



Forecasting The Number of Tourist Trips in Pangkal Pinang City: ARIMA vs. LSTM vs. PROPHET

Muhammad Irfan Hidayat^{1*}, Mei Dita Kumala²

¹Department of Mathematics, Universitas Bangka Belitung, Indonesia

²Bangka Belitung, Indonesia

* Corresponding author: irfan-hidayat@ubb.ac.id

A B S T R A C T

Accurate forecasting of tourism demand is crucial for regional economic planning, yet existing studies often rely on univariate time-series models or single forecasting techniques without considering spatial interdependence among regions. This study proposes a correlation-based comparative forecasting framework to evaluate the performance of ARIMA, Long Short-Term Memory (LSTM), and Prophet in predicting domestic tourist trips in Pangkal Pinang City, Indonesia. Monthly data from January 2019 to July 2025 were obtained from Statistics Indonesia, and highly correlated neighboring regions were systematically selected using a heatmap correlation analysis to enhance model input relevance. After applying Min–Max normalization and an 80:20 train–test split, all models were evaluated using Root Mean Squared Error (RMSE) under identical experimental settings. The results indicate that Prophet consistently achieves the lowest RMSE, demonstrating superior capability in capturing non-linear dynamics, seasonal variability, and abrupt structural changes in tourism demand compared to ARIMA and LSTM. These findings provide empirical evidence that decomposable time-series models with automatic trend and seasonality handling offer distinct advantages over both classical statistical and deep learning approaches in medium-term tourism forecasting. The proposed framework contributes a concise, data-efficient, and replicable methodology that supports evidence-based tourism planning and strategic decision-making.

Keywords: Forecasting, Long Short-Term Memory (LSTM), ARIMA, Prophet, RMSE.

Received : 24-11-2025;

Revised : 02-02-2026;

Accepted : 11-03-2026;

Published : 13-03-2026;

DOI: <https://doi.org/10.29303/emj.v9i1.345>



This work is licensed under a [CC BY-NC-SA 4.0 International](https://creativecommons.org/licenses/by-nc-sa/4.0/) license

1. Introduction

Domestic tourist trips refers to the movement of Indonesian residents who visit destinations within Indonesia for leisure or tourism purposes. One of the main destinations for domestic tourists is the Bangka Belitung Islands Province. As an island-based region that depends on services and tourism, Bangka Belitung faces challenges in managing the large and changing number of visitors. Data from 2025 show strong fluctuations in domestic tourist trips, with the highest number recorded in April, over half a million trips, while other months saw a clear decline. This unstable pattern shows the need for

a forecasting system that can carefully predict changes in trends and respond to external factors more effectively. Numerous researchers have extensively investigated the prediction and modeling of tourist movements as well as travel behavior patterns in the Bangka Belitung Islands. These studies have employed various analytical approaches and methodologies to understand tourist mobility, destination preferences, and spatial-temporal travel patterns within the region, as in [1], [2], and [3].

Choosing the right forecasting method can be quite challenging. Accuracy and small error values play an important role in determining the most effective approach. Among the time series methods often used in forecasting are ARIMA and LSTM. Traditionally, ARIMA has been widely applied in time series studies, especially in economic and business analysis. However, because tourism data are usually non-linear, diverse, and affected by seasonal changes, the ARIMA model often performs less effectively. With the advancement of deep learning techniques, the LSTM model was developed to handle these limitations by learning long-term dependencies and recognizing complex patterns that traditional linear models cannot capture, for example, in 2018, Namin et al. conducted a study comparing the two methods for forecasting economic and financial data [4]. The study concluded that the LSTM method performed significantly better than the ARIMA method when applied to economic and financial datasets. In the present research, the author aims to perform forecasting on the number of tourists in Pangkal Pinang City, as the number of visitors directly affects the city's economic activities. Since tourism has a substantial impact on the economy, this study seeks to determine whether these two methods exhibit similar behavior when implemented on tourism data.

