

**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

Declaration	
Abstract	
Dedication	
Acknowledgement	
Table of Contents	
List of Figures	
List of Tables	
List of Abbreviations	
List of Appendices	
CHAPTER 1: INTRODUCTION	01
1.1 Overview of Tourism in Sri Lanka	01
1.2 Impact of Tourism Arrivals to Sri Lanka	01
1.3 Forecasting Tourist Arrivals	03
1.4 Main Research Problem	04
1.5 Objectives of the Research	05
1.6 Significance of the Study	05
1.7 Out Line of the Research	06
CHAPTER 2: LITERATURE REVIEW	07
2.1 Related Studies in Sri Lanka	07
2.2 Related Studies in Other Countries	10
2.3 Selection of Forecasting Models	14
2.4 Evaluation of Forecasting Accuracy	15
2.5 Summary of Chapter 2	16
CHAPTER 3: MATERIALS AND METHODOLOGY	17
3.1 Source of the Data	17
3.2 Basic Definition of Time Series	17
3.2.1 Time Series	17

3.2.2 Stationary Process	17
3.2.3 Strictly Stationary	17
3.2.4 Weekly Stationary	18
3.2.5 Purely Random Process	18
3.3 Transformation of Non Stationary Series to Stationary Series	18
3.4 Unit Root Test	19
3.5 Kruskal Wallis Test	19
3.6 Box Jenkins ARIMA Models	20
3.7 ARMA Models	20
3.8 Auto correlation Function and Partial Auto Correlation Function	22
3.9 Exponential Smoothing Techniques	23
3.9.1 Single Exponential Smoothing	24
3.9.2 Double Exponential Smoothing	24
3.9.2 Holt Winters Three parameter model	25
3.10 Determination of values of smoothing parameters	27
3.11 Diagnostic Error Terms	28
3.11.1 Normality of Error Terms	28
3.11.2 Ljung and Box Q Statistics	28
3.11.3 Serial Correlation	29
3.11.4 Heteroscedascity	29
3.11.5 Information Criteria	30
3.11.6 Percentage Errors	30
3.11.7 Mean Absolute Percentage Error (MAPE)	30
 CHAPTER 4: DEVELOPMENT OF ARIMA MODEL	 32
4.1 Temporal Variability of Tourist Arrivals	32
4.2 Summary Statistics of Tourist Arrivals	34
4.3 Distribution of Tourist Arrivals	35
4.4 Test for Seasonality	36
4.5 ACF and PACF for UK and India	37
4.6 Stationarity of Time Series	39
4.7 Time Series Plots of First Differenced Series	40

4.8 Stationarity of First Differenced Series	41
4.9 Model Identification for UK Series	43
4.10 Identification of Parsimonious Models for United Kingdom	47
4.10.1 Comparison of Parsimonious models for United Kingdom	47
4.10.2 Estimation of Parameters	48
4.10.3 Diagnostic Checking	49
4.10.4 Forecasting for an independent set	53
4.10.5 Forecasting for the training set	55
4.11 Model Identification for India Series	59
4.12 Identification of Parsimonious Models for India	63
4.12.1 Comparison of Parsimonious Models for India	63
4.12.2 Estimation of Parameters	64
4.12.3 Diagnostic Checking	65
4.12.4 Forecasting for an independent set	69
4.12.5 Forecasting for the training set	71
4.13 Summary of Chapter 4	74
CHAPTER 5: DEVELOPMENT OF SMOOTHING TECHNIQUES	75
5.1 Use of Holt Winters Multiplicative Model for UK	75
5.1.1 Diagnostic Checking for Multiplicative model for UK	77
5.2 Use of Holt Winters Additive Model for UK	83
5.2.1 Diagnostic Checking for Additive model for UK	84
5.3 Use of Holt Winters Multiplicative Model for India	90
5.3.1 Diagnostic Checking for Multiplicative model for India	91
5.4 Use of Holt Winters Additive Model for India	97
5.4.1 Diagnostic Checking for Additive model for India	98
5.5 Development of Equivalent ARIMA Models of GES Model	105
5.5.1 ARIMA equivalency of Holt Winters Additive Model for UK data	105

5.5.2 ARIMA equivalency of Holt Winters Additive Model for India data	107
5.6 Comparison of SARIMA model and Holt Winters Multiplicative Model	108
5.7 Summary of Chapter 5	110
CHAPTER 6: CONCLUSION AND FUTURE STUDIES	111
6.1 Conclusion	111
6.2 Future Studies	112
Reference List	113
Appendix A: Monthly tourist arrivals from UK and India	119
Appendix B: ACF plots of residuals for six SARIMA models postulated for UK data	121
Appendix C: ACF plots of residuals for six SARIMA models postulated for India data	124
Appendix D: Residuals of Seasonal ARIMA (0, 1, 1) (0, 1, 1) model – UK	127
Appendix E: Residuals of Seasonal ARIMA (0, 1, 1) (0, 1, 1) model – India	129

