ARIMA Model for Time Series Forecasting: A Study on Tourist Arrivals in Mandalay

Chaw Ei Ei Tun[1]

Abstract

Tourism is one of the most important service export industries and earners of foreign exchange for several countries today. The tourism phenomenon is vital interest not only to government and other organizations but also to other related industries such as hotel industries, transportation services and so on. Especially, for region such as Mandalay, in which tourism region a very important sector of the economy and possess a large part of GDP. It is usually essential to measure seasonal variation of tourism time series data in order to provide relevant information to related organization and department. The objectives of the study are to examine the best-fitted model and to forecast the future of monthly tourist arrival in Mandalay. The number of international monthly tourist's arrival to Mandalay by region from January 2012 to December 2019 was applied to investigate in this study. It was found that this time series was likely to have seasonal cycle, the lowest value of seasonal index is in June and the highest is in November. Theoretical framework that it was used the Box- Jenkins methodology for seasonal ARIMA models. After constructing the appropriate models, SARIMA $(\mathbf{1},\mathbf{1},\mathbf{0})\times(\mathbf{0},\mathbf{1},\mathbf{1})_{\mathbf{12}}$ was occurred as valid and appropriate model can be considered for forecasting. This appropriate model was utilized them to generate the forecasts of demand within Mandalay tourism. The obtain result can provide important information needed for an adequate destination.

Keywords: Tourism, Seasonal Autoregressive Integrated Moving Average(SARIMA), Mandalay, International Tourist Arrivals, Forecasting.

1. Introduction

Tourism is a very important source of income for developing especially in the region of the South East Asia which is characterized from its very important natural and cultural resources. Tourism industry is one of the world's fasted developing industries. In 20th Century, the tourism industry experienced universal expansion that has obvious economic, social and political benefits. The benefits of tourism have been large especially for developing poor countries that have limited sources of foreign currency; it has an essential source of income and employment.

International tourism accounts for about 25-30% of all global services and employs over 100 million people all around the world. Great importance of tourism activity was in the past proved mainly in the 20th and 21th century tourism was seen as a significant potential method to accelerate economic growth in the developing world.

Nowadays tourism becomes an increasingly import sector of the world economy. Tourism is one of the world's booming industries supporting 313 million jobs and generating 10.4% of the world GDP. The future plans for tourism are brighter than ever.

The Government has been making continued efforts to promote Myanmar tourism industry since the year 1996 was designed as the Visit Myanmar, tourism has become a key contributor to the economy. Myanmar's tourism industry is enjoying substantial growth and the best other sectors of the economy. In January 2019, total tourist arrivals to Myanmar reached approximately 4.88 million with 150229 of tourists entering through airports and 11750

Associate Professor, Department of Statistics, University of Co-operative and Management, Sagaing

through border gateways. Tourism development is also analysed one of the first priorities by the current government. In fact, the Ministry of Hotels and Tourism (MOHT) in the tourism mater plan 2013-2020 declaims follows.

Mandalay, the last royal city of Myanmar, invites many local and foreign visitors. Tourist arrivals this year have increased by 30 percent compared with last year data. Among the foreign visitors, Chinese visitors rank highest over others this year according to upper Myanmar Committee of Hotels and Tourism. Mandalay is the economic center of upper Myanmar and examined the center of Burmese culture. Mandalay is not only tourist's interesting places but also earning foreign income from tourism and other industries.

Thus, this research has studied the status of the tourism industry, which considerably supports Myanmar economy, based on past and present situation.

1.1 Objectives of the Study

The objectives of the study are:

1. to examine the best-fitted model of monthly tourist arrival in Mandalay.
2. to forecast the future tourist arrivals in Mandalay.

1.2 Method of Study

The present study makes an extensive use of secondary time series data of tourist arrivals from different countries to Mandalay during the year 2012 to 2019. The different approaches used for forecasting the number of tourist arrivals in Myanmar by using Box and Jenkins Autoregressive Integrated Moving Average Process (ARIMA) method.

1.3 Scope and Limitations

In this study monthly tourist arrivals in Mandalay during the period from 2012-2019 are conducted. The required information and data are obtained by Ministry of Hotels and Tourism. Among various time series analysis SARIMA are used in this study.

2. Literature Review

Msofe & Mbago (2019) studied "Forecasting international tourist arrivals in Zanzibar using box – jenkins SARIMA model". The objective of this paper is to forecast international tourist arrivals in Zanzibar. It is using Seasonal Autoregressive Integrated Moving Average (SARIMA) model. Secondary data are using from January 1995 to December 2017. The acceptable of the fitted model was confirmed by Ljung-Box test statistic. The fitted model was used to create monthly forecasts from January 2018 to December 2019 with 95% confidence interval. The forecasting performances of the models were evaluated on the basis of Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). The forecasts value show that the number of tourists visiting Zanzibar is increasing with seasonal pattern.

Makoni & Chikobvu (2018)studied "Modelling and forecasting of tourist arrivals at one of the Seven Natural Wonders of the World, the Victoria Falls Rainforest". The purpose of this paper is to support quantitative techniques that will aid with accurate tourist arrivals forecasting and other models of tourist arrivals. The secondary data are used from January 2006 to December 2017 gained by the Zimbabwe Tourism Authority and Zimbabwe Parks and Wildlife Management Authority. A time series is an upward trend in tourist arrivals with wide fluctuations. The SARIMA $(2, 1, 0)(2, 0, 0)_{12}$ model fits well to the data and better other SARIMA models and the naïve, seasonal naïve and Holt-Winters exponential smoothing models. A two-year future forecast is done using this model. It is providing reasonable forecasts that evidence a general increase in tourist arrivals.