Meanwhile, in 2017, Meta released a new forecasting method known as the Prophet algorithm. This method is claimed to provide a more intuitive and efficient approach, capable of handling trend patterns, seasonality, and holidays without requiring complex mathematical understanding. Previously, Angelo et al. (2023) published a paper comparing the ARIMA method and the Prophet algorithm for Bitcoin price forecasting [5]. By comparing these three forecasting models, ARIMA, LSTM, and Prophet, this study aims to identify the most suitable model for predicting the number of domestic tourist trips in the Bangka Belitung Islands Province. The findings of this research are expected to contribute to the development of data-driven tourism forecasting approaches that not only improve estimation accuracy but also strengthen evidence-based strategic planning for sustainable tourism development.

2. Research Methods

The forecasting process in this research consists of several stages, which are illustrated in the research flowchart presented in Figure 1.

The process flow in this study consists of several processes, including:

1. Collecting data from Statistics Indonesia Bangka Belitung Province website.
2. Selection of relevant data using a correlation heatmap to identify the optimal parameters for model development.
3. Scaling data using Min-Max scaling also known as data normalization.
4. Splitting data to 80% and 20% data train-data test ratio.
5. Train the data to obtain the model.
6. Compare the predicted data that obtained by model with the real data (data test).
7. Plotting real data (data test) and predicted data.
8. Evaluating error using Root Mean Squared Error (RMSE).

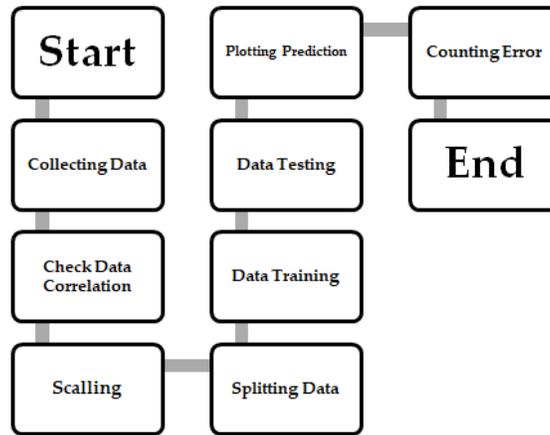


Figure 1. Flowchart

2.1. Heatmap Correlation

The correlation between features in the dataset is analyzed using a heatmap diagram. A heatmap is a tool for visualizing a correlation matrix, illustrating the strength and direction of linear relationships among the numerical variables in the dataset [6].

2.2. Min-Max Scaling

Min-Max scaling is a type of scaling that transform the original data linearly. The resulting scale ranges from 0 to 1. The formula for Min-Max scaling is shown in Equation 1. Where x_i represents the data being scaled, x' denotes the scaled result, $\min(x)$ is the minimum value of a feature, and $\max(x)$ is the maximum value of a feature [7]. Min-Max scaling can be expressed as follows

$$x' = \frac{x_i - \min(x)}{\min(x) - \max(x)}. \quad (1)$$

2.3. ARMA

The Autoregressive Moving Average (ARMA) model is one of the most widely used and popular approaches in time series modeling. ARMA combines the Autoregressive (AR) component, which assumes that the current value is influenced by its past values, and the Moving Average (MA) component, which assumes that the current value is influenced by past residuals. A time series itself refers to a sequence of observations recorded in chronological order. The ARMA model was introduced by George Edward Pelham Box and Gwilym Meirion Jenkins in 1976. Consequently, ARMA modeling is often referred to as the Box-Jenkins time series method. The general form of the Autoregressive model, denoted as AR (p), is expressed as follows

$$y_t = b_0 + b_1y_{t-1} + b_2y_{t-2} + \dots + b_ny_{t-n} + \epsilon_t. \quad (2)$$

The Model Moving Average (MA), represented by MA(q), can be formulated as

$$y_t = a_0 - a_1\epsilon_{t-1} + a_2\epsilon_{t-2} + \dots + a_n\epsilon_{t-n} + \epsilon_t \quad (3)$$

where :

y_t : the data value at time period- t

$y_{t-1}, y_{t-2}, \dots, y_{t-n}$: the data value at the previous time period
 ϵ_t : error at time period-t
 $\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-n}$: error at the previous time period
 $b_0, b_1, b_2, \dots, b_n$: coefficient constan AR
 $a_0, a_1, a_2, \dots, a_n$: coefficient constan MA [8].

