Significant Wave Height Forecasting using Long-Short Term Memory (LSTM) in Seribu Island Waters

Husnul Khatimah*[1], Indra Jaya[2], Agus Saleh Atmadipoera[3]
[1,2,3] Marine Science and Technology Department, IPB University, 16680 Bogor, West Java, Indonesia
e-mail: ***[1]khatimahhusnul@gmail.com**, [2]indrajaya@apps.ipb.ac.id, [3]atmadipoera_itk@apps.ipb.ac.id

Abstrak

Gelombang laut (wind waves) merupakan fenomena alam yang utamanya dibangkitkan oleh angin. Informasi tinggi dan periode gelombang sangat diperlukan dalam berbagai bidang kelautan, seperti rekayasa pantai, perikanan, dan transportasi maritim. Namun, untuk memprediksi tinggi gelombang secara akurat masih menjadi tantangan karena sifat stokastik gelombang laut itu sendiri. Beberapa pendekatan untuk memprediksi tinggi gelombang telah dikembangkan, termasuk model numerik dan metode pembelajaran mesin, seperti algoritma Long-Short Term Memory (LSTM) yang saat ini telah mendapatkan cukup banyak perhatian peneliti. Tujuan penelitian ini adalah untuk mengembangkan model prakiraan tinggi gelombang laut menggunakan algoritma LSTM di kawasan perairan Kepulauan Seribu Jakarta. Dataset dari ERA5 terdiri komponen angin zonal dan meridional $(H_{s(wind)})$ *dan tinggi gelombang signifikan* $(H_{s(ERA5)})$*, yang digunakan untuk melatih dan menguji model LSTM. Hasil penelitian menunjukkan bahwa model LSTM dapat memprediksi tinggi gelombang signifikan dengan bagus. Prediksi menggunakan dataset ERA5 terlihat lebih mendekati data lapang menggunakan serta data pengukuran angin dengan Automatic Weather System (AWS)* $(H_{s(AWS)})$ *dengan nilai RMSE, MAE, MAPE berturut-turut yaitu 0,1535 m, 0,1181 m dan 37,11%. Sehingga hasil evaluasi model menunjukkan kinerja yang baik, dengan nilai RMSE, MAE yang relatif rendah, serta MAPE yang baik. Nilai akurasi prediksi tinggi gelombang yang paling tinggi ditemukan untuk prediksi dalam satu minggu (7 hari) ke depan.*

Kata kunci*—Deep Learning, Peramalan, LSTM, Gelombang Laut, Tinggi Gelombang Signifikan*

Abstract

Ocean waves (wind waves) are natural phenomena primarily generated by the wind. Information about wave height and period is highly crucial in various marine fields such as coastal engineering, fisheries, and maritime transportation. However, accurately predicting wave height remains a challenge due to the stochastic nature of ocean waves themselves. Several approaches to predicting wave height have been developed, including numerical models and machine learning methods, such as the Long-Short Term Memory (LSTM) algorithm, which has currently garnered significant attention from researchers. The objective of this research is to develop a forecast model for ocean wave height using the LSTM algorithm in the waters of the Thousand Islands region in Jakarta. The ERA5 dataset comprises zonal and meridional wind components and significant wave height, along with wind measurement data using the Automatic Weather System (AWS) instrument, which was utilized to train and test the LSTM model. The research results demonstrate that the LSTM model can predict significant wave height effectively. Predictions using the ERA5 significant height dataset are observed to be closer to field data, with RMSE, MAE, and MAPE values of 0.1535 m, 0.1181 m, and 37.11% respectively. Thus, the model evaluation results indicate good performance, with relatively low RMSE and MAE values, and a good MAPE value. The highest accuracy in wave height prediction is found for forecasts one week (7 days) ahead.

