Institute of Research and Publication Indonesia
IJEERE: Indonesian Journal of Electrical Engineering and Renewable Energy
Journal Homepage: https://journal.irpi.or.id/index.php/ijeere
Vol. 4 Iss. 2 December 2024, pp: 77-90
P-ISSN: 2797-1155 E-ISSN: 2797-0868

Multilayer Perceptron Application in Electricity Consumption Forecasting from Wind Power

Meftah Elsaraiti [1*]

[1]Higher Institute of Technical Sciences, Misurata. Libya

Corresponden E-Mail: [1]elsaraitimeftah@gmail.com

Makalah: Diterima 21 June 2024; Diperbaiki 27 June 2024; Disetujui 17 December 2024
Corresponding Author: Elsaraiti

Abstarct

In the context of increasing global demand, and the importance of renewable energy sources. There is a need to focus on renewable energy management and explore more efficient sources of energy generation. In recent years, the rise of the transition from fossil energy to renewable energy and its decreasing cost have posed new forecasting problems of interest to researchers, namely, the knowledge in advance of the electricity consumption that wind power plants may generate in the future. Recognizing the need to improve renewable energy management as a strategy to search for better and more efficient sources of energy generation, a system for forecasting the generation of electricity consumption by wind power using neural networks is proposed that allows maximum utilization of the potential that can produce wind power. A backpropagation neural network is used as an artificial intelligence component to accurately predict electricity consumption. The network is a multi-layer perception (MLP) network that uses a reverse transcription learning algorithm to train a feed-forward neural network to perform a specific task. The Levenberg-Marquardt algorithm is used to improve network training and evaluate its performance. In this study, the suggested approach trains the feed-forward neural network for the specified task using a multi-layer perception (MLP) network and a backpropagation learning method. By modifying the weights of the connections between neurons, the number of neurons in the hidden layers is optimised in order to reduce the output error in comparison to the desired target values and achieve high accuracy. Support vector machine (SVM) and auto-regressive integrated moving average (ARIMA) models' performances are compared to those of the suggested model.

Keyword : Electricity consumption; forecasting; multi-layer perception; neural network

1. Introduction

A new energy community has emerged as a result of the increased interest in the switch to renewable energy sources and the need to lessen the environmental impact of conventional power plants. The chance to make major progress in three critical areas low power prices, supply reliability, and carbon emission reduction is presented by the rising demand for renewable energy sources [1]. Calculation and forecasting of electrical energy consumption are important for improving the efficiency and reliability of energy supply systems. Electricity consumption forecasting has emerged as a significant research field in the electric power sector and is critical for the planning and operational management decisions of electric power systems [2]. The ability of artificial neural networks to handle complicated, nonlinear interactions between variables has led to an increase in their application in the energy sector for forecasting future energy consumption. Numerous research have been done to determine how well ANNs predict energy consumption, and they consistently show that ANNs outperform traditional methods [3]. In the 1970s, the back propagation algorithm was initially proposed. However, it was not until the seminal 1986 work by David Rumelhart, Geoffrey Hinton, and Ronald Williams that its potential for training multi-layer neural networks was fully realised [4]. The back propagation algorithm is a common learning algorithm used in feed-forward multilayer neural networks to improve the efficiency of the network through parallel training. One of the most popular techniques for neural network training and tackling previously unsolvable issues is the backpropagation algorithm. The backpropagation algorithm is the backbone of neural network learning [5]. Backpropagation is a crucial algorithm in training neural networks. It's a method for efficiently computing gradients of the loss function with respect to the weights and biases by exploiting the chain rule of calculus. By calculating this gradient, backpropagation provides a way to update the network's weights and biases in order to minimize the loss function and improve its performance [6]. The Levenberg-Marquardt algorithm is a widely used optimization algorithm, particularly in the context of training feed-forward neural networks. It is a combination aspects of both Newton's method and gradient descent.The algorithm is based on the nonlinear least squares method, which involves minimizing the sum of the squared differences between predictions and actual target values [7]. The backpropagation method effectively trained multilayer perceptrons, which was a significant achievement in the field of neural networks. It has several drawbacks and is not a cure-all. In conclusion, the backpropagation algorithm was a big advancement in the field of neural networks, although it is not a universally applicable technique. One must thoroughly evaluate the issue at hand and experiment with various models and combinations to get the optimal results [8]. The purpose of this study is to contribute to the use of transfer learning to improve the balance between memorization and generalization in training the network to predict the consumption of electrical energy from wind energy. The objective is to train the network to predict electrical energy consumption from wind energy in the United States in a way that strikes a balance between the ability of the network to give the correct response to the input data used in the learning process (memorization) and the ability to produce the correct results in response to the input data that are similar but not identical to those used in training (generalization principle). It sounds like an interesting and relevant study. Overall, designing a neural network to predict energy consumption from wind energy in the United States can have significant implications for the energy sector, and the approach outlined in this study can serve as a useful framework for similar applications in other domains. Training the network by the error backpropagation method to perform a given task based on the most widely used optimisation algorithm, Levenberg-Marquardt, for the ANN training to increase the speed of convergence instead of the conventional backpropagation (BP) method, the model was used to analyse monthly data corresponding to electrical energy usage for the years 2011 to 2016 and to assess the effectiveness of the network trained. By contrasting the predictions of the ANN algorithm with those of SVM and ARIMA, the effectiveness of the algorithm was evaluated. The organisation of the current work is as follows: The past research relevant to the topic is briefly explained in Section 2, the research technique is discussed in Section 3, and the results are analysed in Section 4.simulation, and training of the network and evaluates them by comparing them with the objective results based on the training algorithm used; and finally, Sections 5 and 6 present the conclusion and appropriate references used in the work.

