FORECASTING MONTHLY TOURIST ARRIVALS TO SRI LANKA

Chapa Lakmali Jayasundara

158880F

Degree of Master of Science in Business Statistics

Department of Mathematics

University of Moratuwa
Sri Lanka

June 2019

FORECASTING MONTHLY TOURIST ARRIVALS TO SRI LANKA

Chapa Lakmali Jayasundara

158880F

Thesis/Dissertation submitted in partial fulfillment of the requirements for the Degree of Master of Science in Business Statistics

Department of Mathematics

University of Moratuwa
Sri Lanka

June 2019

DECLARATION

I declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles and books).

Signature Date

The above candidate has carried out research for the Masters Dissertation under my supervision.

Name of the supervisor:

Signature of the supervisor: Date:

ABSTRACT

The accurate forecasts of tourist arrivals have a significant impact on the economy of a country. The patterns of arrivals from different countries may be varied in a same region. Thus, forecasting the tourist arrivals to Sri Lanka is important for decision making processes. On the view of the above, the study has focused to forecast monthly tourist arrivals from highest tourist generating countries such as United Kingdom and India. In this context three forecasting techniques, namely Seasonal ARIMA, Holt Winters (HW) Multiplicative model and Holt Winters Additive model were employed to find the most appropriate model with least forecasting error. The models were trained from monthly tourist arrivals for the period from November 2010 to August 2017 and validated using data from September, 2017 to February, 2018. The Seasonal ARIMA$(0,1,1,) \times (0,1,1)_{12}$ was identified as the best fitted model for both countries. Among Holt Winters models, Holt Winters multiplicative model with smoothing constants $\alpha = 0.3,\ \beta = 0.1$ and $\gamma = 0.1$ was found to be the most suitable model for both countries. In both models errors were found to be white noise. The forecasts of monthly tourist arrivals from the UK and India from both models have high accuracy as corresponding values of percentage errors were within $\pm 10\ \%$ and MAPE is less than 10% for independent set. By comparing percentage error for both training and validation set it was found that SARIMA is more superior than HW. The percentage changes of monthly tourist arrivals reveal that it can be expected an increment of monthly tourist arrivals in the coming months. The models developed in this study are recommended to use for policy decisions in medium term forecasting and which would be useful for the tourism industry in Sri Lanka.

Keywords: Forecasting, Holt Winters method, SARIMA, Tourist arrivals

DEDICATION

I dedicate this to my father who directed me and made me interesting for Statistics & to my mother and sister who always encourage me to achieve my targets.

Acknowledgement

First and foremost I offer my sincerest gratitude to my supervisor Prof. T. S. G. Peiris, Professor in Applied Statistics, Department of Mathematics, University of Moratuwa and the course coordinator of M.Sc/ PG Diploma in Business Statistics course who has supported me throughout this thesis with his patience and knowledge.

I wish to extend my sincere thanks to all the lecturers who gave me knowledge of different subjects in degree of Business Statistics. Finally, I am thankful to my family members and all my colleagues for their help and encouragement throughout this period.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

Page

LIST OF TABLES

LIST OF TABLES (CONTINUED)

LIST OF TABLES (CONTINUED)

LIST OF ABBREVIATIONS

ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller Test
AIC	Akaike Information Criterion
AR	Autoregressive Components
ARIMA	Autoregressive Integrated Moving Average Model
ARMA	Autoregressive Moving Average Model
BIC	Bayesian Information Criterion
MA	Moving Average Components
MAPE	Mean Absolute Percentage Error
PACF	Partial Autocorrelation Function
SACF	Sample Autocorrelation Function
SPACF	Sample Partial Autocorrelation Function

LIST OF APPENDICES

CHAPTER 1
INTRODUCTION

This chapter discusses the background, main research problem, objectives, the significance of the study and the outline of the research.

1.1 Overview of Tourism in Sri Lanka

Tourism is a fastest and growing industry in Sri Lanka as well as in many economies of the present world. To pursue an efficient and effective planning strategies and to avoid the imbalances between demand and supply, forecasting of tourism demand is a necessity. Tourism demand can be measured with regard to tourist arrivals and/or departures, tourist expenditures and/or receipts, travel exports and/or imports, tourist length of stay, nights spent at tourist accommodation etc. (Malhotra and Yadav, 2017). When investigating the tourism demand literature, most of the researchers have concerned on tourist arrivals as a tourism demand measure. The persistence of tourism sector totally depends on the process of continuous arrivals of tourists.

1.2 Impact of Tourism Arrivals to Sri Lanka

After the end of civil war in 2009, Sri Lanka has become a safe tourist destination as well as a global tourist hub. Tourist arrivals have been increased significantly in the past few years. Tourism plays an important role in Sri Lankan economy as a main source of foreign exchange earnings in the economy (SLTDA, 2016). It also contributes to create new employment opportunities, to initiate business opportunities and for government revenue (Kodituwakku, *et al*., 2015).

According to Sri Lanka Tourism Development Authority (SLTDA), total tourist arrivals in Sri Lanka reached over 2 million in 2016 which is an increase of 14% when it compares with the previous year. Many tourists are coming to Sri Lanka from different countries in different time periods in the year. Mainly, the highest number of tourists is coming to Sri Lanka from Western Europe and South Asia region (Table 1.1).

Table 1.1: Distribution of tourist arrivals in Sri Lanka

Regions	Percentage Share		
	2014	2015	2016
North America	4.8	4.7	4.8
Western Europe	31.4	30.7	31.4
Eastern Europe	10.1	8.3	7.9
Middle East	5.8	5.6	5.2
South Asia	24.2	25.5	25
North East Asia	12.2	15.2	16.1
South East Asia	5.8	4.7	4
Australasia	4.3	4	4.1
Others	1.4	1.3	1.5
Total	100	100	100

Source: SLTDA (2016)

According to the Table 1.1, tourists from Western Europe region have become the major source of tourism in Sri Lanka in the years 2014-2016, accounting for about 31 percent and the number of tourist arrivals from the South Asia region have become the second source of tourism with an average of 24.9 percent of the total shares.

Furthermore, according to SLTDA (2016) Western Europe was the major source market, which recorded 16.5 percent growth in 2016 compared to the previous year. South Asia region was the second major source market, which recorded an increase of 11.8 percent of growth in 2016 compared to the year 2015 (Table 1.2).

Table 1.2: Percentage Changes of tourist arrivals by regions

Market Region	Change (%)		
	2014/2013	2015/2014	2016/2015
America-North	10.7	16.9	15.8
Asia-North east	80.8	47.3	20.2
Asia-South east	17.9	-5.9	7.8
Asia-South	13.4	24.1	11.8
Australasia	7.3	9.8	17
Europe-West	13.8	15.3	16.5
Europe-East	22.6	-3.7	8.6
Middle East	10.5	13.6	6.5
Others	44.8	7.6	1.4
Total	19.8	17.8	14

Source: SLTDA (2016)

1.3 Forecasting Tourist Arrivals.

Many researchers have taken efforts to forecast the total number of international tourist arrivals to a particular destination while some of the researchers have predicted tourist arrivals from several individual countries to a particular destination country. At present the growth rate and the patterns of international tourist arrivals may be affected by many global factors such as the global financial crisis and the ongoing European debt crisis. On the other hand the different origin countries in a same region under similar economic and social conditions may be affected by uncertain political and terrorist activities and environmental conditions (Preez and Witt, 2003). Thus, the patterns of tourist arrivals from different countries may be varied in a same region.

Therefore, modeling and forecasting of the tourist arrivals country wise is essential to reduce the imbalances between demand and supply and in decision making processes of tourism stakeholders in both public and private sectors.

In regard to Sri Lanka, considering the tourist arrivals from Asian countries and Western European countries is important. The highest number of tourists coming to Sri Lanka is from the Asian region. Sri Lanka can gain more benefits from the Asian tourist market due to its highest head count (Konarasinghe, 2016a). Since Western Europe countries have high purchasing power and per capital income, arrivals of Western Europe countries provide more benefits. (Konarasinghe, 2016b; Pasquali, 2016). Also, Kurukulasooriya and Lewala (2014) recommended to do further research on the arrivals of Western Europe and South Asia regions as those regions are major tourist generating sources. They further suggest to study the tourist arrivals country wise. Accordingly, Table 1.3 shows the percentage shares of the major top ten source markets.

Table 1.3: Tourist arrivals by country in 2016.

Market region	Total Arrivals	Percentage Share
India	356729	17.39
China	271577	13.24
United Kingdom	188159	9.17
Germany	133275	6.5
France	96440	4.7
Maldives	95167	4.64
Australia	74496	3.63
Russia	58176	2.84
USA	54254	2.65
Canada	44122	2.15
Total	1372395	66.92

Source: SLTDA (2016)

It can be observed from Table 1.3 that India was the largest tourist producer in Sri Lanka with a share of 17.39%, while the United Kingdom has become the third largest source of tourism in the country with a share of 9.17% in 2016. The countries such as Germany and France in Western Europe also shows a significant amount of contribution in total tourist arrivals.

1.4 Main Research Problem

International tourists are coming to Sri Lanka from all the regions of the world. Though the different origin countries in a same region under similar economic and social conditions may be affected by uncertain political and terrorist activities and environmental conditions (Preez and Witt, 2003). Thus, forecasting and modeling the tourist arrivals country wise provide more benefits to Sri Lanka and it facilitate the development programs of Government compared to forecasting region wise. Thus, this study focuses to find suitable models to forecast the arrivals from major tourist generating countries. At present there is no scientific method is used to predict future tourist arrivals in Sri Lanka.

Kurukulasuriya and Lewala (2014) mentioned that there exists a linear trend pattern with an obvious seasonal pattern in arrivals of tourists in Sri Lanka. Also, seasonality is a dominant feature of tourism sector considered by decision makers. It can be seen that vast number of researchers take into consideration the seasonal component. Accordingly, this study has conducted a comparative analysis of three time series techniques to capture the trend and seasonal patterns.

1.5 Objectives of the Research

In view of the above, the objectives of the study are to

- Identify the trend and seasonal patterns of tourist arrivals from United Kingdom and India.
- Develop the appropriate models to forecast monthly tourist arrivals from UK and India using different forecasting techniques and validate the models.
- Forecast the monthly tourist arrivals from two countries for six months ahead using the most appropriate model

1.6 Significance of the Study

At present the growth rate and the patterns of international tourist arrivals may be affected by many global factors such as the global financial crisis and the ongoing European debt crisis. On the other hand the different origin countries in a same region under similar economic and social conditions may be affected by uncertain political and terrorist activities and environmental conditions (Preez and Witt, 2003). Thus, the patterns of arrivals from different countries may be varied in a same region. Therefore, this study is with paramount importance as it takes into consideration the monthly tourist arrivals country wise. The inference derived here can be effectively used by the decision makers in the Tourist Industry for efficient management.

In this study, monthly data were used to forecast. Monthly data provide more detailed information by taking into consideration the seasonal and trend components compared to annual data. Also, with regard to tourism industry, short term forecasts help to decision makers for scheduling, staffing and planning tour operator brochures (Preez and Witt, 2003).

1.7 Outline of the Study

Chapter 1 explains the introduction which consists of the background of the study, main research problem, objectives and the significance of the research. Chapter 2 discusses literature review related forecasting and modeling tourist arrivals from particular destinations to particular origin country. In Chapter 3, materials and methodologies are explained. Some definitions related to time series and theoretical background of three forecasting techniques are explained. Chapter 4 brings out the results and discussions related to Box Jenkins methodology. It includes statistical analysis of the monthly tourist arrivals from United Kingdom and India for the period from November 2010 to August 2017. Chapter 5 brings out the development of smoothing techniques. Chapter 6 includes conclusions, recommendations and future work.

CHAPTER 2

LITERATURE REVIEW

Many researches have been carried out various studies on forecasting tourism with different solution techniques in recent years. This chapter covers a wide range of knowledge about past work related to forecasting techniques of tourism arrivals.

2.1 Related Studies in Sri Lanka

Gnanapragasm and Cooray (2016b) have forecasted the monthly tourist arrivals in Sri Lanka using Holt Winters smoothing method for the period from June 2009 to December 2015. In this analysis, grid search and auto search procedures were carried out to determine the smoothing constants. The data in this study were analyzed by using the software called STATISTICA. ADF statistic and the Kruskal-Wallis test was used to check the trend and seasonality in the corresponding series respectively. The best model was selected using the minimum MAPE value and the corresponding value is 11.36 percent.

Gnanapragasm and Cooray (2016a) have also attempted to forecast the tourist arrivals using Dynamic Transfer Function (DTF) model for the period of June 2009 to December 2015. The fitted model showed MAPE value of 8.63% in terms of 90% forecasting accuracy.

Kurukulasooriya and Lelwala (2014) have employed Classical Decomposition Approach to forecast the tourist arrivals in the post war period in Sri Lanka for the period from July 2009 to June 2013. The Mann Kendall test and the Kruskal Wallis test were applied to check the trend and seasonal patterns in the data series respectively. By considering MAPE criterion, multiplicative model was identified as the best model compared to the additive model. In their study, it was found that the additive model overestimate the forecast values.

Kurululasooriya (2008) has used SARIMA model with interventions to study the tourist demand for the Southern coastal area which is the highest number of tourist visited destinations in Sri Lanka. In their study monthly foreign guest nights was taken as the demand variable. The data series covers the period from January 1999 to July 2006. The terrorist attack at Colombo international airport on July 2001 and the tsunami disaster in December 2004 were two events occurred during the chosen time period which affects for tourism demand. However, he has attempted to use intervention analysis to quantify these non-random changes in the time series before modeling the time series data. The stationarity of the data series was determined by applying ADF statistic. The Ljung-Box Q statistic was used to test the adequacy of the model and normal probability quantiles were used to check the normality of residuals. In this study ARIMA $(1, 1, 1) \times (1, 0, 0)_{12}$ with two interventions was selected as the most suitable model. MAPE criteria and Theil's U statistic were used as measures of forecasting accuracy.

Ishara and Wijekoon (2017) have applied Auto Regressive Integrated Moving Average (ARIMA) and multiplicative decomposition approach to forecast the monthly tourist arrivals to Sri Lanka. Their analysis consists of two sets of data which are long term (from January, 2000 to February, 2016) and post war period (from January 2010 to February 2016). Accordingly, multiplicative decomposition models have higher forecasting accuracies for both long term and post war data analysis due to its minimum MAPE values. For the long term period ARIMA and multiplicative decomposition model showed MAPE values of 11 % and 8.25% respectively, and the corresponding values of that is for the post war period are 8 % and 6.6 %, respectively.

Konarasinghe (2017) has used two univariate techniques, namely decomposition additive and multiplicative models and SARIMA model to predict tourist arrivals from the Western Europe region for the period from January 2008 to December 2015. The results revealed that decomposition models didn't satisfy the model validation criteria while SARIMA satisfied the model validation criteria which also has a low relative measurement value.

Also Konarasinghe (2016b) has conducted a study to fit a suitable decomposition model to forecast the arrivals from Western Europe countries (UK, Germany, France, Netherlands and Italy) for the period from January 2008 to December 2014. According to the results of his study both additive and multiplicative model satisfied the assumptions of normality and independence. However, he concluded that additive model is the most suitable among two models to forecast arrivals from Western Europe countries under consideration of least errors. Since the data shows a wave-like pattern, he recommended the circular model to see whether the forecasting accuracy increases.

Moving Average, Exponential Smoothing and Holt Winters method were employed to identify a suitable short term forecasting technique to forecast the arrivals from Asian region to Sri Lanka based on the monthly data for the period from January 2008 to December 2014 (Konarasinghe, 2016a). Out of all models, double exponential smoothing model with a MAPE value of 1.5% was identified as the most suitable model which is statistically significant.

Dias, *et al*. (2016) have conducted a study to select the best model among SARIMA, Holt Winter's additive and multiplicative model and artificial neural networks in forecasting the total monthly tourist arrivals and arrivals from six countries (India, United Kingdom, Germany, Maldives, France and China) to Sri Lanka for the time period from January 2010 to August 2014. Holt Winters additive model showed the best results in forecasting tourist arrivals from Germany and China and Multiplicative model showed the best results for India, Maldives and United Kingdom while SARIMA has shown the best results for France. All models were selected by comparative analysis of the minimum errors of RMSE, MPE, MAPE and Theil's U statistic. The analysis of their study confirmed that a linear trend and seasonal fluctuations are significant features of international visitor arrivals to Sri Lanka during that period. In this study, neural network was identified as an alternative approach when the required data are available.

Peiris (2016) has found a suitable SARIMA model to predict total tourist arrivals to Sri Lanka for the period from January 1995 to July 2016. SARIMA $(1, 0, 16)\times(36, 0, 24)_{12}$ was chosen as the best model by considering the lowest value of AIC and

SBC criteria while forecasting accuracy was evaluated by using MAPE, RMSE and MAE. In this study HEGY test was employed to identify the deterministic seasonality in the data series.

2.2 Related Studies in Other Countries

Mamula (2015) has examined the different forecasting techniques, i.e. Holt Winters triple exponential smoothing, the seasonal naïve model, the seasonal ARIMA model and the multiple regression model in forecasting the international tourism demand in Croatia. The quarterly German tourist arrivals in the Republic of Croatia for the period from the first quarter, 2003 to last quarter, 2012 were employed in the study. All the models were statistically significant and forecast values were compared based on the MAPE values of both in sample and the out of sample data. It was found that multiple regression model was the most suitable model. Based on the selected model, forecasts were provided for a one year period ahead.

Hassani, *et al*. (2015) have investigated a number of parametric and non-parametric forecasting methods to forecast tourism demand in European countries. Exponential Smoothing, ARIMA, Neural Networks, Trigonometric Box-Cox ARMA Trend Seasonal, Fractionalized ARIMA and both Singular Spectrum Analysis Algorithems, i.e. recurrent and vector SSA were employed to examine short, medium and long run forecasts in terms of total tourist arrivals in ten European countries: Germany, Greece, Spain, Netherlands, Cyprus, Austria, Sweden, Portugal and United Kingdom. The comprehensive analysis of their study highlighted that there is no specific model which provides best forecasts regarding to the country, forecasting horizon and direction of change criteria for forecasting tourism demand.

Cuhadar (2014) has used 168 observations of monthly international tourist arrivals to Istanbul for the period from January 2010 to December 2013 to model and forecast the tourism demand. Five time series models namely, simple seasonal, multiplicative and additive, Holt Winter's exponential smoothing, multiplicative and additive seasonal ARIMA (SARIMA) model were considered in the study. He used MAPE measure to evaluate the forecasting performance and further states that "MSE and MAPE may give misleading measures due to the cancellation of positive and negative errors''. SARIMA $(2, 0, 0) \times (1, 1, 0)_{12}$ was selected as the most suitable model which recorded the lowest MAPE of 3.42%. The validity of the model was tested by using AIC, BIC criteria and Ljung-Box test.

Chaitip, *et al*. (2008) investigated the forecasting performance of SARIMA, ARIMA, Holt Winter- additive, Holt Winters multiplicative, Holt Winter's-no seasonal, neutral network, VAR, GMM methods in forecasting international tourism arrivals to Thailand from the period from 2006 to 2010. The results of the study revealed that the best forecasting method based on first concept is SARIMA $(0, 1, 1) \times (0, 1, 4)_{12}$ and the best forecasting method based on second concept is VAR model.

Papic-Blagojevic, *et al*. (2016) examined the seasonality of monthly tourist presence in Serbia using different exponential smoothing methods. Their study was conducted on monthly data of number of overnight stays in three different cities in Serbia such as Belgrade, Novi Sad and Nis for the time period from January, 2010 to December, 2013. In their study, the performance of Simple Seasonal, Holt Winters Multiplicative and Holt Winters Additive smoothing methods were compared regarding to these cities. Based on the BIC and RMSE criterion, Holt Winters multiplicative model showed the best results for series in Novi Sad and Nis while simple seasonal model outperformed other two models for series in Belgrade.

Lwesya and Kibambila (2017) also attempted to compare the performance of SARIMA and Holt Winters multiplicative and additive smoothing methods in forecasting the tourist arrivals in Tanzania. The time span covered the period from January, 2000 to December, 2009. Among Holt Winters multiplicative and additive model, Holt Winters multiplicative smoothing model recorded the minimum values of the Sum of Error (SSE) and Mean Squared Error (MSE). On the other hand, ARIMA $(4, 1, 4) \times (3, 1, 4)_{12}$ model showed better results compared to several SARIMA models based on the MAPE, RMSE, BIC and MAD criterion. At last, among Holt Winters multiplicative model and ARIMA $(4, 1, 4) \times (3, 1, 4)_{12}$ model, Holt Winters multiplicative model with α (0.01), β (0.11) and γ (0.11) was selected as the best model in forecasting tourist arrivals in Tanzania using the criteria of MAPE, RMSE, BIC and MAD. According to their findings, they suggested that the seasonal variations of the tourist arrivals data are changing in proportional to the level of the series in Tanzania.

Petrevska (2017) also applied Box Jenkins methodology to forecast the tourism demand of F. Y. R. Macedonia. The data covers the period from 1956 to 2013. In this study, the L.B statistic was used to check the stationarity. Although, Petrovska stated that L.B statistic takes a low power as the significant coefficient can be neutralized by the insignificant ones. Thus, he additionally tested the L.B Statistic by employing the ADF and Phillips Perron test. After that, ARIMA (1, 1, 1) model was identified as the most appropriate model. Accordingly, he found that the increasing trend in international tourism will continue by 2018. Further, he states that the ARIMA model regarding to this study is not a highly accurate one as several structural breaks have occurred during that period.

Lognathan and Ibrahim (2010) used Box Jenkins ARIMA model to forecast quarterly tourism demand in Malaysia for the period from 1995:Q1 to 2008:Q2. They found that quarterly international tourist arrivals to Malaysia don't depend on seasonal effects and found that the time series has a deterministic pattern of long term upward trend. Augmented Dickey fuller and Phillip-Perron test were applied to check the stationarity. ARIMA (1, 0, 1) was selected as the best model to forecast.

Borhan and Arsad (2014) have forecasted monthly tourist arrivals to Malaysia from three selected countries, namely Japan, US and South Korea by using Box Jenkins seasonal ARIMA model. The study covers the period from January 1999 to December 2012. The best model was selected using BIC criterion. By using the selected model, forecasts were provided for six years ahead. This study also presented the percentage changes in the forecast values of the number of tourist arrivals for three countries. Accordingly, they found that growth of tourist arrivals from South Korea and US will continue to rise in the coming months while arrivals from Japan will continue to decrease in coming months.

The three forecasting techniques, namely ARIMA, exponential smoothing and ANN were employed to forecast monthly tourist arrivals from six countries (UK, USA, Singapore, Taiwan, Japan and Korea) to Hong Kong (Cho, 2002). The ANN model outperformed other models in forecasting arrivals from all countries except the UK. He pointed out that ANN model is appropriate when there is a less obvious pattern in the series unless ARIMA and exponential smoothing models are sufficiently adequate.

Cho (2002) has used six time series models(Naïve 1, Naïve 11, Liner trend, Holt Winters method, Sine wave and ARIMA) to predict tourism demand in ten countries (Japan, South Korea, Taiwan, Hong Kong, Philippines, Singapore, Indonesia, Thailand, Australia and New Zealand). By considering the minimum MAPE value, ARIMA model was selected to forecast for nine of ten countries. The study of Kumari (2015) revealed that SARIMA and Holt Winters multiplicative model outperform the Grey model in forecasting the tourist arrivals in India for the period from January 2000 to October 2015 using MAPE criterion. In this study, turning point analysis and Theils U statistic were used to evaluate the performance accuracy of SARIMA and Holt Winters models, while Posterior variance ratio test was used to check the accuracy of grey model. Chu (2009) has examined three univariate ARMA based model to forecast the international tourist arrivals to nine major tourist destinations in Asian- Pasific region, including Hong-Kong, Korea, Japan, Taiwan, Singapore, the Philippines, Thailand, New Zealand and Australia. Both monthly and quarterly time series were considered in this study.

Researchers have recently attempted to improve the forecasting accuracy of the SARIMA/ ARIMA model by using different techniques (Goh and Law, 2002; Lim and McAleer, 2002). Goh and Law (2002) used the SARIMA (Seasonal ARIMA) and MARIMA (Multiplicative seasonal ARIMA with intervention) models to study the monthly tourist arrivals from ten highest tourist generating countries to Hong Kong. The time span of the study was affected by many external events such as the Asian financial crisis and bird flu epidemic in Hong Kong. Accordingly, they concluded that MARIMA outperforms other forecasting models when there exist any significant interventions in the data series. However, they pointed out that when there isn't any obvious intervention in the series, SARIMA model provides more accurate forecasts. The forecasting performance was evaluated by using MAPE criteria.

Lim and McAleer (2002) employed Box Jenkins methodology to estimate the tourist arrivals to Australia from three Asian countries: Hong Kong, Malaysia and Singapore for the period from 1st quarter, 1975 to 4th quarter, 1989. During the period, Australia experienced oil price crisis, air pilot strike and the Bicentennial celebration in 1988 of European settlement in Australia. They analyzed these one-off events with intervention analysis. MAPE and RMSE criteria were used to evaluate the forecasting performance. Their results revealed that the ARIMA model outperforms seasonal ARIMA models for Malaysia and Hong Kong while SARIMA model provides better forecasts for Singapore.

To improve the forecasting accuracy, some of the scholars have suggested to combine the forecasts generated from different models. Wong, *et al*. (2007) examined the forecasting accuracy of combined forecasts by utilizing the tourist arrival data in Hong Kong from top ten tourist generating countries/regions. The forecasts were obtained by applying different forecasting methods, namely Autoregressive Distributed Lag Model (ADLM), ARIMA model, Vector Autoregressive (VAR) model and Error Correction Model (ECM). They have found from their study that combined forecasts don't always outperform best individual model forecasts. Even though, they concluded that combined forecasts can perform better than the worst single model forecasts to reduce the forecast failures.

2.3 Selection of Forecasting Models.

Cho (2008) pointed out that selection of method should mainly be depended on the accuracy of the forecasts generated, ease of use, the cost of the process and the running speed. Obviously, the cost associated with data collection and model estimation process of time series models is low as it only requires the historical data of the corresponding variable (Goh and Law, 2002; Kodithuwakku, *et al*., 2015). Thus, time series techniques have gained more popularity in past researches and studies.

Most of the studies in previous literature based on time series techniques were comparative analysis. However, Dias, *et al*. (2016) pointed out that “the comparative forecasting results are not meant to be conclusive in term of model choice, but rather help to illustrate the potential of origin country based forecasting”. When investigating past literature regarding the selection of a better model for a data series, contradictory findings can be seen between the conclusions. Different studies show that different models outperform other competing models in forecasting tourist arrivals.

Song and Li (2008) have provided a most comprehensive review of forecasting international tourism demand. This review brings out the published studies on tourism forecasting and modeling since 2000. The study revealed that there is no single model that outperforms other forecasting models in all situations. Forecasting performance of the models varies with the data frequencies used in the model, the destination, the origin country and the length of the forecasting horizons concerned. Thus, with regard to the empirical literature on tourism demand forecasting, it can’t be found a single model which has a superior forecasting performance ability in all situations. Thus, this study focuses to choose the most appropriate model which has a lowest forecasting error in forecasting tourist arrivals in Sri Lanka for the corresponding time span.

2.4 Evaluation of Forecasting Accuracy

Regarding to previous researches and studies, various measures of forecasting accuracy have been used. MAPE, MAE, RMSE, MPE and MSE are some of the widely used measures in past tourism literature. On the other hand, Chuhadar (2014) stated that “MSE and MPE may give misleading measures due to the cancellation of positive and negative errors’’. Most of the researchers have used the MAPE criterion (Cuhadar, 2014; Gnanapragasm and Cooray, 2016b; Kurukulasooriya and Lelwala, 2014a). Lewis (1982) pointed out that since MAPE is a relative measure and it is most useful in comparing the accuracy of forecasts.

2.5 Summary of Chapter 2

The forecasting performance of the models varies with the data frequencies used in the model estimation, the destination, the origin country and the length of the forecasting horizons. It is difficult to find the best model which outperforms all other models in all situations. Thus, it can be recommended to select the most appropriate model which has a lowest forecasting error. The information gathered from this literature survey is useful to carry out this study.

CHAPTER 3

MATERIALS AND METHODS

In this chapter, data source of which data are acquired, the fundamental concepts of time series analysis, theoretical background of three time series techniques: i.e. Box Jenkins SARIMA, Holt Winter's additive and multiplicative model are discussed.

3.1 Source of the Data

The historical data on monthly tourist arrivals to Sri Lanka were acquired from annual reports of Sri Lanka Tourism Development Authority (SLTDA). The in-sample period for developing the models was from November, 2010 to August, 2017. The out of sample period for validating models was from September, 2017 to February, 2018. Raw data are shown in Appendix A.

3.2 Basic Definitions in Time Series

3.2.1 Time Series

A time series is a sequence of data points taken at successive equally spaced points in time. It is a sequence of discrete-time data.

3.2.2 Stationary Process

A stochastic process which is considered as families of random variables which are functions of time $\{Y_t, t \geq 1\}$ is said to be a stationary process if for arbitrary points $t_1, t_2 \ldots t_n$, , the joint distribution of the random variables $\{Y_{t1}, Y_{t2}, \ldots \ldots .. Y_m\}$ and $\{Y_{t1+h}, Y_{t2+h}, \ldots \ldots Y_{tn+h}\}$ are the same (Peiris, 2017).

3.2.3 Weekly Stationary

In time series analysis, we generally dilate stationary series into weekly stationary series. A stochastic process $\{Y_t, t \geq 1\}$ with finite second order moments is called weekly (covariance) stationary, if $E(Y_t) = constant,\ V(Y_t) = \sigma^2$ and $Cov(Y_t, Y_s)$ depends on the difference $t - s$ for all t and s.

3.2.4 Purely Random Process (White Noise)

A time series $\{e_t\}$ is called a purely random process if it consists of a sequence of random variable which are mutually independent and identically distributed, in particular if $\{e_t\}$ is normally distributed with mean zero and variance σ^2 , the series is called Gaussian white noise.

That is $Cov(e_t e_s) = \begin{cases} \sigma^2 & if \ t = s \\ 0 & if \ t \neq s \end{cases}$

3.3 Transformation of Non Stationary Series to Stationary Series

Non stationary time series are transformed into stationary time series by taking the first differences of the non stationary time series. The first difference of a time series can be written as

$$x_t = y_t - y_{t-1} = (1 - B)y_t \tag{3.1}$$

If the first difference series is also non stationary, it is transformed into stationary by taking the first difference of first difference series. First difference series of the first differenced series can be written as

$$\{x_t - x_{t-1}\} = \{y_t - y_{t-1} - y_{t-1} + y_{t-2}\} \tag{3.2}$$

$$= (1 - 2B + B^2)y_t$$

The second difference of the original series can be taken to make original series stationary and it can be written as

$$x_t = y_t - y_{t-2} = (1 - B^2)y_t \tag{3.3}$$

where $B^l y_t = y_{t-l}$

Also, when time series has a seasonal pattern, it is better to get long term difference as well, in addition to short term differences to make the series stationary. When the seasonal length is 12, the stationary is achieved by taking one short term difference and one long term difference. That is,

$$Z_t \quad = (y_t - y_{t-1}) - (y_{t-12} - y_{t-13}) \tag{3.4}$$

$$= (1 - B - B^{12} + B^{13})y_t \text{ , where } B^l y_t = y_{t-l}$$

3.4 Unit Root Test (Augmented Dickey Fuller test) for Stationary

The time series properties of stationary and non stationary are checked by applying Augmented Dickey Fuller Test (Dickey Fuller, 1979). The hypotheses tested under unit root test are:

Ho: The series is not stationary or series has a unit root ($\phi_1 \geq 1$).
H1: The series is stationary or series hasn't a unit root ($\phi_1 < 1$).

where ϕ_1 is the parameter of AR(1).

3.5 Kruskal- Wallis test

To confirm the seasonality, Kruskal- Wallis test can be applied. The corresponding hypotheses are:

H_0: series has no seasonality Vs H_1: series has seasonality

Under H_0, the test statistic of Kruskal-Wallis test is defined as follows.

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{n_i} \frac{R_i^2}{n_i} - 3(N+1) \chi_{L-1}^2 \tag{3.5}$$

When N is the number of total rankings, n_i is the number of the ranking in a particular season, R_j is the sum of the ranking in a particular season and L is the length of the season (Gnanapragasam & Cooray, 2016b).

3.6 Box Jenkins Models

The Box Jenkins methodology is used to model a univariate time series as well as multivariate time series. Box Jenkins model assumes that the time series is weekly stationary. If the time series is not stationary, it is recommended to difference the non stationary series one or more times to achieve stationarity.

