Accredited Ranking SINTA 2
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 6 No. 5 (2022) 776 - 781 ISSN Media Electronic: 2580-0760

Time Series Forecasting of Significant Wave Height using GRU, CNN-GRU, and LSTM

Cornelius Stephanus Alfredo[1], Didit Adytia[2]
[1,2]Informatics, School of Computing, Telkom University
[1]sinambelax@student.telkomuniversity.ac.id, [2]adytia@telkomuniversity.ac.id*

Abstract

Predicting wave height is essential to reduce significant risks for shipping or activities carried out at sea. Waves inherit a stochastic nature, mainly generated by wind and propagated through the ocean, making them challenging to forecast. In this paper, we design time series wave forecasting using a deep learning model, which is a hybrid Convolutional Neural Network (CNN)-Gated Recurrent Unit (GRU) or CNN-GRU. We use two time series of wave data sets, i.e., reanalysis data from ERA5 by ECMWF and GFS from NOAA. As a study area, we choose Pelabuhan Ratu, located in the south of West Java which is connected to the open Indian Ocean. Moreover, we also compare the results by using other deep learning models, i.e., the Long Short-Term Memory (LSTM) and GRU. We evaluated these models to forecast 7, 14, and 30 days. Models' performance is assessed using RMSE, MAPE, and Correlation Coefficient (CC). For predicting 30 days, using the ERA5 data, the CNN-GRU model produces relatively accurate results with an RMSE value of 1.8844 and CC of 0.9938, whereas for the GFS data, results in RMSE value of 1.8852 and CC of 0.9915.

Keywords: *wave height, pelabuhan ratu, cnn-gru, long short-term memory, gated recurrent unit*

1. Introduction

Indonesia is the world's biggest archipelagic country, located between two continents, Asia and Australia, and links the two major oceans, the Indian and Pacific oceans[1]. More than 17,000 islands are connected by the sea, which covers 70% of the total area of Indonesia. Wave height information is critical for designing and planning many sea-related operations in maritime constructions [2]. Forecasting sea wave height is particularly vital for ship navigation, loading and unloading big ships in tidal zones, forecasting and mitigating flood damage in low-lying coastal communities, and early warning of disasters along the shore [3]. As a result, wave height prediction is necessary; sea wave height prediction may be produced using wind information; wind pressure on ocean waves can create fluctuations in wave height [4].

There are many variations in wave height, one of which is caused by wind pressure which makes the waves have different variations every time, making it difficult to predict [5]. Numerous techniques to wave height prediction have been presented, including machine learning, soft computing, and numerical methods [6]. Unlike machine learning, the numerical method's approach usually requires a high computational cost to produce good wave prediction, which requires an affordable computational cost [7]. Several works on wave height prediction have been published in the literature, employing a machine learning approach because it can efficiently map large data sets to suitable forecasts and has been widely used for forecasting in recent times [8]. Mandal et al. [9], analyzing wave forecasting in the ocean off the west coast of India using an artificial neural network with the RPROP update algorithm. The wave forecasting carried out is 3, 6, 12, and 24 hours, producing a correlation coefficient of 0.95, 0.90, 0.87, and 0.73. This study shows that iterative neural networks' wave forecasting produces better results.

Anggraeni et al. [10] compared the XGBoost and AdaBoost methods for wave height predictions for the next seven days, 14 days, 30 days, 45 days, and 60 days in the Pangandaran region of Indonesia. The data was used for five years with training data for 4.5 years and six months as test data. They concluded that the XGBoost produces better results than the AdaBoost, with predictive evaluation metrics RMSE value of 0.093 and correlation coefficient (CC) value of 0.989, whereas AdaBoost produces RMSE value of 0.110 and

Accepted: 21-06-2022 | Received in revised: 26-09-2022 | Published: 31-10-2022

CC value of 0.985. Kaloop et al. [11], Predicting wave height on the southeast coast of the US using the Wavelet, Particle Swarm Optimization (PSO), and Extreme Learning Machine (ELM) methods with the results of prediction evaluation metrics in the next 12 hours get RMSE 0.0495 and MAE 0.363. In their paper, the results obtained using the WPSO-ELM method get better results for predicting wave heights for coastal and deep-sea areas up to a grace time of 36 hours. M. Sajjad et al. [12] predicted electricity consumption in residential buildings with the CNN-GRU hybrid model. They conclude that the CNN-GRU model performed better than the LSTM model.

