Journal of Tourism, Hospitality and Sports www.iiste.org
ISSN (Paper) 2312-5187 ISSN (Online) 2312-5179 An International Peer-reviewed Journal
Vol.29, 2017

Forecasting Tourist Arrivals to Sri Lanka Using Seasonal ARIMA

Asitha Kodippili
Department of Mathematics and Computer Science, Fayetteville State University, USA.
akodippili@uncfsu.edu

Deepthika Senaratne
Department of Mathematics and Computer Science, Fayetteville State University, USA.
dsenaratne@uncfsu.edu

Abstract
The benefits of accurate forecasts of international tourist arrivals in short and long term policy planning are well documented in tourism research literature. Forecast of the tourist arrivals can provide reliable estimates of tourist related product and services such as hotel rooms, car rentals, and a number of visitors to places of tourist attractions, etc. In particular, short term tourist arrival forecasts are very vital for optimal pricing of tourist related services such as air tickets and hotel rooms. However, we can find only a few research articles dealing with short term forecasting of international tourist arrivals to Sri Lanka. Therefore, we attempted to forecast monthly tourist arrivals to Sri Lanka using customized model building procedure, statistical time series modeling techniques. We found that general seasonal ARIMA model ARIMA$(4,0,0)\times(0,1,0)_{12}$ as a suitable model to forecast monthly international tourist arrivals to Sri Lanka.
Keywords: tourism demand, Sri Lanka, forecasting, SARIMA

1. Introduction

The tourism sector has experienced rapid growth and gained importance for the economies of some developed countries as well as many developing countries. The number of tourists or tourist arrivals has significant impact on Sri Lankan economy too (WTTC, Sri Lanka 2015). The tourists' demand or tourist consumption contributes to Gross Domestic Product (GDP), increasing the employment rate, making new source of revenue for local people, public and private sector participants and the government of Sri Lanka. For instance, according to World Travel & Tourism Council (WTTC), in 2014 the total contribution of Travel & Tourism to GDP was LKR 1,067.4bn (11.1% of GDP) and the total contribution to employment, including jobs indirectly supported by the industry was 10.0% of the total employment (819,500 jobs). This significant impact of tourism industry to Sri Lankan economy is more than enough to encourage researches to investigate on the number of international tourist arrivals and attempt to make a more accurate prediction for future planning. However, attempts at comprehensive research with the aim of modeling and forecasting monthly international tourist arrivals to Sri Lanka is very rare in the literature (Lelwala et al., 2008).

Tourism in Sri Lanka expanded rapidly after 1966. The main attractions were beach resorts followed by ancient heritage sites, wild life sanctuaries and mountainous region dominated by tea plantations. From 1966 to 1982, the number of tourist arrivals to Sri Lanka had been growing steadily following the trend of worldwide tourism. There were approximately 407,230 tourist arrivals in 1982. Almost three decades of separatist war combined with political violence adversely affected tourism in Sri Lanka, with arrival figures declining year on year. The total arrivals were 184,732 in 1989, down 54.6% from 1982. From 1990 to 2009, tourist arrivals increases slowly but in very unpredictable manner due random terrorist activities and Indian Ocean earthquake and tsunami in 2004. Sri Lanka witnessed a strong upsurge in tourism after the end of civil war in 2009 (Fernando et al., 2012). The data provides the tourist arrivals from 1999 to 2016 (Figure 1), clearly showing the rise in tourist arrivals since 2009 over the rest of the period. The tourist industry of Sri Lanka expects this trend to continue for many more years and has set a target of over 4 million tourist arrivals in year 2020.

