# Agent-based simulation for aircraft stand operations to predict ground time using machine learning

Mingchuan Luo
*Institute of Logistics and Aviation*
*Dresden University of Technology*
Dresden, Germany
mingchuan.luo@tu-dresden.de

Michael Schultz
*Institute of Logistics and Aviation*
*Dresden University of Technology*
Dresden, Germany
michael.schultz@tu-dresden.de

Hartmut Fricke
*Institute of Logistics and Aviation*
*Dresden University of Technology*
Dresden, Germany
hartmut.fricke@tu-dresden.de

Bruno Desart
*EUROCONTROL*
Brussels, Belgium
bruno.desart@eurocontrol.int

Floris Herrema
*EUROCONTROL*
Brétigny-sur-Orge, France
floris.herrema@eurocontrol.int

Rocío Barragán Montes
*EUROCONTROL*
Brussels, Belgium
rocio.barragan-montes@eurocontrol.int

***Abstract*—Punctual and reliable aircraft ground handling operations at the airports significantly contribute to efficient traffic flows in the air traffic network. Any improved prediction of aircraft ground times can help to reduce local delays and delay propagation in the network by taking into account the forecast of future operational states for adjusted planning and delay mitigation strategies. In our work, we target to predict aircraft ground times at their stands by machine learning algorithms, where the complete turnaround sub-processes and domain knowledge are input for the models. We develop two types of models, the first type is regression-oriented that intends to forecast the exact aircraft ground time. And the second one is classification-oriented, which attempts to confirm aircraft off-block time adherence. An agent-based approach is applied to generate some synthetic data, besides, we also obtain an actual aircraft ground handling dataset from a certain European airport to validate our models. Finally, the interpretable method for the machine learning models is used to analyse the feature importances, and the feature affections on the prediction results. The results show that our classification model is capable to predict accurate aircraft off-block time adherence.**

***Index Terms*—aircraft stand operations, airport collaborative decision making, agent-based model, interpretable machine learning, aircraft ground time prediction**

## I. Introduction

Efficient aircraft ground handling at airports is important to ensure performance operations in the overall air traffic network. Close cooperation between all involved stakeholders (e.g., airport operators, airlines, ground handlers, air traffic service providers) positively impacts the punctuality and predictability of the aircraft turnaround process. In the aviation industry, it is consensus that each aircraft earns revenue only when they are en-route [1]. The airlines thereby always expect a maximum air time and minimal time for stand operations at airports. A reliable turnaround process with a short duration also assists the airports to arrange all the ground resources smoothly.

Airport collaborative decision making (A-CDM) is a process that consists of sharing information between stakeholders of the complex airport system to provide a common situational awareness and to enable mutual strategies to solve operational challenges. A-CDM was developed in establishing operational milestones method for every joint arrival and departure aircraft activity to improve the efficiency of airports and the air traffic network [2]. With a focus on airports, A-CDM provides a solution, which is generating cost reductions, environmental benefits, capacity optimization, and efficiency improvements. A performance-based airport environment enables to get full A-CDM benefits (e.g. enhanced use of airport resources or reliable scheduling) since airport stakeholders can collectively work on (dynamically) the agreed performance targets during the day of operations [3]. By giving airport stakeholders access to data from different sources, airports are able to make more accurate predictions about their operational progress in the next planning horizon [4]. An integrated management is embodied in an airport operations centre (APOC), where all stakeholder operators co-ordinate tasks to monitor and maintain the agreed performance targets in their respective areas of responsibility [5].

The COVID-19 pandemic strongly influences the entire air transportation industry. As the data collected by EUROCONTROL [6], yearly flights of 2020 fell 55% in comparison to 2019. Meanwhile, the overall average delay per flight decreased by 5.6 mins to 7.4 mins, where the reactionary delay and primary delay per flight are 2.8 mins and 4.6 mins respectively. After the COVID-19 impacted aviation, the unprecedented operational conditions appeared with the en-route ATFM delays basically vanishing and COVID-19 related delay causes becoming conspicuous. However, the abnormality in air transportation is not durable and the aviation industry is recovering from its former prosperity gradually. According to the statistic of EUROCONTROL, it is estimated that 660 million euros are lost every year over their areas due to delays generated by the turnaround.

The aircraft turnaround, as part of the aircraft trajectory over the day of operations, has to be part of optimization strategies for minimizing flight delays and ensuring flight connection considering operational uncertainties. In this context, reliable turnarounds depended on buffer time can absorb inbound delays and could enhance slot adherence at airports or mitigate problems of push-back scheduling [7]–[9]. Previous research focuses on the critical path of the aircraft turnaround and exhibits that both land- and airside processes can be bottlenecks [10]. Whenever these processes are part of the critical turnaround path, the effects could also propagate an accumulating delay through the ATM network [11,12]. Investigations on turnaround reliability show significant improvement potentials in standardization, data quality and availability, process design, integrated planning, and optimization, even under COVID-restrictions [13]. The speed and extent, with which data is shared, have massively increased over the last years as well as the need for implementation of new methods to evaluate this data that will further guarantee the sustainability of the air traffic network. In the course of increasing digitalization in almost all areas, airports are trying to implement innovative approaches in their current operations [14].