In this study, we choose Pelabuhan Ratu, in West Java, Indonesia, as a study case. The harbor has a complex geometry that facing directly to the Indian Ocean. Moreover, the harbor is densely used by fisherman, and logistic shipping, making the harbor an active port in Southern Java Island. Since the port is facing the Indian Ocean, the wave height in Pelabuhan Ratu is relatively high. Therefore, wave forecasting is crucial for scheduling of marine activities in the port.

This paper aims to build a time series forecasting architecture using a deep learning hybrid Convolutional Neural Network (CNN)-Gated Recurrent Unit (GRU) model to predict wave heights of 7, 14, and 30 days at Pelabuhan Ratu. This research will compare the performance of Hybrid CNN-GRU to GRU and LSTM to show how well the performance of Hybrid CNN-GRU is in making wave height predictions with the other two methods. The purpose of using GRU and LSTM is because GRU is the base model of Hybrid CNN-GRU, and GRU is an improvement from LSTM. Moreover, we use two wave data, i.e., the ERA5 from ECMWF [13] and GFS from NOAA [14].

The structure of this research is as follows. Section 2 describes the prediction system methodology used in this study, and Section 3 presents the results and discussion of the research conducted. Furthermore, in Section 4, conclusions and suggestions are given based on the results of our experiment.

2. Research Methods

The research procedure carried out in this paper is shown in Figure 1. First, we included significant wave height data from two sources, ERA5 from ECMWF and NOAA's GFS. Second, preprocessing is carried out to convert the time data on the wave height into an index. The data is divided into two parts, 20% test data and 80% training data, from the data collection period. Third, the selection of the best parameters to be implemented with Hyperparameter Tuning to improve performance on the CNN-GRU, LSTM, and GRU models by using the grid search method. Parameter value in the batch and epoch uses multiples of 10 starting from 10 to 100 for the unit we try with values in multiples of 50 ranging from 1 to 100. CNN-GRU, LSTM, and GRU models can have different layers, units, solids, Batch Size, Epoch, and Activation. Fourth, the model predicts the wave height and then looks at the expected accuracy. If the results are not good, then the model is re-implemented, but if it is good, then a performance evaluation is carried out using RMSE, MAPE, and CC to find the best machine learning model.

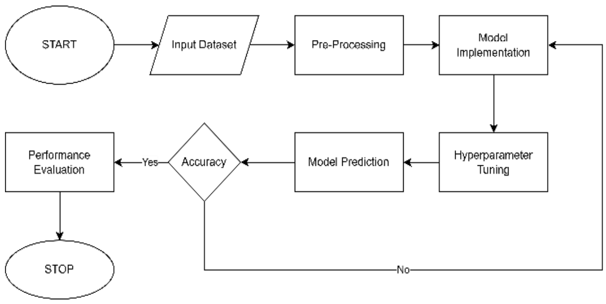

Figure 1. Methodology for Wave Height Forecasting

2.1 Dataset

The research location is in Pelabuhan Ratu, West Java, Indonesia. The two data sets differ, i.e. the ERA-5 with a seven-year wave height from 2014 to 2021, with hourly wave height information. The dataset derived from GFS has five-year wave height data from 2014 to 2019, with wave height information every three hours. Both data sets have the right attributes, namely time and wave height. Figures 2 and 3 are wave height visualizations from the ERA-5 and GFS datasets.

