



# Comparison of Three Different Learning Methods of Multilayer Perceptron Neural Network for Wind Speed Forecasting

Mehmet BULUT<sup>1,\*</sup> , Hakan TORA<sup>2</sup> , Magdi BUAISHA<sup>3</sup> 

<sup>1</sup>Electricity Generation Company, Ankara, Turkey

<sup>2</sup>Atılım University, Department of Avionics, Ankara, Turkey

<sup>3</sup>University of Benghazi, Environment Health Department, Benghazi, Libya

## Highlights

- Study focuses on wind speed forecasting with measured meteorological data.
- Approach is proposed for comparison of three different learning methods in the study.
- Highly precise and more accurate results were obtained.

## Article Info

Received: 05 July 2020

Accepted: 08 Nov 2020

## Keywords

Wind speed forecasting  
Artificial neural network  
Renewable energy  
Energy resources

## Abstract

In the world, electric power is the highest need for high prosperity and comfortable living standards. The security of energy supply is an essential concept in national energy management. Therefore, ensuring the security of electricity supply requires accurate estimates of electricity demand. The share of electricity generation from renewables is significantly growing in the world. This kind of energy types are dependent on weather conditions as the wind and solar energies. There are two vital requirements to locate and measure specific systems to utilize wind power: modelling and forecasting of the wind velocity. To this end, using only 4 years of measured meteorological data, the present research attempts to estimate the related speed of wind within the Libyan Mediterranean coast with the help of ANN (artificial neural networking) with three different learning algorithms, which are Levenberg-Marquardt, Bayesian Regularization and Scaled Conjugate Gradient. Conclusions reached in this study show that wind speed can be estimated within acceptable limits using a limited set of meteorological data. In the results obtained, it was seen that the SCG algorithm gave better results in tests in this study with less data.

## 1. INTRODUCTION

Electric power is necessary for high prosperity and comfortable living standards in the world. Since electricity is an indispensable need in daily life, one of the goals of all countries is to provide consumers with continuous, reliable, and high-quality electrical energy. That is only possible through the proper planning electricity generation and demand side management. Effective planning and management of the electricity supply system require effective modelling and forecasting capability. Forecasting for the future is not an easy process and requires very intricate details to take into account. One of the essential forecasting processes in the energy sector is electricity demand forecasts. Electricity demand forecasts can be made in the short-term and long term. Short-term estimates are usually one hour to one week. Due to the unique characteristics of the electricity sector, it may be challenging to develop estimation models that are suitable for the different subsystems it carries. So far, many estimation methods have developed to estimate electricity demand.

A commonly used renewable energy source is wind power plants that generate electricity from the wind. It is of great importance to make more realistic production planning for these wind farms and to prepare wind speed estimates for the preparation of maintenance and control timelines. However, due to the irregular nature of the wind speed at any point and its discrete nature, it is not easy to estimate a high precision wind speed. Considering the many benefits of wind power – namely, in social as well as economic and ecological

terms – it has turned into a rapidly growing source of re-usable tools to produce power. Across the globe, there are more cases of wind-use as other sources like oil become quickly depleted, ecologies polluted and climates warmer. The countless merits of wind include being clean, economical, and highly available; plus, there is no need to move the power source or any advanced means to make use of it. To forecast wind energy output, we use a variety of models for data as a mixture of meteorology and background or past generation. Yet, to obtain the best prediction, these models need to incorporate certain other factors for upcoming patterns. The spatial change of wind energy also occurs in a large amount. The topography and the roughness of the soil significantly affect the distribution and speed of the winds. In order to estimate the wind speed in the region where the plant is located in a short time range, that is, for the next 1 hour to a 24 hours, reliable and high accuracy estimation tools need to be developed and used to predict the electrical energy to be taken from the power plant.

Numerous studies have tried to come up with appropriate models to forecast wind patterns, some of which are related to physical, statistical, hybrid physical-statistical, and AI or artificial intelligence, and other approaches [1]. In case of the physical ones like Numerical Weather Forecast (NWF) and mesoscale models [2], they often prove to be useful as they merge many such factors in the form of autoregressive integrated moving average models to determine the connection in the estimated wind speed time series [3]. Other works address both measurement and forecast of the energy generated by turbines. As far as the literature goes, AI methods of neural networking, fuzzy logic, and others have provided promising results as opposed to conventional statistical approaches. In most cases, as stated before, physical and statistical data are combined alongside NWF estimates as input variables and also historical data to train the system as per the regional status and by statistical theories.

