Air Passenger and Freight Demand Forecast for Ireland

MSc Research Project
Data Analytics

Nivriti Verma
Student ID: 21201421

School of Computing
National College of Ireland

Supervisor: Dr. Hicham Rifai

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Nivriti Verma
Student ID:	21201421
Programme:	Data Analytics
Year:	2023
Module:	MSc Research Project
Supervisor:	Dr. Hicham Rifai
Submission Due Date:	14/12/2023
Project Title:	Air Passenger and Freight Demand Forecast for Ireland
Word Count:	8089
Page Count:	20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Nivriti Verma
Date:	14th December 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Air Passenger and Freight Demand Forecast for Ireland

Nivriti Verma
21201421

Abstract

This study addresses the dynamic challenge of predicting variable air transport demand in Ireland and its partner nations. Investigating the repercussions of the COVID-19 pandemic on freight and passenger movement within Ireland, the research underscores the critical need for precise predictions to enhance operational efficiency and strategic foresight. Utilizing a comprehensive approach, the paper employs four distinct time series forecasting models, such as ARIMA/SARIMA, Simple Exponential Smoothing, Double Exponential Smoothing (Holt's Linear), and Triple Exponential Smoothing (Holt-Winters) and meticulously evaluates their efficacy in forecasting both passenger and freight demand. Notably, the findings indicate the superior performance of the Holt Linear model in predicting passenger demand, while the SARIMA model stands out for its accuracy in freight demand forecasting. This study provides valuable insights into forecasting methodologies within the aviation industry, establishing the foundation for technological advancements in capacity optimization and informed decision-making.

1 Introduction

The air transport industry is pivotal in facilitating freight and mail movement and passenger travel, both domestically and internationally. The geographical location of Ireland makes it a critical hub for air transportation, connecting it with its partner countries. However, aviation sector navigation is a dynamic and complex challenge influenced by various factors Standfuss et al. (2021). According to the Central Statistics Office (CSO), more than 38 million passengers and 146,000 tonnes of freight in 2019 were handled by Ireland's main airports, before the COVID-19 pandemic severely disrupted the aviation industry[1]. In 2020, the passenger numbers dropped by almost 80%, while the freight volumes decreased by 4%[2]. In 2021, the passenger numbers recovered slightly, reaching 9.1 million, but still remained far below the pre-pandemic levels, while the freight volumes increased by just 2%[3]. The demand for air transport forms the bedrock for airlines crafting their flight schedules and ground support strategies for passenger travel and freight flow. Disruptions and delays in schedules, particularly in crowded airports, create a ripple effect that causes operational bottlenecks and discontent among passengers and congestion which spans the entirety of the air transport system Zografos et al. (2017).

[1] https://www.cso.ie/en/releasesandpublications/ep/p-as/aviationstatisticsquarter22022/
[2] https://www.cso.ie/en/releasesandpublications/er/as/aviationstatisticsquarter32021/
[3] https://www.cso.ie/en/releasesandpublications/ep/p-tranom/transportomnibus2021/aviation/

Confronted with the unpredictable stemming from various unforeseen factors the quest for more reliable forecasting methods for aviation demand emerges as a pivotal challenge for the civil this industry.
Passenger or freight loading delays can have negative consequences for airports, airlines, and customers, as they can affect the profitability, operational efficiency, and satisfaction of the air transport sector. The average cost of delay per minute for airlines in Europe was estimated at 76.5 euros in 2016, translating into a total cost of 17.6 billion euros for the whole year. It was found that the main causes of delay were airline-related (aircraft, crew and baggage issues), airport-related (security, capacity and infrastructure issues), and en-route-related (weather issues and air traffic control). Reducing delays could generate significant benefits for the industry and society, such as lower emissions and operating costs, higher revenues and passenger satisfaction[4]. Zografos et al. (2017)examined the impact of passenger boarding and freight loading on the turnaround time of aircraft, which is the time between the arrival and departure of an aircraft at an airport. The study considers various factors, such as aircraft type, passenger profile, baggage and cargo volume and ground handling equipment to optimise resource allocation. It is extracted that technological advancements can aid decision-making to pave the path of recovery for the aviation sector from such incidences Kitsou et al. (2022). Irish airports need to have accurate and reliable forecasts of the air transport demand to adapt to the changing market conditions, optimize operations and increase passenger travel for revenue generation. In airports with restricted access, (many of the busiest European ones), limitations can lead to either loss of demand or a shift in demand to less favourable times of the day or alternative airports. Conversely, at less busy airports, airline demand may significantly lag behind the available capacity, resulting in the underutilization of infrastructure resources Jacquillat and Odoni (2018). Therefore, predicting future air traffic demand is formidable given the influence of seasonal variations, economic shifts, and ever-changing travel preferences.

1.1 Research Question

In addressing the challenges faced by the aviation sector, this research addresses the following question: "How can time series forecast models adapt to the high fluctuating passenger and freight demands?". The current work seeks to fill the gap in the literature on the application of time series models for air transport demand forecasting, especially in the Irish context. The objective is to predict the maximum passenger and freight traffic, along with comparing the accuracy of different time series forecasting models such as ARIMA/SARIMA, simple exponential smoothing, double exponential smoothing and triple exponential smoothing based on obtained evaluation metrics (MAE, MSE, RMSE). This will provide a quantitative measure of alignment with the actual data, which is crucial for identifying the highly reliable forecasting approach. Visual comparison of forecasted values with actual data will help the research in analyzing each model's adaptation to the fluctuations in passenger and freight flow.

The research is organised as follows: Section 2 includes a review of the existing literature on the current and future trends of air transport demand, understanding the methods and models already used for air transport demand forecasting, and their benefits and

[4]https://www.independent.ie/business/world/europes-flight-delays-cost-100-a-minute/37934906.html

limitations. Section 3 describes the methodology used for this study. In addition to the description of the dataset, section 4 involves the design specification followed by Section 5 which describes the implementation steps of the time-series forecasting models. Evaluation of the models using statistical values of MAE, MSE and RMSE is discussed in section 6. Section 7 does the contrast and comparison of the results with the existing literature. Lastly, section 8 concludes the main findings and contributions of the study and restates the research question and objectives.

2 Related Work

Predicting the demand for air transport is a vital task in the aviation sector, facilitating the strategic planning and efficient utilization of capacities at different airports. Accurate and reliable forecasts aid in decision-making for airport and airline authorities. However, it is a complex challenge due to the various factors and uncertainties influencing the travel choices and inclinations of both passengers and freight shippers. Suh and Ryerson (2019) delves into the forecasting of aviation demand and considers two factors that contribute to the demand uncertainties: dynamic socioeconomic changes and industry trends. The findings from publicly available data indicate that in regions where passenger numbers show a consistent increase over time, the likelihood of demand reduction in airports within those areas is diminished. Moreover, the air transport industry is subject to frequent disruptions, such as economic crises, natural disasters, pandemics, and environmental regulations, that can have a strong impact on demand trends and patterns. This section critically analyses four models for predicting air transport demand forecast from existing studies.

