Indonesian Tourism Demand Forecasting Using Time Series Approach to Support Decision Making Process

[1]Putu Bella Ayastri Friscintia, [2]Andry Alamsyah

[1]Telkom University

bellafriscintia@gmail.com

Abstract

Tourism industry is always growing dan uphold an important role in national economy as the second largest portion of foreign exchange contributor, as well as its role in national employment. In improving tourism industry, forecasting is needed to anticipate the perishable nature of tourism. Therefore, an accurate forecasting is needed as the baseline of strategic resource planning in order to maximize the utilization dan efficiency of the available resources. The objective of this research is to build an accurate model that is able to forecast Indonesian tourism demand. This research use ARIMA algorithm to forecast the arrivals of tourist. The result of this paper is a time-series model for tourist arrivals in Indonesia.

Keywords: Tourism, Demand Forecasting, ARIMA, Indonesia.

JEL : M3

DOI : 10.24002/kinerja.v22i2.2970

Received : 02/10/2019 Reviewed: 03/15/2019 Final Version: 08/16/2019

1. INTRODUCTION

Indonesian tourism industry has a major role in national economy. In 2017, tourism industry becomes the second largest contributor to Indonesian foreign exchange by successfully generating 14.07 million tourist and 13.5 billion rupiah to Indonesia foreign exchange (Indonesian Ministry of Tourism, 2017). Tourism industry great contribution is also seen in employment. In 2016, 9% of total national employment is from tourism industry (World Travel & Tourism Council, 2017). These phenomenon shows that tourism industry holds an important role in developing the country. Similar opinion also stated by the president of Indonesia who established tourism industry as one of Indonesia's leading economy (Indonesian Statistic Bureau (BPS), 2018). In the future, tourism industry also shows a positive potential. Sandra Horbach, Cohead of Carlyle Group's US through Business Insider stated that global travel is becoming an interesting investment area. Global consumers show a shifting patter of consumption from a tangible product to an intangible experience (Butt, 2016). In the last five years, global tourism shows a positive growth. The increasing number of middle-class families and the decreasing number of unemployment around the world are the triggers of this growth. Here we can conclude that Indonesian tourism has a great potential to grow and develop in the future.

Considering the financial profits and potential growth offered by the tourism industry, a proper strategic planning is needed to allocate financial and other sources. The baseline in an efficient resource planning is forecasting (Chopra & Meindl, 2016). Forecasting is especially important in tourism management because of its perishable nature (Frechtling, 2001). Forecasting is a method to effectively estimate future occurrence. One of the key factors of an efficient resource planning is an accurate forecast. An accurate forecast will help the other following phase in resource planning such as demand planning for market share. Unlike manufacturing product, tourism product is unable to be stored. Exceeds in capacity and failure to accommodate demand is a major problem (Frechtling, 2001). Therefore, demand forecast becomes crucial to prevent future financial loss and high opportunity cost. In this process, an accurate forecast of tourism volume in terms of arrivals and destination is important as an indicator of future demand (Witt & Witt, 1995). There are several proxies to measure tourism demand. One of the most reliable is tourist arrivals (Claveria & Torra, 2014). Tourist arrivals is counted by the authority through arrival gates inbound. Therefore, tourist arrivals is considered an accurate proxy.

This study aims to create a forecast model that best describes and forecast the international tourism demand in Indonesia. To reach this objective, Indonesian monthly tourist arrivals data from 2008-2017 is gathered. A time series method to measured the tourism demand is used to forecast the future arrivals of international tourists.

This study contributes to enrich the research of tourism forecasting through a time-series model. Through this model, the government can estimates the future demand of tourism and prepare a proper strategic plan to accommodate future demand. This study also offers a new insight to the tourism stakeholders about the future potential and pattern of tourism demand in Indonesia.

2. LITERATURE REVIEW

2.1. Tourism Forecasting

Tourism involves every travelling and living activity outside of one's living environment for entertainment, business and other purposes in a span of less than a year (Goeldner & Ritchie, 2009). Tourism industry refers to a corporation, institution, and other organization with a main activity of providing product and service for tourists. There are three main elements in tourism which are the tourists itself, geographic elements, and tourism industry. The interaction between those three will shape transaction and impact (Pender & Sharpley, 2005).

