An Analysis on Wind Speed Forecasting Result with the Elman Recurrent Neural Network Method

Martaleli Bettiza[1*] *, Nola Ritha*[2] *, Nanang Gusti Rama*[3]

[1]Department of Informatics, Maritime University of Raja Ali Haji, Tanjungpinang, Indonesia

Abstract. Weather factors in the archipelago have an important role in sea transportation. Weather factors, especially wind speed and wave height, become the determinants of sailing permits besides transportation's availability, routes, and fuel. Wind speed is also a potential source of renewable energy in the archipelago. Accurate wind speed forecasting is very useful for marine transportation and development of wind power technology. One of the methods in the artificial neural network field, Elman Recurrent Neural Network (ERNN), is used in this study to forecast wind speed. Wind speed data in 2019 from measurements at the *Badan Meteorologi Klimatologi dan Geofisika* (BMKG) at Hang Nadim Batam station were used in the training and testing process. The forecasting results showed an accuracy rate of 88.28% on training data and 71.38% on testing data. The wide data range with the randomness and uncertainty of wind speed is the cause of low accuracy. The data set is divided into the training set and the testing set in several ratio schemas. The division of this data set considered to have contributed to the MAPE value. The observation data and data division carried out in different seasons, with varying types of wind cycles. Therefore, the forecasting results obtained in the training process are 17% better than the testing data.

*Corresponding author: mbettiza@umrah.ac.id

1 Introduction

Indonesia is a maritime country in Southeast Asia located between two continents, Asia and Australia, and two oceans, Pacific Ocean and Indian Ocean. Indonesia's land area is 1,922,570 km², and its water area is 3,257,483 km². In addition, Indonesia is an archipelagic country with the largest number of islands in the world. Based on data from the Central Statistics Agency (BPS) published in 2017, Indonesia has 17,504 islands spread across 34 provinces [1]. Indonesia's ocean area is much wider than its land area, make it further strengthens Indonesia's position as the largest archipelagic country in the world.

Riau Islands Province (Kepri) is an area with a sea area of 96% of its total area, 241,215 km2. Riau Islands Province has the largest number of islands with a total of 2,408 islands [1]. In archipelagic areas, solving transportation problems is a solution for improving logistics distribution networks and shipping marine products from certain regions. Sea transportation is not only influenced by the availability of means of transportation, routes, and fuel requirements, but also influenced by weather factors, such as wind speed and wave height.

The energy sources widely used today, oil and coal, derived from fossils, which are shrinking in number. There have been many studies on other energy sources, classified as new and renewable energy, such as the use of water, solar radiation, wind, and ocean energy. Riau Islands is an area that is rich in potential renewable energy. In archipelagic country, studying the possibility of renewable energy, such as wind speed, becomes very important. The study of wind speed will support further analyses on transportation issues and studies of new and renewable energy potentials.

Several algorithms have been applied on artificial neural networks to forecast wind speed and produce training models with certain accuracy. In this study, the Elman Recurrent Neural Network (ERNN) method used to forecast wind speed.

Zhang [6] has developed an adaptive hybrid model for short-term wind speed forecasting as a result before the accuracy of the single model needs to improve. An adaptive hybrid model based on Variation Mode Decomposition (VMD), Fruit Fly Optimization Algorithm (FOA), Autoregressive Integrated Moving Average Model (ARIMA) and Deep Belief Network (DBN) is proposed [6]. This adaptive hybrid makes the accuracy of the proposed model better than the other models. Xiao [7] also proposed a combined model to improve accuracy in short-term wind speed prediction based on ARIMA-GARCH and Elman method, and these models significantly improved the accuracy. Zhang [8] also performed a study in short-time wind speed prediction using a combined model to obtain better accuracy.

1.1 Methodology

1.1.1 Data

This study aims to predict short-term wind speed using data over a certain period of observations from the Hang Nadim BMKG station, Batam. Data is recorded every 3 (three) hours for 365 days, in 2019 and not all informatio is filled in or complete. To meet data processing rules, data imputation carried out using the mean, median, and mode values of the original data. A comprehensive data is needed in the prediction process to ensure the accuracy of the results.

