Machine learning methods, including linear regression, k-nearest-neighbor (KNN), multi-layer-perceptron (MLP), random forest, and XGBoost, are then applied to the simulation data for analysis. As a result, a model that can be used to predict boarding time for a given context of passenger behavior is discussed. Although random forest and XGBoost showed the highest R-squared values, they presented overfitting. Linear regression, with an R-squared close to 0.5, reveals that boarding strategy and bag distribution are the most influential variables, consistent with the literature. Steffen's strategy provides the lowest boarding time, averaging 12 ± 0.02 minutes to board 180 passengers.

Keywords: Machine learning; Airport; Airline; Strategy; Simulation.

INTRODUCTION

Among all airport processes, boarding is the one that most depends on the way passengers behave and their willingness to follow the established procedures. The *FAA* (2022) estimates the pre-pandemic costs of airline delays at over $30 billion and, as Neumann (2019) illustrates, "most of primary delays occurred at the gates." Also, Efthymiou *et al.* (2018) highlight that over 80% of passengers are only informed of delays at the boarding gate or just before boarding, often leading to negative emotions and contributing significantly to passengers' dissatisfaction, which can affect customer loyalty.

Regarding the boarding process itself, several variables contribute to it, including priority fares, whether the passenger is a frequent flyer and thus used to the airport procedures, the number of hand-on luggage, whether the passenger is traveling in groups or alone, the passenger's agility, whether the passenger is delayed or not, and simply the way the passenger behaves while in the queue. The literature around boarding procedures is mainly focused on determining which are the variables that most impact boarding time and which is the best boarding strategy (the one that minimizes boarding time). However, there is still no consensus on this, as it depends on each context (airplane and passenger characteristics).

Received: Dec. 13, 2024 | **Accepted:** Apr. 01, 2025
Peer Review History: Single Blind Peer Review.
Section editor: Leonardo Marini-Pereira

Based on that, this study aims to contribute to answering the question: which strategy should be set, and how can boarding time minimized considering the passenger behavior? Answering this question can help improve the perceived level of service (LoS), which is related to loyalty and helps avoid costs while reducing delays. To do that, two objectives are set: 1) establishing an environment based on a screening design of experiments (DOE) in which different scenarios can be simulated to generate the database around boarding strategies, boarding time and passenger behavior aiming to understand which strategy is the one that minimizes boarding time for a given passenger' behavior context; and 2) based on the dataset, using machine learning techniques to develop a model that can be used to predict boarding time for a given context and understand the influence of behavior variables. The techniques chosen are among the most used: linear regression, k-nearest-neighbor (KNN), multi-layer-perceptron (MLP), random forest, and XGBoost. Achieving these objectives can help airline managers to adapt boarding procedures dynamically (based on the passenger average profile for specific flights), thereby reducing boarding time.

The structure of the study is presented in four sections after this introduction: the literature around airport boarding procedures focusing on boarding time modeling and boarding strategies is presented considering machine learning techniques. The methodology shows the steps that are followed to achieve the objective of this study. The results of the simulation (database generation) and the model developed are presented and can be replicated in different contexts. Following the results, the conclusion discussion and next steps are presented.

LITERATURE REVIEW

Simulation and models

Milne and Kelly (2014) developed a discrete event simulation based on Steffen's boarding strategy to optimize boarding time by assigning seats to passengers based on the amount of luggage they were carrying. Steffen's method assigns passengers to specific numerical positions (seat locations) in an optimized way, reducing boarding time. The authors added luggage storage to this method, increasing boarding efficiency, however, their method lacks consideration of personal factors such as seat preference, special needs, and cabin segmentation (e.g., first class). This highlights a gap in addressing individual passenger characteristics, which is critical for a predictive model of boarding time based on passenger behavior.

Notomista *et al.* (2016) advanced Steffen's method by introducing a dynamic seat allocation system based on optical sensor data. These sensors measured each passenger's agility and estimated the size of their hand-luggage using computer vision techniques to obtain real-time data, which served as inputs for their algorithm. This study provides a significant contribution by incorporating real-time passenger behavior data, which aligns well with the goal of creating a predictive model based on such data. Bidanda *et al.* (2017) reviewed numerous studies on operations research and physical models aimed at reducing boarding time, including Markov decision processes, queue theory, and simulation techniques such as discrete-event simulation, Monte Carlo, and agent-based modeling. They concluded that fixed boarding strategies lack flexibility; for example, Steffen's method is optimal for an A380 but not for a B777. They suggest developing a dynamic tool to tailor boarding strategies to specific contexts, considering variables such as the type of airplane and passenger characteristics. This underscores the necessity for adaptable models that can predict boarding time in various contexts, reinforcing the need for a flexible, passenger-behavior-based approach.

Kisiel (2020) used Monte Carlo simulation to analyze the impact of priority fares on boarding time across different boarding strategies. To validate the discrete event-based simulation model, 62 flights with 180 passengers each were observed through two cameras inside an A320 operated by a low-cost carrier. The study found that even the best boarding strategies can be disrupted by priority fares, as priority passengers board randomly, affecting the boarding sequence. Consequently, the number of priority passengers should be considered in any boarding strategy, and developing hybrid strategies that account for delayed passengers and sequence disruptions could enhance efficiency. This emphasizes the importance of accounting for variability in passenger behavior in predictive models.

In the context of COVID-19, Milne *et al.* (2020) conducted over 10,000 simulations using stochastic simulation and agent-based modeling (NetLogo software to model passenger mobility) to evaluate the risk transmission impact of different boarding strategies for a single-door Airbus A320. They used the reverse pyramid boarding strategy, which boards passengers in groups, as it minimizes risk transmission. Variations of the reverse pyramid strategy were evaluated based on the number of boarding groups, luggage volume, and aisle social distancing. This study illustrates the adaptability of boarding strategies to different scenarios, which is crucial for developing a robust predictive model.