Keywords*—Deep Learning, Forecast, LSTM, Ocean Wave, Significant Wave Height*

1. INTRODUCTION

Ocean waves can be defined as the rising and falling process of seawater molecules forming crests and troughs [1]. Waves located on the ocean's surface are commonly caused by wind transferring its energy to the water, and big waves, or swells, can travel over long distances. Wind waves are the most dominant waves in the marine meteorological field. The factors determining wind-wave growth are 1) wind speed; 2) The distance over which the wind blows across the water in the same direction (Fetch); 3) the duration or period of wind blowing; 4) water depth (for shallow water). These factors work together to determine the wave's growth. As the factors increase, so does the size, length, and speed of the resulting waves.

Wave height is an important design parameter in coastal and marine engineering design, fisheries, exploration, power generation, and maritime transportation [2,3]. Information about wave height can be used in determining the layout of ports, shipping lanes, and coastal structures for disaster mitigation such as coastal abrasion, thus having significant implications for national security. The limited in-situ observations carried out to observe the ocean in general, as well as the limited and discontinuous data from measurements and observations of sea waves, make wave height models providing current wave prediction information [4]. However, in reality, making accurate predictions for sea wave height information is very difficult due to the stochastic nature of ocean waves themselves [2].

Several methods have been developed for wave height prediction using empirical methods, such as the Sverdroup-Munk-Bretscheider (SMB) method and the Darbyshire method [5] used for wave height significant prediction. However, empirical methods cannot describe sea surface wave conditions in detail. Numerical models have been proven effective in predicting wave height over large spatial and temporal ranges. However, when computing large models, high-performance computers and significant time costs are required due to the large amount of data and the complexity of the calculations, especially for higher-resolution grid computations in nearshore zones with complex seafloor topography [6,7,8]. The principle of numerical models is to obtain wave height, period, and other information by solving the wave spectrum equations from marine physical processes. Third-generation models like the Wave Model (WAM) [9], WAVEWATCH-III [10,11], and Simulating Waves Nearshore (SWAN) [12,13] are among the most advanced numerical models used. The WAM and WAVEWATCH-II models have similar structures, using wave spectrum equations to simulate waves. Some studies show that the SWAN model accurately simulates waves in coastal regions when boundary conditions are accurate. However, the fixed energy spectrum equations with fixed expressions hardly characterize the complex and changeable ocean environment in a comprehensively fine way [22].

With the advancement of technology in artificial intelligence, especially in machine learning, many machine learning models are used for wave height prediction by learning the data variability used, such as the Backpropagation Neural Network–Mind Evolutionary algorithm [8], JST backpropagation [14] and Regression Neural Network algorithm [15]. Research on wave height prediction using Artificial Neural Networks (ANNs) continues to evolve, from simple machine learning algorithms to deep learning with more complex algorithms. Recently, the Long Short-Term Memory (LSTM) algorithm, which is a reconstruction of the Recurrent Neural Network (RNN), has attracted much attention from researchers for predicting time series data in weather [16,17] and metocean forecasting, including significant wave height prediction [2,18,19]. Some studies also compare the performance of LSTM with other RNN algorithms [20] and even combine LSTM with other models to enhance its performance [21,22,23]. The goal of this research is to develop a significant wave height forecasting model using the LSTM algorithm.

The Seribu Islands Jakarta waters were opted as a study area to perform sea wave height predictions using the LSTM deep learning method. The Seribu Islands' waters are open to the Java Sea and Natuna Sea through the Karimata Strait, so monsoonal wind conditions and sea surface waves vary depending on the monsoonal period and other weather factors. The peak of the northwest monsoon (NWM) period occurs in December-February when strong and stable northwest monsoonal winds blow. The sea waves and swells are generally high during this period. The maximum wave height from the swell generally occurs in this period, related to the development of tropical cyclones in the South China Sea and Natuna Sea [24,25]. In contrast, during the southeast monsoon (SEM) period, easterly – southeasterly monsoonal winds are fully developed during July-September, generating much higher mean wave height since wave fetch is much farther across the Java Sea than during the NWM period [24]. Information on surface waves and direction in the study area is needed by the Seribu islands community, such as for small fisheries activity, sea transportation between islands and Jakarta ports, as well as coastal islands recreations. The objectives of this study are to apply the LSTM method for developing surface wave prediction in the Seribu islands and to evaluate the predicted surface waves in terms of statistical quantification.

