Vol. 8, No. 2, April 2026

ISI Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

Data-Driven Traffic for Infrastructure Planning: An LSTM Approach Using Indonesian Road-Vehicle Trends

Aria Hendrawan[1], **Nabilah Putri Wijaya**[2]

[1,2]Department of Information Technology and Communication, Universitas Semarang, Semarang, Indonesia

Received:
October 11, 2026
Revised:
March 3, 2026
Accepted:
March 26, 2026
Published:
April 12, 2026

Corresponding Author:
Author Name*:
Aria Hendrawan
Email*:
ariahendrawan@usm.ac.id

DOI:
10.63158/journalisi.v8i2.1516

Abstract. The rapid growth of motorized vehicles in Indonesia, unmatched by proportional expansion in road infrastructure, has intensified pressure on the national transportation system. This study examines the application of a Long Short-Term Memory (LSTM) model to analyze and forecast the national traffic load ratio, defined as the ratio of total motorized vehicles to total road length. Annual aggregate data from the Indonesian Central Bureau of Statistics (BPS) for the period 2016–2023 were used in the analysis. The results indicate that the model achieved a strong fit on the training data, with RMSE = 0.3652 and MAE = 0.3617, but performed substantially worse on the test data, with RMSE = 1.7585 and MAE = 1.7585. This discrepancy suggests overfitting, largely attributable to the extremely limited sample size. As such, the findings should be interpreted as exploratory rather than as evidence of reliable forecasting performance. Despite these limitations, the model projects a continued upward trend in national infrastructure pressure over the next five years. These findings provide an initial data-driven indication that transportation infrastructure demand in Indonesia is likely to intensify, while also underscoring the need for future research using larger datasets and baseline model comparisons before policy-level application can be justified.

Keywords: Exploratory forecasting, Infrastructure planning, LSTM, Time series forecasting, Traffic load ratio, Transport infrastructure

1. INTRODUCTION

Indonesia has experienced rapid growth in motorized vehicle ownership over the past decade, while road infrastructure expansion has progressed at a slower pace. This imbalance has intensified pressure on the transportation system and contributed to congestion, travel inefficiency, environmental burdens, and broader infrastructure strain [1]-[3]. At the macro level, these conditions indicate that transportation planning must increasingly rely on measurable indicators that can capture how strongly vehicle growth is outpacing road development. A useful way to represent this infrastructure pressure is through the traffic load ratio, defined as the proportion of total motorized vehicles to total road length. This ratio offers a compact national-level indicator of how heavily the available road network is being burdened. When observed over time, it can provide an interpretable basis for evaluating whether infrastructure development is keeping pace with mobility demand [4]-[6].

In countries with rapid urbanization and continuing motorization growth such as Indonesia, this type of indicator can support longer-term infrastructure planning and policy reflection. Previous traffic forecasting studies have often focused on high-frequency urban traffic flow data, such as hourly or daily observations from sensors, cameras, or specific corridors [7]-[9]. In such settings, deep learning models, including Long Short-Term Memory (LSTM) networks, have shown strong capacity to model temporal dynamics [10]-[12]. However, there remains a gap in the literature regarding the use of LSTM for macro-level national traffic load forecasting based on annual aggregate data [13]-[15]. This is an important distinction because national planning problems are structurally different from short-term operational traffic forecasting: they rely on longer-horizon indicators, far fewer observations, and broader infrastructure interpretations [16], [17].

Accordingly, this study formulates the national traffic load ratio as a time-series forecasting problem using annual aggregate data from Indonesia for the period 2016–2023. The study does not claim predictive superiority over simpler forecasting approaches. Rather, its contribution lies in two aspects: first, in formalizing the national traffic load ratio as a macro-level indicator of infrastructure pressure; and second, in examining, in an explicitly exploratory manner, whether an LSTM framework can capture its annual trend under severe data limitations. This positioning is important because the available dataset contains only eight annual observations, which substantially constrains generalization and requires cautious interpretation. Therefore, the objective of this research is to explore the feasibility of using LSTM to model and project the national traffic load ratio in Indonesia, while transparently acknowledging that the resulting forecasts should be interpreted as exploratory signals rather than as definitive policy-grade predictions. In this sense, the study is intended to provide an initial data-driven reference for infrastructure planning and to motivate future work using larger datasets and direct baseline comparisons.

2. METHODS

2.1. Research Design and Workflow

This study adopted an exploratory quantitative forecasting design to examine whether an LSTM model could capture the annual trajectory of Indonesia's national traffic load ratio. The workflow consisted of five stages: (1) data collection, (2) data preprocessing and national aggregation, (3) sequence generation, (4) LSTM model development, and (5) evaluation and exploratory future forecasting as shown in Figure 1.

