FORECASTING DOMESTIC GUEST NIGHTS IN ANCIENT CITIES OF SRI LANKA: HYBRID APPROACH

KMUB Konarasinghe
Institute of Mathematics & Management, Ranala, Sri Lanka
udaya35@yahoo.co.uk

Abstract- The Ancient Cities are highly occupied by domestic tourists after 2009. The high occupancy increases the demand for accommodation. Hence, the hotel industry should adopt various practices to maximize profits and minimize the risk. This can be achieved by accurate forecasting. But, there were least attempts on forecasting occupancy guest nights of domestic tourist in Ancient Cities of Sri Lanka. Therefore, this study was focused on forecasting occupancy guest nights of domestic tourist in Ancient Cities of Sri Lanka. Monthly data of domestic guest nights for the period of January 2008 to December 2016 were obtained from Sri Lanka Tourism Development Authority (SLTDA). Descriptive statistics were obtained. The trend models; Linear, Quadratic, Growth Curve and S-Curve models were tested. The Anderson-Darling test revealed the residuals of Linear and Quadratic were normally distributed, but Ljung-Box Q-test and Auto-Correlation Function (ACF) does not confirm the independence. Therefore the de-trended data were further analyzed; the stationary of the series was tested by Augmented Dickey-Fuller (ADF) test and ACF. Then the Auto-Regressive Integrated Moving Average (ARIMA) model was tested on each series. The ARIMA model was well fitted for de-trended series of Linear trend and Growth Curve models. Hence, the residuals of two hybrid models; Linear trend-ARIMA and Growth Curve trend-ARIMA models were tested for model assumptions. It was concluded that both hybrid models; Linear trend-ARIMA and Growth Curve-ARIMA are suitable for forecasting occupancy guest nights of domestic tourist in Ancient Cities of Sri Lanka.

Keywords - Trend, ARIMA, Ancient Cities, Occupancy

I. INTRODUCTION

Sri Lanka is a home to eight UNESCO world heritage sites. Kandy, Sigiriya, Anuradhapura, Polonnaruwa, and Dambulla are named as Ancient Cities of world heritage sites in Sri Lanka. It is flooded with opportunities for experiencing ancient history. There is evidence left by a proud civilization stretching back more than two thousand years. Today these cities are highly visited and occupied by domestic tourists. The domestic travel refers to travel within their own country, the word domestic originating from the Latin word " Domus" meaning home. The domestic tourist travels from one to another place during their leisure time or other purposes. After 2008, there is an improvement of traveling culture of domestic travellers in Sri Lanka. The Ancient Cities are one of the highly visited and occupied regions by domestic tourist (SLTDA, 2016).

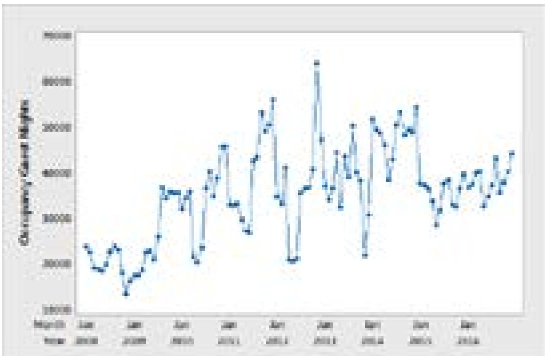

Figure 1. Time series plot of domestic occupancy in Ancient Cities

Figure 1 clearly shows an increasing trend with fluctuations of occupancy of domestic tourist in Ancient Cities over the years.

Problem Statement

The high occupancy increases the demand for accommodation and other tourism-related business. Hence, the hotel industry and other business should adopt various business practices to maximize benefits and minimize the risk. This can be achieved by accurate forecasting of occupancy by tourist (Schwartz and Hiemstra 1997). It is a well-known fact that accurate forecasting is a critical component of efficient business operations. But it was hard to find attempts of forecasting occupancy guest nights of domestic tourist in Ancient Cities of Sri Lanka.

The objective of the study

To forecast occupancy guest nights of domestic tourist in Ancient Cities of Sri Lanka

I. The significance of the study

The results of this study can be used for strategy and policy development of tourism-related business and local government authorities of Ancient Cities of Sri Lanka.

