



PLANNING MALAYSIA:

Journal of the Malaysian Institute of Planners

VOLUME 23 ISSUE 6 (2025), Page 348 – 360

TOURISM DEMAND FORECASTING IN MALAYSIA USING SARIMA AND LONG SHORT-TERM MEMORY (LSTM) MODEL

**Siti Aishah Tsamienah Taib¹, Noratikah Abu², Nor Azuana Ramli³,
Nur Arina Bazilah Kamisan⁴, Nurul Amira Zainal⁵, Fiki Hidayat⁶**

^{1,2,3}Centre for Mathematical Sciences,

UNIVERSITI MALAYSIA PAHANG AL-SULTAN ABDULLAH

⁴Department of Science Mathematics, Faculty of Science,

UNIVERSITI TEKNOLOGI MALAYSIA

⁵Faculty of Mechanical Engineering,

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

⁶Faculty of Engineering,

UNIVERSITAS ISLAM RIAU

Abstract

This study investigates a robust forecasting framework for predicting Malaysia's international tourist arrivals by comparing the performance of a classical Seasonal Autoregressive Integrated Moving Average (SARIMA) model and a deep learning-based Long Short-Term Memory (LSTM) model. Unlike previous studies that applied these models individually, this research provides a direct empirical comparison using recent historical data on international tourist arrivals to evaluate their suitability for Malaysia's tourism context. Model accuracy was evaluated using two statistical performance metrics. The results demonstrate that both models effectively capture tourism demand patterns, with SARIMA achieving superior forecasting accuracy (MAPE = 5.31%). The findings offer new empirical evidence on model performance in tourism demand forecasting and provide practical insights to support data-driven decision-making and strategic planning within Malaysia's tourism sector.

Keywords: Tourism Demand Forecasting, Time Series, Seasonal Autoregressive Moving Average, Long Short-Term Memory.

¹ Corresponding author. Email: tsamienah@gmail.com

INTRODUCTION

Malaysia, one of Southeast Asia's countries, comprises West Malaysia (Peninsular Malaysia) and East Malaysia. Malaysia is recognized as a fascinating and affordable tourism destination due to its well-developed tourism infrastructure, price competitiveness, diversity of races, cultures, and religions, as well as a variety of food and beverages. The tourism sector is known as a sector that significantly contributes in stimulating economic growth and indirectly has made invaluable contributions to the nation's development particularly after the outbreak of COVID-19.

Nowadays, Malaysian government is exerting significant efforts to redevelop the tourism sector and improve the quality of tourism services, in line with their long-term goal to make Malaysia a top-of-mind destination. Moreover, with strong government support and incentives, the tourism sector has become significantly more productive than the agricultural sector (Othman & Jizan, 2020). All involved practitioners, as well as the government, must play a vital role in strategic planning and development. They should also be well prepared to address the needs of future tourists while ensuring their satisfaction. Tourism forecasting is an important and essential component in the planning and decision-making processes, both in business and non-business organizations.

Forecasting can be defined as a statistical process to generate future values or outcomes by using historical time series data as inputs (Mohd Lip et al., 2020). A reliable tourism demand forecasting model that provides evidence-based insights plays an important role in supporting the Malaysian government's strategic objectives, as outlined in the Twelfth Malaysia Plan (2021–2025) or known as 12MP, the National Tourism Policy 2020–2030, and the Shared Prosperity Vision 2030 (SPV 2030). One of the main focuses in the 12MP is to restore and revitalize the tourism sector after the impact of the COVID-19 pandemic (MEA, 2019; MOTAC, 2020; EPU, 2021). Through the identification of patterns and trends within time series data, forecasting serves as a potential tool for informed decision-making and developing a strategic plan to boost the performance of the tourism industry and achieve higher profit returns.