2. METHODS

2.1 Datasets and Area of Study

The dataset was collected from the official Copernicus Climate Data Store (CDS) website implemented by the ECMWF (European Center for Medium Range Forecast). The dataset used for predicting significant wave height consists of zonal and meridional components (U10 and V10 components) and significant wave height data from ERA5 reanalysis. ERA5 is the fifth generation ECMWF reanalysis for the global climate and weather for the past 4 to 7 decades and ERA5 is the fifth-generation reanalysis data from ECMWF for global climate and weather over 4 to 7 decades [26]. Wind component and significant wave height data were collected from ERA5 ECMWF data every hour from 1997 to 2022 located in the North Java Waters. This dataset was chosen because ERA5 data is an approximation of actual sea conditions and provides comprehensive global data publicly available, it allows any user to test the present method and provides long-term data that provides comprehensive and open global data [19].

Figure 1. Research Location

As for field data for comparison, we used Automatic Weather Station (AWS) instrument data installed in the Semak Daun Island, Seribu Islands, DKI Jakarta during the period from March 2022 to June 2022. AWS data used is wind direction and speed data. The research locations can be seen in Figure 1.

2.2 Methodology

In general, the step for significant wave height forecasting using Long-Short Term Memory (LSTM) can be seen in Figure 2.

Figure 2. Research flowchart

2. 3 Preprocessing Data

2. 3.1 UV Komponent Conversion

For datasets derived from wind components, Data preprocessing started with converting wind components (U, V) into significant wave height $(H_{s(wind)})$ data using the SPM (1984) formula. Additionally, field data from AWS is transformed into significant wave height $(H_{s(AWS)})$. The overall process of converting significant wave height data based on wind components can be visualized in Figure 3.

Figure 3. Converting process for zonal and meridional components

2.3.2 Data Normalization

Data normalization is used to prevent data with large values from dominating data with a small value range. The normalization method used is min-max normalization with a range of 0-1. The min-max normalization formula can be seen in the following equation [27]:

Where :

x = original value of the data

$\min(x)$ = minimum value of the data

$\max(x)$ = maximum value of the data

2. 4 LSTM

Long Short-Term Memory (LSTM) is a variant of the Recurrent Neural Network (RNN) unit [28]. Like RNN, LSTM networks also consist of modules with recurrent processing and the main application of RNN is in the processing and prediction of sequence data. However, RNNs suffer from the problems of exploding and vanishing gradients [2]. The LSTM network has a significant advantage over the traditional RNN network in finishing the problem by replacing the hidden layer with a memory cell that can store information for a long period [29]. The architecture of LSTMs is built as follows: inside the memory block, there are one or more central cells that are self-looped into three multiplicative units called input, output and the forget gates [19]. The input gate shields the memory block from irrelevant inputs, while the output gates protect other units from irrelevant information within the current block [28].

Figure 4. Schematic Diagram of LSTM unit [2]

For a sequence of input x_i, the value of x can influence LSTM output or spatial distance h_j. The training process learns the quantitative relationship of this influence. The LSTM structure resembles a square matrix with the central part being the LSTM units except for input x_t and output h_t. LSTM units are connected end-to-end. The same layer uses the output from the previous unit as input to the last unit. The output from the previous layer is used as input to the next layer, and finally, the output is obtained. A schematic diagram of the LSTM unit can be described in Figure 4.