Schultz and Soolaki (2021) employed a genetic algorithm to allocate seats to passengers during COVID-19. They modeled passenger behavior using a stochastic cellular automata model, incorporating a transmission risk model. Their findings suggested that group boarding reduces both boarding time and transmission risk. This aligns with the objective of using passenger behavior data to optimize boarding strategies. More recently, Kobbaey *et al.* (2023) used autonomous agent-based simulation to compare the efficiency of different boarding strategies, considering luggage, walking speed, and passenger behavior (including non-compliance). They discussed which strategies minimize boarding delays and enhance the passenger experience while improving sustainability at airports. Their work supports the need for comprehensive models that integrate various aspects of passenger behavior to predict boarding time accurately. Rajarajeswari *et al.* (2023) highlight how disruptions can affect subsequent flights and impact operational efficiency, leading to increased costs and changes to itineraries. Machine learning methods, including naïve Bayes, neural networks, and decision trees, are used to build prediction models for delays. By identifying patterns in input data, such as weather and operational parameters, algorithms like decision trees and neural networks can forecast delays and improve operational efficiency. The use of machine learning in this context helps airlines optimize scheduling, reduce costs, and enhance the passenger experience by anticipating disruptions, which directly connects to the goal of this study.

Pedestrian behavior

According to Schultz and Fricke (2011), terminal handling progress depends on individual passenger behavior. They developed an agent-based model to analyze passenger movements and decision-making for route choice, validated with real and virtual terminal data. Schultz *et al.* (2013) further modeled passenger motions to understand the impact of behavior on boarding strategies, using three different airplanes (A320, B777, and A380) for validation with AirBerlin data. This study highlighted the influence of different aircraft types on boarding efficiency.

Schultz (2017) emphasized that aircraft boarding is driven by passengers, not airport or airline employees. The author provided a dataset including passenger motion, luggage storage time, and interactions to calibrate boarding simulation models. This underscores the need to model individual behaviors accurately in predictive models. Schultz (2018) developed a sensor environment model using seats as sensors to evaluate boarding progress in real time, allowing airlines to adjust processes dynamically. Schultz and Reitmann (2019) improved this model with machine learning (long short-term memory) to predict boarding time, incorporating passenger interactions.

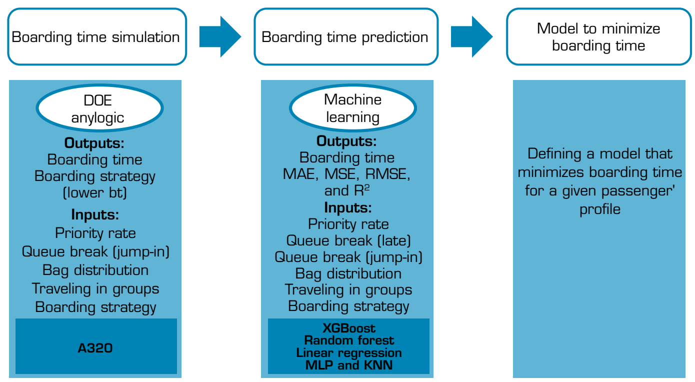
Gadaleta and Rossi (2018) introduced IDNet, a gait recognition model using smartphone motion signals (accelerometer and gyroscope) and convolutional neural networks for user authentication. This demonstrates the potential of using personal motion data to enhance boarding models. Similarly, Gjoreski *et al.* (2020) employed machine learning and hidden Markov models to classify human activities using smartphone sensor data, highlighting the feasibility of using such data to model passenger behavior. Kececi *et al.* (2020) also used machine learning for gait recognition with a large dataset, reinforcing the application of motion data in predictive models.

Rodríguez-Sanz *et al.* (2021) studied queue patterns in airports, using simulation and real data from a European airport to predict queue behaviors with random forest algorithms. This study illustrates the importance of understanding and predicting passenger flow in enhancing boarding efficiency. Sadou and Njoya (2023) explore the use of artificial intelligence in the air transport industry, highlighting the use of machine learning in improving efficiency, safety, and customer experience to explore the application of technological tools in different aspects of air transport, which, in the case of this study, is setting boarding strategies in airports.

Fabrin *et al.* (2024) focused on individual passenger metrics in an agent-based model to reduce boarding time while considering passenger experience, aligning closely with the objective of integrating passenger behavior into predictive models.

METHODOLOGY

The methodology consists of three parts: the first is generating the database with simulations to estimate the boarding time for a given passenger behavior and boarding strategy. The second part is the application of machine learning techniques to the dataset to predict which boarding time is the lowest for a given passenger behavior. The third part is defining a model that can suggest the definition of a boarding strategy for a given passenger behavior to reduce boarding time. Figure 1 shows a summary.



Source: Elaborated by the authors.

Figure 1. Methodology steps.

Database generation (simulation)

A screening DOE is set for a pre-defined agent-based model from AnyLogic software that is used. The simulation data were generated using an agent-based model for an A320, based on AnyLogic's predefined passenger flow library. The model incorporates a trajectory model and a social force model, which considers interactions between people. The screening DOE technique was employed to develop and run the simulations within this environment defining scenarios to look for the boarding strategy that minimizes boarding time (dependent variable). The independent variables include factors related to passenger behavior, such as boarding strategy, priority fare rate, queue breaks, bag distribution, and the airline company. Based on the DOE, the variables assume discrete values (-1 or 1). Specifically, when the variable priority rate assumes the value "1," it means that 25% of passengers have some kind of priority. The same for the variable "queue break," where assuming the value "1" means that there is a jump-in rate that represents 5% of passengers interfering in queues. Those values aim to represent what is expected in a boarding process behavior, and so they are assumed empirically considering a priority rate of 25% and 5% of jump-in interferences. Those respective rates could be different, leading to other results and the impact of their change is not evaluated in the scope of this study.

