INSTITUTO
UNIVERSITÁRIO
DE LISBOA

FORECASTING MODEL DEVELOPMENT AND APPLICATION IN THE AVIATION INDUSTRY

Bernardo Maria Marçal Grilo de Sousa Pereira

MSc in Management of Services and Technology

Prof. Sofia Kalakou, Assistant Professor at ISCTE Business School
Department of Marketing, Operation and Management

Prof. José Joaquim Dias Curto, Associate Professor at ISCTE Business School
Department of Quantitative Methods for Management and Economics

November 2021

Department of Marketing, Strategy and Operations

FORECASTING MODEL DEVELOPMENT AND APPLICATION IN THE AVIATION INDUSTRY

Bernardo Maria Marçal Grilo de Sousa Pereira

MSc in Management of Services and Technology

Prof. Sofia Kalakou, Assistant Professor at ISCTE Business School
Department of Marketing, Operation and Management

Prof. José Joaquim Dias Curto, Associate Professor at ISCTE Business School
Department of Quantitative Methods for Management and Economics

November 2021

Acknowledgements

I would like to thank my mother, father, grandparents and girlfriend for the constant support and encouragement through the development of this work.

I would also like to thank my thesis supervisors Dr. Sofia Kalakou and Dr. José Dias Curto the crucial guidance and feedback in every step of this work. Their constant presence was extremely helpful and necessary.

Abstract:

Forecasting models have been applied to many industries as a decision-making tool for over 100 years. Their application in the aviation industry benefits a wide variety of stakeholders such as airliners and airport authorities, who use past data to predict demand and passenger choices so that they can better define fares, manage their fleet and make decisions on the airport layout and future expansions, among others.

The main objective of this dissertation is the development of a forecasting model capable of predicting the number of flight movements at Lisbon Airport. The model was based on an autoregressive model, which uses past data in order to forecast future figures. Weekly data regarding the flight movements at Lisbon Airport was the sample for this study, which was processed through RStudio programming software.

Once the Autoregressive Moving Average (ARIMA) models were defined, the forecasting data was created and further tested for accuracy using extant data. The impact of COVID-19 had to be considered in this situation, leading to the breakdown of the original time-series into three different samples. The forecasting models were subsequently created through each of these models.

The results were expressed through the three different models, and since two of them have extant data, meaning an existing sample to compare the predicted data, it was possible to determine the accuracy level. However, these models cannot be applied immediately since the impact of COVID-19 is still present and flights have not resumed normality. Once this normality resumes, the models can be applied.

Keywords: Forecasting Methodologies, ARIMA Models, Airline Industry, COVID-19

JEL Classification:

C53: Forecasting and Prediction Methods - Simulation Method

C55: Large Data Sets: Modeling and Analysis

L93: Air Transportation

Resumo:

Modelos preditivos têm sido aplicados a variados setores como ferramenta de tomada de decisão há mais de 100 anos. A sua aplicação na indústria aeronáutica beneficia uma ampla variedade de interessados, como companhias aéreas e autoridades aeroportuárias que utilizam dados para prever a procura, definir preços, gerir frotas e tomar decisões relativas ao layout do aeroporto, expansões futuras, entre outros.

O principal objetivo desta dissertação é o desenvolvimento de um modelo de previsão capaz de prever o número de movimentos de voos no Aeroporto de Lisboa. O modelo foi baseado num modelo autorregressivo, que utiliza dados passados para prever valores futuros. O Aeroporto de Lisboa foi o objeto escolhido para esta dissertação. Dados semanais relativos aos movimentos aéreos no Aeroporto de Lisboa consistiram na amostra para este estudo, os quais foram processados através do software de programação RStudio.

Assim que os modelos *Autoregressive Moving Average* (ARIMA) foram definidos, os dados de previsão foram criados e testados quanto à precisão usando os dados existentes. O impacto do COVID-19 teve que ser considerado nesta situação, levando à divisão da série temporal original em três amostras diferentes. Os modelos de previsão foram posteriormente criados através de cada um desses modelos. Os resultados foram expressos através dos três modelos, e como dois deles possuem dados existentes para comparação com dados previstos, foi possível determinar o nível de precisão. No entanto, os modelos não podem ser aplicados imediatamente, uma vez que o impacto do COVID-19 ainda está presente e os voos não voltaram à normalidade. Uma vez resumida essa normalidade, os modelos podem ser aplicados.

Palavras Chave: Metodologias de Previsão, Modelos ARIMA, Indústria Aeronáutica, COVID-19

Classificação JEL:

C53: Métodos de previsão e predição - Método de simulação

C55: Grandes conjuntos de dados: modelagem e análise

L93: Transporte Aéreo

Table of Contents

Table of Figures

List of Abbreviations:

ACF - Autocorrelated Function

ANSP - Air Navigation Service Provider

ARIMA -AutoRegressive Integrated Moving Average

ATC - Air Traffic Control

ATFM - Air Traffic Flow Management

DF - Dickey Fuller

EMPBFE - Enhanced Mean Peer-Based Forecast Error

GA - General Aviation

GARCH - Generalized AutoRegressive Conditional Heteroskedasticity

HHI - Herfindahl-Hirschman Index

IFR - Instrumental Flight Rules

KELM - Kernel Extreme Learning Machine

KPSS - Kwiatowski, Phillips, Schmidt and Shin

LSSVR - Least Squares Support Vector Regression

LSTM - Long Short-Term Memory

MAE - Mean Absolute Error

MAPE - Mean Absolute Percent Error

MFE - Mean Forecast Error

MGBFE - Mean Growth-Based Forecast Error

MPBFE - Mean Peer-Based Forecast Error

OD - Origin-Destination

PACF - Partial Autocorrelated Function

PP - Phillips Perron

RNN - Recurrent Neural Network

SEM - Socioeconomic Mobility

TRAMO - Time series Regression with ARIMA noise, Missing values, and Outliers

VMD - Variational Mode Decomposition

1. Introduction

For any given industry, forecasting is a useful tool that can help business gaining business advantage and make data-driven strategic decisions based on past consumer behavior. Anticipating a market's movement and pattern allows a firm or company to gain advantage compared to its competitors as the more accurate the forecast is, the bigger chances of thriving a certain player has in said market.

Ever since the air travel boom in the 1950s and 60s, airlines have developed predictive strategies in order to adapt themselves to the constant changing environment, aiming at maximizing profits and client satisfaction. Many studies regarding forecasts within the aviation industry have been conducted throughout the years, with most considering past data and behaviors as sample and indicators for the predictive model creation, whilst others additionally consider external factors such as the economic impact of new routes, increasing passenger movements or even local demographics.

In the early 2020s, the most important economic factor to consider in any industry is the impact of COVID-19, since it has led to lockdowns, travel restrictions and overall halt of non-necessary services. The aviation industry was one of the most impacted by the virus since most international travel was banned due to governments restrictions, leaving only minimum services and rare domestic flights untouched. The creation of a forecasting model in this period is challenging due to the uncertainty of COVID-19's consequences and how long will it take for air travel to resume its normality.

1.1 Objectives

The main objective of this dissertation was the development of a forecasting model to predict the number of flight movements at any given airport. Lisbon International Airport was the chosen case for this study. Another important aspect was the analysis of COVID-19 and its direct impact on the airport's flight numbers.

1.2 Methodology

Since forecasting models are usually based on past data and behaviors, it was necessary to obtain flight movements regarding Lisbon International Airport. Once the data was obtained through EUROCONTROL (2020), RStudio was used to analyze it by conducting Stationarity and Unit

Root tests, and subsequently create the different ARIMA models. This type of model was chosen due to its use of past data in order to predict future values by using lagged moving averages to smooth time series data, consisting in a more flexible model when compared to exponential smoothing or simple linear regression. These ARIMA models were the basis for the future forecasting models, and the consequent results were tested for accuracy based on extant data for the respective periods.

Whilst this model was directly applied to Lisbon International Airport, most commercial airport behave similarly in terms of seasonality, with higher peaks during the summer and lower peaks during the winter. This implies that different datasets can originate forecasting models with similar level of accuracy.

1.3 Structure

This thesis will be subdivided into five chapters:

- Introduction – consisting of a general context, objective, an initial approach to the methodology, scope and structure.
- Literature review – with the aim of providing the foundation for the research, this chapter provides a synopsis of the existing literature on factors affecting airport demand, different forecasting methodologies and time series models, and the impact of COVID-19 in the aviation industry.
- Methodology – which incorporates the different methods and principles used to fulfil the objectives of this dissertation, with focus on the sample data, the used software and proposed models.
- Results and Discussion – where the model development is exposed. The different models were then compared in terms of accuracy and further discussed on how can be applied to distinct situations.
- Conclusion – a segment which contains a simplified analysis of the study results and how these relate to the overall objective of this dissertation. Furthermore, the model's limitations and possible applications of this project are also dissected.

2. Literature Review

In order to fulfil the objectives of this dissertation, a literature review is conducted to review relevant topics associated with this study's subject. This section includes existing literature on airport operations, determinant factors on passenger demand, forecasting methodologies, and the impact of COVID-19 on the aviation industry. The development of the forecasting model will be influenced by previous studies along with present economic factors which will be further dissected in this section. The highly uncertain impact of the recent COVID-19 pandemic also plays a relevant role in the model creation and therefore is also being explored in this section.

2.1 Factors Affecting Airport Demand

Socio-economic characteristics have been shown to affect demand for health care (Celik and Hotchkiss, 2000), broadband connectivity (Dwivedi and Lal, 2007), fuel (Wadud et al., 2009), gambling (Layton and Worthington, 1999), and alcoholic beverages (Johnson and Oksanen, 1974), to name just a few examples. Hence, there is ample and long-term research documenting that socio-economic characteristics are important determinants of demand for a variety of goods and services (Hofer *et al*, 2018). Since the emergence of railways and airports, many questions have been raised regarding the impact of economic factors on passenger demand. This research focuses on the airline industry, therefore, it's important to reference how air traffic is influenced by external factors, and which of these have a greater impact on airline and airport operations. Both airlines and airports must be appropriately synchronized, since they depend on each other for the aviation industry's well-being, leading to, on one hand, airports ensuring the new routes sustainability and satisfying the primary needs of the passengers in the catchment area, and on the other hand, public stakeholders asking airports management to measure the economic impact of active flights in order to grant a financial support (Perboli *et al.*, 2011b). The studies that have dealt with this topic are presented in the section.

The first study worth analysing dates back to 2012, when Benedetti *et al* developed a logistic regression model to estimate the passenger flows in *The Cagliari Airport impact on Sardinia tourism: a Logit-based analysis,* while considering the airport schedule, the accessibility of the tourist destination from the airport area, the cost of the flight and the attractiveness of the destination region for the tourists. Given the model, a change in a target airport schedule was introduced, and the new tourist flows, as well as their expected economic impact, were forecasted.

Additionally, researchers intended to predict the effect of the opening of one or more new routes on the flow of passengers from chosen origins to Cagliari Airport and to measure the economic impact on the close area.

The economic and spatial interaction logistic regression model was based on the forecasting analysis, which was mostly focused on a two-level Logit Analysis that can represent intermodal or transhipment transportation networks (Benedetti *et al*, 2012). It includes characteristics such as identifier of flight origins, identifier of intermediate points, identifier of destination, observed flows, generalized flight cost matrix, generalized travel cost matrix and the average total cost. The intention is to define a valid model for the estimated flows matrix considering both the airport and flight characteristics, as well as the odd features of the destination.

Consequently, to calibrate the model and to simulate the changes due to the introduction of a new schedule, which is the basic information needed to estimate the economic impact of this schedule change the model must include the following input data: number of arrivals registered in tourist facilities; cost of the flight; cost of accommodation and cost for the rental car; presence of the direct flight.

After the model was created, a what-if scenario was designed - the introduction of a route Cagliari-Russia represents a further and more interesting development for the Italian airport. Providing this direct connection, Cagliari could reach a larger range of potential customers, distinguishing itself from the airports of Olbia and Alghero, which do not handle any connections with any of Russian airports (Benedetti *et al*, 2012). The results estimate that an increase of 90% in terms of Russian tourists in the region and a significant revenue increment around 12 million euros, which leads to the conclusion that thoroughly understanding the economic and demographic characteristics of a specific location can bring benefits to players in the airline industries.

The relationship between economic factors and air travel is symbiotic: The prediction of future air travel demand is of critical importance in the aviation industry and the basis for policy and managerial decision-making related to infrastructure and production planning (Carson *et al*., 2011), since it’s determined by a set of factors such as ticket prices, income, and population, with characteristics such as quality of air service, consumer wealth and flight delays found to have a partial influence (Hofer *et al*, 2018)

Socio-economic mobility and passenger demand in the U.S. (Hofer *et al*, 2018), studies the impact of socioeconomic mobility (SEM) on passenger volume in the United States by applying recent findings on SEM research. Socioeconomic mobility in the United States refers to the ascendent or descendant movement of Americans from one social class or economic level to another, through job changes, inheritance, marriage, etc. The United States faced record levels of income inequality and one of the lowest rates of SEM among industrial nations (Piketty and Saez, 2003) since it is usually associated with higher diversity in a specific region. In regions with a diversified business base, workers with variable skills can be more successfully matched up with a wider range of employment opportunities, hence generating economic outcomes for both businesses and individuals that otherwise would be impossible to achieve, but at an individual level it's not quite clear how SEM affects air traffic demand (Hofer *et al*, 2018)

In order to conduct this research, different sets of passenger data were used along with income and population statistics. The dependent variables included Interest and Total Passenger Enplanements, while independent variables included Socioeconomic Mobility, Income and Population, AltAirports – substitute commercial airports for a given destination (Morrison, 2001). – and Yield – distance adjusted air fares.