The paper is organized as follows. Section 2 provides insight into the state-of-the-art and research objective, describing the general aircraft stand operations and reviewing the previous research. Section 3 explains the methodology used to realize this research, introduces the available turnaround data and the designed model in detail. In Section 4, the simulation results and analysis are presented. Finally, Section 5 gives the main conclusions and discusses potential future work on the subject, as well as the utility of the study.

## II. Background and motivation

The stand occupancy time of an aircraft is the duration that the aircraft stops at its parking position, i.e., occupies a remote apron position or a nearby gate at an airport. And the turnaround time refers to the sum of all the activities of the turnaround critical path and subjects to the ground handling (GH) activities, which are recognized as the turnaround sub-processes [15]. Turnaround is considered to finish when all doors of the aircraft are closed, all ground support equipment (GSE) is disconnected, the aircraft is ready to leave and the chocks are removed [16].

### *A. Previous research*

The critical path of the turnaround depicts the combination of all the sub-processes that determines the minimum turnaround time. At tight flight schedules, where this ground time is close to the minimal duration, mostly all processes are parts of the aircraft turnaround which typically consists of the parallel processes of cleaning, catering, fueling, and the sequential processes of the passenger deplaning, boarding as well as the baggage unloading and loading, as shown in Fig. 1. Based on the structure of the turnaround critical path, a model called GMAN (known as "Ground MANager") was proposed in [17] to obtain the total turnaround time. Monte Carlo simulation was chosen to generate samples from the distributions of the stochastic turnaround sub-processes. A review of the conducted studies on aircraft turnaround operations and simulations was summarized in [18], which highlighted some key challenges for the current aircraft operators, such as, the airport capacity restrictions, the schedule interference, and the increasing fee stress, and the corresponding solution attempts were also introduced to improve the potential process.

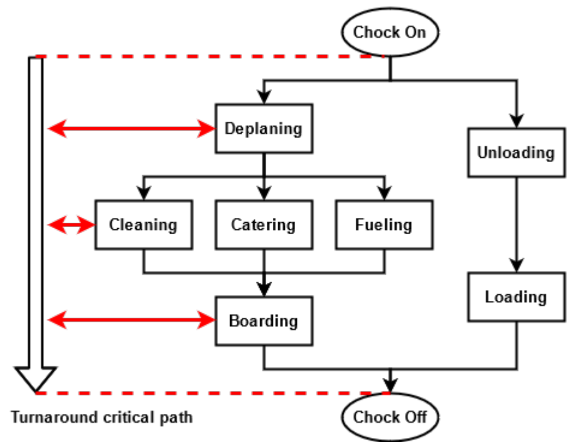


Fig. 1. Simple turnaround process.

The article [19] exhibited a multi-time scale ground handling management structure for the automation of the aircraft stand operations in order to improve the working efficiency of the GSE. From the perspective of total airport management (TAM), it provided two instructional heuristics of the GSE assignment within the typical turnaround time. Another research [20] built an agent-based model (ABM) to optimize the number of GSE allocated for airport service planning. In the real manufacturing environment, the multi-agent-based approach can also set feasible working schedules using negotiation/bidding mechanisms between agents [21,22].

At the current stage, data can be captured and analyzed from many aspects of airport operations (i.e. weather impact [23]) which are used to monitor the system performance and to identify areas of improvement. It provides the opportunity to get deeper insights into airport management by data analysis with advanced machine learning methods. Some regression models can estimate the turnaround duration, e.g., in [24] a flight turnaround time prediction model was established based on neural networks with these factors (including the aircraft stand, the aircraft type, the domestic or international flight, airline agent, the flight arrival time, the flight arrival and departure passenger number) that quantitatively and qualitatively impact the flight turnaround time, and the work [25] adopted support vector machine regression to forecast the aircraft turnaround time, which additionally considered the arrival delay and used the Gaussian Radial Basis Function to select the optimal parameters. Furthermore, a specific prediction model only focusing on the turnaround sub-process, i.e., aircraft boarding was implemented based on the Long Short-Term Memory model. Since no operational data of the specific passenger behavior is available, a reliable validated boarding simulation environment was applied to provide data about the aircraft boarding events [26]. With the interpretable Shapley values, the full turnaround time prediction was able to be analyzed on the importance of the available data, which would infer the contribution of each sub-process to the total duration [27].

