PRE-PROCESSING AND ANALYSIS OF RECORDED AIRPORT DATA FOR MORE REALISTIC SIMULATIONS

M. Schaper, L. Nöhren, N. Wüstenbecker, A. Pick, A. Papenfuß, M. Helms
Deutsches Zentrum für Luft- und Raumfahrt e.V., Lilienthalplatz 7, 38108 Braunschweig

Abstract

In the air traffic control domain simulations are a well-known method for concept validations. Each simulation shall be as realistic as needed. Traditionally, expert knowledge is used to shape the simulation environment to meet the requirements for validation. Since air traffic data becomes more and more available, the question arose, how this air traffic data can be processed, if it is of the needed precision and if it can be used to shape simulations. The paper describes the simulation environment targeted and the methods applied to process and filter data, recorded at the Airport Research and Innovation Facility Hamburg. Some exemplary results are presented and discussed, showing that the presented methodology can be used to improve realism of simulation environments.

1. INTRODUCTION

New concepts for Air Traffic Control (ATC) need to be tested and validated before introducing them to real operations. The NASA Technology Readiness Level (TRL) [1] defines 9 levels. Level 4 addresses components and level 5 prototypes tested in a realistic simulation environment. This is a very important step on the way to the introduction of systems in the operational environment due to safety concerns and costs. Simulations allow extensive testing in a realistic environment that would not be possible in the real-world without substantially disrupting air traffic. How realistic a simulation environment has to be depends on the type and stage of the prototype to be validated. For an early prototype, a simplified simulation might be enough to validate first functionalities. A high-fidelity prototype needs a more complex simulation environment, for example with a human-in-the-loop setup. Which requirements the simulation environment has to fulfil is determined through the E-OCVM (European Operational Concept Validation Methodology) process [2]. If the simulation environment has to be shaped near to reality, expert knowledge is needed to replicate the precise details according to the real-world environment. Especially if air traffic controllers from the replicated ATC environment participate in the simulations, small inaccuracies can have a major impact on the validation results. Although ATC environments have some commonalities, each environment has its specific characteristics. For example, in the airport environment, which this work focuses on, there are always taxiways for flights moving on the airport surface. However, each airport has its own specifics and procedures that are not inherently obvious without extensive knowledge of the operations at that airport.

Air traffic data gets more and more available. On the one hand due to the rising number of airports employing A-SMGCS (Advanced Surface Movement Guidance and Control Systems) which are able to collect and record large amounts of data. On the other hand, there are some publicly available sources of flight data, usually based on ADS-B (Automatic Dependent Surveillance Broadcast) data, e.g. FlightRadar [3] or OpenSky [4]. The question emerged, if this traffic data could be used to shape simulation environments and to increase realism without or with less expert knowledge. The goal of the paper is to identify data that can be applied to different airports, e.g. behaviour of certain aircraft types or airlines, as well as data relevant for a specific airport. Our specific questions are:

1. How can recorded data be processed and is it of the required precision?
2. Can the recorded data be used to find out
    a. Taxi velocities in specific areas?
    b. If different airlines taxi with different speeds?
    c. How pushback behaviour differs when using different aircraft types?
3. Is it possible to use recorded traffic data to make simulations more realistic without using expert knowledge?

Of course, many more questions could be raised. For this work we focus on these three as they are most relevant for the simulation of current DLR projects like THOR (Towards Zero Emission Airports) [5]. One part of THOR dealt with towed/sustainable taxi operations which was to be validated using simulated Hamburg airport with runway configuration 05 for departures and 15 for arrivals.

2. RELATED WORK

High-fidelity airport surface simulations are commonly used to validate new operational concepts, for example, trajectory prediction and conflict-free ground movement planning [6]. As the calculated trajectories for all aircraft need to be conflict-free, the planning is required to be adaptive, as the executed trajectories might not be exactly the planned ones, as e.g. taxi speed differs. The realistic modelling of these taxi speeds is typically described in an aircraft performance model. Traditional performance databases such as EUROCONTROL's Base of Aircraft Data (BADA) [7] provide standardized aircraft performance parameters; however, these models often rely on nominal values that fail to capture operational variability. For instance, BADA specifies constant ground speeds of 15 knots for all aircraft types during taxi operations, regardless of aircraft category or operational context. Consequently, surveillance data has enabled data-driven approaches to aircraft performance modelling. Sun et al. developed OpenAP [8], an open-source aircraft performance model