II. LITERATURE REVIEW

The studies focusing on forecasting number of occupancy guest nights are very limited across the world. Brannas, and Nordstrom (2000) model the number of Norwegian guest nights in Swedish hotels and cottages and demand analysis. The study used Integer-valued autoregressive model. Autoregressive Moving Average (ARMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) used to examine and forecast tourist accommodation demand in New Zealand using hotel-motel room nights by Lim, Chang and McAleer (2009). In Sri Lanka, SARIMA and Decomposition methods used for forecasting foreign guest nights in Colombo and Greater Colombo by Konarasinghe (2017-a).SARIMA used for forecasting foreign guest nights in Southern Cost and Ancient Cites of Sri Lanka (Konarasinghe, 2017-b) and (Konarasinghe, 2017-c). Univariate and Multivariate time series techniques were used in forecasting hotel room occupancy rates and guest nights. Brannas and Nordstrom (2000), Lim, Chang and McAleer (2009) confirmed that the Integer-valued autoregressive, ARMA and SARIMA models are suitable for forecasting guest nights. In Sri Lankan context SARIMA performed extremely well in forecasting occupancy of foreign guest nights. It was confirmed by Konarasinghe (2017-a), (2017-b) and (2017-c).

III. METHODOLOGY

Monthly domestic Occupancy data for the period of 2008 to 2016 were obtained from annual reports of Sri Lanka Tourism Development Authority (SLTDA). At first, the Descriptive statistics of occupancy were obtained. Figure 1 is the time series plot of domestic occupancy guest night in Ancient Cites. It is clear that there is an increasing trend. Therefore the trend models; Linear, Quadratic, Growth Curve and S- Curve were tested on log-transformed series. Four trend models were tested with log transformation. They are;

Linear trend model: ln $Y_t = \alpha + \beta t + \varepsilon$ (1)

Quadratic trend model: ln $Y_t = \alpha + \beta_1 t + \beta_2 t^2 + \varepsilon$ (2)

Growth curve model: ln $Y_t = \alpha(\beta^t) + \varepsilon$ (3)

S-Curve model: ln $Y_t = \dfrac{A}{\alpha + \beta(\gamma^t)} + \varepsilon$ (4)

Trend analysis revealed that the residuals of models 1 and 3 are normally distributed but not independent. The residuals of models 2 and 4 did not follow the normally distribution and independent.

Therefore de-trended data were analysed and it was intended to model them. The Auto-Regressive Integrated Moving Average (ARIMA) model was tested for the purpose. An ARIMA model is given by;

$$\varphi(B)(1-B)^d y_t = \theta(B)\varepsilon_t \quad (5)$$

Where;

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 \dots \varphi_p B^p$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 \dots \theta_p B^q$$

ε_t = Error term

D = Differencing term

B = Backshift operator ($B^a Y_t = Y_{t-a}$)

The time series plots used for pattern recognition. The Anderson-Darling test, Ljung-Box Q (LBQ) test, Auto-Correlation Function (ACF) and Augmented Dickey-Fuller (ADF) test were used to test normality, independence of residuals and stationary of the series. Forecasting ability of the models was assessed by considering Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE) and Mean Absolute Deviation (MAD). This approached applied for Sri Lankan tourism industry by Konarasinghe (2017-d).

IV. RESULTS

Data analysis is organized as follows;

(1). Descriptive Statistics

(2). Trend analysis

(3). Model de-trended series

(4). Testing the validity of hybrid models

Outliers are the extremely large or small values of a data set. They were identified with the help of Box Plot (Figure 2) and replaced by moving an average of order three. The study adopted the technique used by Konarasinghe, Abeynayake, and Gunaratne (2016), Konarasinghe (2016) and Konarasinghe (2017-d) for outlier adjustment. That is; if the ith value of a series is an outlier;

$$i^{th}\text{value} = [(i-1)^{th}\text{value} + (i-2)^{th}\text{value} + (i-3)^{th}\text{value}]/3 \quad (6)$$