2. Literature Review

Predicting electrical loads is a challenging problem that requires extensive statistical analysis and modeling. For load forecasting, a variety of techniques are utilised, including regression-based techniques, ANN), and (SVM), as well as statistical techniques like ARIMA. Research studies have shown that statistical methods such as ARIMA are well-suited for short-term forecasting where the time series data is stationary and the patterns are relatively simple. In contrast, machine learning methods such as ANN and SVM are better suited for long-term forecasting, where the data is non-stationary and the patterns are more complex [9, 10]. ARIMA and ANN models were compared to predict the electrical load over three different horizons: the first 24 hours, the first 48 hours, and a whole week (168 hours). The ARIMA model performed best with a 24 and 48-hour horizon on the MAPE error scale, while ANN performed better with a 1:168-hour horizon [11]. The ARIMA statistical model outperformed the MLP network with the backpropagation algorithm for short-term time periods to forecast wind speed, as the calculation time was less. This was shown by contrasting the performance of the ARIMA statistical model and the backpropagation neural network with the Levenberg-Marquardt learning algorithm [12]. In some instances, SVM have been found to perform better than conventional statistical models and other machine learning methods in prediction tasks [13, 14]. The application of the SVM model demonstrated its capacity to forecast day-ahead wind speed based on past data. The model's statistical performance was assessed using the mean absolute error percentage (MAPE) and correlation coefficient [15]. ANNs are statistical techniques that are frequently employed in wind energy forecasts. ANNs have the capacity to learn data according to the solution of difficult and time-consuming mathematical equations and can be generated with simpler models. This strategy is suited for wind energy systems with noisy data because of ANNs' fault tolerance and success modelling nonlinear systems [16]. Due to their capacity to simulate nonlinear relationships between time series, ANNs, and SVMs are among the most efficient algorithms. When compared to other techniques, a prediction model cannot be said to have the highest accuracy. Alternatively, depending on numerous variables like the prediction horizon, data properties, and the amount of training data, each forecasting approach may produce reasonable results. For effectively estimating energy, ANNs have proven to be very competitive and highly accurate algorithms [17]. Choosing the right number of hidden neurons for ANN is essential for creating a model for wind speed prediction that is accurate and trustworthy. Depending on the particular data and task, the ideal number of hidden neurons can change, necessitating careful consideration and selection. Trial and error is a popular method for figuring out the ideal number of hidden neurons. This involves evaluating the performance of the model using an appropriate evaluation metric while iteratively training and testing the ANN with various numbers of hidden neurons. The ideal number of hidden neurons for the task is the number that results in the lowest error on the test dataset [18]. In order to anticipate wind speed and select a model that can monitor the real direction of wind speed, the autoregressive integrated moving average (ARIMA) was utilised [19]. ARIMA, ANN, and the hybrid technique (ARIMA and ANN) were the three models employed in wind speed forecasting. Using numerical error evaluation techniques like mean absolute percentage error (MAPE), mean square error (MSE), and mean absolute error (MAE), prediction accuracy results for each model were compared [20]. In recent years, ANNs and SVMs, two well-liked machine learning algorithms, have been extensively applied for wind energy prediction [21]. Despite significant progress in the development of methods for forecasting wind energy, there is still room for improvement in accuracy. One challenge of using neural networks for time series analysis is that the gradient information can disappear or explode during the training process, particularly in the case of recurrent neural networks (RNNs) used for long-term forecasting. To address this issue, long short-term memory (LSTM) neural networks have been developed as a solution to short-term memory [22]. Several ANN models have been proposed in renewable energy prediction applications, including the recurrent neural network (RNN), the convolutional neural network (CNN), the multilayer perception (MLP), and the long-term memory (LSTM) [23]. Creating a noniterative image reconstruction method by training the parameters of the Levenberg-Marquardt back-propagation algorithm [24]. The performance and output of an ANN model are largely dependent on the characteristics and functions of the hidden layers. The hidden layers are responsible for transforming the input data into a representation that the output layer can use to make predictions [26]. By using the Levenberg-Marquardt (LM) back propagation approach instead of the conventional back propagation method, the two hidden layer ANN model was trained more quickly [27]. The Levenberg-Marquardt Backpropagation (LMBP) algorithm is a popular optimization algorithm used to train multilayer artificial neural networks (MLANNs) [28]. A model for predicting air temperature has been developed based on a multi-layer ANN with the LMBP training algorithm [29]. A perceptron multilayer neural network (MLP) has been developed based on the radial basis function (RBF) to predict daily sunlight [30]. To predict short-term wind speed, various models based on the ANN algorithm were applied. The results showed that the proposed model, which contains 60 neurons, achieved a distinguished performance with a low mean square error (MSE) [31]. The recursive radial basis function neural network (RRBFNN) model, containing 44 hidden neurons, achieved a successful prediction of wind speed with minimal statistical errors [32]

3. Methodology

3.1. Data Collection

Time series forecasting is a technique for predicting the future values of a variable based on its past values. Thus, in this study, the database was used for covering the period from January 2011 to December 2016 in the United States, as a source of information to develop and test a time series forecasting model, by analyzing the historical data in the database. Figure 1 shows a time series chart of data for monthly electricity consumption generated by wind energy used in this analysis.

