Economic and Technology Management Review, Vol. 19(2022): 55 – 65

Statistical modelling and forecasting of tourist arrivals in MARDI Langkawi Agro Technology Park using SARIMA

(Pemodelan statistik dan ramalan kedatangan pelancong di Taman Agroteknologi MARDI Langkawi menggunakan model SARIMA)

Aimi Athirah Ahmad[1]*, Nik Rozana Nik Mohd Masdek[1] and Zawiyah Pono[1]

Keywords: trend analysis, statistical modelling, forecasting, tourist, SARIMA

Abstract

The tourism industry recorded a growth rate of 11.2% (RM82.6 billion) to the country's gross domestic product (GDP) in 2017 compared to RM76.6 billion in 2016. Malaysia's diverse agricultural resources are seen as a critical factor in accelerating growth in the agro-tourism industry. Due to that, MARDI has established several agro-tourism centres in Langkawi, Cameron Highlands and Kuala Kangsar. The most significant contributor is MARDI Langkawi Agro Technology Park (TATML). This research aims to forecast the eco-tourism demand based on the number of tourist arrival at TATML. Time series forecasting (SARIMA) was conducted on tourist arrival in TATML. The forecasting models are compared and assessed using Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). This study shows a higher number of foreign visitors in July and August, whereas hike in local visitors are observed in November and December, being the school holiday season. July to August are holidays for regions like Europe and the Middle East. Time series forecasting analysis estimates the number of foreign visitors to TATML in 2019 using SARIMA (9,0,0) x $(0,1,0)_{12}$ is 69,170. The findings of this study are applicable in other agro-tourism centres. In addition, promotion and management strategies must be focused on the season and peak times to facilitate travel and attract foreign visitors.

Introduction

Malaysia's tourism growth has reached a point of excellence. The government has provided various facilities to enhance the tourism industry's performance. As a result, tourism has become profitable and benefitted the country's economy. Besides, the tourism sector can also increase domestic income through foreign currency exchange (Kumar et al. 2015). The tourism sector has various contributions that can increase the country's economic growth (Abdul Latiff et al. 2020).

Agro-tourism is a sub-sector in the tourism industry that includes tourism activities related to the agriculture sector based on crops, livestock, fisheries and food-based industry. The country's focus on developing agricultural activities has opened a new chapter for tourism, especially agro-tourism activities. Agro-tourism has provided employment opportunities and increased the income of farmers and rural residents. Agro-tourism is the concept of

[1]Socio-Economy, Market Intelligence and Agribusiness Research Centre, MARDI Headquarters, Persiaran MARDI-UPM, 43400 Serdang, Selangor, Malaysia
E-mail: aimiathirah@mardi.gov.my

visiting agricultural or livestock areas that offer opportunities for tourists to participate in activities carried out by agropreneurs. Furthermore, some agro-tourism centres provide accommodation facilities, giving visitors more opportunities to explore the whole area. Among the exciting activities often done is visiting cattle farms, where visitors will be given information on how to raise cattle (MOA 2011). The increase in domestic and foreign tourist arrivals indicates a healthy economy for Malaysia. Based on the view of the monetary economy, an increase in domestic tourists ensures fewer currency outflows, and an increase in international tourists will increase foreign currency inflow into the country (Mansor and Ishak 2015).

In this study, the analysis of tourist arrivals focuses on foreign tourists who visited MARDI Langkawi Agro Technology Park (TATML), one of the frequently visited agro-tourism destination in Langkawi Island. The location of TATML is very suitable because Langkawi Island provides an incredible way out for business and leisure. Based on tourist arrival statistics in Langkawi, the data shows an increasing trend from 2005 till now, where tourist arrivals in 2005 were recorded at 1.84 million and this number increased to 3.57 million in 2014, an increase by almost 94% (LADA 2015). As the trend grows, it is vital to accurately predict the number of tourist arrivals as it will benefit the direct and indirect activities related to the tourism industry. Thus, governments or related organisations and agencies can use forecast figures to drive social, economic and physical development, to create development scenarios such as the conservation of natural resources, to develop a conducive environment and also to create attractive opportunities for foreign investors (Mahendran 2008); (Marzuki 2011).