The study results show that greater SEM is widely associated with lower yields, and there's also evidence that passenger volume decreases as SEM increases. This means SEM has a clear impact on air travel demand, and airlines should be cautious when operating in markets with higher socioeconomic mobility, regarding the impact of yields and this indicator should be considered when forecasting demand (Hofer *et al*, 2018).

2.2 Short-term Forecasting

Throughout aviation history, forecasting has been used as tool for mostly every player in the industry, with both airline companies and airports relying on a wide variety of models to enhance and optimize their operations. Forecasting the demand for aviation activities is an important task in economic planning (Ghobrial, 1997). Although the definition of short-term forecasting is not exact, most researchers acknowledge the time frame as up to two years.

In the early 1990s, Atef Ghobrial developed a multiple regression model using both dependent and independent variables. The dependent variable in the model is the number of annual aircraft operations at a given general aviation (GA) airport (Ghobrial, 1997), whilst the independent variables in the model include a set of descriptors of the nature and level of the socioeconomic activities in the county where the airport is located, and another set of supply variables that affect the levels of service at an airport (Ghobrial, 1997).

The results showed that the existence of Air Traffic Control (ATC) and runway length are the most prominent factors in the 82 GA airports in terms of demand. Other services such as avionics, charter flights and rentals, aircraft repair, and crop dusting are shown to have an impact on the increased activity of such airports. However, the most important conclusion lies on the fact that there's a need to analyse more variables and sensitivity of the model specifications and correlations between some variables in the model are also drawbacks to this particular of analysis. A method to overcome this problem would be to develop separate models for both local and itinerant aircraft operations.

Multiple regression models are not the only used for short-term forecasting in the aviation industry. A combination of different models can be useful than other time series models, indicating that they are promising tools to predict complex time series with high volatility and irregularity (Xie *et al*, 2018).

A relevant example of a hybrid model applied to short-term forecasting is the work developed by Xie *et al* in 2018, in which a Least Squares Support Vector Regression (LSSVR) was combined with a seasonal decomposition method in order to predict air passenger movements at Hong Kong International Airport.

The Least squares support vector regression model transforms the regression model in an optimization model. There are two distinct seasonal decomposition methods: X-12-ARIMA, which decomposes time series into three components - trend-cycle component, seasonal factor and irregular component that can be combined into the original data in additive and multiplicative forms. TRAMO (Time series Regression with ARIMA noise, Missing values, and Outliers) is a program for estimation and forecasting of regression models with errors that follow mostly nonstationary ARIMA processes. Consequently, two distinct hybrid approaches were built which include X-12-LSSVR (combination of X-12- ARIMA and LSSVR) and TS-LSSVR (combination of TRAMO and LSSVR), (Xie *et al*, 2018).

Two hybrid approaches are developed for the comparison with other time series methods. The investigation suggests that seasonal decomposition is an effective way to air passenger forecasting. It is important to describe the seasonal characteristic and nonlinear nature of air passenger series for better forecasting performance.

Another relevant of a hybrid model being used for the purpose of short-term forecasting in aviation is the work developed by Jin *et al* in 2018, entitled *Forecasting air passenger demand with a new hybrid ensemble approach*, which focused on analysing and modelling air passenger dynamic, concentrating on its impact on management and operation across the whole aviation industry. The combination of models was based on a Variational Mode Decomposition (VMD), an Autoregressive Moving Average (ARMA) and a Kernel Extreme Learning Machine (KELM).

The hybrid model was then applied to three distinct data sets – Beijing Airport and Guangzhou to test the performance and Shanghai Airport to test the robustness and applicability. Monthly passenger demand from January 2006 to November 2017 from the three airports was collected, from which January 2006 to July 2015 was considered the training set and from August 2015 to November 2017 was considered the out-of-sample data (Jin *et al*, 2019).

The main conclusions taken from this research were the following:

- Internal characteristics can be more efficient to extract in the original air passenger demand by adopting VMD.
- The stationary and non-stationary series are predicted respectively by the compatible models based on the results of the stationarity test, and the unique characteristics of each subseries can more be captured completely.
- By taking advantage of different forecasting models, the proposed approach can obtain more effective and convincing results, which are proved by different evaluation criteria.
- The hybrid approach VMD-ARMA/ KELM-KELM is initially developed, and its stability is demonstrated from various aspects including running times, the number of iterations and so on (Jin *et al*, 2019).

Short-term forecasting can also be used to predict variables with direct impact on passenger demand and aircraft movements. Robust and adaptive statistical models were previously developed in order to explore the effect of intermediate weather variable related to accuracy prediction using single layer LSTM - Long Short-Term Memory - Multi Layers memory block (Salman *et al*, 2018). LSTM is a specific recurrent neural network (RNN) architecture that designed a model temporal sequences with their long-range dependencies, which is widely used for time series prediction (Salman *et al*, 2018).

Dataset for this research was obtained from Weather Underground which collects weather data including temperature, dew point, humidity, and visibility from many weather stations all over the world. The range of data for this study was from year 2012 to year 2016 comprise of 40,025 time series data at Hang Nadim Airport Indonesia (Salman *et al*, 2018). The proposed model is a stacked LSTM with subsequent layers having 200, 100, 90, and 50 nodes of hidden layers.

Despite many models have been proposed for weather prediction, most of these models used the same input and output variables. The result of this study, which exploited LSTM model variant, showed that intermediate variables can improve prediction capability of the model. The LSTM model is feasible and suggested to be implemented in predicting weather with the addition of intermediate data in order to improve the accuracy. The best model of LSTM model in this experiment is multiple layers LSTM and the best intermediate data is pressure variable.

2.3 Long-Term Forecasting

Long-term forecasting is likely to be dominated by trend curves, particularly the simple linear and exponential trends (Granger and Jeon, 2007). Nevertheless, players in the airline industry benefit from long-term forecasting and this subsection provides an insight on a few conducted studies regarding longer time frames of predicting methodologies within the aviation industry. Similarly to short-term models, there is a wide variety of models which can be applied to accurately predict long term indicators.

The application of trends is very popular in fields such as business planning, financial market (Visser and Dangendorf, 2015), and the aviation industry can also be included. Trend is conceived as the part of the series which, when extrapolated, provides the clearest indication of the future long-term movements in the series (Visser and Dangendorf, 2015). Consequently, univariate and multivariate trend models are good methodologies to apply when predicting indicators such as demand levels. Andreoni and Postorino (2006) have calibrated and compared this type of model in order to estimate the passenger demand at Reggio Calabria regional airport in Italy by using the annual passenger number between 1989 and 2004.

The developed ARIMA model was based on the Autocorrelated Function (ACF) and a Partial Autocorrelated Function (PACF). The Box-Jenkins methodology was also applied to identify the model, estimate the parameters and diagnostic checking (Andreoni and Postorino, 2006)

In 2004, during the months of March, April and May, the runway was on maintenance which means the full year has been considered an outlier. This led the researchers to create two separate univariate models – one in which the outlier had been removed and another in which the outliner's values had been slightly altered in order to follow the trend (Andreoni and Postorino, 2006). Both results have been described as acceptable since they fit a time series model. The multivariate model was based on the univariate and includes two variables – income per capita and the number of movements in and out of Reggio Calabria airport – used to calculate the airport demand in the subsequent years.

After the results were obtained, it was concluded that both models had satisfactory levels of accuracy, however, the univariate model had a better performance every time the function contained peaks. Despite this, it was not concluded that univariate was better than the multivariate models, since univariate can only forecast the demand level if all the underlying conditions are the same and cannot be used to simulate the effects of different policies. Contrarily, data for the independent variables is always more difficult to obtain, which compromises the structure of a time-series model and consequently, its validity (Andreoni and Postorino, 2006).

Demand models also fit the purpose of long-term forecasting, as the German Aerospace Centre (DLR) developed a four-step model to forecast passenger and flight volume at German airports, whilst considering the impacts of Brexit.

The traditional DLR-Demand Model used the phases of trip generation, trip distribution, modal split and trip assignment to follow the traditional four-step algorithm of models used for simulating and forecasting traffic (Gelhausen *et al*, 2018) This gravity model, where the interaction between two places was determined by the product of the population of both places, divided by the square of their distance from one another, was divided in four distinct steps:

- First step - the number of journeys generated in and out of Germany, as the study region, are forecasted by using different approaches depending on the trip purpose

- Second step – the spatial division of Germany's trip volume to the originating airports inside the country and destination zones outside Germany.
- Third step - trip assignment by which the O-D flows were assigned to the routes served by direct flights.
- Fourth step - vehicle assignment by which route specific travel volumes were converted to the number of flights (Gelhausen *et al*, 2018).

While the four-step approach to demand modelling has a strong microeconomic foundation based on individual choice behaviour, this approach is not only very data-intense, but also needs rather well-defined, unswerving sub-models (Gelhausen *et al*, 2018).

Another important subject to highpoint is that many socioeconomic and traffic time series are non- stationary because of a stochastic trend, which can produce a spurious regression (Gelhausen *et al*, 2018). Variables explained are annual passenger and flight volume growth at German airports and the growth rate of passengers per flight serves as a variable to model the supply side of the air transport market. Air fares and jet fuel prices are expected to have a significant effect on passenger and flight volume growth over time, since they are important parameters of the demand and supply side of air transport.

The complete application of the model yields for each forecast year quite detailed forecast results, such as:

- Passenger volumes of the study area (i.e. Germany) by trip purpose
- OD-Passenger flows between the study area zones (in Germany) and foreign zones
- Passenger volumes on routes served by scheduled and charter flights
- Passenger volumes of airports of the study area (Germany)
- Passenger flight volumes in scheduled and charter traffic by route
- Passenger flight volumes of airports of the study area (Germany)

The new model has been employed to estimate the effects of Brexit on passenger and flight volumes at German airports. This case study focused on two scenarios, which were written by the IMF and HM Treasury, and compared them to a baseline scenario of no Brexit. Then, a distinction between effects that were due to a decrease in the UK GDP and a devaluation of GBP (Gelhausen *et al*, 2018).

Compared to the classical DLR model, the new direct demand model employs a far lower number of input variables. Moreover, the forecast values for the input variables are easier to obtain for the new direct demand model because they are less specific and more commonly available, e.g. GDP forecast for the European Union. This holds both for the demand as well as the flight forecast. The new direct demand model combines step one and two of the classical DLR model and omits step three ("route-specific passenger volumes"). Instead, based on the results of the demand forecast, the new direct demand model proceeds to the flight forecast (step four) (Gelhausen *et al*, 2018).

Direct demand prediction is not the sole objective of long-term forecasting models, since there is a significant number of variables which need to be understood in order to explain fluctuations. Suh and Ryerson (2019) attempted to understand the explanatory variables that impact the outcome of a severe reduction in passenger volumes, rather than to achieve the highest predictive performance, by developing a predictive model using binary logistic regression which estimated the probability of an airport experiencing a dramatic contraction in passenger demand over a period of 10 years. This time period is crucial for airport authorities since it allows the perception of need in terms of building new infrastructures such as runways or terminals (Suh and Ryerson, 2019).

The data revolves around 64 airports around the United States, located in the top 50 metropolitan areas of the country. Then, a binary logistic regression model was built using several airport and MSA explanatory variables to predict the binary outcome. These variables included airport competition, connecting passenger share, avg. number of seats per aircraft, avg. ticket price, HHI, HHI change, population change, per capita income and service sector employment, and were all standardized in order to fit the proposed model.

The results expressed an increased importance of variables such as:

- HHI – since most airports with large airline presence (hubs) are likely to suffer a decline in passenger movements if operations are ceased in the respective location.

- Per capita income – noted as one of the most important demand influencer in terms of passenger movements.
- Connecting passenger share – this factor correlates with HHI, since most hubs are prone to have a high number of passengers on connecting flights and a "de-hubbing" can bring significant impact to the airport's demand.

Another topic focused on this research was the optimism bias regarding forecasting in the aviation industry. To overcome this problem, the researchers have proposed a reference class forecasting, which was tested for feasibility and can be manifested in four different models:

- Mean Forecast Error (MFE): which uses each airport's own past empirical forecast errors.
- Mean Growth-Based Forecast Error (MGBFE): To reflect the observation that there may exist a correlation between forecasted growth percentage and forecast error, the empirical forecast errors of the past forecasts (of any airport) were used with forecasted growth percentage, meaning within a range of the forecasted growth percentage of interest.
- Mean Peer-Based Forecast Error (MPBFE): airports with similar socioeconomic and airport characteristics are compared and their past forecast errors to adjust the current forecast are used.
- Enhanced Mean Peer-Based Forecast Error (EMPBFE): the predicted probabilities of a severe contraction in passenger were incorporated to adjust the MPBFE in the previous method. Because the predicted probabilities provide the information on how likely it is for an airport to experience a dramatic drop in the passenger volumes (and thus, a potentially larger forecast error), we use this additional information to calibrate the MPBFE and name this approach Enhanced Peer-Based Forecast Error (EPBFE).

