Department of Digital Business
Journal of Artificial Intelligence and Digital Business (RIGGS)
Homepage: https://journal.ilmudata.co.id/index.php/RIGGS
Vol. 5 No. 2 (2026) pp: 3949-3955
P-ISSN: 2963-9298, e-ISSN: 2963-914X

Tourist Arrival Forecasting in Wakatobi Using LSTM with Comparative Analysis

La Ode Alvin Rahmat Saputra, La Ode Muh. Golok Jaya, Ilham Julian Efendi, Isnawaty
Department of Informatics, Faculty of Engineering, Halu Oleo University, Kendari, Indonesia
laodealvinrahmats@gmail.com*, laodemgj@uho.ac.id, ilham.julian.efendi@uho.ac.id, isnawaty@uho.ac.id

Abstract

Wakatobi Regency, located in Southeast Sulawesi, Indonesia, is widely recognized as one of the world's leading marine tourism destinations due to its rich biodiversity, coral reef ecosystems, and growing ecotourism sector. Despite its international reputation, the region continues to experience substantial fluctuations in monthly tourist arrivals. These variations are influenced by seasonal tourism cycles, weather conditions, transportation accessibility, economic dynamics, and the long-term structural impacts caused by the COVID-19 pandemic. Such instability creates significant challenges for regional authorities in designing sustainable tourism policies, allocating resources, and planning infrastructure development. Therefore, accurate tourism forecasting has become increasingly important for supporting evidence-based decision making and improving destination management strategies. This study develops and evaluates a Long Short-Term Memory (LSTM)-based forecasting model to predict monthly tourist arrivals in Wakatobi using 192 monthly observations collected from January 2010 to December 2025. The proposed LSTM model is compared with three benchmark models, namely SARIMAX, XGBoost, and Transformer, to assess comparative forecasting performance. Hyperparameter optimization is conducted through a systematic grid search process, resulting in the optimal configuration consisting of a window size of 6, 128 LSTM units in the first layer, 64 units in the second layer, and a dropout rate of 0.1. Experimental results indicate that the LSTM model achieves the highest predictive accuracy with a MAPE of 8.98%, categorized as High Accuracy, alongside an MAE of 104.65 and RMSE of 143.20. These results outperform Transformer, XGBoost, and SARIMAX models. Statistical validation using the Diebold-Mariano test further confirms the superiority of LSTM forecasting performance. Additionally, the developed model is implemented in a Flask-based web application that enables regional government officials to conduct interactive forecasting with multiple prediction horizons and SHAP-based interpretability analysis.

Keywords: LSTM, Tourist Arrival Forecasting, Wakatobi, Deep Learning, Time Series

1. Introduction

Tourism is a major economic driver globally, contributing to GDP, employment, and regional development. In Indonesia, tourism has been recognized as a strategic sector within the national development framework. Wakatobi Regency, an archipelagic region in Southeast Sulawesi, holds international recognition as a UNESCO Man and the Biosphere destination due to its exceptional marine biodiversity, hosting 750 coral species and 942 fish species within the global Coral Triangle [1].

Historical data from the Wakatobi Statistics Agency (BPS Wakatobi) and the Wakatobi Tourism Office show that monthly tourist arrivals grew steadily from 6,793 annual visitors in 2010 to a peak of 28,857 in 2019. The COVID-19 pandemic caused a dramatic collapse, with annual arrivals plummeting to 3,511 in 2020, representing an 87.8% decline from the 2019 peak. Although recovery has been observed, with arrivals reaching 15,378 in 2025, the data exhibit a standard deviation of approximately 907 tourists per month across the full observation period, reflecting persistent inter-monthly and inter-annual variability [2].

Despite these challenges, tourism planning in Wakatobi continues to rely heavily on manual estimations and simple trend analyses, which lack the accuracy required for data-driven decision-making. This gap necessitates the adoption of advanced forecasting models capable of capturing the nonlinear, seasonal, and structurally disrupted patterns present in tourism data [3].

