A comparative study of ARIMA-GARCH model and artificial neural network model for wind speed forecasting

Conference Paper *in* AIP Conference Proceedings · February 2023

DOI: 10.1063/5.0116189

3 authors:

Nor Hafizah Hussin
Technical University of Malaysia Malacca

Fadhilah Yusof
Universiti Teknologi Malaysia

Siti Mariam Norrulashikin
Universiti Teknologi Malaysia

A comparative study of ARIMA-GARCH model and artificial neural network model for wind speed forecasting

Cite as: AIP Conference Proceedings **2500**, 020009 (2023); https://doi.org/10.1063/5.0116189
Published Online: 08 February 2023

Nor Hafizah Hussin, Fadhilah Yusof and Siti Mariam Norrulashikin

A Comparative Study of ARIMA-GARCH Model and Artificial Neural Network Model for Wind Speed Forecasting

Nor Hafizah Hussin[1, 2,b)], Fadhilah Yusof [1, a)], and Siti Mariam Norrulashikin[1, c)]

[1]*Department of Mathematical Sciences, Faculty of Science, UniversitiTeknologi Malaysia, 81300 Skudai, Johor, Malaysia*
[2]*Fakulti TeknologiKejuruteraan Elektrik &Elektronik, UniversitiTeknikal Malaysia Melaka, Hang Tuah Jaya, 76100, Durian Tunggal, Melaka, Malaysia.*

[a)] *Corresponding author: fadhilahy@utm.my*
[b)]*norhafizah.hussin@graduate.utm.my*
[c)]*sitimariam@utm.my*

Abstract. Wind energy is a noteworthy alternate energy in times of energy crisis. An accurate wind speed forecasting model are essential in determining the suitable location for wind energy harvesting. However, the intermittency and nonlinearity of a wind speed make it difficult to obtain an accurate prediction and may cause several operational challenges to grid interfaced the wind energy system. In this study, the time series model and artificial neural network (ANN) model was applied on the daily wind speed data in Senai and Mersing, Johor to forecast future wind speed series. For the time series model, the daily wind speed data was initially been modelled using theARIMA model. However, due to the presence of heteroscedastic effect in the residuals of ARIMA model, GARCH model was introduced to handle the nonlinearity criteria. On the other hand,the Multilayer Perceptron (MLP)model which is in the class of feed-forward ANN was developed with different configurations based on selected hyperparameters. The MLP model configuration with the lowest RMSE value was selected as the best MLP model. A comparison has been made between the ARIMA-GARCH model and theMLPmodel. Results indicate that the MLP model was found to outperform the ARIMA-GARCH model by providing the lowest value of root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) in both training and testing data sets. Thus, the artificial neural network can be concluded as the best method to provide a good forecasting model in predicting the daily wind speed data.

INTRODUCTION

Wind energy has become one of the most vital and efficient renewable energy sources due to the numerous benefits in term of economy and environmentally friendly. Not only due to the availability and free-cost nature, it also plays a large contribution especially in preserving the natural non-renewable source, diminishing the environmental pollution, and handling the greenhouse effect [1].

Likewise, Malaysia is battling with the depletion of natural resource due to the lack of renewable energy sources. A highly dependencies on fossil fuel, which is over 90% to support the electricity generation not only will leads towards depletion, it also will increase the occurrence of air pollution. Therefore, the Malaysian government is intensifying on the usage of renewable energy for energy generation, for instance wind energy projects [2]. To obtain a maximum wind power generation, the site selection for potential wind farm is a very crucial step. The

The 5th ISM International Statistical Conference 2021 (ISM-V)
AIP Conf. Proc. 2500, 020009-1–020009-10; https://doi.org/10.1063/5.0116189
Published by AIP Publishing. 978-0-7354-4310-5/$30.00

potential location should have a higher wind speed than the cut-in wind speed needed to operate the wind turbine. Since the wind speed and wind power are known to have a cubic proportional relation, this means that any slight changes that occur in the wind speed will results in relatively higher wind power (cubic) [3].

Thus, it is vital to obtain a suitable and accurate model to represent a wind speed data that can provide a long-term prediction of wind speed. By providing a suitable wind speed prediction model, it will be helpful to minimize the risk of selecting a less effective location for wind farm installation. It can also be used to assist the authorities in determining the suitable location and planning for electrical grid operation for wind energy conversion system in Malaysia.