- Bag distribution – 1 for random distribution, -1 for even distribution.
- Priority rate – 1 meaning at least 25% of passengers have any type of priority and -1 meaning no one has.
- Queue break (late passengers' rate) – Variable related to interferences; 1 means a 5% rate of passengers are late, and -1 means no interference (natural flow of the queue with no source of interruption).
- Queue break (jump-in rate) – Variable related to interferences; 1 means a 5% rate of passengers interfere by jumping-in the queues, and -1 means no interference (natural flow of the queue with no source of interruption).
- Passenger company – 1 meaning the passengers are traveling in groups, and -1 meaning they are alone.
- Boarding strategy – Three of the most discussed boarding strategies in the literature are simulated: random, block, and Steffen.
  Random: as the name says, passengers are boarded randomly with no pre-defined sequence.
  Blocks (back to front): the airplane rows are divided into sections, and each section is boarded in order (back to front).
  Steffen's: as Kisiel (2020) describes, each passenger is boarded separately according to a sequence of seats that is set based on "back-to-front"; starting from the back, a passenger seats every second row and the sequence is repeated for the columns until everyone is seated.

To determine which and how many simulations to run, a DOE is set. Vanaja and Rany (2008) describe the Plackett-Burman technique that is used to screen variables that can influence a determined output (in this case boarding time). According to the authors, "Statistical experimental design, also known as DOE, is the methodology of how to conduct and plan experiments in order to extract the maximum amount of information with the lowest number of analyses." It helps to measure interactions and it allows extrapolation of data. The experiments to simulate boarding time are based on five variables of passengers' behavior: priority rate, queue break, bag distribution, company, and three different boarding strategies (random, block, and Steffen). The boarding strategies represent different approaches to boarding and the time to board everyone from the gate check until the last passenger is seated. The output of the simulation is the boarding time for each strategy. It results in eight runs for each of the three boarding strategies that are simulated, resulting in 24 samples for each boarding strategy as shown in Table 1.

Table 1. Design of experiments results for the five variables (first eight samples).

Run order	Bag distribution	Priority rate	Passenger company (if in groups)	Queue break late passenger rate	Queue break jump-in rate
1	1	-1	1	-1	-1
2	1	1	-1	-1	1
3	1	1	-1	1	-1
4	-1	1	1	1	-1
5	1	-1	1	1	1
6	-1	1	1	-1	1
7	-1	-1	-1	1	1
8	-1	-1	-1	-1	-1

Source: Elaborated by the authors.

The Plackett-Burman design with 24 combinations was chosen due to computational resource constraints in the simulation tool AnyLogic. While a full factorial design with thirty-two combinations would provide more detailed analysis, including interactions between variables, Plackett-Burman efficiently identifies the main effects with fewer experiments. As a limitation of this approach, second-order interactions are not assessed, but this trade-off is justified by the reduced number of experiments. Regarding the passenger operational variables, only five were considered. In future work, behavioral and demographic data can be included, such as age, mobility issues, travel purposes, elderly passengers, and babies, which could improve this analysis.

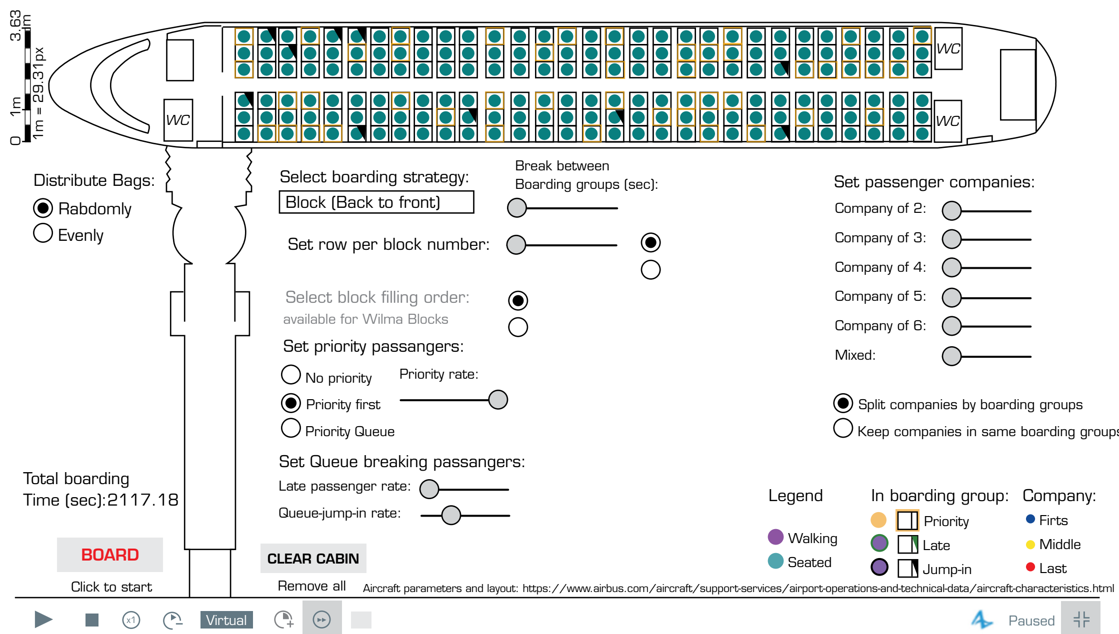
The simulation based on an AnyLogic application was developed by Nadtochiy (2020), and its interface used for the simulation is presented, where it is possible to visualize inputs and outputs in a user-friendly interface (Fig. 2).