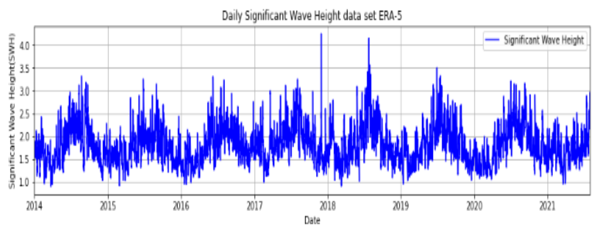

Figure 2. Significant wave height data from ERA5 ECMWF.

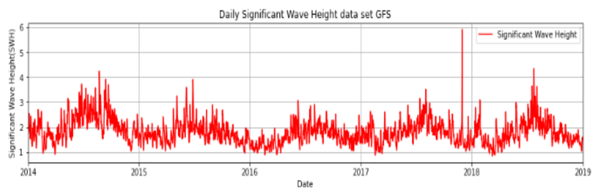

Figure 3. Significant wave height data from GFS NOAA.

2.2 Long Short-Term Memory

Sepp Hochreiter and Jürgen Schmidhuber first suggested the concept of Long Short-Term Memory (LSTM) in 1997 [15]. LSTM is a type of Recurrent Neural Network (RNN) that can remember values from an early stage for future use [16]. The LSTM has a hidden layer consisting of one or more memory cells,

DOI: https://doi.org/10.29207/resti.v6i5.4160

with each memory cell including a forget gate, input gate, and output gate [17]. As seen in Figure 4, each memory cell in the LSTM employs 3 components gate unit that seeks to eliminate disparities in the weight of input values [15].

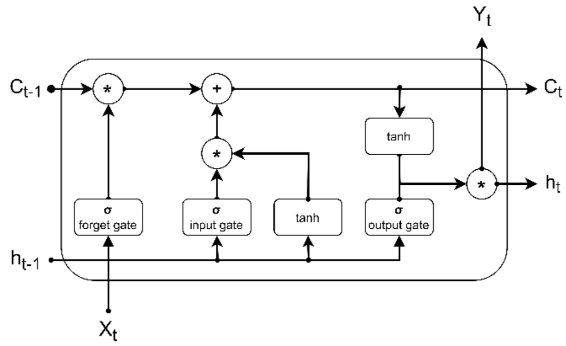

Figure 4. Long Short-Term Memory Architecture.

The first component is the forget gate, which determines which information to store and which to discard using the sigmoid function (σ); if f_t = 0, the data is discarded, but if f_t > 0, the information is saved based on the ratio of memory. The forget gate is characterized by Equation (1).

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

where X_t and h_{t-1} indicate the current input value. W_f and b_f are training coefficients whose values will be calculated iteratively. The process then proceeds to input gates. In this step, decide and save the information from the new input (X_t) in the cell state and update the cell state. This stage consists of the sigmoid layer and the tanh layer. First, the sigmoid layer determines if the new information should be updated or ignored (0 or 1) can be seen on Equation (2). Then the tanh function assigns weight to the data that passed by, determining their level of importance (1 to 1) as seen in Equation (3). The two values are multiplied to update the cell state. This new memory is subsequently added to the existing memory C_{t-1}, resulting in C_t will be found from Equation (4).

$$i_t = \sigma(W_t.[h_{t-1}, X_t] + b_t) \tag{2}$$

$$N_t = \tanh(W_t.[h_{t-1}, X_t] + b_t) \tag{3}$$

$$C_t = C_{t-1}f_t + N_t i_t \tag{4}$$

where C_{t-1} and C_t represent the cell states at times $t-1$ and t, whereas W_t and b_t represent the cell state's weight matrices and bias, respectively. Finally, the last component of a gate unit is the output gate, which determines which data can travel through the output using the sigmoid function.

$$O_t = \sigma(W_o.[h_{t-1}, X_t] + b_o) \tag{5}$$

$$h_t = O_t \tanh(C_n) \tag{6}$$

Where in Equation (5) W_o the weight matrices and b_o bias of the output gate.

