Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo, and M. Loper, eds.

PREDICTING RUNWAY CONFIGURATION TRANSITION TIMINGS USING MACHINE LEARNING METHODS

Max En Cheng Lau
Andy Jun Guang Lam
Sameer Alam

Saab-NTU Joint Research Lab
School of Mechanical and Aerospace Engineering
Nanyang Technological University
Singapore, 639798, SINGAPORE

ABSTRACT

Runway configuration change is one of the major factors effecting runway capacity. The transition-time required to change from one runway configuration to another is a key concern in optimising runway configuration. This study formulates prediction of runway transition timings as machine learning regression problem by using an ensemble of regressors which provides continuous estimates using flight trajectories, meteorological data, current and past runway configurations and active STAR routes. The data consolidation and feature engineering convert heterogeneous sources of data and includes a clustering-based prediction of arrival runways on with an 89.9% validity rate. The proposed model is applied on PHL airport with 4 runways and 23 possible configurations. The 6 major runways configuration changes modelled using Random Forest Regressor achieved R^2 scores of at least 0.8 and median RMSE of 18.8 minutes, highlighting the predictive power of Machine Learning approach, for informed decision-making in runway configuration change management.

1 INTRODUCTION

The runway system at an airport is one of the most important resource at an airport and is often deemed as the bottleneck in airport capacity (Idris et al. 1998). To maximise the capacity of the runway system, proper management of the runways is required. One important field of interest in runway management is the selection of the optimal runway configuration. Runway configuration refers to the combination of runways used for arrivals and departures at an airport. The active runway configuration is updated based on dynamic conditions such as prevailing traffic and weather conditions. The optimal runway configuration is the combination which maximises the capacity of the runway system while not violating any safety requirements or any other policies (for e.g. environmental). These conditions increases the complexity of the decision-making process of the air traffic control officer (ATCO), where the ATCO has to decide the most appropriate runway configuration to use under the given operational conditions (Yin et al. 2020).

The decision-making process in a runway configuration change is associated with many factors such as international regulations, prevailing wind and traffic conditions, degree of delays imposed as well as subjective and administrative factors such as workload and manpower allocation respectively. One central factor which is correlated with such factors is the transition-time required for a change between two runway configurations. This transition-time is the time required to fully complete the runway configuration change and resume operations at expected capacity, measured by the difference between the time of initiation and time of completion of the runway configuration change. For every runway configuration change with unique conditions, the transition-time is a key variable that needs to be weighed in the decision-making

978-1-6654-3311-2/21/$31.00 ©2021 IEEE

process as it has a potential ripple effect given the complexity of the terminal airspace. Therefore, it is imperative that this information be quantified as part of the decision support system for the ATCO before initiation of runway configuration change, rather than just relying on experience.

This study aims to formulate a machine learning model to predict the transition-time required for the runway configuration change process, considering environmental and operational factors, based on historical data. Innovative feature engineering methods are employed to transform the raw data from heterogeneous sources into useful features that are utilized by the model to predict the runway configuration transition times. The model is then trained and validated on Philadelphia International Airport (PHL), which includes flight information, meteorological and runway configuration data for the months of December 2019 and January 2020. These predicted transition-times represent the expected and average time period for the runway configuration change to be completed based upon historical data patterns and dynamic conditions experienced at the airport. However, this transition time should not always be considered as the most efficient or minimum timings for a runway configuration change to be accomplished. Regardless, this information aids the PHL ATCO in making better decisions in the runway configuration change by gauging the average efficiency in terms of transition-time required. This will allow the runway and airport system to operate with optimal capacity with the least disruptions, regardless of the configuration decision.

2 BACKGROUND

Recent research have formulated various models to predict the runway configurations based on common dynamic conditions.

Descriptive models instead used data-heavy scenarios in order to determine the optimal runway configuration, using historical data and case studies. Examples include a discrete-choice model (Ramanujam and Balakrishnan 2011). The discrete choice model utilize factors that include wind speed, wind direction, demand, current meteorological conditions, noise abatement procedures and switch proximity to determine the runway configuration. Another example is a Multi-layer Artificial Neural Network (Ahmed et al. 2018) that predicts the runway configurations using traffic demand, wind conditions and other weather conditions. These descriptive models predict the runway configurations based on past decisions made by ATCOs and the decisions made by them might not be the most optimal. However, these descriptive models are useful to understand the factors that influence runway configuration selection.

Prescriptive models place focus on optimization, using constraints and demand-capacity balancing to select the ideal runway configuration. These models include a Dynamic Control Model (Jacquillat et al. 2017), Mixed Integer Programming (Bertsimas et al. 2011), Min-Max Regret approach (Ng et al. 2017) which utilizes most of the factors mentioned in the descriptive models above and recommend a runway configuration based upon the operational constraints.

