Wave Height Prediction based on Wind Information by using General Regression Neural Network, study case in Jakarta Bay

Vita Juliani
School Of Computing
Telkom University Bandung, Indonesia
vitajuliani@student.telkomuniversity.ac.id

Didit Adytia
School Of Computing
Telkom University Bandung, Indonesia
adytia@telkomuniversity.ac.id

Adiwijaya
School Of Computing
Telkom University Bandung, Indonesia
adiwijaya@telkomuniversity.ac.id

Abstract—Information about ocean wave is very important for naval navigation, port operations, offshore or nearshore activities around the sea waters. Moreover prediction of wave condition is necessary for design of harbour, coastal and offshore structures. Variations in wave heights are caused by wind pressure on free waves which make it random and uncertain, so that become difficult to predict. In previous studies, wave prediction have been carried out by using semi-empirical methods and conventional methods that require high resolution simulations and high computation. In this paper, we propose a method for prediction wave height from wind data by using a variant of Artificial Neural Network (ANN) with single pass associative memory-forward, so called General Regression Neural Network (GRNN). To obtain a set of training data, we perform numerical wave simulation by using SWAN (Simulating Wave Nearshore) model by using wind data obtained from ECMWF ERA-5. As a study area, we choose a rather shallow bathymetry and complex geometry, in Jakarta Bay, Indonesia. Results of prediction by using GRNN show a good agreement with wave data.

Keywords – General Regression Neural Network, Wave Prediction, Wind Waves, SWAN model.

I. INTRODUCTION

Wave height is a one significant wave parameter that must be predicted when designing coastal and offshore structures. Not only that, for daily operations in port, offshore platform, as well as naval navigation, prediction of wave height is necessary [2]. For construction and operation in nearshore areas, wave prediction is also required, especially for sediment transport estimation. The wave condition is strongly affected by wind pressure on the sea surface, therefore, the wave prediction can be derived from wind information [1], [3].

Several approaches have been proposed for wave prediction. In general, these approaches can be categorized into three approaches. The first approach is a traditional approach by using semi-empilical relation between wind and wave such as the SMB method [4], Coastal Engineering Manual [6], and Shore Protection Manual [5]. In this simplified method, the method can be computationally cheap and quick, since wave height is approximated from the wind speed and duration data, and fetch length.

More advanced approaches are by performing numerical simulation using the third-generation wave model, in which the wave spectrum evolution is simulated by taking into account effects of nonlinearity, breaking, refraction, diffraction, dissipation by bottom roughness. Among these models are SWAN model [7], and the WAM model [8]. This second approach requires high computational cost, especially for simulating a complex geometry and bathymetry which requires high resolution numerical grid. The third approach is via soft computing approach where relation between wind and wave information is learned by using soft computing approach. Among these approaches, the Artificial Neural Network (ANN) is the most favorite tool to obtain nonlinear relation between wind and wave [1], [9]. The prediction with ANN can be quite time consuming, since one needs to obtain best fit parameters for the ANN such as number of hidden layers, neurons, which in general can be found by trial and error [10]. For that reason, there is a need to find an cheap but yet accurate method for obtaining wave prediction.

In this paper, we propose a variant of ANN with single pass associative memory-forward, so called General Regression Neural Network (GRNN). In some cases such as regression, classification, approximation and fitting problems GRNN provides accurate and relatively fast solutions [12]. Liu et al. 2014 uses the GRNN to predict coeffient of sound absorbtion for a type of absorber structure. The GRNN is optimized in term of spread parameter to give best prediction [19]. Chen et al. 2009 [17] applies the GRNN to design a model for controller and simulation in ultrasonic motors. They use the GRNN to obtain relation between parameters on ultrasonic motor, which can replace traditional design of controller and software simulation.

In this study, wave height prediction is produced by using numerical simulation using SWAN model [7] by using wind data from ECMWF ERA-5 [11]. The obtained data set, i.e. wind and wave data, are then used for training data for the GRNN model. To test the proposed method, we choose as a study case in Jakarta Bay, Indonesia. Not only representing a complex geometry and bathymetry, the Jakarta Bay is the busiest port in Indonesia.

