Journal of Open Aviation Science (2023), Vol.**1**
doi:10.59490/joas.2023.7199

JOAS

POSTER | *The 11th OpenSky Symposium*

Classification of Holdings in Flights Arriving at Dubai International Airport (DXB) in One Year

Luiz Pradines de Menezes Junior*

Toulouse, France

*Corresponding author: pradines_family@hotmail.com

(This poster paper is not peer reviewed.)

Abstract

The purpose of this study was to use a machine-learning approach for the classification of holdings for all flights arriving at Dubai International Airport (DXB) in a period of one year. The study used data from ADS-B for all flights arriving and departing DXB for the period 15 February 2018 and 15 February 2019. For all 189,999 arrivals analyzed, it was identified the Standard Terminal Arrival Route (STAR) flown, the occurrence of holdings, the aircraft flying (for the determination of wake turbulence category), the visibility, and data related to the interception of a circle centered around DXB with a radius of 50 NM (geographical position, number of aircraft entering the circle for several time windows, number of aircraft flying within this circle etc), along with the number of aircraft taking-off and landing during a given time period. These features were tested, in different combinations, to determine their impact on metrics used to evaluate the classifier output quality. After running classifiers designed with different algorithms and combinations of features, it was identified the one which, with the CatBoost algorithm, gave the best F1 (0.777), Precision (0.847) and Accuracy (0.912) for a period of 10 minutes after interception. The F1 increased to 0.85 when the altitude and the speed at the interception of the circle were included as features, but they were discarded as they introduced an undesirable bias in the final model.

Keywords: ATC; Machine-learning; holdings

Abbreviations: JOAS: Journal of Open Aviation Science, OFP: Operational Flight Plan, EASA: European Aviation Safety Agency, ERA: En-Route Alternate, SCF: Statistical Contingency Fuel, TMA: Terminal Maneuvering Areas, DXB: Dubai International Airport, FIR: Flight Information Region, STAR: Standard Terminal Arrival Route, DWC: Al-Maktoum International Airport, SHJ: Sharjah International Airport, ATC: Air Traffic Control

1. Introduction

Airlines use to uplift additional fuel to cater for holdings in congested airports, which are more frequent during peak arrival times. The additional fuel is included by the Flight Dispatcher in the OFP (Operational Flight Plan) based on airlines' policies. This fuel is added to the minimum required by operational regulations, which includes the Contingency Fuel (according to EASA regulations, a minimum of 5% of trip fuel, or 3% provided an ERA – En-Route Alternate is available). Some airlines calculate their contingency fuel with quantities different from the percentages defined in the operational regulations, in a method called "Statistical Contingency Fuel", also represented by the acronym SCF. As the name suggests, this method relies on the measurement of the contingency fuel burnt by all aircraft of a given model, operating in that city pair [1].

The contingency fuel provides a hedge against any possible deviation from the hypothesis used to generate the OFP. Thus, when used, it may be the result of changes in the forecasted weather (winds and temperatures), the need to avoid regions with turbulences, unaccounted differences in the weight of the passengers plus carry-on baggage, and holdings at the destination. Thus, the contingency fuel used encompasses several factors, and the breakdown is not easy to do. Considering that some airlines, for the sake of conservativism, decide to uplift additional fuel due to the possibility of a holding at the destination, chances are that this fuel can be double counted. This inaccurate fuel prediction leads to unnecessary and costly fuel burns.

Based on this, it is necessary to have a better comprehension of the factors that result in holdings on congested TMA (Terminal Maneuvering Areas) and their duration. In this paper, we propose a machine learning approach to carry out a classification of holdings on all arrival flights at Dubai International Airport (DXB) in a period of one year. The model thus derived can be used not only for flight planning purposes, but also to optimize the flow of aircraft arriving at Dubai TMA, reducing the occurrences of holdings.

2. Method

2.1 Description of the airspace around DXB

The Emirates FIR borders Tehran FIR to the North, Muscat FIR to the East, Doha FIR to the West, and Jeddah FIR to the South. After a restructuring of the Dubai TMA that took place in December 2017 (which introduced some features to optimize the airspace, like the "trombone" represented by the light blue curve), the STARs were grouped around four different entry points: PUVAL to the Northeast, VUTEB/DATOB to the Northwest, LORID to the Southwest and IMPED to the Southeast. These points, along with an arrival trajectory and the corresponding STAR, are shown in Figure 1 (STAR entry points are represented by green dots).