2.4. ARIMA

If the data are not stationary, differencing must be applied until stationarity is achieved. The ARIMA model extends the ARMA framework by incorporating a differencing component (d). This model, expressed as ARIMA (p, d, q), is widely used in tourism research to forecast various indicators such as tourist arrivals, tourism demand, and destination visitation patterns. In general, the ARIMA model is expressed as

$$\phi_p(B)(1 - B)^d Z_t = \theta_q(B)\epsilon_t \tag{4}$$

where :

- p : autoregressive order
- d : differencing order
- q : moving average operator order
- Z_t : time series data at time- t
- B : backshift operator
- ϕ, θ : AR and MA parameters [9].

The ARIMA method is advantageous due to its flexibility in managing non-stationary data and its ability to produce highly accurate short-term forecasts.

2.5. Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) is a type of machine learning algorithm based on processing information from data and is a form of deep learning algorithm. RNN retains patterns of past information by performing iterations within its architecture, allowing this information to be preserved. RNN is an extension of Artificial Neural Networks (ANN) and its architecture closely resembles that of a Multilayer Perceptron (MLP) [10].

Recurrent Neural Networks (RNN) fundamentally share the same parameters as standard neural networks, with the primary difference being the underlying concept of RNNs. In an RNN, the hidden layer from the previous timestep is used as input for the next process. This hidden layer retains information from the initial feature extraction stage, so when it is utilized as input for the subsequent stage, it stores the memory of the previous input [10].

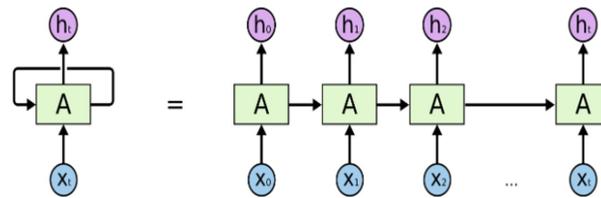


Figure 2. Recurrent Neural Network (RNN) Architecture

Figure 2 illustrates the RNN process both before and after it is unrolled into a complete network. X_t is the input at time t and h_t is the output from adalah output RNN process. Once unrolled, the full network is visible in its entirety. The RNN process starts by taking X_0 as input, generating h_0 and simultaneously receiving X_1 as input for the next step. Hence, h_0 and X_1 are used as inputs for the

following step, and this sequence continues. This mechanism allows the RNN to remember context during training. The RNN performs the same task iteratively, with its output being dependent on prior computations [10].

2.6. Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) represents an advanced form of the Recurrent Neural Network (RNN) architecture, first introduced by Hochreiter and Schmidhuber. Compared to a conventional RNN, LSTM offers significant advantages through its memory and forget mechanisms, which allow outputs to be fed back as inputs and enable more accurate information processing. Moreover, LSTM has the ability to preserve error signals during backpropagation, preventing them from compounding over time [10].

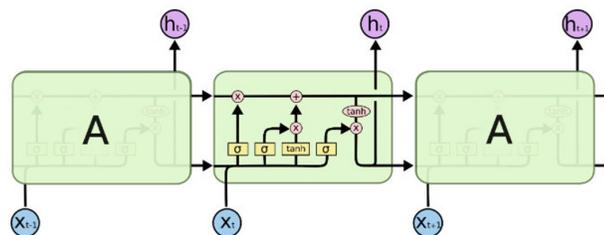


Figure 3. Long Short-Term Memory (LSTM) Architecture

As shown in Figure 3, the general architecture of Long Short-Term Memory (LSTM), which consists of three gates: the forget gate, the input gate, and the output gate. The forget gate determines which information should be discarded using a sigmoid activation function. An output of 1 indicates that all data are retained, while an output of 0 means that all data are discarded [10].