A prescriptive model by Duarte et al. (2010) introduced the idea of a transition penalty matrix to indicate the relative cost to throughput capacity during a switch in runway configurations, presenting a more realistic scenario where different configuration changes results in different effects. The transition penalty is formulated based on factors such as repositioning of taxiing and inbound aircraft and movement of equipment and persons through an approximate effective percentage of inactivity. The varying transition-time for each runway configuration change is an important distinction to make as airports frequently cycle through multiple configurations depending on demand and other factors.

These prescriptive models provide robust predictors in many dynamic conditions. However, the heavy focus on the optimisation and predictions of the runway configurations also highlight a research gap into transition-time. The prescriptive optimisation models often present a simplifying assumption or baseline levels regarding configuration change timings, such as a constant time, in the stochastic model (Li and Clarke 2010) and Mixed Integer Programming (Bertsimas et al. 2011).

The determination of transition timings is also dependent on the workload of the ATCO. In dynamic situations, ATCOs commonly rely on their past experiences for runway configuration change decisions which might not be always optimal. Runway configuration change planning can happen strategically and can be hours in advance of the actual cause of change like poor wind conditions. These cannot be accurately decided without an understanding of transition timings, which can potentially limit the runway capacity.

This research thus proposes a machine learning model that could predict the transition-time required for an airport, with multiple runways, to make specific runway configuration changes while considering environmental and operational factors. The predictions of transition times between runway configurations are never predicted in any of the literature and never used in the prediction of runway configurations. Hence, the contributions of this work can further improve on future runway configuration prediction models.

3 PROBLEM FORMULATION

The problem is formulated as a machine learning regression problem, where the prediction model provides an estimate of transition times between two runway configurations for 6 unique runway configuration changes. Figure 1 shows a concept diagram of the proposed framework which illustrates the flow of data, learning model and prediction.

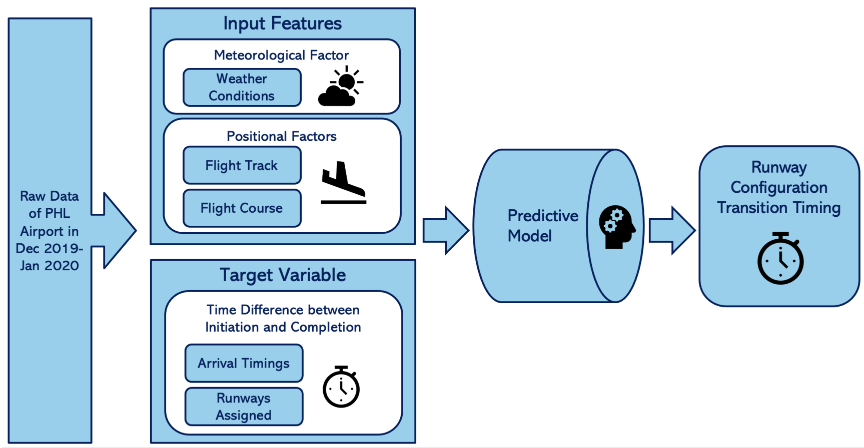

Figure 1: Proposed framework illustrating the flow of data, learning model, prediction of runway configuration transition-time.

As illustrated in figure 1, there are two aspects required for the machine learning model, input features and the target variable of transition-time. The input features refer to the positional factors like flight direction and location during transition and landing, and meteorological factors, including wind and cloud information. The target variable, transition-time, is the time difference between initiation and completion of a runway configuration change. The time when a runway configuration change is realized is provided in the airport data. Since the time of initiation of runway configuration change is not available, arrival timings of flights which have a change of arrival runway is used as a proxy. When such flights are landing on these changed runway, it signifies that the transition has begun. Hence the arrival timings for these flights are used as approximations to the time of initiation. The arrival runway used by each flight, which is not provided in the airport data, is first determined in order to identify flights landed on the changed runway. After which, the complete set of input features and target variable data is used to develop the model to predict transition timings for the 6 major runway configuration changes at PHL airport.

4 METHODOLOGY

The methodology of the study is shown in Figure 2. The specific data processing steps are sequentially explained in this chapter.