Zahedjahromi, M. (2018) studied “Forecasting Tourist’s Arrivals to the USA with SARIMA models”. The objective of the study is to identify a model best fitting the tourist in the USA data and to forecast the number of tourists entering the United States for 2018. The method of maximum likelihood was used to estimate the parameters and to forecast the number of tourists in the future. The secondary data from 1998 to 2011, which have been recorded annually by Rachel Passmore at Census School of New Zealand, show that the SARIMA model may fit most adequately. The fitted model is forecast that the number of tourists who enter the U.S. will arrive at approximately 540,000 in 6 years, which describes 2.6 times increase, evaluated to the number of tourists entering the U.S. in 1998.

Borhan & Arsad (2014) studied “Forecasting International Tourism Demand from the US, Japan and South Korea to Malaysia: A SARIMA Approach”. The objective of the study is to contrast the model and to forecast tourism demand for Malaysia by three selected countries: the US, Japan and South Korea. This study analysed monthly time series data for the period from January 1999 to December 2012. The well-known Box-Jenkins seasonal ARIMA modeling procedures is applying. The best model for the number of tourist arrivals from the US is SARIMA $(1,1,1)\ (1,0,1)_{12}$.

The best model for the number of tourist arrivals from Japan and South Korea, are indicated a SARIMA $(1,0,0)\ (1,0,0)_{12}$ model and SARIMA $(0,1,1)\ (0,1,1)_{12}$ model. The results show the number of tourist arrivals from the three countries consist of strong seasonal component as the arrivals strongly dependent on the season in the country of origin. The findings of the study also show that the number of tourist arrivals from the US and South Korea will remain to increase in the near future. Meanwhile the arrivals from Japan are forecasted to show decline in the near future. Therefore, tourism authorities in Malaysia need to enhance the promotional try to attract more tourists from Japan to visit Malaysia.

Brida & Garrido (2011) studied “Tourism Forecasting using SARIMA models in Chilean Regions”. The objective of the study is search for the best SARIMA specification for forecasting tourist arrivals in thirteen regions of Chile. Monthly secondary data of the arrivals from January 2004 to Marco 2009 was applied by the National Institute of Statistics (INE). The forecasting interpretation is assessed using data for the limit October 2008 to March 2009. The Box-Jenkins method, the method of minimizing the forecast error and the reg ARIMA method of the X12-ARIMA package are using performance of specification model. It is comparing the performance of the three methods according to their forecast determinations. Regions have different SARIMA specifications, liking the underlying differences in tourism infrastructure and capacities available within each Region.

3. Methodology

3.1 Seasonal Autoregressive Integrated Moving Average, SARIMA$(\mathbf{p},\mathbf{d},\mathbf{q}) \times (\mathbf{P},\mathbf{D},\mathbf{Q})_s$ Model

The ARIMA model is for non-seasonal non-stationary data. Box and Jenkins have derived this model to deal with seasonality. The theoretical justification for modeling univariate time series of traffic flow data as seasonal ARIMA processes is constructed in the time series theorem known as the world decomposition. Therefore, it is also necessary to support a claim that an appropriate seasonal difference will induce stationarity. Therefore, it is also necessary to support an assertion that an appropriate seasonal difference will induce stationarity.

The generalized form of SARIMA $(\mathbf{p},\mathbf{d},\mathbf{q}) \times (\mathbf{P},\mathbf{D},\mathbf{Q})_s$model can be written as:

$$\boldsymbol{\varphi_p(B)\emptyset_p(B)^s(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\theta_Q(B^s)a_t} \qquad (1)$$

The seasonal components are:

$$\boldsymbol{AR(P)}: \emptyset_p(B)^s = \mathbf{1} - \emptyset_1 \boldsymbol{B^s} - \emptyset_2 \boldsymbol{B^{2s}} - \cdots - \emptyset_p \boldsymbol{B^{Ps}}$$

$$\boldsymbol{MA(Q)}: \boldsymbol{\theta_Q(B)^s = 1 - \theta_1 B^s - \theta_2 B^{2s} - \cdots - \theta_Q B^{Qs}}$$

The non-seasonal components are:

$$\boldsymbol{AR(P)}: \boldsymbol{\varphi_p\ (B) = 1 - \varphi_1 B - \varphi_2 B^2 - \cdots - \varphi_p B^p}$$

$$\boldsymbol{MA(Q)}: \boldsymbol{\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_p B^Q}$$

whereB is the backshift operator such that $\mathbf{BZ_t = Z_{t-1}}$

$(\mathbf{1-B})^{\mathbf{d}}$ = non-seasonal difference, $(\mathbf{1-B^s})^{\mathbf{D}}$ = seasonal difference, p= non-seasonal AR order, d= degree (order) of non-seasonal differencing, q= non-seasonal MA order, P= seasonal AR order, D= degree (order) of seasonal differencing, Q= seasonal MA order, s= the number of seasons per year, $\mathbf{a_t}$ =the white noise process at period t.

It is identically and normally distributed with mean zero, variance $\boldsymbol{\sigma^2}$; and $\mathbf{cov}\ (\mathbf{e_t}, \mathbf{e_{t-k}}) = \mathbf{0} \forall \mathbf{k} \neq \mathbf{0}$, that is, $\{\mathbf{e_t}\} \sim \mathbf{WN}\ (\mathbf{0}, \boldsymbol{\sigma^2})$.

From a practical point of view, fitted seasonal ARIMA models give linear state transition equations that can be applied recursively to make single and multiple interval forecasts.