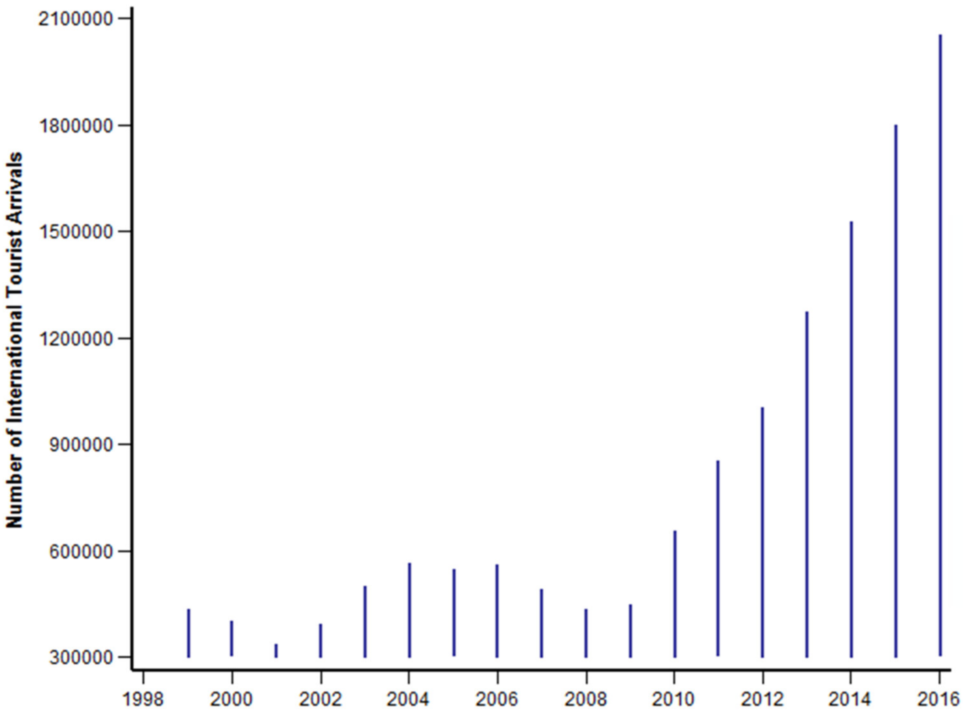

Figure 1: International Tourist Arrivals to Sri Lanka 1999-2016

2. Objective of the Study

This research is conducted to determine a suitable statistical model to predict short term forecasts of international tourist arrivals to Sri Lanka. A number of quantitative techniques have been applied to tourism demand forecasting over the past few decades. These methods include economic models, time series analysis, and non-linear modeling approaches. We explore the use of time series model auto regressive integrated moving average (ARIMA) a generalization of the ARMA (Box et al., 1970) model to forecast monthly international tourist arrivals to Sri Lanka. We selected the ARIMA model due its wide use in tourist demand forecasts and relatively reliable performance (Vu and Turner, 2006; Chu, 2009).

3. Literature Review

Tourism demand is usually measured in terms of the number of international tourist arrivals, in terms of tourist night spends, or official tourist expenditure (receipts), etc. (Josip and Ivan, 2015). Data generally used are monthly, annual arrivals, or cross-sectional data.

Quantitative methods that have been applied to tourism demand forecasting include econometric and time series methods. Most of the econometric studies involve the use of regression analysis to estimate the quantitative relationship between tourism demand and multitude of explanatory variables such as exchange rates between the two counties (tourist destination and country of the tourist origin), income (of origin country), cost of living in two countries etc. Econometric models have increased our understanding of relationship between the tourism demand and its determinants. However, econometric models have been shown to generate relatively less accurate forecast of tourist demand. This poor performance of econometric models may be a result of various factors; tourism demand is complex phenomenon which is very sensitive to catastrophic influences (terrorist activities, natural disasters, etc.).

Time series models use pattern (trend, seasonality etc.) in historical data to predict future values of a time series variable. Time series models have been widely used for tourism forecasting with the dominance of integrated autoregressive moving average (ARIMA) proposed by Box and Jenkins (Box et al., 1970) and seasonal ARIMA (i.e. SARIMA) later gaining an increasing popularity over the last decade. These methods are often more accurate in forecasting tourist demand than econometric models (Vu and Turner, 2006; Chu, 2009; Kulendran and Witt, 20011; Lim and McAleer, 2002).

