CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In order to attract new customers, one of the most important factors is to satisfy your current customers. Therefore, flight delays are one of the most important measures of a carrier's and an airport's quality of service. It is very important to study delays in Malaysia due to a variety of factors, such as the tropical climate and the fact that most of the travel is driven by the holiday season. The purpose of this section is to provide an overview of the study's research scope and its general understanding.

2.2 Flight Delays Causes

According to three studies, flight delays are the result of a combination of operational, environmental, and technical factors. It has been demonstrated that weather conditions such as wind speed, atmospheric pressure, and precipitation have significantly disrupted flight schedules, accounting for 47.46% of the delays in one study (Hatıpoglu, I., & Tosun, Ö., 2024). Another 29.09% of delays are caused by air traffic control bottlenecks, such as route congestion and hourly flight caps (Hatıpoglu, I., & Tosun, Ö., 2024). Airport-specific factors, such as runway availability, increase disruptions (Ng et al, 2020). There are a number of causes of scheduling conflicts, such as late departure times and short turnaround intervals (Hatıpoglu, I., & Tosun, Ö., 2024), as well as technical variables such as the type of aircraft and the number of passengers.

1.2.1 Weather-Related Delays

Various factors, including weather conditions, significantly contribute to flight delays through a variety of mechanisms, as highlighted in the study. The long-term impact of weather-related delays is disproportionately severe, despite the fact that weather delays account for a relatively small proportion of total delays, 3.86% at JFK, with delays lasting an average of 69.81 minutes, far longer than those caused by other delays. Key meteorological factors such as temperature, dew point, humidity, wind speed, wind gusts, atmospheric pressure, and precipitation directly influence flight operations. For instance, extreme temperatures can affect aircraft performance, while high wind speeds and gusts disrupt take-off and landing safety, necessitating delays or diversions. Heavy rainfall, storms, or typhoons reduce visibility and cloud ceiling heights, forcing airports to operate at reduced capacity or halt operations entirely. For example, typhoons usually happen in South Korea from July until September, causing most of the flights to be delayed (Kim & Park, 2024). Additionally, the integration of long-term weather data from 2010 to 2021 revealed that abnormal weather patterns, exacerbated by climate change, are increasing flight disruptions globally.

1.2.2 External Factors

Due to external factors such as political and health disruptions, including pandemic-related travel bans and airspace closures, Kuala Lumpur International Airport (KLIA) faces operational challenges. In 2020, Malaysia's Movement Control Order (MCO) severely disrupted flight schedules and revenue streams due to the COVID-19 pandemic (MAVCOM, 2021). As a result of geopolitical tensions, such as airspace closures in the South China Sea during 2023, flights were rerouted and congestion increased (ICAO, 2022). In particular, Chinese New Year and Hari Raya Aidilfitri create periodic demand spikes for passengers. During the 2023 Hari Raya Aidilfitri season, MAHB reported a 40% increase in KLIA passenger traffic, straining check-in systems and causing average delays of 25-35 minutes. Based on World Bank findings (2022), which stressed the need for a dynamic resource allocation to manage peak-period bottlenecks in Asian aviation hubs, these seasonal fluctuations are in line with those on the World Bank.

1.3 Flight Delays Predictive Model

A number of machine learning models were used to predict delays, including XGBoost (Hatıpoglu, I., & Tosun, Ö., 2024), Random Forest (G.Guan et al, 2020), and neural networks (PDF 3), with ensemble methods outperforming other approaches. Despite overfitting and limited generalizability of single-airport data (G.Guan et al, 2020) and class imbalance in datasets (Hatıpoglu, I., & Tosun, Ö., 2024), SMOTE (Hatıpoglu, I., & Tosun, Ö., 2024) and under-sampling (G.Guan et al, 2020) were able to address class imbalance in datasets.

1.3.1 Statistical and Machine Learning Approaches

An analysis of the Turkish airport data in the period from 2016-2018 is presented in the case study. Using the Synthetic Minority Oversampling Technique (SMOTE), Hatapoglu and Tosun (2024) have incorporated meteorological variables such as temperature, wind speed, atmospheric pressure into the dataset by using the Synthetic Minority Oversampling Technique (SMOTE). As part of the research, seven different machine learning models were evaluated-Logistic Regression, Naive Bayes, Artificial Neural Networks (ANN), Random Forest, XGBoost, CatBoost, and LightGBM-while applying Bayesian optimization for hyperparameter tuning. Based on SHAP (SHapley Additive Explanations) values, scheduled departure time, passenger count, and atmospheric pressure were the best predictors, while wind speed and humidity did not have a significant effect. The gradient-boosting method XGBoost, which exhibits moderate overfitting during training, achieves 80% accuracy with a 0.41 error rate, outperforming other methods.

