

Forecasting Foreign Guest Nights in Southern Coast of Sri Lanka: An Application of SARIMA Model

KMUB Konarasinghe

Institute of Mathematics and Management, Ranala, Sri Lanka

Abstract: The Southern Coast of Sri Lanka has been flooded with natural beauty with rich biodiversity attracted to the tourists for centuries. There is an uptrend of occupancy by foreign guest up to Galle and beyond after 2008. The increasing of night occupancy will increase the demand for tourism-related business. It creates competition of business and it causes the risk for their business. Therefore, this study was focused on forecasting foreign guest nights in Southern Coast of Sri Lanka to minimize the risk. Monthly data of foreign guest nights for the period of January 2008 to December 2016 were obtained from annual reports of 2008 -2016 published by SLTDA. At first descriptive statistics obtained and SARIMA model was tested for forecasting. The normality and randomness of residuals tested for model validation. Forecasting ability of the models was assessed by relative and absolute measurements of errors. On average night occupancy up to Galle and beyond are 131000 and 40298 respectively. ARIMA (1,0,0)(1,0,1)₆ and ARIMA (1,0,0)(1,2,1)₆ models satisfied the validation criterion. The study concluded that SARIMA performs well in forecasting night occupancy up to Galle and beyond in the Southern coast of Sri Lanka. Occupancy series follows wave-like pattern with trend. It may contain both seasonal and cyclical variation, but the SARIMA is unable to separate them. Therefore, it is recommended to test the Sama Circular Model, in order to see whether it improves the forecasting.

Keywords: Foreign tourist, Normality, Occupancy, Randomness, SARIMA.

I. INTRODUCTION

International tourism industry in Sri Lanka shows an increasing trend after 2008. It increases the occupancy of various regions. Southern coast is the second highest region occupied by foreign tourist while Colombo is the highest. Hikkaduwa, Unawatuna, Bentota, Mirissa, Kogalla and Tangalle are the famous beaches in Southern coast of Sri Lanka. Sunbathing, water sport activities, boat rides, scuba diving, snorkeling, surfing are the popular leisure activities attracted by foreign tourist to Southern coast. In addition the reef and coral gardens, natural swimming pools, witness whales, and dolphins are some other attractions to tourist (SLTDA, 2016). Galle is the capital of the Southern province and its fortifications is one of the UNESCO world heritage sites, located in the Southern Coast. Night occupancy of Southern coast has divided into two parts; as “Up to Galle and Beyond Galle” (SLTDA, 2016). This study focused on occupancy of foreign guest nights up to Galle and beyond.

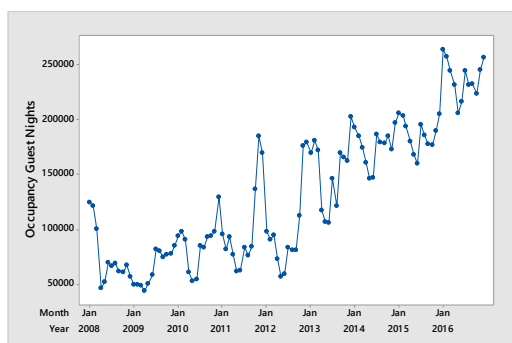


Figure 1: Time Series plot of occupancy of up to Galle

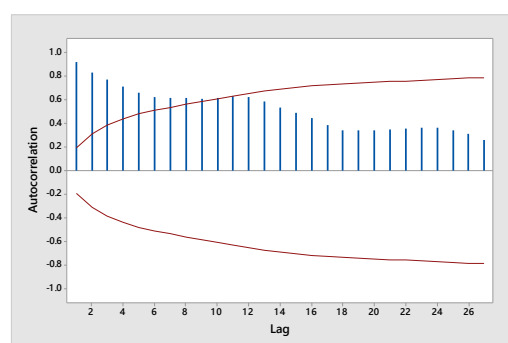


Figure 2: Autocorrelation function occupancy of up to Galle

Fig 1 shows that there is an increasing trend of night occupancy up to Galle by a foreign tourist. Fig 2 confirms the same and shows that the series is non-stationary with existing seasonal behavior.