Tourism industry held a significance role in social welfare. The growth of tourism industry will improve the social aspect. Therefore, an effort to averse risk is needed. Forecasting is a process to organize past information of a phenomenon to predict the future (Chopra & Meindl, 2016). Forecasting is particularly important for tourism industry because of the following characteristics as stated by Frechtling (2001). The first one is that tourism product is perishable. This puts a premium on shaping demand in the short run and anticipating it in the long run, to avoid both unsold 'inventory' on the one hand and unfulfilled demand on the other. The second one is that people are inseparable from the production-consumption process. Interaction between suppliers and consumers shapes a tourism product, which makes it happen simultaneously. The third one is customer satisfaction depends on complementary services. Forecasting can help ensure these complementary services are available when and where future visitors need them. The last one is that tourism supply requires large, long lead-time investments in plant, equipment, and infrastructure. Future demand must be anticipated correctly if suppliers are to avoid the financial costs of excess capacity or the opportunity costs of unfilled demand.

Forecasting tourism demand has attracted much interest in academia. A variety of econometric models have been applied. In this regard, according to the comprehensive study of Song and Li (2008), the forecasting methodology is very diverse since researchers use both time series and econometric models in estimating tourism demand. Although the basic indicator describing tourism demand has gradually been modified, tourist arrivals is still the most applicable.

2.2. Time Series Approach

Time series approach considers the internal structure of the variable with respect to their own past data and the random variation. Historical trends and seasonal patterns are explored in time series modelling to predict the future of the variable based on the identified trends and seasonal components (Peiris, 2016). In time series approach there are several algorithm models that can be applied, such as ARIMA (Autoregressive moving-average models), SARIMA (Seasonal Autoregressive moving-average models), and SETAR (Self-exciting threshold autoregression models). Claveria and Torra (2014) compares multiple time series model to artificial neural network and found in several cases ARIMA is capable to outperform other models such as ANN and SETAR, especially in short-term stationary data. Gunter and Onder (2015) also compares multiple univariate and multivariate model to forecast tourist arrivals in Paris. The result shows that ARIMA can provide a reliable predictions in touris arrivals. Several studies to forecast touris arrivals is shown in table 1.

Table 1. The Study of Tourism Demand Forecast

Authors	Topics	Results
C. J. S. C. Burger, M. Dohnal, M. Kathrada, R. Law (2001)	time-series methods for tourism demand forecasting: a case study of Durban, South Africa	Time series analysis can be a valuable tool for tourism forecasters at the beginning of a forecasting project. It allows the forecaster to view trends in visitor behavior, both long-term and cyclical.
Oscar Claveria, Salvador Torra (2014)	Comparison of artificial neural network and time series model to predict tourist arrivals in Catalonia	ARIMA models outperformed SETAR and ANN models, especially for shorter horizons.
Biljana Petrevska (2017)	Predicting F.Y.R. Macedonia tourist arrivals with ARIMA models	The accuracy of the proposed A.R.I.M.A. model can be regarded as good, valid and satisfactory.
H. Rangika Iroshani Peiris (2016)	Predicting tourist arrivals in Sri Lanka with seasonal ARIMA models	Seasonal ARIMA is appropriate to capture patterns of international tourist arrival and to forecast the international tourist arrivals in Sri Lanka with a high accuracy level.
Ida Bagus Kade Puja Arimbawa K, Ketut Jayanegara, I Putu Eka Nila Kencana (2013)	Comparing ANFIS and time series method to predict Australian tourist in Bali	A result of a proper model to predict Autralian tourist in Bali
Nofinda Lestari dan Nuri Wahyuningsih (2012)	Predicting tourists visits in a destination. Case study of Kusuma Agrowisata with ARIMA	A result of a proper model to predict tourist visit in a destination.

3. METHODOLOGY

The object of this research is Indonesian monthly tourist arrivals in the year of 2008-2017. The data is obtained from Indonesian Statistics Bureau (BPS). The methods of this study involve data collection to gather tourism data of Indonesia, literature study to see the significance of tourist arrivals in tourism decision making and model building with ARIMA algorithm.

3.1. ARIMA algorithm

The research is based on the Box–Jenkins methodology (Box & Jenkins, 1976) as a suitable technique for short-run forecasting. It is an algebraic model that is commonly applied in forecasting and is known as autoregressive integrated moving average (ARIMA) models. The ARIMA model works in a formula as stated in equation 1 (Petrevska, 2017).