This methodology involves four major steps of model identification, model estimation, diagnostic checking and validation and forecasting of the model. In the first step of model identification, the time series plot is examined to identify the nature of the data series such as trend and seasonality. To make data series stationary, differencing can be applied. After the stationary is attained, by simultaneous inspection of SACF and SPACF, the order of AR and MA is determined. By simultaneous inspection of the SACF and SPACF, parsimonious models are postulated and the best fitted model is selected using various diagnostic measures (Peiris, 2017).

3.7 ARIMA Models

Autoregressive process of order p – AR (p)

An autoregressive model is expressed by sum of prior values and present disturbance. AR model of order p can be represented by

$$Y_t = \mu + e_t + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots \ldots + \phi_p Y_{t-p} \tag{3.6}$$

where e_t is purely random process with mean zero and constant variance.

Moving Average process of order q – MA (q)

The moving average model is expressed by sum of past disturbances and present disturbance. MA (q) model is represented by

$$Y_t = \mu + e_t - \theta_{t-1} e_{t-1} - \theta_{t-2} e_{t-2} - \cdots \ldots - \theta_q e_{t-q} \tag{3.7}$$

ARMA process of order p and q – ARMA (p, q)

The Autoregressive Moving Average model ARMA (p, q) is a combination of AR and MA terms which were discussed previously. It can be represented as

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + e_t - \sum_{i=1}^{q} \theta_i e_{t-i} \qquad (3.8)$$

Autoregressive Integrated Moving Average Models – ARIMA (p, d, q)

In this case y_t in (3.8) is replaced by $z_t = y_t - y_{t-d}$, where d is the number of differencing for stationary.

Seasonal Autoregressive Integrated Moving Average Model - SARIMA

Seasonal Autoregressive Integrated Moving Average (SARIMA) model is an extension of the ordinary ARIMA model which captures the seasonality in the data series. Seasonal ARIMA model incorporates both non seasonal and seasonal factors in a multiplicative model and corresponding model is formed as $ARIMA\,(p, d, q) \times (P, D, Q)_S$. P, D, Q are the corresponding parameters in seasonal part and S is the seasonal length.

The multiplicative seasonal ARIMA model can be represented as

$$\Phi(B^s)\phi(B)\Delta_s^D \Delta^d\, y_t = \Theta(B^s)\theta(B)e_t \qquad (3.9)$$

where y_t is present value of the series and e_t is a white noise process .

Non seasonal AR and MA polynomials can be defined respectively as follows. $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$

$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$

Seasonal AR and MA polynomials can be defined respectively as follows.

$\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \cdots - \Phi_p B^{ps}$

$\Theta(B^s) = 1 + \Theta_1 B + \Theta_2 B^2 + \cdots + \Theta_q B^{qs}$

3.8 Auto Correlation Function and Partial Auto Correlation Function.

The auto correlation function and partial auto correlation functions are useful tools in determining the order of non seasonal and seasonal terms. The auto correlation function measures the degree of correlation between neighboring observations in time series. The auto correlation coefficient is estimated from sample observations using the following formula.

$$\rho_k = \frac{\sum_{r=2}^{n} (Y_t - \mu_y)(Y_{t+k} - \mu_y)}{\sum_{r=1}^{n} (Y_t - \mu_t)^2} \tag{3.10}$$

The auto correlation at lag k is defined as

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad , k = 0, \pm 1, \pm 2, \dots \dots$$

where $\gamma_k =$ auto correlation at lag k.

An estimated PACF is also a graphical representation of statistical relationship between sets of ordered pairs drawn from a single time series. The partial autocorrelation coefficient is estimated from sample observations using following formula (Annorzie, et al., 2018).

$$\hat{\phi}_{kk} = \frac{r_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_j} \quad , \quad (k = 2,3, \dots) \tag{3.11}$$

where $\hat{\phi}_{ij} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk} \, \hat{\phi}_{k-1,k-j} (k = 3,4, \dots,; j = 1,2, \dots., k-1)$

and $\widehat{\phi}_{11} = r_1$

The primary tools for developing the ARIMA class of time series models are auto correlation and partial auto correlation plots. The sample ACF and PACF are compared with the theoretical behavior of ACF and PACF in the identification stage of time series modeling. The Table 3.1 indicates some guidelines to identify the basic ARMA models.

Table 3.1: Identification of AR and MA terms using ACF and PACF

Process	ACF	PACF
AR(p)	Tails off towards zero (exponential decay or damped sine wave)	Cuts off to zero after lag p
MA(q)	Cuts off to zero after lag q	Tails off towards zero (exponential decay or damped sine wave)
ARMA(p, q)	Tails off towards zero (exponential decay or damped sine wave)	Tails off towards zero (exponential decay or damped sine wave)

Source: Spyros, *et al*. (1998)

3.9 Exponential Smoothing

Exponential Smoothing is a time series forecasting technique applying for univariate data (Dias, *et al*., 2016; Chuhadar, 2014; Cho, 2002; Papic-Blagojevic, *et al*., 2016; Lwesya and Kibambila, 2017). There are three major exponential smoothing time series methods, namely single exponential smoothing, double exponential smoothing and triple exponential smoothing method. Also, an equivalent ARIMA model can be defined for any General Exponential Smoothing model (Mckanzied, 1984). In this study, Holt Winters multiplicative and additive smoothing models are employed.

3.9.1 Single Exponential smoothing

Simple exponential smoothing is applied when the data series hasn't any trend or seasonal pattern. The equation for single exponential smoothing model can be represented as follows.

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

The k-step prediction is $\hat{Y}_t(k) = L_t$

ARIMA model equivalency to single exponential smoothing is the ARIMA (0, 1, 1) model. ARIMA (0, 1, 1) model can be represented as follows (SAS Institute, n. d.).

$$(1-B)y_t = (1-\theta B)e_t \text{ , where } \theta = 1-\alpha$$

3.9.2 Double Exponential Smoothing method (DES)

Double exponential smoothing technique is appropriate when the observations have a trend and no seasonality. This model includes two smoothing parameters. The formulas for DEM method are

$$a_t = \alpha Y_t + (1-\alpha)(a_{t-1} + b_{t-1}) \tag{3.13}$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1} \tag{3.14}$$

$$\hat{y}_t(m) = a_t + m b_t \tag{3.15}$$

where, a_t is smoothed level of the series, which computed after y_t is observed, b_t is smoothed trend at the end of period t, $\hat{y}_t(m)$ is the forecast value in m period and m is the number of periods in the forecast lead-time (Ersen, et al., 2017).

ARIMA model equivalency to linear exponential smoothing is the ARIMA (0, 2, 2) model. ARIMA (0, 2, 2) model can be represented as follows (SAS Institute, n. d.).

$$(1-B)^2 y_t = (1-\theta_1 B - \theta_2 B^2)e_t \text{ , where } \theta_1 = 2-\alpha-\alpha\beta \text{ and } \theta_2 = \alpha - 1$$

3.9.3 Holt Winters (HW) smoothing method

The Holt Winters exponential smoothing technique takes into account both seasonal changes and trends. Thus, when the time series data exhibits seasonality, Holt Winters exponential smoothing method can be recommended. It incorporates three smoothing parameters for the level, for the trend and for the seasonality. HW model is an extension of single exponential smoothing model.

This method depends on the historical data and gives more weight to the recent values. There are two types of Holt Winters models, depending on the type of seasonality namely additive and multiplicative model. The additive model is preferred when the seasonal variations in data series exhibits a constant pattern, while the multiplicative model is used when the seasonal variations are changing proportional to the level of the series (Celebi, *et al*., 2017).

The Holt Winters additive model can be presented as follows (Celebi, *et al*., 2017).

$$a_t = \alpha\left(Y_t - s_{t-p}\right) + (1-\alpha)(a_{t-1} + b_{t-1}) \tag{3.16}$$

$$b_t = \beta(a_t - a_{t-1}) + (1-\beta)b_{t-1} \tag{3.17}$$

$$s_t = \gamma(Y_t - a_t) + (1-\gamma)s_{t-p} \tag{3.18}$$

a_t: the smoothed level at time t

b_t: the change in the trend at time t

s_t: the seasonal smoothing parameter at time t

P: the number of seasons per year

α, β, γ : smoothing parameters

α: the weighting factor for the level

β: the weighting factor for the trend

γ: the weighting factor for the seasonality

The forecast for time period $T + \tau$:

$$\hat{y}_{T+\tau} = a_T + \tau b_T + s_{T+\tau-s} \tag{3.19}$$

Where a_T is the smoothed estimate of the level at time T, b_T is the smoothed estimate of the change in the trend at time T, s_T is the smoothed estimate of the appropriate seasonal component at time T. τ is the τ-step-ahead forecast (Celebi, *et al*., 2017).

The ARIMA model equivalency to Holt Winter's Additive model is the ARIMA (0, 1, p+1) × (0, 1, 0) $_p$ model. ARIMA (0, 1, p+1) × (0, 1, 0)$_p$ model can be written as follows (SAS Institute, n. d.).

$$(1 - B)(1 - B^p)y_t = \left[1 - \sum_{i=1}^{p+1} \theta_i B^i\right]\varepsilon_t \tag{3.20}$$

$$\text{where } \theta_j = \begin{cases} 1 - \alpha - \alpha\gamma & , \ j = 1 \\ \alpha\gamma & , 2 \le j \le p - 1 \\ 1 - \alpha\gamma - \delta(1 - \alpha), & j = p \\ (1 - \alpha)(\delta - 1) & , j = p + 1 \end{cases}$$

Holt Winters multiplicative model is given in following equations (Celebi, *et al*., 2017)

$$a_t = \alpha \frac{Y_t}{s_{t-p}} + (1 - \alpha)(a_{t-1} + b_{t-1}) \tag{3.21}$$

$$b_t = \beta(a_t - a_{t-1}) + (1 - \beta)b_{t-1} \tag{3.22}$$

$$s_t = \gamma \frac{Y_t}{a_t} + (1 - \gamma)s_{t-p} \tag{3.23}$$

Forecast for time period $T + \tau$ is given by

$$\hat{y}_{T+\tau} = (a_T + \tau b_T)s_{T+\tau-s} \tag{3.24}$$

The smoothing parameter α can smooth the level equation. When α is small, it indicates that the series is stable during the corresponding time period. High values of α indicate large fluctuations of the data. For small β, more weight is given to previous slope estimate and large β gives more weight to the most recent estimate of the slope (Cho, 2002). Likewise, small γ would give more weight to previous estimates of the seasonal factor and large γ gives more weight to the most recent estimates of seasonal factor. Each of the smoothing parameter is updated by its own exponential smoothing equation. To find the optimum values for $\alpha,\ \beta$ and γ different combinations of these values are checked using trial and error method. The Holt Winters multiplicative model doesn't have an ARIMA equivalent (SAS Institute, n. d.).

3.10 Determination of Values of Smoothing Constants.

Exponential smoothing assigns greater weight to more recent observations and takes into account all previous data. Weight is given by the exponential smoothing constants and the forecast values differs with the value of smoothing constants. Thus, the error of forecasts depends on smoothing constants. Thus, it is important to choose the optimal values of corresponding constants to minimize the forecasting errors.

Many researches have used numerous methods to select the appropriate values of smoothing constant. There is no exact method to determine the optimal values of smoothing constants. Karmakar, *et al*. (2017) conducted nine trials to determine the optimum smoothing constant in single exponential smoothing with varying the smoothing constant from 0.1 to 0.9. In the double exponential smoothing method, nine trials were used varying the smoothing constants from 0.1 to 0.3. In Holt Winters method, 27 trials were performed by varying parameters from 0.1 to 0.3. MAPE, MAD and MSD were used to compare the forecasting error of different combinations of smoothing constants.

Ersen, *et al*. (2017) have used Minitab 15 programme to find optimum value of α in single exponential smoothing method. For double exponential smoothing and Holt Winters method, smoothing parameters were determined using the lowest value of MAPE. Paul (2011) used trial and error method to choose the optimal value of smoothing constants. He considered around 20 trials by varying the values to minimize the Mean Absolute Deviation (MAD) and Mean Square Error (MSE).

In this study, by using the trial and error method, 27 trials were performed in Holt Winters method by varying smoothing parameters from 0.1 to 0.3. MAPE was used to compare the forecasting error of different combinations of smoothing constants.

3.11 Diagnostic of Error Terms

The residuals of the fitted model should be distributed normally and independently with constant variance. The following tests are performed to check the residuals of the fitted model.

3.11.1. Jarque & Bera (JB) Test for Normality

JB test is applied to test the null hypothesis of error terms are not significantly deviated from normal distribution. That is

H_0: Error terms are not significantly deviated from normal distribution

H_1: Error terms are significantly deviated from normal distribution

The corresponding test statistic of JB test is

$$JB = \frac{Skewness^2}{6/n} + \frac{(kurtosis-3)^2}{24/n} \sim \chi_2^2 \quad (3.25)$$

where n is sample size.

3.11.2 Ljung and Box (LB) Q statistic

The LB Q statistic is used to check the hypothesis that error terms are independently distributed and the corresponding hypotheses tested can be defined as

H_0: $\rho_1 = \rho_2 = \ldots \rho_m = 0$

H_1: There exist at least m such that $\rho_m \neq 0$

under H_0, $\quad Q_m \sim \chi_{m-1}^2$

The corresponding test statistic can be represented as follows (Peiris, 2017).

$$Q_m = n(n+2)\sum_{i=1}^{m} \frac{{r_i}^2}{(n-i)} \quad (3.26)$$

3.11.3 Serial Correlation

Breusch-Godfrey (BG) test is used to check the higher order serial correlation that involves higher order autocorrelation estimators. It is more useful than DW test mainly due to the fact that it allows for higher order autoregressive processes or higher order moving average processes.

Suppose the error terms are AR (p) for p > 1 i.e

$$\varepsilon_t = +\rho_1\varepsilon_1 + \cdots + \rho_p\varepsilon_{t-p} + \nu_t$$

and $\nu_t \sim i.i.d(0, \sigma^2)$

The hypotheses tested here can be defined as

$$H_0: \rho_1 = \rho_2 = \ldots \rho_m = 0$$

$$H_1: \text{There exist at least } m \text{ such that } \rho_m \neq 0$$

Q statistic of squared residuals is defined as

$$Q_m = n(n+1)\sum_{k=1}^{p}\frac{\rho_k^2}{n-k} \sim \chi^2_\alpha, m-p \qquad (3.27)$$

(Makatjane & Moroke, 2016)

3.11.4 Heteroscedasticity

The ARCH LM test is used to test for Heteroscedasticity in residuals. The test statistic of ARCH LM test which can be employed for higher order $ARCH_p$ effects which is presented as:

$$Var(\varepsilon_t) = \gamma_0 + \gamma_1 \varepsilon^2{}_{t-1} + \cdots + \gamma_{p.}$$

The Lagrange Multiplier (LM) for Heteroscedasticity can be presented as

$$\text{LM statistic} = (n-p)R^2 \sim \chi^2_p \qquad (3.28)$$

where n is the sample size, p is the number of parameters and R^2 is the adjusted R^2. Thus, the hypothesis tested here can be defined as

$$H_0: Var(\varepsilon_t) = \sigma_t^2$$

$$H_\alpha: Var(\varepsilon_t) \neq \sigma_t^2$$

When the null hypothesis isn't rejected, it can be concluded that the error term is constant over time (Makatjane and Moroke, 2016).

3.11.5 Information Criteria

The final model can be selected using AIC and BIC criterion. The model with the lowest information criterion value is considered as the best model. These measures are calculated using the following equations.

Akaike Information Criterion (AIC)

$$AIC = -2log(likelihood) + 2k, k = \text{no of parameters} \qquad (3.29)$$

Schwarz Bayesian Information Criterion (BIC)

$$BIC = -\frac{2}{T}log(likelihood) + log(n) * k \qquad (3.30)$$

3.11.6 Percentage Error

When y_t is the value of actual observation for time t and $\hat{y}_t$ is forecast for the same period, the percentage error can be defined as follows.

$$PE_t = \frac{(y_t - \hat{y}_t)}{y_t} \times 100 \qquad (3.31)$$

3.11.7 Mean Absolute Percentage Error (MAPE)

The mean absolute percentage error (MAPE) is a measure of accuracy for selecting the most appropriate fitted model and it is defined by the following formula.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100 \qquad (3.32)$$

where y_t and $\hat{y}_t$ are the observed and forecasted value at the time t respectively. Criteria developed by Lewis (1982) have been used in many research studies to measure the accuracy of the models. The table 3.2 shows some guidelines of the measure of MAPE.

Table 3.2: Details of measure of MAPE

	Evaluation
$MAPE \quad \leq 10\%$	High accuracy forecasting
$10\% < MAPE \leq 20\%$	Good forecasting
$20\% < MAPE \leq 50\%$	Reasonable forecasting
$MAPE > 50\%$	Inaccurate forecasting

Source: Lewis (1982)

CHAPTER 4

DEVELOPMENT of ARMA MODEL

4.1 Temporal Variability of Tourist Arrivals from UK and India

The preliminary understanding about the nature of data series can be identified from the graphical display. The time series plots of monthly tourist arrivals from the United Kingdom and India for the period from November 2010 to August 2017 are shown in Figure 4.1 and Figure 4.2 respectively.

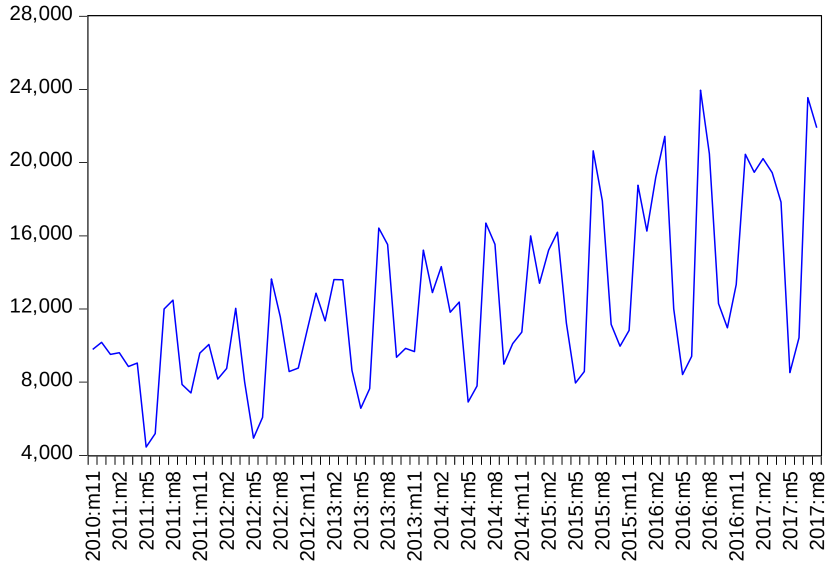


Figure 4.1: Time series plot of monthly tourist arrivals from United Kingdom (from November 2010 to August 2017)

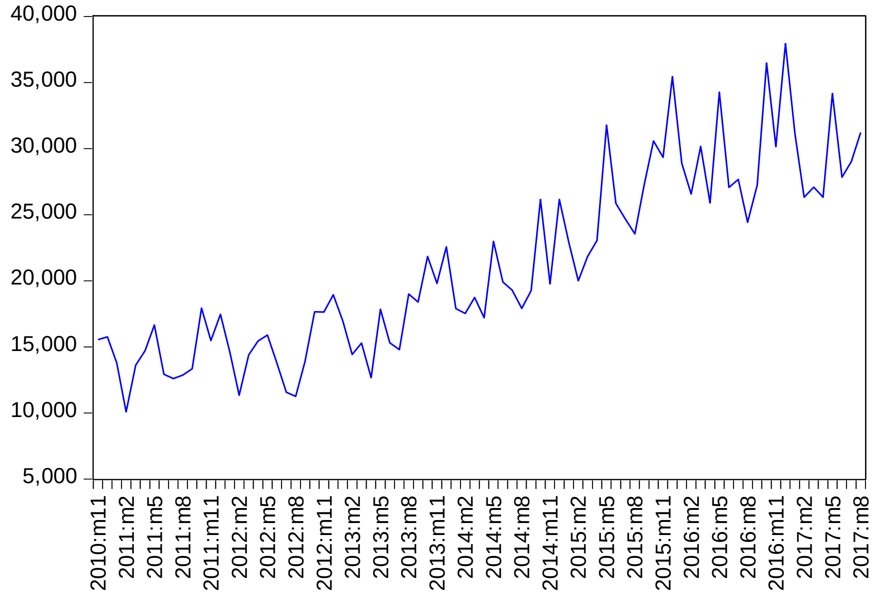


Figure 4.2: Time series plot of monthly tourist arrivals from India (from November 2010 to August 2017)

The time series plot of monthly tourist arrivals from the United Kingdom in Figure 4.1 shows an upward trend with a consistent pattern of monthly changes indicating the existence of seasonal fluctuations. It depicts that the original series is non stationary with seasonality with seasonal lengths of 12.

A different scenario can be recognized from the arrivals from India (Figure 4.2). Seasonality is less obvious for the arrivals from India. It seems that there is a peak in July and trough in May for United Kingdom. With regard to India, highest number of tourists were recorded in December and lowest number of tourists were recorded in August. Moreover, the time series plots of two countries in Figure 4.1 and Figure 4.2 seem that the seasonal pattern is not constant over the time and there are shifts in mean in different time periods. Thus, data series follows a stochastic form more than a deterministic one. Seasonality is considered either as a deterministic component or a stochastic component in tourism demand forecasting. When the seasonality is considered as stochastic, seasonal differencing is required to account for seasonal unit roots in the time series. When the seasonality is considered as deterministic, it is needed to introduce seasonal dummies into the time series models to account for the seasonal variations (Song and Li, 2008).

4.2 Summary Statistics of Tourist Arrivals from UK and India

The summary statistics of the tourist arrivals from the UK and India are shown in Table 4.1.

Table 4.1: Summary statistics of tourist arrivals

Statistic	UK	India
Mean	12356	21179
Median	11197	19261
Maximum	23948	37945
Minimum	4452	10071
Standard Deviation	4606.106	6902.572
Skewness	0.7010	0.4886
Kurtosis	2.6918	2.2548
Coefficient of Variance (%)	37.28	32.59

As indicated in the study of Hassani, *et al*. (2015), when the data series are skewed it is more applicable to use the median and interquartile range whilst for symmetric distribution it is appropriate to consider the mean and standard deviation to explain the central tendency and variation. Also, it was recommended that it is better to consider the coefficient of variance criterion to compare the variability between countries when the most of the series are skewed.

According to the summary statistics in Table 4.1, monthly tourist arrivals from United Kingdom vary between 4452 (min) and 23948 (max) with a median of 11197. The monthly tourist arrivals from India vary between 10071 (min) and 37945 (max) with a mean of 21179 and standard deviation of 6902.572. According to the coefficient of variance, the dispersion of monthly tourist arrivals of India is lower than the dispersion of monthly tourist arrivals of the UK.

4.3 Distribution of Tourist Arrivals from Both Countries

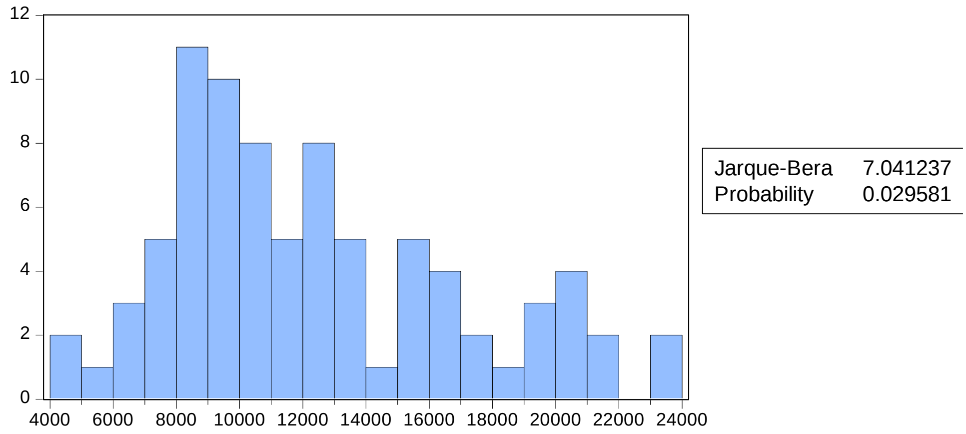


Figure 4.3: Normality plot for arrivals from UK

The normality plot for UK series (Figure 4.3) shows that the data aren't normal as it skewed to left. The significance of JB test (p = 0.0295) also confirmed that data series is significantly deviated from normality.

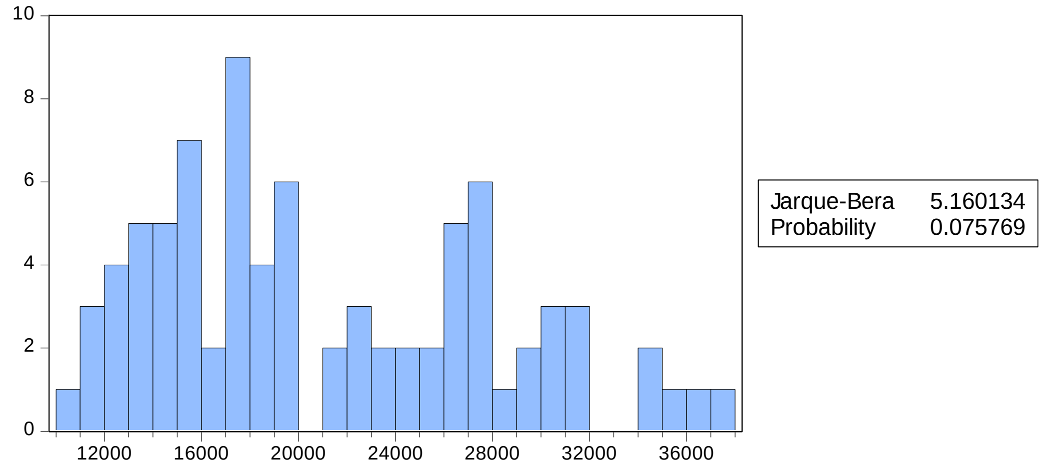


Figure 4.4: Normality plot for arrivals from India

The normality plot for India series (Figure 4.4) shows that data looks normal. The non significance of JB test ($p = 0.0757$) also confirmed that data series is not significantly deviated from normality.

4.4 Test for Seasonality

To confirm the seasonality of two series, the Kruskal Wallis test was carried out and the corresponding results are indicated in Table 4.2. The corresponding hypotheses of the this test are

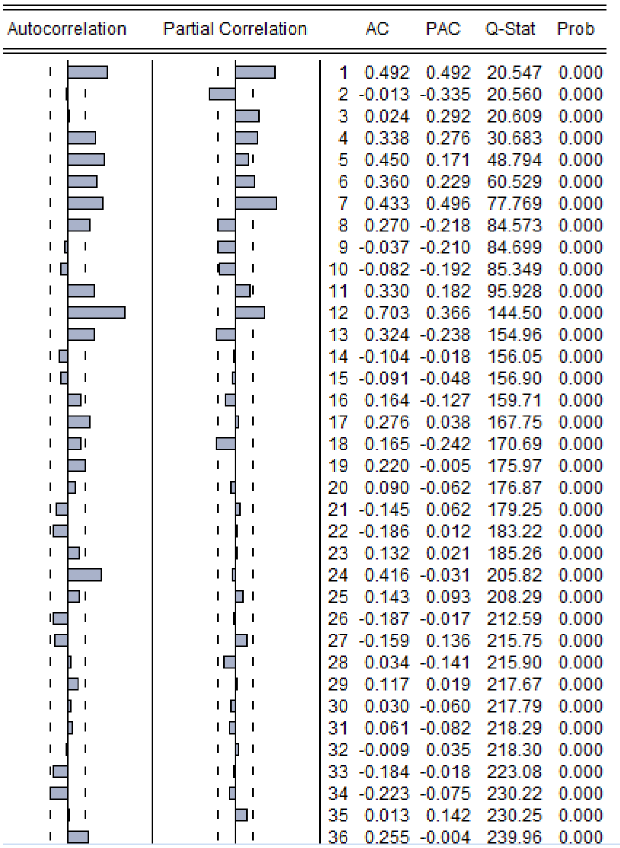
H_0 : series has no seasonality Vs H_1: series has seasonality

Table 4.2: Results of Kruskal Wallis test for data series in UK and India

	UK	India
N	82	82
Chi-Square	47.810[a]	2.762[a]
Df	11	11
Asymp. Sig.	.000	.994

Based on the results in Table 4.2, it can be concluded with 95% confidence that tourist arrivals from the United Kingdom is not significantly different from seasonality as corresponding p value is less than 0.05 while tourist arrivals from India is significantly different from seasonal pattern as p value is greater than 5%.

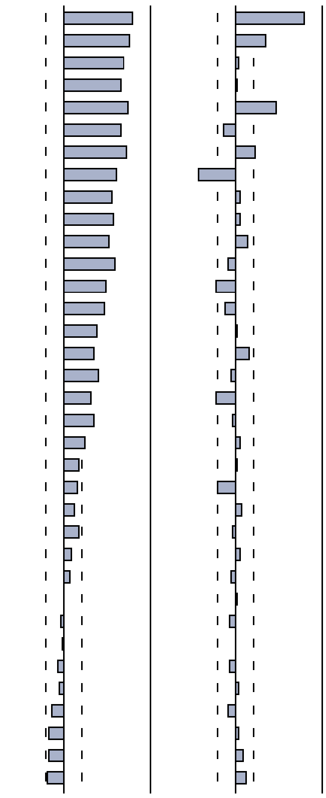
4.5 ACF and PACF for Data Series in United Kingdom and India



	AC	PAC	Q-Stat	Prob
1	0.492	0.492	20.547	0.000
2	-0.013	-0.335	20.560	0.000
3	0.024	0.292	20.609	0.000
4	0.338	0.276	30.683	0.000
5	0.450	0.171	48.794	0.000
6	0.360	0.229	60.529	0.000
7	0.433	0.496	77.769	0.000
8	0.270	-0.218	84.573	0.000
9	-0.037	-0.210	84.699	0.000
10	-0.082	-0.192	85.349	0.000
11	0.330	0.182	95.928	0.000
12	0.703	0.366	144.50	0.000
13	0.324	-0.238	154.96	0.000
14	-0.104	-0.018	156.05	0.000
15	-0.091	-0.048	156.90	0.000
16	0.164	-0.127	159.71	0.000
17	0.276	0.038	167.75	0.000
18	0.165	-0.242	170.69	0.000
19	0.220	-0.005	175.97	0.000
20	0.090	-0.062	176.87	0.000
21	-0.145	0.062	179.25	0.000
22	-0.186	0.012	183.22	0.000
23	0.132	0.021	185.26	0.000
24	0.416	-0.031	205.82	0.000
25	0.143	0.093	208.29	0.000
26	-0.187	-0.017	212.59	0.000
27	-0.159	0.136	215.75	0.000
28	0.034	-0.141	215.90	0.000
29	0.117	0.019	217.67	0.000
30	0.030	-0.060	217.79	0.000
31	0.061	-0.082	218.29	0.000
32	-0.009	0.035	218.30	0.000
33	-0.184	-0.018	223.08	0.000
34	-0.223	-0.075	230.22	0.000
35	0.013	0.142	230.25	0.000
36	0.255	-0.004	239.96	0.000

Figure 4.5: Correlogram of original series for United Kingdom

The plot of ACF of UK series in Figure 4.5 shows that first (r_1)and several auto correlations (r_4, r_5, r_6, r_7, r_8) are significantly different from zero at 5% significance level. Thus, ACF of the original series of United Kingdom doesn’t decay exponentially, which suggests that the series is not stationary. Furthermore, highly significant auto correlations can be seen at lag 11, 12 and 13 while the highest auto correlation can be seen at lag 12 ($r_{11} < r_{12} > r_{13}$) which suggests seasonal differencing is also required to make the series stationary, in addition to short term differences.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.824	0.824	57.791	0.000
		2	0.797	0.366	112.45	0.000
		3	0.728	0.036	158.66	0.000
		4	0.681	0.014	199.61	0.000
		5	0.765	0.486	251.95	0.000
		6	0.690	-0.152	295.12	0.000
		7	0.754	0.246	347.28	0.000
		8	0.623	-0.443	383.45	0.000
		9	0.572	0.051	414.32	0.000
		10	0.585	0.060	447.07	0.000
		11	0.540	0.147	475.38	0.000
		12	0.619	-0.084	513.02	0.000
		13	0.501	-0.231	538.06	0.000
		14	0.476	-0.125	560.97	0.000
		15	0.399	0.021	577.36	0.000
		16	0.357	0.164	590.66	0.000
		17	0.414	-0.057	608.80	0.000
		18	0.329	-0.242	620.44	0.000
		19	0.357	-0.047	634.35	0.000
		20	0.248	0.064	641.16	0.000
		21	0.188	0.009	645.17	0.000
		22	0.163	-0.209	648.20	0.000
		23	0.129	0.070	650.14	0.000
		24	0.184	-0.030	654.16	0.000
		25	0.084	0.058	655.02	0.000
		26	0.069	-0.063	655.61	0.000
		27	0.005	0.011	655.61	0.000
		28	-0.039	-0.071	655.81	0.000
		29	-0.022	-0.002	655.87	0.000
		30	-0.082	-0.074	656.75	0.000
		31	-0.053	0.038	657.14	0.000
		32	-0.148	-0.083	660.17	0.000
		33	-0.179	0.035	664.68	0.000
		34	-0.181	0.083	669.36	0.000
		35	-0.198	0.115	675.10	0.000

Figure 4.6: Correlogram of original series for India

The plot of ACF for India in Figure 4.6 also exhibits that most of the auto correlations at many number of lags are significantly different from zero at 5% level. It suggests that the series for India is also non stationary and only short term difference may be sufficient to make the series stationary.