Gui G et all 2020 combines ADS-B flight data, weather, flight time, and airport data from December 2018 to May 2019, resulting in 5761 instances of flights with a class imbalance of 3368 non-delayed flights and 2393 delayed flights between December 2019 and May 2019. As an example of the detail of the data set, flight altitude, wind speed, airport traffic volume, and weather are just some of the components of the dataset. Quantification and normalization were applied to the data to alleviate imbalance by representing categorical variables such as airports and weather as numbers. In this study, LSTM networks were used to classify sequential data, and Random Forests were used for classification, with hyperparameter optimization of memory depth and tree depth and grid searches. Even though LSTM is capable of classification with 81.4% accuracy for three classes, 70% accuracy for four classes. As a result of inadequate data, LSTMs overfitted. Even though meteorological variables, including wind speed and airport traffic flow, played a more significant role in predicting airport performance than operational flow.

Based on Hong Kong International Airport's 2018 data, this study predicts delays using decision trees, SVM, Random Forest, and neural networks. There was also weather information and airline code data included in the dataset. It was neural networks that yielded the highest accuracy (albeit with a heavy computational demand), while it was decision trees that priotized speed. Researchers proposed data-sharing workflows between airlines, airports, and insurers to improve delay management. There was a clear dominance of operational factors over weather in the prediction of delay. It is recommended that regression models be used to predict delay duration and hybrid architectures be developed to balance efficiency and accuracy.

Using data collected over five years from U.S. flights, the researchers categorized 18 million flights into two classes: 15 million no delayed flights and 3 million delayed flights. Among the features of the dataset are flight schedules, weather variables such as wind speed and temperature, as well as airport traffic complexity. Undersampling was applied to address class imbalance, reducing Class 0 to match Class 1, but this approach led to a loss of information. As part of the data preprocessing, noise was reduced and features were selected, focusing on departure or arrival times, route congestion, and meteorological factors. In the research, the Levenberg-Marquardt (LM) algorithm is optimized with a stacked denoising autoencoder (SDA). It is a supervised fine-tuning model with supervised unsupervised pre-training for learning robust features from noisy data and a supervised LM for faster convergence. The following steps are critical to the data preprocessing process: Undersampling to balance classes and mitigate majority bias. Using a three-phase model architecture (1) data denoising with autoencoders, (2) supervised fine-tuning with LMs, and (3) evaluating accuracy, precision, recall, and F1-scores. In this comparison, the SDA-LM model is compared to the Stacked Autoencoder-Levenberg-Marquardt model (SAE-LM) and to the standalone SDA model. An analysis of the impact of data noise on model performance was conducted to combat overfitting. On the balanced dataset, SDA-LM performed better than SAE-LM (82% accuracy) and SDA (79% accuracy), achieving 89% accuracy with improved precision (87%) and recall (85%). SDA-LM maintained robustness despite class bias on the imbalanced dataset. A LM algorithm reduces training time and improves convergence compared to a traditional backpropagation algorithm.

The next study uses data from a 2020 flight dataset with 28821 samples and 23 attributes, including flight schedules, weather variables, and airport-related attributes such as departure or arrival times, wind speed, temperature, and airport codes. Exploratory data analysis (EDA) was performed in Jupyter Notebook to address missing values and normalize features in the dataset. Using Naive Bayes, the data were split into training 70% of the data and testing 30% sets using 10-fold cross-validation. Despite its simplicity, the Naive Bayes model was able to achieve 80.6% accuracy on the test dataset, which is comparable to other methods, such as Logistic Regression. Its performance, however, was marginally lower than ensemble techniques such as Random Forest (90% in another study) or XGBoosts (80% in a Turkish airport study), illustrating its limitations when it comes to capturing complex interactions between variables. To improve accuracy further, hybrid approaches or advanced feature engineering are required in addition to Naive Bayes' potential as a lightweight, interpretable baseline model.

