Available online at
https://jcrinn.com/
https://crinn.conferencehunter.com/

Journal of Computing Research and Innovation 11 (1) 2026

Journal of Computing Research and Innovation

Deep Learning-Based Forecasting of Significant Wave Height in Padang Coastal Waters: A Comparative Study of RNN, LSTM, and GRU Architectures

Fajri Rinaldi Chan[1*], Rahma Yanti[2], Aldi Hidayat[3], Agung Ramadhanu[4], Firdaus Annas[5]

[1,3]*Torkata Research, Torkata Tech Solution, PT Torkata Jaya Persada, Padang, Indonesia.*
[2,4]*Universitas Putra Indonesia YPTK, Padang, Indonesia.*
[1,5]*Universitas Islam Negeri Sjech M. Djamil Djambek, Bukittinggi, Indonesia.*

ARTICLE INFO

Article history:
Received 23 November 2025
Revised 14 January 2026
Accepted 14 January 2026
Online first
Published 1 March 2026

Keywords:
Significant Wave Height
Deep Learning
Recurrent Neural Network
Long Short-Term Memory
Gated Recurrent Unit
Time Series Forecasting

DOI:
10.24191/jcrinn.v11i1.587

ABSTRACT

This study investigates deep learning–based forecasting of significant wave height (SWH) in Padang coastal waters, western Sumatra, a region exposed to wave-induced hazards that affect navigation, fisheries, and coastal infrastructure. SWH data are obtained from the ERA5 reanalysis single-levels product using the variable *significant_height_of_combined_wind_waves_and_swell* at hourly resolution from January to November 2025. After conversion from NetCDF to CSV, the time series is cleaned, normalized using Min-Max scaling, and transformed into supervised samples through a sliding-window approach. The dataset is split chronologically into training (January–September), validation (October), and testing (November) subsets. Three recurrent neural architectures are compared under a consistent experimental setup: a vanilla Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU). All models use the same input window length, optimization settings, and early-stopping criterion, and are evaluated with Mean Squared Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination R^2. Results show that all architectures capture SWH dynamics with very high skill, yielding test R^2values above 0.997 and MAPE below 0.7%. The RNN achieves the lowest test RMSE (0.007551) and MAE (0.003894), while the GRU delivers slightly higher errors but a modeling compromise between accuracy and training time. The LSTM attains the largest error among the three models (RMSE 0.008308, MAE 0.004423) yet is the most computationally efficient. Residual analyses indicate that the largest errors occur during sharp transitions and high-energy events. Overall, the study demonstrates that recurrent deep learning models driven solely by ERA5 reanalysis can provide accurate short-term SWH forecasts to support coastal and maritime decision-making in Padang and other data-sparse regions.

[1*] Corresponding author. *E-mail address*: fajri@torkataresearch.org
https://doi.org/10.24191/jcrinn.v11i1.587

1. INTRODUCTION

Coastal regions along the eastern Indian Ocean are increasingly exposed to ocean-related hazards, including high waves that threaten maritime navigation, coastal infrastructure, fisheries, and port operations (Eluri et al., 2025). Significant wave height (SWH), defined as the average height of the highest one-third of waves, is a key parameter for representing sea state and for designing and operating coastal and offshore structures (Li et al., 2025). Accurate SWH forecasts are therefore essential to support maritime safety, offshore engineering activities, and disaster risk reduction, particularly under the influence of a changing climate and increasingly frequent extreme events (Ma et al., 2025).

Padang, located on the west coast of Sumatra and directly facing the Indian Ocean, is an important maritime and fisheries hub in western Indonesia. Its coastal waters are strongly influenced by monsoonal winds and long-range swell systems that propagate across the eastern Indian Ocean (Paul et al., 2025). These ocean–atmosphere interactions often generate energetic sea states that can disrupt port operations, endanger small-scale fisheries, and increase the vulnerability of coastal communities (Haditiar et al., 2024). In this context, reliable, high-resolution SWH forecasts for Padang coastal waters are crucial to support local port management, fishing activities, early warning services, and coastal planning.

Operational wave predictions are traditionally produced using physics-based spectral wave models such as WAVEWATCH III or SWAN, which are driven by atmospheric reanalyses or numerical weather prediction outputs (Yao et al., 2025). In recent years, the ERA5 global reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF) has emerged as a widely used and well-validated dataset, providing hourly fields of the "significant height of combined wind waves and swell" on a global grid (Campos-Caba et al., 2024). This makes ERA5 particularly suitable as a consistent, long-term source of SWH information, especially in regions like Padang where in situ wave observations are sparse or unavailable. In this study, ERA5 "reanalysis-era5-single-levels" data at hourly resolution from January to November 2025 are used to construct a high-frequency SWH time series for the Padang coastal area (Bruno et al., 2020).

In parallel, recent advances in deep learning have significantly improved the forecasting of complex geophysical time series, including ocean wave characteristics (Fang & Li, 2025). Sequence-based architectures such as Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Units (GRU) are particularly well suited to modeling temporal dependencies and nonlinear dynamics in one-dimensional data (Zelios et al., 2025). While vanilla RNNs provide a basic framework for time series modeling, LSTM and GRU architectures incorporate gating mechanisms that mitigate vanishing gradient problems and enhance the ability to capture long-term dependencies (Yunita et al., 2025). Although deep learning has been successfully applied to wave forecasting in various regions, systematic comparative studies of RNN, LSTM, and GRU architectures for coastal SWH forecasting in Indonesian waters remain limited (Sathe et al., 2025).

Motivated by these gaps, this study aims to evaluate the capability of three recurrent deep learning architectures RNN, LSTM, and GRU for forecasting significant wave height in Padang coastal waters using ERA5 reanalysis as the primary data source. The specific objectives are to: (i) construct an hourly SWH dataset for the Padang coastal area from ERA5 "reanalysis-era5-single-levels" for January–November 2025; (ii) develop and train RNN, LSTM, and GRU models for short-term SWH forecasting; and (iii) compare their forecasting performance using standard error and skill metrics to assess their suitability for operational support in the region. The main contributions of this work are the provision of one of the first deep learning-based SWH forecasting studies focused on Padang, and a fair, unified comparison of three widely used recurrent architectures that can serve as a reference framework for future coastal wave forecasting applications

2. LITERATURE REVIEW

https://doi.org/10.24191/jcrinn.v11i1.587

Significant Wave Height (SWH) forecasting is a vital component for coastal disaster mitigation and the safety of maritime operations, as emphasized in a comprehensive review by Singh et al. (2023). In research conducted by Hao et al. (2023), it was found that traditional numerical models are often constrained by high computational costs when handling complex and nonlinear ocean dynamics. Conversely, deep learning approaches have proven to be more efficient in processing large-scale oceanographic data, as demonstrated in a study by Lee and Ahn (2025) comparing various data-driven models. Specifically, research by Alqushaibi et al. (2021) highlighted that Recurrent Neural Networks (RNN) possess advantages in modeling time-series data, yet they frequently struggle to capture long-term dependencies.

