Open Access Repository
www.ssoar.info

Modeling and predicting foreign tourist arrivals to Sri Lanka: A comparison of three different methods

Diunugala, Hemantha Premakumara; Mombeuil, Claudel

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:
Diunugala, H. P., & Mombeuil, C. (2020). Modeling and predicting foreign tourist arrivals to Sri Lanka: A comparison of three different methods. *Journal of Tourism, Heritage & Services Marketing*, *6*(3), 3-13. https://doi.org/10.5281/zenodo.4055960

Nutzungsbedingungen:
Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
https://creativecommons.org/licenses/by-nc-nd/4.0/deed.de

Terms of use:
This document is made available under a CC BY-NC-ND Licence (Attribution-Non Comercial-NoDerivatives). For more Information see:
https://creativecommons.org/licenses/by-nc-nd/4.0

gesis
Leibniz-Institut
für Sozialwissenschaften

Mitglied der
Leibniz-Gemeinschaft

Diese Version ist zitierbar unter / This version is citable under:
https://nbn-resolving.org/urn:nbn:de:0168-ssoar-69874-0

Modeling and predicting foreign tourist arrivals to Sri Lanka: A comparison of three different methods

Hemantha Premakumara Diunugala
Beijing Institute of Technology, China

Claudel Mombeuil
Beijing Institute of Technology, China

Abstract:
Purpose: *This study compares three different methods to predict foreign tourist arrivals (FTAs) to Sri Lanka from top-ten countries and also attempts to find the best-fitted forecasting model for each country using five model performance evaluation criteria.*
Methods: *This study employs two different univariate-time-series approaches and one Artificial Intelligence (AI) approach to develop models that best explain the tourist arrivals to Sri Lanka from the top-ten tourist generating countries. The univariate-time series approach contains two main types of statistical models, namely Deterministic Models and Stochastic Models.*
Results: *The results show that Winter's exponential smoothing and ARIMA are the best methods to forecast tourist arrivals to Sri Lanka. Furthermore, the results show that the accuracy of the best forecasting model based on MAPE criteria for the models of India, China, Germany, Russia, and Australia fall between 5 to 9 percent, whereas the accuracy levels of models for the UK, France, USA, Japan, and the Maldives fall between 10 to 15 percent.*
Implications: *The overall results of this study provide valuable insights into tourism management and policy development for Sri Lanka. Successful forecasting of FTAs for each market source provide a practical planning tool to destination decision-makers.*

Keywords: *foreign tourist arrivals, winter's exponential smoothing, ARIMA, simple recurrent neural network, Sri Lanka*

JEL Classification: *C5, Z32,C45*

Biographical note: Hemantha Premakumara Diunugala (3820182090@bit.edu.cn) is affiliated to School of Management and Economics of Beijing institute of technology. He is a senior lecturer attached to the department of social statistics at the University of Sri Jayewardenepura Sri Lanka where he has been working teaching for more than 18 years. His area of expertise includes tourism economics, regression analysis, time series analysis, and modeling tourist arrivals.Claudel Mombeuil (claudelmombeuil@bit.edu.cn) is affiliated to the School of Management and Economics of Beijing institute of technology. He is particularly interested in interdisciplinary study in the field of management and marketing. His publications appeared in the Journal of Cleaner Production, Social Responsibility Journal and Journal of Retailing & Consumer Services. Corresponding author: Claudel Mombeuil (claudelmombeuil@bit.edu.cn).

1 INTRODUCTION

Tourism is one of the world's most complex industries that has a productive relationship directly or indirectly with other industries. Also, the tourism industry has a major growth potential with significant multiplying effects, through earning foreign exchange, encompassing tax revenues, and generating employment. It is, therefore, important to have accurate tourism demand forecasts for strategic, tactical, and operational decision-making processes (Song & Li, 2008; Volgger et al., 2017; Law, Li, Fong, & Han, 2019). Forecasting models such as time series models, econometric models, artificial intelligence-based models, and judgmental models are commonly used to predict tourism demand (Song, Qiu, & Park, 2019). A broad combination of forecasting models are already available in the tourism literature (See, for example, Chu, 2009; Claveria, Monte, & Torra, 2015; Coshall, 2009; Goh & Law, 2002; Li, Wong, Song, & Witt, 2006; Salman, Shukur, & Bergmann-Winberg, 2007; Wong, Song, Witt, & Wu, 2007; Nanthakumar & Ibrahim, 2010; Sun, Sun, Wang, Zhang, & Gao, 2016; Thushara, Su, & Bandara, 2019; Van Truong et al., 2020).

In selecting from the existing forecasting methods, tourism managers need to consider, among other things, the accuracy of the forecasts generated, ease of use of the forecasting model, cost associated with the production of the forecasts, and speed with which the forecasts can be produced (Martin & Witt, 1989). Since "accuracy remains the most important characteristic of a forecast" (Martin & Witt, 1989, p. 408), finding accurate forecasting models for international tourist arrivals remains important for managers and decision-

© 2020 Authors. Published by International Hellenic University
ISSN: 2529-1947. UDC: 658.8+338.48+339.1+640(05)
http://doi.org/10.5281/zenodo.4055960
Published online: 30 October 2020
www.jthsm.gr

Some rights reserved.
Except otherwise noted, this work is licensed under
https://creativecommons.org/licenses/by-nc-nd/4.0

makers. The fact that the literature offers no evidence that one model can outperform all the other ones in all contexts and situations (Song & Li, 2008; Song et al., 2019) suggests the need to find the best forecasting models for international tourist arrivals that fits the context of tourism development in a country like Sri Lanka. While combining traditional and modern forecasting techniques to improve forecasting accuracy has become widespread in the tourism management literature (Sotiriadis & Shen, 2017; Smyl, 2020), there is an absence of studies that combine traditional and modern forecasting techniques to predicts international tourism demand for Sri Lanka.

The objective of this study is two-fold. First, three different methods to forecast inbound tourist arrivals to Sri Lanka from the top ten market sources, namely India, China, the UK, Germany, France, Maldives, Australia, Russia, the USA, and Japan are compared using monthly individual datasets spanning from 2000 to 2018. Second, the best-fitted forecasting model for tourist arrivals for each of the top ten (10) market sources is estimated. To the best of our knowledge, this is the first empirical attempt that compares Winter's exponential smoothing, ARIMA, and model of Elman and Jordan simple recurrent neural network (SRNN) methods to model FTAs to Sri Lanka on top-ten countries. As Sri Lanka had been recovering from the terrorist attacks after the war until the beginning of 2019, and becoming itself as an important hub of sustainable tourism, finding the best methods for tourist arrivals, therefore, has significant theoretical and practical implications. In the remainder of this paper, a brief literature review, the data and methods, results and discussion, and conclusions and policy implications are successively presented.

2 RESEACH PROBLEM AND METHODS

Data

To meet the objectives of this study, we use data on monthly FTAs to Sri Lanka from 2000 to 2018. We retrieved the data from the monthly tourist arrivals reports of the Sri Lanka Tourism Development Authority (See: http://www.sltda.lk/node/765) for the top-ten source markets in descending order: India, China, UK, Germany, France, Maldives, Australia, Russia, USA, and Japan, counting altogether 67% of FTAs to Sri Lanka (See Table 2). These ten source markets are of significance because they have been consistent in ranking order for the past five years running from 2014 to 2018.

Methodology

We started the analysis studying the time series plots of the FTAs of top-ten source markets by considering both, the sample selection periods and model estimation periods. The time series plots are particularly important as they provide insights on the trends of FTAs for the top-ten source markets and also help determine the model estimating periods of each source market taking into consideration the period running from 2000 to 2018. Then, we conducted the Augmented Dickey-Fuller (ADF) test (Dickey & Fuller, 1979) to confirm the stationarity of the FTAs for each of the top ten source markets. In principle, the ADF test follows a unit-root process and indicates whether a unit root exists or not. If the series does not have a unit root, the FTAs series of any country can be taken as stationary. The ADF test with the optimal lag length is determined by the Akaike information criteria (AIC).

