An automated Multi-Model approach for Time Series Forecasting

by

Kiran Anwar

A Dissertation Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in

Computational Sciences and Engineering

KOÇ ÜNİVERSİTESİ

January, 2020

An automated Multi-Model approach for Time Series Forecasting

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Kiran Anwar

and have found that it is complete and satisfactory in all respects,

and that any and all revisions required by the final

examining committee have been made.

Committee Members:

Prof. Metin Turkay (Advisor)

Assoc. Prof. Fadime Üney-Yüksektepe

Assist. Prof. Barış Yıldız

Date:

[To my husband, daughter, mother and my late father!]

ABSTRACT

An automated Multi-Model approach for Time Series Forecasting
Kiran Anwar
Master of Science in Computational Sciences and Engineering
January, 2020

In this thesis, we present a multi-model approach for time series forecasting. Our framework is automated, easy to use and spans around short-medium terms forecast horizons. We focus on multiple basic and advance time series forecasting models. Our basic models include simple moving average, simple exponential smoothing, Holt's and winter's model. We optimize the parameters used in these models, and for advance models we extend Box-Jenkins methodology to automated Auto-Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Average (SARIMA) models. The ARIMA and SARIMA models are complex models and need expert judgement and iterative procedures to select a best fitting model.We substitute expert judgement with statistical tests and iterative procedures with automated enumeration technique. The best fitting model is selected based on the results of a number of statistical and error estimation tests. We tested our system on M-competition data set provided by International Institute of Forecasters. The dataset is comprised of multiple time series from social and economic backgrounds. From our experimental work, we conclude that our SARIMA model outperforms all other models with an average MAPE of $\leq$1% for 1-period ahead and approximately 9% for 6 periods-ahead forecast horizons.

Key Words: ARIMA, SARIMA, Time Series, Automated.

ÖZETÇE

Yü Zaman serisi tahmini için otomatik çok modelli yaklaşımşlığı
Kiran Anwar
Hesaplamalı Bilimler ve Mühendisliği, Yüksek Lisans
Ocak 2019

Bu tezde, zaman serisi tahmini için çok modelli bir yaklaşım sunuyoruz. Çerçevemiz otomatiktir, kullanımı kolaydır ve kısa-orta vadeli tahmin ufuklarını kapsar. Çoklu temel ve ileri zaman serisi tahmin modellerine odaklanıyoruz. Temel modellerimiz basit hareketli ortalama, basit üstel yumuşatma, Hotl ve kış modelini içerir. Bu modellerde kullanılan parametreleri optimize ediyoruz ve Box-Jenkins metodolojisini otomatik Otomatik Regresif Entegre Hareketli Ortalama (ARIMA) ve Mevsimsel Otomatik Regresif Entegre Hareketli Ortalama (SARIMA) modellerine genişletiyoruz. ARIMA ve SARIMA modelleri en uygun modeldir. Alt uzman muhakememiz ve istatistiksel testlerimiz ve sayımla yinelemeli prosedürlerimiz var. En uygun model, bir dizi istatistiksel ve hata tahmin testine göre seçilir. Sistemimizi, Uluslararası Tahminciler Enstitüsü tarafından sağlanan M-rekabet veri seti üzerinde test ettik. Veri seti, sosyal ve ekonomik arka planlardan gelen çoklu zaman serilerinden oluşur. Deneysel çalışma, SARIMA modelimizin, 1 dönem öncesinde tahmin için ortalama$\leq$1% ve 6 dönem öncesinde tahmin için yaklaşık7% ortalama MAPE olan diğer tüm modellerden daha iyi performans gösterdiğini öne sürmektedir.

Anahtar Kelimeler: ARIMA, SARIMA, Zaman Serisi, Otomatik.

ACKNOWLEDGMENTS

I am very thankful to Prof. Metin Turkay for giving me an opportunity to work on time series forecasting and for his endless guidance and support during my research work. I am thankful to all my colleagues in Systems Lab especially Banu Ulosoy.

I am very thankful to Irfan Azhar in particular, my better-half, not only he supported me through my Masters journey but also encouraged me and has given valuable advice for my research work. I am thankful that we have a beautiful daughter Meryem, with her tantrums the research journey has not been easy but its surely been very pleasant. She kept me focused, movtivated and made me learn how to achieve work-family balance.

I am also very grateful for my friends and family especially my sister Aasma, my friend Farhana, and my Turkish neigbors Isma and Anne who listened to my endless whining during my tough times and still supported me. Without my neighbors support I could not have completed my research work, they are the reason I feel like I have a family outside my home country.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

AR	Auto Regression
MA	Moving Average
ARMA	Auto-Regressive Moving Average
ARIMA	Auto-Regressive Integrated Moving Average
SARMA	Seasonal Auto-Regressive Moving Average
SARIMA	Seasonal Auto-Regressive Integrated Moving Average
MAPE	Mean Absolute Percentage Error
RMSE	Root Mean Squared Error
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
AC	Auto Correlation
PAC	Partial Auto Correlation
TS	Tracking Signal

Chapter 1

INTRODUCTION

Forecasting future demands accurately plays an important role in supply-chain planning. Accurate forecasting is very helpful in utilizing the resources efficiently and prohibiting any catastrophic situation in advance. Time series analysis is used for forecasting future values when there is availability of past data and trends. It is an ordered sequence of quantitative random values observed over equally spaced intervals of time. Thus a time series is generally modeled via a stochastic procedure $\mathbf{X(t)}$. For forecasting, we are keen in estimating $\mathbf{X(t+h)}$, utilizing the data at time $\mathbf{t}$ [Davide Burba, 2019]. There should be enough values in the time series data to estimate forecasting model parameters. With a cyclic or seasonal series, there should be enough cycles or seasons to model them accurately. Therefore, seasonal series needs more values than non-seasonal ones [Yaffee and McGee, 2000]. The missing data in series should be replaced using some algorithms. Usually the missing value is replaced by the average of last values before it. If there is a lot of missing values in the social or economic data then the data is not suitable for time series analysis, as the forecast generated will not be an accurate forecast.

Forecasts are not always accurate, especially short term forecast is more accurate than long term forecast [Chopra and Meindl, 2016]. With demand or prices predicted ahead of time, a decision-makers can make important decisions about its growth strategies and pricing. Without it, the chances of making poor decisions increase. For instance, if a business already knows approximately how many products are going to be sold in future; the related department can order as many units close to demand. An accurate estimation of a product's annual demand is interesting for the business experts as it leads to better decision-making. If the forecast turns out to be accurate, the business will have a profit. Otherwise, if the demand turns out to be less than predicted, the business will have loss. By forecasting future demands we are trying to reduce inventory holding cost and increase customer-satisfaction and profitability. Therefore, an accurate future demand or number of sales is critical in good decision making and efficient supply chain system. In this thesis, we focus on devising an automated forecasting system for the time series analysis

This thesis studies time series forecasting models. We use the quantitative forecasting method. We consider simple moving average, exponential smoothing, holt's-Winter's model and Box-Jenkins ARIMA and SARIMA models. The aim is to devise an automated, accurate and fast forecasting system to facilitate the use of forecasting for future demand planning. Automated implies that all the necessary steps to generate a forecast are automatically processed by the system. No human judgment or modification is needed. The system aims to make accurate predictions in a reasonable time. Our system is user-friendly; it only asks the user for input data, processes the data automatically and output future forecasts. The general processing of data involves following steps:

(1) Data-checking for negative values,
(2) Apply transformations if necessary,
(3) Testing all the forecasting models on data,
(4) Comparison of the results of all models,
(5) Selection of best fitting model,
(6) Generating future forecasts

Our forecasting system is divided in to two categories: Basic models and Advance models. We conidser simple moving average, exponential smoothing and holt-winter's model as basic models, whild Box-Jenkins ARIMA and SARIMA as advance models for our system. For our basic models we consider our data as demand comprised of systematic and random components. The systematic components are

(1) Level,
(2) Trend, and
(3) Seasonality.

Level is the current deseasonalized demand in the series [Chopra and Meindl, 2016].

Trend refers to the rate at which the next period of demand grows or declines while seasonality shows predictable seasonal fluctuations in demand. The random component is opposite of the systematic component. The forecast should never be in the direction of the random component, it is the noise that is supposed to be filtered out in forecasting.

The ARIMA and SARIMA models are unable to produce good results with non-stationary data. There is no specific rule about the minimum number of values in a series but there should be at least 30 values for ARIMA and SARIMA models. It is a complex forecasting model and need a lot of expert judgement involvement to work properly thus takes long time to implement it. The implementation time is long because the Box-Jenkins methodology makes use of multiple repetitive procedures to choose a best-fitting model. These facts necessitate the need of an an automated Box-Jenkins methodology.

To the best of our knowledge [Anvari et al., 2016] research work deals with data-stationary automatically. We take inspiration from her work and make our own data-stationary identification and stationary testing algorithm. Also, in the literature the model order selection is chosen using ACF and PACF plots. ACF plots are use to choose the order of Auto-Regressive (AR) and Seasonal Auto Regressive (SAR) order models, while PACF and time plots are used to choose the order of Moving average (MA) , Seasonal Moving Average (SMA) Models. However, even human judgement can misread these plots and select wrong order for the modedls. We avoid the use of human judgement and process this step with enumeration technique using coputer programs.

Our integrated algorithm is based on the following steps:

(1) Multiple statistical tests and mathematical modifications applied to data to generate stationary data,

(2) Enumeration process for combinations of model's order parameters and coefficient estimation,

(3) Selection of best fitting models on the basis of diagnostic checks,

(4) Final model selection on the basis of performance accuracy parameters, and

(5) forecasting and applying needed transformations.

The rest of the thesis is as follows: We discuss literature review for time series forecasting models and innovative approaches in chapter 2. Detailed explanation of our forecasting models and algorithmic details is provided in chapter 3. It also includes our step-by-step detailed approach for automated Box-Jenkins ARIMA and SARIMA models. In 4 we describe our system's platform, datasets and analyze the results obtained by testing M-competiton data-set's time series on our forecasting system. A small proportion of each time series is provided in the Appendix section of thesis. Note that our system is general and is able to generate forecast for any time series data-set.

Chapter 2

LITERATURE REVIEW

There are many studies on forecasting time series data with so many focus on forecasting road traffic, tourism demand, number of air passenger, predicting inflation, electricity demand and predicting gasoling prices. In this chapter we review the time series forecasting models used for different types of demands. We focus on finding the relations between demand and certain or uncertain factors, and comparing the models based on performance statistics. This chapter is divided in to three sections: the first section reviews the research work based on univariate time series models, the second section discusses forecast combination techniques, and finally we discuss some of the research work based on automated forecasting systems.

2.1 Univariate Time series Models

Forecasting methods are divided into qualitative and quantitative forecasting methods. The latter is used: when authentic data concerning the events to be anticipated is not accessible at all or are rare. For instance, when there is a launch of new product. In case of quantitative forecasting methods various techniques are utilized for the estimation of future demands; multi-mode models, time series models, market surveys, multiple regression and neural networks [Anvari et al., 2016]. In addition, to choose the best model, the specific plan of forecasting, for instance, tactical or strategic decisions and medium-long term or short-medium term estimates must be considered [KRASIC and GATTI, 2009]. For instance, for short-term forecasts of number of passengers, the univariate time series analysis is useful[Uysal M, 1985].

ARIMA and SARIMA has been widely used in air passenger forecasting. [J., 1996] used Holt's, Winter's and ARIMA models along with Naive 1 and Naive 2 models to forecast outbound air travelers form United Kingdom. The study concluded that ARIMA model outperformed all other models except in one case. [Oh C-o, 2005] studied the models proposed by [F-L, 1998] and[Y-M., 1993], alongside two of their own on inbound Singapore tourists arrival. They concluded that ARIMA models shows better results as compared to Winter's model when used for post-sample forecasts. Whereas, Winter's model show better results as compared to ARIMA when used for within-sample forecasts.

[Emiray and Rodríguez, 2003] applied six time series models on Canadian air passengers data. The models used were linear model with seasonal unit roots, Auto regressive AR(p) and Auto Regressive AR(p) with seaonal unit roots, SARIMA model, structural time series model (STSM) and periodic autoregressive model (PAR). The smoothing models, for instance, exponential smoothing were not considered. The authors concluded that forecasting horizon and performance parameters plays an important role in accuracy of the forecast.

There have been studies where quantitative statistical and non-statistical methods are combined to predict future values. [Mwakalonge, 2014] stated that if only recent data is considered to forecast, it can cause variations of low and high frequency values than expected. This is due to instability of economic conditions and travel behavior. They focused on an urban region and collected multiple cross sectional data sets for it at different periods. They made forecasts using two categories: In the first category, all the model parameters were combined from different periods, and in second category all the data from different periods was combined. The research concluded that using multiple data sets there is a chance to develop better models.

[Alysha M. De Livera, 2011] proposed a state space modeling framework for dealing with complex seasonal patterns, dual calendar effect and multiple seasonal periods. Along with Exponential smoothing, the framework incorporates ARMA error correction, Box-Cox transformations and time varying coefficients by Fourier representations. The authors claimed that their framework can model both linear and non-linear time series. The framework is able to produce accurate out of sample forecasts with minimum estimations of parameters as compared to holt-Winer's exponential smoothing model.

[Omane-Adjepong and et. all., 2013] obtained the monthly data of almost 40 years from Ghana Statistical Service. The authors used Holt-Winters and SARIMA models to obtain out of samples forecast. MAE, MAPE, RMSE, and MASE were used to determine the performance of forecasting models. The authors concluded that SARIMA model outperforms the other models and is best for short-term out of sample forecast for Ghana's inflation. [Nguyen and Hansen, 2017] generated short-term forecasts for load demand using ARIMA and SARIMA models. The hourly data was collected from Electric Reliability Council of Texas. The authors concluded that SARIMA models outperform ARIMA models. However, with the increasing number of periods, the accuracy of SARIMA models aslo reduce.