2.7. Prophet

The Prophet model, introduced by Taylor and Letham (2017) at Facebook, represents a contemporary forecasting framework that emphasizes adaptability and practical usability. It is based on a decomposable time-series structure in which observed values are expressed through the integration of three fundamental components: a long-term trend, recurrent seasonal fluctuations, and the influence of holidays or exceptional events. This structure is formalized as

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t \quad (5)$$

where $g(t)$ captures systematic trend movements, $s(t)$ reflects periodic seasonal patterns typically modeled with Fourier expansions, and $h(t)$ accounts for event-driven variations, while ϵ_t represents irregular disturbances. Such a formulation allows Prophet to effectively address nonlinear patterns and shifting temporal dynamics, making it particularly suitable for complex and highly variable time-series data [11].

Prophet's primary strength lies in its capacity to streamline the forecasting process by automating several steps that often require advanced statistical proficiency. The model incorporates mechanisms for identifying changepoints, adjusting for outliers, and estimating multiple seasonalities without substantial manual parameter tuning. Additionally, Prophet integrates robust evaluation procedures, including cross-validation and performance measures such as RMSE, MSE, MAE, MAPE, and coverage, enabling comprehensive assessment of predictive accuracy. Its interpretability, flexibility in incorporating custom seasonal and holiday effects, and straightforward implementation in both R and Python have contributed to its widespread application in empirical forecasting studies. As a result,

Prophet has emerged as a reliable and scalable approach for modeling time series characterized by pronounced trends and seasonal structures [11].

2.8. Root Mean Squared Error (RMSE)

The Root Mean Squared Error (RMSE) value is used to measure the extent of deviation between predicted and actual values. This is the formula of RMSE,

$$RMSE = \sqrt{\frac{\sum_{t=1}^N (X_t - Y_t)^2}{N}}, \quad (6)$$

where X_t denotes the actual value at time- t , Y_t is the predicted value at time- t , and N refers to the total number of predicted data [12].

3. Result and Discussion

In this section, the author presents the results of predicting the number of domestic tourist trips in Pangkal Pinang City, Bangka Belitung Province, using the ARIMA method, Long Short-Term Memory (LSTM) method, and Prophet's Algorithm. The data employed in this study were obtained from the Statistics Indonesia, specifically the dataset on the number of domestic tourist trips by destination City. The dataset consists of monthly records of domestic tourist trips in the Bangka Belitung Islands Province from January 2019 to July 2025. The total number of domestic tourists in the province is illustrated in Figure 4.

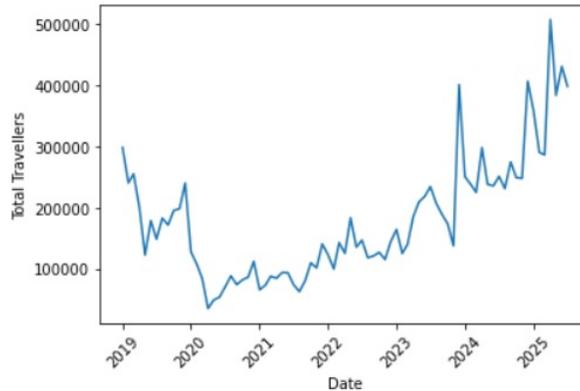


Figure 4. The total number of tourists in Bangka Belitung Province

Since the prediction is focused on Pangkal Pinang City, the data on domestic tourist trips in Pangkal Pinang City are presented in Figure 5.

Upon completion of the dataset collection, this study does not utilize data from all districts/cities in the Bangka Belitung Islands Province, but only from those with a strong correlation to Pangkal Pinang City. The method employed to determine the correlation analysis is the heatmap correlation. The results of this analysis are presented in Figure 6.

In the heatmap, color gradients illustrate both the magnitude and direction of correlations, whereas the numerical values within each cell denote the correlation coefficients, ranging from -1 to $+1$. Based on Figure 6, it can be observed that the larger the coefficient value for a given regency, the stronger its correlation with Pangkal Pinang City. For this reason, only regencies with correlation coefficients greater than or equal to 0.9 were selected, namely Bangka, West Bangka, and Central Bangka.

