IOP Conf. Series: Materials Science and Engineering 1024 (2021) 012106 doi:10.1088/1757-899X/1024/1/012106

Applying Machine Learning Modeling to Enhance Runway Throughput at A Big European Airport

G Stempfel[1], I De Visscher[2], M Ellejmi[3], V Brossard[1], A Bonnefoy[1] and V Treve[4]

[1] EURANOVA, Marseille, France

[2] Wake Prediction Technologies (WaPT), Louvain-la-Neuve, Belgium

[3] EUROCONTROL, Brétigny-sur-Orge, France

[4] EUROCONTROL, Brussels, Belgium

guillaume.stempfel@euranova.eu, ivan.devisscher@wapt.be

Abstract. One of the factors limiting busiest airport's runway throughput capacity is the spacing to be applied between landing aircraft to ensure that the runway is vacated when the follower aircraft reaches the runway threshold. Today, because the Controller is not able to always anticipate the runway occupancy time (ROT) of the leader aircraft, significant spacing buffers are added to the minimum required spacing in order to cover all possible cases, which negatively affects the resulting arrival throughput. The present paper shows how a Machine Learning (ML) analysis can support the development of accurate, yet operational, models for ROT prediction depending on all impact parameters. Based on Gradient Boosting Regressors, those ML models make use of flight information (such as aircraft type, airline, flight data) and weather information to model the ROT. This paper shows how it can be used operationally to increase runway capacity while maintaining or reducing the risk of delivery of separations below runway occupancy time. The methodology and related benefits are assessed using three years of field measurements gathered at Zurich airport.

1. Introduction

The traffic growth in the recent years put significant pressure on the European Airport's capacity [1]. In today's efficient operations, landing aircraft are separated either by wake turbulence separation rules or by runway and surveillance separation rules that apply when wake turbulence separation minima are not required. For arrivals on final approach, the wake turbulence and surveillance separation requirements are expressed by distance minima to be applied at a separation delivery point which is usually defined as the runway threshold.

The SESAR2020 [2] Project 02 EnhAnced Runway Throughput (EARTH) aims to enhance airport runway throughput addressing situations of over-demand on capacity constrained airports related to separation minima, considering constraints such as weather, runway configuration and traffic mix. The final objective is to optimise traffic throughput with existing infrastructure, and to improve safety and environment.

When reducing wake separation minima (through the use of, e.g., Time-Based Separation, Weather Dependent Separation) or surveillance minima (based on Required Surveillance Performance (RSP) [3]), the most limiting constraint might become the time required for the leader aircraft to vacate the runway, i.e. the Runway Occupancy Time (ROT). Delivering time separations below the required ROT could lead to a missed approach procedure or even to a loss of runway separation if not detected by the controller. However, the ROT depends on several factors such as aircraft type, runway configuration, runway status, landing profile or weather conditions, making it difficult for the Controller to anticipate. Without a model able to accurately predict the ROT for each landing aircraft, significant buffers have to be taken into account so as to cover all possible cases. The objective of the present paper is to show how a Machine Learning analysis can support the development of an accurate, yet operational, model for ROT prediction depending on all impact parameters. This model can then directly be used to provide the required ROT spacing to the Controller using a separation delivery tool (simple or more advanced such as the ORD tool developed in SESAR2020 PJ02 [4]). An accurate prediction of the required ROT allows a separation delivery tool to provide, on the Controller Working Position (CWP), an optimised spacing distance to be applied at threshold between a leader and a follower aircraft in order to ensure that the runway will be vacated when the follower will overfly the runway threshold.

In previous work by one of the authors and the University of TU Delft, a detailed literature was performed on runway capacity enhancements and methods to predict Arrival ROT [5]. In that latter paper a ROT prediction using ML that allows to predict the exit of the runway was described. In the present paper the ML approach is used to provide several ROT predictors with various prediction time horizons useful at different stages of the approach traffic control.

