Vol. 4 No. (1) (June 2026) 81-89

Journal of Data Insights

http://journalnew.unimus.ac.id/index.php/jodi

Comparative Machine Learning Models for Water-Level Forecasting and Tidal Flood Mitigation: A Case Study of Tanjung Priok Indonesia

Sugiarto*[1], Riduan[2], Linus Yoseph Wawan Rukmono[3]
[1,2,3,4]Doctoral Program, Indonesia Defense University, Indonesia

DOI: https://doi.org/10.26714/jodi.v4i1.1167

Article Info

Article history:
Received June 04, 2026
Revised June 12, 2026
Accepted June 22, 2026

Keywords:
Long Short-Term Memory (LSTM); Machine learning; Random forest; Tidal flood mitigation; Water level.

Abstract

Coastal water level variability poses significant risks to maritime logistics and increases the potential for tidal flooding in strategically important coastal areas such as Tanjung Priok. This study develops and compares Random Forest, Long Short-Term Memory (LSTM), and a simple averaging ensemble for short-term water level forecasting to support technology based tidal flood mitigation. The dataset was obtained from the Tanjung Priok Maritime Automatic Weather Station in August 2024 and consisted of one-minute meteorological and oceanographic observations. The predictors included rainfall, air temperature, air pressure, wind speed, and seawater temperature, while water level was used as the prediction target. Model performance was evaluated using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), residual analysis, and daily error dynamics. Random Forest achieved an RMSE of 0.2957 and the lowest MAE of 0.2130, indicating stable overall predictive performance. LSTM produced an RMSE of 0.3156 and an MAE of 0.2339, while showing greater flexibility in following short-term temporal variations. The ensemble achieved the lowest RMSE of 0.2954, with an MAE of 0.2132 that remained comparable to Random Forest. These findings indicate that Random Forest and the averaging ensemble provide comparable baseline performance, while LSTM offers complementary temporal characteristics. The models may support the further development of data-driven tidal-flood early warning systems at Tanjung Priok, although broader spatial coverage, longer observation periods, and systematic hyperparameter optimisation are required before operational implementation.

✉ Correspondence Address:
E-mail: giarto94@gmail.com

e-ISSN: 2988 - 2109

INTRODUCTION

As an archipelagic nation, Indonesia's extensive coastline is increasingly vulnerable to the long-term impacts of climate change, specifically the accelerated rise in sea levels [1]. However, the dynamics of sea-level change currently have significant impacts that threaten the region's stability, including tidal floods that inundate low-lying residential and infrastructure zones[2]. In the conceptualisation of national resilience, threats that damage vital infrastructure and paralyse coastal economic activity due to environmental impacts are categorised as non-military threats. To anticipate these impacts, accurate water-level predictions are a crucial component of disaster management, as they enable early detection of flood events and informed planning of coastal infrastructure to minimise economic and social disruptions[3], [4]. Therefore, providing a reliable, technology-based early warning system is no longer merely a technical necessity but part of a non-military defence strategy to safeguard Indonesia's maritime logistics sovereignty and socio-geographical safety.

Changes in sea level are influenced by complex multi-parameter drivers, including lunar cycles, precipitation levels, and meteorological variability, necessitating advanced computational approaches to accurately isolate tidal components from non-tidal environmental factors[5], [6]. The meteorological variables include rainfall, air temperature, air pressure, and wind speed, while seawater temperature represents the oceanographic variable. The complexity of these dynamic interactions among variables often results in nonlinear and stochastic patterns of water-level changes, posing unique challenges for precise modelling. Traditional hydrodynamic modelling, while foundational, often encounters limitations in developing countries due to restricted computational capacity and sparse observational data[7]. Moreover, conventional Numerical Weather Prediction methods, which rely on solving complex partial differential equations, often suffer from extreme latency, which precludes their use in real-time decision-making for rapid flood responses. To support responsive mitigation strategies in the face of these non-military threats, advanced modelling approaches are needed that can quickly and accurately unravel the nonlinear relationships in multi-parameter data.