Table 1. Wind Speed Data Preparation

Date	UTC	Wind Speed (Km/h)	Wind Speed (Km/h)
01/01/2019	00Z	5.6	5.6
01/01/2019	03Z	14.8	14.8
01/01/2019	06Z		12.05
01/01/2019	09Z	9.3	9.3
01/01/2019	12Z		9.3
01/01/2019	15Z		9.3
01/01/2019	18Z	9.3	9.3
01/01/2019	21Z	5.6	5.6
02/01/2019	00Z		18.6
02/01/2019	03Z	13.0	13
02/01/2019	06Z	14.8	14.8
02/01/2019	09Z	14.8	14.8
02/01/2019	12Z	5.6	5.6
02/01/2019	15Z	9.3	9.3

The application was built by applying the Elman Recurrent Neural Network algorithm for wind speed forecasting. This process used 2916 wind speed data in 2019. The reliability of the model carried out with variations in the distribution of the dataset. The dataset divided to the training set and testing set to the following scheme: 60%: 40%, 70%: 30%, and 80%: 20%. The learning rate parameter tested varies in value from 0.1 to 0.9, with an error tolerance of 0.001 and iterations of 50 and 100.

1.1.2 MAPE

The robustness of forecasting results of the three models used in this study will be evaluated using MAPE value as an error indicator.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{y_t - f_t}{y_t} * 100\% \qquad (1)$$

Intuitively, MAPE represents a mean absolute error between the prediction results and the actual data. Therefore, it is easier to evaluate the result and calculate the accuracy.

2 Elman Recurrent Neural Network

Elman Recurrent Neural Network (ERNN) first developed by Jeff Elman in 1990. ERNN has advantages in predicting time series data. The output from the hidden layer in ERNN performs a feedback process to itself through the recurrent layer, which is known as the context layer. This process strengthens the network to learn, recognize and generate trained data patterns. ERNN consists of four layers: input layer, hidden layer, context layer, and output layer.
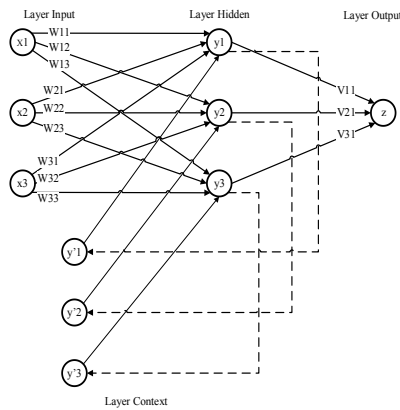The figure below shows ERNN network architecture.

Figure 1. ERNN network architecture

ERNN algorithm steps:

1. Initialize the weights between the input and hidden layers and the weights between the hidden output layers randomly. Specifies the values for the maximum epoch, learning rate and error tolerance.
2. The input data is sent to all hidden layer units, by multiplying the input value by the weight, and combined with the weight in the context layer, adding up the bias.

$$net_j(t) = \left(\sum_i^n x_i(t)v_{ji} + \sum_h^m y_h(t-1)u_{jh}\right) + \theta_j \qquad (2)$$

$$y_j(t) = f\big(net_j(t)\big) \qquad (3)$$

3. The activation function used in the output layer is a binary sigmoid function

$$f\big(net_j\big) = \frac{1}{1+e^{-net_j}} \qquad (4)$$

4. The output of the hidden layer is multiplied by the weight and added by the bias. The network output calculated using the activation function.

$$net_k(t) = \left(\sum_j^m y_j(t)w_{kj}\right) + \theta_k \qquad (5)$$

$$y_k(t) = g\big(net_k(t)\big) \qquad (6)$$

3 Results

The following is the output of the data training process using the ERNN algorithm.
Using scheme dataset division, 60%, 70%, and 80% training data; error tolerance: 0.001.

Table 2. Output for 60% training data

Learning Rate	Epoch = 50	Epoch = 100
	MAPE %	MAPE %
0.1	13.988	14.058
0.2	13.71	13.774
0.3	13.824	13.29
0.4	13.581	12.585
0.5	13.612	13.36
0.6	13.356	11.797
0.7	12.614	13.412
0.8	13.171	14.058
0.9	13.119	11.755

Table 3. Output for 70% training data

Learning Rate	Epoch = 50	Epoch = 100
	MAPE %	MAPE %
0.1	14.005	14.014
0.2	13.801	12.998
0.3	13.755	12.653
0.4	13.481	12.513
0.5	13.554	12.438
0.6	13.208	12.464
0.7	13.214	12.459
0.8	12.582	12.414
0.9	12.804	12.385