### B. Research objectives

Actual A-CDM milestones do not provide any further information about turnaround processes during the time between in-block and off-block, only the target off-block time (TOBT) updating by the aircraft operator or the ground handler until the final TOBT confirmation before the target start-up approval time (TSAT) issue [2]. We target our research on the prediction of aircraft ground times at their stands using machine learning (ML) approaches, where the complete turnaround sub-processes and some domain knowledge are the model inputs. The final TOBT update indicates directly the off-block time and its timestamp is ahead of the milestone of passenger boarding so that if we can obtain the TOBT adherence information at this time, it provides the notable time horizon for the airports to improve the process management on the operational/tactical level [3].

Many computer vision techniques have been used at airports to monitor and collect information about the ongoing turnarounds, which provides a huge convenience to all the stakeholders on the airport real-time manual surveillance and the collaborative decision making in APOC. Alternatively, for the automatic turnaround sub-process identification, there are still massive false detections and blank records made by machines due to the nature of the computer vision techniques. Thus, our research will firstly generate one synthetic dataset by means of the generalized ABM for the aircraft's ground handling at their stand according to the actual data and the domain knowledge. After that, we intend to develop and train ML models for the prediction of the aircraft ground time using the synthetic datasets and also validate the same ML models against the actual turnaround data at one certain European airport. We only focus on the performance of the real dataset due to data confidentiality. Finally, a TOBT adherence prediction will be implemented with all the available information before the A-CDM milestone of the final TOBT update.

## III. Research Design and Modeling

In this section, the available turnaround data from a certain European airport is introduced firstly. Then we generate the comprehensive, synthetic duration data for the aircraft ground handling activities by means of the ABM. In the following, the used machine learning algorithms and the features importances interpreting methods are described. Finally, the concrete model building is shown.

### A. Data structure

Nowadays, many airports have equipped the aircraft stand monitoring system so that all the operations around the aircraft can be supervised and recorded in APOC, which includes the total ground handling activities. The actual dataset utilizing for our research is collected between a busy time span of 2019. It provides the normal operations data at a common airport without the COVID-19 negative effects. We only focus on the turnaround process, therefore, the data that presents the stand occupancy overnight is omitted. Besides the actual turnaround process records, the corresponding domain knowledge of each aircraft is also included that differentiates between several input factors, such as scheduled in-block and off-block time (SIBT and SOBT), estimated in-block and off-block time (EIBT and EOBT), TOBT, flight properties (e.g., the aircraft type, the airline), arrival delay, or time conditions (e.g. time of the day, month).

The actual available dataset collects 22,620 aircraft turnaround processes, in which the sub-processes are formatted to the start and end timestamps. Due to technical reasons, some records are lacking. In Fig. 2, the nodes present the specific number of each turnaround sub-process, and the edges describe the connections between them, where one connection is built when two sub-processes jointly belong to an exact turnaround process. The strongest connections exist between deplaning and boarding with 20,135 records. Contrarily, the 2,218 links between unloading and catering are considered as the weakest connections in this real dataset.

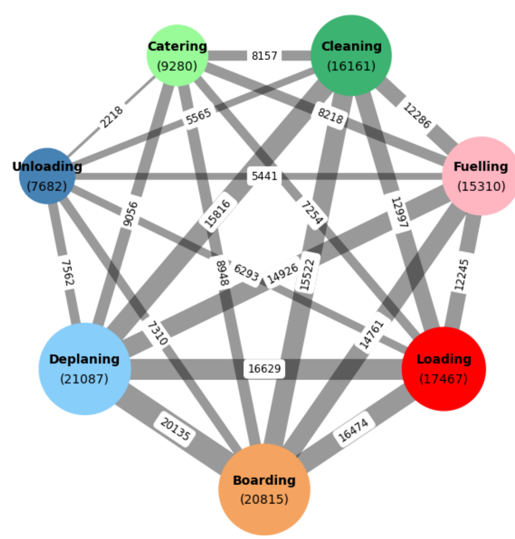


Fig. 2. Actual available turnaround dataset.

### B. Agent-based model

Even though the real dataset missed some records, it still displayed the turnaround sub-processes distribution. Combining with the common structure of the turnaround, the agent types within the proposed agent-based approach are designed according to the GSE respective services. In this model, there are seven types of agents that represent the seven recorded sub-processes. They interact with each other under the fixed sequence restrictions to rebuild the whole turnaround process and generate the relative synthetic data.

The agent types and the reactive restrictions are presented in Table I, which basically comply with the simple turnaround process in Fig. 1. Two serial lines, which are from deplaning to boarding and from unloading to loading, occur independently. The total turnaround time $T$ containing a set of GH activities can be expressed as:

$$T = \max\ (T_{l1}, T_{l2}) + \Delta t, \tag{1}$$

where $T_{l1}$ and $T_{l2}$ are the operation time of the two serial lines, start from deplaning and from unloading, respectively. $\Delta t$ represents the waiting time between the sub-processes due to the increasing time available between the scheduled (actual) operations, additional or idle processes can take place. The detailed sub-process times are deriving from the available real turnaround dataset, whose duration accords its own Normal Distribution ($\mu$, $\sigma$).