Traditional methods such as ARIMA and SARIMAX have been widely used for tourism demand forecasting due to their interpretability and statistical rigor. However, they rely on linear assumptions and struggle with complex nonlinear patterns [4]. Machine learning approaches such as XGBoost have demonstrated strong predictive performance on tabular data but lack inherent mechanisms for modeling sequential temporal dependencies [5]. Deep learning methods, particularly Long Short-Term Memory (LSTM) networks introduced by Hochreiter and Schmidhuber [6], have shown superior capability in capturing long-term temporal dependencies in time series data. Several studies confirm that LSTM outperforms traditional statistical models in tourism demand forecasting [7][8][9].

Existing studies on LSTM-based tourism forecasting primarily focus on national or city-level destinations with large visitor volumes. No prior research has applied LSTM to Wakatobi, a small archipelagic marine destination with distinct seasonality and structural COVID-19 disruption. This study addresses this gap by developing an optimized LSTM model for Wakatobi, comparing it against SARIMAX, XGBoost, and Transformer baselines, and deploying results in a Flask-based web application designed for regional government use.

2. Research Methods

2.1. Dataset

The dataset comprises 192 monthly observations of tourist arrivals in Wakatobi Regency from January 2010 to December 2025, obtained through direct field consultation with BPS Wakatobi and the Wakatobi Tourism Office. The dataset includes eight variables: Total Tourist Arrivals (target output), Departing Passengers, Arriving Passengers, Total Passengers, Monthly Rainfall (mm), Public Holidays, Tourism Event Count, and a COVID-19 Restriction Dummy. The data were confirmed to be consistently recorded by the relevant institutions, ensuring that variable selection is grounded in official data availability rather than researcher assumptions.

Table 1. Dataset Description

No.	Variable	Type	Description	Source
1.	Tourist_Arrivals	Target	Monthly total tourist arrivals	BPS / Dinas Pariwisata
2.	Passenger_Depart	Feature	Monthly departing passengers	Dinas Pariwisata
3.	Passenger_Arrive	Feature	Monthly arriving passengers	Dinas Pariwisata
4.	Passenger_Total	Feature	Monthly total passenger movement	Dinas Pariwisata
5.	Rainfall_mm	Feature	Monthly rainfall in millimeters	BPS Wakatobi
6.	Public_Holidays	Feature	Number of national public holidays per month	National Holiday Calendar
7.	Tourism_Event	Feature	Number of tourism events held per month	Dinas Pariwisata
8.	Google_Trends	Feature	Monthly total google search	Google Trends
9.	COVID_Dummy	Feature	Binary restriction indicator (1 = restriction period, 0 = normal)	Government Regulation

2.2. Data Preprocessing and Feature Engineering

Prior to model training, outlier handling was applied using winsorization at the 1st and 99th percentiles to mitigate the influence of extreme values without data removal. Seven additional features were engineered from the Total Tourist Arrivals variable: lag features at 1, 3, and 12 months; rolling mean features over 3, 6, and 12 months; a year-over-year ratio (capped at 3); and cyclical month encoding using sine and cosine transformations. The final feature set contains 15 variables. All variables were normalized to the range [0,1] using Min-Max Scaling applied only to the training partition to prevent data leakage.

2.3. Model Architectures

Four forecasting models were developed and compared. The LSTM model consists of two stacked LSTM layers followed by a Dense output layer, with dropout regularization applied after each layer. The Transformer model employs a multi-head self-attention architecture adapted for univariate time series forecasting. XGBoost is applied

DOI: https://doi.org/10.31004/riggs.v5i2.8788

as a gradient boosting baseline using lag features as inputs. SARIMAX serves as the statistical baseline, incorporating seasonal autoregressive and moving average components with exogenous variables.

