MATEMATIKA, MJIM, Volume 38, Number 2, 125–139

Data-driven Models for Wind Speed Forecasting in Malacca State

[1]Basri Badyalina, [1]Nurkhairany Amyra Mokhtar*, [2]Ahmad Izul Fakhruddin Azimi, [3]Muhammad Majid, [3]Mohamad Faizal Ramli and [3]Fatin Farazh Yaa'coob

[1]Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA Cawangan Johor, Kampus Segamat, 8500 Segamat, Johor, Malaysia

[2]Centre for Mathematical Sciences, Universiti Malaysia Pahang Lebuhraya Tun Razak, 26300 Gambang, Kuantan, Pahang, Malaysia

[3]Faculty of Business and Management, Universiti Teknologi MARA Cawangan Johor Kampus Segamat, 8500 Segamat, Johor

*Corresponding author: nurkhairany@uitm.edu.my

Article history
Received: 28 February 2022
Received in revised form: 28 July 2022
Accepted: 30 July 2022
Published online: 31 August 2022

Abstract Wind energy is a type of renewable energy that has received much attention in the electricity market. This study focuses on forecasting wind speed at Malacca state's Station. This paper proposes using a Group Method of Data Handling (GMDH) model with a hyperbolic tangent transfer function in forecasting wind speed data in Malacca state. The performance of this model is compared with three other models: Artificial neural network (ANN) hyperbolic tangent, ANN log sigmoid and GMDH model. In addition, this study investigates the impact of using different transfer functions in data-driven models for forecasting wind speed. Usually, the application of the GMDH model only implements the second-order polynomial as a partial description. However, in this study, we implement the hyperbolic tangent transfer function in the GMDH model. The forecasting model's accuracy is evaluated using mean absolute percentage error (MAPE), root mean square error (RMSE) and mean absolute error (MAE). The results show that the GMDH model with hyperbolic tangent is highly accurate with more negligible computational power than other models. The GMDH-hyperbolic tangent model managed to improve the forecasting performance of the conventional GMDH model by 11.21% in mean absolute percentage error and outperforms other ANN models.

Keywords Artificial Neural Network, wind, forecasting, GMDH, tanh, data-driven model

Mathematics Subject Classification 37M10, 68T05

1 Introduction

Researchers from all around the globe are now studying the possibilities of harnessing renewable energy resources, for example, sunlight, wind, and geothermal heat. Renewable energies contributed an average of 26.2 per cent of the planet's energy demand in 2018, with that figure expected to rise to 45 per cent by 2040 [1]. The goal is to get the most out of these resources, especially when it comes to creating electricity [2, 3]. Certain countries are putting in effort to achieve this goal. In Malaysia, particularly in the Peninsular Malaysia, plans to increase the renewable energy capacity to 31% in 2025 and the country is expected to attain 40% in 2035. To achieve this goal, wind energy should be utilized as it plays an essential role in the market of electricity. It is one of the potential sources of renewable energy for economic growth. It does not release any harmful products that cause environmental pollution, such as the greenhouse gasses or trigger adverse effects on human health like air pollution [4]. Wind power is precisely correlated to the wind speed cube. Wind power increases dramatically when wind speed changes even a little. As a result, wind speed forecasting is critical in energy scheduling and dispatching and the wind energy market [5]. Aside from that, the rising cost of petroleum-based fuels makes wind power one of the most cost-effective renewable resource options. However, wind speed is unpredictable and sporadic [1]. Therefore, wind power forecasting is vital for ensuring power system efficiency and minimizing power system expenses. In addition, accurate forecasts will enable the government, policymakers, and other relevant authorities to take appropriate actions. In relation to this, data driven model has been widely used for forecasting and prediction in various field [6, 7, 8]. The ANN model is one type of data driven model. In this regard, Artificial Neural Networks (ANN) were discovered to be quite effective among wind power forecasting researchers [9]. Pillai *et al.* [10] stated that the ANN model could provide a solution for a complex mathematical model in wind forecasting which is prone to errors. As a result, ANN techniques have been used to estimate and forecast wind speed and power [11]. In addition, given the extremely nonlinear and seasonal nature of the data, the ANN model showed to be more efficient than other conventional autoregressive predictors [12]. Other than that, for various time spans ranging from relatively short to medium data, the ANN model consistently outperforms linear time series-based models [13].

Furthermore, ANN methods are also known to be robust in forecasting [14]. Most studies on the ANN model consider integrating neural networks with another computing method to utilize the interpretation ability of data [15]. Several short-term wind speed forecasting research explore diverse models, such as two-layers ANN with varied time horizons and distinct wind variability [16]. In 2010, Li and Shi [17] assessed the ANN model's effectiveness in forecasting wind speeds using hourly mean wind speeds gathered at two observation locations in North Dakota. Fazelpour *et al.* [18] utilized ANN to forecast the short-term wind speed in Tehran, Iran, using one-hour intervals data supplied by the Iran Renewable Energy Organization. Navas *et al.* [19] explored ANN in finding the possibility of predicting wind speed models in Coimbatore, Tamil Nadu, India.