2.3 Gated Recurrent Unit

Chung conceptualized the Gated Recurrent Unit (GRU) model in 2014 to create various units with various time scales. The GRU is a gate unit that modulates the flow of information within the unit without a separate memory cell [18]. GRU is a method of Recurrent Neural Network (RNN) [19]. GRU has a more straightforward underlying structure than LSTM, making it easier to train and requiring fewer computations [20]. GRU does not save information using cell state but instead concealed condition. The reset gate of the GRU controls whether new information should be lost, whereas the update gate is for remembering [21]. The GRU model architecture is shown in Figure 5.

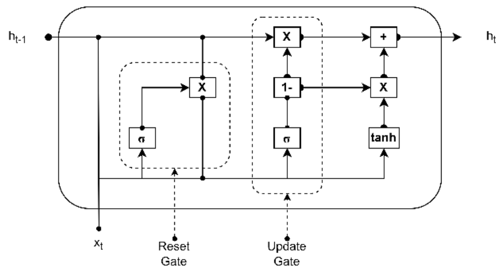

Figure 5. Gated Recurrent Unit Architecture.

The update gate helps the GRU model determine how much past information needs to be passed into the future, as shown in Equation (7).

$$z_t = \sigma\left(W^{(z)}x_t + U^{(z)}h_{t-1}\right) \tag{7}$$

where x_t is the input value, it is multiplied by its own weight $W^{(z)}$. The same holds true for h_{t-1}, which stores information about the preceding $t-1$ units and is multiplied by its own weight $U^{(z)}$. The reset gate selects how much of the previous knowledge to forget. This formula is the same as one for update gate; the distinction is in weights and gate usage expressed by Equation (8).

$$r_t = \sigma\left(W^{(r)}x_t + U^{(r)}h_{t-1}\right) \tag{8}$$

$$h_t = tanh(W_t + r_t \oplus U h_{t-1}) \tag{9}$$

where in Equation (9) h_{t-1} is hidden state, U is weight matrix.

2.3. CNN-GRU

The CNN-GRU method is a hybrid framework of the Convolutional neural network (CNN) and the GRU methods. The benefits of CNN-GRU are that the GRU method has the advantage of processing time well in sorting data. The advantages of the CNN method are ideal for handling high-dimensional data [22]. The structure of the CNN-GRU method is that CNN makes two-layer convolutions and smooths the data into one dimension. While the GRU method can learn old

DOI: https://doi.org/10.29207/resti.v6i5.4160

training information through memory cells, useless training information will be forgotten by the forget gate.

2.4. Performance Evaluation

In this study, to compare the prediction accuracy values, performance evaluation was carried out on each hybrid CNN-GRU, GRU, and LSTM model using Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Correlation Coefficient (CC) and CPU time. In Equations (10), (11), and (12) x_i is the observed value, y_i is the predicted value, and n is the amount of data. The results of the RMSE calculation are considered good if the expected value is close to 0, MAPE calculation results are considered good if the prediction results are low, less than 50% MAPE, whereas CPU time is how long the computer runs while the model is running.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)}{n}}. \tag{10}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \frac{|y_i - x_i|}{x_i} \times 100. \tag{11}$$

$$CC = \frac{\sum_{i=1}^{n}(x_i - \tilde{x})(y_i - \tilde{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \tilde{x})^2 (y_i - \tilde{y})^2}}. \tag{12}$$

3. Results and Discussions

To obtain the best performance for each model, we performed hyperparameter tuning for each model. Here, we use the grid search method for hyperparameter tuning. Table 1 shows the best parameter setting obtained from the process of hyperparameter tuning. The three models have the same batch size but differ in other parameters. For example, the unit parameter in the CNN-GRU model is 100 units, while the LSTM has 100,50,1 units, and the GRU has 100,50 units. For activation, CNN-GRU uses relu while LSTM tanh.