3.2 Stationarity of the Time Series

SARIMA models are defined for stationary time series, thus there was a need to check whether the data are stationary or nonstationary. Time plot and Augmented Dickey Fuller (ADF) test were used to test for stationarity of the series of tourist arrival to Mandalay.

3.2.1 Differencing Technique

The differencing technique with both seasonal and non-seasonal differencing were used to transform the series from non-stationary to stationary. Seasonal differencing of the first order was employed to remove seasonality in the given time series data while non-seasonal differencing was employed to get rid of the trend. Seasonal and non-seasonal differencing of the first order can be stated as presented in equations (1).

$(1 - B^{12}) Z_t = Z_t - Z_{t-12}$ (2)

$(1 - B) Z_t = Z_t - Z_{t-1}$ (3)

$B^j Z_t = Z_{t-j}$, j=0, 1, 2, …

3.3 Box-Jenkins Procedure

Box and Jenkins model building procedure which covered four steps was utlized in order to find the best model to fit within the class of SARIMA models. There are four steps in Box and Jenkins procedure for time series analysis.

3.3.1 Model Identification

In time series analysis, the most fundamental steps are to identify and build a model based on the available data. Model identification refers to the methodology in identifying the needed transformations. The following useful steps are used to identify a tentative model for a given time series.

(1) Plot the data from entire time series data and select the correct transformations.

(2) Compute and examine the sample ACF and sample PACF of the original series to further confirm a necessary degree of differencing. Some general rules are; If the sample ACF decays very slowly, (the individual ACF may not be large) and the sample PACF cuts off after lag 1, it indicates that differencing is needed.

(3) On the basis of the findings of the ACF and PACF measurements of the appropriate transformed series, the orders of p and q are determined.

3.3.2 Model Selection Criteria

Model identification tools such as ACF and PACF used only for identifying adequate models. Some model selection criteria based on residuals are: to assess the quality of the model fitting Bayesians introduced an information criterion that is called BIC (Bayesian's Information Criteria), the optimal order of the model is chosen by the value of M, which is a function of p and q, so that BIC (M) is imposes a greater penalty for the number of estimated model parameters than does AIC (Akaike's Information Criteria). Use of minimum BIC for model selection appears in a decided model whose number of parameters is less than that chosen under AIC.

3.3.3 Model Diagnostic Checking

After parameter estimation, it has to assess model adequacy by checking whether the model assumptions are satisfied. The basic assumption is that white noise is the $\{\boldsymbol{a_t}\}$.For any estimated model, the residuals $\boldsymbol{\hat{a}_t}$ \`s are estimates of these unobserved white noise $\boldsymbol{a_t}$ \`s.

To check whether the errors are normally distributed, one can construct a histogram of the standardized residuals $\boldsymbol{\hat{a}_t} / \boldsymbol{\hat{\sigma}_t}$ and compare it with the standard normal distribution using the chi-square goodness of fit test or even Tukey\`s simple five-number summary.

$$\boldsymbol{Q = n(n+2)\sum_{k=1}^{k}(n-k)^{-1}\hat{\rho}_k^2} \quad (4)$$

The modified Q statistic originally proposed by Box and Pierce is this test statistic (1970). Under the null hypothesis of model fit, Ansley and Newbold (1979) show that the Q statistics approximately follow the distribution of $\boldsymbol{X^2}$ (k-m) where m is denotes the number of parameters estimated in the model. Ljung and Box (1978).

If the entertaining model is insufficient, a new model can be easily generated based on the results of these residual analyses.

3.3.4 Forecasting

This involved predicting future number of international tourist arrivals using the fitted times series model selected. In this case one years in future monthly forecasts were generated. That is the forecasts from January 2020 to December 2020.

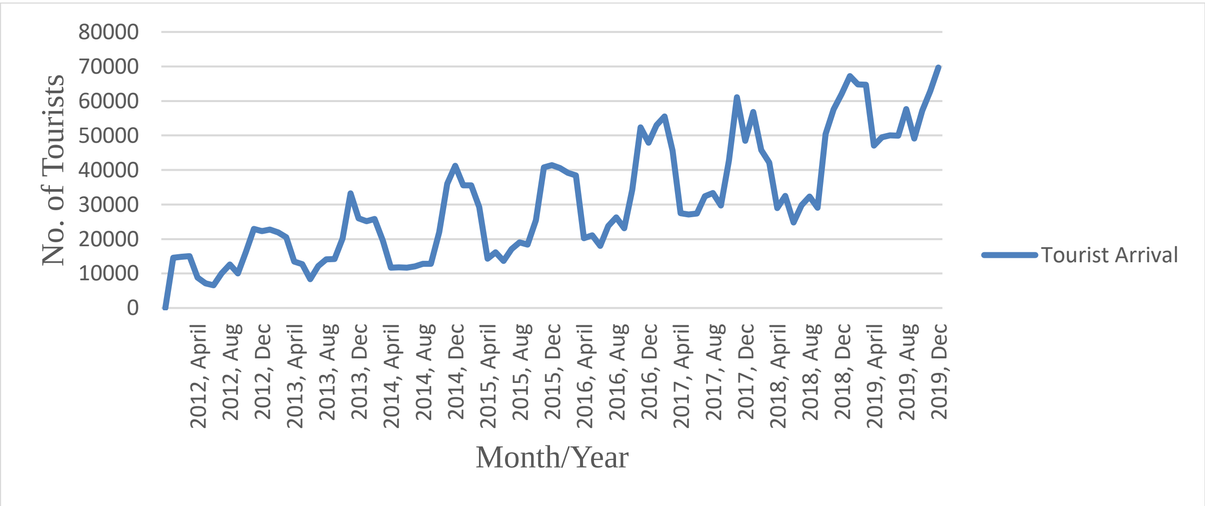
4. Data Analysis of Tourist Arrivals in Mandalay

4.1 Tourist Arrivals in Mandalay

Mandalay is popular among foreign travelers for its cultural heritage.The significant tourist attractions of the region include Bagan Cultural Zone, Golden Palace Monastery, Kuthodaw Pagoda, Mahi Myatmuni Pagoda, Mandalay Hill, Pahtodawgyi Pagoda, Mandalay Palace and Mandalay Fort.