Estimation of models

We generally use time series models when there exists a large amount of data on the variable that we want to forecast (Chu, 2004) as this often provides a more accurate forecast for tourist arrivals (Akuno, Otieno, Mwangi, & Bichanga, 2015; Chandra & Kumari, 2018; Gnanapragasam, 2018; Saayman & Saayman, 2010). Therefore, this study employs two different univariate-time-series approach and one Artificial Intelligence (AI) approach to develop models that best explain the tourist arrivals to Sri Lanka from the top-ten tourist generating countries. The univariate-time series approach contains two main types of statistical models, namely Deterministic Models and Stochastic Models. The Deterministic time series models consist of several simple extrapolation techniques, which include various types of trend models, moving average models, smoothing techniques, and seasonal adjustment models. There are, in principle, four types of Stochastic time series models, namely AR, MA, ARMA, and ARIMA (Box & Jenkins, 1970). This study applies exponential smoothing, ARIMA and SRNN techniques, which are the most commonly, used deterministic, stochastic and ANN models that come under the univariate-time series approach for forecasting tourist arrivals (Cho, 2003b; Gnanapragasam, 2018; Nanthakumara & Ibrahim, 2010; Teixeira & Fernandes, 2012).

Estimation of exponential smoothing models

The exponential smoothing methods (ESM) are average data (observation) calculated using the exponential technique. However, not all observations entering into the averaging process receive equal weight as the ESM suggests that we attribute higher weight to the most recent observations. That is to say, the weighting of the observations is exponentially declined with the most recent observation having the highest weight while those further back having progressively lesser weight. In principle, the various type of ESM is differentiated by the treatment attributed to the trend and seasonality. For example, the underlying assumption for the simple exponential smoothing method (SESM) is that there is no trend and the forecast from this method is almost horizontal after adjusting the weighted-average of the observations. Holt's exponential smoothing method is an extension of the SESM, which is calculated by adding a growth factor (a trend factor) to the smoothing equation to adjust the trend. Winters' multiplicative exponential smoothing model (WMESM) is another extension of the SESM that can be used when the observations that exhibit both trend and seasonality. Therefore, this study applied WMESM as an exponential smoothing method of estimating the models because these series were contained both trend and seasonality which could be found studying the time series plots for all series of selected estimation periods. To capture the multiplicative effects in the levels of the series, it is important to apply logarithmic transformations for each series (Lim & McAleer, 2001a). To find the optimal values and logarithms for α, γ, and δ parameters of WMESM we used E-views (version 10).

$$\hat{Y}_{t+1}(t) = \left[a(t) + b(t)\right] * Sn_{n+1}(t+1-L) + \varepsilon_t, \qquad (1)$$

$\hat{Y}_{t+1}(t)$ is the forecast for the period denotes *t+1*; *a*(*t*) is the smoothed estimate for the level at period *t*; *b*(*t*) the smoothed estimate for the slope at period *t*; $Sn_{n+1}(t+1-L)$ is the smoothed estimate for the $(t+1)^{th}$ season for the period denotes *t*+1-*L*, and *L* represents the periodicity of the seasonality. To update the smoothing components of the series, we use, in general, the following expressions by adding the main components of equation (1).

$$\left.\begin{aligned} a(t) &= \alpha\left[\frac{Y_t}{Sn_t(t-L)}\right] + (1-\alpha)\left[a(t-1)+b(t-1)\right], \\ b(t) &= \gamma\left[a(t)-a(t-1)\right] + (1-\gamma)b(t-1) \text{ and} \\ Sn_{n+1}(t) &= \gamma\left(\frac{Y_t}{a(t)}\right) + (1-\delta)Sn_t(t-L), \end{aligned}\right\} \qquad (2)$$

where *α* represents the weighting factor for the level; *γ* the weighting factor for the trend; and *δ* the weighting factor for seasonal components. The Holt-Winters' Multiplicative model has two characteristics, namely a linear trend and a multiplicative seasonal factor. The underlying assumption of the Holt-Winters' Multiplicative model is that the seasonal swings are proportional to the level, but the errors are not. In that case, the most recent observation must be adjusted using its seasonal factor denoted as Sn_{t+1-L} when the value of *α* is large, we need to give more weight to the most recent observation to allow the level to update faster. Also, when the value of *γ* is large, we need to give more weight to the most recent estimate of the slope. Inversely, for a small value of *γ* one needs to give more weight to the previous slope estimate. Similarly, we need to give more weight to the most recent estimate of the seasonal factors for a large value of *δ*. However, we need to give more weight to previous estimates of the seasonal factor when the value of *δ* is small. In the exponential smoothing system equation given above, it is important to update each of the three smoothing parameters using its exponential smoothing equation. Hence, these smoothing equations can be adjusted by combining the components of the prediction equation from the values of the previous components.

To find the optimal values for *α*, *γ*, and *δ*, we used the 'grid search method' to retrieve the value that provides the smallest sum of squared errors (SSE). Then, the grid search selects combination *α*, *γ*, and *δ* using the method of trial and error (Cho, 2003a). In this method, the grid values start with the value of 0 and end with 1 and increase by adding the value of 0.01. The grid will thus generate 101*101*101=1,030,301 models, where the values of each parameter are ranging from 0, 0.01, 0.02, and so on until reaching 1. Based on the Akaike Information Criterion (AIC), we registered the parameter that produces the smallest SSE. Thereby, we used E-views to formulate the model for the series.

A. Estimation of ARIMA models

The ARIMA models, developed by Box and Jenkins (1970), are widely used in time series data analysis. The ARIMA combines three types of processes: auto-regression (AR), moving averages (MA) differencing to strip off the integration (I) of the series are three types of processes offered by the ARIMA models. Each of these three processes responds differently to a random disturbance. Pankratz (1983, p. 183) presented the general linear model resulting from the ARIMA as follow:

$$\phi_p(B)\Phi_{sp}(B^L)\nabla^d\nabla^{sd}Z_t = \Theta_{sq}(B^L)\theta_q(B)\varepsilon_t, \qquad (3)$$

where Z_t is the stationary data point at time point *t*; *B* the backshift operator, with $B(Z_t) = Z_{t-1}$; the seasonal periodicity; ε_t is the present disturbance at time point *t*; $\phi_p(B)$ is $(1-\phi_1 B-\phi_2-\cdots-\phi_p B)$, non-seasonal operator; $\Phi_{sp}(B^L)$ is $(1-\Phi_{1L}B-\Phi_{2L}B^L-\cdots-\Phi_{sp}B^L)$, seasonal operator; $\theta_q(B)$ is $(1-\theta_1 B-\theta_2 B-\cdots-\theta_q B)$; non-seasonal operator; $\Theta_{sq}(B^L)$ is $(1-\Theta_{1L}B-\Theta_{2L}B^L-\cdots-\Theta_{sqL}B^L)$; seasonal operator; ∇^d is $(1-B)^d$; non-seasonal differencing operator; and ∇_L^{sd} is $(1-B^L)^{sd}$; seasonal differencing operator.

The ARIMA model is represented by ARIMA (p, d, q). In this notation, the parameters inside the parentheses represent the order of (p) auto-regression, (d) for differences, and (q) for moving average in the model. Just as ARIMA notation has regular parameters, it may also have seasonal parameters. In a multiplicative seasonal ARIMA (SARIMA) model, the seasonal components of the ARIMA model are denoted by ARIMA $(P, D, Q)^s$; where capitalized letters respectively represent the seasonal components of the model and 's' indicates the order of periodicity or seasonality. Equivalently, SARIMA is an extension of ARIMA, where seasonality in the data is accommodated using seasonal difference. The full formulation of a multiplicative seasonal ARIMA model as the general form ARIMA (p, d, q) $(P, D, Q)^s$. The parentheses enclose the non-seasonal and seasonal factor parameters, respectively. The parameters enclosed indicate the order of the model.

The logarithmic transformations are also used to enable the log-linear model to capture the multiplicative effects in the levels of each series (Lim & McAleer, 2001b),. Therefore, we estimate the appropriate ARIMA model using levels data series and logarithmic data series. It is important to note that the suitable model specification varies with the characteristics of the time series under investigation. To search for a suitable time series stochastic scheme, a stepwise specification process is needed. In this analysis, at first, a series is transformed into a condition of stationary and then in four steps i.e. identification, estimation, diagnostic checking, and forecasting were completed. This is a procedure suggested by Box and Jenkins (1970). Through these procedures, one can estimate an ARIMA model using econometrics packages like in E-views version 10. The best-fitted model is selected based on the AIC.