Wind speed prediction can be performed based on several approaches, such as, physical approach, statistical approach, and hybrid methods [4]. The commonly used approach to build a prediction model for wind speed data is the statistical approach, which consist of time series models, spatial correlation and artificial intelligence (AI) methods. The time series models include, the Box-Jenkins method such as, auto regressive moving average (ARMA) model and auto regressive integrated moving average (ARIMA) model. While the AI methods are based on the artificial neural network (ANN), such as Multilayer Perceptron (MLP), Support Vector Machines (SVM), and Long Short-Term Memory (LSTM).

An extensive comparative study between the classical time series model and AI models have been performed by the previous researchers to determine the best model for wind speed forecasting. Most of the comparative analysis are made between the ARIMA and ANN models. ARIMA is a well-known model used for wind speed forecasting through analyzing historical data [5-6]. While ANN is amodel that are commonly used due to the ability on capturing the complex nonlinear criteria exist in the wind speed data [7-9].

A study conducted by [1] introduced ARIMA and ANN model to forecast the wind speed data in Hubei province of China. The ANN model was found to give a better performance than ARIMA model in forecasting short term hourly wind speed. While [10] performed a comparative study between the feed forward back propagation neural network (ANN) and ARIMA for wind speed and direction forecasting in Turkey. The finding shows that ANN method outperformed ARIMA by giving a better solution in MSE. A study from [11] examined three forecasting techniques to model and forecast the wind speed data in Kiribati. The empirical results reveal that the ANN provides a more efficient and accurate forecasting model as compared to the regression and time series ARIMA model.

Despite the fact that the ARIMA is powerful and flexible time series model, there are cases where ARIMA was unable to provide a stable model due to the presence of heteroscedastic effect in the wind speed data. Hence, many researchers combine the ARIMA with GARCH model to capture the nonlinearity effects. A study by [12] suggested a method using ARIMA-GARCH model to forecast short-term wind speed data. It is proven that the proposed model has successfully handle the heteroscedastic effect presence in the wind speed data, where the accuracyof the predictionby the proposed model is higher than the ARIMA model. While [13] proves that ARIMA-GARCH have successfully captured the nonlinear criteriaexist in the daily wind speed data in Peninsular Malaysia.

The primary target of this study is to compare the performance of time series model and artificial neural network model in providing the best wind speed forecasting model. The daily wind speed data in Mersing and Senai stations are analyzed using the ARIMA-GARCH and ANN model, where the forecasting performance will be measured based on the value of root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). The model with a better forecasting performance measure will be selected as the best daily wind speed forecasting model.

METHODOLOGY

Study Area and Data

In this study, data measured in two meteorological stations which are Mersing and Senai, Johorwhich is located in the southern region of the Peninsular Malaysia, were used. These stations were chosen since it has the highest mean of the wind speed. Mersing station is located in the coastal area of Mersing district with a geographical coordinate between 2°26'N and 103°49'E. While Senai station is located in the Senai International Airport with a geographical coordinate between 1°38'N and 103°39'E.This data used a in this research are the daily wind speed dataobtained from Malaysian Meteorological Department (MMD). The data from Mersing station was taken from January 1990 to December 2019, while Senai station was taken from January 1985 to December 2019.The statistical description for the data in these stations are provided in Table 1.

TABLE1. The descriptive statistics of daily wind speed data

Station	Mean	Standard Deviation	Min	Max
Mersing	9.8923	2.7838	3.5	36.5
Senai	9.2283	2.3539	1.1	28.6

For the modelling purpose, the data will go through a pre-processing to determine the existence of missing values before modelling is performed. Data splitting will also be performed witha training data having a higher allocation as compared to the testing data; i.e.; 90% to 10%, 80% to 20% and 70% to 30% [14]. In this study, the data will be split into 80% training data and 20% testing data. The testing data will then be used to compare with the forecasted daily wind speed dataobtained from each model.