TABLE I
AGENT TYPES AND SEQUENCE RESTRICTIONS

| Agent type | Sequence restrictions |
|---|---|
| Deplaning agent (Ramp) | Start after in-block time |
| Unloading agent (Tractor) | Start after in-block time, roughly parallel with deplaning |
| Cleaning agent (Cleaning vehicle) | Start after deplaning |
| Catering agent (Catering vehicle) | Start after deplaning, almost parallel with cleaning |
| Fuelling agent (Fuelling truck) | Start after deplaning, almost parallel with cleaning |
| Loading agent (Tractor) | Start after unloading independently |
| Boarding agent (Ramp) | Start after the longest sub-process (Catering, Cleaning, Fuelling) |

### C. Machine learning with interpretable methods

Thanks to the tremendous enhancement from the modern hardware computing power, the machine learning algorithms are applied to predict future data and behavior, though the underlying mechanisms are not understood fully [28]. To handle the stochastic turnaround process, three ML techniques have been adopted and implemented in Python with the two powerful machine learning libraries, *scikit-learn* [29] and *XG-Boost* [30]. These three algorithms are all tree-liked models, which are Decision Trees, Random Forests, and XGBoost.

A decision tree is a supervised learning method that tests each attribute under a series of given conditions, shunts to different branches, and finally shunts to the leaf nodes of the decision tree to get the final result, whose basic process follows the divide and conquer strategy [31]. Random forest is an ensemble algorithm, which belongs to the Bagging type. By combining multiple weak classifiers, the final result is voted or averaged, so that the result of the overall model has higher accuracy and generalization performance. It can achieve good results mainly due to "random" and "forest", one makes it resistant to overfitting, and the other makes it more accurate [32]. XGBoost (eXtreme Gradient Boosting) is an optimized distributed gradient tree boosting system designed to be highly efficient, flexible, and commodious. This gradient boosting framework means multiple models are built sequentially that each model tries to improve the performance of the previous one [33].

The interpretable ML can prevent the model biases and help the decision-makers understand how to use the model correctly. In some rigorous scenarios, such as the aircraft ground time prediction, the model needs to provide evidence to prove how it works and avoid errors. The traditional feature importances assessment only tells which feature is important, but it does not show how the feature affects the prediction result. Besides the feature importance assessment, we choose the SHapley Additive exPlanations (SHAP) as the variable evaluation method additionally in our research, whose biggest advantage is that SHAP can have an influence on reflecting the characteristics of each sample, and it also shows the positive and negative characteristics of the influence [34].

### D. Model description

In this paper, we build the prediction models in two types. The first type is regression-oriented that intends to forecast the aircraft ground time. And the second one is classification-oriented, which attempts to confirm the TOBT adherence. Table II summarizes the created six datasets for both regression and classification models, which comprise one synthetic dataset and five real datasets.

*1) Aircraft ground time prediction:* In Table II, the first four datasets are the synthetic dataset, Dataset 1, Dataset 2, and Dataset 3 that will be the inputs of the regression models, where the synthetic dataset has 30,000 groups of the generated turnaround data according to our designed agent-based approach. Each turnaround data only contains the duration of the sub-processes. The predicted off-block time (OBT) will be compared with the synthetic OBT. Dataset 1 represents the available actual turnaround records that have the same input types as the synthetic dataset. Dataset 2 is Dataset 1 with the corresponding domain knowledge, i.e., airline, aircraft type, scheduled stand occupancy, and arrival delay, where the categorical data (airline and aircraft type) has been transferred to the numerical data, the scheduled stand occupancy is known as SOBT - SIBT, and the arrival delay can be obtained by AIBT - SIBT, where AIBT is the actual in-block time. Due to that the sub-process data in the first three datasets only presents the duration values, it doesn't provide any turnaround process sequential information to the model. However, the relative start and end of each sub-process can be calculated by choosing AIBT as the starting point. Then these calculated data can describe the sequential information of the turnaround process. For example, in Dataset 3 the deplaning is not the duration value anymore and it is divided into two values, which are the "First PAX Off" and the "Last PAX Off".

In order to ensure the reasonable prediction based on the aircraft stand operations as much as possible, we only add these four domain data inputs from the real turnaround records, which imply the OBT information rarely.