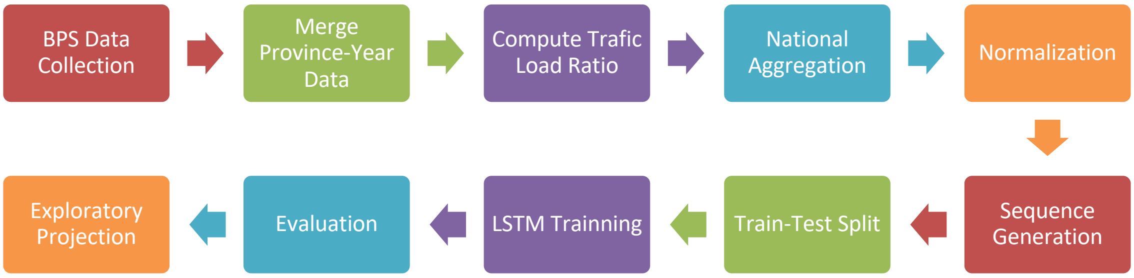

Figure 1. Research Design and Workflow

This structure follows the same core workflow already used in the manuscript, but the methodological positioning in the revised version is made more explicit: the present study should be interpreted as exploratory proof-of-concept work, not as a definitive comparative forecasting study.

2.2. Data Source

The data were obtained from the Indonesian Central Bureau of Statistics (BPS), which served as the official source for both vehicle and road infrastructure information [18]. Two datasets were used. The first dataset contained the annual number of motorized vehicles in each Indonesian province for the period 2016–2023, including passenger cars, buses, trucks, and motorcycles. The second dataset contained annual road length data for each province over the same period, categorized by road status. These two datasets were merged by *Province* and *Year* to construct a unified dataset containing vehicle counts, road length, and derived traffic load ratio values. A sample of the merged dataset is shown in Table 1.

Table 1. Data Head of Merged Dataset

Province	Year	Total Number of Vehicles	Total Road Length (km)	Traffic Load Ratio
Aceh	2016	1,800,552	22,810	78.94
North Sumatra	2016	5,706,756	40,773	139.96
West Sumatra	2016	2,021,827	23,434	86.28
Riau	2016	3,097,414	27,040	114.55
Jambi	2016	1,900,208	13,731	138.39

2.3. Traffic Load Ratio Formulation and National Aggregation

The traffic load ratio for each province-year observation was defined as the number of motorized vehicles divided by total road length in kilometers. At the national level, the ratio was computed from annual aggregate totals as follows Equation 1.

$$T\,L\,R_t = \frac{V_t}{R_t} \tag{1}$$

where TLR_t is the national traffic load ratio in year t, V_t is the total number of motorized vehicles in Indonesia in year t, and R_t is the total road length in Indonesia in kilometers in year t. This formulation yields the average number of vehicles per kilometer of road and serves as a macro-level indicator of infrastructure pressure [19]. After aggregating all provinces by year, the final national time series consisted of eight annual observations covering 2016 to 2023.

Table 2. Data Head of National Aggregate Data

Year	Total Number of Vehicles (National)	Total Road Length (National) (km)	National Traffic Load Ratio
2016	112,205,711	537,837	208.62
2017	118,922,708	540,490	220.03
2018	126,702,280	540,252	234.52
2019	133,617,012	542,160	246.45
2020	136,137,735	545,155	249.72

2.4. Data Normalization

To prepare the annual traffic load ratio series for neural-network training, the values were normalized using MinMax scaling to the range 0,1 [20]. Normalization was applied to improve numerical stability and to prevent the magnitude of the original values from dominating the training process [21]. The normalized series was then used for sequence generation and model input.

2.5. Sequence Preparation and Data Splitting

A fixed-length sliding-window approach was used to transform the annual time series into a supervised learning format [22], [23]. Each input sequence consisted of five consecutive yearly observations, and the target output was the traffic load ratio for the following year. With eight annual observations in total, this procedure produced exactly three supervised sequences:

Table 3. Sequence Generation from Annual Traffic Load Ratio Data

Sequence Id	Input (Years)	Output (Prediction Year)
Seq 1	[2016, 2017, 2018, 2019, 2020]	2021
Seq 2	[2017, 2018, 2019, 2020, 2021]	2022
Seq 3	[2018, 2019, 2020, 2021, 2022]	2023

Because the data volume was extremely limited, this study explicitly treats the LSTM modeling exercise as exploratory. With only eight annual observations and three derived sequences, any deep learning model is at substantial risk of overfitting. Under such constraints, simpler baselines such as linear regression or ARIMA would typically be preferred for formal comparative forecasting. However, LSTM was intentionally retained here to test whether it could still capture the non-linear macro-trend of infrastructure pressure in a proof-of-concept setting [24], [25]. The generated sequences were divided into training and testing subsets using a minimal split strategy consistent with the limited sample size. Given the very small number of sequences, the testing subset contained only one sequence, while the remaining sequences were used for model training.