Time Series Model

ARIMA Model

In statistical time series modelling, the Box-Jenkins method or autoregressive integrated moving average (ARIMA) model is widely used in building a wind speed forecasting model. The general form of ARIMA (*p, d, q*) can be expressed in the following mathematical equation:

$$\varphi(B)(1-B)^d y_t = \theta(B)\varepsilon_t \tag{1}$$

where y_t and ε_t represents the observed value of wind speed data and the random error term with respect to time t, respectively, φ_1, φ_2, φ_3,…, φ_p denotes the autoregressive coefficient of order p, d denotes the differencing order, while θ_1, θ_2, θ_3,…, θ_q denotes the moving average coefficient of order q. B represents the backward shift operator, while, $\varphi(B) = 1 - \sum_{i=1}^{p} \varphi_i B^i$ and $\theta(B) = 1 - \sum_{j}^{q} \theta_j B^j$ are the polynomials with respect to B with order p and q respectively.

Therefore, ARIMA (*p, d, q*) model that mixed the autoregressive (AR) and moving average (MA) can be expressed in the following mathematical equation:

$$v_t = \varphi_1 v_{t-1} + \varphi_1 v_{t-2} + \ldots + \varphi_p v_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \tag{2}$$

There are four methodology procedure for this method, which are the model identification, parameter estimation, diagnostic checking and model forecasting. These methodology aims to obtain a parsimony model that contain the smallest number of parameters needed to adequately fit the pattern in the observed data. A stationary test will first be conducted to measure the stationarity of the time series. A non-stationary series data has to be transformed using the suitable order of differencing, *d*, into a stationary time series.

Once the data are concluded to be stationary, the auto correlation function (ACF) and partial auto correlation function (PACF) plot will be used to identify the order of the time series ARIMA model. In the parameter estimation stage, the maximum likelihood estimator (MLE) will be used. While the model selection will be performed by referring to the value of Akaike Information Criterion (AIC), where the model that has thesmallest AIC value will better estimation model.

The diagnostic checking on the residuals and squared residuals of the model can be performed using the Ljung-Box Q statistics, to determine the presence of autocorrelation and heteroscedasticity effect, respectively [15]. The model is considered as adequate if the *p*-value associated with the Q statistics test are greater than 5% significance level. At this point, if the residuals are relatively small and does not contain any unnecessary information, such as serial correlation and heteroscedastic effect, the model can proceed for forecasting. Nonetheless, if the model is inadequate, the GARCH model will be introduced towards the residuals in order to capture the nonlinearity in the wind speed data.

GARCH Model

Generalized Autoregressive Conditional Heteroscedastic (GARCH) is the generalized ARCH model that was developed by [16]. This model is known to help the ARIMA model to treat the presence of heteroscedastic in a time series data. To model a univariate data, let $y_t = \mu_t + z_t$ represents the mean equation at time t, where the conditional mean of y_t is denoted by μ_t, z_t represents the shock with respect to time t, and the equation is $z_t = \sigma_t \varepsilon_t$, $\varepsilon_t \sim iid\ N(0,1)$. Then, the conditional variance of y_t denoted by σ_t^2, that follows a GARCH (*p, q*) model can defined as follows:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i z_{t-i}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2 \tag{3}$$

where $\propto_0 > 0$ and $\sum_{i=1}^{\max(p,q)} (\propto_i + \beta_i) < 1$. The parameterscoefficient for ARCH and GARCH are denoted by $\propto_i$andβ_i respectively. In a situation where all the coefficient of β gives a value of zero, the GARCH model will then change to the ARCH model.

ARIMA-GARCH Model

There are two procedures in modelling using the hybrid of ARIMA and GARCH model. In the first step, ARIMA model will performed the modelling forthe linear part of the wind speed data, while, the residual will just consist of the nonlinear data [17]. It is then followed by the second step which is modelling the residual that contain only the nonlinear pattern using the GARCH model. The hybridization of ARIMA modelwith the error component from GARCH model yields amodel that are known to be able to handle the dynamics in wind speed data and to predict the values of future wind speed data. According to [18-19], the easiest GARCH model is GARCH (1, 1). Therefore, this study will apply the standard GARCH (1,1) as the benchmark model to be used in capturing the heteroskedastic effects of the time series process.

Artificial Neural Network

ANN is one of the commonly used models which belongs to the class of artificial intelligence method for time series analysis. A simple feed forward neural network model, also known as the multilayer perceptron (MLP) consist of three layers and should be evaluated before going into a more complex model.