Table 4. Output for 80% training data

Learning Rate	Epoch = 50	Epoch = 100
	MAPE %	MAPE %
0.1	13.982	14.015
0.2	13.205	13.551
0.3	13.384	12.831
0.4	12.875	12.853
0.5	12.807	12.824
50	12.872	12.845
76	12.805	12.807
42	12.22	12.823
23	12.814	12.81

The results of the system training process were carried using variations of training data 60%, 70%, and 80%. Experiment with learning rate 0.1 to 0.8, error tolerance 0.001, and epochs 50 and 100, obtained the best training results with the smallest MAPE error value of 11.751% and the highest accuracy of 88.249% on 60% training data experiment with learning rate 0.8,

error tolerance 0.001 and the process stops at 94th iteration. Parameters obtained from system training with the smallest MAPE value or highest accuracy used to predict wind speed.

The following is the output of the data testing process using the ERNN algorithm.

Table 5. Output for 40% testing data

| Learning Rate | Epoch = 50 | Epoch = 100 |
	MAPE %	MAPE %
0.1	39.115	39.433
0.2	38.67	38.781
0.3	38.981	37.029
0.4	36.802	32.87
0.5	38.741	35.948
0.6	37.11	28.355
0.7	32.243	37.539
0.8	36.02	28.622
0.9	34.807	29.48

Table 6. Output for 30% testing data

| Learning Rate | Epoch = 50 | Epoch = 100 |
	MAPE %	MAPE %
0.1	38.86	39.188
0.2	39.681	34.467
0.3	38.905	32.583
0.4	36.842	32.029
0.5	37.7	29.743
0.6	36.111	31.421
0.7	36.02	32.387
0.8	33.502	31.921
0.9	35.7	32.311

Table 7. Output for 20% testing data

| Learning Rate | Epoch = 50 | Epoch = 100 |
	MAPE %	MAPE %
0.1	43.741	43.938
0.2	41.512	43.358
0.3	44.542	40.574
0.4	40.502	41.376
0.5	41.21	40.869
0.6	43.02	42.555
0.7	41.184	41.722
0.8	43.417	43.814
0.9	41.525	43.426

4 Conclusions

The best accuracy result of 71.38% is considerednotgood enough for this prediction process. The distribution of data in a fairly far range is analyzed to be the cause. Wind speed in archipelagic areas, such as data collection areas, varies greatly depending on the season. So the use of a fairly long period, one year of data, in this case, affects the results obtained. The distribution of data in the training process and the test process considered to have contributed to the magnitude of the MAPE score. The observation data and data division carried out in different seasons, with varying types of wind cycles, so that the parameters obtained in the training process provide an accuracy difference of up to approximately 17% in the test process. The results will be better if the data can be first clustered according to the season. The training and testing process carried out on the data in the same season. The number of input variables used considered affecting the results. This research uses time series data with 3 (three) input variables in the prediction process

References

1. Indonesian Statistic Bureau (BPS), Official Statistical News, Jakarta : Badan Pusat Statistik, 2018.

2. Department of Economic and Social Affairs, Population and Vital Statistics Report, New York : United Nations, 2020.

3. J. Zhang, Y. Wei, Z. Tan, "An Adaptive Hybrid Model for Short-term Wind Speed Forecasting", Elsevier Publisher, 2019.

4. H. Xiao, H. Tang, J. Zhou, "Short-term Wind Speed Combined Prediction Model Based on ARIMA-GARCH and Elman", 2020 Asia Pacific Conference on Image Processing, Electronics, and Computers, 2020

5. H. Liu, Xi-wei Mi, Yan-fei Li, "Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network", Elsevier Publisher, 2018

6. J. Wang, et all.," Forecasting wind speed using empirical mode decomposition and Elman neural network", Applied Coft Computing, Elsevier Publisher, 2014

7. Y. Zhang, C. Zhang, Y. Zhao, S. Gao, "Wind speed Prediction with RBF neural network based on PCA and ICA", Journal of Electrical Engineering, Vol 69 (2018), No 2, 148-155.

8. J. L. Elman, "Finding Structure in Time", Cognitive Science 14, 179-211, 1990.

9. J. Wang, Jun Wang, Wen Fang, and Hongli Niu, "Financial Time Series Prediction Using Elman Recurrent Random Neural Networks", Computational Intelligence and Neuroscience, 2016.