2.6. LSTM Model Development

The final model used in this study was a sequential neural network composed of one LSTM layer with 32 hidden units, followed by a Dense output layer with one neuron as shown in Figure 2. This configuration was selected from several tested alternatives because it provided the most stable behavior within the available experimental setting [26]-[28]. The selected architecture contained 4,513 trainable parameters in total. The model was trained using the Adam optimizer, Mean Squared Error (MSE) as the loss function, and Mean Absolute Error (MAE) as an additional metric. Training was conducted for 100 epochs with a batch size of 1 [29], [30].

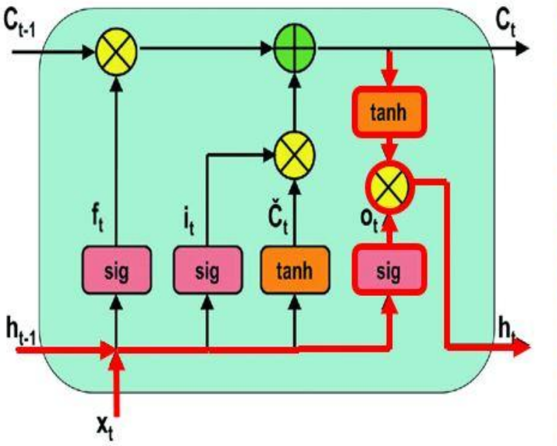

Figure 2. Architecture of LSTM Model

To improve readability and reproducibility, the core LSTM equations are presented below.

1) Forget Gate: Determining how much information from the previous state cell should be forgotten is shown in Equation 2.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{2}$$

2) Input Gate: Controlling what new information will be stored in the cell state is shown in Equation 3 below.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

3) Candidate Cell State: The new candidate value that will be added to the cell state is shown in Equation 4.

$$\tilde{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{4}$$

4) Update Cell State: The combination of forget gate and input gate to update the cell state is shown below in Equation 5.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{5}$$

5) Output Gate: Determining what to output from the cell to the output is shown in Equation 6.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$$

6) Hidden State (Output of LSTM Unit) is shown below in Equation 7.

$$h_t = o_t * \tanh(C_t) \tag{7}$$

where x_t denotes the input at time t, h_{t-1} is the previous hidden state, C_t is the cell state, W and b are trainable weights and biases, σ is the sigmoid activation function.

2.7. Evaluation Metrics

Model performance was assessed using Root Mean Squared Error (RMSE) equation (8) and Mean Absolute Error (MAE) equation (9), both on the training and testing subsets. These metrics were selected because they provide complementary information regarding prediction quality [31], [32].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (8)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (9)$$

where y_i is the observed value, $\hat{y}_i$ is the predicted value, and n is the number of observations. Lower values of RMSE and MAE indicate better predictive fit. In this study, these metrics were used not only to evaluate fit, but also to assess whether the model generalized beyond the training sequences.

2.8. Exploratory Forecasting

After training, the model was used to generate a five-year forward projection of the national traffic load ratio for 2024–2028. Given the size of the dataset and the observed generalization gap, these projections were interpreted cautiously as exploratory indications of trend direction, rather than as precise policy-level forecasts.

3. RESULTS AND DISCUSSION

3.1. Historical Trend of National Traffic Load Ratio

The annual national traffic load ratio increased consistently during the observation period, indicating a continuing rise in infrastructure pressure as shown in Figure 3. The ratio increased from 208.62 in 2016 to 285.22 in 2023, showing that growth in motorized vehicles continued to outpace growth in road infrastructure. This long-run upward pattern provides the empirical basis for treating the indicator as a meaningful measure of macro-level transportation strain in Indonesia [33]. This upward trajectory suggests that the imbalance between mobility demand and road supply has not been temporary, but structural. In practical terms, the pattern implies that without stronger infrastructure expansion or transport demand management, pressure on the national road system is likely to continue accumulating.