With technological advancements, water level prediction modelling techniques have shifted from traditional statistical methods to sophisticated data-driven architectures, such as Artificial Neural Networks and Long Short-Term Memory networks, which excel at capturing complex, non-stationary temporal dynamics[8], [9]. The Random Forest model, based on an ensemble of decision trees, has proven reliable in capturing nonlinear relationships between input and target variables , effectively leveraging atmospheric indicators such as barometric pressure and humidity to enhance predictive accuracy in harsh maritime environments [10]. Meanwhile, Long Short-Term Memory (LSTM), a variant of Recurrent Neural Network (RNN), is specifically designed to address the problem of vanishing gradients and is highly effective at managing time-series dependencies by maintaining long-term memory of previous input states[11].

Although global studies have demonstrated the potential of deep learning for modelling complex hydrological and water-level dynamics [12], its application for tidal-flood forecasting in Indonesian coastal areas still requires further empirical evaluation. Previous Indonesian studies have applied recurrent neural networks, LSTM, TCN, and other deep-learning architectures to forecast sea levels in locations such as the Sunda Strait, Surabaya, and Pekalongan [1]–[3]. Random Forest has also been applied to environmental and hydrometeorological data in Indonesia, including water-level prediction at Tanjung Priok [10] and flood-disaster mitigation in Demak [13]. However, these studies generally evaluated individual model architectures, used different predictor combinations, or focused primarily on historical water-level patterns.

The present study differs from previous Indonesian water-level forecasting studies in three specific aspects. First, it compares Random Forest and LSTM using the same dataset and chronological train-test split. Second, it integrates meteorological variables, comprising rainfall, air temperature, air pressure, and wind speed, with the oceanographic variable of sea-water temperature in a single multivariate forecasting framework. Third, it combines Random Forest and LSTM predictions through a simple averaging ensemble and evaluates their performance using RMSE, MAE, residual distributions, and daily error dynamics. These elements provide a more direct assessment of model stability and temporal responsiveness for tidal-flood early warning at Tanjung Priok.

METHOD

This study used time-series data obtained from meteorological and oceanographic sensors installed at the Tanjung Priok Maritime Automatic Weather Station during August 2024. The observations were recorded at a temporal resolution of one minute. The input variables were divided into meteorological and oceanographic categories. The meteorological variables included rainfall, air temperature, air pressure, and wind speed, while seawater temperature represented the oceanographic variable. Water level was defined as the prediction target.

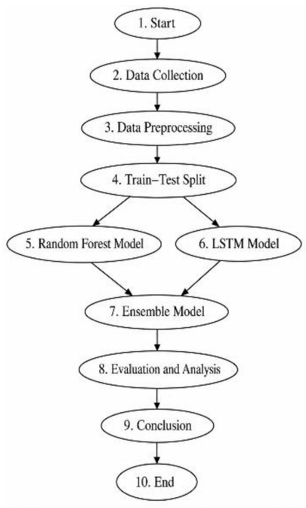

Figure 1. Research methods

The raw sensor data were preprocessed to improve data consistency and model reliability. First, the time variable was converted into a datetime format and arranged chronologically. Missing values caused by sensor interruptions were handled using time-based linear interpolation for both the input variables and the prediction target. Observations that still contained missing values after interpolation were removed to preserve the chronological structure of the dataset. The dataset was divided chronologically into training and testing sets using an 80:20 ratio. Randomisation was not applied because it could disrupt the temporal dependence of the observations and introduce information leakage from future observations into the training data. The first 80% of the observations were used to train the models, while the remaining 20% were used exclusively for performance evaluation.

The Random Forest model was selected as a representative tree-based ensemble algorithm because it can capture nonlinear relationships among multiple environmental variables and handle multivariate data without extensive feature transformation [14]. The model used 100 decision trees. This configuration was selected as a moderate ensemble size that could reduce prediction variance and maintain stable performance without imposing excessive computational requirements on the high-frequency dataset. The random state was fixed to ensure the reproducibility of the experimental results. The remaining Random Forest parameters were maintained at their standard settings. All model configurations were established before evaluating the testing dataset to prevent test-data leakage.