There are also some latest AI-based approaches, such as multilayer perceptrons ANN [4] radial basis function [5], repetitive neural networks [6], and fuzzy logic [7]. The essential estimation of local speed is possible using time series analysis and statistics [8]. Oztopal et al. [9] applied neural networks to estimate velocity and interpolation. Kalogirou et al. [10] suggested that it is necessary that the sector for forecast different fluctuations in terms of weather factors like the velocity of wind, average moisture, radiation from the sun, overall temperatures, and others to accurately set the aim, analyze operations, and estimate the expenditure associated with renewable energy. Sahin and Sen [11] developed a system to determine wind velocity in the west of Turkey based on typical distribution graphics. Sirdas [12] proposes a time series for the sea Marmara, again in Turkey, to predict wind with Fourier analysis. There is a grouping based on topography in Sen [13], which ends up with various clusters. Others, like Matheron [14], introduce useful modeling of in an ore reserve study. To make local estimates, semi-variogram, cumulative semi-variogram, and point cumulative semi-variogram (PCSV) are among the other techniques suggested by Krige [15]. Sen and Sahin [16], studies the Turkish Marmara region for wind velocity using the PCSV approach to determine the spatial dependence function.

Wherever there is a shortage of data, experts might have to arrange inventories of wind power accessibility. Under such circumstances, the ANN method can be the right solution. Apart from this, one needs to know about stability in terms of statistical features, presence of wind, diurnal changes, and speed estimates. The power distribution figures help in terms of choosing the right spot, forecasting, and designing turbines. In addition, prior assessments are to be done for wind velocity in case of making inventories, thus necessitating specific models to be in place for regional correlations so that both the resources can be managed well and additional studies made in related fields of renewable energy [17].

Libya is a North African country. As an alternative to hydrocarbon-based sources, the country has convenient wind speeds that can be effectively utilized for electricity production. Libya is exposed to atmospheric winds in winter and north-eastern winds in summer and is also affected by Ghibli winds that bring hot, dry, usually south to south-easterly dust-bearing desert [18]. Wind speed measurement data made in Libya showed that there is a high potential of wind energy that can be evaluated for electricity production in Libya. The study states that the average wind speed at altitudes of 40 meters is ranging from 6 to 7.5 m/s in Libya [19]. Tripoli, which is the city with the sea port of the country, is rich in wind energy and it can be used for other applications such as electrical energy, sea water desalination etc. Sources with current measurements show that the annual wind speed in Tripoli, above 6 m / s, is sufficient to produce 2303 MWh

in this region [20]. Libya has so far not been able to take full advantage of wind energy potential as renewable energy compared to traditional energy resources. Hifa Salah Adeen Embirsh et al., in the objective of their study, evaluates the wind energy in the capital city of Tripoli, Libya. They presented the assessment of wind power. Libya's first wind power farm is to be set up in Msallata city very soon, and it has components of 16 wind turbines. This wind power plant will be installed in Shafeen Park, located 130 km east of Tripoli, and the plant will have a power of 27 MW, corresponding to 70% of the city's electricity needs [21]. In Libya, it foresees that 10% of Libya's energy needs will be provided by renewable energy by 2020, and wind energy will generate 1000 MW by 2020. This is expected to result from an increase in electricity demand in Libya's commercial and industrial sectors. Libya has a chance to meet the growing demand for electricity from renewable energy sources and has the potential to utilize renewable energy [22].

Today, artificial neural networks have begun to be applied successfully in all areas of life. Instead of normal learning, these artificial neural networks now use algorithms that we call deep learning. Deep learning has solved many complex problems that have existed in the field of artificial intelligence (AI) for many years. In fact, Deep Learning models are deeper variants of artificial neural networks (ANNs) with multiple layers, whether linear or nonlinear. Each layer is attached to its lower and upper layers with different weights. The ability of Deep Learning models to learn hierarchical properties from various data types such as numerical, image, text, and sound makes them powerful in solving problems of recognition, regression, semi-supervised or unsupervised problems [23]. For the smallness of the training data set used in this study, it is not necessary to use it if it has a deep learning algorithm. Recent studies suggest the development of combined ANN solutions that include more than one algorithm to solve the problem instead of a single algorithm [24,25,26].