B. Estimation of neural network models

An Elman network (EN) is a Neural Networks (NN) based architecture developed by Elman (1990), which is often employed to forecast tourist arrivals (Cho, 2003a). Similarly to the other types of network, the Elman network is also constructed on a three-layer network but it possesses an additional set of context units that allow the output of the hidden layers is feedback to itself (Claveria et al., 2015) which can, therefore, use time-series analyses (Trippi & Turban, 1996). The hidden layer is connected to the context units with a fixed weight value of 1. At each time step, the input is transmitted in a normal feed-forward mode which then allows the application of a backpropagation type of learning rule. The output of the network, presented in equation 4, is a scalar function of the output vector of the hidden layer:

$$y_t = \beta_0 + \sum_{j=1}^{q} \beta_j z_{j,t} \tag{4}$$

$$z_{j,t} = g(\sum_{i=1}^{p} \phi_{ij} x_{t-i} + \delta_{ij} z_{j,t-1})$$

$$\{x_{t-i} = (1, x_{t-1}, x_{t-2}, ..., x_{t-p}), i = 1, ..., p\}$$

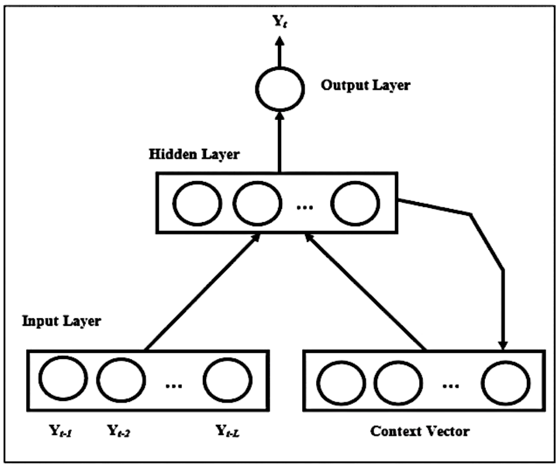
$$\{\phi_{ij}, i = 1, ..., p, j = 1, ...q\}$$

$$\{\beta_j, \mathrm{j} = 1, ..., \mathrm{q}\}$$

$$\{\delta_{ij}, i = 1, ..., p, j = 1, ..., q\}$$

where y_t represents the output vector of the EN at time t and $z_{j,t}$ the output of the hidden layer neuron j at time t. In the equation 4, g denotes the non-linear function of the neurons in the hidden layer and x_{t-i} the value of the input at time $t-i$, where i indicates the memory that is to say the number of lags that are used to introduce the context of the current observation. Then, ϕ_{ij} symbolizes the weights of neuron j that link the input to the hidden layer q the number of neurons in the hidden layer. The symbol denotes β_j stand for the weights of neuron j that connect the hidden layer with the output, and the symbol denotes δ_{ij} represent the weights that are connected to the output layer and to the activation at time t.
In the study at hand, the input vector to the neural network (NN) contains the dimensionality denoted as p+1, and the value of the output denoted as y_t, represents the estimate of the time series at time t+1. We performed the training of the network using backpropagation through time, which allows the generalization of backpropagation for feed-forward networks (Claveria et al., 2015). We estimated the parameters of the Elman NN by minimizing an error cost function, which allows us to consider the entire time series. To minimize total error, we employ a gradient descent to change each weight in proportion to its derivative in line with the error by taking into consideration whether the non-linear activation functions are differentiable or not. The main issue of the gradient descent for standard recurrent architectures is the fact that the error gradients decrease exponentially with the size of the time lag, which makes it difficult to develop simple recurrent neural networks for large numbers of neuron units as this may result into scaling issues (Claveria et al., 2015). Jordan network is a three-layer recurrent neural network similar to the Elman network (Jordan, 1997). In this network, the context units (also denoted as the state layer) are fed from the output layer rather from the hidden layer and also have a recurrent connection to themselves. Both the Elman and Jordan networks are often referred to as 'Simple Recurrent Neural Networks'.

Figure 1: Elman network



Time Span for Estimation of Models

Having studied different cyclical and trend patterns through the descriptive analysis, this study selected the periods given in Table 1 for the estimation of models for tourist arrivals series from top-ten countries.

Table 1: Model estimation and prediction periods for FTAs of top-ten countries

Country	Model estimation period
India	2012(01) - 2018(09)
China	2012(04) - 2018(09)
UK	2011(10) - 2018(09)
Germany	2011(04) - 2018(09)
France	2010(01) - 2018(09)
Maldives	2009(01) - 2018(09)
Australia	2008(01) - 2018(09)
Russia	2011(04) - 2018(09)
USA	2009(07) - 2018(09)
Japan	2010(04) - 2018(09)

The FTAs series from top-ten countries to Sri Lanka during the period starting from 2000 was used as the base in selecting the periods for estimation of models.

Evaluation of competing models

Various criteria of model selection, diagnostic checking, and forecasting accuracy measurement have been applied in many past tourist arrival forecasting studies. However, our

study uses the AIC (Akaike, 1974) as the model selection criterion for WMESM and ARIMA. RMSE as the model selection criteria for SRNN. The main importance of forecasting remains accuracy (Lichtendahl & Winkler, 2020). The accuracy of the forecasting model needs to be evaluated to identify the procedure with the smallest error and the model that generally works as the best. We, therefore, used the following performance evaluation criteria to evaluate competing models: Mean Absolute Deviation (MAD), Mean-Squared Error (MSE), Root-Mean-Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and Root-Mean-Squared Percentage Error (RMSPE). The most accurate tourist arrivals forecasting model was selected by ranking the above criteria according to the order from minimum to maximum errors and obtaining the overall average rank.
The FTAs series from top-ten countries to Sri Lanka during the period starting from 2000 was used as the base in selecting the periods for estimation of models.

$$AIC = -2\ln(L) + k \quad (5)$$

Where: L is the value of the likelihood function, ln is the natural logarithm and k is the number of estimated parameters

$$MAD = \frac{\sum |Y_t - F_t|}{n} \quad (6)$$

$$MSE = \frac{\sum (Y_t - F_t)^2}{n} \quad (7)$$

$$RMSE = \sqrt{\frac{\sum (Y_t - F_t)^2}{n}} \quad (8)$$

$$MAPE = \frac{\sum |(Y_t - F_t) / \mathrm{Y}_t|}{n} * 100 \quad (9)$$

$$RMSPE = \sqrt{\frac{\sum (Y_t - F_t)^2 / F_t}{n}} \quad (10)$$

Where Yt is the actual value in period t, Ft is the forecast value in period t, and n is the number of periods used in the calculation.

Forecasting accuracy

Forecasting, in principle, is conducted to provide insights on a subject matter on the short-term, medium-term, and long-term time horizon. The short-term forecasts often cover a period of fewer than three months; the medium-term forecasts may have a time horizon ranging from three (3) months to two (2) years, and the long-term forecasts may cover a time horizon beyond two (2) years. The accuracy of the forecasted results is measured based on the value of the MAPE. Further, to evaluate the forecast on an objective basis, we used the scaled developed by Lewis (1982), which is a tool that is commonly used forecasting studies. This scale of the MAPE can have four possible levels measured in percentage, where MAPE less than 10% is considered as highly accurate, between 10-20% is considered as good; between 20-50% reasonable, and greater than 50% is considered as inaccurate.

3 EMPIRICAL RESULTS AND DISCUSSION

The results of the study are important especially for policymaking for the development and planning of the tourism sector. To have maximum economic benefits from the tourism sector, decision-makers such as managers and policymakers need accurate forecasting of foreign tourist arrivals from the top market sources. Table 2 shows the number, percentage, and rank of foreign tourist arrivals in Sri Lanka from 2014 to 2018. India and China occupy the first two places in FTAs, counting 17.6 and 11.7 percent respectively.

Table 2: The annual average of FTAs of top-ten countries from 2014 to 2018

Country	2014	2015	2016	2017	2018	Total	Annual Average	FTAs %	Rank
India	242734	316247	356729	384628	424887	1725225	345045	17.6	1
China	128166	214783	271577	268952	265965	1149443	229889	11.7	2
UK	144168	161845	188159	201879	275205	971256	194251	9.9	3
Germany	102437	115868	133275	130227	156888	638695	127739	6.5	4
France	78883	86122	96440	97282	106449	465176	93035	4.7	5
Maldives	86359	90617	95167	79371	76108	427622	85524	4.4	6
Australia	57940	63554	74496	81281	110928	388199	77640	4	7
Russia	69718	61846	58176	59191	64497	313428	62686	3.2	8
USA	39371	47211	54254	57479	75308	273623	54725	2.8	9
Japan	39136	39358	43110	44988	49450	216042	43208	2.2	10
Total Arrivals	1507153	1798380	2050832	2116407	2333796	9806568	1961314	**67%**	

Table 3 summarizes the results of the Augmented Dickey-Fuller (ADF) test with trend effects for model estimation period of Level series and logarithmic series of foreign tourist arrivals (FTAs) to Sri Lanka from the top-ten source markets.