*2) TOBT adherence confirmation:* We target to forecast the TOBT accuracy at the A-CDM milestone of the final TOBT

TABLE II
DATASETS OF THE MODELS

| Feature variables | Datasets for aircraft ground time prediction | | | | Datasets for TOBT adherence confirmation | |
|---|---|---|---|---|---|---|
| | Synthetic dataset[1] | Dataset 1[1] | Dataset 2[1] | Dataset 3[2] | Dataset 4[1] | Dataset 5[2] |
| Deplaning | x | x | x | x | x | x |
| Unloading | x | x | x | x | x | x |
| Cleaning | x | x | x | x | x | x |
| Catering | x | x | x | x | x | x |
| Fuelling | x | x | x | x | x | x |
| Loading | x | x | x | x | x | x |
| Boarding | x | x | x | x | | |
| Airline | | | x | x | x | x |
| Aircraft type | | | x | x | x | x |
| Scheduled stand occupancy | | | x | x | x | x |
| Arrival delay | | | x | x | x | x |
| Estimated arrival difference | | | | | x | x |
| Estimated stand occupancy | | | | | x | x |
| Daytime of AIBT | | | | | x | x |
| Absolute AIBT | | | | | x | x |
| Month | | | | | x | x |

[1]Un-sequential sub-processes; [2]Sequential sub-processes.

update and any aircraft stand operations data obtained before this final update can be set as the model inputs. In the A-CDM implementation manual [2], this milestone is determined ahead of the boarding starting and there are no concrete relations with other turnaround sub-processes. Thus, at this milestone except boarding, we consider other stand operations done. Dataset 4 and Dataset 5 in Table II include all the feature variables for the TOBT adherence confirmation. Currently, the sub-process boarding is excluded because it occurs after this moment, whereas all the domain information has been already received at the TOBT update milestone. The estimated arrival difference is calculated by AIBT - EIBT, and the estimated stand occupancy means EOBT - AIBT. Furthermore, the AIBT has changed into the accumulated value in minutes counting from 0 o'clock. The daytime of AIBT represents the aircraft landing in the early morning, morning, afternoon, evening, and night of the day.

Dataset 4 and Dataset 5 are differentiated with the unsequential and sequential sub-processes, whose domain knowledge is equal. The TOBT accuracy is obtained from TOBT - AOBT, which we further split the values according to Table III. Then the seven TOBT adherence levels are the designed outputs for the classification models, in which the level interval is 10 mins.

TABLE III
TOBT ADHERENCE LEVELS

| TOBT adherence level | Value interval (min) |
|---|---|
| Ultra delay | $(-\infty, -25]$ |
| High delay | [-25, -15) |
| Low delay | [-15, -5) |
| Good match | [-5, 5) |
| Low forward | [5, 15) |
| High forward | [15, 25) |
| Ultra forward | $[25, +\infty)$ |

## IV. SIMULATION RESULTS AND ANALYSIS

The simulation results of the regression and classification models are discussed separately in this section. We always split 30% of the total dataset as the test data in every prediction, therefore, the remaining 70% is the train data.

### A. *Results of the aircraft ground time prediction*

In Fig. 3, we adopt the standard deviation to estimate the aircraft ground prediction. It shows that the XGBoost based on the four datasets always owns the minimal prediction standard deviation. The data types of the synthetic dataset and Dataset 1 are the same, but the synthetic data performs much better. This is because the synthetic data is more ordered, though there is a similar data structure supporting the real stand operating dataset, the realistic turnaround process is still stochastic that is affected by many factors at the airport, such as weather or limited service resources. After adding the domain knowledge with the unsequential sub-process duration data in Dataset 2, and then with the sequential turnaround data in Dataset 3, the standard deviation value is decreasing from 7 min to a minimum of nearly 4.5 min.

Table IV records the feature importances derived from XGBoost concretely, which has the best prediction results in regression. Comparing these values in these four datasets, we discover that boarding is often weighted in the highest, and loading also has the notable feature importance. Therefore, it's inferred that these last two sub-processes express more information than others because they are commonly closed to the end of the turnaround process. In Dataset 3, the feature importance of the boarding end is 0.625, which can be understood that this time point depicts the OBT basically so that the prediction model relies on this data most. Besides boarding and loading in Dataset 2, the scheduled stand occupancy and arrival delay also impact the prediction strong, where the sum of these two values is the available aircraft ground time that sometimes the airport coordinators expect the actual

TABLE IV
FEATURE IMPORTANCES DERIVED FROM XGBOOST IN REGRESSION AND RANDOM FORESTS IN CLASSIFICATION

| Feature variables | Importances from XGBoost regression | | | | Importances from random forests classification | |
|---|---|---|---|---|---|---|
| | Synthetic dataset | Dataset 1 | Dataset 2 | Dataset 3[1] | Dataset 4 | Dataset 5[1] |
| Deplaning | 0.148 | 0.071 | 0.054 | (0.012, 0.014) | 0.059 | (0.039, 0.042) |
| Unloading | 0.133 | 0.062 | 0.036 | (0.016, 0.013) | 0.040 | (0.030, 0.028) |
| Cleaning | 0.068 | 0.094 | 0.052 | (0.014, 0.014) | 0.058 | (0.029, 0.041) |
| Catering | 0.069 | 0.113 | 0.032 | (0.016, 0.018) | 0.032 | (0.019, 0.024) |
| Fuelling | 0.073 | 0.118 | 0.057 | (0.016, 0.023) | 0.062 | (0.033, 0.043) |
| Loading | 0.118 | 0.213 | 0.169 | (0.017, 0.080) | 0.078 | (0.055, 0.070) |
| Boarding | 0.387 | 0.326 | 0.122 | (0.032, 0.625) | | |
| Airline | | | 0.072 | 0.023 | 0.035 | 0.028 |
| Aircraft type | | | 0.080 | 0.027 | 0.033 | 0.026 |
| Scheduled stand occupancy | | | 0.157 | 0.014 | 0.023 | 0.018 |
| Arrival delay | | | 0.164 | 0.018 | 0.134 | 0.112 |
| Estimated arrival difference | | | | | 0.067 | 0.049 |
| Estimated stand occupancy | | | | | 0.182 | 0.167 |
| Daytime of AIBT | | | | | 0.024 | 0.016 |
| Absolute AIBT | | | | | 0.105 | 0.076 |
| Month | | | | | 0.062 | 0.043 |