As a result, the simulation presents the boarding time in minutes for different passengers' behavior. It is composed of autonomous agents of human behavior that interact with each other and simulate a single-aisle aircraft boarding using various boarding policies. The aircraft type is an Airbus A320 (one of the most used and referred to in the literature review) with 180 passengers split into 20 rows. There is no cabin segmentation in this simulation, meaning that all passengers are assigned to the same kind of seats (there is only economy class) and the airport gate infrastructure is kept unchanged for the simulations.

Boarding time prediction

By using the dataset generated from the AnyLogic simulation environment, different machine learning techniques are applied in order to predict boarding time for a specific boarding strategy, considering the passengers' behavior for a specific flight. Given those results, a training database is available, resulting in a boarding time based on passengers' behavior. The techniques are chosen as they are among the most used ones (Gjoreski *et al.* 2020; Kececi *et al.* 2020; Rodríguez-Sanz *et al.* 2021): linear regression, KNN, MLP, random forest, and XGBoost. Inputs are passengers' behavior variables and the output is boarding time. The independent variables and the values they assume are the same as in the simulation.

Linear regression and random forest give insights into variables that have the most impact on the boarding time, which can be useful for understanding which passenger attributes are impacting boarding time. Also, they allow a level of interpretation



Source: Retrieved from Nadtochiy (2020).

Figure 2. AnyLogic simulation tool interface (https://www.anylogic.com).

through feature importance scores and decision path analysis. K-nearest-neighbors (KNN) is an instance-based learning algorithm that predicts the output (boarding time) based on the nearest neighbors in the feature space. The predictions are based on similar historical instances, which can be intuitively explained (e.g., "this passenger has similar features to another who boarded in X minutes"). Multi-layer-perceptron (MLP) is a type of artificial neural network composed of multiple layers of neurons, where each layer performs a nonlinear transformation of the input data, making MLP a complement to the other tools for modeling complex relationships (non-linear relationships between features and the target variable, which might be important in predicting boarding time based on passenger characteristics).

XGBoost is a boosting algorithm that builds a series of weak learners (typically decision trees) where each tree corrects the errors of the previous one. It is known for its performance on structured data, which aligns well with passenger operational data.

The models were trained with predefined hyperparameters and then used to predict values, which were compared and used to evaluate the model's performance. The dataset was split into training and testing sets using train_test_split, with 30% of the data reserved for testing and 70% for training.

In the linear regression, KNN, random forest, MLP, and XGBoost models, the hyperparameters were selected with default values for parameters such as learning_rate, max_depth, and n_estimators in XGBoost, or n_jobs=None in random forest and KNN. The hyperparameters for the linear regression model were set with copy_X=True to ensure the input data is copied, fit_intercept=True to include an intercept term, n_jobs=None to use a single central processing unit (CPU) core, and positive=False to allow both positive and negative coefficients.

The KNN model uses the kd_tree algorithm for efficient distance calculation, with a leaf_size of thirty, balancing memory and accuracy. It employs Minkowski distance (Euclidean with p=2), considers five nearest neighbors, and uses uniform weights for neighbors.

The hyperparameters for the random forest model were set with bootstrap=True to use bootstrap sampling, criterion='squared_error' for splitting nodes based on squared error, and n_estimators=100 for 100 trees. Other parameters like max_depth=None, min_samples_split=2, and min_samples_leaf=1 allow the trees to grow to their maximum depth and split nodes as needed, while n_jobs=None ensures a single CPU core is used during training.

For the MLP model hyperparameters were set with activation='relu' for the activation function, hidden_layer_sizes=(10, 100) for two hidden layers with 10 and 100 neurons, and solver='adam' for the optimization algorithm. Other settings like learning_rate_init=0.001, batch_size='auto', and max_iter=200 control learning behavior, with early_stopping=False and n_iter_no_change=10 preventing premature stopping, while momentum=0.9 and nesterovs_momentum=True are used to accelerate convergence.

The XGBoost model was configured with objective='reg:linear' for regression tasks, using squared error as the loss function, max_depth=16, n_estimators=500 and random_state=2 for reproducibility. Learning_rate and booster used its default tree boosting method.

Defining a model (machine learning)

Developing a model that can predict boarding time based on the previous simulations and analyzing the impact of the passengers' behavior variables on it by analyzing the performance of different machine learning techniques.

These techniques are compared using four different metrics: mean absolute error (MAE), mean square error (MSE), and R-squared. The MAE measures the difference between the predicted boarding time and the real boarding time (residual). It does this for all boarding time values and divides it by the number of simulations. The MSE squares the difference of the MAE before summing it allowing them to identify outliers. R-squared varies from 0 to 1, and it measures how accurate a regression line is to predict each output (in this case, boarding time). Overfitting analysis was performed by comparing R-squared and MSE for the training and testing databases. When models are more adjusted to the training database, it shows overfitting.

The lower the boarding time, the better the LoS perception by passengers. Defining which boarding strategy to use based on boarding time, considering passengers' behavior, can not only reduce delays but it can also increase the passengers' perceived LoS.

Airports and airlines willing to increase LoS can use this model to choose which boarding strategy to set for a specific flight depending on the average passenger profile characteristics for that flight. Although airports and airlines could use paid tools to simulate boarding process strategies, these solutions are often expensive and require training and licensing fees.