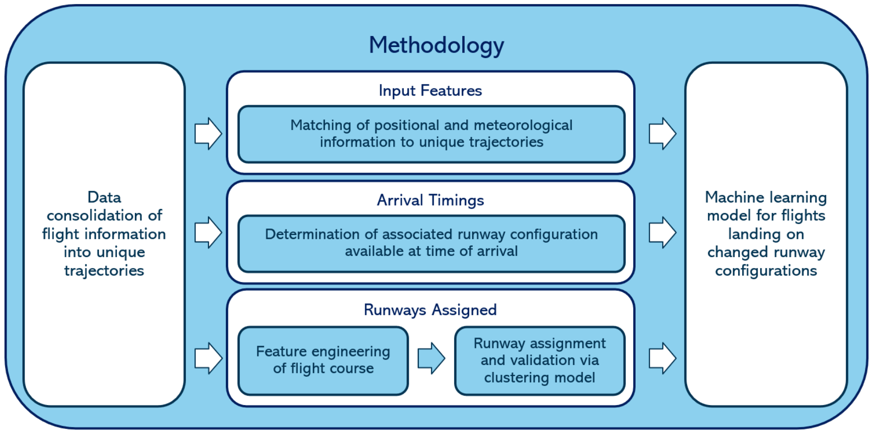

Figure 2: Proposed Methodology Determining Input Features and Target Variable for Predictive Modelling.

4.1 Data Preparation

The data set used in this paper is for a period of 62 days, during the months of December 2019 and January 2020 for Philadelphia International Airport (PHL), provided by Saab AB. The data set consists of flight trajectories, meteorological data, current and past runway configurations and active STAR routes. The data is pre-processed for input into the machine learning model.

4.1.1 Consolidation of Flight Positions into Trajectories

The first step in data preparation is to consolidate the individual flight position into trajectories. These trajectories are critical in deciding which runway the flight lands on as the flight positions and course in the trajectory provides information about the flight approach towards the runway.

The bulk of the data within the data set is the flight track data, consisting of positional information of each flight. These positional information are Cartesian coordinates with respect to the center of the airport. Each data point corresponds to a particular position of an aircraft at a fixed time and is reported every 1-5 seconds. Each flight has around 200 or more of such positional information. The positional information are grouped according to unique flights to produce a unique 'trajectory' class for each flight, consisting of the start and end times, unique call sign and positions of the flight. The final trajectory data set consists of 32,197 unique trajectories with a total of 36,587,053 flight positions over the two months period.

4.1.2 Associating Trajectories with Meteorological Data

The data preparation for input features included the consolidation of weather information for each of the flights. The raw data source provides METAR and TAF data. The relevant weather data is matched with each flight based on the time of the meteorological reports and the time of the flights. Table 1 shows the weather data extracted from the reports.

4.1.3 Matching Trajectories with Arrival Times and Runway Configuration

With the flights grouped into unique trajectories, the next step is to match each flight with their arrival times from the flight information dataset, the current active runway configuration and the upcoming runway configuration. The upcoming runway configuration is important for subsequent feature engineering to determine if each unique flight landed on a runway that is in the current and/or in the changed runway configurations. Each runway configuration is associated with its time of activation. The flight information and the runway configurations comes in separate data tables from the trajectories and they have to be matched accordingly. The matching that is done to combine the datasets is described below.

Table 1: Weather Information Extracted from METAR and TAF.

Weather Data	Description
Altimeter in Hg	Atmospheric pressure based on reference in inch of Mercury
Flight Category	Regime of flight rule (Visual/Marginal/Instrument) based on certain meteorological conditions
METAR Type	Type of report (METAR/SPECI) indicating a standard observation or an unscheduled report
Visibility in Statute Miles	Prevailing visibility in Statute Miles
Wind Direction in Degrees	Direction of wind measured
Wind Speed in Knots	Speed of wind measured
Sky Cover	Cloud coverage based on fraction of sky obscured by clouds
Cloud Base Above Ground Level (ft)	Height of the base of specified cloud cover

For arrival times, the flights are matched according to the date of flight and the call sign. From the 32,197 unique flights, 30,380 had successful matches to arrival times and runway configurations, with 1817 flights having no landing times determined, an overall matching rate of 94% . An example of Flight AAL655 with data consolidation can be observed in Figure 3.

traj_id	landing_time	Current_RC	Upcoming_RC	Current_RC_Active_Time	Upcoming_RC_Active_Time
AAL655_2	2019-12-06 13:04:00-05:00	27R,17	27L,17	2019-12-06 10:52:00-05:00	2019-12-06 13:28:00-05:00

Figure 3: AAL655_2 Runway Configuration.

For the flight illustrated in Figure 3, the active arrival runways are 27R and 17 (‘Current RC’). The changed runway configuration is expected to be 27L and 17 (‘Upcoming RC’), and the time of completion of the runway configuration change is at 13:28 local time(‘Upcoming RC Active Time’). If this flight are to land on runway 27L (a runway used only in the future runway configuration) at 13:04 local time(‘landing time’), it would be hypothesised that the transition had already begun and hence the difference in landing time (13:04) and time of completion (13:28) would result in a total transition time of 24 minutes.