doi: 10.25967/650191

constructed from open aircraft surveillance data. While the developed models primarily focus on airborne performance, their methodologies demonstrate the feasibility of deriving accurate, aircraft-specific parameters from surveillance data. Most works for ground operations utilize such data to evaluate airport performance; for example, Balakrishnan et al. [9] use Airport Surface Detection Equipment, Model-X (ASDE-X) data to analyze average taxi speeds as time series data. Similarly, other studies have characterized ground traffic at Zurich airport based on ADS-B data [10], utilizing the data to examine ground traffic patterns in complex environments. To facilitate these ground operations analyses, the traffic library for Python provides a toolkit for processing ADS-B trajectories [11]. This library supports the research community by offering filtering methods, such as Kalman filters and trajectory mapping to airport layouts, to mitigate noise and data gaps. Beyond general data processing, these capabilities enable specific infrastructure assessments, as evidenced by recent work that applies trajectory analysis to measure how Rapid Exit Taxiways affect airport capacity [12].

3. METHODS

This chapter first introduces the simulation environment we are targeting before describing the used dataset and the methods applied to filter, process and extract the data.

3.1. Simulation Environment

The motivation for this work was to improve the realism of human-in-the-loop simulations for a tower environment. To give an example, a currently used simulation setup of DLR is presented here. Figure 1 shows the ATS360 [13]: A tower simulator with 360-degree outside view and multiple configurable working positions. The NARSIM software platform [14] is used as traffic simulator. The movement of aircraft is controlled by simulation pilots in a separate room. The reaction of flights to given commands (e.g. acceleration) is modelled in NARSIM with a physical model based on BADA data [7]. This model already ensures a high degree of realism for the aircraft behaviour. Nevertheless, some standard parameters need to be set as basis for this model and default behaviour of flights if the simulation pilots don't specify the target behaviour. One example for that is the default taxi speed. This speed is set to 15 knots. During acceleration, breaking and curves the speed is reduced according to the physical model. The pilots could enter different speeds manually, but in most circumstances, the standard speed is used on straight taxiways. Factors such as aircraft type, airline, aircraft weight or taxiway location do not have an influence on this value. The pushback procedure, i.e. when flights are moved backwards out of their parking position by a separate vehicle, also relies on some default values to model the behaviour. The simulation pilots can select a target end point of the pushback in the topology. The pushback is then performed with a speed of 8 knots. At the end of the pushback, the flight has to wait some time until the engines are started and the flight is ready for taxi. This time is configurable per engine. Currently 90 seconds are used as startup time of one engine. This time is independent from the type of engine or aircraft. That means all aircraft with two engines will have the same startup time, no matter their size or type. There are other examples of such standard values, but in this analysis, we will mainly focus on startup, push and taxi as a first step. Of course, all of these default values can be changed via the configuration of the simulator, or in some cases be overridden by the simulation pilots. But to achieve a higher degree of realism, realistic values need to be known for these parameters.

Figure 1: The ATS360 tower simulator at the DLR Braunschweig.

3.2. Example Dataset

The data used for this work was received from the ARIF (Airport Research and Innovation Facility) Hamburg [15]. ARIF was created in 2006 as collaboration between DLR, DFS (Deutsche Flugsicherung GmbH) and FHG (Flughafen Hamburg GmbH). It can record real world A-SMGCS (Advanced Surface Movement Guidance and Control System) data from Hamburg airport and provide it for research and development projects without disturbing the operational systems. The data is collected by performing a fusion of multiple data sources like SMR (Surface Movement Radar), ASR (Airport Surface Radar), MLAT (Multilateration) and ADS-B (Automatic Dependent Surveillance – Broadcast). It has therefore high quality and directly mirrors the data that is available for the operational systems at the airport. Data is collected and used compliant to the cooperation framework of DFS, FHG and DLR.
Six months of 2024 were selected for this analysis (February, March, May, June, August, September). This way the dataset contains data from each season of the year, ensuring that different weather conditions, different traffic loads and different runway configurations are included. After filtering and pre-processing (see following sections) the dataset contains 57.422 flights divided in 29.229 arrivals and 28.193 departures. The data had a wide variety, containing numerous different aircraft types and airlines.