Consequently, four different forecasts were compared to the actual forecast in order to determine which one has a more approximated value. The study concluded that only the first method (MFE) did not bring any significant difference to the actual forecast values, and that the remaining three can be used as tools to minimize this optimism regarding aviation forecast and provide a more realistic set of results.

It is fair to assume the aviation industry is progressively becoming performance driven, with KPIs gaining importance from a strategic point of view. Delgado et al (2020) have presented a new holistic, microscopic model for air transport management system with the objective of predicting how projected changes in 2035 and 2050 will likely interact by using a set of parameters such as macroeconomic, technological or regulatory and it is aimed at five different stakeholders in the aviation industry, including ANSPs (Air Navigation Service Provider), airlines, airports, passengers and the environment. It modelled the three temporal phases of ATM (strategic, pre-tactical and tactical) for each considered scenario, with the objective of generating a representative day of operations for each given situation.

Large airports' current business models rely heavily on non-aeronautical revenues (parking, shopping, etc.) (D'Alfonso et al., 2013). Congestion is a major issue for most, and different strategies are implemented to increase their capacity, such as soft management procedures or heavy changes in infrastructure (Berster et al., 2013), or improvements from airport expansion programs and technological enhancements.

From a strategic point of view, the economic block was the first block of this layer and had the objective of creating accurate supply-demand prediction for any given scenario. It implemented agents such as airliners, flights, passengers and ANSPs (Delgado et al, 2020). Supply and demand interact in this network in a complex mode. On one hand, the supply is leg-based, each airline creating its own capacity for each leg and on the other hand, demand reacts to the prices of itineraries. The schedule mapper is the second block in the strategic layer (Delgado et al, 2020). It converts the high-level flows of the economic model into individual schedules to be used by the flight plan generator, which is a very demanding task with several constraints. It goes through the following steps:

- load data on airports, historical schedules, pattern and strategic flows;
- compute average travelling times between every OD pair;
- compute likely departure times;
- load the decision tree for the turnaround times;
- for each airline:

- trim its network by removing aircraft which are in excess;
- grow the network by adding aircraft to meet demand;
- compute the new schedules and add them to the database.

The pre-tactical layer aimed at creating the required level of detail in order to create each flight and passenger's routes by generating outputs such as individual flight plans, ATFM regulations and probabilities of being assigned delay due to a regulation, and individual passengers' itineraries. Flight plan generator transforms origin-destination schedules into actual flight plans through a route generator, a trajectory generator and a fuel estimation; ATFM regulation generator estimates the probability of a route being affected by ATFM regulations and the consequent delay, and is divided into two sections – capacity issues and remaining; and passenger itinerary generator creates the passenger flow and respective schedule by computing the possible options available for the passengers in each flow considering the minimum connecting times at the airports, optimizing the assignment of passengers among their options considering aircraft capacities and minimum connecting times at airports and Creating additional passengers' itineraries to ensure that the load factors of the aircraft are realistic.

The tactical layer models delayed propagation between flights and the flexibility of the system during disruptions such as cancellations or delays and with limited resources such as airports and en-route capacity. It is the result of two processes:

- Gate-to-gate simulation – based on flight plan submission, previous aircraft ready, push back process and arrival processes:
- Door-to-door simulation (Delgado et al, 2020).

In order to create the model, the following data was used extensively:

- set the initial state of the economic model;
- extract the distribution of delays for airports and ANSPs, which helped to:
- infer delay-traffic relationships (for airports);
- perform a mean–variance analysis on airport delay;
- perform an analysis of ATFM regulations;

- compute the length of trajectories in each ANSP's area;
- cluster possible routes between origin and destination airports;
- model flight plan preferences (flight level and speed requests);
- model flight trajectories (characteristics of climb and descent phases);
- estimate average wind distributions between regions by comparing ground speed with requested air speed.

Additionally, in order to calibrate the model, a few analyses such as the pairs of variables analysis, route clustering, ATFM probabilities and passenger itineraries were performed. In terms of results, it is easier to demonstrate the obtained results through each of the previously mentioned layer.

For the strategic layer, the strategic layer shows a tendency for airlines to increase the average size of their aircraft when capacity is scarce, whilst the operational cost decreases from 2035 to 2050. Additionally, delays increase throughout this period, driven by the large increase in traffic, barely mitigated by the increase in airport capacity

For the pre-tactical layer, the computation of flight plans indicates that the average flight plan distance tends to increase in the period from 2014 to 2050, leading to higher fuel burn per flight, accompanied by an increase in the percentage of connecting passengers, since the number of flights increase thus widening the range of connecting possibilities. Also, different operating costs might shift demand for particular airspace areas, resulting in variations in revenues per ANSP.

For the tactical layer, average gate-to-gate times do not increase, despite the increment in delays. This is a volume effect, due to the fact that airlines tend, on average, to operate longer routes (in time) in 2035 and 2050.

To conclude, the model was run based on a series of different factors and scenarios in order to consider distinct outputs, by using two different baselines which would somewhat regulate the uncertainties and inconsistencies. Additionally, it was proven that this model is actually effective when working with a large quantity of metrics.

2.4 COVID-19 in the aviation industry

Since the outbreak of COVID-19 in late 2019, the aviation industry has gone through a multitude of changes which have impacted airlines, airports and consequently passengers and local economies. The huge uncertainty regarding the consequences of this pandemic and the concerns over when "normality" will resume are probably one of the biggest challenges of this study, since the used data is highly affected by the enormous drop in passenger movements registered in the early weeks of 2020. Nonetheless, a few researchers have already developed papers and studies regarding the impact of COVID-19, which will be explored in the following section.

In early 2021, Sun, Zhang, Zheng and Wandelt have released a paper entitled COVID-19 pandemic and air transportation: Successfully navigating the paper hurricane, which focus not only on the impact COVID- 19 has had in the aviation industry, but also how the industry has impacted the pandemic itself based on the analysis of the global air transportation system during COVID-19, the impacts on the passenger-centric flight experience, and the long-term impacts on broad aviation.

COVID-19 is the most severe pandemic in recent decades, with highly contagious indexes and consequently very fast spreading rates. Airports are critical points since they aggregate people from all over the world in relatively small, shared environments such as gates, lounges, shops or baggage claims. The network perspective has forced many airports to close in the early stages of the pandemic, where the level of uncertainty was considerably high, and the infrastructures were not satisfactory for travellers to use safely. As a consequence, most airports have experienced a massive drop in traffic in the period ranging from February to June 2020, to either domestic or international destinations. However, it's not possible to equally compare these two categories of destinations, since domestic flights were shortly resumed after the mention period, and at the time, governments had many traveling restrictions between countries and in many cases, international flights were very limited.

Since the virus is highly contagious, it was fundamental to ensure travellers felt safe and had the lower possible risk of contracting the disease whilst on the airport or inside the aircraft. There was a need for a major infrastructure redesign, so that the safety measures could be applied. COVID-19 testing stations were one of the first features to be implemented, which obligated travellers to arrive at the airport earlier and restricted those whose temperature did not meet the requirements. Despite being a debate on whether these tests were effective or not, it did not stop most airports adhering to this particular measure. In-airport shopping was severely restricted since it is one of the most crucial points of converging passengers from all over the world. The boarding process was also altered, with many considering it one of the most challenging adaptations of the COVID-19 era. Whilst before the pandemic each airliner had its predefined boarding method, with aircraft configuration, presence of hand baggage, load factor and many others being the determinant factors of the considered technique, during the pandemic airliners were recommended not to operate their aircrafts at full capacity, and therefore distribute passengers in compliance with safe distancing. The more common techniques include an empty middle seat, boarding in small groups of people or by seat number, with all relying on a passenger self-will to practice social distancing as much as possible. In-flight practices were highly debated since its implied dozens or even hundreds of people enclosed inside an aircraft cabin, however, there's little evidence that those cabins can accelerate the virus propagation. Nevertheless, airliners have acted by implementing measures such as the previously mentioned operation at lower-than-full capacity, mandatory usage of mask, limited food and drinking services and reinforced sanitary disinfection.

Despite some of these measures can still be applied post-COVID-19, it's fair to say they are short-term since once the virus is somewhat controlled, there will no longer be the need for social distancing and extreme sanitization.

However, it's also interesting to predict the long-term impact of the pandemic in the aviation industry. The aviation industry was the most affected by the travel restrictions, which is mirrored by the enormous financial losses airliners have suffered since the beginning of the pandemic. Due to the economic impact air travel has on a particular country, many companies will require a financial aid from government entities. Whereas in the past, many airlines were aided by private investors, the uncertainty of the future economic situation has created some reluctance among these, shifting the recovery plan to a more political point of view, with government help being the most reliable alternative for airline companies to pursue. Passenger demand is also facing a high level of uncertainty, with experts such as the Oxford Analytica predicting that air traffic will resume the normal growing path observed before the pandemic in about 2 to 4 years. It is expected that the Asian market bounces back faster than others such as the European, American or Middle Eastern, with China as the main contributor to this rise. Overall, the future of aviation cannot be accurately predicted with sample data collected in early 2021, but a few studies have reached potential solutions to the COVID-19 problem. For instance, Chen et al. (2020) discussed the possibility of creating an immunity passport which would lead to less enforced screening at airports by only allowing travelers with a clean bill of health (in terms of COVID and other potential future diseases); Tardivo et al. (2020) has recommended that sustainable European mobility would be made through rail and not air; and Hendrickson and Rilett (2020) who have debated the role of air travel when society has the necessary tools to operate via telecommunication, and the emergence of more and more automated vehicles.

To conclude, in early 2021 the insufficient available information, alongside the high levels of uncertainty difficult the task of accurately predicting the long-term consequences of the COVID-19 pandemic on the aviation industry. For now, measures such as the usage of protective masks, social distancing and disinfection practices will continue to be standard at airports, alongside limited services such as shopping and lounges. Additionally, in order to minimize economic losses, alternatives to air transport have been emerging such as the creation of upgraded rail systems, capable of transporting people as fast as air travel does.

2.5 Time-Series Models

Whilst economic factors and previous forecasting studies will influence the model development, understanding time series models is also crucial for constructing an accurate predictive model. This section will briefly clarify the meaning of Auto Regressive models and how they can be used to predict future values.

Autoregressive Integrated Moving Average models are widely used in statistics and econometrics, specifically to study events happening over a time period and consequently forecast future values. Applied to stationary time series, the models are given by ARIMA (p,d,q), of which p represents the order of autoregression, d the differentiation order, and q the moving average order.

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) is a statistical model used time-series data analysis in which the variance error is believed to be serially autocorrelated. Heteroskedasticity describes the irregular pattern of variation of an error term, or variable, in a statistical model.

The hybrid model was put in practice by using real data from the Bureau of Transportation Statistics, who publishes the United States monthly aggregated pax at national level from January 1990 to April 2016. The data was then divided in two sets: The first from January 1990 to April 2013, and the second from May 2013 to April 2016. The first data set was used for the assessment of four-time series methods: Holt-Winters, ARIMA, DTGM and the ARIMA + GARCH + Bootstrap, whilst the second data set was used as out of sample data to compare the time series methods one-step ahead forecasts with the original data for pax demand, and to validate them.

3. Methodology

The following section encompasses the different methods used to conduct this study by describing the various stages of research, in which the use data and software will be broken-down. A well-structured and concise method description is fundamental to expose the research's objectives and the numerous stages it goes through in order to reach the goals.

Developing a forecasting model using real data requires a thorough data analysis and comprehension due to the multitude of existing prediction models, since the slight misinterpretation can lead to erroneous results.

3.1 Data

The research aims at creating an accurate prediction model in order to forecast the number of aircraft movements at Lisbon's Humberto Delgado International Airport, located 7 km north of the Portuguese capital's city centre. Data collected from EUROCONTROL displayed the total number of daily Instrument Flight Rules (IFR) arrivals and departures – flights using on-board instruments and electronic signals as opposed to a visual flight plan (VFR) – between the 3rd of January 2016 and 31st of December 2020. In this period, Lisbon Airport registered 912,913 flight movements, of which 456,320 were departures and the remaining 456,593 were arrivals.

The data was then aggregated in weekly batches, in order to reduce the period sample and still embrace seasonality. For instance, if the number of daily movements was considered as the data unit, the sample would involve 1827 periods instead of the 261 periods extracted from the weekly reports.

3.2 Software

RStudio was the selected tool to analyse the data and subsequently create the accurate model for air traffic movement prediction through R language, a form of programming language useful for statistical computing and graphic display.

The wide variety of packages in this software - collections of *R* functions, data, and compiled code in a well-defined format, created to add specific functionality – allowed the performance of several tests on the time series data, which were fundamental to create an optimal Auto Regressive Integrated Moving Average model and consequent forecasting technique.

3.3 Model

To estimate the model, the data has been classified as a time series and several tests have been computed in order to highlight the characteristics of data. The first step was the Autocorrelation Function and Partial Autocorrelation Function, which allowed to determine the correlation coefficient within the series' data sets and graphically represent it. The functions compared values within the series and described how the earlier values of the series are related to the most recent ones.