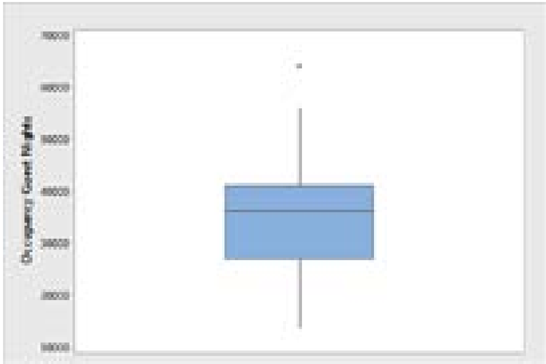

Figure 2. Box plot of occupancy guest nights

D. Descriptive Statistics

Graphical summary of descriptive statistics is shown in Figure 3.

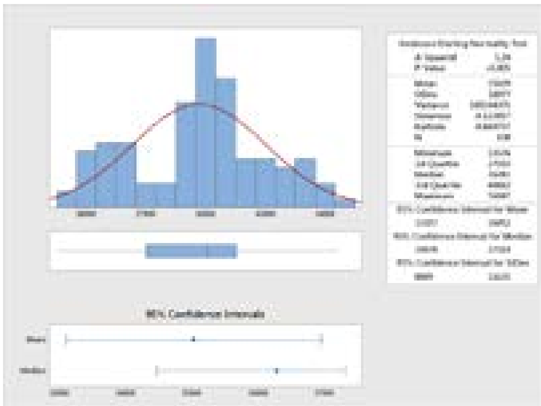

Figure 3. Graphical summary of occupancy guest nights

Minimum occupancy recorded from domestic tourist was 13576 whereas maximum was 56047 during the period. The first quartile of occupancy is 27015. It means ¼ of the months had at most 27015. A median occupancy is 36281 and the third quartile of occupancy is 40662. Histogram of the arrivals does not look symmetrical.

The p-value of the Anderson-Darling test is less than the significance level (p-value <0.005). As such, the number of arrivals does not follow the normal distribution. They are negatively skewed.

E. Trend analysis

All four trend models were tested. The relative and absolute measurements of errors of all fitted models were small. The Anderson-Darling test confirmed the normality of residuals of Linear and Growth Curve models. The residuals of Quadratic and S-Curve trend models were not normally distributed. The LBQ-test and ACF did not confirm the random of residuals of all four trend models. Therefore none of the trend models can be recommended for forecasting. Two trend models were selected in model de-trended series based on the normality of residuals. Table 1 is the results of selected trend models.

Table 1. Model summary of selected trend models

Model	Model Fitting	
Linear trend model	MAPE	1.86092
	MAD	0.19182
	MSE	0.06139
	Normality	P = 0.355
	Random of Errors	No
Growth curve model	MAPE	1.86457
	MAD	0.19226
	MSE	0.06171
	Normality	P =0.347
	Random of Errors	No

F. Model de-trended series

The de-trended data of Linear and Growth Curve trend models were extracted and plotted.

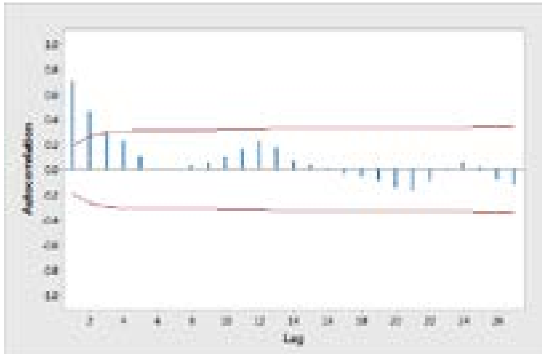

Figure 2. The Auto Correlation Function for de- trended data from Linear trend model

Figure 3. The Partial Auto Correlation Function for de-trended data from Linear trend model

Figure 2 and Figure 3 are the ACF and Partial Autocorrelation function (PACF) of de-trended series of Linear Trend model. They were tested for stationery by ADF test, ACF, and PACF. The ADF test, ACF, and PACF confirmed the stationary of the series.

