Long-term and Short-term Forecasting Techniques for Regional Airport Planning

ROBIN WARGENTIN
Long-term and Short-term Forecasting Techniques for Regional Airport Planning

ROBIN WARGENTIN

Master’s Thesis in Mathematical Statistics (30 ECTS credits)
Master Programme in Applied and Computational Mathematics (120 credits)
Royal Institute of Technology year 2016
Supervisor at University of Bologna: Luca Mantecchini, Antonio Danesi
Supervisor at KTH: Timo Koski
Examiner: Timo Koski

TRITA-MAT-E 2016:45
ISRN-KTH/MAT/E--16/45-SE

Royal Institute of Technology
SCI School of Engineering Sciences
KTH SCI
SE-100 44 Stockholm, Sweden
URL: www.kth.se/sci
Abstract

The aim of this thesis is to forecast passenger demand in long term and short term perspectives at the Airport of Bologna, a regional airport in Italy with a high mix of low cost traffic and conventional airline traffic. In the long term perspective, a time series model is applied to forecast a significant growth of passenger volumes in the airport in the period 2016-2026. In the short term perspective, time-of-week passenger demand is estimated using two non-parametric techniques; local regression (LOESS) and a simple method of averaging observations. Using cross validation to estimate the accuracy of the estimates, the simple averaging method and the more complex LOESS method are concluded to perform equally well. Peak hour passenger volumes at the airport are observed in historical data and by use of bootstrapping, these are proved to contain little variability and can be concluded to be stable.
Sammanfattning

Målet med denna uppsats är att prognosticera passagerarefterfrågan i lång- och kortsiktigt perspektiv på Bologna Flygplats, en regional flygplats i Italien med hög mix av lågkostnadsbolag och konventionella flygbolag. I det långsiktiga perspektivet appliceras en tidsseriemodell som prognosticera hög tillväxt i passagerarvolymer på flygplatsen under perioden 2016-2026. I det korta perspektivet uppskattas efterfrågan utefter tid i veckan med hjälp av två icke-parametriska modeller; local regression (LOESS) och en simpel metod som beräknar medelvärdet utav observationer. Med cross validation uppskattas precisionen i modellerna och det kan fastställas att den simpla medelvärdesmetoden och den mer avancerade LOESS-metoden har likvärdig precision. Passagerarvolymer på flygplatsen under högtrafik observeras i historisk data och med hjälp av bootstrapping visas att dessa volymer har låg variabilitet och det kan fastställas att de är stabila.
Acknowledgements

I would like to send special thanks to my supervisors Luca Mantecchini and Antonio Danesi at the University of Bologna, without whose help and ideas this thesis would not have been possible. Further I wish to thank Timo Koski, my supervisor at the department of statistics at KTH, for his support through the work.
1 Introduction

The annual number of passengers traveling with commercial air transport has increased substantially in recent years and is expected to continue increasing, with regional airports experiencing extra strong growth. Both the number of flight movements and the average load factor of each flight are increasing. In the Airbus forecast of 2015-2034, the global number of revenue passenger kilometers (RPK), is expected to double between 2014 and 2034, while the intra-Central European market is forecast to experience 4.4% annual growth[1]. The growth in demand for air traffic is partly driven by macroeconomic factors such as increased globalization and the change of travel behaviour following from demographic changes during economic upswings, particularly in Asian and eastern European economies. Another factor is the introduction in the 1990s of Low Cost Carriers (LCC) such as Ryanair and Easyjet, which has stimulated demand by introducing low fare flights. The price pressure has proved challenging to the established airlines, often referred to as Former Flag Carriers (FFC) or Legacy Carriers, and has lead to an industry-wide lowering of fares. As airlines search to reduce costs, regional airports have experienced an increase in attractive power; since smaller and less used airports don’t experience the congestion found at bigger airports, operating at these often increases productivity for the airlines. For example, in the Frankfurt-London route, Ryanair flying between Stansted-Hahn has 33% better productivity of aircraft and crew than Lufthansa has flying between the bigger airports Heathrow-Frankfurt. This is due to the less time spent being idle in queues, both on ground and in the air[6]. From the perspective of the management of a regional airport, the fast growth in number of passengers puts pressure on an effective planning of the capacity of the airport. Capacity improvements in airport infrastructure represents large and lumpy capital investments and long-term forecasts of passenger volumes and peak hour volumes are therefore of high importance[9].