4.1.4 Modal Flight Course Feature Engineering

With knowledge of the active and future runway configuration and the time of landing, the omission of the actual runway the flight landed on meant necessary feature engineering and imputation have to be carried out to predict the actual arrival runway.

The aircraft flight course is defined as the direction of the aircraft’s track representing the actual movement of the aircraft. During landings with little crosswind, the pilot aligns the aircraft as close to the centre line of the runway. Hence, determining the flight course will provide a good indicator of the heading of the runway in which the flight landed on. With the positional data at multiple instances of time, given as x and y coordinates, the authors defined each aircraft’s course at every instance.

In chronological order, the differences in x and y coordinates are evaluated and processed using trigonometric functions to determine the angle clockwise from True North (360°). This engineered value is saved as a new feature termed flight course (‘Course Deg’) in the data set. An example for flight PDT4997 is seen in Figure 4. Using the difference in x and y coordinates (-93 and +254 respectively), the appropriate arctangent function is applied to determine a flight course of 339.89°clockwise from True North, under the ‘Course Deg’ column. For each flight, there is a specific flight course computed for each instance based on consecutive flight positions.

The next feature engineered would be the modal flight course(‘Course Mode’). The modal flight course defined in this paper is flight course with the highest occurrence during the landing phase in the airport proximity. This proximity is set as (-7000, -1500) to (4500, 5000) on the x-y coordinate plane using the

timestamp	call_sign	x_position	y_position	traj_id	Course_Deg
2019-11-30 19:00:00.069000-05:00	PDT4997	2120	-3874	PDT4997_0	339.890213
2019-11-30 19:00:04.199000-05:00	PDT4997	2027	-3620	PDT4997_0	340.685339

Figure 4: PDT4997 Flight Course.

density of flight paths. The arithmetic mode of the flight course('Course Deg') for each flight is processed as the modal flight course('Course Mode'). This reduced the tracking errors at particular instances of time.

One such tracking deviation faced by Flight BAW69V is shown to be corrected in Figure 5. Flight BAW69V have a flight path that is observed to follow a general Southwestern track. However, from the figure, the flight course based on the two positions are 159.7°, which indicate the flight flies in the Southeastern direction during these 2 instances. This could be due to a suspected tracking error in the flight positions. However, the flight course based on the other positions are 260°. Hence, by taking the arithmetic mode of the flight course, a modal flight course of 260°is assigned for the entire flight.

timestamp	call_sign	x_position	y_position	traj_id	Course_Deg	Course_Mode
2020-01-12 19:44:37.046000-05:00	BAW69V	1394	2008	BAW69V_29	159.718987	260.0
2020-01-12 19:44:48.770000-05:00	BAW69V	1656	1299	BAW69V_29	159.718987	260.0

Figure 5: BAW69V Flight Course.

4.1.5 Runway Assignment

Since information regarding the arrival runway for each individual flight is unavailable in the data set, a hybrid circumscribed rectangular assignment method and clustering model is developed to assign the flights to their respective runways. The arrival runway for each flight is required to determine if the flight has landed on the changed runway. With the landing time matched with trajectories, the landing position of each aircraft can be determined. Using the 30,380 matched data points, the flight positions at the closest time of landing around the airport vicinity are plotted out. Based on the coordinate map of the landing positions, 4 distinct lines representing the 4 runways in PHL are determined to be the possible runway positions and plotted out. Figure 6 shows the landing positions of the flights and the possible runway positions.

The coordinates of the runways are determined iteratively to match the straight sections of the landing positions. Following which, 4 bounding boxes are created around the runways, representing the 4 runways. With these circumscribed rectangles in place and the previously engineered modal flight course for each flight on hand, the two-fold circumscribed rectangular approach assigned eligible flights to the arrival runways. Only flights with landing coordinates inside the rectangular bounding boxes are considered. The modal flight course of these flights must lie within the ±40°of the possible runway heading associated with that bounding box to be assigned as the runway of which the flight landed on. The result of this assignment is saved as the runway assigned for the 11,935 flights and plotted in Figure 7 out of the data set of 27,972 flights after removing outliers.

For the remaining unassigned flights, clustering models are applied to assign flights to the runways. As there is no ground truth label with the actual arrival runways provided, clustering provided the best alternative in grouping and sorting these flights into the arrival runways. The circumscribed rectangle assignment will be further used as a validation tool to interpret each of the derived clusters. The features used for clustering of flights to runways are the x-positions, y-positions and modal flight course of the flights.