Equation 1. ARIMA formula

$$\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \ldots + \phi_p Z_{t-p} + a_t = Z_t = a_t - \theta_1 e_{t-1} - \theta_1 e_{t-1} - L - \theta_q e_{t-q}$$

In which p is the order of the autoregressive (AR) process, d is the number of differences or integrations and q is the order of the moving average (MA) process. ARIMA models can be used for short-run estimation based on annual, quarterly, monthly or even weekly, daily or hourly data. The parameter p and q are determined by examining the behavior of auto-correlation function (ACF) and partial autocorrelation function (PACF).

3.2. Model Evaluation

To evaluate if the model is appropriate, several measurements are used. Based on literature review, there are three indicators to measure the errors of forecasts (Claveria & Torra, 2014, Constantino, Fernandes, & Teixeira, 2016, Peiris, 2016, Burger, Dohnal, Kathrada, & Law, 2001). They are Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and Mean Absolute Error (MAE). The calculation of these indicators is stated in equation 2, 3, and 4

Equation 2. RMSE formula

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2}$$

RMSE is a quadratic scoring rule that also measures the average magnitude of the error. It’s the square root of the average of squared differences between prediction and actual observation. In this formula y is the predicted results, while Ŷ is the actual result.

Equation 3. MAPE formula

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_1 - \hat{Y}_1}{Y_1}\right|$$

The MAPE (Mean Absolute Percent Error) measures the size of the error in percentage terms. In this formula y is the predicted results, while Ŷ is the actual result.

Equation 4. MAE formula

$$MAE = \frac{\sum_{t-1}^{n} |e_t|}{n}$$

MAE measures the average magnitude of the errors in a set of predictions, without considering their direction. It's the average over the test sample of the absolute differences between prediction and actual observation where all individual differences have equal weight. Here, e_t is the difference between predicted results and the actual result.

4. RESULT AND DISCUSSION

This study uses monthly tourist arrivals historical data by Indonesian Statistic Bureau from the year 2008-2017 which has the total of 120 data. All this data is processed in Minitab 18.

One of the requirements in implementing ARIMA is to make sure the data is stationary. Therefore, a correlogram test is needed to decide whether the data has a seasonality pattern or not. ACF (autocorrelation function) and PACF (partial autocorrelation function) is shown to see if the data is appropriate for ARIMA. First, in figure 1 is shown the graph of Indonesian monthly tourist arrivals for the year of 2008-2017.

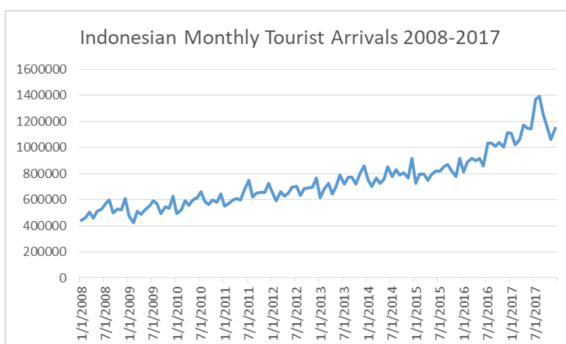

Figure 1. Monthly Tourist Arrivals in Indonesia

In figure 1, the movement of monthly tourist arrivals shows a potential stationarity as a basic assumption. The increasing trend is shown in the graph and relatively stable. Nevertheless, further test is needed to suspect any seasonality in the data and determine if tourist arrivals data is stationary and appropriate for ARIMA model. Figure 2 and 3 shows the correlogram of ACF and PACF. Table 2 shows the correlogram analysis.