3.3. Processing

The following sections describe the processing of the dataset to extract the desired information. A first step was a rough filtering, afterwards a pre-processing was performed with different algorithms and finally the information was extracted using a publicly available python library.

3.3.1. Preparation and Filtering

The recorded raw data used for this study are track data (fused from different sensors) in the EUROCONTROL ASTERIX [16] format. It consists data of all movements of aircraft and equipped vehicles in and around the airport and is recorded every second. In a first step, they were transformed into tabular datasets with interpretable data fields, like callsigns, geo-coordinates and timestamps. Second, data logs were summarized per month and the datasets were checked if recordings are available for all days. Data logs were sorted to form consistent tracks per flight and the quality of these flight tracks were checked before conducting further analysis. Tracks were excluded in case they contained time lapses of more than 300 second / five minutes (1.2%), unplausible gaps in geo-coordinates with a distance per second above 260 meters (6.4%) or tracks which were shorter than 60 entries (< 1%). The criteria for this filtering were developed iteratively and checked with subject matter experts (SME) for plausibility. The percentage values show that around 8% of flight tracks were excluded.

For each monthly dataset, performance values like movements per hour, movements per month and movements per airline were determined. Those were compared with published data about EDDH's traffic. The goal was to ensure that the analysis of taxi times uses representative data samples regarding traffic numbers. Lastly, for each hour within the collected data, based on the runways used for take-off and landing, runway configurations were determined.

To simplify further analysis that focuses on aircraft types and airlines, the aircraft types and airlines with a small number of data points (less than 1000) were ignored, resulting in a list of 50 aircraft types and 65 airlines. For analysis in which these factors play no role, all aircraft types and airlines are considered. The most used aircraft type was the Airbus A320 with a total of 618.469 data points in the dataset. The airline with the most data points had 397.512 in the dataset.

3.3.2. Pre-Processing

Using recorded real data, means coping with uncertainties due to measurement inaccuracies, jitter, and interpolation artifacts. Therefore, AI (Artificial Intelligence) techniques have been used to smooth the trajectories on ground. The approach was to use only the raw movement data and appropriate filters to produce realistic representations of aircraft paths [17]. First a Gaussian filter was used. It reduces minor anomalies and high-frequency noise by emphasizing the central data points and gradually diminishing the influence of neighbouring points (cp. Figure 2). At the seconds step a Hatch filter was applied to the data. This filter is effective in smoothing geospatial data. Also, it is less sensitive to outliers compared to the Gaussian filter. The Gaussian filter parameters were set to sigma = 1, and a gap_threshold = 5 seconds, to avoid smoothing across larger gaps. The Hatch filter parameters were set to alpha = 0.25 (balancing jitter in straight parts and preserving curvature in turns), and again a gap_threshold = 5 seconds. Both filters were successively applied forward and backward to preserve both ends of a trajectory properly. The following figures show the single effect of each filter (cp. Figure 2 and Figure 3).

3.3.3. Information Extraction

Aircraft ground trajectory processing is performed using the Python library traffic [11], which provides comprehensive functionality for aviation data analysis. Airport infrastructure data, including detailed runway and taxiway designations, are obtained from OpenStreetMap to establish the spatial reference framework. To account for inherent positioning uncertainties and map trajectories to the underlying taxiway network, a Kalman filter incorporating taxiway geometric constraints is employed, effectively projecting observed positions onto the nearest taxiway segment (see Figure 4).

Figure 2: Comparison of raw data (black points) and Gaussian filter applied (green and blue points).

Figure 3: Comparison of raw data (black points) and Hatch filter applied (green and blue points).

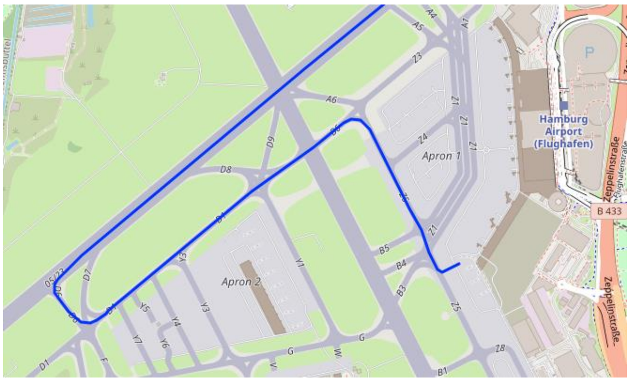

Figure 4: Trajectory generated by the traffic library using the Kalman filter with taxiway constraints.