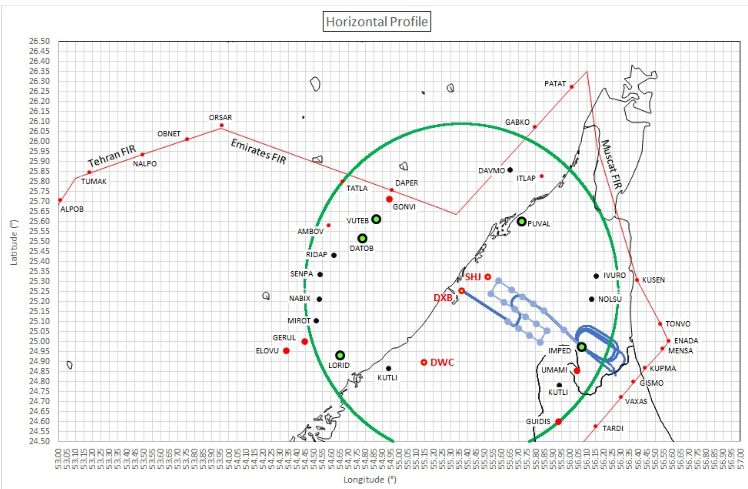

Figure 1. Emirates FIR with borders and a flight trajectory.

The origin of the flight normally dictates the STAR entry point. For example, flights coming from Asia and Australia fly via IMPED, whilst the ones coming from Europe fly via VUTEB and DATOB. Most of the traffic arriving at DXB comes from these two waypoints. In the figure, the landing runway was 30L. Runway 30R is used mostly for takeoffs. The opposite runways are 12L (landings) and 12R (takeoffs). The distance between the runways precludes simultaneous landings, imposing a limitation on the maximum number of operations per hour.

Holdings are mostly confined within Emirates FIR. The published holding fixes are VUTEB, IMPED and ITLAP. They are virtually non-existent South of DXB, and flights coming via LORID, when the airspace is saturated, are sent to ITLAP.

DXB has two airports nearby. Sharjah International (SHJ) is located around 16 km to the Northeast, and Al-Maktoum International (DWC), mostly used for cargo flights, is located around 45 km to the Southwest. The air traffic is dominated by DXB: there are as many as five times more landings in DXB than SHJ, and 55 times more landings in DXB than DWC. Due to the proximity, these three airports constitute what is called a "multiplex" [2].

2.2 Data extraction, cleaning, and features identification

ADS-B files for all flights inbound and outbound to DXB were extracted for the period 15th of February 2018 to 15th of February 2019, amounting to a total of 380329 flights (190001 arrivals and 190328 departures). One year of flights encompasses seasonal variations in weather at the airport. All these trajectories were filtered to remove data repetition and outliers. Data gaps, however, were not filled. After data extraction and cleaning, every file was matched to an aircraft type and tail-number, to allow the identification of the aircraft's wake turbulence category.

After this phase, several features were identified for each individual flight. For example, the landing or takeoff runway, the touchdown or takeoff time, the STAR flown and the visibility at the operation time. A dedicated algorithm was created for the identification of a holding pattern. Due to the variety of trajectories that could be interpreted as a holding, which do not necessarily reflect the traditional racetrack, the algorithm labelled a flight as having a holding pattern if a segment of the horizontal trajectory crossed a previous one. This differed from the strategy for holding identification adopted at [3]. Out of the 190001 arrivals, around 40424 (or 21% of them) had holdings.

It was inferred that the number of aircraft flying within Emirates FIR could be a very important feature for the classification. Moreover, other features associated with the interception of the FIR boundaries, like time, altitude, and speed, could be relevant as well. Because some STAR entry points are more distant from the Emirates FIR borders than others, this asymmetry could introduce irregularities in the model. Thus, it was decided to define a circle with a radius equal to 50 NM centered at DXB (green circle in Figure 1). By doing so, every inbound flight could have a common reference for the definition of relevant features, like the geographical location of the interception point, time, altitude, and speed. The 50 NM radius was chosen since 95% of the holdings took place within this distance.