2.5 Model Validation

To quantify the accuracy and performance of the built model in predicting significant wave height, we used three metric evaluations, i.e. root mean square error (RMSE), mean absolute error (MAE), and mean absolute error percentage (MAPE) [30]. RMSE is chosen because it is considered more sensitive to large deviations between predicted values and actual results. Each metric is presented in the following formulas:

Where :

x_i = the actual value.

$\hat{x}_i$ = the predicted value.

n = the total number of samples.

3. RESULT AND DISCUSSION

3.1 Preprocessing Data

Before entering the LSTM process, U10 and V10 components from ECMWF-ERA5 data are converted into significant wave height data $(H_{s(wind)})$ based on the equation in Figure 1. Then, it becomes the input data for training and validating the LSTM model using wind datasets$(H_{s(wind)})$. Next, both the significant wave height from wind $(H_{s(wind)})$ and significant wave height data from ERA5 $(H_{s(ERA5)})$ are normalized by scaling the datasets to the range (-1,1). This rescaling is important because LSTM cells using the sigmoid function are sensitive to a relatively large number of changes, and rescaling is necessary to avoid issues during the training process [31]. Figure 4 describes the difference between the data before normalization (4a) and after normalization (4b). There is no change in the data pattern; only changes in values occur within the range of -1 to 1.

(a)

(b)

Figure 4. Before and after data normalization

2.1 Significant Wave Height Prediction

The LSTM model is optimized using the Adam optimizer function [32]. Several hyperparameters are utilized for applying the LSTM model, such as the number of hidden layers, epochs, batch size, and learning rate. The number of hidden layers represents the number of computations in the training process, the batch size is used to control how often the network's weights are updated, the learning rate determines the learning speed [33] and epochs refer to the number of iterations or training cycles performed by the algorithm [34]. It is important to note that too many epochs can lead to overfitting [35].

The training $(H_{s(wind)})$ dataset for 1997-2021 using 100 epochs shows a final training loss of 0.00064. Training loss measures how the model predicts the given training datasets. Then, the trained data is tested using the 2022 data. Figure 5 shows a comparison between the training dataset (black), significant wave height data $(H_{s(wind)})$ in 2022 (red), and the LSTM prediction results (blue). Visually, the LSTM prediction results can follow the same trendline pattern as the significant wave height data converted from wind speed data in 2022. This indicates that the prediction results can follow the patterns formed from the training dataset and learn event patterns over time. LSTM models were designed to learn and remember over a long sequence and do not depend on pre-specified lagged observations. Hence, it can produce consistently good results for higher lead times [36].

The training $(H_{s(ERA5)})$ dataset using the same hyperparameters shows a final training loss of 0.0017. These training loss and validation loss values indicate that the model built with the specified hyperparameters performs well in predicting the datasets. Figure 6 shows a comparison between the training dataset (black), actual significant wave height data converted from wind speed in 2022 (red), and LSTM prediction results (blue). Visually, it can be seen that the LSTM prediction results can follow the same pattern as the actual significant wave height data in 2022. This indicates that the prediction results can follow the patterns formed from the training dataset and learn event patterns over time

Figure 5. Significant wave height ($H_{s(wind)}$) forecasting performance using LSTM

As mentioned in the introduction, wind significantly influences wave height conditions in a region. The Asian-Australian monsoon system influences the climate in Indonesia, where the wind blows from Indonesia to Australia during the northwest monsoon (NEM) from December to February, causing a rainy season in Indonesia. In contrast to the southeast monsoon (SEM) during June-August, the wind blows from Australia to Indonesia, and Indonesia has a dry season. Monsoon winds blow periodically and vary in direction and intensity each season [37]. This monsoon system will further affect the variability of the oceanographic parameters [38].

From the forecasting of the LSTM model results, for both the ($H_{s(wind)}$) and the ($H_{s(ERA5)}$) dataset is not significantly different from the actual data (Table 1 and Table 2). Whereas, based on the table, we can observe the seasonal variation of significant wave heights. The average significant wave height from all datasets shows that during the southeast monsoon, the average significant wave height is higher than during the northwest monsoon. This is consistent with the research [4,39] using ECMWF-ERA5 data, which stated that the seasonal average significant wave height during SEM in the North of Java is higher than during the NWM.