Therefore, motivated by the previous research, this study aims to analyse a time series modelling of foreign tourist arrivals to TATML (January 2016 to March 2019) using the SARIMA model and to estimate foreign tourist arrivals from April to December 2019. In order to determine the appropriate model, some models (candidate models) from the SARIMA group were selected based on the Akaike Information Criteria (AIC) method. First, the models were selected based on the smallest AIC values. Then, to see the model's accuracy, the value of the Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) were calculated by comparing the actual and predicted data.

Literature review

Previous research on tourist arrival forecast to Malaysia used various forecasting techniques such as Autoregressive Integrated Moving Average (ARIMA) forecasting method, naive method, the exponential smoothing method, Holt and Holts-Winter's linear method, as well as trend and seasonal analysis considered (Chuah 2011). The study found that the Moving Average 10 or MA (10) model is the best forecast model because MA can produce the smallest root mean square error (RMSE). In subsequent research, Shitan and Pauline (2003a) compared and investigated the performance of the Autoregressive (ARAR) and Autoregressive Fractionally Integrated Moving Average (ARFIMA) models with the previously proposed MA (10) model. Another study by (Shitan and Pauline 2003b) also made a comparison of the ARFIMA model and the Autoregressive Moving Average (ARMA) model and suggested that the ARFIMA model (7, 0.42, 10) is more accurate than ARMA. In this case, some unforeseen circumstances, such as natural disasters, disease outbreaks or pandemics, need to be taken into account as they can affect tourist arrivals immediately (Shitan and Fei 2004).

However, in the last 10 years, Seasonal ARIMA or SARIMA has been more widely used in predicting tourist arrival data because many studies found that tourist arrival data has a seasonal trend (Yu-wei and Meng-yuan, 2010); (Loganathan and Ibrahim 2010); (Loganathan et al. 2012); (Rusyana et al. 2016). The SARIMA model is an extension of the ARIMA model, which consists of seasonal and non-seasonal data. Chang and Liao (2010) used this model to forecast the monthly outbound tourist departures from Taiwan to the three major destinations of Hong Kong, Japan and the USA, respectively. The Hylleberg Engle-Granger-Yoo (HEGY) test is applied to determine the seasonality of the data. Meanwhile, the Mean Absolute Percentage Error (MAPE) is adopted to measure the accuracy of the prediction. The result shows that the proposed models were SARIMA (0,1,1) $(1,0,1)_{12}$ for Hong Kong, SARIMA (0,1,1) $(1,0,0)_{12}$ for Japan, SARIMA (1,1,0) $(0,0,1)_{12}$ for the USA, and SARIMA (1,1,1) $(1,0,0)_{12}$ for total outbound tourism in Taiwan.

Current research also has been using SARIMA in predicting tourist arrival. For example, Noratikah Abu et al. (2021) forecasted eco-tourism demand in National Park Kuala Tahan, Pahang using SARIMA (1,0,0) $(1,0,1)_{12}$. In this study, the SARIMA model was chosen based on smaller MAPE, MAE and RMSE compared to Exponential Smoothing. Besides that, the SARIMA model was also used by Nurhasanah et al. (2022) to predict the arrival of foreign tourists in Indonesia. The performance of the model was also evaluated using MAPE value. This study concluded that SARIMA (0,1,0) $(1,10)_{12}$ has a statistically significant model coefficient and successfully predicted that the arrival of foreign tourists increased in each period. This model is also commonly used and applicable in various situations, shown in *Table 1*.

Methodology

Data collection

MARDI Langkawi Agro Technology Park (TATML) (6.3612 ° N, 99.7927 ° E) is located at Langkawi Island, Kedah (*Figure 1*). Data on local and foreign tourist