A robust forecasting framework can contribute significantly to sustaining long-term tourism growth across both urban and rural regions. Accurate projections enable tourism stakeholders to identify growing tourist hotspots and anticipate future demands, thereby facilitating the planning and development of essential infrastructure, accommodations, transport networks, healthcare facilities, community spaces, and other tourism-related amenities. Moreover, these insights support the goals of SPV 2030, particularly in empowering B40 communities in urban areas by creating job opportunities and fostering tourism-related entrepreneurship. In addition, the predictive capacity of a reliable forecasting model can aid in the formulation of comprehensive

contingency plans, enabling stakeholders to better manage crises and employ effective recovery procedures for the tourism sector.

In tourism demand forecasting, there are several variables that can be used as indicators such as the number of tourist arrivals or number of inbound tourists, the total receipts, the number of days of stay, and total expenditure. According to Song and Li (2007), the variable that is commonly used is the volume of tourist arrivals. Hence, this study utilizes historical data on the volume of inbound international tourists in Malaysia and applies the Box-Jenkins and LSTM modelling approaches for forecasting purposes. The main objective of this study is to identify the best model for forecasting tourism demand in Malaysia. All analyses are conducted using Python. The predictive accuracy of each model was computed using RMSE and MAPE to determine the best forecasting model for the tourism sector.

LITERATURE REVIEW

Future decision-making process for tourism-related businesses are always correlated with an accurate forecasting model (Witt & Witt, 1995). The stakeholders involved in these businesses such as governments, regulatory bodies, researchers, tourism operators, or developers, can make informed decision, manage financial risks effectively and distribute resources more efficiently. A review of the literature reveals introduction of numerous forecasting models that have been improved over time to enhance the effectiveness of existing forecasting models.

Box-Jenkins model remains widely used and still relevant as an appropriate forecasting approach. For instance, Peiris (2016) conducted tourism demand forecasting using the seasonal ARIMA model and found that SARIMA(1,0,16)(32,0,24)₁₂ outperformed other candidate models based on the forecast accuracy. Similarly, Msofe & Mbago (2019) identified that the SARIMA(1,1,1)(1,1,2)₁₂ model performed effectively in forecasting international tourist arrivals to Zanzibar. In the context of Malaysia tourism, Mohd Lip et al. (2020) employed both Holt-Winter model and seasonal ARIMA model for international tourist arrivals prediction in Malaysia. The researchers used secondary time series data, which was tourist arrival volumes from Singapore, Korea, and United Kingdom. Their findings indicated that the SARIMA model was significantly effective at forecasting tourist arrivals from Singapore.

In addition, Abu et al. (2021) conducted a comparative analysis between the Holt-Winters and SARIMA model to identify the most accurate demand forecasting model for eco-tourism in Pahang. Based on the empirical results, they concluded that the SARIMA(1,0,0)(1,0,1)₁₂ model outperformed the Holt-Winters model in forecasting tourist arrivals to National Park Kuala

Tahan, Pahang, Malaysia. Nurhasanah et al. (2022) forecasted the inbound international tourist's volume in Indonesia and the results showed that the SARIMA model provided the most accurate forecasts, as indicated by the lowest percentage error. Kamisan et al. (2024) compared the performance of SARIMA, MLP, and a hybrid model to forecast vehicle purchase loan approvals in Malaysia. Their findings showed that the MLP model outperformed the others, offering higher accuracy than traditional models like SARIMA, with significant implications for banking decisions. However, despite its strengths, the SARIMA model, which is commonly used in time series forecasting, often struggles to capture complex and nonlinear features in tourism data influenced by behavioural factors or external shocks, as it assumes a linear relationship within the data (Wu et al., 2020).

In contrast, the LSTM model, a subclass of AI-based forecasting methods, is widely used in time series forecasting, particularly for handling nonlinear patterns and long-term dependencies in the data. In 2018, Li and Cao employed the LSTM model with two other baseline models to forecast tourism flow. The outcomes of their study indicated that the LSTM model outperformed the baseline models. Laaroussi et al. (2020) used the volume of inbound Moroccan tourists in Morocco as the input for tourism demand forecasting using LSTM, gated recurrent unit (GRU), support vector regression (SVR) and artificial neural network (ANN) models. They concluded that the LSTM and GRU models demonstrated superior performance compared to the other models in this study. In addition, Hsieh (2021) applied the LSTM, Bidirectional LSTM (Bi-LSTM) and GRU for tourism demand forecasting in Taiwan. The study identified LSTM with its variants as the best forecasting model as it provided the lowest MAPE value. Finally, the study conducted by Mukhtar et al. (2023) also revealed the LSTM model as an appropriate forecasting model for tourism industry in Indonesia.