2.4. Hyperparameter Optimization

A systematic grid search across 16 configurations was conducted for the LSTM model, varying window size (6, 12), LSTM units in Layer 1 (64, 128), LSTM units in Layer 2 (32, 64), and dropout rate (0.1, 0.2). The data split strategy follows a temporal partition: training on data up to December 2023, validation on 2024, and testing on 2025, preserving chronological order and preventing data leakage. The Adam optimizer with a learning rate of 0.001 and MAE loss function were used across all LSTM configurations.

2.5. Evaluation and Statistical Testing

Model performance was evaluated using three metrics: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). MAPE classification follows Lewis (1982): below 10% indicates High Accuracy, 10–20% Good, 20–50% Reasonable, and above 50% Inaccurate. To assess whether differences in forecasting performance between LSTM and each baseline model are statistically meaningful, the Diebold-Mariano (DM) test was applied at a 5% significance level [10].

a) MAE (Mean Absolute Error)

$$MAE = \frac{1}{n}\sum_{t=1}^{n} | y_t - \hat{y}_t | \quad (1)$$

b) RMSE (Root Mean Squared Error)

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}.7 \quad (2)$$

c) MAPE (Mean Absolute Percentage Error)

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} | \frac{y_t - \hat{y}_t}{y_t} | \times 100 \quad (3)$$

3. Results and Discussions

3.1. Grid Search Results

Table 1 presents the top results from the 16-configuration grid search. The optimal LSTM configuration uses a window size of 6, 128 units in Layer 1, 64 units in Layer 2, a dropout of 0.1, batch size 8, and learning rate 0.001. This configuration achieved a validation MAPE of 7.25% and a test MAPE of 9.67% on the 2025 test set.

Table 1. Top LSTM Grid Search Results

WS	Units (L1, L2)	Dropout	Val MAPE (%)	Test MAPE (%)
6	128, 64	0.1	7.25	9.67
6	64, 64	0.1	9.49	8.18
6	64, 64	0.2	9.68	8.59
6	128, 32	0.1	10.20	9.17
12	64, 32	0.1	12.02	10.03

3.2. Model Comparison

Table 2 presents the comparative evaluation results of all four models on the 2025 test set. The LSTM model achieves the lowest error across all metrics, with a MAPE of 8.98%, classified as High Accuracy according to Lewis (1982). The Transformer model ranks second with a MAPE of 16.43% (Good), followed by XGBoost at 23.10% (Good) and SARIMAX at 43.49% (Inaccurate). The large error of SARIMAX is attributable to the

DOI: https://doi.org/10.31004/riggs.v5i2.8788

structural break caused by the COVID-19 pandemic, which disrupts the stationarity assumptions underlying the statistical model.