To address the limitations of RNNs, research by Minuzzi and Farina (2023) introduced the application of Long Short-Term Memory (LSTM) networks, which utilize gating mechanisms to retain relevant historical information over longer periods. A study conducted by Zhou et al. (2021) proved that LSTM models consistently outperform conventional methods in predicting wave height with lower error rates. Furthermore, research by Feng et al. (2022) showed that LSTM is highly effective in handling the non-stationary fluctuations of wind and wave data across various water conditions.

As a more efficient alternative, research by Munyao et al. (2025) evaluated the Gated Recurrent Unit (GRU) architecture, which simplifies the gating structure without sacrificing prediction accuracy. In a comparative study by Alqushaibi et al. (2021), it was found that GRUs are capable of achieving faster training convergence compared to LSTMs, making them an ideal choice for early warning systems requiring real-time processing. Additionally, recent research by Ji et al. (2023) proposed the use of hybrid models combining signal decomposition techniques with GRU or LSTM to improve prediction stability in extreme sea states. Finally, a study by Yadav et al. (2025) concluded that integrating global reanalysis data, such as ERA5, into deep learning models can significantly enhance forecasting accuracy in regions with limited observational data.

3. METHODS AND MATERIAL

The overall methodological workflow adopted in this study is summarized in Fig. 1. The process begins with the acquisition of ERA5 reanalysis data from the Copernicus Climate Data Store (CDS) and the extraction of the significant height of combined wind waves and swell (SWH) variable for the Padang coastal area. The raw data in NetCDF format are converted to CSV, then subjected to basic cleaning and normalization. Subsequently, the normalized SWH time series is transformed into supervised learning samples using a sliding-window approach and split chronologically into training and testing subsets.

In the next stage, three recurrent deep learning architectures—simple RNN, LSTM, and GRU—are trained on the prepared data under comparable settings. The trained models are evaluated on the test set using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and the coefficient of determination (R^2). The resulting performance metrics form the basis for a comparative analysis to determine the most effective architecture for short-term forecasting of significant wave height in Padang coastal waters.

https://doi.org/10.24191/jcrinn.v11i1.587

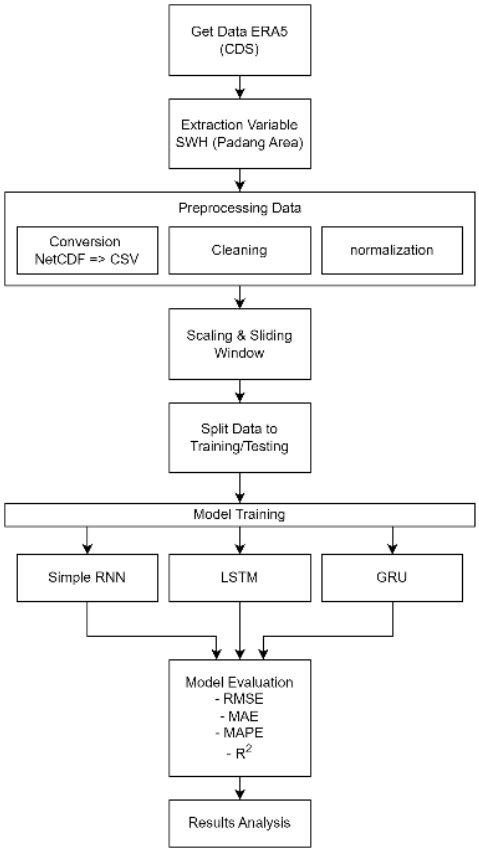

Fig. 1. Overall methodological workflow for deep learning-based forecasting of significant wave height in Padang coastal waters, including data acquisition, preprocessing, model training (RNN, LSTM, GRU), and performance evaluation

Source: Author's illustration

3.1 Study area and data description

The study focuses on coastal waters adjacent to Padang, located on the west coast of Sumatra and directly facing the eastern Indian Ocean. In this region, sea state conditions are dominated by a combination of local wind waves and remotely generated swell, which together determine the significant wave height (SWH) observed at the coast. In this work, SWH is represented by the "significant height of combined wind waves and swell" variable from the ERA5 reanalysis, which provides a consistent, gap-free estimate of hourly sea state conditions over the study period (Hao et al., 2023).

The wave data used in this study are obtained from the ERA5 reanalysis-single-levels product, The extraction is performed using the Copernicus Climate Data Store (CDS) API, selecting the grid point (or a small average of nearby grid points) that best represents Padang coastal waters. This results in an hourly SWH time series from 1 January 2025 00:00 to 15 November 2025 23:00, which forms the primary input dataset for all forecasting experiments.

Fig. 2 presents the hourly time series of significant wave height for Padang coastal waters over the study period. The SWH values generally range between about 0.35 m and 1.4 m, with several pronounced peaks associated with energetic wave events. Higher and more variable wave conditions can be observed

during certain months (e.g., late boreal autumn), while relatively calmer conditions appear in other parts of the record. This variability indicates the presence of both short-term fluctuations and longer-period modulations in the sea state, highlighting the need for forecasting models capable of capturing nonlinear temporal dependencies in the SWH signal (Fang & Li, 2025).

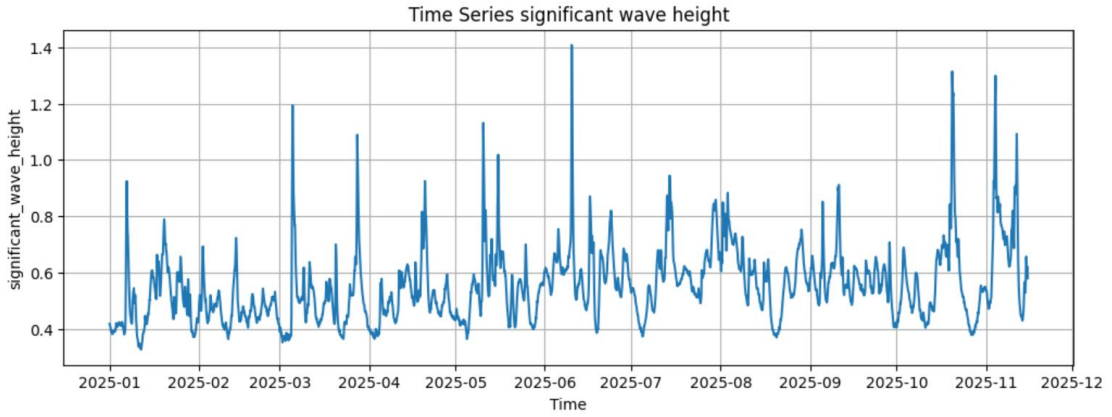

Fig. 2. Hourly time series of significant wave height (m) in Padang coastal waters from January to November 2025 derived from ERA5 reanalysis