MLPModel

The MLP modelis a feed-forward ANN which are commonly used in the wind speed forecasting area. It is consist of units arranged in layers where it is composed of a minimum of three layers which are the input layer, one or more hidden layer, and the output layer. The first layer which is the input layer receives the input values which then get processed by each nodes of the hidden layer. Then the output layer will generate the desired output [20]. Each processing elements in the MLP model are known to be interconnected nodes that are linked by the adaptable weights. The input signals received by each nodes are basically the output of other nodes. Here, the output of each nodes is a function of the weighted input, activation function, and bias. This can be expressed as follows:

$$y_t = f\left(\sum_{i=1}^{n} (w_i . I_i) + b\right) \tag{4}$$

wherey_t is the output from the nodes and $n$is the number of inputs. While*f*, w_i,I_i, and b represents the activation function, theconnection weight of the *i*th input, the*i*th input to the nodes, and the bias, respectively.

Building a model for a univariate time series using MLPmodel can be performed by setting a certain number of lagged terms in the series as the input, while the forecasted results as the output [21]. For the application of ANN in time series, [22-23] have established a methodology to determine the lagged term for the number of inputs. Since the data used are the daily wind speed data, this study will carried out a forecast of each time point from the 30 lagged term which means that the prediction will be made based on the last 30-day data. Here, the input carries the first 30 days and the output will be the daily wind speed data for the next day since we are interested in developing a one-step forecast model.

The rules of thumb in determining the number of hidden nodes in the hidden layers is, it should be in the range of input layer size and output layer size, two thirds of the input layer size plus the output layer size, and less than twice the size of input layer [24]. As for the activation function, this study used the rectified linear activation function (ReLU) for the hidden layer. While for the output layer, linear activation function will be use since we are predicting a continuous value. The equations and plots for the ReLU and linear activation functions are shown in Table 2.

The parameters for learning rates, batch size, number of epochs, and early stopping criterion were determinedby referring to the previous research and will then be investigated based on trial-and-error basis [25]. The training parameters used in modelling using the MLPis summarized in Table 3.

TABLE2. The activation function for the MLP model

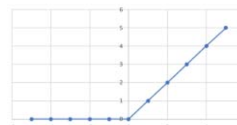
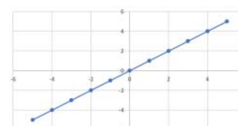
Activationfunction	Equation	Plot
ReLU	$f(x) = \max(0, x)$	
Linear	$f(x) = x$	

TABLE3. Thetraining parameters for the MLP model

Parameter	Descriptions
Number of Input	30 lagged terms
Number of Hidden Nodes	20
Activation function	Rectified linear unit (ReLU)
Loss function	Mean squared error loss
Optimizer	Adam flavor of Stochastic gradient descent
Learning Rates	0.001
Batch Size	500, 1000
Epochs	100, 300

Model Performance Measurement

The performance for each model used in this study will be measured based on the calculation of difference between the observed wind speed data, and the model predicted wind speed data. The forecasting performance will be evaluated using the RMSE, MAE, and MAPE. The equations can be expressed as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \tag{5}$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|y_t - \hat{y}_t\right| \tag{6}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100 \tag{7}$$

where y_t and $\hat{y}_t$ are the observed and predicted wind speed data, while n is the number of data. Here, the model with the smallest value of RMSE, MAE, and MAPE will be concluded as the best wind speed forecasting model

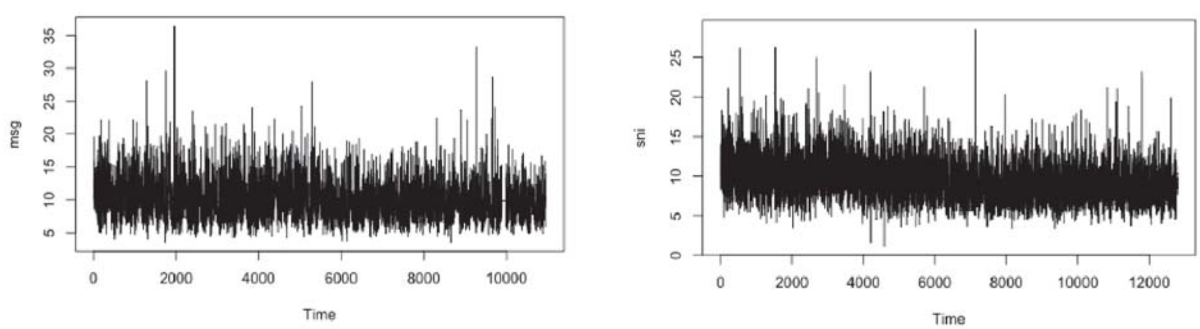
RESULT AND DISCUSSION

Data Description

This research used a daily wind speed data for stations in Johor; Mersing and Senai. The data from Mersing station was taken from January 1990 to December 2019, with a total of 10,958 observations. While Senai station was taken from January 1985 to December 2019, with a total of 12,783 observations. The data are divided into 80% of training data and 20% of testing data. The analysis of the model selected in this study are based on the same data proportion.