Another data-driven model is introduced for wind forecasting in this study, which is the group method of data handling (GMDH) model. The GMDH model is widely applied in other fields such as hydrology, crude oil pyrolysis, electrical power insulators' prediction, and rock mass deformation [20, 21, 22, 23]. However, the GMDH model has rarely been used for wind speed forecasting, especially in Malaysia. The GMDH model was initially developed by Ivakhnenko [24] to identify the complex nonlinear system. The main characteristics of the GMDH model are its self-organizing method and combination of second-order polynomials. Using the second-order polynomials, the GMDH model fuses every possible combination of the input variables and automatically selects the best combination of the input variables based on a particular criterion [25]. The advantage of GMDH model is it still can perform very well without knowing any information of the mechanism generating the original data and very efficient compared to other artificial intelligence models [26]. Other than that. GMDH model are known for solving complex problems in nonlinear system with large degrees of complexity and producing effective result compare to other artificial intelligence model [27].

By simulating the reactions of variable-density flow and solute transport, Lal and Datta [28] utilize the GMDH model to forecast the concentration of salt in coastal area. The result shows that the GMDH model is a more precise and effective coastal saltwater intrusion prediction model than support vector machine regression and genetic programming-based models. Furthermore, in terms of short forecasting, the GMDH model performs better than traditional forecasting approaches such as the Auto-Regressive Integrated Moving Average [21]. Ebtehaj *et al.* [29] studied the forecasting of the streamflow coefficient in a side weir using the GMDH model. In this study, two sets of data were used to train and test the model. Cross-validation was used on the collected results which revealed that the GMDH model is more precise than the ANN model and the current nonlinear regression model for forecasting. In 2008, Kondo and Euno [30] suggested that the effectiveness of the GMDH model can be improved by implementing a radial basis function (RBF). The resulting RBF-GMDH model was applied in image recognition of the blood vessels in the liver. The study's finding shows that the proposed RBF-GMDH model is useful for pattern recognition for medical image recognition. Besides the RBF transfer function, other transfer functions are established in neural network models, such as the hyperbolic tangent transfer function. However, this function is rarely given attention to be used in the GMDH model. Kalman *et al.* [31] assert that the hyperbolic tangent function is superior to the RBF transfer function. The hyperbolic transfer function is also more efficient than the step and linear transfer function for training due to the nonlinear behaviour [32]. Therefore, implementing the hyperbolic tangent function is proposed to be utilized in the GMDH model for wind speed forecasting problems in Malaysia. The objective of the study is to assess the forecasting performance of the proposed GMDH hyperbolic tangent in comparison ANN Hyperbolic Tangent, ANN Sigmoid, and Conventional GMDH using mean absolute percentage error (MAPE), root mean square error (RMSE) and mean absolute error (MAE).

2 Methodology

2.1 Data Description

In this paper, our aim is to forecast the wind speed in Malacca state, Malaysia. Malacca is a state located on the south-western coast of Peninsular Malaysia with a coordinate of 2.29 °N, 102.30 °E [33]. Malacca state is included as a UNESCO World Heritage Site. Malaysia has two yearly monsoons with a local tropical climate, the southwest and northeast [34, 35]. This study's daily wind speed data is dated from January 2016 until November 2020. The data on wind speed is obtained from Malaysia Meteorological Department. Figure 1 shows the location of Malacca state in Peninsular Malaysia.

Table 1 shows the descriptive statistics for wind speed in Malacca state. The notation m/s is referred to meter per second which is the velocity of the wind. The wind speed is measured daily. The mean and median of the wind speed are quite close, 8.72 m/s and 8.30 m/s, respectively. This shows that the data is not highly skewed and no obvious outlier is present in the data. The kurtosis is below 2, at 0.48 and the data tends to have a light tail. The data is moderately skewed to the right with the value of 0.79. This means that the distribution is not symmetric. The standard deviation of the wind speed is 2.28 which shows that the speed varies throughout the years. The minimum speed of the wind is 4.20 m/s and the maximum speed is 19.10 m/s. Figure 2 shows the wind speed plot for Malacca state wind speed data. This plot is supported by the information given in Table 1 which shows that the value of the standard deviation is quite high at 2.28. This is due to the existence of two different monsoon seasons in Malaysia, namely, the southwest monsoon from May to September, and the northeast monsoon from November to March [36]. The plot in Figure 2 shows that the data is non-linear. Therefore, non-linear model is needed to model and forecast the data.

Figure 1: Location of the Study

Table 1: Descriptive Statistics for Wind Speed at Malacca State

Mean(m/s)	Median (m/s)	Standard Deviation	Kurtosis	Skewness	Minimum	Maximum
8.72	8.30	2.28	0.48	0.79	4.20	19.10

2.2 ANN Model

ANN model is one of the tools in computational intelligence that has been widely used in many fields. The inspiration of the ANN model comes from the structure of the human brain, which has interconnected natural networks (neurons) that receive signals (information) and take actions on the basis of the best fit of the data [37]. ANN models are well-known models for forecasting due to their capacity to handle the non-linearity of the time series data [1]. The architecture of the ANN model used in this study is a three-layer feedforward ANN model known as multilayer perceptron (MP) which consists of the input layer, hidden layer, and output layer. The feedforward ANN model received the signals (information) from an input layer and routed the signals across the hidden layer to the output layer.