According to the Department of Hotels and Tourism Growth of the Mandalay Region, tourist arrivals have been on the rise in Mandalay since the start of 2016. Tourist arrivals have increased from October 2017 to March 2018. Most of the tourists are American, French, German, Dutch, Japanese, Thai and Chinese. The department expects 3.5 million tourists to come to Myanmar in 2018 and about 500000 of them will travel to Mandalay and the Bagan Cultural Zone. Mandalay international Airport is directly with Thailand, India, and China. Over 483780 tourists entered Mandalay in 2017 on direct flights from those countries.

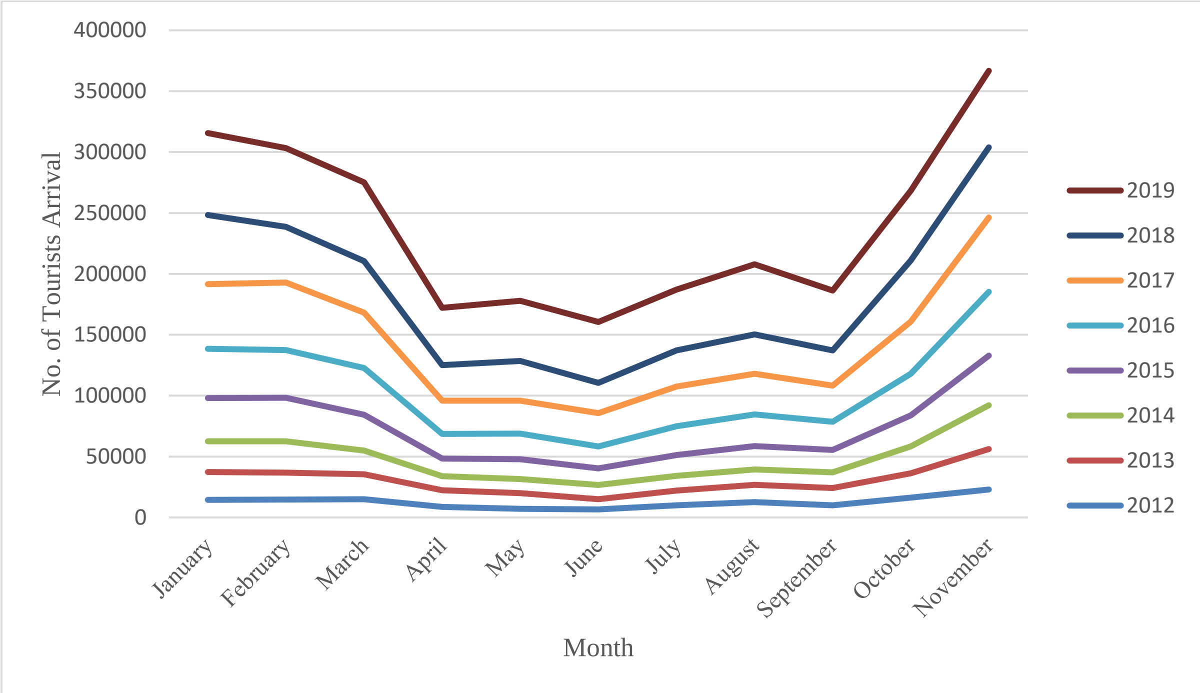
By summarizing the number of tourist visitors (monthly data) arrived in Mandalay during the period 2012 to 2019 are shown in Figure 1. These data are obtained from Myanmar Tourism Statistics, Ministry of Hotels and Tourism (MOHT) in Mandalay. In the Figure 1, it was found that the time series is likely to have seasonal cycle.



Source*:* Ministry of Hotels and Tourism (MOHT)

Figure 1. Monthly Distribution of Tourist Arrivals in Mandalay

Figure (1) shows that the monthly distribution of the number of tourist arrivals in Mandalay during the year 2012 to 2019. Obviously, it shows that, the highest value occurred in November and the lowest value occurred in June in each year. The highest occurred in November, the month with good weather and the month of June is rainy season in Myanmar.



Source: Ministry of Hotels and Tourism (MOHT)