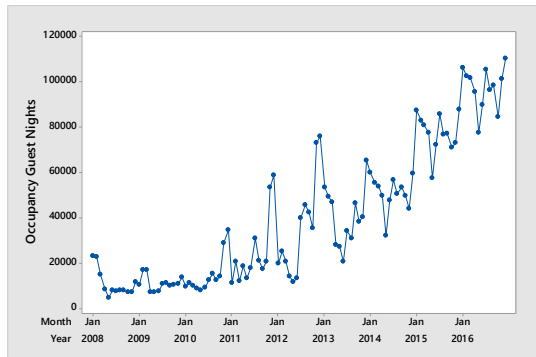


Figure 3: Time Series plot of occupancy of beyond Galle

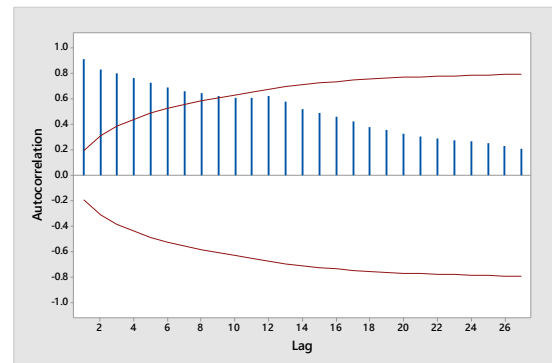


Figure 4: Auto correlation function of occupancy of beyond Galle

Fig 3 shows that there is an increasing trend of night occupancy beyond Galle by a foreign tourist. The series is non-stationary. Figure 4 confirmed the non-stationary of the series and seasonal behavior as well.

Research Problem:

The increasing of night occupancy will increase the demand for tourism-related and other business located up to Galle and beyond. It creates competition of business and it causes the risk for their business and their stakeholders. Then the businesses have to find out the possibilities to create demand for their products and services. 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. Previous studies focus on forecasting occupancy of foreign guest nights of Southern coast, but not separated as up to Galle and beyond (Konarasinghe, 2018). But it was hard to find attempts of forecasting occupancy guest nights focusing up to Galle and beyond in Southern coast of Sri Lanka. Hence, there exists a knowledge gap, which needs to be addressed without delay.

The objective of the Study:

To forecast foreign guest nights up to Galle and beyond in Southern coast of Sri Lanka

II. LITERATURE REVIEW

The literature review based on forecasting guest nights in various destinations. Brannas, and Nordstrom (2000) model the number of Norwegian guest nights in Swedish hotels and cottages and did the demand analysis. They used the integer-valued autoregressive model for the study. Lim, Chang and McAleer (2009) focus on Autoregressive Moving Average (ARMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) approach to examine and forecast tourist accommodation demand in New Zealand using hotel-motel room nights. In Sri Lankan context; Konarasinghe, (2017) used SARIMA to forecast foreign guest nights in ancient cities of Sri Lanka, Konarasinghe (2017) used SARIMA and decomposition techniques for forecasting foreign guest nights in Colombo and Greater Colombo of Sri Lanka. Konarasinghe (2018) used SARIMA, decomposition techniques and Holts Winters three parameter techniques for forecasting foreign guest nights in Hill Country of Sri Lanka. Konarasinghe (2018) used SARIMA, decomposition techniques and Holts Winters three parameter techniques for forecasting foreign guest nights in Southern Coast of Sri Lanka. Hybrid trend - ARIMA used by Konarasinghe (2018) on forecasting domestic guest nights in Ancient Cities of Sri Lanka. Konarasinghe (2018) used SARIMA on forecasting domestic guest nights in Hill Country of Sri Lanka. As concluding remarks of literature, univariate time series techniques are heavily used in guest nights forecasting. The Integer-valued autoregressive, ARMA and SARIMA models are suitable for forecasting guest nights. SARIMA is the most suitable model in forecasting guest nights in Sri Lankan tourism industry.