Movement detection is implemented through a velocity-based threshold, classifying aircraft as moving when maintaining a minimum ground speed of 2 knots for at least

30 seconds. Ground speeds are computed from pre-processed position data with 1 Hz temporal resolution, enabling consistent velocity estimation throughout the trajectory. Additionally, pushback operations are automatically identified by analysing the combined characteristics of ground speed and heading angle, distinguishing these rearward movements from normal taxiing patterns. This multi-stage processing pipeline ensures robust extraction of aircraft ground movement features while maintaining consistency with actual airport topology.
The trajectories generated this way contain all relevant information like the speed at each point, aircraft type, airline and are mapped to the taxiways of the airport. Therefore, this trajectory data structure can be used to extract the desired information in a simplified manner. The extracted data can then be filtered and cumulated as desired for the analysis. This means, once the trajectories are generated and stored, the extraction step is very simple and computationally inexpensive. This allows a wide variety of different analyses.

4. RESULTS

In the following sections some results that were extracted from the ARIF Hamburg dataset are presented. The research questions guided the analysis and the evaluation serves as example for what kind of results can be expected using the described analysis process.

4.1. Relevant Airport Areas

For our current simulation setup, specific areas of the airport are especially relevant. Therefore, we focused our analysis on these areas and ignored the remaining parts of the airport in this evaluation. Figure 5 shows five groups that were selected as most relevant. The groups consist of different taxiways that lead from or to the active runways and the apron.

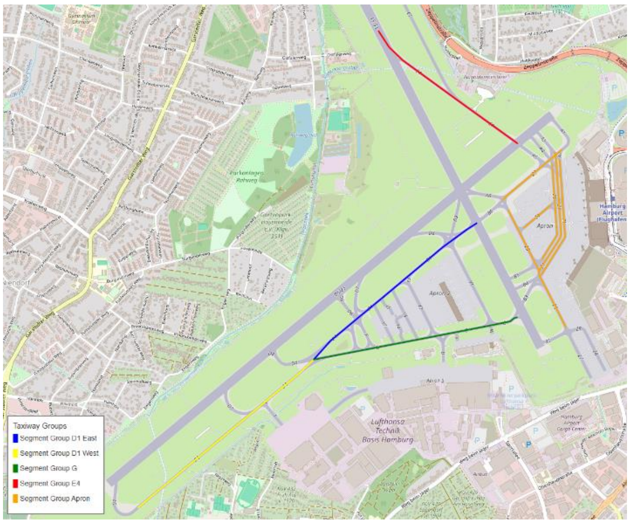

Figure 5: The most relevant taxiways are grouped into 5 areas.

Differentiating the results in the different groups was necessary, as we expected different results for some of these groups. For example, we expected a different average taxi speed on the apron compared to other taxiways. Grouping multiple parts of the topology into an area (e.g. apron), makes the comparison simpler by reducing the number of values to compare.

4.2. Taxi Speed

Figure 6 shows the average taxi speeds on the five areas of the airports for arrivals and departures as a box plot. The boxes contain the median as centreline and depict the upper and lower quartile of the data. For the length of the whiskers, 3.5 times the interquartile range (IQR) was used. The currently used standard speed of 15 knots is displayed as a dotted grey line. On most areas, the usual taxi speed is higher than 15 knots averaging between 23 and 19 knots. On the apron the average speed is considerably lower with 12 knots for arrivals and 8 knots for departures. That shows that 15 knots are overall a relatively good approximation for the average speed over all airport areas. But when focusing on specific areas, the realism could be increased by using multiple standard values depending on the area of the airport (e.g. apron vs. taxiways). To find the reason for the different average taxi speeds for arrivals and departures on the apron, it is possible to analyse the results in this area in more detail. The histogram in Figure 7 shows the number of data points for the different speed values on the apron.