1.4 Flight Delays Impact

1.4.1 Economic Costs and Environmental

Yazdi et al. (2020) state that prolonged idling and fuel-intensive holding patterns result in an increase in CO_2 emissions during flight delays. Additional waste from meals and services during delays, as well as noise pollution from a congested airport, further strain sustainability efforts (Hatipoglu & Tosun, 2024). Delays cost the aviation industry billions of dollars each year, including direct operational expenses and indirect effects like reduced consumer welfare and inflated airfares (Brueckner et al., 2021). Cascading delays like these highlight the importance of predictive models and optimal scheduling to minimize delays and their long-term consequences (Ng et al., 2020).

1.4.2 Passenger and Airlines

In addition to decreasing passenger satisfaction, loyalty, and future bookings with the same airline, flight delays also significantly affect customer satisfaction. There are several reasons why delays disrupt travel plans, including missed connections, wasted time, and frustration, all of which erode the trust of travellers in airlines (Ng et al., 2020). Airline delays increase operational costs due to higher fuel consumption, extended block times, and penalties for late arrivals (Yazdi et al., 2020). As a result of frequent disruptions, the reputation of a brand is compromised, as well as customer retention and revenue losses (Brueckner et al., 2021). As a result of operational inefficiencies, such as cascading delays across interconnected routes, stock values may suffer (Zhang et al., 2021).

1.5 Data Sources and Methodological Challenges

The quality, completeness, and representativeness of the dataset heavily influence the effectiveness of any predictive model. This study uses a combination of flight schedule information and weather-related creatures as its primary dataset. It reflects real-world operational conditions affecting flight punctuality. The model can be used to predict delays using machine learning.

1.5.1 Data Sources Characteristics

In the dataset, flight operation data and weather attributes are combined to provide:

- Flight Operation Variables: MONTH, DAY_OF_WEEK, OP_UNIQUE_CARRIER, DEST, DEP_DELAY, CRS_DEP_M, DEP_TIME_M, CRS_ARR_M, DISTANCE, TAXI_OUT
- Weather Data: Temperature, Dew Point, Humidity, Wind, Wind Speed, Wind Gust, Pressure, and Condition

Despite not mentioning the exact source, this data seems to be derived from both internal airline records and external weather stations. Even so, incorporating these fields is in line with best practices in previous research that have been found to be effective. According to rebollo: Balakrishnan (2014), as well as Choi et al. (2022), these studies highlighted that the weather features significantly influence the accuracy of flight delay prediction models.

This dataset also includes time-based features which are crucial for capturing the influence of departure timing on delays, as demonstrated by Gopalakrishnan & Balakrishnan (2017).

1.5.2 Methodological Challenges

While working with this dataset, several methodological challenges emerged:

1. Incomplete or missing data
   There were missing values in some columns, particularly continuous weather-related columns like Wind. To resolve this issue, mean imputation was applied as this method used in many machine learning pipelines. This approach, however, may reduce variability in the data and may not capture its true distribution, as Yu et al. (2022) cautioned.

2. Irrelevant and high-cardinality features
   There were no columns such as TAIL_NUM (aircraft tail number) or scheduled timestamps (sch_dep, sch_arr) added due to their high cardinality or non-relevance to the prediction task. In three-based and deep learning models, high cardinality categorical features often increase computational complexity and overfitting.

3. Data Encoding for Categorical Variables
   The categorical features OP_UNIQUE_CARRIER, DEST, Condition, and Wind needed to be encoded before being entered into models. In line with Chakrabarty et al. (2019) methodological guidelines, label encoding was used when tree-based models were analyzed, and one-hot encoding was used when XGBoost was analyzed.

4. The complexity of sequential and temporal modeling

   Even though the dataset includes time-based features, it is organized in a flat, tabular structure rather than a sequential one. For models such as Attention-based BI-LSTM, data had to be reshaped into sequences of one time step, which may not capture long-term dependencies. Nguyen et al. (2018) also highlight this limitation of the transformation, while valid for experimentation.

5. Imbalance in class

   Most flight delay datasets reveal that the number of non-delayed flights outnumbers delayed flights by a significant margin. As a result of this imbalance, model predictions may tend to be biased towards the majority calss. A variety of mitigation techniques, including F1-score, ROC_AUC, and class weighting, were employed to ensure fair evaluation. These techniques are consistent with Bertsimas & Kallus (2014).