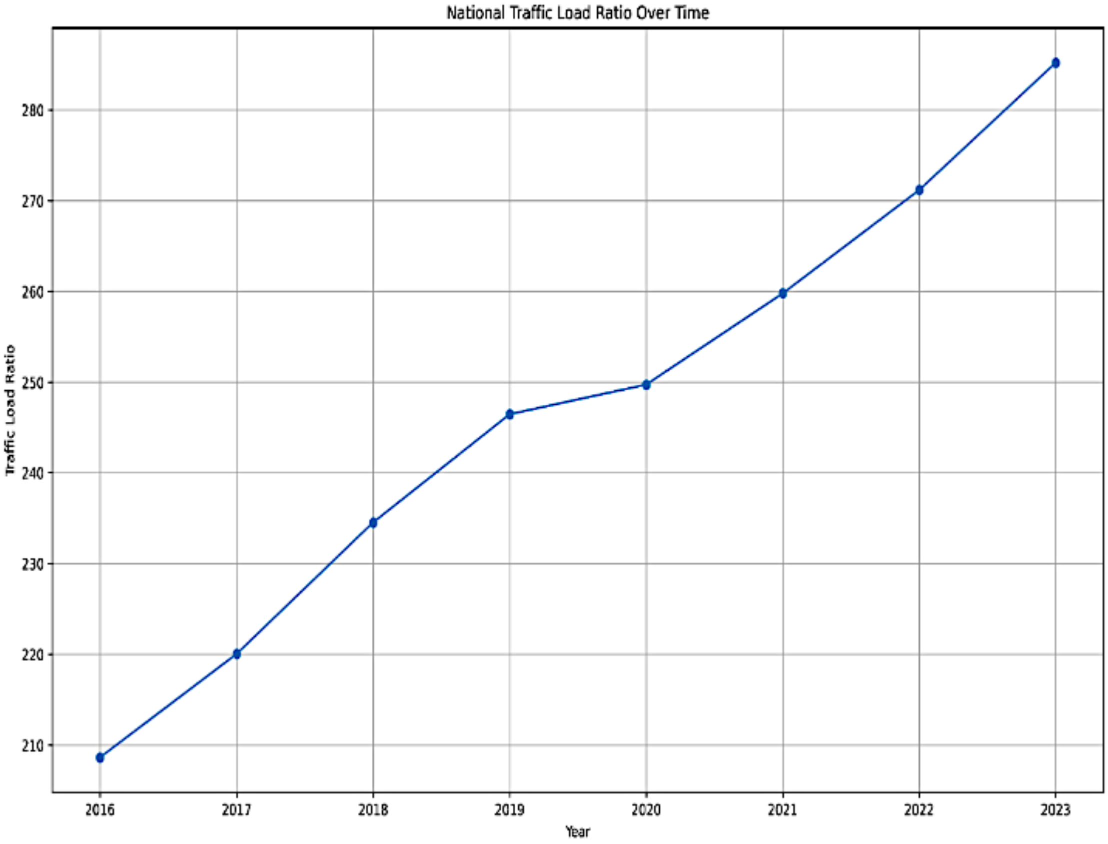

Figure 3. Graph of Traffic Load Ratio in Indonesia from 2016 to 2023

3.2. LSTM Performance

The selected LSTM model achieved an RMSE of 0.3652 and an MAE of 0.3617 on the training set. On the testing set, however, both RMSE and MAE increased to 1.7585. These results indicate that the model was able to fit the training data closely but generalized poorly to the unseen test sequence. The discrepancy between training and testing performance points to clear overfitting. Given the extremely small number of annual observations and the resulting three supervised sequences, this outcome is methodologically unsurprising [34]-[36]. The model had sufficient flexibility to reproduce the training pattern, but insufficient information to support robust out-of-sample forecasting.

Table 4. Model Evaluation

Dataset	RMSE	MAE
Training Set	0.3652	0.3617
Testing Set	1.7585	1.7585

3.3. Prediction versus Actual Visualization

The comparison between actual and predicted values shows that the model followed the historical pattern closely on the training portion of the series as shown in Figure 4. On the testing portion, the prediction still appeared to follow the general upward direction of the series, but the error metrics reveal that its numerical generalization was weak. Therefore, visual resemblance between the predicted and actual curves should not be interpreted as evidence of reliable predictive accuracy. This distinction is important. In short annual series, a model may reproduce the overall direction of change while still failing to estimate the actual magnitude of future values with adequate precision. Accordingly, the test prediction in this study is better interpreted as directional consistency rather than accurate point forecasting.