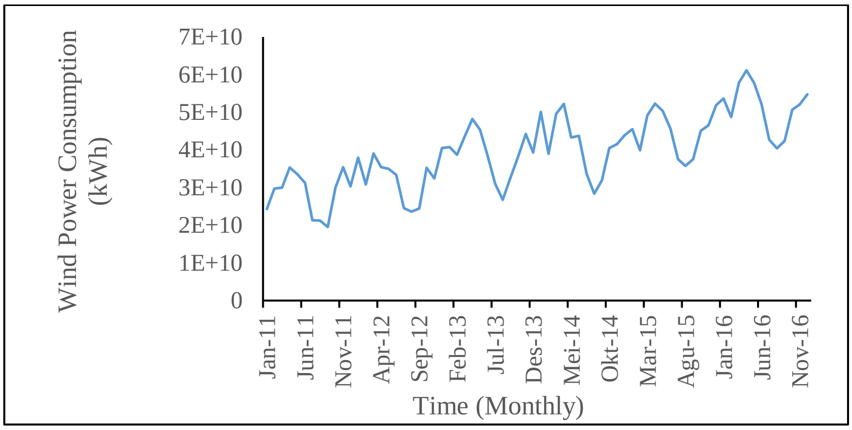

Figure 1. Actual Monthly Wind Power Consumption in the United States from January 2011 to December 2016.

Figure 2 shows the electricity generated from wind energy during the months of the year. Obviously, there are different values throughout the year, as well as different values for the same month in different years. These changes in the monthly electricity consumption generated from wind energy are due to the variation and uncertainty in the wind forecast.

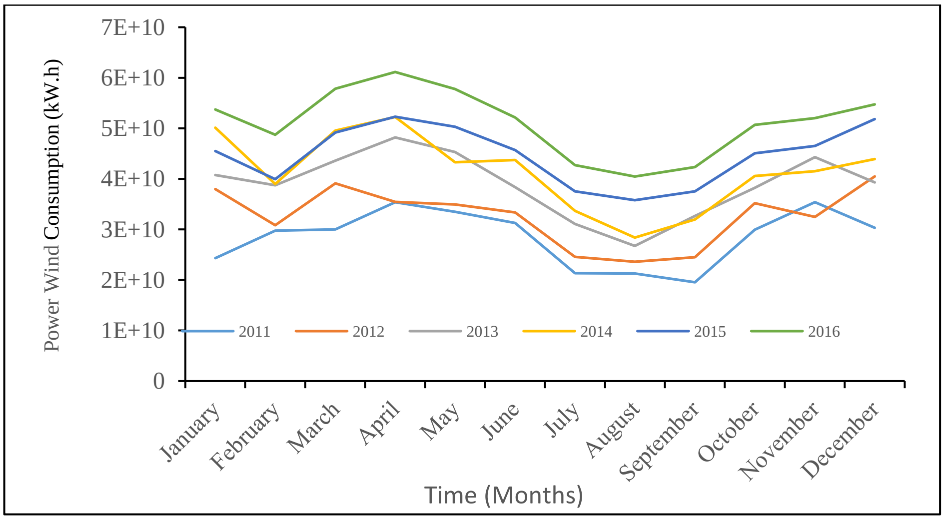

Figure 2. Actual U.S. Wind Power Consumption Values of Each Month for 6 Years.

3.2. Artificial neural networks (ANN)

Recent developments in machine learning and data analytics technology, including the use ANNs, have enabled the analysis of massive volumes of data and led to significant improvements in classification and forecasting problems. One of the key advantages of ANNs is their ability to model both linear and nonlinear systems without making implicit assumptions. This makes them well-suited for solving complex problems that traditional techniques may struggle with [33]. ANNs enable the construction of arbitrary complex models by scaling the number of model layers and utilizing various data processing components. This allows ANNs to detect complex dependencies and patterns in the data, making them well-suited for a wide range of tasks, including classification, regression, and prediction [34]. While ANNs vary in their specific architectures and applications, they share some basic and common features. One of these features is parallelism, where the neurons in an ANN perform their computations simultaneously. This allows ANNs to process large amounts of data quickly and efficiently. Moreover, ANNs are resilient to neuron failures, meaning that even if some neurons are not functioning properly, the network can still produce useful outputs. Another important feature of ANNs is their ability to generalize. During the training process, ANNs learn to recognize and extract meaningful features from the input data, which can then be used to make predictions or classifications on new, unseen data. This ability to generalize is what makes ANNs so powerful and versatile, as they can be trained on one set of data and then applied to new data that they have not seen before. [35]. ANNs are composed of simple processing units called neurons, which receive input information, process it, and produce output. Each neuron is connected to other neurons via synapses, which are essentially connections that have a weight associated with them. During the training process of an ANN, the synaptic weights are optimized through an algorithm called backpropagation. Backpropagation works by computing the error, or the difference between the desired output and the network's output, and comparing it to that output. The weights are changed based on how much each neuron contributed to the error once the fault has been transmitted backwards through the network. Iteratively repeating this procedure till the mistake is minimized [36]. Figure 3 shows an ANN model consisting of 3 input variables (x1, x2, and x3), one hidden layer containing an artificial neuron, and one output variable (y).

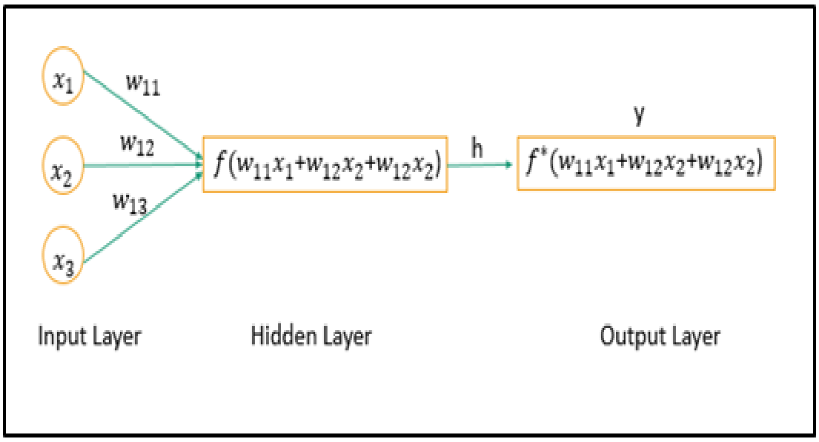

Figure 3. An example of a simple ANN model.