2.2 Forecast Combinations

An alternative approach to single model's forecast is 'Forecast Combinations'. Forecast combination is needed when the past data is not reliable; the number of observations are very few or the individuals forecasts are not specified properly [Timmermann, 2013]. This concept was first introduced by [Bates and Granger, 1969] to increase the accuracy of forecasts. The forecasting performance is more likely to improve with this concept, if the forecasting models combined are uncorrelated [Brown and Murphy, 1996].

[Bougas, 2013] apply simple averaging and variance co-variance methodologies as combination forecasts to air-passenger data in Canada. These techniques uses single model forecast weighted average, but the computation procedure of weights is different. The predictions derived from single model combination are given equal weight in simple averaging. Whereas, un-equal weights are assigned to each single model combined according to its past performance in variance co-variance methodology. The author claims that the forecast accuracy is improved when using combination models as compared to single models. In some cases ARIMA, SARIMA outperforms. However, combination model always performed better than its single model worse performance.

2.3 Automated Forecasting

In this section we will review some of the research work based on automatic forecasting system. [SHARDA and PATIL, 1992] compared neural network approach with Box-Jenkins methodology. For box-jenkins methodology, the authors used a commercial program AUTOBOX [Inc, 1998]. The program is able to take the data in, iterate through stages of Box-Jenkings methodology (Identificaiton, Estimation and Diagnostic Checking) and provide the best model. As per results, the simple training neural network methodology had better result for 50 of the series, while for 22 series AUTOBOX outperformed. For long-term future horizons, the Box-Jenkins methodology performed slightly less better than the short term future horizon. However, the program AUTOBOX is not open source, and can not be compared for academic purposes.

[Hwang et al., 2012] selected best fitted model using automated procedures for selecting best model and Box-Jenkins approach. Thus developed an automated time-series material cost forecasting (ATMF) system. The authors claim that their proposed system (ATMF) is helpful in construction industry for the decision makers as it is generating accurate out of samples forecast and trend prediction in construction material costs. This system is helpful only in predicting materials cost accurately. No other timeseries were tested for it.

[Burkom et al., 2007] compared multiple time series models for bio-surveillance forecasting. These models comprises of non-adaptive regression model, adaptive regression model, generalized exponential smoothing and Box-Jenkins ARIMA model. The research states that the comparison among models is automated. However, for ARIMA automated refers to a pre-selected model. Also for smoothing constants, in case of Holt-Winter's approach, specific values were set based on a grid search using clustering approach.

Using R-trees and delay coordinate embedding technique [Chakrabarti and Faloutsos, 2002] build an automated forecasting system for the predictions of future time series data. The authors used fractals to develop their system. One of the novelty proposed is the use of lag plots. The authors claim that their method is fast, reliable and and accurate. There is a built in package 'forecast package for R' in R by [Hyndman and Khandakar, 2007]. This package is able to forecast in state space models and ARIMA models. However, the R package is unable to select and test all models by itself. The user has to manually code to test each model and compare the results.

Chapter 3

METHODOLOGY

In this chapter we describe the techniques used for devising our automated forecasting system. We also include the detailed implementation of time series forecasting models. Before explaining the methodology of our basic and advance models, we give a brief overview of the pre-processing involved in our data modification and error estimation methods. The overall flow of our forecasting system is described in fig 3.1.

3.1 Dealing With Null or Negative values

Before sending the data to our forecasting models, the first step is to check for null or negative values and length of the series. We find the null and negative values and length of the series using built in commands in Python. If there is any negative value or missing value, it is replaced by a 0. Considering the fact that all null values do not necessarily mean missing data, we replace the null values with a number between 0-1.

The length of series is an important factor for dividing the series in to train and test data. For our forecasting models we divide the data values to train and test data. As mentioned earlier, the minimum values for the train data should be at least 30 observations. If the data is less than 30 values, we test our system for only Simple Moving average, Exponential smoothing and Holt's and Winter's model. We send train data to our models, estimate the optimized parameters associated with the models, and compare the performance of these models on our test data. These forecasting models compute predictions of the same length as of test data. Later the predictions of each model is compared with the test data and forecasting error is computed. The model which produces the minimum forecasting error is the best fitting model. We choose this model and forecast future demand values.

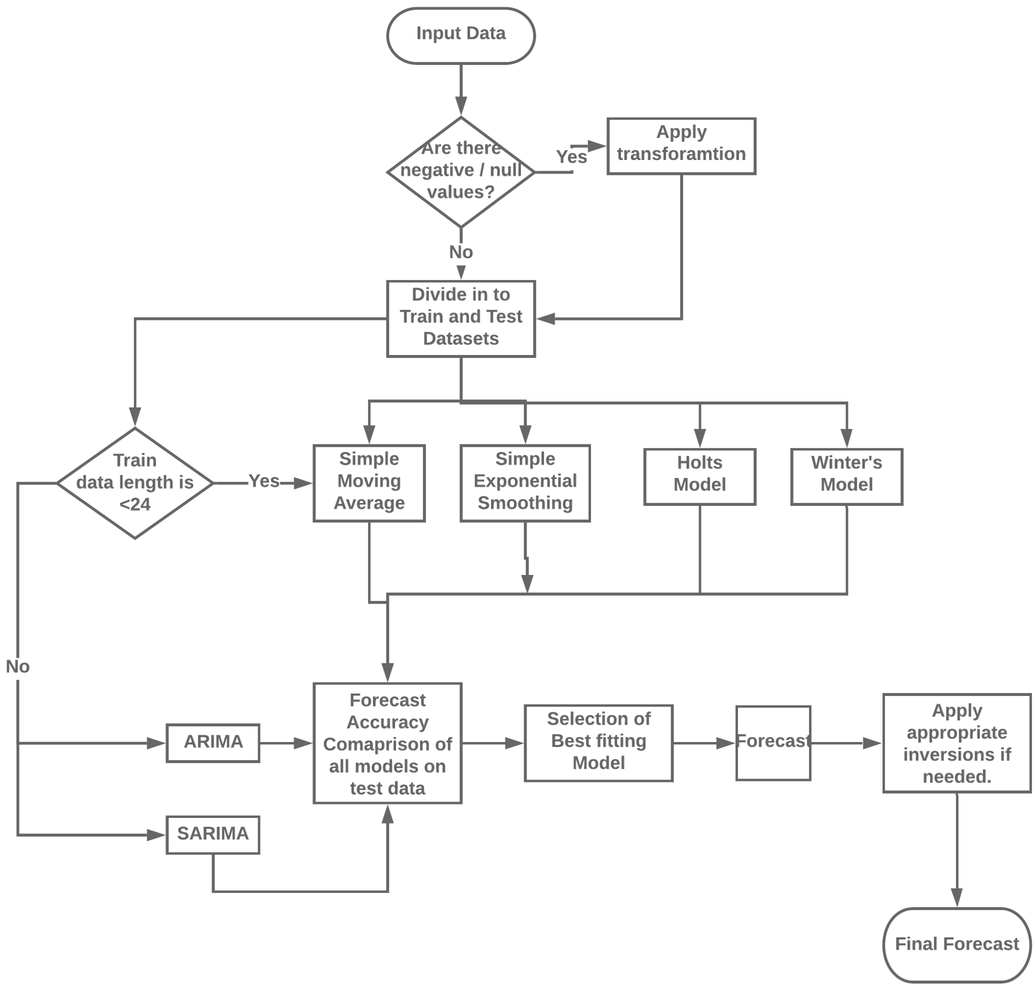


Figure 3.1: Our automated multi-model forecasting System

3.2 *Error Estimation Methods*

There are several measures of forecasting errors in theory to check the model's accuracy, with no generally favored one. Selection of models vary depending on the selection criteria [C-o and BJ, 2005].

The most common error estimation methods are mean absolute error, mean absolute percentage error, mean absolute deviation error, mean squared error, RMSE, BIC and AIC. Mean absolute error (MAE) shows the magnitude of error. However, when compared with RMSE, it is less sensitive to infrequent very large errors. The reason is beause RMSE square the errors but MAE does not. In case of RMSE, the negative and positive errors do not balance each other as it is more sensitive to frequent large errors [Anvari et al., 2016]. [Chu, 2009] and [SF Witt, 1991] confirmed that the squared errors is better than the absolute errors in terms of accuracy criterion.

To choose the best estimated model for ARIMA and SARIMA, we use RMSE. For reporting and comparison purposes and comparisons of the models we use MAPE error. The percentage term makes it easy to understand the error's magnitude. AIC and SBC are minimum information criteria, used to assess parsimonious fit [Yaffee and McGee, 2000]. These measures of accuracy over-parameterize and penalize the number of terms in the model. Therefore, we prefer to use RMSE for model selection and MAPE for comparison and reporting purposes. RMSE is calculated as follows:

$$RMSE = \sum_{t=1}^{i} \sqrt{\bar{e}_t} \tag{3.1}$$

3.3 Basic Time Series Forecasting Models

We will be using following notations for our basic model components explanation: after each demand observation we update these components.

L_t *Level estimate at the end of Period t*

T_t *Trend estimate at the end of Period t*

S_t *Seasonal factor estimate for Period t*

F_t *Forecast for Period t*

D_t *Actual demand at Period t*

E_t *Foreacast error for Period t*

The detailed methodology of the forecasting models is as follows:

3.3.1 Simple Moving Average

We use simple moving average when the demand has no observable seasonality or trend. The systematic component of demand is **level** only in this case. The average demand over most recent **N** periods is calculated to estimate the level at period t. It is evaluated as follows:

$$L_t = (D_t + D_{t-1} + - - - + D_{t-N+1})/N \tag{3.2}$$

The forecast at period t is the level calculated for period t. All further forecasts are calculated by revising the estimate for level t+1 , t+2 etc as shown in equation 3.4.

$$F_{t+1} = L_t \tag{3.3}$$

$$L_{t+1} = (D_{t+1} + D_t + - - - + D_{t-N+2})/N \tag{3.4}$$

$$F_{t+2} = L_{t+1} \tag{3.5}$$

We optimize N for our train data by checking the performance of our model for N=[2,8]. The N for which the RMSE values are minimum is the best N. We save the results and send the train data to Simple Exponential Smoothing model.

3.3.2 Simple Exponential Smoothing

The systematic component of demand is **level** only in this case. The first value for level is calculated by taking an average of all the observations of our train data. Suppose we have **n** periods starting from 1, then:

$$L_0 = 1/n \textstyle\sum_{k=1}^{n} D_k \tag{3.6}$$

The first forecast for period **t+1** will be equal to level estimated in equation 3.6.

$$F_{t+1} = L_t \tag{3.7}$$

All the future forecasts are calculated by revising the estimates of Level at period **t+1** as follows

$$L_{t+1} = \alpha D_{t+1} + (1 - \alpha) L_t \tag{3.8}$$

α is called the smoothing constant for level. Its value ranges between 0 to 1. We tr optimize the value of α by training our train data for all the possible values of α. The value of α for which RMSE is minimum is the best fitting simple exponential smoothing model. The forecast for period >**t+1** is revised as follows:

$$F_{t+2} = L_{t+1} \tag{3.9}$$

3.3.3 Holt's Model

Holt's model is also generally known as trend corrected exponential smoothing model [Chopra and Meindl, 2016]. We have trend and level as systematic component of demand. We run a linear regression between Demand D_t and the time period t to calculate the initial estimate of trend T and level L. The linear regression is of the form:

$$D_t = at + b \tag{3.10}$$

The initial estimate of Level L_0 is equal to the constant b at Period t=0. Whereas the initial trend T_0 is the slope a. It measures the rate of change in demand per period. The forecast for period t and further periods is calculated as follows.

$$F_{t+1} = L_0 + T_0 \qquad\qquad F_{t+n} = L_t + nT_t \tag{3.11}$$

After observing the further values of our train data, the level and trend values are re-estimated and forecast is calculated as follows:

$$L_{t+1} = \alpha D_{t+1} + (1-\alpha)(L_t + T_t) \tag{3.12}$$

$$T_{t+1} = \beta(L_{t+1} - L_t) + (1-\beta)T_t \tag{3.13}$$

$$F_{t+2} = L_{t+1} + T_{t+1} \tag{3.14}$$

The α and β are smoothing constants for Holt's model and their values range between 0 and 1. We try to optimize our Holt's model by training our train data for all the possible values of α and β. The model for which the RMSE values are minimum is the best fitting Holt's model.