Itinerary scheduling and congestion planning are also essential aspects for the airport management. Traditionally, airports have been separated into hubs and spokes and this has determined much of the scheduling for regional airports, which are generally considered as spokes. In recent development, however, the separation between hubs and spokes has become less distinct. Within the hub and spoke-paradigm, passengers who wish to traverse between two spoke airports that are not directly connected to each other, are directed to a hub to take an interconnecting flight. In effect, hubs collects passenger demand from its connected spokes and redirects it to the desired spoke destinations. To synchronize transfers, hub scheduling is organized such that flights from spokes arrive simultaneously in a small time window and then depart in another small time window. This results in planned waves of arrivals and departures at the hubs with very concentrated passenger flows and risk of congestion. Because flight scheduling in this system is done with prioritization on time of arrival at the hub, the wave dynamics of passenger flow are less pronounced at spoke airports. This hub and spoke system used to be the system maintained by national flag carriers, as they centered their operations
around one hub airport. However, LCC airlines tend not to use the paradigm of hub and spoke scheduling for cost reasons[7]. Since LCC is growing its share of the market, the hub and spoke separation is becoming less distinct. In light of this, it is of growing interest for airport management to understand how passenger demand varies during the week and how concentrated the passenger flows are, in order to plan operations.

Another topic of high importance for the aviation industry is its impact on the environment. Due to its international nature, the aviation industry is generally exempt from national CO2-targets established in the Kyoto Protocol and other agreements. In combination with the heavy growth of the industry, air transport poses a serious threat to the 2°C target on global warming which has been set by IPCC. Although several international organizations, for example International Civil Aviation Organization (ICAO) and EU, work towards implementing measures such as CO2 emission trading and carbon neutral growth, the process is slow. And while a lot is invested in developing more efficient technology solutions for the industry, technological progress in itself is unlikely to improve the situation to a satisfying level. Bows-Larkin et. al [3] make the conclusion that ”the aviation industry’s current projections of the sector’s growth are incompatible with the international community’s commitment to avoiding the 2°C characterization of dangerous climate change”. They further argue that there is a clear role for demand management in aviation, i.e. attempting to reduce demand by increasing fares throughout the industry.

1.1 Objectives

This thesis provides techniques for forecasting passenger demand at a regional airport on long-term and short-term basis. A long term forecast of passenger demand on a quarterly level is obtained using a seasonal ARIMA time series model. A non-parametric predictive model of passenger demand during the times of the week is created by local regression technique (LOESS) as well as by a simpler average value technique. Further, estimates of the annual peak-hour passenger flow (Standard Hour Rate and Busy Hour Rate) are obtained, and the variability in these estimate is analyzed using bootstrapping. The techniques are applied on data from Bologna Guglielmo Marconi Airport, a large size regional airport in the Emilia-Romagna region of Italy that handled 6.9 million passengers in 2015. The airport has a high mix of LCC and former flag carrier traffic. In 2015, 37% of the flights were operated by LCC airlines, and they carried 52% of the passengers at the airport. While intergovernmental demand management aimed at reducing the aviation industry’s environmental impact would be highly relevant to the topic of passenger forecasts, its impact is out of the scope of this thesis.
2 Literature Review

2.1 Forecasting Techniques

Because of its high economic relevance, the field of forecasting air traffic demand is widely explored. However, no single technique holds the place as a standard method for forecasting. For example, executive judgement, the judgement of a person with some specific knowledge of the route or market in question, is still one of the techniques most widely used\cite{7}. Academic research tends to focus on statistical methods but also here the approaches differ. For example, Xie, Wang & Lai obtain a short-term forecast of passengers by using hybrid seasonal decomposition and support vector regression\cite{16}. Profiliidis uses traditional and fuzzy regression models to forecast the passenger demand\cite{12}. Andreoni & Postorino produce a multivariate ARIMA model with GDP per capita and number of flight movements as explanatory variables in order to forecast demand at a regional airport\cite{2}.