Table 3: Results of ADF unit root tests for level and logarithmic data of FTAs

Variable	Level series of FTAs			Logarithmic series of FTAs		
	Level	First difference	Second difference	Level	First difference	Second difference
FTAs-*India*	-0.7657	-3.3978 ***	-6.4950 *	-0.4064	-3.3978 ***	-6.3560 *
FTAs-*China*	0.0003	-4.9494 *		0.0003	-4.9494 *	
FTAs-*UK*	-1.8262	-9.7533 *		-0.4267	-12.3625 *	
FTAs-*Germany*	-0.5038	-15.562 *		-0.5038	-8.4354 *	
FTAs-*France*	-2.6826	-5.3758 *		-2.6826	-5.3758 *	
FTAs-*Maldives*	0.5391	-3.6166 **		0.5392	-3.6166 **	
FTAs-*Australia*	-1.3902	-4.4210 *		-1.3902	-4.4210 *	
FTAs-*Russia*	-1.8359	-2.4846	-15.5357 *	-1.6912	-2.4846	-15.5357 *
FTAs-*USA*	-2.6138	-17.7254 *		-2.6138	-17.7254 *	
FTAs-*Japan*	-0.6121	-11.2307 *		-0.4267	-12.3625 *	

*, ** and *** denote the significance of the null hypothesis at 1%, 5%, and 10% respectively.

Under the ADF test (Table 3), we need to test the null hypothesis for the existence of a unit root in the FTAs series against the alternative hypothesis. We choose the lag-length based on the Akaike Information Criterion (AIC) after testing for first and higher-order serial correlation in the residuals. The (ADF) test indicates that the null hypothesis has a unit root that cannot be rejected on the level series and logarithmic series of FTAs at levels I(0) on top- ten countries. However, after applying the first difference, the ADF test rejects the null hypothesis for all level series and logarithmic series of FTAs at first difference except for the FTAs series of two countries: India and Maldives. Since the FTAs series of both countries do not appear to be stationary by applying the ADF test in the first differences, we perform, therefore, further ADF test at second differences. Then, rejection of the null hypothesis at second differences confirms the stationarity of both FTAs series of these two countries.

Figure 2 shows the time series plots of the FTAs of top-ten countries and the trends in FTAs for each of the top-ten countries. The time series plots were also served to model FTAs for each market source during the period running from 2000 to 2018.

Figure 2: Time series plots of FTAs for the sample period

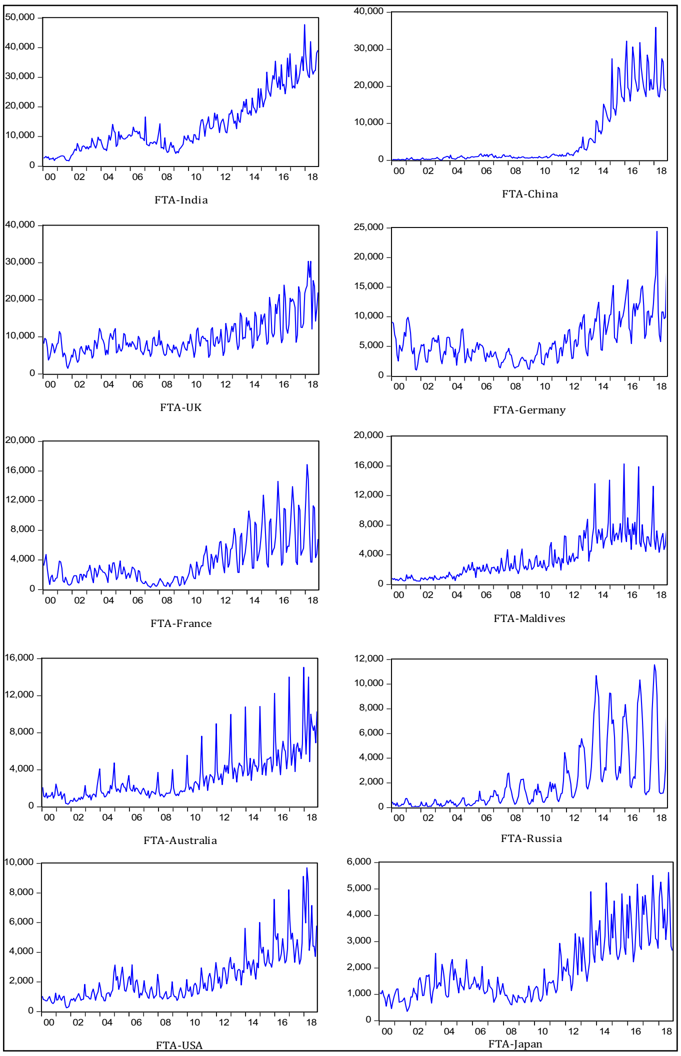
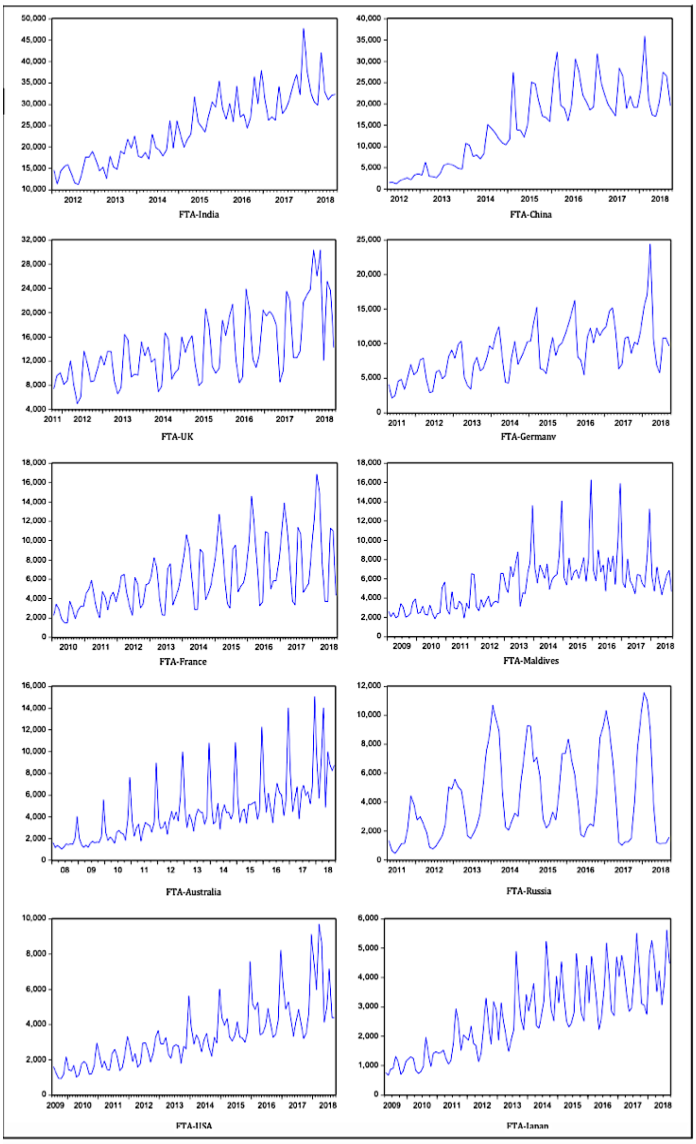


Figure 3 shows the time series plots of FTAs of top-ten countries to Sri Lanka for model estimation periods. The time series plots were also served to explain the trends of FTAs of each top-ten countries during the period for model estimation periods. Further, we can observe that FTAs series of each country gradually increase initially with lower rate and variations. However, at the end of these FTAs series, an increase in the rate and variations were expanded for all top-ten countries except the FTAs series of Maldives.

Figure 3: Time series plots of FTAs for the model estimation periods



Winter's exponential smoothing models

Under the E-views software package, we ran the Holt-Winter exponential smoothing to assess the time series behavior for FTAs to Sri Lanka. The parameters , , and are used to estimate the exponential smoothing for the level series and logarithmic series, the FTAs in our case (see Table 4). The parameter smooths the level equation with low value, suggesting that the corresponding series is quite steady over the model estimation period. On the other hand, a high value indicates a large fluctuation in FTAs. As shown in Table 4, the parameters of trend exponential smoothing, denoted as , for all series equal to 0 except for the case of FTAs-China (= 0.60) and the FTAs-Maldives (= 0.06) in the Level series and FTAs-China (= 0.10), FTAs-German (= 0.13), and the FTAs-Maldives (= 0.09) in the Logarithmic series. The 0 values indicate that the slopes of the series are constant over the model estimation period, or the trend can be represented by a straight line. Seasonal effects are captured by the seasonal parameter denoted in the seasonal equations and all the values are equal to 0, except for FTAs-Australia where equals 0.48 in the Level series. These parameters , and are all interrelated. For example, a large value will tend to have low value and vice versa.