3.3.4 Winter's Model

Winter's model is also generally known as trend and seasonality corrected exponential smoothing model [Chopra and Meindl, 2016]. In this model we have all three systematic components of demand. Another parameter that is used in this model is the periodicity **p**. It represents the number of periods after which there is a repetition of seasonal cycles. If the demand is measured in yearly basis, the p=12. To estimate level and trend in winter's model, we first de-seasonalize our train data. We first detect if our periodicity is an even or odd number and apply equation 3.15.

$$\bar{D}_t = \begin{cases} D_{t-(p/2)} + D_{t+(p/2)} + \sum_{k=t+1-(p/2)}^{t-1+(p/2)} 2D_k \text{ for p is even} \\ \sum_{k=t-[(p-1)/2]}^{t+[(p-1)/2]} D_k/p \text{ for p is odd} \end{cases} \tag{3.15}$$

We estimate the initial value of level and trend by running a linear regression between the de-seasonalized demand and time period t. The original train data is not linear and contains seasonality in it. Therefore, we run the linear regression on de-seasonalized train data. The relationship is of the form:

$$\bar{D}_t = L + Tt \tag{3.16}$$

Now we calculate the seasonal factors from the de-seasonalized train data 3.17. Then we calculate final seasonal factors using periodicity and seasonal factors of de-seasonalized train data 3.18. These seasonal factors are later used for the calculations of forecast.

$$\bar{S}_t = D_t/\bar{D}_t \tag{3.17}$$

$$S_k = \sum_{j=0}^{r-1} \bar{S}_{(j+p+k)} \text{ where } 1 \leq \text{ k} \leq \text{ p} \tag{3.18}$$

The r in equation 3.18 refers to seasonal cycles. Using equation 3.16 and 3.18 we calculate the forecast in period t as follows:

$$F_{t+1} = (L_t + T_t)S_{t+1} \qquad F_{t+n} = (L_t + nT_t)S_{t+n} \tag{3.19}$$

The level, trend and seasonal forecasts are re-estimated once more train data values are observed. The revised edition is as follows:

$$L_{t+1} = \alpha(D_{t+1}/S_{t+1}) + (1-\alpha)(L_t + T_t) \tag{3.20}$$

$$T_{t+1} = \beta(L_{t+1} - L_t) + (1 - \beta)T \tag{3.21}$$

$$S_{t+p+1} = \gamma(D_{t+1}/L_{t+1}) + (1 - \gamma)S_{t+1} \tag{3.22}$$

The α, β and γ are the smoothing constants for level, trend and seasonal factors respectively. The value of these constants range between 0 and 1. We try to find the best fitting Winter's model by optimizing these constants. The values for which α, β and γ produce minimum RMSE are the best values for Winter's model.

3.4 Advance Time Series Forecasting Models

We categorize Box-Jenkins ARIMA and SARIMA models as advance time series forecasting models.

3.4.1 Box-Jenkins ARIMA

The Box-Jenkins is helpful in predicting future values of time series. This time series values are related to each other or are statistically dependent. [Bowerman and Connell, 1979]. Box-Jenkins ARIMA methodology works best on stationary data. ARIMA is a combination of AR, MA models and a parameter **d** called differencing. There are four basic steps in applying Box-Jenkins ARIMA models: **(1)** Identification**(2)** Estimation **(3)** Diagnostic checking **(4)** Forecasting.

The first step 'Identification' checks if the data is stationary or not. If the data is not stationary, it is converted to stationary data else it decides the Moving Average (MA) and Auto Regressive (AR) order of the model. The second step 'Estimation' estimates the coefficients for AR and MA components. The third step 'Diagnostic checking' checks if the model selected in earlier steps is accurate and best fitted. If the model is best fitted then Box-Jenkins ARIMA model forecasts the future values.

The mathematical formula for ARIMA model is as follows:

$$X_t = \sigma + \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \in_t - \theta_1 \in_{t-1} - \theta_2 \in_{t-2} - \theta_q \in_{t-q} \tag{3.23}$$

X_t is the value to be forecast-ed at time **t**. The AR model defines the term **p**. $X_{t-1}, X_{t-2}, ..., X_{t-p}$ are the previous observed value at time t-1, t-2 till t-p. The AR coefficients ϕ_1,ϕ_2...ϕ_p define the term σ. The MA coefficients are the terms θ_1, θ_2...

The q term is defined by the MA model. $\in_t, \in_{t-1}...$ are the residuals terms. The original Box Jenkins methodology is shown in figure 3.4.1.

We propose an automated forecasting system that uses Box-Jenkins ARIMA methodology. Our main purpose here is to avoid human's logical reasoning dependency with the help of computer algorithms. Instead of choosing one model parameters **p, d and q** based on ACF and PACF plots, we increase the accuracy by enumerating through all the possible models. We check for the performance of all models based on characteristics of best model and find the best fitting model. Our algorithm proceeds as follows:

3.4.1.1 Stationary Data Identification and Differencing:

Our original data has already been modified for missing and null values. The non-stationary data is often checked via human graphical judgment of ACF and PACF plots. For instance, if ACF plot reaches unity with a larger spike or is decaying slowly then the data is considered non-stationary [MM, 1999]. Our automated process checks data stationary using following tests and transformation.

- We remove any variance from the train data using natural logarithm. It helps to remove high variance and even if there is low variance, we have observed that the forecast is still accurate

- We determine if the seasonal differencing is needed using OSCB test [Bougas, 2013]. This test gives 0 or 1 as output. An output of 1 shows that the seasonal pattern changes are sufficient enough to apply appropriate seasonal differencing. An output of 0 suggests that no seasonal differencing is needed.

- A stationary data doesn't exhibit any trend. To check if our train data has any trend we use Kwiatkowski, Phillips, Schmidt, and Shin test [Kwiatkowski et al., 1992]. We check for the significance of p-value. If the p-value of the test result is greater than 0.05, then normal differencing is applied. Else, we move to our next test.

- We use Mean stationary test to check if the mean is time dependent. T-test is useful for this purpose. If dependency is detected from t-test result: normal differencing is applied.

- We apply unit root test to check if the data is stochastically non stationary. We use Augmented Dickery Fuller test for this purpose. If the p-value turns out to in-significant, we apply normal differencing.

All the tests are proceeded automatically and does not involve human judgment. The output is a stationary train data. We keep track of how many times we applied the differencing transformations. We make sure that we don't over-difference the data by setting the maximum number of difference transformation to 3.

3.4.1.2 Identification and Estimation:

In many studies its emphasized that the identification and estimation of ARIMA model should be an iterative process of selection according to selection via human judgment, estimation of coefficients, judgement and re-selection [Anvari et al., 2016]. This method is prone to errors and difficult to implement as a computer program [Y and SM, 2009].

In order to avoid uncertainty and error caused by human judgement, we use enumeration to choose order of AR and MA components (p and q values). We use the fact that order is usually less than or equal to two [Y and SM, 2009] and order of more than 2 causes over-fitting of series. We investigate till order 3 and go over all possible 15 combinations of MA and AR model.

3.4.1.3 Estimation of AR and MA coefficients

Once the order has been finalized we estimate the AR and MA coefficients using yule-walker equation. Using statsmodels library in python, we use built in python function for this purpose. We keep in check if the AR and MA model estimated coefficients are stationary and invertible respectively, this suggests:

- To be stationary: the sum of AR coefficients should always be less than one [Y and SM, 2009].

- To be invertible: the sum of MA coefficients should always be less than one [Y and SM, 2009]

If in a case the sum is not less than 1, we estimate the parameters again with the lagged version of train data. The constant term σ in equation 3.23 is calculate as $1 - \phi_1 - \phi_2 ...$

3.4.1.4 Diagnostic Checking

In diagnostic checking we run some tests to check how good the selected ARIMA model fits the data. Each ARIMA model is tested for randomness and correlation test of residuals. If a model satisfies these tests, its added to pool of good models.

- The residuals of the best model are expected to be distributed randomly [MM, 1999]. We use Runs test for this purpose using python's designed library stastmodel [Josef Perktold et.al, 2019].

- There should be no autocorrelation among the residuals to be a good fitted model [MM, 1999, KRASIC and GATTI, 2009]. To check autocorrelation, we use the Ljungand Box test (lbq test) designed in python.

If the tests are not rejected we put the model in the pool of good models. Else, we choose another combination of p q model. This process continues, until all the 15 combinations are checked. No we check all the models from the pool on our test data. The final model is the one which has minimum RMSE value. We use this model to forecast future periods.

The forecast goes through inversion process before making it to the final forecast. In case there is normal and seasonal differencing we apply inverse differencing to the forecast. Lastly, we apply inverse log transformation to our forecast.

3.4.2 Box-Jenkins SARIMA

A seasonal ARIMA model is known as SARIMA. Additional seasonal terms $(P, D, Q)_s$ are added to ARIMA model. The number of observations per year is denoted by s in seasonal terms of the model. The SARIMA model also has the same steps: **(1)** Identification**(2)** Estimation **(3)** Diagnostic checking **(4)** Forecasting.

SARIMA $(p, d, q, P, D, Q)_s$ has the form :

$$\Phi_P(B^s)\phi_p(B)(1-B^s)^D(1-B)^dX_t = \Theta_Q(B)^s\theta_q(B) \in_t \tag{3.24}$$

where

$$\theta_q(B) = 1 + \theta_1(B) + ... + \theta_q(B)^q \tag{3.25}$$

$$\Theta_Q(B)^s = 1 + \Theta_1(B)^s + \Theta_1(B)^{2s} + ... + \Theta_Q(B)^{Qs} \tag{3.26}$$

$$\phi_q(B) = 1 - \phi_1(B) - \phi_2(B)^2 - ...... - \phi_p(B)^p \tag{3.27}$$

$$\Phi_Q(B)^s = 1 - \Phi_1(B)^s - \Phi_1(B)^{2s} - .... - \Phi_p(B)^{ps} \tag{3.28}$$

SARIMA $(p, d, q, P, D, Q)_s$ has two parts: Non seasonal part (p,d,q) and seasonal parts (P,D,Q,s)

- p- order of non-seasonal AR terms
- d- order of non-seasonal differencing
- q- order of non-seasonal MA terms
- P- order of seasonal AR terms (i.e., SAR)
- D- order of seasonal differencing
- Q- order of seasonal MA terms (i.e., SMA)

Just like our automated ARIMA model, the SARIMA model steps are also automated. The only difference is the addition of seasonal terms and lack of constant term. We expand the back-shift operator in equation 3.24 automatically. After applying appropriate differencing, we enumerate the selection of p, q, P and Q up-to order three. For automated SARIMA model we have 81 combinations. Testing this many combination seems resource consuming, but from our experimental work its proven that accuracy is improved and error caused by human judgement is reduced using this approach.Our detailed automated SARIMA algorithm is shown in figure 3.4.2.3

3.4.2.1 Non-Stationary Data Identification and Differencing:

To identify non-stationary data we apply the statistical test mentioned in section 3.4.1.1. The difference here is the addition of seasonal differencing order of SARIMA model. We first apply our seasonal Differencing order **D** then check for non-stationary data, if data is still non-stationary the we apply normal differencing **d**.

3.4.2.2 Estimation of AR, SAR, MA and SMA coefficients

Once the order has been finalized we estimate the AR, SAR, MA and SMA coefficients using yule-walker equation. Using statsmodels library in python, we use built in python function for this purpose. We keep in check if the AR and MA model estimated coefficients are stationary and invertible respectively, this suggests:

- To be stationary: the sum of AR and SAR coefficients should always be less than one [Y and SM, 2009].
- To be invertible: the sum of MA and SMA coefficients should always be less than one [Y and SM, 2009]

3.4.2.3 Diagnostic Checking

We run diagnostic checking for SARIMA using the same tests as mentioned in 3.4.1.4 i.e, Runs test for checking random distribution of the residuals and Ljungand Box test for checking if there is auto-correlation among the residuals. Thee models for which tests are not rejected are placed in the pool of good models. If the tests are rejected we choose another combination of p, P, q and Q model. The model with minimum RMSE is the best fitting model, we use this model to predict SARIMA forecast. Appropriate inversions are applied afterward. The forecast after inversion is the final forecast, with all the trends and seasonality in it as of original data.

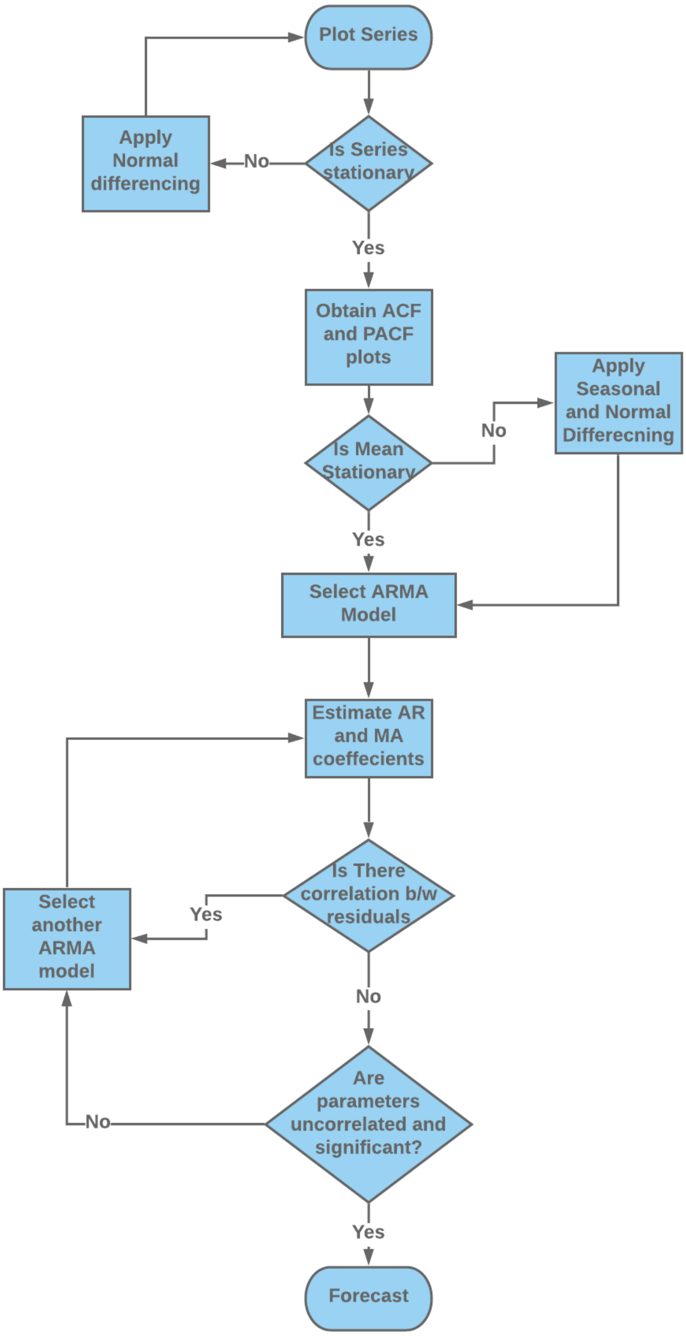


Figure 3.2: The conventional Box-Jenkins ARIMA model.