The methodology has been tested using 3 years of field measurement data gathered at Zurich airport. This paper also shows the benefits related to the use of a Machine Learning model compared to those obtained using a simpler data mining approach.

2. ROT definition

The ROT is defined as the difference between the time the aircraft is observed to overfly the runway threshold and the time at which it has vacated the runway. The aircraft is considered to have vacated the runway if either it is on the runway exit or it is still on the runway but sufficiently far from the threshold. According to ICAO doc 4444, this minimum distance is 2400 m. It should be noted that the definition of the reference distance of the aircraft on the runway exit for runway vacation varies from one airport to another. Note also that some airports do not consider the 2400 m distance criterion prescribed by ICAO or use a different locally-defined distance value.

The definition of the ROT obviously impacts the absolute numbers obtained from data analysis. Yet, when comparing different prediction approaches, the definition does not affect their relative assessment.

3. Database description and processing

To support the investigation of Runway Occupancy Time, a database covering 3 years of operations at Zurich airport (LSZH) has been used. It contains the surveillance radar tracks of aircraft flights landing at LSZH airport between 2015 and 2017. The radar sampling rate is 4 seconds when the aircraft is on the glide (using approach radar data) and is reduced to 1 second when in ground proximity (using A-SMGCS data). The METAR met data were also made available providing wind speed and direction, wind gust, visibility, temperature, dew point and pressure with a refresh rate of 30 minutes. Finally, the corresponding runway surface wind speed and direction measurements were also available for each runway threshold with a sampling rate of 3 s.

In Zurich airport, the main arrival runway is runway 14 followed by runway 28 and runway 34. Runway 14 significantly differs from the two other main arrival runways by the locations and types of its exits. Indeed, for runway 14, the first runway exit is located at more than 2200 m from the runway threshold whereas the second and third are located further than 2400 m away from the threshold. They are all high-speed runway exits. On the contrary, on runways 28 and 34, several exits are distributed along the runway length. Moreover, for the period of operation considered in the database (i.e. 2015-2017), there were no high-speed runway exits on those runways. For runway 14, we can thus anticipate that the ROT will be mainly driven by the time required for the aircraft to roll down to the first runway exit as most of the aircraft will be able to vacate at the first runway exit or alternatively will quickly reach the 2400 m threshold distance. On the contrary, for runways 28 and 34, the ROT will be mainly driven by the runway exit used by the aircraft and thus by the aircraft brake capacity.

The main challenge in the data pre-processing relies in correctly detecting the entrance and exit times from RADAR records for each flight. Due to missing or inconsistent data, slightly over 85% of the database is suitable for performing this task.

Since ROT is a limiting factor only when an aircraft is closely followed by another aircraft, we focus on peak operations periods. Out of these periods, the aircraft may have inefficient behavior due to low runway releasing time constraints; this could bias the models towards non-interesting cases in our context. Then, we filtered out all flights for which the time separation at runway threshold with their follower is larger than a limit value. That threshold has been fixed at 180 seconds for Heavy and Super aircraft types, and at 120 seconds for all other categories. After having applied this rule on the full dataset, we observe that 60% of the flights can be considered as constrained. We have also discarded a few dozens of flights for which METAR raw data are incomplete, or for which departure airport is not defined. Finally, helicopters are not considered in our study, neither are a few dozens of flights having landed on runway 16. The final dataset is then composed of slightly more than 233,000 flights.

4. Machine Learning modelling

The present study considers Machine Learning (ML) ROT predictors with various prediction time horizons. Since the positioning of the aircraft on the glide must be performed by the Air Traffic Controller (ATCO) several minutes before its landing, the first predictor considers a 10-minutes time horizon (i.e. only data available 10 minutes before the landing are used as model features). Nevertheless, later on, it is still interesting for the ATCO to know if the ROT estimation of the predictor can be improved by using updated data, or even data which were not available in the "10-minutes" setup. To do so, we here also investigate three additional setups at three shorter time horizons.