Table 1. Hyperparameter Tuning of dataset ERA-5 and GFS

Parameter	CNN-GRU	LSTM	GRU
Epoch	170	100	100
Activation	relu	tanh	-
Batch Size	32	32	32
Units	100	100,50,1	100,50
Filters	128	-	-

3.1 Result of dataset ERA-5

Using the ERA-5 dataset, the evaluation of the LSTM model in Table 2 with MAPE and RMSE continues to increase for forecasts from 1 to 30 days. There is no significant increase in CC because the error value that decreases is not too large. At the same time, the GRU model predicts similar results in Table 3, where the resulting error value increases in the wave height prediction for the next 30 days with an error of 0.0364 RMSE, 1.8512 MAPE, and a correlation coefficient of 0.9941. As presented in Table 4, the accuracy of results of the CNN-GRU for predicting 30 days ahead is lower than results of 14 days prediction, especially in term of RMSE and MAPE. Nevertheless the value of correlation coefficient of 30 days predictions remains relatively the same as 14 days predictions, i.e., 0.9938.

Table 2. Wave height prediction model LSTM dataset ERA-5.

Prediction	Performance Evaluation		
	RMSE	MAPE	CC
7 Days	0.0367	1.7439	0.9958
14 Days	0.0290	1.4020	0.9958
30 Days	0.0305	1.5306	0.9958

Table 3. Wave height prediction model GRU dataset ERA-5.

Prediction	Performance Evaluation		
	RMSE	MAPE	CC
7 Days	0.0429	2.0845	0.9941
14 Days	0.0341	1.6872	0.9941
30 Days	0.0364	1.8512	0.9941

Table 4. Wave height prediction model CNN-GRU dataset ERA-5.

Prediction	Performance Evaluation		
	RMSE	MAPE	CC
7 Days	0.0439	2.1215	0.9938
14 Days	0.0349	1.7150	0.9938
30 Days	0.0373	1.8844	0.9938

Figure 6 plots predictions for wave heights of 7, 14, and 30 days generated by CNN-GRU, LSTM, and GRU. The results given on January 30, 2020, are significant predictions far from the test data results.

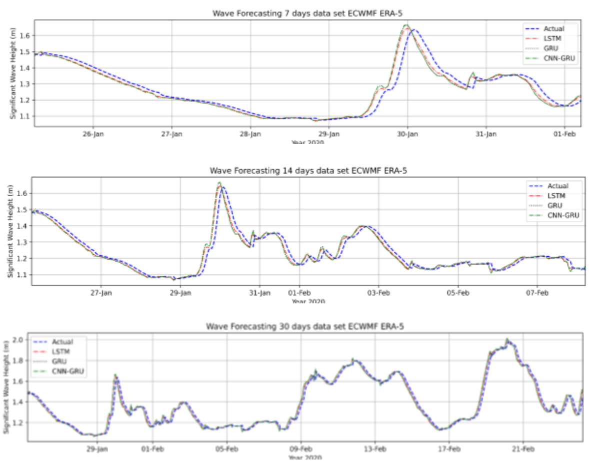

Figure 6. Plot of wave height prediction dataset ERA-5.

3.2 Result of dataset GFS

Prediction results using datasets derived from GFS are shown in Tables 5, 6, and 7 for LSTM, GRU, and CNN-GRU, respectively. The training data given to the GFS source indicates that the prediction results increase at 14

DOI: https://doi.org/10.29207/resti.v6i5.4160

days and decrease at 30 days with a significant error rate in the LSTM, GRU, and CNN-GRU models.

Table 5. Wave height prediction model LSTM dataset GFS.

Prediction	Performance Evaluation		
	RMSE	MAPE	CC
7 Days	0.0233	1.2832	0.9683
14 Days	0.0644	2.2950	0.9873
30 Days	0.0769	2.7353	0.9823

Table 6. Wave height prediction model GRU dataset GFS.