2.1 Econometric models

Such models are widely used in aviation demand forecasts, as they have the capability of capturing the causal relationships between demand and explanatory variables, such as income, price, GDP, exchange rate, and other socio-economic and demographic factors. Econometric models have two types, that is, aggregate and disaggregate. Aggregate models estimate the demand at a macro level (region, country, or market). In contrast, disaggregate models estimate the demand at a micro level (household, individual, or segment). The gravity model is one of the most sort aggregate models, which assumes that two location's demand is directly proportional to their size and inversely proportional to their distance. For example, in a 2018 study, Matsumoto and Domae (2018) the gravity model was applied to a study that examined the impact of international air traffic on East and Southeast Asian city's hub status from 1982 to 2012. Surprisingly, results revealed a positive correlation between distance and air cargo volumes for specific cities. A subsequent 2019 study using a similar gravity model identified a decline in generative factors like GDP and population post-2000. This decline may result from shifts in GDP components, indicating the need for future air traffic gravity models to incorporate more direct measures, such as consumer spending on goods and transportation services.

The vector autoregressive (VAR) model is another aggregate model, which can capture the interdependencies and feedback effects among multiple variables. For example, Gunter and Zekan (2021) utilized a Global Vector Autoregression (GVAR) model to analyze passenger figures from the foremost 20 global airports, those in the Asia-Pacific region, and those in Latin America-Caribbean areas. In considering air passenger numbers as an indicative measure of demand, the GVAR framework incorporated country-specific proxies for key economic determinants such as fare price and income of passengers at the airport level. Additionally, foreign variables were derived by establishing a weighting matrix grounded in the volume of flight connections among the selected airports. Although the study is rooted in air transportation, the modelling methodology demonstrated can be applied to demand forecasting in any tourism sector where the emphasis is on interdependencies. However, it was also revealed that the VAR model requires an enormous amount of data and may suffer from overfitting and multicollinearity issues.

An example of a disaggregate model is the discrete choice model, designed to grasp the inclinations and decisions of individual travellers or shippers based on the utility or appeal of each available option. In their work, Zhu et al. (2019) utilized a nested logit model to predict passenger demand for high-speed rail and air transport in China, utilizing survey data from 2015. They took into account four attributes for each mode: travel time, travel cost, service frequency, and comfort level. The researchers determined that the nested logit model effectively captured the diversity and interconnectedness of preferences among various traveller segments, such as business and leisure. However, they acknowledged that the discrete choice model is unanticipated on the presence and quality of survey data and may not entirely mirror the real behaviour and fluctuations in the market.

2.2 Time Series Models

Another widely utilized method for forecasting air transport demand is time series models, which prove effective in capturing temporal patterns and trends within the demand, including seasonality, cyclicity, and irregularities. These models are typically categorized into two types: univariate and multivariate. Univariate models solely rely on the historical data of the demand itself, whereas multivariate models incorporate historical data from other correlated variables in addition to the demand.

One of the most widely used univariate models is the ARIMA model, it can capture linear and stationary relationships among past, present, and future values of demand. In the study Madhavan et al. (2023) both ARIMA and BSTS models were used to forecast a decade's worth of data (2009–2018) for air passenger and freight demand in the Indian aviation industry. It was found that the ARIMA model was useful in capturing the seasonal variations and multiplicative nature of the air passenger data. Along with ARIMA, BSTS (Bayesian Structural Time Series) proved suitable for short-term forecasting of four (international and domestic passenger, international and domestic air cargo) commercial aviation sectors. Although the research found that ARIMA has good forecast accuracy, however, it was less accurate in accommodating the uncertainty aspects, which may include over-fitting risks.

The study Jamil and Akbar (2017) forecasts hotspots of passengers using an Automatic ARIMA Model for conducting time-series analysis based on spatio-temporal data derived from Bandung's local taxi company. The Automatic ARIMA model, applied in this research, reveals its capability to accurately predict spatio-temporal time-series data, particularly when dealing with substantial datasets. The Mean Absolute Scaled Error (MASE) values of 0.8797 for the New York City dataset and 0.6338 for the Bandung dataset support the findings. However, the reliability of the analysis lowers when the actual demand approaches zero. According to the research, this is attributed to data quality issues rather than deficiencies in the prediction model. Another common univariate model

is the exponential smoothing model, which can capture the level, trend, and seasonality of the demand, by assigning different weights to the past observations. For the first time, Dantas et al. (2017) introduced a novel forecasting approach, Bagging Holt-Winters, combining Bagging and the Holt-Winters exponential smoothing method to predict future air transportation demand. Empirical results demonstrate consistently high forecast accuracy, outperforming established benchmarks like SARIMA, Holt-Winters, ETS, Seasonal Naive, and Bagged.BLD.MBB.ETS across 13 of 14 time series. The method's simplicity, parallelizability, and reduced sMAPE highlight its significance as an effective forecasting tool for the aviation sector. Additionally, the proposed methodology offered two notable advantages, that is, it is straightforward to implement, and despite potential slowness in Bagging due to Bootstrap replicates, the method is easily parallelizable. These outcomes underscore the significance of the Bagging Holt-Winters methodology as a valuable forecasting tool for the air transportation industry.
A typical multivariate model is the seasonal ARIMA or SARIMA model, which can capture the linear and stationary relationships among the demand and other variables, as well as the seasonal patterns and trends. For example Nai et al. (2017) used the SARIMA model to predict passenger demand at a regional airport in China where monthly air passenger data spanning from 2004 to 2015 was used. This model incorporated four external variables, namely GDP, population, income, and price. The study revealed that the SARIMA model enhanced the predictive accuracy compared to the ARIMA model, particularly during peak and off-peak seasons. Nevertheless, the researchers cautioned that the SARIMA model demands meticulous parameter selection and might not effectively address the non-stationary and non-linear attributes inherent in demand patterns.

2.3 Machine Learning Models

Machine learning models represent a modern and dynamic approach to predicting air transport demand. They possess the ability to comprehend complex and non-linear relationships between demand and explanatory variables, revealing hidden patterns in the data. These models are broadly categorized as supervised and unsupervised. Supervised models leverage labelled data, incorporating historical demand and explanatory variables, while unsupervised models work with unlabeled data, handling raw images and texts. In the context of forecasting air transport demand, supervised models are commonly favoured.
One of the popular supervised models is the Artificial Neural Network (ANN) model, which can approximate any nonlinear function by using a network of interconnected nodes and layers. For example, Baxter and Srisaeng (2018) proposes using an Artificial Neural Network (ANN) to predict Australia's annual export air cargo demand from 1993 to 2016. The ANN model incorporates input parameters like world real merchandise exports, global population growth, world jet fuel prices, air cargo yields worldwide, outbound flights from Australia, and the Australian/United States dollar exchange rate. 2003 and 2015 cargo demands dummy variables address the cyclical fluctuations. The multi-layer perceptron (MLP) architecture yields a high R-value of 0.97, indicating strong predictive capability. Key variables impacting demand include global population growth, global air cargo yields, and the inclusion of a dummy variable for fluctuations in 2015. The results highlight the interconnectedness between Australia's export air cargo demand, global population growth, and world trade volumes. However, challenges such as data requirements and the risk of overfitting and underfitting are noted in the ANN

model.
The utilization of artificial intelligence (AI) methods in predicting air cargo demand has witnessed a growing trend in recent years primarily attributed to their robust capabilities in nonlinear mapping and autonomous learning from historical data. A notable instance of this application was found in Sun et al. (2019), where a sophisticated nonlinear vector auto-regression (NVAR) neural network method was employed for forecasting air passenger flow at Beijing International Airport. The results of this study underscored the effectiveness of NVAR neural networks in handling the inherent fluctuations present in air passenger flow data. When compared to alternative forecasting methods, the NVAR exhibited superior performance due to its stability, as well as its accuracy in both directional forecasting and forecasting at the level of detail.