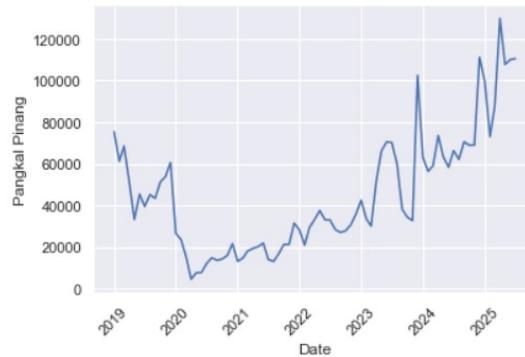


Figure 5. The total number of tourists in Pangkal Pinang City

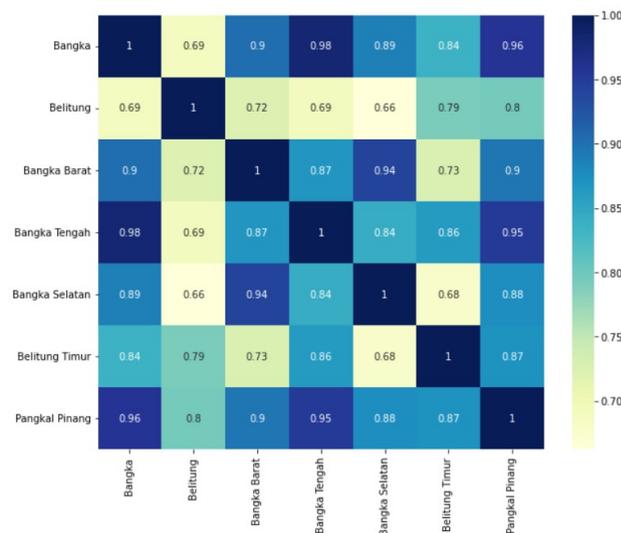


Figure 6. Correlation heatmap of the number of tourists in the cities of Bangka Belitung Province

After selecting the data to be used as input, this study applies data scaling, also known as normalization. This procedure is performed to accelerate the convergence of the algorithm and to improve model accuracy. Next, the dataset was split into two subsets: 80% for training and 20% for testing.

The splitted data were then processed in the training stage. The model training process was iterated according to the predetermined number of epochs. After several experimental trials, the optimal training configuration was found at 500 epochs. Based on the training results, it can be observed that the LSTM method is capable of handling a large number of epochs as well as a large number of samples. After that, the predicted would be obtained automatically using python command as below.

```
predictions = model.predict_generator(test_generator)
```

In the other hand the command to forecast in ARIMA using python as below.

```
model = ARIMA(history, order=order)
model_fit = model.fit(dispatch=-1)
predictions_f_ms = model_fit.forecast(steps=len(test))[0]
predictions_p_ms = model_fit.predict(start=len(history),
```

```
end=len(history)+len(test)-1)
```

The command to forecast in Prophet using python as below.

```
my_model = Prophet(interval_width=0.95)
forecast = my_model.predict(future_dates)
```

The results of predicted data, for example the model of LSTM are presented in Figure 7.

	Bangka	Bangka Barat	Bangka Tengah	Pangkal Pinang_pred	Pangkal Pinang
0	56903.0	26133.0	43590.0	60927.421875	63355.0
1	57921.0	28344.0	44800.0	58514.933594	58380.0
2	60783.0	27625.0	46786.0	59100.882812	66339.0
3	54439.0	26240.0	42302.0	59527.019531	62159.0
4	61379.0	28646.0	47151.0	58034.085938	70547.0
5	53274.0	24266.0	45001.0	59716.996094	68942.0
6	53255.0	24452.0	44071.0	58186.507812	69021.0
7	87225.0	39059.0	74673.0	58016.710938	111175.0
8	79110.0	31627.0	64939.0	61170.460938	99639.0
9	61523.0	37910.0	50927.0	61339.007812	73069.0
10	59511.0	29005.0	48628.0	60761.335938	87512.0
11	118667.0	46948.0	77132.0	59827.792969	129795.0
12	87993.0	34694.0	65108.0	60570.351562	107712.0
13	102366.0	43337.0	74793.0	61347.273438	110021.0
14	90662.0	38066.0	68588.0	60927.035156	110486.0

Figure 7. Predicted data

When displayed in graphical form, the comparison can be observed in Figure 8.