4.1. ROT Prediction at a 10-minutes horizon

In this standard setup, we focus on the ROT prediction using only features available 10 minutes before the actual landing time, extracted from RADAR, METAR and anemometer recordings:

- the aircraft type and RECAT category
- the landing runway, landing week, weekday and hour
- the departure airport, its ICAO area and its country
- the airline code, plus a low-cost airline flag
- surface wind data (computed from runway anemometers recordings):
  - head- and cross-wind speeds averaged over 15 minutes and over 3 minutes
  - head- and cross-wind speed variances over 15 minutes
- METAR information (recorded between 10 and 40 minutes before the landing time):
  - pressure in hPa, visibility in meters, temperature and dew point in °C
  - weather flags: rain, brume, shower, fog, snow, storm, drizzle, haze, convective clouds
  - weather events intensity (strong or weak)
  - ceiling clouds altitude

All numerical data, but landing week, day and hour, are standardized. All categorical data (except Boolean ones) are impact coded. Impact coding consists in replacing an instance categorical feature with a blend of posterior probability of the target given the instance feature value and the prior probability of the target over all the training data. Categorical data can then be processed as continuous data.

The problem is modelled as a classic regression problem, where the prediction target is the expected value of the ROT. To avoid introducing bias in the model ability evaluation, the full dataset is split in two subsets, one learning set (representing two third of the full database) and one testing set (the remaining third).

The improvement brought by the machine learning algorithm will be assessed by confronting the results to a baseline. The baseline decision is defined as the average ROT observed by type and by runway. The weather conditions are not taken in account for establishing this baseline.

Selecting the most appropriate algorithm is performed by using a 5-folds cross-validation process, among a set of models including decision trees, random forest regressors, gradient boosting regressors [6], support vector regressors and XGBoost [7]. For all these algorithms, extensive sets of parametrizations have been tested. By basing the selection on the higher R2-score (a classic indicator when coming to regression), the process finally selected XGBoost, parametrized with a maximal number of 5000 estimators, with a maximum depth of 8. Note that XGBoost can be ran on GPU, which makes the learning process very cheap in terms of computation time, roughly three minutes for three years of historical data on a GeForce GTX 1070.

To assess the ML model performance, we here first compare the R2-score to the baseline. Based on the full test dataset, the use of XG-Boost leads to a R2-score of 0.516 whereas the baseline R2-score is 0.352, showing that the ML model manages to exploit information from input features since it allows an improvement of the baseline R2-score of almost 50%.

Since we know (see Section 3.) that the three main runways have very dissimilar configurations, we are interested in estimating the improvements brought by the ML model for each runway independently (see Table 1). Both baseline and XG-Boost model perform way better on runway 14 which seems easier to predict than the other ones. This could be the consequence of a much larger amount of available data, but also to the very specific configuration of this runway. Indeed, on runway 14, an aircraft must travel around 2300m before being able to reach an exit. The ROT is then mainly linked to landing speed on this runway. The number of exit possibilities is much more important on the other runways; the exit choice, and then the ROT, is related to brake ability and to the pilot's own decision. This makes the runway 14 problem sensibly easier than the other ones.

Table 1: R2-scores of ML model and baseline per runway on the full test dataset

Runway\Model	XG-Boost	Baseline	# Data
14	0.518	0.300	54504
28	0.222	0.146	11161
34	0.340	0.207	4546

Gradient Boosting Regressor algorithm makes use of a large set of decision trees. It is interesting to assess the feature importance of the results for ROT prediction. This was computed using SHAP library [13] and is displayed in Figure 1.