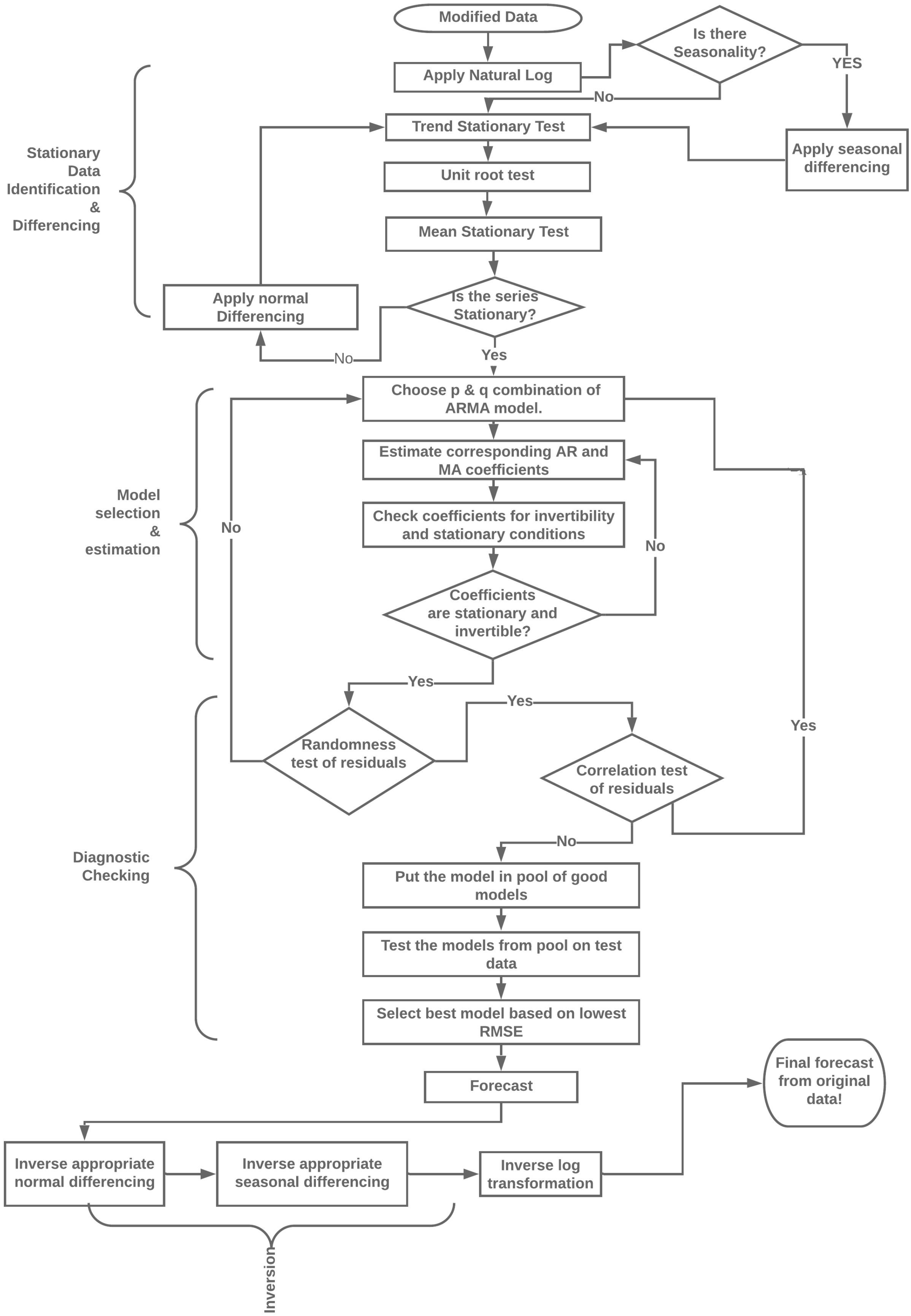


Figure 3.3: Our Automated Box-Jenkins ARIMA forecasting model.

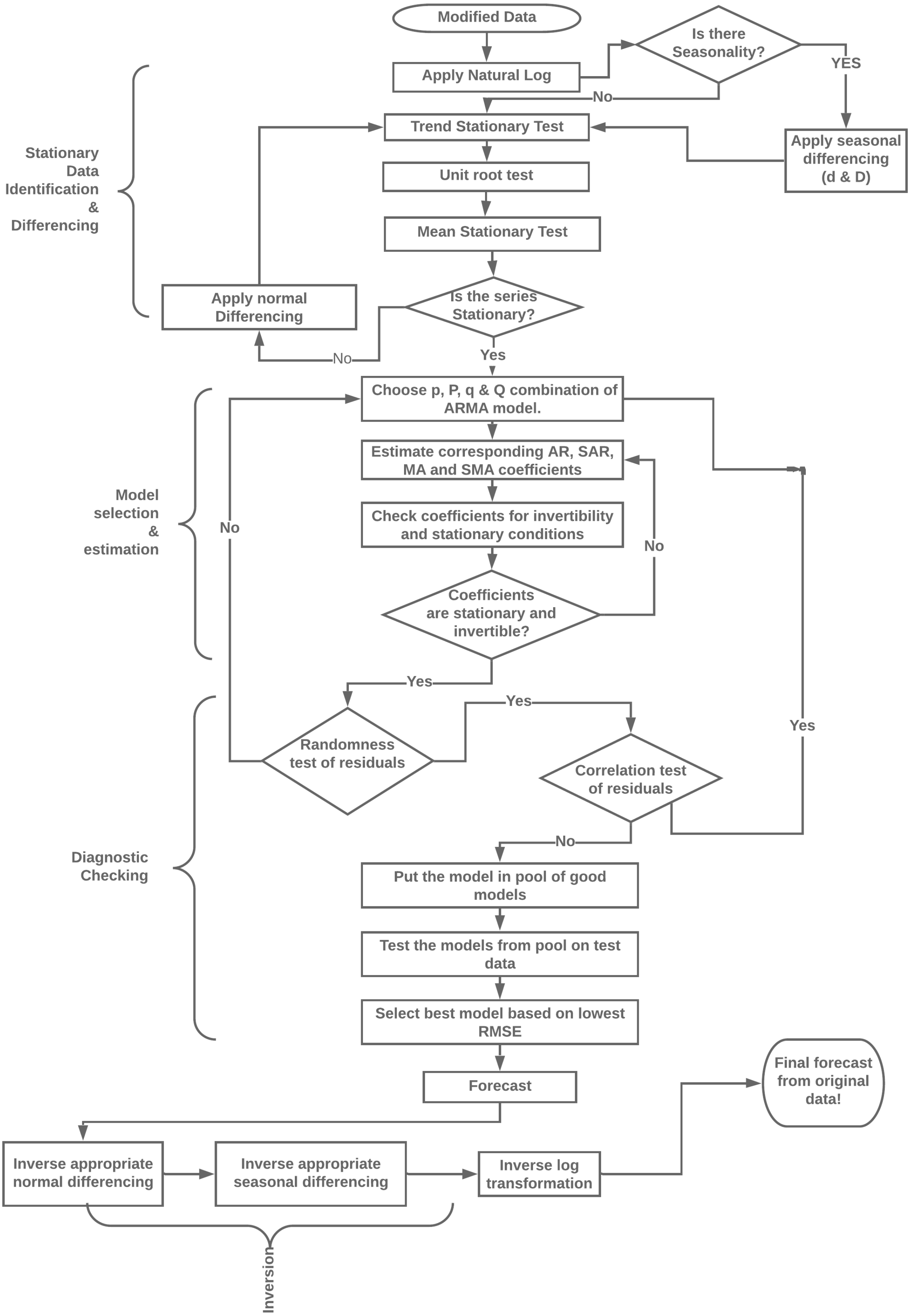


Figure 3.4: Our Automated Box-Jenkins SARIMA forecasting model.

Chapter 4

COMPUTATIONAL RESULTS

In this chapter, we compare the results of all the models used in our forecasting system. We describe the computational environment used to build our system and the dataset for testing our system in sections 4.1 and 4.2 respectively. Section 4.3 describes all the results gathered from testing our forecasting system.

4.1 Computational Platform

We develop our forecasting system in Python 3.7 platform. For our algorithms we use some of the statistical tests already available in Python. All the models are integrated in to one single file. The user is only supposed to provide data file. The system does all the computation automatically, there is no involvement of human judgement at any step of computation. Our computer is equipped with 2.60GHz processor speed, Core(TM) i7-6600U CPU and 8 GB RAM.

4.2 Dataset

We get our test dataset from International Institute of Forecasters [IIF, 2019]. We use the dataset made for M2-Competition by actual companies [Makridakis and et. all., 1993]. The dataset is divided in to 29 acutal time series, of which 23 are from 4 real companies and rest are of macro-economic nature. The dataset comprises of monthly and quarterly values ranging from Jan 1947-Jan 1981. We take 8 Series randomly from the dataset. The number of observation ranges from 82-225 values. We also take 23 monthly series from the dataset and compare the average MAPE of these series with the average MAPE of forecasting models available in [Makridakis and et. all., 1993].

4.3 Computational Results

We judge the performance of our forecasting system on the basis of following criteria:

- Accuracy of our forecasting models on the basis of MAPE and Tracking signal values for different future time horizons.

- The average MAPE of our forecasting models compared with the average MAPE of same models for the complete dataset, available in [Makridakis and et. all., 1993]

- Testing for the conversion of non-stationary data to stationary data in case our automated ARIMA and SARIMA models.

4.3.1 Forecast Accuracy

Changing the forecasting horizon can affect the accuracy of forecast [Chu, 2009]. We test our methods with different forecasting horizons. We mentioned forecasting error comparison criteria in chapter 3. Along with MAPE, we are using Tracking signals to compare accuracy of our forecasting models. Equation 4.1, and equation 4.2 shows the formula for these error measurements. The Tracking Signal (TS) will check how biased our forecast is i.e., if at any period, the model underestimates (TS<-6) or overestimates (TS>+6) [Anvari et al., 2016]. The criteria used here are mostly for reporting the accuracy of our model (MAPE), while to compare and select the models we use RMSE.

$$MAPE = \sum_{t=1}^{i} |e_t|/y_t * (100/i) \tag{4.1}$$

$$TS = Bias_i/MAD_i,\ where\ Bias_i = \sum_{t=1}^{i}(e_t),\ MAD = \sum_{t=1}^{i} |e_t| * (1/i) \tag{4.2}$$

We test our accuracy for up to 6 periods ahead for different series. Table 4.1 and table 4.2 shows MAPE values for 1-3 and 4-6 periods ahead forecast values of all our models available in our forecasting system. Table 4.3 and table 4.4 shows Tracking signal values for 1-3 and 4-6 periods ahead forecast values of all our models available in our forecasting system. We also calculated the total MAPE of all the time series in our dataset for 1-6 periods ahead forecast in table 4.5.

One Period Ahead Forecast						
Series Name	S.M.A.	Ex.S.	Holt's	Winter's	ARIMA	SARIMA
PANTER	2.52	30.67	7.17	4.95	0.3866	0.00439
CHEETAH	13.21	2.191	10.56	30.53	3.924	0.0032
REALGNP	4.10	42.44	0.854	0.115	0.0208	0.0025
IPDGNP	3.679	57.59	1.62	2.76	0.045	0.00012
OEMUSAL	15.42	50.40	32.95	33.06	19.32	0.0054
TRADUSAL	71.81	61.09	3.86	22.79	4.862	0.0037
UNCPAPUSA	4.7945	32.37	10.18	5.299	2.40	0.0276
CPAPUSA	4.099	31.08	3.684	9.177	0.206	0.0011
Two Period Ahead Forecast						
Series Name	S.M.A.	Ex.S.	Holt's	Winter's	ARIMA	SARIMA
PANTER	10.51	18.55	12.38	10.16	1.225	0.427
CHEETAH	32.90	9.089	29.786	23.94	9.362	1.286
REALGNP	4.284	21.83	3.19	0.6254	0.753	0.0881
IPDGNP	3.45	29.92	2.136	4.232	0.749	0.9032
OEMUSAL	28.92	50.54	38.68	16.88	31.64	0.0037
TRADUSAL	83.35	28.42	68.58	35.41	18.31	0.7816
UNCPAPUSA	4.035	26.38	3.496	4.922	2.501	0.0172
CPAPUSA	3.728	24.85	0.51	3.955	1.884	0.0982
Three Period Ahead Forecast						
Series Name	S.M.A.	Ex.S.	Holt's	Winter's	ARIMA	SARIMA
PANTER	17.46	17.85	28.56	13.49	10.37	0.4877
CHEETAH	26.72	13.47	28.52	24.42	5.564	1.879
REALGNP	4.10	16.62	3.55	3.04	3.278	0.405
IPDGNP	2.915	21.19	1.736	5.448	2.07	1.614
OEMUSAL	36.38	44.91	33.44	11.39	21.35	7.15
TRADUSAL	87.79	66.30	95.57	31.87	73.80	5.81
UNCPAPUSA	5.35	27.83	5.609	3.200	4.163	0.294
CPAPUSA	5.455	19.88	4.469	27.34	2.951	0.374

Table 4.1: 1, 2 and 3 periods-ahead MAPE values of our Forecasting Models, Simple Moving Average(S.M.A),Exponential Smoothing (Ex. S.)

