

EDITORIAL

Reviews and responses for

Using ADS-B Trajectories to Measure How Rapid Exit Taxiways Affect Airport Capacity

Authors: Manuel Waltert and Benoit Figuet

Reviewers: Luis Delgado and Xavier Olive

Editor: Junzi Sun

1. Original paper

DOI for the original paper: <https://doi.org/10.59490/joas.2023.7207>

2. Review - round 1

2.1 Reviewer 1

The paper uses ADS-B data to estimate the number of departures and arrivals at Lisbon airport for two time periods: pre and post-installation of a rapid exit taxiway (RET). The data is then grouped in 15-minute intervals to compute the relationship between departures and arrivals, and this information is used to fit a capacity envelope.

If I am not mistaken, the methodology to estimate the capacity envelope from the observed arrivals and departures is not new. Therefore, the contributions are two-fold: the use of (open) ADS-B data for this purpose (and the required data preparation needed) and the estimation of the benefit of the rapid exit taxiway on these airports' capacities.

In the first part, the authors provide an XGboost method to estimate the remaining duration of the arrival and departure trajectory missing due to the poor low-altitude coverage of ADS-B traces.

The results shown in Figure 3 are very impressive and highly accurate. As reported, the mean square error is 21.15 seconds.

However, it's unclear why two very distinct patterns are observed in the data (Figure 3). What are the characteristics of the flights with the much steeper time-to-fly prediction with respect to the distance from the runway threshold?

Secondly, I understand that the authors train the XGBoost models using only the subset of flights for which their trace is 'complete'; how many flights are these?

Six features are used for these models: distance to threshold, aircraft type, ground speed, altitude, track angle and rate of climb. The traces of these flights are full to the runway, so how the 'distance to the threshold' is computed should be clarified. Do the authors sample the trace from the ground up to 12 NM?

It would be interesting to understand the importance of the six features used. The results shown seem very linear, so it would be useful to understand the benefit of these additional features with respect to a simple linear fit as a function of the distance. Computing the difference in prediction between these two models (the ML XGBoost and a simple linear regression) and the error of each model would be useful.

Finally, these models provide an estimation between 0 and around 300 seconds (5 minutes), but as shown in Figure 2 and Figure 4, the distance to/from the threshold seems to be usually lower than 5 NM and 1 NM for landings and departures, respectively. Therefore, the 'time to fly' estimation would range mostly between 0 and 200 seconds (3 minutes). Considering that the authors will then aggregate the data (arrivals and departures) every 15 minutes to compute the departure/arrival relationship, how important is this machine learning model for this particular problem? How would the results change if, instead of this, a simple linear approximation is used to estimate the time remaining/time before as a function of the distance of the aircraft with respect to the threshold? This might be estimated by recomputing the values used to generate Figure 6 with this simpler approach.

Then, the authors rightly identify the need for 'good' coverage and use that to filter the case study to LPPT from the potential 7 airports which implemented some RET in the time period considered. However, not much detail is provided on the specific steps for this process. Please describe this in more detail.

The second contribution is estimating the benefit of building a rapid exit taxiway. The authors mention that they discarded Zurich as a candidate airport for the analysis as their discussion with experts led them to identify that the rapid taxiway developed there would be used to allow 'soother' operations. One could argue that that's the impact of all of these infrastructures. Wouldn't smoother operations on delay-prone operations not lead to higher airport usage? How is this different in the Lisbon case? It might be useful to provide a diagram chart of Lisbon indicating the location of the RET to understand its potential usage better.

Finally, I wonder if there are theoretical studies on the expected benefit of implementing a rapid exit taxiway on the runway occupancy time (ROT). If this is the case, the authors could compute the expected benefit on runway occupancy for Lisbon due to the RET. This will provide further insight into the results obtained, assessing how the actual observed capacity increase relates to the theoretical expected benefits. It would be a way to support the theoretical model or, at least, to identify the gap between the two.

2.2 Reviewer 2

This paper is overall very well written. I have a few comments:

- line 132: according to Skyguide, RET at LSZH is not used to increase capacity. So how about the other airports? Could it be that RET is never used to increase capacity? If you do the same analysis in LSZH and get similar results, then what you did in LPPT would be less convincing wouldn't it?

- what about the other airports? When you say "exclusively Lisbon airport" it is not clear whether we could apply similar approaches to other airports in the Table.