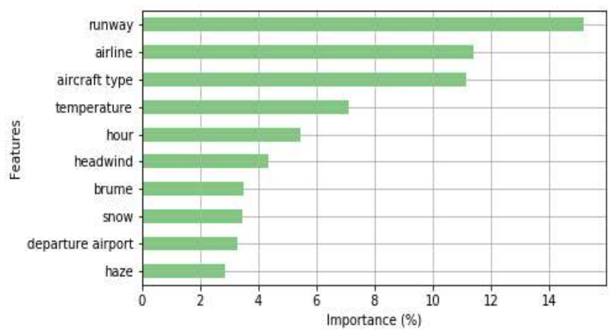

Figure 1: Features importance in the ML model decision

There are four major important features. First one is the runway, as it accounts for 15.20% of the overall importance. As stated in the previous sections, runway configurations are disparate. This variety is well captured by the model, since the largest importance is attributed to this feature. Second in terms of importance, the airline accounts for 11.40% of the importance. Third and fourth are the aircraft type and the temperature (11.16% and 7.10% respectively). The influence of the temperature models a decrease of ROT when the temperature increases. Data exploration exhibits an average difference of 5 seconds on runway 14 between cases where the temperature is below 0 and cases where the temperature is higher than 20 °C. We also note the presence of weather features such as headwind, snow, rain and haze.

A feature that is absent but that could have been expected is RECAT aircraft wake turbulence category. Instead, the model seems to rely on the type of the aircraft. It may mean that there is a too wide behaviour spectrum in each RECAT category, and then that relying on the aircraft type is much more accurate.

4.2. Shorter-terms ROT decisions

Three other set-ups of ROT prediction are also explored at shorter term than ten minutes, to assess whether additional information gathered in the interval helps in improving the prediction. The three horizons are defined as the time when an aircraft arrives at 8NM from runway threshold, at 2NM from runway threshold, and at runway threshold.

All the features described in Section 4.1. are used but are updated if fresher measurements are available (this concerns wind and METAR data). The encoding and standardization strategies still hold. Additionally, we take in account the following information:

- aircraft speed (only for 2NM and threshold setups, since the 8NM from threshold is located before the usual aircraft deceleration fix)
- presence of an aircraft on the runway (2NM only)
- time elapsed since leader landing (2NM only)

Note that the amount of available data slightly varies (by a few hundreds of flights) among set-ups, due to some missing data in the raw recordings. These three problems are modelled as regressions using the same learning algorithm as in Section 4.1.

Table 2 compares the R2-score of the four predictors. It shows that the 8NM setup does not allow any improvement in the prediction. The weather conditions update does not seem to bring any exploitable information. Note that the wind is usually relatively low in Zürich, and the conclusion may be different for an airport facing frequent strong wind conditions. On the other hand, we observe a net increase of R2-score on the 2NM and threshold prediction set-ups.

Table 2: R2-scores of ML models in the 4 set-ups, computed on the test datasets

Setup	R^2 Score
10 Minutes	0.524
8 Nautical Miles	0.528
2 Nautical Miles	0.554
Threshold	0.582

The feature importance is here assessed for the 2NM and threshold setups, since prediction qualities at 8NM and 10 minutes horizon are of the same order of magnitude.

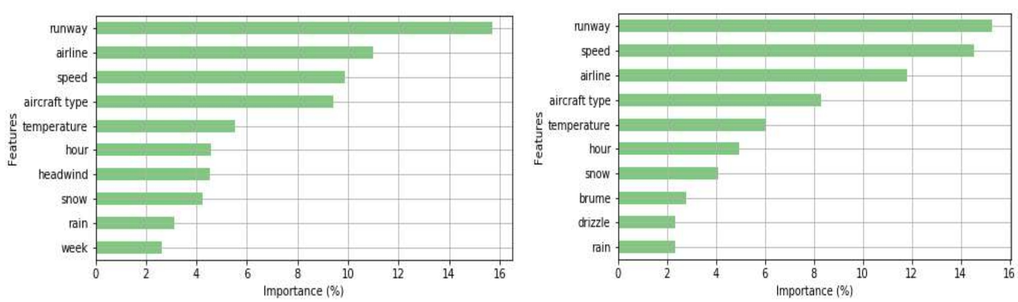

Figure 2: Features importance in the ML model decision in 2NM set-up (left) and in the threshold set-up (right)

In Figure 2, we observe that the features importance is roughly the same than in 10-minutes set-up. The main difference lies in the presence of the speed. The knowledge of the ground speed is then exploited by the model and explains the improvement of the R2-score in these setups. Additional experiments show that, in the threshold setup, the use of ground speed allows to increase the R2-score by more than 12%. Note that its importance is slightly bigger in the threshold setup than in 2NM. This can be explained by the fact that in the threshold set-up, the current speed is the actual final ground speed whereas this is only an estimation of it at 2NM. On the other hand, we observed that the leader position does not bring any improvement to the model R2-score.