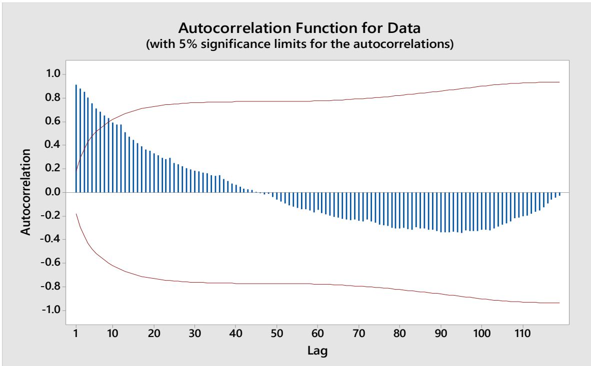

Figure 2. ACF plot

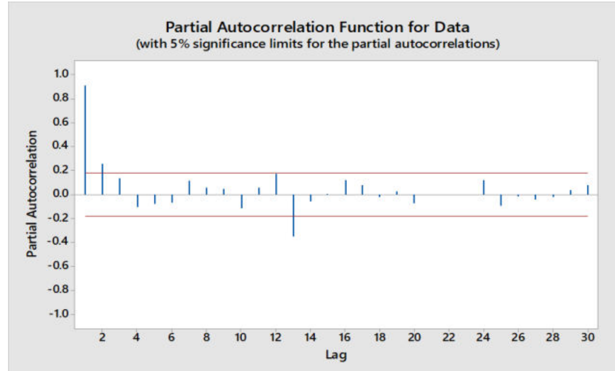

Figure 3. PACF plot

Table 2. Correlogram Analysis

Lag	ACF	UL	LL	PACF	UL	LL
1	91.64%	17.89%	-17.89%	94.84%	17.97%	-17.97%
2	88.06%	29.29%	-29.29%	33.57%	18.04%	-18.04%
3	85.53%	36.80%	-36.80%	30.85%	18.12%	-18.12%
4	80.58%	42.69%	-42.69%	7.11%	18.20%	-18.20%
5	75.78%	47.31%	-47.31%	20.17%	18.28%	-18.28%
6	70.95%	51.05%	-51.05%	16.46%	18.36%	-18.36%
7	68.18%	54.11%	-54.11%	18.09%	18.44%	-18.44%
8	65.14%	56.80%	-56.80%	7.94%	18.52%	-18.52%
9	62.46%	59.14%	-59.14%	24.96%	18.60%	-18.60%
10	58.68%	61.21%	-61.21%	-12.80%	18.69%	-18.69%

Figure 2 shows the ACF correlogram of monthly tourist arrivals. The data is particularly significance in the first lag and continue doing so in the following lags. Figure 3 shows the PACF result. Here, the data in the first lag is significant and then relatively not significant for the others. Both ACF and PACF result shows that there is not any seasonality pattern or repetition in the data. From Table 2, it is also obvious that the data starts at a very high correlation of 91.64% and 94.84%. This number is slowly decay and therefore implies non-stationary properties. To transform these non-stationary data, differencing is needed. After differencing, ACF plot and PACF plot is shown in Figure 4 and 5.

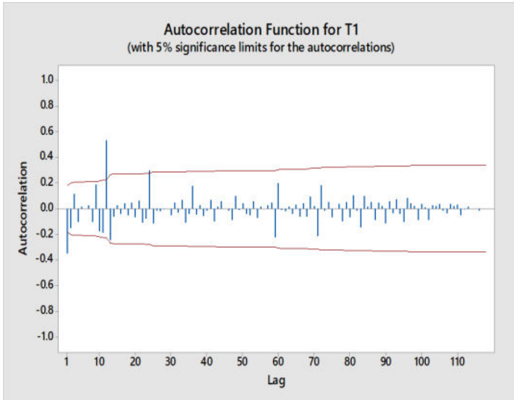

Figure 4. ACF plot after regular differencing

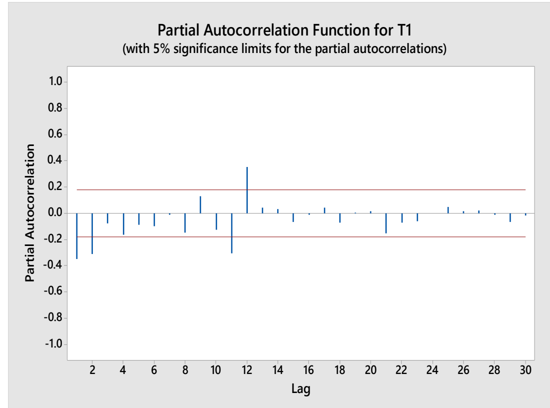

Figure 5. PACF plot after regular differencing

In figure 4 and 5, ACF plot and PACF plot shows that data is now stationary and therefore ARIMA is applicable. By seeing ACF plot which shows a dies-down pattern in lag 3 and PACF that shows dying-down patters as it cuts off in lag 8, therefore ARIMA (0,1,1) $(0,1,1)^{12}$ or ARIMA (0,1,2) $(0,1,2)^{12}$ will be a proper model. By looking at AIC value, ARIMA (0,1,2) $(0,1,2)^{12}$ shows lesser AIC. The significance of parameter is shown in table 2. And table 3 shows the residual test of ARIMA (0,1,2) $(0,1,2)^{12}$