Source: ERA5 reanalysis-era5-single-levels dataset, variable significant_height_of_combined_wind_waves_and_swell (January–November 2025), Copernicus Climate Change Service (C3S), European Centre for Medium-Range Weather Forecasts (ECMWF)

3.2 Data preprocessing

Prior to model development, several preprocessing steps are applied to the raw ERA5 SWH time series. First, the data are checked for missing values and obvious outliers. Because ERA5 is a reanalysis product, missing values are not expected; however, any anomalous spikes due to interpolation or download issues are removed or replaced using short-window linear interpolation. If desired, the time series can be smoothed using a moving average 3–5 hours to reduce high-frequency noise, although the main experiments in this study are performed on the original hourly resolution to preserve short-term variability.

The cleaned SWH time series is then normalized to facilitate efficient training of deep learning models. A Min–Max scaling is applied to map the data into the [0, 1] interval using the minimum and maximum values of the training set only. Let H_t denote the original SWH at time t; the scaled value $\widetilde{H_t}$ is computed as (Han et al., 2025).

$$\widetilde{H_t} = \frac{H_t - H_{min}}{H_{max} - H_{min}} \tag{1}$$

After scaling, the time series is transformed into supervised learning samples using a sliding-window approach. For a chosen input window length w, each input sequence consists of w consecutive SWH values $[\widetilde{H}_{t-w+1}, \dots, \widetilde{H_t}]$, and the corresponding target is the SWH at a future time step $\widetilde{H}_t + h$ where h is the forecast horizon. In this study, a one-step-ahead forecasting strategy is adopted and multi-step forecasts are produced recursively by feeding previous predictions back into the models (Omar et al., 2024).

To respect temporal dependence and avoid information leakage, the dataset is split chronologically. Data from January to September 2025 are used for training, data from October 2025 for validation (hyperparameter tuning and early stopping), and data from November 2025 for independent testing. The scaling parameters are derived from the training set and then applied to the validation and test sets.

https://doi.org/10.24191/jcrinn.v11i1.587

3.3 Deep learning architectures

This study compares three recurrent deep learning architectures for one-step-ahead forecasting of significant wave height: a vanilla Recurrent Neural Network (RNN), a Long Short-Term Memory (LSTM) network, and a Gated Recurrent Unit (GRU) network. All models are designed as sequence-to-one regressors, where an input window of past SWH values is mapped to the next SWH value. To ensure a fair comparison, the three architectures share a similar overall design: (i) an input layer that receives sequences of normalized SWH values, (ii) one or two recurrent layers (RNN, LSTM, or GRU units), optionally followed by dropout, and (iii) a fully connected output layer with linear activation that predicts the next value in the time series.

The models are implemented using the same train/validation split, scaling procedure, and loss function (mean squared error, MSE). Hyperparameters such as the number of units, number of layers, learning rate, and dropout rate are kept within comparable ranges across architectures and are tuned based on validation performance (Chan et al., 2024). The following subsections describe each architecture in more detail.

3.3.1. Recurrent Neural Network (RNN)

The vanilla Recurrent Neural Network (RNN) is the simplest recurrent architecture used in this study and serves as a baseline deep learning model. An RNN processes a time series sequentially by maintaining a hidden state that is updated at each time step based on the current input and the previous hidden state (Munyao et al., 2025). For an input sequence $\{x_1, x_2, \dots, x_T\}$, the hidden state h_t at time t is computed as:

$$h_t = \phi(W_{xh}x_t + W_{hh}h_{t-1} + b_h), \tag{2}$$

$$\widehat{y_t} = W_{hy}h_t + b_y \tag{3}$$

Where W_{xh}, W_{hh}, W_{hy} are weight matrices, b_h and b_y are bias vectors, $\phi(\cdot)$ is a nonlinear activation function (tanh in this study), and $\widehat{y_t}$ is the model output at time t.

In this work, the RNN receives as input a sliding window of length w containing normalized SWH values $\left[\widetilde{H_{t-w+1}}, \dots, \widetilde{H_t}\right]$. The final hidden state h_T (corresponding to the last time step in the window) is passed to a dense layer to produce a single forecast $\widehat{H}_{t+1}$. The RNN architecture used here typically consists of one recurrent layer with 32–64 units, followed by a dropout layer (e.g., dropout rate 0.2) to reduce overfitting, and a linear output layer. Although vanilla RNNs are conceptually simple and computationally efficient, they are known to suffer from vanishing and exploding gradient problems when modeling long sequences, which can limit their ability to capture long-term dependencies in SWH time series (Unlu, 2025).

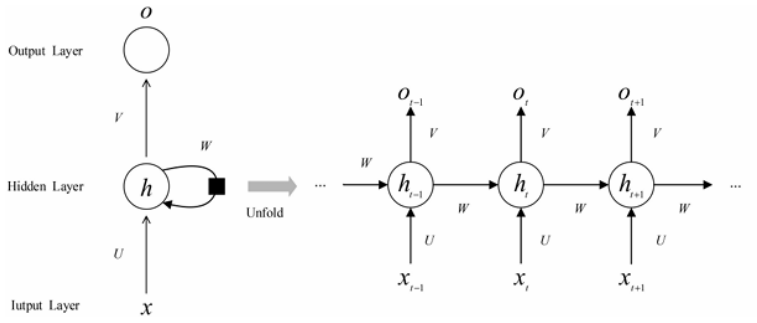

Fig. 3. Schematic architecture of a vanilla recurrent neural network (RNN) in folded (left) and unfolded (right) form along the time axis

Source: Author's illustration based on standard RNN formulation in the literature

3.3.2. Long Short-Term Memory (LSTM)

The Long Short-Term Memory (LSTM) network is an extension of RNN designed to overcome the vanishing gradient problem and to better capture long-range temporal dependencies. LSTM introduces a memory cell and a set of gating mechanisms that control the flow of information through time (Yunita et al., 2025). For each time step t, the LSTM computes:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \ \textit{(forget gate)} \tag{4}$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \ \textit{(input gate)} \tag{5}$$

$$\tilde{c_t} = \tanh(W_c x_t + U_c h_{t-1} + b_c) \ \textit{(candidate cell state)} \tag{6}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c_t} \ \textit{(updated cell state)} \tag{7}$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \ \textit{(output gate)} \tag{8}$$

$$h_t = o_t \odot \tanh(c_t) \tag{9}$$

where $x_t$is the input at time t, $h_t$is the hidden state, $c_t$is the cell state, $\sigma(\cdot)$is the sigmoid activation function, and $\odot$denotes element-wise multiplication.