4.6 Stationarity of the Series

The time series properties of stationary and non stationary can be determined by applying the Augmented Dickey Fuller test (Dickey Fuller, 1979). The corresponding hypotheses are:

Ho: The series is not stationary or series has a unit root

H1: The series is stationary or series hasn't a unit root

The results of unit root tests for two series are shown in Table 4.3 and Table 4.4 respectively.

Table 4.3: Results of unit root test for original series of UK

Null Hypothesis: UK has a unit root Exogenous: Constant Lag Length: 11 (Automatic - based on SIC, maxlag=11)			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		0.754519	0.9925
Test critical values:	1% level	-3.527045	
	5% level	-2.903566	
	10% level	-2.589227	

Table 4.4: Results of unit root test for original series of India

Null Hypothesis: INDIA has a unit root Exogenous: Constant Lag Length: 7 (Automatic - based on SIC, maxlag=11)			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-0.713331	0.8364
Test critical values:	1% level	-3.521579	
	5% level	-2.901217	
	10% level	-2.587981	

Results of Augmented Dickey Fuller test on the original series of UK in Table 4.3 confirm that the series isn't stationary as corresponding p value is greater than 0.05 at 5% significance level (ADF = 0.7545, p = 0.9925). Similarly, results of the Augmented Dickey fuller test on the original series of India in Table 4.4 confirm that the series is also non stationary as corresponding p value is greater than 0.05 at 5% significance level (ADF = -0.7133, p = 0.8364).

Thus, the results of unit root tests shown in the Table 4.3 and Table 4.4 are consistent with the non stationary patterns of ACF plots in Figure 4.5 and Figure 4.6. Hence, the first difference is taken to eliminate the trend and achieve the stability of both series.

4.7 Time Series Plots of First Differenced Series.

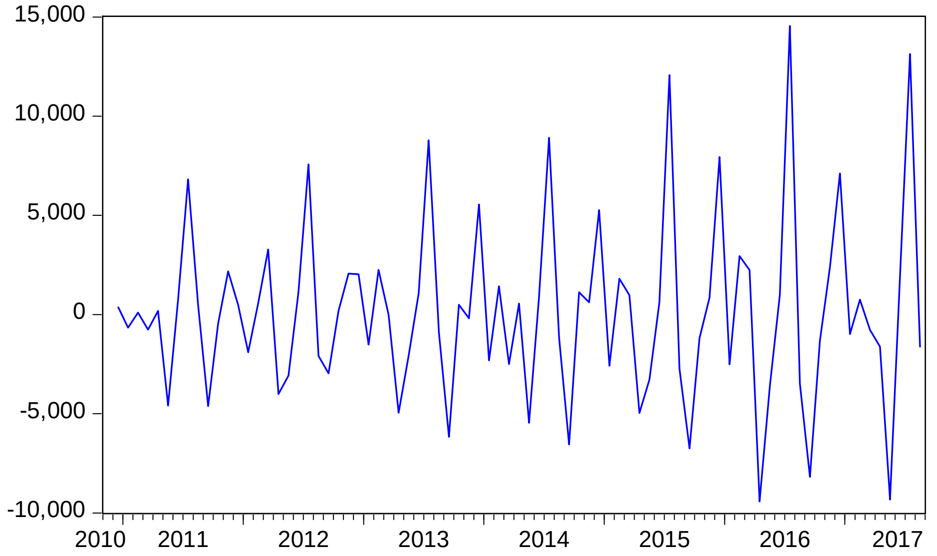


Figure 4.7: Time series plot of 1st differenced series of UK

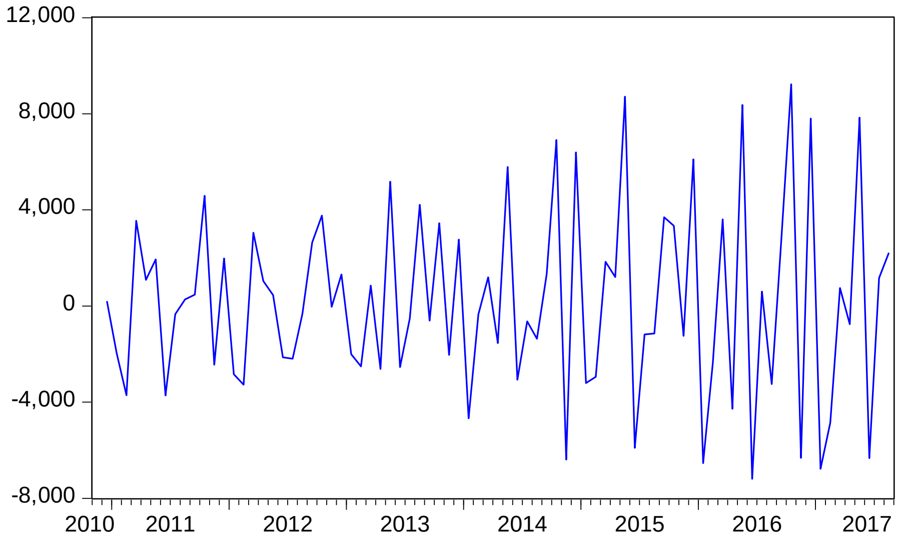


Figure 4.8: Time series plot of 1st differenced series of India

Time series plots of 1st differenced series in Figure 4.7 and Figure 4.8 don't exhibit any trend pattern. The Augmented Dickey Fuller test was carried out again to confirm the stationarity of the 1st differenced series.

4.8 Stationarity of First Differenced Series.

Table 4.5: The results of unit root test for 1st difference series of UK

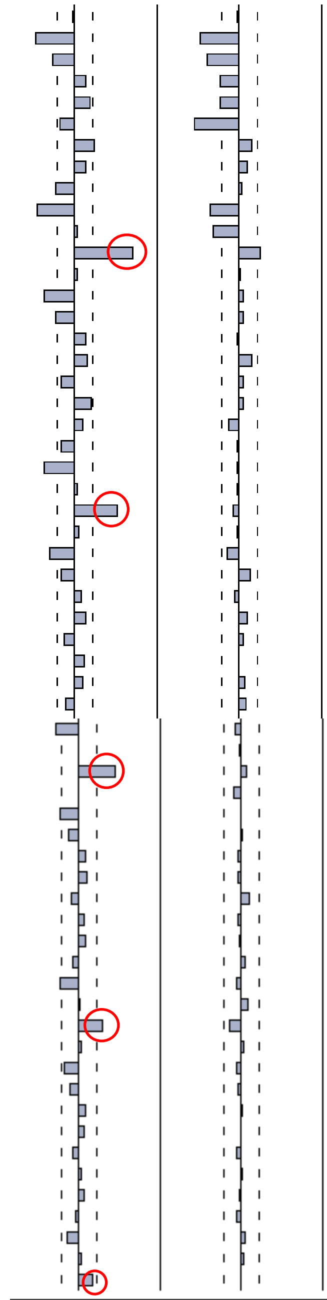
Null Hypothesis: UK1 has a unit root Exogenous: Constant Lag Length: 10 (Automatic - based on SIC, maxlag=11)			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-8.854333	0.0000
Test critical values:	1% level	-3.527045	
	5% level	-2.903566	
	10% level	-2.589227	

Table 4.6: The results of unit root test for 1st difference series of India

Null Hypothesis: I1 has a unit root			
Exogenous: Constant			
Lag Length: 11 (Automatic - based on SIC, maxlag=11)			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.991126	0.0407
Test critical values:	1% level	-3.528515	
	5% level	-2.904198	
	10% level	-2.589562	

According to the Table 4.5 and Table 4.6, the corresponding p values of Augmented Dickey fuller tests on 1st difference series of UK and India are 0.000 and 0.0407 respectively. The results confirm that both series are stationary as corresponding p values are less than 0.05 at 5 % significance level. However, it should be noted when there is seasonality pattern, the ADF test leads to wrong conclusions (Peiris, 2017) unless we do not look at ACF.

4.9 Models Identification for UK Series



	AC	PAC	Q-Stat	Prob
1	-0.015	-0.015	0.0179	0.894
2	-0.489	-0.489	20.343	0.000
3	-0.279	-0.390	27.030	0.000
4	0.155	-0.238	29.116	0.000
5	0.201	-0.233	32.687	0.000
6	-0.184	-0.551	35.710	0.000
7	0.251	0.160	41.413	0.000
8	0.154	0.104	43.604	0.000
9	-0.230	0.030	48.563	0.000
10	-0.472	-0.355	69.656	0.000
11	0.033	-0.327	69.761	0.000
12	0.738	0.269	122.87	0.000
13	0.030	0.019	122.96	0.000
14	-0.379	0.050	137.34	0.000
15	-0.245	0.063	143.48	0.000
16	0.134	-0.011	145.32	0.000
17	0.169	0.161	148.32	0.000
18	-0.165	0.046	151.23	0.000
19	0.210	0.051	156.00	0.000
20	0.106	-0.124	157.24	0.000
21	-0.160	-0.020	160.11	0.000
22	-0.375	-0.012	176.09	0.000
23	0.028	-0.023	176.18	0.000
24	0.544	-0.081	211.11	0.000
25	0.045	-0.014	211.36	0.000
26	-0.305	-0.142	222.76	0.000
27	-0.165	0.146	226.16	0.000
28	0.096	-0.061	227.34	0.000
29	0.139	0.111	229.83	0.000
30	-0.120	0.058	231.74	0.000
31	0.116	0.003	233.53	0.000
32	0.108	0.075	235.13	0.000
33	-0.113	0.088	236.92	0.000
34	-0.293	-0.072	249.24	0.000
35	0.002	-0.019	249.24	0.000
36	0.461	0.066	280.93	0.000
37	-0.000	-0.091	280.93	0.000
38	-0.235	-0.005	289.54	0.000
39	-0.122	0.022	291.92	0.000
40	0.089	-0.038	293.23	0.000
41	0.101	-0.029	294.95	0.000
42	-0.090	0.099	296.34	0.000
43	0.067	-0.042	297.14	0.000
44	0.087	-0.016	298.53	0.000
45	-0.076	0.064	299.60	0.000
46	-0.232	-0.058	309.93	0.000
47	0.021	0.084	310.02	0.000
48	0.308	-0.152	329.38	0.000
49	0.027	0.030	329.53	0.000
50	-0.178	-0.061	336.44	0.000
51	-0.108	-0.038	339.03	0.000
52	0.098	0.012	341.26	0.000
53	0.069	0.001	342.40	0.000
54	-0.078	-0.058	343.91	0.000
55	0.031	0.012	344.16	0.000
56	0.074	-0.023	345.63	0.000
57	-0.049	-0.056	346.29	0.000
58	-0.148	0.056	352.69	0.000
59	0.026	0.037	352.89	0.000
60	0.183	-0.003	363.61	0.000

Figure 4.9: ACF of 1st differenced series for UK (1 ST)

According to the ACF of 1st differenced series for UK data in Figure 4.9, the auto correlations at lag 12, 24 and 36 are significantly greater than zero. Also, these seasonal auto correlations decay exponentially at lag 12, 24, 36, 48, 60 as $r_{12} > r_{24} > r_{36} > r_{48} > r_{60}$. Moreover, it depicts that the series has a seasonal pattern with length 12. Thus, the seasonal difference of length 12 of UK series is taken to eliminate the seasonality. The plot of 12th difference of 1st differenced series in UK is shown in Figure 4.10.

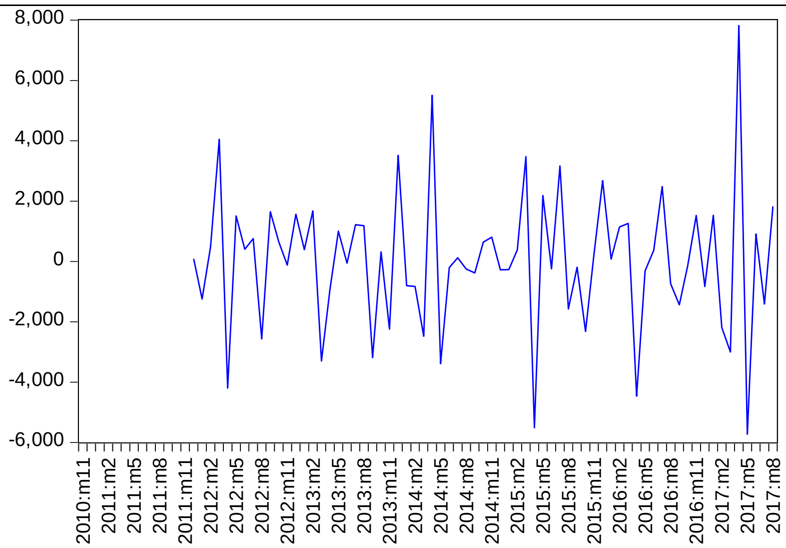


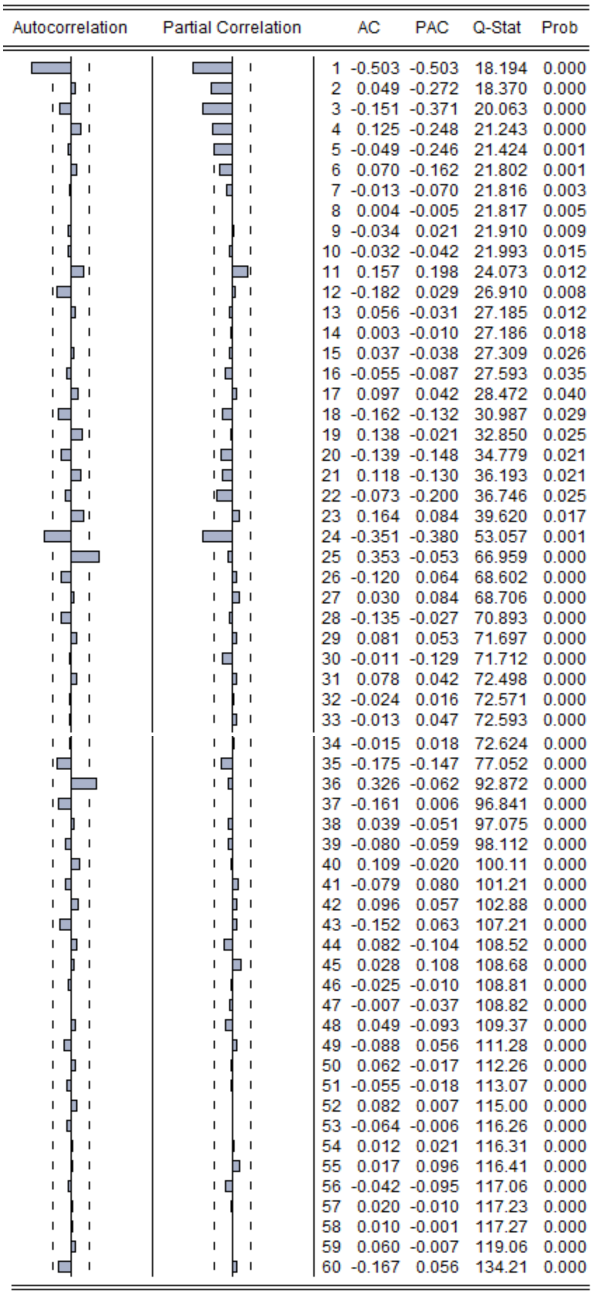
Figure 4.10: Time series plot of 12th difference of 1st differenced series for UK

The unit root test was carried out and the corresponding results were given in Table 4.7.

Table 4.7: Results of ADF test for 12th difference of 1st differenced series of the UK

Null Hypothesis: UK12 has a unit root			
Exogenous: Constant			
Lag Length: 3 (Automatic - based on SIC, maxlag=10)			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.400338	0.0000
Test critical values:	1% level	-3.534868	
	5% level	-2.906923	
	10% level	-2.591006	
*MacKinnon (1996) one-sided p-values.			

The results in Table 4.7 confirmed that the series is stationary as corresponding p value is less than 0.05 at 5 % significance level (p = 0.000). Corrlegram of the 12th difference of 1st differenced series of UK $\{(y_t - y_{t-1}) - (y_{t-12} - y_{t-13})\}$ is plotted to identify the patterns of auto correlations.



	AC	PAC	Q-Stat	Prob
1	-0.503	-0.503	18.194	0.000
2	0.049	-0.272	18.370	0.000
3	-0.151	-0.371	20.063	0.000
4	0.125	-0.248	21.243	0.000
5	-0.049	-0.246	21.424	0.001
6	0.070	-0.162	21.802	0.001
7	-0.013	-0.070	21.816	0.003
8	0.004	-0.005	21.817	0.005
9	-0.034	0.021	21.910	0.009
10	-0.032	-0.042	21.993	0.015
11	0.157	0.198	24.073	0.012
12	-0.182	0.029	26.910	0.008
13	0.056	-0.031	27.185	0.012
14	0.003	-0.010	27.186	0.018
15	0.037	-0.038	27.309	0.026
16	-0.055	-0.087	27.593	0.035
17	0.097	0.042	28.472	0.040
18	-0.162	-0.132	30.987	0.029
19	0.138	-0.021	32.850	0.025
20	-0.139	-0.148	34.779	0.021
21	0.118	-0.130	36.193	0.021
22	-0.073	-0.200	36.746	0.025
23	0.164	0.084	39.620	0.017
24	-0.351	-0.380	53.057	0.001
25	0.353	-0.053	66.959	0.000
26	-0.120	0.064	68.602	0.000
27	0.030	0.084	68.706	0.000
28	-0.135	-0.027	70.893	0.000
29	0.081	0.053	71.697	0.000
30	-0.011	-0.129	71.712	0.000
31	0.078	0.042	72.498	0.000
32	-0.024	0.016	72.571	0.000
33	-0.013	0.047	72.593	0.000
34	-0.015	0.018	72.624	0.000
35	-0.175	-0.147	77.052	0.000
36	0.326	-0.062	92.872	0.000
37	-0.161	0.006	96.841	0.000
38	0.039	-0.051	97.075	0.000
39	-0.080	-0.059	98.112	0.000
40	0.109	-0.020	100.11	0.000
41	-0.079	0.080	101.21	0.000
42	0.096	0.057	102.88	0.000
43	-0.152	0.063	107.21	0.000
44	0.082	-0.104	108.52	0.000
45	0.028	0.108	108.68	0.000
46	-0.025	-0.010	108.81	0.000
47	-0.007	-0.037	108.82	0.000
48	0.049	-0.093	109.37	0.000
49	-0.088	0.056	111.28	0.000
50	0.062	-0.017	112.26	0.000
51	-0.055	-0.018	113.07	0.000
52	0.082	0.007	115.00	0.000
53	-0.064	-0.006	116.26	0.000
54	0.012	0.021	116.31	0.000
55	0.017	0.096	116.41	0.000
56	-0.042	-0.095	117.06	0.000
57	0.020	-0.010	117.23	0.000
58	0.010	-0.001	117.27	0.000
59	0.060	-0.007	119.06	0.000
60	-0.167	0.056	134.21	0.000

Figure 4.11: Correlogram of the 12th difference of 1st differenced series of UK (1 ST and 1 LT)

In Figure 4.11, it can be seen that among non seasonal auto correlations, only first auto correlation is significantly different from zero and among seasonal auto correlations, the auto correlation at lag 24 and 36 are significantly different from zero. Thus, it suggests a seasonal ARIMA model with non seasonal MA (1) component and seasonal component of MA (2) or MA (3). Also, it can be seen that PACF decays exponentially towards to zero, indicating a non seasonal MA component of order 1. Thus, it is possible to have a model like ARIMA $(0, 1, 1) \times (0, 1, 1)_{12}$.

In the plot of PACF, it can be also observed that first three non seasonal partial auto correlations are significantly different from zero and then it slowly diminish to zero. This suggests that there may have been higher orders of non seasonal AR components in the model. Thus, models with non seasonal AR components of order 1, 2 and 3 were considered.

However, it is not possible to identify a unique model with the inspection of sample ACF and sample PACF. Thus, based on the simultaneous inspection of the plots of ACF and PACF the following six models were postulated as follows.

$$\text{ARIMA } (0, 1, 1) \times (0, 1, 1)_{12}$$
$$\text{ARIMA } (0, 1, 1) \times (0, 1, 0)_{12}$$
$$\text{ARIMA } (1, 1, 1) \times (0, 1, 0)_{12}$$
$$\text{ARIMA } (1, 1, 0) \times (0, 1, 0)_{12}$$
$$\text{ARIMA } (2, 1, 0) \times (0, 1, 0)_{12}$$
$$\text{ARIMA } (3, 1, 0) \times (0, 1, 0)_{12}$$

The details of the six models are shown in Table 4.8.

4.10.1 Comparison of parsimonious models for United Kingdom.

Table 4.8: Comparison of postulated models for the United Kingdom.

Model	Parameter		p value	ARCH-LM test for heteroscedasticity	Breusch-Godfrey test for serial correlation	Normality
1	C MA(1) SMA(12)	9.03520 -0.98563 -0.41499	0.2290 0.0000 0.0016	1.1375 (0.2901)	0.2347 (0.7915)	3.5346 (0.1707)
2	C MA(1)	5.98533 -0.99951	0.5530 0.0000	2.2243 (0.1406)	0.3942 (0.6758)	2.4415 (0.2950)
3	C AR(1) MA(1)	8.05301 0.04816 -0.99996	0.2781 0.7035 0.0000	2.9923 (0.0884)	1.3494 (0.2668)	2.1559 (0.3402)
4	C AR(1)	14.06777 -0.50702	0.9322 0.0000	2.7932 (0.0995)	8.2193 (0.0007)	1.9885 (0.3699)
5	C AR(1) AR(2)	22.85178 -0.64363 -0.27380	0.8579 0.0000 0.0265	2.3737 (0.1283)	6.0552 (0.0040)	0.6413 (0.7256)
6	C AR(1) AR(2) AR(3)	21.30859 -0.74861 -0.51474 -0.37702	0.8085 0.0000 0.0004 0.0022	10.1468 (0.0022)	3.0411 (0.0552)	0.6509 (0.7221)

According to the results in Table 4.8, the model (3) is not suitable as AR (1) parameter is not significant (p = 0.7035). According to the ACF plots of residuals (Appendix B), the auto correlations of model (1), model (2), model (3) and model (6) are within 95% confidence intervals indicating residuals are not correlated. Also, the high probability values of L-jung Box Q statistc confirmed that residuals of these models are not correlated. The low probability values ($p < 0.05$) of Q statistic of residuals of the model (4) and model (5) confirmed that the residuals of the model (4) and model (5) are correlated.

Results of Breusch-Godfrey test indicates that there is a serial correlation in the residuals of the model (4) and model (5) as the corresponding p values of F statistic are less than 0.05 at 5% significance level. Among other models, model (6) can be rejected as the null hypothesis of homoscedasticity is rejected ($p = 0.0022$). Accordingly, the model (3), model (4), model (5) and model (6) can't be considered as suitable models.

The residuals of the model (1) and model (2) model don't indicate any arch effect and serial correlation as corresponding p values of F statistic are greater than 0.05. Furthermore, residuals of these two models don't significantly deviate from normality as corresponding p values are greater than 0.05. Thus, based on the minimum values of AIC and SIC criteria the model (1) and model (2) are compared and the corresponding results are shown in Table 4.9.

Table 4.9: Comparison of suggested ARIMA models.

Model	AIC	SIC	R square
ARIMA (0, 1, 1) $\times(0, 1, 1)_{12}$ (model 1)	17.642	17.739	54.22
ARIMA (0, 1, 1) $\times(0, 1, 0)_{12}$ (model 2)	17.732	17.797	48.46

According to the results in Table 4.9, ARIMA $(0, 1, 1)\times(0, 1, 1)_{12}$ model has the lowest values of AIC and SIC criterion. Meanwhile, R square of ARIMA $(0, 1, 1)\times(0, 1, 1)_{12}$ model is better than the R square of ARIMA $(0, 1, 1) \times (0, 1, 0)_{12}$ model. Thus, ARIMA $(0, 1, 1)\times(0, 1, 1)_{12}$ is recommended for further analysis.

4.10.2 Estimation of parameters

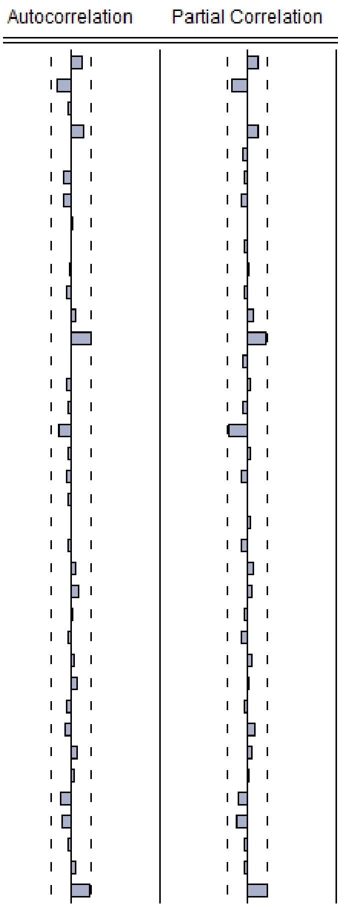
It was observed from Table 4.8, the coefficient for constant parameter of ARIMA $(0, 1, 1)\times(0, 1, 1)_{12}$ model is not significant ($p = 0.2290$). The coefficient for MA (1) parameter is -0.98563 which is statistically significant as the corresponding p value is less than 0.05 (p = 0.000). Also, the coefficient for SMA (12) parameter is -0.41499 which is significant (p = 0.0016).

4.10.3 Diagnostic Checking

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.044	0.044	0.1412	
		2	-0.073	-0.075	0.5263	
		3	-0.210	-0.205	3.7984	0.051
		4	0.040	0.054	3.9195	0.141
		5	0.105	0.077	4.7650	0.190
		6	0.151	0.113	6.5427	0.162
		7	0.095	0.124	7.2582	0.202
		8	-0.035	0.009	7.3544	0.289
		9	-0.146	-0.099	9.0911	0.246
		10	-0.111	-0.099	10.118	0.257
		11	0.061	0.010	10.430	0.317
		12	0.039	-0.056	10.562	0.393
		13	0.030	-0.011	10.642	0.474
		14	-0.108	-0.073	11.680	0.472
		15	-0.156	-0.122	13.880	0.382
		16	-0.152	-0.129	16.022	0.312
		17	-0.051	-0.099	16.264	0.365
		18	-0.134	-0.242	17.985	0.325
		19	0.056	-0.007	18.291	0.371
		20	-0.131	-0.162	20.010	0.332
		21	-0.037	-0.040	20.147	0.386
		22	-0.102	-0.048	21.228	0.384
		23	0.040	0.030	21.401	0.435
		24	-0.083	-0.121	22.148	0.451
		25	0.215	0.244	27.275	0.244
		26	-0.024	-0.035	27.341	0.289
		27	-0.067	-0.076	27.864	0.314
		28	-0.153	-0.122	30.656	0.241
		29	0.075	0.025	31.344	0.257
		30	0.095	-0.104	32.485	0.255
		31	0.165	0.107	35.979	0.174
		32	0.021	-0.086	36.039	0.207
		33	0.000	-0.002	36.039	0.245
		34	-0.053	-0.091	36.438	0.270
		35	-0.017	-0.054	36.480	0.310
		36	0.291	0.129	49.060	0.046
		37	0.033	-0.068	49.223	0.056

Figure 4.12: ACF of residuals for United Kingdom

The ACF plot in of Figure 4.12 implies that residuals are independently distributed as all auto correlation estimates are within 95% confidence intervals. Also, all the corresponding p values of Ljung-Box Q statistic are greater than 0.05 confirming residuals are random. The correlogram of squared residuals was plotted at different lags to check the heteroscedasticity and the corresponding plot is shown in Figure 4.13.



	AC	PAC	Q-Stat	Prob
1	0.130	0.130	1.2106	0.271
2	-0.157	-0.176	3.0054	0.223
3	-0.045	0.002	3.1555	0.368
4	0.135	0.120	4.5387	0.338
5	-0.001	-0.050	4.5387	0.475
6	-0.087	-0.043	5.1314	0.527
7	-0.096	-0.080	5.8629	0.556
8	0.018	0.007	5.8897	0.660
9	-0.006	-0.035	5.8928	0.751
10	-0.008	0.012	5.8985	0.824
11	-0.053	-0.044	6.1397	0.864
12	0.062	0.069	6.4718	0.890
13	0.226	0.205	10.939	0.616
14	-0.001	-0.063	10.939	0.691
15	-0.060	0.027	11.265	0.734
16	-0.036	-0.051	11.385	0.785
17	-0.151	-0.226	13.519	0.701
18	-0.031	0.034	13.610	0.754
19	-0.051	-0.082	13.866	0.791
20	-0.039	-0.001	14.018	0.830
21	-0.006	0.038	14.022	0.869
22	-0.045	-0.072	14.234	0.893
23	0.044	0.070	14.444	0.913
24	0.091	0.061	15.347	0.910
25	0.008	-0.047	15.355	0.933
26	-0.032	-0.067	15.475	0.948
27	0.027	0.050	15.561	0.961
28	0.079	0.025	16.305	0.961
29	-0.056	-0.048	16.690	0.967
30	-0.072	0.091	17.347	0.968
31	0.081	0.057	18.189	0.967
32	0.035	0.009	18.355	0.974
33	-0.116	-0.100	20.180	0.961
34	-0.101	-0.128	21.607	0.951
35	-0.037	-0.043	21.803	0.960
36	0.058	-0.032	22.304	0.964
37	0.212	0.237	29.212	0.816

Figure 4.13: Correlogram of squared residuals.

It can be seen from the correlogram of squared residuals in Figure 4.13 that the auto correlations and partial auto correlations aren't significantly different from zero implying the constant variance. Further, high probability values ($p > 0.05$) confirm the homoscedasticity of the residuals. Relatively, correlogram of residual squared of the fitted model suggests the residuals of mean equation have a constant variance. Further, the ARCH LM test is employed to confirm the presence of ARCH effect and the results are indicated in Table 4.10.

Table 4.10: Results of Heteroscedasticity Test (ARCH - LM test)

F-statistic	1.137457	Prob. F(1,66)	0.2901
Obs*R-squared	1.152071	Prob. Chi-Square(1)	0.2831

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 04/12/19 Time: 11:13
Sample (adjusted): 2012M01 2017M08
Included observations: 68 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2158633.	521240.7	4.141337	0.0001
RESID^2(-1)	0.130444	0.122309	1.066516	0.2901

R-squared	0.016942	Mean dependent var	2484522.
Adjusted R-squared	0.002047	S.D. dependent var	3485800.
S.E. of regression	3482229.	Akaike info criterion	32.99321
Sum squared resid	8.00E+14	Schwarz criterion	33.05849
Log likelihood	-1119.769	Hannan-Quinn criter.	33.01908
F-statistic	1.137457	Durbin-Watson stat	1.936738
Prob(F-statistic)	0.290076		

The results in Table 4.10 clearly indicates that there is no significant arch effect in the residuals of the fitted model as the corresponding p value of F statistic is greater than 0.05 (p = 0.2901). Results of Breusch-Godfrey Serial Correlation LM Test are indicated in Table 4.11.

Table 4.11: Breusch-Godfrey Serial Correlation LM Test

F-statistic	0.234675	Prob. F(2,64)	0.7915
Obs*R-squared	0.154228	Prob. Chi-Square(2)	0.9258

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 04/12/19 Time: 11:13
Sample: 2011M12 2017M08
Included observations: 69
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.229849	7.542235	0.030475	0.9758
MA(1)	0.001643	0.027521	0.059699	0.9526
SMA(12)	-0.016137	0.129685	-0.124437	0.9014
RESID(-1)	0.053365	0.127609	0.418190	0.6772
RESID(-2)	-0.071637	0.130104	-0.550613	0.5838

R-squared	0.002235	Mean dependent var	111.6233
Adjusted R-squared	-0.060125	S.D. dependent var	1577.274
S.E. of regression	1623.999	Akaike info criterion	17.69288
Sum squared resid	1.69E+08	Schwarz criterion	17.85477
Log likelihood	-605.4042	Hannan-Quinn criter.	17.75710
F-statistic	0.035843	Durbin-Watson stat	2.016026
Prob(F-statistic)	0.997481		

The non significance of the Breusch-Godfrey test of serial correlation ($p = 0.9258$) indicates that there is no significant serial correlation in the residuals. The normality plot of the UK is presented in Figure 4.14.