The content of this paper is as follows. Section II discusses some brief details regarding SWAN model and GRNN that are used in this paper. It is then followed by Section III that discusses the methodology that is used for obtaining wave prediction by using GRNN. Results and discussions are described in detail in Section IV. The paper is concluded in the last section.

II. LITERATURE REVIEW

As mentioned in the previous section, in this paper, the wave height data is obtained by performing numerical simulation using SWAN model [7] with wind data from ECMWF ERA-5 as wind field input. The wind and wave data set are then used for training data of the GRNN model. In this section, we describe briefly the basic idea of the SWAN model and the GRNN model.

A. SWAN Model

The software Simulating WAve Nearshore model or SWAN is wave model that based on the third generation 2D spectral wave model, that simulate the evolution and propagation of the wave energy spectrum under wind forcing and bathymetry effects [13]. The model is implemented numerically by using Finite Different method, and it has unconditionally stable numerical scheme [7]. The model takes into account effects of nonlinearity, refraction, diffraction, wave breaking due to whitecapping and bottom dissipation. Effect of dissipation such as whitecapping have been implemented in the SWAN model to improve the accuracy of the model [14]. The evolution equation of wave spectrum in the model is represented in Eulerian formulation as the balance of wave energy density [15]. The spectral action balance equation in SWAN model can be desribed as follows.

$$\frac{S_{total}}{\sigma} = \frac{\partial}{\partial t}N + \frac{\partial}{\partial x}C_x N + \frac{\partial}{\partial y}C_y N + \frac{\partial}{\partial \sigma}C_\sigma N + \frac{\partial}{\partial \theta}C_\theta N \quad (1)$$

The right hand side (RHS) terms shows the temporal change of the action density N, and its propagation in horizontal x and y direction, respectively. The other terms in RHS represent the action propagation in frequency domain, and polar coordinate, respectively. Here $N(\sigma, \theta, x, y, t)$ is the the action density that is the function of horizontal coordinates x and y, and time t, frequency σ, and direction θ, [13], [15]. The source term of the equation S_{total} is described as follows.

$$S_{total} = S_{in} + S_{nl3} + S_{nl4} + S_{ds,wcap} + S_{ds,bot} + S_{ds,br} \quad (2)$$

S_{in} represents the source term that is forced by the wind, whereas $S_{nl3}, S_{nl4}, S_{ds,wcap}$ denote the nonlinear wave-wave interaction due to whitecapping, respectively, and followed by the bottom dissipation and the wave breaking due to bottom [13].

In this paper, the SWAN model is used for simulating wind datasets based on global wind data from ECMWF ERA-5. In the subsection, brief description of the General Regression Neural Network (GRNN) is described.

B. General Regression Neural Network

The General Regression Neural Network (GRNN) is introduced in 1991 by [16]. The GRNN is a type of Neural Network with single pass architecture with four layers with normalized Gaussian activation function [12]. The GRNN is a Neural Network (NN) architecture that can be used for solving function approximation problems and can be used for continuous variables estimation. The GRNN has smoothing parameter called spread parameter [19]. The performance of the GRNN model is determined by the spread parameter [20]. In the GRNN, there are four layers, i.e. the input, hidden or pattern, summation, and output layer [12]. Its architecture is illustrated in Fig. 1.

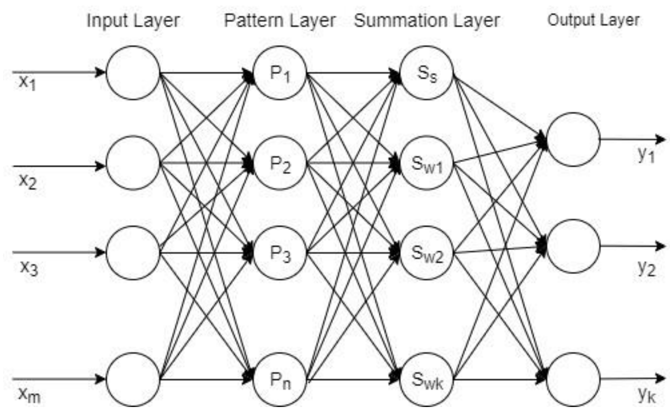

Fig. 1. The architecture of the Generalized Regression Neural Network (GRNN).