5. Application: Definition of ROT-based MRS

In today's regulation, Minimum Radar Separation (MRS) is strongly related to ROT. Indeed, in ICAO PANS ATM doc 4444, the first condition to allow for reduced MRS from 3.0 NM to 2.5 NM between succeeding aircraft which are established on the same final approach track within 10 NM of the runway threshold is that "the average runway occupancy time of landing aircraft is proven, by means such as data collection and statistical analysis and methods based on a theoretical model, not to exceed 50 seconds". This definition establishes a direct link between applicable MRS and ROT observations. It accounts for the variabilities, of the follower time-to-fly 2.5 NM and of the runway occupancy time around the 50 s average, in the reference speed (180 kts, which is far above the final landing speed of commercial aircraft) used to convert the reference distance (2.5NM) into a reference time (50 s).

Following the same logic as for MRS=2.5 NM, the applicable ROT-based MRS (i.e. applicable only provided that the RSP are met) can then be obtained per aircraft type based on its average ROT, measured locally in peak conditions, and multiplied by 180kts. For example, if the average ROT for a given aircraft is 40s the ROT-based MRS will be 2.0 NM, if it is 55 s, the applicable ROT-based MRS will be 2.75 NM. The characterization of the average ROT shall be performed per aircraft type and per runway direction (i.e. QFU) and for a period covering at least one year of operation in order to account for seasonal effect. This approach is thus based on a "naïve" predictor (i.e. the average ROT) and is denoted naïve approach in what follows.

When using a ML-based predictor, a similar logic can be applied but using the ML ROT predictor in place of the naïve ROT estimator. The ROT-based MRS is computed for each aircraft pair using the leader ROT prediction and a reference speed. Since the ROT ML prediction is more accurate than the naïve predictor, if using the same reference speed as for the naïve approach, the risk of provision of separation delivery below ROT is decreased. Alternatively, one can also reduce the reference speed (e.g. to 175 kts) while maintaining the same level of risk compared to the naïve approach, which allows further separation reductions and hence runway capacity.

Operationally, the ROT-based MRS reductions can be applied with a simple separation delivery tool, providing, for MRS pairs only, one static chevron corresponding to the applicable distance-based ROT-based MRS for the considered leader.

Using the LSZH database the application of naïve compared to ML ROT prediction in an operational environment is assessed. The naïve approach ROT predictor corresponds to the mean ROT value computed per aircraft type and per runway QFU from the learning dataset observations. The ML approach uses the 10 minutes ML ROT predictor described in Section 4.1. and calibrated through ML using the same learning dataset.

The assessment of the two methods is performed considering only Medium-Medium and Medium-Heavy aircraft pairs of the test dataset, since no wake separation applies for those pairs. For each leader aircraft flight, the time separation obtained if delivering exactly the predicted MRS_{naive} or MRS_{ML} is computed. The time separations T_{sep} corresponding to the MRS_{naive} and MRS_{ML} are computed using the measured time-to-fly profiles (i.e. the observed time required to travel the computed MRS distances). The MRS value only depends on the leader aircraft whereas the obtained time separation depends on the MRS value and the follower speed profile. The analysis has thus to be performed based on a database of aircraft pairs. For this analysis, in order to increase the pair data sample, artificial aircraft pairs are built assuming that each aircraft flight could have been the leader (i.e. preceding aircraft) of any other aircraft flight landing within +/- 5 min. This approach allows us to increase the number of pairs in the sample (and hence to better cover rare cases) while still combining pairs observed in similar conditions. Finally, for each pair, T_{sep} is compared to the actual ROT observed for the leader flight. The results are quantified in terms of: (i) error rate defined as the fraction of pairs with separation delivery below the actual observed ROT; (ii) mean time separation.