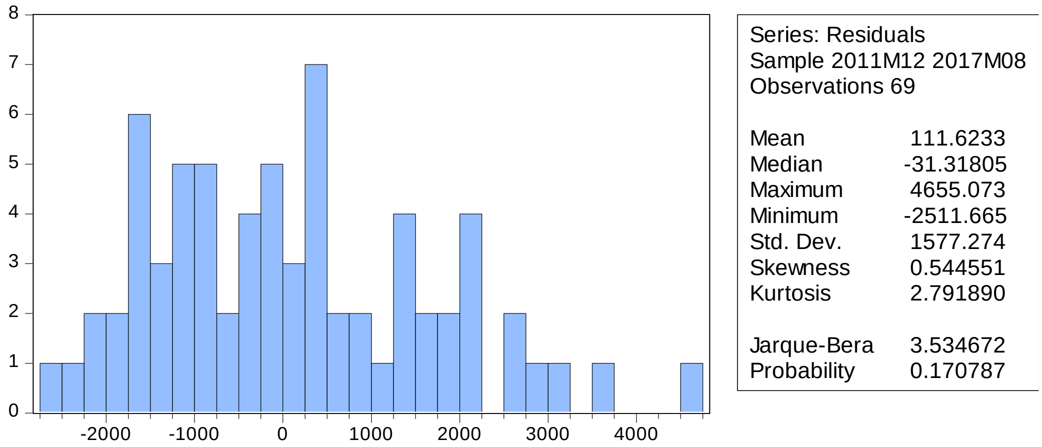


Figure 4.14: Normality plot for residuals of UK

Results in Figure 4.14 confirm that residuals are not significantly deviated from a normal distribution as the corresponding p value of the Jarque-Bera statistic is greater than 5%. The above results concluded that the residuals are white noise. Thus, ARIMA $(0, 1, 1) \times (0, 1, 1)_{12}$ model is advocated to forecast the monthly tourist arrivals from the UK.

4.10.4 Forecasting monthly tourist arrivals from the UK for an independent set.

The best fitted ARIMA $(0, 1, 1) \times (0, 1, 1)_{12}$ model for United Kingdom can be written as:

$$(1 - B)(1 - B^{12})y_t = 9.035 + (1 - 0.414994B^{12})(1 - 0.985634B)e_t$$

The above equation can be expanded as follows.

$$y_t - B^{12}y_t - By_t + B^{13}y_t = 9.035 + e_t - 0.985634Be_t - 0.414994B^{12}e_t$$
$$+0.409032B^{13}e_t$$

$$y_t - y_{t-12} - y_{t-1} + y_{t-13} = 9.035 + e_t - 0.985634e_{t-1} - 0.414994e_{t-12}$$
$$+ \ 0.409032B^{13}e_t$$

$$y_t = 9.035 + y_{t-12} + y_{t-1} - y_{t-13} + e_t - 0.985634e_{t-1} - 0.414994e_{t-12} + 0.409032e_{t-13} \quad \text{(A)}$$

The forecast values can be calculated from the formula (A) using past values (Appendix A) and the residuals (Appendix D).

y_{83} = value for September, 2017 (Table 4.12)

$$y_{83} = 9.035 + y_{71} + y_{82} - y_{70} + e_{83} - 0.985634e_{82} - 0.414994e_{71} + 0.409032e_{70}$$

= 14089.3909

Similarly, monthly values were predicted up to February, 2018 and the values are as follows.

y_{84} = value for September, 2017

= 13175

y_{85} = value for November, 2017

= 14907

y_{86} = value for December, 2017

=21813

y_{87} = value for January, 2018

=20245

y_{88} = value for February, 2018

=21699

4.10.5: Forecasting for the training set

Table 4.12: Forecasting for the training set (from November, 2010 to August, 2017)

Time period	UK data	Forecasts	Percentage error (%)
2010M11	9788		
2010M12	10176		
2011M01	9518		
2011M02	9614		
2011M03	8852		
2011M04	9038		
2011M05	4452		
2011M06	5188		
2011M07	12003		
2011M08	12486		
2011M09	7871		
2011M10	7408		
2011M11	9589		
2011M12	10063	11103	-10.34
2012M01	8162	10242	-25.48
2012M02	8746	10529	-20.39
2012M03	12032	10204	15.19
2012M04	8019	9695	-20.91
2012M05	4940	5360	-8.5
2012M06	6076	6161	-1.41
2012M07	13643	13251	2.87
2012M08	11558	13356	-15.56
2012M09	8586	8781	-2.27
2012M10	8767	8447	3.65
2012M11	10828	10998	-1.57
2012M12	12861	12156	5.49
2013M01	11350	11303	0.41
2013M02	13604	11599	14.74
2013M03	13590	11283	16.97
2013M04	8642	10784	-24.78
2013M05	6567	6457	1.67
2013M06	7642	7268	4.9

Table 4.12 (Continued)

Time period	UK data	Forecasts	Percentage error (%)
2013M07	16424	14367	12.53
2013M08	15519	14480	6.69
2013M09	9356	9915	-5.97
2013M10	9850	9590	2.64
2013M11	9663	12149	-25.73
2013M12	15209	13316	12.45
2014M01	12896	12473	3.28
2014M02	14316	12778	10.74
2014M03	11823	12471	-5.48
2014M04	12380	11981	3.23
2014M05	6918	7663	-10.77
2014M06	7790	8483	-8.89
2014M07	16692	15591	6.6
2014M08	15532	15713	-1.17
2014M09	8983	11157	-24.2
2014M10	10112	10841	-7.21
2014M11	10730	13409	-24.97
2014M12	15996	14585	8.82
2015M01	13410	13751	-2.54
2015M02	15212	14065	7.54
2015M03	16191	13767	14.97
2015M04	11233	13286	-18.27
2015M05	7954	8977	-12.86
2015M06	8580	9806	-14.29
2015M07	20643	16923	18.02
2015M08	17908	17055	4.77
2015M09	11160	12507	-12.07
2015M10	9970	12200	-22.37
2015M11	10822	14778	-36.55
2015M12	18762	15963	14.92
2016M01	16253	15138	6.86
2016M02	19194	15461	19.45
2016M03	21430	15172	29.2
2016M04	12006	14699	-22.43

Table 4.12 (Continued)

Time period	UK data	Forecasts	Percentage error (%)
2016M05	8412	10400	-23.63
2016M06	9406	11238	-19.47
2016M07	23948	18364	23.32
2016M08	20475	18504	9.62
2016M09	12288	13966	-13.65
2016M10	10964	13668	-24.66
2016M11	13337	16255	-21.88
2016M12	20446	17449	14.66
2017M01	19468	16633	14.57
2017M02	20218	16965	16.09
2017M03	19451	16685	14.22
2017M04	17841	16221	9.08
2017M05	8520	11931	-40.03
2017M06	10424	12778	-22.58
2017M07	23553	19913	15.46
2017M08	21903	20063	8.4

The forecasts for training set were given in Table 4.13. The error percentages for training set vary between -40.03 % to 29.20% and MAPE value for the model is 13.43%.

Table 4.13: Comparison of actual values and forecasts for independent set

Period	Actuals	Forecasts	95% Confidence Interval		Error %	MAPE
September, 2017	12593	14089	10776	17402	-11.88	7.89
October, 2017	12518	13175	9793	16557	-5.25	
November, 2017	13634	14907	11578	18235	-9.33	
December, 2017	21756	21813	18495	25130	-0.26	
January, 2018	22940	20245	16890	23600	11.75	
February, 2018	23817	21699	18369	25029	8.89	

The error percentages for independent set vary between -11.88% to 11.75% and the value of MAPE is 7.89% (Table 4.12). The plot of forecasts of monthly tourist arrivals from the UK for the period from September, 2017 to February, 2018 is shown in Figure 4.15.

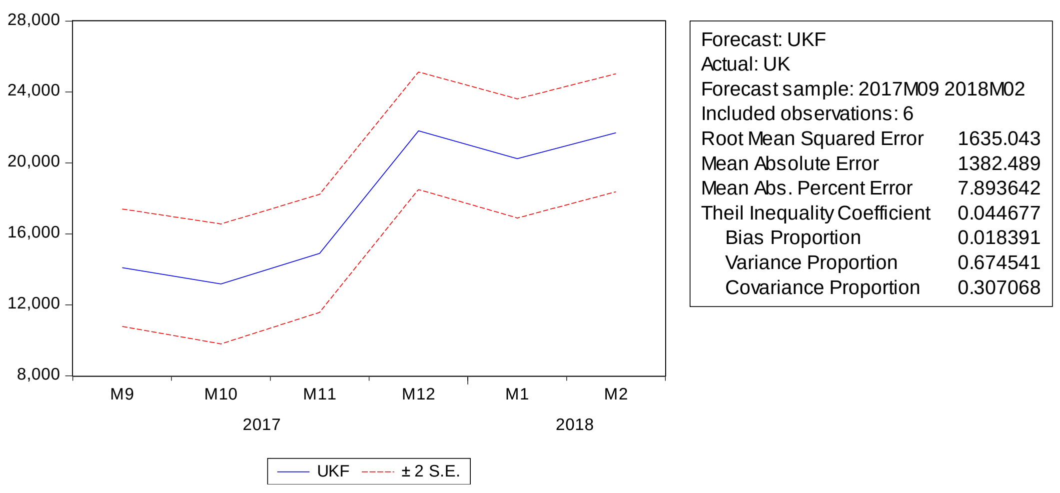
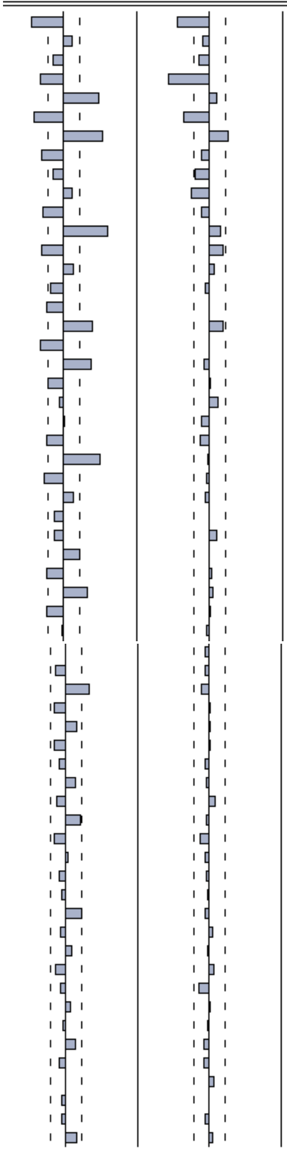


Figure 4.15: The plot of forecasts of monthly tourist arrivals from UK

4.11 Models Identification for India Series



	AC	PAC	Q-Stat	Prob
1	-0.451	-0.451	17.106	0.000
2	0.128	-0.094	18.510	0.000
3	-0.142	-0.154	20.249	0.000
4	-0.335	-0.590	30.061	0.000
5	0.499	0.102	52.055	0.000
6	-0.424	-0.356	68.160	0.000
7	0.553	0.268	95.898	0.000
8	-0.311	-0.109	104.80	0.000
9	-0.142	-0.196	106.68	0.000
10	0.130	-0.261	108.27	0.000
11	-0.295	-0.103	116.61	0.000
12	0.630	0.164	155.34	0.000
13	-0.312	0.191	164.99	0.000
14	0.151	0.078	167.29	0.000
15	-0.186	-0.054	170.82	0.000
16	-0.232	-0.001	176.41	0.000
17	0.414	0.189	194.42	0.000
18	-0.326	-0.007	205.77	0.000
19	0.400	-0.076	223.14	0.000
20	-0.225	0.020	228.74	0.000
21	-0.061	0.115	229.15	0.000
22	0.024	-0.104	229.22	0.000
23	-0.240	-0.137	235.87	0.000
24	0.519	-0.024	267.70	0.000
25	-0.272	-0.033	276.60	0.000
26	0.135	-0.063	278.81	0.000
27	-0.130	-0.006	280.93	0.000
28	-0.118	0.102	282.69	0.000
29	0.238	0.008	290.03	0.000
30	-0.241	0.040	297.69	0.000
31	0.352	0.057	314.33	0.000
32	-0.234	0.012	321.87	0.000
33	-0.022	-0.027	321.93	0.000
34	-0.008	-0.063	321.94	0.000
35	-0.148	-0.052	325.14	0.000
36	0.339	-0.099	342.27	0.000
37	-0.158	0.023	346.10	0.000
38	0.163	0.022	350.23	0.000
39	-0.166	0.016	354.64	0.000
40	-0.085	-0.062	355.81	0.000
41	0.153	-0.031	359.76	0.000
42	-0.135	0.083	362.91	0.000
43	0.209	-0.026	370.62	0.000
44	-0.165	-0.129	375.55	0.000
45	0.034	-0.057	375.77	0.000
46	-0.083	-0.046	377.09	0.000
47	-0.053	-0.012	377.65	0.000
48	0.241	-0.052	389.48	0.000
49	-0.075	0.053	390.67	0.000
50	0.087	-0.018	392.32	0.000
51	-0.140	0.070	396.72	0.000
52	-0.070	-0.145	397.85	0.000
53	0.076	0.010	399.26	0.000
54	-0.038	-0.015	399.61	0.000
55	0.136	-0.067	404.41	0.000
56	-0.087	-0.076	406.44	0.000
57	0.004	0.070	406.44	0.000
58	-0.054	0.007	407.29	0.000
59	-0.052	-0.055	408.12	0.000
60	0.163	0.063	416.64	0.000

Figure 4.16: ACF of 1st difference series for India (1 ST)

The ACF of 1st difference series for India series (Figure 4.16) appears that several non seasonal and seasonal auto correlations at different lags are considerably high and are significantly different from zero suggesting seasonal differencing. Thus, the seasonal difference of length 12 is taken to make the series stationary.

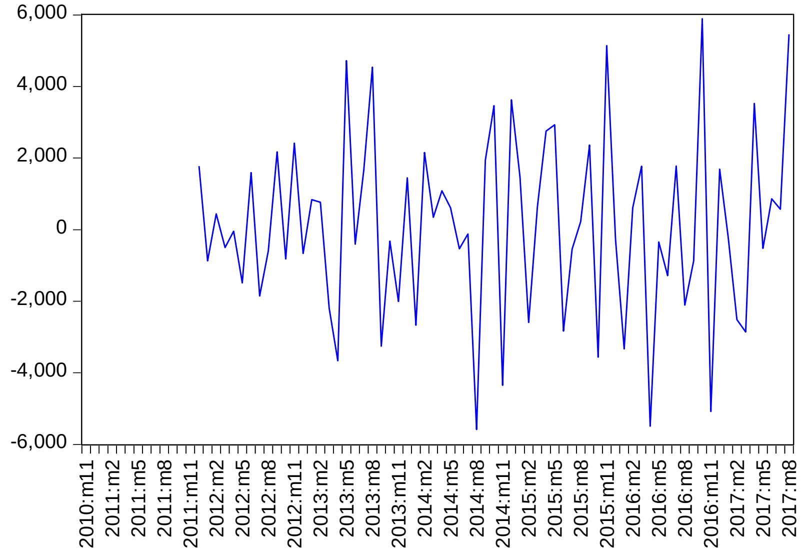


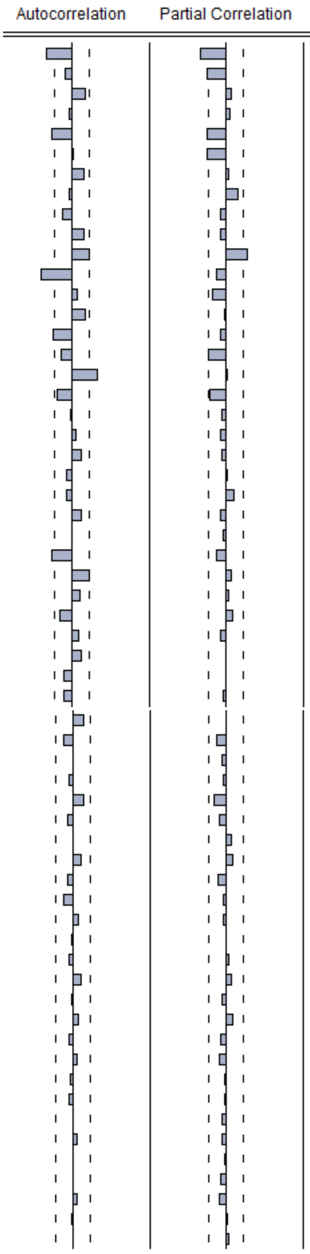
Figure 4.17: Time series plot of 12th difference of 1st differenced series for India

The time series plot of 12th difference of 1st differenced series for India was shown in Figure 4.17. The unit root test was used to confirm the stationarity and the relevant results are indicated in Table 4.14.

Table 4.14: Results of ADF test of 12th difference of 1st differenced series of India

Null Hypothesis: I12 has a unit root			
Exogenous: Constant			
Lag Length: 1 (Automatic - based on SIC, maxlag=10)			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-8.486680	0.0000
Test critical values:	1% level	-3.531592	
	5% level	-2.905519	
	10% level	-2.590262	
*MacKinnon (1996) one-sided p-values.			

The results of Augmented Dickey fuller tests in Table 4.14 confirm that the 12th difference in 1st differenced series of India is stationary as corresponding p value is less than 5 % (p = 0.000). The corrlegram of the 12th difference in 1st differenced series of India $\{(y_t\text{-}y_{t-1})\text{-}(y_{t-12}\text{-}y_{t-13})\}$ is plotted to identify the patterns of auto correlations.



	AC	PAC	Q-Stat	Prob
1	-0.349	-0.349	8.7921	0.003
2	-0.095	-0.247	9.4462	0.009
3	0.183	0.067	11.920	0.008
4	-0.027	0.064	11.974	0.018
5	-0.267	-0.256	17.411	0.004
6	0.019	-0.256	17.440	0.008
7	0.170	0.031	19.713	0.006
8	-0.027	0.169	19.770	0.011
9	-0.139	-0.075	21.340	0.011
10	0.167	-0.074	23.643	0.009
11	0.222	0.294	27.816	0.003
12	-0.415	-0.121	42.627	0.000
13	0.067	-0.188	43.024	0.000
14	0.174	-0.016	45.716	0.000
15	-0.258	-0.069	51.747	0.000
16	-0.153	-0.244	53.908	0.000
17	0.340	0.017	64.767	0.000
18	-0.200	-0.225	68.615	0.000
19	-0.012	-0.051	68.629	0.000
20	0.050	-0.070	68.883	0.000
21	0.118	-0.062	70.294	0.000
22	-0.079	0.023	70.952	0.000
23	-0.077	0.099	71.585	0.000
24	0.123	-0.068	73.224	0.000
25	-0.005	-0.027	73.226	0.000
26	-0.268	-0.133	81.434	0.000
27	0.243	0.069	88.336	0.000
28	0.103	0.028	89.603	0.000
29	-0.157	0.098	92.630	0.000
30	0.098	-0.066	93.824	0.000
31	0.120	-0.007	95.675	0.000
32	-0.106	0.002	97.155	0.000
33	-0.099	-0.045	98.477	0.000
34	0.135	0.003	101.02	0.000
35	-0.128	-0.124	103.38	0.000
36	0.005	-0.051	103.39	0.000
37	-0.056	-0.036	103.86	0.000
38	0.132	-0.159	106.60	0.000
39	-0.070	-0.097	107.41	0.000
40	-0.001	0.072	107.41	0.000
41	0.114	0.089	109.68	0.000
42	-0.077	-0.106	110.75	0.000
43	-0.119	-0.041	113.44	0.000
44	0.075	-0.048	114.53	0.000
45	-0.021	0.003	114.62	0.000
46	-0.063	0.035	115.47	0.000
47	0.110	0.070	118.18	0.000
48	-0.017	-0.064	118.25	0.000
49	0.067	0.093	119.35	0.000
50	-0.051	-0.066	120.02	0.000
51	0.058	-0.090	120.92	0.000
52	-0.044	-0.024	121.48	0.000
53	-0.050	-0.018	122.26	0.000
54	0.001	-0.055	122.26	0.000
55	0.042	-0.060	122.88	0.000
56	-0.006	-0.018	122.89	0.000
57	-0.006	-0.069	122.91	0.000
58	0.049	-0.090	124.00	0.000
59	-0.012	0.018	124.07	0.000
60	0.002	0.039	124.07	0.000

Figure 4.18: Corrlegram of 12th difference of 1st differenced series of India (1 ST & 1 LT)

From the plot of ACF in Figure 4.18, it can be seen that among non seasonal auto correlations in the plot of ACF, the first auto correlation is significantly different from zero and among seasonal auto correlations, the auto correlation at lag 12 is significantly different from zero. Thus, it suggests a seasonal ARIMA model with non seasonal MA (1) and seasonal MA (1) components. Also, the auto correlation at the first lag in the ACF is cut off to zero and PACF decays exponentially towards zero depicts non seasonal MA component of order 1. Thus, it is possible to suggest ARIMA $(0, 1, 1) \times (0, 1, 0)_{12}$ model.

From PACF (Figure 4.18), it can be seen that no seasonal partial auto correlations are significant. Although, it can be seen that first two non seasonal partial auto correlations are significantly different from zero, which suggest models with possible non seasonal AR components of order 1 and 2. There by based on the simultaneous inspection of the plots of ACF and PACF, the following five models were postulated as follows.

$$\text{ARIMA}(0, 1, 1) \times (0, 1, 0)_{12}$$

$$\text{ARIMA}(0, 1, 1) \times (0, 1, 1)_{12}$$

$$\text{ARIMA}(1, 1, 1) \times (0, 1, 0)_{12}$$

$$\text{ARIMA}(2, 1, 0) \times (0, 1, 0)_{12}$$

$$\text{ARIMA}(1, 1, 0) \times (0, 1, 0)_{12}$$

The details of the five models are shown in Table 4.15.

4.12.1 Comparison of parsimonious models for India

Table 4.15: Comparison of postulated models for the India.

Model	Parameter		p value	ARCH-LM test for heteroscedasticity	Breusch-Godfrey test for serial correlation	Normality
1	C MA(1)	40.72326 -0.544581	0.7580 0.0000	0.1683 (0.6830)	0.1844 (0.8320)	0.3042 (0.8589)
2	C MA(1) SMA(12)	26.7612 -0.62887 -0.8306	0.5953 0.0000 0.0000	1.3606 (0.2476)	0.3523 (0.7044)	0.7397 (0.6908)
3	C AR(1) MA(1)	32.48767 0.253759 -0.752254	0.7482 0.2216 0.0000	0.1354 (0.7141)	0.2063 (0.8141)	0.1199 (0.9418)
4	C AR(1) AR(2)	51.02842 -0.481672 -0.284835	0.7576 0.0003 0.0260	0.6878 (0.4100)	0.1391 (0.8704)	0.1630 (0.9217)
5	C AR(1)	60.26742 -0.373646	0.7802 0.0024	0.0303 (0.8623)	2.4310 (0.0960)	0.1747 (0.9163)

According to the results in Table 4.15, the model (3) is not suitable as the coefficient for AR (1) parameter is not significant (p = 0.2216). The parameters of other models except constants are statistically significant as the corresponding p values are less than 0.05. The plots of ACF of residuals (Appendix C) show that residuals of all models except the model (5) are uncorrelated as all auto correlations are within 95 % confidence interval except few auto correlations. Any of the model doesn’t reject the null hypothesis of no serial correlation and the null hypothesis of homoscedasticity as corresponding p values of F statistic are greater than 5%. Furthermore, residuals of all five models do not significantly deviate from normality as corresponding p values are greater than 0.05. Thus, except the model (3) and model (5), the most accurate model among other three models is identified based on the minimum values of AIC and SIC criterion as illustrated in Table 4.16.

Table 4.16: Comparison of suggested ARIMA models

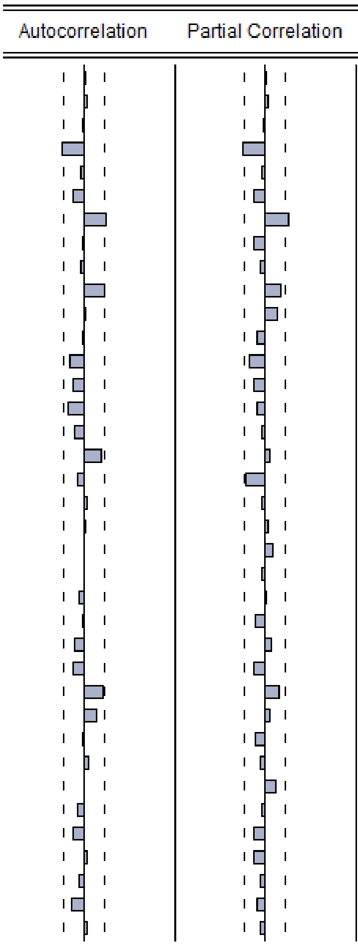
Parameter	AIC	SIC	R square
ARIMA(0, 1, 1)× $(0, 1, 0)_{12}$ (model 1)	18.38794	18.45269	18.69
ARIMA(0, 1, 1)× $(0, 1, 1)_{12}$ (model 2)	18.0204	18.11754	45.3
ARIMA(2, 1, 0)× $(0, 1, 0)_{12}$ (model 4)	18.43031	18.52903	19.52

According to the results in Table 4.16, ARIMA (0, 1, 1) × $(0, 1, 1)_{12}$ model has the lowest values of AIC and SIC criterion (AIC = 18.0204, SIC = 18.45269). Meanwhile, R square of ARIMA (0, 1, 1) × $(0, 1, 1)_{12}$ model is better than the R square of other models. Thus, ARIMA (0, 1, 1) × $(0, 1, 1)_{12}$ is recommended for further analysis.

4.12.2 Estimation of parameters

Table 4.15 shows that the coefficient for constant parameter of the model (1) is not significant (p = 0.5953). The coefficient for MA (1) is significant as the corresponding p value is less than 0.05 (p = 0.000) and the corresponding coefficient for MA (1) parameter is -0.62887. The coefficient of SMA (12) parameter is significant (p = 0.000) and the corresponding value is -0.8306.

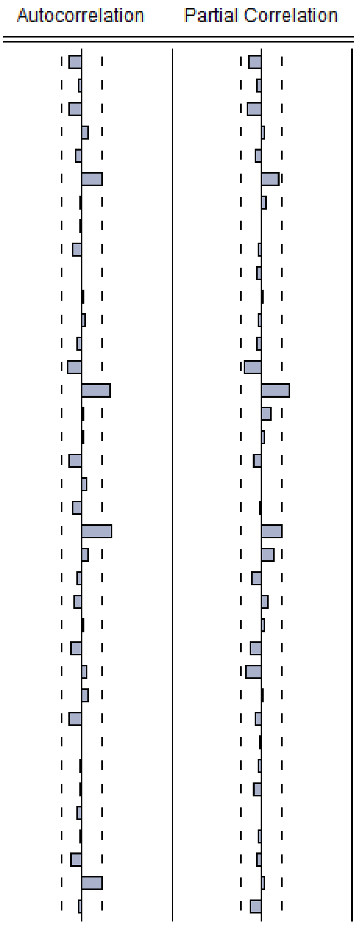
4.12.3 Diagnostic Checking



	AC	PAC	Q-Stat	Prob
1	0.024	0.024	0.0414	
2	0.035	0.035	0.1315	
3	-0.020	-0.022	0.1617	0.688
4	-0.261	-0.262	5.2842	0.071
5	-0.040	-0.030	5.4075	0.144
6	-0.134	-0.121	6.7962	0.147
7	0.255	0.274	11.942	0.036
8	-0.024	-0.116	11.990	0.062
9	-0.027	-0.057	12.049	0.099
10	0.224	0.187	16.201	0.040
11	0.012	0.152	16.213	0.063
12	-0.019	-0.088	16.245	0.093
13	-0.167	-0.183	18.695	0.067
14	-0.126	-0.116	20.111	0.065
15	-0.185	-0.095	23.228	0.039
16	-0.107	-0.043	24.280	0.042
17	0.203	0.049	28.155	0.021
18	-0.073	-0.219	28.661	0.026
19	0.036	-0.034	28.791	0.037
20	0.024	0.039	28.851	0.050
21	0.001	0.087	28.851	0.068
22	0.001	-0.030	28.851	0.091
23	-0.052	0.016	29.134	0.111
24	-0.022	-0.103	29.188	0.140
25	-0.111	0.069	30.554	0.134
26	-0.130	-0.131	32.484	0.115
27	0.207	0.164	37.464	0.052
28	0.144	0.054	39.943	0.040
29	-0.013	-0.108	39.963	0.052
30	0.064	-0.065	40.477	0.060
31	-0.004	0.122	40.479	0.076
32	-0.081	-0.029	41.349	0.081
33	-0.117	-0.121	43.196	0.071
34	0.030	-0.118	43.321	0.087
35	-0.065	-0.051	43.923	0.097
36	-0.155	-0.094	47.494	0.062
37	0.026	-0.053	47.598	0.076

Figure 4.19: ACF for residuals of India

Results of Figure 4.19 show that all the auto correlations are approximately within 95% confidence intervals. Thus, it can be concluded that residuals are random. The high probability values of the L-jung Box statistic also confirmed that residuals are random.



	AC	PAC	Q-Stat	Prob
1	-0.141	-0.141	1.4393	0.230
2	-0.030	-0.051	1.5063	0.471
3	-0.143	-0.158	3.0156	0.389
4	0.074	0.028	3.4244	0.489
5	-0.072	-0.075	3.8173	0.576
6	0.222	0.195	7.6627	0.264
7	-0.017	0.052	7.6854	0.361
8	-0.014	-0.002	7.7002	0.463
9	-0.101	-0.044	8.5373	0.481
10	-0.007	-0.055	8.5414	0.576
11	0.018	0.019	8.5690	0.662
12	0.025	-0.035	8.6225	0.735
13	-0.050	-0.062	8.8424	0.785
14	-0.162	-0.190	11.173	0.672
15	0.325	0.336	20.743	0.145
16	0.014	0.100	20.760	0.188
17	0.010	0.028	20.770	0.237
18	-0.150	-0.089	22.933	0.193
19	0.046	-0.007	23.140	0.231
20	-0.102	-0.011	24.180	0.235
21	0.344	0.236	36.223	0.021
22	0.068	0.132	36.706	0.025
23	-0.062	-0.111	37.121	0.032
24	-0.095	0.066	38.110	0.034
25	0.019	0.039	38.152	0.045
26	-0.124	-0.119	39.903	0.040
27	0.050	-0.177	40.201	0.049
28	0.079	0.012	40.937	0.054
29	-0.144	-0.067	43.480	0.041
30	-0.001	-0.014	43.480	0.053
31	-0.015	-0.028	43.509	0.067
32	-0.015	-0.094	43.540	0.084
33	-0.062	0.007	44.056	0.095
34	-0.019	-0.047	44.105	0.115
35	-0.139	-0.061	46.900	0.086
36	0.233	0.037	54.988	0.022
37	-0.038	-0.122	55.213	0.027

Figure 4.20: Correlogram of squared residuals.

Correlogram of residual squared of the fitted model in Figure 4.20 shows that all the auto correlations except few auto correlations are between 95% confidence intervals which depicts that residuals of mean equation have constant variance. The high probability values of the L-jung Box statistic also confirmed the constant variance of residuals. The results of the ARCH LM test are indicated in Table 4.17.

Table 4.17: Results of ARCH LM test

Heteroskedasticity Test: ARCH

F-statistic	1.360587	Prob. F(1,66)	0.2476
Obs*R-squared	1.373503	Prob. Chi-Square(1)	0.2412

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 04/12/19 Time: 13:52
Sample (adjusted): 2012M01 2017M08
Included observations: 68 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4100012.	740142.0	5.539493	0.0000
RESID^2(-1)	-0.142896	0.122505	-1.166442	0.2476

R-squared	0.020199	Mean dependent var	3594849.
Adjusted R-squared	0.005353	S.D. dependent var	4962761.
S.E. of regression	4949460.	Akaike info criterion	33.69643
Sum squared resid	1.62E+15	Schwarz criterion	33.76171
Log likelihood	-1143.678	Hannan-Quinn criter.	33.72229
F-statistic	1.360587	Durbin-Watson stat	1.997152
Prob(F-statistic)	0.247633		

The results in Table 4.17 confirm that there is no significant ARCH effect in the model estimated above as the corresponding p value of F statistic is greater than 0.05 ($p = 0.2476$). The results of the Breusch Godfrey test are indicated in Table 4.18.