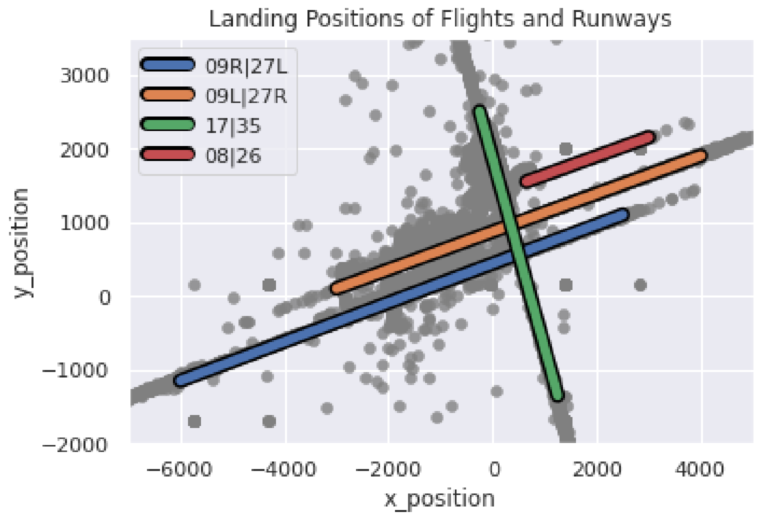

Figure 6: Plot of Theorised Runway Positions.

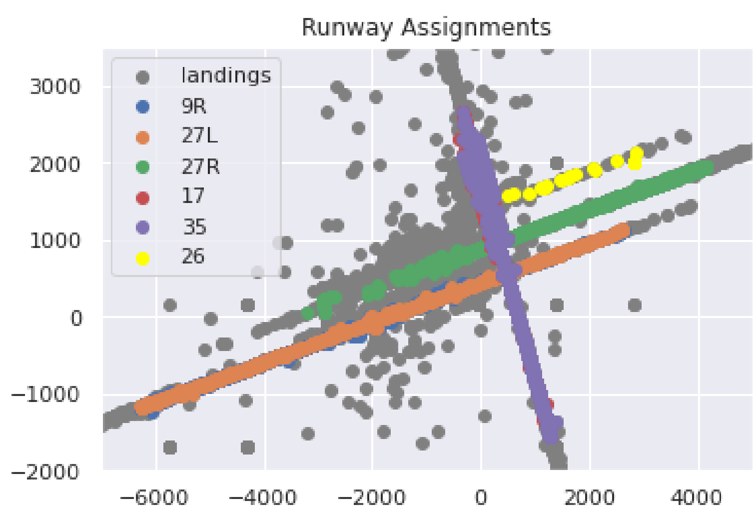

Figure 7: Plot of Circumscribed Rectangle Runway Assignments and Flight Landing Positions.

A state-of-the-art Gaussian Mixture Model (GMM) is selected as the clustering algorithm as it is able to form stretched and elliptical clusters, which are closer to the straight-line runways than all other types of clustering algorithms tested. The GMM assumes that each cluster is a Gaussian Distribution of the flight positions and courses, and the cluster should represent a runway. For each flight, the GMM computes the probabilities of each flight belonging to the hypothesized clusters based on the features. The flight is then assigned to the cluster with the highest probability. Since the GMM is unable to cluster distinct runways in one try, an iterative process is done until all flights are assigned a cluster that represents a runway.

The runway assignments are summarised in Table 2 and shown visually in Figure 8. The flights in each cluster are assigned to a runway and corroborated with the circumscribed rectangle assignment (CR).

The final runway assignments predicted the runways for an additional 12,444 flights for a total of 24,379 flights, dropping 3593 flights that are unassigned due to noise. Within the 11,935 flights assigned using the CR method, 96.9 %, or 11,566, of the flights had similar runway assignments using the clustering

Table 2: Cluster Assignment and Circumscribed Rectangle (CR) Evaluation.