III. METHODOLOGY

Monthly data of foreign guest nights for the period of January 2008 to December 2016 were obtained from annual reports of Sri Lanka Tourism Development Authority (SLTDA). At first, the Descriptive statistics of arrivals were obtained. Time

series plots were used for pattern identification and Seasonal Autoregressive Integrated Moving Average (SARIMA) model was tested for forecasting. The Anderson–Darling test, Auto-Correlation Function (ACF), and Ljung-Box Q (LBQ)-test were used to validate and fit the model. Forecasting ability of the models was assessed by three measurements of errors; Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE) and Mean Absolute Deviation (MAD).

The Box-Jenkins Method:

This study follows the Box-Jenkins methodology for modeling. The following conceptual framework proposed by Box, Jenkins, and Reinsel (1976) is considered in this study.

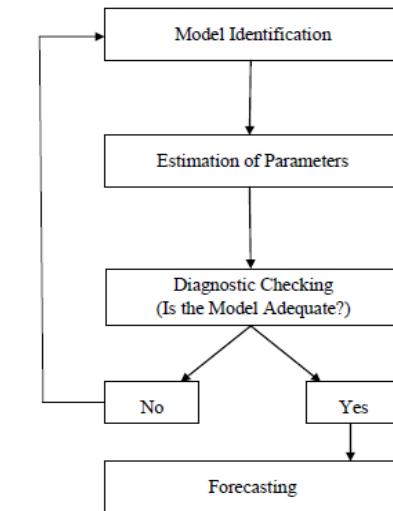


Figure 5: Box- Jenkins Method

Autoregressive Integrated Moving Average (ARIMA):

ARIMA modeling can be used to model much different time series, with or without trend or seasonal components, and to provide forecasts. The model as follows;

An ARIMA model is given by:

$$\phi(B)(1 - B)^d y_t = \theta(B)\varepsilon_t$$

Where; $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 \dots \phi_p B^p$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 \dots \theta_q B^q \tag{1}$$

ε_t = Error term

D = Differencing term

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

Analogous to the simple SARIMA parameters, these are:

Seasonal autoregressive - (Ps)

Seasonal differencing - (Ds)

Seasonal moving average parameters - (Qs)

Seasonal models are summarized as ARIMA (p, d, q) (P, D, Q)_s

Number of periods per season - S

$$(1 - \phi_1 B)(1 - \phi_1 B^s)(1 - B)(1 - B^s)Y_t = (1 - \theta_1 B)(1 - \theta_1 B^s)\varepsilon_t \tag{2}$$

Anderson–Darling test:

This test compares the empirical cumulative distribution function of your sample data with the distribution expected if the data were normal. If this observed difference is sufficiently large, the test will reject the null hypothesis of population normality. For the normality test, the hypotheses are,

H_0 : data follow a normal distribution

H_1 : data do not follow a normal distribution

Auto-Correlation Function (ACF) and Partial Autocorrelation Function (PACF)

Autocorrelation computes and plots the autocorrelations of a time series. Autocorrelation is the correlation between observations of a time series separated by k time units. The plot of autocorrelations is called the autocorrelation function. Partial autocorrelation computes and plots the partial autocorrelations of a time series. Partial autocorrelations, like autocorrelations, are correlations between sets of ordered data pairs of a time series. As with partial correlations in the regression case, partial autocorrelations measure the strength of the relationship with other terms being accounted for. The partial autocorrelation at a lag of k is the correlation between residuals at time t from an autoregressive model and observations at lag k with terms for all intervening lags present in the autoregressive model. The plot of partial autocorrelations is called the partial autocorrelation function. The correct model for the series is identified by analyzing the ACF and PACF. They reflect how the observations in a time series are related to each other. It is useful that the ACF and PACF are plotted against consecutive time lags for the purposes of modeling and forecasting (Brockwell and Davis 2002). The order of the Auto-Regressive (AR) and Moving Average (MA) are determined by these plots. The ACF and PACF plots are used to identify the terms of the SARIMA model.