2.2 Time-of-Day Demand

Previous research on time-of-day demand has been made by for example Koppelman et al\cite{10}, who construct a model for the desirability of a flight itinerary based on qualitative factors of the flight including time of departure. In this model, the time of departure is modeled both as a dummy-variable for every hour of the day, and as a continuous combination of sine- and cosine-functions with estimated parameters. In short, these models indicate that mid-morning and late-afternoon flights are preferred, midday flights are moderately preferred while early-morning and late-evening flights are unpreferred by passengers. They find that the model based on sine- and cosine-functions significantly rejects the model with hour dummies as the true model. The authors also go on to present a schedule delay model that values the attractiveness of an itinerary based on how much it differs from assumed ideal departing times. The model gives however no insight in how the day of the week impacts the desirability of a flight.

2.3 Defining a Measure for Peak Hour

In order to translate a long term forecast into layout plans regarding size and function- alities of an airport, it is of interest to know the concentration of passengers in the peak hours. However, there are various ways to define the peak hour. The US Federal Aviation Administration (FAA) suggests a measure called Typical peak Hour Passengers (TPHP) which is calculated at a flat-rate based on annual traffic at the airport (Table 1). For a calculated typical peak hour rate, FAA then advises dimensions for different functions of the airport. The British Airport Authority (BAA), on the other hand, uses the Standard Busy Rate (SBR) as well as the Busy Hour Rate (BHR) as a measure of
peak hour passenger volumes. SBR is defined as the hour with the 30th highest passenger flow in a normal year, i.e. the hourly passenger flow that is surpassed only in 29 hours in a year. BHR is similarly defined as the hourly passenger flow that is surpassed only by 5% of the hours in a year[15]. It is also possible to modify the measure to the hour that is surpassed by 2.5% of the hours in a year. The SBR and BHR can be calculated directly from passenger volumes, or with other techniques. For example, Jones and Pitfield suggests a technique to estimate the BHR by using the average passenger load factors of flights and assuming that the number of hourly flight movements follow a normal distribution[9].

<table>
<thead>
<tr>
<th>Total Annual Passengers</th>
<th>TPHP as percentage of annual passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>More than 20 million</td>
<td>0.030%</td>
</tr>
<tr>
<td>10-20 million</td>
<td>0.035%</td>
</tr>
<tr>
<td>1-10 million</td>
<td>0.040%</td>
</tr>
<tr>
<td>0.5-1 million</td>
<td>0.050%</td>
</tr>
<tr>
<td>100,000-500,000</td>
<td>0.065%</td>
</tr>
<tr>
<td>Less than 100,000</td>
<td>0.12%</td>
</tr>
</tbody>
</table>

Table 1: Typical peak Hour Passenger rate (TPHP) as suggested by FAA

3 Methodology and Theory

3.1 Time Series Models

A time series is a data series \(\{y_t\}_{t=1}^T\) collected with equal time steps \(t = 1, \ldots, T\). By fitting a model, such as the ARIMA, to the data, forecasts of future values of \(y_t, t > T\) can be obtained. Below, important concepts in the analysis of time series are introduced, closely following Tsay[14].

3.2 Stationarity

In the analysis of time series, stationarity and weak stationarity are two properties of a time series that are of high importance. A time series \(\{y_t\}_{t=1}^T\) is said to be strictly stationary if the joint distribution of \(y_{t_1}, \ldots, y_{t_k}\) is invariant under time shifts, i.e. that the joint distribution of \(y_{t_1}, \ldots, y_{t_k}\) is identical to \(y_{t_1+t}, \ldots, y_{t_k+t}\) for all \(t\) and \(k > 0\). Strict
stationarity is a strong condition which is hard to verify in practice. In applications, it often suffices to verify that a time series is weakly stationary. A time series \( \{y_t\}_{t=1}^T \) is said to be weakly stationary if the mean and autocovariance of \( y_t \) are time-invariant, i.e. if

\[
E(y_t) = \mu, \quad t = 1, \ldots, T
\]

\[
\text{Cov}(y_t, y_{t-l}) = \gamma_l, \quad t = 1, \ldots, T
\]

If \( y_t \) is strictly stationary and \( E(y_t) < \infty \) and \( E(y_t^2) < \infty \) respectively, then \( y_t \) is also weakly stationary. The converse is not true in general, but holds in the special case when \( y_t \) is normally distributed.