Subsequently, Unit Root Tests were performed in order to check stationarity of the time series. Statistical properties such as mean, variance and autocorrelation were constant over time if the series was stationary and the opposite if the series was non-stationarity. Generally, Unit Root Tests imply that the null hypothesis is the presence of Unit Roots or non-stationarity and the alternative hypothesis is stationarity (the exception is KPSS test). This research has encompassed the three following Unit Root Tests:

- The Dickey-Fuller Test - based on a simple autoregression with or without a constant or time trend;
- Phillips-Perron Test;
- Kwiatowski, Phillips, Schmidt and Shin Test – uses stationarity as a null hypothesis.

Since the three tests treat serial correlation in the errors of the auxiliar regression in different ways, all of them were computed to check if all corroborate the same conclusion.

As previously mentioned in the literature review section, COVID-19 brought noteworthy implications to the aviation sector, which were reflected by the significant drop in air traffic movements in February 2020. This factor massively impacted the time series in terms of stationarity, which required the performance of two different tests in order to check the existence of a break within the series. Breaks occur when there is a change in the standards for defining and observing a variable over time, and can be substantiated by:

- Chow Test – checks if regression coefficients are different for split data sets;
- Bai-Perron Test - estimates multiple structural shifts in a linear model estimated by least squares.

As the structural breaks can affect the unit tests results, the previously mentioned Unit Root Tests were applied to the newly defined time series periods (pre and post break) in order to check stationarity once again.

Furthermore, the proposed ARIMA models were built through different periods, determined by the break, which meant the data set was different for each period, and the optimal ARIMA model was different for each case. Nonetheless, several ARIMA models were estimated for each period, serving as a base for the forecasting method.

In order to create a realistic timeline for the model, all forecasts were aimed at the end of 2021, which meant the number (*n)* of periods ahead would correspond to the number of weeks until the 31st of December 2021. A practical example can be demonstrated by the previously mentioned time series period between the 3rd of January 2016 and the 31st of December 2020, which would imply a prediction sample of n=60 – the number of weeks between 1st of November 2020 and the 31st of December 2021.

The *forecast* function in RStudio enabled to convert the estimated ARIMA models into a wider time series, according to the number of samples, and subsequently plot the results. The function was applied to every ARIMA model, corresponding to each of the periods determined by the existence of a break.

Since it was not possible to classify a model as either correct or incorrect, accuracy tests were fundamental in order to estimate which of the models is predicts better the dependent variable. The *accuracy* function in R determined the level of "agreement" between the observed and the predicted values from the models, based on loss functions such as Mean Squared Error (MSE), Mean Absolute Percent Error (MAPE), Mean Absolute Error (MAE), among others. Each prediction model was submitted to this accuracy test and the optimal result corresponded to a lower value of loss function, meaning the cost or error margin between the previously mentioned indicators.

4. Results and Discussion

4.1 Results

The data retrieved from EUROCONTROL, 2020 has provided the necessary basis to conduct this study. As previously mentioned in the methodology section, Lisbon International Airport has registered over 900,000 flight movements between January 2016 and December 2020. The database compiled daily information about IFR departures and arrivals in Lisbon, which gave around 1766 samples for the mentioned time period. In order to reduce the sample size and furthermore increase the chances of creating a precise model, the daily data was transformed into weekly values, reducing the number of observations to 261. This transformation was vital in order

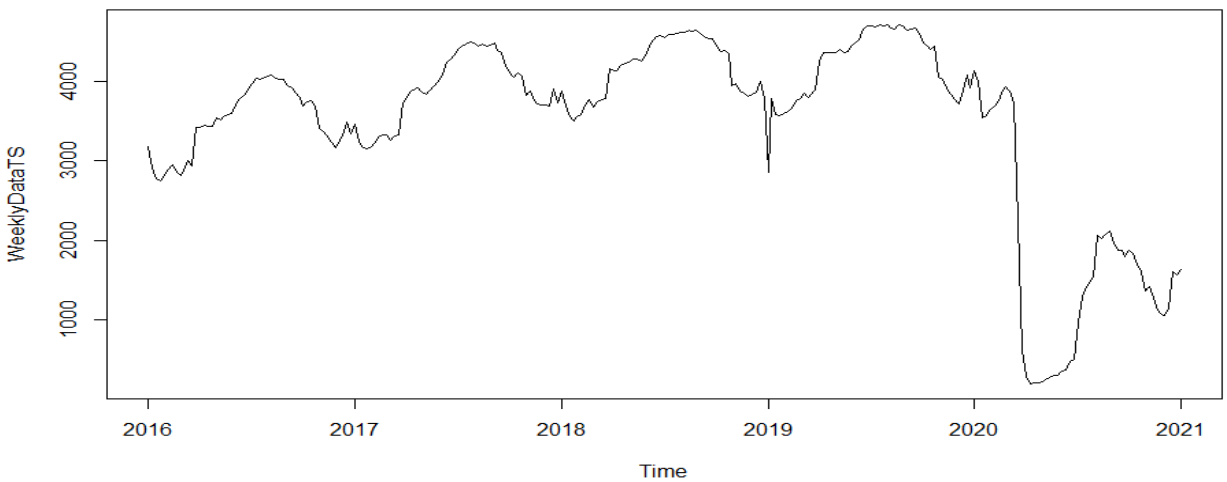

Figure 1 - Time Series for Weekly Flights at Lisbon Airport from January 2016 to December 2020

to attenuate daily fluctuations which could massively impact the final outcome of the predictive model.

Once the data set was well established, it was necessary to classify the weekly flights into a time series, which was performed by using the *ts* function in RStudio. The plotted time series is visible in figure 1.

As it is perceivable in fig.1, during the first weeks of 2020 the world was struck by the COVID-19 pandemic, which brought severe implications for a wide variety of sectors, where the aviation industry is well included. Closed borders and travel restrictions heavily struck both airliners and airports, leaving enormous question marks regarding the future of aviation in the short and medium term. To put in perspective, on the 15th of March 2020, Lisbon airport had registered 518 movements, compared to just 5 on the 12th of April of the same year, which corresponds to a 96,5% drop in just 28 days. The impact of COVID-19 consisted in one of the biggest challenges to the creation of the forecasting model.

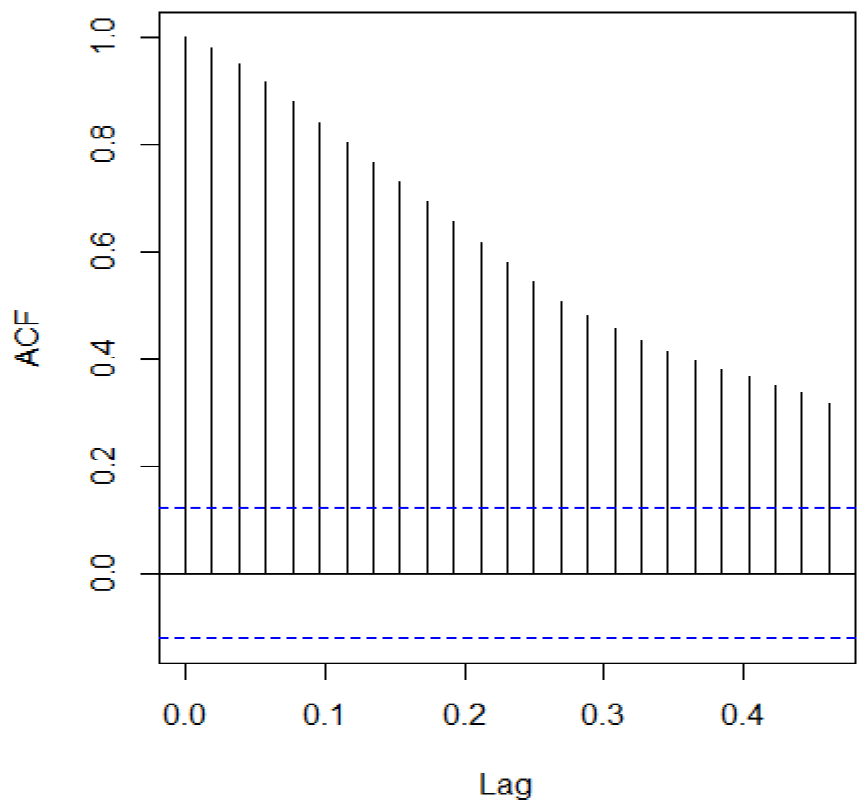

Figure 2 - Autocorrelated function for Lisbon Weekly Data

To test correlation between observations within the time series, Autocorrelation Function and Partial Autocorrelation tests were performed. The results can be seen on fig. 2 and fig. 3 respectively.

The plotted Autocorrelated function shows strong correlation indexes in lower lag – fixed amount of passing time – values. The autocorrelation function shows a linear decay point to a non-stationary series. The Partial Autocorrelated Function show significant correlation level only for the first lag. For higher lag values, the correlation index is also substantially lower.

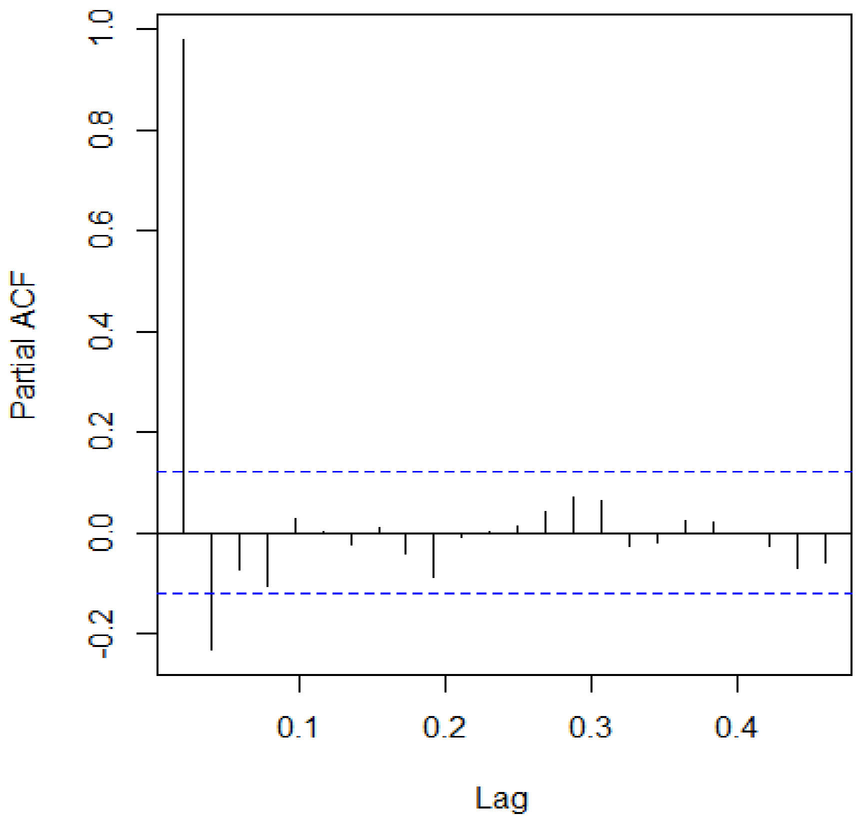

Figure 3 - Partial Autocorrelated function for Lisbon Weekly Data

Furthermore, Unit Root Tests were performed in order to check the series' stationarity, as three distinct types provided the necessary accuracy to reach the optimal outcome. The results obtained from Dickey-Fuller (DF), Phillips-Perron (PP) and Kwiatowski, Phillips, Schmidt and Shin (KPSS) tests could be explained as following:

- DF Test – the null hypothesis (H0) was not rejected and the presence of Unit Root in this time-series was confirmed. This hypothesis could be obtained by comparing the calculated values (value of test statistic) with the tabulated values (tau1). Since the value of test statistic (0.8878) was in fact lower than any percentile of tabulated values (2.58 for 1pct, 1.95 for 5pct, and 1.62 for 10pct), the Dickey-Fuller Test revealed non-stationarity within the time series.

- PP Test – follows the same hypothesis principle as the DF test, and in this particular case, the calculated value is -10.5128 compared to the 2.341 calculated value. Therefore, the null hypothesis can’t be rejected, and the series is non-stationary.

- KPSS Test – contrarily to the DF and PP tests, this test follows a reverse methodology, in which the null hypothesis is the non-presence of Unit Root, and since the calculated value for this time series (0.836) is higher than the tabulated values for each percentile (0.347 for 1pct, 0.463 for 2.5pct, 0.574 for 5pct, and 0.739 for 10pct) the null hypothesis was rejected, and the non-stationarity also confirmed by using this test.

The impact of COVID-19 and consequent decrease in weekly flights has implied the presence of one or more breaks within the series. In order to test the existence and consequent period in which the time series breaks, Chow and Bai Perron tests were applied. Through the *breakpoints* functionality in R, it was determined that the series registered the presence of breaks in two distinct periods – 66 and 213 – corresponding to the weeks starting on the 26th of March 2017 and the 27th of January 2020 respectively.