2.4 Hybrid Models

Artificial Neural Network (ANN) and Support Vector Machine (SVM) stand out as the most commonly employed hybrid AI models, often fine-tuned through evolutionary algorithms for aviation forecasting as indicated by Kern et al. (2004). Alekseev and Seixas (2009) employed a back-propagation neural network, to predict air passenger throughput in Brazil by decomposing raw time series into multiple input variables. This method enhanced the forecasting performance. While artificial neural networks (ANN) often outperform traditional models, they have drawbacks like susceptibility to local minima, sensitivity to parameters, over-fitting, and structural definition challenges. In contrast, support vector machines (SVM) are rarely used independently for statistical indicator forecasting. The common approach involved creating a hybrid model to capture linear and non-linear patterns separately and then integrating them for the final forecast. For instance, Xie et al. (2014) devised a hybrid method integrating seasonal decomposition and the Least Squares Support Vector Regression (LSSVR) to forecast air passenger traffic at Hong Kong International Airport. They segmented the raw data into trend-cycle, seasonal factor, and irregular components through x-12-ARIMA. These components were individually forecasted and then combined through additive or multiplicative methods to yield the ultimate result. The authors advocated for seasonal decomposition as an effective means to forecast air passenger traffic. In Xu et al. (2019) study, the proposed SARIMA-SVR model demonstrated superior accuracy compared to alternative methods, reinforcing the idea that the inclusion of Gaussian White Noise contributes to enhanced forecasting precision.

In conclusion, the literature review highlights the critical role of forecasting in the air transport sector for effective planning of capacity across airports. Econometric model like the gravity model, offers valuable insights into the causal relationships between demand and socioeconomic factors. Time series models, particularly ARIMA and exponential smoothing, prove effective in capturing temporal patterns. The study on ARIMA and BSTS demonstrates their usefulness but highlights the challenge of addressing uncertainty aspects, including overfitting risks. Furthermore, because econometric models often assume stable relationships between time series, it is not feasible to forecast the complex and nonlinear air cargo series Li et al. (2020). The time series approach emerges as a viable option for forecasting air passenger traffic because of the intricacies involved in selecting and testing variables in econometric modelling. However, its limitation lies in often being nonlinear and nonstationary and its incapacity to distinctly identify the

reasons behind the growth in air passenger traffic Bao et al. (2012).
Artificial Neural Network (ANN) model on the other hand showcases reliability in capturing nonlinear relationships. Despite having high predictive capability, certain challenges like the need for large and diverse data and the risk of overfitting possess potential drawbacks. Hybrid models such as ANN and SVM, display improved forecasting performance, but the sensitivity to parameter selection, and the necessity of defining network structure raise challenges in association with ANN. Overall, critical analysis of the related studies suggests that no single model is universally superior and the choice of model selection depends on the specific context, data availability, and forecasting requirements. This study will focus on time series forecasting using ARIMA, SARIMA, autoARIMA and Exponential Smoothing of air passenger and freight and mail across Ireland and partner countries.

3 Methodology

This study utilises the Knowledge Discovery in Databases (KDD) methodology to achieve the aim of forecasting air transport data for Ireland and partner countries. Figure 1 displays the methodology steps involved in this study which are described below:

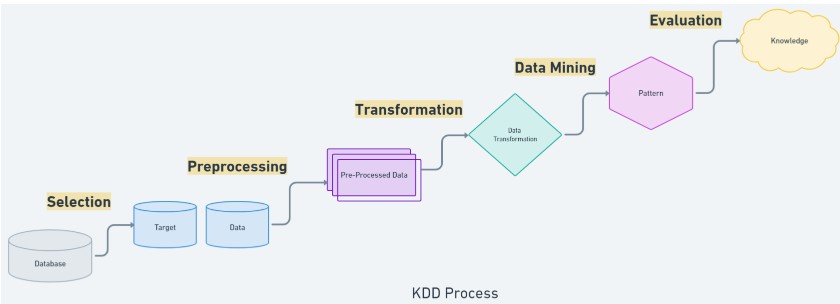

Figure 1: Methodology for Air Transport Demand Forecast

3.1 Data Collection

Step one involves the data collection process which forms the foundational step for subsequent predictive analysis of passenger and freight demand. The utilization of robust passenger Eurostat (2023a) and freight Eurostat (2023b) datasets sourced from Eurostat's air transport data section enhances the methodological integrity of the study. These publicly available datasets, serve as a foundation for conducting in-depth analysis for obtaining metrics to identify the best prediction methods.

3.2 Data Pre-Processing

Following the initial data collection, the second pivotal step involves a comprehensive process of data preprocessing, this stage significantly influences the accuracy and effectiveness of subsequent forecasting. Loaded datasets are cleaned and prepared to ensure their quality and reliability in the modelling phase. To facilitate the accurate temporal alignment which is essential for time-series forecasting, variables from the dataframe are selected to convert into datetime format. Subsequent measures involve sorting, as well as the determining of unique values along with the length of the dataframe, which is essential for gaining insights into the data structure. Addressing null values is an important step of data preprocessing, which aims to eliminate potential outliers that could introduce deviations and compromise the reliability of predicted values. By systematically handling missing data, the preprocessing stage enhances the robustness of the dataset. An additional step involves the creation of separate dataframes for both passenger and freight individually to store the top three airports grouped values and target variables for further time series forecasting. This approach to data preprocessing ensures that the dataset is refined, organized, and free from irregularities. These steps help to build a groundwork for the development of reliable time series predictive models.

3.3 Data Transformation

Once the data is preprocessed, the third step involves data transforming into a format that data mining techniques can analyze. The aggregated data's temporal trends are visualized through a line plot, which provides an overview of the passenger and freight selected airports' overall patterns. Using the additive and multiplicative decomposition of time series data, each dataset is deconstructed into its fundamental components—trend, seasonality, and residuals. This provides an understanding of inherent patterns present in the data and contributes to improved data modelling, anomaly detection, and forecasting accuracy. A critical stationarity test is executed on both datasets utilizing the Dickey-Fuller test. This test serves as a pivotal tool in determining the stationarity of time series data by assessing the presence of a unit root. The obtained p-value is compared against a significance level of 0.05. Rejection of the null hypothesis and achieving a p-value below the significance level indicates that the time series is stationary. A more negative test statistic value proves in favour of stationarity. The comparison of the test statistic value with critical values further informs the stationarity determination. Diverse data transformation methods are applied, including differencing, log transformation, and moving averages, each contributing to attaining stationarity through distinct mechanisms. Differencing involves determining the variance between successive data points, while log transformation stabilizes variance. The application of moving averages entails computing means within defined windows of data points. the inversion of transformed data must be performed to restore accurate values for subsequent data forecasting. The effectiveness of each transformation method is assessed through the Dickey-Fuller test, ensuring the selection of the most suitable approach for achieving stationarity. Throughout this process, seasonal decomposition plots offer valuable insights into the impact of each transformation method on the data's underlying patterns. Data transformation establishes a foundation for further analysis of data and helps in achieving the research objectives.