For the Long Short-Term Memory model, all input features and the target variable were standardised using parameters calculated only from the training dataset. The standardised data were subsequently transformed into temporal sequences. Each input sequence consisted of 24 consecutive observations and was used to predict the water level at the following observation. This formulation allowed the model to learn short-term temporal dependencies from preceding water-level conditions and environmental variables.

The LSTM architecture consisted of one LSTM layer with 64 hidden units, followed by a dropout layer with a rate of 0.2 and a dense output layer containing one neuron. The 64-unit configuration was selected as a moderate model capacity that could capture temporal dependencies without creating excessive architectural complexity. A dropout rate of 0.2 was applied to reduce the risk of overfitting by randomly deactivating a proportion of the hidden units during training. The model was trained for 30 epochs with a batch size of 32. This configuration was used to balance training stability, computational efficiency, and the risk of excessive fitting to a dataset covering a limited observation period. The Adam optimiser and Mean Squared Error loss function were used during model training. After prediction, the outputs were converted back to the original water-level measurement scale through an inverse transformation.

The selected numbers of trees, LSTM units, epochs, and dropout rate were used as reproducible baseline configurations rather than fully optimised hyperparameters. Exhaustive hyperparameter optimisation was not performed because the primary objective of this study was to compare the predictive characteristics of Random Forest and LSTM and assess the performance of their ensemble under the same dataset and chronological data-splitting procedure. Therefore, systematic sensitivity analysis and hyperparameter tuning should be conducted in future studies to examine the effects of alternative model configurations.

Predictions generated by Random Forest and LSTM were combined through a simple averaging ensemble. For each corresponding observation in the testing dataset, the ensemble prediction was calculated as the average of the Random Forest and LSTM predictions. This approach was intended to integrate the nonlinear modelling stability of Random Forest with the temporal learning capability of LSTM.

The predictive performance of Random Forest, LSTM, and the ensemble model was evaluated using Root Mean Square Error and Mean Absolute Error [15], [16]. RMSE was used to identify the influence of relatively large prediction errors, while MAE measured the average absolute difference between predicted and observed water levels. Residual analysis was also performed by examining the error distribution and plotting residuals over time. In addition, daily MAE was calculated to assess changes in model performance across the testing period and evaluate the responsiveness of each model to short-term water-level fluctuations.

RESULTS AND DISCUSSION

Performance Analysis of the Random Forest Model in Addressing Nonlinear Patterns

The implementation of the Random Forest model in this study successfully provided water level estimates, with evaluation metrics including a Root Mean Square Error (RMSE) of 0.2957 and a Mean Absolute Error (MAE) of 0.2130. These metric values indicate that this decision-tree-based algorithm can adequately capture the nonlinear interaction patterns between meteorological and oceanographic features and sea level. Visually, the trend of the model's projections to real-world data can be observed in Figure 2, which demonstrates good agreement with the actual data, although some extreme fluctuations remain that are not fully compensated for by the model.

