Vol 2, No 2, July - December 2017, pp: 1 - 10

Faculty Publications
ISSN: 3008-1017
Title DOI: https://doi.org/10.62458/021024

Forecasting International Tourist Arrivals in Cambodia Using Holt-Winters Methods

Mara Mong*

ABSTRACT

The ability to forecast something is what most people wish. The prediction of tourist time series data is very important in tourism management. There are many methods developed to analyze and forecast univariate time series. This study is mainly focused on forecasting quarterly international tourist arrivals in Cambodia. Holt-Winters method with additive and multiplicative seasonality is used. Diagnostic checking reveals that both models are appropriate but based mean absolute percentage error(MAPE), the model with additive seasonality is better and hence, chosen.

keywords: Tourist arrivals, Holt-Winters' method, seasonality, trend, ACF, MAPE, Ljung-Box test, augmented Dickey-Fuller test, forecast.

* Mara Mong, PhD
CamEd Business School
Email: mmara@cam-ed.com

INTRODUCTION

The tourism which is often referred to as the non-smoke industry is one of the most important sectors for a successful economy in the country. It accounts for a large part of Gross Domestic Product (GDP) of Cambodia. Tourism sector is the third largest sector of Cambodia economy after the agriculture and garment sector, and the second contributor of foreign earning after the garment sector. In 2012, the revenue from the sector represented 15.7% of total GDP of Cambodia (PPS, 2013). This key component of the service sector accounted for 17% of GDP in 2013 (ADB, 2014).

Few studies on the modelling and forecasting of time series of international tourist arrivals to Cambodia have been encountered. (Mong, 2015) used regression equations of various forms such as simple linear, quadratic, exponential, and power forms with time indices as independent variables to forecast the time series value based on yearly international tourist arrivals for the period 1997 to 2013. (Ou, 2016) used SARIMA model to forecast monthly tourism demand in Cambodia, implemented with R function, named auto.arima() based on the data from period January 2000 to December 2015.

Since it is simple but powerful for modelling the trend, seasonality, and random errors, Holt-Winters method is widely used. Here are some examples. (Amanda and Qian, 2008) used Holt-Winters time series model for population forecasting using a series of annual population estimates from 1960 to 2005 to construct population forecasts for states in the U.S. and for Virginia's planning districts and counties. (Gnanapragasam and Cooray, 2016) used Holt-Winters method to analyze and forecast tourist arrivals after the internal conflict in Sri Lanka.

METHODOLOGY

In this study, we use quarterly tourist arrival data from the period first quarter of 2000 to fourth quarter of 2016. The data were extracted from the annual report on tourism statistics of Statistics &Tourism Information Department (MOT 2008, MOT 2010, and MOT 2016). Time series plot, decomposition, and ACF plot are used to observe the behavior of the series. All statistical hypothesis tests were conducted at 5% significance level and interval estimates were constructed with 95% confidence level.

We use both Holt-Winters model with additive seasonality and Holt-Winter method with multiplicative seasonality to fit the data. The mathematical equations for the additive model are as follow (Stephen and Kenneth, 2014)

$$S_t = \alpha\left(x_t - C_{t-p}\right) + \left(1-\alpha\right)\left(S_{t-1} + T_{t-1}\right) \quad (1)$$

$$T_t = \beta\left(S_t - S_{t-1}\right) + \left(1-\beta\right)T_{t-1} \quad (2)$$

$$C_t = \gamma\left(x_t - S_t\right) + \left(1-\gamma\right)C_{t-p} \quad (3)$$

$$F_{t+1} = \left(S_t + T_t\right) + C_{t+1-p} \quad (4)$$

The multiplicative model is described by mathematical equations as (Stephen and Kenneth, 2014)

$$S_t = \alpha \frac{x_t}{C_{t-p}} + (1-\alpha)(S_{t-1} + T_{t-1}) \quad (5)$$

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} \quad (6)$$

$$C_t = \gamma \frac{x_t}{S_t} + (1-\gamma)C_{t-p} \quad (7)$$

$$F_{t+1} = (S_t + T_t)C_{t+1-p} \quad (8)$$

where S_t is the smoothed value, T_t is the estimated trend, C_{t+1-p} is the estimated cyclical factor for the period $(t+1)$, and F_{t+1} is the forecast value.