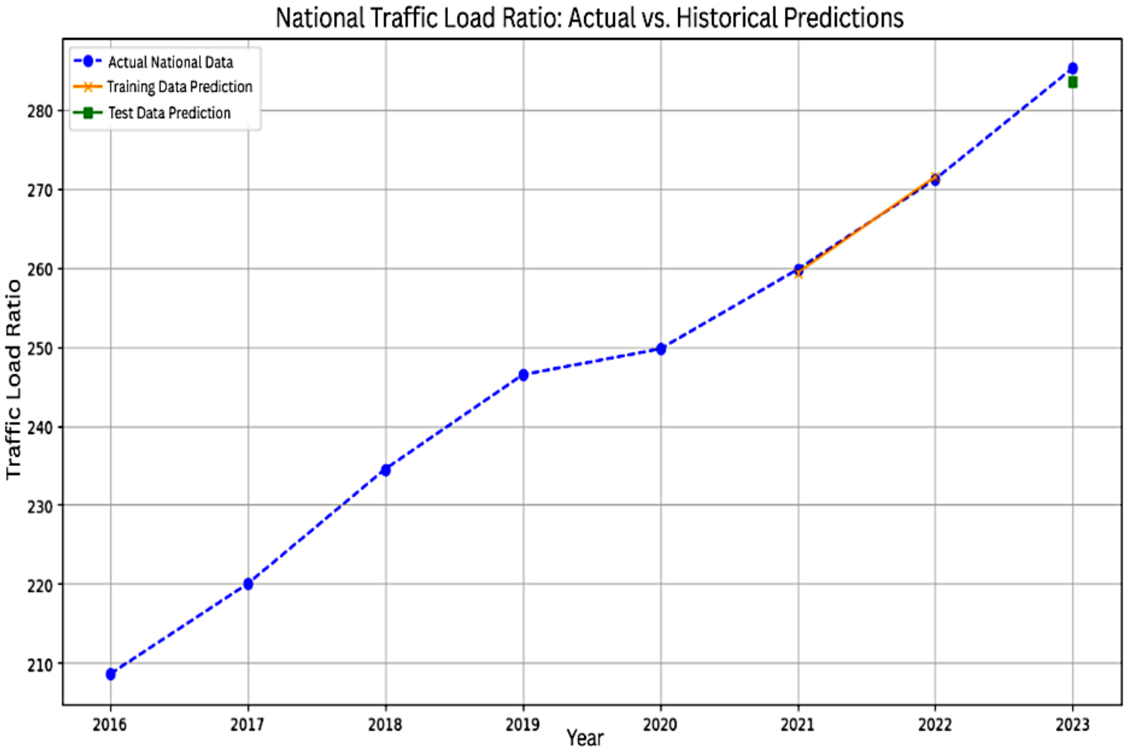

Figure 4. Comparison of Actual Value of National Traffic Load Ratio with LSTM Model

3.4. Exploratory Five-Year Projection

Using the trained LSTM model, the national traffic load ratio was projected for the next five years.

Table 5. Prediction of Future Traffic Load

Year	Traffic Load
2024	294.83
2025	305.58
2026	317.84
2027	329.84
2028	341.13

The resulting values were 294.83 for 2024, 305.58 for 2025, 317.84 for 2026, 329.84 for 2027, and 341.13 for 2028 as shown in Figure 5. These values indicate a continued upward trend in infrastructure pressure over the forecast horizon. However, this projection should be interpreted cautiously. The forecast does not establish that the numerical values themselves are highly reliable; rather, it suggests that, under the historical pattern observed in the annual aggregate data, the burden on the national road system is likely to continue increasing. In this sense, the forecast is more appropriately understood as an exploratory warning signal than as an exact planning target.