Previous studies have typically evaluated these two baseline models separately, leaving a gap in understanding how classical linear and modern deep learning models perform under identical historical data conditions. Addressing this gap is crucial to determine whether the nonlinear learning capability of the LSTM model can translate into measurable forecasting improvements over SARIMA in real-world tourism applications. Therefore, this study conducts a systematic comparative analysis of SARIMA and LSTM models for forecasting international tourist arrivals in Malaysia, providing new empirical evidence to support both academic research and practical forecasting strategies.

RESEARCH METHODOLOGY

Data

This study utilized monthly historical data on the international tourist arrivals volume in Malaysia, before the emergence of the COVID-19 outbreaks. The data, consisting of 242 observations (spanning from January 2000 to February 2020), was sourced from the Malaysia Tourism Satellite Account, obtained from MOTAC.

Box-Jenkins Model

In 1970, the Box-Jenkins model was developed by Box and Jenkins and is recognized as a classical forecasting model. This classical model is divided into two groups: stationary models (e.g., AR model, MA model, and ARMA model) and non-stationary models (e.g., ARIMA model and seasonal ARIMA model) (Montgomery et al., 2015).

The Box-Jenkins model is broadly utilized in the forecasting process and in line with the current situation. Several modified forecasting models were designed by maintaining ARIMA model as the base forecasting model (Kumar & Anand, 2014). According to Hyndman and Athanasopoulos (2018), the ARIMA model is highly effective for forecasting annual time series data, while the seasonal ARIMA (SARIMA) model is better suited for high-frequency time series data (such as daily, weekly and monthly). Since our time series data consisted of monthly tourist arrivals volume which normally exhibit seasonality behavior (Kumar & Sharma, 2016), this study focused on utilizing the SARIMA model for tourism demand forecasting in Malaysia.

This model is an extension of the ARIMA model that incorporates seasonality (Kumar & Anand, 2014). The seasonal ARIMA model is denoted as $SARIMA(p, d, q)(P, D, Q)_m$ where (p, d, q) and $(P, D, Q)_m$ represent the non-seasonal and seasonal parts, respectively. The parameters p , d and q indicate the order of autoregressive, degree of non-seasonal differencing and the order of moving average, respectively. Meanwhile, the parameters P , D and Q denote the order of seasonal autoregressive, degree of seasonal differencing and the order of seasonal moving average, respectively, while m represents the seasonal cycle length.

In this study, the SARIMA model followed an iterative approach comprising four stages to yield the best fitted model for forecasting. The first stage is model identification where, three major behaviours of time series data were identified: stationarity in-mean, stationarity in-variance and seasonality. A Box-Cox transformation was implemented to check the stationarity in-variance while the Augmented Dickey Fuller (ADF) test and time plot was used to check stationarity in-mean. Differencing process was carried out for non-stationary to

stationary time series transformation. Finally, the Ljung-Box Q test was used to verify the presence of serial correlation in the time series data.

Next, in stage 2, the model with the smallest AIC value is acknowledged as the most appropriate forecasting model. In stage 3, diagnostic checking, the residuals were evaluated using the Ljung-Box Q. Once the final model was obtained, it was used to forecast the future time series data.

LSTM Model

Achieving a robust forecasting model in tourism demand remains a crucial challenge for most researchers, particularly when dealing with the non-linear features of time series data. Based on previous studies, a deep learning approach, specifically the long short-term memory (LSTM) model can be adopted to handle the non-linear relationship (Laaroussi et al., 2023). The LSTM model is a specialized type of recurrent neural networks (RNN) and was designed by Hochreiter and Schmidhuber in 1997.