Four Period Ahead Forecast						
Series Name	S.M.A.	Ex.S.	Holt's	Winter's	ARIMA	SARIMA
PANTER	24.34	26.07	32.92	15.16	12.06	3.578
CHEETAH	23.20	19.81	15.46	26.12	13.76	4.816
REALGNP	3.731	13.49	2.44	3.16	3.70	0.312
IPDGNP	2.751	16.60	0.593	5.843	2.71	2.50
OEMUSAL	43.30	45.58	32.40	18.43	30.64	4.074
TRADUSAL	89.07	77.54	93.32	40.06	91.24	10.56
UNCPAPUSA	5.964	26.25	7.21	7.827	2.240	1.179
CPAPUSA	6.259	20.780	1.994	14.61	4.323	0.707
Five Period Ahead Forecast						
Series Name	S.M.A.	Ex.S.	Holt's	Winter's	ARIMA	SARIMA
PANTER	23.195	24.55	27.87	6.70	20.97	7.234
CHEETAH	15.86	14.86	12.84	13.95	11.58	8.10
REALGNP	3.513	10.90	2.142	2.261	3.627	1.083
IPDGNP	2.804	13.35	0.560	5.838	2.754	0.8209
OEMUSAL	47.54	46.04	26.86	37.10	30.30	13.69
TRADUSAL	84.84	233.83	88.97	30.77	63.16	23.27
UNCPAPUSA	6.110	26.06	3.917	10.26	4.85	1.990
CPAPUSA	5.404	17.01	4.90	10.40	6.175	1.087
Six Period Ahead Forecast						
PANTER	24.39	16.75	32.51	9.11	13.68	7.319
CHEETAH	11.78	12.49	12.93	18.93	20.32	6.350
REALGNP	3.70	9.641	2.508	1.954	3.566	0.84
IPDGNP	3.184	11.94	0.880	6.874	2.845	2.20
OEMUSAL	39.88	38.40	31.09	20.08	35.41	9.04
TRADUSAL	108.35	253.09	95.23	25.92	139.13	31.37
UNCPAPUSA	4.890	15.18	9.322	306.79	4.602	2.03
CPAPUSA	5.406	23.47	10.19	10.63	7.039	1.492

Table 4.2: 4, 5 and 6 periods-ahead MAPE values of our Forecasting Models, Simple Moving Average(S.M.A),Exponential Smoothing (Ex. S.)

One Period Ahead Forecast TS						
Series Name	S.M.A.	Ex.S.	Holt's	Winter's	ARIMA	SARIMA
PANTER	1.0	1.0	1.0	1.0	-1.0	-1.0
CHEETAH	1.0	-1.0	1.0	-1.0	-1.0	-1.0
REALGNP	-1.0	-1.0	-1.0	1.0	-1.0	-1.0
IPDGNP	-1.0	-1.0	-1.0	-1.0	-1.0	1.0
OEMUSAL	-1.0	-1.0	-1.0	1.0	-1.0	-1.0
TRADUSAL	-1.0	-1.0	-1.0	1.0	-1.0	-1.0
UNCPAPUSA	1.0	-1.0	1.0	1.0	1.0	1.0
CPAPUSA	1.0	-1.0	-1.0	-1.0	1.0	1.0
Two Period Ahead Forecast TS						
Series Name	S.M.A.	Ex.S.	Holt's	Winter's	ARIMA	SARIMA
PANTER	2.0	2.0	2.0	2.0	-0.941	-2.0
CHEETAH	2.0	2.0	2.0	1.0	-0.13	1.21
REALGNP	-2.0	-2.0	-2.0	-2.0	-2.0	1.722
IPDGNP	-2.0	-2.0	-2.0	-2.0	-2.0	0.405
OEMUSAL	-2.0	-2.0	-2.0	-0.859	-2.0	-2.0
TRADUSAL	-2.0	-1.67	-2.0	2.0	-2.0	2.0
UNCPAPUSA	0.285	-2.0	-1.77	2.0	-0.613	-2.0
CPAPUSA	0.171	-2.0	0.161	2.0	0.743	-0.496
Three Period Ahead Forecast TS						
PANTER	3.0	3.0	3.0	3.0	0.908	-3.0
CHEETAH	3.0	1.0	3.0	3.0	1.293	0.260
REALGNP	-3.0	-3.0	-3.0	-3.0	-3.0	-3.0
IPDGNP	-3.0	-3.0	-3.0	-3.0	-3.0	-1.53
OEMUSAL	-3.0	-3.0	-3.0	-1.236	-1.501	-0.92
TRADUSAL	-3.0	-3.0	-3.0	-3.0	-3.0	-2.057
UNCPAPUSA	-1.35	-3.0	-3.0	0.616	-1.08	-0.4795
CPAPUSA	1.728	-3.0	-3.0	-3.0	3.0	3.0

Table 4.3: 1, 2 and 3 periods-ahead forecasts TS values of our Forecasting Models, Simple Moving Average(S.M.A),Exponential Smoothing (Ex. S.),Tracking Signal (TS)

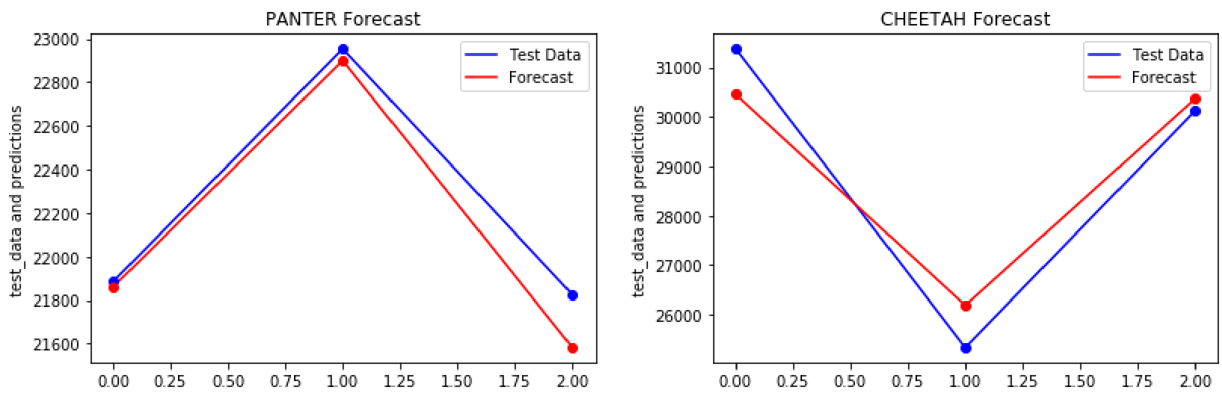
Four Period Ahead Forecast						
Series Name	S.M.A.	Ex.S.	Holt's	Winter's	ARIMA	SARIMA
PANTER	4.0	4.0	4.0	4.0	4.0	0.3276
CHEETAH	3.0	2.0	3.0	4.0	4.0	-1.834
REALGNP	-4.0	-4.0	-4.0	-4.0	-4.0	-0.354
IPDGNP	-4.0	-4.0	-3.016	-4.0	-4.0	-3.771
OEMUSAL	-4.0	-4.0	-4.0	2.337	-0.747	-4.0
TRADUSAL	-4.0	-3.717	-4.0	-4.0	-4.0	-1.33
UNCPAPUSA	0.0268	-4.0	-4.0	-4.0	-0.893	-3.329
CPAPUSA	2.90	-4.0	1.98	-4.0	3.622	1.9515
Five Period Ahead Forecast						
Series Name	S.M.A.	Ex.S.	Holt's	Winter's	ARIMA	SARIMA
PANTER	5.0	5.0	5.0	4.04	4.429	1.334
CHEETAH	1.823	0.588	0.445	-0.552	-0.952	-2.188
REALGNP	-5.0	-5.0	-5.0	-4.233	-4.705	4.367
IPDGNP	-5.0	-5.0	-4.486	-5.0	-4.086	-2.522
OEMUSAL	-5.0	-5.0	-5.0	-5.0	-5.0	-4.19
TRADUSAL	-4.0	-3.937	-4.0	-4.0	-5.0	1.76
UNCPAPUSA	-1.19	-5.0	-3.340	-3.116	-0.648	-2.426
CPAPUSA	4.841	-5.0	5.0	5.0	1.71	0.3611
Six Period Ahead Forecast						
PANTER	6.0	6.0	6.0	6.0	4.4298	1.334
CHEETAH	-1.66	-3.53	0.049	-4.79	-0.9522	-2.188
REALGNP	-6.0	-6.0	-6.0	-4.936	-6.0	2.046
IPDGNP	-6.0	-6.0	-6.0	-6.0	-6.0	-6.0
OEMUSAL	-6.0	-5.492	-4.120	-4.168	-1.242	-3.36
TRADUSAL	-5.54	-4.359	-5.647	-5.659	-5.423	2.12
UNCPAPUSA	-0.288	-6.0	-6.0	6.0	-3.41	-4.29
CPAPUSA	-0.345	-6.0	-6.0	5.90	0.373	1.827

Table 4.4: 4, 5 and 6 periods-ahead forecast TS values of our Forecasting Models, Simple Moving Average(S.M.A),Exponential Smoothing (Ex. S.),Tracking Signal(TS)

Total Average MAPE of Dataset						
Periods Ahead	S.M.A.	Ex.S.	Holt's	Winter's	ARIMA	SARIMA
One	14.95	38.47	8.85	13.58	3.89	0.0061
Two	21.39	26.19	19.84	12.51	8.303	0.45
Three	23.27	28.50	25.18	15.024	15.44	2.25
Four	24.83	30.76	23.29	16.40	20.08	3.46
Five	23.66	48.35	21.00	14.70	17.92	7.159
Six	25.20	47.62	24.33	50.04	30.21	7.58

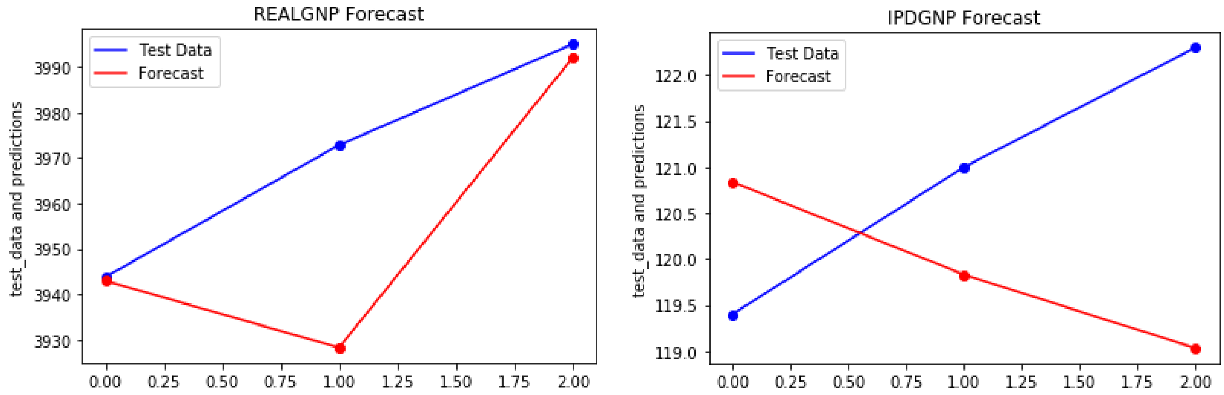
Table 4.5: 1-6 periods-ahead forecast total MAPE of the Dataset, Simple Moving Average(S.M.A),Exponential Smoothing (Ex. S.),Tracking Signal(TS)

We show results of 3 periods and 6 periods ahead in graphical format. For 3 periods ahead forecast, SARIMA outperform all other models.Figures 4.1, 4.2, 4.3 and 4.4 show best 3 peiods-ahead forecast for all the series. As we move towards 6 periods ahead forecast, the accuracy of models reduce. For 6 periods ahead forecast, our SARIMA model worked best for all of the series except 1 (TRADUSAL). The winter's model gives lower MAPE than SARIMA in this case. Figures 4.5, 4.6, 4.7 and 4.8 show best 6 peiods-ahead forecast for all the series.