Table 4.18: Results of Breusch Godfrey test of serial correlation

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.352263	Prob. F(2,64)	0.7044
Obs*R-squared	0.568294	Prob. Chi-Square(2)	0.7527

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 04/12/19 Time: 13:51
Sample: 2011M12 2017M08
Included observations: 69
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.680703	51.08433	0.111202	0.9118
MA(1)	-0.166099	0.223326	-0.743752	0.4597
SMA(12)	-0.002964	0.035750	-0.082904	0.9342
RESID(-1)	0.194077	0.256598	0.756344	0.4522
RESID(-2)	0.147479	0.192333	0.766792	0.4460

R-squared	0.008236	Mean dependent var	-98.07575
Adjusted R-squared	-0.053749	S.D. dependent var	1907.875
S.E. of regression	1958.478	Akaike info criterion	18.06743
Sum squared resid	2.45E+08	Schwarz criterion	18.22932
Log likelihood	-618.3262	Hannan-Quinn criter.	18.13165
F-statistic	0.132873	Durbin-Watson stat	1.976170
Prob(F-statistic)	0.969749		

The non significance of Breusch Godfrey test of serial correlation ($p = 0.7527$) in Table 4.18 confirms that there is no significant serial correlation in the residuals. The normality plot is presented in Figure 4.21.

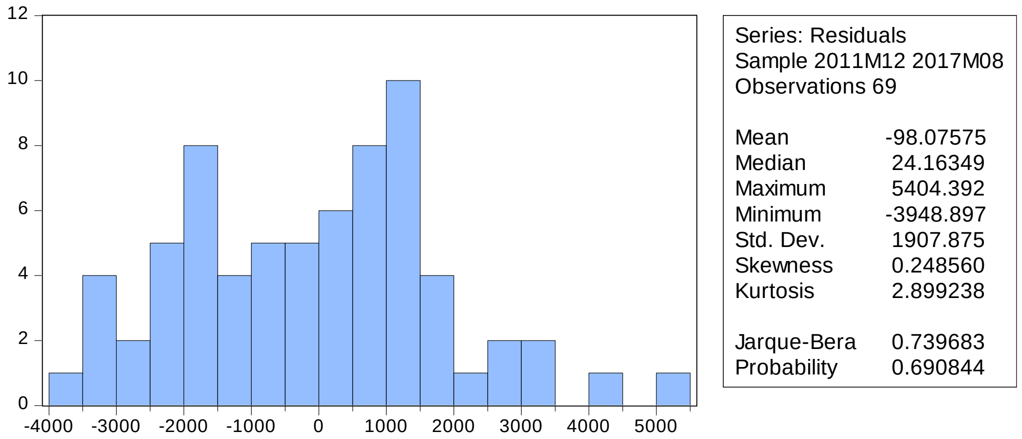


Figure 4.21: The normality plot of residuals.

According to the results of the Jarque-Bera test in Figure 4.21, residuals are not deviated from a normal distribution as the corresponding p value of the Jarque-Bera test is greater than 5% (p = 0.6908, JB= 0.7396). Therefore, according to the results of diagnostic tests of residuals, it can be concluded that residuals of fitted model are distributed normally and independently with constant variance. Thus, residuals are white noise.

4.12.4 Forecasting monthly tourist arrivals from India for an independent set

The best fitted ARIMA (0, 1, 1) × (0, 1, 1) $_{12}$ model for India can be written as

$$(1-B)(1-B^{12})y_t = 26.76120 + (1-0.830598B^{12})(1-0.628868B)e_t$$

The above equation can be expanded as follows.

$$y_t - B^{12}y_t - By_t + B^{13}y_t = 26.76 + e_t - 0.628868Be_t - 0.830598B^{12}e_t + 0.522336B^{13}e_t$$

$$y_t - y_{t-12} - y_{t-1} + y_{t-13} = 26.76 + e_t - 0.628868e_{t-1} - 0.830598e_{t-12} + +0.522336e_{t-13}$$

$$y_t = 26.76 + y_{t-12} + y_{t-1} - y_{t-13} + e_t - 0.628868e_{t-1} - 0.830598e_{t-12}$$

$$+0.522336e_{t-13} \qquad \text{(B)}$$

The forecast values can be calculated from the formula (B) using past values (Appendix A) and residuals (Appendix E).

y_{83} = the forecast for September, 2017 (Table 4.20)

$$y_{83} = 26.76 + y_{71} + y_{82} - y_{70} + e_{83} - 0.628868e_{82} - 0.830598e_{71} + 0.522336e_{70}$$

=31516

Similarly, monthly values were predicted up to February, 2018 and values are as follows.

y_{84} = value for September, 2017

= 37449

y_{85} = value for November, 2017

= 34061

y_{86} = value for December, 2017

=38748

y_{87} = value for January, 2018

=34298

y_{88} = value for February, 2018

=31303

4.12.5 Forecasting for the training set

Table 4.19: Forecasts for the training set (From November, 2010 to August, 2017)

Time period	India series	Forecasts	Percentage error (%)
2010M11	15550		
2010M12	15753		
2011M01	13786		
2011M02	10071		
2011M03	13619		
2011M04	14705		
2011M05	16649		
2011M06	12927		
2011M07	12587		
2011M08	12857		
2011M09	13329		
2011M10	17915		
2011M11	15474		
2011M12	17455	19387	-11.07
2012M01	14615	15402	-5.38
2012M02	11342	11898	-4.9
2012M03	14391	14018	2.59
2012M04	15432	13689	11.3
2012M05	15888	18456	-16.16
2012M06	13758	13991	-1.69
2012M07	11564	13755	-18.94
2012M08	11242	14400	-28.09
2012M09	13888	16039	-15.49
2012M10	17654	22008	-24.66
2012M11	17625	18742	-6.34
2012M12	18941	22333	-17.91
2013M01	16938	18374	-8.48
2013M02	14429	14897	-3.25
2013M03	15281	17044	-11.54
2013M04	12657	16741	-32.27
2013M05	17834	21536	-20.76
2013M06	15297	17097	-11.77

Table 4.19 (Continued)

Time period	India series	Forecasts	Percentage error (%)
2013M07	14783	16888	-14.24
2013M08	18999	17560	7.57
2013M09	18389	19226	-4.55
2013M10	21833	25221	-15.52
2013M11	19796	21982	-11.04
2013M12	22559	25600	-13.48
2014M01	17886	21668	-21.15
2014M02	17534	18218	-3.9
2014M03	18734	20391	-8.84
2014M04	17192	20115	-17
2014M05	22981	24936	-8.51
2014M06	19911	20524	-3.08
2014M07	19277	20342	-5.52
2014M08	17912	21041	-17.47
2014M09	19244	22734	-18.13
2014M10	26148	28756	-9.97
2014M11	19762	25543	-29.26
2014M12	26153	29188	-11.6
2015M01	22944	25283	-10.19
2015M02	19999	21859	-9.3
2015M03	21838	24059	-10.17
2015M04	23048	23810	-3.31
2015M05	31764	28658	9.78
2015M06	25860	24273	6.14
2015M07	24681	24117	2.28
2015M08	23540	24843	-5.54
2015M09	27233	26562	2.46
2015M10	30574	32611	-6.66
2015M11	29329	29426	-0.33
2015M12	35437	33097	6.6
2016M01	28895	29219	-1.12
2016M02	26559	25822	2.77
2016M03	30170	28049	7.03
2016M04	25890	27827	-7.48

Table 4.19 (Continued)

Time period	India series	Forecasts	Percentage error (%)
2016M05	34259	32701	4.55
2016M06	27068	28343	-4.71
2016M07	27665	28214	-1.98
2016M08	24418	28967	-18.63
2016M09	27241	30712	-12.74
2016M10	36471	36788	-0.87
2016M11	30148	33629	-11.55
2016M12	37945	37327	1.63
2017M01	31173	33476	-7.39
2017M02	26320	30106	-14.38
2017M03	27075	32359	-19.52
2017M04	26323	32164	-22.19
2017M05	34167	37065	-8.48
2017M06	27836	32734	-17.59
2017M07	29006	32631	-12.5
2017M08	31220	33411	-7.02

The error percentages for training set vary between -32.27 % to 11.30 % and MAPE value for the model is 10.58%. The error percentages for independent set vary between -5.64 % to 9.59 % with the exception of the forecast in December and the corresponding value of MAPE is 8.14% (Table 4.20).

Table 4.20: Comparison of actual values and forecasts for independent set

Month	Actuals	Forecasts	95% Confidence Interval		Error %	MAPE
September,2017	34481	31516	27516	35516	8.60	8.14
October, 2017	36996	37449	33101	41797	-1.23	
November, 2017	32243	34061	29437	38686	-5.64	
December, 2017	47788	38748	33810	43687	18.92	
January, 2018	37936	34298	29028	39568	9.59	
February, 2018	32914	31303	25714	36891	4.90	

The plot of forecasts of monthly tourist arrivals from the India for the period from September, 2017 to February, 2018 is shown in Figure 4.22.

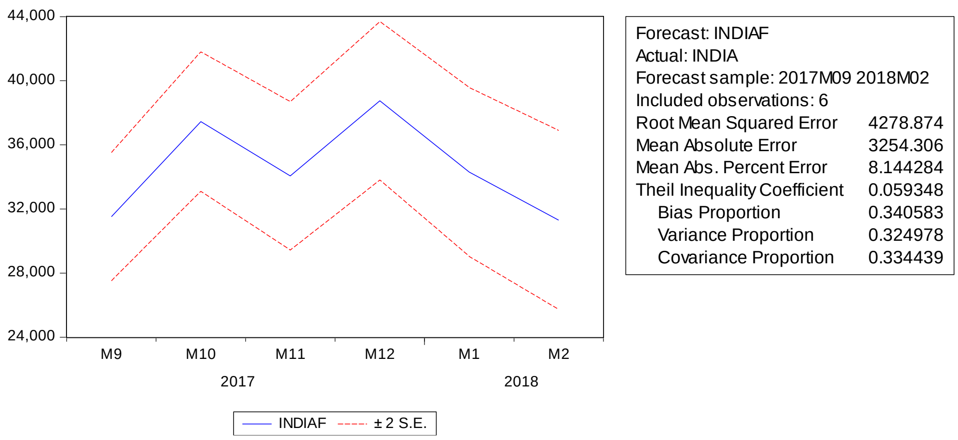


Figure 4.22 :Forecasts of monthly tourist arrivals from India

4.13 Summary of Chapter 4

The best fitted model among suggested ARIMA models for the UK is as follows.

$$(1-B)(1-B^{12})y_t = 9.035 + (1-0.414994B^{12})(1-0.985634B)e_t$$

The best fitted model among suggested ARIMA models for India is as follows.

$$(1-B)(1-B^{12})y_t = 26.761 + (1-0.830598B^{12})(1-0.628868B)e_t$$

The errors of both models found to be white noise. The models were tested for the training data set as well for the independent data set. The percentage errors for the independent set in both models vary between -12 % to 12%. The error percentages for the independent set in UK model vary between -11.88% to 11.75% and the error percentages for the independent set in India vary between -5.64 % to 9.59 %. MAPE values for training set and validation set were found to be as illustrated in Table 4.21.

Table 4.21: Value of MAPE criterion

	UK		India	
	Training set	Validation set	Training set	Validation set
MAPE	13.43	7.89	10.58	8.14

CHAPTER 5

DEVELOPMENT OF SMOOTHING TECHNIQUES

Seasonal ARIMA models are somewhat complicated to identify and estimate in many business environments (Athanasopulos and De Silva, 2010). However, exponential smoothing technique is simple to implement and generates relatively accurate forecasts. The major advantages of the exponential smoothing method are simplicity, easily understood, less computing requirements and low data storage requirements (Zhi-Peng, *et al*., 2008). Accordingly, in this study Holt Winters method of exponential smoothing technique also applied to capture the trend and seasonality in the monthly tourist arrivals from the UK and India and to forecast.

5.1 Use of Holt Winters Multiplicative Model for UK

To find the optimum values of smoothing parameters, a simulation study with 27 trials were conducted with varying the smoothing constants from 0.1 to 0.3 using MINITAB software. The results of Holt Winters Multiplicative model for the arrivals from the UK are illustrated in Table 5.1.

Table 5.1: Determination of smoothing constants of the model for UK

value of smoothing parameter			MAPE criteria
α	β	γ	
0.1	0.1	0.1	9.2356
0.1	0.1	0.2	9.4293
0.1	0.1	0.3	9.5653
0.1	0.2	0.1	9.1726
0.1	0.2	0.2	9.4009
0.1	0.2	0.3	9.5251
0.1	0.3	0.1	9.0888
0.1	0.3	0.2	9.2837
0.1	0.3	0.3	9.3822
0.2	0.1	0.1	8.8078
0.2	0.1	0.2	9.2100
0.2	0.1	0.3	9.3962
0.2	0.2	0.1	8.9237
0.2	0.2	0.2	9.3838
0.2	0.2	0.3	9.6023
0.2	0.3	0.1	9.1971
0.2	0.3	0.2	9.6819
0.2	0.3	0.3	9.9325
0.3	0.1	0.1	8.7993
0.3	0.1	0.2	9.3050
0.3	0.1	0.3	9.6785
0.3	0.2	0.1	9.1345
0.3	0.2	0.2	9.6384
0.3	0.2	0.3	10.0344
0.3	0.3	0.1	9.4735
0.3	0.3	0.2	10.0434
0.3	0.3	0.3	10.5452

According to the results in Table 5.1, it can be concluded that MAPE is minimum at $\alpha: 0.3,\ \beta: 0.1$ and $\gamma = 0.1$. Thus, the Holt Winters Multiplicative model for UK series can be written as follows.

$$a_t = 0.3 + 0.7(a_{t-1} + b_{t-1})$$

$$b_t = 0.1 + 0.9b_{t-1}$$

$$s_t = 0.1\frac{Y_t}{a_t} + 0.9\, s_{t-12}$$

$$\hat{y}_{T+\tau} = (a_T + \tau b_T)s_{T+\tau-s}$$

5.1.1 Diagnostic tests for Holt Winters multiplicative model for UK

The plot of ACF, the normality probability plot of residuals and the plot of standardized residuals versus standardized predicted values are carried out to check the model adequacy. The ACF plot of residuals is shown in Figure 5.1.

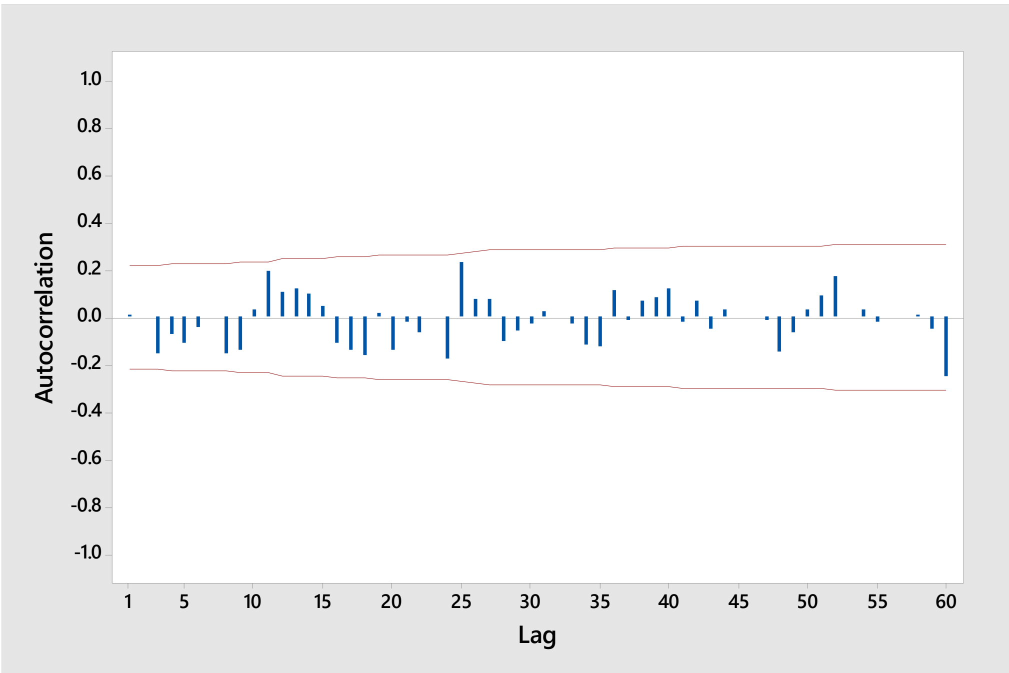


Figure 5.1: ACF of residuals for UK (with 5% significance limits for the auto correlations)

According to the plot of ACF in Figure 5.1, it can be clearly seen that all the auto correlation estimates are within 95% confidence intervals. Thus, it can be concluded that the residuals are not significantly deviated from random. The normality plot of residuals of the selected model is shown in Figure 5.2.

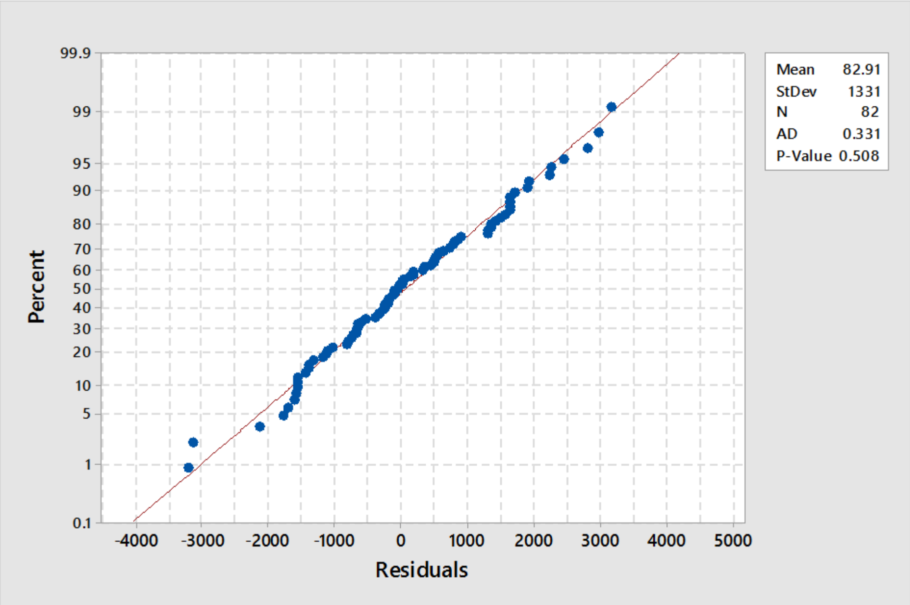


Figure 5.2: Normality plot of residuals for UK model

Results in Figure 5.2 confirm that residuals are not significantly deviated from a normal distribution as the corresponding p value of the AD statistic is greater than 5% ($p = 0.508$, $AD = 0.331$).

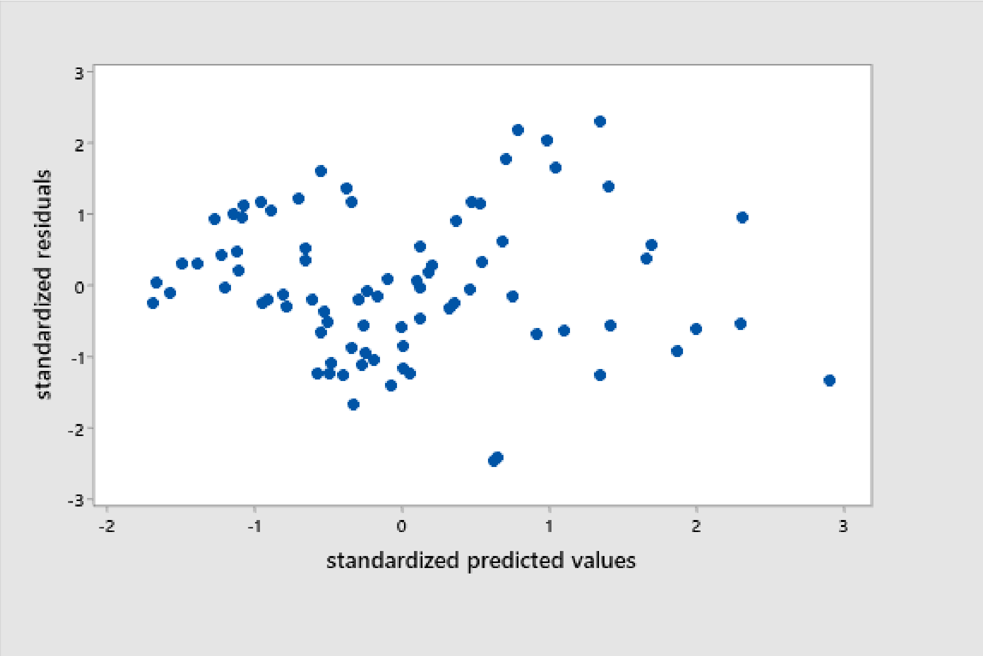


Figure 5.3: The plot of standardized residuals vs standardized fitted values

The plot of the standardized residuals verses standardized predicted values doesn't exhibit any systematic pattern. It seems that the plot is symmetric about zero and within ± 3 limits. Hence, the variance of the residuals is constant. Therefore, it can be concluded that residuals are independent and normally distributed with homoscedasticity. Table 5.2 brings out the forecasts obtained from the MINITAB software for training set using Holt Winters Multiplicative model with smoothing constants $\alpha = 0.3,\ \beta = 0.1$ and $\gamma = 0.1$.

Table 5.2: Forecasts for training set using Holt Winters multiplicative method

	UK	Smoothing values	Level	Trend	Seasonal	Forecasts	Percentage error (%)
2010:m11	9788	8388.590	9947.740	-53.249	0.890	8297	15.23
2010:m12	10176	12017.700	9453.130	-97.385	1.195	11953	-17.47
2011:m1	9518	10025.300	9241.450	-108.810	1.057	9922	-4.24
2011:m2	9614	10841.700	8851.330	-136.950	1.164	10714	-11.44
2011:m3	8852	10626.300	8312.090	-177.180	1.187	10462	-18.19
2011:m4	9038	7637.620	8645.280	-126.140	0.932	7475	17.3
2011:m5	4452	4788.460	8374.750	-140.580	0.552	4719	-5.99
2011:m6	5188	5344.570	8202.740	-143.720	0.638	5255	-1.29
2011:m7	12003	12027.500	8097.130	-139.910	1.468	11817	1.55
2011:m8	12486	10788.100	8381.510	-97.481	1.348	10602	15.09
2011:m9	7871	6644.160	8777.560	-48.127	0.803	6567	16.57
2011:m10	7408	6815.150	8972.950	-23.776	0.781	6778	8.51
2011:m11	9589	7981.610	9498.410	31.148	0.902	7960	16.98
2011:m12	10063	11349.800	9197.140	-2.094	1.185	11387	-13.16
2012:m1	8162	9725.680	8752.070	-46.392	1.045	9723	-19.13
2012:m2	8746	10191.500	8347.200	-82.239	1.153	10137	-15.91
2012:m3	12032	9907.920	8826.480	-26.087	1.205	9810	18.46
2012:m4	8019	8221.990	8742.850	-31.842	0.930	8198	-2.23
2012:m5	4940	4823.020	8784.170	-24.525	0.553	4805	2.72
2012:m6	6076	5600.850	8990.570	-1.433	0.641	5585	8.08
2012:m7	13643	13197.100	9080.690	7.723	1.471	13195	3.28
2012:m8	11558	12241.400	8934.010	-7.718	1.343	12252	-6
2012:m9	8586	7175.050	9455.660	45.219	0.814	7169	16.51
2012:m10	8767	7388.130	10016.700	96.805	0.791	7423	15.33
2012:m11	10828	9030.300	10682.700	153.723	0.913	9118	15.8
2012:m12	12861	12657.300	10841.900	154.268	1.185	12839	0.17
2013:m1	11350	11329.600	10955.800	150.227	1.044	11491	-1.24
2013:m2	13604	12629.700	11314.500	171.074	1.158	12803	5.89
2013:m3	13590	13629.300	11424.400	164.963	1.203	13835	-1.81
2013:m4	8642	10625.700	10900.100	96.031	0.916	10779	-24.73
2013:m5	6567	6024.740	11261.600	122.582	0.556	6078	7.45
2013:m6	7642	7223.500	11543.100	138.478	0.643	7302	4.45
2013:m7	16424	16983.900	11525.900	122.908	1.467	17188	-4.65
2013:m8	15519	15475.100	11621.800	120.201	1.342	15640	-0.78

Table 5.2: (Continued)

	UK	Smoothing values	Level	Trend	Seasonal	Forecasts	Percentage error (%)
2013:m9	9356	9455.550	11669.200	112.924	0.812	9553	-2.11
2013:m10	9850	9227.220	11984.500	133.165	0.794	9317	5.42
2013:m11	9663	10938.600	11658.500	87.242	0.904	11060	-14.46
2013:m12	15209	13815.000	12072.400	119.916	1.192	13918	8.49
2014:m1	12896	12604.600	12240.100	124.692	1.045	12730	1.29
2014:m2	14316	14171.000	12365.000	124.708	1.158	14315	0
2014:m3	11823	14876.200	11690.900	44.834	1.184	15026	-27.09
2014:m4	12380	10713.100	12268.000	98.060	0.926	10754	13.13
2014:m5	6918	6818.160	12390.600	100.508	0.556	6873	0.66
2014:m6	7790	7973.200	12375.500	88.952	0.642	8038	-3.18
2014:m7	16692	18151.200	12139.300	56.437	1.458	18282	-9.52
2014:m8	15532	16289.800	12009.400	37.802	1.337	16366	-5.37
2014:m9	8983	9756.730	11750.200	8.097	0.808	9787	-8.96
2014:m10	10112	9327.860	12052.200	37.487	0.798	9334	7.69
2014:m11	10730	10899.300	12022.300	30.748	0.903	10933	-1.89
2014:m12	15996	14336.100	12461.400	71.586	1.202	14373	10.15
2015:m1	13410	13022.500	12622.700	80.562	1.047	13097	2.33
2015:m2	15212	14614.000	12834.100	93.640	1.161	14707	3.32
2015:m3	16191	15194.400	13152.200	116.085	1.189	15305	5.47
2015:m4	11233	12174.100	12928.400	82.101	0.920	12282	-9.33
2015:m5	7954	7188.480	13398.900	120.941	0.560	7234	9.05
2015:m6	8580	8603.260	13472.700	116.226	0.642	8681	-1.18
2015:m7	20643	19636.900	13761.100	133.447	1.462	19806	4.05
2015:m8	17908	18399.300	13744.300	118.420	1.334	18578	-3.74
2015:m9	11160	11100.300	13849.400	117.084	0.807	11196	-0.32
2015:m10	9970	11056.900	13522.900	72.730	0.792	11150	-11.84
2015:m11	10822	12213.300	13111.700	24.333	0.895	12279	-13.46
2015:m12	18762	15754.700	13879.500	98.686	1.217	15784	15.87
2016:m1	16253	14528.600	14442.800	145.146	1.055	14632	9.97
2016:m2	19194	16761.000	15173.400	203.687	1.171	16929	11.8
2016:m3	21430	18035.500	16172.700	283.252	1.202	18278	14.71
2016:m4	12006	14878.200	15434.400	181.091	0.906	15139	-26.09
2016:m5	8412	8639.890	15439.000	163.445	0.558	8741	-3.91
2016:m6	9406	9905.080	15320.000	135.204	0.639	10010	-6.42
2016:m7	23948	22394.700	15733.500	163.027	1.468	22592	5.66

Table 5.2 (Continued)

	UK	Smoothing values	Level	Trend	Seasonal	Forecasts	Percentage error (%)
2016:m8	20475	20982.700	15733.400	146.715	1.330	21200	-3.54
2016:m9	12288	12703.900	15681.500	126.861	0.805	12822	-4.35
2016:m10	10964	12423.800	15217.600	67.778	0.785	12524	-14.23
2016:m11	13337	13625.500	15168.400	56.079	0.894	13686	-2.62
2016:m12	20446	18453.800	15698.900	103.521	1.225	18522	9.41
2017:m1	19468	16556.300	16599.600	183.241	1.066	16666	14.4
2017:m2	20218	19437.300	16927.900	197.744	1.173	19652	2.8
2017:m3	19451	20351.900	16841.500	169.333	1.198	20590	-5.85
2017:m4	17841	15254.200	17816.800	249.933	0.915	15408	13.64
2017:m5	8520	9946.960	17225.000	165.757	0.552	10086	-18.39
2017:m6	10424	11003.400	17068.900	133.575	0.636	11109	-6.57
2017:m7	23553	25054.200	16855.600	98.887	1.461	25250	-7.21
2017:m8	21903	22424.900	16807.200	84.152	1.328	22556	-2.98

The error percentages for the training data set of UK vary between -27.09 % to 18.46%. The value of MAPE criterion is 8.7 (Table 5.2). The error percentages for independent set vary between -11.83 % to 19.91 %. and MAPE value is 10.727% (Table 5.3).

Table 5.3: Comparison of actual values versus forecasts for arrivals from UK

Period	Actuals	Forecasts	95% Confidence Interval		Percentage error (%)	MAPE
September, 2017	12593	13599	11022	16176	-7.99	10.727
October, 2017	12518	13327	10658	15997	-6.46	
November, 2017	13634	15247	12471	18024	-11.83	
December, 2017	21756	21004	18108	23900	3.46	
January, 2018	22940	18373	15346	21399	19.91	
February, 2018	23817	20312	17145	23480	14.72	

5.2 Use of Holt Winters Additive Model for UK

Table 5.4: Determination of smoothing constants of additive model for UK

value of smoothing parameter			MAPE criteria
α	β	γ	Additive model
0.1	0.1	0.1	13.5507
0.1	0.1	0.2	13.6888
0.1	0.1	0.3	13.6355
0.1	0.2	0.1	13.2937
0.1	0.2	0.2	13.3686
0.1	0.2	0.3	13.2486
0.1	0.3	0.1	13.2552
0.1	0.3	0.2	13.2369
0.1	0.3	0.3	13.0889
0.2	0.1	0.1	13.1604
0.2	0.1	0.2	13.3365
0.2	0.1	0.3	13.3926
0.2	0.2	0.1	13.2771
0.2	0.2	0.2	13.3571
0.2	0.2	0.3	13.4698
0.2	0.3	0.1	13.5349
0.2	0.3	0.2	13.6012
0.2	0.3	0.3	13.8471
0.3	0.1	0.1	13.2988
0.3	0.1	0.2	13.4415
0.3	0.1	0.3	13.6249
0.3	0.2	0.1	13.6335
0.3	0.2	0.2	13.7844
0.3	0.2	0.3	13.9502
0.3	0.3	0.1	13.9918
0.3	0.3	0.2	14.2070
0.3	0.3	0.3	14.4473

The above simulation confirmed that MAPE is minimum at $\alpha: 0.1$, $\beta: 0.3$ and $\gamma = 0.3$. Thus, the Holt Winters Additive model for UK can be written as follows.

$$a_t = 0.1(Y_t - s_{t-p}) + 0.9(a_{t-1} + b_{t-1})$$

$$b_t = 0.3(a_t - a_{t-1}) + 0.7b_{t-1}$$

$$s_t = 0.3(Y_t - a_t) + 0.3s_{t-12}$$

Forecast for time period $T + \tau$:

$$\hat{y}_{T+\tau} = a_T + \tau b_T + s_{T+\tau-s}$$

5.2.1 Diagnostic tests of residuals for Holt Winters additive model for UK

The plot of ACF, the normality probability plot of residuals and the plot of standardized residuals versus standardized fitted values are carried out to check the model adequacy of Holt Winters additive model. The ACF plot of residuals is shown on Figure 5.4.