In this paper, using only 4 years of measured meteorological data, the speed of wind within the Libyan Mediterranean coast is predicted by using artificial neural networks (ANN) to reveal its effectiveness in an accurate estimation of the stated element. The results of artificial neural networks trained with different learning algorithms were compared to estimate wind speed. Levenberg-Marquardt (LM), Bayesian Regularization (BR), and Scaled Conjugate Gradient (SCG) methods were chosen for the training method of neural network forecasting system. For each algorithm, the wind speed is separately estimated for the city of Tripoli using four years of limited meteorological data for training and testing.

## **2. WIND SPEED ESTIMATION FOR PRODUCING ENERGY**

Wind energy is the movement energy of wind due to the air flow. The source of wind energy is basically the sun. It is assumed that 1-2% of solar energy turns into wind energy. Wind turbines can only start generating electricity at certain wind speeds. Advanced wind turbines have cut-in speeds of 2-4 m/s, nominal speeds of 10-15 m/s, and cut-out speeds of 25-35 m/s. It is significant for the electricity market that wind power plants can predict the electricity power they will produce with high accuracy from the day ahead. But the wind's stochastic, unstable, and discontinuous nature and having a nonlinear relationship between wind power and wind speed increases the error rate in predictions. The wind speed estimation methods with high accuracy are a useful tool to minimize problems resulting from variability and uncontrollability of wind speed. Meteorological variables required as input for wind speed estimation are basically temperature, pressure, and humidity. Generally, historical data is used to fine-tune estimation models. In this way, power generation for similar weather conditions can be included in the forecast. One prediction calculation model gives the best results at high temperatures, while another gives the best results in a given region when the winds are strong. Just as for a weather forecast, many models can be combined in a power prediction for a more accurate prediction.

This study employs the ANN method to forecast the velocity of wind-based on atmospheric pressure, temperature, and relative humidity in Tripoli City on the Mediterranean coast of Libya. The wind formed anywhere is heavily dependent on local topography and other factors, and the instantaneous wind speed and direction varies more than hourly averages. During the year, significant seasonal changes occur at the average wind speed in Tripoli, which is measured at 10 m from the ground surface. Wind speed estimation

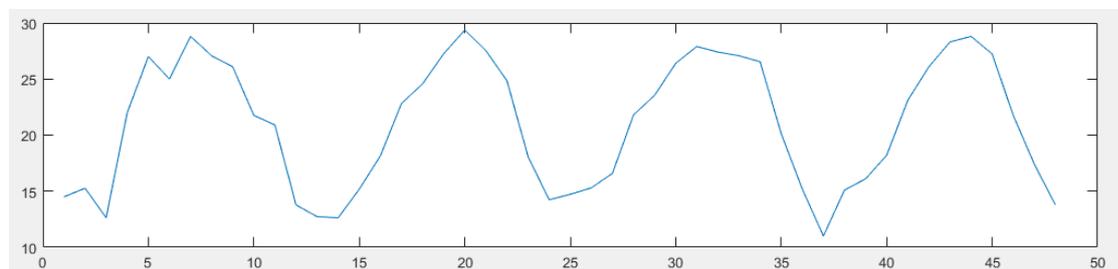
is carried out in short, medium, and long terms. Our study relies on monthly averaged wind speed observations.

## 2.1. Materials and Methods

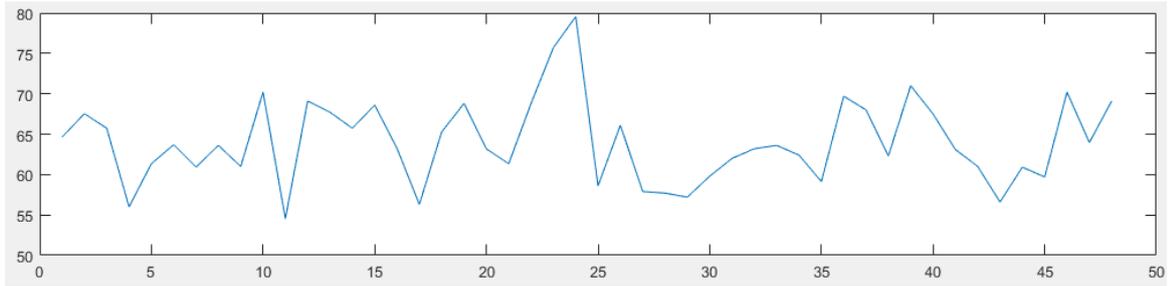
The ANN methods are arranged upon modelling the learning patterns of humans, with neurons at their core and positioned within repeated layers as an input layer, a hidden layer, and an output layer. Data arrives at the input layer, through the hidden one, and on to the output. This mechanism helps in modelling the patterns of structures with noisy and incomplete data which can be otherwise too complicated to set up algorithms, as well as to select a certain structure from within the available data. Learning through examples characterizes the neural network processes as they are not programmable for specific jobs. Practically, the system comprises statistical information that is non-linear with algorithms, among which backpropagation (BP) is the most common and applied to keep errors at the minimum between the output and expected values upon adjusting the neuron weights.