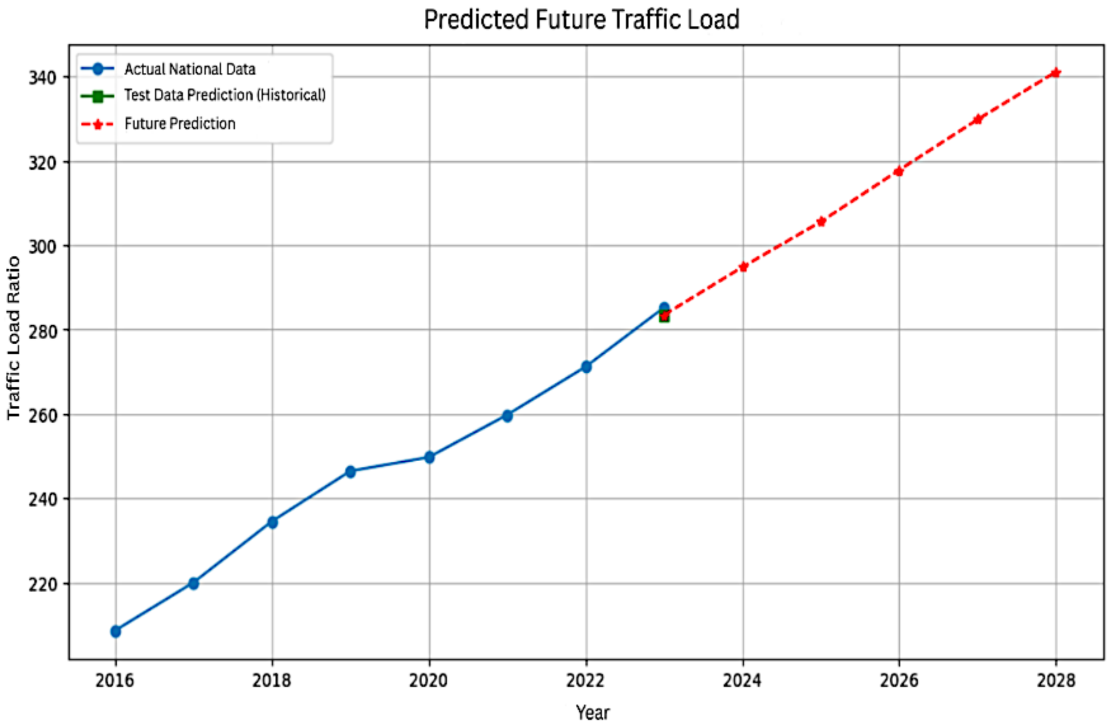

Figure 5. National Traffic Load Ratio Projection Using LSTM Models

3.5. Discussion

The findings of this study highlight the tension between the theoretical flexibility of LSTM and the practical limitations of macro-level annual transportation data. The model achieved low training errors, indicating that it could fit the observed historical pattern. However, the much larger testing error reveals weak out-of-sample generalization and confirms that the model overfit the available data. This result is consistent with the structure of the dataset itself, which contains only eight annual observations and yields only three supervised sequences for learning.

One important implication is that visual similarity between predicted and actual values should not be overinterpreted. In the present case, the testing prediction appears directionally aligned with the upward historical trend, yet the error metrics remain poor. This apparent contradiction can occur because directional pattern tracking and precise numerical forecasting are not equivalent. A model may capture the broad direction of growth while still producing unreliable point estimates. The revised interpretation therefore avoids any claim that the test predictions are “relatively accurate,” because such wording would not be consistent with the reported generalization error.

The present findings should also be interpreted relative to prior traffic forecasting studies. Many previous LSTM-based traffic studies use larger and denser datasets, such as hourly or daily traffic flow observations collected from sensor networks, road segments, or urban corridors. In those settings, deep learning models can learn temporal variation from hundreds or thousands of observations. By contrast, the present study uses annual aggregate national data, which substantially limits learnable variation and makes direct performance comparison inherently asymmetric. Therefore, this study should not be read as evidence that LSTM performs strongly in a conventional forecasting sense; instead, it should be read as an exploratory test of whether the model can still recover a broad macro-trend under highly constrained conditions.

From an infrastructure planning perspective, the most useful contribution of this study is not the exact five-year numeric forecast, but the formalization of the national traffic load ratio as a compact indicator of infrastructure pressure. The upward historical series, together with the exploratory projection, suggests that vehicle growth is continuing to outpace road expansion at the national level. This insight supports the argument that infrastructure expansion, sustainable transportation policy, and better long-term monitoring remain urgently needed. At the same time, the study has clear limitations. The sample size is extremely small, the LSTM architecture is complex relative to the available data, and no direct empirical baseline was included in the current experiment. For these reasons, the present paper is best positioned as exploratory feasibility work rather than comparative forecasting evidence. Future research should incorporate higher-frequency datasets, test simpler statistical baselines such as linear regression or ARIMA, and examine whether national forecasting can be improved through hybrid or multi-level models that combine aggregate and provincial information.

4. CONCLUSION

This study investigated the use of an LSTM model to analyze Indonesia's national traffic load ratio using annual aggregate data on motorized vehicles and road infrastructure from 2016 to 2023. The results show that the model fit the training data well, but generalized poorly to the test data, indicating substantial overfitting caused by the extremely limited sample size. Accordingly, the study should be interpreted as an exploratory proof-of-concept rather than as confirmation of robust forecasting performance. Even with these limitations, the historical series and the exploratory five-year projection consistently indicate increasing pressure on Indonesia's road infrastructure. Thus, the main value of this research lies in demonstrating the analytical usefulness of the national traffic load ratio as a macro-level indicator of infrastructure strain, and in providing an initial data-driven signal that this strain is likely to continue increasing if vehicle growth keeps outpacing road development. Future work should use larger and higher-frequency datasets, include direct comparison with simpler baseline models, and evaluate whether broader forecasting reliability can be improved before the model is considered for policy-level planning use.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the Indonesian Central Bureau of Statistics (BPS) for providing the essential datasets on motorized vehicles and road infrastructure used in this study. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] C. W. Sunie Rahardja, "Kuantifikasi Parameter Lalu Lintas Kendaraan Bermotor Di Jalan Jenderal Sudirman Jakarta Pusat," *FaST-Sains dan Teknologi*, vol. 6, no. 1, 2022.

[2] G. Yannis and A. Chaziris, "Transport System and Infrastructure," *Transportation Research Procedia*, vol. 60, no. 2021, pp. 6–11, 2022, doi: 10.1016/j.trpro.2021.12.002.

[3] N. P. S. E. Setyarini and F. I. Taubi, "Evaluasi Kondisi Ruas Jalan Tomang Raya Dengan Akj Untuk Mencapai Jalan Berkeselamatan," *Jurnal Muara Sains, Teknologi, Kedokteran dan Ilmu Kesehatan*, vol. 6, no. 2, pp. 291–300, 2022, doi: 10.24912/jmstkik.v6i2.18453.

[4] M. Chang, Z. Ding, L. Guo, and Z. Zhao, "Traffic Propagation in Road Network from a Data-Driven Analysis Perspective," *Journal of Transportation Engineering, Part A: Systems*, vol. 149, no. 2, Feb. 2023, doi: 10.1061/JTEPBS.0000758.