3.3 ARIMA Models

The autoregressive integrated moving average model, ARIMA\((p,D,q)\), is formulated as

\[
\left(1 - \sum_{i=1}^{p} \phi_i B^i\right) \left(1 - B\right)^D y_t = \left(1 + \sum_{i=1}^{q} \theta_i B^i\right) a_t,
\]

where \( B \) is the lag operator, i.e. an operator that returns the previous element in the time series,

\[ B y_t = y_{t-1} \]

The parameter \( p \) represents the number of lags present in the autoregressive part of the model, \( D \) represents the order of integration and \( q \) represents the order of the moving average part of the model. \( \{a_t\} \) is assumed to be a white noise series with mean zero and variance \( \sigma_a^2 \).

3.4 Testing For Stationarity

Several alternative methods exist to test for stationarity in an observed time series. Examples include Augmented Dickey-Fuller (ADF) test, Phillips-Perron (PP) test and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test. In this work, we will stick to the ADF test. The model with lag \( p \) is formulated as

\[
y_t = \phi y_{t-1} + \beta_1 \delta y_{t-1} + \cdots + \beta_p \delta y_{t-p} + a_t
\]
The ADF tests the hypothesis

\[ H_0 : \phi = 0 \]
\[ H_1 : \phi < 0 \]

In other words, the null hypothesis is that the time series has a unit root while the alternative hypothesis states that it is a stationary process. The test should be performed with various values of \( p \) in order to account for different autoregressive lags in the model. The actual test is performed with a calculated test statistic versus tabulated values of a nonstandard distribution[14].

### 3.5 Akaike Information Criterion

The Akaike Information Criterion (AIC) is an information criterion used to determine the optimal setup of coefficients in a regressive model. For a model with \( k \) estimated parameters, used on a sample of \( T \) observations, AIC is defined as

\[
AIC = -\frac{2}{T} \ln(\text{likelihood}) + \frac{2}{T} \times k
\]

Here, likelihood is the maximum likelihood for parameters of the model. Based on the AIC criteria, the model that corresponds to the lowest value of AIC should be selected to represent the data.

### 3.6 Seasonal Time Series

Time series that are measured on a cyclical basis over the year typically follow a heavy seasonal pattern and the time series model needs to be adjusted to capture this. A common method to handle the serial correlation of the time series \( y_t \) is to use differentiation, i.e. \( \Delta y_t = y_t - y_{t-1} = (1 - B)y_t \). However, when the time series has a seasonal pattern of \( s \) steps, it will also have a high autocorrelation at lags \( k \cdot s \) for \( k = 1, 2, ... \). To adjust the time series for this behaviour, a further seasonal differencing can be applied as

\[
\Delta_s(\Delta y_t) = (1 - B^s)\Delta y_t = \Delta y_t - \Delta y_{t-s} = y_t - y_{t-1} - y_{t-s} + y_{t-s-1}
\]

With seasonal differencing as well as seasonal autoregressive and moving average terms, the seasonal model ARIMA\((p, d, q) \times (P, D, Q)_s\) is formulated as
The airline model is a special case of seasonal time series ARIMA$(0, 1, 1) \times (0, 1, 1)_4$, which is used as an example by Box, Jenkins and Reinsel[4]. It is formulated as

\[(1 - B^s)(1 - B) y_t = (1 - \theta B)(1 - \Theta B^s) a_t\]

where $a_t$ is a white noise with variance $\sigma^2_a$, $\theta$ and $\Theta$ are constants such that $|\theta| < 1$ and $|\Theta| < 1$. This is however not necessarily the model that best fits the data in this thesis.

### 3.7 Estimating Time Series Models

Maximum likelihood is commonly used to estimate time series models[14]. Assume that we wish to estimate some model parameters, which we collectively call $\theta$. We define the Likelihood function

\[F(y_1, \ldots, y_T; \theta) = F(y_1|\theta)F(y_2|y_1, \theta)\cdots F(y_T|y_{T-1}, \ldots, y_1, \theta) = F(y_1|\theta) \prod_{t=2}^{T} F(y_t|y_{t-1} \ldots y_1, \theta)\]  

(1)

By varying the parameters in $\theta$, we chose the model (parameter set) that corresponds to maximizing the likelihood function (1). In practice, it is computationally easier to find the logarithm of the likelihood function since it is additive instead of multiplicative. The maximum of the likelihood and the logarithm of the likelihood are obtained for the same value of $\theta$.