4. Model of Tourist Arrivals

The ARIMA procedure analyzes and forecasts equally spaced univariate time series data, transfer function data, and intervention data by using the autoregressive moving-average (ARMA) model or autoregressive integrated moving average $\mathrm{ARIMA}(p, d, q)$ model given by:

$$\varphi(B)(1-B)^d y_t = c + \theta(B)z_t,$$

where $\{z_t\}$ is a white noise process with mean 0 and variance σ^2, B is the backshift operator, and $\emptyset(B)$ and $\theta(B)$ are polynomials of orders p and q with no common factors and d is the order of differencing. To deal with series containing seasonal fluctuations, Box-Jenkins recommended the seasonal ARIMA $(p, d, q) \times (P, D, Q)_s$ process given by:

$$\Phi(B^s)\ \varphi(B)(1-B^s)^D(1-B)^d y_t = c + \Theta(B^s)\ \theta(B)z_t,$$

where $\Phi(z)$ and $\Theta(z)$ are polynomials of orders P and Q respectively, each containing no roots inside the unit circle.

The main task in ARIMA forecasting is to select the appropriate model order; that is, the values of d, D, p, q, P and Q. We use SAS Enterprise (Flynn et al., 2014) Guide to forecast international tourist arrivals to Sri Lanka. We manually select the best values for d and D using auto-correlation function (ACF) and augmented Dickey-Fuller unit root test. Once d and D are known, we select the orders p, q, P and Q using the ACF, partial autocorrelation function (PACF), and the Akaike's Information Criterion.

5. Tourism Demand Forecasting

We use monthly international tourist arrivals data publish by the Sri Lanka Tourism Development Authority (http://www.sltda.lk) in its various monthly bulletins.

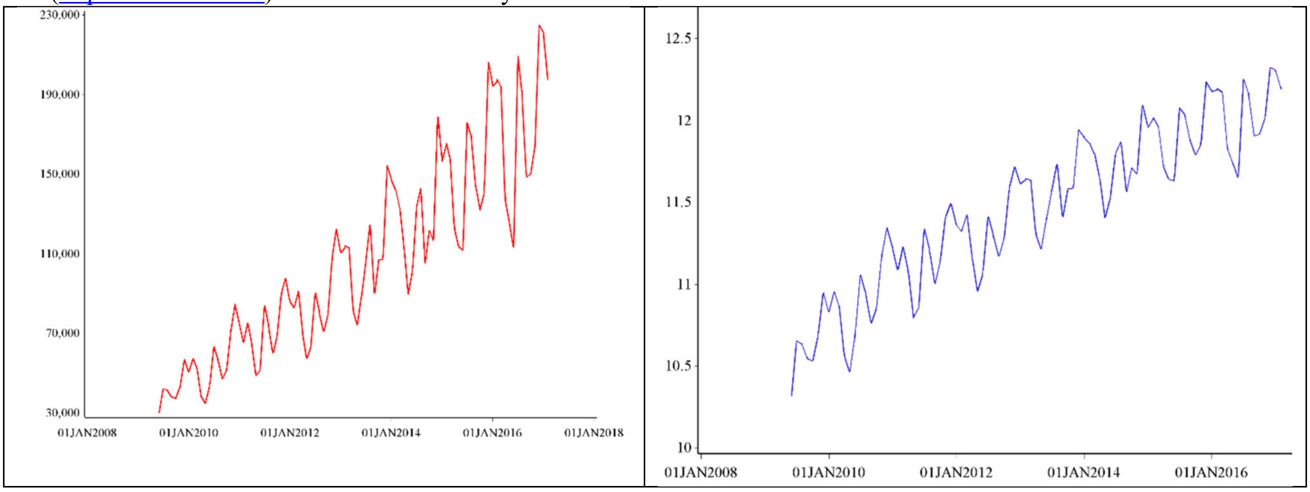

Figure 2: Number of International Tourist Arrivals x_t to Sri Lanka from June 2009 – March 2017

It appears from the graph (Figure 2 Left) that the tourist arrivals have an upward trend and a seasonal pattern with a peak in December and trough in May. The magnitude of the fluctuations appear to grow roughly linearly with the level of the series suggesting natural logarithm transformation. Let us denote the transformed series by $y_t = Ln(x_t)$. Transformed series (Figure 2 Right) shows that the increasing variability with the arrival level is reduced by natural logarithm of the original series.