As indicated in Table 4, the pattern of FTAs on levels series for the UK, France, Russia, and the USA are mostly similar. The values are high (close or equal to 1) and the values are zero. FTAs from these countries are in sync with the most recent figure of arrivals, with the constant trend and steady seasonal effect. Further, this is supported only for the pattern of FTAs on logarithmic series for Russia. On the other hand, the exponential smoothing model with low (close to 0) and higher (close to 1) values indicate that the tourist arrivals series from Australia are strongly affected by seasonal factors. In the case of FTAs for China, the model with low (0.09) and (0.60) values specify that the FTAs series cannot be represented by a straight line. Table 4 also shows that FTAs on both level and logarithmic series for countries India, Germany, Maldives, and Japan have low values and the values of and fall in the neighborhood of zero, indicating that the corresponding FTAs series are quite steady over the model estimation period.

Table 4: Parameters of Winter's exponential smoothing on the level and Logarithmic series of FTAs

Variable	Level series of FTAs			Logarithmic series of FTAs		
	α (Level)	γ (Trend)	δ (Seasonal)	α (Level)	γ (Trend)	δ (Seasonal)
FTAs-*India*	0.29	0.00	0.00	0.36	0.00	0.00
FTAs-*China*	0.09	0.60	0.00	0.04	0.10	0.00
FTAs-*UK*	0.90	0.00	0.00	0.05	0.00	0.00
FTAs-*Germany*	0.49	0.00	0.00	0.01	0.13	0.00
FTAs-*France*	1.00	0.00	0.00	0.35	0.00	0.00
FTAs-*Maldives*	0.19	0.06	0.00	0.17	0.09	0.00
FTAs-*Australia*	0.21	0.00	0.48	0.33	0.00	0.00
FTAs-*Russia*	1.00	0.00	0.00	1.00	0.00	0.00
FTAs-*USA*	0.61	0.00	0.00	0.13	0.00	0.00
FTAs-*Japan*	0.10	0.00	0.00	0.25	0.00	0.00

ARIMA models

The estimation of the parameters as described in methodology is executed using the automatic ARIMA selection available in E-views. Using the model estimation period, the estimation of ARIMA models using the described software facility is shown in Table 5 in standard form. The level series and logarithmic series of top-ten countries are commonly supported by both the autoregressive and moving average processes. This supposes that the previous month or months affect the current figure of FTAs. And also, previous month or months random errors (white noises) affect the current figure of FTAs. Except for the FTAs series from Russia, Table 5 shows that all the level series and logarithmic series of FTAs need differencing to produce stationary series for analysis. Table 5 also shows that none of the FTAs series need seasonal differencing to produce seasonal stationary series for analysis. As shown in Table 5, the best multiplicative SARIMA model for the FTAs series for each country has a minimum value of AIC comparing with different SARIMA models. The best ARIMA model is produced in running the number of iterations by the automatic ARIMA model facility available. Further, we can observe that the adjusted R2 value of FTAs level series of all the countries, except for Germany, is greater than 0.72 with an adjusted R2 value of FTAs logarithmic series are greater than 0.77. These results thus suggest that the relevant selected models for FTAs series of each country, are at a higher significance level.

Table 5: Estimated the best ARIMA models in standard form for both level and logarithmic series

Variable	Model for Level Series	Model Selection Criteria		Model for Logarithmic Series	Model Selection Criteria	
		Adjusted R^2	AIC		Adjusted R^2	AIC
FTAs-*India*	ARIMA(1,1,1)(1,0,0)^{12}	0.7272	18.6956	ARIMA(4,1,4)(1,0,1)^{12}	0.8460	0.0927
FTAs-*China*	ARIMA(5,1,4)(1,0,1)^{12}	0.9177	18.1555	ARIMA(5,1,5)(1,0,1)^{12}	0.9177	0.1366
FTAs-*UK*	ARIMA(2,1,5)(1,0,0)^{12}	0.7659	18.8585	ARIMA(1,1,1)(1,0,1)^{12}	0.8317	0.5167
FTAs-*Germany*	ARIMA(1,1,4)(0,0,1)^{12}	0.6309	18.0922	ARIMA(5,1,1)(1,0,1)^{12}	0.8959	0.2234
FTAs-*France*	ARIMA(2,1,5)(0,0,1)^{12}	0.8441	17.1833	ARIMA(2,1,3)(1,0,0)^{12}	0.9537	0.2225
FTAs-*Maldives*	ARIMA(4,1,5)(1,0,0)^{12}	0.8345	17.0983	ARIMA(2,1,5)(1,0,1)^{12}	0.8173	0.5519
FTAs-*Australia*	ARIMA(2,1,3)(0,0,1)^{12}	0.7962	17.2084	ARIMA(2,1,3)(1,0,1)^{12}	0.9223	0.2709
FTAs-*Russia*	ARIMA(4,0,3)(1,0,1)^{12}	0.9368	16.2576	ARIMA(5,0,3)(1,0,1)^{12}	0.9758	0.2761
FTAs-*USA*	ARIMA(3,1,3)(1,0,0)^{12}	0.7876	15.9723	ARIMA(2,1,2)(1,0,1)^{12}	0.7785	0.3884
FTAs-*Japan*	ARIMA(5,1,3)(1,0,1)^{12}	0.8289	15.1118	ARIMA(1,1,1)(1,0,1)^{12}	0.7825	0.3817

Simple recurrent neural network (SRNN) models

As shown in Table 6, all the SRNN models for the top-ten series have the same structure as stated above in methodology. These models were implemented with the help of computer software python and its Simple Recurrent Neural Networks (SRNN) module. First, we arranged each series into the stationary series, then, normalized to the range (0,1) before feeding into the networks. Normalized series for each country was divided into two sets, namely training series, and test series. We considered 80% of the observation in each series as the training series and the remaining 20% as series as test series. The training series of the recurrent networks were processed using a maximum of 30 hidden inputs and the number of epochs (cycles) between 50 and 150. Finally, we selected the best model based on the RMSE criterion. The following table 6 provides the above information to select each model on both level and logarithmic series for top-ten countries.

Table 6: The SRNN models and RMSE values for FTA of top-ten countries

Variable	Model	Number of Hidden inputs	Number of Epochs	RMSE
FTAs-India	SRNN-LD	26	150	6495.61
	SRNN-LTD	23	100	6655.55
FTAs-China	SRNN-LD	5	100	4246.46
	SRNN-LTD	30	100	3241.50
FTAs-UK	SRNN-LD	4	100	7128.73
	SRNN-LTD	4	100	6639.08
FTAs-Germany	SRNN-LD	3	110	3177.75
	SRNN-LTD	13	100	3345.27
FTAs-France	SRNN-LD	5	100	1273.27
	SRNN-LTD	29	100	1154.18
FTAs-Maldives	SRNN-LD	4	100	3147.64
	SRNN-LTD	20	100	1994.01
FTAs-Australia	SRNN-LD	30	80	1676.37
	SRNN-LTD	2	100	1530.89
FTAs-Russia	SRNN-LD	7	100	1148.17
	SRNN-LTD	30	100	894.91
FTAs-USA	SRNN-LD	3	100	2034.92
	SRNN-LTD	10	100	1643.20
FTAs-Japan	SRNN-LD	2	100	635.83
	SRNN-LTD	9	100	619.21

Comparisons among the three forecasting methods

The best-fitted models for FTAs on both level and logarithmic series in the model estimation periods on top-ten countries are done using the three different methods.

Table 7 illustrates the results of five measures of accuracy as model evaluation criteria: MAD, MSE, RMSE, MAPE, and RMSPE to evaluate the best-fitted model for each country from three different methods on the level and logarithmic series. Finally, we compare the following six models: 1. Winter's Multiplicative Exponential Smoothing Method on Levels Series (WMESM-LS), 2. Winter's Multiplicative Exponential Smoothing Method on Logarithmic Series (WMESM-LOS), 3. Multiplicative Seasonal ARIMA Method on Levels Series (MSARIMAM-LS), 4. Multiplicative Seasonal ARIMA Method on Logarithmic Series (MSARIMAM-LOS), 5. Simple Recurrent Neural Network Method on Level Series (SRNNM-LS), and 6. Simple Recurrent Neural Network Method on Logarithmic Series (SRNNM-LOS) using the aforementioned evaluation criteria to find the best for each series.