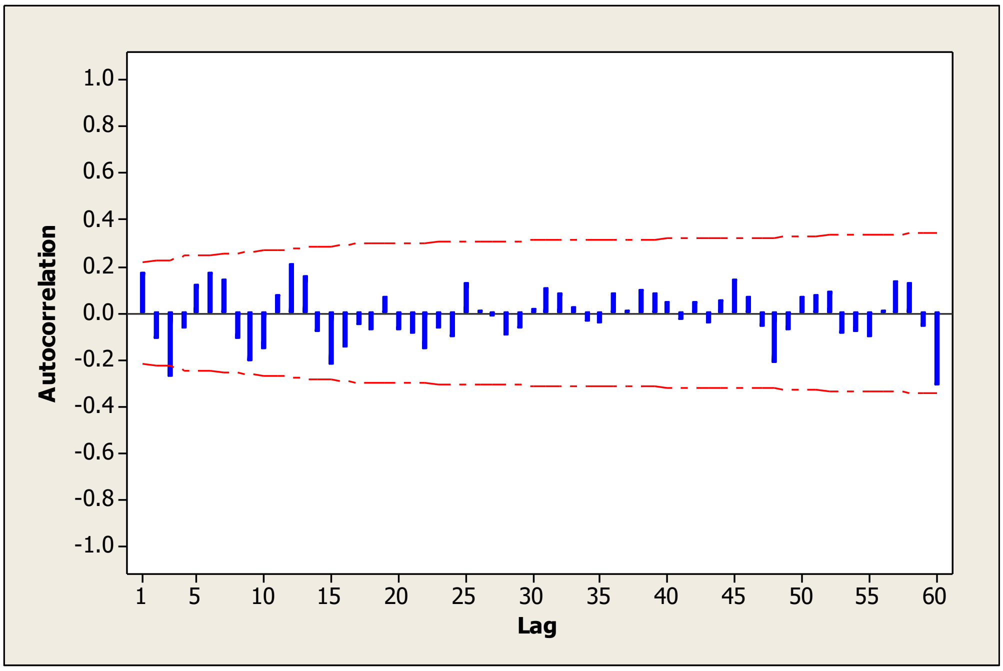


Figure 5.4: ACF of residuals for model of UK(with 5% significance limits for the autocorrelations)

According to the plot of ACF in Figure 5.4, it can be clearly seen that all the auto correlation estimates except third auto correlation are within 95% confidence intervals. Thus, residuals are random. The normality plot of residuals of the selected model is shown in Figure 5.5.

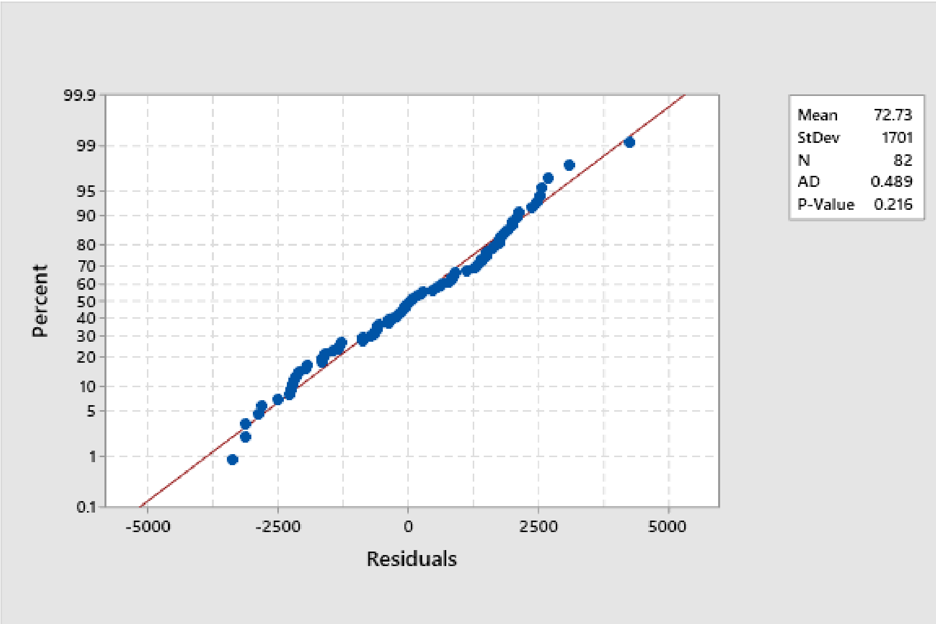


Figure 5.5: Normality plot of residuals for model of UK

Results in Figure 5.5 confirm that residuals are not deviated from a normal distribution as the corresponding p value of the AD statistic is greater than 5% ($p = 0.216$, $AD = 0.489$). The plot of the standardized residuals verses standardized predicted values is used to check the constant variance of residuals and relevant graph is shown in Figure 5.6.

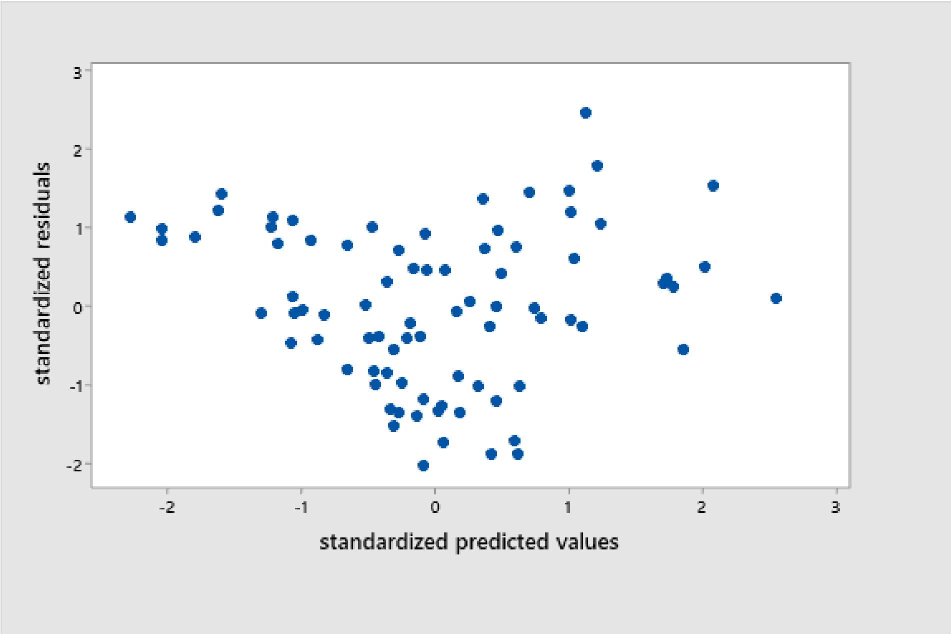


Figure 5.6: The plot of standardized residuals vs standardized predicted values

The plot of the standardized residuals versus standardized predicted values doesn't exhibit any systematic pattern. It seems that the plot is symmetric about zero and within ± 2 limits indicating constant variance of the residuals. Thus, above results concluded that the residuals are independent and normally distributed with homoscedasticity. Thus, Holt Winters Additive model with smoothing constants α: 0.1, β: 0.3 and $\gamma = 0.3$ is advocated to forecast the monthly tourist arrivals from the UK. Table 5.5 brings out the forecasts of the training set using Holt Winters Additive model with smoothing constants α: 0.1, β: 0.3 and $\gamma = 0.3$

Table 5.5: Forecasts for the training set for Holt Winters additive model (UK)

	UK	Smoothing values	Level	Trend	Seasonal	Forecasts	Percentage error (%)
2010:m11	9788	8384.666	9589.644	-58.909	-751.348	8281	15.40
2010:m12	10176	12424.050	9311.821	-124.583	2243.339	12365	-21.51
2011:m1	9518	10251.570	9126.339	-142.853	775.324	10127	-6.40
2011:m2	9614	11357.580	8823.414	-190.874	1799.041	11215	-16.65
2011:m3	8852	11291.570	8407.671	-258.335	1861.004	11101	-25.40
2011:m4	9038	7302.021	8348.767	-198.506	-567.184	7044	22.07
2011:m5	4452	2642.746	8351.037	-138.273	-5163.930	2444	45.10
2011:m6	5188	3578.788	8387.512	-85.848	-4300.430	3441	33.68
2011:m7	12003	13757.180	8134.831	-135.898	4919.217	13671	-13.90
2011:m8	12486	11742.840	8086.838	-109.527	3845.355	11607	7.04
2011:m9	7871	5495.323	8225.832	-34.970	-1920.510	5386	31.57
2011:m10	7408	5323.588	8402.800	28.611	-2330.010	5289	28.61
2011:m11	9589	7651.452	8622.305	85.879	-235.935	7680	19.91
2011:m12	10063	10865.640	8619.332	59.223	2003.438	10952	-8.83
2012:m1	8162	9394.655	8549.367	20.467	426.516	9454	-15.83
2012:m2	8746	10348.410	8407.547	-28.219	1360.865	10369	-18.56
2012:m3	12032	10268.550	8558.494	25.531	2344.755	10240	14.89
2012:m4	8019	7991.310	8584.241	25.596	-566.601	8017	0.03
2012:m5	4940	3420.316	8759.246	70.418	-4760.520	3446	30.24
2012:m6	6076	4458.818	8984.340	116.821	-3882.800	4529	25.46
2012:m7	13643	13903.560	9063.424	105.500	4817.325	14020	-2.77
2012:m8	11558	12908.780	9023.296	61.812	3452.160	13014	-12.60
2012:m9	8586	7102.785	9227.248	104.454	-1536.730	7165	16.55
2012:m10	8767	6897.237	9508.232	157.413	-1853.380	7002	20.14
2012:m11	10828	9272.297	9805.474	199.362	141.603	9430	12.91
2012:m12	12861	11808.910	10090.110	224.943	2233.674	12008	6.63
2013:m1	11350	10516.630	10375.900	243.196	590.793	10742	5.36
2013:m2	13604	11736.760	10781.500	291.918	1799.357	11980	11.94
2013:m3	13590	13126.250	11090.600	297.073	2391.149	13418	1.26
2013:m4	8642	10524.000	11169.760	231.701	-1154.950	10821	-25.21
2013:m5	6567	6409.241	11394.070	229.482	-4780.490	6641	-1.13
2013:m6	7642	7511.267	11613.680	226.520	-3909.460	7741	-1.29
2013:m7	16424	16431.000	11816.840	219.514	4754.274	16658	-1.42
2013:m8	15519	15269.000	12039.410	220.429	3460.390	15489	0.20

Table 5.5 (Continued)

	UK	Smoothing values	Level	Trend	Seasonal	Forecasts	Percentage error (%)
2013:m9	9356	10502.680	12123.130	179.416	-1905.850	10723	-14.61
2013:m10	9850	10269.750	12242.620	161.441	-2015.150	10449	-6.08
2013:m11	9663	12384.230	12115.800	74.961	-636.717	12546	-29.83
2013:m12	15209	14349.470	12269.220	98.498	2445.507	14424	5.16
2014:m1	12896	12860.010	12361.460	96.623	573.916	12959	-0.48
2014:m2	14316	14160.820	12463.940	98.379	1815.167	14257	0.41
2014:m3	11823	14855.090	12249.270	4.465	1545.923	14953	-26.48
2014:m4	12380	11094.320	12381.860	42.902	-809.023	11099	10.35
2014:m5	6918	7601.374	12352.130	21.113	-4976.580	7644	-10.50
2014:m6	7790	8442.670	12305.870	0.900	-4091.390	8464	-8.65
2014:m7	16692	17060.140	12269.860	-10.172	4654.633	17061	-2.21
2014:m8	15532	15730.250	12240.880	-15.814	3409.608	15720	-1.21
2014:m9	8983	10335.040	12091.450	-55.901	-2266.630	10319	-14.87
2014:m10	10112	10076.300	12044.710	-53.152	-1990.420	10020	0.91
2014:m11	10730	11407.990	11929.070	-71.898	-805.424	11355	-5.82
2014:m12	15996	14374.580	12026.510	-21.098	2902.703	14303	10.59
2015:m1	13410	12600.420	12088.480	3.822	798.199	12579	6.19
2015:m2	15212	13903.640	12222.750	42.958	2167.392	13907	8.58
2015:m3	16191	13768.670	12503.650	114.339	2188.352	13812	14.70
2015:m4	11233	11694.620	12560.390	97.060	-964.533	11809	-5.13
2015:m5	7954	7583.809	12684.760	105.254	-4902.840	7681	3.43
2015:m6	8580	8593.377	12778.150	101.695	-4123.420	8699	-1.38
2015:m7	20643	17432.790	13190.700	194.951	5493.933	17534	15.06
2015:m8	17908	16600.310	13496.930	228.333	3710.047	16795	6.21
2015:m9	11160	11230.300	13695.400	219.374	-2347.260	11459	-2.68
2015:m10	9970	11704.980	13719.340	160.744	-2518.090	11924	-19.60
2015:m11	10822	12913.910	13654.810	93.164	-1413.640	13075	-20.82
2015:m12	18762	16557.520	13959.110	156.504	3472.759	16651	11.25
2016:m1	16253	14757.310	14249.530	196.679	1159.780	14914	8.24
2016:m2	19194	16416.920	14704.250	274.091	2864.099	16614	13.44
2016:m3	21430	16892.600	15404.670	401.990	3339.445	17167	19.89
2016:m4	12006	14440.140	15523.050	316.906	-1730.290	14842	-23.62
2016:m5	8412	10620.220	15587.440	241.153	-5584.620	10937	-30.02
2016:m6	9406	11464.030	15598.680	172.177	-4744.200	11705	-24.44

Table 5.5 (Continued)

	UK	Smoothing values	Level	Trend	Seasonal	Forecasts	Percentage error (%)
2016:m7	23948	21092.610	16039.180	252.674	6218.400	21265	11.20
2016:m8	20475	19749.230	16339.160	266.867	3837.785	20002	2.31
2016:m9	12288	13991.900	16408.950	207.744	-2879.370	14259	-16.04
2016:m10	10964	13890.860	16303.240	113.706	-3364.440	14099	-28.59
2016:m11	13337	14889.590	16250.310	63.717	-1863.540	15003	-12.49
2016:m12	20446	19723.070	16379.950	83.493	3650.747	19787	3.22
2017:m1	19468	17539.730	16647.920	138.836	1657.870	17623	9.48
2017:m2	20218	19512.020	16843.470	155.851	3017.228	19651	2.81
2017:m3	19451	20182.920	16910.540	129.218	3099.748	20339	-4.56
2017:m4	17841	15180.260	17292.920	205.164	-1046.780	15309	14.19
2017:m5	8520	11708.300	17158.730	103.360	-6500.850	11913	-39.83
2017:m6	10424	12414.540	17052.700	40.543	-5309.550	12518	-20.09
2017:m7	23553	23271.100	17117.380	47.783	6283.565	23312	1.02
2017:m8	21903	20955.170	17255.170	74.785	4080.799	21003	4.11

The error percentages for training set vary between -39.83 % to 45.10 % and the MAPE value is 13.08 % (Table 5.5). The error percentages for the independent data of the UK vary between -14.75 % to 15.92 %. The corresponding value of MAPE criterion is 12.15 (Table 5.6).

Table 5.6: Comparison of actual values and forecasts for independent data

Period	Actuals	Forecasts	95% Confidence Interval		Error %	MAPE
September, 2017	12593	14451	17886	11015	-14.75	12.15
October, 2017	12518	14040	17599	10482	-12.16	
November, 2017	13634	15616	19317	11915	-14.54	
December, 2017	21756	21205	25065	17345	2.53	
January, 2018	22940	19287	23322	15252	15.92	
February, 2018	23817	20721	24944	16499	13	

5.3 Use of Holt Winters Multiplicative Model for India

Table 5.7: Determination of smoothing constants of multiplicative model for India

value of smoothing parameter			MAPE criteria
α	β	γ	
0.1	0.1	0.1	7.3200
0.1	0.1	0.2	7.3886
0.1	0.1	0.3	7.4573
0.1	0.2	0.1	7.0219
0.1	0.2	0.2	7.1233
0.1	0.2	0.3	7.1967
0.1	0.3	0.1	7.0164
0.1	0.3	0.2	7.0834
0.1	0.3	0.3	7.1223
0.2	0.1	0.1	6.7210
0.2	0.1	0.2	6.8190
0.2	0.1	0.3	6.8911
0.2	0.2	0.1	6.7634
0.2	0.2	0.2	6.8616
0.2	0.2	0.3	6.9978
0.2	0.3	0.1	6.9356
0.2	0.3	0.2	7.0488
0.2	0.3	0.3	7.1829
0.3	0.1	0.1	6.6387
0.3	0.1	0.2	6.7615
0.3	0.1	0.3	6.8947
0.3	0.2	0.1	6.7958
0.3	0.2	0.2	6.9102
0.3	0.2	0.3	7.0072
0.3	0.3	0.1	7.0100
0.3	0.3	0.2	7.1210
0.3	0.3	0.3	7.2545

The optimum values for the smoothing parameters of Multiplicative model regarding to India are $\alpha: 0.3$, $\beta: 0.1$ and $\gamma = 0.1$(MAPE=6.63867). Thus, the Holt Winters Multiplicative model for India can be written as follows.

$$a_t = 0.3 + 0.7(a_{t-1} + b_{t-1})$$

$$b_t = 0.1 + 0.9b_{t-1}$$

$$s_t = 0.1\frac{Y_t}{a_t} + 0.9\, s_{t-12}$$

Forecast for time peiriod $T + \tau$:

$$\hat{y}_{T+\tau} = (a_T + \tau b_T)s_{T+\tau-s}$$

5.3.1: Diagnostic tests of residuals for Holt Winters multiplicative model for India

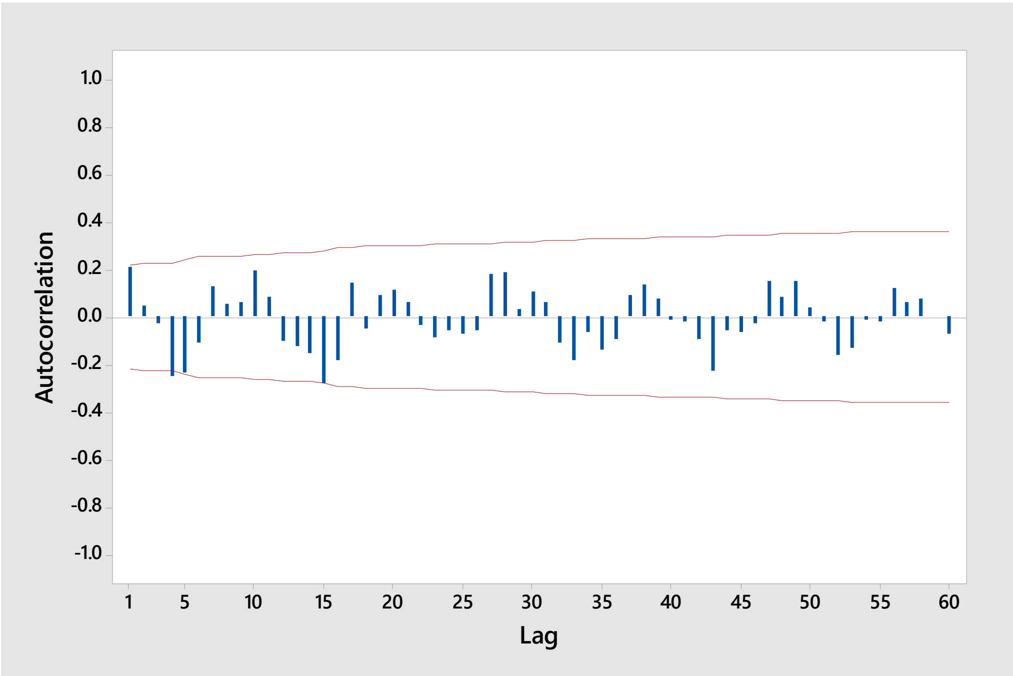


Figure 5.7: ACF of residuals for model of India

According to the plot of ACF in Figure 5.7, it can be clearly seen that all the auto correlation estimates are within 95% confidence intervals. Thus, residuals are random. The normality plot for residuals of the selected model is shown in Figure 5.8.

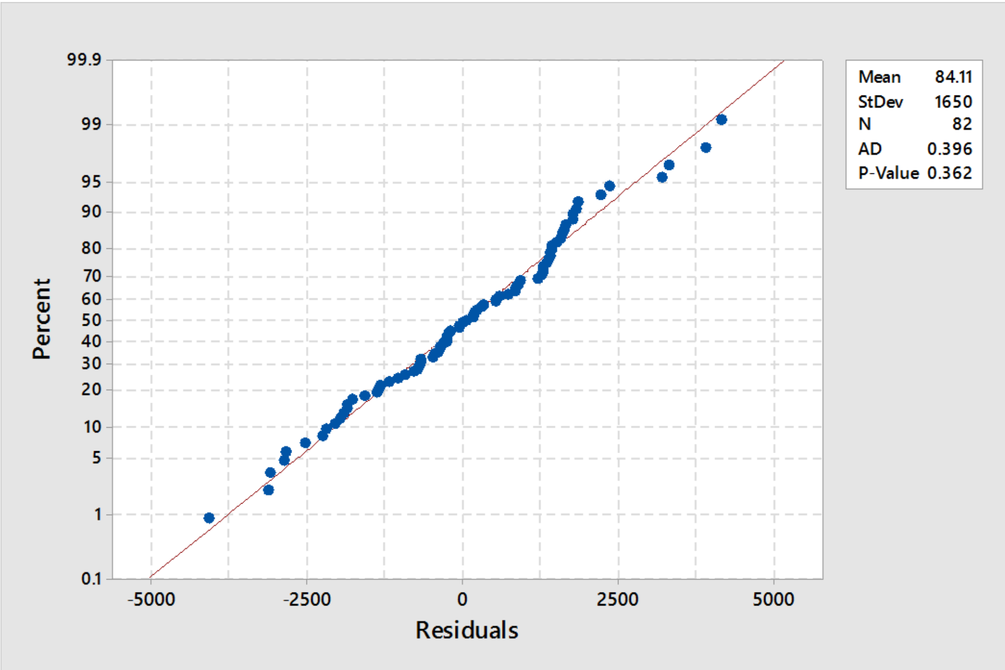


Figure 5.8: Normality plot of residuals for model of India

Results in Figure 5.8 confirm that residuals are not significantly deviated from normal distribution as the corresponding p value of the AD statistic is greater than 5% ($p = 0.362$, $AD = 0.396$).

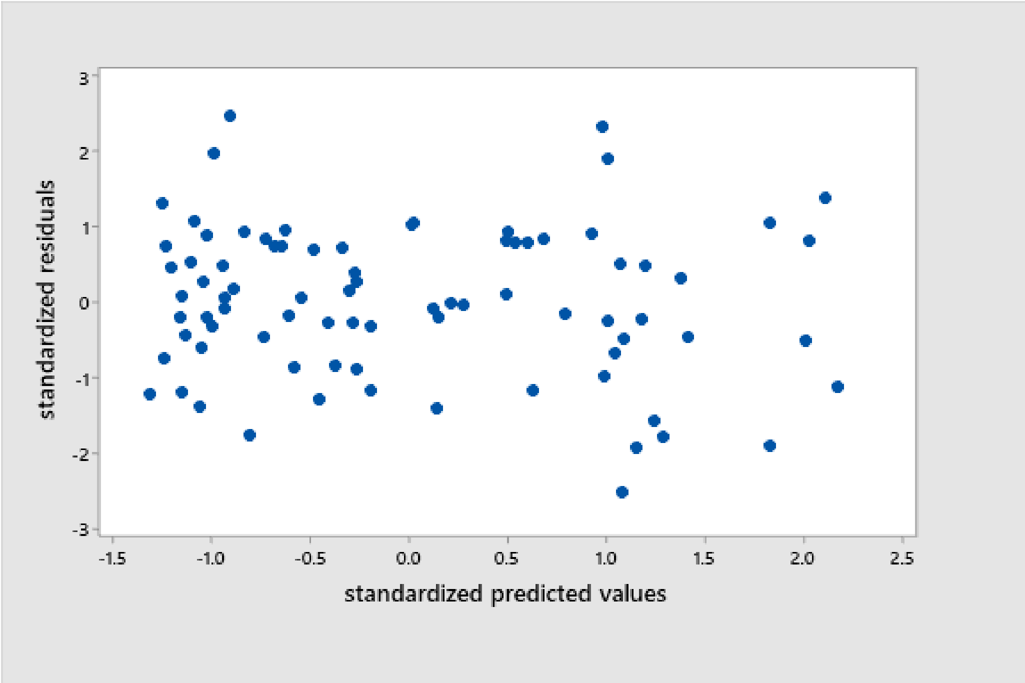


Figure 5.9: The plot of standardized residuals vs standardized predicted values

The plot of standardized residuals versus standardized predicted values doesn't exhibit any systematic pattern. It seems that the plot is symmetric about zero and within ± 3 limits implying the constant variance of the residuals. Thus, above results concluded that residuals are independent and normally distributed with homoscedasticity. Table 5.8 brings out the forecasts for the training set using Holt Winters Multiplicative model with smoothing constants α: 0.3, β: 0.1 and $\gamma = 0.1$.

Table 5.8: Forecasts for training set using Holt Winters multiplicative model for India

	India	Smoothing values	Level	Trend	Seasonal	Forecasts	Percentage error (%)
2010:m11	15550	14000.330	14402.020	80.612	1.013	14036	9.74
2010:m12	15753	17000.540	14141.390	46.488	1.174	17096	-8.52
2011:m1	13786	13993.390	14111.060	38.806	0.988	14039	-1.84
2011:m2	10071	12003.660	13456.640	-30.517	0.840	12037	-19.52
2011:m3	13619	12805.680	13691.670	-3.962	0.956	12777	6.19
2011:m4	14705	12484.280	14419.550	69.222	0.923	12481	15.13
2011:m5	16649	16787.240	14432.380	63.583	1.163	16868	-1.31
2011:m6	12927	13789.500	14206.080	34.594	0.951	13850	-7.14
2011:m7	12587	13224.020	14025.000	13.027	0.928	13256	-5.32
2011:m8	12857	13100.870	13955.790	4.804	0.933	13113	-1.99
2011:m9	13329	13124.230	14024.480	11.192	0.941	13129	1.5
2011:m10	17915	16615.780	14361.300	43.754	1.191	16629	7.18
2011:m11	15474	14554.920	14663.980	69.648	1.018	14599	5.65
2011:m12	17455	17212.300	14774.760	73.761	1.175	17294	0.92
2012:m1	14615	14601.560	14830.480	71.956	0.988	14674	-0.41
2012:m2	11342	12463.990	14480.340	29.747	0.835	12524	-10.43
2012:m3	14391	13842.220	14673.390	46.077	0.958	13871	3.62
2012:m4	15432	13537.860	15321.550	106.285	0.931	13580	12
2012:m5	15888	17821.080	14897.360	53.238	1.153	17945	-12.95
2012:m6	13758	14165.990	14805.910	38.769	0.949	14217	-3.33
2012:m7	11564	13732.930	14131.530	-32.546	0.917	13769	-19.07
2012:m8	11242	13182.230	13484.760	-93.968	0.923	13152	-16.99
2012:m9	13888	12694.740	13799.240	-53.123	0.948	12606	9.23
2012:m10	17654	16435.420	14068.990	-20.836	1.197	16372	7.26
2012:m11	17625	14317.420	15029.460	77.295	1.033	14296	18.89
2012:m12	18941	17652.760	15412.610	107.880	1.180	17744	6.32
2013:m1	16938	15227.610	16007.480	156.579	0.995	15334	9.47
2013:m2	14429	13361.680	16500.690	190.242	0.839	13492	6.49
2013:m3	15281	15814.490	16466.860	167.835	0.955	15997	-4.68
2013:m4	12657	15331.850	15722.490	76.614	0.918	15488	-22.37
2013:m5	17834	18135.490	15697.700	66.474	1.152	18224	-2.19
2013:m6	15297	14893.010	15871.980	77.255	0.950	14956	2.23
2013:m7	14783	14548.400	16002.850	82.616	0.917	14619	1.11
2013:m8	18999	14769.180	17435.620	217.631	0.940	14845	21.86

Table 5.8 (Continued)

	India	Smoothing values	Level	Trend	Seasonal	Forecasts	Percentage error (%)
2013:m9	18389	16527.490	18177.090	270.016	0.954	16734	9.00
2013:m10	21833	21765.540	18383.010	263.606	1.196	22089	-1.17
2013:m11	19796	18992.620	18800.810	279.025	1.035	19265	2.68
2013:m12	22559	22184.620	19091.310	280.173	1.180	22514	0.20
2014:m1	17886	18996.050	18952.750	238.299	0.990	19275	-7.76
2014:m2	17534	15895.440	19705.670	289.762	0.844	16095	8.21
2014:m3	18734	18826.240	19879.540	278.173	0.954	19103	-1.97
2014:m4	17192	18258.730	19725.840	234.985	0.914	18514	-7.69
2014:m5	22981	22718.960	19958.580	234.761	1.152	22990	-0.04
2014:m6	19911	18965.470	20421.430	257.569	0.953	19189	3.63
2014:m7	19277	18733.090	20779.600	267.630	0.918	18969	1.60
2014:m8	17912	19524.220	20452.180	208.125	0.933	19776	-10.40
2014:m9	19244	19517.310	20511.950	193.289	0.953	19716	-2.45
2014:m10	26148	24541.340	21050.110	227.776	1.201	24773	5.26
2014:m11	19762	21789.790	20621.870	162.174	1.027	22026	-11.45
2014:m12	26153	24336.850	21197.070	203.477	1.186	24528	6.21
2015:m1	22944	20982.560	21933.950	256.817	0.995	21184	7.67
2015:m2	19999	18507.840	22643.880	302.129	0.848	18725	6.37
2015:m3	21838	21603.890	22928.990	300.426	0.954	21892	-0.25
2015:m4	23048	20951.960	23827.430	360.228	0.919	21226	7.90
2015:m5	31764	27442.200	25205.350	461.997	1.163	27857	12.30
2015:m6	25860	24013.590	26110.160	506.279	0.956	24454	5.44
2015:m7	24681	23978.570	26694.020	514.037	0.919	24444	0.96
2015:m8	23540	24911.040	26613.090	454.540	0.928	25391	-7.86
2015:m9	27233	25353.740	27523.050	500.082	0.956	25787	5.31
2015:m10	30574	33055.600	27253.230	423.092	1.193	33656	-10.08
2015:m11	29329	28001.480	27937.010	449.161	1.030	28436	3.04
2015:m12	35437	33119.690	28837.830	494.327	1.190	33652	5.04
2016:m1	28895	28707.990	29240.220	485.133	0.995	29200	-1.06
2016:m2	26559	24788.060	30206.520	533.249	0.851	25199	5.12
2016:m3	30170	28814.200	31006.190	559.891	0.956	29323	2.81
2016:m4	25890	28498.630	30546.660	457.950	0.912	29013	-12.06
2016:m5	34259	35512.220	30543.830	411.872	1.158	36045	-5.21
2016:m6	27068	29214.810	30158.800	332.182	0.951	29609	-9.39

Table 5.8 (Continued)

	India	Smoothing values	Level	Trend	Seasonal	Forecasts	Percentage error (%)
2016:m7	27665	27715.470	30374.860	320.569	0.918	28021	-1.29
2016:m8	24418	28198.160	29377.670	188.793	0.919	28496	-16.7
2016:m9	27241	28095.550	29241.760	156.323	0.954	28276	-3.8
2016:m10	36471	34888.290	29749.150	191.430	1.196	35075	3.83
2016:m11	30148	30632.470	29742.000	171.572	1.028	30830	-2.26
2016:m12	37945	35388.380	30506.710	230.886	1.195	35593	6.2
2017:m1	31173	30347.070	30917.410	248.867	0.996	30577	1.91
2017:m2	26320	26307.300	31096.100	241.849	0.850	26519	-0.76
2017:m3	27075	29722.240	30434.510	151.505	0.949	29953	-10.63
2017:m4	26323	27755.360	30069.380	99.842	0.908	27894	-5.97
2017:m5	34167	34834.290	29966.460	79.566	1.157	34950	-2.29
2017:m6	27836	28485.840	29817.070	56.670	0.949	28561	-2.61
2017:m7	29006	27376.980	30389.010	108.197	0.922	27429	5.44
2017:m8	31220	27916.030	31543.740	212.851	0.926	28015	10.26

The error percentages for training set vary between -22.37 % to 21.86 % and the corresponding value of MAPE criterion is 6.63 (Table 5.8). According to Table 5.9, the error percentages for the independent data set of India vary between -3.38 % to 18.98 %, whereas Mean Absolute Percentage Error (MAPE) is 11.116 (Table 5.9).