Table 1. Summary of SARIMA model application in other areas

No.	Objectives	Year	SARIMA model	Author
1.	To model the monthly passenger flow on Serbian Railways	January 2004 – June 2014	SARIMA (0,1,0) $(0,1,1)_{12}$	(Milenkovic et al. 2015)
2.	To predict the unemployment rate of Greece.	April 1998 – September 2015.	SARIMA (0,2,1) $(1,2,1)_{12}$	(Dritsaki C. 2016).
3.	To forecast the monthly maximum and minimum temperatures in Bangladesh	January 2000 – December 2017	Minimum temperature = SARIMA (2,0,1) $(2,1,0)_{12}$; Maxiumum temperature = SARIMA (1,0,0) $(1,1,0)_{12}$	(Doulah M. 2018).
4.	To develop a forecasting model for hospitalised and non-admitted patients.	January 2009 – October 2013	Hospitalised patients = SARIMA (2,0,3) $(1,0,0)_{52}$; Non-admitted patients = SARIMA (0,1,0) $(0,0,2)_{52}$	(Song et al. 2015)

arrivals and monthly ticket details were obtained from MARDI Langkawi. MARDI Langkawi Agrotechnology Park is one of the famous tourist attraction in Langkawi. It often receives visitors from within and outside the country. The 27.39 ha park is open to public and provides exposure to modern technology in the agricultural industry of Malaysia. It is famous for its fruit farms comprising various types with commercial properties (LADA 2020).

Unit root test

Many time series data in economics and finance are usually non-stationary. Therefore, a crucial econometric task is determining the data's most appropriate trend shape. In ARMA modelling, data must be converted to stationary before being analysed. If the data has a trend, then the data needs to undergo several differencing methods.

Two standard trend removal (de-trending) procedures are first-order differencing and time trend regression. The unit root test can be used to determine whether transformed data is stationary or non-stationary. The Augmented Dickey-Fuller (ADF) root unit test was used in this study. The ADF test is the most popular test for determining root units (Dickey and Fuller 1979). The null hypothesis of ADF tests is that the series is non-stationary or has a unit. The ADF can be explained by using equation 1. For the data with a trend and unit root, differencing methods must be carried out to make the data stationary.

$$\Delta y_t = \alpha + \beta t + (p - 1)\, y_{(t-1)} + \delta_1 \Delta y_{t-1} + \ldots + \delta_{p-1} \Delta y_{t-p-1} + \varepsilon_t \quad (1)$$

where

α = a constant term

βt = the coefficient of a simple time trend

p = parameter of interest

Δ = the first difference operator parameters

p = the lag order of the autoregressive process

Source: Google maps (2022)

Figure 1. Location *of Taman Agroteknologi MARDI Langkawi*

SARIMA model

The Auto-Regressive Integrated Moving Average Model (ARIMA) is a frequently used forecasting method for univariate time series data (Brockwell and Davis 2002). As the name suggests, it supports both autoregressive and moving average elements. Integrated element refers to the method of differentiation (differencing) that may support time series data with a trend. Among the main problems of using the ARIMA model is that this model does not support seasonal data (has a recurring cycle) and data with a seasonal effect. However, research suggests that international tourist arrivals exhibit seasonal patterns (Baldigara and Mamula 2015; Chang and Liao 2010).

For this reason, when modelling international tourist arrivals, the SARIMA model has been used to yield an appropriate model to forecast tourist arrivals. Thus, ARIMA has been modified to address this problem, including identification, estimation, and diagnostic examination (Box and Jenkins 1976). The components in the ARIMA model are as follows (*p*: Trend autoregression order, *d*: Trend difference order, *q*: Trend moving average order).

Then, the SARIMA model can be explained in equation 2

$$\Phi_P (B^s)\ \phi\ (\mathrm{B})\ \nabla^d \nabla_S^D x_t = \Theta_Q (B^s)\ \theta\ (B) \varepsilon_t \ (2)$$

In the above equation;
B = the backward shift operator,
ε_t = the residual at time t, the mean of ε_t is zero and the variance of ε_t is constant,
x_t = observed value at time t ($t = 1, 2 \ldots k$),
$\emptyset$(B) and θ (B) = The ordinary autoregressive and moving average components of orders *p* and *q*,
$\Phi_P(B^s)$ and $\Theta_Q(B^s)$ = The seasonal autoregressive and moving average components of orders P and Q,
$\nabla^d \nabla_S^D$ = The ordinary and seasonal difference components.

Therefore if the SARIMA model is used, the components are similar to ARIMA with additional *P*, *Q*, *D* and *m* components as shown below:

p : is the parameter for the non-seasonal autoregressive order AR(p) component.
q : is the parameter for the non-seasonal moving average order MA(q) component.
d : non-seasonal difference order.
P : is the parameter for the seasonal autoregressive order AR(P) component
Q : is the parameter for the seasonal moving average order MA(Q) component
D : seasonal difference order
m : The number of time steps for a single seasonal period.

and written as SARIMA (*p, d, q*) (*P, D, Q*) *m* (Pankratz 1983).