To proceed the calculation using the equation (1) through (8), the initial value for the smoothed value S_t, the initial value for trend T_t and the initial value for cyclical factor C_t are required. How to determine these values varies depending on the authors or the forecasters. The choice for the values for smoothing constants $\alpha, \beta,$ and γ is based on the minimum forecast error measure MAPE. In this study, we implemented these works with R 3.3.3 and Excel 2016. We used mean absolute percentage error (MAPE) as the measure of accuracy.

For the model diagnostic checking, correlogram of the residuals is constructed and Ljung-Box test is conducted to check the goodness-of-fit of the model.

RESULTS AND DISCUSSION

Figure 1 shows the time series plot for international tourist arrivals in Cambodia. Figure 2 shows the components of the time series after decomposing. It is obvious about upward trend and seasonality and random errors. Moreover, since the p-value in augmented Dickey-fuller test is of 0.5263 > 5%, it can be concluded with 95% confidence that there is a trend in the series. Thus, Holt-Winters method is appropriate.

For Holt-Winters with additive seasonality model, we obtain the smoothing constants $\alpha = 0.3290045, \beta = 0.1334714$ and $\gamma = 1$. As the measure of errors in fitting from period 2001Q1-2016Q4, we have mean absolute percentage error (MAPE) equal to 7.377987. The optimal values for the smoothing constants in multiplicative seasonality model are $\alpha = 0.3309891$, $\beta = 0.08719083$, and $\gamma = 0.9276871$ with MAPE equal to 7.558564 which is very slightly higher than that from the additive model. This information is summarized in Table 1.

The plot of residuals from additive model and multiplicative model suggest that they have means of zero and constant variances (Figure 3 and Figure 4). Also, the ACF of both models reveals the residuals do not exceed the significant bound through lags1-10 (Figure 5 and Figure 6).

Figure 1 International Tourist Arrivals

Furthermore, p-value of Ljung-Box test for first 10 lags in additive model is 0.2388 and in multiplicative model 0.3567. These p-values indicate that there is insufficient evidence of non-zero autocorrelation at lags 1-10. Hence, both models satisfy the diagnostic checking. The forecasts by each model for 3 years (12 quarters) ahead with 95 confidence intervals for the forecast values are shown in Table 2 and Table 3, respectively. Figure 7 and Figure 8 show the graphs of fitted and forecasting values with 95% confidence intervals of both models.

Table 1 Summary of models

Type of Mode	Smoothing constants	MAPE
Additive	$\alpha = 0.3290045, \beta = 0.1334714\ \gamma = \ 1$	7.377987
Multiplicative	$\alpha = 0.3309891\ ,\ \beta = 0.08719083\ , \gamma = 0.9276871$	7.558564

Table 2 Forecasts of International tourist arrivals Using Holt-Winters Method with Additive Seasonality

Year-Quarter	Forecasts	Lower Limits	Upper Limits
2017Q1	1,421,094	1,335,429	1,506,758
2017Q2	1,114,877	1,023,450	1,206,305
2017Q3	1,236,970	1,138,817	1,335,123
2017Q4	1,588,222	1,482,430	1,694,013
2018Q1	1,506,018	1,359,954	1,652,082
2018Q2	1,199,802	1,046,365	1,353,239
2018Q3	1,321,894	1,160,281	1,483,508
2018Q4	1,673,146	1,502,585	1,843,707
2019Q1	1,590,943	1,384,809	1,797,077
2019Q2	1,284,727	1,069,459	1,499,994
2019Q3	1,406,819	1,181,719	1,631,919
2019Q4	1,758,071	1,522,468	1,993,675

Table 3 Forecasts of International tourist arrivals Using Holt-Winters Method with Multiplicative Seasonality

Year-Quarter	Forecasts	Lower Limits	Upper Limits
2017Q1	1,437,013	1,356,735	1,517,292
2017Q2	1,102,745	1,019,259	1,186,231
2017Q3	1,224,524	1,133,545	1,315,504
2017Q4	1,603,150	1,531,970	1,674,330
2018Q1	1,531,136	1,395,845	1,666,428
2018Q2	1,173,810	1,043,317	1,304,303
2018Q3	1,302,186	1,162,431	1,441,941
2018Q4	1,703,238	1,586,574	1,819,903
2019Q1	1,625,259	1,439,630	1,810,888
2019Q2	1,244,875	1,071,416	1,418,334
2019Q3	1,379,848	1,194,623	1,565,073
2019Q4	1,803,326	1,640,422	1,966,231