In the pattern/hidden layer, the Gaussian function P_i is defined as follows.

$$P_i = \exp \frac{(X - X_i)^T (X - X_i)}{2\sigma^2} \qquad (i = 1, 2, ...n) \quad (3)$$

where X, X_i, σ, and T denote the input variable, neuron-i in hidden layer, a smoothing parameter, and training data $T_i = [X_i, Y_i](i = 1...n)$, and $Y_i = [y_1, y_2, ...y_k]$, respectively.

The summation layer consists of two different calculations. The first one computes weighted output sum in hidden layer, and the second computes the output in hidden neurons [18]. S_s and $S_w t$ calculate the sum of Gaussian function and weighted sum in the pattern output, respectively.

$$S_s = \sum_{t=1} P_t, t = 1...n \quad (4)$$

$$S_{wt} = \sum_{t=1} W_t P_t, t = 1...n \quad (5)$$

The weight of hidden neuron w_t connected with the summation layer [19]. Here, the output neuron Y is computed by using

$$\hat{Y}_o = \frac{S_s}{S_{wo}}, o = 1...k. \quad (6)$$

III. METHODOLOGY

In this section, the methodology for obtaining wave prediction by using GRNN is described.

A. Design system

In this subsection we describe the design system of the process of the implementation of the wave height prediction based on wind data by using the General Regression Neural Network (GRNN) approach. The system is aimed to make predictions of wave height based on wind input data. The system flow is shown in Fig. 2.

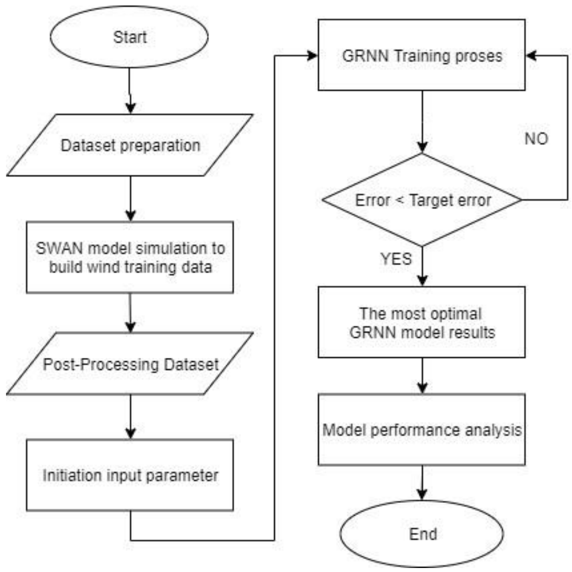

Fig. 2. Flowchart of prediction system using General Regression Neural Network.

Fig. 2 shows the prediction system flow using the GRNN method. Firstly, the wind data set are prepared for performing the wave simulation by using SWAN model. In this step appropriate numerical setting for SWAN is prepared. Since we require a high resolution grid model, the SWAN simulation is calculated in three nested domains, i.e. Domain-1 as the global domain, Domain-2 as the intermediate domain, and Domain-3 as the smallest domain in the Jakarta Bay. The boundary of these three domains are described in Table I. The simulation is run for 2 years, i.e. 1 January 2017 - 31 December 2018. The resulting SWAN simulation in Domain-2 is shown in Fig. 3.

TABLE I
BOUNDARY OF COMPUTATION DOMAIN FOR SWAN SIMULATION.