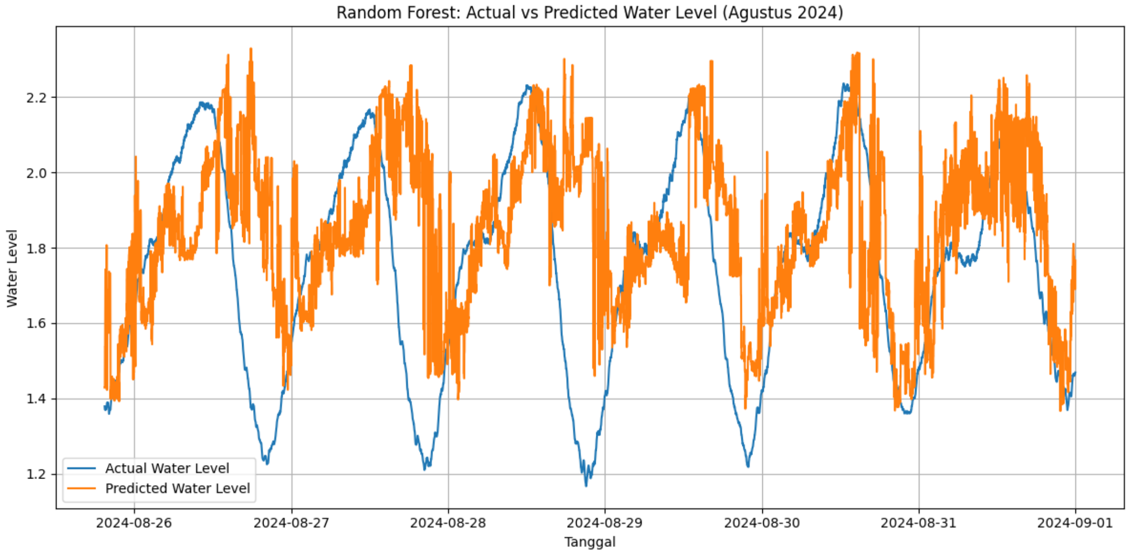

Figure 2. Observed and Random Forest-Predicted Water Levels

Through an in-depth analysis of its prediction errors, the model's error characteristics are visually depicted through the residual frequency graph in Figure 3. Based on the distribution graph, the error distribution is relatively symmetric around zero, but a slight negative bias indicates a tendency for the model to project values slightly below actual conditions (underestimate). In the context of non-military defence strategies, this tendency to underestimate requires vigilance, as it can reduce operational awareness of the risk that tidal flooding may exceed predictions.

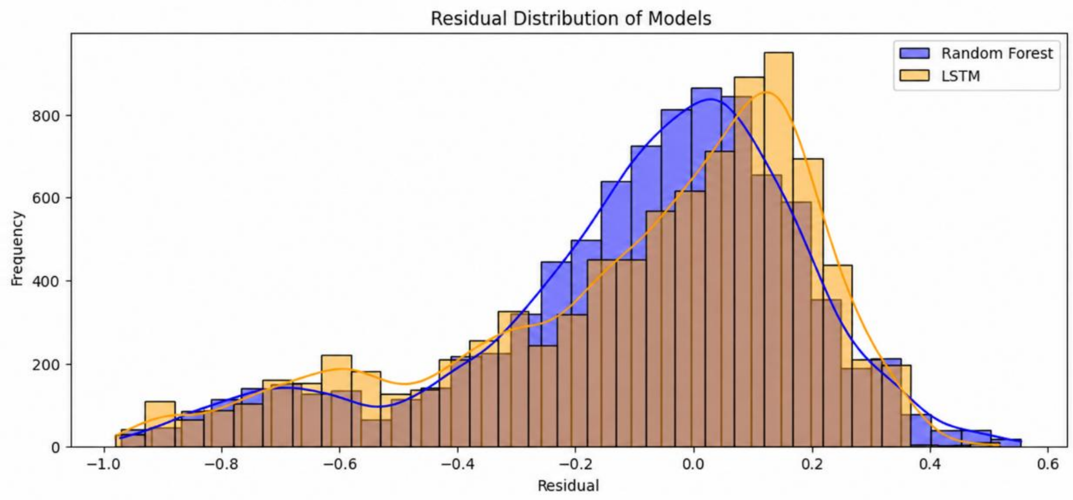

Figure 3. Residual Distributions of Random Forest and LSTM

Temporal Sensitivity Analysis of a Long Short-Term Memory (LSTM)

The Long Short-Term Memory model, configured with a single layer of 64 units, a dropout rate of 0.2, and 30 training epochs, yielded an RMSE of 0.3156 and an MAE of 0.2339. These error values were slightly higher than those obtained by the Random Forest model. Nevertheless, the temporal projection presented in Figure 4 indicates that the LSTM model followed several short-term variations in water level more continuously. This visual pattern should not be interpreted as an overall performance advantage because the aggregate error metrics remained less favourable than those of Random Forest. Instead, the result suggests that LSTM and Random Forest exhibit different predictive characteristics, with LSTM showing greater temporal flexibility and Random Forest providing lower overall prediction errors.