Figure 6. Significant wave height $(H_{s(ERA5)})$ forecasting performance using LSTM

The difference in average significant wave height occurs because, in the NWM, there is an active Asian Monsoon, where areas with high average wave height are generally located to the north or in areas bordering open seas, such as the Karimata Strait, Natuna Sea, and the western region of Sumatra Island. Conversely, in the SEM, there is an active Australian Monsoon, and high waves occur in the sea to the south. This is related to fetch length, wind speed, and wind duration. In the NWM, the wind blows from Asia towards Australia; north of the equator, the wind blows strongly for a long duration, resulting in a longer fetch and thus higher waves. Meanwhile, south of the equator, the wind blows weaker due to obstruction (deflection) when passing through areas caused by the force generated by the Earth's rotation. Therefore, when entering waters to the south of the equator, the fetch formed is shorter compared to the north, while in the SEM where the Australian Monsoon blows towards Asia, the southern equatorial region has a longer fetch, resulting in high waves occurring in the southern waters

Table 1. The average of significant wave height from $(H_{s(wind)})$ dataset

Table 2. The average of significant wave height from ($H_{s(ERA5)}$) dataset

The field data used is the significant wave height data from the instrument Automatic Weather Station (AWS) installed at KJA Balai Sea Farming, Semak Daun Island, during March-June 2022. Figure 7 shows the visualization of the comparison between significant wave heights from $H_{s(AWS)}$ and significant wave height from $H_{s(wind)}$ and $H_{s(ERA5)}$. The LSTM predictions for significant wave height $H_{s(ERA5)}$. using the ERA5 dataset are influenced by global factors and data resolution differences. The reanalysis results from global data account for complex weather patterns and ocean circulation, offering predictions closer to actual conditions due to higher spatial resolution compared to field data. However, the location of the AWS instrument and data retrieval from ECMWF-ERA5 can impact results, especially in sheltered areas like Semak Daun Island's shallow water, potentially leading to discrepancies in wind speed data recorded by AWS compared to global reanalysis data. Therefore, while LSTM predictions exhibit similar patterns across datasets, variations stemming from global factors, data resolution, and data retrieval locations should be considered when interpreting significant wave height predictions.

Figure 7 Visualization of the comparison between significant wave height from $H_{s(AWS)}$ and significant wave height from $H_{s(wind)}$ and $H_{s(ERA5)}$

2.3 Model Validation

The LSTM model's performance is evaluated using RMSE, MAPE, and MAE metrics. Table 3 displays differences in these metrics for $H_{s(wind)}$ and $H_{s(ERA5)}$ datasets across various prediction timeframes (1 year, 6 months, 3 months, 1 month, and 7 days). The model's accuracy is evident from the low RMSE, MAE, and MAPE values, indicating precise predictions of significant wave heights. The highest accuracy in wave height prediction is found for forecasts one week (7 days) ahead. Furthermore, comparing the model's predictions with field data, the ERA5 dataset $H_{s(ERA5)}$ closely matches with field data ($H_{s(AWS)}$), with RMSE, MAE, and MAPE of 0.1535, 0.1181, and 37.11, respectively.

Tabel 3 LSTM Model Validation

4. CONCLUSION

The LSTM model constructed was trained using different datasets, namely based on the significant wave height conversion results from ERA5 wind component (Hs) and significant wave height data from ERA5. The predicted significant wave height based on the LSTM model showed good performance with the actual data. The seasonal mean significant wave height from LSTM forecasting also matches the average actual significant wave height data. The model performance was evaluated using RMSE, MAE, and MAPE metrics. The evaluation results indicated that the LSTM model provided good results in predicting significant wave height, with the lowest RMSE, MAE, and MAPE values for the 7-day prediction being 0.0095, 0.0071, and 2.052, respectively. Thus, the use of the LSTM model for predicting significant wave height yielded satisfactory results, although the result can be influenced by various factors such as data quality and assumptions used in the model. Furthermore, comparison with field data ($H_{s(AWS)}$) also indicated a good fit, with predictions using Hs $H_{s(ERA5)}$ data showing the highest accuracy, resulting in RMSE, MAE, and MAPE values of 0.1535, 0.1181, and 37.11, respectively.