Figure 2 .Yearly Distribution of Tourist Arrivals to Mandalay

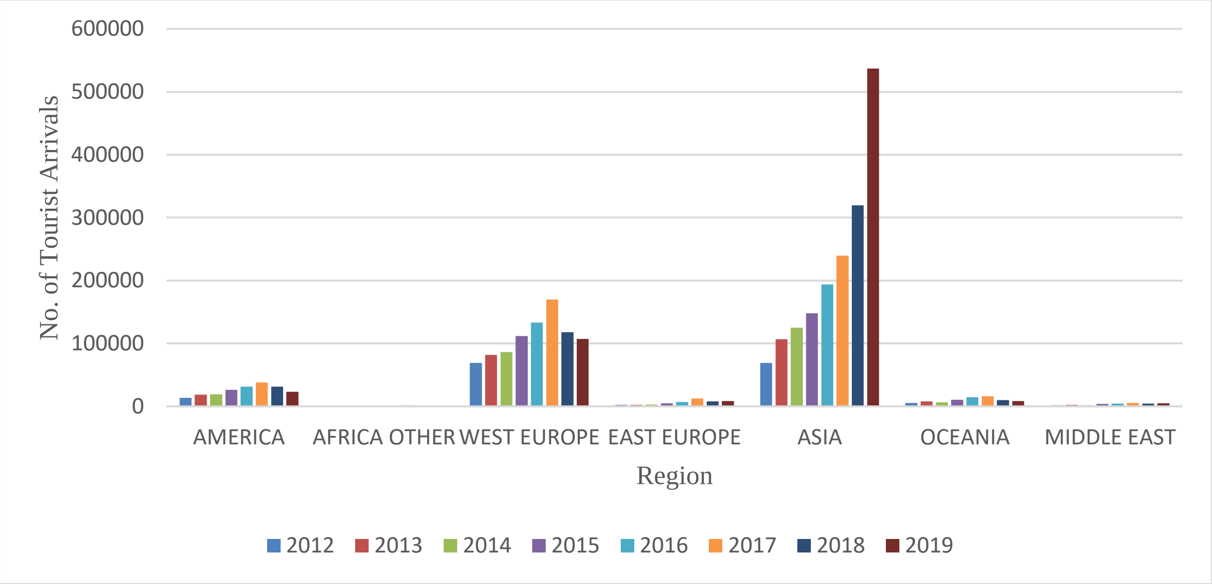
Figure (2) shows that the yearly distribution of the number of tourist arrivals to Mandalay during the year 2012 to 2019. Obviously, the above chart has an upward trend for tourist arrival to Mandalay and there is also seasonal component.

Table 1 . Monthly Tourist Arrivals Series in Mandalay Chaw

Year / Month	2012	2013	2014	2015	2016	2017	2018	2019	%
January	14564	22715	25195	35511	40533	53051	56793	67204	10.59
February	14824	21914	25841	35588	39150	55482	45747	64819	10.17
March	15023	20515	19509	29282	38384	45592	42152	64735	9.23
April	8778	13478	11705	14274	20262	27496	28972	47105	5.77
May	7105	12749	11748	16184	21080	27076	32510	49518	5.97
June	6611	8387	11700	13621	17988	27365	24828	50030	5.38
July	10010	12180	12034	17048	23748	32432	29805	49983	6.28

August	12596	14098	12761	19020	26228	33309	32299	57676	6.97
September	9998	14157	12805	18392	23131	29672	29024	49104	6.25
October	16246	20104	22000	25438	34207	42746	50362	57139	9.00
November	22945	33216	36000	40719	52375	61067	57560	62876	12.31
December	22273	26005	41214	41355	47945	48496	62045	69682	12.05

Source: Ministry of Hotels and Tourism (MOHT)



Source: Ministry of Hotels and Tourism (MOHT)

Figure 3 .Yearly Distribution of Tourist Arrivals to Mandalay by Regions

Figure (3) shows that the yearly distribution of the number of tourist arrivals in Mandalay by regions during the year 2012 to 2019. The above chart show that America, Africa, west Europe, east Europe, Asia, Oceania and Middle East regions came to ancient cities. In 2019, Asia is the highest arrived region and West Europe is the second arrived region.

4.2 Stationarity of the Time series

The upward trend and seasonality as shown in Figure (1) suggests that the time series is not stationary. The ADF test which was used to supplement the graphical methods, had a p-value of 0.05553at lag order 7 which is not less than α=0.05, indicating that the series is non-stationary. Both seasonal and non-seasonal difference of first order were made to transform the series from non-stationary to stationary. The ADF test, after differencing, had the p-value of 0.01 which is less than α=0.05 indicating a stationary series as shown in Table 2.

Table 2. ADF test for Original Series

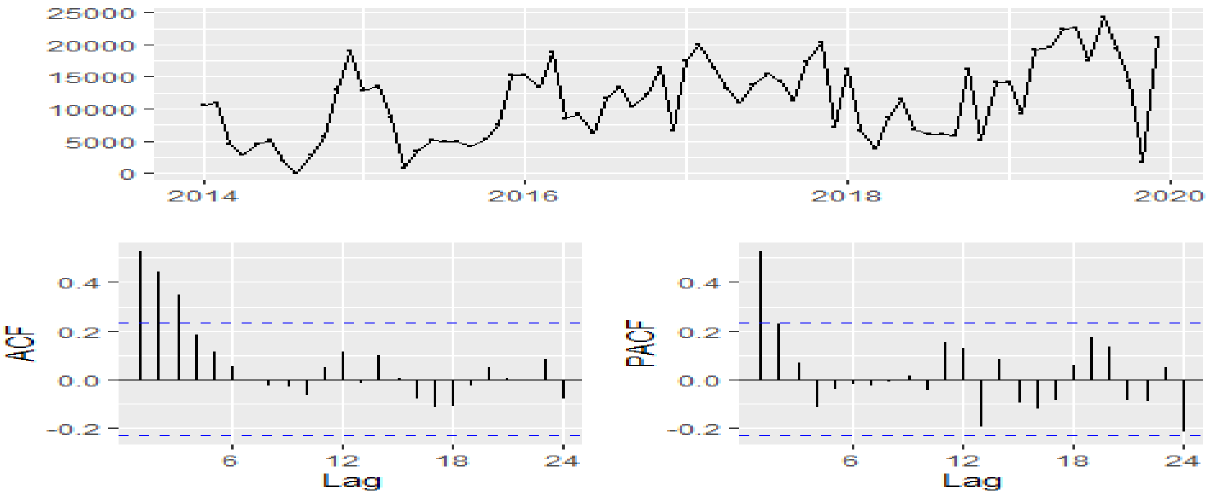
	Original Series	First Differencing
Lag order	7	7

ADF Test Statistics	-3.5054	-7.7745
P- value	0.05553	0.01
Significance level		1%

Source: Ministry of Hotels and Tourism (MOHT), Mandalay

4.2.1 Model Identificationof Best Fit Seasonal ARIMA Model

The plots of ACF and PACF of the stationary time series were examined in order to identify the values of p, q, P and Q. The values of d and D which are the number of non-seasonal and seasonal differencing respectively are both equal to one. That is d=1 and D=1 transformed the time series from non-stationary to stationary.