As can be seen in Figure 3, the neuron receives a series of inputs x_1, x_2, and x_3, which are multiplied by their respective weights w_{11}, w_{12}, and w_{13}, and then the resulting products are summed up. This sum is then passed through an activation function, which helps determine the neuron's output signal. The weights associated with each input are adjusted during the learning phase of the neural network, using techniques such as backpropagation. The neurons in the network are typically organized into layers, with the input layer receiving the initial inputs and the output layer producing the final outputs. In between these two layers, there may be one or more hidden layers that help process the inputs and produce more complex outputs [37]. The final output of the network is typically produced by combining the outputs of the neurons in the output layer, using their respective weights. This can be expressed mathematically as follows:

$$y = h * f(w_{11}x_1 + w_{12}x_2 \ + w_{13}x_3) \qquad (1)$$

In general ANN model can be as:

$$y = \sum_{M=1}^{N} h_M * f\left(\sum_{j=1}^{M}\sum_{i=1}^{N} w_{i,j}\, x_j\right) \qquad (2)$$

The activation function used in the backpropagation algorithm needs to have certain properties in order for the algorithm to work properly. One of the most important properties is differentiability, as this is necessary for calculating the gradients of the error with respect to the weights. Continuity, monotonicity, and monotonically non-decreasing properties are also desirable, as they help to ensure that the optimization process is well-behaved and converges to a good solution. It is also preferable for the activation function to have a derivative that is easy to compute, as this can make the training process more efficient. Saturation functions, which have a range that saturates at a certain value, are often used as activation functions in neural networks. The sigmoid function, with a range of 0 to 1, is a typical activation function utilised in ANN model and is defined as:

$$f(Z) = \frac{1}{1 + e^{-Z}} \qquad (3)$$

Where $z = \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_M X_M$ and $\beta_1, \beta_2, \dots, \beta_M$ are weights for the respective inputs $X_1, X_2, \dots, X_M$.

3.3. Multilayer perceptron networks

MLPs have played a foundational role in the development and popularity of neural networks, and they continue to be a fundamental model in the field of deep learning. The information flows strictly in one direction, from the input layer through the hidden layers to the output layer. There are no cycles or loops in the connections. The MLP is well-known for its ability to approximate any type of continuous, bounded, and differentiable nonlinear function defined in a compact space, making it a versatile tool for a wide range of applications [38]. One of the main limitations of the MLP is that it can only solve linearly separable problems, which means that it may struggle with more complex and nonlinear problems. However, the use of multiple hidden layers and more advanced optimization algorithms, such as backpropagation, can help to mitigate this limitation. The MLP can be either locally or fully connected. In a locally connected MLP, each output of a neuron in layer "i" is an input to all the neurons in layer "i+1", while in a fully connected MLP, each neuron in layer "i" is an input to all the neurons in layer "i+1". The choice of connectivity depends on the specific problem being solved and the available computational resources.

3.4. Network training

An ANN can be trained with a wide variety of algorithms. Because this paper simulates an algorithm for predicting electricity consumption, we can use previous periods' electricity consumption values as training samples and historical data. The training sample is understood as a data set consisting of input parameters and the corresponding known value of the predicted parameter; therefore, we consider the method of error back propagation related to the supervised learning model. In order for a neural network to be able to perform a given task, it must be trained. In order for a neural network to be able to perform a given task, it must be trained. The backpropagation algorithm is one of the methods for training multilayer feedforward neural networks. The backpropagation algorithm involves three stages: forward propagation, error calculation, and backward propagation. During forward propagation, the input data is fed into the network, and the output values are computed layer by layer until the final output is obtained. Then, the error between the predicted output and the actual output is calculated using a suitable loss function. [5]. To train a neural network with new information, need to modify the weights. The systematic method of modifying the weight is known as the "Learning Rule." The error between the expected output and the current output is calculated with a loss function (or cost function). In other words, the loss function determines how well the neural network can predict a given input. Express the process mathematically as follows:

$$wij(updated\ weight) \leftarrow wij(previous\ weight) + \alpha eixj$$

Where xj is the output from node j (where j =1, 2, 3, …..,etc), ei is the error of ‘i’ and α is the learning rate (which determines how much weights are changed each time). Itis value is between 0 and 1.

The initial weights in a neural network play a critical role in determining the performance of the network during training. If the initial weights are too large, the input signal to each neuron may saturate, causing the network to be unable to learn effectively. On the other hand, if the initial weights are too small, the input signal to the neurons may be close to zero, leading to a slow learning rate. Additionally, a multi-layer neural network can learn more complex input patterns than a single-layer network because it can capture non-linear relationships between the input and output. However, training a multi-layer neural network can be more challenging than training a single-layer network because of the increased number of parameters that need to be learned and the potential for overfitting if the network is too complex. Recently, backpropagation neural networks were used as an artificial intelligence component, and they have been successfully applied in several areas with excellent results. The training process of backpropagation neural networks can be computationally expensive and time-consuming, especially for complex tasks such as forecasting power consumption that require large amounts of data and many hidden layers. The slow convergence of the algorithm is one of the main factors that contributes to the long training times [39].