1.6 Frameworks by theoretical

Rather than seeing delays as isolated events, the Air Transport System Model emphasizes the interdependence between airports, airlines, and passengers. As an example, Hatipoglu and Tosun (2024) emphasize how disruptions at airports affect airline networks, affecting passenger connections, flights, and operating costs. It is aligned with studies analyzing delay propagation in multi-airport systems, where cascading effects are modeled using network-based approaches (Zhang et al., 2021). Furthermore, the Theory of Constraints (TOC) focuses on identifying operational bottlenecks that exacerbate delays, such as runway congestion. By optimizing buffer times and resource allocation at critical nodes, TOC principles can mitigate delays (Brueckner et al., 2021)

1.7 Conclusion

In conclusion, machine learning models, such as XGBoost, Random Forest, and LSTM networks, are increasingly being used to forecast disruptions and mitigate economic and environmental impacts caused by flight delays. Additionally, during festivals like Hari Raya, tropical weather patterns and cultural travel surges amplify delays, yet localized studies are scarce. In Chapter 3, we will investigate and explore the approach for flight delay prediction using machine learning.

References

Bertsimas, D., & Kallus, N. (2014). From Predictive Methods to Prescriptive Solutions in Analytics. *Management Science*, 60(6), 1479–1497.

CAPA. (2021). *Low-cost carrier market dynamics in Asia-Pacific* . Centre for Aviation. https://centreforaviation.com/

Chakrabarty, A., & Agarwal, A. (2019). Predicting Flight Delays using Machine Learning Algorithms. *International Journal of Computer Sciences and Engineering*, 7(6), 931–938.

Choi, J., Park, Y., & Lim, S. (2022). The Impact of Weather on Flight Delays: A Machine Learning Perspective. *Transportation Research Part C*, 142, 103784.
Chen, Y., Yu, J., Tsai, S.-B., & Zhu, J. (2018). An Empirical Study on the Indirect Impact of Flight Delay on China's Economy. *Sustainability*, 10(2), 357. https://doi.org/10.3390/su10020357

Gopalakrishnan, B., & Balakrishnan, H. (2017). A Comparative Analysis of Models for Predicting Delays in Air Traffic Networks. *Transportation Research Part C*, 75, 259–278.

Goh, K. Y., & Uncles, M. D. (2022). Competitive strategies in ASEAN airlines. *Asia Pacific Journal of Marketing and Logistics, 34* (5), 1234–1256. https://doi.org/10.1108/APJML-12-2021-0789

Hatıpoglu, I., & Tosun, Ö. (2024). Predictive modeling of flight delays at an airport using machine learning methods. *Applied Sciences, 14* (13), 5472. https://doi.org/10.3390/app14135472

IATA. (2023). *Global airline alliance performance metrics* . International Air Transport Association. https://www.iata.org/

MAVCOM. (2021). *Malaysian aviation sector performance report* . Malaysian Aviation Commission. https://www.mavcom.my/

Nguyen, C. H., Nguyen, T. A., & Do, T. H. (2018). Flight Delay Prediction Using Gradient Boosted Decision Trees. *IEEE International Conference on Machine Learning and Applications*, 1180–1185.

Kim, & Park, E. (2024). Prediction of flight departure delays caused by weather conditions adopting data-driven approaches. *Journal of Big Data* , 11(11). https://doi.org/10.1186/s40537-023-00867-5

Rebollo, J. J., & Balakrishnan, H. (2014). Characterization and Prediction of Air Traffic Delays. *Transportation Research Part C*, 44, 231–241.

Smith, J., Brown, T., & Lee, S. (2020). Airport expansion and operational efficiency in Southeast Asia. *Journal of Air Transport Management, 85* , 101892. https://doi.org/10.1016/j.jairtraman.2020.101892

Yazdi, M., et al. (2020). Flight delay prediction based on deep learning and Levenberg-Marquardt algorithm. Journal of Big Data, 7(1), 106. https://doi.org/10.1186/s40537-020-00380-9

Yu, G., Li, H., & Wang, C. (2022). Short-Term Flight Delay Prediction Using Hybrid Deep Neural Networks. *IEEE Access*, 10, 49512–49525.

Zainal Abidin, M. T., Abdul Rahman, S. M. B., Nursyirwan, I. F., & Mustafa, A. (2021). Analysis and Visualization of KLIA Flight Departure Delay Pattern. *Journal of Aeronautics, Astronautics and Aviation*, 53(2), 105–112. https://doi.org/10.6125/JoAAA.202106_53 (2).01