### 3.8 Local Regression Models (LOESS)

Local regression models can be used in order to find the relationship between a dependent variable $y$ and independent variables $t$ in a setting where it’s not practically possible to find a closed form function to describe the relationship. The following theory follows Cleveland, 1979[5]. We let $y_i$ for $i = 1, \ldots, n$ be observations of a dependent variable and let $(t_i, \ldots, t_p)$ for $i = 1, \ldots, n$ be corresponding independent variables. We further
assume that the data $y_i$ has a relationship to $t_i$ that can be expressed as $y_i = g(t_i) + \epsilon_i$ and that the errors $\epsilon_i$ are assumed to be independently identically normally distributed with mean 0 and variance $\sigma^2$. The difference from classical regression models, however, is that $g(t)$ does not need to belong to a parametric class of functions such as polynomials, but it suffices that $g(t)$ is a smooth function of the independent variables $t$.

We let $b(t)$ be a vector of polynomial terms in $t$ of degree $d$. At each query point $t_0 \in \mathbb{R}^d$, we estimate the fit

$$\hat{f}(t_0) = b(t_0)^T \hat{\beta}(t_0)$$

This is done by solving the minimization problem

$$\min_{\hat{\beta}(t_0)} \sum_{i=1}^{n} K_h(t_0, t_i) \left( y_i - b(t_i)^T \hat{\beta}(t_0) \right)^2$$

(2)

Where $K$ is a weight function, or kernel. It is defined as

$$K_h(t_0, t) = \frac{1}{h} W\left(\frac{||t_0 - t||}{h}\right)$$

where $||.||$ is the Euclidean norm and $h$ is a distance parameter which has to be chosen. $W(t)$ is a weight function that satisfies the properties

$$\begin{align*}
W(t) &> 0 & |t| < 1 \\
W(-t) &= W(t) \\
W(t) &\text{ is nonincreasing for } t \geq 0 \\
W(t) &= 0 & |t| \geq 1
\end{align*}$$

Commonly, the "Tricube" function is used as weight function:

$$W_{\text{tricube}}(u) = \begin{cases} 
(1 - |u|^3)^3 & \text{if } |u| < 1 \\
0 & \text{otherwise}
\end{cases}$$

In effect, this means that all points within the distance of $h$ are given weights $K_h(t_0, t) > 0$ with diminishing weights the farther the point is from the evaluated point $t_0$. Zero weight is given to all points beyond the distance $h$. 8
The parameter $h$ is called the bandwidth and is a free parameter in the local regression model. A large bandwidth makes the regression average over more observations, which implies lower variance but a higher bias. $h$ can either be the distance of the $k$ nearest neighbour to $t_0$, or a specified metric distance window around $t_0$. In this thesis, we will use a constant metric distance as parameter (see further motivation in Section 4.3). The parameter $h$ can be found using the Nelder-Mead method combined with LOOCV for measuring the error. Further reading on the Nelder-Mead method can be found in [11].

### 3.9 Leave-One-Out Cross Validation (LOOCV)

Leave-One-Out Cross Validation is a resampling method that may be used in order to measure the performance of a given statistical learning method (model assessment) or in order to select an appropriate level of flexibility in a statistical method (model selection). A common measure of error rate is Mean Square Error (MSE)

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2$$

where $n$ is the number of observations, $y_i$ is an observation with input $t_i$ and $\hat{y}_i$ is the estimated response for input $t_i$.

Commonly, model performance evaluation makes a distinction between training error rate and test error rate. Training error rate is the error rate that is obtained by running the model on the same data by which it has been created (or trained). Test error rate is the error rate obtained by running the model on a data set separate from the training set called a test set.

In the process of cross validation the separation of the data into a training set and test data (or validation set) is made multiple times with the use of resampling, i.e. the complete data set is randomly divided into a training set and a validation set. The model is estimated based on the training set and a measure of the performance of the model is obtained from the validation set.