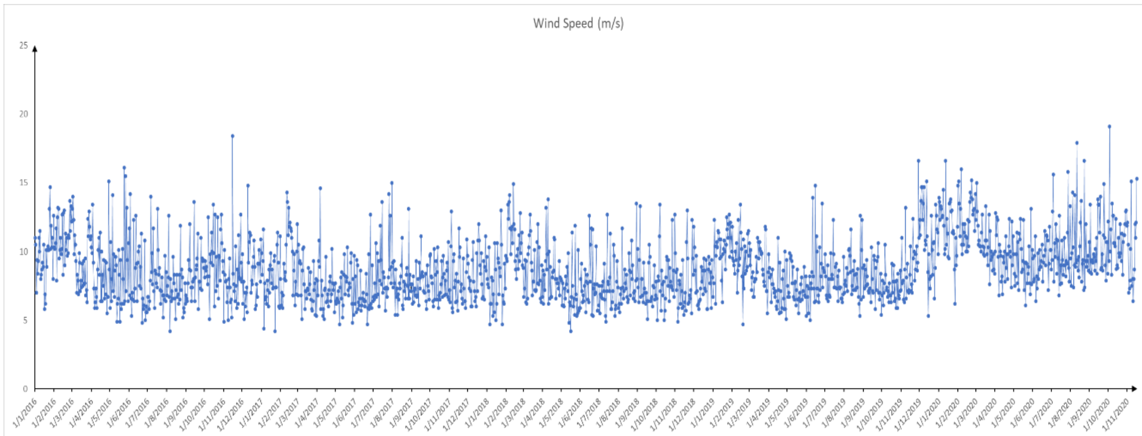

Figure 2: Wind Speed Plot

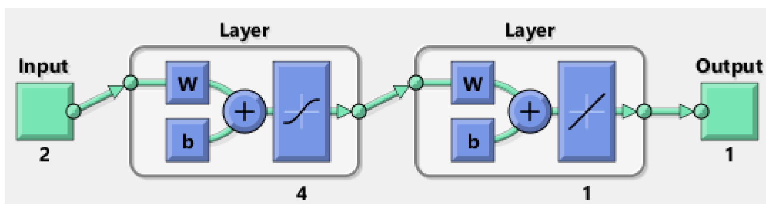

Figure 3: Three-layer Feedforward ANN Model

MP is defined as being interconnected, with each node linked to every node in the preceding layer. It is widely used due to its practicality and simplicity of structure. The architecture of MP is shown in Figure 3. The MP can be expressed as follows:

$$\hat{y}_t = g\left(\sum_{i=1}^{r} z_i^\circ f\left(\sum_{j=1}^{r} z_i^\circ x_i\right) + \omega_t\right) \tag{1}$$

where y_t is the output layer, x_i is the input layer, z_i° is the link between input and hidden layers, z_j° is the link between hidden and output layer, and ω_t is bias. $g(.)$ and $f(.)$ is the activation function, respectively. One of the challenges in ANN architecture is to determine the number of hidden layers. The hidden layer generally interconnected the input and output layer. The number of hidden layers varies depending on the problem. In this study, there are three approaches to determine the number of hidden layers. The approaches are shown in Table 2.

Table 2 shows the guidelines to determine the number of hidden variables based on input variables [23]. If the number of input variables is 2, the hidden layer of ANN model is 2 based on Tang and Fishwick, 4 for based on Wong and 5 based on Hect-Nielsen. Other than the architecture of the ANN model, the critical factor for a successful ANN is the training algorithm. The training algorithm is essential to train the network. Therefore, in this study, a backpropagation algorithm with a variable learning rate is used. The backpropagation algorithm is the most common and highly utilized learning algorithm in many fields [38]. He *et al.* [39] stated that the advantages of the backpropagation algorithm are it is fast, flexible and straightforward. In backpropagation, the weight will fine-tune

Table 2: Guidelines to Determine the Number of Hidden Variables

Approach	Number of Hidden Layers
Tang and Fishwick	r
Wong	$2r$
Hecht-Nielsen	$2r + 1$

r is the total number of inputs in the ANN model

using variable learning rates. There are three types of activation or transfer functions for the ANN model implemented in this study. The activation functions are sigmoid, hyperbolic tangent and linear function.

There are two different combination of transfer function for ANN model used in this study for forecasting wind speed at the Malacca state's station. The first combination is sigmoid transfer function and linear transfer function. The other combination is hyperbolic tangent transfer function and linear transfer function. Hyperbolic tangents and sigmoid connect the input layer with the hidden layer, while linear function connects the hidden layer with the output layer. The sigmoid and hyperbolic tangent transfer function are defined as follows:

$$\varsigma(f) = \frac{1}{1 + e^{-f}}, \tag{2}$$

$$\lambda(f) = \frac{2}{(1 + e^{-2n})} - 1. \tag{3}$$

Equation (2) defines sigmoid transfer function meanwhile Equation (3) define the hyperbolic tangent transfer function.

2.3 GMDH-Hyperbolic Tangent Model

GMDH model is one of the sub-models of the neural network model. Usually, the neural network incorporates the input and output layers, where the input layer's information is analyzed in the hidden layer. The hidden layer features the activation function and algorithm of the network. Input, hidden, and output layers are all featured in the GMDH model which makes it a complete neural network model. Moreover, the GMDH model is a well-known self-organizing process where it autonomously chooses the best predictive factors based on the objective [40]. Other than that, Ivakhnenko [24] established the GMDH model as a multivariate analytical tool for modelling and classifying complex systems. The GMDH algorithm utilizes the Volterra function series to describe and build a relationship between input and output variables. However, most applications of the GMDH model only use second-order polynomial or partial polynomial. The partial polynomial or partial description form of Volterra series or Kolmogorov Gabor polynomial series is as follows:

$$\hat{y} = L(x_i, x_j) = w_0 + w_1 x_i + w_2 x_j + w_3 x_i^2 + w_4 x_j^2 + w_5 x_i x_j. \tag{4}$$