3.5. Levenberg–Marquardt algorithm

A well-liked optimisation technique for training neural networks is the Levenberg-Marquardt (LM) algorithm, especially for regression tasks where the objective function is the sum of squares of nonlinear functions. Due to its effectiveness and efficiency in training neural networks for such tasks, the LM method is frequently employed in Matlab. The algorithm is based on a modified form of Newton's method that incorporates a damping parameter to improve the convergence properties of the algorithm. The Levenberg-Marquardt algorithm uses an approximation of the Hessian matrix, which describes the curvature of the objective function, to determine the search direction during each iteration of the optimization process. Unlike traditional Newton's method, the Levenberg-Marquardt algorithm does not require the explicit computation of the Hessian matrix, which can be computationally expensive and memory-intensive for large networks. The hessian approximation can be expressed as:

$$H = J^T J \qquad (4)$$

And the gradient calculation is expressed as:

$$g = J^T e \qquad (5)$$

Where J is the Jacobian matrix, which contains the first derivatives of the errors of the network with respect to the weights and bias, and e is the vector of errors of the network. The Levenberg-Marquardt algorithm approximates the Hessian matrix by combining the Jacobian matrix with a damping parameter. The damping parameter is a scalar value that controls the size of the steps taken during the optimization process, and it can be adjusted dynamically to balance the speed of convergence with the stability of the algorithm. The Levenberg-Marquardt algorithm updates the weights and biases of the neural network using the following equation:

$$X_{K+1} = X_K - [J^T J + \mu I]^{-1 J^T e} \qquad (6)$$

where X_{K+1} is a new weight calculated using gradient function and X_K is the current weight calculted using Newton algorithm.When the value of the scalar μ is zero, the algorithm is identical to the Newton method using the Hessian matrix approximation. When the value of μ increases, the gradient descent method is used. Newton's method is faster and more accurate in minimizing error, so the goal will be to switch to Newton's method as quickly as possible. As a result, μ decrease following each successful step (decrease in the activation function) and increase between steps.

3.6. Performance Metrics

The absolute fraction of variance values (R^2) and the ideal number of hidden neurons for each training function were obtained and presented in Table 1 to help determine whether an ANN network is producing accurate predictions. The root mean square error (RMSE), mean absolute percentage error (MAPE), and absolute fraction of variance (R^2) of seven different ANN training models have been compared. The following is a list of the equations used to calculate the metrics.

3.6.1. *Mean Absolute Percentage Error (MAPE)*

MAPE is a measure of the prediction accuracy of a forecasting method in statistics. It has been used as a mean of comparison between several algorithms for electricity consumption predictions [40]. It is considered highly accurate when the value is less than 10% and a reasonable forecast when it is value 11–20% [41]. It measures this accuracy as a percentage and can be calculated as the average absolute percent error for each time period minus actual values divided by actual values. It usually expresses the accuracy as a ratio defined by the formula:

$$MAPE = \frac{100}{N}\sum_{i=1}^{N}\frac{|y_a - y_f|}{y_a} \tag{7}$$

3.6.2. *Root Mean Square Error (RMSE)*

The standard deviation of the residuals (prediction errors) is known as root mean square error (RMSE). The RMSE is a measurement of how evenly distributed these residuals are; residuals are a measure of how distant the data points are from the regression line. In other words, it provides information on how tightly the data is clustered around the line of best fit. In climatology, forecasting, and regression analysis, root mean square error is frequently used to validate experimental results. The sample size (N) and observation scale are the foundations of the RMSE, which shows the amount of the estimation error [42]. The official definition is as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_a - y_f)^2}{N}} \tag{8}$$

3.6.3. *Absolute Fraction of Variance (R Squared)*

The coefficient of determination, or R-squared, is a statistical indicator that shows how well a regression model fits the observed data. The percentage of the dependent variable's variance that the independent variable(s) in the model accounts for is expressed as a number between 0 and 1. R-squared can be used as an evaluation metric in the context of neural networks to evaluate how well the network performs in predicting a dependent variable based on one or more independent factors. For example, in the case of forecasting electricity consumption, R-squared could be used to measure how well the neural network predicts future electricity consumption based on historical data. A high R-squared value would indicate that the network is able to accurately capture the patterns and trends in the data and make accurate predictions [41].

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_a - y_f)^2}{\sum_{i=1}^{N}(y_a - \bar{y}_a)^2} \tag{9}$$

where y_a is the actual output value, y_f is the forecast output value, $\bar{y}_a$ is the average of the actual output value and N is the number of samples. The network design process plays an important role in network performance.

4. Result and Discussion

Artificial intelligence (AI) models, such as neural networks, offer a data-driven approach to modeling complex systems, including forecasting models for electricity generation and consumption in micro-grids, as well as forecasting models for solar, wind, and other renewable energy sources [43, 44]. One of the more popular contemporary techniques for forecasting is the use of neural networks. This is because neural networks can be used to explain the behaviour of non-linear data and do not require strict and exact conditions for the purpose of prediction. The performance of the neural network is significantly influenced by the number of hidden layers present. The network performs better when there are many hidden layers because of the over-fitness effect, but when there are few hidden layers, the network performs better for out-of-sample prediction. The number of suitable concealed layers would not be predetermined by any hard and fast rules; instead, it would often depend on the type of data being used. By adjusting the number of input and hidden layers, a trial-and-error approach was utilised to find the optimum ANN model with improved accuracy. With the right amount of hidden layers, inputs, and nodes, the model can predict when its error value will be at its lowest. The learning algorithm used to choose weights is called back-propagation [45]. In this study, the forecasting of the electricity consumption from wind energy in the United States was done using feedforward backpropagation artificial neural networks with the Levenberg-Marquardt (LM) algorithm. The data was split into a training group and a test group in order to confirm the model's accuracy. The neural network was created and optimised using the training group, and its correctness was verified using the test group. The neural network was trained using the cross-validation method with several structural variations, and the performance measurements were utilised to estimate the ideal number of hidden layers and neurons. As shown in Table 1, the attributes number, training percentage, testing, and the number of hidden neurons were determined by experiment based on the accuracy of performance measures RMSE, MAPE and R^2

Table 1. Selection of the number of neurons.