Lastly, it was considered that the number of aircraft that have entered the circle prior to the interception, during a given time window (say, 10 minutes), could also be a relevant feature. The number of aircraft landing and taking off was also identified. Several time windows were studied, ranging from a minimum of 10 minutes up to 60 minutes, and their influence on the classifier's output quality was assessed as well.

3. Discussions

The results indicate that the number of aircraft flying within the circle is the strongest feature in the classification process. Taking as an example the time window equal to 10 minutes and running the Decision-Tree classifier with features like Landing runway/latitude and longitude at the circle interception/number of aircraft flying within the circle, we managed to obtain an F1 equal to 0.770. The insertion of features like the number of aircraft taking off and landing during the time window/wake turbulence category/number of aircraft entering the circle during the time window raised the F1 to just 0.773. However, when introducing features like altitude and speed at the interception of the circle, the F1 jumped to 0.828. Despite this fact, a close analysis of a graph showing these parameters and the occurrence of holdings suggests they should not be used for the classification.

Table 1 shows, for a time window equal to 10 minutes and for features Landing runway/latitude and longitude at the circle interception/number of aircraft flying within the circle, the output quality metrics with different algorithms for the classifiers.

Table 1. Example table

Model	Precision	Recall	Accuracy	F1
CatBoost [4]	0.847	0.718	0.912	0.777
Random Forest [5]	0.844	0.717	0.912	0.775
Gradient Boosting	0.840	0.718	0.911	0.774
Decision Tree	0.834	0.716	0.909	0.770
K Neighbors	0.819	0.682	0.900	0.744
Multi Layer Perceptron	0.819	0.668	0.898	0.736

4. Conclusion

The number of aircraft flying within the circle of 50 NM radius is the main feature for the classification of holdings in Dubai. The best performance was obtained with a classifier based on the CatBoost algorithm (F1 equal to 0.777). Curiously, a classifier based on neural networks (Multi Layer Perceptron) showed the worst performance, which suggests that further refinements in the hyper-parameters for the classification are needed. The inclusion of features like visibility, the number of aircraft entering the circle, number of aircraft taking off and landing, all during a given period, improved the F1 just slightly. This increases unnecessarily the complexity of the model without tangible benefits.

Parameters like altitude and speed, at the interception of the circle, increased the F1 by around 6%. However, the altitude is already an indication that a holding is being acted upon by ATC due to the concentration of aircraft within the circle, thus rendering this parameter invalid for the purpose of classification.

Future investigations will be conducted based on the redefinition of circles encompassing the holdings around each one of the three usual waypoints (VUTEB, ITLAP and IMPED). Additional features associated with the number of aircraft expected to arrive in each time period (after the interception of the circle), along with the possible effect of adjacent traffic to other airports close to DXB (like SHJ and DWC), will be investigated as well.

Acknowledgement

The author acknowledges the fundamental contributions from Cédric Campguilhem and Guido Knigge.

Open data statement

The original data can be downloaded from:
https://github.com/LPradines/11th_Open_Sky_Symposium

Reproducibility statement

The Jupyter Notebooks can also be downloaded from:
https://github.com/LPradines/11th_Open_Sky_Symposium

Instructions to run the analysis are described in the *README.md* file.

References

[1] International Civil Aviation Organization. *Flight Planning and Fuel Management (FPFM) Manual (9976)*. 1rd. DOC-09976-001-01. International Civil Aviation Organization. 2015.

[2] R.L. de Neufville, A.R. Odoni, P. Belobaba, and T.G. Reynolds. *Airport Systems, Second Edition: Planning, Design and Management*. McGraw Hill LLC, 2013. ISBN: 9780071770590. URL: https://books.google.fr/books?id=gfHnA9W5sJkC.

[3] Imen Dhief, Zhi Jun Lim, Sim Kuan Goh, Duc-Thinh Pham, Sameer Alam, and Michael Schultz. "Speed control strategies for e-aman using holding detection-delay prediction model". In: *Proc. 10th EUROCONTROL SESAR Innovation Days*. 2020, pp. 1–10.

[4] Anna Veronika Dorogush, Vasily Ershov, and Andrey Gulin. "CatBoost: gradient boosting with categorical features support". In: *arXiv preprint arXiv:1810.11363* (2018).

[5] Tin Kam Ho. "Random decision forests". In: *Proceedings of 3rd international conference on document analysis and recognition*. Vol. 1. IEEE. 1995, pp. 278–282.