Ljung-Box Q (LBQ)-test

The Ljung–Box Q test is a type of statistical test of whether any of a group of autocorrelations of a time series is different from zero. Instead of testing randomness at each distinct lag, it tests the "overall" randomness based on a number of lags and is, therefore, a portmanteau test. The Ljung–Box Q test may be defined as:

H_0 : The data are independently distributed

H_1 : The data are not independently distributed

The test statistic is:

$$Q = n(n+2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n-k}$$

Where n is the sample size, $\hat{\rho}_k$ is the sample autocorrelation at lag k , and h is the number of lags being tested. Under H_0 the statistic Q follows a $\chi_{(h)}^2$. For significance level α , the critical region for rejection of the hypothesis of randomness is:

$$Q > \chi_{1-\alpha, h}^2$$

Where $\chi_{1-\alpha, h}^2$ is the $1-\alpha$ -quantile of the chi-squared distribution with h degrees of freedom (Ljung, and Box, 1978).

IV. ANALYSIS

Dataset was outlier free. Analysis of the study consists of three parts;

- i. Descriptive Statistics
- ii. Forecasting foreign guest nights up to Galle.
- iii. Forecasting foreign guest nights beyond Galle.

Descriptive Statistics:

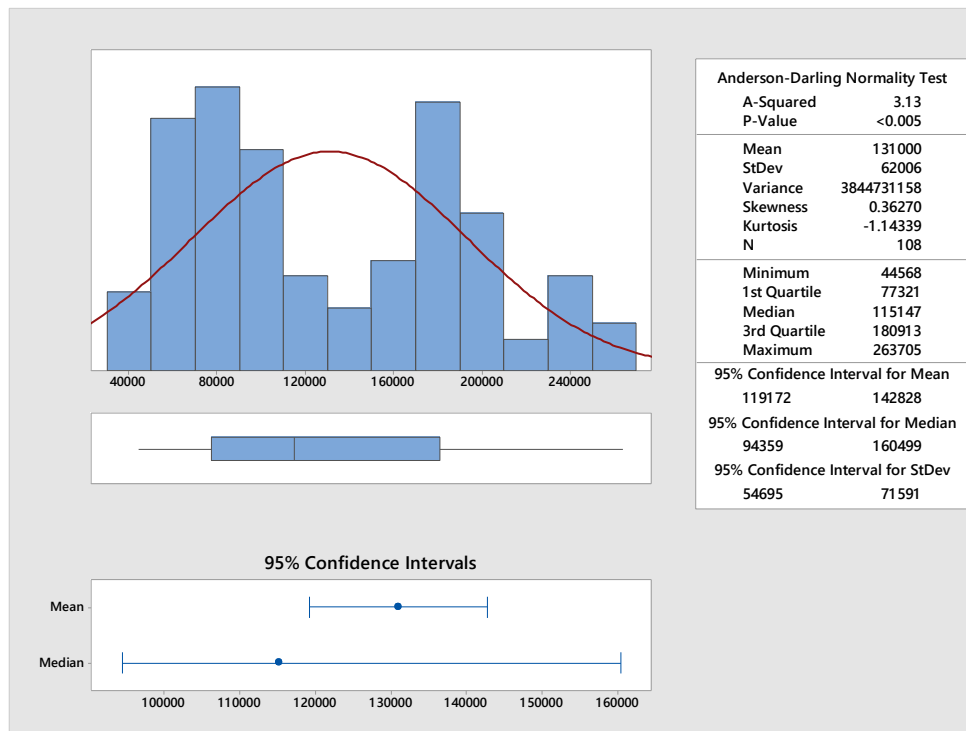


Figure 6: Graphical Summary of Occupancy up to Galle

According to Fig 6, the minimum occupancy by a foreign guest was 44568 whereas the maximum was 263705 during the period. On average night occupancy up to Galle is 131000. The first quartile of occupancy is 77321. It means ¼ of the months were occupied by at most 77321 foreign guests. A median occupancy is 115147 and the third quartile of occupancy is 180913.