3.4 Data Mining

The fourth step is to build and fit time series models for both passenger and freight data to generate forecasts. Optimal model parameters are established by generating Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) graphs, which help in the identification of lag values crucial for model accuracy. This study adopts an approach with an attempt at auto ARIMA, allowing for the automatic exploration of the best parameters suitable for model fitting, denoted as (p, d, q) and (P, D, Q).

The common obstacle in ARIMA forecasting is the parameter selection or order value determination, which is overcome by implementing the automatic ARIMA model proposed by one of the studies Hyndman and Khandakar (2008) as well, as it determines the order value automatically. Following parameter determination, the time series data undergoes training and testing data split. The training set spans from 2020 to 2022, while the subsequent 6 months of 2023 are reserved for testing the models. The auto ARIMA model is then fitted to the training data, providing a summary that contains configuration details and performance metrics. Further, diagnostic plots are generated and scrutinized to assess the model's validity, identifying any potential issues that may impact its reliability. Next, a crucial step is employed to generate forecasts for the specified number of periods determined by the length of the test data. For assessing the model's performance and generalization to unseen data predictions on the test data are performed. Performance of evaluation metrics becomes important with a focus on the model's accuracy, reliability, and ability to capture underlying patterns and variations. This is achieved through a comparison of the obtained predictions to the actual values within the test data. The outcome of this comparative analysis tells about the model's efficacy in generalizing to unseen data. Three additional time series models are applied to obtain predictions of passenger and freight predicted values, these are, Simple Exponential Smoothing (SES), Holt's Linear Exponential Smoothing (HL), and Holt-Winters' Exponential Smoothing (HW). Each model undergoes a similar process, contributing to a comprehensive evaluation of their predictive capabilities in forecasting both passenger and freight traffic.

3.5 Data Evaluation

Step five involves the interpretation and evaluation of the results obtained from models and forecasts. The data evaluation part is a crucial step in checking the robustness of the time series forecasting models. The evaluation metrics calculated for comparison are the Mean Absolute Error (MAE), Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). These metrics help in identifying how close the predicted values are to the actual data. All four forecasts (ARIMA/SARIMA, Simple Exponential Smoothing, Holt's Linear and Holt-Winters) are shown together with the actual data on a line graph for ease of visual comparison. The generated graph reveals the extent to which models capture the patterns in the data. The evaluation metrics table helps in understanding the best forecasting model and the subsequent model's predicted value is also displayed. The reliability and accuracy of these models are calculated based on obtained evaluation metrics and visual graph representation.

3.6 Knowledge Representation

This final step involves representing the knowledge extracted from the passenger and freight data in a way that airports and airlines can comprehend. This is done through visualizations of different plots showcasing the maximum and least traffic flow from Ireland useful for allocating resources efficiently and selecting forecasting models for future planning of passenger and freight demand fluctuations based on obtained evaluation metrics. Knowledge representation can be understood further from the evaluation part and the discussions mentioned.

4 Design Specification

High demand for passenger or freight in selected airports can lead to uneven resource allocation within the airport and the aviation sector constantly adapts to new models and solutions that can ease the distress and provide smooth operations to the passenger travelling 1. Therefore, this research is proposed to predict the demand across passenger and freight traffic in Ireland and partner airports and obtain metrics which help to analyse the best-fit model for those accurate predictions. The data that has been used for the current research is obtained from Eurostat's air transport public dataset. The research was conducted using the Python 3.10.9 programming language within the Jupyter Notebook 6.5.4 environment. Python's versatility, extensive libraries, and ecosystem made it an ideal choice for implementing and analyzing complex algorithms in the context of time series forecast. Jupyter Notebook, with its interactive and dynamic nature, provided a well-suited platform for developing, documenting, and presenting the code, facilitating a transparent and reproducible research process. The combination of Python and Jupyter Notebook empowers this study with the flexibility and efficiency needed to explore, model, and interpret intricate datasets related to air transport demand forecasts. Data loading, pre-processing, analysing and visualising are performed on the aforementioned. For the implementation essential Python libraries were used, such as Pandas and NumPy for data handling and manipulation, Matplotlib and Seaborn for creating visualizations, Statsmodels for time series, ACF and PACF analysis and Adfuller from same to test stationarity, PMDARIMA provided an automated parameter and model selection, Scikit-Learn for evaluating performance of models based on metrics (MSE, RMSE, MAE). Lastly, the Tabulate library gave a structured format of the predicted values and evaluation metrics for easy understanding of the outcomes.

Post preprocessing of passenger and freight data, the Dickey-Fuller test assesses time series stationarity. Rejecting the null hypothesis relies on a p-value below 0.05 and a more negative test statistic. Various data transformations, including differencing, log transformation, and moving averages, are tested, and the best Dickey-Fuller values determine optimal stationarity. Differencing captures successive data distinctions, log transformation stabilizes variance, and moving averages compute means within defined data windows. Reverting transformed data is crucial for accurate forecasting. Systematic illustration, stationarity tests, and seasonal decomposition plots aid in evaluating each transformation's impact. Next, the AutoCorrelation Function (ACF) and the Partial AutoCorrelation Function (PACF) plots are generated to identify lag values, which are crucial for configuring model parameters and understanding the temporal relationships within the data. Following this initial exploration, the dataset is split into training set spanning from 2020 to 2022 and testing set covering the subsequent 6 months of 2023. Time series forecasting models used in this study are ARIMA/SARIMA, Simple Exponential Smoothing, Double Exponential Smoothing (Holt's Linear) and Triple Exponential Smoothing (Holt-Winters). These models have been selected with due consideration of previous work performed using them. As discussed in related work section2 under time series models, it was inferred that the seasonal variations were captured well by the ARIMA model. For implementation purposes on both the datasets of passenger and freight, the auto ARIMA model is used which generates the best forecasting model. This not only gave better results but also eased out the process of implementing multiple models to find the best one. Exponential smoothing has also been used in previous studies to forecast passenger demand.

After splitting the data into training and testing, first, an Auto ARIMA model is chosen, and the model is subjected to a detailed examination through diagnostic plots, providing insights into its configuration and performance metrics. The next step involves generating forecasts based on the trained models and assessing their accuracy and reliability. This evaluation is carried out by comparing the model's predictions with the actual values in the test set. To quantify the model's performance, three key evaluation metrics are calculated: Mean Absolute Error (MAE), which tells on average, how far off the predictions are from the actual values, also better accuracy is achieved with a smaller MAE. Next, Mean Squared Error (MSE), emphasizes larger errors even more than MAE, it calculates the average of squared differences between the predicted and the actual value. Lastly, Root Mean Squared Error (RMSE) is useful for undermining the typical size of errors and it provides an easily interpretable metric as the target variable has the same unit. These metrics offer a comprehensive understanding of how well the model captures patterns and variations in the data. This entire process is then repeated for three other forecasting models: Simple Exponential Smoothing(SES) which assigns decreasing weights to past observations exponentially, assuming a constant level. Holt's Linear or Double Exponential Smoothing(HL) extends Simple Exponential Smoothing to accommodate data with a linear trend by incorporating two smoothing parameters for both the level and trend components. Holt-Winters' or Triple Exponential Smoothing(HW) further enhances Double Exponential Smoothing by incorporating a seasonality component, making it suitable for time series data with both trend and seasonality. Holt-Winters model uses three smoothing parameters for level, trend, and seasonality and is hence able to capture complex patterns in the data. Each model undergoes the same rigorous evaluation process, and the evaluation metrics results are displayed. Each model serves specific requirements such as SES handles the basic time series, DES addresses the trends, and TES accommodates both trends and seasonality. Additionally, line graphs are plotted to visualize the forecasts generated by each model against the actual data. This section thus provides an overview of systematically following these steps and thoroughly evaluating each model's predictions using standardized metrics for assessing the effectiveness of various time series forecasting methods on passenger and freight traffic data. The graphical representations and numerical metrics both facilitate the objectives of this research.