REFERENCE

[1] B. Utoyo, "Geografi Membuka Cakrawala Dunia," PT. Setia Purna Inves, Bandung, 2007.

[2] S. Fan, N. Xiao, and S. Dong, “A novel model to predict significant wave height based on long short-term memory network,” *Ocean Eng.*, vol. 205, Jun. 2020, doi: 10.1016/j.oceaneng.2020.107298.

[3] I.R. Young and A. Ribal, "Multiplatform evaluation of global trends in wind speed and wave height," *Science*, vol. 364, pp. 548–552, 2019, doi:10.1126/science.aav9527

[4] A.W. Pramita, D.N. Sugianto, I.B. Prasetyawan, R. Kurniawan, A.S. Praja, "Pola Tinggi Gelombang di Laut Jawa Menggunakan Model Wavewatch-III," *Jurnal Meteorologi dan Geofisika*, vol. 21, no. 1, pp. 21-28, 2020.

[5] P. Purwanto, R. Tristanto, G. Handoyo, M. Trenggono, and A. A. D. Suryoputro, "Analisis Peramalan dan Periode Ulang Gelombang di Perairan Bagian Timur Pulau Lirang, Maluku Barat Daya," *Indonesian Journal of Oceanography*, vol. 2, no. 1, pp. 80-89, Mar. 2020. [Online]. Available: https://doi.org/10.14710/ijoce.v2i1.7481

[6] H. Yoon, S.C. Jun, Y. Hyun, G.O. Bae, K.K. Lee, "A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer," *J. Hydrol.*, vol. 396, pp. 128–138, 2011. [Online]. Available: https://doi.org/10.1016/j.jhydrol.2010.11.002

[7] H.A.A. Guner, Y. Yuksel, and E.O. Cevik, "Estimation of wave parameters based on nearshore wind-wave correlations," *Ocean Eng.*, vol. 63, pp. 52–62, 2013. https://doi.org/10.1016/j.oceaneng.2013.01.023

[8] W. Wang, R. Tang, C. Li, P. Liu, L. Luo, "A BP neural network model optimized by Mind Evolutionary Algorithm for predicting the ocean wave heights," *Ocean Eng.*, vol. 162, pp. 98–107, 2018. [Online]. Available: https://doi.org/10.1016/j.oceaneng.2018.04.039.

[9] A.B. Bottcher, B.J. Whiteley, A.I. James, J.G. Hiscock, "Watershed Assessment Model (WAM): Model Use, Calibration, and Validation," *Trans. ASABE*, vol. 55, pp. 1367–1383, 2012. [Online]. Available: doi: 10.13031/2013.42248

[10] H.L. Tolman, "User manual and system documentation of WAVEWATCH III TM version 3.14," *Technical note, MMAB Contrib.*, no. 276, pp. 220, 2009.

[11] L. Mentaschi, G. Besio, F. Cassola, A. Mazzino, "Performance evaluation of Wavewatch III in the Mediterranean Sea," *Ocean Model.*, vol. 90, pp. 82–94, 2015. https://doi.org/10.1016/j.ocemod.2015.04.003

[12] A. Akpınar, G.P. van Vledder, M.I. Kömürcü, M. Özger, "Evaluation of the numerical wave model (SWAN) for wave simulation in the Black Sea," *Cont. Shelf Res.*, vol. 50, pp. 80–99, 2012. [Online]. Available: https://doi.org/10.1016/j.csr.2012.09.012