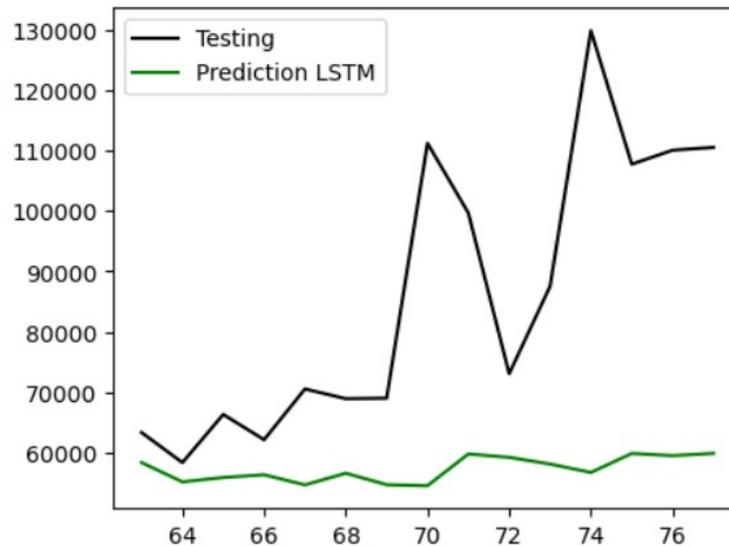


Figure 8. Comparison between predicted data of LSTM and real data

In Figure 8, the black line represents the actual data, while the green line represents the predicted data of LSTM. After evaluating the error using the RMSE method, an error value of 35929.32 was obtained. While using batch=3 and epoch=500 as the best parameters, in ARIMA and Prophet, author

use the confidential rate of 95%. We get the value of RMSE in ARIMA and Prophet consecutively 27720.44 and 20509.83. The RMSE comparison clearly indicates that Prophet achieves the lowest forecasting error (20509.83), followed by ARIMA (27720.44) and LSTM (35929.32). The superior performance of Prophet can be attributed to its decomposable structure, which explicitly models trend, seasonality, and potential structural shifts. Tourism demand in Pangkal Pinang exhibits strong seasonal fluctuations and abrupt changes, particularly during peak holiday periods and post-pandemic recovery phases. Prophet's automatic changepoint detection enables the model to adapt to these structural variations more effectively than ARIMA, which relies on linear differencing assumptions, and LSTM, which requires large datasets to generalize complex nonlinear patterns reliably. The plot to compare the result of ARIMA and Prophet is in Figure 9 and Figure 10 below.