5 Implementation

Passenger and freight demand forecast for Ireland and partner airports is obtained by using four different time series models. Using the passenger[5] and freight[6] datasets obtained from Eurostat in (.csv) format covering monthly statistical data from January 2020 to June 2023, two dataframes are created to load them with the number of passengers (in millions) travelling and the number of freight and mail (in 1000 tonnes) carried from Ireland to each reporting country. The following Table 1 describes the column headings present in the dataframes. After loading, the time period column values in both datasets are converted to datetime format with the help of the inbuilt pandas function. This is to ensure correct temporal alignment with the time-series forecasting. Sorted values for

[5]https://ec.europa.eu/eurostat/databrowser/view/avia_par_ie__custom_8975440/default/table?lang=en
[6]https://ec.europa.eu/eurostat/databrowser/view/avia_gor_ie__custom_8975476/default/table?lang=en

Table 1: Passenger and freight and Mail Dataset Description

Coloumns	**Description**
DATAFLOW	Data flow information
LAST UPDATE	The last update time of the data
freq	Frequency of the data
unit	Unit of measurement
tra_meas	Transport measure
airp_pr	Partner airport
TIME_PERIOD	Time period of the data
OBS_VALUE	Observed value
OBS_FLAG	Observation flag

both dataframes, count of unique values and length of dataframe are examined thereafter. To forecast the passenger and freight demand values, target variables, which are, partner airport and observed values (passenger and freight count) are selected about the time period variable. The highest number of passenger and freight flows from Ireland are identified using the inbuilt value counts Python function. Further, the null values are checked and due to no relevance to the forecast the observation flag column, which showed the presence of null values is ignored.
Further, the total passenger and cargo traffic generation for Ireland and partner airports is analyzed through Figure 2 line plot from 2020 to 2023. The first figure displays the total count of passengers who travelled from Ireland to other airports, while the second figure displays the total count of freight carried from Ireland to other airports. Newly

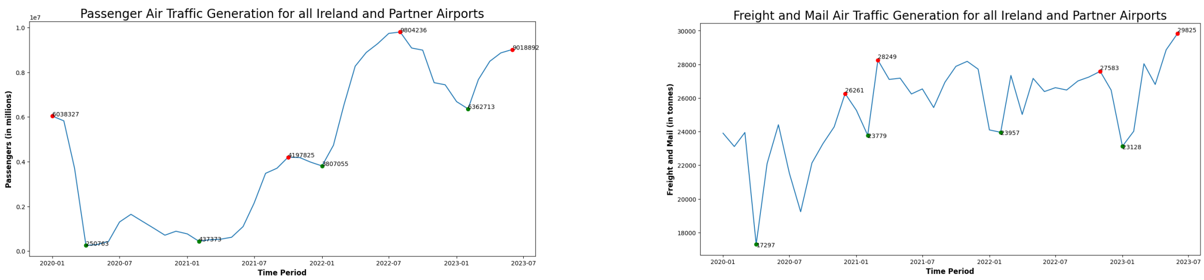

Figure 2: Total count of passenger and freight traffic flow from Ireland to partner airports

created dataframes store the value of the top three airports having the highest passenger and freight traffic. Further, the top three airport pairs having maximum and minimum passenger and freight flow are calculated. The calculation of the least passenger and freight carrying airport pairs' shares and mean of values is essential for airports and airlines. This helps them to make decisions for seizing operations on low revenue generation routes or diverting them for effective outcomes. A line graph created next displays the share of passenger and freight traffic contributed by the top three airports over three years. Passenger and freight data is aggregated over time periods and the total count is visualized through a line plot which gives an overview of the overall trend and pattern for the selected airports. Using the additive and multiplicative decomposition of time series data is decomposed into its three components—trend, seasonality, and residuals. This helps in understanding patterns present in data, thus improving data modelling, detecting anomalies and forecasting.Figure 3 displays the multiplicative decomposition of time series for passenger traffic and Figure 4 displays freight and mail traffic. After

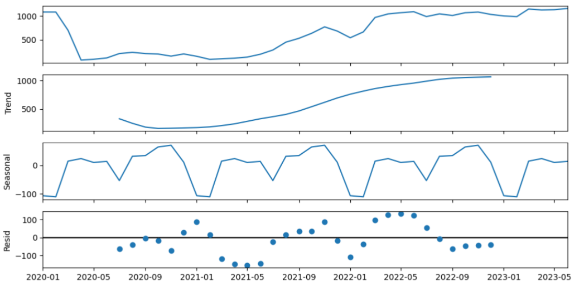

Figure 3: Decomposition of passenger traffic

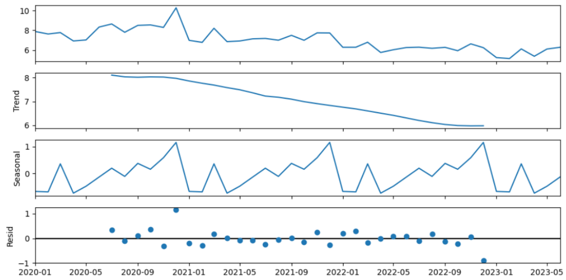

Figure 4: Decomposition of freight traffic

checking the stationarity of the datasets across different transformations and verifying the results using the Dickey-Fuller test, the time series data is split into training and testing sets. Data from 2020 to 2022 is used for training, and the subsequent 6 months of 2023 are used for testing the models for passenger and freight demand forecast. The auto ARIMA model is fitted to the training data and it provides a comprehensive summary, including configuration details and performance metrics. Diagnostic plots are generated and displayed to assess the model's validity and identify potential issues. Next, a crucial step is employed to generate forecasts for the specified number of periods determined by the length of the test data. For assessing the model's performance and generalization to unseen data predictions on the test data are performed. By comparing these obtained predictions to the actual values in the test set, the model's accuracy, reliability, and ability to capture patterns and variations can be assessed. Following the same steps, Simple Exponential Smoothing (SES), Holt's Linear Exponential Smoothing (HL), and Holt-Winters' Exponential Smoothing (HW) forecasting models are applied to predict passenger and freight traffic. The evaluation metrics obtained assess the accuracy of the time series model's predictions. In the end, all four forecasts are plotted against actual data in a line graph for visual comparison of their performance. An evaluation metrics table helps in understanding the best forecasting model and the subsequent model's predicted value is also displayed. The reliability and accuracy of these models are calculated based on obtained evaluation metrics and visual graph representation. The evaluation section next discusses the results obtained in detail for predicting air passenger and freight demand.