For example, if the seasonal period of the series $s = 12$, then equation 2 can be rewritten as:

$$\Phi_P (B^{12})\ \phi\ (\mathrm{B})\ \nabla^d \nabla_{12}^D x_t = \Theta_Q (B^{12})\ \theta\ (B) \varepsilon_t \ (3)$$

Similarly, $d = 1$ will calculate the first-order seasonal differencing while $p = 1$ and $q = 1$ indicate the seasonal autoregressive and seasonal moving average order 1, respectively.

It may be helpful to observe the pattern of the autocorrelation function (ACF) and partial autocorrelation function (PACF) as shown in *Table 2* when determining the values of p and q, and P and Q (Nurhasanah et al. 2022).

Model estimation

To determine if the forecasting model is good, it is necessary to evaluate it. Model evaluation can be calculated from the resulting error. For example, the best model could be determined by using criteria like the Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and the Akaike Information Criterion (AIC).

Forecasting accuracy

In comparing the model efficiencies of several selected forecasting models, several criteria, such as Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE), need to be evaluated. The smaller the values of the three criteria, the better the forecasting model used (Ahmad et al. 2018). All analyses performed in this study used the R Studio software.

The following equation describes the criteria used:

$$MAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_t|}{n} \tag{4}$$

$$MAPE = \frac{\sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_t}{y_i} \right|}{n} \; x \; 100\% \tag{5}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_t)^2}{n}} \tag{6}$$

Results and discussion

Based on *Figure 2*, in 2016, the number of visitors and ticket sales revenue was 100,014 people and RM2,177,916. There was a slight decrease in visitors in 2018, which recorded 96,463 visitors, but ticket sales increased by RM30,084. Despite the reduction, the number of foreign tourists is higher than that of local visitors, which can positively affect ticket prices for foreign tourists as the prices are higher than that of local visitors. *Figure 2* shows that the density of foreign visitors was in July and August, while

Table 2. Order identification of ACF and PACF plots

Order	Description	Patterns in the ACF and PACF plots
AR(p)	An autoregressive model of order p.	On the PACF plot, there is a significant spike in lag p (if the line slowly decays to 0 exponentially, then AR(0).
MA(q)	A moving average model of order q.	On the ACF plot, there is a significant spike in lag q.
AR(P)	A seasonal autoregressive model of order P.	On the PACF plot (with m being a seasonal period), there is a significant spike in lag p.
MA(Q)	A seasonal moving average model of order Q.	On the ACF plot (with m being a seasonal period), there is a significant spike in lag q.

Source: Nurhasanah et al. 2022

Malaysians are in November and December, which is the school holiday season. July to August is a holiday for most regions, such as Europe and the Middle East.

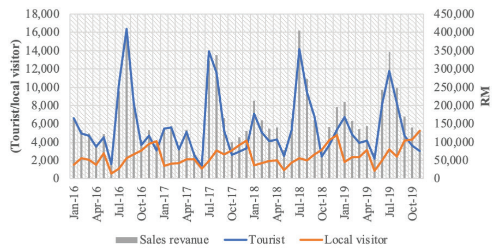
Table 3 summarises the ADF (Augmented Dickey-Fuller) test on foreign tourist arrival data to TATML. The H_0 tested was that the investigated variable had a unit root and was non-stationary, while H_1 had no unit root, and the data were stationary. Therefore, in *Table 2*, the ADF test rejected H_0 and concluded that the data is stationary.

Although the data is stationary, the data shows a seasonal component as in *Figure 2*. In order to ensure the data has a seasonal pattern, there are repeated fluctuations in a certain period by looking at the ACF and PACF plots in *Figure 3*.