Table 3. Final Estimates of Parameters of ARIMA (0,1,2) $(0,1,2)^{12}$

Type	Coef	SE Coef	T-Value	P-Value	Decision
MA 1	0.471	0.105	4.47	0.000	Significant
MA 2	0.135	0.111	1.21	0.128	Significant
SMA 12	0.460	0.140	3.29	0.001	Significant
SMA 24	0.351	0.148	2.38	0.019	Significant
Constant	854	602	1.42	0.159	Significant

Table 4. White noise residual model ARIMA (0,1,2) $(0,1,2)^{12}$

Modified Box-Pierce (Ljung-Box) Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	11.99	18.45	23.97	41.11
DF	7	19	31	43
P-Value	0.101	0.493	0.812	0.554
Decision	White noise	White noise	White noise	White noise

In Table 3 it shows that in ARIMA (0,1,2) $(0,1,2)^{12}$ the parameter is significant in building ARIMA model. In table 3 the number of tourist monthly arrivals' residual data is considered White noise with p-value greater than 0.05. Other normality test through Kolmogorov-Smirnov test with significance of 5% shows that p-value is greater than 0.05 and therefore concluded as normally distributed as seen in Figure 6

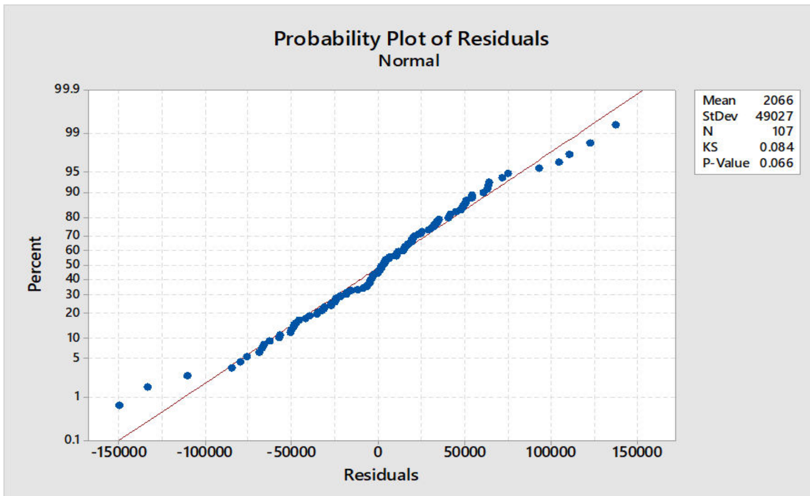

Figure 6. Normality Test Result of ARIMA0,1,2)$(0,1,2)^{12}$

After building the model, we can finally create a forecast for Indonesian monthly tourist arrivals. The forecasting result of 2017 is shown in Table 5.

Table 5. Forecasting Result of model ARIMA (0,1,2) $(0,1,2)^{12}$

Month	Actual Number	Prediction
1/1/2017	1107968	1003568
2/1/2017	1023388	1089497
3/1/2017	1059777	1086760
4/1/2017	1171386	1060658
5/1/2017	1148588	1146990
6/1/2017	1144001	1133295
7/1/2017	1370591	1233450
8/1/2017	1393243	1300062
9/1/2017	1250231	1307140
10/1/2017	1161565	1295121
11/1/2017	1062030	1211578
12/1/2017	1147031	1257284

As the model discern about an event in the future, the most important criterion is how accurately the model is capable of doing so. We measure three parameters: MAE, MAPE and MSE. The result is shown in Table 6.

Table 6. Evaluation of model ARIMA (0,1,2) $(0,1,2)^{12}$

Model	RMSE	MAPE	MAE
ARIMA (0,1,2) $(0,1,2)^{12}$	46,119.80603	4.1395%	32,582

Table 6 presents the evaluation of proposed model. Overall, the model generates good accuracy. MAE shows that there are differences of ±32,582 between the predicted tourist arrivals and the actual number of tourist arrivals. The MAE value of 4.14% shows that the model generates high accuracy forecasting. These three parameters indicate that the model is proper and therefore appropriate to be used in decision making process.