6 Evaluation

Two types of experiments were conducted in this study for Irish airports, that is, passenger demand forecast and freight and mail demand forecast. The research aimed to predict maximum passenger and freight flow across Ireland and partner airports. For ease of forecasting, the top three airport pairs were selected for predicting passenger and freight demand. The best time series forecasting model selection based on the obtained evaluation metrics was essential for fulfilling research objectives. The following sub-sections discuss the experiments in detail.

6.1 Passenger Demand Forecast

After preprocessing the passenger dataset and using the value counts Python inbuilt function to find the airport pair with maximum passenger flow, it was revealed that Dublin, Ireland airport to Heathrow, London airport had a share of 39.2%. The share of each airport is calculated and plotted using a line graph. Table 2 has the top three partner airports with the maximum passenger traffic share (in %) along with the mean number of passengers per year. Each airport's traffic share in the overall total is computed in a table. Figure 5 displays the top 3 airports with maximum passenger traffic. Next, the

Table 2: Top 3 airport pairs with maximum passenger traffic

Airport Pair	**Mean Pax/Year**	**Share**
Dublin - London Heathrow	249601	39.2
Dublin - London Gatwick	196773	30.9
Dublin - Amsterdam/Schiphol	190658	29.9

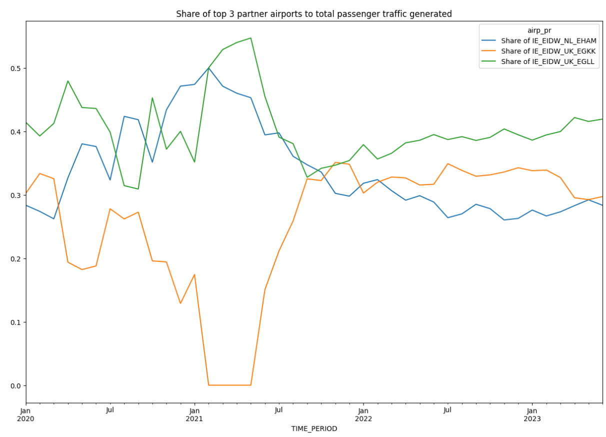

Figure 5: DUBLIN-HEATHROW airport shown in Green has top shares for total passengers generated.

stationarity test was performed on a new dataframe containing only the top 3 airports' passenger traffic. The examination of stationarity in the passenger time series data was conducted through the application of the Dickey-Fuller test. The acquired test statistic, registering at -1.62, underwent comparison with critical values corresponding to various significance levels. The associated p-value, determined to be 0.46, was juxtaposed with a critical value of -2.93 at a 5% significance level. Given that the test statistic exceeds this critical value and the p-value surpasses the significance threshold, the null hypothesis cannot be rejected. hence, the data was non-stationary. Furthermore, the test incorporated 1 lag and drew insights from 40 observations. Cumulatively, these findings imply that the transformation of data was required to establish stationarity before the application of time series models. Next, to convert the non-stationary data to stationary, first-order differencing is performed. The stationarity test gave a test statistic value of -7.351515e+00 and the p-value dropped significantly at 1.003112e-10. Therefore, the test statistic significantly deviating from the critical values at 1%, 5% and 10% suggests a rejection of the null hypothesis and establishes the presence of stationarity in the time series data, validating its suitability for time series modelling and analysis. Next, forecasting models are implemented and the auto ARIMA model reveals a configuration of SARIMAX(2, 1, 0), representing a seasonal autoregressive integrated moving average model with two autoregressive lags and a first-order differencing to achieve stationarity. The coefficients for the autoregressive terms, ar.L1 and ar.L2, are 0.5310 and -0.3198, respectively. These values suggest a positive impact of the first lag and a negative impact of the second lag on the dependent variable. The significant p-values associated with these coefficients (both ¡ 0.05) support their relevance in the model. The estimated sigma2 of 1.749e+04 indicates the variance of the model's error term. The Ljung-Box test reveals a low Q statistic (0.01) with a high p-value (0.94), suggesting that the model adequately captures temporal patterns. However, the Jarque-Bera test indicates non-normality in the residuals, emphasizing the need for caution in interpreting the model. The Heteroskedasticity test yields a p-value of 0.05, suggesting potential heteroskedasticity. The evaluation metrics obtained from the ARIMA/SARIMA, SES, DES and TES forecasting models and the model comparison graph are analysed in the upcoming discussion section.

6.2 Freight and Mail Demand Forecast

For freight and mail traffic it was revealed that Dublin, Ireland airport to Cologne Bonn, Germany airport had a share of 37.8%. The share of each airport is calculated and plotted using a line graph. Table 3 has the top three Ireland and partner airports with the maximum freight and mail traffic share (in %) along with the mean number of freight per year. Each airport's traffic share in the overall total is computed in a table. Next, Figure 6 displays the top 3 airports having maximum freight and mail traffic from Ireland to partner countries. Next, the stationarity test was performed on a new dataframe containing only the top 3 airports' freight traffic. The acquired test statistic, registering at -1.03, underwent comparison with critical values corresponding to various significance levels. The associated p-value, determined to be 0.74, was juxtaposed with a critical value of -2.93 at a 5% significance level. Given that the test statistic exceeds this critical value and the p-value surpasses the significance threshold, the null hypothesis cannot be rejected. hence, the data was non-stationary. Furthermore, the test incorporated 2 lags and drew insights from 39 observations. Cumulatively, these findings imply that the transformation of data was required to establish stationarity before the application of time series models. Next, to convert the non-stationary data to stationary, first-order differencing is performed. The stationarity test gave a test statistic value of -7.351515e+00 and the p-value dropped significantly at 1.003112e-10. Therefore, the test statistic significantly deviating from the critical values at 1%, 5% and 10% suggests a rejection of the null hypothesis and establishes the presence of stationarity in the time series data, validating its suitability for time series modelling and analysis. Next, forecasting models are implemented and the auto ARIMA model reveals a model configuration of SARIMAX(2, 1, 0)x(1,

Table 3: Top 3 airport pairs with maximum freight traffic

Airport Pair	**Mean Freight/Year**	**Share**
Dublin - Cologne Bonn	2645	37.8
Dublin - London Stansted	2212	31.6
Dublin - East Midlands	2139	30.6

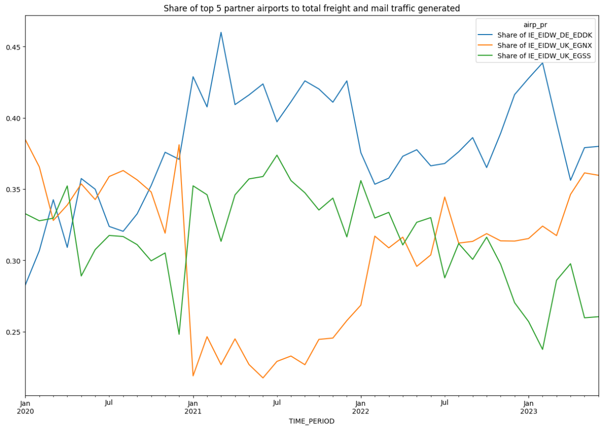

Figure 6: DUBLIN-KOELN/BONN airport shown in Blue has top shares for total freight generated.