[5] I. Stan, D. A. Ghere, P. I. Dan, and R. Potolea, "Urban Congestion Avoidance Methodology Based on Vehicular Traffic Thresholding," *Applied Sciences (Switzerland)*, vol. 13, no. 4, Feb. 2023, doi: 10.3390/app13042143.

[6] M. Syahbandi, A. N. R. Mardiah, and S. E. Wijaya, "Designing Sustainable Transportation Strategy in Covid-19: Jabodetabek Commuter Community Movement in Indonesia," *International Journal for Disaster and Development Interface*, vol. 2, no. 1, Jul. 2022, doi: 10.53824/ijddi.v2i1.20.

[7] D. Li and J. Lasenby, "Mitigating urban motorway congestion and emissions via active traffic management," *Research in Transportation Business and Management*, vol. 48, Jun. 2023, doi: 10.1016/j.rtbm.2022.100789.

[8] D. L. Guidoni, G. Maia, F. S. H. Souza, L. A. Villas, and A. A. F. Loureiro, "Vehicular Traffic Management Based on Traffic Engineering for Vehicular Ad Hoc Networks," *IEEE Access*, vol. 8, pp. 45167–45183, 2020, doi: 10.1109/ACCESS.2020.2978700.

[9] T. Dongare, D. Huljute, P. Jadhav, A. Lad, and P. S. Marawar, "A Review on Traffic Management and Road Analysis of Porwal Road," *International Journal for Research in Applied Science and Engineering Technology*, vol. 11, no. 1, pp. 881–884, Jan. 2023, doi: 10.22214/ijraset.2023.48508.

[10] E. Prih Raharjo, K. Risma Devi Kartika, and I. Kadek Surya Putra Adidana, "Application of Transport Demand Management in Addressing Economic Losses Due to Traffic Congestion (Case Study of Gilimanuk Port During The 2024 Eid Holiday)," *Jurnal Teknologi Transportasi dan Logistik*, vol. 5, no. 2, 2024.

[11] T. Afrin and N. Yodo, "A Survey of Road Traffic Congestion Measures towards a Sustainable and Resilient Transportation System," *Sustainability*, vol. 12, no. 11, p. 4660, Jun. 2020, doi: 10.3390/su12114660.

[12] P. K. Baskar and H. Kaluvan, "Long short-term memory (LSTM) recurrent neural network (RNN) based traffic forecasting for intelligent transportation," 2022, p. 020039. doi: 10.1063/5.0083590.

[13] K. Elmazi, D. Elmazi, E. Musta, F. Mehmeti, and F. Hidri, "An Intelligent Transportation Systems-Based Machine Learning-Enhanced Traffic Prediction Model using Time Series Analysis and Regression Techniques*," in *2024 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, IEEE, Sep. 2024, pp. 1–6. doi: 10.1109/INISTA62901.2024.10683864.

[14] T. Lv, Y. Wang, J. Yuan, Y. Wu, and L. Zhang, "Prediction and Application of Network Business Traffic based on LSTM," in *Journal of Physics: Conference Series*, Institute of Physics, 2022. doi: 10.1088/1742-6596/2289/1/012011.

[15] G. Van Houdt, C. Mosquera, and G. Nápoles, "A review on the long short-term memory model," *Artificial Intelligence Review*, vol. 53, no. 8, pp. 5929–5955, Dec. 2020, doi: 10.1007/s10462-020-09838-1.

[16] S.-H. Noh, "Analysis of Gradient Vanishing of RNNs and Performance Comparison," *Information*, vol. 12, no. 11, p. 442, Oct. 2021, doi: 10.3390/info12110442.

[17] Q. Wang, R.-Q. Peng, J.-Q. Wang, Z. Li, and H.-B. Qu, "NEWLSTM: An Optimized Long Short-Term Memory Language Model for Sequence Prediction," *IEEE Access*, vol. 8, pp. 65395–65401, 2020, doi: 10.1109/ACCESS.2020.2985418.

[18] A. Isah *et al.*, "Digital Twins Temporal Dependencies-Based on Time Series Using Multivariate Long Short-Term Memory," *Electronics*, vol. 12, no. 19, p. 4187, Oct. 2023, doi: 10.3390/electronics12194187.

[19] D. H. Shin, K. Chung, and R. C. Park, "Prediction of Traffic Congestion Based on LSTM through Correction of Missing Temporal and Spatial Data," *IEEE Access*, vol. 8, pp. 150784–150796, 2020, doi: 10.1109/ACCESS.2020.3016469.

[20] K. M. Hussain, "Optimizing Urban Traffic Flows Using Advanced Data-Driven Machine Learning Models", Artificial Intell. Cyb., vol. 2, pp. 37–49, Aug. 2025, doi: 10.61356/j.aics.2025.2570.