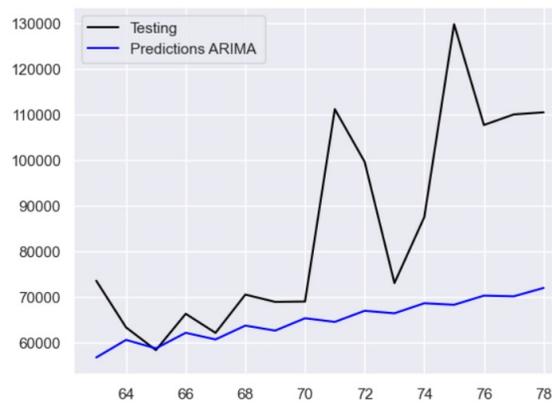


Figure 9. Comparison between predicted data of ARIMA and real data

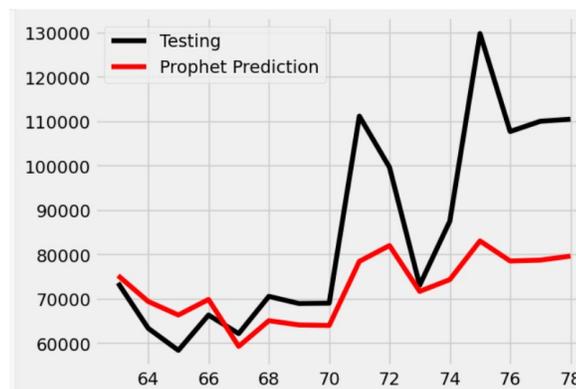


Figure 10. Comparison between predicted data of Prophet and real data

In Figure 9 and Figure 10, the black line represents the actual data, while the blue line represents the predicted data of ARIMA, and the red line represents the predicted data of Prophet. A closer inspection of Figure 8 reveals that the LSTM predictions tend to underestimate sharp increases in tourist arrivals and produce smoother trajectories compared to the actual data. This behavior suggests that, despite its theoretical capacity to capture nonlinear dependencies, the LSTM model may have been constrained by the relatively limited sample size (monthly data from 2019–2025). Deep learning models generally require large-scale datasets to achieve stable generalization; therefore, in this context, LSTM may have exhibited underfitting toward extreme seasonal peaks.

In contrast, ARIMA demonstrates improved alignment with the overall upward trend but still struggles to fully capture sudden spikes and irregular seasonal amplitudes. This limitation arises from its linear structure and reliance on differencing to achieve stationarity, which may suppress important nonlinear dynamics inherent in tourism flows. Prophet, as illustrated in Figure 10, shows a closer alignment with both the seasonal cycles and abrupt increases observed in the test data. Its additive decomposition framework allows flexible trend adjustments through automatically selected change-points, while Fourier-based seasonal components capture periodic patterns more explicitly. Consequently, Prophet achieves a more balanced bias–variance tradeoff, explaining its lower RMSE value relative to ARIMA and LSTM.

Overall, these findings suggest that for medium-term tourism forecasting characterized by moderate sample size, pronounced seasonality, and structural shifts, decomposable time-series models such as Prophet may provide a more robust and data-efficient alternative compared to purely statistical linear models or data-intensive deep learning architectures.

4. Conclusions

This study presents a comparative evaluation of ARIMA, Long Short-Term Memory (LSTM), and Prophet models for forecasting monthly domestic tourist trips in Pangkal Pinang City using data spanning January 2019 to July 2025. By integrating a correlation-based spatial feature selection process, only regions exhibiting strong relationships with Pangkal Pinang were incorporated as model inputs, thereby enhancing the relevance and robustness of the forecasting framework. The empirical results indicate that the Prophet model consistently outperforms ARIMA and LSTM in terms of RMSE, demonstrating superior capability in capturing seasonal patterns and abrupt trend changes that characterize tourism demand dynamics.

From a practical perspective, these findings highlight the suitability of decomposable time-series models for tourism forecasting in regions experiencing high volatility and structural shifts. The proposed approach offers a data-driven basis for policymakers and tourism stakeholders to improve planning accuracy, optimize resource allocation, and support strategic decision-making at the city level.

Despite its contributions, this study is subject to several limitations. The analysis relies exclusively on monthly domestic tourist data and does not account for exogenous factors such as economic conditions, policy interventions, or transportation accessibility. Moreover, forecasting performance was assessed using a single evaluation metric, which may not fully reflect all dimensions of predictive accuracy.

Future research is encouraged to incorporate external variables, explore hybrid or ensemble modeling strategies, and extend the analysis to higher-frequency data or different tourism destinations in order to further validate and generalize the proposed framework.