0, 0, 12). This suggests a seasonal autoregressive integrated moving average model with two autoregressive lags, one seasonal autoregressive lag with a 12-month seasonality, and a first-order differencing required for achieving stationarity. The coefficients associated with the autoregressive terms (ar.L1 and ar.L2) and the seasonal autoregressive term (ar.S.L12) indicate their impact on the dependent variable. The negative coefficient for ar.L1 and ar.L2, significant at a 5% level, suggests a negative influence on the variable. Additionally, the positive coefficient for ar.S.L12, also significant, implies a positive seasonal impact. The diagnostic statistics such as the Ljung-Box test and Jarque-Bera test provide insights into the model's goodness-of-fit, suggesting adequacy for time series forecasting. The small p-values associated with these tests indicate that the model adequately captures the temporal patterns and exhibits normality. Moreover, the Heteroskedasticity test indicates homoskedasticity, further validating the model's reliability. The evaluation metrics obtained from the SARIMA, SES, DES and TES forecasting models and the graph displaying a comparison of them are analysed in the upcoming discussion section.

6.3 Discussion

This section discusses the findings from the above two experiments. For freight and mail peak demand forecast, models are plotted together to compare the best out of them. Figure 6.3 compares the performance of four forecasting models against actual data over six months of 2023. The models compared are ARIMA/SARIMA, Simple, Double and Triple Exponential Smoothing. In both graphs, the x-axis represents the time period from January 2023 to June 2023 and the y-axis represents the passenger (in millions) and freight (in 1000 tonnes) count. The blue line in both graphs represents the actual data, while other lines represent the forecasted values from four models. The corresponding line to the model is mentioned in the index. In the first graph for passenger predictions, the actual data displays a fluctuating trend with a peak around March 2023. Holt's Linear model follows the actual data relatively closely, suggesting that the seasonality and trend are captured effectively. The ARIMA/SARIMA model also appears to track the actual data well during the first couple of months, indicating that underlying patterns in the data were being captured. SES and Holt-Winters models do not capture the fluctuations as effectively, this may be due to the incompetence of these models to handle seasonal-

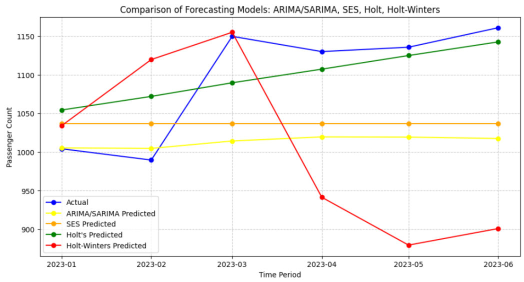

Figure 7: Passenger Data

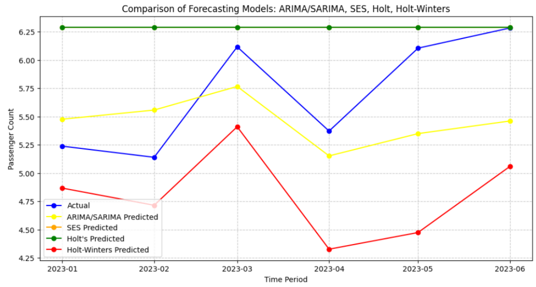

Figure 8: Freight and Mail Data

ity, which is present in the data. In the second graph for freight predictions, the actual data shows a sharp increase in freight count in March, followed by a decrease and then a steady increase in values. Here, the Holt-Winters model tracks the actual data in trend and seasonality peaks, however, lacks in predicting values close to the actual data. The SARIMA model does not follow the sharp increase reliably, which might suggest that the seasonal component is not able to fully capture the actual data. SES and Holt's models do not capture the trend or seasonality at all, which is indicated by their relatively flat lines compared to the actual data.
Figure 6.3 displays results of all evaluation metrics of time series models applied for maximum passenger and freight forecast. For passenger demand, the Holt Linear's model ex-

Model	MAE	MSE	RMSE
ARIMA	87.05	10831.5	104.07
SES	84.96	8356	91.41
Holt Linear	40.79	2318.7	48.15
Holt-Winters	145.13	31159.4	176.52

The best model is: Holt Linear with RMSE: 48.15

Figure 9: Passenger Evaluation metrics

Model	MAE	MSE	RMSE
SARIMA	0.47	0.27	0.52
SES	0.58	0.56	0.75
Holt Linear	0.58	0.56	0.75
Holt-Winters	0.9	1.01	1.01

The best model is: SARIMA with RMSE: 0.52

Figure 10: Freight and Mail Evaluation metrics

hibited the lowest MAE(40.79), MSE(2318.7) and RMSE(48.15), indicating that it is the best-performing model. The ARIMA model has a high MAE (87.05), indicating a notable deviation between predicted and actual values. Additionally, the ARIMA model demonstrated a substantial MSE (10831.5) and a corresponding RMSE (104.07), highlighting the model's challenges in accurately capturing the variability in passenger demand. In comparison, the SES model has lower MAE (84.96) and MSE values, suggesting improved accuracy. The Holt-Winters model, however, exhibited the highest MAE (145.13), MSE (31159.4), and RMSE (176.52), which reveals a significant discrepancy between forecasted and observed passenger demand. These results collectively suggest that the Holt Linear model performed relatively well among the evaluated models, providing more accurate forecasts for passenger demand.
In the study Nai et al. (2017) SARIMA model outperformed for passenger forecast, however, the SARIMA model emerged as the best-performing model in this study for freight demand, with the lowest evaluation metrics. This indicates that the SARIMA model holds a superior ability to predict demand accurately. In close pursuit are the SES (Simple Exponential Smoothing), Holt Linear or DES, and Holt-Winters or TES models. Although SES, Holt Linear, and Holt-Winters show comparable MAE and MSE metrics, indicating

similar accuracy levels, the Holt-Winters model exhibits a slightly higher RMSE (1.01), suggesting that it is less precise in capturing the fluctuations. These evaluation metrics collectively favour the SARIMA model as the optimal choice for forecasting freight demand, as it consistently outperforms its counterparts.
Figure 6.3 displays the maximum passenger and freight predicted values for 2023. Based

Time Period	HL Predicted
2023-01-01	1054.408214
2023-02-01	1072.056282
2023-03-01	1089.704351
2023-04-01	1107.352419
2023-05-01	1125.000488
2023-06-01	1142.648556

Figure 11: Passenger Forecast (millions)

Time Period	SARIMA Predicted
2023-01-01	5.476169
2023-02-01	5.557281
2023-03-01	5.765798
2023-04-01	5.151896
2023-05-01	5.348787
2023-06-01	5.460460

Figure 12: Freight Forecast (1000 tonnes)

on Holt's Linear passenger predicted values for January 2023 to June 2023 show that June 2023 has the maximum prediction of passengers which is approximately 1142.64 (in millions). Whereas, based on SARIMA freight predicted values, March 2023 has the maximum prediction of freight traffic which is approximately 5.76 (in 1000 tonnes). These values give an insight into the data that fluctuating demands could arise from these partner airports for future travel season and pre-planning for tackling such sudden demand would keep airlines and airports from crashing under pressure 1. Concerning the objectives of this research maximum passenger and freight traffic demand has been computed, as well as best forecasting model based on calculated evaluation metrics is present. Further visual comparisons of the forecasted and actual values provide an analysis of each model's adaptation to fluctuations.