It is indicated that all ACF and PACF values should extend within the confidence limits in stationary series (Singh et al. 2011). If ACF values were high at specific lags for this series, these values were determined as making sudden peaks especially at periodic lags of 12 months (*Figure 3*). This demonstrates that the series has a seasonal structure. In order to provide the stationary of the data, seasonal differences should be taken. *Figure 3* also suggested that it is appropriate to consider an order m = 12 fitting the model to the data. For the PACF plot, there is a significant spike at lag p = 1,2,7 and 9. Thus, the non-seasonal model of AR(1), AR(2), AR(7) and AR(9) can be considered.

Since the data has a seasonal component, it takes a seasonal difference at m = 12. The ACF and PACF plot after the differencing at m = 12 are further examined to check the stationarity of the data. Based on *Figure 4*, the ACF and PACF plot seems stationary. However, the ADF test indicates no evidence to reject H_0. Thus, it shows that the unit root is present and the data is non-stationary (*Table 4*). Therefore, the additional differencing at lag 1 needs to be applied. After additional differencing at lag 1, the ADF test results show that the data is stationary (*Table 5*). Thus the seasonal component of (P = 0, D = 1, Q = 0) and m = 12 will be included in the model.

Having found that the data are stationary, examination of the autocorrelation function plot (ACF) and



Source: TATML (2019)

Figure 2. *The number of visitors and sales ticket revenue in TATML (Jan 2016 – Nov 2019).*

Table 3. Unit root test for tourist arrival data to TATML

	Level: $I(0)$	p-value
ADF test (τ)	–4.4129	<0.01***

*** significance at .01

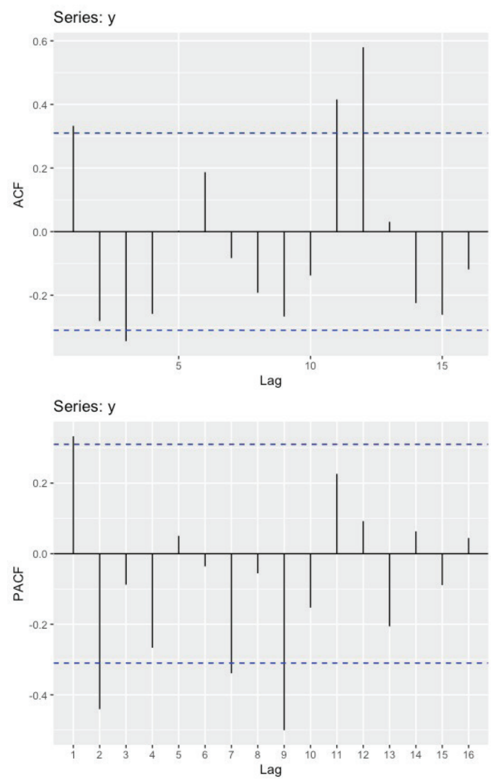


Figure 3. *ACF and PACF plot of the data*

partial autocorrelation function plot (PACF) in *Figure 3* showed that the p = 1,2,7 and 9 non-seasonal autoregressive (AR) components should be included in the ARIMA model. Therefore, foreign tourist arrival data to TATML is more suitable with the ARIMA combination (p, d, q) to obtain a good SARIMA model.

Standard procedures for more accurate identification, estimation, diagnostic inspection and installation in Box-Jenkins analysis of time series are performed. The parameter estimation method is performed using the maximum likelihood method. The determination of a good SARIMA model is based on AIC. The smaller the AIC value, the better the data is modelled using SARIMA. After conducting a modelling series, this study found that two models will be used to estimate the number of tourists to TATML as shown in *Table 5*.