Four MRS and ROT-based MRS values are tested: (i)MRS=3 NM as baseline corresponding to what is applied today at Zurich airport (ii) MRS_{naive} computed using a reference speed of 180 kts and the mean ROT per aircraft type and per runway (iii) $MRS_{ML,180kts}$ computed using a reference speed of 180 kts and the ROT as predicted by the ML model; and (iv) $MRS_{ML,175kts}$ computed using a reference speed of 175 kts and the ROT as predicted by the ML model. The last ROT-based MRS definition benefits from the higher accuracy and ability of the ML model to detect outliers allowing a reduction of the over-conservatism introduced when considering a 180 kts reference speed.

Table 3: Error rate and mean time separation buffer compared to ROT for various MRS definitions

MRS	Error rate			Buffer		
	RWY14	RWY28	RWY34	RWY14	RWY28	RWY34
3 NM	0.10%	0.20%	0.10%	27.8 s	23.9 s	27.9 s
Naïve	0.90%	1.10%	0.90%	16.7 s	14.3 s	16.7 s
ML 180 kts	0.60%	0.60%	0.60%	16.7 s	14.1 s	16.9 s
ML 175 kts	1.00%	0.90%	1.00%	14.8 s	12.3 s	15.1 s

The error rates for the three main arrival runways are provided in Table 3 also detailing the mean time separation buffer compared to actual ROT when using the various MRS definitions. Note that, the error rate value (of about 1% for the naïve approach and the ML predictor with 175 kts) is obtained assuming that all pairs are delivered at minima and without accounting for any ATC action. Most of the pairs will indeed be delivered with a separation buffer and if adding only 5 s buffer (i.e. about 0.2NM) the error rate falls to about 0.25%.

Note also that the obtained separation buffer of about 16s corresponds to more than two standard deviations of ROT for a given aircraft type on a given runway. It also corresponds to the margin compared to a ROT of 50 s obtained if a follower is flying 2.5 NM at a typical approach speed of 135 kts. It is thus similar to what would be obtained if applying 2.5 NM with an average ROT of 50 s, as allowed by ICAO PANS ATM doc 4444.

For the considered pairs and for the three main arrival runways, the use of $MRS_{ML,180kts}$ and MRS_{naive} are seen to lead to equivalent mean time separation, both solutions providing significant benefits (10-11 s) compared to the MRS=3NM baseline. However, the ML predictor is observed to lead to a lower error rate. Given this lower error rate, some of the conservatism introduced by the use of a high 180 kts speed is removed by using 175 kts as reference speed. With 175 kts, the ML approach leads to an error rate equivalent to that of the naïve approach results. However, this definition allows a reduction of the mean delivered time separation by about 2 s. Note also that even with a 175 kts reference speed, the ML predictor leads to slightly less pairs delivered 5 s below the observed ROT. It thus still better mitigates the extreme cases compared to the naïve approach with 180 kts reference speed.

ML approach thus enables to further reduce the MRS compared to the naïve approach by considering a lower reference speed (e.g. 175 instead of 180 kts) while maintaining the same under-spacing rate leading to benefits reaching 13 s on average compared to an MRS=3 NM baseline. It is interesting to note that this conclusion is observed for all investigated runways.

6. Conclusion

The present paper shows how a Machine Learning (ML) analysis can support the development of accurate, yet operational, models for runway occupancy time (ROT) prediction depending on all impact parameters. The method has been developed and assessed based on three years of field measurements gathered at Zurich airport.

Based on a Gradient Boosting Regressor algorithm, the ML model is able to predict the ROT using only data available 10 minutes before the landing. The prediction is then exploitable by the ATC to position the aircraft on the glide. Experiments show that, by using both flight and weather conditions information, the model outperforms substantially an average-based baseline.