[13] B. Liang, H. Gao and Z. Shao, "Characteristics of global waves based on the third-generation wave model SWAN," *Mar. Struct.*, vol. 64, pp. 35–53, 2019. [Online]. Available: https://doi.org/10.1016/j.marstruc.2018.10.011

[14] N. Rahmadani, B. Darma Setiawan, and S. Adinugroho, "Prediksi Ketinggian Gelombang Laut Menggunakan Metode Jaringan Saraf Tiruan Backpropagation," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 7, pp 6517-6525, 2019.[Online]. Available: http://j-ptiik.ub.ac.id.

[15] V. Juliani, D. Aditya and Adiwijaya, "Wave Height Prediction based on Wind Information by using General Regression Neural Network, Study Case in Jakarta Bay," *IEEE Explorer*, doi: 10.1109/ICoICT49345.2020.9166305, 2020.

[16] A. Huang, B. Vega-Westhoff and R.L. Sriver, "Analyzing El Niño–Southern Oscillation Predictability Using Long-Short-Term-Memory Models," *Earth Sp Sci*., vol. 6, no. 2, pp. 212–221, 2019. [Online]. Available: https://doi.org/10.1029/2018EA000423.

[17] M. Rizki, S. Basuki and Y. Azhar, "Implementasi Deep Learning Menggunakan Arsitektur Long-Short Term Memory Untuk Prediksi Curah Hujan Kota Malang," *Repositor*, vol. 2, no. 3, pp. 331–338, 2020.

[18] S.B. Park, S.Y. Shin, K.H. Jung, B.G. Lee, "Prediction of Significant Wave Height in Korea Strait Using Machine Learning," *J Ocean Eng Technol*., vol. 35, no. 5, pp. 336–346, 2021. [Online]. Available: doi: 10.26748/ksoe.2021.021.

[19] F.C. Minuzzi and L. Farina, "A deep Learning Approach to Predict Significant Wave Height using Long Short-Term Memory," *Ocean Modelling*, vol. 181, 2023. [Online]. Available: https://doi.org/10.1016/j.ocemod.2022.102151.

[20] P. Hao, S. Li, and Y. Gao, "Significant Wave Height Prediction Based on Deep Learning in The South China Sea," *Front. Mar. Sci*., vol. 9, pp. 1113788, 2023. [Online]. Available: doi: 10.3389/fmars.2022.1113788.

[21] T. Song, J. Jiang, W. Li, D. Xu, "A Deep Learning Method With Merged LSTM Neural Networks for SSHA Prediction," *IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, vol. 13, pp. 2853–2860, 2020.

[22] X. Zhang, Y. Li, S. Gao, P. Gen, "Ocean Wave Height Series Prediction with Numerical Long Short-Term Memory," *Marine Science and Engineering*, vol. 9, no. 5, pp. 514, 2021. [Online]. Available: https://doi.org/10.3390/jmse9050514.

[23] A. Ahmed, J.J. Jui, M.S. Al-Musaylh, N. Raj, R. Saha, R.C. Deo, S.K. Saha, "Hybrid Deep Learning Model for Wave Height Prediction in Australia’s Wave Energy Region," *Applied Soft Computing Journal*, vol. 150, pp. 111003, 2023. [Online]. Available: https://doi.org/10.1016/j.asoc.2023.111003.

[24] M. N. Habibie and W. Fitria, “Kajian Indeks Variabilitas Tinggi Gelombang Signifikan Di Indonesia,” *J. Segara*, vol. 14, no. 3, pp. 159–168, 2019, [Online]. Available: 10.15578/segara.v14i3.6650.

[25] Purwanto *et al.*, “Seasonal variability of waves within the Indonesian seas and its relation with the monsoon wind,” *Ilmu Kelaut. Indones. J. Mar. Sci.*, vol. 26, no. 3, pp. 189–196, 2021, [Online]. Available: 10.14710/ik.ijms.26.3.189-196.