The model development was based on each of the periods, which meant three distinct approaches for each time series. Figures 4, 5 and 6 represent the plotted graph for each time series data set.

- Pre-break – From the 3rd of January 2016 unit the 25th of March 2017 (week 1 to 65);
- Break 1 – From the 26th of March 2017 until the 26th of January 2020 (week 66 to 212);
- Break 2 – From the 27th of January 2020 until the 31st of December.

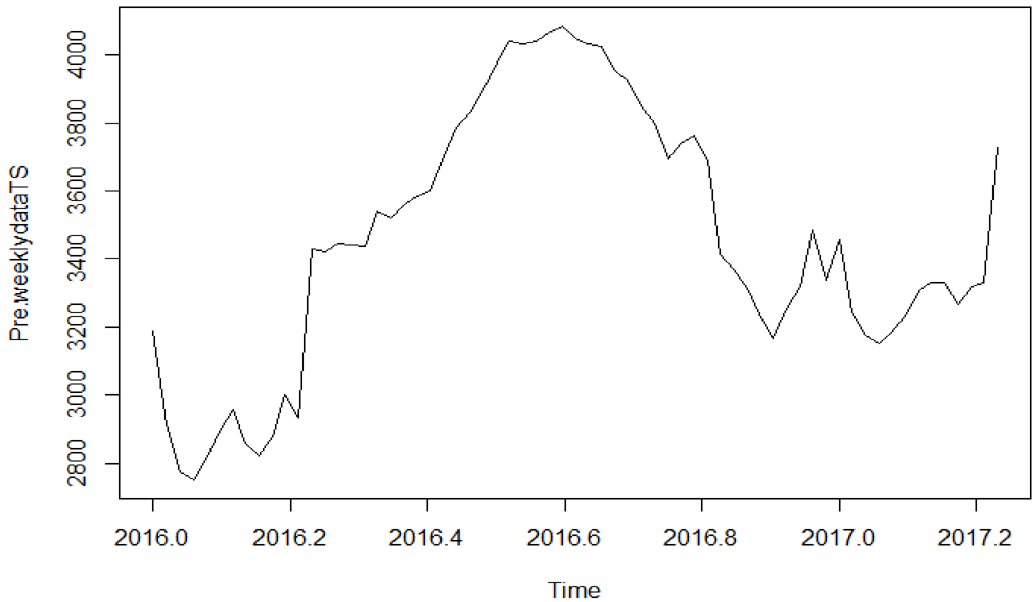

Figure 5 - Lisbon Weekly Flights from January 2016 until March 2017

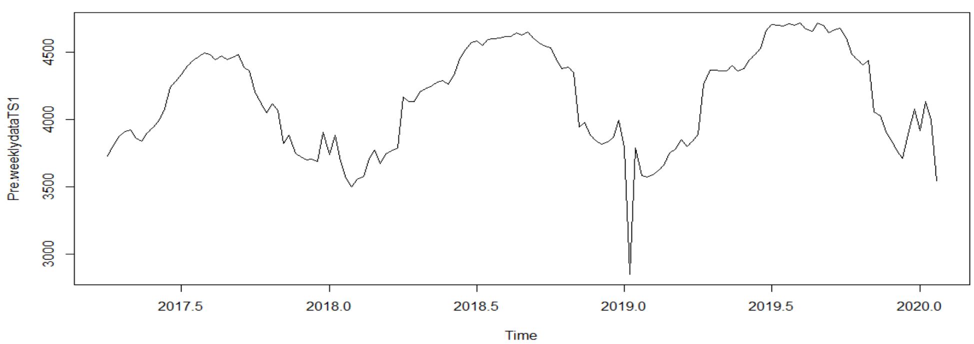

Figure 6 - Lisbon Weekly Flights after the 1st Break

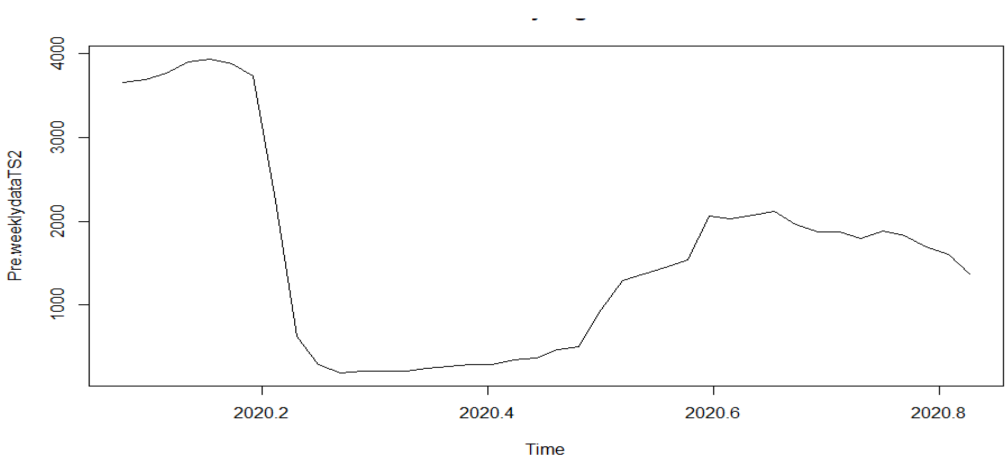

Figure 4 - Lisbon Weekly Flight Data after the 2nd Break

Once the new time series were defined, both Autocorrelation and Unit Root Tests were computed once again for each period in order to check its correlation and stationarity. The same method was applied to the new periods, which meant ACF and PACF were the designated tests for correlation; and ADF, KPSS and PP tests were performed to check the presence of Unit Root.

The pre-break period was the first to be submitted to the tests, with ACF and PACF represented in figure 7. Similarly to the whole time series, the pre break period also expresses stronger correlation levels in lower lag values, with progressively weaker correlation as the number of lags increase.

In terms of Unit Root Tests, the decision of the Dickey-Fuller points to the non-stationarity of the series since the calculated values are lower than the tabulated values. The presence of Unit Root may indicate the need for differencing the data in order to transform the series into a stationary one. By taking the first difference through the *diff* function in R, the data was again submitted to a Dickey Fuller Test, with the results expressing stationarity within the series since the calculated values were higher than the tabulated ones, thus rejecting the null hypothesis stating the existence of Unit Root. The PP test has brought the same conclusions regarding the existence of Unit Root in the time series, as the first test registered lower calculated compared to tabulated values and thus not rejecting the null hypothesis, with the first differenced data pointing to a stationary series by rejecting the existence of a Unit Root. Differencing has proven to be a useful manner to eliminate trends and enable stationarity within the time series.

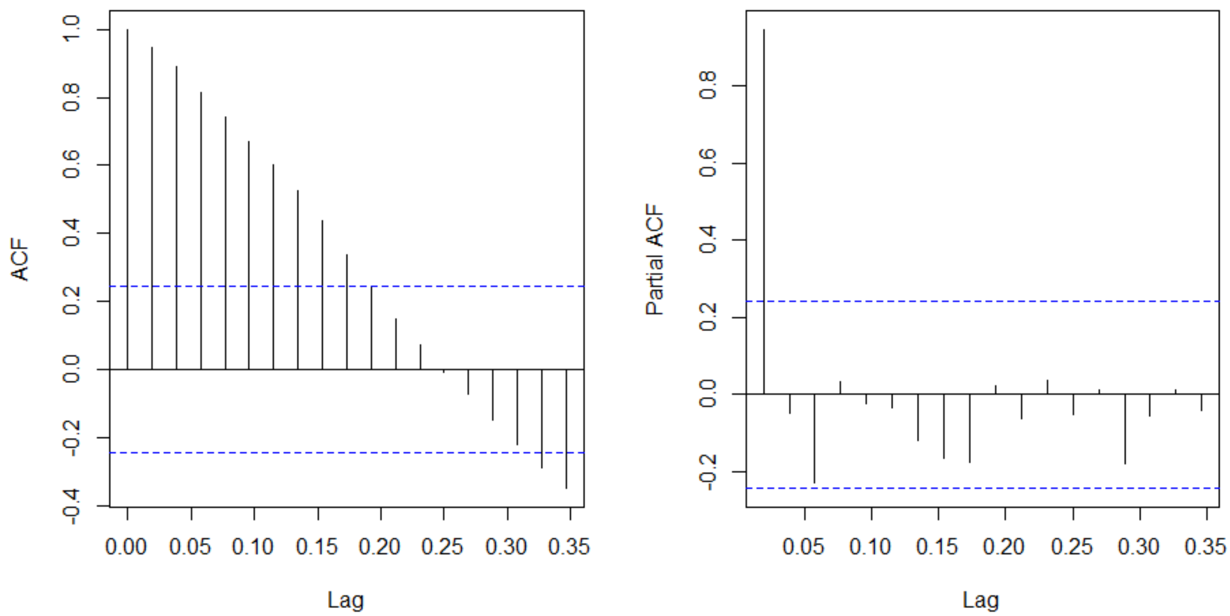

Figure 7 - ACF and PACF tests for Pre-break

On the other hand, the KPSS test rejected the hypothesis of stationarity as the calculated values exceeded the tabulated, leading to a first difference application of the test in which the null hypothesis could not be rejected, and the series was considered stationary.

The same methodology was applied to the second sample, containing the time series between the two breaks – from week 66 to week 212. The ACF and PACF plots are represented in figure 8. Similarly to the pre-break data the Break 1 time series, also demonstrates stronger correlation indexes in lower lag values and decreasing values as the number of lags increases.

Equivalently to the results expressed by the applied Unit Root Tests on the Pre-Break period, Break 1 only rejected the null hypothesis of non-stationarity for the first differenced data through ADF and PP tests. Additionally, the KPSS test rejected the stationarity hypothesis for the original time series and the contrary result for the differenced data.

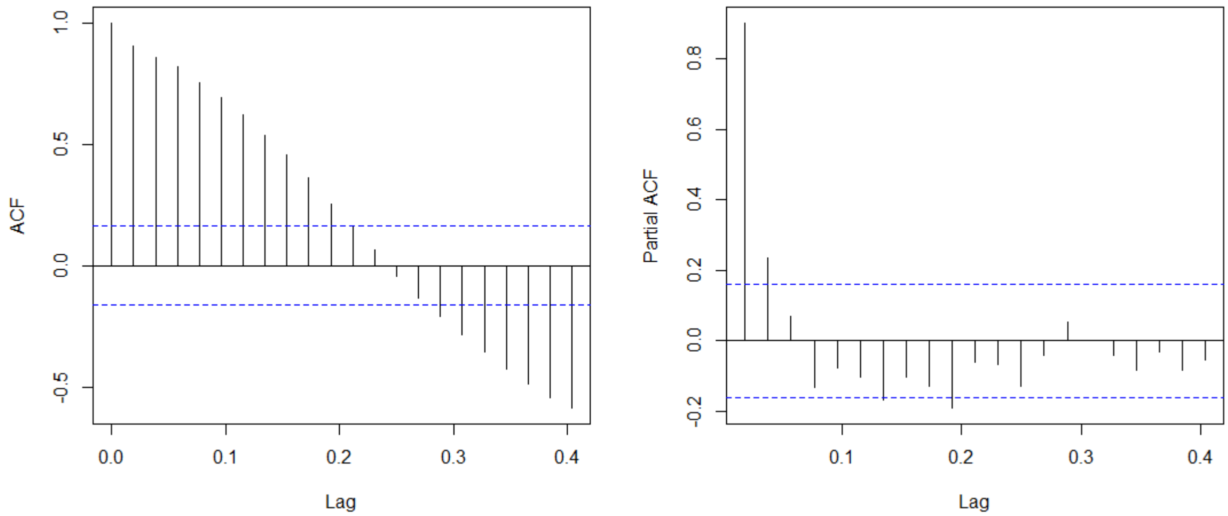

Figure 9 - ACF and PACF for Break

The third time series, designated as Break 2 which comprises weeks 213 to 252, was also put through the exact same process with similar results. The ACF and PACF tests are shown on figure 9.

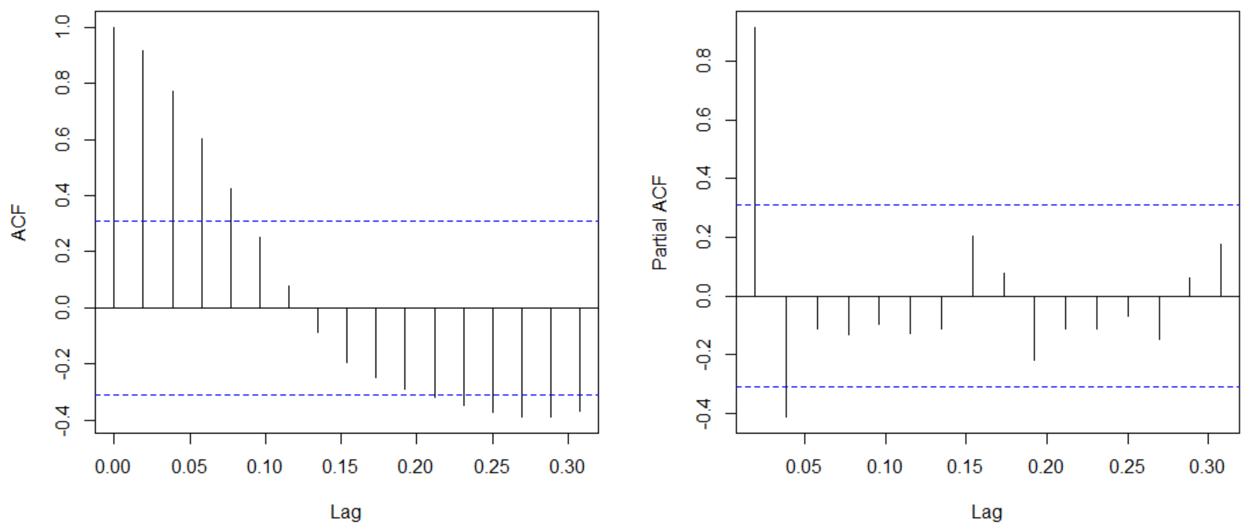

Figure 8 - ACF and PACF for Break 2

The Unit Root Tests also required a transformation into a first order differenced data in order to meet the time series stationarity requirements. Differenced data influenced the estimation of each ARIMA model since the level of differentiation is determined by the presence of Unit Root. The three time series could be considered stationary if the first difference was calculated.