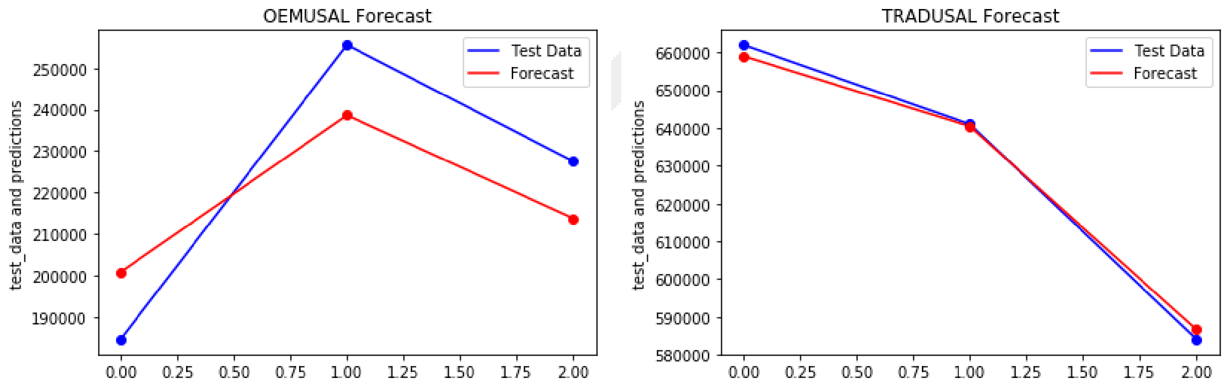
(a) (b)

Figure 4.1: 3 periods-ahead Test data and Predictions for (a) PANTER and (b) CHEETAH Series



(a) (b)

Figure 4.2: 3 periods-ahead Test data and Predictions for (a) REALGNP and (b) IPDGNP Series

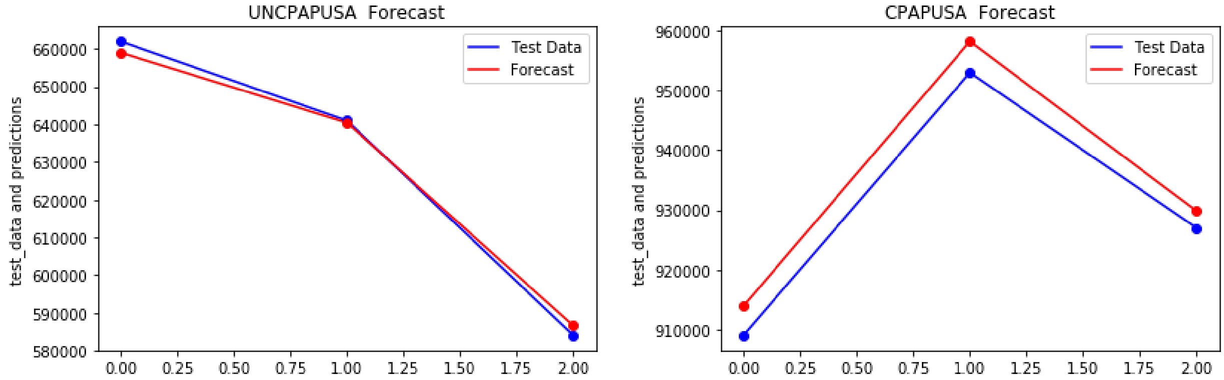


(a) (b)

Figure 4.3: 3 periods-ahead Test data and Predictions for (a) OEMUSAL and (b) TRADUSAL Series

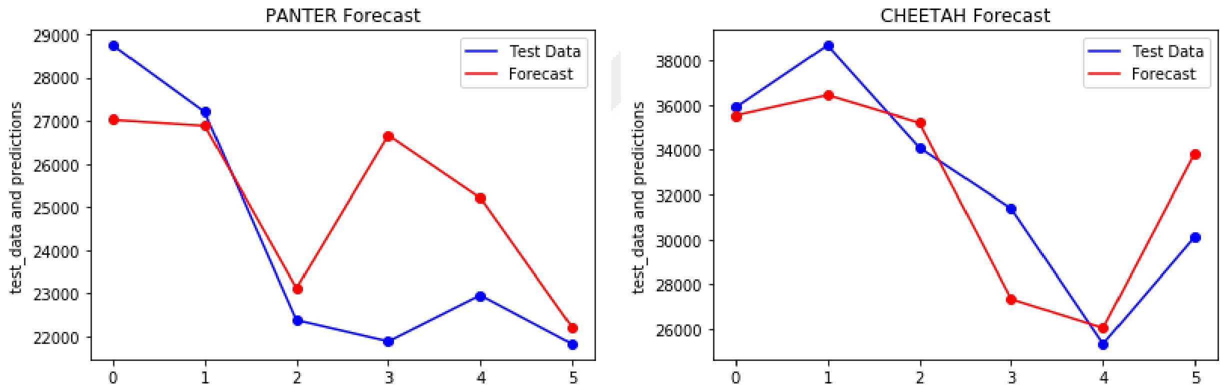
4.3.2 Comparison of Average MAPE of Dataset

[Makridakis and et. all., 1993] provides the average MAPE of 23 monthly series for M2-competition dataset for multiple forecasting models. In this section we compare this information with our models average MAPE for our models. As concluded earlier that our Box-Jenkins SARIMA outperforms all other models, so our focus will be comparing SARIMA results with the results provided in [Makridakis and et. all., 1993]. Table 4.6 provides the comparison between average 1-6 periods ahead



(a) (b)

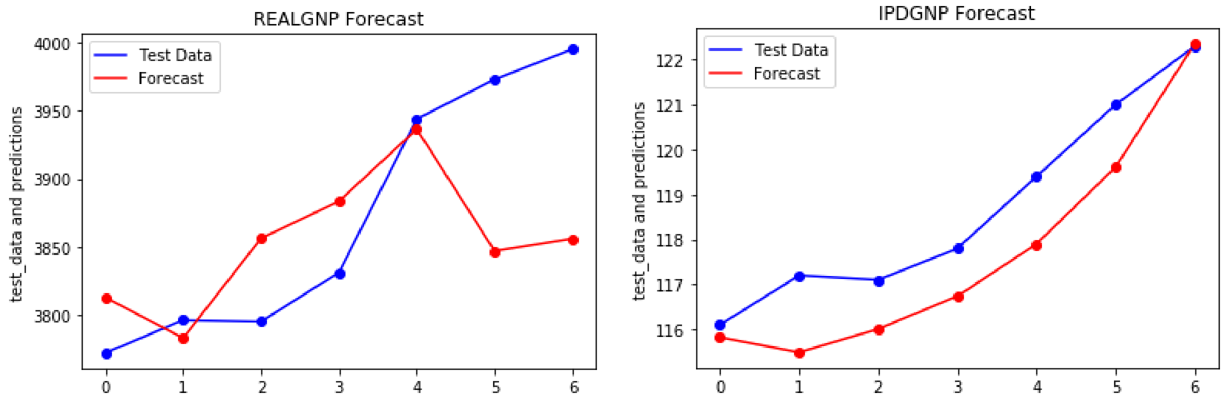
Figure 4.4: 3 periods-ahead Test data and Predictions for (a) UNCPAPUSA and (b) CPAPUSA Series



(a) (b)

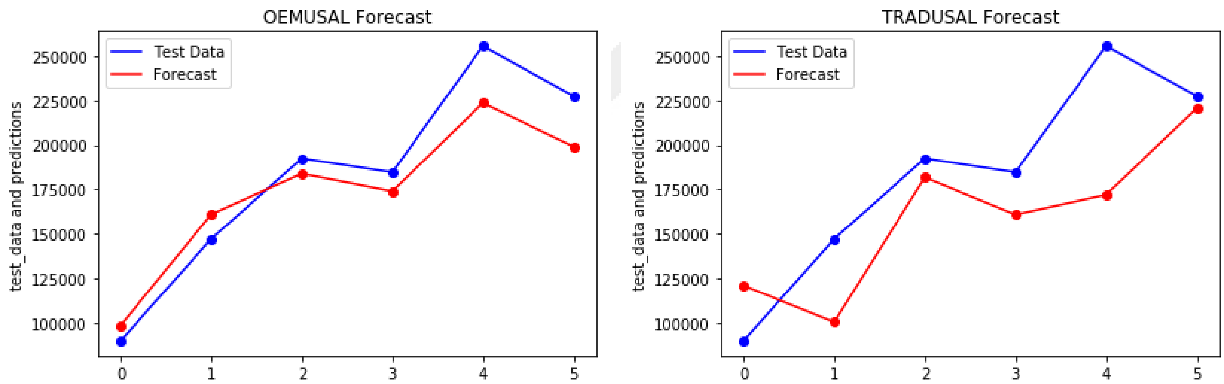
Figure 4.5: 6 periods-ahead Test data and Predictions for (a) PANTER and (b) CHEETAH Series

ARIMA and SARIMA forecast MAPE of 23 time series with the average MAPE of Box-Jenkins methodology provided by [Makridakis and et. all., 1993]. Table 4.7 provides the comparison between basic model approach Naive 1 and our automated optimized Moving Average. Whereas, Table 4.8 provides the comparison between basic model approach Holt and Winters and our automated optimized Holts and Winters models.



(a) (b)

Figure 4.6: 6 periods-ahead Test data and Predictions for (a) REALGNP and (b) IPDGNP Series

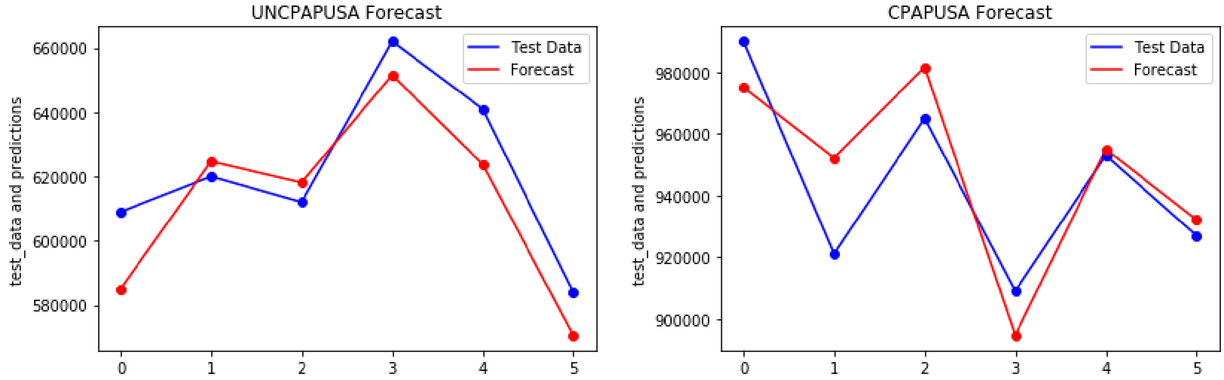


(a) (b)

Figure 4.7: 6 periods-ahead Test data and Predictions for (a) OEMUSAL and (b) TRADUSAL Series

4.3.3 Accuracy of Non-Stationary Series Identification and Transformation

For ARIMA and SARIMA models, one of the most important step is identification of non-stationary data and its conversion. We achieve this by the aid of statistical tests and differencing on the data. We test our method accuracy by comparing the plots of our transformed data with original data. Figure 4.9a and figure 4.13a shows our input data from two different time series. From the plots, we can see that there is trend, level changing in former and increasing variance in latter; confirming



(a) (b)

Figure 4.8: 6 periods-ahead Test data and Predictions for (a) UNCPAPUSA and (b) CPAPUSA Series

MAPE comparison of Box-Jenkins and our automated ARIMA SARIMA						
Model	One	Two	Three	Four	Five	Six
Box-Jenkins	14.4	7.6	15.6	12.4	15.9	22.2
Automated ARIMA	5.68	14.22	20.287	20.50	23.01	17.94
Automated SARIMA	0.013	1.54	3.367	5.68	8.65	9.36

Table 4.6: MAPE comparison of Box-Jenkins and our automated ARIMA SARIMA models

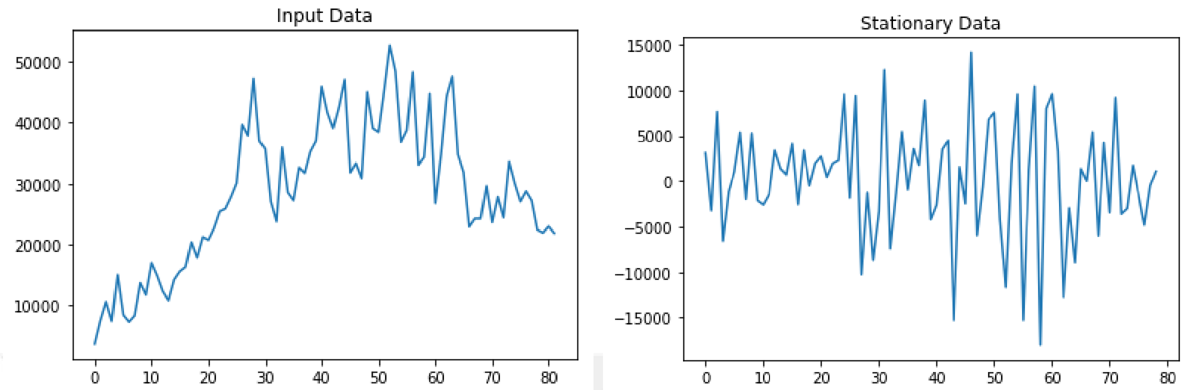
that the data is non-stationary. We identify this non-stationary data automatically and apply transformations. Figure 4.9b and figure 4.13b shows the transformed stationary data plots. This is also confirmed by the ACF plots in figure 4.10 and 4.14. The train data ACF plots decays slowly, hence confirming non-stationary data. While the transformed plots in 4.10b and 4.14b shows that plots reaches 0 quickly, thus confirming stationary data.