[26] H. Hersbach, B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, et al., "The ERA5 global reanalysis," *Q J R Meteorol. Soc*., vol. 146, no. 730, pp. 1999–2049, 2020. [Online]. Available: doi: 10.1002/qj.3803.

[27] Y.K. Jain and S.K. Bhandare, "Min-max normalization based data perturbation method for privacy protection," *International Journal of Computer & Communication Technology (IJCCT)*, vol. 2, no. 8, pp. 45–50, 2011. [Online]. Available: 10.47893/ijcct.2013.1201

[28] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput*., vol. 9, pp. 1735–1780, 1997. [Online]. Available: https://doi.org/10.1162/neco.1997.9.8.1735.

[29] N. K. Manaswi, "Deep Learning with Applications Using Python," Apress, 2018.

[30] J. Shi et al., “A Machine-Learning Approach Based on Attention Mechanism for Significant Wave Height Forecasting,” J. Mar. Sci. Eng., vol. 11, no. 9, Sep. 2023, doi: 10.3390/jmse11091821.

[31] I. Delgado and M. Fahim, "Wind turbine data analysis and LSTM-based prediction in SCADA system," *Energies*, vol. 14, no. 1, pp. 125-140, 2021. [Online]. Available: https://doi.org/10.3390/en14010125.

[32] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *ICLR (International Conference for Learning Representations)*. [Online]. Available: http://arxiv.org/abs/1412.6980

[33] D. Z. Haq, DCR. Novitasari, A. Hamid, N. Ulinnuha, Arnita, Y. Farida, D. Nugraheni, R. Nariswari, Ilham, H. Rohayani, and others, "Long Short-Term Memory algorithm for rainfall prediction based on El-Nino and IOD data," *Procedia Comput. Sci*., vol. 179, pp. 829–837, 2021. doi: 10.1016/j.procs.2021.01.071. [Online]. Available: https://doi.org/10.1016/j.procs.2021.01.071.

[34] Y. J. Wai, Zulkarnain, S. Irwan, and LK. Chuan, "Fixed point implementation of Tiny-Yolo-v2 using opeCL on FPGA," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 10, pp. 506-512, 2018. [Online]. Available: 10.14569/IJACSA.2018.091062

[35] Y. Bai, Y. Guo, Q. Zhang, B. Cao, and B. Zhang, "Multi-network fusion algorithm with transfer learning for green cucumber segmentation and recognition under complex natural environment," *Computers and Electronics in Agriculture*, vol. 194, pp. 106789, 2022. doi.org/10.1016/j.compag.2022.106789. [Online]. Available: https://doi.org/10.1016/j.compag.2022.106789.

[36] M. M. Puspham and F. Enigo, "Forecasting Significant Wave Height using RNN-LSTM Models," *Proceedings of the International Conference on Intelligent Computing and Control Systems*, pp. 1141-1146, 2020. ISBN: 978-1-7281-4876-2. [Online]. Available: 10.1109/ICICCS48265.2020.9121040

[37] P. J. Webster, "Dynamics of the Tropical Atmosphere and Oceans (Advancing Weather and Climate Science): 1st Edition," Georgia Institute of Technology, 2020.

[38] T. Qu, Y. Dua, J. Strachan, G. Meyers, and J. Slingo, "Sea surface temperature and its variability in the Indonesian region," *Oceano*, vol. 18, pp. 50-6, 2005. [Online]. Available: 10.5670/oceanog.2005.05

[39] I. A.H Aswad, H. D. Armono, S. Rahmawati, A. Rodlwan, and R. M. Arifieanto, "Modeling Wave Height for Wave Energy Study in the Western Waters of Lampung Province," *J. Ilm. Teknol. Marit*., vol. 15, no. 2, pp. 75-84, 2021. [Online]. Available: 0.29122/jurnalwave.v15i2.4958