According to Li and Cao (2018) and Law et al. (2019), the network in LSTM consists of a cell and a gating mechanism composed forget gate, input gate, input modulation gate and output gate which interact with each other. The three gates (forget, input and output gate) responsible for the signals transfer into and out of it while the input modulation gate memorizes the values over time intervals.

The unique function of the input modulation gate enables the LSTM to implicitly learn repeating or cyclical patterns inherent in the data. In the context of tourism demand forecasting, the seasonality such as annual peaks or periodic fluctuations can be captured by the LSTM through its long-term memory state. As the network processes multiple seasonal cycles during training, it gradually identifies consistent temporal patterns and learns to associate specific time intervals with expected changes in the data. Law et al. (2019) describing the general LSTM equations introduced by Hochreiter and Schmidhuber are given as follows:

Input gate:	$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i)$	(1)
Forget gate:	$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f)$	(2)
Output gate:	$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o)$	(3)
Memory cell:	$c_t = f_t \times c_{t-1} + i_t \times \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c)$	(4)
Output:	$h_t = o_t \times \tanh(c_t)$	(5)
Input vector:	x_t	(6)
The activation functions:	$\tanh, \sigma$	(7)
Learning parameters:	W, b	(8)

In order to produce an accurate tourism demand forecasting model in this study, five stages were involved in the implementation of the LSTM. The first stage was data preparation, during which the historical time series data was

filtered before being used in the LSTM training process. The second stage involved data normalization, a widely used technique to improve the LSTM learning acceleration rate and convergence level. In this process, the input data was rescaled within the range $[0,1]$ using MinMaxScaler. Next, the data was split into two parts during the third stage: training data and testing data, based on a ratio of 80:20 respectively. An 80:20 ratio is commonly adopted as a standard time series split ratio for datasets of moderate size, as allocating a larger portion for training ensures that the model sufficiently learns from historical trends to predict future outcomes. This proportion has been widely applied and recommended in previous studies on time series forecasting and machine learning applications, as demonstrated by Zhang et al. (2018) and Makridakis et al. (2021). Training data is allocated for model training, while testing data is used to evaluate the model's performance. The fourth stage involved the training process of the LSTM model.

In order to prevent under-fitting or over-fitting, the hyper-parameters were modified along the process. Each hyperparameter play an important role in the operation of each gate, as they determine how effectively and how fast the gates learn to control the flow of information within the LSTM network. The LSTM architecture employed in this study was adapted from the preliminary experiments conducted by Li and Cao (2018) and Law et al. (2019). The optimal configuration of hyperparameters, which yielded the lowest forecasting error, was determined through a manual tuning process, commonly referred to as trial-and-error experimentation. Several combinations of hyperparameters, including the number of LSTM units, learning rates, and epochs were tested, while the number of inputs was fixed at one, as this study utilized univariate time series data. The number of output nodes was also fixed at one, since only one-step-ahead forecasting was considered in this study. The LSTM model employed the sigmoid activation function and the Adam optimizer with adaptive moment estimation were used to facilitate efficient gate learning and gradient stability. The final configuration, consisting of one input, four LSTM units, and 50 epochs, produced the most stable training performance. Based on the trial-and-error results, increasing the number of LSTM units did not significantly enhance validation accuracy, suggesting that a smaller number of units is sufficient for datasets with a moderate number of observations (242 data points).

Forecast Accuracy

To validate the performance of SARIMA and LSTM, the Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) values are utilized. The model with the smallest percentage error is considered the most appropriate for tourism demand forecasting.

ANALYSIS AND DISCUSSION

The volume of inbound tourists in Malaysia before the COVID-19 pandemic was used as the input for both models in this study. There were 242 data; divided into two categories, 192 data for the training data and 50 data for the testing data. The training data of the volume of inbound tourists in Malaysia before COVID-19 pandemic is illustrated in Figure 1.