Since ARIMA (p,d,q) is given by the Autoregressive (p), Moving Average (q), and the level of differencing (d), the models we consider are ARIMA (p,1,d) in order to match the required 1st order differencing to make the series stationary.

The ARIMA (p,d,q) model is restricted to non-seasonal data, and since the weekly flight data is highly seasonal, the regular part of the model is not enough to capture all the dependence in our data. By using a SARIMA (p,d,q) (P,D,Q) model, it is possible to incorporate also the seasonal components and improve the model accuracy.

The ARIMA models developed though the pre-break period were based on the previous autocorrelation functions and the results of the Unit Root tests. The list is presented below:

- ARIMA (1,1,2) $(0,1,1)_{52}$
- ARIMA (1,1,2) $(0,1,2)_{52}$
- ARIMA (1,1,3) $(0,1,1)_{52}$
- ARIMA (1,1,3) $(0,1,2)_{52}$

The pre-break period proved to be the most challenging for creating the forecasting model, as the number of periods to predict is far superior when compared to the other two periods, and the level of accuracy of latter stages can be impacted by the long-term prediction.

The second sample comprises the period between the first and second breaks, ranging from week 66 to week 66. The order of the models was found in the same way as in the previous period, were not only chosen manually, but also through the *auto.arima* function in R. This function returns the best ARIMA model for each time series, but it should not limit our model development to the assigned ARIMA. The list of models for Break 1 are listed below.

- ARIMA (0,1,1) $(0,1,0)_{52}$ – developed through the *auto.arima* function
- ARIMA (1,1,1) $(0,1,0)_{52}$
- ARIMA (2,1,1) $(0,1,0)_{52}$
- ARIMA (3,1,1) $(0,1,0)_{52}$
- ARIMA (2,1,2) $(0,1,0)_{52}$

- ARIMA (1,1,2) $(0,1,0)_{52}$
- ARIMA (1,1,1) $(0,1,1)_{52}$
- ARIMA (1,1,2) $(0,1,1)_{52}$
- ARIMA (1,1,3) $(0,1,1)_{52}$
- ARIMA (1,1,3) $(0,2,1)_{52}$
- ARIMA (1,1,3) $(0,2,1)_{52}$
- ARIMA (3,1,2) $(0,1,1)_{52}$

The last sampled period was modelled through the same process as the remaining two, including the *auto.arima* function and the variety of ARIMA models. The period comprises week 213 until week 252, leading the forecasting period to be the one with fewer samples, if we consider the end of 2021 as the timeline objective. The only difference is the impossibility of creating a seasonal ARIMA or SARIMA(p,d,q) (P,D,Q) model, as the minimum number of observations for this model is 50, and this time-series only comprises 39. The list for Break 2 is displayed below:

- ARIMA (1,0,1) – chosen through the *auto.arima* function
- ARIMA (1,1,1)
- ARIMA (3,1,1)
- ARIMA (2,1,2)
- ARIMA (1,1,2)
- ARIMA (1,1,3)
- ARIMA (3,1,2)
- ARIMA (1,1,1)
- ARIMA (1,2,1)
- ARIMA (2,2,1)
- ARIMA (2,2,2)

Once each ARIMA model was estimated, the forecasting was available through another function in R. The *forecast* command allowed to predict future values for the ARIMA models by computing the respective model and the *n* number of periods ahead. Comparably to the remaining processes, each of the forecasting models was based on the three distinct time series, given *by forecast (model, h=n).*

The first forecasting models were applied to the Pre-Break time series and subsequent ARIMA models. The data sample for this time series comprises the weekly flights from the 3rd of January 2016 until the 25h of March 2017, which meant the highest number *n* of predicted samples. Until the end of 2021, 248 periods had to be predicted in order to meet the forecasting timeline, represented in figure 10.

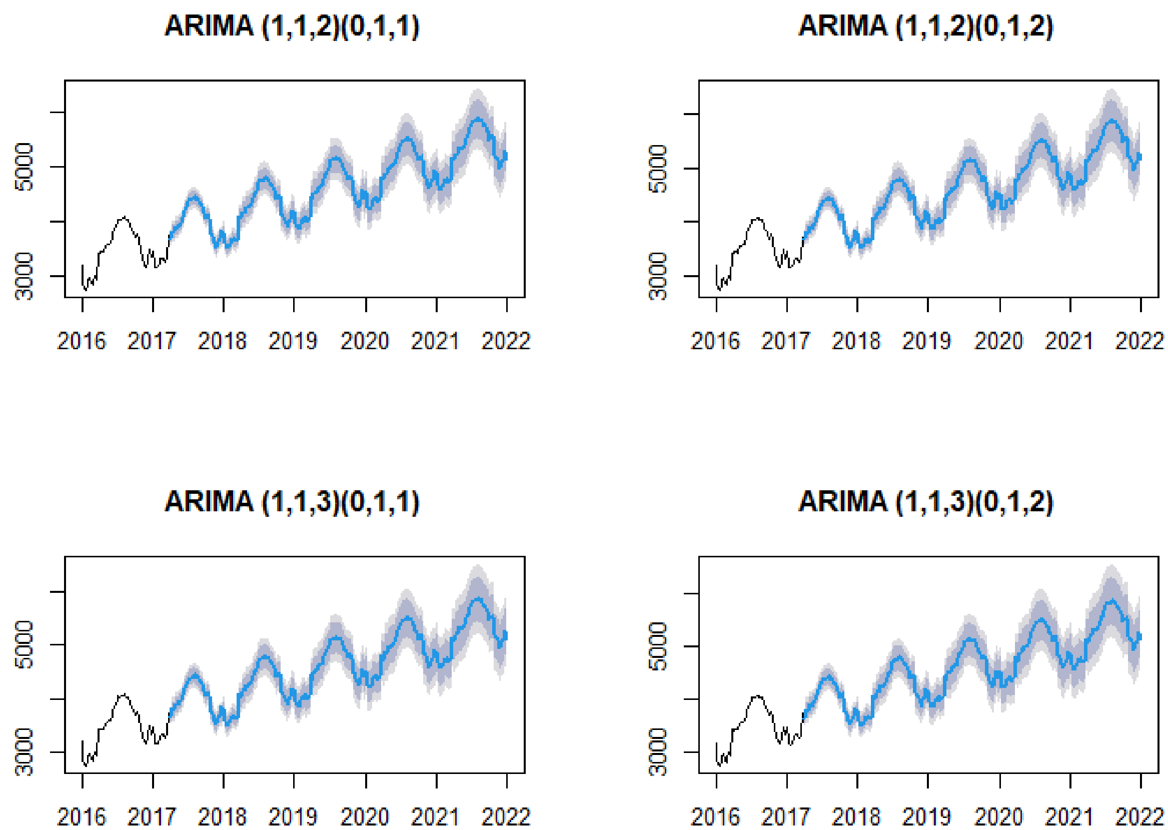

Figure 10 - Forecasting Models for Pre-Break period

Afterwards, the forecasting models for Break 1 were also developed, by determining the number of periods ahead. In this case, the models were estimated with 100 weeks ahead, corresponding to the number of weeks between the 27th of January 2020 and the last week of 2021. Ten different models were developed for the forecasting estimation, each with different parameters and outcomes. Figure 11 gives a graphic representation of each model.

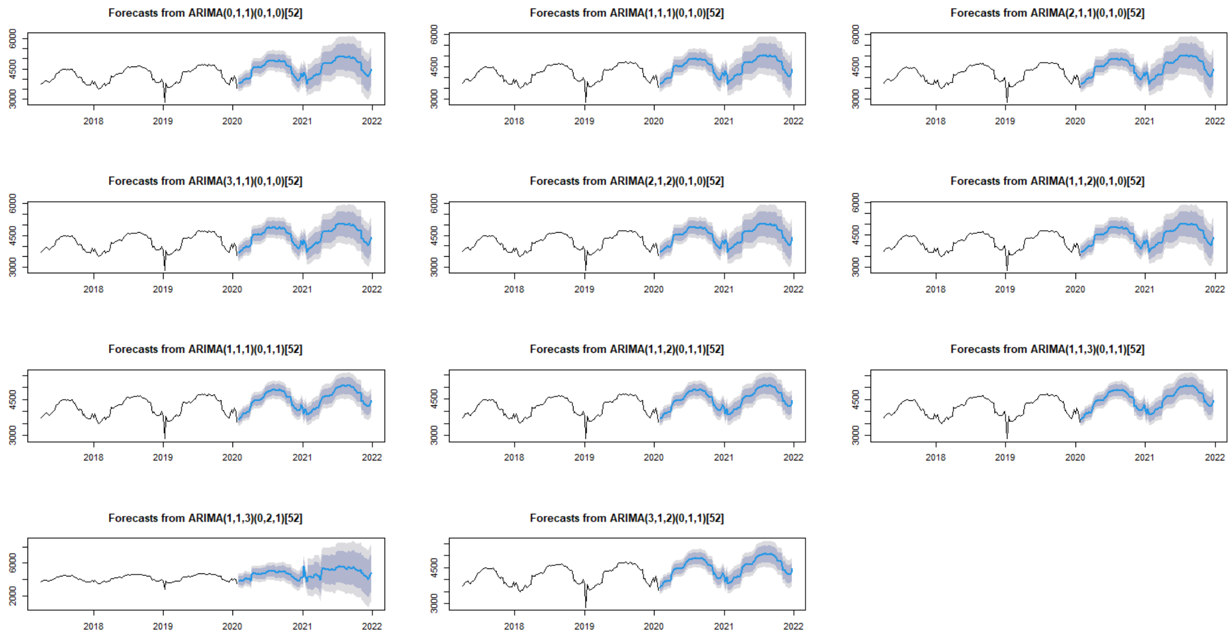

Figure 11 - Forecasting models for Break 1

The forementioned period encompasses the largest sample of data, as it comprises weekly data between the 26th of March 2017 and the 26th of January 2020, accounting in almost 3 years of information.

The last forecasting models were based on the Break 2 time series, comprising 49 samples between 28th of January 2020 and 31st of December 2020. Similarly to the remaining two periods, the forecasting model aims at the end of 2021, consisting in 52 periods to predict. Figure 12 represents the models for Break 2.

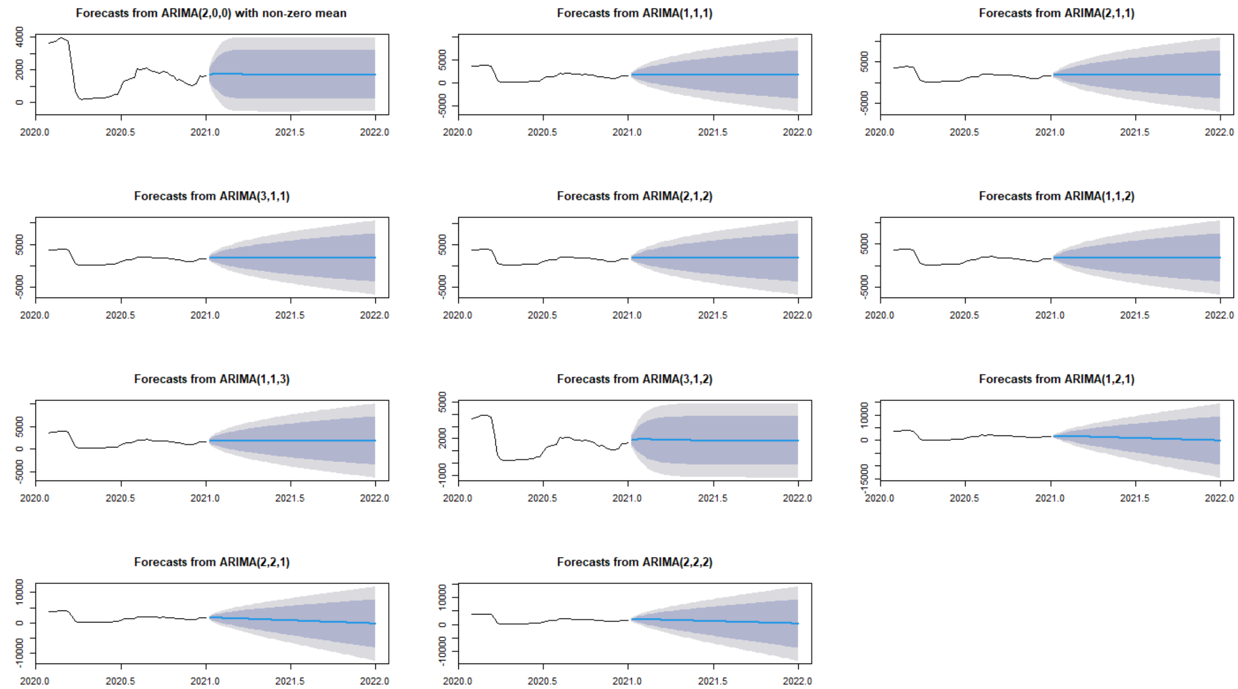

Figure 12 - Forecasting models for Break 2

At first sight, the graphs for Break 2's models are vastly different from Pre Break and Break 1 models. This difference results from the fact Break 2 models could not be predicted as SARIMA, due to the lack of observed samples. As previously mentioned, Break 2 consists in 49 periods, which remains below the recommended 50-100 samples in order to compute a seasonal model such as Pre Break and Break 1.