[21] S. Reza, M. C. Ferreira, J. J. M. Machado, and J. M. R. S. Tavares, "Traffic State Prediction Using One-Dimensional Convolution Neural Networks and Long Short-Term Memory," *Applied Sciences (Switzerland)*, vol. 12, no. 10, May 2022, doi: 10.3390/app12105149.

[22] C. Ma, G. Dai, and J. Zhou, "Short-Term Traffic Flow Prediction for Urban Road Sections Based on Time Series Analysis and LSTM_BILSTM Method," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 6, pp. 5615–5624, Jun. 2022, doi: 10.1109/TITS.2021.3055258.

[23] J. Li, G. Xu, Y. Cheng, and B. Huang, "Short-Time Traffic Flow Prediction Based on K-means++ and LSTM," in *2022 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*, IEEE, Dec. 2022, pp. 286–293. doi: 10.1109/ISPA-BDCloud-SocialCom-SustainCom57177.2022.00043.

[24] B. Fernandes, F. Silva, H. Alaiz-Moreton, P. Novais, J. Neves, and C. Analide, "Long Short-Term Memory Networks for Traffic Flow Forecasting: Exploring Input Variables, Time Frames and Multi-Step Approaches," *Informatica (Netherlands)*, vol. 31, no. 4, pp. 723–749, 2020, doi: 10.15388/20-INFOR431.

[25] D. Liu and A. Wei, "Regulated LSTM Artificial Neural Networks for Option Risks," *FinTech*, vol. 1, no. 2, pp. 180–190, Jun. 2022, doi: 10.3390/fintech1020014.

[26] S. Ranjan, Y. C. Kim, N. Ranjan, S. Bhandari, and H. Kim, "Large-Scale Road Network Traffic Congestion Prediction Based on Recurrent High-Resolution Network," *Applied Sciences (Switzerland)*, vol. 13, no. 9, May 2023, doi: 10.3390/app13095512.

[27] X. Ma, Y. Li, Z. Cui, and Y. Wang, "Forecasting Transportation Network Speed Using Deep Capsule Networks with Nested LSTM Models," 2020. doi: 10.1109/TITS.2020.2984813.

[28] M. Srinath, "Vehicular Traffic Flow Prediction Model Deep Learning," *International Journal for Research in Applied Science and Engineering Technology*, vol. 11, no. 7, pp. 109–112, Jul. 2023, doi: 10.22214/ijraset.2023.54576.

[29] Q. Wu, Q. Fu, and M. Nie, "Graph Wavelet Long Short-Term Memory Neural Network: A Novel Spatial-Temporal Network for Traffic Prediction," *Journal of Physics: Conference Series*, vol. 1549, no. 4, p. 042070, Jun. 2020, doi: 10.1088/1742-6596/1549/4/042070.

[30] Z. Cui, K. Henrickson, S. A. Biancardo, Z. Pu, and Y. Wang, "Establishing Multisource Data-Integration Framework for Transportation Data Analytics," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 5, May 2020, doi: 10.1061/JTEPBS.0000331.

[31] S. Ma and M. Zhao, "Traffic Flow Prediction and Analysis in Smart Cities Based on the WND-LSTM Model," *Computational Intelligence and Neuroscience*, vol. 2022, 2022, doi: 10.1155/2022/7079045.

[32] Z. Wang, X. Su, and Z. Ding, "Long-Term Traffic Prediction Based on LSTM Encoder-Decoder Architecture," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 10, pp. 6561–6571, Oct. 2021, doi: 10.1109/TITS.2020.2995546.

[33] L. Qu, M. Zhang, Z. Li, and W. Li, "Temporal Backtracking and Multistep Delay of Traffic Speed Series Prediction," *Journal of Advanced Transportation*, vol. 2020, 2020, doi: 10.1155/2020/8899478.

[34] H. Xiao, J. Xiao, Y. Shi, X. Deng, and Y. Yang, "Traffic Speed Sequence Prediction by Adaptive Weighted Long Short-Term Memory With Classification-Type Loss," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2677, no. 8, pp. 219–233, Aug. 2023, doi: 10.1177/03611981231155911.

[35] J. Bi, X. Zhang, H. Yuan, J. Zhang, and M. Zhou, "A Hybrid Prediction Method for Realistic Network Traffic With Temporal Convolutional Network and LSTM," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 3, pp. 1869–1879, Jul. 2022, doi: 10.1109/TASE.2021.3077537.

[36] H. N. Bhandari, B. Rimal, N. R. Pokhrel, R. Rimal, and K. R. Dahal, "LSTM-SDM: An integrated framework of LSTM implementation for sequential data modeling," Software Impacts, vol. 14, p. 100396, Dec. 2022, doi: 10.1016/j.simpa.2022.100396.