Figure 4 and Figure 5 are the ACF and PACF of de-trended series of Growth Curve trend model. They were tested for

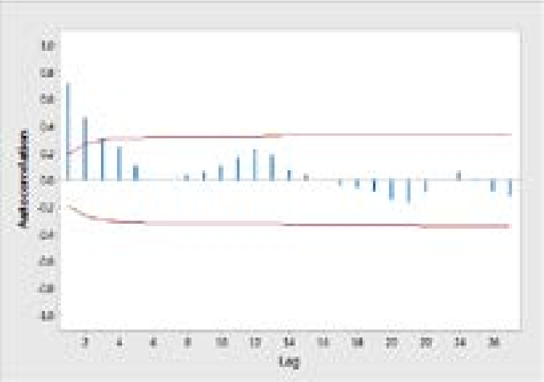

Figure 4. The ACF for de-trended data from Growth – Curve trend model

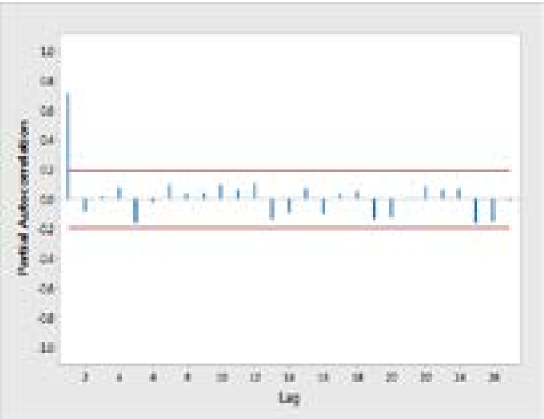

Figure 5. The PACF for de-trended data from Growth – Curve trend model

stationary by ADF test, ACF, and PACF. The ADF test, ACF, and PACF confirmed the stationary of the series.

Then the ARIMA (p, d, q) model was tested on them. ARIMA (1, 0, 0) was the suitable model for de-trended series of; Linear and Growth Curve trend models. Table 2 is the summary of best-fitting ARIMA models. The residuals of ARIMA (1, 0, 0); fitted to de-trended data of Linear and Growth Curve trend models were normality distributed and random. The results of the table 2 confirmed that the ARIMA model can be used to model the de-trended series of Linear and Growth Curve trend models. Measurements of errors of these two models were also satisfactory small.

Table 2. Summary of ARIMA (1,0,0) models

De-trended data of Model	Model Fitting		Model Verification	
	MSE	MAD	MSE	MAD
1	0.037507	0.144494	0.056832	0.223278
3	0.037489	0.144450	0.0628412	0.235893

G. Testing the validity of hybrid models

Based on the results of E and F of the analysis, following two hybrid models were selected for forecasting;

Hybrid Linear –ARIMA model

$$Y_t = \alpha_1 + \beta^t + \alpha_2 + \sum_{i=1}^{p} \varphi_i G_{t-i} + \varepsilon_t \quad (7)$$

Where; L_t = De –trended Series of the Linear Model

Hybrid Growth Curve -ARIMA model:

$$Y_t = \alpha_1 + \beta^t + \alpha_2 + \sum_{i=1}^{p} \varphi_i G_{t-i} + \varepsilon_t \quad (8)$$

Where; Gt = De –trended Series of the Growth Curve Model

Forecasts were estimated for both hybrid Linear – ARIMA and Growth Curve -ARIMA models. According to table 3, the residuals of both models were normally distributed and random. Measurements of errors of both models were very small. Fitted models as follows:

Hybrid Linear- ARIMA model:

$$\ln Y_t = 9.97026(100100^t) + 0.6947G_{t-1} \quad (9)$$

Hybrid Growth Curve -ARIMA model:

$$\ln Y_t = 9.96833 + 0.01030t + 0.6953L_{t-1} \quad (10)$$

Table 3. Summary of hybrid models

Model	Model Fitting		Model Verification	
9	MAPE	1.47749	MAPE	4.88911
	MAD	0.15519	MAD	0.51324
	MSE	0.04295	MSE	0.27002
	Normality	P =0.066		
	Random	Yes		
10	MAPE	1.48410	MAPE	5.05858
	MAD	0.15595	MAD	0.53108
	MSE	0.04329	MSE	0.28870
	Normality	P = 0.073		
	Random	Yes		

Figure 4 and figure 5 are the time series plot of actual vs fits of Hybrid Linear - ARIMA and Hybrid Growth Curve – ARIMA models. These figures clearly show that the actual data and fits are very close. Hence, these two models are selected as suitable models for forecasting domestic guest nights in Ancient Cities of Sri Lanka.