In the proposed framework, the LSTM takes the same input sequences as the RNN (windows of length $w$of normalized SWH) and outputs a single forecast from the final hidden state. The LSTM architecture typically uses 32–64 units in a single LSTM layer, followed by dropout and a dense output layer with linear activation. Fig. 4 shows the internal structure of an LSTM cell, including the forget, input, and output gates, as well as the update of the cell state and hidden state. Due to its gating mechanisms and explicit cell state, the LSTM is expected to better learn medium- to long-range dependencies in the SWH time series, such as patterns associated with swell propagation or persistent meteorological conditions (Zelios et al., 2025).

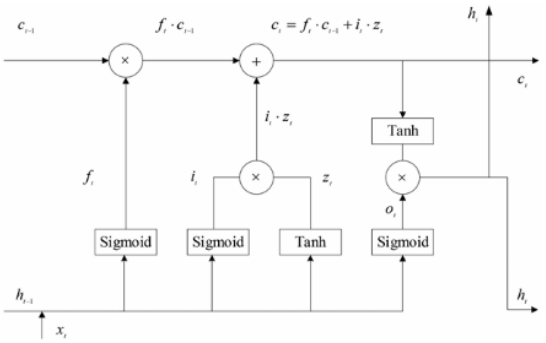

Fig. 4. Internal structure of a Long Short-Term Memory (LSTM) unit, showing forget, input, and output gates and the update of the cell state

Source: Author's illustration based on standard LSTM formulation in the literature

3.3.3. Gated Recurrent Unit (GRU)

The Gated Recurrent Unit (GRU) is another gated recurrent architecture that simplifies the LSTM design while retaining the ability to model long-term dependencies. GRU combines the functions of the forget and input gates into a single update gate and uses a reset gate to control how new information is incorporated (Sathe et al., 2025). For each time step t, the GRU equations are:

$$z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \ \ (update\ gate) \tag{10}$$

$$r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \ \ (reset\ gate) \tag{11}$$

$$\widetilde{h_t} = \tanh(W_h x_t + U_h(r_t \odot h_{t-1}) + b_h) \ \ (candidate\ hidden\ state) \tag{12}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \widetilde{h_t} \tag{13}$$

In this study, the GRU model is configured analogously to the LSTM: it receives windows of normalized SWH values as input, and the final hidden state is passed to a dense layer with linear activation to predict the next SWH value. The GRU layer typically uses 32–64 units and may be followed by dropout to mitigate overfitting. Fig. 5 illustrates the internal structure of a GRU unit, highlighting the update and reset gates and the computation of the new hidden state. Because GRU has fewer gates and parameters than LSTM, it is computationally more efficient and can converge faster, which is advantageous when training on relatively long time series or when deploying models in operational settings, while still providing improved modeling capability over vanilla RNN (Maria et al., 2025).

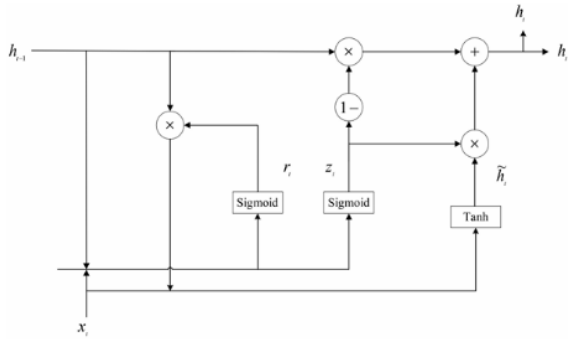

Fig. 5. Internal structure of a Gated Recurrent Unit (GRU), showing update and reset gates and the computation of the new hidden state

Source: Author's illustration based on standard GRU formulation in the literature

3.4 Training procedure and experimental design

The training process is carried out using mini-batch gradient descent with batch sizes 64 samples per batch, over a maximum number of epochs (e.g., 100–200 epochs). Early stopping is employed based on the validation loss: training is halted if the validation loss does not improve after a fixed number of consecutive epochs (patience), and the model parameters from the epoch with the lowest validation loss are retained. This strategy helps to prevent overfitting and ensures that the selected model generalizes well to unseen data.

https://doi.org/10.24191/jcrinn.v11i1.587

All three architectures (RNN, LSTM, GRU) are trained under identical conditions: same input window length w, same train/validation/test splits, same normalization scheme, and similar hyperparameter ranges. Hyperparameters (number of units, number of layers, dropout rate, learning rate, window length) are tuned using a combination of manual search and validation performance. In addition to the deep learning models, a simple persistence model is used as a baseline, in which the forecast is simply equal to the most recent observation $\widehat{H}_{t+1} = H_t$. This provides a reference level of skill that any advanced model should exceed.

3.5 Evaluation metrics

Model performance is evaluated on the independent test period (November 2025) using several commonly used error and skill metrics. The Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) quantify the average magnitude of prediction errors, while the coefficient of determination R^2measures the proportion of variance in the observations explained by the model (Arleth et al., 2024). These metrics are defined as follows, for Ntest samples, observed values H_i, and predictions $\widehat{H}_i$ (J. Zhou et al., 2025):

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(H_i - \widehat{H_\iota}\right)^2} \tag{14}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|H_i - \widehat{H_\iota}\right| \tag{15}$$

$$\text{MAPE} = \frac{100\backslash\%}{N}\sum_{i=1}^{N}\left|\frac{H_i - \widehat{H_\iota}}{H_i}\right| \tag{16}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(H_i - \widehat{H_\iota}\right)^2}{\sum_{i=1}^{N}(H_i - \bar{H})^2} \tag{17}$$

where $\bar{H}$is the mean of the observed SWH values in the test set.

For each architecture (RNN, LSTM, GRU), these metrics are computed and compared against each other and against the persistence model. The comparative analysis focuses on identifying which architecture provides the best trade-off between accuracy and model complexity for short-term SWH forecasting in Padang coastal waters, thereby informing the selection of an appropriate deep learning model for future coastal wave forecasting applications.