LIST OF FIGURES

	Page
Figure 4.1: Time series plot of monthly tourist arrivals from United Kingdom	32
Figure 4.2: Time series plot of monthly tourist arrivals from India	33
Figure 4.3: Normality plot for UK	35
Figure 4.4: Normality plot for India	36
Figure 4.5: Correlogram of original series for United Kingdom	37
Figure 4.6: Correlogram of original series for India	38
Figure 4.7: Time Series plot of first differenced series of UK	40
Figure 4.8: Time Series plot of first differenced series of India	41
Figure 4.9: ACF of first difference series for UK	43
Figure 4.10: ACF of 12 th difference of 1 st difference series for UK	44
Figure 4.11: Corrologram of 12 th difference of 1 st difference series for UK	45
Figure 4.12: ACF of residuals for United Kingdom	49
Figure 4.13: Correlogram of Squared residuals	50
Figure 4.14: Normality plot for residuals of UK	53
Figure 4.15: The plot of forecasts of monthly tourist arrivals from UK	58
Figure 4.16: ACF of first difference series for India	59
Figure 4.17: ACF of 12 th difference of 1 st difference series for India	60
Figure 4.18: Corrologram of 12 th difference of 1 st difference series for India	61
Figure 4.19: ACF of residuals for India	65
Figure 4.20: Correlogram of Squared residuals	66
Figure 4.21: Normality plot for residuals of India	69
Figure 4.22: The plot of forecasts of monthly tourist arrivals from India	74

LIST OF FIGURES (CONTINUED)

	Page
Figure 5.1.: ACF of residuals of multiplicative model for UK	77
Figure 5.2.: Normality plot of residuals for model of UK	78
Figure 5.3: The plot of standardized residuals vs standardized predicted values	79
Figure 5.4: ACF of residuals of additive model for UK	84
Figure 5.5: Normality plot of residuals for model of UK	85
Figure 5.6: The plot of standardized residuals vs standardized predicted values	86
Figure 5.7: ACF of residuals for multiplicative model of India	91
Figure 5.8: Normality plot of residuals for model of India	92
Figure 5.9: The plot of standardized residuals vs standardized predicted values	93
Figure 5.10: ACF of residuals for additive model of India	98
Figure 5.11: Normality plot of residuals for model of India	99
Figure 5.12: The plot of standardized residuals vs standardized predicted values	100

LIST OF TABLES

	Page
Table 1.1: Composition of tourist arrivals in Sri Lanka	02
Table 1.2: Percentage changes of tourist arrivals by region	03
Table 1.3: Tourist arrivals by country in 2016	04
Table 3.1: Identification of AR and MA terms using ACF and PACF	23
Table 3.2: Details of measure of MAPE	31
Table 4.1: Summary statistics of tourist arrivals	34
Table 4.2: Results of Kruskal-Wallis test	36
Table 4.3: Results of unit root test for original series of UK	39
Table 4.4: Results of unit root test for original series of India	39
Table 4.5: The results of unit root test for 1 st difference series of UK	41
Table 4.6: The results of unit root test for 1 st difference series of India	42
Table 4.7: Results of ADF test for 12 th difference of 1 st differenced series - UK	44
Table 4.8: Comparison of postulated models for the United Kingdom	47
Table 4.9: Comparison of suggested ARIMA models	48
Table 4.10: Results of Heteroskedasticity test (ARCH LM test)	51
Table 4.11: Breusch-Godfrey Serial Correlation LM test	52
Table 4.12: Forecasting for training set	55
Table 4.13: Comparison of actual values and forecasts	57
Table 4.14: Results of ADF test for 12 th difference of 1 st differenced series –India	60
Table 4.15: Comparison of postulated models for the India	63
Table 4.16: Comparison of suggested ARIMA models	64
Table 4.17: Results of Heteroskedasticity Test (ARCH LM test)	67

LIST OF TABLES (CONTINUED)

	Page
Table 4.18: Breusch-Godfrey Serial Correlation LM Test	68
Table 4.19: Forecasts for training set	71
Table 4.20: Comparison of actual values and forecasts	73
Table 4.21: Values of MAPE criterion	74
Table 5.1: Determination of smoothing constants of the model for UK	76
Table 5.2: Forecasts for training set using multiplicative method	80
Table 5.3: Comparison of actual values vs forecasts for arrivals from UK (Holt Winters multiplicative model)	82
Table 5.4: Determination of smoothing constants of additive model for UK	84
Table 5.5: Forecasts for training set using additive model for UK	87
Table 5.6: Comparison of actual values vs forecasts for arrivals from UK (Holt Winters additive model)	89
Table 5.7: Determination of smoothing constants of multiplicative model for India	90
Table 5.8: Forecasts for training set using multiplicative model for India	94
Table 5.9: Comparison of actual values versus forecasts for India	96
Table 5.10: Determination of smoothing constants of additive model for India	97
Table 5.11: Forecasts for training set using additive model for India	101
Table 5.12: Comparison of actual values vs forecast values for India	103
Table 5.13: Comparison between multiplicative and additive model- UK	104
Table 5.14: Comparison between multiplicative and additive model-India	104
Table 5.15: Estimation of parameters	106
Table 5.16: Estimation of parameters	107
Table 5.17: Actuals and forecasts for UK series using two models	108

LIST OF TABLES (CONTINUED)

	Page
Table 5.18: Actuals and forecasts for India series regarding two models	109
Table 5.19: Forecasts of monthly tourist arrivals using SARIMA model (From March, 2018 to July, 2018)	110
Table 5.20: Percentage Changes in future monthly tourist arrivals	110

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

Appendix	Description	Page
Appendix A:	Monthly tourist arrivals from UK and India	117
Appendix B:	ACF plots of residuals for six SARIMA models postulated for UK data	119
Appendix C:	ACF plots of residuals for six SARIMA models postulated for India data	122
Appendix D:	Residuals of Seasonal ARIMA (0, 1, 1) (0, 1, 1) model – UK	125
Appendix E:	Residuals of Seasonal ARIMA (0, 1, 1) (0, 1, 1) model – India	130