Prediction	Performance Evaluation		
	RMSE	MAPE	CC
7 Days	0.0199	1.1448	0.9802
14 Days	0.0451	1.6446	0.9942
30 Days	0.0561	1.9200	0.9914

Table 7. Wave height prediction model CNN-GRU dataset GFS.

Prediction	Performance Evaluation		
	RMSE	MAPE	CC
7 Days	0.0195	1.1126	0.9802
14 Days	0.0439	1.6007	0.9943
30 Days	0.0548	1.8852	0.9915

The plot results using a dataset derived from GFS for 7, 14, and 30 days predictions with LSTM, GRU, and CNN-GRU can be seen in Figure 7. Significant prediction results occurred on January 6, 2020, and January 13, 2020, where the CNN-GRU model performs well compared to LSTM and GRU models.

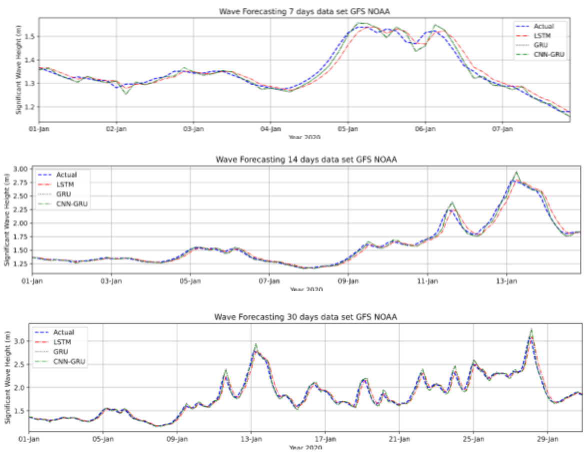

Figure 7. Plot of wave height prediction dataset GFS

3.3 Analysis

The results of prediction by using the CNN-GRU models with both ERA5 and GFS datasets produced relatively accurate results, with low errors in terms of MAPE and RMSE, as well as in correlation coefficient. From Table 14, the CNN-GRU hybrid model did not require large computations to get a reasonably good prediction, with CPU Time values of 440.3968 seconds for the ERA-5 dataset and 85.2198 seconds for the GFS dataset. However, when compared to the performance of the LSTM and GRU models on the ERA-5 dataset, the CNN-GRU hybrid model produces a slightly higher error value and a slightly low correlation coefficient value compared to the two deep learning models in this study. Using the GFS data set, the performance of the CNN-GRU model is very good, with the highest correlation coefficient value of up to 0.9943 and a low average error in the prediction of wave height of 14 days. The average error value generated by the CNN-GRU hybrid model has different variations in the two datasets, especially for predicting 7, 14, and 30 days ahead, which can be caused by the different characteristics between these two datasets. The ERA 5 has an hourly dataset, whereas the GFS is three hourly dataset. A comparison between these datasets is shown in Figure 8.

Table 8. CPU time

Source Dataset	CPU Time (s)		
	CNN-GRU	LSTM	GRU
ERA-5	440.3968	552.0879	373.2863
GFS	85.2198	121.1475	90.0790