Domain Number	Longitude		Latitude	
	From	To	From	To
1	0.500	175.500	-69.500	30.500
2	96.000	118.000	-9.500	9.500
3	106.650	107.050	-6.122	-5.858

After the numerical simulation, obtained wind and wave data sets are prepared as training data for the GRNN model. We extracted wind and wave data at 6 point locations, as

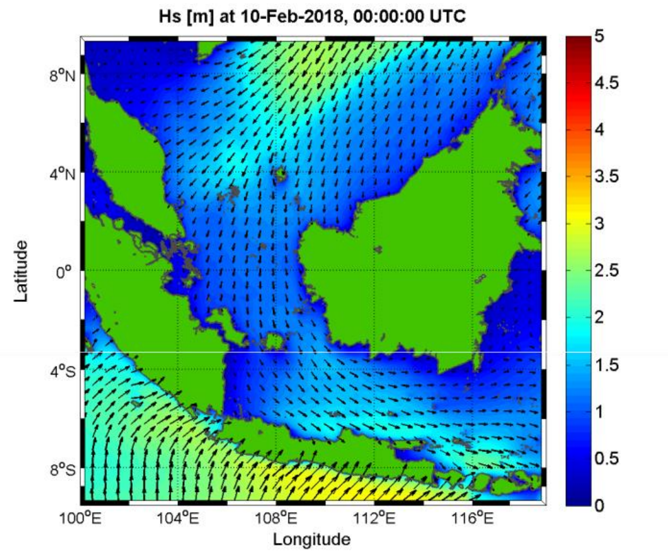

Fig. 3. Result of wave simulation by using SWAN model in Domain-2.

shown in Fig. 4. The location of these points are shown in Table II.

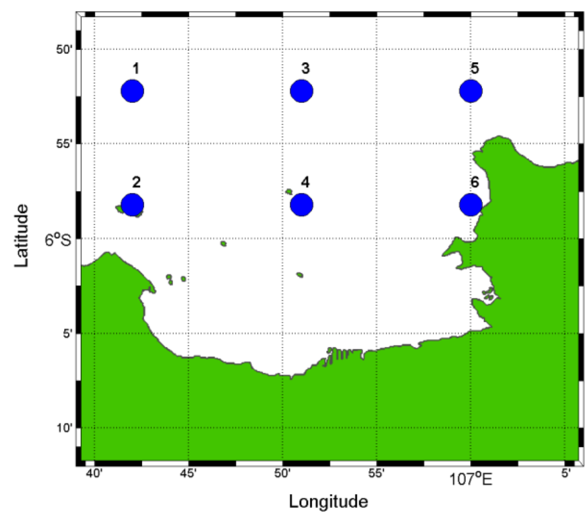

Fig. 4. Location of extracted wind and wave data in Jakarta Bay, Indonesia.

TABLE II
LONGITUDE AND LATITUDE LOCATION OF POINT

Point No.	Location	
	Longitude	Latitude
1	106.70	-5.870
2	106.70	-5.970
3	106.85	-5.870
4	106.85	-5.970
5	107.00	-5.870
6	107.00	-5.970

The extracted data from wave simulation is then used for the

GRNN model. In the training process of the GRNN, an optimal value of parameter spread of GRNN is estimated. In the next section, we show the sensitivity of the parameter spread with respect to the accuracy of the resulting wave prediction.

IV. Result and Discussion

In this section, results of wave prediction by using the GRNN is analyzed qualitatively as well as quantitatively. To obtain quantitative results between the wave prediction and the wave data, we use the Correlation Coefficient (CC) and the Root Mean Square Error (RMSE) that are defined as follows.

$$CC(y, \hat{y}) = \frac{< y, \hat{y} >}{|y|\,|\hat{y}|}, \quad RMSE = \sqrt{\frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{N}}. \tag{7}$$

where y and $\hat{y}$ denote the wave data and the wave prediction, respectively. Here, $|.|$ and $< .,. >$ denote the L_2 norm and inner product, respectively.

To show sensitivity of the length of data training, we perform various scenarios to investigate the proportion of training-testing data for the GRNN model. Table III and IV describe various proportions of training and testing data, with resulting CC and RMSE values. It can be concluded that the longer the training data resulting better wave height prediction, i.e. higher CC value and lower RMSE value. These values are obtained by using spread parameter is 1. It is found that by using 95% of training data and 5% of testing data obtained the most optimal CC and RMSE that is 0.92 and 0.13 .