Table 7: Model evaluation criteria for all applied models for FTAs of top-ten countries

Variable	Model	Model Performance Evaluation Criteria				
		MAD	MSE	RMSE	MAPE	RMSPE
FTAs-India	WMESM-LS	1390.75	3245504.61	1801.53	6.00	11.38
	WMESM-LOS	1404.59	3312143.8	1819.93	6.00	11.43
	MSARIMAM-LS	2029.38	6880666.84	2623.10	8.00	15.77
	MSARIMAM-LOS	1276.20	2979313.18	1726.07	5.00	10.25
	SRNNM-LS	5584.68	42192932.57	6495.61	15.84	37.03
	SRNNM-LOS	5432.51	44296371.54	6655.55	14.88	39.23
FTAs-China	WMESM-LS	1537.19	4959242.76	2226.94	13.00	16.7
	WMESM-LOS	1686.61	7019080.09	2649.35	12.00	17.74
	MSARIMAM-LS	1271.61	2665479.47	1632.63	6.00	10.96
	MSARIMAM-LOS	1084.43	2247031.31	1499.01	5.00	9.82
	SRNNM-LS	3111.21	18032443.97	4246.46	13.03	26.49
	SRNNM-LOS	2175.64	10507334.06	3241.50	9.02	20.74
FTAs-UK	WMESM-LS	1641.89	7431179.08	2726.02	12.00	20.02
	WMESM-LOS	1455.90	6914808.84	2629.60	10.00	22.05
	MSARIMAM-LS	1855.94	6934178.75	2633.28	12.00	20.33
	MSARIMAM-LOS	1575.20	8461852.38	2908.93	10.00	23.26
	SRNNM-LS	4501.62	50818838.59	7128.73	18.61	65.83
	SRNNM-LOS	3801.77	44077405.48	6639.08	15.06	62.11
FTAs-Germany	WMESM-LS	793.36	1259409.95	1122.23	9.00	11.03
	WMESM-LOS	805.02	1430765.73	1196.15	9.00	11.16
	MSARIMAM-LS	1311.82	3479902.73	1865.45	14.00	17.93
	MSARIMAM-LOS	763.03	1199560.76	1095.24	7.00	10.12
	SRNNM-LS	1990.25	10098085.20	3177.75	14.91	27.61
	SRNNM-LOS	2278.75	11190844.75	3345.27	17.22	29.84
FTAs-France	WMESM-LS	580.03	555275.14	745.17	11.00	9.56
	WMESM-LOS	587.08	561715.10	749.48	11.00	9.24
	MSARIMAM-LS	882.73	1384557.45	1176.67	15.00	14.24
	MSARIMAM-LOS	589.39	610890.48	781.59	8.00	8.55
	SRNNM-LS	1039.00	1621209.23	1273.27	15.84	13.85
	SRNNM-LOS	883.46	1332132.54	1154.18	11.79	11.81
FTAs-Maldives	WMESM-LS	783.52	1106240.14	1051.78	15.00	13.29
	WMESM-LOS	782.77	1127400.91	1061.79	15.00	13.52
	MSARIMAM-LS	823.36	1252179.73	1119.01	15.00	14.64
	MSARIMAM-LOS	859.33	1435432.62	1198.10	13.00	13.27
	SRNNM-LS	2590.16	9907625.68	3147.64	45.41	31.56
	SRNNM-LOS	1705.59	3976057.90	1994.01	27.92	26.63
FTAs-Australia	WMESM-LS	442.94	627912.58	792.41	10.00	9.97
	WMESM-LOS	459.32	722339.38	849.91	9.00	10.25
	MSARIMAM-LS	945.74	1533728.9	1238.44	19.00	16.62
	MSARIMAM-LOS	445.59	614837.96	784.12	9.00	9.56
	SRNNM-LS	1131.62	2810222.44	1676.37	13.60	19.24
	SRNNM-LOS	1152.27	2343613.26	1530.89	13.99	16.37
FTAs-Russia	WMESM-LS	471.58	436556.93	660.72	12.00	9.00
	WMESM-LOS	492.75	468571.73	684.52	13.00	9.27
	MSARIMAM-LS	540.25	511842.62	715.43	15.00	10.81
	MSARIMAM-LOS	403.18	333279.76	577.3	8.00	7.26
	SRNNM-LS	934.46	1318293.46	1148.17	33.35	16.87
	SRNNM-LOS	722.59	800856.86	894.91	27.15	13.45
FTAs-USA	WMESM-LS	382.81	373322.22	611.00	11.00	8.67
	WMESM-LOS	358.98	376355.89	613.48	10.00	9.00
	MSARIMAM-LS	425.83	414332.68	643.69	11.00	9.44
	MSARIMAM-LOS	421.57	458190.24	676.90	10.00	9.68
	SRNNM-LS	1485.16	4140916.04	2034.92	25.76	30.50
	SRNNM-LOS	1251.91	2700116.33	1643.20	24.17	22.23
FTAs-Japan	WMESM-LS	285.88	167369.27	409.11	11.00	7.26
	WMESM-LOS	310.18	215937.05	464.69	11.00	7.86
	MSARIMAM-LS	324.08	162578.39	403.21	11.00	7.07
	MSARIMAM-LOS	340.05	198015.21	444.99	11.00	7.75
	SRNNM-LS	510.59	404280.15	635.83	13.92	10.25
	SRNNM-LOS	423.43	383422.98	619.21	11.11	9.54

Table 7 additionaly demonstrates ranking from minimum to maximum values by each criterion for the Models with different approaches on the level series and logarithmic series for FTAs of top-ten countries and the overall rank. The overall rank for different methods on level series and logarithmic series directs us to find the best-fitted model for each country.

As shown in Table 8, the results of three different approaches, with more criteria values of minimal errors, provide the best performance or fitted model to the level of logarithmic series. The results show that the MSARIMAM-LOS is the best-fitted model for the FTAs series for India, China, Germany, Australia, and Russia while the WMESM-LS remains the best-fitted model for FTAs series for France, Maldives, and USA. However, the best-fitted model for the FTAs series for the UK and Japan are WMESM-LOS and MSARIMAM-LS respectively. Except for the SRNN, all the other two methods for any series were best fitted for at least an FTAs series of one of these top-ten countries.

Table 8: Ranking for Models with different methods on the level and logarithmic series for FTAs of top-ten countries