The main objective of developing a model was to create an accessible and reusable tool to support decision-making. The dataset was generated using AnyLogic, a paid tool, but it was only used to train the model, which itself, being free, can be adapted to different scenarios, allowing for adjustments to parameters and the addition or removal of variables as needed. This makes it a plug-and-play solution, replicable in different contexts without the need to rely on paid tools.

RESULTS

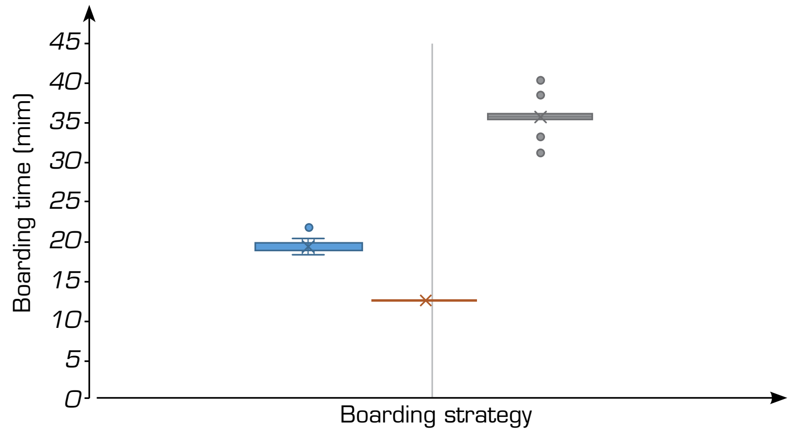
The strategy that shows the lowest boarding time is Steffen's, which presents an average of 12.25 minutes to board 180 passengers with a low standard deviation of 0.02 for the 24 runs. Following Steffen's, the random strategy presents an average of 19.12 minutes to board, while back to front resulted in an average of 35.38 minutes with a standard deviation of 2 minutes as Table 2 and Fig. 3 show. These are results from the AnyLogic simulation.

To assess if the averages differ across the three boarding scenarios, an analysis of variance (ANOVA) test was conducted to assess the significance of the samples. Following this, pairwise comparisons between the three boarding strategies were performed. The resulting p-value for each comparison was below 0.05, indicating that the null hypothesis (which states that there is no difference in means) was rejected. This suggests that there are statistically significant differences between the average values of the scenarios.

Table 2. Boarding strategies scenarios results.

Boarding time (min)	Average (min)	Median (min)	Standard deviation (min)	p-value for average comparison
Steffen	12.24	12.24	0.02	< 0.05
Random	19.12	18.82	0.73	< 0.05
Block (back to front)	35.38	35.62	2.55	< 0.05

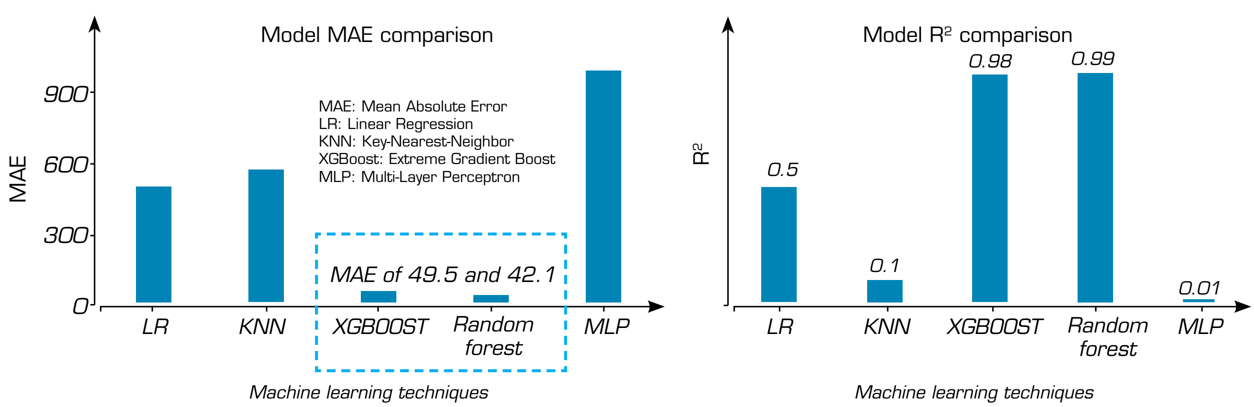
Source: Elaborated by the authors.



Source: Elaborated by the authors.

Figure 3. Boarding time vs. boarding strategy: random in blue, Steffen in orange, and grey back to front.

These findings validate the effectiveness of Steffen's method, aligning with previous studies that highlight its efficiency. The random strategy (19.12 minutes) and back to front (35.38 minutes) are consistent with the literature that shows fixed strategies often fall short of dynamic solutions. Regarding the prediction results of this study, Fig. 4 shows how accurate each of the machine learning regressors is when comparing MAE and R-squared.



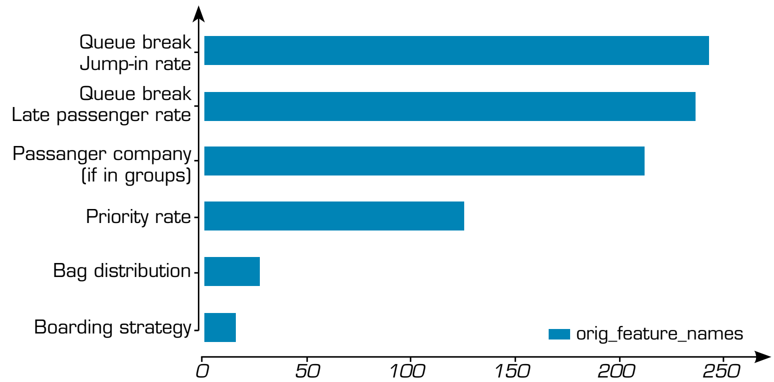
Source: Elaborated by the authors.