**Leave-One-Out Cross Validation (LOOCV)** estimates a test error rate that is independent of how we decide to split the data set into training set and test set[8]. The LOOCV method splits the data set into two sets, where the validation set only includes one observation $(y_i, x_i)$ and the other observations make up the training set. The statistical model is estimated on the training set and a prediction $\hat{y}_i$ is made for the excluded observation. The $MSE_i = (y_i - \hat{y}_i)^2$ is then computed. This procedure is repeated for each of the observations $(y_i, x_i)$ in the entire dataset, and we end up with $MSE_i, i = 1, \ldots, n$ corresponding to each data point. The LOOCV estimate for the test MSE of the entire data set is then estimated as
\[ CV_{(n)} = \frac{1}{n} \sum_{i=1}^{n} MSE_i \]

### 3.10 k-fold Cross Validation

k-fold cross validation is an alternative method to LOOCV that is less computationally heavy. The original data set is divided into k sets or *folds*. The folds then take turn to act as validation set one at a time, while the other folds are treated as one training set. For each fold acting as validation set, we obtain an error measure \( MSE_i \) and from the average of these we get an estimate of the true test error.

\[ CV_{(k)} = \frac{1}{k} \sum_{i=1}^{k} MSE_i \]

Typically, this is performed with \( k = 5 \) or \( k = 10 \) folds[8].

### 3.11 Bootstrapping

Bootstrapping is a statistical method that can be used to quantify the uncertainty of a statistic estimation by repeatedly resampling data from the original data set. From a data set with \( n \) observations, a new data set of \( n \) observations is sampled randomly with replacement and from the new data set, an independent estimate of the statistic can be made. Repeating the procedure of resampling and estimating the statistic, the uncertainty in the statistic can be measured as the standard deviation between the estimated statistic over many resamplings.

### 4 Problem Formulation

#### 4.1 Data

Two different datasets have been used in this thesis. The first dataset consists of passenger volumes at the Airport of Bologna on a quarterly basis from 2007 until 2015. The start year 2007 is chosen since this was the year that LCC-airlines started operating on the airport, which had a major impact on passenger volumes. This dataset is an aggregate of arrivals and departures, and does not separate LCC-airlines from other airlines. This dataset is used to calibrate the time series and form a long term forecast. The second dataset consists of passenger figures for every flight at the Airport of Bologna during
2013-2015. This dataset enables separation between arriving and departing flights, separation between LCC and FFC airlines, as well as a seasonal separation between summer and winter. Following the standard at the Airport of Bologna, summer scheduling takes place between 15th of April and 15th of October while the rest of the year has winter scheduling.

4.2 Long Term Forecast

In order to forecast passenger demand between 2016 and 2026, an ARIMA($p,d,q \times (P,D,Q)_4$) time series model is estimated on quarterly passenger data from 2007-2015. The ARIMA($p,d,q \times (P,D,Q)_4$)-model includes seasonal lags and the number of lags $p$, $P$, $q$ and $Q$ are determined by the Akaike Information Criterion (AIC).

4.3 Time-of-Week Demand

With the objective to find the relative density of departing passenger demand during the time of the week, a predictive model is trained from intraday data in the detailed data set covering all flights from 2013-2015. A significant portion of the passengers at Airport of Bologna travel with LCC airlines. These depart at fix hours during the day (typically close to 6 a.m., 10 a.m. and 8 p.m.). It can be assumed that these passengers chose their flights because of its ticket price rather than that it suits their preferred embarking time. Passengers who chose to travel with the FFC-airlines are generally less price sensitive, and therefore their chosen departure time has a closer relation to the actual demand of flights. For this reason, LCC-airlines have been removed from the data set so that it only covers FFC-airlines.

Some modification has been done to the data set. The data set covers 159 weeks, over a period where the annual number of passengers has grown by 11%. This implies that the weeks will not have identically distributed numbers of passengers. To account for this, the passengers at each hour are measured as percentage of the number of passengers during the week, which will neutralize the trend from the data. In order to avoid heteroskedasticity, logarithmic values of the observations are used. Due to this, observations with zero passengers are removed. To account for this, the values at each time point are scaled proportionally to how large fraction of the observations that are zero in that time.

In order to obtain the predictive model, two methods are used and evaluated. One model estimates the mean number of passengers per hour at each fifteen minutes interval of the week, and the mean is then used for prediction. Another model is obtained with the LOESS method to predict new data points for each fifteen minute period. The $k$-nearest neighbour approach is not suitable to our dataset which has roughly equal number of observations on each time slot of the week. The distance to neighbour $k$ would be the same in nearly all points $t_0$. It is therefore more straightforward to use a metric space
\( h \) as parameter. A test MSE is estimated with LOOCV using different values of the parameter \( h \), and \( h \) is then chosen such that the test MSE is minimized. As kernel function in the LOESS, the tricube function is used and the polynomial in the regression is set to order 2.