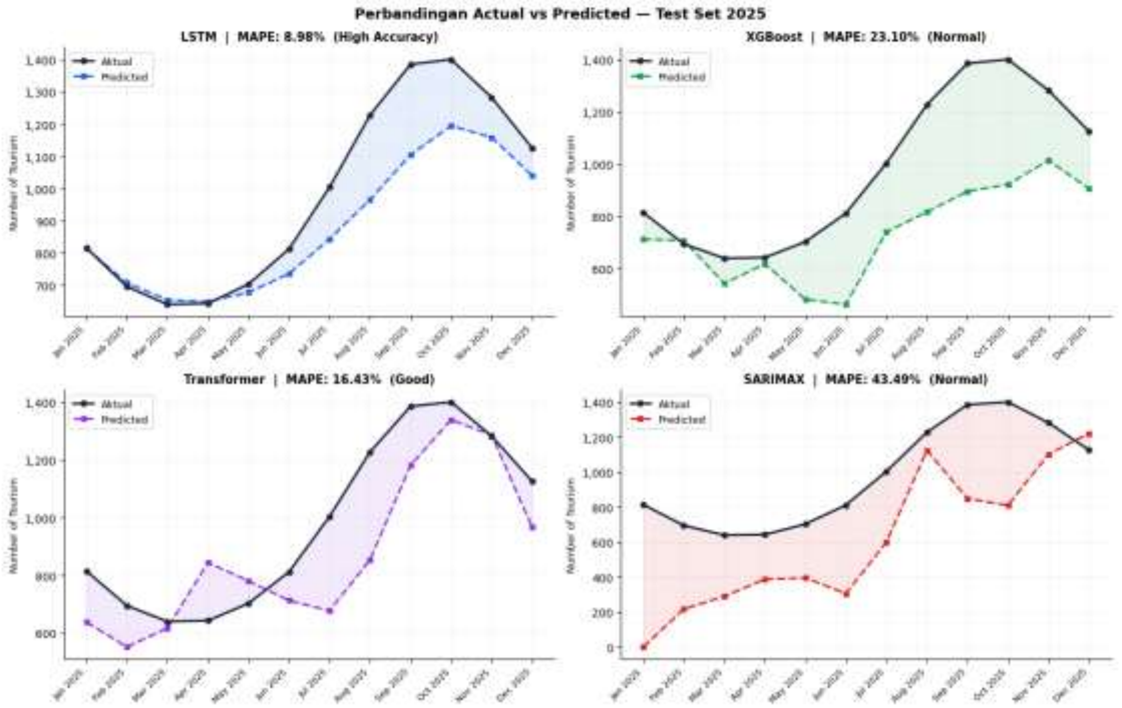

Figure 1. Comparison Model

Table 2. Comparative Model Evaluation Results

Model	MAE	RMSE	MAPE (%)	Category
LSTM	104.65	143.20	8.98	High Accuracy
Transformer	154.35	188.54	16.43	Good
XGBoost	244.64	291.26	23.10	Good
SARIMAX	385.62	436.52	43.49	Inaccurate

3.3. Diebold-Mariano Statistical Test

Table 3 presents the Diebold-Mariano test results comparing LSTM against each baseline. LSTM significantly outperforms XGBoost (DM = −3.54, p = 0.0004) and SARIMAX (DM = −3.10, p = 0.0019) at the 5% significance level. The comparison between LSTM and Transformer yields p = 0.1637, indicating that the performance difference is not statistically significant despite LSTM's numerically lower MAPE. These results confirm that LSTM provides a statistically meaningful improvement over machine learning and statistical baselines for this dataset.

Table 3. Diebold-Mariano Test Results

Comparison	DM Statistic	p-value	Result
LSTM vs XGBoost	−3.5385	0.0004	Significant
LSTM vs Transformer	−1.3928	0.1637	Not Significant
LSTM vs SARIMAX	−3.0978	0.0019	Significant

DOI: https://doi.org/10.31004/riggs.v5i2.8788

3.4. Prediction Results on Test Set (2025)

Table 4 presents the monthly prediction results of the LSTM model on the 2025 test set. The model demonstrates strong accuracy in the first half of the year, with APE values consistently below 10%. A moderate increase in error is observed during the mid-year high season months (July–September), where MAPE reaches approximately 20%, reflecting the inherent difficulty of predicting peak-season surges in post-pandemic recovery conditions. Overall MAPE of 8.98% confirms the model's suitability for practical deployment.

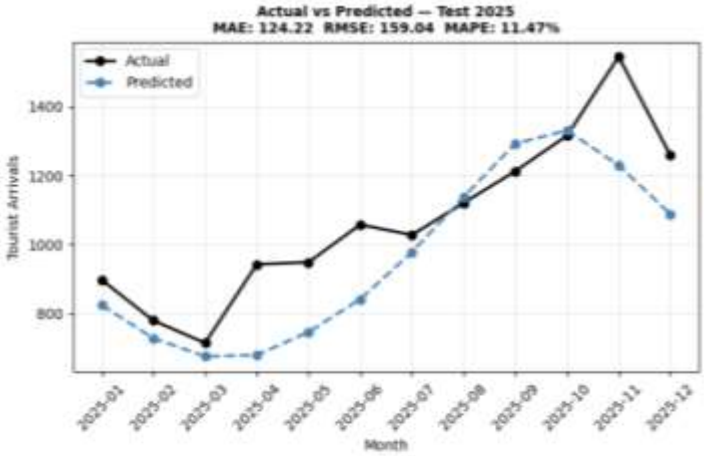

Figure 2. LSTM Monthly Prediction Visualization

Table 4. LSTM Monthly Prediction Results (2025)