Here, as stated earlier, we use the ANN approach to determine wind velocity; to do so, we apply a feed-forward neural network with backpropagation algorithm with the help of the Matlab [27]. We research performance of learning methods for with 4 years limited data; used learning methods are Levenberg-Marquardt, Bayesian regularization, and Scaled conjugate gradient. Regarding the computational procedure, we adopted a computational-intelligence model using a Levenberg-Marquardt, Bayesian Regularization, and Scaled Conjugate Gradient Backpropagation algorithms and a training algorithm for short-term wind speed forecasting. As ANN is a model that can tolerate faulty or missing data depending on its generalization ability, it is a very suitable method for wind energy estimation. However, the better the data set used for learning, the more successful the network results. For this reason, the data set to be used first should be examined and arranged accordingly. Developing an ANN model needs two-parameter as an input parameter data set and an output parameter data set. These sets are subdivided into two different parts: training of the network and verification of generated estimates.

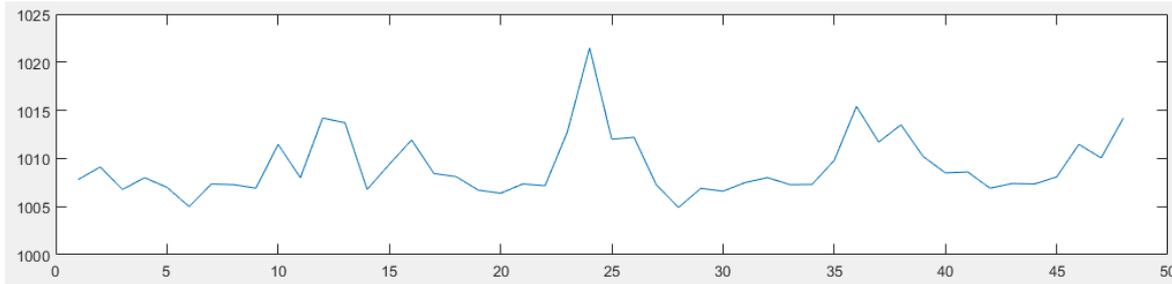
Input parameters are chosen as the monthly average relative humidity values, monthly average atmospheric pressure values, and monthly average atmospheric temperature values for the neural network. The output parameters are the monthly average wind speed values at the location. For this purpose, we use the 4-year monthly basis meteorological data provided by the Libyan State Meteorological Service as to the January 2014 - December 2017 period. The monthly patterns of weather temperature, weather humidity, weather pressure are depicted in Figure1, respectively. Figure 2 shows the measured wind speed for 48 months. Additionally, the average wind speed per month over 4 years is calculated and is shown in Figure 3. The proposed different neural network models use previously enrolled monthly average temperature, humidity, and pressure values together to forecast the future value of the wind speed.



(a) Monthly average atmospheric temperature (C°)

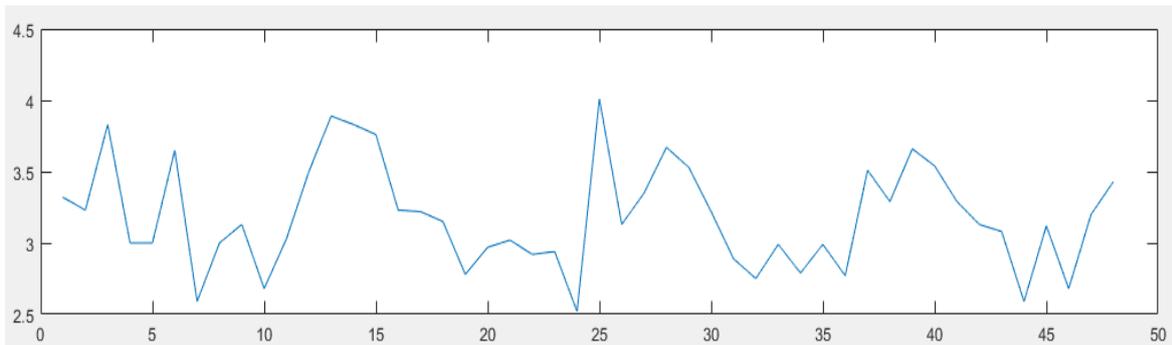


(b) Monthly average relative humidity (%)