4.4 Busy Hour Rate

An estimate of the Standard Busy Rate (SBR) and the Busy Hour Rate (BHR) are obtained from data for Airport of Bologna from 2015. Hours with no scheduled flights have been excluded from the data. In order to evaluate the uncertainty of the estimates, the rates are estimated 100 times using bootstrap resampling of the data. The standard deviation of the estimates is then observed in the variability in the bootstrap estimates.

5 Results

5.1 Long Term Forecast

5.1.1 Model Selection

The observed volume of passengers at the Airport of Bologna can be seen in Appendix A.1 and the differenced passenger volumes in Appendix A.2. There is a clear trend in the non-differenced data but the differenced data appears to have mean zero and constant variance. In order to validate this, the ADF test is performed for various lags in the autoregressive term \( p \) to show stationarity. The results are shown in Table 2. The ADF test concludes that the non-differenced data series has a unit root but that the differenced value of passengers are stationary. In light of this, a difference operator \( d = 1 \) will be included in the model.
Table 2: ADF Test results for AR-lags 0-2.

<table>
<thead>
<tr>
<th>Model</th>
<th>Result of ADF test</th>
<th>p-value of test</th>
</tr>
</thead>
<tbody>
<tr>
<td>passenger volume, p=0</td>
<td>Unit root</td>
<td>0.52</td>
</tr>
<tr>
<td>passenger volume, p=1</td>
<td>Unit root</td>
<td>0.50</td>
</tr>
<tr>
<td>passenger volume, p=2</td>
<td>Unit root</td>
<td>0.97</td>
</tr>
<tr>
<td>differenced passenger volume, p=0</td>
<td>Trend stationary</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>differenced passenger volume, p=1</td>
<td>Trend stationary</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>differenced passenger volume, p=2</td>
<td>Trend stationary</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

A seasonal ARIMA\((p, d, q) \times (P, D, Q)\)\_s model with degree of seasonal component \(s = 4\) and seasonal and non-seasonal integration \(d = 1\) and \(D = 1\) is formulated. In order to determine the optimal number of lags in the model, an exhaustive comparison between models with all possible combination of lags \(p, q, P, Q\) up to 6 steps is made. The models are fitted on the data, and corresponding AIC is calculated. The model with the lowest AIC is found to be of the form ARIMA\((0, 1, 1) \times (0, 1, 1)\)_4,

\[
(1 - B)^1(1 - B^4)y_t = (1 - 0.2048B)(1 - 0.5119B^4)a_t \tag{3}
\]

In order to test the performance of the model ARIMA\((0, 1, 1) \times (0, 1, 1)\)_4, it is set to forecast 2014 and 2015 after being estimated from the data points 2007-2013. The forecast is shown in Figure 1 along with the actual values for 2014 and 2015. As a measure of the accuracy of the forecast, the absolute percentage forecasting error is calculated as

\[
PFE = \text{mean}\left(\frac{\hat{y}_t - y_t}{y_t}\right)
\]

where \(\hat{y}_t\) is the predicted value at time \(t\) and \(y_t\) is the observed value at time \(t\). For the model the PFE is 2.90%, meaning that the forecast on average differs 2.90% from the observed value.
Figure 1: Performance of time series model when it forecasts 2014-2015, shown in red. The observed outcome of 2014-2015 is shown in blue.
5.1.2 Forecast

A forecast between 2016 and 2026 is made from the estimated model (3), and is shown in Figure 2. The model suggests a heavy growth of passenger demand at the airport, reaching above 11 million annual passengers in 2026, corresponding to an annual growth rate of 4.77%.

![Forecast of passengers at the Airport of Bologna on a quarter basis between 2016 and 2026. The blue line indicates the sum of the previous 4 quarters on a rolling basis.](image)

Figure 2: Forecast of passengers at the Airport of Bologna on a quarter basis between 2016 and 2026. The blue line indicates the sum of the previous 4 quarters on a rolling basis.

5.2 Time-of-Week Demand

In order to find the density of passenger demand during the week, two non-parametric models are estimated. The first model is computed by taking the average passenger volume at each hour of the week and the second model is estimated with local regression. The resulting models are shown in Fig. 3-6. The test error rate in the models is estimated using 10-fold cross validation. As can be seen in Table 3, the MSE is fairly similar for the two methods, even though very different techniques are used in order to obtain them. A possible source for under performance of the LOESS method is the asymmetry shown in the distribution of passengers (see QQ-plot in Appendix).
Figure 3: Time-of-week demand model estimated with the average method. Estimated on observations from winter.