Equation (4) is a partial description that linked an input variable and output variable for the GMDH model. w_0 is a parameter that is estimated from the least square method, $\hat{y}$ is a forecasted value or forecasted output from the GMDH model and x is an input variable for the GMDH model. In addition to the GMDH model for this study, a hyperbolic tangent function was added to the GMDH model. The hyperbolic tangent function is defined as follows:

$$h_k = \left(\frac{2}{1+e^{-2y_k}}\right) - 1. \tag{5}$$

In order to solve Equation (5) using the least square method, a set of a linear system can be illustrated as follows:

$$Y = Xw. \tag{6}$$

The illustration of Equation (5) is shown as follows:

$$\mathbf{Y}=\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{1i} & x_{1j} & x_{1i}x_{ij} & x_{1i}^2 & x_{1j}^2 \\ 1 & x_{2i} & x_j & x_{2i}x_{2j} & x_{2i}^2 & x_{2j}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{li} & x_{lj} & x_{li}x_{lj} & x_{li}^2 & x_{lj}^2 \end{bmatrix} \quad \text{and} \quad w = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_5 \end{bmatrix}.$$

Here, Y denotes a vector of output variables, w represents the parameter of partial description and X is an expression of input variables for Equation (4). In order to implement a hyperbolic tangent function to the GDMH model, the output variables Y will utilize Equation (5). The new output variable based on a hyperbolic tangent is shown in Equation (7). Transformation of the real needed as the hyperbolic tangent function domain is $0 < x < 1$.

$$s_i = -\frac{1}{2}\ln\left(\frac{2}{y_i+1} - 1\right). \tag{7}$$

If the conventional polynomial function is used, therefore $s_i = y_i$. Equation (7) will be utilized to implement the hyperbolic tangent function. After utilizing the hyperbolic tangent function, the parameter of partial description can be estimated using the least square method. The least-square method for parameter estimation is expressed as follows:

$$w = \left(X^T X\right)^{-1}\left(X^T Y\right). \tag{8}$$

Equation (6) provides the best parameter for Equation (4) for the given data. The GMDH model will pair every possible combination of the input variable. The number of partial descriptions constructed can be calculated using Equation (9).

$$L = \frac{r(r+1)}{2}, \tag{9}$$

where r is the total number of input variables and L is the number of partial descriptions produced. After every partial description is built, each partial description will be evaluated. The best partial description will be selected as the new input variables for the next layer of the GMDH model. The optimal partial description was assessed using mean square error (MSE). The MSE is defined as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(s_i - \hat{y}_i)^2. \tag{10}$$

This process is iterated until the realization of the network is achieved. The realization of the network is achieved when the MSE of the current GMDH layer is greater or equal to the MSE of the previous layer. As a result, the partial description that produces the lowest MSE is selected as the output. Equation (3) will be applied to the output to obtain the actual domain of the output.

2.4 GMDH Model

The difference between conventional GMDH model and GMDH-Hyperbolic tangent model is the type of transfer function used in the GMDH model. Conventional GMDH model uses polynomial function as the transfer function whereas GMDH-Hyperbolic tangent model uses hyperbolic tangent function as the transfer function. The polynomial transfer function for conventional GMDH model is defined in Equation (4). The architecture of conventional GMDH model and GMDH-Hyperbolic tangent model is the same which it starts with determination of input variable and split the data into training and forecasting data set. Equation (4) or knowns as partial description is fitted into the training data set. The weight of Equation (4) is using the least square method. The least square method is defined in Equation 8. However, for polynomial, transfer function will be defined as follows:

$$s_i = y_i. \tag{11}$$

The number of partial descriptions constructed can be calculated using Equation (9). The best partial description output will be selected to become an input variable for the next layer. The selection for best partial description is the partial description must have the lowest MSE. The MSE is defined in Equation (10). This process is repeated until the realization of the network is achieved. The realization of the network is achieved when the MSE of the current GMDH layer is greater or equal to the MSE of the previous layer.

2.5 Model Evaluation and Validation

The ratio of training and forecasting data set is 70% and 30%, respectively [41]. The training data set starts from January 2016 until May 2019. The forecasting data set starts from June 2019 until November 2020. The training data set is used to developed the model and the forecast data set is used to evaluate the forecasting performance of the forecast model. The forecasting models are evaluated using three statistical measures which are commonly used for forecasting problems. Each of these measures the disparity in wind speed between the observed and forecast values. The definition of mean absolute percentage error (MAPE), root mean square error (RMSE) and mean absolute error (MAE) are as follows:

$$MAPE = \left(\frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right) \times 100\%, \tag{12}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}, \tag{13}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i|, \tag{14}$$

where, y_i is the observed wind speed $\hat{y}_i$ is the forecasted wind speed and n is the number of flow series that has been modelled.

3 Result and Discussion

In this study, four forecasting models have been developed to forecast the wind speed in Malacca state's station. The study's main objective is to investigate the GMDH- hyperbolic tangent model forecasting performance compared to other machine learning models that are commonly used for wind forecasting, such as the ANN model. In order to evaluate the model, the wind speed data is split into training and forecasting data set. Three statistical measures are used to assess the forecasting model's performance, which are MAPE, RMSE and MAE. Each forecasting model is fitted with five different input variables (M1, M2, M3, M4 and M5). Normalization of the data is essential as the function used in this study is defined from 0 to 1 and this applies to both ANN and GMDH models. The five different input variables implemented in this study are represented by M1, M2, M3, M4 and M5. The inputs are obtained from auto-correlation function (ACF) plot, auto-correlation function (PACF) plot and stepwise regression. The inputs are defined in Table 3.