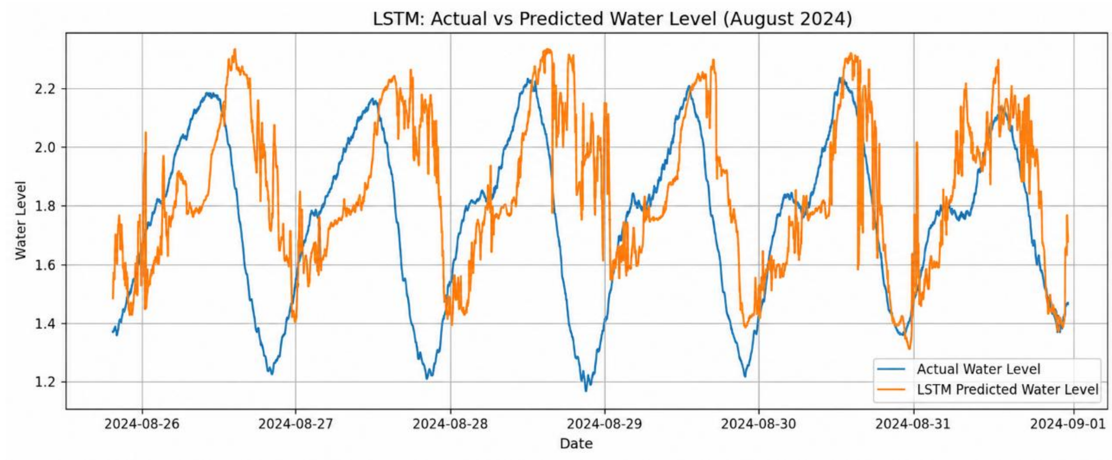

Figure 4. Observed and LSTM-Predicted Water Levels

The real-time error characteristics of this deep learning model were then evaluated comparatively through the residual scatterplot shown in Figure 5. Based on this visual comparison, the LSTM model exhibits slightly more dynamic error variations, but does not manifest significant systematic bias. The LSTM's ability to minimise this bias confirms its capability to capture complex long-term, time-dependent patterns, making it a highly adaptive component for monitoring ocean water movements influenced by dynamic weather changes.

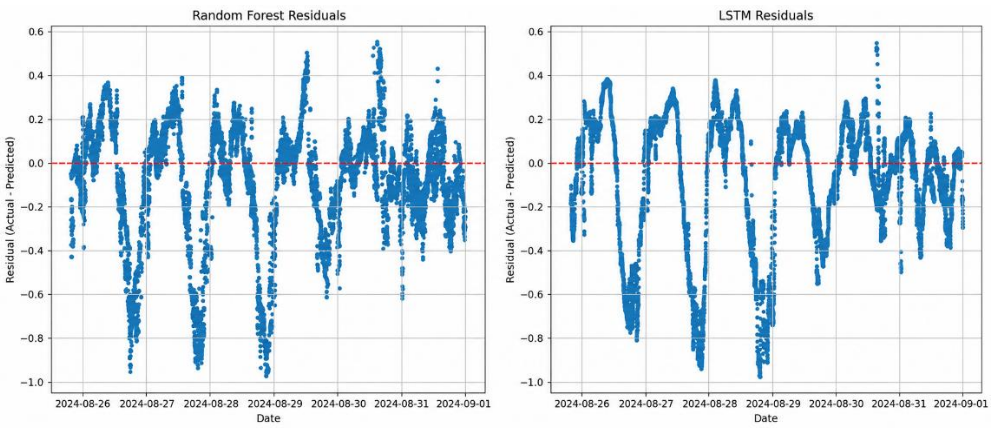

Figure 5. Random forest and LSTM comparation

Integration of a Simple Ensemble Model for Prediction Optimisation

As a tactical effort to combine the advantages of both approaches, an ensemble model based on the simple averaging method was integrated into the system. This ensemble model demonstrated a slight but significant performance improvement, achieving the lowest RMSE value of 0.2954 and MAE of 0.2132. This combined approach successfully combined the structural stability of Random

Forest with the temporal sensitivity of LSTM, resulting in much more robust and consistent predictions. Validation through a plot of the ensemble model's predictions against real data is clearly depicted in Figure 6, shows that the ensemble predictions generally follow the observed water-level pattern. However, deviations remain visible during several rapid fluctuations and peak conditions, making it an ideal option for application in coastal disaster early warning systems.