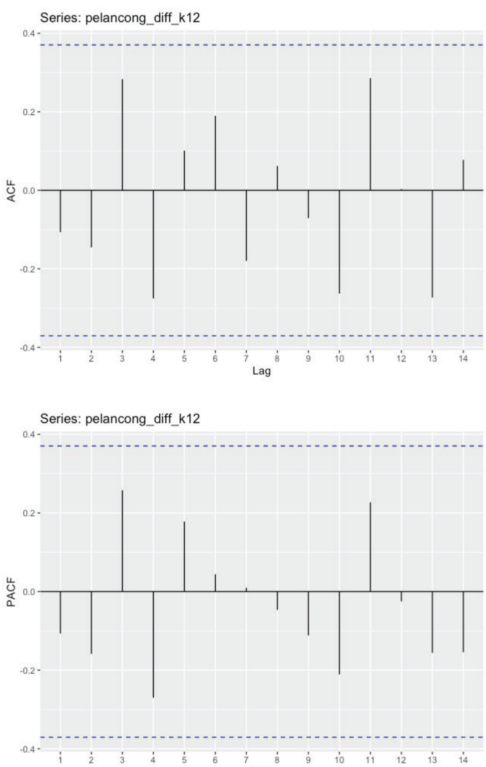


Figure 4. ACF and PACF plot of the data after differencing at m = 12

Table 6 shows the AIC values, and the decision to select the most suitable model is made by comparing the AIC values. The smaller the value, the better the suitability of the SARIMA model used. Furthermore, based on Lewis's (1982) definition, the forecasting performance of these models is very accurate because the MAPE of all models is smaller than 10%. Therefore, it can be concluded that these models have good forecasting characteristics. However, the best model was selected in this study

Table 4. Unit root test for tourist arrival data to TATML after differencing at lag 12

	Level: $I(0)$	p-value
ADF test (τ)	–3.4509	0.070

Table 5. Unit root test for tourist arrival data to TATML after additional differencing at lag 1

	Level: $I(0)$	p-value
ADF test (τ)	–3.6713	0.045**

** significance at .05

based on the smaller RMSE . Thus, the time series forecasting analysis estimates the number of tourist to TATML in 2019 using SARIMA (9,0,0) x $(0,1,0)_{12}$ is 69,170 and shows the same trend where the peaks are in July and August as in *Figure 5*. *Table 7* shows the predicted tourist arrival results in TATML from May 2019 to December 2019.

Conclusion

Data on arrival of foreign tourists to TATML can be modelled using SARIMA (9,0,0) x $(0,1,0)_{12}$. This model's projection show high accuracy with a small RMSE value of 1972.861. The findings of this study can be applied to other agro-tourism centres. In addition, promotional and management strategies should be focused on peak seasons and times to facilitate travel and attract more overseas visitors.

Table 6. Summary of forecasting accuracy method

Model SARIMA	AIC	MAE	MAPE	RMSE
$(1,0,0) \times (0,1,0)_{12}$	765.94	2230.016	50.204%	3225.758
$(2,0,0) \times (0,1,0)_{12}$	759.61	1989.756	47.085%	2892.025
$(7,0,0) \times (0,1,0)_{12}$	761.39	1730.066	44.825%	2570.356
$(9,0,0) \times (0,1,0)_{12}$	748.85	1554.313	39.942%	1972.861

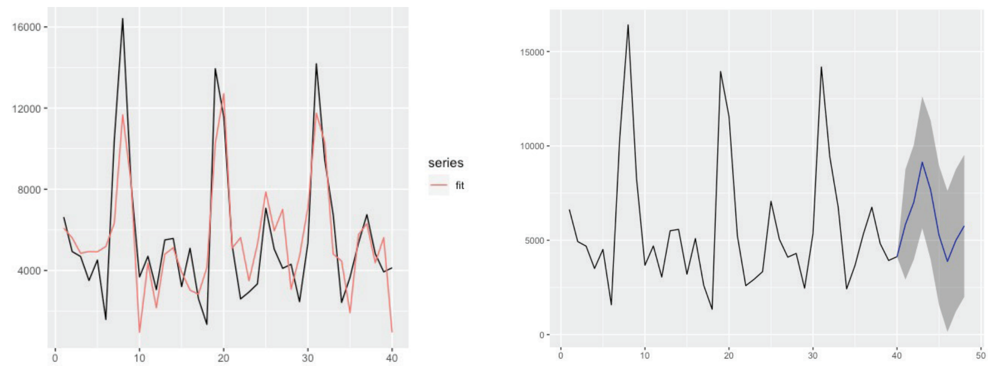


Figure 5. Actual and Forecasted plot of tourist arrival in TATML based on SARIMA $(9,0,0) \times (0,1,0)_{12}$

Table 7. Forecasted value of tourist arrival in TATML based on SARIMA $(9,0,0) \times (0,1,0)_{12}$