Source: Ministry of Hotels and Tourism (MOHT), Mandalay

Figure 4. Seasonally Differenced Tourist Arrivals Series

There are enough spikes in the plots outside the insignificant zone (dotted horizontal lines) it can conclude that the residuals are not random. This means that juice or data is available in residuals to be processed by models of AR and MA. A seasonal component is also available in the residuals at lag 12 at lag 12 (represented by spikes at lag 12). This makes sense are analyzing monthly data that tends to have seasonality of 12 months because of patterns in tourist arrived.

In Fig (5) ACF has significant spike at non-seasonal lag 1 suggesting the possible non-seasonal moving average of the first order, AR (1) to be included in the model. In addition, ACF has a large seasonal first seasonal lag spike (at 1.0), indicating the first order seasonal moving average SMA (1).

4.2.2 Model Selection

To find out the best fitted model for international tourist arrival to Bagan, the following SARIMA models are used. The identified model SARIMA $(\mathbf{1},\mathbf{1},\mathbf{0})\times(\mathbf{0},\mathbf{1},\mathbf{1})_{\mathbf{12}}$ is compared with other stationary first order regular differenced and first seasonal differenced models. The model of the least residuals is selected. The estimated values of MAPE and BIC are pointed out in the following Table 3.

Table 3. MAPE and BIC Values for Different SARIMA Models

SARIMA	MAPE	AIC	Normalized BIC
SARIMA $(\mathbf{1},\mathbf{1},\mathbf{0})\times(\mathbf{1},\mathbf{1},\mathbf{0})_{\mathbf{12}}$	9.4707	1636.62	1643.88
SARIMA $(\mathbf{1},\mathbf{1},\mathbf{0})\times(\mathbf{0},\mathbf{1},\mathbf{1})_{\mathbf{12}}$	9.2939	1632.47	1639.73
SARIMA $(\mathbf{0},\mathbf{1},\mathbf{0})\times(\mathbf{1},\mathbf{1},\mathbf{0})_{\mathbf{12}}$	9.5936	1645.17	1650.00

Source*:* Ministry of Hotels and Tourism (MOHT), Mandalay

According to the above table, SARIMA $(\mathbf{1},\mathbf{1},\mathbf{0})\times(\mathbf{0},\mathbf{1},\mathbf{1})_{\mathbf{12}}$ is ignored because of the principle of parsimony, which says that given any two models, the model with the least number ofparameters is more preferred. Comparing the rest of the models based on AIC and BIC, SARIMA $(\mathbf{1},\mathbf{1},\mathbf{0})\times(\mathbf{0},\mathbf{1},\mathbf{1})_{\mathbf{12}}$ has the least error value. Hence, the most preferred model is SARIMA $(\mathbf{1},\mathbf{1},\mathbf{0})\times(\mathbf{0},\mathbf{1},\mathbf{1})_{\mathbf{12}}$.

4.2.3 Model Estimation

The SARIMA (1,1,0) × $(0,1,1)_{12}$ model was selected as the best model amongst the candidate SARIMA models. The selected model also performs best in term of forecasting with the lowest RMSE and MAPE compared to other candidate SARIMA models as shown in Table 3. The coefficients of the best model estimated by using the method of Maximum Likelihood (ML) and all the coefficients are statistically significant at 5% as shown in Table (4).

Table 4. Coefficients and Standard Error (SE) of the Best Model

Best Fitted model: ARIMA (1,1,0) (0,1,1) [12]		
	AR1	SMA1
Coefficient	-0.3489	-0.5408
S.E	0.1045	0.1497
t-value	-3.3387	-3.6126
Log Likelihood =-813.23		
AIC = 1632.47		

Source*:* Ministry of Hotels and Tourism (MOHT), Mandalay

The equation of SARIMA $(\mathbf{1},\mathbf{1},\mathbf{0})\times(\mathbf{0},\mathbf{1},\mathbf{1})_{\mathbf{12}}$model is

$$(\mathbf{1}-\emptyset\mathbf{B})\mathbf{Z_t}=(\mathbf{1}-\mathbf{B^{12}})(\mathbf{1}-\mathbf{\Theta_1 B^{12}})\mathbf{a_t}$$

The estimated model is

$$(\mathbf{1}+\mathbf{0.3489B})\mathbf{Z_t}=(\mathbf{1}-\mathbf{B^{12}})(\mathbf{1}+\mathbf{0.5408B^{12}})\mathbf{a_t}$$

$Z_t$is the number of international tourist arrivals in month t

a_t is the white noise

B*i*s the backshift operator such that:

$B^j Z_t = Z_{t-}$, j=0, 1, 2, …………..