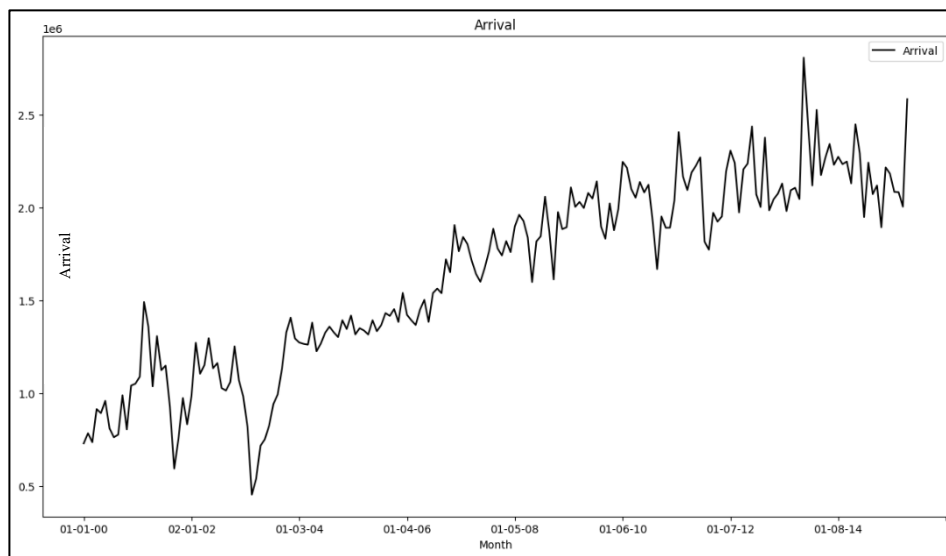


Figure 1: The Training Data of the Volume of Inbound Tourists in Malaysia before COVID-19 Pandemic

Source: Annual Report of the Malaysian Promotion Board Statistics (2021)

Based on Figure 1, the data plotting exhibits both upward and downward movements over the given period. Therefore, this time series data can be classified as non-stationary. The data series also display a linear trend since the time series data tend to increase for the same month over time. Additionally, the pattern of this data set exhibits seasonality, characterized by recurring pattern of fluctuations during specific months each year. Hence, the SARIMA model, a specific Box-Jenkins model is identified as suitable for this study.

In the parameter estimation process, 31 tentative models of SARIMA were generated using Python software. $SARIMA(2,1,1)(0,1,1)_{12}$ was chosen as the best model to forecast the number of tourist arrivals in Malaysia due to its lowest AIC value. The chosen SARIMA model is formulated as follows:

$$\begin{aligned} & (1 + 0.6618B + 0.3217B^2)(1 - B)(1 - B^{12})y_t \\ & = (1 + 0.6057B^{12})(1 - 0.4908B)a_t \end{aligned} \quad (9)$$

For the LSTM model analysis, the process of training optimized at 46th epoch and the value of loss function obtained was 0.5%. The model was over-fitting starting from 47th epoch and onwards. Based on the observations, the hyper-parameters able to effectively control the training process.

The performance of SARIMA and LSTM models was evaluated using RMSE and MAPE. The summary of performance is tabulated in Table 1. According to [20], both models are classified as highly accurate prediction model since the MAPE values obtained for both models are smaller than 10%. This is verified from the MAPE values presented in Table 1. It can be observed that the SARIMA model performed better than the LSTM model.

Table 1: The performance evaluation of SARIMA and LSTM models

Forecast evaluation	SARIMA (2,1,1) (0,1,1) ₁₂	LSTM
RMSE	152539.8	182382.0
MAPE	5.3120	6.8123

Source: Author's Calculation

Figure 2 illustrates the comparison between the actual and predicted values generated by the SARIMA model. It is evident that the actual values exhibit a pattern that is nearly identical to the predicted values. As depicted in Figure 3, the LSTM performed well in fitting the time series data, as the pattern of predicted values closely matched the pattern of the actual values.