Period	Forecasted value	Lower confidence limit	Upper confidence limit
May 2019	5834	1369.54	10299.39
June 2019	7019	2369.36	11667.72
July 2019	9133	3794.05	14472.32
August 2019	7664	2036.70	13291.12
September 2019	5260	–386.62	10905.64
October 2019	3879	–1832.64	9590.71
November 2019	4988	–768.04	10744.27
December 2019	5761	1.04	11520.3

References

Abdul Latiff, A., Mohd, S., Allah Pitchay, A. and Daud, L. (2020). Kesan Pertumbuhan Industri Pelancongan Kepada Pendapatan Negara Dan Peluang Pekerjaan: Analisis Input-Output. *International Journal of Social Science Research*, 2(4), 44 – 61

Ahmad, A.A., Jamaluddin, J.A., Yusof, F., Safari, S. and Yusof, R.M. (2018). Maximising the benefit of domestic and export markets scenario: Predicting models for durian production. *Economic Management and Technology Management Review*. Vol.14. 19 – 31

Box,G.E. and Jenkins,G.M. (1976). *Time Series Analysis: Forecasting and Control San Francisco*. Calif: Holden-Day

Brockwell, P.J. and Davis, R.A. (2002). *Introduction to Time Series and Forecasting, 2nd Edition*, Springer-Verlag, New York

Chuah, S.C. (2001). Forecasting Tourist Arrivals to Malaysia, Masters Project, Universiti Putra Malaysia, Malaysia

Dickey, D. and Fuller,W. (1979). Distribution of the Estimator for Autoregressive Time series with a Unit Root, *Journal of the American Statistical Association*, 74, pp. 427 – 431

Langkawi Development Authority (2015). Tourist Statistics. Retrieved from http://www.lada.

gov.my/
Langkawi Development Authority (2020). Taman Agroteknolgi MARDI Langkawi. Retrieved from http://www.lada.gov.my/
Lewis, C. D. (1982). *International and Business Forecasting Methods*. London: Butterworths
Loganathan, N. and 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
Loganathan, N., Thirunaukarasu, S. and Mori, K. (2012). Is 'Malaysia Truly Asia'? Forecasting Tourism Demand from Asean using SARIMA Approach. *TOURISMOS: An International Multidisciplinary Journal of Tourism*. Volume 7, Number 1. (367 – 381)
Mansor, K. A. and Ishak, W. I. (2015). Forecasting tourist arrivals to Langkawi island Malaysia. *Cross-Cultural Management Journal*, *17*(1), 69 – 76
Marzuki, A. (2011). Residents Attitudes Towards Impacts from Tourism Development in Langkawi Islands, Malaysia. *World Applied Sciences Journal 12* (Special Issue of Tourism & Hospitality). pp. 25 – 34
Ministry of Agricultre and Agro-based Industry, Malaysia (MOA) (2011). *Dasar Agromakanan Negara 2011 – 2020*
Nuratikah, A., Syahidah, W. N., Afif, M. M. and Nordin, S. Z. (2021 July). SARIMA and Exponential Smoothing model for forecasting eco-tourism demand: A case study in National Park Kuala Tahan, Pahang. In Journal of Physics: Conference Series (Vol. 1988, No. 1, p. 012118). IOP Publishing
Nurhasanah, D., Salsabila, A. M. and Kartikasari, M. D. (2022). Forecasting International Tourist Arrivals in Indonesia Using SARIMA Model. Enthusiastic. *International Journal of Applied Statistics and Data Science, 2(1), 19 – 25. https://doi.org/10.20885/enthusiastic.vol2.iss1.art3*
Pankratz, A. (1983). *Forecasting with Univariate Box-Jenkins Models*. Wiley & Sons, Inc., New York. http://dx.doi.org/10.1002/9780470316566
Shitan, M. (2008). Time series modelling of tourist arrivals to Malaysia. *InterStat*, 1 – 12
Rusyana. A, Nurhasanah, Marzuki and Flancia.M, (2016). SARIMA model for forecasting foreign tourists at the Kualanamu International Airport. *12th International Conference on Mathematics, Statistics, and Their Applications (ICMSA)*, Banda Aceh, 2016, pp. 153 – 158, doi:10.1109/ICMSA.2016.795432
Shitan, M and Tan.Y.F. (2004), A Comparative Study be- tween ARFIMA Models and Intervened Time Series Model for Tourist Arrivals to Malaysia. Paper presented at the Asia Pacific and Business Conference 2004, 7 – 9 October, 2004, Penang
Shitan, M. and Pauline,M.J.W. (2003a), ARAR and Long Memory Modelling of Tourist Arrivals to Malaysia, *Proceedings of National Conference on Management Science and Operations Research*, Vol. 1, 258 – 271, 24 – 25 June, 2003
Shitan, M. and Pauline, M.J.W. (2003b), A comparison between ARFIMA Model and Regression Model with ARMA Errors for Tourist Arrivals in Malaysia, *Proceedings of National Seminar on Decision Science*, 109 – 114, 15 – 16 October, 2003, Langkawi, Kedah
Yu-Wei Chang and Meng-Yuan Liao (2010) A Seasonal ARIMA Model of Tourism Forecasting: The Case of Taiwan. *Asia Pacific Journal of Tourism Research*, 15:2, 215 – 221, DOI: 10.1080/10941661003630001
Milenkovic, M., Syadlenka, L., Melichar, V., Bojovic, N. and Avramovic, Z. (2015). SARIMA modelling approach for railway passenger flow forecasting. Transport, 33(5), 1113 – 1120
Dritsaki, C. (2016). Forecast of Sarima Models: An Application to Unemployment Rates of Greece. Journal of Applied Mathematics and Statistics, 4(5), 136 – 148
Doulah, M. (2018). Forecasting Temperatures in Bangladesh: An Application of SARIMA Models. International Journal of Statistics and Mathematics, 5(1), 108 – 118
Song, Y., Wang, F., Wang, B., Tao, S., Zhang, H., Liu, S., Ramirez,O. and Zeng, Q. (2015). Time Series Analyses of Hand, Foot and Mouth Disease Integrating Weather Variables. Plos One, 1 – 18
Chang, Y. W. and Liao, M. Y. (2010). A Seasonal ARIMA Model of Tourism Forecasting: The Case of Taiwan. Asia Pacific Journal of Tourism Research, 5(2), 215 – 221
Singh, M., Singh, R. and Shinde, V. 2011. Application of Software Packages for Monthly Stream Flow Forecasting of Kangsabati River in India, International Journal of Computer Applications (0975 – 8887), Volume 20 – No.3, 7 – 14. [http://www.ijcaonline.org/volume20/number3/pxc3873231.pdf]