Table 3: Input Variables

Input Variables	Input
M1	y_{t-19}, y_{t-20}
M2	$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}$
M3	$y_{t-6}, y_{t-11}, y_{t-18}, y_{t-19}, y_{t-20}$
M4	$y_{t-6}, y_{t-8}, y_{t-11}, y_{t-16}, y_{t-17}, y_{t-19}, y_{t-20}$
M5	$y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}, y_{t-10}, y_{t-12}, y_{t-15}$

As stated earlier, the percentage of the training data is 70%. The training data set is used to construct the time series model for four forecasting models used in this study. In terms of ANN model architecture, the epoch number is set to 1000 for every ANN model. Then, the number of the hidden layer is selected based on Table 2. Each input variable for forecasting models is defined by Table 3. The best architecture that produces the best forecasting performance for each input is shown in Table 4.

The forecasting of wind speed is evaluated using three statistical measures: MAPE, RMSE and MAE. Table 4 shows the forecasting performance of four forecasting models with various input variables. There are two distinct ANN models with different transfer functions. Based on our observations, the ANN model with hyperbolic tangent transfer function produced the best prediction performance when M3 was used as the input variable, while the ANN log sigmoid transfer function produced the best prediction performance when M2 was used as the input variables. The best ANN-hyperbolic tangent produced MAPE, RMSE and MAE that are 12.0150 m/s, 1.5024 m/s and 1.2101 m/s, respectively. Then best ANN log sigmoid produced of MAPE, RMSE and MAE 12.7501 m/s, 1.6551 m/s and 1.2808 m/s, respectively. The best GMDH-Polynomial model produces the best forecasting performance when M3 was used as the input variable. It produced MAPE, RMSE and MAE that are 11.7817 m/s, 1.4809 m/s and 1.2180m/s, respectively. The m/s notation is the international system unit for velocity of wind speed. Figure 4 shows the forecasting accuracy evaluation based on MAPE, RMSE and MAE.

Figure 4 shows the graphical evaluation of MAPE, RMSE and MAE for four forecasting models with five different input variables. In addition, the lowest MAPE, RMSE and MAE values are 10.4614 m/s, 1.2818 m/s and 1.0310 m/s. The values indicate a good forecasting performance by the GMDH-HT model when using M4 as input variables. Based on Table 4 and Figure 4, the best GMDH-Hyperbolic Tangent model produced the best forecasting performance when M4 was used as the input variable. It produced MAPE, RMSE and MAE that are 11.7817 m/s, 1.4809 m/s and 1.2180m/s,

Table 4: Evaluation of Forecasting Performance of Data-driven Models

	MAPE	RMSE	MAE
	ANN- Hyperbolic Tangent		
M1	12.5099	1.2020	1.1850
M2	13.4868	1.7312	1.3428
M3	**12.0150**	**1.5024**	**1.2101**
M4	13.5528	1.5667	1.2388
M5	16.6260	1.9969	1.6101
	ANN-Log sigmoid		
M1	13.2170	1.5767	1.2404
M2	**12.7501**	**1.6551**	**1.2808**
M3	16.8817	2.0599	1.6378
M4	15.6081	1.9007	1.5240
M5	13.9595	1.7830	1.3825
	GMDH-Polynomial		
M1	13.2605	1.6073	1.3281
M2	14.6179	1.7713	1.4479
M3	**11.7817**	**1.4809**	**1.2180**
M4	13.0003	1.5979	1.3121
M5	12.8582	1.5855	1.2993
	GMDH-Hyperbolic Tangent		
M1	14.9478	1.5133	1.2672
M2	12.1405	1.3017	1.0776
M3	11.8870	1.2969	1.0627
M4	**10.4614**	**1.2818**	**1.0310**
M5	12.1817	1.4657	1.1783

respectively. MAPE error indices measure the forecast accuracy of the forecasting model. The lower number of MAPE show a great forecasting accuracy of the forecasting model. MAPE with a value of 0 signifies perfect forecasting. Based on Table 4, the proposed GMDH model with hyperbolic tangent has the lowest MAPE value of 10.4614 m/s. This indicates that the GMDH-hyperbolic tangent outperformed other forecasting models. The forecasting model with the highest MAPE value is the ANN model with log sigmoid as a transfer function. Therefore, it indicates that the ANN-log sigmoid model has the poorest forecasting performance. Other than that, incorporating hyperbolic tangent as a transfer function for the GMDH model has improved the prediction performance of the conventional GMDH model, which utilizes polynomial as a transfer function. In order to determine the best model in terms of magnitude error, the forecast model should have the lowest RMSE and MAE value. Based on Table 4, the empirical result clearly shows that the GMDH-Hyperbolic Tangent model outperformed other forecasting models in forecasting performance. Furthermore, based on Figure 5, the forecasted data did not deviate much from the observation data. Implementing hyperbolic tangent transfer function manages to enhance the GMDH model performance by 11.21% for MAPE,