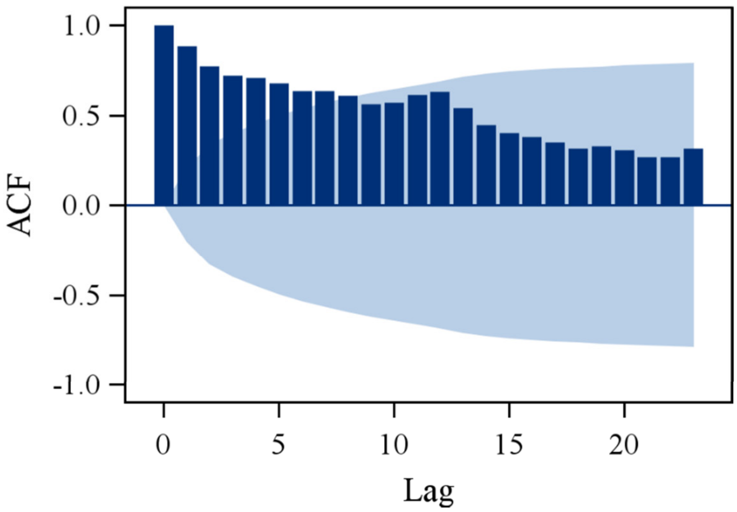

Figure 3: Auto Correlation Function of natural logarithm of Tourist arrivals

Table 1: Seasonal Unit Root Test for natural logarithm of tourist arrivals

Seasonal Augmented Dickey-Fuller Unit Root Tests						
Type	Lags	Rho	Pr < Rho	Tau	Pr < Tau
Zero Mean	0	1.4220	0.6487	15.61	0.9999
	1	4.2135	0.8300	3.52	0.9997
	2	3.0702	0.7618	1.86	0.9730
Single Mean	0	-11.9590	0.0725	-7.19	<.0001
	1	-12.1511	0.0695	-2.69	0.0113
	2	-7.9614	0.1697	-2.00	0.0516

Visual inspection of **auto correlation function plot of y_t (Figure 3)** indicates the natural logarithm of arrivals series is non-stationary (ACF decays very slowly). Also Seasonal Augmented Dickey-fuller test (Table 1) fail to reject the unit root hypothesis. Therefore, we need to consider seasonal differencing $(1-B^s)y_t$ with $s = 12$, where $s \equiv$ period of seasonality.

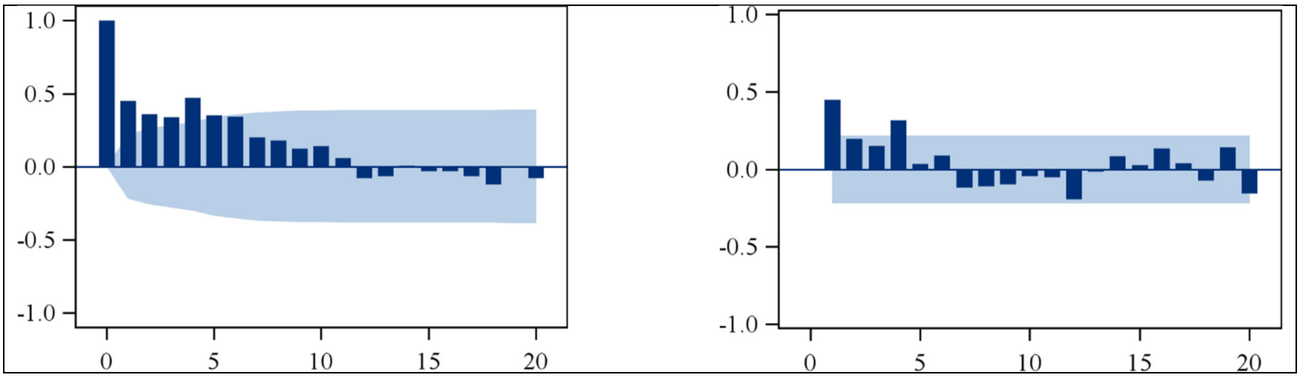

Figure 4: ACF & PACF of Tourist arrivals after taking natural logarithm and differencing at lag 12