Runway Assignment	17	35	27R	27L	9R	26	Total
Number of CR Flights Assigned	524	1950	5178	938	3302	43	11935
Number of CR Flights with Identical Clustering Assignments	521	1733	5053	937	3295	27	11566
CR Flights Assignment Accuracy	99.4%	91.6%	97.6%	99.9%	99.8%	62.8%	96.9%
Number of Flights Predicted	1050	1695	7672	207	1476	344	12444
Total	1574	3645	12850	1145	4778	387	24379

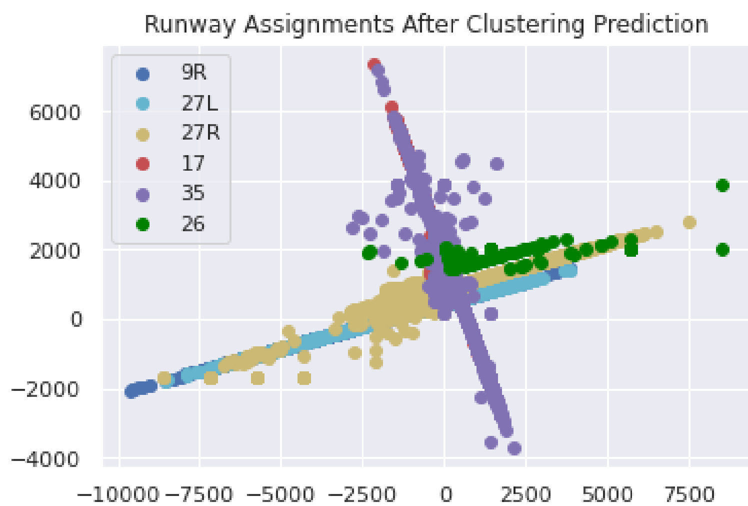

Figure 8: Runway Assignments for 24,379 Predicted Flights by Clustering.

approach. For the 24,379 predicted flights, 89.9%, or 21,917 flights, had valid assignments, where the flights landed either in the active runway configuration or changed runway configuration.

4.1.6 Transition-time Computation

Due to the methodology of the study, flights that are landing on the same arrival runway in both active and changed runway configurations are not relevant in the data set as there is no change in runways they landed on. The data set is filtered to consider flights that landed only in the changed configuration but not in the current active configuration. Based on the clustering, the 21,917 flights with valid runways included 1024 flights that landed exclusively in the changed runway configuration, signalling changes due to the runway configuration transition. For this relevant data set, the transition-time required is obtained by finding the difference between the time of completion and time of initiation, which is approximated by the time of landing. Time of completion is provided and processed in Section 4.1.3 during data consolidation.

All flights are also organised into their unique configuration changes, such as grouping all flights that changed from Arrival Runways 27L,17 to 27R,17 together. The median transition time for each of the runway configuration changes is calculated and subsequently capped at 360 minutes, the median value, based on the background knowledge and feasibility. Table 3 shows the 6 major configuration changes selected in this study, based on the number of flights.

Table 3: Runway Configurations Modelled.

Runway Configuration Change (Configuration Index)	Number of Flights
27L, 17 to 27R, 17 (1)	68
27L, 17 to 27R, 35, 26 (2)	49
27R, 35, 26 to 9R, 35 (3)	43
9R, 17 to 27R, 35, 26 (4)	38
9R, 35 to 27R, 35, 26 (5)	33
9R to 27R, 35, 26 (6)	28

4.2 Machine Learning Model

Several machine learning models are evaluated using a single runway configuration change and the root mean square errors (RMSE) is shown in table 4. Random Forest Regressor (RF) is shown to have the best result and therefore selected for the regression problem. RF is an ensemble model adopts an averaging approach, where decision trees are combined and the prediction of the ensemble is then the averaged prediction of all the randomised decision trees. The randomness in each of the decision trees trained independently and considered in the prediction reduces the variance of the model as a whole, improving the overall performance (Pedregosa et al. 2011).

Table 4: RMSE of Regression Models for Arrival Runways Change from 27L, 17 to 27R, 35, 26

Model	Root Mean Squared Error (RMSE)
Linear Regression	56.2
Ridge Regression	55.3
Bayesian Ridge	50.0
Support Vector Regression	73.3
K-Nearest Neighbour	65.4
Random Forest Regressor	**20.9**
AdaBoost	31.9
GradientBoost	33.9

The input features selected involved both the positional flight information and the weather information and is outlined in Table 5. The target variable is calculated for each of the relevant flights as the time difference between the time of completion of runway configuration change and time of landing.

Table 5: Regression Analysis Input Features

Input Features						
Positional Factors			Meteorological Factors			
X position	Y position	Modal flight course	Wind Direction	Wind Speed	Visibility	First Cloud Base

5 EXPERIMENT AND RESULTS

5.1 Training and Evaluation

For each individual runway configuration, the regression analysis is carried out first by splitting the training and test set with a train/test mix of 0.8/0.2. This allowed each of the models to learn from the training set and then predict the transition timings on the testing set separately, without data leakage. For a supervised learning prediction, the first evaluative metric chosen to determine the performance for tuning is root mean square error (RMSE) as defined in Equation 1, where N is the number of data points and the y value is the target variable (transition time).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_{pred} - y_{actual})^2} \quad (1)$$