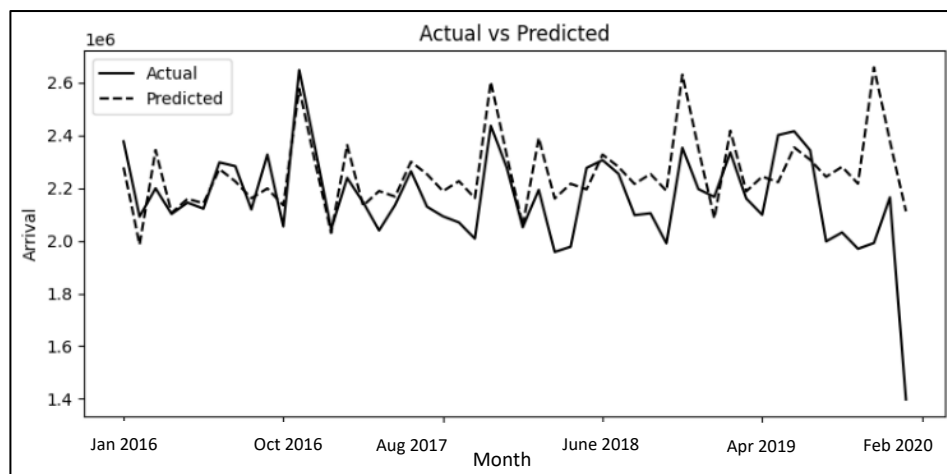


Figure 2: Actual Value versus Predicted Value using SARIMA Model

Source: Author's Calculation (2023)

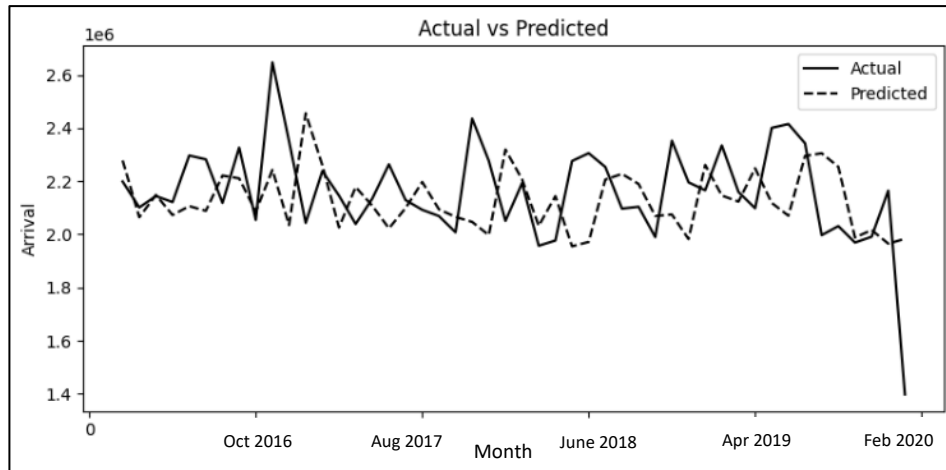


Figure 3: Actual Value versus Predicted Value using LSTM Model
Source: Author's Calculation (2023)

The characteristics of the tourism data used in this study contributed to the superior performance of the SARIMA model over the LSTM model. The data exhibited a smooth linear relationship and a clear seasonal pattern, which consistent with SARIMA's linearity assumption and its ability to capture periodic fluctuations through autoregressive (AR) and moving average (MA) components. Moreover, the nonlinear learning advantage of the LSTM model was not fully exploited in this study due to the moderate size of the tourism dataset, as LSTM models typically require a large dataset to perform effectively. The findings of this study align well with existing studies conducted by Zhang et al. (2003), Song and Li (2007), and Hyndman & Athanasopoulos (2018), which reported that classical forecasting models, such as ARIMA and SARIMA, often outperform neural network models when the time series data exhibit stable and linear seasonal patterns.

CONCLUSION

Based on the results, the SARIMA model produced a smaller percentage error compared to the LSTM model. This confirms that the traditional SARIMA model remains a reliable choice for prediction in time series analysis, particularly when dealing with stable, small to moderate-sized datasets. Although the percentage error produced by the LSTM model is larger than the SARIMA model, the LSTM model is still suitable for predicting tourist arrivals in Malaysia, as it achieved an error value of less than 10%, qualifying it as a highly accurate forecasting model according to its MAPE value (Lewis, 1982).