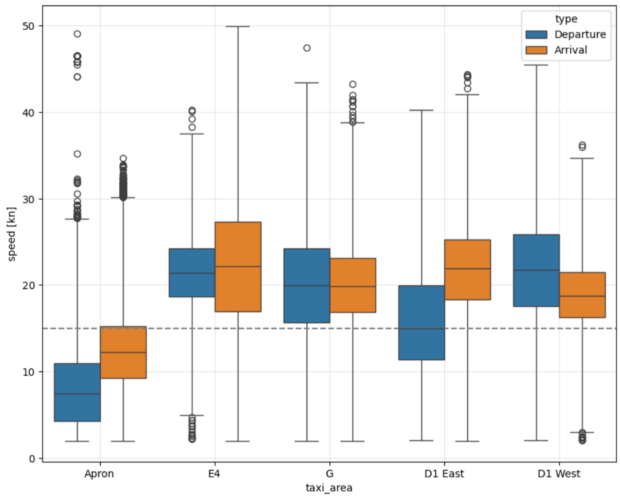

Figure 6: The taxi speeds of arrivals and departures on different airport areas.

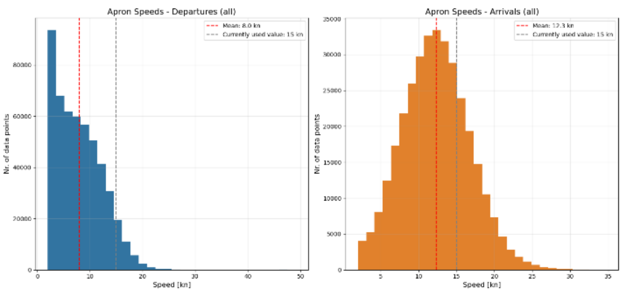

Figure 7: The taxi speeds of arrivals and departures on the apron.

As mentioned above, any data point with a speed less than two knots is not classified as movement and therefore not contained in that diagram. For departures, the highest point of the histogram is at the lowest possible speed of two knots. An explanation for this is, that most departures perform a pushback on the apron. During the pushback and in the acceleration phase when starting to taxi, the aircraft will start moving with a lower speed than in average over all speeds leading to the higher concentration of low speeds and the low average speed of 8 knots. Arrivals that do not perform such a manoeuvre show a typical normal distribution of speeds with an average speed value of 12 knots. Taxi speeds are also influenced by other factors than the aircraft type or the position on the airport. One of those factors is the airline. Figure 8 contains the taxi speeds of four selected airlines showing significant differences. The average taxi speed can differ by more than 5 knots between certain airlines. Another factor influencing the taxi speed is the aircraft type. Figure 9 shows that the widebody aircraft B77W (Boeing 777-300ER) taxis on average 2 to 5 knots slower than smaller aircraft. The engine type seems to have no significant impact on the taxi speeds as the A320 (Airbus A320) and A20N (Airbus A320-Neo), which are similar aircraft with different engines, have very similar speeds. The turboprop aircraft DH8D (De Havilland DHC-8) also moves with a similar speed to the medium sized jet aircraft from the Airbus A320 family.