Visual inspection of **auto correlation function plot of** $(1-B^s)y_t$ (Figure 4 center) and results of Augmented Dickey-fuller test (results not included) confirm that the further differencing is not required.

Table 2: Auto correlation check for White Noise for the $(1-B^{12})Y_t$

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	79.50	6	<.0001	0.449	0.360	0.338	0.471	0.351	0.343
12	90.65	12	<.0001	0.204	0.179	0.126	0.142	0.062	-0.075
18	93.18	18	<.0001	-0.064	0.010	-0.029	-0.029	-0.062	-0.119

Auto correlation check for white noise (Table 2) of the seasonally difference series reject the white noise hypothesis and therefore ARMA modeling is possible. The ACF gradually damp to zero and PCFA has spikes at lag 1 and 4 (see Figure 4), hence seasonally difference series $(1-B^{12})Y_t$ suggests a non-seasonal autoregressive component. Adding suitable autoregressive terms, we obtained the seasonal ARIMA model with auto regressive factors $1 - 0.3496B - 0.43664B^4$. The parameter estimate of the fitted model is given in the Table 3 and clearly all of them are significantly different from zero.

Table 3: Estimated parameters of model fitted to $(1-B^{12})Y_t$

Conditional Least Squares Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > |t|	Lag
MU	0.22496	0.03722	6.04	<.0001	0
AR1,1	0.34960	0.09853	3.55	0.0007	1
AR1,2	0.43664	0.10343	4.22	<.0001	4

The correlation and normal probability plot of residuals are given in Figure 5 and Figure 6,respectively. They satisfy all the requirements of residuals (white noise and normality) of an excellent fitted model to the data with a very low **AIC = -160.63** value.