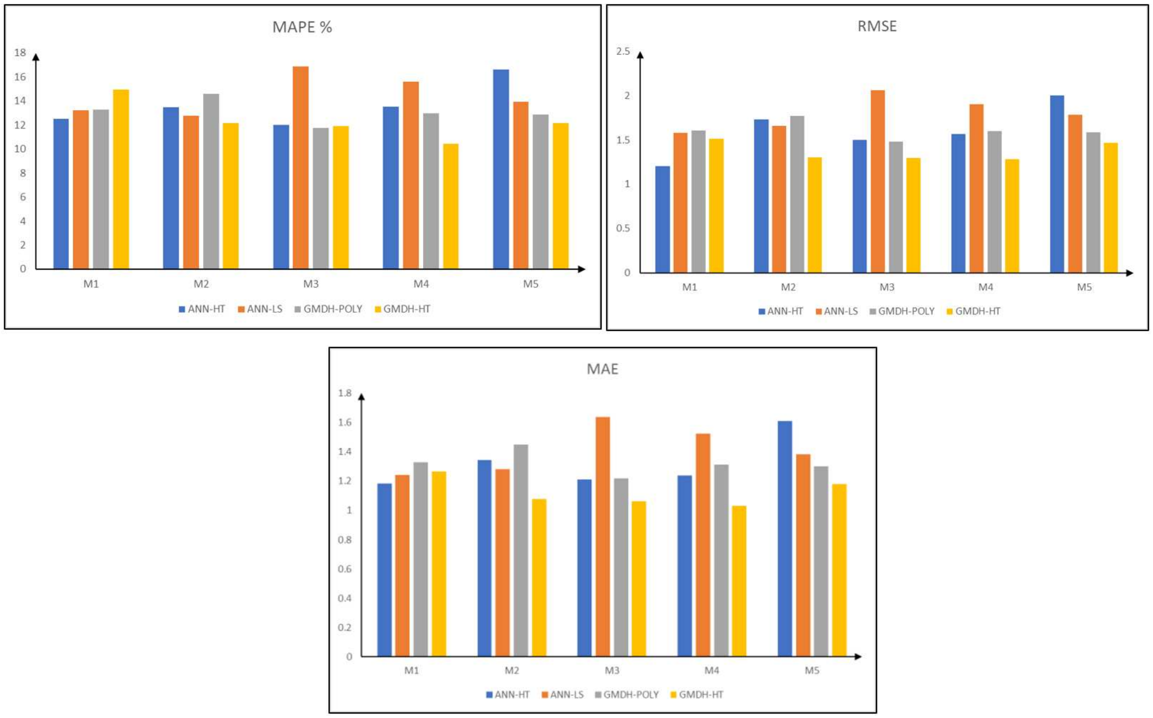

Figure 4: Graphical Evaluation of MAPE, RMSE and MAE

13.44% for RMSE, and 18.7% for MAE. The results showed that the proposed GMDH-Hyperbolic Tangent model could produce high accuracy at forecasting wind speed. Other than that, the results showed significant forecast when using longer lags rather than shorter lag for time series model input variables. The empirical results on the four models showed that GMDH-Hyperbolic Tangent outperformed other data-driven models.

4 Conclusion

In this study, four data-driven models were developed to forecast the wind speed at Malacca state's Station. The models are the ANN-Hyperbolic Tangent, ANN-Log Sigmoid, GMDH and GMDH Hyperbolic Tangent model. This study also investigates the impact of using different transfer functions in data-driven models. Usually, the application of the GMDH model only implements the second-order polynomial as a partial description. However, the implementation of the hyperbolic tangent transfer function in the GMDH model is proposed in this study. The proposed model is compared with the conventional GMDH and ANN models with two different transfer functions: hyperbolic tangent and sigmoid transfer. The proposed model shows highly accurate wind speed forecasting with less computational compare to the ANN model. Based on the results, the proposed model managed to improve the forecasting performance of the conventional GMDH model by 11.21% for MAPE. Furthermore, the proposed model also outperformed the other two ANN models. For future studies, other transfer functions such as radial basis function and sigmoid can be implemented in the GMDH model to enhance the performance of the conventional GMDH model.