Variable	Model	Model Evaluation Criteria					Overall Rank
		MAD	MSE	RMSE	MAPE	RMSPE	
FTAs-India	WMESM-LS	2	2	2	2	2	2
	WMESM-LOS	3	3	3	2	3	3
	MSARIMAM-LS	4	4	4	3	4	4
	MSARIMAM-LOS	1	1	1	1	1	1
	SRNNM-LS	6	5	5	5	5	5
	SRNNM-LOS	5	6	6	4	6	6
FTAs-China	WMESM-LS	3	3	3	4	3	3
	WMESM-LOS	4	4	4	3	4	4
	MSARIMAM-LS	2	2	2	2	2	2
	MSARIMAM-LOS	1	1	1	1	1	1
	SRNNM-LS	6	6	6	6	6	6
	SRNNM-LOS	5	5	5	5	5	5
FTAs-UK	WMESM-LS	3	3	3	2	1	2
	WMESM-LOS	1	1	1	1	3	1
	MSARIMAM-LS	4	2	2	2	2	2
	MSARIMAM-LOS	2	4	4	1	4	3
	SRNNM-LS	6	6	6	4	6	5
	SRNNM-LOS	5	5	5	3	5	4
FTAs-Germany	WMESM-LS	2	2	2	2	2	2
	WMESM-LOS	3	3	3	2	3	3
	MSARIMAM-LS	4	4	4	3	4	4
	MSARIMAM-LOS	1	1	1	1	1	1
	SRNNM-LS	5	5	5	4	5	5
	SRNNM-LOS	6	6	6	5	6	6
FTAs-France	WMESM-LS	1	1	1	2	3	1
	WMESM-LOS	2	2	2	2	2	2
	MSARIMAM-LS	4	5	5	3	6	4
	MSARIMAM-LOS	3	3	3	1	1	3
	SRNNM-LS	6	6	6	5	5	6
	SRNNM-LOS	5	4	4	4	4	5
FTAs-Maldives	WMESM-LS	2	1	1	2	2	1
	WMESM-LOS	1	2	2	2	3	2
	MSARIMAM-LS	3	3	3	2	4	4
	MSARIMAM-LOS	4	4	4	1	1	3
	SRNNM-LS	6	6	6	4	6	6
	SRNNM-LOS	5	5	5	3	5	5
FTAs-Australia	WMESM-LS	1	2	2	2	2	2
	WMESM-LOS	3	3	3	1	3	3
	MSARIMAM-LS	4	4	4	5	5	4
	MSARIMAM-LOS	2	1	1	1	1	1
	SRNNM-LS	5	6	6	3	6	6
	SRNNM-LOS	6	5	5	4	4	5
FTAs-Russia	WMESM-LS	2	2	2	2	2	2
	WMESM-LOS	3	3	3	3	3	3
	MSARIMAM-LS	4	4	4	4	4	4
	MSARIMAM-LOS	1	1	1	1	1	1
	SRNNM-LS	6	6	6	6	6	6
	SRNNM-LOS	5	5	5	5	5	5
FTAs-USA	WMESM-LS	2	1	1	2	1	1
	WMESM-LOS	1	2	2	1	2	2
	MSARIMAM-LS	4	3	3	2	3	3
	MSARIMAM-LOS	3	4	4	1	4	4
	SRNNM-LS	6	6	6	4	6	6
	SRNNM-LOS	5	5	5	3	5	5
FTAs-Japan	WMESM-LS	4	2	2	1	2	2
	WMESM-LOS	1	4	4	1	4	4
	MSARIMAM-LS	2	1	1	1	1	1
	MSARIMAM-LOS	3	3	3	1	3	3
	SRNNM-LS	6	6	6	3	6	6
	SRNNM-LOS	5	5	5	2	5	5

Moreover, the results show that the SRNN does not outperform in forecasting all series. These top-ten (10) FTAs series in Figure 3 shows obvious pattern within the model estimation periods of each country, where usual time series methods; exponential smoothing and ARIMA can do much on such kind of series for best performance in forecasting. However, SRNN outperforms in forecasting well especially for the FTAs series with less obvious patterns (Cho, 2003b). Further, the performance of SRNN increases in series with a large number of observations or big data (Song et al., 2019; Sheresheva et al., 2020).

4 CONCLUSIONS AND IMPLICATIONS

This study aimed to compare three different methods to predict foreign tourist arrivals to Sri Lanka from top-ten countries and also attempts to find the best-fitted forecasting model for each country. After the end of the civil war in 2009, the FTAs of these countries gradually increased with lower seasonal impacts. By the end of 2018, FTAs of these countries have increased at a rate that has never experienced and with higher seasonal impacts. The highest annual average percentage (17.6%) of FTAs is generating from India while Japan represents the lowest annual percentage (2.2%) FTAs to Sri Lanka among top-ten countries. We estimated the best fitted time series model for each country that best describes the FTAs in Sri Lanka based on overall rank calculated from ranks of five model selection criteria with minimum errors. We achieved this objective by applying three selected models: WMESM, MSARIMA, and SRNN models on level series and logarithmic series. The results of our study showed that the best-fitted model for FTAs series in the case of India China, Germany, Australia and Russia are the MSARIMA models which can be respectively denoted as ARIMA (4,1,4) (1,0,1) 12, ARIMA (5,1,5) (1,0,1) 12, ARIMA (5,1,1) (1,0,1) 12, ARIMA (2,1,3) (1,0,1) 12 and ARIMA (5,0,3) (1,0,1) 12. The results show that the WMESM-LS most appropriate model for tourist arrival from France, Maldives, and the USA. However, the best-fitted models recorded for UK and Japan were WMESM-LOS and MSARIMA-LS. Moreover, the results show that MSARIMA-LS for Japan is ARIMA (5,1,3) (1,0,1) 12.

The results of this study revealed that the MAPE value of the MSARIMA forecast model is generally less than the other two models for seven out of ten countries while equal with MAPE value of the WMESM for three countries; UK, USA, and Japan. MSARIMA model showed the best forecasting ability for FTAs of top-ten countries. The MAPE of SRNN had the highest values indicating that its forecasting ability is the worst for all top-ten countries except China. However, the second minimum value for MAPE (9.02%) was observed from the SRNN model for only FTAs of China. The MAPE of WMESM was between the other two models, indicating that its forecasting ability is normal comparatively. The selected best-fitted models based overall rank of five model performance evaluation criteria can be applied for forecasting tourist arrivals from each of the top-ten countries. When one of the above main criteria used to measure the forecasting accuracy, MAPE, values obtained for the best-fitted models range from 5 percent to 15 percent between models of top-ten countries. Among these, the highest accuracy of 5 percent is shown from the model estimated for India and China FTAs. The accuracy rates of the models of another country Germany FTAs are 7 percent while this value for Russia and Australia FTAs are 8 percent and 9 percent respectively. According to the Lewis criteria on MAPE, the accuracy rate for the models of India, China, Germany, Russia, and Australia are relatively high (Lewis, 1982). The accuracy levels of models for countries; UK, France, USA, Japan, and the Maldives, respectively range between 10 percent and 15 percent which remain at a good level of accuracy. However, the diverse characteristics of tourists from various countries can explain different results of the above time-series models.

Implications for decision-makers

Forecasting of FTAs for each market source has a practical implication to decision-makers. To increase FTAs, the government needs to tailor better promotion and advertising strategies. They also need to invest more in modern and sustainable infrastructure facilities that meet international quality and standards. The government is encouraged to develop strategies that also boost local tourism especially during off-seasons of foreign tourists. However, it is important that government authorities ensure the capacity building of tourism development officers through intensive training so that they can have adequate skills and knowledge of tourism development as suggested by Briedenhann (2007). More importantly, they need to improve the security strategy, provide more training to the security forces, and invest in modern technology to avoid random shocks to tourism. However, it is encouraged that governments develop more cooperation with more technologically advanced countries to track the early sign of attacks to mitigate or eliminate casualties (terrorist attacks, contagious diseases) that can hurt the tourism sector. It is also encouraged that governments have an open channel that can help whistleblowers to inform security forces of suspected activities and behavior that threaten national security. Further, the evidence of the study confirms that there is a dearth in FTAs during internal crisis periods. On this basis, we can conclude that a stable economic and political environment has the capability of attracting a higher number of tourists, particularly if there are policies that encourage sustainable tourism development. In working towards meaningful tourism development, governments or policymakers must integrate diverse cultural views into their planning as suggested by Higgins-Desbiolles (2018) and Pirnar et al. (2019).

Limitations and suggestions for future studies

There are several limitations to this study. First, the fact that it focuses on Sri Lanka makes the results unlikely to be generalized as this is case of most tourist forecasting arrivals that are country specific. Second, this study is also limited as only a set of forecasting models have been selected, leaving the possibility to derive tourism forecasting to Sri Lanka using other models. Beyond these limits, some important suggestions for further studies can also be given through the study. First, the factors affecting the FTAs should be identified through a specially designed study for that purpose. A causal model that considers more than one independent variable affecting tourist arrivals is encouraged in future studies. Second, further studies may be directed towards identifying seasons where FTAs are relatively low and the strategies that can be adopted to boost local tourist arrivals during the off-seasons of the FTAs. Future study is also encouraged to use Winter's exponential smoothing method, Box-Jenkin models ARIMA, and AI methods like SRNN, for example, to forecast domestic tourist arrival during the off-seasons of foreign tourists.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. In Selected Papers of Hirotugu Akaike (pp. 215-222): Springer.

Akuno, A. O., Otieno, M. O., Mwangi, C. W., & Bichanga, L. A. (2015). Statistical Models for Forecasting Tourists' Arrival in Kenya. Open Journal of Statistics, 05(01), 60-65. doi:10.4236/ojs.2015.51008

Avdimiotis, S. & Christou, E. (2004). GIS applications in tourism planning: A tool for sustainable development involving local communities. Journal of Environmental Protection & Ecology, 5(2), 457-468

Box, G. E. P., & Jenkins, G. M. (1970). Time series analysis: forecasting and control. San Francisco: Holden-Day.