The findings of this study demonstrate that a well-structured tourism demand forecasting model can be a valuable tool in supporting Malaysia's future national development. The insights generated from the model can significantly contribute to the achievement of goals outlined in the country's key development plans. These include promoting sustainable and long-term tourism growth, advancing digital transformation in the tourism sector or known as smart tourism (Amir et al., 2020), supporting environmentally responsible planning through net-zero initiatives, preserving Malaysia's cultural and historical heritage, and strengthening the niche of tourism markets. Reliable forecasts can help tourism practitioners effectively anticipate peak seasons by managing workforce distribution, allocating resources wisely, and optimizing transportation and accommodation services to meet the needs of future tourists and visitors. As a direction for future research, this study proposes extending the model by combining SARIMA and LSTM to enhance forecasting effectiveness, particularly for tourism demand in Malaysia.

ACKNOWLEDGMENT

The authors would like to express gratitude for all motivation and guidance provided throughout this research work. This study was funded by a "GERAN KERJASAMA STRATEGIK MTUN" under project code RDU233501.

REFERENCES

- Abu, N., Syahidah, W. N., Afif, M. M., & Nordin, S. Z. (2021). SARIMA and Exponential Smoothing model for forecasting ecotourism demand: A case study in National Park Kuala Tahan, Pahang. *Journal of Physics Conference Series*, 1988(1), 012118. <https://doi.org/10.1088/1742-6596/1988/1/012118>
- Amir, S., Dura, N. H., Yusof, M. A., Nakamura, H., & Nong, R. A. (2020). Challenges of smart tourism in Malaysia eco-tourism destinations. *Journal of the Malaysian Institute of Planners*, 18(4), 442-451. <https://doi.org/10.21837/pm.v18i14.844>
- EPU. (2021). Twelfth Malaysia Plan 2021-2025. https://pulse.icdm.com.my/wp-content/uploads/2021/09/Twelfth-Plan-Document_compressed-1.pdf
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term memory. *Neural Computation*, 9(8), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Law, R., Li, G., Fong, D. K. C., & Han, X. (2019). Tourism demand forecasting: A deep learning approach. *Annals of Tourism Research*, 75, 410-423. <https://doi.org/10.1016/j.annals.2019.01.014>
- Hsieh, S. (2021). Tourism demand forecasting based on an LSTM network and its variants. *Algorithms*, 14(8), 243. <https://doi.org/10.3390/a14080243>
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice*. OTexts.
- Kamisan, N. a. B., Jafry, N. A., & Norrulashikin, S. M. (2024). Forecasting loan consented for vehicle purchase in Malaysia. *Malaysian Journal of Fundamental*