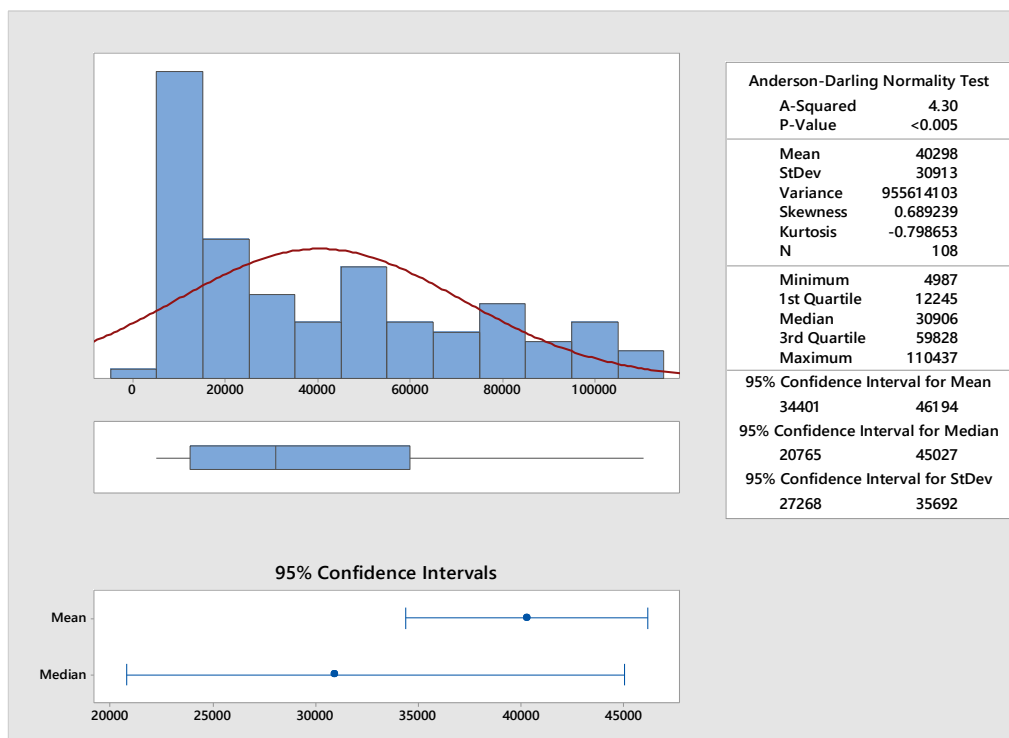


Figure 7: Graphical Summary of Occupancy beyond Galle

According to Fig 7, the minimum occupancy beyond Galle was 4987 whereas maximum was 110437. On average night occupancy beyond Galle is 40298. The first quartile of occupancy is 12245. A median occupancy is 30906 and the third quartile of occupancy is 59828. The p-value of the Anderson-Darling test of both series is less than the significance level (p-value <0.005). As such, the number of occupancy of foreign guest up to and beyond Galle does not follow the normal distribution. They are positively skewed.

FORECASTING FOREIGN GUEST NIGHTS UP TO GALLE:

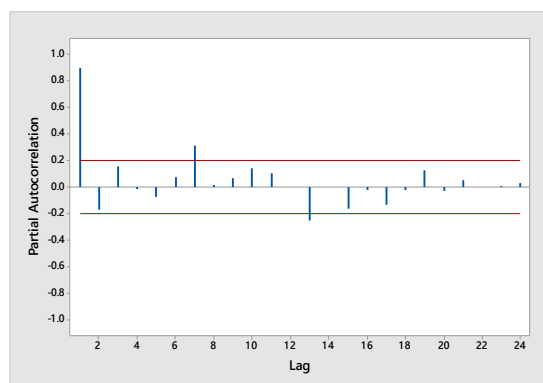
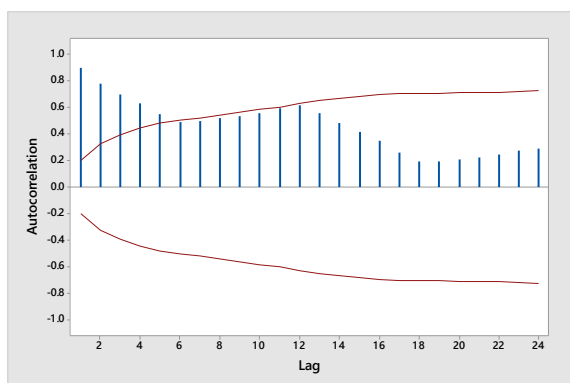