Abstrak
Industri pelancongan telah merekodkan kadar pertumbuhan tahunan sebanyak 11.2% (RM82.6 billion) kepada Keluaran Dalam Kasar Negara (KDNK) pada tahun 2017 berbanding dengan RM76.6 billion pada tahun 2016. Malaysia yang kaya dengan hasil pertanian dan kepelbagaian sumber pertanian dilihat sebagai faktor utama dalam kepesatan pertumbuhan industri agropelancongan. MARDI telah membuka beberapa pusat agropelancongan seperti di Langkawi, Cameron Highlands dan Kuala Kangsar. Penyumbang terbesar ialah Taman Agroteknologi MARDI Langkawi (TATML). Analisis unjuran siri masa dijalankan bagi melihat trend dan melihat unjuran kehadiran pelancong luar negara ke TATML. Hasil kajian menunjukkan kepadatan pengunjung luar negara adalah pada bulan Julai dan Ogos, manakala pengunjung tempatan adalah pada bulan November dan Disember yang merupakan musim cuti sekolah. Bulan Julai sehingga Ogos pula adalah cuti untuk kebanyakan negara seperti Eropah dan Timur Tengah. Analisis unjuran siri masa pula menganggarkan jumlah pengunjung luar negara ke TATML pada tahun 2019 dengan menggunakan model (*Seasonal Autoregressive Integrated Moving Average*) SARIMA (9,0,0) x $(0,1,0)_{12}$ adalah 69,170 orang dan menunjukan trend kepadatan pengunjung yang sama iaitu pada bulan Julai dan Ogos. Dapatan kajian ini dilihat dapat diaplikasikan pada pusat agropelancongan yang lain. Selain itu, strategi promosi dan pengurusan harus ditumpukan kepada musim dan waktu puncak untuk melancarkan perjalanan dan menarik lebih ramai pengunjung luar negara.