4. RESULTS AND DISCUSSION

4.1 Overall model performance

The overall predictive performance of the three architectures is summarized in Table 1. All models achieve very small errors, with test MSE values on the order of $10^{-5}$and coefficients of determination $R^2 > 0.997$, indicating an excellent fit to the observed SWH time series. Among the three architectures, the RNN provides the best accuracy on the test set with a test MSE of 0.000057, RMSE of 0.007551, MAE of 0.003894, and MAPE of about 0.59%. The GRU model ranks second (test RMSE 0.007722, MAE 0.004186, MAPE ≈ 0.65%, $R^2 = 0.997485$), while the LSTM yields the largest errors (test RMSE 0.008308, MAE 0.004423, MAPE ≈ 0.68%, $R^2 = 0.997089$).

https://doi.org/10.24191/jcrinn.v11i1.587

In terms of computational cost, the LSTM is the most efficient, requiring approximately 117 s of training time, followed by the GRU (≈319 s), whereas the RNN demands the longest training time (≈394 s). This highlights a clear trade-off between accuracy and training efficiency: the RNN achieves the lowest error but at the highest computational cost, while the LSTM sacrifices a small amount of accuracy in exchange for substantially faster training.

Table 1. Validation and test performance of RNN, LSTM, and GRU models for significant wave height forecasting in Padang coastal waters (validation MSE/MAE; test MSE, RMSE, MAE, MAPE, R^2; and total training time)

Architecture	Evaluation	Value	Training Time (SEC)
RNN	Val Loss (MSE)	0.000024	394.245021
	Val MAE	0.002914	
	Test MSE	0.000057	
	Test RMSE	0.007551	
	Test MAE	0.003894	
	Test MAPE (%)	0.594253	
	R^2	0.997596	
LSTM	Val Loss (MSE)	0.000030	117.226785
	Val MAE	0.003284	
	Test MSE	0.000069	
	Test RMSE	0.008308	
	Test MAE	0.004423	
	Test MAPE (%)	0.684989	
	R^2	0.997089	
GRU	Val Loss (MSE)	0.000028	319.351064
	Val MAE	0.003153	
	Test MSE	0.000060	
	Test RMSE	0.007722	
	Test MAE	0.004186	
	Test MAPE (%)	0.652050	
	R^2	0.997485	

The evolution of validation loss during training is depicted in Fig. 6. All three models exhibit a rapid decrease in MSE during the first 10–20 epochs, followed by a gradual convergence towards very small loss values. The RNN curve converges to the lowest validation loss, closely followed by the GRU, whereas the LSTM stabilizes at a slightly higher level. No strong signs of overfitting are observed, as the validation curves flatten rather than diverge in the later epochs, confirming that the chosen architectures and regularization settings are adequate for this dataset.

https://doi.org/10.24191/jcrinn.v11i1.587

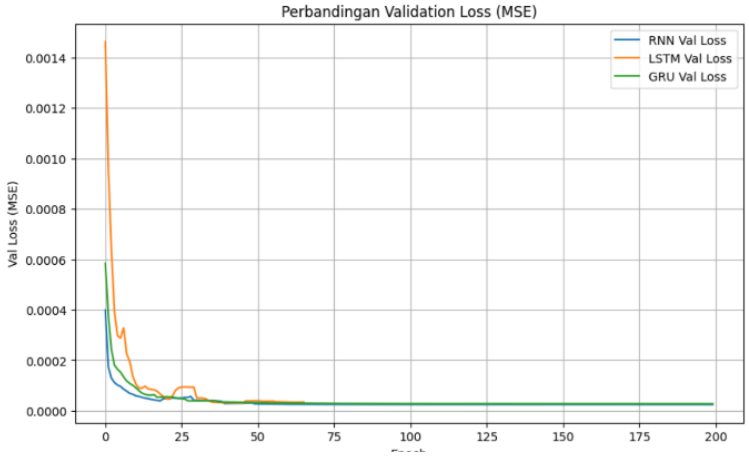

Fig. 6. Comparison of validation loss (MSE) as a function of epoch for RNN, LSTM, and GRU models

4.2 Accuracy on the test set

The predictive accuracy on the test set is illustrated in Fig. 7, which shows the last 200 points of the SWH series together with the corresponding forecasts from each model. Visually, the three predicted curves almost overlap with the observed series, accurately capturing both gradual variations and relatively sharp rises and falls in wave height. Small differences between models appear mainly around local maxima: the GRU and LSTM occasionally over- or under-estimate peak values relative to the RNN, but the discrepancies remain very small (on the order of a few millimetres to centimetres).

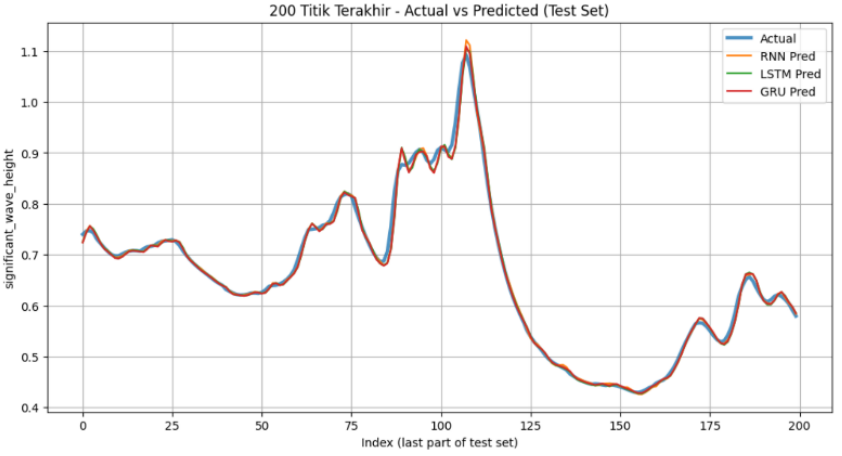

Fig. 7. Comparison between observed significant wave height and predictions from RNN, LSTM, and GRU models for the last 200-time steps of the test set.