Neurons	**MAPE**	**RMSE**	$\mathbf{R^2}$
10	22.571	0.1053	0.962
16	14.464	0.0945	0.9917
20	4	0.0245	0.999
24	10.8	.0746	0.9903
30	4.6	.0292	0.999
40	7.3	0.0583	0.999
50	20.4	0.126	0.976

Using a single hidden layer and varying the number of neurons in that layer is a common approach in neural network analysis. By doing so, we can test the impact of different levels of complexity on the performance of the model. The number of neurons in the hidden layer was changed in this investigation to identify the neural network layout that produces the lowest statistical errors (10, 16, 20, 24, 30, 40, and 50 neurons). Cross-validation, a method for assessing the model's generalizability, was used to compare the model's performance across various structures. The behaviours of networks with various selections of neurons are shown in Table 1. The ANN model with a single hidden layer of 20 neurons appears to have produced the best results based on the table, having the lowest values for RMSE and MAPE. As a result, it is decided to be the best for the model structure. Up to 70% of the total data were utilised for training, 15% for testing, and the remaining data were used to assess how well the ANNs performed. To avoid "overfitting," which occurs when a model performs well on training data but badly on fresh, untried data, this is a typical practise in machine learning. Figure 4, shows the forecasted monthly values for 2016 of electricity consumption from wind power in the U.S. using the ANN method with different numbers of hidden neurons versus the actual consumption values. From the figure, it appears that the accuracy of the predictions varied depending on the number of hidden neurons used in the model. The study found that using 10 or 16 hidden neurons may not be sufficient to meet the requirements for capacity and error accuracy. On the other hand, using 50 hidden neurons resulted in increased error values and a trade-off in stability between the input and the hidden output connection. It is important to note that there is a trade-off between the size and number of hidden neurons. Too few neurons might result in the model being unable to capture the complexity of the data, leading to underfitting. Conversely, too many neurons can lead to overfitting, where the model essentially memorizes the training data and fails to generalize well to new, unseen data.

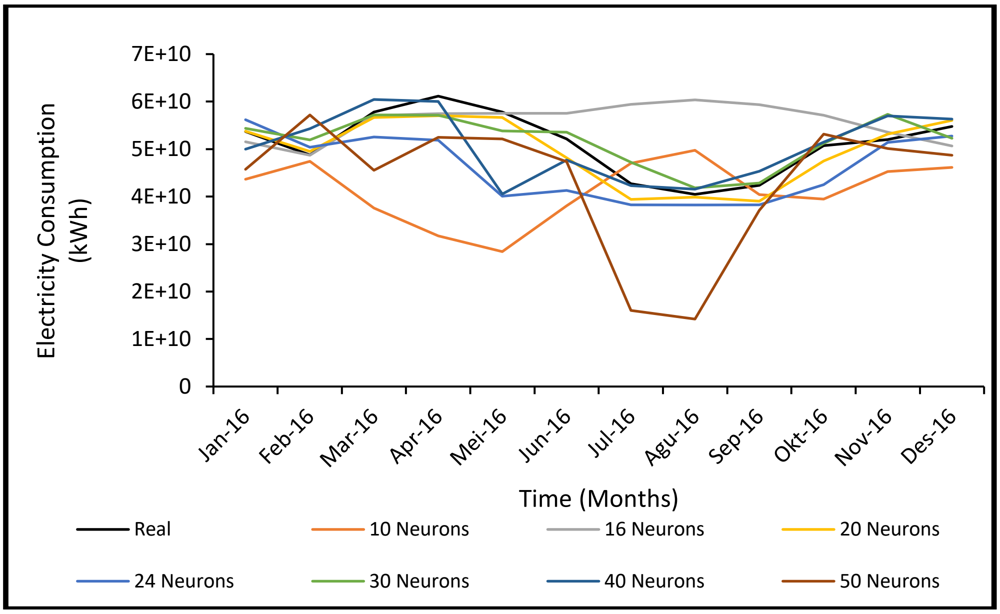

Figure 4. The Plot of 2016 Real Consumption Values vs. Forecasting from ANN with different neurons.

The proposed model was evaluated for forecasting electricity consumption from wind power in the US using projected monthly values for the year 2016. The study compared the performance of the proposed ANN model with SVM and ARIMA models. Based on Figure 5, it appears that all three models were able to track the actual trend of electricity consumption, suggesting that they are effective tools for predicting electricity consumption. However, the accuracy of the predictions for each model varied, and the proximity of the predicted values to the actual values is a measure of their reliability. Overall, these results suggest that the proposed ANN model has the potential to be a useful tool for forecasting electricity consumption from wind power in the US, and it compares favorably with other popular machine learning models such as SVM and ARIMA

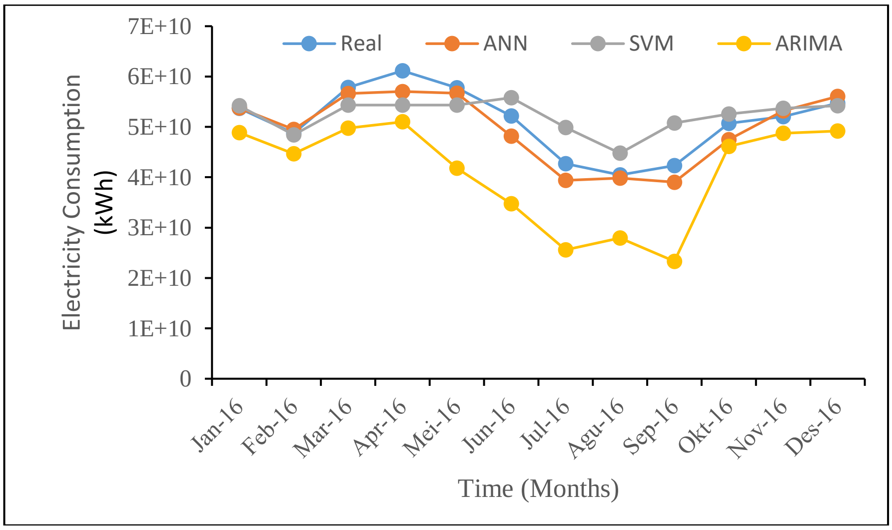

Figure 5. The Plot of 2016 Real Consumption Values vs. Predictions from ANN, SVM and ARIMA.