[1]In Dataset 3 and Dataset 5, the two recorded feature importance values for the sub-processes are formatted by (process start, process end).

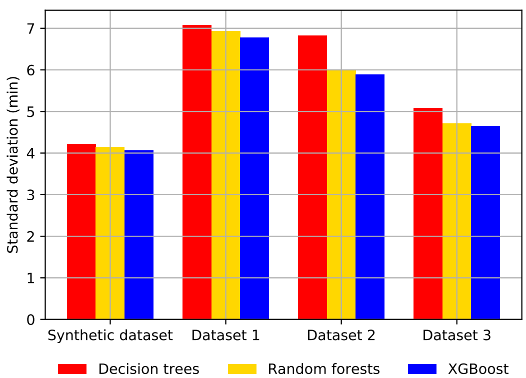


Fig. 3. Standard deviations for the regression models.

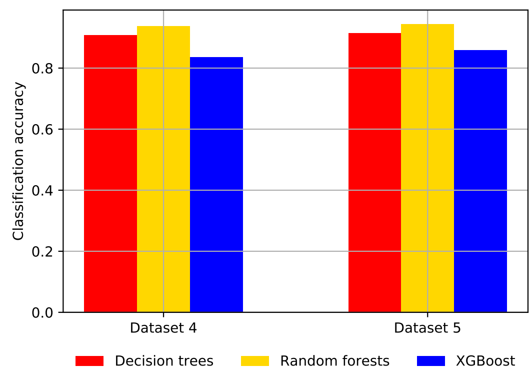


Fig. 4. Accuracy for the classification models.

turnaround duration able to match with this available value by adjusting the airport resources.

### B. Results of the TOBT adherence confirmation

As we designed, all the available data before the A-CDM milestone of the TOBT final update is inputted to predict the TOBT adherence, which results in the prediction model absorbing the maximal accessible information. In Fig. 4, we note that all the classification accuracy values are at least 83%. The best accuracy based on the random forests method is nearly 95%. The similar good results of Dataset 4 and Dataset 5 indicate that in the classification model the unsequential or sequential sub-processes are not the critical input data. Making a detailed review on the feature importances of the best performed random forests model for both datasets in Table IV, we observe that the important values about the sub-processes are totally small. On the contrary, the available time-related information shows high importance, which are the arrival delay, the estimated stand occupancy, and the AIBT.

In addition, we analyze the SHAP values of the model XGBoost based on Dataset 4. Fig. 5 visualizes the TOBT adherence related to the total feature importances and feature effects, and it exhibits the same results of the high importance features as that shown in Table IV. But in the model, the maximum mean SHAP value referring to the estimated stand occupancy is at least two times bigger than the second value. The following Fig. 6 plots the SHAP values only for the category of Good match [-5, 5). In this SHAP summary plot, the SHAP value of each feature for each sample is in a density scatter graph, which can help to understand the overall pattern and allow the discovery of the predicted outliers. The x-axis is the SHAP value for every single feature, and the color represents the characteristic value (red is high, and blue is low). In this picture, we observe that the scheduled stand occupancy is positively related to the TOBT good adherence. When the scheduled time is loose, all the activities can follow the operation process easily. However, the first three main features are negatively related to this TOBT adherence. The longer the estimated stand occupancy time is, the harder the turnaround process finishes in time, which means that there are too many operations needing to process, thus, it's difficult for the aircraft to take off as planned. The long arrival delay refers to the short available ground handling time reflecting that the operation schedule is too tight to follow. Big absolute AIBT states that the plane arrives late in the time of day when the airport crew is off work that there are not many human resources serving. Last but not least, the turnaround sub-process fuelling also shows the strict negative relationship, because fuelling is normally part of the turnaround critical path that the short fuelling duration always ensures efficient aircraft stand operations.