5. CONCLUSION

Tourism in Indonesia holds an important role in strengthening the national economy, especially in national foreign exchange and employment. The number of international tourists visiting Indonesia shows a positive trend over the years and therefore anticipating their demand and providing necessary service to accommodate this demand is crucial. This research provides a model to forecast the future demand of Indonesian tourism industry by predicting the number of tourist arrivals. From this research, a time-series approach model of ARIMA conveys a promising accuracy which is considered high-accuracy. By applying ARIMA to forecast future tourist arrivals, the government and tourism stakeholders can foresee the opportunity and averse the risks of tourism industry. Even though the accuracy is considered high, the suggested model is unable explain the underlying factors of these trends. But nevertheless, the model can provide a solid base to anticipate any potential negative impact and prepare for an effective strategic plan in Indonesia.

Future aspect of this research may focus on micro forecasting that not only forecast the arrival of tourist on the national level but also specific on certain destination. Application of other model such as hybrid models of time-series and econometric is also advised to increase the accuracy of the forecast.

This research can be the baseline of further managerial planning from different levels. As for higher level such as the government, this forecast model can used as a benchmark to see the development of tourism and integrate it with the infrastructure planning. As for the middle level, such as regional government, this model can be integrated with its destination planning. As for lower levels, such as local businesses, flight company, and hotel owners, this model can be integrated with its operational planning in providing service for tourists.

REFERENCE

Box, G. E., & Jenkins, G. M. (1976). Time series analysis: Forecasting and control. San Francisco: Holden-Day Inc.

Burger, C. J., Dohnal, M., Kathrada, M., & Law, R. (2001). A practitioners guide t time-series methods for tourism forecasting - a case study of Durban, South Africa. Tourism Management, 403-409.

Butt, R. (2016, 7 31). Wall Street's Big Investors are Pilling into Tourism. Retrieved from Business Insider: http://www.businessinsider.sg/private-equity-firms-piling-into-tourism-investment-2016-7/

Chen, C. F., & Lai, M. C. (2012). Forecasting Tourism Demand based on Empirical Mode Decomposition and Neural Network. Knowledge-Based Systems, 281-287.

Chopra, S., & Meindl, P. (2016). Supply Chain Management: Strategy, Planning, and Operation. Boston: Pearson.

Claveria, O., & Torra, S. (2014). Forecasting tourism demand to Catalonia: Neural networks vs. time series model. Economic Modelling, 220-228.

Constantino, H., Fernandes, P., & Teixeira, J. (2016). Tourism demand modelling and forecasting withartificial neural network models: The Mozambiquecase study. TÉKHNE - Review of Applied Management Studies, 14(2), 113-124.

Contantino, J. P., Fernandes, H., & Teixeira, P. (2016). Tourism Demand Modelling and Forecasting with Artificial Neural Networks Models: The Mozambique Case Study. Tekhne, 113-124.

Frechtling, D. C. (2001). Forecasting Tourism Demand: Methods and Strategies. Oxford: Butterworth-Heinemann.

Goeldner, C. R., & Ritchie, J. R. (2009). Tourism Principles; Practices; and Philosophies. New Jersey: John Wiley & Sons.

Indonesian Ministry of Tourism. (2017). The Ministry of Tourism Performance Accountability Report. Jakarta: Indonesian Ministry of Tourism.

Indonesian Statistic Bureau (BPS). (2018). Indonesian Economic Report 2018. Jakarta: Badan Pusat Statistik.

K, I. B., Jayanegara, K., & Kencana, I. P. (2013). Komparasi Metode ANFIS dan Fuzzy time series Kasus Peramalan Jumlah Wisatawan Australia ke Bali. E-Jounal Matematika, 18-26.

Lestari, M., & Wahyuningsih, N. (2012). Peramalan Kunjungan Wisata dengan Pendekatan Model SARIMA. Jurnal Sains dan Seni ITS, 29-33.

Peiris, H. R. (2016). A Seasonal ARIMA Model of Tourism Forecasting: The Case of Sri Lanka. Journal of Tourism, Hospitality and Sports, 98-109.

Pender, L., & Sharpley, R. (2005). The Management of Tourism. London: SAGE Publication.

Petrevska, B. (2017). Predicting tourism demand by A.R.I.M.A. models. Economic Research-Ekonomska Istraživanja, 939-950.

Witt, S. F., & Witt, C. A. (1995). Forecasting Tourism Demand: A Review of Empirical Research. International Journal of Forecasting, 11, 447-475.

World Travel & Tourism Council. (2017). Travel & Tourism Economic Impact 2017: Indonesia. New York: World Travel & Tourism Council.

.