Table 2. Summary of test statistical errors.

Model	MAPE	RMSE
ANN	4	0.0245
SVM	7	0.0439
ARIMA	20.2	0.1169

From Table 2, it appears that the proposed ANN model outperformed the SVM and ARIMA models in terms of MAPE and RMS values. This suggests that the ANN model is better suited for predicting electricity consumption from wind power in the US, and its evaluation is consistent with the predictions generated using time series algorithms. Overall, the results presented in Table 2 support the conclusion that the proposed ANN model is a powerful tool for forecasting electricity consumption from wind power in the US and provides more accurate predictions compared to SVM and ARIMA models.

5. Conclusions

In this study, a backpropagation neural network was used as a component of artificial intelligence and specifically employed a multi-layer perception network (MLP) with a backpropagation learning algorithm to train a feed-forward neural network for predicting electricity consumption from wind power in the U.S. The Levenberg-Marquardt algorithm was used to evaluate the performance of the trained network, and various numbers of neurons in the hidden layer were selected for comparison. According to the results, the proposed ANN model containing 20 neurons in the hidden layer achieved the most accurate predictions of electricity consumption with minimal statistical errors. This suggests that ANN is a reliable and powerful tool for predicting electricity consumption from wind power in the U.S. Overall, choosing the appropriate size of an ANN structure is critical for achieving accurate predictions and optimal performance in solving specific tasks. The study's results indicate that ANN can be a valuable technique for forecasting electricity consumption in the renewable energy sector.

References

[1] Baldwin S, Bindewald G, Brown A, Chen C, Cheung K, Clark C, Cresko J, Crozat M, Daniels J, Edmonds J, Friley P. Quadrennial technology review: an assessment of energy technologies and research opportunities. Technical Report; 2015.

[2] Novas N, Alcayde A, Robalo I, Manzano-Agugliaro F, Montoya FG. Energies and Its Worldwide Research. Energies. 2020 Jan;13(24):6700.

[3] Arafat M, Sjafrizal T, Anugraha RA. An artificial neural network approach to predict energy consumption and surface roughness of a natural material. SN Applied Sciences. 2020 Jul;2(7):1-1.

[4] Pinjare SL, Kumar A. Implementation of neural network back propagation training algorithm on FPGA. International journal of computer applications. 2012 Jan 1;52(6).

[5] Zhao Z, Xin H, Ren Y, Guo X. Application and comparison of BP neural network algorithm in MATLAB. In2010 International Conference on Measuring Technology and Mechatronics Automation 2010 Mar 13 (Vol. 1, pp. 590-593). IEEE.

[6] Alsmadi, M., Omar, K.B. and Noah, S.A., 2009. Back propagation algorithm: the best algorithm among the multi-layer perceptron algorithm.

[7] Yu H, Wilamowski BM. Levenberg–marquardt training. InIntelligent systems 2018 Oct 3 (pp. 12-1). CRC Press.

[8] Li G, Shi J. On comparing three artificial neural networks for wind speed forecasting. Applied Energy. 2010 Jul 1;87(7):2313-20.

[9] Monteiro C, Bessa R, Miranda V, Botterud A, Wang J, Conzelmann G. Wind power forecasting: State-of-the-art 2009. Argonne National Lab.(ANL), Argonne, IL (United States); 2009 Nov 20.

[10] Negnevitsky M, Potter CW. Innovative short-term wind generation prediction techniques. In2006 IEEE Power Engineering Society General Meeting 2006 Jun 18 (pp. 7-pp). IEEE.

[11] Friedrich L, Afshari A. Short-term forecasting of the Abu Dhabi electricity load using multiple weather variables. Energy Procedia. 2015 Aug 1;75:3014-26.

[12] Palomares-Salas JC, De La Rosa JJ, Ramiro JG, Melgar J, Aguera A, Moreno A. ARIMA vs. Neural networks for wind speed forecasting. In2009 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications 2009 May 11 (pp. 129-133). IEEE.

[13] Yu L, Zhang X, Wang S. Assessing potentiality of support vector machine method in crude oil price forecasting. EURASIA Journal of Mathematics, Science and Technology Education. 2017 Nov 21;13(12):7893-904.

[14] Risse M. Combining wavelet decomposition with machine learning to forecast gold returns. International Journal of Forecasting. 2019 Apr 1;35(2):601-15.

[15] Sangita B P, Deshmukh SR. Use of Support Vector Machine, decision tree and Naive Bayesian techniques for wind speed classification. In2011 International Conference on Power and Energy Systems 2011 Dec 22 (pp. 1-8). IEEE.

[16] Ata, Rasit. "RETRACTED: Artificial neural networks applications in wind energy systems: A review." (2015): 534-562.

[17] Al-Roomi AR, El-Hawary ME. Optimizing Load Forecasting Configurations of Computational Neural Networks. In IEEE Conference on Electrical & Computer Engineering (CCECE), Canada13 May 2018, 1-6.

[18] Sozen A, Arcaklioglu E, Ozkaymak M. Modelling of Turkey's net energy consumption using artificial neural network. International Journal of Computer Applications in Technology. 2005 Jan1;22(2-3):130-6.

[19] Elsaraiti M, Merabet A, Al-Durra A. Time Series Analysis and Forecasting of Wind Speed Data. In2019 IEEE Industry Applications Society Annual Meeting 2019 (pp. 1-5). IEEE.