MAPE comparison of Navie 1 and our optimized Moving Average						
Model	One	Two	Three	Four	Five	Six
Naive 1	9.2	16.1	21.4	20.5	29.9	34.3
Our MA	29.840	33.77	29.35	25.766	24.276	25.105
MAPE comparison of Single and our optimized Exponential Smoothing						
Model	One	Two	Three	Four	Five	Six
Single	11.1	9.4	9.7	10.1	17.8	14.9
ES	37.89	33.70	32.98	25.89	27.34	29.09

Table 4.7: MAPE comparison of Naive 1 and our automated Moving Average model, Single and our automated Exponential smoothing model for 1-6 periods ahead forecast, Moving Average (MA), Exponential Smoothing (ES)

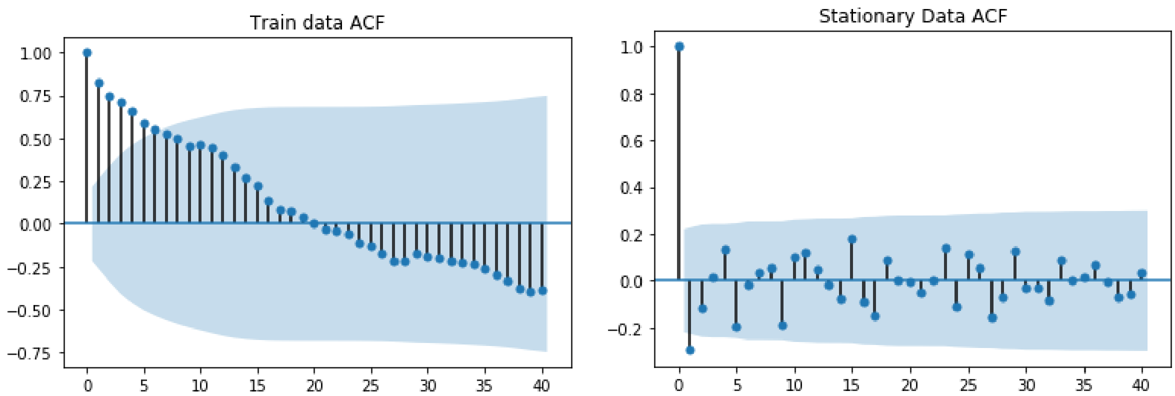
MAPE comparison of Holts and our optimized Holts model						
Model	One	Two	Three	Four	Five	Six
Holts	23.59	24.64	26.38	22.07	24.28	24.23
Our Holts	29.840	33.77	29.35	25.766	24.276	25.105
MAPE comparison of Winters and our optimized Winters						
Model	One	Two	Three	Four	Five	Six
Winters	6.5	8.1	8.4	11.5	17.20	16.1
Our Winters	26.86	21.24	29.76	92.48	21.62	60

Table 4.8: MAPE comparison of Holts and our automated Holts, Winters and our automated optimized Winters model for 1-6 periods ahead forecast



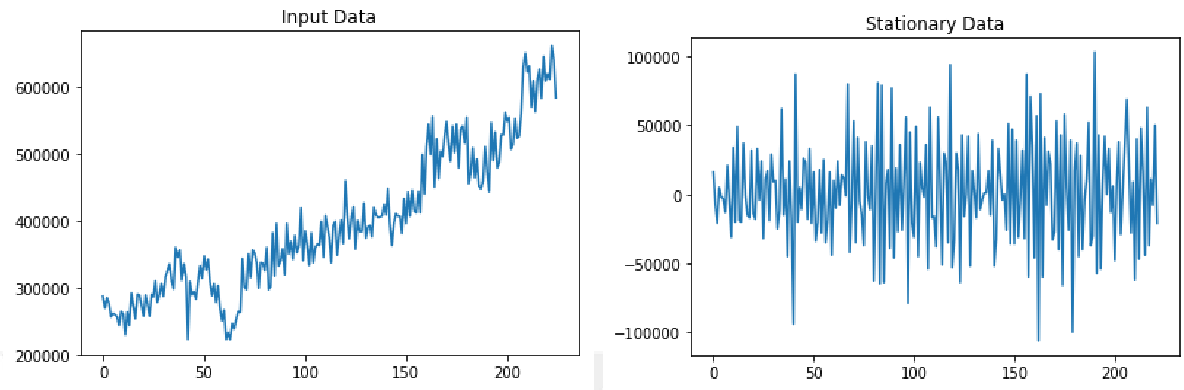
(a) (b)

Figure 4.9: PANTER SERIES: (a) Non-stationary Input Data plot (b) Stationary Data after Transformation plot.



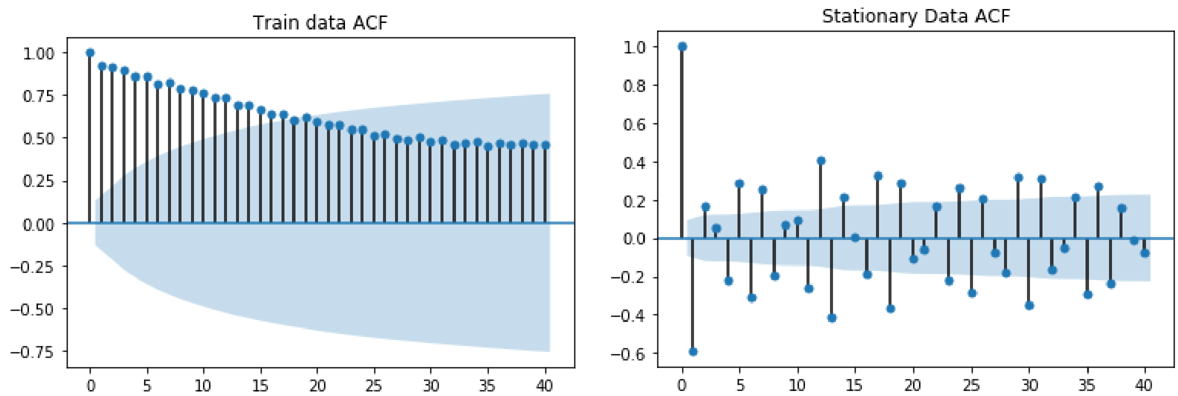
(a) (b)

Figure 4.10: PANTER SERIES: (a) Non-stationary Input Data ACF plot (b) Stationary Data after Transformation ACF plot.



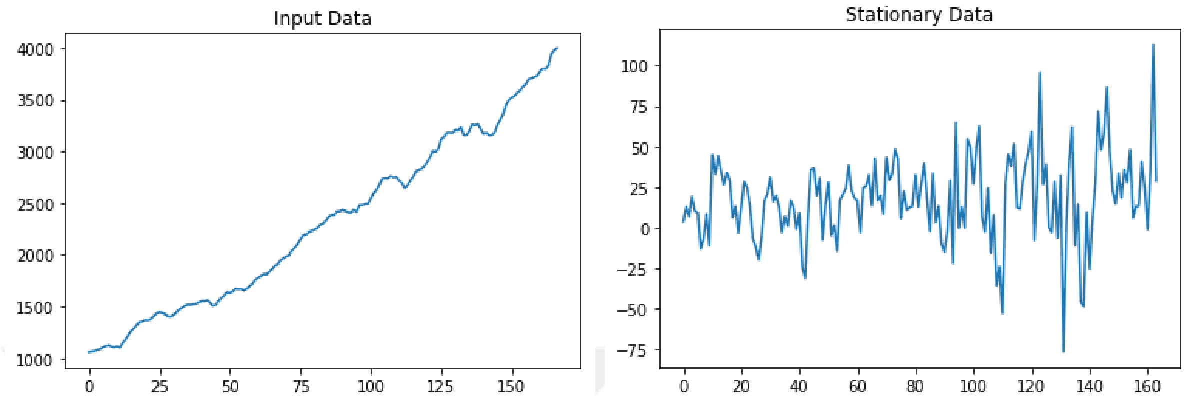
(a) (b)

Figure 4.11: UNCPAPUSA SERIES: (a) Non-stationary Input Data (b) Stationary Data after Transformation.



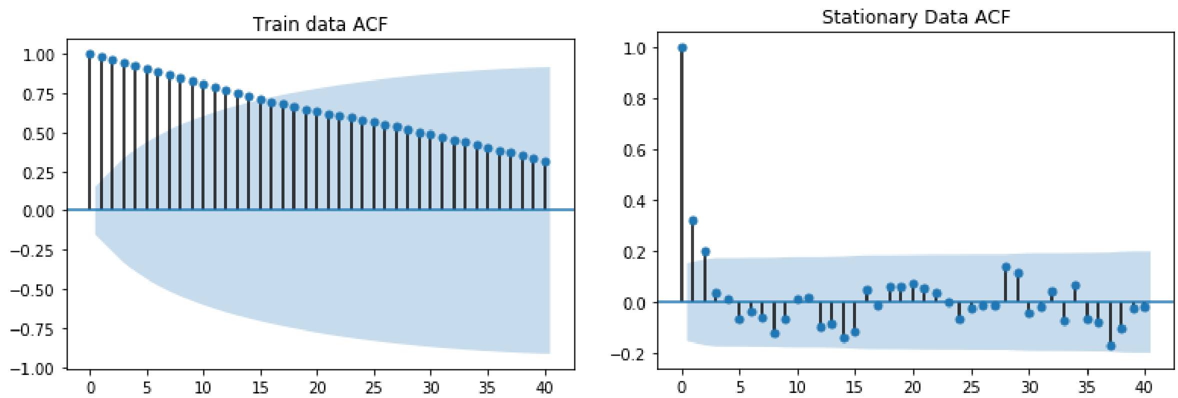
(a) (b)

Figure 4.12: UNCPAPUSA SERIES: (a) Non-stationary Input Data ACF plot (b) Stationary Data after Transformation ACF plot.



(a) (b)

Figure 4.13: REALGNP SERIES: (a) Non-stationary Input Data (b) Stationary Data after Transformation.



(a) (b)

Figure 4.14: REALGNP SERIES: (a) Non-stationary Input Data ACF plot (b) Stationary Data after Transformation ACF plot.

Chapter 5

CONCLUSION

Time series forecasting is a well known subject in supply-chain planning. There has been many studies in literature about it but not a single one which implements, compares multiple models, and selects the best fitting model automatically. There is always a need of manual input or expert judgement involved. In this thesis, we consider multiple basic and advance time series models and form an automated multi-model approach for time series forecasting. For our ARIMA and SARIMA models, the Box–Jenkins framework is extended to an integrated and automated framework. Our frameworks can identify the non stationary data and transform it to stationary data, and can provide reliable information about the best fitting model in reasonable amount of time

The objective of our approach is to select the best fitting model for time series data automatically and thus increase the accuracy of forecast. We tested our system on multiple time series from M-competition data set and had some interesting results:

- As we increase the forecasting horizons, the MAPE values also increases.
- For 7 out of 8 series, SARIMA outperforms other forecasting models.
- For 1 period-ahead forecast, our SARIMA model is able to forecast with $\leq$1% MAPE.
- For 6 period-ahead forecast, SARIMA model outperforms with average MAPE of 7%. However, in only 1 series Winter's model outperforms all other models.
- For accuracy purposes, we checked out non-stationary identification and transformation via ACF plots. The results confirmed that our ARIMA SARIMA

model is able to identify and convert non-stationary series to stationary series.

- We tested our forecast for Tracking signals accuracy as well. We set the threshold to (TS<-6) or (TS>+6) to check forecast bias. None of the TS values exceed -6 or +6.

- We also compared our models results with the forecasting results available in [Makridakis and et. all., 1993]. Our automated SARIMA performs better than Box-Jenkins methodology with an average MAPE of 9% for 6 perios-ahead forecast and $\leq$1% for 1 period-ahead forecast for 23 monthly series dataset.

We thus conclude that our framework is able to forecast upto 6 period ahead accurately with an average MAPE of $\leq$ 7%. Our framework is a user-friendly system and is very easy to use. The user only needs to provide input data and the system generates forecast on the basis of best fitting model.

5.1 Future Work

Time series forecasting is a deep subject and there are many dimensions to it. We covered a part of it in our thesis. There is still possibility of adding more knowledge based dimensions to it. The accuracy of our forecasting system can further be increased by learning more about the input data, dealing with anomalies or outliers of input data and applying some machine learning concepts to clean the data before using it in forecasting. Our system compares only a few basic time series models, addition of more models may increase the accuracy of forecast further. Addition of neural network models can also be helpful in further comparison of time series models.We aim to extend our work to knowledge intensive models, for instance leading indicators and bass diffusion.

BIBLIOGRAPHY

[Alysha M. De Livera, 2011] Alysha M. De Livera, Rob J. Hyndman, R. D. S. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, 106(496):513–1527.

[Anvari et al., 2016] Anvari, S., Tuna, S., Canci, M., and Turkay, M. (2016). Automated box–jenkins forecasting tool with an application for passenger demand in urban rail systems. *JOURNAL OF ADVANCED TRANSPORTATION*, 1(50):25 – 49.

[Bates and Granger, 1969] Bates and Granger (1969). The combination of forecast. *Operational Research Quarterly*, 20(4):451–468.

[Bougas, 2013] Bougas, C. (2013). Forecasting air passenger traffic flows in canada: an evaluation of time series models and combination methods. Master's thesis, Laval University.

[Bowerman and Connell, 1979] Bowerman, B. L. and Connell, R. T. (1979). *Forecasting and time series.* Duxbury Press.

[Brown and Murphy, 1996] Brown and Murphy (1996). Improving forecasting performance by combining forecasts: The example of road-surface temperature forecasts. *Meteorological Applications*, 3(3):257–265.

[Burkom et al., 2007] Burkom, H. S., Murphy, S. P., and Shmueli, G. (2007). Automated time series forecasting for biosurveillance. *STATISTICS IN MEDICINE*, 26:4202–4218.

[C-o and BJ, 2005] C-o, O. and BJ, M. (2005). Evaluating time-series models to forecast the demand for tourism in singapore comparing within-sample and post sample results. *Journal of Travel Research*, 4(43):404–413.

[Chakrabarti and Faloutsos, 2002] Chakrabarti, D. and Faloutsos, C. (2002). F4: Large-scale automated forecasting using fractals. In *CIKM '02: Proceedings of the eleventh international conference on Information and knowledge management.*

[Chopra and Meindl, 2016] Chopra, S. and Meindl, P. (2016). *Supply chain management, Strategy Planning and Operation.* Pearson Education.

[Chu, 2009] Chu, F.-L. (2009). Forecasting tourism demand with arma-based methods. *Tourism Management*, 30(5):740–751.

[Davide Burba, 2019] Davide Burba (2019 (accessed December 6, 2019)). Towards data science tutorial. `https://https://towardsdatascience.com/an-overview-of-time-series-forecasting-models-a2fa7a358fcb/`.