After identifying the best-fit SARIMA model for international tourist arrivals in Bagan, the best- fit model is estimated based on the OLS estimation procedure. The results are given in table 4, providing that of the moving average and seasonal moving average order one, AR (1) term $\emptyset_{\mathbf{12},\mathbf{1}}$is estimated to be (SE = 0.1045) and and SMA (1) term $\mathbf{\Theta_{12,1}}$ is estimated to be (SE = 0.1497) is statistically significant at least < 0.01 significance level. The estimated parameter is less than one, supporting the required stationarity and invertibility conditions.

4.2.4 Diagnostic Checking

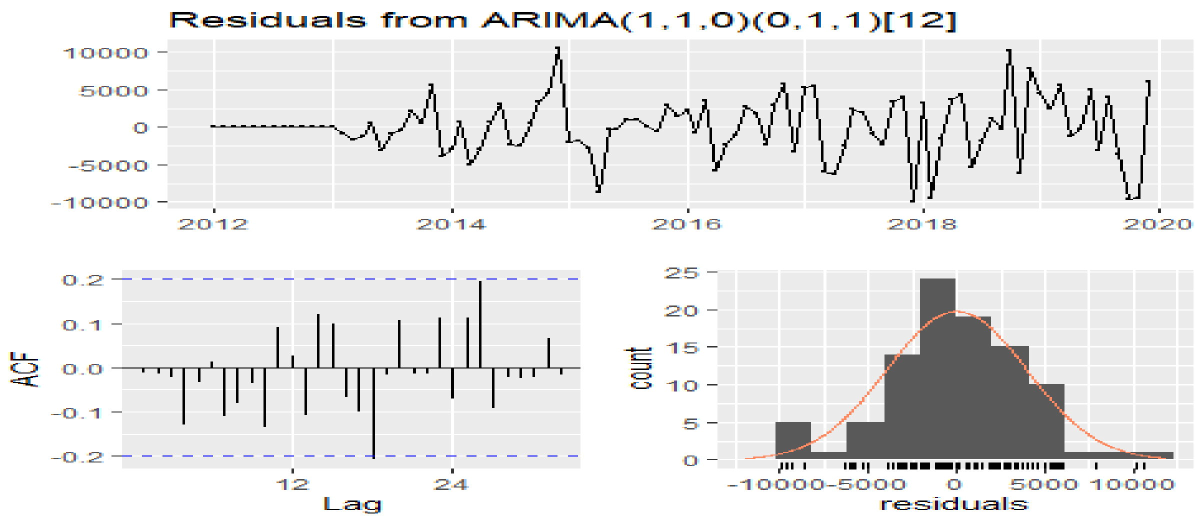
This involved checking whether the fitted model has adequately captured the information in the data. Residual analysis which involves graphical procedures and the Ljung-Box statistical test were used as shown in Table(5) and Figure(6) respectively.

Table 5. Ljung-Box test

Ljung-Box Test
data: Residuals from ARIMA (1,1,0) (0,1,1) [12]
Q* =22.049, df = 22, p-value = 0.457
Model df: 2. Total lags used: 24

Source: Ministry of Hotels and Tourism (MOHT), Mandalay

The residuals from the fitted model shown in Figure 6 seem to be random as they have nearly constant variance and zero mean implying that the fitted model is adequate. The Ljung-Box statistical value, Q*=22.049 is not significant because the p-value shown in Table 5 is which is greater than 0.05 level of significance indicating that the residuals of the best model are not statistically significantly distinguishable from white noise. That means the model SARIMA (1, 1, 0) × (0, 1, 1)$_{12}$ is adequate.

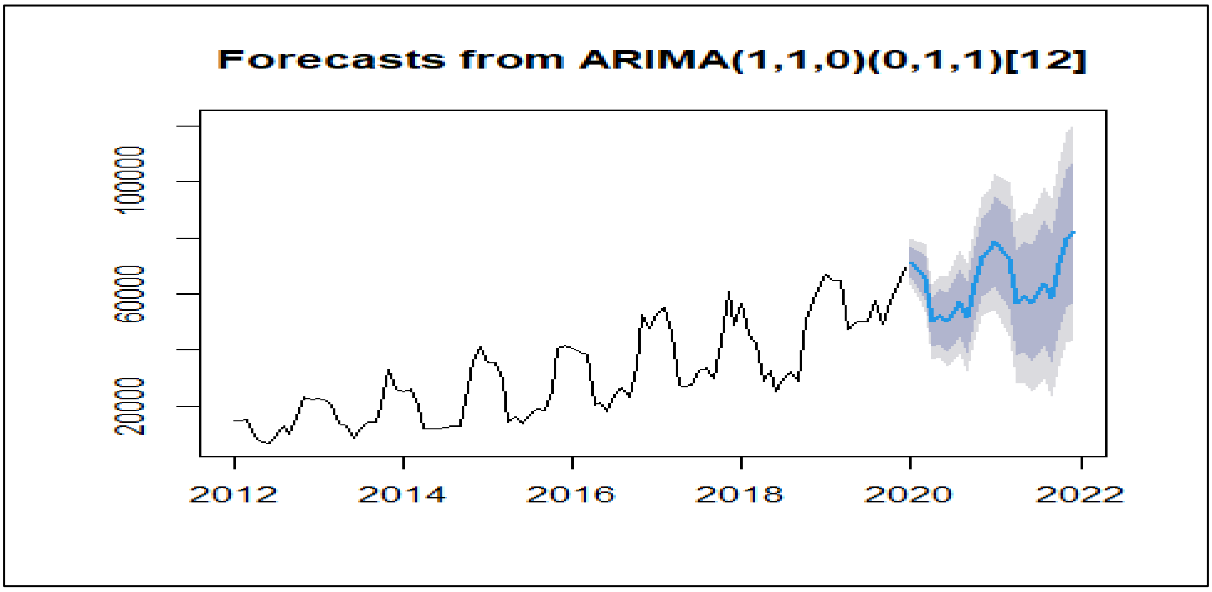


Source: Ministry of Hotels and Tourism (MOHT), Mandalay

Figure 5. Residual Analysis from Model ARIMA (1,1,0) (0,1,1) [12]