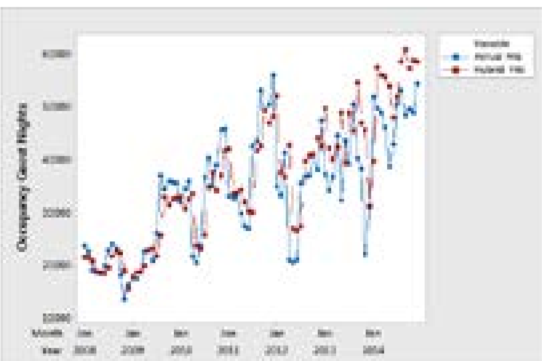

Figure 4. Actual Vs Fits of Hybrid Linear -ARIMA

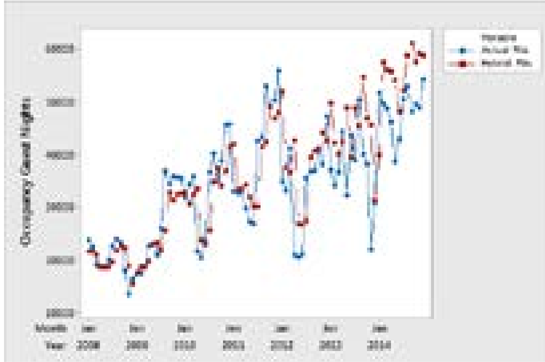

Figure5. Actual Vs Fits of Hybrid Growth -ARIMA

Estimated domestic guest nights of Ancient Cities for the period of January 2019 to December 2019 show in table 4.

Table 4. Summary of hybrid models

Year	Month	Hybrid Linear -ARIMA	Hybrid GC -ARIMA
2019	January	59745	60909
	February	60171	61373
	March	60597	61839
	April	61026	62308
	May	61457	62779
	June	61890	63253
	July	62325	63731
	August	62763	64212
	September	63204	64696
	October	63647	65184
	November	64093	65675
	December	64542	66170

V. CONCLUSION AND DISCUSSION

The study concluded that hybrid models; Hybrid Linear trend - ARIMA and Hybrid Growth Curve trend – ARIMA models are suitable for forecasting domestic guest nights in Ancient Cities of Sri Lanka.

The results of this study can be used to forecast the number of domestic occupancies in Ancient Cites of Sri Lanka. This is useful for strategy development of tourism-related business located in Ancient Cities. These strategies can be used as a lighthouse of the business journey of the tourism industry. Businesses can decide various product developments and increasing or decreasing their production volume by observing the occupancy behaviour. They can plan for various promotional offers, pricing of accommodation and other products and services based on the results of this study.

Financial controllers in the hotel need information from forecasting to understand cash/ credit flow for the hotel as that needs to be considered for multiple expenses will be generated from different departments in hotels such as food and beverages, laundry, transport etc., and including rooms. In hotels, requirements of food and beverages, purchasing decisions of perishable and non- perishable items, hiring employment, and maintenance decisions can be decided by accurate forecasting of occupancy guest nights. During high occupancy period, the host community can be a focus on new business opportunities for their additional revenue. Accommodation, food and beverages and transport are some of the business avenues for host communities. Therefore, the local government can provide guidance, facilities, and employment opportunities to host communities for business avenues based on forecasting occupancy guest nights.