Figure 8: Autocorrelation function of occupancy of up to Galle **Figure 9: Partial Autocorrelation function of occupancy of up to Galle**

The ARIMA process begun by observing ACF and PACF of occupancy of up to Galle has shown in Fig 8 and 9. The SARIMA model runs for night occupancy by the foreign guest up to Galle with two seasons. The results of SARIMA are given in TABLE 1. The model ARIMA(1,0,0)(1,0,1)₆ describes a model that includes one autoregressive parameter, one seasonal autoregressive parameter, and one seasonal moving average parameter and these parameters were computed for the series after no differenced. Both relative and absolute measurements of errors were very low in model fitting and verification. The Anderson-Darling normality test confirmed the normality of residuals (P=0.353) the LBQ- test and ACF confirmed the independence of residuals (h=0).

TABLE 1: Model Summary of SARIMA

Model	Model Fitting		Verification	
	ARIMA (1,0,0)(1,0,1) ₆	MAPE	1.03755	MAPE
MSE		0.0245768	MSE	0.331973
MAD		0.119518	MAD	0.565323
Normality		P=0.353		
Randomness of Residuals		h=0		

FORECASTING FOREIGN GUEST NIGHTS BEYOND GALLE:

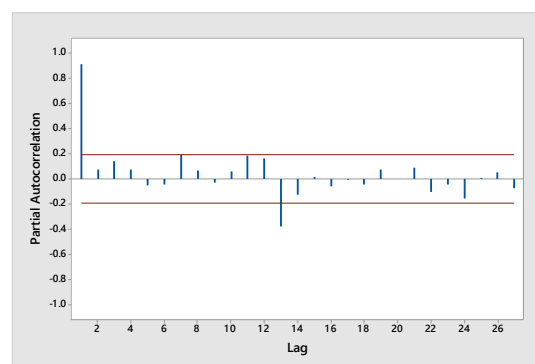
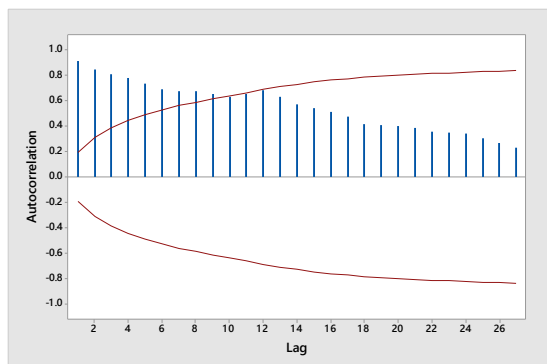


Figure 10: Autocorrelation function of occupancy of beyond Galle **Figure 11: Partial Autocorrelation of occupancy of beyond Galle**

The ARIMA process begun by observing ACF and PACF of occupancy of beyond Galle has shown in Fig 10 and 11. The SARIMA model runs for night occupancy by the foreign guest beyond Galle with two seasons. The results of SARIMA are given in TABLE 2. The model ARIMA(1,0,0)(1,2,1)₆ describes a model that includes one autoregressive parameter, one seasonal autoregressive parameter, and one seasonal moving average parameter and these parameters were computed for the series after no differenced with, and two seasonally differenced. Both relative and absolute measurements of errors were very low in model fitting and verification. The Anderson-Darling normality test confirmed the normality of residuals (P=0.334) the LBQ- test and ACF confirmed the independence of residuals (h=0).