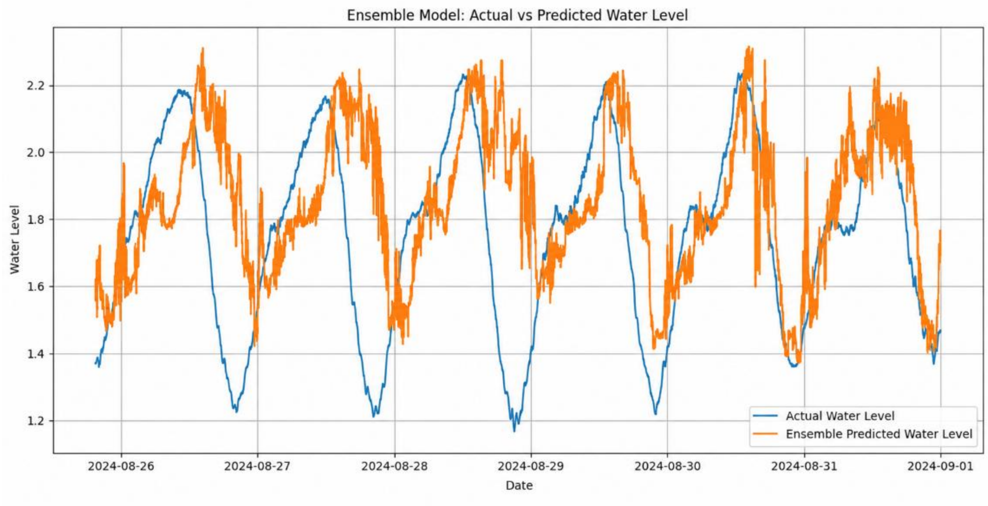

Figure 6. Observed and Ensemble Predicted Water Levels

Evaluation of Daily Error Dynamics and Implications for Mitigation Strategies

A more in-depth evaluation was conducted by analysing daily error dynamics, using the daily average absolute error (daily MAE), to assess model performance at a finer time scale. Based on the graphical trends presented in Figure 7, both Random Forest and LSTM generally exhibit similar daily error trends. However, LSTM proved much more responsive and flexible in adapting to sharp changes in water levels on certain days. Conversely, Random Forest provided highly consistent predictions throughout the testing period but tended to stiffen when faced with large, sudden data fluctuations, whereas the ensemble model successfully combined the positive characteristics of both methods.

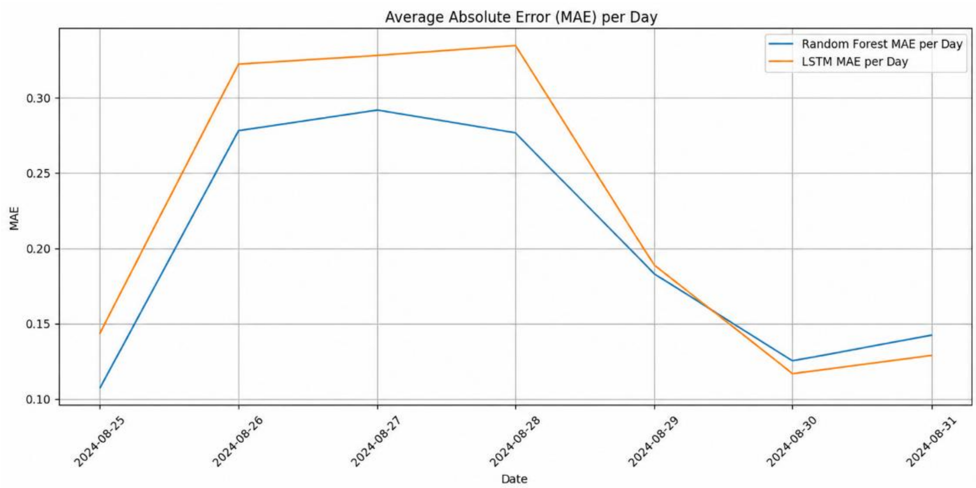

Figure 7. Daily error analysis

The quantitative performance and operational characteristics of all models tested in this study are comprehensively summarised in Table 1. A synthesis of all evaluation metrics indicates that quantitative differences in RMSE and MAE between Random Forest and the ensemble model were small ; therefore, determining the best model should be tailored to the specific objectives of the implemented non-military defence strategy. If operational objectives focus on mapping long-term macro trends with a high degree of computational stability, a single Random Forest model is considered adequate. However, for tactical disaster mitigation applications in coastal areas that require high responsiveness to real-time changes and precise temporal pattern modelling, the use of LSTM models or ensemble models is highly recommended to minimise the risk of early warning system failure at major national ports.