The residual distributions for each model are presented in Fig. 8. All histograms are centred around zero and approximately symmetric, indicating the absence of strong systematic biases. The RNN residuals exhibit the narrowest spread, consistent with its lower RMSE and MAE. The LSTM and GRU show slightly wider tails, reflecting a higher frequency of moderate residuals, but still within a very small error range.

https://doi.org/10.24191/jcrinn.v11i1.587

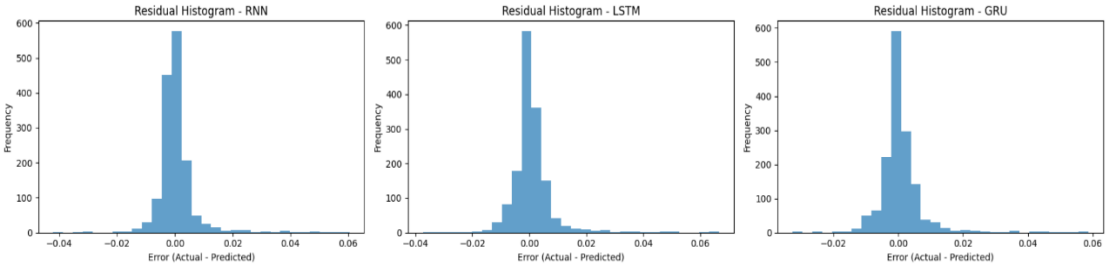

Fig. 8. Residual histograms (error = actual − predicted) for the (a) RNN, (b) LSTM, and (c) GRU models on the test set

Residuals plotted as a function of time Figure 9 reveal that errors for all models generally remain within ±0.01 m, with a few larger spikes coinciding with periods of more energetic sea states. The three architectures show similar timing and magnitude of these spikes, implying that the main difficulty lies in reproducing very rapid transitions or sharp peaks in SWH rather than the more regular background variability.

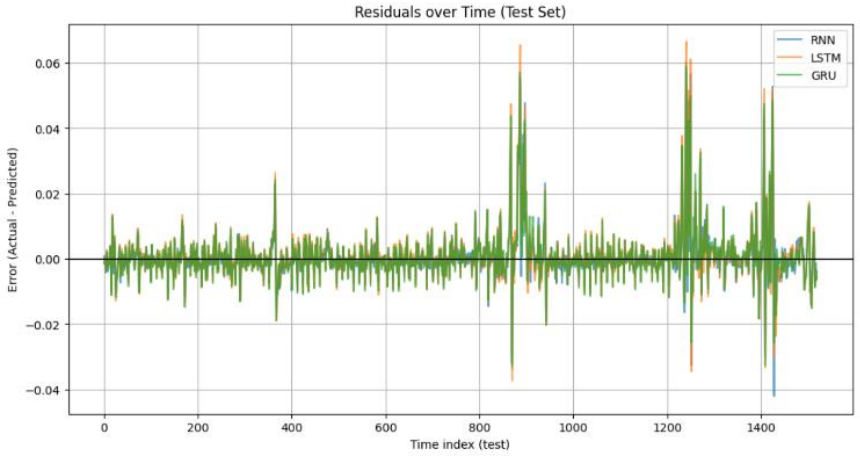

Fig. 9. Residuals over time for RNN, LSTM, and GRU models on the test set

4.3 Multi-step forecasting behaviour

The multi-step forecasting behaviour of the models is illustrated in Fig. 10, which compares the historical SWH time series with forward projections produced by the three architectures at the end of the study period. All models generate smooth forecast trajectories that continue the prevailing trend of the most recent observations. Differences become visible as the forecast horizon increases: the GRU tends to project a slightly stronger increase in SWH, while the LSTM yields more conservative changes; the RNN generally lies between the two.

https://doi.org/10.24191/jcrinn.v11i1.587

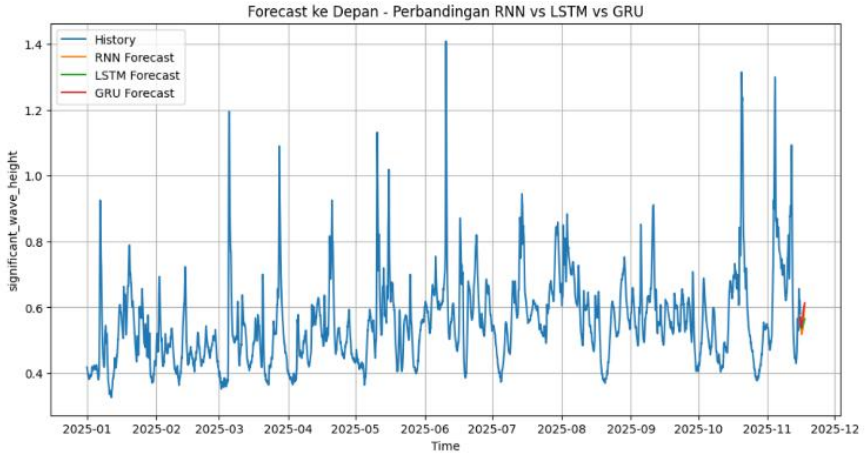

Fig. 10. Historical SWH time series and forward forecasts generated by RNN, LSTM, and GRU at the end of the analysis period

A zoomed view of the last 20 historical points and subsequent forecasts is shown in Fig. 11, where the vertical dashed line marks the boundary between observed data and predictions. Immediately after the boundary, all three models extend the observed decreasing trend in SWH fairly consistently. Further into the forecast horizon, however, the GRU predicts a faster recovery and increase in wave height, the RNN shows a moderate upturn, and the LSTM remains the most conservative. From an operational standpoint, this implies that the GRU could be more sensitive to potential increases in sea state (at the risk of occasionally over-predicting), whereas the LSTM emphasises stability and may underestimate rapid transitions.

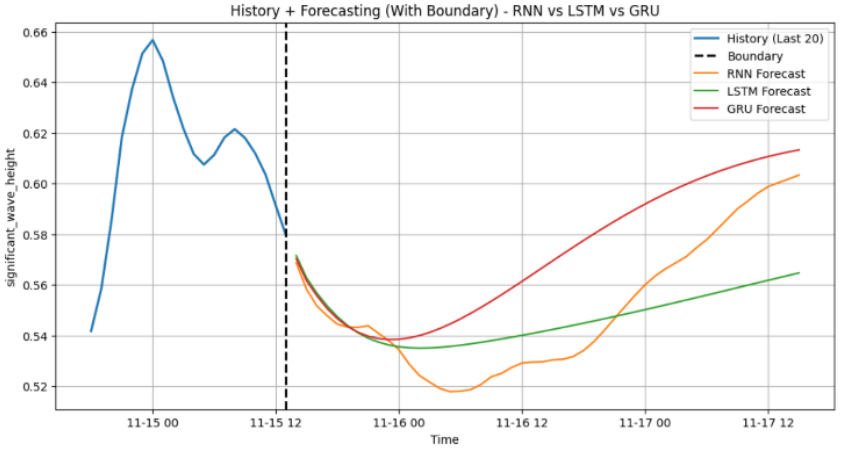

Fig. 11. Zoomed-in view of the last 20 observed points and subsequent multi-step forecasts from RNN, LSTM, and GRU models; the dashed line indicates the boundary between history and forecast horizon

4.4 Discussion

Overall, the results demonstrate that all three recurrent architectures RNN, LSTM, and GRU are capable of modelling the temporal dynamics of significant wave height in Padang coastal waters with very high accuracy. Numerically, the RNN achieves the best performance on the test set but at the cost of the

longest training time. The GRU model offers a modeling balance between accuracy and computational efficiency, while the LSTM is the most computationally efficient but slightly less accurate.