Table 5.9: Comparison of actual values versus forecasts for India

Period	Actual	Forecast	95% Confidence Interval		Error%	MAPE
September, 2017	34,481	30292	27072	33512	12.15	11.1162
October, 2017	36,996	38248	34913	41583	-3.38	
November, 2017	32,243	33086	29618	36555	-2.62	
December, 2017	47,788	38720	35102	42338	18.98	
January,2018	37,936	32481	28700	36263	14.38	
February, 2018	32,914	27912	23955	31870	15.2	

5.4 Use of Holt Winters Additive Model for India

Table 5.10: Determination of smoothing constants of additive model for India

value of smoothing parameter			MAPE criteria
α	β	γ	
0.1	0.1	0.1	7.3328
0.1	0.1	0.2	7.4052
0.1	0.1	0.3	7.4501
0.1	0.2	0.1	7.0023
0.1	0.2	0.2	7.1477
0.1	0.2	0.3	7.2657
0.1	0.3	0.1	7.0264
0.1	0.3	0.2	7.1542
0.1	0.3	0.3	7.2641
0.2	0.1	0.1	6.7165
0.2	0.1	0.2	6.7481
0.2	0.1	0.3	6.8464
0.2	0.2	0.1	6.7916
0.2	0.2	0.2	6.8715
0.2	0.2	0.3	6.9506
0.2	0.3	0.1	6.8935
0.2	0.3	0.2	6.9942
0.2	0.3	0.3	7.0353
0.3	0.1	0.1	6.7269
0.3	0.1	0.2	6.7581
0.3	0.1	0.3	6.7678
0.3	0.2	0.1	6.8613
0.3	0.2	0.2	6.8743
0.3	0.2	0.3	6.8837
0.3	0.3	0.1	7.0053
0.3	0.3	0.2	7.0164
0.3	0.3	0.3	7.0099

According to the results of MAPE criteria in Table 5.10, the least value of MAPE criteria was obtained at $\alpha: 0.2$, $\beta: 0.1$ and $\gamma = 0.1$. Thus, the Holt Winters Additive model for India can be written as follows.

$$a_t = 0.2\left(Y_t - s_{t-p}\right) + 0.8(a_{t-1} + b_{t-1})$$

$$b_t = 0.1(a_t - a_{t-1}) + 0.9b_{t-1}$$

$$s_t = 0.1(Y_t - a_t) + 0.9s_{t-12}$$

Forecast for time period $T + \tau$:

$$\hat{y}_{T+\tau} = a_T + \tau b_T + s_{T+\tau-s}$$

5.4.1 Diagnostic tests of residuals for Holt Winters additive model for India

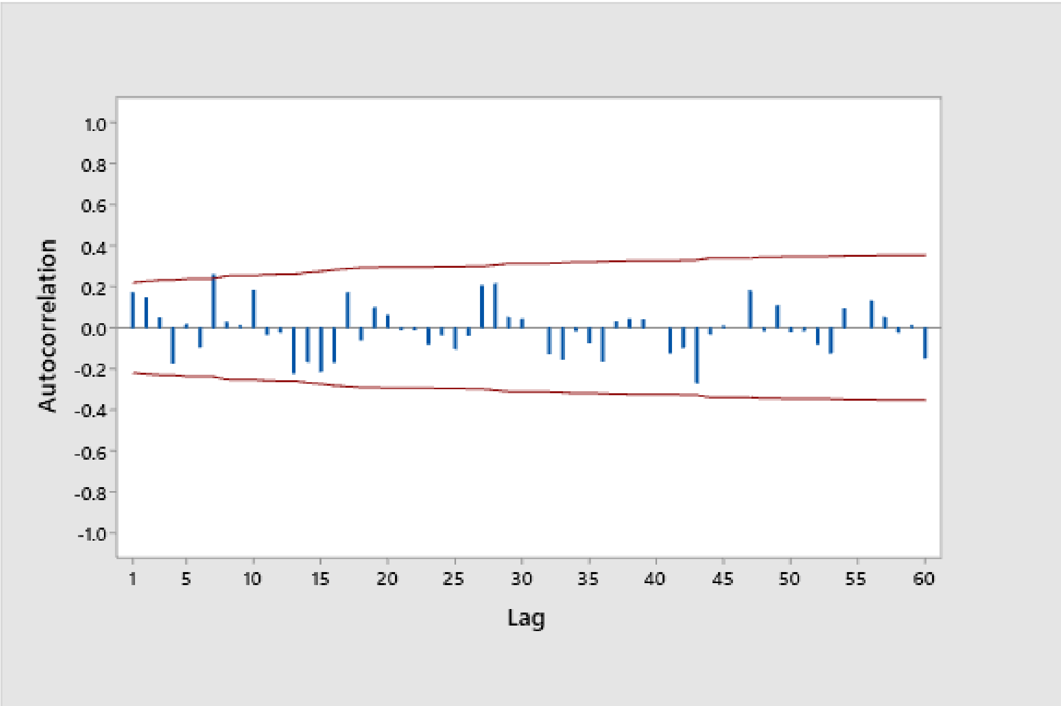


Figure 5.10: ACF of residuals for model of India

According to the plot of ACF in Figure 5.10, it can be seen that all the auto correlation estimates are within 95% confidence intervals. Thus, residuals are random. The normality plot for residuals of the selected model is shown in Figure 5.11.

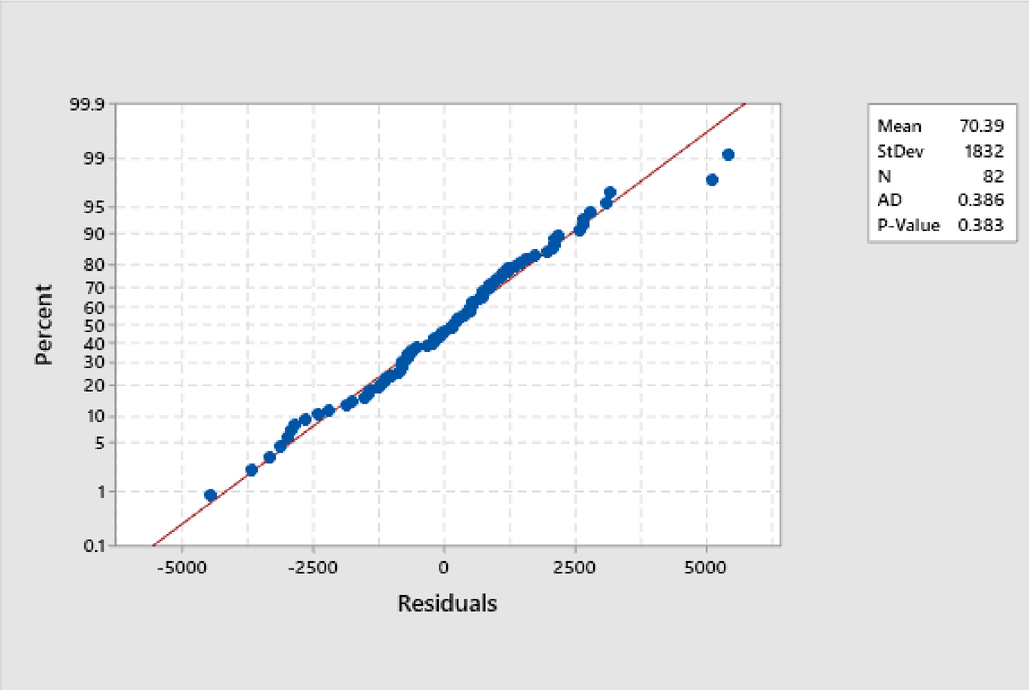


Figure 5.11: Normality plot of residuals for model of India

Results in Figure 5.11 confirm that residuals are not significantly deviated from normality as the corresponding p value of the AD statistic is greater than 0.05 at 5% significance level (p = 0.383, AD = 0.386).

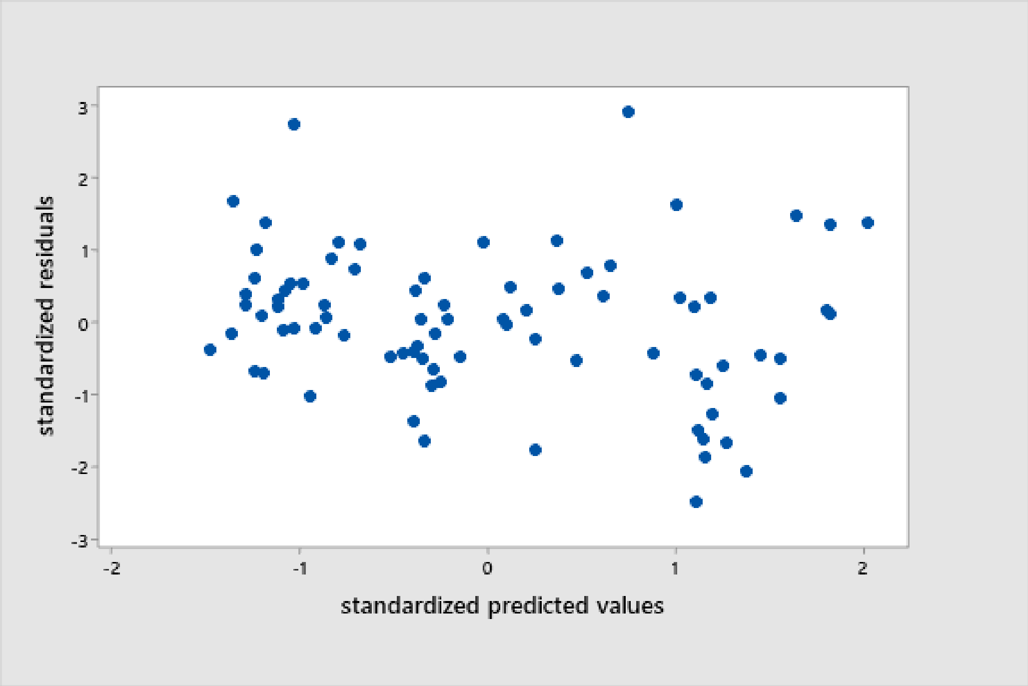


Figure 5.12: The plot of standardized residuals vs standardized predicted values

The plot of standardized residuals verses standardized predicted values in Figure 5.12 doesn't exhibit any systematic pattern. It seems that the plot is symmetric about zero and within ± 3 limits indicating the constant variance of residuals. Therefore, it can be concluded that the residuals are independent and normally distributed with homoscedasticity. Table 5.11 brings out the forecasts for out of sample using Holt Winters Additive model with smoothing constants α: 0.2, β: 0.1 and $\gamma = 0.1$.

Table 5.11: Forecasts for training set using Holt Winters additive model for India

	India	Smoothing values	Level	Trend	Seasonal	Forecasts	Percentage errors (%)
2010:m11	15550	14984.370	14056.620	46.072	1111.658	15020	3.41
2010:m12	15753	18664.440	13511.190	-13.078	4371.211	18711	-18.77
2011:m1	13786	13862.570	13485.420	-14.348	346.299	13849	-0.46
2011:m2	10071	10726.510	13342.840	-27.171	-2810.200	10712	-6.37
2011:m3	13619	12450.350	13554.830	-3.255	-796.822	12423	8.78
2011:m4	14705	11569.480	14179.330	59.521	-1734.250	11566	21.34
2011:m5	16649	17409.120	14074.920	43.128	3164.218	17469	-4.92
2011:m6	12927	12636.990	14167.430	48.065	-1418.180	12680	1.91
2011:m7	12587	12031.920	14316.890	58.206	-2094.940	12080	4.03
2011:m8	12857	12015.100	14531.840	73.880	-2239.100	12073	6.1
2011:m9	13329	13368.400	14583.060	71.614	-1172.500	13442	-0.85
2011:m10	17915	18375.880	14548.180	60.964	3750.218	18448	-2.97
2011:m11	15474	15659.840	14559.780	56.028	1091.914	15721	-1.59
2011:m12	17455	18930.990	14309.410	25.388	4248.650	18987	-8.78
2012:m1	14615	14655.710	14321.580	24.066	341.012	14681	-0.45
2012:m2	11342	11511.370	14306.950	20.197	-2825.680	11535	-1.71
2012:m3	14391	13510.130	14499.290	37.411	-727.968	13530	5.98
2012:m4	15432	12765.040	15062.610	90.002	-1523.880	12802	17.04
2012:m5	15888	18226.820	14666.840	41.425	2969.912	18317	-15.29
2012:m6	13758	13248.660	14801.850	50.783	-1380.750	13290	3.4
2012:m7	11564	12706.910	14613.900	26.910	-2190.440	12758	-10.32
2012:m8	11242	12374.800	14408.860	3.715	-2331.880	12402	-10.32
2012:m9	13888	13236.360	14542.160	16.674	-1120.670	13240	4.67
2012:m10	17654	18292.380	14427.830	3.573	3697.814	18309	-3.71
2012:m11	17625	15519.740	14851.740	45.606	1260.049	15523	11.92
2012:m12	18941	19100.390	14856.340	41.507	4232.250	19146	-1.08
2013:m1	16938	15197.360	15237.680	75.489	476.943	15239	10.03
2013:m2	14429	12412.000	15701.470	114.320	-2670.360	12487	13.46
2013:m3	15281	14973.500	15854.430	118.183	-712.514	15088	1.26
2013:m4	12657	14330.540	15614.260	82.349	-1667.220	14449	-14.16
2013:m5	17834	18584.180	15530.110	65.698	2903.310	18667	-4.67
2013:m6	15297	14149.360	15812.190	87.337	-1294.190	14215	7.07
2013:m7	14783	13621.750	16114.310	108.815	-2104.530	13709	7.26
2013:m8	18999	13782.440	17244.680	210.970	-1923.260	13891	26.88

Table 5.11 (Continued)

	India	Smoothing values	Level	Trend	Seasonal	Forecasts	Percentage errors(%)
2013:m9	18389	16124.010	17866.450	252.051	-956.346	16335	11.17
2013:m10	21833	21564.270	18121.840	252.384	3699.148	21816	0.08
2013:m11	19796	19381.890	18406.570	255.619	1272.987	19634	0.82
2013:m12	22559	22638.820	18595.100	248.910	4205.415	22894	-1.49
2014:m1	17886	19072.040	18557.020	220.211	362.147	19321	-8.02
2014:m2	17534	15886.660	19062.660	248.754	-2556.190	16107	8.14
2014:m3	18734	18350.140	19338.430	251.456	-701.705	18599	0.72
2014:m4	17192	17671.210	19443.750	236.842	-1725.670	17923	-4.25
2014:m5	22981	22347.060	19760.010	244.784	2935.078	22584	1.73
2014:m6	19911	18465.820	20244.880	268.792	-1198.160	18711	6.03
2014:m7	19277	18140.350	20687.240	286.149	-2035.100	18409	4.5
2014:m8	17912	18763.980	20745.760	263.387	-2014.310	19050	-6.35
2014:m9	19244	19789.420	20847.390	247.210	-1021.050	20053	-4.2
2014:m10	26148	24546.540	21365.450	274.296	3807.489	24794	5.18
2014:m11	19762	22638.440	21009.600	211.281	1020.929	22913	-15.94
2014:m12	26153	25215.010	21366.220	225.815	4263.552	25426	2.78
2015:m1	22944	21728.370	21790.000	245.611	441.332	21954	4.31
2015:m2	19999	19233.810	22139.530	256.003	-2514.620	19479	2.6
2015:m3	21838	21437.820	22424.360	258.886	-690.171	21694	0.66
2015:m4	23048	20698.690	23101.340	300.695	-1558.440	20958	9.07
2015:m5	31764	26036.410	24487.410	409.233	3369.229	26337	17.09
2015:m6	25860	23289.250	25328.950	452.463	-1025.240	23698	8.36
2015:m7	24681	23293.850	25968.350	471.157	-1960.320	23746	3.79
2015:m8	23540	23954.040	26262.460	453.453	-2085.120	24425	-3.76
2015:m9	27233	25241.410	27023.540	484.216	-898.000	25695	5.65
2015:m10	30574	30831.030	27359.510	469.391	3748.189	31315	-2.42
2015:m11	29329	28380.440	27924.730	478.974	1059.262	28850	1.63
2015:m12	35437	32188.290	28957.660	534.369	4485.131	32667	7.82
2016:m1	28895	29398.990	29284.350	513.602	358.263	29933	-3.59
2016:m2	26559	26769.730	29653.090	499.115	-2572.570	27283	-2.73
2016:m3	30170	28962.920	30293.800	513.274	-633.534	29462	2.35
2016:m4	25890	28735.360	30135.350	446.102	-1827.130	29249	-12.97
2016:m5	34259	33504.580	30643.110	452.268	3393.895	33951	0.9
2016:m6	27068	29617.870	30494.950	392.225	-1265.410	30070	-11.09

Table 5.11 (Continued)

	India	Smoothing values	Level	Trend	Seasonal	Forecasts	Percentage errors (%)
2016:m7	27665	28534.630	30634.810	366.988	-2061.270	28927	-4.56
2016:m8	24418	28549.680	30102.060	277.015	-2445.020	28917	-18.42
2016:m9	27241	29204.060	29931.060	232.213	-1077.210	29481	-8.22
2016:m10	36471	33679.250	30675.180	283.404	3952.952	33911	7.02
2016:m11	30148	31734.440	30584.620	246.007	909.675	32018	-6.2
2016:m12	37945	35069.750	31356.470	298.592	4695.471	35316	6.93
2017:m1	31173	31714.740	31487.000	281.786	291.037	32013	-2.7
2017:m2	26320	28914.430	31193.540	224.261	-2802.670	29196	-10.93
2017:m3	27075	30560.010	30675.950	150.076	-930.275	30784	-13.7
2017:m4	26323	28848.820	30290.850	96.558	-2041.200	28999	-10.17
2017:m5	34167	33684.740	30464.540	104.272	3424.751	33781	1.13
2017:m6	27836	29199.140	30275.340	74.924	-1382.800	29303	-5.27
2017:m7	29006	28214.060	30493.660	89.264	-2003.910	28289	2.47
2017:m8	31220	28048.650	31199.340	150.906	-2198.450	28138	9.87

According to the results in Table 5.11, it indicates that the error percentages for training set vary between -18.77 % to 26.88 % and the value of MAPE criterion is 6.71. It can be noted that the error percentages for the independent set of India vary between -0.99 % to 23.62 %, whereas Mean Absolute Percentage Error (MAPE) is 11.16 (Table 5.12).

Table 5.12: Comparison of actual values versus forecast values for India

Period	Actual	Forecast	95% Confidence Interval		Percentage error (%)	MAPE
September, 2017	34481	30273	33685	26861	12.2	11.16
October, 2017	36996	35454	38920	31988	4.17	
November, 2017	32243	32562	36087	29036	-0.99	
December, 2017	47788	36498	40089	32908	23.62	
January, 2018	37936	32245	35907	28583	15	
February, 2018	32914	29302	33040	25564	10.97	

Moreover, Table 5.13 brings out a comparative analysis between the Holt Winters Multiplicative model and the Additive model with respect to the series of UK.

Table 5.13: Comparison between Holt Winters multiplicative and additive model- UK

	Range of percentages errors (%)		MAPE criterion	
	Training set	Validation set	Training set	Validation set
Multiplicative model	-27.09% -18.46 %	-11.83% -19.01%	8.7	10.72
Additive model	-39.83 % -45.10%	-14.75% -15.92%	13.08	12.15

As shown in Table 5.13, the range of percentage errors of Holt Winters multiplicative model is lower than the range of percentage errors of Holt Winters additive model. Also, the value of MAPE criteria is lower in Holt Winters multiplicative model in both training set and validation set. Thus, among two types of Holt Winters models, Holt Winters Multiplicative model is the most suitable model to forecast the monthly tourist arrivals from the UK. Table 5.14 brings out a comparative analysis between the Holt Winters multiplicative model and the additive model with respect to the series of India.

Table 5.14: Comparison between Holt Winters Multiplicative and Additive model-
India

	Range of percentages errors (%)		MAPE criterion	
	Training set	Validated set	Training set	Validated set
Multiplicative model	-22.37% - 21.86%	-3.38% - 18.98%	6.67	11.11
Additive model	-18.77% - 26.88%	-0.99% - 23.62%	6.71	11.16

As indicated in Table 5.14, the range of percentage errors of Holt Winters Multiplicative model is lower than the range of percentage errors of Holt Winters Additive model. Also, the value of MAPE criteria is lower in Holt Winters Multiplicative model in both training set and validated set. Thus, among two types of Holt Winters models, Holt Winters Multiplicative model is the most appropriate to forecast the monthly tourist arrivals from the UK.

5.5 Development of Equivalent ARIMA Models of GES Models

The equivalent ARIMA model can be defined for any GES model (Mckanzied, 1984). Mckanzied (1984) stated that the minimum mean square error forecasts of the ARIMA processes are identical to the forecasts generated by the GES models for all lead times. Further, he revealed that one can obtain the forecasts using the equivalent ARMA process instead of generating forecasts from the GES methods directly. In this study some sort of attempt is made to develop the ARIMA equivalent models.

5.5.1 ARIMA equivalency of Holt Winters additive model for UK data

The ARIMA equivalent of Holt Winters Additive model for UK data can be represented as $\text{ARIMA}(0, 1, 13) \times (0, 1, 0)_{12}$.

Holt Winters Additive model (UK) $\equiv$ $\text{ARIMA}(0, 1, 13)\times(0, 1, 0)_{12}$

Results in Table 4.7 confirmed that 12th difference of 1st differenced series is stationary. Thus, there is a possibility that $\text{ARIMA}(0, 1, 13) \times (0, 1, 0)_{12}$ model can be developed for the transformed data of the UK. The estimated parameters of relevant model are illustrated in Table 5.15.

Table 5.15: Estimation of parameters

Dependent Variable: UK12
Method: Least Squares
Date: 03/16/19 Time: 18:28
Sample (adjusted): 2011M12 2017M08
Included observations: 69 after adjustments
Failure to improve SSR after 14 iterations
MA Backcast: 2010M11 2011M11

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.554699	13.23838	0.646204	0.5208
MA(1)	-0.722876	0.145349	-4.973396	0.0000
MA(2)	-0.199713	0.157975	-1.264207	0.2115
MA(3)	-0.176135	0.164010	-1.073931	0.2875
MA(4)	0.139758	0.163033	0.857242	0.3950
MA(5)	0.028412	0.177341	0.160214	0.8733
MA(6)	0.099064	0.187786	0.527536	0.5999
MA(7)	0.038518	0.183823	0.209536	0.8348
MA(8)	-0.015536	0.183919	-0.084472	0.9330
MA(9)	-0.114058	0.185154	-0.616013	0.5404
MA(10)	-0.029242	0.183882	-0.159028	0.8742
MA(11)	0.176718	0.184428	0.958194	0.3422
MA(12)	-0.384062	0.176110	-2.180804	0.0335
MA(13)	0.159734	0.163182	0.978869	0.3319

R-squared	0.561729	Mean dependent var	23.57971
Adjusted R-squared	0.458137	S.D. dependent var	2337.250
S.E. of regression	1720.481	Akaike info criterion	17.91762
Sum squared resid	1.63E+08	Schwarz criterion	18.37092
Log likelihood	-604.1578	Hannan-Quinn criter.	18.09746
F-statistic	5.422544	Durbin-Watson stat	2.460759
Prob(F-statistic)	0.000004		

Inverted MA Roots	1.00	.85+.38i	.85-.38i	.51
	.43-.69i	.43+.69i	.05-.89i	.05+.89i
	-.45+.82i	-.45-.82i	-.80-.49i	-.80+.49i
	-.91			

Results in Table 5.15 indicate that the coefficients of several parameters are not significant as corresponding p values are greater than 5%. Thus, the time series of monthly tourist arrivals of UK data fails to develop an equivalent ARIMA model for Holt Winters additive model.

5.5.2 ARIMA equivalency of Holt Winters additive model for India data

The ARIMA equivalent of Holt Winters additive model for India data can be represented as $\text{ARIMA}(0, 1, 13) \times (0, 1, 0)_{12}$.

Holt Winters Additive model (India) ≡ $\text{ARIMA}(0, 1, 13) \times (0, 1, 0)_{12}$

Results in Table 4.13 confirmed that 12^{th} difference of 1^{st} differenced series is stationary. Thus, there is a possibility that $\text{ARIMA}(0, 1, 13) \times (0, 1, 0)_{12}$ model can be developed for the transformed data of India. The estimated parameters of relevant model are illustrated in Table 5.16.

Table 5.16: Estimation of parameters

Method: Least Squares
Date: 03/23/19 Time: 12:24
Sample (adjusted): 2011M12 2017M08
Included observations: 69 after adjustments
Convergence achieved after 70 iterations
MA Backcast: OFF (Roots of MA process too large)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	146.2138	49.15694	2.974428	0.0044
MA(1)	-0.729751	0.161627	-4.515024	0.0000
MA(2)	-0.162621	0.145010	-1.121447	0.2670
MA(3)	-0.003956	0.133465	-0.029640	0.9765
MA(4)	-0.373484	0.124377	-3.002846	0.0040
MA(5)	-0.139154	0.127098	-1.094859	0.2784
MA(6)	-0.077099	0.127544	-0.604488	0.5480
MA(7)	0.428861	0.114244	3.753900	0.0004
MA(8)	-0.153626	0.129564	-1.185716	0.2408
MA(9)	0.037393	0.127147	0.294093	0.7698
MA(10)	0.269253	0.130832	2.058010	0.0443
MA(11)	0.044355	0.145252	0.305366	0.7612
MA(12)	-0.932649	0.156784	-5.948620	0.0000
MA(13)	-0.110855	0.181154	-0.611940	0.5431

R-squared	0.656894	Mean dependent var	99.68116
Adjusted R-squared	0.575796	S.D. dependent var	2583.232
S.E. of regression	1682.483	Akaike info criterion	17.87295
Sum squared resid	1.56E+08	Schwarz criterion	18.32625
Log likelihood	-602.6169	Hannan-Quinn criter.	18.05279
F-statistic	8.100031	Durbin-Watson stat	2.208377
Prob(F-statistic)	0.000000		

Inverted MA Roots	1.18	.94-.37i	.94+.37i	.54-.85i
	.54+.85i	.11-1.01i	.11+1.01i	-.12
	-.49-.84i	-.49+.84i	-.77-.46i	-.77+.46i
	-1.00			

Results in Table 5.16 indicate that the coefficients of several parameters are not significant as corresponding p values are greater than 5%. Thus, the time series of monthly tourist arrivals of India fails to develop an equivalent ARIMA model for Holt Winters Additive model. It can be concluded that it is impossible to develop ARIMA equivalent models for any type of exponential smoothing model with respect to time span considered in the study. Finally, the forecasting accuracy between Seasonal ARIMA$(0, 1, 1) \times (0, 1, 1)_{12}$model and Holt Winters Multiplicative model were compared and the results were indicated in Table 5.17.

5.6 Comparative Analysis of SARIMA Model and Holt Winters Multiplicative Model.

Table 5.17: Comparison of actuals and forecasts for UK series using two models

Period	Actuals	Forecasts (SARIMA)	Error %	MAPE	Forecasts (HWM)	Error %	MAPE
September, 2017	12,593	14089	-11.88	7.89	13599	-7.99	10.727
October, 2017	12,518	13175	-5.25		13327	-6.46	
November, 2017	13,634	14907	-9.33		15247	-11.83	
December, 2017	21,756	21813	-0.26		21004	3.46	
January, 2018	22,940	20245	11.75		18373	19.91	
February, 2018	23,817	21699	8.89		20312	14.72	

As indicated in Table 5.17, the range of percentage errors for validated set of SARIMA model and Holt Winters multiplicative model varies between -11.88% to -11.75% and -11.83% to 19.91%, respectively. Also, the value of Mean Absolute Percentage Error (MAPE) obtained for SARIMA model and Holt Winters multiplicative model are 7.89 and 10.727, respectively. It is evident that the accuracy of the SARIMA model is high as the corresponding MAPE value is less than 10%. Thus, it can be concluded that forecasting performance of SARIMA model is better than the Holt Winters multiplicative model regarding to the data considered in this study. Thus, it can be recommended that Box Jenkins Seasonal ARIMA$(0, 1, 1) \times (0, 1, 1)_{12}$ model is the best model to forecast the monthly tourist arrivals from the UK. Table 5.18 shows comparative analysis between actuals and forecasts of India regarding two models.

Table 5.18: Comparison of actuals and forecasts for India series regarding two models

Period	Actuals	Forecasts (SARIMA)	Error %	MAPE	Forecasts (HWM)	Error %	MAPE
September, 2017	34481	31516	8.6	8.14	30292	12.15	11.11
October, 2017	36996	37449	-1.23		38248	-3.38	
November, 2017	32243	34061	-5.64		33086	-2.62	
December, 2017	47788	38748	18.92		38720	18.98	
January, 2018	37936	34298	9.59		32481	14.38	
February, 2018	32914	31303	4.9		27912	15.2	

The range of percentage errors of SARIMA model and Holt Winters multiplicative model for the validated set varies between -5.64% to 9.59% and –3.38% to 14.38%, respectively with the exception of the forecast in December. The values of Mean Absolute Percentage Error (MAPE) obtained for the SARIMA model and Holt Winters multiplicative model are 8.14 and 11.116 respectively. The accuracy of the SARIMA model is high as MAPE value is less than 10%. Accordingly, it is evident that SARIMA model outperforms the Holt Winters multiplicative model. Thus, Box Jenkins Seasonal ARIMA$(0, 1, 1) \times (0, 1, 1)_{12}$ model is recommended to forecast the monthly tourist arrivals from India. The future monthly tourist arrivals from the UK and India were forecasted using the SARIMA model for the time period from March, 2018 to July, 2018 and indicated in Table 5.19.

Table 5.19: Forecasts of monthly tourist arrivals using SARIMA model

(From March, 2018 to July, 2018)

Month	United Kingdom	India
March, 2018	21825	33345
April, 2018	18017	32463
May, 2018	11068	38670
June, 2018	12569	33890
July, 2018	25626	33813
August, 2018	23478	34102

Table 5.20: Percentage Changes in future monthly tourist arrivals

Month	Monthly tourist arrivals, 2018		Monthly tourist arrivals, 2017		Percentage Change (%)	
	UK	India	UK	India	UK	India
March	21825	33345	19451	27075	12.2	23.16
April	18017	32463	17841	26323	0.99	23.33
May	11068	38670	8520	34167	29.9	13.18
June	12569	33890	10424	27836	20.57	21.75
July	25626	33813	23553	29006	8.8	16.57
August	23478	34102	21903	31220	7.19	9.23

The percentage changes in future monthly tourist arrivals indicated in Table 5.20 reveal that it can be expected an increment of monthly tourist arrivals in the coming months in the year, 2018 as monthly percentage changes are positive.

5.7 Summary of Chapter 5

Both additive and multiplicative models can be easily forecast the monthly tourist arrivals from both countries. These models are easier to apply than SARIMA model. Out of these two Holt Winters models, multiplicative model is recommended for both countries. The errors of the models found to be white noise. However, according to the details of validated set, it was found that SARIMA model is better than the Holt Winters multiplicative model. Thus, SARIMA model can be recommended to forecast.

CHAPTER 6

CONCLUSIONS AND FUTURE STUDIES

6.1 Conclusions

- Seasonal ARIMA models are better than smoothing models in forecasting monthly tourist arrivals from UK and India.

- The best fitted model for UK series is as follows.

$$(1-B)(1-B^{12})y_t = 9.035 + (1-0.414994B^{12})(1-0.985634B)e_t \qquad (6.1)$$

The predicted values for independent set are from September, 2017 to February, 2018. The corresponding percentage errors for independent set vary between -11.88% to 11.75% and the value of MAPE is 7.89%.

- The best fitted model for India is as follows.

$$(1-B)(1-B^{12})y_t = 26.76120 + (1-0.830598B^{12})(1-0.628868B)e_t \qquad (6.2)$$

The error percentages for independent set vary between -5.64 % to 9.59 % with the exception of the forecast in December and the value of MAPE is 8.14%.

- Accordingly, the two models represented by the equation (6.1) and (6.2) are recommended to forecast the monthly tourist arrivals from UK and India to Sri Lanka, respectively. These models can be easily used by policy makers.

6.2 Future Studies

- One of the drawbacks of applying time series techniques in forecasting the monthly tourist arrivals is that it does not identify the external factors. The arrivals of tourists may be affected by numerous factors such as social conflicts, wars, politics and economic factors. Thus, it is recommended to develop forecasting methods which quantify these non random events.
- Only two countries were considered in this study. Moreover, this study can be extended to investigate other high tourist generating countries.
- This study will provide useful information to increase the effectiveness of strategic plans to develop the tourism industry. However, it should be updated continuously with the incorporation of recent data as the nature of stochastic processes change with time.
- Further studies can be done on tourism demand forecasting using multivariate forecasting techniques which consider the inter-series dependencies.

References

Annorzie, M.N., Shittu, O.I., Iwu, H.C., and Yaya, O.S. (2018). Effects of Residuals of Autocorrelation Function and Partial Autocorrelation Function in Long-Range Department Market Analysis, The International Journal of Engineering and Science (IJES) , Vol. 7, pp. 31-40

Athanasopoulos, G., and De Silva, A. (2010). Multivariate Exponential Smoothing for Forecasting Tourist Arrivals to Australia and New Zealand, Department of Econometrics and Business Studies, ISSN 1440-771

Borhan, N., and Arsad, Z., (2014). Forecasting International Tourism demand from the US, Japan and South Korea to Malaysia: A SARIMA approach, AIP Conference Proceedings 1605, 955-960

Chaitip, P., Chaiboonsri, H., and Mukhjang, R. (2008). Time series methods for forecasting International visitor arrivals to Thailand, International Conference on Applied Economics-ICOAE 2008

Cho, V., (2002). A comparison of three different approaches to tourist arrival forecasting, Department of Management, The Hong Kong Polytechnic University, Tourism Management 24, 323-330

Chu, F. (2009). Forecasting tourism demand with ARMA based methods, Tourism Management, Vol 30, No 5, pp. 740-751

Coshall, J.T. (2009). Tourism Management, The Business School, London Metropolitan University, Tourism Management 30, 495-511

Cuhadar, M. (2014). Modeling and Forecasting inbound tourism demand to Istanbul: A Comparative Analysis, European Journal of Business and Social Sciences, Vol 2, No 12, pp 101-119.