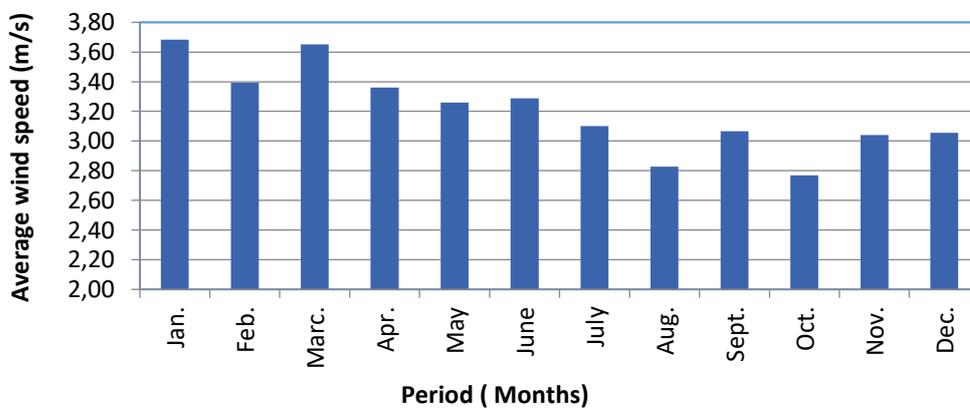


(c) Monthly average atmospheric pressure (m-bar)

**Figure 1.** Used data set for between 2014-2017 years (a) Monthly average atmospheric temperature (b) Monthly average relative humidity (c) Monthly average atmospheric pressure



**Figure 2.** Monthly average wind speed (m/s) for 4 years (2014-2017)



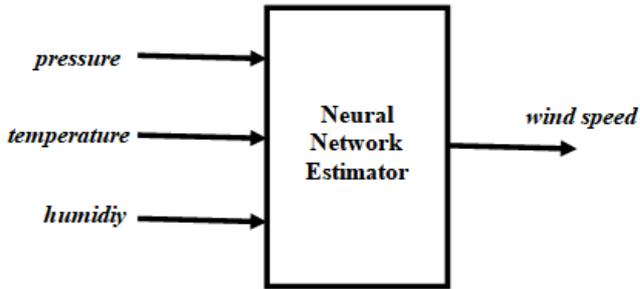
**Figure 3.** Tendency monthly average wind speed (m/s) using 4 years data

## 2.2. ANN Design for Wind Speed Prediction

Figure 4 shows the schematic diagram for the proposed network structure for this study. The NN is fed by three inputs: atmospheric pressure, temperature, and relative humidity. MLP network is trained to predict wind speed. Our neural network approximates a function defined as

$$f_k = f(p_i, t_i, h_i) \quad (1)$$

where  $f_k$  is the  $k$ th predicted wind speed,  $p_i$  is the  $i$ th measurement of pressure, and  $t_i$  is the  $i$ th measurement of temperature, and  $h_i$  is the  $i$ th measurement of humidity.



**Figure 4.** The proposed schematic diagram for the network structure

The proposed artificial neural network (ANN), which use local meteorological measurements together, is trained by three different learning algorithms. This study investigates the best learning method based on the performance of each algorithm.

The activation function used for the hidden layer is the tangent sigmoid function and the pure linear function is employed in the output layer. The sigmoid function used in the proposed neural network is defined as below.