TABLE 2: Model Summary of SARIMA

Model	Model Fitting		Verification	
	ARIMA (1,0,0)(1,2,1) ₆	MAPE	2.19776	MAPE
MSE		0.0817454	MSE	0.105761
MAD		0.220782	MAD	0.278378
Normality		P= 0.334		
Randomness of Residuals		h=0		

The fitted lines in Fig 12 and Fig 13 follow the patterns of the series. Actual occupancy and fits are closer to each other. The deviation of the actual occupancy and fits are very less.

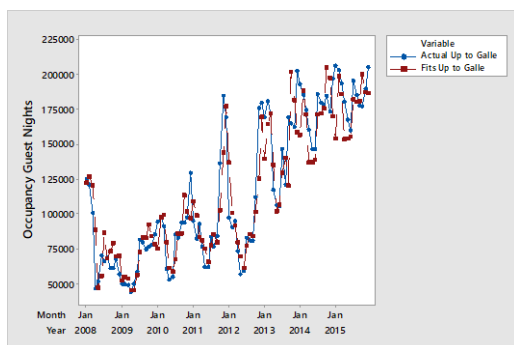


Figure 12: Actual Vs fits of occupancy Up to Galle

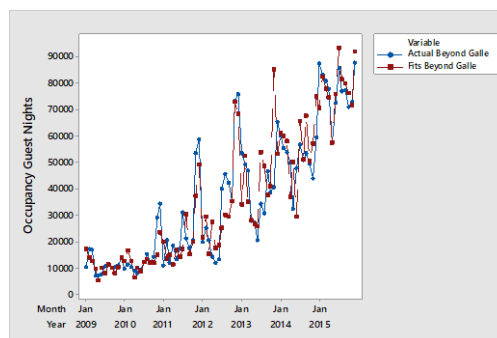


Figure 13: Actual Vs fits of occupancy Beyond Galle

Hence, the SARIMA is capable in forecasting occupancy to the Southern coast. In other words, future night occupancy by the foreign guest up to Galle and beyond can be forecasted by past night occupancy, past errors and seasonal components. Estimated night occupancy by foreign guest up to Galle beyond for the period of January 2019 to December shown in TABLE 3. Accordingly, more guests can be expected beyond Galle. Estimated guest nights of beyond Galle shows an increasing trend and up to Galle shows the decreasing trend.

TABLE 3: Estimated Night Occupancy

Year	Month	Estimated Guests Nights up to Galle	Estimated Guests Nights beyond Galle
2019	January	181228	232048
	February	178965	276918
	March	174694	290355
	April	168572	351979
	May	164810	314104
	June	165156	323609
	July	162933	248472
	August	162563	354415
	September	164101	363428
	October	167587	449528
	November	169065	426096
	December	166603	448991

V. DISCUSSION AND CONCLUSION

The results of this study can be used to forecast the number of occupancy guest nights up to and beyond Galle in the Southern coast of Sri Lanka. This is useful for strategy development and management to take correct decisions in business and authorities to the betterment of the tourism industry in the Southern coast. Forecasting number of occupancy plays an important role in various fields in the business concern. Production planning related to tourism products are the most important area of the tourism industry. The results of the study are of utmost importance in setting up new business opportunities for the host community; such as accommodation, food and beverages, transport and other beach tourism activities. Therefore, the local government can provide guidance, facilities, and employment opportunities for new business avenues to host communities based on forecasting occupancy guest nights. Forecasting occupancy is useful for expansion of existing business in tourism. And assembling the strategic resources (Men, Money, Machines, Materials, Methods, and Knowledge) effectively and efficiently for the development of the tourism industry up to Galle and beyond in the Southern coast of Sri Lanka. Forecasting occupancy needs to estimate the financial requirements of a business. Utilization of capital for various departments is important for the smooth function of the business. Occupancy forecasting provides a guideline of sales, multiple expenses cash and credit flows etc. It is useful for financial controllers for the preparation of their financial budget. In addition, businesses will be able to control their resources in their business. The significant increase in occupancy resulted in a heavy volume of garbage. Then the local government can maintain and improve the standards of formal waste collection services. The negative impacts of tourism occupancy can be minimized by a dedicated environmental impact assessment and monitoring, qualified guides with an enhanced responsibility to enforce regulations and precautionary actions based on the results of this study. High occupancy gives an insight into the necessity of security services implementation, practice, and monitoring, in order to protect the tourism industry and the goodwill of the country.