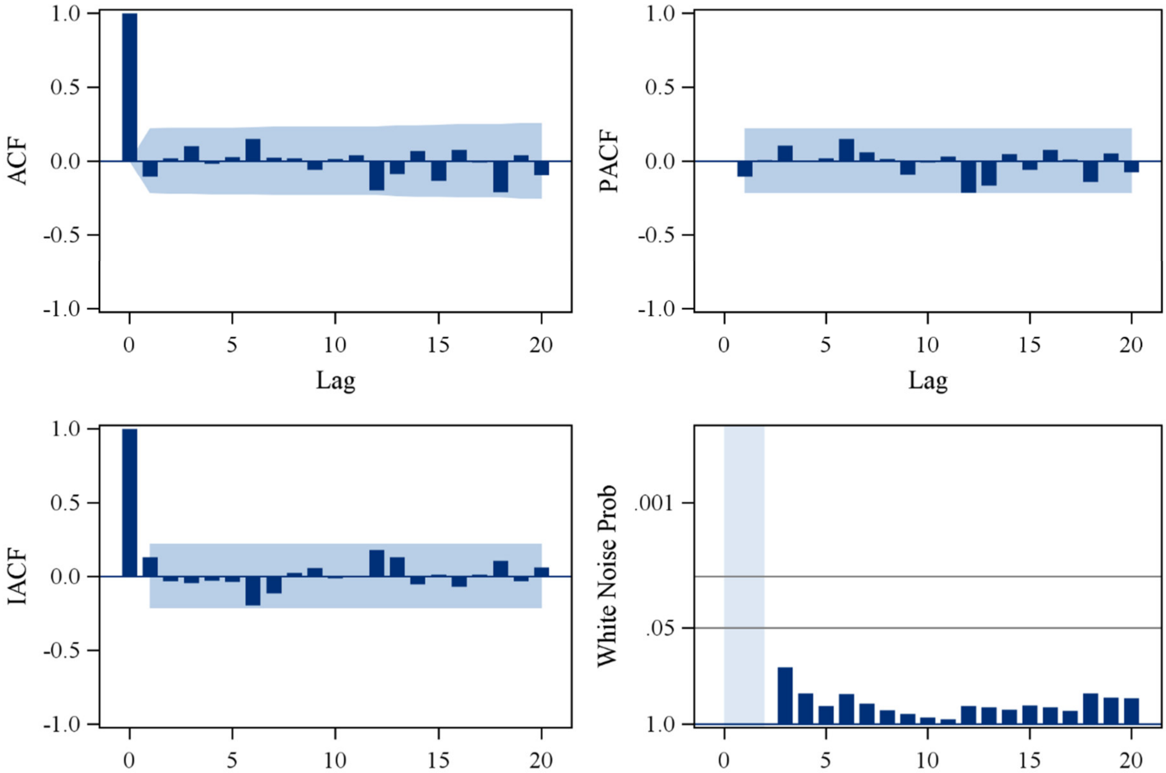

Figure 5: Residual correlation diagnostics for the fitted model

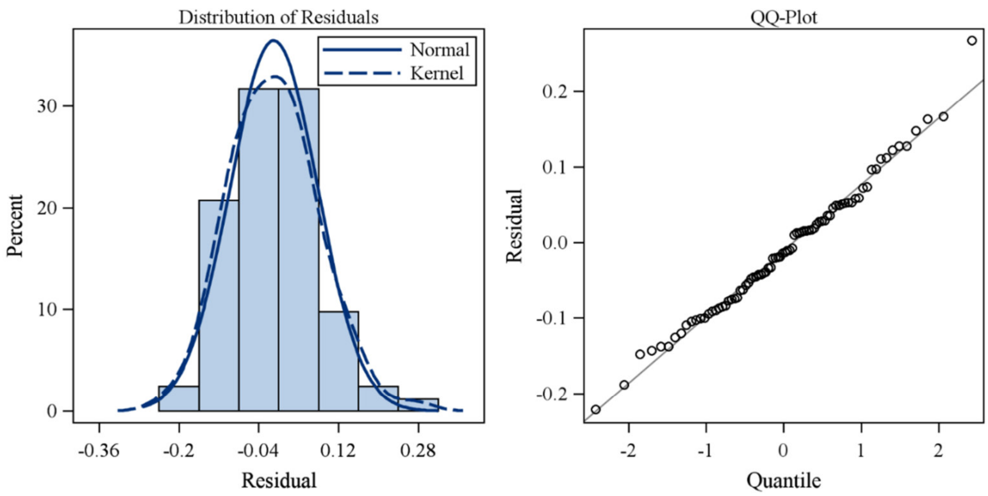

Figure 6: Residual Normality Diagnostics for the fitted model

6. Conclusion

In this study, we investigate seasonal ARIMA model and its usefulness in forecasting monthly international tourist arrivals to Sri Lanka. We performed the analysis using monthly tourist arrival data from June 2009 to March 2017 and predict the monthly tourist arrivals for the one year period April 2017 – March 2018. The residual diagnostic tests shows that the model $\text{ARIMA}(p, d, q) \times (P, D, Q)_s$ or precisely $\text{ARIMA}(4, 0, 0) \times (0, 1, 0)_{12}$ is an excellent fit to the data and therefore highly reliable forecasts. The forecast and 90% confidence intervals for international tourist arrivals to Sri Lanka for the next 12 months based on the fitted model is given in Table 4 and Figure 7

.