4.2 Discussion

Three different sets of models have produced three distinct results in terms of forecasting values within the same time frame. However, the first two sets of models – Pre-Break and Break 1- are partially extant, since a portion of the forecasted samples can be verified with real data. For instance, the Pre-Break forecasted sample ranges from March 2017 until December 2022, which divides the data into extant – from March 2017 until December 2020 – and expost – from December 2020 onwards. The extant data serves as an indicator for the differences between real and predicted data. Nevertheless, the impact of COVID-19 could not be mirrored in any of the extent data since the model was not capable enough to consider this variable.

One of the objectives of this study is the model application by players in the airline industry such as airliners and airport authorities in order to predict future flight movements in a short to medium term.

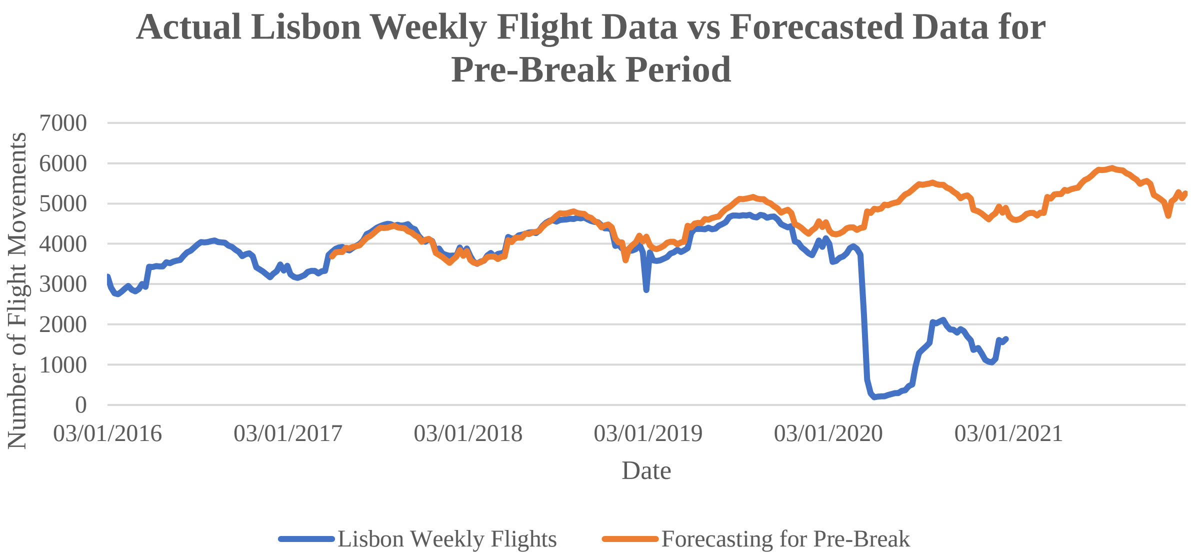

Figure 13 - Comparison between actual and predicted data for Pre-Break period

Extant data can be a good indicator of forecasting accuracy since it's possible to compare predicted values with real data. For this study, both Pre-Break and Break 1 periods will be compared with the actual data in order to observe major differences between the two sets.

Firstly, the Pre-Break forecasted sample ARIMA (1,1,2) $(0,1,1)_{52}$ was compared with the actual data as seen in Figure 13. If we consider the period between March 2017 and January 2020, we can clearly substantiate the similarities between the predicted and actual values. This similarity, however, is weaker over time, since ARIMA models are more accurate in short term predictions rather than long-term. In addition, the occurrence of COVID-19 clearly diverges the two sets of data, making the predicted values after January 2020 not possible to predict with this model.

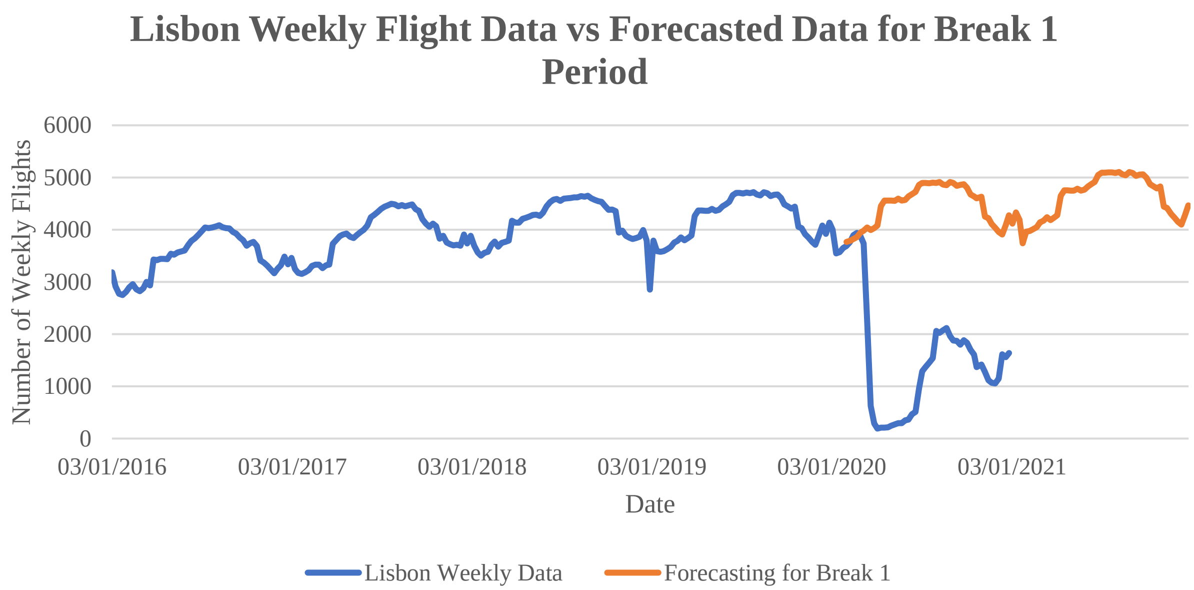

Figure 14 - Comparison between actual and predicted data for Break 1 period

When comparing Break 1's *auto.arima* function forecast with the actual data between January 2020 and the end of the same year, it's possible to conclude two different points regarding the visual representation:

- The effect of COVID-19 was registered almost immediately after the forecasting period;
- The forecasted values follow the same pattern as the previous actual data.

This analysis suggests that if it hadn't been for COVID-19, the model could actually perform significantly well on a short-term range, much like Pre-Break's extant forecasting period, whose predicted values were extremely close to the real flight data, as shown in Figure 13. Break 1's comparison with the real data is represented in Figure 14.

In order to better analyse the two forecasted samples and the actual data, a graphic representation of the three data sets is shown in Figure 15. The comparison allows a deeper understanding of not only how the forecasted values relate with real data, but with each other as well.

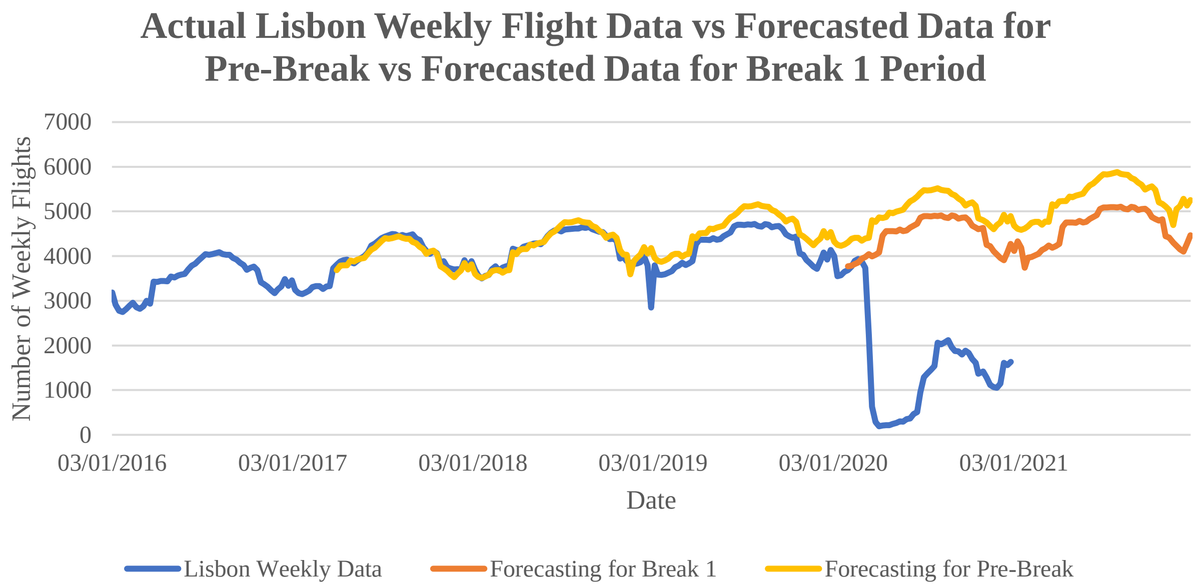

Figure 15 - Comparison between real data and the two forecasted periods

When considering the first two years of forecasted data on the Pre-Break period it's possible to verify the strong correlation with real data, since the two lines follow the same pattern with overlapping values. However, after 2019 the lines diverge, and the forecasting accuracy subsequently reduces. Whilst the forecast for Break 1 cannot be accurately compared with existing data because of COVID-19, it's possible to overview how the two distinct forecasts have extremely similar behaviors, with Pre-Break registering higher values for the same time periods. This can be explained by the fact Pre-Break forecast values tend to increase more than the actual data over time, meaning when Break 1 data becomes forecasted values in January 2020, Pre-Break has slightly higher values for the same period. Since both models were constructed based on similar methodologies, it was expected their behavior would be somewhat similar and matching the profile of an airport's yearly flight evolution – seasonal with higher peaks during the summer.

Finally, Break 2' forecasting model was the most difficult to evaluate in terms of accuracy, mostly because of the lack of extant data, leading to the impossibility of comparing real and forecasted values for the year 2021. Additionally, the period sample does not contain sufficient observations in order to construct a seasonal model and match it with the traditional behavior of an airport's yearly flight movement data. The first weeks following the outbreak of COVID-19 were extremely uncertain and difficult to predict, since most national governments had banned travelling between countries, and air travel was extremely restricted, and since Break 2 period only contains values inside the COVID-19 time frame, it was not possible to establish a trend and predict the expost period comparable to the other two models. The collected data for this study did not allow the construction of an accurate model for post-COVID period, as data from at least the first six months of 2021 is required to understand the rate at which airline movements return to normality.

However, it's quite important to understand how the pandemic impacted airport operations since it's not empirically correct to assume the model can be immediately applied after the last period data (end of 2020). Whilst the *forecast* function on R does not allow the prediction of a post-COVID model, if the trendline function is applied on Excel, it gives an estimation of future values, based on past observations. For this matter, the sample corresponding to Break 2 – from the end of January 2020 until the end of December 2020 – was selected and applied the trendline function, and the results are expressed on Figure 16.