Dias, G.K.A., Hettiarachchi, C., Karunaratne., D.D., Kodituwakku., W.H., and Wijesundara., D.M. (2016). Modelling and Forecasting Tourism Demand for Sri Lanka, University of Colombo School of Computing, Sri Lanka

Dickey, D.A., and Fuller, W.A. (1979). Distribution of Estimators of Autoregressive Time series With a Unit Root, Journal of American Statistical Association, Vol. 74, Issue 366, 427-431

Ersen, N., Akyuz, I., Bayram, B.C., and Ucuncu, T. (2017). Modeling and Forecasting of Log Production in Turkey, International Hjournal of Ecosystems and Ecology Science, Vol. 7(4): 791-796

Ganapragasam S.R., and Cooray.T.M.J.A. (2016a). Forecasting post war tourist arrivals to Sri Lanka using Dynamic Transfer function modeling method, International Journal of Multidisciplinary Studies.

Ganapragasam S.R., and Cooray.T.M.J.A. (2016b). Tourist Arrivals after the international conflict in Sri Lanka: A Time Series Analysis using Holt- Winter's method, Conference Proceeding of the Annual Academic Sessions, The Open University of Sri Lanka.

Goh, C., and Law, R. (2002). Modelling and Forecasting Tourism demand for arrivals with stochastic non stationary seasonality and Intervention, Department of Tourism Management, The Hong Kong Polytechnic University, Tourism Management 30, 495-511

Hassani, H., Silva, E.S., Antonakakis, N., Filis, G., and Gupta, R. (2015). Forecasting Accuracy Evaluation of Tourist Arrivals: Evidence from Parametric and Non-parametric Techniques, Department of Economics Working paper series, University of Pretoria, Working paper 2015-52

Ishara, D.K., and Wijekoon, P. (2017). Time series analysis of Modeling and Forecasting international tourist arrivals in Sri Lanka, British Journal of Mathematics and Computer Science, Vol 20(5)

Karmaker,C.L., Halder, P.K., and Sarker, E. (2017). A Study of Time Series Model for Predicting Jute Yarn Demand: Case study, Journal of Industrial Engineering, Volume 2017, Article ID 2061260, 8 pages.

Kodippilli, A. (2017). Forecasting tourist arrivals to Sri Lanka using Seasonal ARIMA, Journal of Tourism, Hospitality and Sports, Vol 29.

Kodituwakku, W., Wijesundara., & Hettiarchchi, C. (2015). Modelling and Forecasting Tourism Demand for Sri Lanka, University of Colombo School of Computing.

Konarasinghe, K.M.U.B. (2016a). Short - term Forecasting of Tourist Arrivals to Sri Lanka from Asian Region, Institute of Mathematics and Management, Nugegoda, Sri Lanka.

Konarasinghe, K.M.U.B. (2016b). Decomposition techniques of forecasting tourist arrivals from Western European countries to Sri Lanka, 13th International Conference on Business Management, Faculty of Management Studies and Commerce, University of Sri Jayewardenepura, Sri Lanka, p31. ISSN: 2235- 9773

Konarasinghe, K.M.U.B. (2017). Model Comparison on Forecasting Tourist Arrivals to Sri Lanka from Western Europe, International Journal of Novel Research in Physics, Chemistry and mathematics, Vol. 4, Issue 2, pp. (28-35)

Kumari, K. (2015). Forecasting foreign tourist arrivals in India using different time series models, International Journal of Emerging Technologies in Computational and Applied Sciences 15(1), pp 38-43

Kurukulasooriya, N., and Lelwala, E. (2014). Time Series Behavior of Burgeoning International Tourist arrivals in Sri Lanka: The post - war experiences, Ruhuna Journals of Management and Finance, Vol. 1 (1)

Lewis, C. (1982). International and business forecasting methods, London: Butterworths.

Lim, C., and McAleer, M. (2001). Time Series Forecasts of International Tourism Demand for Australia, The Institute of Social and Economic Research, Osaka University, Discussion paper no. 533

Loganathan, N., and Ibrahim, Y. (2010). Forecasting International Tourism Demand in Malaysia using Box Jenkins SARIMA Application, South Asian Journal of Tourism and Heritage, Vol.3, Number 2.

Lwesya, F., and Kibambila, V. (2017). A Comparative Analysis of the Application of Seasonal ARIMA and Exponential Smoothing methods in short run Forecasting Tourist Arrivals in Tanzania, European Journal of Business and Management, Vol .9, No.10, 56-69

Makatjane, K.D., and Moroke, N.D. (2016). Comparative study of Holt Winters Triple Exponential Smoothing and Seasonal ARIMA: Forecasting short term seasonal car sales in South Africa, Risk Goverance and Control: Financial Markets and Institutions/ Vol 6, Issue 1.

Malhotra, N., and Yadev, P. (2017). Forecasting the Growth of Indian Tourism using Exponential Smoothing.

Mamula, M. (2015). Modelling and Forecasting International Tourism Demand - Evaluation of Forecasting Performance, International Journal of Business Administration, Vol. 6, No.3, 102-112

Mckenzie, E.D. (1984). General Exponential Smoothing and Equivalent ARMA Process, Journal of Forecasting, Vol 3, 333-344.

Nishantha, K. A. (2008). Forecasting International Tourist Demand for Southern Coastal Area in Sri Lanka, Journal of Economic Studies.

Papic-Blagojevic, N., Vujko, A., and Gajic, T. (2016). Comparative analysis of exponential smoothing models to tourist arrivals in Serbia, Journal of Economic of Agriculture, 757-767.

Pasquali, V. (2016). “The World’s Richest and Poorest Countries. Global Finance”, July/ August 2016.

Paul, S.K. (2011). Determination of exponential smoothing constant to minimize Mean Square Error and Mean Absolute Deviation, Global Journal of Research in Engineering, Vol.11, Issue 3 version 1.0.

Peiris, H. R. I. (2016). A Seasonal ARIMA Model of Tourism Forecasting: The Case of Sri Lanka, Journal of Tourism, Hospitality and Sports, Vol. 22, 98-109.

Peiris, T. S. G. (2017). Time Series Analysis for Business Forecasting [Lecture notes].

Petrevska, B. (2017). Predicting tourism demand by ARIMA Models, Faculty of Tourism and Business Statistics, Goce Delcev University, Economic Research, Vol. 30, No. 1. 939-950

Preez, J. D., and Witt, S.F. (2003). Univariate versus multivariate time series forecasting: An application to international tourism demand, International Journal of forecasting, 19, 436-451

SAS Institute Inc. (n. d.). Equations for Smoothing Models. Retrieved from: http://support.sas.com/documentation/cdl/en/etsug/60372/HTML/default/viewer.htm #etsug_tffordet_sect014.htm

Sri Lanka Tourism Development Authority (2016). Annual Statistical Report 2016.

Sri Lanka Tourism Development Authority (2017). Annual Statistical Report 2017.

Song, H., and Li, G. (2008). Tourism Demand Modeling and Forecasting - A review of recent research, Tourism Management

Tirkes, G., Giiray, C., and Celebi, N. (2017). Demand Forecasting: A Comparison between the Holt-Winters, Trend Analysis and Decomposition Models

Wong, K.K.F., Song, H., Witt, S.F., and Wu, D.C. (2007), Tourism Forecasting: To combine or not to combine, School of Hotel and Tourism Management, The Hong Kong Polytechnic University, Tourism Management 28, 1068-1078

Zhi-Peng, L., Yun-Cai, L., and Fu-Qiang, L. (2008). An Improved Adaptive Exponential Smoothing Model for Short-term Travel Forecasting of Urban Arterial Street, Acta Automatica Sinica, Vol.34, No.11, pp. 1404-1409

Appendix A: Monthly tourist arrivals from UK and India

	UK	India
2010:m11	9788	15550
2010:m12	10176	15753
2011:m1	9518	13786
2011:m2	9614	10071
2011:m3	8852	13619
2011:m4	9038	14705
2011:m5	4452	16649
2011:m6	5188	12927
2011:m7	12003	12587
2011:m8	12486	12857
2011:m9	7871	13329
2011:m10	7408	17915
2011:m11	9589	15474
2011:m12	10063	17455
2012:m1	8162	14615
2012:m2	8746	11342
2012:m3	12032	14391
2012:m4	8019	15432
2012:m5	4940	15888
2012:m6	6076	13758
2012:m7	13643	11564
2012:m8	11558	11242
2012:m9	8586	13888
2012:m10	8767	17654
2012:m11	10828	17625
2012:m12	12861	18941
2013:m1	11350	16938
2013:m2	13604	14429
2013:m3	13590	15281
2013:m4	8642	12657
2013:m5	6567	17834
2013:m6	7642	15297
2013:m7	16424	14783

	UK	India
2013:m8	15519	18999
2013:m9	9356	18389
2013:m10	9850	21833
2013:m11	9663	19796
2013:m12	15209	22559
2014:m1	12896	17886
2014:m2	14316	17534
2014:m3	11823	18734
2014:m4	12380	17192
2014:m5	6918	22981
2014:m6	7790	19911
2014:m7	16692	19277
2014:m8	15532	17912
2014:m9	8983	19244
2014:m10	10112	26148
2014:m11	10730	19762
2014:m12	15996	26153
2015:m1	13410	22944
2015:m2	15212	19999
2015:m3	16191	21838
2015:m4	11233	23048
2015:m5	7954	31764
2015:m6	8580	25860
2015:m7	20643	24681
2015:m8	17908	23540
2015:m9	11160	27233
2015:m10	9970	30574
2015:m11	10822	29329
2015:m12	18762	35437
2016:m1	16253	28895
2016:m2	19194	26559
2016:m3	21430	30170
2016:m4	12006	25890

Appendix A (Continued)

	UK	India
2016:m5	8412	34259
2016:m6	9406	27068
2016:m7	23948	27665
2016:m8	20475	24418
2016:m9	12288	27241
2016:m10	10964	36471
2016:m11	13337	30148
2016:m12	20446	37945
2017:m1	19468	31173
2017:m2	20218	26320
2017:m3	19451	27075
2017:m4	17841	26323
2017:m5	8520	34167
2017:m6	10424	27836
2017:m7	23553	29006
2017:m8	21903	31220

Appendix B: ACF plots of residuals for six SARIMA models postulated for UK data

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.044	0.044	0.1412	
		2	-0.073	-0.075	0.5263	
		3	-0.210	-0.205	3.7984	0.051
		4	0.040	0.054	3.9195	0.141
		5	0.105	0.077	4.7650	0.190
		6	0.151	0.113	6.5427	0.162
		7	0.095	0.124	7.2582	0.202
		8	-0.035	0.009	7.3544	0.289
		9	-0.146	-0.099	9.0911	0.246
		10	-0.111	-0.099	10.118	0.257
		11	0.061	0.010	10.430	0.317
		12	0.039	-0.056	10.562	0.393
		13	0.030	-0.011	10.642	0.474
		14	-0.108	-0.073	11.680	0.472
		15	-0.156	-0.122	13.880	0.382
		16	-0.152	-0.129	16.022	0.312
		17	-0.051	-0.099	16.264	0.365
		18	-0.134	-0.242	17.985	0.325
		19	0.056	-0.007	18.291	0.371
		20	-0.131	-0.162	20.010	0.332
		21	-0.037	-0.040	20.147	0.386
		22	-0.102	-0.048	21.228	0.384
		23	0.040	0.030	21.401	0.435
		24	-0.083	-0.121	22.148	0.451

Appendix B_1: ACF of residuals for UK- model 1

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.021	0.021	0.0317	
		2	0.012	0.012	0.0423	0.837
		3	-0.089	-0.089	0.6244	0.732
		4	0.133	0.138	1.9539	0.582
		5	0.084	0.081	2.4958	0.645
		6	0.136	0.124	3.9342	0.559
		7	0.056	0.077	4.1863	0.651
		8	0.009	0.003	4.1933	0.757
		9	-0.064	-0.067	4.5260	0.807
		10	-0.059	-0.094	4.8196	0.850
		11	0.010	-0.030	4.8281	0.902
		12	-0.227	-0.287	9.2652	0.597
		13	-0.099	-0.136	10.116	0.606
		14	-0.076	-0.078	10.634	0.641
		15	-0.050	-0.089	10.861	0.697
		16	-0.119	-0.055	12.160	0.667
		17	-0.085	-0.014	12.845	0.684
		18	-0.231	-0.160	17.974	0.390
		19	-0.063	-0.008	18.361	0.432
		20	-0.156	-0.126	20.793	0.348
		21	0.030	0.002	20.884	0.404
		22	-0.037	-0.013	21.029	0.457
		23	0.040	0.075	21.196	0.509
		24	-0.218	-0.224	26.375	0.283

Appendix B_2: ACF of residuals for UK- model 2

Appendix B (Continued)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.046	-0.046	0.1510	
		2	0.001	-0.001	0.1510	
		3	-0.086	-0.086	0.6901	0.406
		4	0.117	0.110	1.6997	0.427
		5	0.067	0.078	2.0397	0.564
		6	0.127	0.130	3.2739	0.513
		7	0.050	0.086	3.4718	0.628
		8	-0.006	0.003	3.4746	0.747
		9	-0.066	-0.065	3.8275	0.799
		10	-0.057	-0.096	4.0982	0.848
		11	0.023	-0.027	4.1428	0.902
		12	-0.217	-0.273	8.1478	0.614
		13	-0.074	-0.142	8.6269	0.656
		14	-0.048	-0.065	8.8342	0.717
		15	-0.042	-0.078	8.9952	0.773
		16	-0.120	-0.072	10.314	0.739
		17	-0.070	-0.021	10.765	0.769
		18	-0.226	-0.180	15.645	0.478
		19	-0.036	-0.028	15.769	0.540
		20	-0.140	-0.141	17.708	0.475
		21	0.035	-0.028	17.835	0.533
		22	-0.043	-0.035	18.024	0.586
		23	0.036	0.061	18.161	0.639
		24	-0.227	-0.237	23.733	0.361

Appendix B_3: ACF of residuals for UK- model 3

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.137	-0.137	1.3381	
		2	-0.359	-0.385	10.611	0.001
		3	-0.131	-0.301	11.860	0.003
		4	0.082	-0.207	12.358	0.006
		5	0.030	-0.232	12.426	0.014
		6	0.089	-0.071	13.032	0.023
		7	0.032	-0.022	13.112	0.041
		8	-0.029	0.028	13.178	0.068
		9	-0.076	0.003	13.641	0.092
		10	0.025	0.063	13.694	0.134
		11	0.122	0.191	14.945	0.134
		12	-0.156	-0.070	17.011	0.108
		13	-0.029	0.006	17.083	0.147
		14	0.059	-0.041	17.394	0.182
		15	0.026	-0.076	17.457	0.233
		16	0.007	-0.027	17.461	0.292
		17	0.028	-0.004	17.534	0.352
		18	-0.119	-0.124	18.889	0.335
		19	0.024	-0.005	18.945	0.395
		20	-0.066	-0.185	19.372	0.433
		21	0.046	-0.121	19.584	0.484
		22	0.072	-0.086	20.127	0.513
		23	-0.003	-0.060	20.128	0.575
		24	-0.238	-0.364	26.230	0.290

Appendix B_4: ACF of residuals for UK- model 4

Appendix B (Continued)

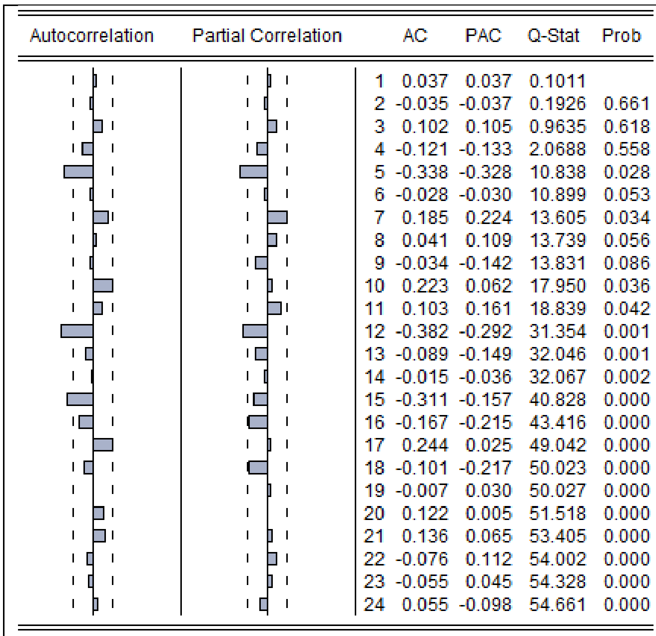
Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.104	-0.104	0.7603	
		2	-0.201	-0.214	3.6378	
		3	-0.341	-0.412	12.043	0.001
		4	0.086	-0.107	12.590	0.002
		5	0.050	-0.164	12.777	0.005
		6	0.137	-0.039	14.207	0.007
		7	0.028	0.045	14.268	0.014
		8	-0.033	0.010	14.354	0.026
		9	-0.048	0.059	14.536	0.042
		10	0.007	0.074	14.540	0.069
		11	0.104	0.170	15.437	0.080
		12	-0.139	-0.087	17.060	0.073
		13	-0.004	0.010	17.061	0.106
		14	0.003	-0.002	17.062	0.147
		15	0.065	-0.063	17.435	0.180
		16	-0.015	-0.027	17.456	0.233
		17	0.037	0.008	17.583	0.285
		18	-0.141	-0.161	19.466	0.245
		19	0.025	-0.021	19.525	0.299
		20	-0.110	-0.210	20.707	0.294
		21	0.107	-0.092	21.850	0.292
		22	-0.031	-0.147	21.947	0.343
		23	0.060	-0.066	22.325	0.381
		24	-0.211	-0.317	27.109	0.207

Appendix B_5: ACF of residuals for UK- model 5

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.079	-0.079	0.4343	
		2	-0.120	-0.127	1.4405	
		3	-0.174	-0.199	3.5948	
		4	-0.183	-0.251	6.0154	0.014
		5	0.078	-0.035	6.4595	0.040
		6	0.164	0.078	8.4763	0.037
		7	0.059	0.025	8.7387	0.068
		8	0.038	0.066	8.8503	0.115
		9	-0.047	0.054	9.0239	0.172
		10	-0.042	0.052	9.1636	0.241
		11	0.098	0.150	9.9470	0.269
		12	-0.146	-0.124	11.724	0.229
		13	0.002	-0.034	11.725	0.304
		14	0.045	0.018	11.900	0.371
		15	-0.020	-0.047	11.936	0.451
		16	0.031	-0.054	12.023	0.526
		17	0.002	-0.027	12.024	0.604
		18	-0.164	-0.165	14.526	0.486
		19	-0.008	-0.073	14.531	0.559
		20	-0.071	-0.146	15.019	0.594
		21	-0.033	-0.166	15.127	0.653
		22	0.030	-0.145	15.220	0.708
		23	0.079	0.036	15.874	0.724
		24	-0.160	-0.246	18.611	0.610

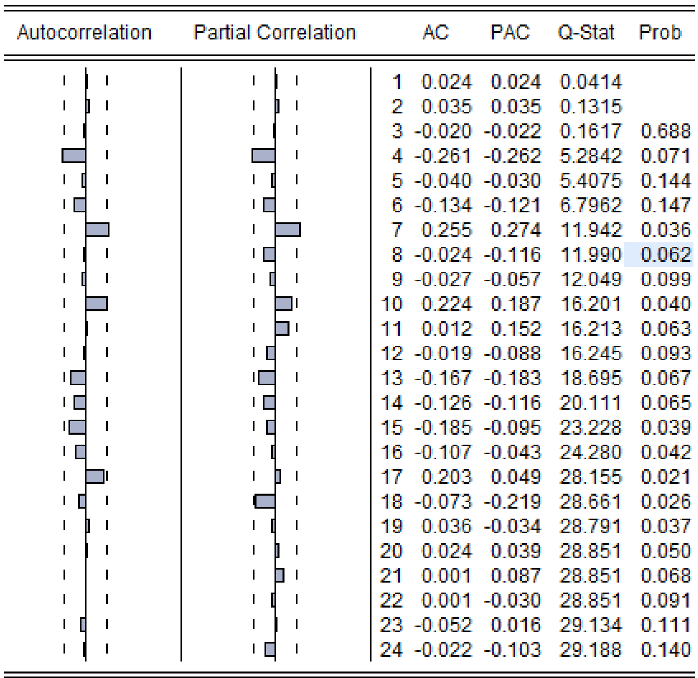
Appendix B_6: ACF of residuals for UK- model 6

Appendix C: ACF plots of residuals for six SARIMA models postulated for India data



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.037	0.037	0.1011	
		2	-0.035	-0.037	0.1926	0.661
		3	0.102	0.105	0.9635	0.618
		4	-0.121	-0.133	2.0688	0.558
		5	-0.338	-0.328	10.838	0.028
		6	-0.028	-0.030	10.899	0.053
		7	0.185	0.224	13.605	0.034
		8	0.041	0.109	13.739	0.056
		9	-0.034	-0.142	13.831	0.086
		10	0.223	0.062	17.950	0.036
		11	0.103	0.161	18.839	0.042
		12	-0.382	-0.292	31.354	0.001
		13	-0.089	-0.149	32.046	0.001
		14	-0.015	-0.036	32.067	0.002
		15	-0.311	-0.157	40.828	0.000
		16	-0.167	-0.215	43.416	0.000
		17	0.244	0.025	49.042	0.000
		18	-0.101	-0.217	50.023	0.000
		19	-0.007	0.030	50.027	0.000
		20	0.122	0.005	51.518	0.000
		21	0.136	0.065	53.405	0.000
		22	-0.076	0.112	54.002	0.000
		23	-0.055	0.045	54.328	0.000
		24	0.055	-0.098	54.661	0.000

Appendix C_1: ACF of residuals for India- model 1



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.024	0.024	0.0414	
		2	0.035	0.035	0.1315	
		3	-0.020	-0.022	0.1617	0.688
		4	-0.261	-0.262	5.2842	0.071
		5	-0.040	-0.030	5.4075	0.144
		6	-0.134	-0.121	6.7962	0.147
		7	0.255	0.274	11.942	0.036
		8	-0.024	-0.116	11.990	0.062
		9	-0.027	-0.057	12.049	0.099
		10	0.224	0.187	16.201	0.040
		11	0.012	0.152	16.213	0.063
		12	-0.019	-0.088	16.245	0.093
		13	-0.167	-0.183	18.695	0.067
		14	-0.126	-0.116	20.111	0.065
		15	-0.185	-0.095	23.228	0.039
		16	-0.107	-0.043	24.280	0.042
		17	0.203	0.049	28.155	0.021
		18	-0.073	-0.219	28.661	0.026
		19	0.036	-0.034	28.791	0.037
		20	0.024	0.039	28.851	0.050
		21	0.001	0.087	28.851	0.068
		22	0.001	-0.030	28.851	0.091
		23	-0.052	0.016	29.134	0.111
		24	-0.022	-0.103	29.188	0.140

Appendix C_2: ACF of residuals for India- model 2

Appendix C (Continued)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.011	-0.011	0.0091	
		2	0.021	0.021	0.0406	
		3	0.161	0.162	1.9452	0.163
		4	-0.051	-0.049	2.1386	0.343
		5	-0.267	-0.283	7.5047	0.057
		6	-0.005	-0.040	7.5066	0.111
		7	0.176	0.242	9.9311	0.077
		8	0.029	0.155	9.9958	0.125
		9	-0.073	-0.154	10.422	0.166
		10	0.175	-0.016	12.925	0.114
		11	0.088	0.137	13.569	0.138
		12	-0.401	-0.305	27.224	0.002
		13	-0.083	-0.190	27.823	0.003
		14	-0.016	-0.041	27.847	0.006
		15	-0.304	-0.129	36.137	0.001
		16	-0.172	-0.185	38.841	0.000
		17	0.224	0.070	43.544	0.000
		18	-0.135	-0.174	45.273	0.000
		19	-0.013	0.065	45.289	0.000
		20	0.080	0.054	45.919	0.000
		21	0.137	0.108	47.806	0.000
		22	-0.063	0.128	48.210	0.000
		23	-0.035	0.069	48.337	0.001
		24	0.080	-0.123	49.024	0.001

Appendix C_3: ACF of residuals for India- model 3

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.011	0.011	0.0082	
		2	0.028	0.028	0.0646	
		3	-0.027	-0.028	0.1174	0.732
		4	-0.145	-0.146	1.6697	0.434
		5	-0.323	-0.327	9.4639	0.024
		6	-0.041	-0.050	9.5931	0.048
		7	0.171	0.204	11.834	0.037
		8	0.069	0.065	12.207	0.058
		9	0.012	-0.115	12.219	0.094
		10	0.198	0.081	15.412	0.052
		11	0.110	0.193	16.406	0.059
		12	-0.347	-0.279	26.547	0.003
		13	-0.075	-0.124	27.034	0.005
		14	-0.070	-0.025	27.460	0.007
		15	-0.236	-0.154	32.432	0.002
		16	-0.141	-0.215	34.246	0.002
		17	0.226	0.017	38.966	0.001
		18	-0.101	-0.264	39.927	0.001
		19	0.022	-0.044	39.974	0.001
		20	0.063	-0.015	40.365	0.002
		21	0.144	0.043	42.446	0.002
		22	-0.070	0.079	42.942	0.002
		23	-0.035	0.053	43.073	0.003
		24	0.007	-0.136	43.078	0.005

Appendix C_4: ACF of residuals for India- model 4

Appendix C (Continued)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.097	-0.097	0.6737	
		2	-0.202	-0.214	3.6230	0.057
		3	0.187	0.150	6.1702	0.046
		4	-0.075	-0.091	6.5866	0.086
		5	-0.362	-0.335	16.461	0.002
		6	-0.001	-0.148	16.461	0.006
		7	0.240	0.152	20.963	0.002
		8	-0.017	0.114	20.987	0.004
		9	-0.129	-0.146	22.334	0.004
		10	0.272	0.088	28.403	0.001
		11	0.148	0.212	30.244	0.001
		12	-0.415	-0.224	44.882	0.000
		13	-0.010	-0.124	44.891	0.000
		14	0.136	-0.027	46.522	0.000
		15	-0.315	-0.145	55.449	0.000
		16	-0.133	-0.174	57.078	0.000
		17	0.308	-0.013	65.912	0.000
		18	-0.137	-0.277	67.699	0.000
		19	-0.082	-0.021	68.357	0.000
		20	0.109	-0.105	69.536	0.000
		21	0.142	0.000	71.582	0.000
		22	-0.099	0.081	72.605	0.000
		23	-0.059	0.057	72.970	0.000
		24	0.112	-0.098	74.331	0.000

Appendix C_5: ACF of residuals for India- model 5

Appendix D: Residuals of Seasonal ARIMA (0, 1, 1) (0, 1, 1) model - UK

	month	actual	Residual
14	2011M12	10063	-1040.3007
15	2012M01	8162	-2065.0400
16	2012M02	8746	-1738.3005
17	2012M03	12032	1897.7243
18	2012M04	8019	-1634.0975
19	2012M05	4940	-353.9658
20	2012M06	6076	-14.5578
21	2012M07	13643	462.7825
22	2012M08	11558	-1733.4741
23	2012M09	8586	-105.6265
24	2012M10	8767	410.5972
25	2012M11	10828	-84.8397
26	2012M12	12861	1400.2489
27	2013M01	11350	1329.6351
28	2013M02	13604	3094.7820
29	2013M03	13590	1239.8520
30	2013M04	8642	-1176.3656
31	2013M05	6567	357.0036
32	2013M06	7642	420.5816
33	2013M07	16424	1818.5109
34	2013M08	15519	2054.6769
35	2013M09	9356	-509.6635
36	2013M10	9850	15.2231
37	2013M11	9663	-2445.1862
38	2013M12	15209	1709.7030
39	2014M01	12896	853.1500
40	2014M02	14316	738.3107
41	2014M03	11823	-2511.6653
42	2014M04	12380	2025.0582
43	2014M05	6918	-770.7433
44	2014M06	7790	-943.1931
45	2014M07	16692	-236.0389
46	2014M08	15532	-387.8341
47	2014M09	8983	-1829.2338
48	2014M10	10112	-962.2040

Appendix D (Continued)

	Month	actual	Residual
49	2014M11	10730	-1173.3804
50	2014M12	15996	264.1173
51	2015M01	13410	-366.9836
52	2015M02	15212	-31.3181
53	2015M03	16191	2087.7780
54	2015M04	11233	-1598.5115
55	2015M05	7954	-549.7502
56	2015M06	8580	-873.0483
57	2015M07	20643	2579.3003
58	2015M08	17908	893.8096
59	2015M09	11160	72.4496
60	2015M10	9970	-1907.7198
61	2015M11	10822	-1748.7221
62	2015M12	18762	1530.9222
63	2016M01	16253	1316.5653
64	2016M02	19194	2564.7277
65	2016M03	21430	4655.0730
66	2016M04	12006	-1404.1778
67	2016M05	8412	-1282.3410
68	2016M06	9406	-1042.3984
69	2016M07	23948	2870.0403
70	2016M08	20475	1397.6830
71	2016M09	12288	-405.9620
72	2016M10	10964	-1364.4915
73	2016M11	13337	221.6852
74	2016M12	20446	729.0722
75	2017M01	19468	2160.7335
76	2017M02	20218	455.4861
77	2017M03	19451	-1680.3216
78	2017M04	17841	3661.9829
79	2017M05	8520	-2084.4701
80	2017M06	10424	-1061.6302
81	2017M07	23553	-850.9903
82	2017M08	21903	381.2911

Appendix E: Residuals of Seasonal ARIMA (0, 1, 1) (0, 1, 1) model - India

	Month	actual	Residual
14	2011M12	17455	-1932.1400
15	2012M01	14615	-69.7868
16	2012M02	11342	186.7761
17	2012M03	14391	1046.9187
18	2012M04	15432	2028.5524
19	2012M05	15888	-3035.8342
20	2012M06	13758	426.2916
21	2012M07	11564	-1689.9698
22	2012M08	11242	-2030.4298
23	2012M09	13888	-269.7884
24	2012M10	17654	-2372.2411
25	2012M11	17625	1744.9238
26	2012M12	18941	-850.1315
27	2013M01	16938	1226.8811
28	2013M02	14429	1700.3732
29	2013M03	15281	-382.4421
30	2013M04	12657	-2794.1986
31	2013M05	17834	-644.0894
32	2013M06	15297	1100.9963
33	2013M07	14783	719.2663
34	2013M08	18999	4159.8240
35	2013M09	18389	169.7011
36	2013M10	21833	-2071.5007
37	2013M11	19796	-649.0223
38	2013M12	22559	-605.4661
39	2014M01	17886	-1614.4192
40	2014M02	17534	1886.4644
41	2014M03	18734	301.7528
42	2014M04	17192	-876.0917
43	2014M05	22981	958.8257
44	2014M06	19911	1294.1307
45	2014M07	19277	689.4068
46	2014M08	17912	-2094.7717
47	2014M09	19244	-1433.9713
48	2014M10	26148	722.2340

Appendix E (Continued)

	month	actual	Residual
49	2014M11	19762	-3378.6274
50	2014M12	26153	1312.6373
51	2015M01	22944	1238.0377
52	2015M02	19999	568.9655
53	2015M03	21838	235.3088
54	2015M04	23048	1987.9202
55	2015M05	31764	5404.3919
56	2015M06	25860	1111.9608
57	2015M07	24681	24.1635
58	2015M08	23540	-1887.5819
59	2015M09	27233	1050.3210
60	2015M10	30574	-1580.3458
61	2015M11	29329	936.8784
62	2015M12	35437	3134.4669
63	2016M01	28895	-1045.9219
64	2016M02	26559	-249.5987
65	2016M03	30170	1486.5297
66	2016M04	25890	-3053.6776
67	2016M05	34259	1156.3940
68	2016M06	27068	-2485.8614
69	2016M07	27665	-374.7876
70	2016M08	24418	-3948.8970
71	2016M09	27241	-1521.7480
72	2016M10	36471	3044.0065
73	2016M11	30148	-1586.8408
74	2016M12	37945	2778.4425
75	2017M01	31173	-1015.4758
76	2017M02	26320	-2843.3543
77	2017M03	27075	-3305.7720
78	2017M04	26323	-1890.5038
79	2017M05	34167	814.9082
80	2017M06	27836	-1323.0709
81	2017M07	29006	701.3606
82	2017M08	31220	2791.1201