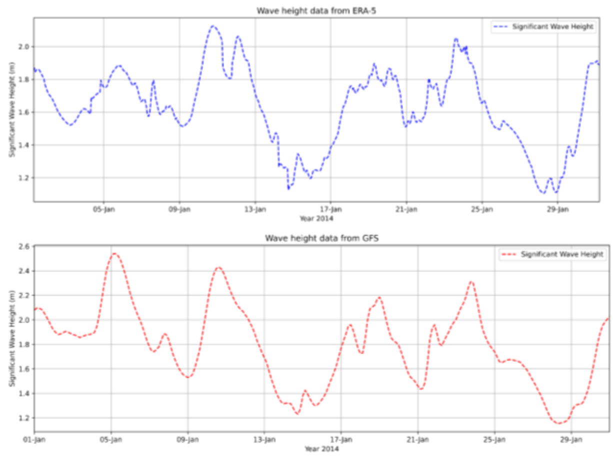

Figure 8. Comparison of dataset characteristics

4. Conclusion

In this study, we build a time series forecasting architecture using the CNN-GRU hybrid deep learning model, where we use wave data from the reanalysis wave data of ERA5 from ECMWF and the GFS data from NOAA. In addition, we also compared the CNN-GRU model with other deep learning models, such as LSTM and GRU. From our numerical experiments, we conclude that for ERA-5 data, the CNN-GRU hybrid model performs slightly lower than LSTM and GRU for predicting wave height. For the GFS wave data, the CNN-GRU gives the best results compared to LSTM

DOI: https://doi.org/10.29207/resti.v6i5.4160

and GRU, with the highest correlation coefficient value and a relatively low average error. The character of the dataset is very influential in predicting the wave height in the deep learning model, where the GFS dataset has wave height data every 3 hours, making the wave height data very volatile, whereas the ERA5 is hourly data. The CNN-GRU model produces relatively accurate predictions with reasonable computational time, especially for the three hourly data GFS.

Reference

[1] R. Arto, L. Prakoso, D. S.-S. P. Laut, and undefined 2021, "Strategi Pertahanan Laut Indonesia dalam Perspektif Maritim Menghadapi Globalisasi," *jurnalprodi.idu.ac.id*, Accessed: Apr. 14, 2022. [Online]. Available: http://jurnalprodi.idu.ac.id/index.php/SPL/article/view/644

[2] J. Berbić, E. Ocvirk, D. Carević, and G. Lončar, "Application of neural networks and support vector machine for significant wave height prediction," *Oceanologia*, vol. 59, no. 3, pp. 331–349, Jul. 2017, doi: 10.1016/J.OCEANO.2017.03.007.

[3] X. H. Le, H. V. Ho, G. Lee, and S. Jung, "Application of Long Short-Term Memory (LSTM) Neural Network for Flood Forecasting," *Water 2019, Vol. 11, Page 1387*, vol. 11, no. 7, p. 1387, Jul. 2019, doi: 10.3390/W11071387.

[4] M. C. Deo, A. Jha, A. S. Chaphekar, and K. Ravikant, "Neural networks for wave forecasting," *Ocean Engineering*, vol. 28, no. 7, pp. 889–898, Jul. 2001, doi: 10.1016/S0029-8018(00)00027-5.

[5] V. Juliani, D. Adytia, and Adiwijaya, "Wave Height Prediction based on Wind Information by using General Regression Neural Network, study case in Jakarta Bay," *2020 8th International Conference on Information and Communication Technology, ICoICT 2020*, Jun. 2020, doi: 10.1109/ICOICT49345.2020.9166305.

[6] C. Jörges, C. Berkenbrink, and B. Stumpe, "Prediction and reconstruction of ocean wave heights based on bathymetric data using LSTM neural networks," *Ocean Engineering*, vol. 232, p. 109046, Jul. 2021, doi: 10.1016/J.OCEANENG.2021.109046.

[7] F. A. R. Abdullah, N. S. Ningsih, and T. M. Al-Khan, "Significant wave height forecasting using long short-term memory neural network in Indonesian waters," *Journal of Ocean Engineering and Marine Energy 2022*, pp. 1–10, Mar. 2022, doi: 10.1007/S40722-022-00224-3.

[8] X. Feng, G. Ma, S. F. Su, C. Huang, M. K. Boswell, and P. Xue, "A multi-layer perceptron approach for accelerated wave forecasting in Lake Michigan," *Ocean Engineering*, vol. 211, p. 107526, Sep. 2020, doi: 10.1016/J.OCEANENG.2020.107526.