Table 4: Prediction of tourist arrivals to Sri Lanka based on the fitted model

Month	Forecast	90% Confidence Limits	
April 2017	147902	127200	170707
May 2017	142769	121662	166048
June 2017	130062	110712	151415
July 2017	225437	191873	262480
Aug. 2017	208249	174735	245552
Sep. 2017	171987	143390	203936
Oct. 2017	173652	144533	206218
Nov. 2017	190901	158830	226774
Dec. 2017	259798	215463	309482
Jan. 2018	258562	213821	308786
Feb 2018	234090	193328	279882
March 2018	222237	183460	265812

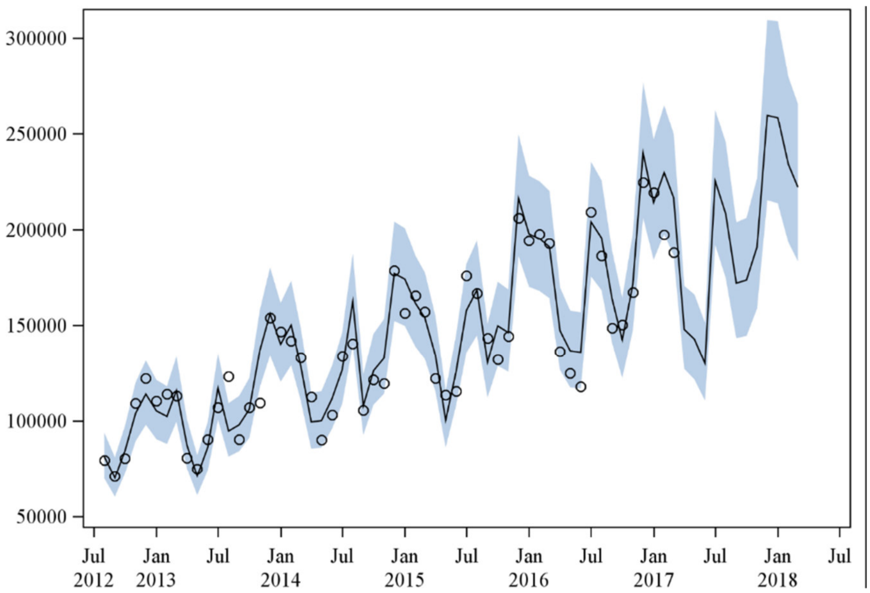

Figure 7: Monthly Tourist Arrival forecast and 90% Confidence Interval

Sri Lanka attracts tourists from almost all regions of the world. The expectations and spending habits of tourists highly depend on the region. Therefore, our future research will focus on forecasting monthly tourist arrivals from different regions of the world in order to obtain the optimum economic benefits to the country.

References

Box, G., and G. Jenkins (1970). *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco, CA.

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

Fernando, S., S. B. Jayatilleke, and Smith. C. (2012). Regaining Missed Opportunities: The Role of Tourism in Post-war Development in Sri Lanka. *Asia Pacific Journal of Torisum Research.* 18(7), pp. 685-771.

Flynn, J., C. Smith, and A. Song (2014). "Check It Out! Versioning in SAS® Enterprise Guide®." Proceedings of the SAS Global 2014 Conference, Cary, NC: SAS Institute Inc.

Josip, T., and K. Ivan (2015). Forecasting Croatian inbound tourism demand, *Economic Research*, 28:1, 1046-1062.

Kulendran, N., and S. F. Witt (2001). Cointegration versus least squares regression, Annals of Tourism Research, 28(2), pp. 291-311.

Lelwala, E. I., and L. H. P. Gunaratne (2008). Modelling Tourism Demand using Cointegration Analysis: A

Case Study for Tourists Arriving from United Kingdom to Sri Lanka, *Tropical Agricultural Research*, Vol 20, pp. 50-59.

Lim, C., and M. McAleer (2002). Time series forecasts of international travel demand for Australia, Tourism Management, 23(4), 389-96.

Vu, J.C., and L. W. Turner (2006), "Regional data forecasting accuracy: the case of Thailand", *Journal of Travel Research*, Vol 45, pp. 186-193.

Travel & Tourism, Economic Impact 2015 Sri Lanka, *World Travel & Tourism Council*.