TABLE III
Sensitivity of data composition of training with spread parameter is 1.00

Training Data (%)	TRAINING	
	CC	RMSE
50	0.85	0.08
60	0.84	0.09
70	0.84	0.09
80	0.83	0.09
90	0.82	0.08
95	0.82	0.09

TABLE IV
Sensitivity of data composition of testing with spread parameter is 1.00

Testing Data (%)	TESTING	
	CC	RMSE
50	0.76	0.10
40	0.75	0.10
30	0.75	0.10
20	0.77	0.10
10	0.89	0.12
5	0.92	0.13

We also investigate that the number of wind input data affects significantly the accuracy of the wave prediction. The wave prediction is obtained by using 6 wind input data at position P1 to P6 (see Fig. 4). In Table V, we also calculate the accuracy of the GRNN approach when we modify the number of wind input, i.e. 1, 2, 3 and 6 wind input data. As expected, with 6 wind input data gives more accurate prediction, i.e. CC value of 0.92 and RMSE value of 0.13 compared to 1, 2 and 3 wind input data.

TABLE V
Sensitivity of number of wind input data with respect to correlation coefficient CC and RMSE

Number of Point	Point Name	CC	RMSE
1	1	0.87	0.15
	3	0.76	0.19
	5	0.88	0.16
2	1,3	0.88	0.16
	3,5	0.89	0.14
3	1,3,5	0.89	0.14
6	1,2,3,4,5,6	0.92	0.13

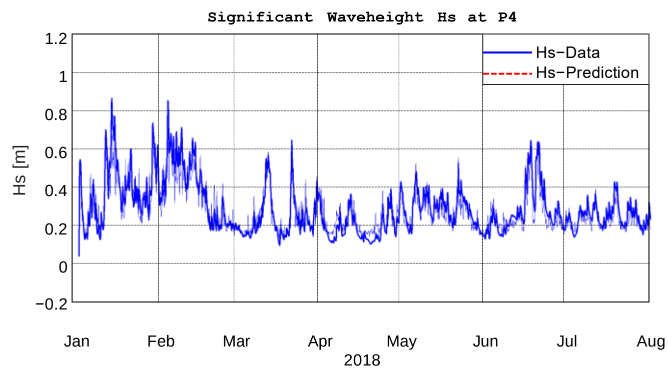

Fig. 5. Time series comparison of wave height prediction (dashed red line) with the wave height data (solid blue line) at P4.

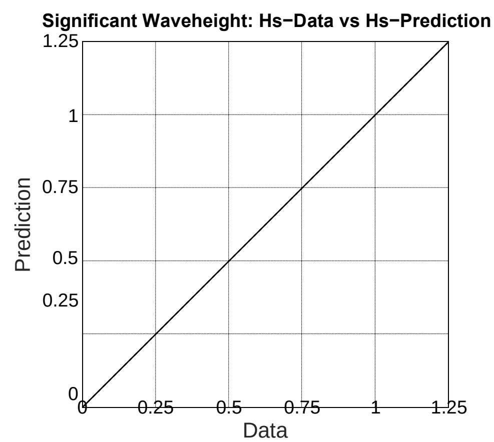

Fig. 6. Scatter plot of wave prediction and wave data at P4.

Results of the wave prediction at position P4 is compared with the wave data time series at Fig. 5. The scatter plot of this comparison between wave prediction and data time series is shown in Fig. 6. It can be seen that the wave prediction obtained from GRNN can follow accurately the wave data.

V. Conclusion

In this paper we predict wave height using the General Regression Neural Network or GRNN model. To that aim, we construct wave data set by performing numerical wave simulation using SWAN model with global wind data ECMWF ERA-5 as the input for the wave model. The obtained data set are used to train the GRNN model. The result of wave height prediction using GRNN method is very satisfactory, with correlation coefficient of 0.92 and RMSE of 0.13. It is found that the accuracy of the prediction is highly influenced by the data, i.e. the more data points of wind input that are used results in more accurate predicted wave height.

REFERENCES

[1] Deo, M.C. and Jha, A, 2001, "Neural networks for wave forecasting." Pergamon, Ocean Engineering, pp. 889–898.