A secondary indicator, R^2, or coefficient of determination, is also calculated to measure goodness-of-fit or how close the predicted data and actual data are by measuring the difference in sum of squares. The maximum score is 1.0 for a perfect fit. Mathematically, it is represented as $R^2 = 1 - \frac{u}{v}$, where

$$u = \sum_{i=1}^{N}(y_{actual} - y_{pred})^2, \text{ and} \quad (2)$$

$$v = \sum_{i=1}^{N}(y_{actual} - \overline{y_{actual_{mean}}})^2 \ . \quad (3)$$

After hyper parameter tuning, the respective RMSE and R^2 scores for each of the 6 configuration changes using a RF model are shown in Table 6. A visual example of the difference in predicted and actual values of such transition timings for runway configuration index 2 is observed in Figure 9. The percentage errors between the predicted and actual values are listed in red above the bars.

Table 6: RMSE and R^2 Scores of RF Model for 6 Runway Configuration Changes.

Runway Configuration (Configuration Index)	RF RMSE	RF R^2 Score
27L, 17 to 27R, 17 (1)	20.9	0.839
27L, 17 to 27R, 35, 26 (2)	16.7	0.922
27R, 35, 26 to 9R, 35 (3)	37.2	0.906
9R, 17 to 27R, 35, 26 (4)	32.7	0.843
9R, 35 to 27R, 35, 26 (5)	15.2	0.844
9R to 27R, 35, 26 (6)	14.6	0.949

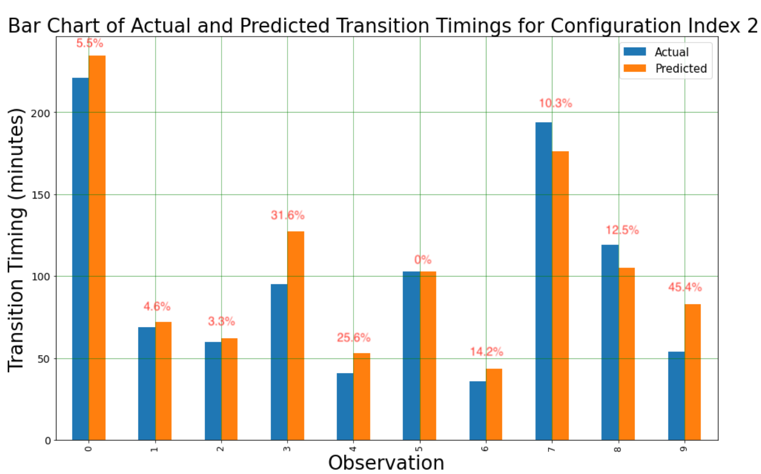

Figure 9: Bar Chart of Predicted and Actual Transition Timings for Runway Configuration Index 2.

With the results for the 6 runway configuration changes as shown in 6, the R^2 scores of at least 0.8 signified each RF model developed could largely explain the variance of transition timings with a good fit of the results. The median RMSE across the 6 models of 18.8 minutes represents the difference in actual and predicted values of the runway configuration transition timings. The results highlight the ability of the model to provide a decent prediction of the transition times as the actual transition times for various configurations may take several hours based on our definition of transition times.

5.2 Feature Importance

Apart from the selection and scoring of the RF model for each runway configuration change, the author conducted a further investigation by observing feature importance. The results can be condensed into Table 7. The top 3 important features for each runway configuration index are highlighted in bold.

Table 7: Feature Importance for Configuration Changes using Random Forest Regressor.

Feature	Feature Importance for Configuration Index					
	1	2	3	4	5	6
X position	0.114	0.047	0.04	0.024	0.081	0.012
Y position	0.107	0.134	0.053	0.033	**0.352**	0.077
Flight course	0.114	**0.235**	**0.066**	**0.11**	0.095	0.019
Wind Direction	**0.233**	**0.161**	0.04	0.045	**0.264**	**0.123**
Wind Speed	**0.26**	0.127	**0.611**	**0.093**	**0.137**	**0.126**
Visibility	0	0	0.002	0.002	0	0
First Cloud Base	**0.171**	**0.296**	**0.188**	**0.693**	0.071	**0.643**

From Table 7, it is observed that wind speed, wind direction and first cloud base height are 3 of the most important factors in almost all the 6 models trained. The predictive power of wind factors in the feature importance of the models corroborate the conclusion reached in related literature. However, cloud base height is an important feature specifically in the case of Philadelphia International Airport. Furthermore, positional factors from the flight information indicated a less important role compared to combined feature importance from meteorological conditions in the determination of transition time. Considering the entire runway configuration change process for an ATCO, a macro view such as weather conditions around the airport vicinity might seem to be a bigger factor than an individual plane's positioning and distance away from the runways. Nonetheless, this might still be an area to consider when thinking about the number of disruptions and holdings that might occur in the terminal airspace.