Figure 4: Time-of-week demand model estimated with the average method. Estimated on observations from summer.
Figure 5: Time-of-week demand model estimated with LOESS. Estimated on observations from winter.

Figure 6: Time-of-week demand model estimated with LOESS. Estimated on observations from summer.
<table>
<thead>
<tr>
<th>Season</th>
<th>Method</th>
<th>Cross Validation MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter</td>
<td>Local Regression</td>
<td>$2.97212 \times 10^{-5}$</td>
</tr>
<tr>
<td>Summer</td>
<td>Local Regression</td>
<td>$1.70344 \times 10^{-5}$</td>
</tr>
<tr>
<td>Winter</td>
<td>Average method</td>
<td>$2.96291 \times 10^{-5}$</td>
</tr>
<tr>
<td>Summer</td>
<td>Average method</td>
<td>$1.70835 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

Table 3: Cross validation error for the time-of-week models.

5.3 Busy Hour Rate

The peak hour rates have been estimated 100 times with bootstrap resampling. The distributions of the resampled observations can be seen in Appendix A.5-A.7. In Table 4, the mean and standard deviation of the resampled observations are shown. It can be concluded that the variability in the rates is negligible. In the three measures, the standard deviation is highest in the SBR (Top 30), where it amounts to 1.04% of the estimated mean.

<table>
<thead>
<tr>
<th>Peak Type</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Mean as % of annual passengers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Top 95%</td>
<td>1932.02</td>
<td>16.74</td>
<td>0.02793%</td>
</tr>
<tr>
<td>Top 97.5%</td>
<td>2157.50</td>
<td>14.97</td>
<td>0.03124%</td>
</tr>
<tr>
<td>Top 30</td>
<td>2516.94</td>
<td>26.15</td>
<td>0.03644%</td>
</tr>
</tbody>
</table>

Table 4: Mean and standard deviation of the peak hours measures, as they are calculated on bootstrap resamples.

6 Conclusions

The time series model forecasts a significant growth of passenger volumes at the Airport of Bologna. In 2026, the number of annual passengers is forecast to succeed 11 million, compared to 6.9 million in 2015, after having grown with an average annual rate of 4.77%. Regarding the time-of-week curves, it can be concluded that the LOESS method does not have any advantage in accuracy over the averaging method. Further, the findings of Koppelman et. al can be confirmed in that that mid-morning and late-afternoon
flights are preferred, midday flights are moderately preferred while early-morning and late-evening flights are unpreferred. The estimates for peak hour passenger volumes at the airport are proved to contain little variability and can be concluded to be stable. The peak hour volumes estimated according to the three different definitions are all lower than the TPHP of 0.04% of annual traffic, which is stipulated by FAA for an airport the size of Airport of Bologna. This indicates a more evenly distributed flow of passengers at the Airport of Bologna than at the model airport of FAA.

7 Limitations of the Research

There are several elements that might limit the accuracy of the forecasts presented in this thesis. In the time series forecast, the data set is limited to nine years of observed passenger levels. This might lead to less accurate parameter estimations and a less accurate forecast. In the LOESS time-of-week model, the slight asymmetry of the data might affect the performance of the model. Further, and of high importance, is the fact that this thesis does not account for the possibility of regulation of the airline industry or other international efforts made to reduce demand for air travel. Efforts to lower the demand for air transport will undoubtedly have an effect on the accuracy of the forecasts presented here. This should be examined further.
A Appendix

Figure A.1: Observed passenger volumes on quarterly basis at the Airport of Bologna 2007-2015.

Figure A.2: Differenced passenger volumes on quarterly basis at the Airport of Bologna 2007-2015.
Figure A.3: QQ-plot of log of passenger volume versus standard normal distribution, winter observations.

Figure A.4: QQ-plot of log of passenger volume versus standard normal distribution, summer observations.
Figure A.5: Distribution of busiest 95% hour of 2015 after bootstrap resampling

Figure A.6: Distribution of busiest 97.5% hour of 2015 after bootstrap resampling
Figure A.7: Distribution of 30th busiest hour of 2015 after bootstrap resampling)

References