[2] B. Kamranzad, A. Etemad-Shahidi and M.H. Kazeminezhad ,"Wave height forecasting in Dayyer, the Persian Gulf." Ocean engineering, vol. 38(1), pp.248-255, 2011.

[3] S. Asma, A. Sezer and O. Ozdemir, "MLR and ANN models of significant wave height on the west coast of India." , Computers geosciences, vol. 49, pp.231-237, 2012.

[4] Bretschneider, C. L., "Wave forecasting relations for wave generation." Look Lab, Hawaii 1.3 (1970).

[5] US Army, "Shore Protection Manual."4th ed., 2vols. US Army Engineer Waterways Experiment Station, US Government Printing Office, Washington, DC. 1984.

[6] US Army, "Coastal Engineering Manual." Chapter II-2, Meteorology and Wave Climate. Engineer Manual 1110-2-1100. US Army Corps of Engineers, Washington, DC. 2006.

[7] Booij, N. R. R. C., Roeland C. Ris, and Leo H. Holthuijsen., "A third-generation wave model for coastal regions: 1. Model description and validation." Journal of geophysical research: Oceans 104.C4: 7649-7666. 1999.

[8] The WAMDI group, "The WAM model—a third generation ocean wave prediction model." Journal of Physics Oceanography 18, 1775–1810. 1988.

[9] Tsai, C.P., Lin, C., Shen, J.N.,, "Neural network for wave forecasting among multi-stations." Ocean Eng. 29, 1683–1695. 2002.

[10] A. Etemad-Shahidi, and J. Mahjoobi, "Comparison between M5' model tree and neural networks for prediction of significant wave height in Lake Superior ," Ocean Engineering, vol. 36, pp. 1175–1181, 2009.

[11] Ersbach, Hans, and D. J. E. N. Dee., "ERA5 reanalysis is in production." ECMWF newsletter 147.7: 5-6. 2016.

[12] Ahmad Jobran Al-Mahasneh, Sreenatha G. Anavatti and Matthew A. Garratt, "Review of Applications of Generalized Regression Neural Networks in Identification and Control of Dynamic Systems," ArXiv, vol. abs/1805.11236, 2018.

[13] K. Amarouche, A. Akpinar, Nour El Islam Bachari, Recep Emre Cakmak, and Fouzia Houma, "Evaluation of a high-resolution wave hindcast model SWAN for the West Mediterranean basin." Applied Ocean Research, vol. 84, 2019, pp. 225–241.

[14] V. Kutupoğlu, R.E. Çakmak, A. Akpınar and G.P. van Vledder, "Setup and evaluation of a SWAN wind wave model for the Sea of Marmara." Ocean Engineering, 165, pp.450-464, 2018.

[15] A. Akpınar, G.P. van Vledder, M.I. Kömürcü and M. Özger, "Evaluation of the numerical wave model (SWAN) for wave simulation in the Black Sea." Continental Shelf Research, 50, pp.80-99, 2012.

[16] D. F. Specht, "A general regression neural network." IEEE transactions on neural networks, 2(6), 568-576., 1991.

[17] Tien-chi Chen and Chih-Hsien Yu, "Generalized regression neural-network-based modeling approach for traveling-wave ultrasonic motors." Electric Power Components and Systems, vol. 37(6), pp.645-657, 2009.

[18] Ö. Eskidere, F. Ertaş and C.Hanilçi, "A comparison of regression methods for remote tracking of Parkinson's disease progression." Expert Systems with Applications, vol. 39(5), pp.5523-5528, 2012.

[19] J. Liu, W. Bao, L. Shi, B. Zuo and W. Gao, "General regression neural network for prediction of sound absorption coefficients of sandwich structure nonwoven absorbers." Applied Acoustics, vol. 76, pp.128-137, 2014.

[20] R. Hu, S. Wen, Z. Zen, and T. Huang, "A short-term power load forecasting model based on the generalized regression neural network with decreasing step fruit fly optimization algorithm." Neurocomputing, vol. 221, pp.24-31, 2017.