There could be negative and positive consequences can exist in Ancient Cities due to high occupancy of tourist. The results of this study could be a light house to minimize the negative impacts and maximize the positive gains. During high or low occupancy periods the visitors could show the little respect for the sanctity of spiritual places, practices, and traditions. Then the local government authorities should impose control methods to avoid such situations. Their concentration should be pre-planned by observing the results of this study. Promotions, distribution, and sales of local crafts and other products could be poorly managed during high occupancy periods. It could be degrading cultural integrity. The results of this study could be useful to overcome these impacts. And integrate education and training opportunities for policymakers, planners, researchers, designers, architects, interpreters and tourism operators to avoid, conflicting issues, opportunities, and problems encountered by host communities. High occupancy will be increased traffic, noise and air pollution of Ancient Cites. Local government can plan and adopt better practices to avoid negative consequences from traffic, noise and air pollution during high occupancy periods with the guidance of the results of this study. Authorities can be introducing local transport systems to overcome these issues during the higher occupancy period and it will be the benefit to the economy and environment of Ancient Cities as well.

A large number of vehicle movements cause significant damages to all kind of roads and culverts at Ancient Cities. Authorities can be classified these roads and restricted for vehicle moments by observing the occupancy rate. The significant increase in occupancy resulted in heavy volume of garbage. Then the local government can maintain and improve the standards of formal waste collection services in Ancient Cites based on the results of this study. Wastewater management is another issue during the high occupancy periods. Authorities should implement, practice and monitor the wastewater management systems to avoid water pollution of Ancient Cites. Increasing the occupancy could result in disturbance of wildlife of Ancient Cites. Therefore, wildlife authorities should implement, practice and monitor their conservation policies following the results of this study.

Various forms of treats could be results due to occupancy of tourists. Security services implement, practice and monitor security and safety measures to protect tourists, host community and wildlife of Ancient Cites by observing the results of this study. The series of occupancy guest nights follow the wave-like pattern; as such may contain both seasonal and cyclical variation. The Decomposition technique is the commonly used technique to capture this variation. But it will be a long process. Therefore, it is recommended to test the Circular Model; in order to see whether the forecasting ability improves.

REFERENCES

Brannas, K., and Nordstrom, J. (2000).A Bivariate Integer Value Allocation Model for Guest Nights in Hotels and Cottages, *Umea Economic Studies*. Sweden.

Lim,C., Chang, C, and McAleer, M. (2009). Forecasting h(m)otel guest nights in New Zealand. *International Journal of Hospitality Management*, 28(2), 228-235.

Schwartz, Z., and Hiemstra,S. (1997).Improving the accuracy of hotel reservations forecasting: Curves Similarity Approach. *Journal of Travel Research*, 36(1), 3-14.

Sri Lanka Tourism Development Authority. (2016).

Tourism Research and Statistics. URL: http://www.sltda.lk/sites/default/files/Annual_Statistical_ Report_2016.pdf. .

Konarasinghe, K.M.U.B. (2017-a).Forecasting Foreign Guest Nights in Colombo and Greater Colombo of Sri Lanka. *Conference Proceedings of the 14th International Research Conference on Business Management (ICBM)* – 2017, Faculty of Management Studies and Commerce, University of Sri Jayawardanapura, Sri Lanka, 22

Konarasinghe, K.M.U.B.(2017-b). Forecasting Foreign Guest Nights in Southern Coast of Sri Lanka Conference. Proceedings of the Peradeniya University International Research Sessions (iPURSE) -2017, University of Peradeniya, Sri Lanka, 21, 40.

Konarasinghe,K.M.U.B.(2017-c).Forecasting Foreign Guest Nights in Ancient Cities of Sri Lanka. *Conference Proceedings of the 22nd International Forestry and Environment Symposium 2017*, Department of Forestry and Environmental Science, University of Sri Jayewardenepura, Sri Lanka,114.

Konarasinghe, K.M.U.B. (2017-d). Hybrid Trend–ARIMA Model for Forecasting Employment in Tourism Industry in Sri Lanka. *Review of Integrative Business and Economics Research, 6*(4), 214-223.

Konarasinghe, W.G.S., Abeynayake, N.R.,and Gunaratne, L.H.P. (2016).Circular Model on Forecasting Returns of Sri Lankan Share Market. *International Journal of Novel Research in Physics, Chemistry, and Mathematics*, 3(1), 49-56.

Konarasinghe, K.M.U.B. (2016). Time Series Patterns of Tourist Arrivals to Sri Lanka. *Review of Integrative Business and Economics Research*, 5(3), 161-172.