Time Series Model

The first step in Box-Jenkins methodology is to observe whether the wind speed data display any trend or seasonality which can be done by plotting a time series plot. Based on the plots in Fig. 1, the daily wind speed does not display any trend or seasonality since it is not varied in a fixed level.



(a) (b)

FIGURE 1. The time series plot for station in (a) Mersing and (b)Senai

The stationarity of the time series is then investigated using the unit root test of Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test. Based on the unit root test for both stations, the p-value of 0.01 for the KPSS test concluded that the series are not stationary. The ACF plot for both stations show a slow decaying which also suggest that the time series are not stationary and should undergo differencing in order to obtain a stationary series. Once the first differencing is performed, the KPSS test for both stations shows a p-value of 0.1 which concluded that the first differenced series is stationary. This result is also supported by the ACF plot that shows a drastic decay to zero. The time series plot for the differenced series shows that most of the data are located around mean zero, which also indicates stationarity.

Once the series are stationary, the ACF and PACF plot for the first differenced series are used to decide the suitable component to be included in the ARIMA model. Based on the plots, the possible combinations are based on p=0,1,2 and q=0,1,2, with d=1. From the analysis, the model with the smallest AIC is ARIMA (1, 1,2) for Mersing station and ARIMA (2,1,1) for Senai station. Therefore, these two models will go through the diagnostic checking to measure the adequacy of the potential model.

TABLE 4. The *p*-value for Ljung-Box test for the potential ARIMA model.

Station	ARIMA Model	Residuals		Squared Residuals	
		Lag 10	Lag 20	Lag 10	Lag 20
Mersing	ARIMA (1,1,2)	0.187	0.137	0.000	0.000
Senai	ARIMA (2,1,2)	0.553	0.262	0.000	0.000

Table 4 presents the results for the Ljung-Box test of the potential ARIMA model in both stations selected in this study. The results show that the wind speed data does not exhibit any serial autocorrelation. However, the *p*-value for Ljung-Box test tested on the squared residuals concluded that the heteroscedastic effect is present in the data series for both stations. Thus, the ARIMA models are not adequate and the GARCH model are needed to handle the heteroscedastic effect presence in the data.

This study applies the standard GARCH (1,1) as the benchmark model to model the variance behavior since it is the easiest model GARCH model [18-19], while the Ljung-Box test was tested on the residuals of the ARIMA-GARCH model. For Mersing station, the ARIMA (1,1,2)-GARCH (1,1) model yields a *p*-value of 0.2156 for the Ljung-Box test. While Senai station yields a *p*-value of 0.4011 for the Ljung-Box test. This shows that the GARCH (1,1) model have successfully captured the dynamics in the residual part of the data. In the next step, the obtained model will be used to forecast the daily wind speed data and will be compared with the out-sample data to obtain the forecasting performance measure. Then, the results will be compared with the MLP neural network model in order to conclude the most appropriate forecasting model for the daily wind speed data.

MLP Model

The MLP model is a model that are known to have a capability on providing an efficient forecasting results when the modelling involves a nonlinearity condition. Based on the list of parameters presented in Table 3, a total of four MLP model configurations to be evaluated for both stations. For all configurations, lagged terms of 30 days were used as inputs. Each network consist of single hidden layer with 20 hidden nodeswhich is two thirds of the input layer size plus the output layer size and are trained based on mean squared error loss function with the Adam flavor of stochastic gradient descent as the optimizer. The ReLU activation function will be used in the hidden layer while the linear activation function will be applied in the output layer for all configurations. The main advantages of using ReLU activation function is that there is a fixed derivative for every input larger than zero.

The statistical results obtained by different configurations were evaluated based on the RMSE value [26]. The MLP model configurations with the lowest RMSE valuewill be selected as theconfiguration that provide the best predictionfor the wind speed in both stations and are presented in Table 5 and 6 below. The configurations with the lowest RMSE values were chosen as the best MLP model.