4.2.5 Forecasting with SARIMA $(1,1,0)\times(0,1,1)_{12}$ Model for Tourist Arrived Series

Figure7 shows the forecast values of tourist arrived of ancient cities of Mandalay for 24 months period from January to December in 2020 and illustrates in Figure 7. The total number of tourists will increase about 51384 in December 2021 compared in December 2019. According to these forecasts, Myanmar's ancient cities of tourist flows will increase in the future; this result indicates that the model provides an acceptable fit to predict the number of tourists flows.



Source: Ministry of Hotels and Tourism (MOHT), Mandalay

Figure 6. Forecast from SARIMA (1,1,0) (0,1,1) [12]

Table 6 .Forecasting with SARIMA $(\mathbf{1},\mathbf{1},\mathbf{0}) \times (\mathbf{0},\mathbf{1},\mathbf{1})_{\mathbf{12}}$ Model

		95% Confidence Intervals	
Month	Forecast	Lower Limit	Upper Limit
Jan 2021	78866.65	54466.51	103266.78
Feb 2021	75731.25	49806.39	101656.12
Mar 2021	72843.93	45324.82	100363.04
Apr 2021	56849.01	27874.93	85823.10
May 2021	58870.67	28494.09	89247.26
Jun 2021	56851.83	25140.45	88563.21
Jul 2021	59392.50	26398.35	92386.64
Aug 2021	63989.76	29761.54	98217.98
Sep 2021	58476.04	23056.50	93895.58
Oct 2021	70656.93	34084.93	107228.94
Nov 2021	80026.27	42337.00	117715.54
Dec 2021	82292.15	43517.82	121066.48

Source: Ministry of Hotels and Tourism (MOHT), Mandalay

5. Conclusion

Tourism is currently one of the important sectors for economic growth in Mandalay. The results show that the SARIMA (1, 1, 0) × (0, 1, 1)$_{12}$ is the best fitted model and the model was used to generate monthly forecasts from January 2021 to December 2021 with 95% confidence interval. The forecasts indicate that the number of tourists visiting Mandalay is likely to keep on increasing with seasonal pattern similar to that of the original data. The government has taken steps to improve infrastructure, more needs to be done to upgrade airports, roads and public transport to ease domestic travel and extend the length of stay for many tourists.

International tourist arrivals are providing vital connectivity on a national and regional developing of resident in Mandalay. Arrival numbers vary greatly in high and low season, but tourism operators have begun to offer green season promotions to attract tourists throughout the year, and travel options have diversified to cater to different needs of tourists. Therefore, to predict future number of international tourists visiting the Mandalay is important for tourism planning and marketing.

Acknowledgements

I would specially like to express my respect to Dr. Moe Moe Yee, Rector of University of Co-operative and Management, Sagaing for her permission and suggestions to write this paper. I would like to profoundly thank Visiting Professor Daw Khin Aye Myint for supporting, guidance and invaluable advice this paper. I specially thanks to Daw Khin San Kyi, Professor, and Head of Department of Statistics for her encouragement. I am also very grateful to all my colleagues at the Department of Statistics for their kind cooperation and to all teachers.

References

Asbollah, A.Z., Hassan,N.,&Idris,H. (2017). THE TOURIST BEHAVIOUR IN DIFFERENT ENVIRONMENTS: A LITERATURE REVIEW. *PLANNING MALAYSIA JOURAL,15(1).*

Borhan, N., & Arsad, Z. (2014). *Forecasting international tourism demand from the US, Japan and South Korea to Malaysia: A SARIMA approach*. 955–960. https://doi.org/10.1063/1.4887719

Brida, J. G., & Garrido, N. (2011). Tourism forecasting using SARIMA models in Chilean regions. *International Journal of Leisure and Tourism Marketing*, *2*(2), 176. https://doi.org/10.1504/IJLTM.2011.038888

G. E. Box & G. M. Jenkins, Time series analysis: forecasting and control, revised ed. Holden-Day, (1976), https://doi.org/10.1177/058310248201400608

Makoni, T., & Chikobvu, D. (2018). Modelling and Forecasting Zimbabwe's Tourist Arrivals Using Time Series Method: A Case Study of Victoria Falls Rainforest. *Southern African Business Review*, *22*. https://doi.org/10.25159/1998-8125/3791

Msofe, Z. A., & Mbago, M. C. (2019). Forecasting international tourist arrivals in zanzibar using box – jenkins SARIMA model. *General Letters in Mathematics*, *7*(2). https://doi.org/10.31559/glm2019.7.2.6

Msofe, Z. A., & Mbago, M. C. (2019). Forecasting international tourist arrivals in zanzibar using box – jenkins SARIMA model. *General Letters in Mathematics*, *7*(2). https://doi.org/10.31559/glm2019.7.2.6

Winters, M., Davidson, G., Kao, D., & Teschke, K. (2011). Motivators and deterrents of bicycling: comparing influences on decisions to ride. *Transportation, 38*(1), 153– 168. http://dx.doi.org/10.1007/s11116-010-9284-y

World Travel and Tourism Council. Travel and tourism economic impact 2018 Myanmar.

Zahedjahromi, M. (2018). Forecasting Tourist Arrivals to the USA With SARIMA Models. *University of Northern Colorado*.