Table 1. Summary of evaluation metrics results (RMSE and MAE)

Model	RMSE	MAE	Remark
Random Forest	0.2957	0.2130	Stable predictions, able to capture nonlinear patterns, slight underestimation bias
LSTM	0.3156	0.2339	More responsive to temporal fluctuations, slightly larger error variance
Ensemble	0.2954	0.2132	Combining the strengths of RF and LSTM, more robust and consistent prediction results

Coastal Resilience Development Recommendations

This study is limited by its reliance on one month of data from a single observation station and by the use of baseline hyperparameter configurations without systematic optimisation. These limitations restrict the spatial and temporal generalisability of the findings. For further research to strengthen the reliability of the national mitigation strategy, it is recommended to expand the spatial coverage of the data, extend the temporal duration of observations, and apply hyperparameter tuning techniques. Furthermore, the exploration of more advanced ensemble methods, such as stacking or weighted averaging, is expected to yield a much more precise coastal disaster prediction system, thereby strengthening the resilience of Indonesia's maritime regions.

CONCLUSION

This study developed and compared Random Forest, Long Short-Term Memory (LSTM), and a simple averaging ensemble for water-level forecasting at the Tanjung Priok Maritime Automatic Weather Station using meteorological and oceanographic observations from August 2024. Random Forest achieved an RMSE of 0.2957 and the lowest MAE of 0.2130, indicating stable overall predictive performance. LSTM produced an RMSE of 0.3156 and an MAE of 0.2339. Although its aggregate errors were higher, the model showed greater flexibility in following several short-term water-level variations. The averaging ensemble achieved the lowest RMSE of 0.2954 and an MAE of 0.2132, which was comparable to the Random Forest result. These findings show that Random Forest and the averaging ensemble provided similar baseline forecasting performance, while LSTM contributed complementary temporal characteristics. The small differences in RMSE and MAE do not establish the statistical superiority of one model. Instead, model selection should consider operational priorities, including prediction stability, temporal responsiveness, and computational requirements. The results demonstrate the potential of data-driven models to support the development of tidal-flood monitoring and early warning systems at Tanjung Priok. However, the findings are limited to one month of observations from a single station and baseline model configurations without systematic hyperparameter optimisation. Further studies should use longer observation periods, additional coastal stations, systematic hyperparameter tuning, and validation during extreme water-level events before the models are considered for operational implementation. Such development may strengthen technology-based tidal-flood mitigation and support the resilience of strategic maritime infrastructure.

REFERENCES

[1] D. Saepudin, E. S. Rabbani, D. Navialdy, and D. Adytia, "Water Level Time Series Forecasting Using TCN Study Case in Surabaya," Jurnal Online Informatika, vol. 9, no. 1, pp. 61–69, Apr. 2024, doi: 10.15575/join.v9i1.1312.

[2] R. S. Pontoh et al., "Deep learning approaches to predict sea surface height above geoid in Pekalongan," International Journal of Data and Network Science, vol. 8, no. 2, pp. 743–752, Jan. 2024, doi: 10.5267/j.ijdns.2024.1.004.

[3] A. W. Ramadhan, D. Adytia, D. Saepudin, S. Husrin, and A. Adiwijaya, "Forecasting of Sea Level Time Series using RNN and LSTM Case Study in Sunda Strait," Lontar Komputer Jurnal Ilmiah Teknologi Informasi, vol. 12, no. 3, pp. 130–130, Oct. 2021, doi: 10.24843/lkjiti.2021.v12.i03.p01.