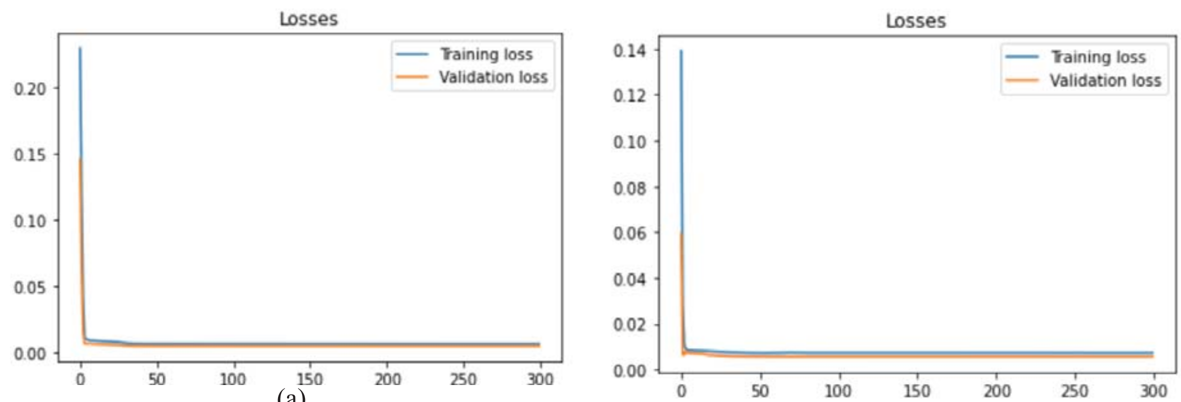
TABLE 5. The MLP model configurations and the forecasting performance measure for Mersing

Model	Batch Size	Epoch	RMSE	Rank
MLP-Msg1	500	100	2.2987	4
MLP-Msg2	500	300	2.2942	1
MLP-Msg3	1000	100	2.2959	3
MLP-Msg4	1000	300	2.2947	2

TABLE 6. The MLP model configurations and the forecasting performance measure for Senai

Model	Batch Size	Epoch	RMSE	Rank
MLP-Sni1	500	100	2.0310	3
MLP-Sni2	500	300	2.0237	1
MLP-Sni3	1000	100	2.0312	4
MLP-Sni4	1000	300	2.0304	2

From the results presented in Table 5 and 6, MLP-Msg2 and MLP-Sni2 models yields the lowest RMSE followed by MLP-Msg4 and MLP-Sni4. Each of these model runs with 300 epochs. Hence, the number of epochs can be considered as one of the influential parameters, where by increasing the iterations, the error become smaller. The learning curves for MLP-Msg2 and MLP-Sni2 models are displayed in Fig. 2. The shape and dynamics of this plot can be used to perform a diagnosis on the behaviour of the neural network model fitting whether the model is underfit, overfit, or good fit. For the learning curve that has a good fit criteria, the validation loss will reduces as the modelling error for the training stage decreases to a point of stability, with a minimal gap between the two graph lines. Based on the plots in Fig. 2, the learning curves for MLP-Msg2 and MLP-Sni2 show a good fit model where the training and validation loss decrease to a point of stability and the distance between the two final loss values at epoch 300 has a minimal gaps.



(a) (b)

FIGURE 2. Learning curves for the best configuration of MLP model in (a) Mersing and (b) Senai

Thus, it can be concluded that MLP-Msg2 and MLP-Sni2 are the best MLP model configurations to provide the best prediction for the wind speed in Mersing and Senai station, respectively. These models will be used to forecast the daily wind speed data and will be compared with thetesting data to obtain the forecasting performance measure.Next, the forecasting performance measure for ARIMA-GARCH and MLP models will be evaluated to conclude the most appropriate wind speed forecasting model.

Model Performance Comparison

For the analysis of the model performance, the value for RMSE, MAE, and MAPE for each model will be measured and compared to select the best model for forecasting the wind speed data with a higher prediction accuracy. The model that yields the lowest value of RMSE, MAE, and MAPE will be concluded as the best wind speed forecasting model. Table 7presents the model performance comparison for the ARIMA-GARCH and the MLP model. The results indicate that the MLPmodel has outperformed the ARIMA-GARCH model by yielding a better forecasting performance measure with lowest value of RSME, MAE, and MAPE for both training and testing data.