Briedenhann, J. (2007). The Role of the Public Sector in Rural Tourism: Respondents' Views. Current Issues in Tourism, 10(6), 584-607. doi:10.2167/cit331.0

Chandra, S., & Kumari, K. (2018). Forecasting Foreign Tourist Arrivals to India using Time Series Models. Journal of Data Science, 16(4), 707-722 doi:10.6339/jds.201810_16(4).00003

Cho, V. (2003a). A comparison of three different approaches to tourist arrival forecasting. Tourism Management, 24, 323-330. doi:10.1016/S0261-5177(02)00068-7

Cho, V. (2003b). A comparison of three different approaches to tourist arrival forecasting. Tourism Management, 24(3), 323-330. doi:10.1016/s0261-5177(02)00068-7

Christou, E. (2002). Examining the impact of tourism destination image and reputation on visitor loyalty likelihood. Tourism Today, 2(1), 34-46

Chu, F.-L. (2004). Forecasting tourism demand: a cubic polynomial approach. Tourism Management, 25(2), 209-218.

Chu, F.-L. (2009). Forecasting tourism demand with ARMA-based methods. Tourism Management, 30(5), 740-751. doi:https://doi.org/10.1016/j.tourman.2008.10.016

Claveria, O., Monte, E., & Torra, S. (2015). Tourism Demand Forecasting with Neural Network Models: Different Ways of Treating Information. 17(5), 492-500. doi:10.1002/jtr.2016

Coshall, J. T. (2009). Combining volatility and smoothing forecasts of UK demand for international tourism. Tourism Management, 30(4), 495-511. doi:https://doi.org/10.1016/j.tourman.2008.10.010

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association, 74(366a), 427-431.

Elman, J. L. (1990). Finding structure in time. Cognitive science, 14(2), 179-211.

Gnanapragasam, S. R. (2018). Tourist Arrivals in Sri Lanka: A Comparative Study of Holt- Winter's versus Box- Jenkin's Modeling Methods. OUSL Journal, 13(1), 65. doi:10.4038/ouslj.v13i1.7395

Goh, C., & Law, R. (2002). Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. Tourism Management, 23(5), 499-510. doi:https://doi.org/10.1016/S0261-5177(02)00009-2

Higgins-Desbiolles, F. (2018). Sustainable tourism: Sustaining tourism or something more? Tourism management perspectives, 25, 157-160.

Jordan, M. I. (1997). Chapter 25 - Serial Order: A Parallel Distributed Processing Approach. In J. W. Donahoe & V. Packard Dorsel (Eds.), Advances in Psychology (Vol. 121, pp. 471-495): North-Holland.

Law, R., Li, G., Fong, D. K. C., & Han, X. (2019). Tourism demand forecasting: A deep learning approach. Annals of Tourism Research, 75, 410-423. doi:https://doi.org/10.1016/j.annals.2019.01.014

Lewis, C. D. (1982). Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting. London: Butterworth Scientific.

Li, G., Wong, K. K. F., Song, H., & Witt, S. F. (2006). Tourism Demand Forecasting: A Time Varying Parameter Error Correction Model. Journal of Travel Research, 45(2), 175-185. doi:10.1177/0047287506291596

Lichtendahl, K. C., & Winkler, R. L. (2020). Why do some combinations perform better than others? International Journal of Forecasting, 36(1), 142-149. doi:https://doi.org/10.1016/j.ijforecast.2019.03.027

Lim, C., & McAleer, M. (2001a). Forecasting Tourist Arrivals. Annals of Tourism Research, 28(4), 965-977. doi:doi:10.1016/s0160-7383(01)00006-8

Lim, C., & McAleer, M. (2001b). MONTHLY SEASONAL VARIATIONS: Asian Tourism to Australia. Annals of Tourism Research, 28(1), 68-82. doi:doi:10.1016/s0160-7383(00)00002-5

Lyons, A. & Branston, C. (2006). Cross cultural change, adjustment and culture shock: UK to USA. Tourism: An International Interdisciplinary Journal, 54(4), 355-365

Martin, C. A., & Witt, S. F. (1989). Accuracy of econometric forecasts of tourism. Annals of Tourism Research, 16(3), 407-428. doi:https://doi.org/10.1016/0160-7383(89)90053-4

Nanthakumar, L., & Ibrahim, Y. (2010). Forecasting International Tourism Demand in Malaysia Using Box Jenkins Sarima Application South Asian Journal of Tourism and Heritage 3(2), 50-60.

Nanthakumara, L., & Ibrahim, Y. (2010). Forecasting International Tourism Demand in Malaysia Using Box-Jenkins Sarima Application. South Asian Journal of Tourism and Heritage, 3(2), 50-60.

Pankratz, A. (1983). Forecasting with univariate Box-Jenkins models : concepts and cases. New York, N.Y.: Wiley.

Pirnar, I., Kurtural, S., & Tutuncuoglu, M. (2019). Festivals and destination marketing: An application from Izmir City. Journal of Tourism, Heritage & Services Marketing, 5(1), 9–14. http://doi.org/10.5281/zenodo.2640987

Saayman, A., & Saayman, M. (2010). Forecasting Tourist Arrivals in South Africa. Acta Commercii, 281-293.

Salman, A. K., Shukur, G., & Bergmann-Winberg, M.-L. v. (2007). Comparison of Econometric Modelling of Demand for Domestic and International Tourism: Swedish Data. Current Issues in Tourism, 10(4), 323-342. doi:10.2167/cit257.0

Sheresheva, M. Y., Polukhina, A. N., & Oborin, M. S. (2020). Marketing issues of sustainable tourism development in Russian regions. Journal of Tourism, Heritage & Services Marketing, 6(1), 33–38. http://doi.org/10.5281/zenodo.3603422

Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. International Journal of Forecasting, 36(1), 75-85. doi:https://doi.org/10.1016/j.ijforecast.2019.03.017

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting—A review of recent research. Tourism Management, 29(2), 203-220. doi:https://doi.org/10.1016/j.tourman.2007.07.016

Song, H., Qiu, R. T. R., & Park, J. (2019). A review of research on tourism demand forecasting: Launching the Annals of Tourism Research Curated Collection on tourism demand forecasting. Annals of Tourism Research, 75, 338-362. doi:10.1016/j.annals.2018.12.001

Sotiriadis, M., & Shen, S. (2017). The contribution of partnership and branding to destination management in a globalized context: The case of the UNWTO Silk Road Programme. Journal of Tourism, Heritage & Services Marketing, 3(2), 8–16. http://doi.org/10.5281/zenodo.1209121

Sun, X., Sun, W., Wang, J., Zhang, Y., & Gao, Y. (2016). Using a Grey–Markov model optimized by Cuckoo search algorithm to forecast the annual foreign tourist arrivals to China. Tourism Management, 52, 369-379. doi:https://doi.org/10.1016/j.tourman.2015.07.005

Teixeira, J. P., & Fernandes, P. O. (2012). Tourism Time Series Forecast -Different ANN Architectures with Time Index Input. Procedia Technology, 5, 445-454. doi:10.1016/j.protcy.2012.09.049

Thushara, S. C., Su, J.-J., & Bandara, J. S. (2019). Forecasting international tourist arrivals in formulating tourism strategies and planning: The case of Sri Lanka. Cogent Economics & Finance, 7(1), 1699884. doi:10.1080/23322039.2019.1699884

Trippi, R. R., & Turban, E. (1996). Neural networks in finance and investing : using artificial intelligence to improve real-world performance. Chicago: Irwin Professional Pub.

Van Truong, N., Shimizu, T., Kurihara, T., & Choi, S. (2020). Generating reliable tourist accommodation statistics: Bootstrapping regression model for overdispersed long-tailed data. Journal of Tourism, Heritage & Services Marketing, 6(2), 30–37. http://doi.org/10.5281/zenodo.3837608

Volgger, M., Pechlaner, H., & Pichler, S. (2017). The practice of destination governance: A comparative analysis of key dimensions and underlying concepts. Journal of Tourism, Heritage & Services Marketing, 3(1), 18–24. http://doi.org/10.5281/zenodo.401371

Wong, K. K. F., Song, H., Witt, S. F., & Wu, D. C. (2007). Tourism forecasting: To combine or not to combine? Tourism Management, 28(4), 1068-1078. doi:https://doi.org/10.1016/j.tourman.2006.08.003

SUBMITTED: DEC 2019
REVISION SUBMITTED: APR 2020
2nd REVISION SUBMITTED: MAY 2020
ACCEPTED: JUL 2020
REFEREED ANONYMOUSLY

PUBLISHED ONLINE: 30 OCT 2020