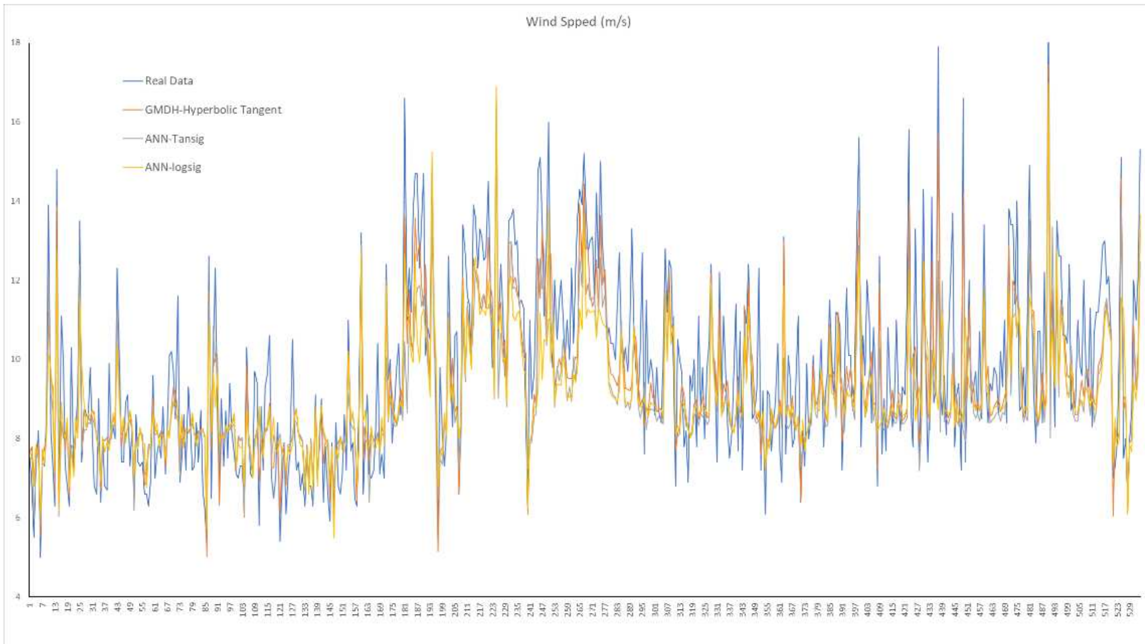

Figure 5: Plot of the Forecasting Data Set for the Individual Forecasting Model

References

[1] Peiris, A. T., J. Jayasinghe, and U. Rathnayake. Forecasting wind power generation using artificial neural network: "Pawan Danawi"—A case study from Sri Lanka. *Journal of Electrical and Computer Engineering.* 2021. 2021: 1-10.

[2] Konneh, K. V., H. Masrur, D. A. Konneh, and T. Senjyu. Independent or complementary power system configuration: A decision making approach for sustainable electrification of an urban environment in Sierra Leone. *Energy.* 2022. 239: 122310.

[3] Ağbulut, Ü. A novel stochastic model for very short-term wind speed forecasting in the determination of wind energy potential of a region: A case study from Turkey. *Sustainable Energy Technologies and Assessments.* 2022. 51:101853.

[4] Navas, R. K. B., S. Prakash, and T. Sasipraba. Artificial Neural Network based computing model for wind speed prediction: A case study of Coimbatore, Tamil Nadu, India. *Physica A: Statistical Mechanics and its Applications.* 2020. 542: 123383.

[5] Zhuo, C., G. Junhong, L. Wei, Z. Fei, X. Chan, and P. Zhangrong. Changes in wind energy potential over China using a regional climate model ensemble. *Renewable and Sustainable Energy Reviews.* 2022. 159:112219.

[6] Badyalina B. and A. Shabri. Streamflow Forecasting at Ungauged Sites using Multiple Linear Regression. MATEMATIKA. 2013. 29.

[7] Duplyakov, V. M., A. D. Morozov, D. O. Popkov, E. V. Shel, A. L. Vainshtein, E. V. Burnaev,... & G. V. Paderin. Data-driven model for hydraulic fracturing design optimization. Part II: Inverse problem. *Journal of Petroleum Science and Engineering.* 2022. 208: 109303.

[8] Ahmad T. and D. Zhang. A data-driven deep sequence-to-sequence long-short memory method along with a gated recurrent neural network for wind power forecasting. *Energy.* 2022. 239: 122109.

[9] Lydia M. and G. E. P. Kumar. Deep Learning Algorithms for Wind Forecasting: An Overview. *Artificial Intelligence for Renewable Energy Systems.* 2022. 129-145.

[10] Pillai, S. R., A. Yadav, and V. K. Vashishtha. Prediction of Wind Power Curve Based on Wind Speed and Direction Utilizing Artificial Neural Network. *Recent Advances in Mechanical Engineering*: Springer. 2021. 515-522.

[11] Dosdoğru A. T. and A. B. İpek. Hybrid boosting algorithms and artificial neural network for wind speed prediction. *International Journal of Hydrogen Energy.* 2022. 47(3): 1449-1460.

[12] Piazza, A. D., M. C. Di Piazza, G. La Tona, and M. Luna. An artificial neural network-based forecasting model of energy-related time series for electrical grid management. *Mathematics and Computers in Simulation.* 2021. 184: 294-305.

[13] Zhang, G. P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing.* 2002. 50: 159-175.

[14] C. Karaman, O. Karaman, P.-L. Show, H. Karimi-Maleh, and N. Zare. Congo red dye removal from aqueous environment by cationic surfactant modified-biomass derived carbon: equilibrium, kinetic, and thermodynamic modeling, and forecasting via artificial neural network approach. *Chemosphere.* 2022. 290: 133346.

[15] Hashim, F. R., P. Nagappan, M. T. Ishak, N. F. Joini, N. F. Makmor, M. S. Saleh, & N. Zolkiply. Solar location estimation using logsig based activation function using artificial neural network approach. *Zulfaqar Journal of Defence Science, Engineering & Technology*. 2021.4(1): 1-7.

[16] Malhan P. and M. Mittal, ”A novel ensemble model for long-term forecasting of wind and hydro power generation,” *Energy Conversion and Management,* vol. 251, p. 114983, 2022.

[17] Li G. and J. Shi. On comparing three artificial neural networks for wind speed forecasting. *Applied Energy. 2010.* 87(7): 2313-2320.

[18] Fazelpour, F., N. Tarashkar, and M. A. Rosen. Short-term wind speed forecasting using artificial neural networks for Tehran, Iran. *International Journal of Energy and Environmental Engineering,* 2016. 7(4) 377-390.

[19] Navas, R. K. B., S. Prakash, and T. Sasipraba. Artificial Neural Network based computing model for wind speed prediction: A case study of Coimbatore, Tamil Nadu, India. *Physica A: Statistical Mechanics and its Applications.* 2020. 542. 123383.

[20] Badyalina B. and A. Shabri. Flood estimation at ungauged sites using group method of data handling in Peninsular Malaysia. *Jurnal Teknologi.* 2015. 76(1): 373–380.