7 Conclusion and Future Work

The primary objective of this research was to predict the best time series forecasting model based on the evaluation metrics that adapt to highly fluctuating passenger and freight demands in the aviation sector. The study successfully addressed this objective by evaluating and comparing the performance of four time series forecasting models, ARIMA/SARIMA, simple Exponential Smoothing (SES), Double Exponential Smoothing (DES) or Holt's Linear, and Triple Exponential Smoothing (TES) or Holt-Winters. The research applied rigorous evaluation metrics (MAE, MSE, RMSE) to assess the accuracy of each forecasting model. The key findings indicate that the Holt Linear model demonstrated superior performance, achieving the lowest error metrics among other evaluated models for passenger demand forecasting. Conversely, for freight demand forecast, the SARIMA model emerged as the optimal choice, consistently outperforming other models in terms of accuracy. The success of this study is evident in its ability to quantitatively measure forecast accuracy and gauge the selection of appropriate models based on performance metrics. The findings contribute to bridging the existing gap in the literature regarding time series forecasting for air transport demand, especially concerning the Irish aviation sector. The research highlights the critical role of accurate forecasts in planning and optimizing passenger and freight capacity across airports. While the research has proven successful in achieving its objectives, certain limitations should be acknowledged, such as addressing uncertainty aspects, of forecasting models remain a challenge, and the incapacity of econometric models to vividly identify the reasons behind air passenger traffic growth adds complexity to the forecasting task.

Future work is recommended to further fine-tune the forecasting methodologies. Exploring the hybrid models or machine learning approaches could bring forward solutions to the mentioned limitations. In conclusion, this research provides valuable insights into time series forecasting for air passenger and freight demand, offering a foundation for future advancements in the aviation field. The proposed models and methodologies hold potential for practical implementation, ultimately contributing to more effective planning, decision-making, and a safer and more efficient air transport system.

References

Alekseev, K. and Seixas, J. (2009). A multivariate neural forecasting modeling for air transport–preprocessed by decomposition: a brazilian application, *Journal of Air Transport Management* **15**(5): 212–216.

Bao, Y., Xiong, T., Hu, Z. et al. (2012). Forecasting air passenger traffic by support vector machines with ensemble empirical mode decomposition and slope-based method, *Discrete Dynamics in Nature and Society* **2012**.

Baxter, G. and Srisaeng, P. (2018). The use of an artificial neural network to predict australia's export air cargo demand, *International Journal for Traffic and Transport Engineering* **8**(1): 15–30.

Dantas, T. M., Cyrino Oliveira, F. L. and Varela Repolho, H. M. (2017). Air transportation demand forecast through bagging holt winters methods, *Journal of Air Transport Management* **59**: 116–123.
URL: *https://www.sciencedirect.com/science/article/pii/S0969699716302265*

Eurostat (2023a). Air passenger transport by reporting country and eu partner country. Accessed on 13 December 2023.
URL: *https://ec.europa.eu/eurostat/databrowser/view/avia_par_ie_custom_8975440/default/table?lang=en*

Eurostat (2023b). Freight and mail air transport by reporting country and eu partner country. Accessed on 13 December 2023.
URL: *https://ec.europa.eu/eurostat/databrowser/view/avia_gor_ie_custom_8975476/default/table?lang=en*

Gunter, U. and Zekan, B. (2021). Forecasting air passenger numbers with a gvar model, *Annals of Tourism Research* **89**: 103252.

Hyndman, R. J. and Khandakar, Y. (2008). Automatic time series forecasting: the forecast package for r, *Journal of statistical software* **27**: 1–22.

Jacquillat, A. and Odoni, A. R. (2018). A roadmap toward airport demand and capacity management, *Transportation Research Part A: Policy and Practice* **114**: 168–185.

Jamil, M. S. and Akbar, S. (2017). Taxi passenger hotspot prediction using automatic arima model, *2017 3rd International Conference on Science in Information Technology (ICSITech)*, pp. 23–28.

Kern, S., Müller, S. D., Hansen, N., Büche, D., Ocenasek, J. and Koumoutsakos, P. (2004). Learning probability distributions in continuous evolutionary algorithms–a comparative review, *Natural Computing* **3**: 77–112.

Kitsou, S. P., Koutsoukis, N. S., Chountalas, P. and Rachaniotis, N. P. (2022). International passenger traffic at the hellenic airports: Impact of the covid-19 pandemic and mid-term forecasting, *Aerospace* **9**(3): 143.

Li, H., Bai, J., Cui, X., Li, Y. and Sun, S. (2020). A new secondary decomposition-ensemble approach with cuckoo search optimization for air cargo forecasting, *Applied Soft Computing* **90**: 106161.

Madhavan, M., Sharafuddin, M. A., Piboonrungroj, P. and Yang, C.-C. (2023). Short-term forecasting for airline industry: The case of indian air passenger and air cargo, *Global Business Review* **24**(6): 1145–1179.
URL: *https://doi.org/10.1177/0972150920923316*

Matsumoto, H. and Domae, K. (2018). The effects of new international airports and air-freight integrator's hubs on the mobility of cities in urban hierarchies: A case study in east and southeast asia, *Journal of Air Transport Management* **71**: 160–166.

Nai, W., Liu, L., Wang, S. and Dong, D. (2017). An emd–sarima-based modelling approach for air traffic forecasting, *Algorithms* **10**(4): 139.

Standfuss, T., Fricke, H. and Whittome, M. (2021). Forecasting european air traffic demand - how deviations in traffic affect ans performance, *Transportation Research Procedia* **59**: 105–116. 10th International Conference on Air Transport – INAIR 2021, TOWARDS AVIATION REVIVAL.
URL: *https://www.sciencedirect.com/science/article/pii/S2352146521008632*

Suh, D. Y. and Ryerson, M. S. (2019). Forecast to grow: aviation demand forecasting in an era of demand uncertainty and optimism bias, *Transportation Research Part E: Logistics and Transportation Review* **128**: 400–416.

Sun, S., Lu, H., Tsui, K.-L. and Wang, S. (2019). Nonlinear vector auto-regression neural network for forecasting air passenger flow, *Journal of Air Transport Management* **78**: 54–62.

Xie, G., Wang, S. and Lai, K. K. (2014). Short-term forecasting of air passenger by using hybrid seasonal decomposition and least squares support vector regression approaches, *Journal of Air Transport Management* **37**: 20–26.

Xu, S., Chan, H. K. and Zhang, T. (2019). Forecasting the demand of the aviation industry using hybrid time series sarima-svr approach, *Transportation Research Part E: Logistics and Transportation Review* **122**: 169–180.

Zhu, X., Zhang, G. and Sun, B. (2019). A comprehensive literature review of the demand forecasting methods of emergency resources from the perspective of artificial intelligence, *Natural Hazards* **97**: 65–82.

Zografos, K. G., Madas, M. A. and Androutsopoulos, K. N. (2017). Increasing airport capacity utilisation through optimum slot scheduling: review of current developments and identification of future needs, *Journal of Scheduling* **20**: 3–24.