TABLE 7. The comparative performance of ARIMA-GARCH and MLP models

Station	**Model**	**Training Data**			**Testing Data**		
		RMSE	**MAE**	**MAPE**	**RMSE**	**MAE**	**MAPE**
Mersing	ARIMA(1,1,2) - GARCH (1,1)	2.6024	1.9455	20.5030	3.6341	3.1362	38.8697
	MLP-Msg2	**2.5773**	**1.9171**	**19.8189**	**2.2942**	**1.5985**	**17.6015**
Senai	ARIMA (2,1,2) - GARCH (1,1)	2.2263	1.7012	19.3187	2.7363	2.3253	31.9202
	MLP-Sni2	**2.2230**	**1.6860**	**18.8375**	**2.0237**	**1.5214**	**18.5576**

Overall, the MLP model for both stations perform well compared to the ARIMA-GARCH model. The MLP model for both stations show a lower value of RMSE and MAPE for training and testing data. The RMSE value for Mersing station is 2.5773 for training data and 2.2942 for testing data which is lower than the ARIMA-GARCH model. While Senai station also show a lower RMSE for MLP model with the value of 2.2230 for training data and 2.02237 for testing data. The error measure for MLP model yields a lower value fortraining and testing data of MAE, with 1.9171 and 1.5985 for Mersing station and 1.6860 and 1.5214 for Senai station. Therefore, the result indicates that the MLP model outperforms ARIMA-GARCH model in providing a better wind speed forecasting result.

As recommended by most of the researchers, the forecasting performance of model can be improvised by performing a hybrid of the time series model and the neural network model [27-28]. A hybrid model can be built by employing the principle of separate modelling of linear and nonlinear components of time series [29]. Therefore, future work can be done by hybridizing the ARIMA-GARCH model with the MLP model to obtain a better forecasting performance. The result of an ARIMA-GARCH model will give the linear forecasted dataset, while the residuals from the ARIMA-GARCH model will be fed to the neural network as the input data for the MLP model which is the nonlinear forecasted dataset.

CONCLUSION

The primary target of this study is to compare the performance of time series model and artificial neural network model in providing the best wind speed forecasting model. Initially, the ARIMA model was used to model the wind speed series in Mersing and Senai, Johor station. However, with a presence of heteroscedastic effect in the residual of the ARIMA model, the GARCH model was introduced to treat the nonlinearity criteria. For Mersing station, ARIMA (1,1,2)-GARCH (1,1) is the best model, while ARIMA (2,1,2)-GARCH (1,1) is the best model for Senai station. For the neural network model, the MLP model was applied by listing four different configurations with different hyperparameter that was obtained using trial and error. The best configuration was selected based on the lowest value of RMSE. Result of the study shows that MLP-Msg2 and MLP-Sni2 models have the best configuration that provide a good fit model with lowest the RMSE value. Thus, it can be concluded that these MLP models can be used to provide an effective daily wind speed modelling.

The ARIMA-GARCH and the MLP model was then used to forecast the daily wind speed data.Three types of forecasting performance measure were being used to evaluate each model which are the RMSE, MAE, and MAPE. The model with the smallest error will provide a more accurate forecasting results and was concluded as the best daily wind speed forecasting model. The MLP model shows a smaller value of RMSE, MAE, and MAPE for the training and testing data in both Mersing and Senai stations as compared to the ARIMA-GARCH model. In conclusion, the artificial neural network model has outperformed the time series model in determining the best method to provide a good forecasting model since it is capable of predicting the data efficiently when it comes to dealing with a nonlinear data.

ACKNOWLEDGMENTS

The authors would like to acknowledge UniversitiTeknikal Malaysia Melaka (UTeM), Ministry of Higher Education (MOHE) for financial funding under the FRGS, Registration Proposal No. : FRGS/1/2018/STG06/UTM/02/2 and UniversitiTeknologi Malaysia (UTM).