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (2)$$

In training process, for each input-output data pair, the error is calculated between the required output and output of neural network using the below formulas,

$$out(j) = \sigma\left(\left(\sum_{i=1}^m w_{i,j}x_i\right) + b_j\right) \quad (3)$$

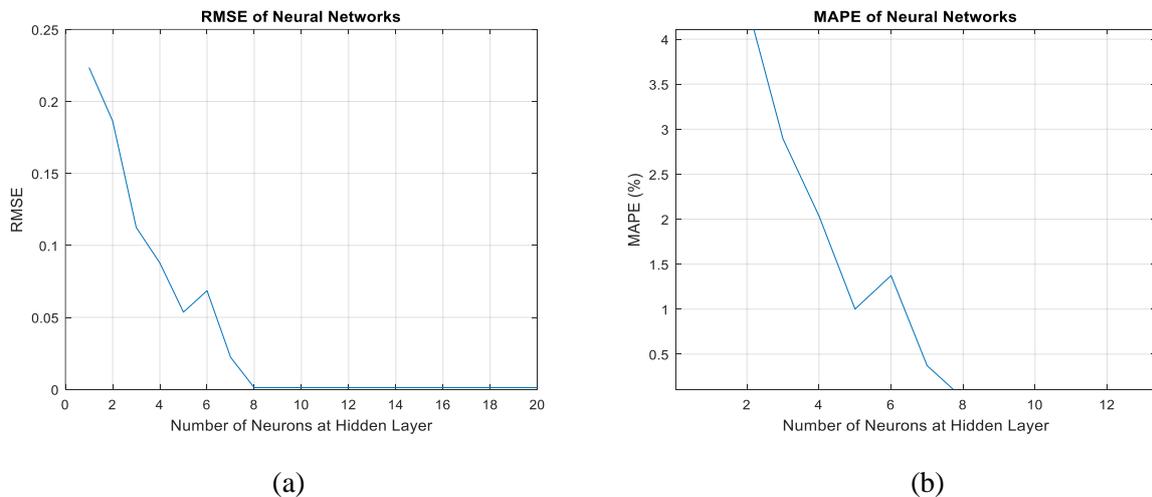
$$out(k) = \sum_{j=1}^n w_{j,k}out(j) + b_k \quad (4)$$

where,  $n$  : number of neurons,  
 $out(j)$  : output of neuron  $j$  in the hidden layer,  
 $out(k)$  : output of neuron  $k$  in the output layer,  
 $\sigma(x)$  : activation function,  
 $m$  : number of input of network,  
 $w_{i,j}$  : weight from input layer node  $i$  to hidden layer node  $j$ ,  
 $w_{j,k}$  : weight from hidden layer node  $j$  to output layer node  $k$ ,  
 $b_j$  : biases of the node  $j$  of hidden layer,  
 $b_k$  : biases of the node  $k$  of output layer.

To accomplish the estimation of wind speed, LM, BR and SCG training algorithms have been separately utilized to train the same ANN structure. In these algorithms, it is aimed to minimize the error by updating the weights and biases of artificial neural networks at every iteration. The proposed model is introduced as

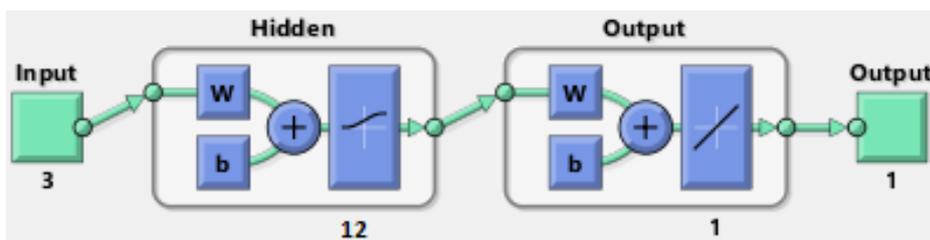
a one-hidden layer network to estimate the velocity of wind in the stated location. The number of the neurons in the hidden layer is selected based on the trial and error approach. In this approach, we fully trained the NN with the neuron numbers starting from 1 through 20. Then, we obtained the corresponding performance of the network for each neuron number. Figure 5 illustrates the RMSE and MAPE of the trained network against the number of the neurons. As clearly seen, performance remains constant after 9 neurons. Therefore, we set the number of neurons to 12 for the hidden layer.

The data has been taken in monthly basis for 4 years. In other words, the total size of the data is 48 measurements. The data studied here are divided into two parts as training data and test data. The first 3-year data is employed for training and the last 1-year is for testing. Prediction performance is evaluated with real data sets using one year of wind speed data. During ANN training, 3 years of data were divided into three groups: training (90%), validation (5%) and testing (5%), and the selection of the data was done randomly.



**Figure 5.** NN performances versus the number of the neurons (a)RMSE values of Networks (b)MAPE values of Networks

The used NN structure throughout our implementations is shown in Figure 6. It is with 3-input, 1-hidden layer with 12 neurons, and 1-output layer.