Figure 4. Comparison of different regressors' performance (MAE and R-squared).

XGBoost and random forest showed R-squared values of 0.9877 and 0.9887, meaning that the variables used in the simulation would explain 98% of the boarding time variance or, in other words, that the variables would be great predictors of boarding time. Also, they showed respectively the lowest values of residuals with a MAE of 49.5 and 42.1 seconds, respectively. The variables related to interferences (queue break) and whether passengers are traveling in groups are the ones with the highest importance for boarding time prediction when using XGBoost, as Fig. 5 shows, while the boarding strategy and bag distribution are the lowest.

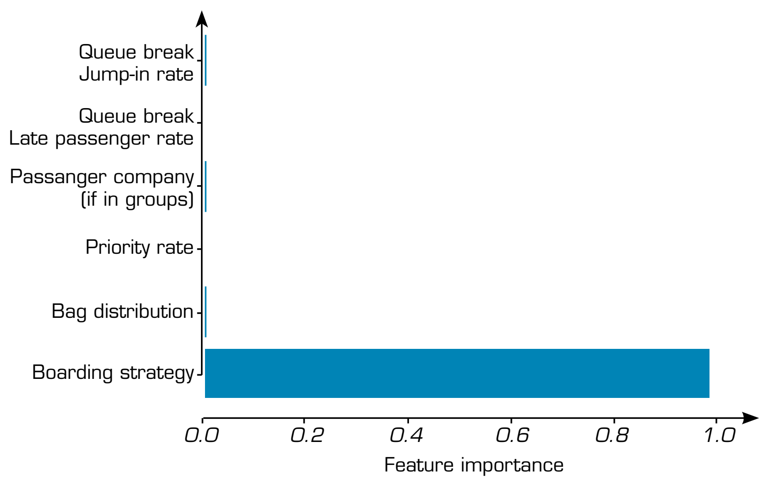
Considering random forest, as Fig. 6 shows, the vast majority of boarding time is explained by the boarding strategy, or in other words, defining the boarding strategy itself as the main predictor for boarding time.

Specifically, XGBoost identified queue breaks and group travel as significant predictors, aligning with studies that emphasize passenger behavior's impact on boarding efficiency. Conversely, random forest highlighted the boarding strategy itself as the primary determinant, underscoring the literature's call for adaptable and context-specific strategies.



Source: Elaborated by the authors.

Figure 5. Coefficients of boarding time XGBoost model.



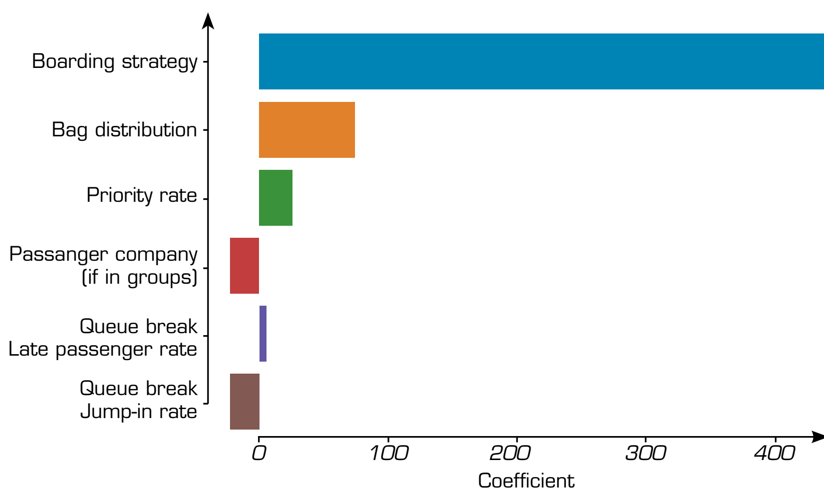
Source: Elaborated by the authors.

Figure 6. Feature importance of boarding time variables using random forest model.

The variables that impact boarding time for different contexts have been discussed in many studies as the literature review presented. Although some studies achieve important levels of prediction, it is known that modeling the passenger behavior and airport operations for different scenarios and contexts is not an easy task. The disagreement and the lack of consensus around which are the variables that most influence the process is a common theme in the literature on the boarding procedure subject which is coherent with the divergence between the importance of the variables achieved in this study. Nevertheless, considering that both random forest and XGBoost achieved 98% R-squared and knowing the limitations of this study (only 72 samples, three boarding strategies and five variables), it would be superficial to affirm that both models are close to 100% as predictors. To check if there was overfitting when using random forest, the R-squared for the training database and the R-squared for the test database were compared. It resulted in 0.99 for both cases, but when the MSE was compared, the training database (MSE = 1,854) was much lower than the test MSE (2,303), showing that the model was more adjusted for the training database than for the test database, which characterizes overfitting and helps to explain why a single variable (boarding strategy) was sufficient to describe the variability of boarding time. Regarding the XGBoost analysis, R-squared is similar (0.99) for both the train and test database, and the MSE is 1,677 for training and 1,829 for test showing that the model is more adjusted to training data than to test data (which could indicate overfitting but lower than random forest). Another factor is the challenge of interpreting machine learning models such as XGBoost and random forest. Although it can be done, it is not an easy task to provide a simple visual way to interpret how each variable relates to each other and how it influences the output (boarding time in this case).