[9] S. Mandal and N. Prabaharan, "Ocean wave forecasting using recurrent neural networks," *Ocean Engineering*, vol. 33, no. 10, pp. 1401–1410, Jul. 2006, doi: 10.1016/J.OCEANENG.2005.08.007.

[10] F. Anggraeni, D. Adytia, and A. W. Ramadhan, "Forecasting of Wave Height Time Series Using AdaBoost and XGBoost, Case Study in Pangandaran, Indonesia," *2021 International Conference on Data Science and Its Applications, ICoDSA 2021*, pp. 97–101, 2021, doi: 10.1109/ICODSA53588.2021.9617524.

[11] M. R. Kaloop, D. Kumar, F. Zarzoura, B. Roy, and J. W. Hu, "A wavelet - Particle swarm optimization - Extreme learning machine hybrid modeling for significant wave height prediction," *Ocean Engineering*, vol. 213, p. 107777, Oct. 2020, doi: 10.1016/J.OCEANENG.2020.107777.

[12] M. Sajjad *et al.*, "A Novel CNN-GRU-Based Hybrid Approach for Short-Term Residential Load Forecasting," *IEEE Access*, vol. 8, pp. 143759–143768, 2020, doi: 10.1109/ACCESS.2020.3009537.

[13] H. Hersbach *et al.*, "The ERA5 global reanalysis," *Q J R Meteorol Soc*, vol. 146, pp. 1999–2049, 2020, doi: 10.1002/qj.3803.

[14] L. Bengtsson *et al.*, "Convectively Coupled Equatorial Wave Simulations Using the ECMWF IFS and the NOAA GFS Cumulus Convection Schemes in the NOAA GFS Model," *Monthly Weather Review*, vol. 147, no. 11, pp. 4005–4025, Nov. 2019, doi: 10.1175/MWR-D-19-0195.1.

[15] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/NECO.1997.9.8.1735.

[16] S. Siami-Namini and A. S. Namin, "Forecasting Economics and Financial Time Series: ARIMA vs. LSTM," Mar. 2018, doi: 10.48550/arxiv.1803.06386.

[17] W. Tuerxun, C. Xu, H. Guo, L. Guo, N. Zeng, and Y. Gao, "A Wind Power Forecasting Model Using LSTM Optimized by the Modified Bald Eagle Search Algorithm," *Energies 2022, Vol. 15, Page 2031*, vol. 15, no. 6, p. 2031, Mar. 2022, doi: 10.3390/EN15062031.

[18] R. Dey and F. M. Salemt, "Gate-variants of Gated Recurrent Unit (GRU) neural networks," *Midwest Symposium on Circuits and Systems*, vol. 2017-August, pp. 1597–1600, Sep. 2017, doi: 10.1109/MWSCAS.2017.8053243.

[19] O. T. BİŞKİN and A. ÇİFCİ, "Forecasting of Turkey's Electrical Energy Consumption using LSTM and GRU Networks," *Bilecik Seyh Edebali University Journal of Science*, vol. 8, no. 2, pp. 656–667, Dec. 2021, doi: 10.35193/BSEUFBD.935824.

[20] S. Zargar, "Introduction to Sequence Learning Models: RNN, LSTM, GRU." May 2021. doi: 10.13140/RG.2.2.36370.99522.

[21] L. Zaman, S. Sumpeno, and M. Hariadi, "Analisis Kinerja LSTM dan GRU sebagai Model Generatif untuk Tari Remo," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, vol. 8, no. 2, p. 142, May 2019, doi: 10.22146/JNTETI.V8I2.503.

[22] L. Wu, C. Kong, X. Hao, and W. Chen, "A Short-Term Load Forecasting Method Based on GRU-CNN Hybrid Neural Network Model," *Mathematical Problems in Engineering*, vol. 2020, 2020, doi: 10.1155/2020/1428104.

DOI: https://doi.org/10.29207/resti.v6i5.4160