[4] N. A. Ramaputra, A. S. Budiman, and W. A. Arifin, "Deep Learning for Tidal Flood Prediction in West Pandeglang Waters, Banten," International Journal of Marine Engineering Innovation and Research, vol. 10, no. 1, pp. 175–181, Jul. 2025, doi: 10.12962/j25481479.v10i1.4759.

[5] K. D. I. Setyaningrum, K. S. Masfufah, E. Rahmawati, and A. Hermanto, "Tidal Flood Prediction in Surabaya Based on Hydrometeorological Data Using Gradient Boosting and Logistic Regression," JURNAL PIJAR MIPA, vol. 20, no. 6, pp. 1176–1181, Oct. 2025, doi: 10.29303/jpm.v20i6.10068.

[6] R. R. Salim and D. Adytia, "Deep Learning-Based Sea Level Forecasting Using Informer in Cilacap, Indonesia," JUITA Jurnal Informatika, pp. 317–327, Nov. 2025, doi: 10.30595/juita.v13i3.27176.

[7] J. Sampurno, V. Vallaeys, R. Ardianto, and E. Hanert, "Integrated hydrodynamic and machine learning models for compound flooding prediction in a data-scarce estuarine delta," Nonlinear processes in geophysics, vol. 29, no. 3, pp. 301–315, Aug. 2022, doi: 10.5194/npg-29-301-2022.

[8] P. Gholap et al., "Machine Learning Models for Early Warning of Coastal Flooding and Storm Surges," Sustainable Marine Structures, Aug. 2025, doi: 10.36956/sms.v7i3.2438.

[9] M. Liu, N. Bao, X. Yan, C. Li, and K. Peng, "A Transformer variant for multi-step forecasting of water level and hydrometeorological sensitivity analysis based on explainable artificial intelligence technology," arXiv (Cornell University), May 2024, doi: 10.48550/arxiv.2405.13646.

[10] A. P. K. Anggraini, T. Trismadi, and A. A. Supriyadi, "Predicting Water Levels from Environmental Parameters Using Random Forest Models," International Journal of Engineering Continuity, vol. 4, no. 1, pp. 39–53, Dec. 2024, doi: 10.58291/ijec.v4i1.314.

[11] S. Giaremis, N. Nader, C. Dawson, C. Kaiser, E. Nikidis, and H. Kaiser, "Storm surge modeling in the AI era: Using LSTM-based machine learning for enhancing forecasting accuracy," Coastal Engineering, vol. 191, pp. 104532–104532, Apr. 2024, doi: 10.1016/j.coastaleng.2024.104532.

[12] X. Mao et al., "Short term prediction of water level based on deep learning during the flood season, in the downstream area of The Three Gorges Reservoir," Mar. 21, 2024. doi: 10.21203/rs.3.rs-4106910/v1.

[13] J. Sutopo, F. A. Rohmatika, M. K. A. Ghani, Aprijanto, M. Pabbajah, and J. Juhansar, "Machine Learning Method with Random Forest Model for Flood Hydrometeorological Disaster Mitigation in Demak," International Journal on Advanced Science Engineering and Information Technology, vol. 15, no. 5, pp. 1475–1484, Oct. 2025, doi: 10.18517/ijaseit.15.5.20753.

[14] S. Yi and J. Yi, "Reservoir-based flood forecasting and warning: deep learning versus machine learning," Applied Water Science, vol. 14, no. 11, Oct. 2024, doi: 10.1007/s13201-024-02298-w.

[15] M. A. Raza, M. Hanzla, M. Mukhtiar, and N. Shaukat, "Advanced Flood Risk Assessment Using a Novel Hybrid Transformer-SE-ANN Framework," Aug. 06, 2025. doi: 10.21203/rs.3.rs-7103826/v2.

[16] S. J. Mohammed, S. L. Zubaidi, S. Ortega-Martorell, N. Al-Ansari, S. Ethaib, and K. Hashim, "Application of hybrid machine learning models and data pre-processing to predict water level of watersheds: Recent trends and future perspective," Cogent Engineering, vol. 9, no. 1, Nov. 2022, doi: 10.1080/23311916.2022.2143051.