Residual analyses show that the largest errors for all models occur during sharp transitions and high-energy events, which are inherently more difficult to predict. For practical coastal and maritime applications, the choice of model can thus be guided by the trade-off between accuracy and efficiency: RNN or GRU may be preferred when maximising accuracy is critical, whereas LSTM may be more appropriate in resource-constrained or real-time settings. Future work could further improve performance by incorporating additional meteorological predictors (e.g., wind speed and direction) and by using loss functions that give greater weight to extreme events, as well as extending the training dataset to multiple years to enhance robustness and generalisation.

5. CONCLUSIONS

This study presented a deep learning–based framework for forecasting significant wave height (SWH) in Padang coastal waters using hourly ERA5 reanalysis data for January–November 2025. The ERA5 *significant height of combined wind waves and swell* variable was preprocessed through cleaning, normalisation, and a sliding-window transformation to construct supervised time-series samples. Three recurrent architectures—vanilla RNN, LSTM, and GRU—were trained and evaluated under a consistent experimental setup using standard error metrics (MSE, RMSE, MAE, MAPE) and the coefficient of determination R^2.

The results show that all three architectures are capable of reproducing the temporal dynamics of SWH with very high skill, achieving test R^2 values above 0.997 and MAPE below 1%. Among them, the RNN model attained the lowest test error and the narrowest residual distribution, indicating the highest point-wise accuracy, but required the longest training time. The GRU model provided slightly lower accuracy than the RNN but with reduced computational cost, whereas the LSTM model was the most computationally efficient yet exhibited the largest (though still small) errors. Residual analyses further revealed that the main forecasting difficulty for all models lies in capturing sharp transitions and high-energy events, while background variability is modelled very well.

From an application perspective, the proposed framework demonstrates that recurrent deep learning models, trained solely on reanalysis-derived SWH, can deliver accurate short-term forecasts suitable to support coastal and maritime activities in Padang, such as port operation planning, small-scale fisheries, and early-warning services. The trade-off profiles observed in this study suggest that RNN or GRU may be preferred when maximising forecast accuracy is critical, whereas LSTM may be advantageous in real-time or resource-limited environments.

Future work will focus on extending the training period to multiple years to improve robustness, incorporating additional atmospheric and oceanic predictors (e.g., wind fields, pressure, or directional wave parameters) to better resolve extreme events, and exploring hybrid or spatial models that combine deep learning with physics-based information. Integrating the developed forecasting system into operational decision-support tools for local stakeholders represents a further practical direction of this research.

6. ACKNOWLEDGEMENTS/FUNDING

The authors would like to express their sincere gratitude to Universitas Putra Indonesia YPTK Padang and UIN Sjech M. Djamil Djambek Bukittinggi for the academic support, institutional encouragement, and research facilities provided during the preparation and completion of this manuscript. The authors also appreciate the valuable academic environment that contributed to the successful conduct of this research

7. CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

8. AUTHORS' CONTRIBUTIONS

Fajri Rinaldi Chan contributed to the conceptualization, methodology, data analysis, and drafting of the manuscript; **Rahma Yanti** contributed to the methodology and data analysis; **Aldi Hidayat** contributed to data collection and preprocessing; **Agung Ramadhanu** contributed to validation and analysis; and **Firdaus Annas** contributed to supervision, manuscript review, and editing. All authors have read and approved the final version of the manuscript.

9. REFERENCES

Alqushaibi, A., Abdulkadir, S. J., Rais, H., & Al-tashi, Q. (2021). Enhanced weight-optimized recurrent neural networks based on sine cosine algorithm for wave height prediction. *Marine Science and Engineering*, *9*(524). https://doi.org/10.3390/jmse9050524

Arleth, S., Ludeña, R., Collazos, S. Z., Antonio, J., & Chavez, C. (2024). Model for reducing mean absolute percentage error through smoothing and time series forecasting in a tourism SME : A case study. *Journal of Machine Intelligence and Data Science (JMIDS)*, *5*, 109–116. https://doi.org/10.11159/jmids.2024.012

Bruno, M. F., Molfetta, M. G., Totaro, V., & Mossa, M. (2020). Performance assessment of ERA5 wave data in a swell dominated region. *Journal of Marine Science and Engineering*, 2–19. https://doi.org/10.3390/jmse8030214

Campos-Caba, R., Alessandri, J., Camus, P., Mazzino, A., Ferrari, F., Federico, I., Vousdoukas, M., Tondello, M., & Mentaschi, L. (2024). Assessing storm surge model performance: What error indicators can measure the model's skill? *Ocean Science*, *20*(6), 1513–1526. https://doi.org/10.5194/os-20-1513-2024

Chan, F. R., Annas, F., Yuspita, Y. E., & Darmawati, G. (2024). Implementation of Convolutional Neural Networks (CNN) in an emotion detection system for measuring learning concentration levels. *Knowbase: International Journal of Knowledge in Database*, *04*(01), 49–61. https://doi.org/10.0.121.7/knowbase.v4i1.8429

Eluri, P. R. R., Annapurnaiah, K., Uday Bhaskar, T. V. S., Kiran Kumar, N., Padmanabham, J., Karthik, T. N. C., Vighneshwar, S. P., Reddem, V. S., Venugopala Rao, V., Murali Krishna, A., Ganesh, M. V, Lotliker, A., Nagaraja Kumar, M., Francis, P. A., Joseph, S., Balakrishnan Nair, T. M., & Srinivasa Kumar, T. (2025). Empowering the Blue Economy: The role of ICT in ocean observations, data, information and advisory services. *CSI Transactions on ICT*, *13*(1), 3–25. https://doi.org/10.1007/s40012-025-00407-x

Fang, H., & Li, T. (2025). Comparative analysis of machine / deep learning models for single-step and multi-step forecasting in river water quality time series. *Water*, *17*(13), 1–14. https://doi.org/10.3390/w17131866