REFERENCES

1. L. Chenand X. Lai, "Comparison between ARIMA and ANN models used in short-term wind speed forecasting",in 2011 Asia-Pacific Power and Energy Engineering Conference, (IEEE, 2011), pp. 1–4.
2. A.Albaniand M. Z. Ibrahim,*Energies*,**10**(3), 307–328 (2017).
3. R.Sarkar, S.Julai, S. Hossain, W. T. Chong, andM. Rahman, *Mathematical Problems in Engineering*. 2019.
4. N. C.De Freitas, M. P. D. S.Silva, andM. S. Sakamoto, *International Journal of Engineering Research and Application*, **8(**1), 4–9(2018).
5. E. Grigonytė, and E. Butkevičiūtė, *Energetika*,**62**, 1–2 (2016).
6. K. Yunus, T.Thiringer, and P. Chen,*IEEE Transactions on Power Systems*,**31**(4), 2546-2556(2015).

7. T. Brahimi, F. Alhebshi, H.Alnabilsi, A.Bensenouci, and M. Rahman, *Procedia Computer Science*. **163**, 41–48(2019).
8. Y. P.Faniband and S. M. Shaahid, "Forecasting Wind Speed using Artificial Neural Networks–A Case Study of a Potential Location of Saudi Arabia" in Proceedings of the E3S Web of Conferences, (EDP Sciences, 2020), pp. 1–4.
9. P.J.Zucatelli, E.G.S. Nascimento, G.Y.R. Aylas, N.B.P.Souza, Y.K.L. Kitagawa, A.A.B.Santos, A.M.G. Arce, andD.M.Moreira,*Heliyon*, **5**(5)(2019).
10. B. S. Çorbaand P.Kasap,*American Journal of Engineering Research,* **7**(11), 240–250 (2018).
11. A. Arzu, S. S.Kutty, M. R. Ahmed, and M. G. Khan, "Wind speed forecasting using regression, time series and neural network models: a case study of Kiribati", in Australasian Fluid Mechanics Conference (Australasian Fluid Mechanics Society, Australia, 2018), pp. 1–4.
12. J. Yan, H. Guoqing, P.Xinyan, and L. Yongle,*Journal of Southwest Jiaotong* University,**51**(4), 663–670 (2016).
13. N. H. Hussin, F. Yusof, S. M. Norrulashikin, and A. R. Jamaludin, *Pertanika Journal of Science & Technology*, **29**(1), 31–58 (2021).
14. S. Ismail, A. Shabriand R.Samsudin, *Hydrology and Earth System Sciences*,**16**, 4417–4433 (2012).
15. G. M. Ljung, and G. E. Box,*Biometrika*. **65(**2), 297–303 (1978).
16. T. Bollerslev, *Journal of econometrics*,**31(**3), 307–327 (1986).
17. S. R. Yaziz, N. A.Azizan, R. Zakaria, M. H. Ahmad, "The performance of hybrid ARIMA-GARCH modeling in forecasting gold price", in 20th International Congress on Modelling and Simulation (Adelaide, 2013),pp. 1–6.
18. I. L. Kane and F. Yusof, *Modern Applied Science*,**7**(12), 78 (2013).
19. H. Chen, Q. Wan, F. Li, and Y. Wang,"GARCH in mean type models for wind power forecasting", in Proceedings of the 2013 (IEEE Power & Energy Society General Meeting2013). pp. 1–5.
20. A. F. Antor, and E. D. Wollega, "Comparison of Machine Learning Algorithms for Wind Speed Prediction", in 5th NA International Conference on Industrial Engineering and Operations Management (Detroit, USA, 2020), pp. 857–866.
21. C. M. Bishop, Neural networks for pattern recognition. (Oxford university press, 1995).
22. T. Masters, Practical neural network recipes in C++. (Morgan Kaufmann, 1993).
23. J. J. M. Moreno, *Psicothema*. **23**(2), 322–329(2011).
24. G. Panchal, A. Ganatra, Y. P. Kosta, and D. Panchal, *International Journal of Computer Theory and Engineering*. **3**(2): 332–337(2011).
25. H. Apaydin, H. Feizi, M. T.Sattari, M. S. Colak, S. Shamshirband, and K. W. Chau,*Water*. **12**(5), 1500(2020).
26. J. Brownlee, "Deep learning for time series forecasting: predict the future with MLPs, CNNs and LSTMs in Python", Machine Learning Mastery, (2018).
27. J. Jiao, MATEC Web of Conferences,**232**, 03013(2020).
28. P. K. Singh, N. Singh, and R. Negi,*Ambient Communications and Computer Systems,*209–220 (2019).
29. A. S. Temür, M.Akgün, and G.Temür, *Journal of Business Economics and Management*, **20(**5), 920–938 (2019).