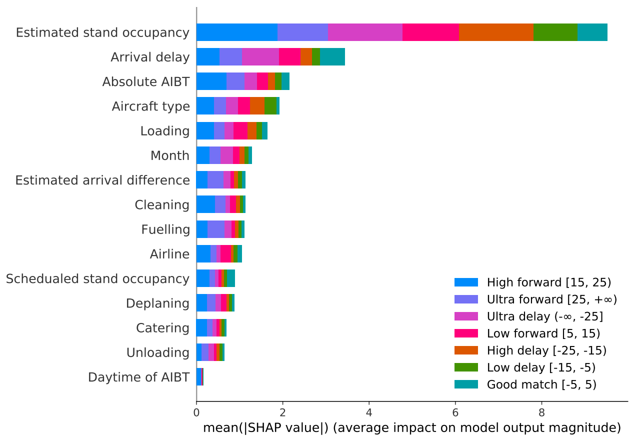


Fig. 5. Summarized mean absolute SHAP values on Dataset 4.

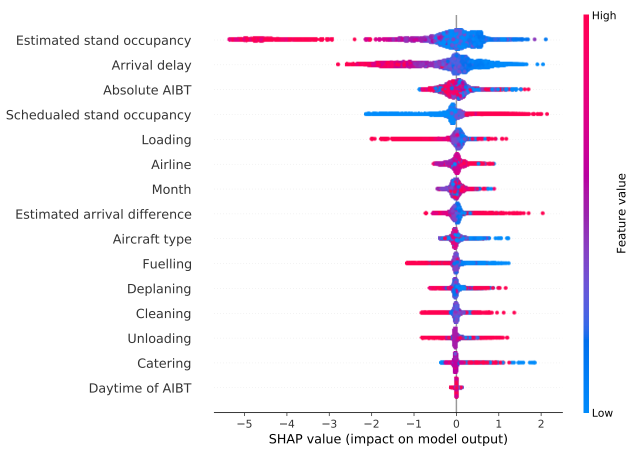


Fig. 6. Density scatter plot of the SHAP values only for the category of good TOBT adherence on Dataset 4.

## V. Conclusions and Future Work

In this paper, we firstly implemented an agent-based approach to generate the comprehensive and synthetic ground handling data that was used to predict the aircraft stand time only based on the clean turnaround sub-processes duration. Then the same machine learning regression models were applied on the actual ground handling datasets with the increasing information inputs. We saw that the prediction results of the regression models also got better accordingly. After that, a TOBT accuracy prediction model by means of the classification methods was developed to confirm the TOBT adherence at the A-CDM milestone of the final TOBT update, which can achieve 95% prediction accuracy in the highest. Meanwhile, the machine learning models were also interpreted with SHAP values. It ensured the model transparency with the available feature importances and the feature affections on the prediction results. This model can be embedded into the APOC so that all the stakeholders at airports are able to benefit from the believable aircraft ground time prediction. And fluent airport operations can also benefit the air traffic network.

The regression models in our research have some shortcomings. Many reasons can contribute to this, such as data lacking, limitations of the models, etc. The issues will be deeply researched in the future. Besides, at the current stage, the generated synthetic dataset is only basically providing the turnaround sub-processes data. We will focus on developing a more efficient and realistic agent-based model to simulate the complex turnaround process, which hopefully could facilitate the total airport management as well.

## Acknowledgment

This publication was supported and funded by EUROCONTROL via contract NO.19-220468-C. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of EUROCONTROL.

## References

[1] H. Wang, M. Wang, and Y. Wu, "Development of an aircraft turnaround time estimation model based on discrete time simulation," in *Information Technology and Intelligent Transportation Systems*. IOS Press, 2017, pp. 29–36.
[2] EUROCONTROL, *Airport CDM implementation manual v5.0*, 2017.
[3] S. Helm, S. Loth, and M. Schultz, "Advancing total airport management–an introduction of performance based management in the airport context," in *19th ATRS World Conference*, 2015.
[4] M. Schultz, S. Reitmann, and S. Alam, "Predictive classification and understanding of weather impact on airport performance through machine learning [accepted]," *Transportation Research Part C: Emerging Technologies*, in press.
[5] EUROCONTROL, *Airport network integration v1.1*, 2018.
[6] ——, *CODA Digest: All-causes delay and cancellations to Air Transport in Europe for 2020*, 2021.
[7] H. Fricke and M. Schultz, "Delay impacts onto turnaround performance," in *ATM Seminar*, 2009.
[8] M. Schultz, T. Kunze, B. Oreschko, and H. Fricke, "Dynamic turnaround management in a highly automated airport environment," in *Proceedings of the 28th International Congress of the Aeronautical Sciences*, 2012, pp. 4362–4371.
[9] H. Fricke and M. Schultz, "Improving aircraft turn around reliability," in *Third International Conference on Research in Air Transportation*, 2008, pp. 335–343.
[10] H. Jin, E. Garcia, and D. N. Mavris, "Simulation of integrated approach for aircraft turnaround process," in *2018 AIAA Modeling and Simulation Technologies Conference*, 2018, p. 0418.
[11] J.-T. Wong and S.-C. Tsai, "A survival model for flight delay propagation," *Journal of Air Transport Management*, vol. 23, pp. 5–11, 2012.
[12] C. Malandri, L. Mantecchini, and V. Reis, "Aircraft turnaround and industrial actions: How ground handlers' strikes affect airport airside operational efficiency," *Journal of Air Transport Management*, vol. 78, pp. 23–32, 2019.