According to TABLE 3, estimate occupancy of up to Galle shows a decreasing trend. It is recommended to find out the reasons for decreasing and plan to increase the occupancy. Further, it is useful to find out the satisfaction of tourist who occupied beyond Galle to improve the standards of products and services consumed by them. Occupancy series follows the trend with a wave-like pattern with trend. It may contain both seasonal and cyclical variation, but the SARIMA is unable to separate them. The Sama Circular Model (SCM) of Konarasinghe (2018) is a recently developed univariate forecasting technique, which can be used to capture; trend, seasonal and cyclical patterns. Therefore, it is recommended to test the Sama Circular Model, in order to see whether it improves the forecasting.

The study concluded that the SARIMA model is suitable for forecasting foreign occupancy guest nights up to Galle and beyond. Results of the study agree with the results of Lim, Chang, and McAleer (2009), Konarasinghe (2017) and Konarasinghe (2018).

REFERENCES

- [1] Peter. J. Brockwell, and Richard. A. Davis. "Introduction to time series and forecasting", 2nd.ed, Springer-Verlang, 2002.
- [2] K. Brannas, and J. Nordstrom. "A bivariate integer value allocation model for guest nights in hotels and cottages", Umea Economic Studies. Sweden. 2000.
- [3] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel. "Time series analysis, forecasting and control", 3rd ed. Prentice Hall, Englewood Clifs, NJ. 1976.
- [4] K.M.U.B. Konarasinghe. "Forecasting foreign guest nights in hill country of Sri Lanka", Review of Integrative Business and Economics Research, Vol. 7, No. 4, pp. 41-56. 2018.
- [5] W.G.S. Konarasinghe. "Improved circular model on forecasting arrivals from Western European countries to Sri Lanka". Proceedings of the 4th ISM International Statistical Conference (ISM-IV) of Sunway University, Malaysia.2018.
- [6] K.M.U.B. Konarasinghe. "Forecasting ability of univariate time series approach in foreign guest nights in the southern coast of Sri Lanka". International Journal of Research and Review. Vol. 5, No. 8, pp.17-25. 2018.

International Journal of Novel Research in Physics Chemistry & MathematicsVol. 5, Issue 3, pp: (22-30), Month: September - December 2018, Available at: www.noveltyjournals.com

- [7] K.M.U.B. Konarasinghe. "Forecasting domestic guest nights in ancient cities of Sri Lanka: Hybrid approach". Conference Proceedings of the 11th International Research Conference, General Sir John Kotelawala Defence University, Sri Lanka, pp. 36. 2018.
- [8] .M.U.B.Konarasinghe. "Forecasting domestic guest nights in hill country of Sri Lanka". Conference Proceedings of the 5th International Conference on Multidisciplinary Approaches -2018, Faculty of Graduate Studies, University of Sri Jayawardanapura. Sri Lanka, pp.131. 2018.
- [9] K.M.U.B. Konarasinghe. "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 Jaywardenepura, Sri Lanka,pp.114. 2017.
- [10] K.M.U.B. Konarasinghe. "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, pp.22.2017.
- [11] Lim, C. Chang, and M. McAleer. "Forecasting h(m)otel guest nights in New Zealand". International Journal of Hospitality Management, Vol. 28, No. 2, pp. 228-235. 2009.
- [12] Sri Lanka Tourism Development Authority (SLTDA). "Tourism Research and Statistics". Retrieved from http://www.slttda.lk/sites/default/files/Annual_Statistical_Report_2016.pdf. 2016.
- [13] Z. Schwartz. and S. Hiemstra. "Improving the accuracy of hotel reservations forecasting: Curves Similarity Approach". Journal of Travel Research. Vol. 36, No. 1, pp. 3-14. 1997.