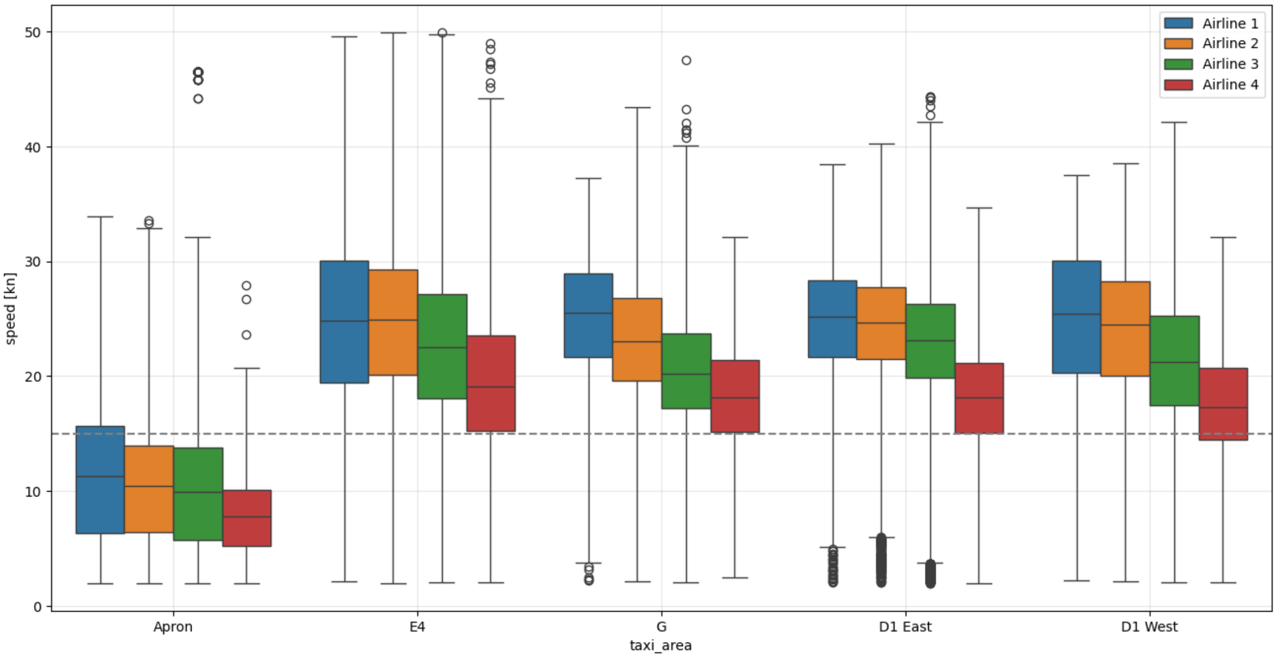

Figure 8: Taxi speeds of different airlines on different airport areas.

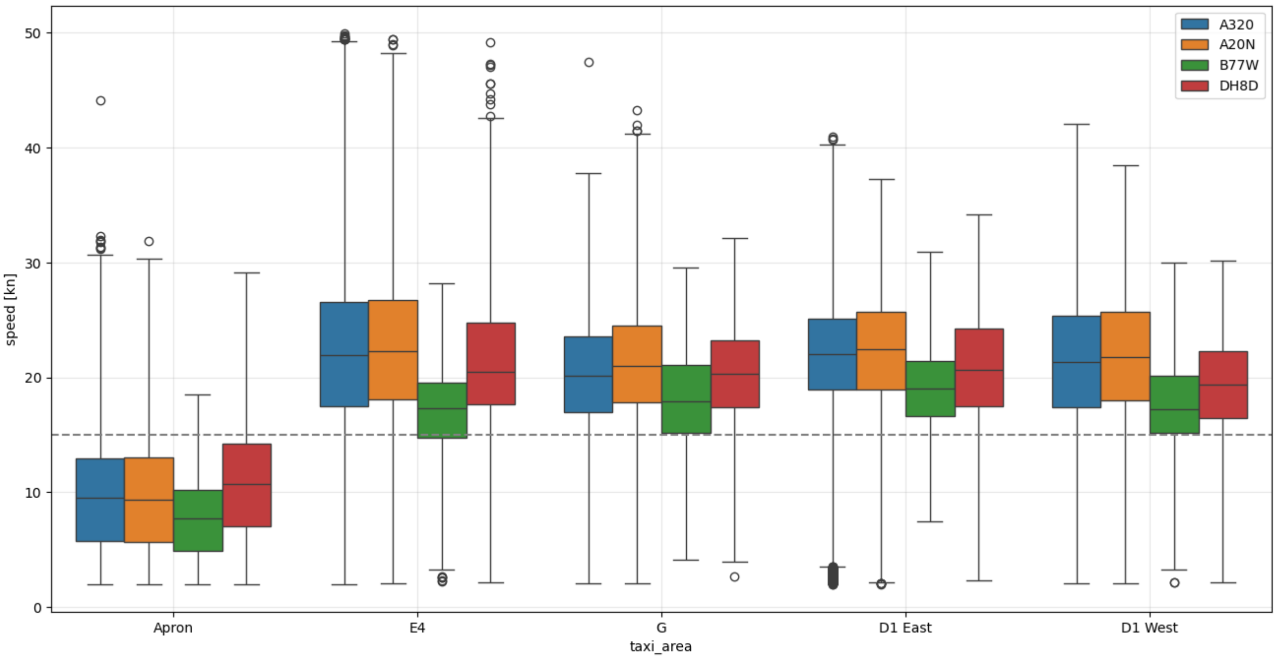

Figure 9: Taxi speeds of different aircraft types on different airport areas.

4.3. Pushback Times

Figure 10 shows the times departures spent during the pushback procedure. The left side shows the time spent during movement in the pushback and the right side shows the time the aircraft stood still after completion of pushback. This is the time when the engine startup and, dependent on airline specific operations, some cockpit preparations are completed. In total the average duration of pushback and hold combined of all aircraft is 261.8 seconds. As mentioned before, the currently used time for that in the DLR simulator is 180 seconds for aircraft with two engines and 360 seconds for aircraft with four engines.