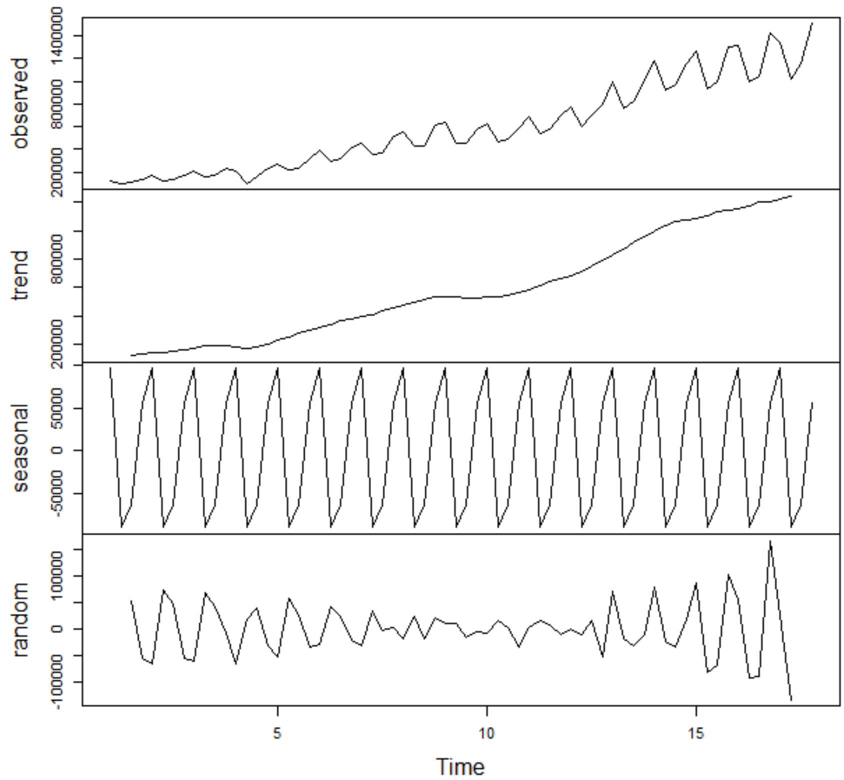

Figure 2 Components of the series: trend, seasonal, and random error

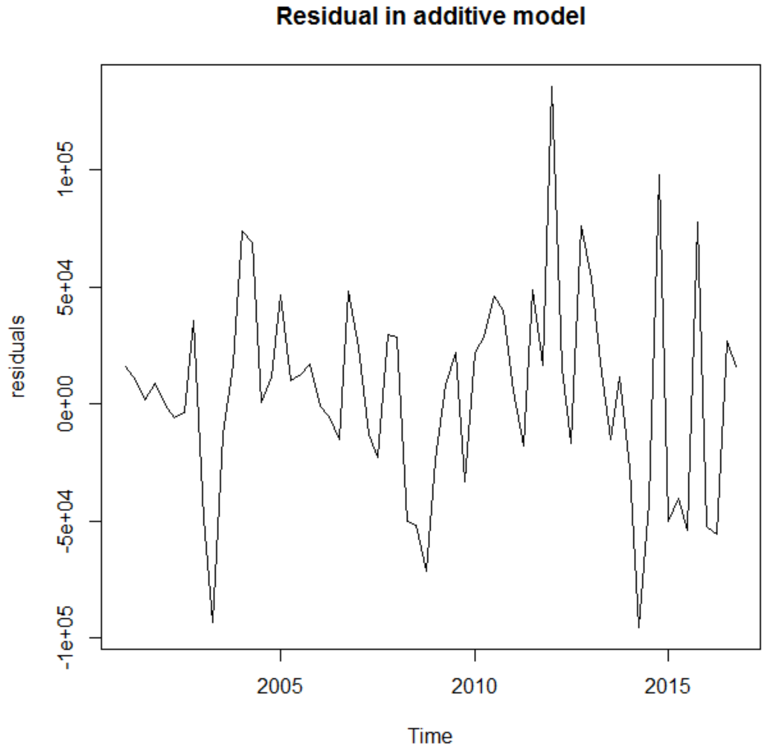

Figure 3 Plot of residuals in Additive model

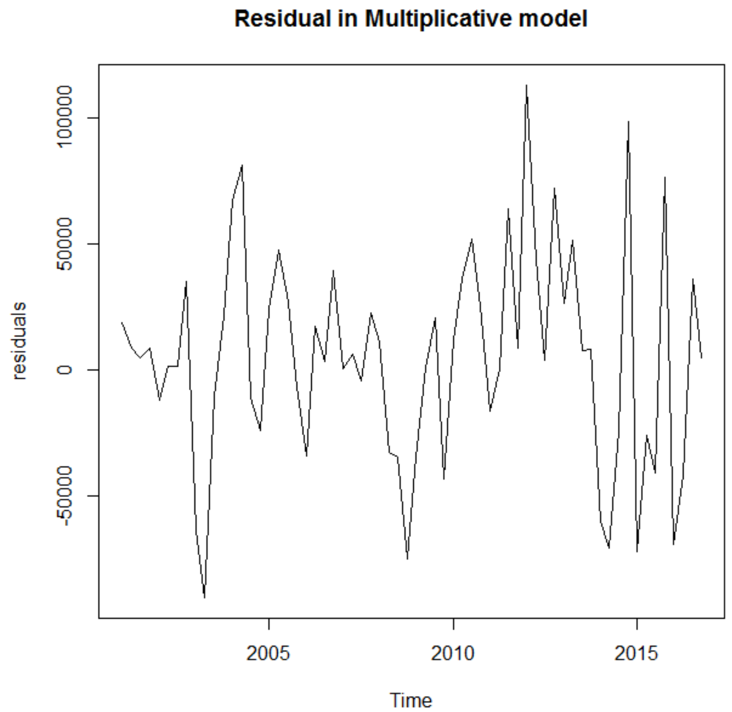

Figure 4 Plot of residuals in Multiplicative Model

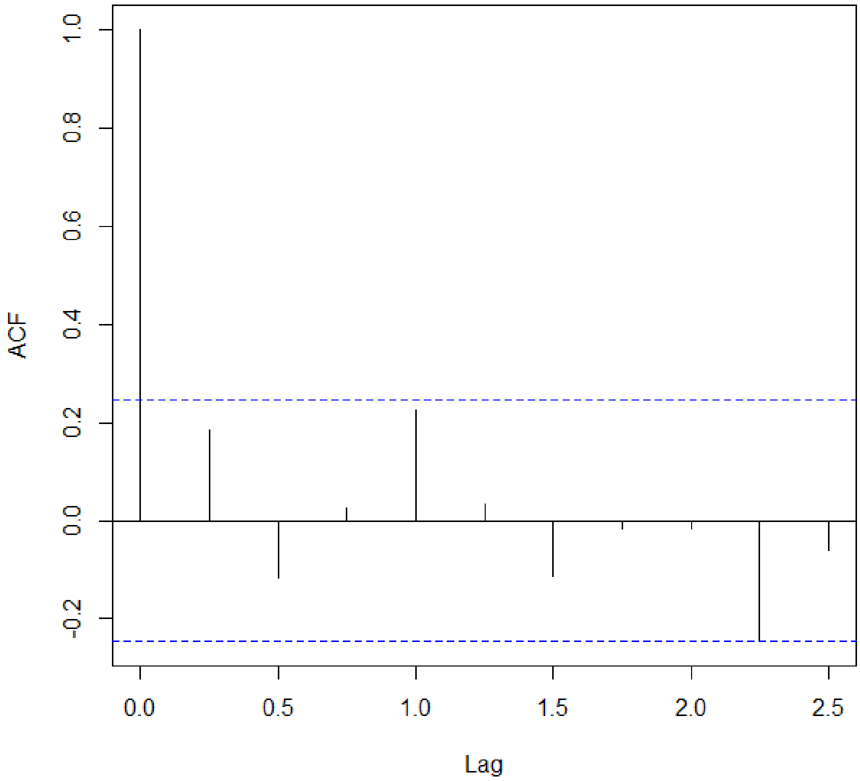

Figure 5 ACF of the residuals in additive model

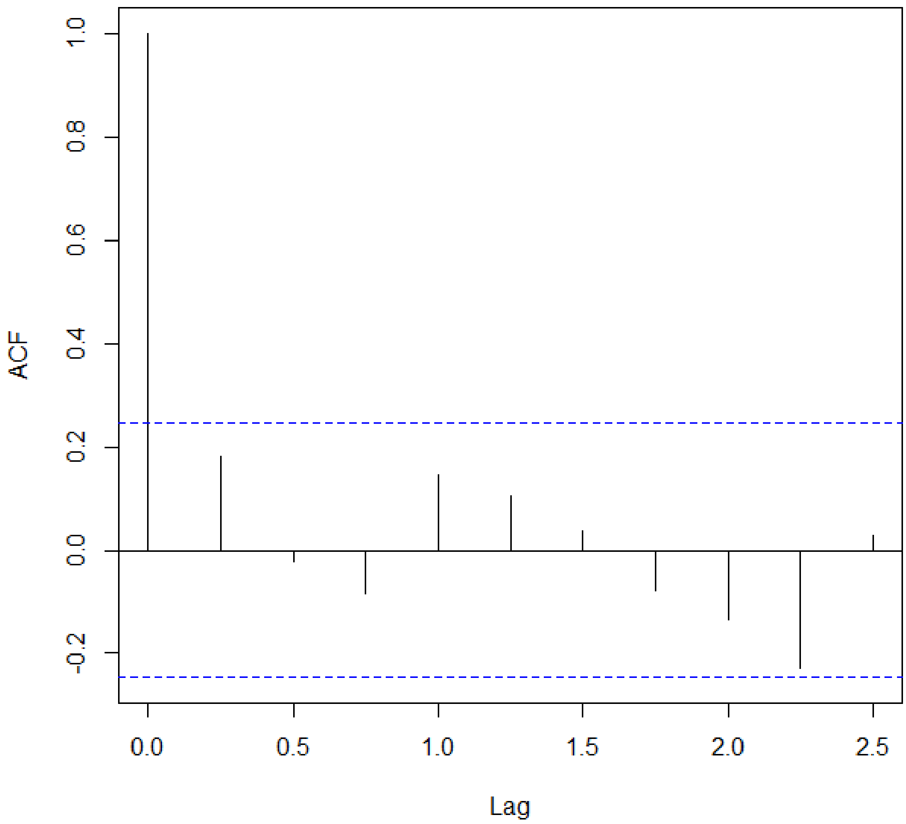

Figure 6 ACF of the residuals in multiplicative model

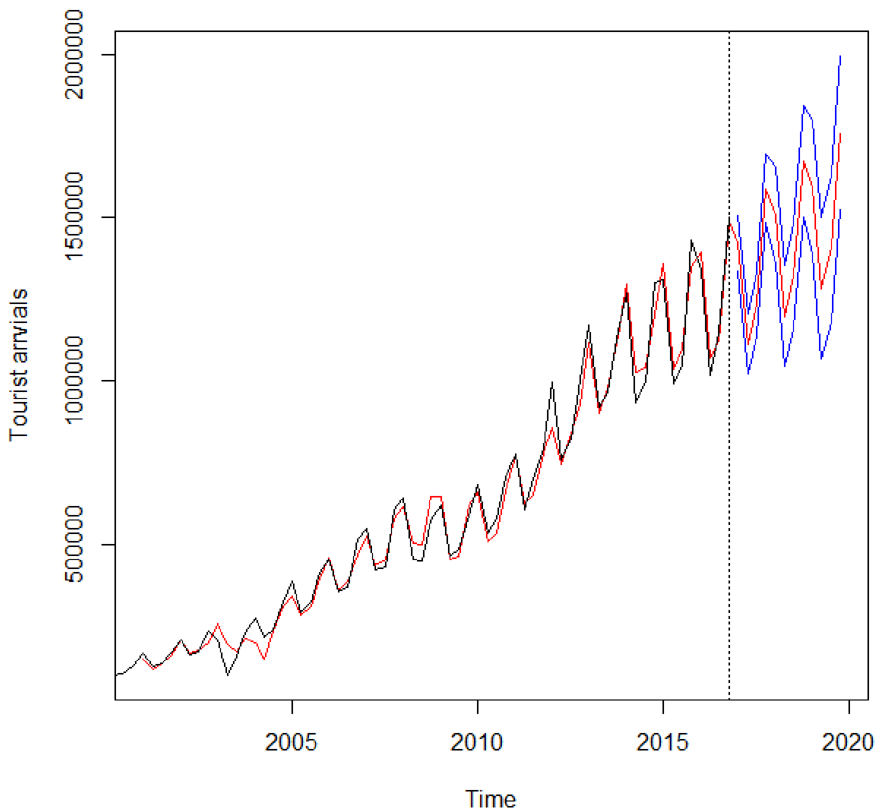

Figure 7 Fit and Forecast by Holt-Winters with additive seasonality

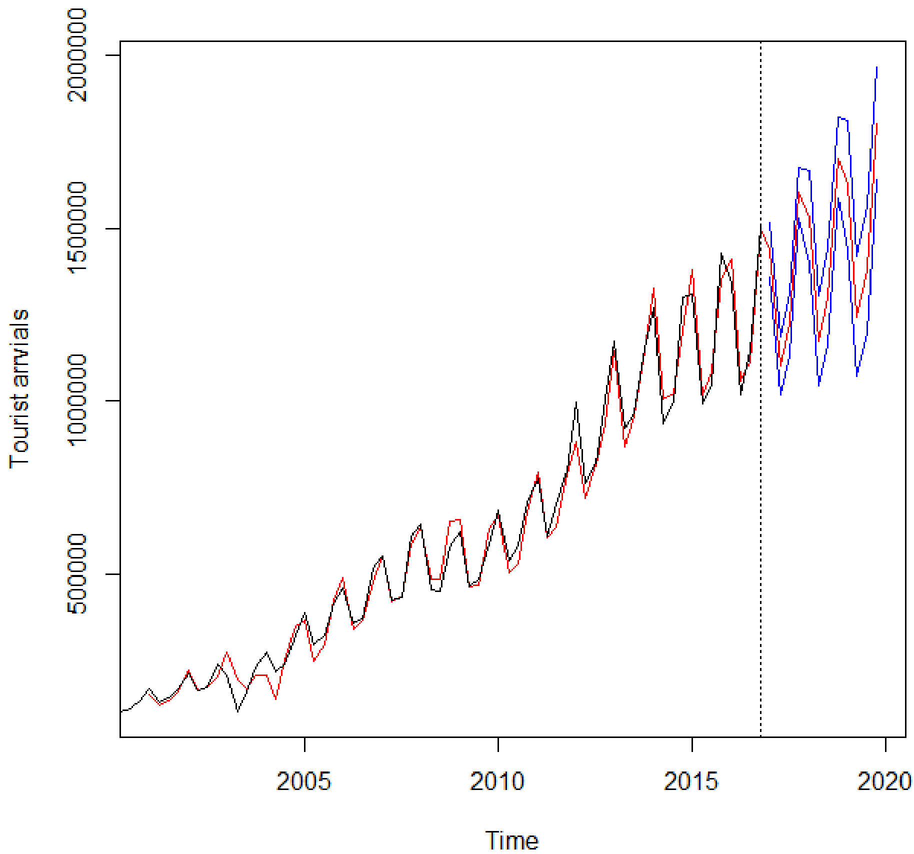