- Section 2. I think the reader would benefit from two sections: one for the methodology, and one for the constitution of the dataset (including preprocessing)

- notebook: please remove the noisy output (warnings and prints during the training with RMSE), but I would encourage you to keep the outputs of plots which do not appear here.

3. Response - round 1

3.1 Response to reviewer 1

In the first part, the authors provide an XGboost method to estimate the remaining duration of the arrival and departure trajectory missing due to the poor low-altitude coverage of ADS-B traces.

The results shown in Figure 3 are very impressive and highly accurate. As reported, the mean square error is 21.15 seconds.

However, it's unclear why two very distinct patterns are observed in the data (Figure 3). What are the characteristics of the flights with the much steeper time-to-fly prediction with respect to the distance from the runway threshold?

Response

The flight with a much slower time-to-fly compared to others is a Helicopter (Type AgustaWestland AW139) which explains the steep curve. To emphasize this fact, we added the following sentence to the caption of Figure 3: "Note: The data points with the considerably steeper slope of the time to fly prediction refers to the trajectory of a helicopter of type AgustaWestland AW139."

Secondly, I understand that the authors train the XGBoost models using only the subset of flights for which their trace is 'complete'; how many flights are these?

Response

A total of 544 flights have been used to train the XGBoost model.

To clarify this point, we added the following sentence in the second paragraph of "Estimation of landing time" in Section 2.1: "In the course of this process, 544 trajectories were identified as being fully covered."

Six features are used for these models: distance to threshold, aircraft type, ground speed, altitude, track angle and rate of climb. The traces of these flights are full to the runway, so how the 'distance to the threshold' is computed should be clarified. Do the authors sample the trace from the ground up to 12 NM?

Response

The distance to threshold refers to the distance between the aircraft and the threshold of runway 20. To train the model, only the flights that are tracked up to at least $d_{LDG} = 0.03NM$ from the runway were used (544 flights). All these 544 flights are re-sampled at 1 second resolution and all the resulting data points of these flights within the 12NM have been used to train the XGBoost model.

We added the following sentences in the second paragraph of part "Estimation of landing time" in Section 2.1: "From these fully covered landings, the time to fly to the threshold of runway 20 is computed within the last 12 NM of the approach. For this purpose, the trajectories are resampled at a resolution of 1 second and the distance between all resulting aircraft positions and the threshold of runway 20 is determined."

It would be interesting to understand the importance of the six features used. The results shown seem very linear, so it would be useful to understand the benefit of these additional features with respect to a simple linear fit as a function of the distance. Computing the difference in prediction between these two models (the ML XGBoost and a simple linear regression) and the error of each model would be useful.

Response

The primary focus of this study was not to develop a machine learning model per se. While we understand the interest in comparing the performance of the XGBoost model with simpler models, e.g., linear regression, such a comparison was not within the scope of our research. We chose gradient boosting techniques, specifically XGBoost, due to their empirically demonstrated effectiveness in various applications. Additionally, XGBoost offers practical advantages such as not requiring feature scaling and the ability to handle categorical variables like aircraft typecodes and missing values effectively.

Finally, these models provide an estimation between 0 and around 300 seconds (5 minutes), but as shown in Figure 2 and Figure 4, the distance to/from the threshold seems to be usually lower than 5 NM and 1 NM for landings and departures, respectively. Therefore, the 'time to fly' estimation would range mostly between 0 and 200 seconds (3 minutes). Considering that the authors will then aggregate the data (arrivals and departures) every 15 minutes to compute the departure/arrival relationship, how important is this machine learning model for this particular problem? How would the results change if, instead of this, a simple linear approximation is used to estimate the time remaining/time before as a function of the distance of the aircraft with respect to the threshold? This might be estimated by recomputing the values used to generate Figure 6 with this simpler approach.

Response

Our data set also contains, among others, categorical values. Subsequently, the application of a simple linear approximation-based model is challenging.

Even if we developed and applied a second model based on an alternative method, we are not sure whether we could answer your question. This is because we do not know the ground truth and therefore do not know conclusively for both the XGboost-based and the alternative model whether the aircraft are assigned to the 'correct' 15-minute windows.

Then, the authors rightly identify the need for 'good' coverage and use that to filter the case study to LPPT from the potential 7 airports which implemented some RET in the time period considered. However, not much detail is provided on the specific steps for this process. Please describe this in more detail.