[21] Stefenon, S. F., M. H. D. M. Ribeiro, A, Nied, V.C. Mariani, L. D. S. Coelho, D. F. M. D. Rocha, ... & A. E. D. B. Ruano. Wavelet group method of data handling for fault prediction in electrical power insulators. *International Journal of Electrical Power & Energy Systems.* 2020. 123: 106269.

[22] Shariq, A., A. Hussain, and M. A. Ansari. Prediction of Discharge Coefficient for Side Rectangular Weir Using Group Method of Data Handling (GMDH). *River Hydraulics*: *Springer*. 2022. 83-95.

[23] Badyalina, B., A. Shabri, and M. F. Marsani, ”Streamflow Estimation at Ungauged Basin using Modified Group Method of Data Handling,” *Sains Malaysiana,* 2021. 50(9): 2765-2779.

[24] Ivakhnenko A. G. The group method of data of handling; a rival of the method of stochastic approximation. *Soviet Automatic Control.* 1968. 13. 43-55.

[25] Amiri M. and S. Soleimani. A Hybrid Atmospheric Satellite Image-Processing Method for Dust and Horizontal Visibility Detection through Feature Extraction and Machine Learning Techniques. *Journal of the Indian Society of Remote Sensing,* 2022. 50(3): 523-532.

[26] Vukovic, D. B., K. Romanyuk, S. Ivashchenko, and E. M. Grigorieva. Are CDS spreads predictable during the Covid-19 pandemic? Forecasting based on SVM, GMDH, LSTM and Markov switching autoregression. *Expert systems with applications.* 2022. 194; 116553.

[27] Özkal C. B. and Ö. ARSLAN. Developing a GMDH-type neural network model for spatial prediction of NOx: A case study of Çerkezköy, Tekirdağ. *Environmental Research and Technology.* 2022. 5(1): 56-71.

[28] Lal, A. and B. Datta, "Application of the group method of data handling and variable importance analysis for prediction and modelling of saltwater intrusion processes in coastal aquifers," *Neural Computing and Applications,* 2021. 33(9): 4179-4190.

[29] Ebtehaj, I., H. Bonakdari, A. H. Zaji, H. Azimi, and F. Khoshbin. GMDH-type neural network approach for modeling the discharge coefficient of rectangular sharp-crested side weirs," *Engineering Science and Technology, an International Journal,* 2015. 18 (4): 746-757.

[30] Kondo, T and J. Ueno. Multi-layered GMDH-type neural network self-selecting optimum neural network architecture and its application to 3-dimensional medical image recognition of blood vessels. *International Journal of innovative computing, information and control.* 2008. 4(1): 175-187.

[31] Kalman B. L. and S. C. Kwasny. Why tanh: choosing a sigmoidal function," in *[Proceedings 1992] IJCNN International Joint Conference on Neural Networks*, IEEE. 1992. 4: 578-581.

[32] Rajput, G., G. Raut, M. Chandra, and S. K. Vishvakarma. VLSI implementation of transcendental function hyperbolic tangent for deep neural network accelerators. *Microprocessors and Microsystems.* 2021. 84:104270.

[33] Mokhtar, N. A., B. Badyalina, K. L. Chang, F. Yaa'cob, A. Ghazali, and P. Shamala. Error-in-Variables Model of Malacca Wind Direction Data with the von Mises Distribution in Southwest Monsoon. *Applied Mathematical Sciences.* 2021. 15(9): 471-479.

[34] Jan, N. A. M., A. Shabri, J. Hounkpè, and B. Badyalina. Modelling non-stationary extreme streamflow in Peninsular Malaysia. *International Journal of Water,* 2018. 12 (2):116-140.

[35] Badyalina, B., N. A. Mokhtar, N. A. M. Jan, N. H. Hassim, and H. Yusop. Flood Frequency Analysis using L-Moment For Segamat River. *MATEMATIKA,* 2021. 37.

[36] Mokhtar, N. A., Y.Z. Zubairi, A.G. Hussin, B. Badyalina, A. F. Ghazali, F. F. Ya'acob, P. Shamala, and L. C. Kerk. Modelling wind direction data of Langkawi Island during Southwest monsoon in 2019 to 2020 using bivariate linear functional relationship model with von Mises distribution. *Journal of Physics: Conference Series*. 2021.1988(1): 012097.

[37] Tian, S., N. I. Arshad, D. Toghraie, S. A. Eftekhari, and M. Hekmatifar.Using perceptron feed-forward Artificial Neural Network (ANN) for predicting the thermal conductivity of graphene oxide-Al2O3/water-ethylene glycol hybrid nanofluid. *Case Studies in Thermal Engineering*. 2021. 26. 101055.

[38] Salehuddin, N. F., M. B. Omar, R. Ibrahim, and K. Bingi, A Neural Network-Based Model for Predicting Saybolt Color of Petroleum Products. *Sensors*. 2022. 22(7): 2796.

[39] He, L., W. Fang, G. Zhao, Z. Wu, L. Fu, R. Li, Y. Majeed and J. Dhupia. Fruit yield prediction and estimation in orchards: A state-of-the-art comprehensive review for both direct and indirect methods. *Computers and Electronics in Agriculture,* 2022. 195: 106812.

[40] Farlow, S. J. *Self-organizing methods in modeling: GMDH type algorithms.* CrC Press. 2020.

[41] López G. and P. Arboleya, Short-term wind speed forecasting over complex terrain using linear regression models and multivariable LSTM and NARX networks in the Andes Mountains, Ecuador. *Renewable Energy*. 2022. 183: 351-368.