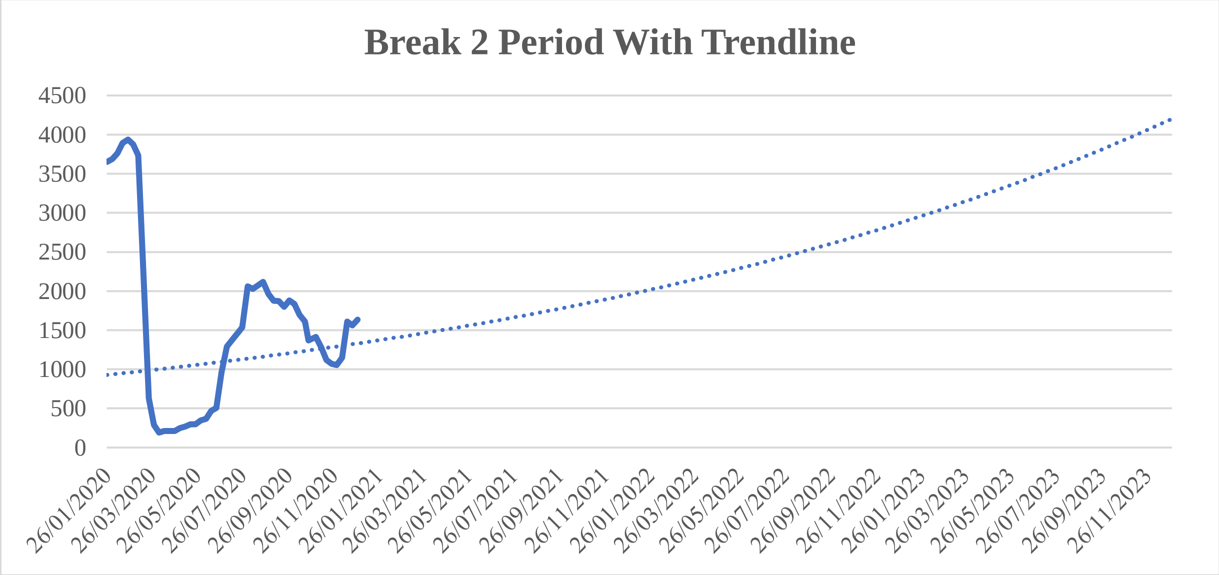

Figure 16 - Break 2 period and subsequent trendline

As it's possible to perceive on the figure above, the trendline estimates that the number of weekly flights can resume the pre-COVID figures around the end of 2023, resulting in a three-year recovery period for most airport and airliners, which coincides with studies conducted by Bloomberg's Angus Whitley, Jason Gale, Tara Patel, and Christopher Jasper; and IATA Director-General Alexandre de Juniac. In this specific situation, the last observations before the major drop are being considered, with around 4000 weekly flight movements. The trendline estimates a return to 4000 weekly flights around October of 2023.

Once the expected normality resumes, the constructed model can be applied to the specific situation and consequently accurately predict the number of weekly flight movements at Lisbon's Humberto Delgado Airport on a short to medium term basis.

5. Conclusions

The objective of this Management of Services and Technology dissertation was to develop an accurate prediction model, capable of forecasting the number of registered flight movements at Lisbon's Humberto Delgado International Airport, based on weekly data ranging from January 2016 until December 2020.

The weekly flight data consisted in 261 different periods, which were then classified as a time-series and subsequently tested for stationarity. The presence of structural breaks was also tested, leading to two distinct breaks being defined within the time series. The two breaks resulted in the dissection of the original time series in three different periods – Pre Break, Break 1 and Break 2 – which were then tested again for stationarity. The three different periods resulted in three different forecasting models being developed and each one's accuracy was determined separately.

Developing prediction models with no extant data to compare leaves question marks regarding its accuracy. Of the three forecasting models, only one of them – Break 2 – did not have any real data to be compared with, which leaves Pre-Break's model in clear advantage for having the largest amount of real data to compare the predicted values with. This analysis has determined that the Pre-Break model performs significantly well on a short-term range, since the first two predicted years – 2017 and 2018 – expressed extremely similar results to the ones collected from the actual data.

Additionally, Break 1's forecasted sample could also be compared with existing data, if it was not for the presence of COVID-19. The model's behavior was extremely similar to Pre-Break's, however, the extant data suffered severe alterations leading to a wide gap between real and forecasted data since the model was unable to predict the occurrence of COVID-19 and its devastating consequences for the industry. The impact of this pandemic was also felt on the last forecasted model – Break 2's – as the data sample for this period was simply not enough to create a seasonal model, and the uncertainty revolving around COVID in the early days of 2021 resulted in a significant accuracy reduction for this model's outcome.

Nevertheless, the devastating impact of this pandemic cannot be overseen since it has been affecting airports, airliners and ultimately every single passenger. Whilst not recurring to this study's basis model, it was still possible to determine a future trend based on the post-COVID outbreak period, which predicted normality to resume in the latter stages of 2023. If we consider the early weeks of 2020 as the critical point regarding the virus' outbreak, it means almost 4 full years where the aviation industry was significantly wounded by the negative impact of government travel restrictions or improvised airport layouts to accommodate testing and distancing, which all connect to the number of flights each airline and subsequently airport will register in this period.

Since most airports have similar behaviors regarding air traffic patterns, the first two models are applicable to most regional, domestic and international airports. The models can be used as a beneficial tool for the majority of players in the airline industry, since the number of flights and consequent passenger volume has a crucial impact in decision making. Accurately predicting the number of flight movements can assist airliners and airports in short to medium term decisions such as terminal expansions, planning new routes or even acquiring new aircraft. The next few years are crucial for this industry in order to recover from the devastating effects of COVID-19 and using an accurate forecasting model can be an interesting tool to rapidly recover.

Lastly, the first two models are applicable to most airports, as long as the traffic patterns has similarities to the one registered in Lisbon – where there's clearly a seasonal component with higher peaks during the summer months, contrasting with a lower peak during winter months. Despite the extant data relevantly regarding the predicted data as accurate, the models cannot correctly reflect the future short-term prediction due to the impact of COVID-19, meaning the values for 2021, 2022 and 2023 will most definitely differ in terms of prediction and reality. The uncertainty regarding the pandemic has consisted in a major limitation for this study due to the irregular patterns of flight movements which consequently interdict the immediate application of a model.

The main conclusion of this study can be mirrored by the possibility of applying an accurate forecasting model to any airport following a standard seasonal trend on a short to medium term basis, whilst at this precise moment in time this application is suspended by the negative impact of COVID-19, which has brough high uncertainty affecting all players in the aviation industry. This effect has consequently restricted a correct prediction starting in early 2021, which implies only once the pre-pandemic figures are restored the model can truly be applied.

6. References

1. Andreoni, A. and Postorino, M.N., 2006. Time series models to forecast air transport demand: A study about a regional airport. IFAC Proceedings Volumes, 39(12), pp.101-106.

2. Belobaba, P., Odoni, A., Barnhart, C., 2015. The Global Airline Industry, second ed. http://dx.doi.org/10.1002/9780470744734. New York.

3. Benedetti, G., Gobbato, L., Perboli, G. and Perfetti, F., 2012. The Cagliari Airport impact on Sardinia tourism: a Logit-based analysis. Procedia-Social and Behavioral Sciences, 54, pp.1010-1018.

4. Berry, W.D., Feldman, S. and Stanley Feldman, D., 1985. Multiple regression in practice (No. 50). Sage.

5. Berster, P., Gelhausen, M., Wilken, D., 2013. Is increasing seat capacity common practice of airlines at congested airports? J. Air Transp. Manage. 46, 1–25.

6. Celik, Y., Hotchkiss, D.R., 2000. The socio-economic determinants of maternal health care utilization in Turkey. Soc. Sci. Med. 50 (12), 1797–1806.

7. D'Alfonso, T., Jiang, C., Wan, Y., 2013. Airport pricing, concession revenues and passenger types. J. Transp. Econ. Policy (1), 71–89

8. De Gooijer, J.G. and Hyndman, R.J., 2006. 25 years of time series forecasting. International journal of forecasting, 22(3), pp.443-473.

9. Delgado, L., Gurtner, G., Cook, A., Martín, J. and Cristóbal, S., 2020. A multi-layer model for long- term KPI alignment forecasts for the air transportation system. Journal of Air Transport Management, 89, p.101905.

10. Dwivedi, Y.K., Lal, B., 2007. Socio-economic determinants of broadband adoption. Ind. Manage. Data Syst. 107 (5), 654–671.

11. Efron, B. and Tibshirani, R.J., 1994. An introduction to the bootstrap. CRC press.

12. Engels, G., Förster, A., Heckel, R., Thöne, S., 2005. Process modeling using UML. In: Dumas, M., van der Aalst, W.M.P., ter Hofstede, A.H.M. (Eds.), Process Aware

Information Systems: Bridging People and Software through Process Technology. Wiley & Sons, Chichester, pp. 85–118.

13. Fattah, J., Ezzine, L., Aman, Z., El Moussami, H. and Lachhab, A., 2018. Forecasting of demand using ARIMA model. International Journal of Engineering Business Management, 10, p.1847979018808673.
14. Gelhausen, M.C., Berster, P. and Wilken, D., 2018. A new direct demand model of long-term forecasting air passengers and air transport movements at German airports. Journal of Air Transport Management, 71, pp.140-152.
15. Ghobrial, A., 1997. A model to forecast aircraft operations at general aviation airports. Journal of Advanced Transportation, 31(3), pp.311-323.
16. Goedeking, P., 2010. Networks in Aviation: Strategies and Structures, Networks in Aviation: Strategies and Structures.
17. Granger, C.W. and Jeon, Y., 2007. Long-term forecasting and evaluation. International Journal of Forecasting, 23(4), pp.539-551.
18. Hamilton, J.D., 2020. Time series analysis. Princeton university press.
19. Hofer, C., Kali, R. and Mendez, F., 2018. Socio-economic mobility and air passenger demand in the US. Transportation Research Part A: Policy and Practice, 112, pp.85-94.
20. Horv, L. and Kokoszka, P., 2003. GARCH processes: structure and estimation. Bernoulli, 9(2), pp.201-227.
21. Jaseena, K.U. and Kovoor, B.C., 2020. Deterministic Weather Forecasting models based on Intelligent Predictors: A Survey. Journal of King Saud University-Computer and Information Sciences.
22. Jin, F., Li, Y., Sun, S. and Li, H., 2020. Forecasting air passenger demand with a new hybrid ensemble approach. Journal of Air Transport Management, 83, p.101744.
23. Kurdel, P., Sedláčková, A.N. and Novák, A., 2020. Analysis of using time series method for prediction of number of passengers at the airport. Journal of KONBiN, 50(3), pp.203-216.

24. Layton, A., Worthington, A., 1999. The impact of socio-economic factors on gambling expenditure. Int. J. Soc. Econ. 26 (1), 430–440.

25. Morrison, S.A., 2001. Actual, adjacent, and potential competition: estimating the full effect of Southwest Airlines. J. Transp. Econ. Policy 35 (2), 239–256.

26. Neapolitan, R., E. (1989), Probabilistic Reasoning in Expert Systems: Theory and Algorithms. A Wiley-Interscience Publication. John Wiley and Sons, Inc., New York

27. Nieto, M.R. and Carmona-BenÃtez, R.B., 2018. ARIMA+ GARCH+ Bootstrap forecasting method applied to the airline industry. Journal of Air Transport Management, 71(C), pp.1-8.

28. Perboli, G. Musso, S., & Perfetti, F (2011b). Analysis of the economic impact of airport services in Sardinia: the cagliari airport case study. Proceedings of Routes Europe 2011, Cagliari, Italy.

29. Piketty, P., Saez, E., 2003. Income Inequality in the United States, 1913–1998. Quart. J. Econ. 118 (1), 1–39.

30. Rodríguez-Sanz, Á., Comendador, F.G., Valdés, R.A. and Pérez-Castán, J.A., 2018. Characterization and prediction of the airport operational saturation. Journal of Air Transport Management, 69, pp.147-172.

31. Rodríguez-Sanz, Á., Fernández, B.R., Comendador, F.G., Valdés, R.A., García, J.M.C. and Bagamanova, M., 2018. Operational Reliability of the Airport System: Monitoring and Forecasting. Transportation research procedia, 33, pp.363-370.

32. SESAR, 2014b. SESAR Performance Framework (Edition 2). Single European Sky ATM Research Consortium, Brussels.

33. Sivrikaya, O. and Tunç, E., 2013. Demand forecasting for domestic air transportation in Turkey. The Open Transportation Journal, 7(1).

34. Stephenson, T.A., 2000. An introduction to Bayesian network theory and usage (No. REP_WORK). IDIAP.

35. 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, pp.400-416.

36. Veerappan, M., Sahu, P.K., Pani, A., Patil, G.R. and Sarkar, A.K., 2020. Analysing and modelling the relationship between air freight movement and airport characteristics in India. Transportation Research Procedia, 48, pp.74-92.

37. Visser, H., Dangendorf, S. and Petersen, A.C., 2015. A review of trend models applied to sea level data with reference to the "acceleration-deceleration debate". Journal of Geophysical Research: Oceans, 120(6), pp.3873-3895.

38. Wadud, Z., Graham, D.J., Noland, R.B., 2009. Modelling fuel demand for different socio-economic groups. Appl. Energy 86 (12), 2740–2749.

39. Wilke, S., Majumdar, A., Ochieng, W.Y., 2014. Airport surface operations: a holistic framework for operations modeling and risk management. Saf. Sci. 63, 18–33. http://dx.doi.org/10.1016/j.ssci.2013.10.015.

40. Williams, B.M. and Hoel, L.A., 2003. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. Journal of transportation engineering, 129(6), pp.664-672.

41. 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, pp.20-26.

42. Ye, B., Liu, B., Tian, Y. and Wan, L., 2020. A methodology for predicting aggregate flight departure delays in airports based on supervised learning. Sustainability, 12(7), p.2749.

43. Zanin, M., Zhu, Y., Yan, R., Dong, P., Sun, X. and Wandelt, S., 2020. Characterization and prediction of air transport delays in China. Applied Sciences, 10(18), p.6165.