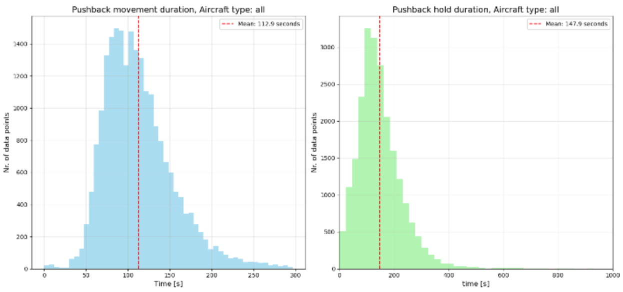

Figure 10: The pushback movement time and pushback hold time.

To receive a more detailed result, another evaluation of these times was done for different aircraft types, shown in Figure 11. It is noticeable that the pushback movement duration of the heavy aircraft A333 (Airbus A330-300) is significantly longer than all other types that were evaluated here, which are all medium sized aircraft. The pushback hold duration of this type is also longer than that of most other types. The two aircraft of the Airbus Neo family (A20N and A21N) also show a similarly long pushback hold duration. The movement duration of these types is closer to the movement duration of other aircraft types of similar size. The reason for the longer hold duration is the different engine type that is used in the Neo family. The engine startup lasts much longer, but the rest of the pushback procedure is not influenced by that.

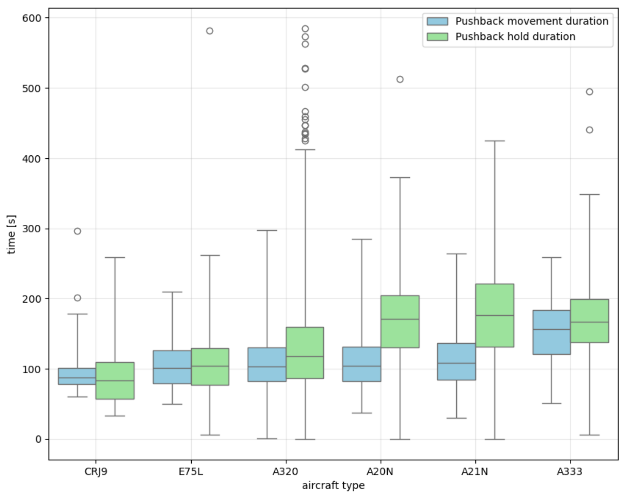

Figure 11: The pushback movement and pushback hold time for different aircraft types.

As only a very small number of aircraft with other than 2 engines that performed pushback was contained in the dataset, it was not possible to validate whether the differentiation of startup times based on number of engines is realistic. Nevertheless, it can be observed that there are even strong differences when just comparing different aircraft types with two engines. A differentiation based on aircraft type seems to be more realistic than based only on the number of engines.

5. DISCUSSION AND CONCLUSION

Section 3.3 describes the steps applied to process the used data from the ARIF platform and shows, that the precision of the data is satisfying. Therewith research question 1 *How can recorded data be processed and is it of the required precision?* is answered in the context of the ARIF dataset. As this dataset is currently the first and only example, the remaining research questions have to be put in the context of this limitation. Research question 2a *Can the recorded data be used to find out taxi velocities in specific areas?* is answered positive as well, as Figure 6 depicts taxi speeds in different airport areas. Figure 8 differentiates taxi speed to different airlines and answers question 2b *Can the recorded data be used to find out if different airlines taxi with different speeds?*. Research question 2c *Can the recorded data be used to find out how pushback behaviour differs when using different aircraft types?* is also answered positive, as Figure 11 depicts. Research question 3 *Is it possible to use recorded traffic data to make simulations more realistic without using expert knowledge?* might be answered in two parts. First, when implementing the findings of the analysed dataset from ARIF platform, the simulations will get more realistic. For the study dealing with towed/sustainable taxi within the THOR project we introduced specifically an Airbus A320-Neo into the traffic scenarios with the pushback movement and hold duration found in this evaluation. The second part, if it is possible to use recorded traffic data to make simulations more realistic *without using expert knowledge* is not trivial to answer. It depends also on the definition of what is "expert knowledge". The authors worked with the simulated Hamburg airport and had some interviews with air traffic controllers. However, the authors might not be experts for Hamburg airport. On the other hand, they are already skilled up to a certain level. We can conclude, that a certain level of knowledge might be helpful to focus on the right questions and to finetune the simulation environment accordingly. In turn, the results of the recorded data's analysis can be used to attest experts' subjective experience by externalisation.