6 CONCLUSION

The main aim of the study is to develop a model to predict the transition-time required for a runway configuration change to be completed. Through the heterogeneous sources of data provided, extensive data preparation and processing is performed, including predictions on runways on which aircraft landed by using a clustering model. The transition times for 6 major runway configurations changes are modelled using Random Forest Regressor based on the training and test of the consolidated data set. The prediction of transition-time for all models achieved R^2 scores of at least 0.8 and a median RMSE of 18.8 minutes, highlighting the predictive capabilities of the machine learning model in a dynamic environment. The machine learning methodology enabled further inspection of the feature importance in transition timings. Wind conditions, as expected, played an important role in determining the transition time required. Cloud base height is surprisingly an important feature in almost all configurations considered in the congested Philadelphia airspace. The quantification of transition-time may provide Philadelphia ATCO more informed choices in determining the right runway configuration to achieve the optimal runway capacity.

For future works regarding transition timings, the machine learning methodology of the study may be refined by considering other features such as regulatory restrictions or explore inter-regional effects in multi-airport scenarios. Research on other airports with published time of initiation of runway configuration changes over extended periods of time may also yield new insights regarding personalised or regional characteristics of the airport to better serve the ATCO and greater goal of air traffic optimisation.

REFERENCES

Ahmed, M. S., S. Alam, and M. Barlow. 2018. "A Multi-Layer Artificial Neural Network Approach for Runway Configuration Prediction". In *Proceedings of the 8th International Conference on Research in Air Transportation (ICRAT 2018)*. June 26th - 29th, Castelldefels, Spain.

Bertsimas, D., M. Frankovich, and A. Odoni. 2011. "Optimal Selection of Airport Runway Configurations". *Operations Research* 59(6):1407–1419.

Duarte, M., C. Weld, and R. Kincaid. 2010. "A Runway Configuration Management (RCM) model with marginally decreasing transition capacities". In *2010 IEEE Systems and Information Engineering Design Symposium*. April 23th, Charlottesville, Virginia, 106-111.

Idris, H., B. Delcaire, I. Anagnostakis, W. Hall, N. Pujet, E. Feron, R. Hansman, J.-P. Clarke, and A. Odoni. 1998. "Identification of flow constraint and control points in departure operations at airport systems". In *Guidance, Navigation, and Control Conference and Exhibit*. August 10th - 12th, Boston, Massachusetts, 947-956.

Jacquillat, A., A. R. Odoni, and M. D. Webster. 2017. "Dynamic Control of Runway Configurations and of Arrival and Departure Service Rates at JFK Airport Under Stochastic Queue Conditions". *Transportation Science* 51(1):155–176.

Li, L., and J.-P. Clarke. 2010. "A Stochastic Model of Runway Configuration Planning". In *AIAA Guidance, Navigation, and Control Conference*. August 2th - 5th, Toronto, Ontario, Canada.

Ng, K. K. H., C. K. M. Lee, and F. T. S. Chan. 2017. "A robust optimisation approach to the aircraft sequencing and scheduling problem with runway configuration planning". In *2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. December 10th - 13th, Singapore, 40-44.

Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. 2011. "Scikit-learn: Machine Learning in Python". *Journal of Machine Learning Research* 12:2825–2830.

Ramanujam, V., and H. Balakrishnan. 2011. "Estimation of maximum-likelihood discrete-choice models of the runway configuration selection process". In *Proceedings of the 2011 American Control Conference*. June 29th - July 1th, San Francisco, California, 2160-2167.

Yin, J., Y. Ma, W. Tian, D. Chen, Y. Hu, and W. Ochieng. 2020. "Impact Analysis of Demand Management on Runway Configuration in Metroplex Airports". *IEEE Access* 8:66189–66212.

AUTHOR BIOGRAPHIES

LAU EN CHENG MAX is a research assistant at Saab-NTU Joint Lab at Nanyang Technological University. He graduated with a Bachelor of Engineering in Aerospace Engineering in 2021. His email address is elau004@e.ntu.edu.sg.

LAM JUN GUANG ANDY is a research engineer at the Saab-NTU Joint Lab at Nanyang Technological University. He is currently a PhD student at School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore. His e-mail address is lamj0013@e.ntu.edu.sg.

SAMEER ALAM is an associate professor and co-director of Saab-NTU Joint Lab at School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore. His email address is sameeralam@ntu.edu.sg.