Figure 8 Fit and Forecast by Holt-Winters with multiplicative seasonality

CONCLUSION AND RECOMMENDATION

Since MAPE (7.38%) in additive model is smaller, we choose Holt-Winters with additive seasonality model which has the smoothing constants $\alpha = 0.3290045, \beta = 0.1334714$ $\gamma = 1$ as the best model. Therefore, this model should be recommended for the short term forecasting the international tourist arrivals. However, no method is perfect in forecasting the time series. A limitation is that the forecasting is done just based on the past data and we assume that the behavior of the data will remain the same in the future. It is recommended that the model be updated when new data are available to enter into the sample and the forecast should be made for short term in the nearest future as possible.

REFERENCES

ADB (2014). Diversifying Beyond Garments and Tourism Country Diagnostic Study: Country Diagnotic Study. Philippines.

Gnanapragasam, S., & Cooray, T. (2016) **Tourist Arrivals After The Internal Conflict in Sri Lanka: A Time Series Analysis Using Holt-Winters' Method.**

Ministry of Tourism (2008), Tourism Statistics Report, Phnom Penh, Cambodia.

Ministry of Tourism (2010), Tourism Statistics Report, Phnom Penh, Cambodia.

Ministry of Tourism (2016), Tourism Statistics Report, Phnom Penh, Cambodia.

Mong Mara (2015), **The Creation of Mechanism for Deciding a Regression Model**, Phnom Penh

Ou Phichhang (2016), **Forecasting of Tourism Demand in Cambodia: Automatic Seasonal ARIMA model**, First Angkor International Conference on Applied Mathematics and Statistics: AICAMS 2016

PPS (2013), **Tourism Sector In Cambodia**, Phnom Penh, Cambodia.

Stephen G. Powell and Kenneth R. Baker (2014), **Management Science: The art of Modeling with Spreadsheets**, fourth edition, John Wiley & Sons, Inc., USA.

Walters, A. and Q. Cai (2008). "Investigating the use of Holt-Winters time series model for forecasting population at the State and sub-State levels." University of Virginia.