6. SUMMARY AND OUTLOOK

The recorded ARIF data required some pre-processing, but could be processed successfully and is precise enough to allow evaluations of specific questions, like taxi speeds for different areas on the airport. This knowledge can be used to shape the configurations of the simulation environment to be more realistic. Nevertheless, it has to be carefully decided, how realistic the simulation environment has to be for the purpose it is going to be used actually. The gained experience of processing the recorded data can be transferred to analyse data from other airports. One future question is, if also publicly available data are of sufficient quality to be assessed. As another future step recorded data might be used to train AI algorithms or to assess airports' performance.

7. REFERENCES

[1] NASA, "Technology Readiness Levels," [Online]. Available: https://www.nasa.gov/directorates/somd/space-communications-navigation-program/technology-readiness-levels/. [Accessed 10 11 2025].

[2] EUROCONTROL, "E-OCVM Version 3.0," 2010. [Online]. Available: https://www.eurocontrol.int/sites/default/files/2019-08/e-ocvm-version-3.0-volume-1.pdf. [Accessed 10 11 2025].

[3] "flightradar24 - live air traffic," [Online]. Available: https://www.flightradar24.com/. [Accessed 10 11 2025].

[4] "OpenSky Network - Open Air Traffic Data for Research," [Online]. Available: https://opensky-network.org/. [Accessed 10 11 2025].

[5] DLR, "THOR," [Online]. Available: https://www.dlr.de/de/fl/forschung-transfer/projekte/thor. [Accessed 10 11 2025].

[6] L. Nöhren, M. Schaper, L. Tyburzy and K. Muth, "REAL-TIME CALCULATION AND ADAPTION OF CONFLICT-FREE AIRCRAFT GROUND TRAJECTORIES," in *33rd Congress of the International Council of the Aeronautical Sciences, ICAS 2022*, Stockholm, 2022.

[7] EUROCONTROL, "Base of aircraft data," [Online]. Available: https://www.eurocontrol.int/model/bada. [Accessed 10 11 2025].

[8] J. Sun, J. Ellerboek and J. M. Hoekstra, "WRAP: An open-source kinematic aircraft performance model," *Transportation Research Part C: Emerging Technologies,* vol. 98, pp. 118-138, 2019.

[9] H. Khadilkar, H. Balakrishnan and B. Reilly, "Analysis of Airport Performance using Surface Surveillance Data: A Case Study of BOS," in *11th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, Virginia Beach, VA, 2011.

[10] X. Olive and L. Basora, "A Python Toolbox for Processing Air Traffic Data: A Use Case with Trajectory Clustering," in *Proceedings of the 7th OpenSky Workshop 2019*, Zurich, Switzerland, 2019.

[11] "traffic - air traffic data procesing with Python," [Online]. Available: https://github.com/xoolive/traffic. [Accessed 10 11 2025].

[12] M. Waltert and B. Figuet, "Using ADS-B trajectories to measure how rapid exit taxiways affect airport capacity," *Journal of Open Aviation Science,* vol. 1, 2023.

[13] DLR, "ATS," [Online]. Available: https://www.dlr.de/de/fl/forschung-transfer/validierungszentrum-luftverkehr/echtzeitsimulation/ats. [Accessed 10 11 2025].

[14] NLR, [Online]. Available: https://www.nlr.org/newsroom/facility/narsim/. [Accessed 10 11 2025].

[15] DLR, "Airport Research and Innovation Facility Hamburg," [Online]. Available: https://www.dlr.de/de/fl/forschung-transfer/validierungszentrum-luftverkehr/externe-testumgebungen/airport-research-and-innovation-facility-hamburg. [Accessed 10 11 2025].

[16] EUROCONTROL, "All-purpose structured EUROCONTROL surveillance information exchange," [Online]. Available: https://www.eurocontrol.int/asterix. [Accessed 10 11 2025].

[17] N. Persigehl, "Data Science Internship – Enhancing Air Traffic Insights Trough Data-Driven Solutions," DLR, 2025.