[20] Nair KR, Vanitha V, Jisma M. Forecasting of wind speed using ANN, ARIMA and Hybrid models. In2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT) 2017 Jul 6 (pp. 170-175). IEEE.

[21] Jørgensen KL, Shaker HR. Wind Power Forecasting Using Machine Learning: State of the Art, Trends and Challenges. In2020 IEEE 8th International Conference on Smart Energy Grid Engineering (SEGE) 2020 Aug 12 (pp. 44-50). IEEE.

[22] Elsaraiti M, Merabet A. Application of Long-Short-Term-Memory Recurrent Neural Networks to Forecast Wind Speed. Applied Sciences. 2021 Jan;11(5):2387.

[23] Rahman MM, Shakeri M, Tiong SK, Khatun F, Amin N, Pasupuleti J, Hasan MK. Prospective Methodologies in Hybrid Renewable Energy Systems for Energy Prediction Using Artificial Neural Networks. Sustainability. 2021 Jan;13(4):2393

[24] Mia MM, Biswas SK, Urmi MC, Siddique A. An algorithm for training multilayer perceptron (MLP) for Image reconstruction using neural network without overfitting. International Journal of Scientific & Technology Research. 2015 Feb;4(02):271-5.

[25] Alaloul WS, Qureshi AH. Data Processing Using Artificial Neural Networks. In Dynamic Data As- similation-Beating the Uncertainties 2020 May 26. IntechOpen.

[26] Feng J, Sun Q, Li Z, Sun Z, Jia K. Back-propagation neural network-based reconstruction algorithm for diffuse optical tomography. Journal of biomedical optics. 2018 Dec;24(5):051407.

[27] Ranjbar M, Soleymani S, Sadati N, Ranjbar AM. Electricity price forecasting using artificial neural network. In2006 International Conference on Power Electronic, Drives and Energy Systems 2006 Dec 12 (pp. 1-5). IEEE.
[28] Lv C, Xing Y, Zhang J, Na X, Li Y, Liu T, Cao D, Wang FY. Levenberg–Marquardt backpropagation training of multilayer neural networks for state estimation of a safety-critical cyber-physical system. IEEE Transactions on Industrial Informatics. 2017 Nov 24;14(8):3436-46
[29] El Houari MB, ZEGAOUI O, ABDALLAOUI A. Prediction of air temperature using multi-layer perceptrons with Levenberg-Marquardt training algorithm. Parameters. 2015;1(11).
[30] Behrang MA, Assareh E, Ghanbarzadeh A, Noghrehabadi AR. The potential of different artificial neural network (ANN) techniques in daily global solar radiation modeling based on meteorological data. Solar Energy. 2010 Aug 1;84(8):1468-80.
[31] Selcuk Nogay H, Akinci TC, Eidukeviciute M. Application of artificial neural networks for short term wind speed forecasting in Mardin, Turkey. Journal of Energy in Southern Africa. 2012;23(4):2-7.
[32] Madhiarasan M. Accurate prediction of different forecast horizons wind speed using a recursive radial basis function neural network. Protection and Control of Modern Power Systems. 2020 Dec;5(1):1- 9.
[33] Demuth, H. Beale, M, Hagan, M. Neural Network Toolbox 5: Users Guide. Natick, MA, The Math-Works Inc, 2007.
[34] Damborg MJ, El-Sharkawi MA, Aggoune ME, Marks RJ. Potential of artificial neural networks in power system operation. InIEEE International Symposium on Circuits and Systems 1990 May 1 (pp. 2933-2937). IEEE.
[35] Fernández, I., Borges, C.E, Penya, Y.K. September. Efficient building load forecasting. In ETFA2011, IEEE 2011, 1-8.
[36] Mehlig BE. Machine learning with neural networks. arXiv preprint arXiv:1901.05639. 2019 Jan 17.
[37] Sevlian. R, Rajagopal, R. A scaling law for short term load forecasting on varying levels of aggregation. International Journal of Electrical Power & Energy Systems 2018, 98:350-361.
[38] Isabona J, Imoize AL, Ojo S, Karunwi O, Kim Y, Lee CC, Li CT. Development of a Multilayer Perception Neural Network for Optimal Predictive Modeling in Urban Microcellular Radio Environments. Applied Sciences. 2022 Jan;12(11):5713.
[39] Wassermann, F. "Neurocomputer Techniques: Theory and Practice." (1992).
[40] Kasemset, C., Sae-Haew, N. and Sopadang. Multiple regression model for forecasting quantity of supply of off-season longan. Chiang Mai University Journal of Natural Sciences (Thailand), 2014.
[41] Fan, C., Xiao, F, Zhao, Y. A short-term building cooling load prediction method using deep learning algorithms. Applied energy 2017, 195: 222-233.
[42] Vakili, M., Sabbagh-Yazdi, S.R., Kalhor, K, et al. Using artificial neural networks for prediction of global solar radiation in Tehran considering particulate matter air pollution. Energy Procedia 2015, 74: 1205-1212.

[43] Aslam S, Herodotou H, Ayub N, Mohsin SM. Deep learning based techniques to enhance the performance of microgrids: a review. In2019 International Conference on Frontiers of Information Technology (FIT) 2019 Dec 16 (pp. 116-1165). IEEE.
[44] Mehrzadi M, Terriche Y, Su CL, Xie P, Bazmohammadi N, Costa MN, Liao CH, Vasquez JC, Guerrero JM. A deep learning method for short-term dynamic positioning load forecasting in maritime microgrids. Applied Sciences. 2020 Jul 16;10(14):4889.
[45] Shahzad MN, Kanwal S, Hussanan A. A New Hybrid ARAR and Neural Network Model for Multi-Step Ahead Wind Speed Forecasting in Three Regions of Pakistan. IEEE Access. 2020 Oct 30;8:199382-92.