Response

We added a footnote in Section 2.1 which specifies what we consider as 'good' coverage. The footnote reads as follows: "The coverage of a trajectory is considered 'good' in this study if both the air and the ground part of the trajectories are included in the data."

The second contribution is estimating the benefit of building a rapid exit taxiway. The authors mention that they discarded Zurich as a candidate airport for the analysis as their discussion with experts led them to identify that the rapid taxiway developed there would be used to allow 'smoother' operations. One could argue that that's the impact of all of these infrastructures. Wouldn't smoother operations on delay-prone operations not lead to higher airport usage? How is this different in the Lisbon case? It might be useful to provide a diagram chart of Lisbon indicating the location of the RET to understand its potential usage better.

Response

Regarding “smoother operations”: We adapted the end of the "Selection of airport" paragraph to better explain why we did not choose Zurich and why we chose Lisbon Airport as a practical example in our study.

Regarding your suggestion of adding an airport chart of Lisbon Airport: Figure 1, which contains an OpenStreetMap-based chart of Lisbon Airport, has been added to the manuscript.

Finally, I wonder if there are theoretical studies on the expected benefit of implementing a rapid exit taxiway on the runway occupancy time (ROT). If this is the case, the authors could compute the expected benefit on runway occupancy for Lisbon due to the RET. This will provide further insight into the results obtained, assessing how the actual observed capacity increase relates to the theoretical expected benefits. It would be a way to support the theoretical model or, at least, to identify the gap between the two.

Response

This is a very interesting question. However, we not aware of any theoretical study or model that could be employed to answer it.

The ROT can be regarded as a random variable, which depends on a large number of factors (e.g., aircraft type, weight, weather situation, runway condition, presence of other traffic, available RETs, crew intentions, ATC instructions, etc.). It is therefore difficult to determine on a theoretical basis what influence one or more additional RETs will have on the ROTs realized in practice.

3.2 Response to reviewer 2

Line 132: according to Skyguide, RET at LSZH are not used to increase capacity. So how about the other airports? Could it be that RET are never used to increase capacity. If you do the same analysis in LSZH and get similar results, then what you did in LPPT would be less convincing wouldn't it?

Response

This is an interesting point you are raising. The main idea behind RETs is to reduce the time aircraft spend on the runway, which, in theory, should help increase an airport's capacity. However, the situation at Zurich Airport, which operates with three runways, presents a unique case. According to Skyguide, the number of arrivals and departures at Zurich is limited more by the surrounding airspace rather than aircraft occupying the runway itself.

We don't think that RET are never used to increase capacity but we believe that the effectiveness of RETs in increasing airport capacity could vary significantly based on specific airport characteristics (number of runways, traffic mix, ...). Therefore, it would be indeed interesting to expand our study to include more airports. By doing so, we can assess the impact of RETs in a variety of operational settings. This broader comparative analysis would allow us to gain deeper insights into how different factors, like the number of runways and traffic mix, can influence the efficiency and capacity of airports.

We have adapted the following sentences in Section 2.1 to draw attention to the crossing procedure-based capacity restriction at Zurich Airport: "According to information provided by Skyguide, the capacity of Zurich Airport operated in the runway configuration in which aircraft land on runway 28 is not limited by the maximum throughput of runway 28, but rather by airspace constraints. Indeed, the two RET B7 and L7 newly installed on runway 28 are used to ensure 'smooth' day-to-day operations only."

What about the other airports? When you say “exclusively Lisbon airport” it is not clear whether we could apply similar approaches to other airports in the Table.

Response

The methodology presented in this paper can be applied to other airports as well. Our application of the method to Lisbon Airport is a practical example only. To clarify this point, we have adapted the following sentence in Section 2.1: “...Lisbon Airport is used in this study as a practical example to measure the influence of RET on the capacity of a runway system.”

Section 2. I think the reader would benefit from a two sections: one for the methodology, and one for the constitution of the dataset (including preprocessing)

Response

We reorganized the methods section: It now consists of only two subsections (2.1 Data collection and pre-processing, 2.2 Determination of envelope function).

Notebook: please remove the noisy output (warnings and prints during the training with RMSE), but I would encourage you to keep the outputs of plots which do not appear here.

Response

We adapted the source code of our software accordingly.