[Emiray and Rodríguez, 2003] Emiray and Rodríguez (2003). Evaluating time series models in short and long-term forecasting of canadian air passenger data. Master's thesis, Department of Economics, University of Ottawa.

[F-L, 1998] F-L, C. (1998). Forecasting tourist arrivals: nonlinear sine wave or arima? *Journal of Travel Research*, 36(3):79–84.

[Hwang et al., 2012] Hwang, S., Park, M., and et. all. (2012). Automated time-series cost forecasting system for construction materials. *Journal of Construction Engineering and Management*, 138(11).

[Hyndman and Khandakar, 2007] Hyndman, R. and Khandakar (2007). Automatic time series for forecasting: The forecast package for r. Technical report, Monash University, Department of Econometrics and Business Statistics.

[IIF, 2019] IIF (2018 (accessed December 20, 2019)). International institute of forecasters. `https://forecasters.org/resources/time-series-data/m-competition/`.

[Inc, 1998] Inc, A. (1998). Autobox software user manual, automatic forecasting systems. Technical report, Hatboro, PA.

[J., 1996] J., C. (1996). Time series analyses of uk outbound travel by air. *Journal of Travel Research*, 44(3):335–347.

[Josef Perktold et.al, 2019] Josef Perktold et.al (2009 (accessed December 14, 2019)). Runs test statsmodels python library. `https://www.statsmodels.org`.

[KRASIC and GATTI, 2009] KRASIC, D. and GATTI, P. (2009). Forecasting methodology of maritime passenger demand in a tourist destination. *PROMET Traffic and Transportation*, 3(21):183–190.

[Kwiatkowski et al., 1992] Kwiatkowski, PCB, P., P, S., and Y, S. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1):159–178.

[Makridakis and et. all., 1993] Makridakis, S. and et. all. (1993). The m2-competition: A real-time judgmentally based forecasting study. *International Journal of Forecasting*, 9(1):5–22.

[MM, 1999] MM, H. (1999). Stochastic modeling of airlines' scheduled services revenue. *Journal of Air Transportation World Wide*, 4(2):32–48.

[Mwakalonge, 2014] Mwakalonge, B. (2014). Trip generation modeling using data collected in single and repeated cross-sectional surveys. *Journal of Advanced Transportation*, 48(4):318–331.

[Nguyen and Hansen, 2017] Nguyen, H. and Hansen, C. K. (2017). Short-term electricity load forecasting with time series analysis. In *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*.

[Oh C-o, 2005] Oh C-o, M. B. (2005). Evaluating time-series models to forecast the demand for tourism in singapore comparing within-sample and postsample results. *Journal of Travel Research*, 43(4):404–413.

[Omane-Adjepong and et. all., 2013] Omane-Adjepong, M. and et. all. (2013). Determining the better approach for short-term forecasting of ghana's inflation: Seasonal-arima vs. holt-winters. *Journal of the American Statistical Association*, 3(1):69–79.

[SF Witt, 1991] SF Witt, C. W. (1991). Forecasting tourism demand with arma-based methodstourism forecasting: error magnitude, direction of change error, and trend change error. *Journal of Travel Research*, 30(2):26–33.

[SHARDA and PATIL, 1992] SHARDA, B. and PATIL, R. B. (1992). Connectionist approach to time series prediction: an empirical test. *Journal of Intelligent Manufacturing*, 3(1):317–323.

[Timmermann, 2013] Timmermann, A. (2013). Forecast combinations. Technical report, Oxford.

[Uysal M, 1985] Uysal M, C. J. (1985). An overview of approaches used to forecast tourism demand. *Journal of Travel Research*, 4(23):7–15.

[Y and SM, 2009] Y, L. and SM, A. (2009). Automated box–jenkins forecasting modelling. *Automation in Construction*, 18(5):547–558.

[Y-M., 1993] Y-M., C. (1993). Forecasting tourism: a sine wave time series regression approach. *Journal of Travel Research*, 32(2):58–60.

[Yaffee and McGee, 2000] Yaffee, R. A. and McGee, M. (2000). *An Introduction to Time Series Analysis and Forecasting: With Applications of SAS and SPSS.* Academic Press, Inc.

Appendix A

All the time series used from M-Competition Dataset are copied here. Due to large number of observations, we have pasted partial timeseries here. The full dataset can be accessed from [IIF, 2019]

Series Name: PANTER Seasonality: 12 Start Date:1982/JAN End Date:1988/OCT		Series Name: PANTER Seasonality: 12 Start Date: 1982/JAN End Date: 1988/OCT	
Index	Observations	Index	Observations
1	3619.8	1	5987.7
2	7466.7	2	13104.9
3	10579	3	16807.4
4	7369.1	4	8265.4
5	14987.7	5	12009.9
6	8398.8	6	6091.4
7	7254.3	7	5948.1
8	8292.6	8	7970.4
9	13658	9	13437
10	11680.2	10	8769.1
11	16948.1	11	12838.3
12	14821	12	10414.8
13	12240.7	13	12503.7
14	10781.5	14	13874.1
15	14179	15	19164.2
16	15539.5	16	23272.8
17	16254.3	17	26069.1
18	20356.8	18	33858
19	17812.3	19	27498.8
20	21209.9	20	22993.8
21	20711.1	21	21445.7
22	22659.3	22	28729.6
23	25397.5	23	27837
24	25854.3	24	26179
25	27785.2	25	28044.4
26	30090.1	26	29770.4
27	39672.8	27	37508.6
28	37833.3	28	39435.8
29	47240.7	29	46766.7
30	36976.5	30	38927.2
31	35719.8	31	39238.3
32	27016	32	34993.8
33	23725.9	33	37118.5
34	35946.9	34	30745.7
35	28540.7	35	30018.5

Table A.1: Part a-M Competition Dataset series PANTER and CHEETAH

Series Name: REALGNP Seasonality: 4 Start Date:1947/I End Date:1947/I		Series Name: IPDGNP Seasonality: 4 Start Date: 1988/III End Date: 1988/III	
Index	Observations	Index	Observations
1	1056.5	1	21.5
2	1063.2	2	21.8
3	1067.1	3	22.1
4	1080	4	22.8
5	1086.8	5	23.2
6	1106.1	6	23.5
7	1116.3	7	23.9
8	1125.5	8	23.8
9	1112.4	9	23.6
10	1105.9	10	23.4
11	1114.3	11	23.4
12	1103.3	12	23.4
13	1148.2	13	23.5
14	1181	14	23.6
15	1225.3	15	24.2
16	1260.2	16	24.5
17	1286.6	17	25.1
18	1320.4	18	25.1
19	1349.8	19	25
20	1356	20	25.2
21	1369.2	21	25.2
22	1365.9	22	25.3
23	1378.2	23	25.5
24	1406.8	24	25.9
25	1431.4	25	25.9
26	1444.9	26	25.9
27	1438.2	27	26
28	1426.6	28	25.8
29	1406.8	29	26.2
30	1401.2	30	26.3
31	1418	31	26.3
32	1438.8	32	26.5
33	1469.6	33	26.8
34	1485.7	34	27.1
35	1505.5	35	27.3

Table A.2: Part a: M-Competition Dataset series REALGNP and IPDGNP

Series Name: REALGNP Seasonality: 4 Start Date:1947/I End Date:1947/I		Series Name: IPDGNP Seasonality: 4 Start Date: 1988/III End Date: 1988/III	
Index	Observations	Index	Observations
106	2741	106	49
107	2738.3	107	50
108	2762.8	108	51.2
109	2747.4	109	51.9
110	2755.2	110	53
111	2719.3	111	54.8
112	2695.4	112	56.3
113	2642.7	113	57.7
114	2669.6	114	58.6
115	2714.9	115	59.9
116	2752.7	116	61
117	2804.4	117	61.7
118	2816.9	118	62.5
119	2828.6	119	63.4
120	2856.8	120	64.5
121	2896	121	65.6
122	2942.7	122	66.9
123	3001.8	123	67.7
124	2994.1	124	68.9
125	3020.5	125	69.9
126	3115.9	126	71.6
127	3142.6	127	72.9
128	3181.6	128	74.4
129	3181.7	129	76.1
130	3178.7	130	77.8
131	3207.4	131	79.4
132	3201.3	132	81
133	3233.4	133	82.7
134	3157	134	84.6
135	3159.1	135	86.5
136	3199.2	136	89
137	3261.1	137	91.3
138	3250.2	138	92.8
139	3264.6	139	94.9
140	3219	140	96.7

Table A.3: Part d: M-Competition Dataset series REALGNP and IPDGNP

Series Name: REALGNP Seasonality: 4 Start Date:1947/I End Date:1947/I		Series Name: IPDGNP Seasonality: 4 Start Date: 1988/III End Date: 1988/III	
Index	Observations	Index	Observations
141	3170.4	141	98.2
142	3179.9	142	99.4
143	3154.5	143	100.8
144	3159.3	144	101.7
145	3186.6	145	102.5
146	3258.3	146	103.3
147	3306.4	147	104.2
148	3365.1	148	105.4
149	3451.7	149	106.5
150	3498	150	107.3
151	3520.6	151	108.2
152	3535.2	152	109
153	3568.7	153	109.9
154	3587.1	154	110.8
155	3623	155	111.6
156	3650.9	156	112.4
157	3698.8	157	112.9
158	3704.7	158	113.7
159	3718	159	114.7
160	3731.5	160	114.9
161	3772.2	161	116.1
162	3796.4	162	117.2
163	3795.3	163	117.1
164	3831.2	164	117.8
165	3943.8	165	119.4
166	3972.8	166	121
167	3995	167	122.3

Table A.4: Part e: M-Competition Dataset series REALGNP and IPDGNP

Series Name: OEMUSAL Seasonality: 12 Start Date:1981/JAN End Date:1988/OCT		Series Name: TRADUSAL Seasonality: 12 Start Date: 1981/JAN End Date: 1988/OCT	
Index	Observations	Index	Observations
71	190324	71	32715
72	99187	72	13574
73	114183	73	10513
74	89600	74	5823
75	86995	75	3741
76	100604	76	5663
77	103692	77	2295
78	78944	78	2869
79	138824	79	33000
80	173460	80	47613
81	192976	81	37421
82	222062	82	53688
83	136549	83	29119
84	154170	84	21234
85	166088	85	23697
86	81809	86	5846
87	87814	87	3608
88	128723	88	4260
89	89886	89	1738
90	147126	90	2106
91	192389	91	14231
92	184566	92	52432
93	255535	93	41283
94	227465	94	50487

Table A.5: Part c: M-Competition Dataset series OEMUSAL and TRADUSAL

Series Name: UNCPAPUSA Seasonality: 12 Start Date:1970/JAN End Date:1988/SEP		Series Name: CPAPUSA Seasoanlity: 12 Start Date: 1970/JAN End Date: 1988/SEP	
Index	Observations	Index	Observations
1	287000	1	464000
2	269000	2	405000
3	285000	3	457000
4	277000	4	459000
5	256000	5	453000
6	261000	6	448000
7	259000	7	422000
8	256000	8	433000
9	243000	9	433000
10	264000	10	460000
11	260000	11	438000
12	229000	12	417000
13	263000	13	436000
14	243000	14	431000
15	292000	15	507000
16	273000	16	484000
17	253000	17	471000
18	290000	18	478000
19	289000	19	419000
20	274000	20	477000
21	257000	21	453000
22	289000	22	474000
23	275000	23	450000
24	257000	24	442000
25	290000	25	471000
26	286000	26	471000
27	310000	27	552000
28	278000	28	489000
29	289000	29	537000
30	306000	30	505000
31	287000	31	446000
32	316000	32	501000
33	325000	33	519000
34	335000	34	554000
35	310000	35	536000

Table A.6: Part a: M-Competition Dataset series UNCPAPUSA and CPAPUSA

Series Name: UNCPAPUSA Seasonality: 12 Start Date:1970/JAN End Date:1988/SEP		Series Name: CPAPUSA Series Name: 12 Start Date: 1970/JAN End Date: 1988/SEP	
Index	Observations	Index	Observations
176	545000	176	724000
177	479000	177	712000
178	537000	178	797000
179	542000	179	733000
180	516000	180	693000
181	555000	181	777000
182	455000	182	763000
183	472000	183	865000
184	509000	184	855000
185	464000	185	827000
186	492000	186	771000
187	452000	187	761000
188	448000	188	808000
189	459000	189	785000
190	511000	190	912000
191	474000	191	869000
192	444000	192	775000
193	547000	193	890000
194	490000	194	827000
195	533000	195	879000
196	479000	196	866000
197	487000	197	904000
198	529000	198	875000
199	529000	199	889000
200	562000	200	922000
201	549000	201	841000
202	555000	202	922000
203	507000	203	839000
204	515000	204	856000
205	553000	205	850000
206	524000	206	844000
207	527000	207	900000
208	561000	208	892000
209	630000	209	902000
210	651000	210	923000

Table A.7: Part f: M-Competition Dataset series UNCPAPUSA and CPAPUSA

Series Name: UNCPAPUSA Seasonality: 12 Start Date:1970/JAN End Date:1988/SEP		Series Name: CPAPUSA Series Name: 12 Start Date: 1970/JAN End Date: 1988/SEP	
Index	Observations	Index	Observations
211	623000	211	945000
212	632000	212	996000
213	570000	213	946000
214	610000	214	1029000
215	563000	215	919000
216	611000	216	978000
217	627000	217	844000
218	583000	218	917000
219	646000	219	1032000
220	609000	220	990000
221	620000	221	921000
222	612000	222	965000
223	662000	223	909000
224	641000	224	953000
225	584000	225	927000

Table A.8: Part g: M-Competition Dataset series UNCPAPUSA and CPAPUSA