There are several reasons that could explain the high values of the residuals and the low R-squared for the other algorithms (MLP, KNN, linear regression). The first is the small sample, as it was based on a screening DOE, aiming to illustrate and take the first steps to investigate the problem of this research (24 runs for each of the three strategies that were simulated). For machine learning techniques, their capability to be good predictors is also associated with higher samples sizes, which can be achieved as this study evolves. The second reason is the small number of variables that were used in the models (only five with two states each). Based on those factors, the linear regression model is further analyzed. It showed a MAE of 486.4 seconds or 8 minutes and an R-squared close to 0.5. The regression model is shown in Eq. 1, and Fig. 7 shows the model coefficients for the boarding time regression model.

Boarding time (s) =457.14*boarding strategy+73.9*bag distribution+26.5
*priority rate+6.3 *queue break (late pax)-21.5
*queue break (jump-in rate)-22.2 pax company



Source: Elaborated by the authors.

Figure 7. Coefficients of boarding time regression model.

Although the R-squared is only close to 50%, and thus the coefficients only explain 50% of the boarding time variance, it can be seen that the boarding strategy has the highest influence on the variance of boarding time: Steffen's was the strategy with the lowest values of boarding time, followed by random. In consensus with the literature that often points to "bags" as a key factor influencing boarding time, bag distribution is the second highest coefficient, meaning that there is an influence on how the bags are distributed in the overhead bins (randomly or evenly), which affects boarding time. The other variables, such as "priority rate," and "queue-breaks due to late passengers," have minimal impact on the boarding time variance. One hypothesis is that, as the priority rate in the simulations were only varied from 0 to 25% of passengers and that interferences was only varied from 0 to 5%, higher variations could provide another effect. Higher levels of interference could lead those variables to be more impactful in the explanation of boarding time variance. Queue jump-in rate and whether passengers are traveling in groups or not impact negatively the variance of boarding time. This effect is probably related to low rates of interference caused by jump-ins and passengers traveling alone (who tend to be faster), which tend to reduce boarding time.

Airlines can use these findings to refine boarding strategies, particularly by adopting dynamic methods. Machine learning models offer a data-driven approach to predict boarding times based on numerous factors, enabling airlines to make informed decisions about boarding procedures. By integrating real-time data and advanced predictive models, airlines can enhance operational efficiency and perceived LoS.

CONCLUSION

This study achieved the objective of taking steps to help answer which boarding strategy to set to minimize boarding time based on passenger behavior contributing by integrating advanced machine learning techniques to model and predict boarding times.

To do that, a screening DOE was developed and applied to a model in AnyLogic to simulate different scenarios based on three boarding strategies, and machine learning techniques were applied to predict boarding time for a given boarding strategy. Results from the AnyLogic simulation show that Seteffen's strategy is the better choice for the context, taking 12 minutes on average and a standard deviation of 0.02 minutes to board 180 passengers in an A320. Among the machine learning techniques, although random forest and XGBoost exhibited the highest R-squared values, they presented overfitting (the random forest model for instance was adjusted to the training database with an MSE of 1,854 compared to the test database of 2,303). Then, a linear regression model was proposed with an R-squared close to 0.5, revealing that the boarding strategy and the way bags are distributed are the variables that most influence boarding time in consensus with the literature.

The study is limited to scenarios based on one airplane (A320) and in three different boarding strategies. As next steps, using real data to improve the model would be a significant contribution, and exploring the use of internet of things (IoT) sensors to monitor real-time boarding dynamics can provide insights to improve the process's efficiency. Finally, modeling the passengers' behavior while boarding has several challenges and it can be improved, as it was limited to five variables in this study (bag distribution, priority rate, interferences, queue breaks & queue jump-in rate, and whether traveling alone or in groups). For the DOE, a Plackett-Burman design with 24 combinations for each of the strategies was chosen due to computational resource constraints in the simulation tool, resulting in seventy-two samples. As a limitation of this approach, second-order interactions are not assessed, but this trade-off is justified by the reduced number of experiments. To improve this study, some other variables to be considered could be passengers' demographics, interferences caused by passengers themselves while looking for seats, gait/biometric recognition, or simply behavior issues when not following the required procedures established by airlines.

CONFLICT OF INTEREST

Nothing to declare.

AUTHORS' CONTRIBUTION

Conceptualization: Gehlen MA and Ronzani GMR; **Methodology:** Gehlen MA and Ronzani GMR; **Software:** Gehlen MA; **Validation:** Gehlen MA and Ronzani GMR; **Formal analysis:** Gehlen MA and Ronzani GMR; **Investigation:** Gehlen MA; **Resources:** Ronzani GMR; **Data Curation:** Gehlen MA; **Writing - Original Draft:** Gehlen MA; **Writing - Review & Editing:** Gehlen MA and Ronzani GMR; **Visualization:** Ronzani GMR; **Supervision:** Ronzani GMR; **Project administration:** Ronzani GMR; **Funding acquisition:** Ronzani GMR; **Final approval:** Gehlen MA.

DATA AVAILABILITY STATEMENT

All data sets were generated or analyzed in the current study.

FUNDING

Conselho Nacional de Desenvolvimento Científico e Tecnológico ROR
Grant No: 501100003593

ACKNOWLEDGMENTS

Not applicable.

REFERENCES

[FAA] *Federal Aviation Administration* (2022) Cost of delays. https://www.faa.gov/data_research/aviation_data_statistics/ [accessed Jan 2024]

Bidanda R, Geng Z, Winakor J, Vidic N (2017) A review of optimization models for boarding a commercial airplane. Paper presented 2017 24th International Conference on Production Research. International Foundation for Production Research; Poznan, Poland.