Month	Actual	Predicted	APE (%)
January 2025	814	814	0.06
February 2025	695	707	1.75
March 2025	639	654	2.20
April 2025	643	649	0.94
May 2025	702	677	3.68
June 2025	811	735	9.43
July 2025	1,004	843	16.02
August 2025	1,227	965	21.38
September 2025	1,385	1,104	20.29
October 2025	1,401	1,194	14.71
November 2025	1,283	1,158	9.68
December 2025	1,126	1,040	7.62
Average	—	—	8.98

3.5. Variable Influence Analysis

Pearson correlation analysis between each input variable and total tourist arrivals reveals that passenger-related variables exert the strongest influence: Arriving Passengers ($r = +0.627$, $p < 0.001$), Total Passengers ($r = +0.610$, $p < 0.001$), and Departing Passengers ($r = +0.589$, $p < 0.001$). Tourism Events show a moderate positive correlation ($r = +0.468$, $p < 0.001$), while Rainfall ($r = -0.282$) and COVID-19 ($r = -0.297$) exhibit weak but significant negative correlations. Public Holidays and Google Trends show very weak and statistically non-significant correlations, indicating limited influence on monthly tourist arrivals in Wakatobi.

DOI: https://doi.org/10.31004/riggs.v5i2.8788

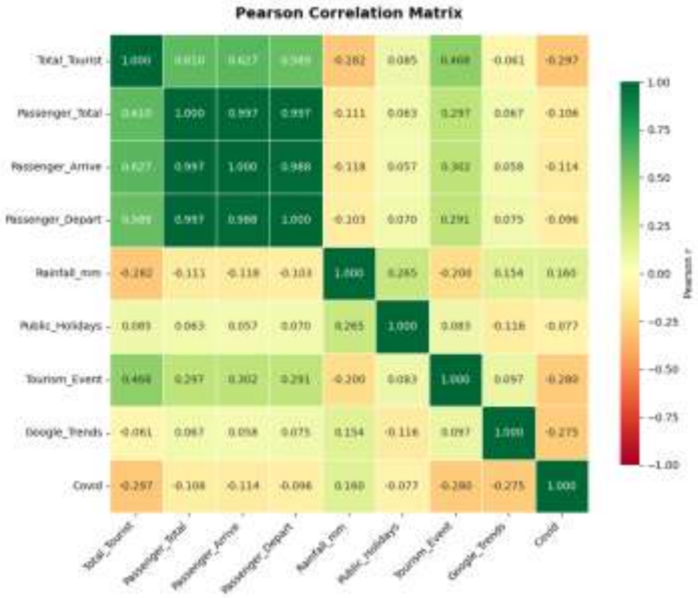

Figure 3. Pearson Correlation Matrix

3.6. System Implementation

The trained LSTM model was deployed in a Flask-based web application designed for use by regional government officials and tourism stakeholders. The system provides five main pages: Dashboard, Prediction (Prediksi), Factor Analysis (Analisis Faktor), Data and Reports, and Model Information. Users can select forecast horizons of 3, 6, or 12 months and view results as prediction tables, line charts comparing historical and forecasted values, and SHAP-based variable influence bar charts. Download functionality for Excel exports and chart images is also provided. Black box testing confirms that all 12 functional test scenarios pass successfully.