Feng, Z., Hu, P., Li, S., & Mo, D. (2022). Prediction of significant wave height in Offshore China based on

the machine learning method. *Marine Science and Engineering*, *10*(836). https://doi.org/10.3390/jmse10060836

Haditiar, Y., Ikhwan, M., Nanda, M., & Haridhi, H. A. (2024). Understanding sea wave height conditions in sumatra waters. *BIO Web of Conferences*, *02014*. https://doi.org/10.1051/bioconf/20248702014

Han, Y., Zhao, R., Wu, F., Yan, J., & Dong, C. (2025). A two channel optimized SWH deep learning forecast model coupled with dimensionality reduction scheme and attention mechanism. *Ocean Engineering*, *330*, 121217. https://doi.org/10.1016/j.oceaneng.2025.121217

Hao, P., Li, S., & Gao, Y. (2023). Significant wave height prediction based on deep learning in the South China Sea. *Frontiers in Marine Science*, *9*(February), 1–12. https://doi.org/10.3389/fmars.2022.1113788

Ji, Q., Han, L., Jiang, L., Zhang, Y., Xie, M., & Liu, Y. (2023). Short-term prediction of the significant wave height and average wave period based on the variational mode decomposition – temporal convolutional network – long short-term memory (VMD – TCN – LSTM) algorithm. *Ocean Science*, *19*, 1561–1578. https://doi.org/10.5194/os-19-1561-2023

Lee, H., & Ahn, Y. (2025). Comparative study of RNN-based deep learning models for practical 6-DOF ship motion prediction. *Journal of Marine Science and Engineering*, *13*(9), 1–35. https://doi.org/10.3390/jmse13091792

Li, G., Gao, Y., & Yang, H. (2025). Multi-step significant wave height prediction model based on feature enhancement compression , mode decomposition , multi-path convolutional recurrent network and regression correction. *Measurement*, *256*(PD), 118401. https://doi.org/10.1016/j.measurement.2025.118401

Ma, J., Song, N., Nie, J., Ye, M., Liu, X., Yuan, Y., & Wei, Z. (2025). MPST : Effective significant wave height prediction method based on multiscale physical space - time - frequency fusion. *Intelligent Marine Technology and Systems*. https://doi.org/10.1007/s44295-025-00080-5

Maria, E., Wahyono, T., Hartomo, K. D., & Arthur, C. (2025). BiLSTM OptiFlow : An enhanced LSTM model for cooperative financial health forecasting. *Bulletin of Electrical Engineering and Informatics*, *14*(3), 2004–2016. https://doi.org/10.11591/eei.v14i3.8653

Minuzzi, F. C., & Farina, L. (2023). A deep learning approach to predict significant wave height using long short-term memory. *Ocean Modelling*, *181*, 102151. https://doi.org/10.1016/j.ocemod.2022.102151

Munyao, J. N., Oluoch, L. A., Iftikhar, H., & Rodrigues, P. C. (2025). Recurrent Neural Networks for hierarchical Time Series Forecasting: An application to the S&P 500 market value. *Physica A: Statistical Mechanics and Its Applications*, *678*, 130869. https://doi.org/10.1016/j.physa.2025.130869

Omar, M., Yakub, F., Abdullah, S. S., Rahim, M. S. A., Zuhairi, A. H., & Govindan, N. (2024). One-step vs horizon-step training strategies for multi-step traffic flow forecasting with direct particle swarm optimization grid search support vector regression and long short-term memory. *Expert Systems with Applications*, *252*, 124154. https://doi.org/10.1016/j.eswa.2024.124154

Paul, A., Maheswaran, P. A., Satheesan, K., & Kottayil, A. (2025). Assessing ocean reanalysis accuracy for marine extremes in the Indian ocean using in-situ observations. *Climate Dynamics*, *63*(9), 343. https://doi.org/10.1007/s00382-025-07780-y

Sathe, A. M., Supraja, R., & Thomas, A. A. (2025). Results in engineering enhancing agricultural sustainability: Time Series Forecasting with ICEEMDAN-VMD-GRU for economic-resilience. *Results in Engineering*, *27*(June), 106423. https://doi.org/10.1016/j.rineng.2025.106423

Singh, L. K., Garg, H., & Khanna, M. (2023). An artificial intelligence-based smart system for early glaucoma recognition using OCT images. *International Journal of E-Health and Medical Communications (IJEHMC)*, *12*(4), 32–59. https://doi.org/10.4018/IJEHMC.20210701.oa3

Unlu, A. (2025). Comparative analysis of hybrid deep learning models for electricity load forecasting during extreme weather. *Energies*, *18*(12), 1–27. https://doi.org/10.3390/en18123068

Yadav, H., Padavale, V., & Parate, H. (2025). Early natural disaster prediction using machine learning : A comprehensive review. *International Journal of Scientific Research in Engineering and Management (IJSREM)*, *09*(04), 1–14. https://doi.org/10.55041/IJSREM44416

Yao, R., Shao, W., Hu, Y., Xu, H., & Zou, Q. (2025). Numeric modeling of sea surface wave using WAVEWATCH-III and SWAN during tropical cyclones : An overview. *Journal of Marine Science and Engineering*, *13*(8), 1–34. https://doi.org/10.3390/ jmse13081450

Yunita, A., Iqbal, M. H. D., Zaki, M., Ramadhan, H., Akashah, E., Akhir, P., Besse, A., Mansur, F., & Hoirul, A. (2025). MethodsX performance analysis of Neural Network architectures for Time Series Forecasting : A comparative study of RNN, LSTM, GRU, and hybrid models. *MethodsX*, *15*, 103462. https://doi.org/10.1016/j.mex.2025.103462

Zelios, V., Mastorocostas, P., Kandilogiannakis, G., Kesidis, A., Tselenti, P., & Voulodimos, A. (2025). Short-Term electric load forecasting using deep learning : A case study in Greece with RNN, LSTM, and GRU Networks. *Electronics*, *14*(14), 1–24. https://doi.org/10.3390/ electronics14142820

Zhou, J., Liu, Z., Li, C., Du, K., & Yang, H. (2025). Cutting-edge approaches to specific energy prediction in TBM disc cutters: Integrating COSSA-RF model with three interpretative techniques. *Underground Space*, *22*, 241–262. https://doi.org/10.1016/j.undsp.2024.11.004

Zhou, S., Bethel, B. J., Sun, W., Zhao, Y., Xie, W., & Dong, C. (2021). Improving significant wave height forecasts using a joint empirical mode decomposition – Long Short-Term Memory network. *Journal of Marine Science and Engineering*, *9*(744). https://doi.org/10.3390/jmse9070744