Efthymiou M, Njoya ET, Lo PL, Papatheodorou A, Randall D (2018) The impact of delays on customers' satisfaction: an empirical analysis of the British airways on-time performance at Heathrow Airport. J Aerosp Technol Manag 11:977. https://doi.org/10.5028/jatm.v11.977

Fabrin BHP, Ferrari DB, Arraut EM, Neumann S (2024) Towards balancing efficiency and customer satisfaction in airplane boarding: an agent-based approach. Oper Res Perspect 12:100301. https://doi.org/10.1016/j.orp.2024.100301

Gadaleta M, Rossi M (2018) IDNet: smartphone-based gait recognition with convolutional neural networks. Pattern Recognit 74:25-37. https://doi.org/10.1016/j.patcog.2017.09.005

Gjoreski M, Janko V, Slapničar G, Mlakar M, Reščič N, Bizjak J, Drobnič V, Marinko N, Mlakar N, Luštrek M, Gams M (2020) Classical and deep learning methods for recognizing human activities and modes of transportation with smartphone sensors. Inf Fusion 62:47-62. https://doi.org/10.1016/j.inffus.2020.04.004

Kececi A, Yildirak A, Ozyazici K, Ayluctarhan G, Agbulut O, Zincir I (2020) Implementation of machine learning algorithms for gait recognition. Eng Sci Technol Int J 23(4):931-937. https://doi.org/10.1016/j.jestch.2020.01.005

Kisiel T (2020) Resilience of passenger boarding strategies to priority fares offered by airlines. J Air Transp Manag 87:101853. https://doi.org/10.1016/j.jairtraman.2020.101853

Kobbaey T, Bilquise G, Naqi AA (2023) A comparative evaluation of airplane boarding strategies with a novel method for sustainable air travel. *Paper presented 2023 9th International Conference on Information Technology Trends*. IEEE; Dubai, United Arab Emirates. https://ieeexplore.ieee.org/document/10184234

Milne RJ, Cotfas L-A, Delcea C, Crăciun L, Molănescu A-G (2020) Adapting the reverse pyramid airplane boarding method for social distancing in the times of COVID-19. PLoS ONE 15(11):e0242131. https://doi.org/10.1371/journal.pone.0242131

Milne RJ, Kelly AR (2014) A new method for boarding passengers onto an airplane. J Air Transp Manag 34:93-100. https://doi.org/10.1016/j.jairtraman.2013.08.006

Nadtochiy D (2020) Aircraft boarding simulation (prototype). AnyLogic Simulation Software. [accessed Jan 2024]. https://cloud.anylogic.com/model/762a238c-2294-498a-88ef-c37878746ddd?mode=SETTINGS

Neumann S (2019) Is the boarding process on the critical path of the airplane turn-around? Eur J Oper Res 277(1):128-137. https://doi.org/10.1016/j.ejor.2019.02.001

Notomista G, Selvaggio M, Sbrizzi F, Di Maio G, Grazioso S, Botsch M (2016) A fast aircraft boarding strategy using online seat assignment based on passenger classification. J Air Transp Manag 53:140-149. https://doi.org/10.1016/j.jairtraman.2016.02.012

Rajarajeswari P, Lalitha P, Kumari S, Lakshmi AS, Muguthan SR, Kumar PR, Reddy AP (2023) Simulation-based predictive analysis of Indian Airport transportation system using computational intelligence techniques. J Aerosp Technol Manag 15. https://doi.org/10.1590/jatm.v15.1300

Rodríguez-Sanz A, de Marcos FA, Pérez-Castán JA, Comendador FG, Valdés RA, Loreiro AP (2021) Queue behavioral patterns for passengers at airport terminals: A machine learning approach. J Air Transp Manag 90:101940. https://doi.org/10.1016/j.jairtraman.2020.101940

Sadou AM, Njoya ET (2023) Applications of artificial intelligence in the air transport industry: a bibliometric and systematic literature review. J Aerosp Technol Manag 15:1312. https://doi.org/10.1590/jatm.v15.1312

Schultz M (2017) Aircraft boarding-data, validation, analysis. Paper presented 2017 12th USA/Europe Air Traffic Management Research and Development Seminar. ATRD; Seattle, USA.

Schultz M (2018) A metric for the real-time evaluation of the aircraft boarding progress. Trans Res C: Emerg Technol 86:467-487. https://doi.org/10.1016/j.trc.2017.11.002

Schultz M, Fricke H (2011) Managing passenger handling at airport terminals: individual-based approach for modeling the stochastic passenger behavior. Air traffic management research and development seminar. [Feb 2024]. https://innaxis-comm.s3.eu-central-1.amazonaws.com/ENGAGE/WIKI/PAPERS/ATMSEMINAR/2011/122_Schultz.pdf

Schultz M, Kunze T, Fricke H (2013) Boarding on the critical path of the turnaround. Paper presented 2013 10th USA/Europe Air Traffic Management Research and Development Seminar. ATRD; Chicago, USA.

Schultz M, Reitmann S (2019) Machine learning approach to predict aircraft boarding. Trans Res C: Emerg Technol 98: 391-408. https://doi.org/10.1016/j.trc.2018.09.007

Schultz M, Soolaki M (2021) Analytical approach to solve the problem of aircraft passenger boarding during the coronavirus pandemic. Trans Res C: Emerg Technol 124:102931. https://doi.org/10.1016/j.trc.2020.102931

Vanaja K, Rani RH (2008) Design of experiments: concept and applications of plackett Burman design. Clin Res Regul Aff 24:1-23. https://doi.org/10.1080/10601330701220520