A second set of ML models has been proposed, with shorter-terms prediction, using updated and additional information. Experiments show that the knowledge of the ground speed in the 2 NM and threshold set-ups allows to refine the 10-minutes horizon prediction.

The paper also presented a methodology to convert the 10-minutes ML ROT predictors in distance indicators to be used in a simple approach ATC tool in order to allow safe ROT-based spacing reductions. The delivery tool shall then provide, for MRS pairs only, one static chevron corresponding to this calculated distance indicator. Compared to the application of fixed MRS=3NM, the use of reduced ROT-based MRS defined using a naïve approach is seen to increase capacity by more than 10%.

The use of a ML ROT predictor rather than the naïve one is seen either to significantly increase safety as leading to a reduced risk of provision of separation below actual ROT or to further increase capacity by 2-3% while still slightly increasing safety (as it provides better prediction of extremely large ROT cases).

Finally, note that the proposed ML ROT predictors can also be used in a more advanced separation delivery tool (such as the Optimised Runway Delivery tool) in which different ROT time-based predictors are used in combination with follower time-to-fly modelling in order to provide to the controller optimised ROT spacing and warning in case extreme large ROT values are expected.

Acknowledgement
The authors want to thank all people who participated to the elaboration of the operational concepts and solution 8 of the project SESAR2020 EARTH and in particular all the Zurich team: Mattes Kettner from Zurich airport, Christian Wyler, Fabian Hummel, Tani Rieben, Bernhard Daenzer, Bernhard Blümel and Thomas Muhl from Skyguide.

References

[1] EUROCONTROL. Seven-Year Forecast – Edition: 19/01/15/01- 21/02/2019

[2] SESAR Joint Undertaking, http://www.sesarju.eu.

[3] Ellejmi, M.; Graham R.; Treve, V.; Toussaint, J. & De Visscher, I.; Technical Feasibility and impacts of reducing standard separation minima in final approach, in *18th Integrated Communications Navigation and Surveillance (ICNS) Conference,* Herndon, VA, USA, April 10-12**,** 2018

[4] Cappellazzo, V.; Treve, V.; De Visscher, I. & Chalon, C.; Design Principles for a Separation Support Tool Allowing Optimized Runway Delivery, in *Aviation Technology, Integration, and Operations Conference, AIAA AVIATION Forum*, (AIAA 2018-4237**),** 2018

[5] Chalon, C; Ellejmi, M; Herrema, F; Curran, R; Validation of the Runway Utilisation concept, a case study for Vienna airport, SID2019, December 2019

[6] J. H. Friedman, "Stochastic gradient boosting," Computational Statistics & Data Analysis - Nonlinear methods and data mining, vol. 38, no. 4, pp. 367 - 378, 28 February 2002.

[7] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), pp. 785-794.

[8] I. De Visscher, G. Stempfel, F. Rooseleer and V. Treve, "Data mining and Machine Learning techniques supporting Time-Based Separation concept deployment," 37th Digital Avionics Systems Conference (DASC), pp. 594-603, 23-27 September 2018.

[9] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal and S.-I. Lee, "Explainable AI for Trees: From Local Explanations to Global Understanding," arXiv 1905.04610, 2019.

[10] J. H. Friedman, "Stochastic gradient boosting," Computational Statistics & Data Analysis - Nonlinear methods and data mining, vol. 38, no. 4, pp. 367 - 378, 28 February 2002.

[11] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), pp. 785-794.

[12] I. De Visscher, G. Stempfel, F. Rooseleer and V. Treve, "Data mining and Machine Learning techniques supporting Time-Based Separation concept deployment," 37th Digital Avionics Systems Conference (DASC), pp. 594-603, 23-27 September 2018.

[13] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal and S.-I. Lee, "Explainable AI for Trees: From Local Explanations to Global Understanding," arXiv 1905.04610, 2019.