- and *Applied Sciences*, 20(4), 728–738. <https://doi.org/10.11113/mjfas.v20n4.2988>
- Kumar, M., & Anand, M. (2014). An application of time series ARIMA forecasting model for predicting sugarcane production in India. *Studies in Business and Economics*, 9(1), 81–94. <http://eccsf.ulbsibiu.ro/articole/vol91/918kumar&anand.pdf>
- Kumar, M., & Sharma, S. (2016). Forecasting tourist in-flow in South East Asia: A case of Singapore. *Tourism & Management Studies*, 12(1), 107–119. <https://doi.org/10.18089/tms.2016.12111>
- Laaroussi, H., Guerouate, F., & Sbihi, M. (2020). Deep learning framework for forecasting tourism demand. *IEEE*. <https://doi.org/10.1109/ictmod49425.2020.9380612>
- Laaroussi, H., Guerouate, F., & Sbihi, M. (2023). A novel hybrid deep learning approach for tourism demand forecasting. *International Journal of Electrical and Computer Engineering*, 13(2), 1989–1996. <http://dx.doi.org/10.11591/ijece.v13i2.pp1989-1996>
- Law, R., Li, G., Fong, D. K. C., & Han, X. (2019). Tourism demand forecasting: A deep learning approach. *Annals of Tourism Research*, 75, 410–423. <https://doi.org/10.1016/j.annals.2019.01.014>
- Lewis, C. D., Schilling, E. E., Panero, J. L., Aranda, P. D., & Crozier, B. S. (1982). *Industrial and business forecasting methods*. https://lib.pdii.lipi.go.id/index.php?p=show_detail&id=5883&keywords=
- Li, Y., & Cao, H. (2018). Prediction for tourism flow based on LSTM neural network. *Procedia Computer Science*, 129, 277–283. <https://doi.org/10.1016/j.procs.2018.03.076>
- Makridakis, S., Spiliotis, E., & Assimakopoulou, V. (2020). The M4 competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*, 36(1), 54–74. <https://doi.org/10.1016/j.ijforecast.2019.04.014>
- MEA. (2019). Shared Prosperity Vision 2030. <https://ekonomi.gov.my/sites/default/files/202002/Shared%20Prosperity%20Vision%202030.pdf>
- Mohd Lip, N., Jumery, N. S., Termizi, F. a. A., Mulyadi, N. A., Anuar, N., & Ithnin, H. (2020). Forecasting international tourist arrivals in Malaysia using SARIMA and Holt-Winters model. *Journal of Tourism Hospitality and Environment Management*, 5(18), 41–51. <https://doi.org/10.35631/jthem.518004>
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to time series analysis and forecasting*. John Wiley & Sons, Inc., Hoboken, New Jersey.
- MOTAC. (2020). Executive Summary: National Tourism Policy 2020–2030. https://www.tourism.gov.my/files/uploads/Executive_Summary.pdf
- 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>
- Mukhtar, H., Remli, M. A., Wong, K. N. S. W. S., Fuad, E., Siregar, J., & Rizki, Y. (2023). Forecasting Tourist Arrivals with Partial Time Series Data Using Long-Short Term Memory (LSTM). *Zenodo (CERN European Organization for Nuclear Research)*. <https://doi.org/10.5281/zenodo.7970542>

- Nurhasanah, D., Salsabila, A. M., & Kartikasari, M. D. (2022). Forecasting international tourist arrivals in Indonesia using SARIMA model. *Enthusiastic International Journal of Applied Statistics and Data Science*, 19–25. <https://doi.org/10.20885/enthusiastic.vol2.iss1.art3>
- Othman, A. G., & Jizan, N. H. (2020). Land use changes in rural town: A case study of Kuala Nerang, Kedah. *Journal of the Malaysian Institute of Planners*, 18(1), 118–130. <https://doi.org/10.21837/pm.v18i1.714>
- Peiris, H. R. I. (2016). A seasonal ARIMA model of tourism forecasting: The case of Sri Lanka. *Journal of Tourism, Hospitality and Sports*, 22, 98–109. <https://www.iiste.org/Journals/index.php/JTHS/article/download/33831/34820>
- Song, H., & Li, G. (2007). Tourism demand modelling and forecasting—A review of recent research. *Tourism Management*, 29(2), 203–220. <https://doi.org/10.1016/j.tourman.2007.07.016>
- Witt, S. F. & Witt, C. A. (1995). Forecasting tourism demand: A review of empirical research. *International Journal of Forecasting*, 11(3), 447–475. [https://doi.org/10.1016/0169-2070\(95\)00591-7](https://doi.org/10.1016/0169-2070(95)00591-7)
- Wu, D. C. W., Ji, L., He, K., Tso, K. F. G. (2021). Forecasting tourist daily arrivals with a hybrid SARIMA-LSTM approach. *Journal of Hospitality & Tourism Research*, 45(1), 52–67. <https://doi.org/10.1177/1096348020934046>
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175. [https://doi.org/10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0)
- Zhang, X., Zhang, Q., Zhang, G., Nie, Z., & Gui, Z. (2018). A hybrid model for annual runoff time series forecasting using Elmann neural network with ensemble empirical mode decomposition. *Water*, 10(4). <https://doi.org/10.3390/w10040416>

Received: 2nd September 2025. Accepted: 14th September 2025