[13] M. Schultz, J. Evler, E. Asadi, H. Preis, H. Fricke, and C.-L. Wu, "Future aircraft turnaround operations considering post-pandemic requirements," *Journal of Air Transport Management*, vol. 89, p. 101886, 2020.

[14] M. Schultz, J. Rosenow, and X. Olive, "A-cdm lite: Situation awareness and decision-making for small airports based on ads-b data," in *9th SESAR Innovation Days: Inspiring Long-Term Research in the Field of Air Traffic Management, SIDs 2019*, 2019.

[15] M. Schultz, T. Kunze, B. Oreschko, and H. Fricke, "Microscopic process modelling for efficient aircraft turnaround management," *Air Transport and Operations*, 2013.

[16] J. Nosedal Sanchez and M. A. Piera Eroles, "Causal analysis of aircraft turnaround time for process reliability evaluation and disruptions' identification," *Transportmetrica B: Transport Dynamics*, vol. 6, no. 2, pp. 115–128, 2018.

[17] B. Oreschko, T. Kunze, M. Schultz, H. Fricke, V. Kumar, and L. Sherry, "Turnaround prediction with stochastic process times and airport specific delay pattern," in *International Conference on Research in Airport Transportation (ICRAT), Berkeley*, 2012.

[18] M. Schmidt, "A review of aircraft turnaround operations and simulations," *Progress in Aerospace Sciences*, vol. 92, pp. 25–38, 2017.

[19] D. A. Tabares, F. Mora-Camino, and A. Drouin, "A multi-time scale management structure for airport ground handling automation," *Journal of Air Transport Management*, vol. 90, p. 101959, 2021.

[20] W. Ip, V. Cho, N. Chung, and G. Ho, "A multi agent based model for airport service planning," *International Journal of Engineering Business Management*, vol. 2, p. 7, 2010.

[21] R. Erol, C. Sahin, A. Baykasoglu, and V. Kaplanoglu, "A multi-agent based approach to dynamic scheduling of machines and automated guided vehicles in manufacturing systems," *Applied soft computing*, vol. 12, no. 6, pp. 1720–1732, 2012.

[22] V. Kaplanoğlu, C. Şahin, A. Baykasoğlu, R. Erol, A. Ekinci, and M. Demirtaş, "A multi-agent-based approach to dynamic scheduling of machines and automated guided vehicles (agv) in manufacturing systems by considering agv breakdowns," *International Journal of Engineering Research & Innovation*, vol. 7, no. 2, pp. 32–38, 2015.

[23] M. Schultz, S. Lorenz, R. Schmitz, and L. Delgado, "Weather impact on airport performance," *Aerospace*, vol. 5, no. 4, p. 109, 2018.

[24] Y. Gao, Z. Huyan, and F. Ju, "A prediction method based on neural network for flight turnaround time at airport," in *2015 8th International Symposium on Computational Intelligence and Design (ISCID)*, vol. 2. IEEE, 2015, pp. 219–222.

[25] C. Wan, C. Zhaoxin, D. Yuxian, L. Yang, P. Ye, and H. Tao, "Dynamic prediction about turnaround time of flight based on support vector machine regression," in *2019 IEEE 1st International Conference on Civil Aviation Safety and Information Technology (ICCASIT)*. IEEE, 2019, pp. 109–113.

[26] M. Schultz and S. Reitmann, "Machine learning approach to predict aircraft boarding," *Transportation Research Part C: Emerging Technologies*, vol. 98, pp. 391–408, 2019.

[27] E. Halmesaari *et al.*, "Master thesis: Interpretable machine learning for prediction of aircraft turnaround times," Master's thesis, Aalto University, 2020.

[28] D. Bzdok, N. Altman, and M. Krzywinski, "Points of significance: statistics versus machine learning," *Nature Methods 2018a*, pp. 1–7, 2018.

[29] Scikit-learn, "see "https://scikit-learn.org/stable/index.html" for details."

[30] XGBoost, "see "https://xgboost.readthedocs.io/en/latest/index.html" for details."

[31] A. J. Myles, R. N. Feudale, Y. Liu, N. A. Woody, and S. D. Brown, "An introduction to decision tree modeling," *Journal of Chemometrics: A Journal of the Chemometrics Society*, vol. 18, no. 6, pp. 275–285, 2004.

[32] G. Biau and E. Scornet, "A random forest guided tour," *Test*, vol. 25, no. 2, pp. 197–227, 2016.

[33] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.

[34] C. Molnar, *Interpretable machine learning*. Lulu. com, 2020.