REFERENCES

- [1] E. Marcelina, T. Agustin, K. Luthfiyaturrohman, J. Octaviani, A. Pramita, I. Monika, D. Y. Dalimunthe, and A. Nasrun, “Peramalan jumlah wisatawan kabupaten belitung menggunakan simulasi monte carlo,” *Euler: Jurnal Ilmiah Matematika, Sains, dan Teknologi*, vol. 12, pp. 57–63, 2024. <https://doi.org/10.37905/euler.v12i1.25153>.
- [2] D. D. Silalahi and N. Mindiyarti, “Understanding tourism behaviour through exploratory data analysis with machine learning on search engine data: Case study in bangka belitung islands, indonesia,” *IJAIR: International Journal of Artificial Intelligence Research*, vol. 8, no. 1.1, 2025. <https://doi.org/10.29099/ijair.v8i1.1.1372>.

- [3] M. I. Hidayat and M. D. Kumala, "Analisis peramalan jumlah kunjungan wisatawan domestik menggunakan model long short-term memory (lstm) di kota pangkal pinang," *Proximal: Jurnal Penelitian Matematika dan Pendidikan Matematika*, vol. 9, no. 1, pp. 1–12, 2026. <https://doi.org/10.30605/proximal.v9i1.7478>.
- [4] S. S. Namin, N. Tavakoli, and A. S. Namin, "A comparison of arima and lstm in forecasting time series," *IEEE*, pp. 1394–1401, 2018. <https://ieeexplore.ieee.org/abstract/document/8614252>.
- [5] M. D. Angelo, I. Fadhiilrahman, and Y. Purnama, "Comparative analysis of arima and prophet algorithms in bitcoin price forecasting," *Procedia computer science* 227, pp. 490–499, 2023. <https://www.sciencedirect.com/science/article/pii/S1877050923017179>.
- [6] A. Syaifudin, R. Risqiati, and H. W. Hapsoro, "Implementasi exploratory data analysis untuk analisis data lemak tubuh," *IC Tech: Majalah Ilmiah*, vol. 20, no. 1, pp. 1–10, 2025. <https://doi.org/10.47775/icttech.v20i1.328>.
- [7] P. P. Allorerung, A. Erna, M. Bagussahrir, and S. Alam, "Analisis performa normalisasi data untuk klasifikasi k-nearest neighbor pada dataset penyakit," *JISKa: Jurnal Informatika dan Sunan Kalijaga*, vol. 9, no. 3, p. 178–191, 2024. <https://doi.org/10.14421/jiska.2024.9.3.178-191>.
- [8] C. A. Melyani, A. Nurtsabita, G. Z. Shafa, and E. Widodo, "Peramalan inflasi di indonesia menggunakan metode autoregressive moving average (arma)," *JAMES: Journal of Mathematics Education and Sciences*, vol. 4, no. 2, p. 67–74, 2021. <https://doi.org/10.32665/james.v4i2.231>.
- [9] N. T. Qurniawan and T. Sukmono, "Peramalan permintaan dengan menerapkan metode autoregressive integrated moving average (arima) pada industri beton," *Jurnal Teknologi dan Manajemen Industri Terapan (JTMIT)*, vol. 4, no. 3, pp. 1024 – 1032, 2025. <https://doi.org/10.55826/jtmit.v4i3.1117>.
- [10] S. S. Nurashila, F. Hamami, and T. F. Kusumasari, "Perbandingan kinerja algoritma recurrent neural network (rnn) dan long short-term memory (lstm): Studi kasus prediksi kemacetan lalu lintas jaringan pt xyz," *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 8, no. 3, p. 864–877, 2023. <https://doi.org/10.29100/jupi.v8i3.3961>.
- [11] M. Navratil and A. Kolkova, "Decomposition and forecasting time series in business economy using prophet forecasting model," *CENTRAL EUROPEAN BUSINESS REVIEW*, vol. 8, no. 4, pp. 26–39, 2019. <https://doi.org/10.18267/j.cebr.221>.
- [12] U. Khaira, M. Alfalah, P. C. S. Gulo, and R. Purnomo, "Prediksi kemunculan titik panas di lahan gambut provinsi riau menggunakan long short term memory," *Jurnal Informatika Jurnal Pengembangan IT*, vol. 5, no. 3, pp. 77–82, 2020. <https://doi.org/10.30591/jpit.v5i3.1931>.