Figure 4. Dashboard Page

Figure 5. Forecasting Page

DOI: https://doi.org/10.31004/riggs.v5i2.8788

4. Conclusion

This study demonstrates that the LSTM model is the most effective approach for forecasting monthly tourist arrivals in Wakatobi, achieving a MAPE of 8.98% (High Accuracy) and outperforming SARIMAX, XGBoost, and Transformer models across all evaluation metrics. The Diebold-Mariano test statistically confirms LSTM's superiority over XGBoost (p = 0.0004) and SARIMAX (p = 0.0019). The optimal model configuration, identified through systematic grid search, uses a window size of 6 months, two LSTM layers with 128 and 64 units respectively, and a dropout rate of 0.1. The deployed Flask web application provides an accessible forecasting tool for regional government officials, enabling data-driven tourism planning with interactive forecast horizons and SHAP-based interpretability. Future research may explore hybrid architectures combining LSTM with attention mechanisms, or incorporate additional variables such as social media sentiment and flight schedules to further improve forecasting accuracy.

Reference

[1] D. Purbani et al., "Marine biodiversity and tourism development in Wakatobi," J. Mar. Environ. Eng., vol. 10, pp. 1–15, 2021.
[2] BPS Kab. Wakatobi, Wakatobi Regency in Figures 2023. Kendari: BPS, 2023.
[3] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 3rd ed. Melbourne: OTexts, 2021.
[4] U. M. Sirisha, M. C. Belavagi, and G. Attigeri, "Profit prediction using ARIMA, SARIMA and LSTM models in time series forecasting: A comparison," IEEE Access, vol. 10, pp. 124715–124727, 2022. https://doi.org/10.1109/ACCESS.2022.3224938
[5] M. P. Anisa, H. Irawan, and S. Widiyanesti, "Forecasting demand factors of tourist arrivals in Indonesia's tourism industry using recurrent neural network," IOP Conf. Ser.: Mater. Sci. Eng., vol. 1077, p. 012035, 2021. https://doi.org/10.1088/1757-899X/1077/1/012035
[6] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997. https://doi.org/10.1162/neco.1997.9.8.1735
[7] S.-C. Hsieh, "Tourism demand forecasting based on an LSTM network and its variants," Algorithms, vol. 14, no. 8, p. 243, 2021. https://doi.org/10.3390/a14080243
[8] S. A. T. Taib et al., "The implementation of long-short term memory for tourism industry in Malaysia," J. Adv. Res. Appl. Sci. Eng. Technol., vol. 46, no. 2, pp. 90–97, 2025. https://doi.org/10.37934/araset.46.2.9097
[9] A. Sofia, Z. Dien, and G. Erda, "Forecasting international tourist arrivals to Indonesia using LSTM: Post-pandemic analysis for 2024–2025," J. Math. Comput. Stat., vol. 8, no. 2, pp. 316–324, 2025. https://doi.org/10.35580/jmathcos.v8i2.7309
[10] L. M. G. Jaya, M. Muchtar, Sarimuddin, and S. H. Idrus, "Enhancing accuracy in detection and counting of islands using object-based image analysis: A case study of Kepulauan Seribu, DKI Jakarta," Indones. J. Geogr., vol. 56, no. 2, pp. 169–179, 2024. https://doi.org/10.22146/ijg.82037
[11] H. Wulandari, Statiswaty, I. J. Effendi, and Sutardi, "Sistem monitoring pintu pembendung air otomatis menggunakan algoritma fuzzy berbasis prototype," ANIMATOR, vol. 1, no. 2, pp. 86–92, 2023.
[12] C. D. Lewis, Industrial and Business Forecasting Methods. London: Butterworth-Heinemann, 1982.
[13] S. Mahjoub et al., "Predicting energy consumption using LSTM, multi-layer GRU and drop-GRU neural networks," Sensors, vol. 22, no. 11, p. 4062, 2022. https://doi.org/10.3390/s22114062
[14] UNWTO, International Tourism Highlights 2023. Madrid: UNWTO, 2023.
[15] Al Dilwan et al., "Tourism development and marine conservation in Wakatobi," Sustainability, vol. 15, 2023.

DOI: https://doi.org/10.31004/riggs.v5i2.8788