**Figure 6.** NN structure for selected neurons number at layers

In order to see the effect of the LM, SCG, and BR algorithmic techniques on the forecasting of the wind speed, the network is trained with each method and the performance metrics are computed. These techniques are discussed in detail in the following subsections.

**Levenberg-Marquardt (LM) Method:** Levenberg Marquardt is basically the least squares calculation method based on maximum neighbourhood. Algorithms that require second order derivatives such as Levenberg-Marquardt (LM) significantly increase learning speed. Combining the Newtonian algorithm speed and the stability of the steepest descent method, the LM algorithm is used effectively in network training today [28]. Today, it has become a standard technique for nonlinear least-squares problems that are widely approved for a wide range of applications.

Generally, the LM algorithm is used for solving nonlinear least-squares problems. Using the LM algorithm, the objective function of the least squares problem is formulated as follows,

$$F(\rho) = \frac{1}{2} \sum_{i=1}^n r_i^2 \quad (5)$$

$$r_i(x) = f(x_i; \rho) - y_i \quad (6)$$

where,

$n$  : total data number,  
 $r_i$  : residual number,  
 $y_i$  : i.th data component of  $y$ ,  
 $x_i$  : i.th data component of  $x$ ,  
 $\rho = [\rho_1 \ \rho_2 \ \dots \ \rho_m]$  : model parameters vector,  
 $m$  : total parameter number

Bayesian regularization (BG) Method: As a second method, Bayesian regularization (BG) was used. This learning method minimizes a linear combination of square errors and weights. It also changes the linear combination so that the resulting network has good generalization characteristics. Bayesian regularization minimizes the combination of quadratic errors and weights and determines the correct combination to produce the weight parameters of the neural network [29].

In the Bayesian regularization method, matrix weight values of neural network are accepted as random variables. When considering the data, the probability density function of the network weights  $w$  sequence is defined as follows:

$$f(w|D, \alpha, \beta, M) = \frac{f(D|w, \beta, M)f(w|\alpha, M)}{f(D|\alpha, \beta, M)} \quad (7)$$

Where;

$M$  : Particular neural network model,  
 $f(w|\alpha, M)$  : Prior density (represents knowledge of the weights before any data)  
 $f(D|w, \beta, M)$  : Likelihood function (the probability of the data),  
 $f(D|\alpha, \beta, M)$  : A normalization factor (guarantees that the total probability is 1).

Scaled conjugated gradient (SCG) Method: SCG is a supervised learning algorithm for feed forward neural networks and is a member of the conjugate gradient methods class. Recently, this algorithms have been used as training algorithms in ANN training. The number of epochs is not relevant when comparing SCG with other algorithms such as known standard backpropagation. In this method, it is necessary to calculate two repetitions of a repetition in SCG and call an error function. In standard backpropagation, an iteration requires a gradient and an error function to be calculated. Because the problem of minimizing the global error function is very common in other science field, generally in the numerical analysis [30].

Conjugated gradient methods are iterative methods for training neural networks because of their simple structure, more efficiency and need low memory requirements. Conjugated gradient method produces a weight sequence using the iterative formula as follow,

$$w_{k+1} = w_k + \eta_k d_k, \quad k = 0, 1, \dots, \quad (8)$$

where  $k$  is the current iteration,  $w_0 \in R^n$  is a given initial point,  $\eta_k > 0$  is the learning rate, and  $d_k$  is a descent search direction defined by

$$d_k = \begin{cases} -g_0, & \text{if } k = 0, \\ -g_0 + \beta_k d_{k-1}, & \text{otherwise,} \end{cases} \quad (9)$$

where,  $g_k$  is the gradient of error  $E$ ,  $w_k$  is weight parameters and  $\beta_k$  is a scalar. Studies in the literature suggest various options for  $\beta_k$ , which lead to different conjugate gradient methods.

### 2.3. Performance Comparison of Algorithms for Training of ANN

In the center of the study, it is aimed to accurately predict wind speed by using ANN based local meteorological measurements trained with different training algorithms. The ANN models proposed in the study were used to predict the future wind speed value using past meteorological values such as humidity, pressure and temperature. The meteorological data set used consists of a total of 4 years of data. Three years of this data were used to train neural network models, and last year was used to measure the wind speed estimation performance of ANN models. In training set, prediction values are almost close to actual ones for LM as seen in Figure 7. On the other hand, the deviations from the actual values become more visible in both BR and SCG.

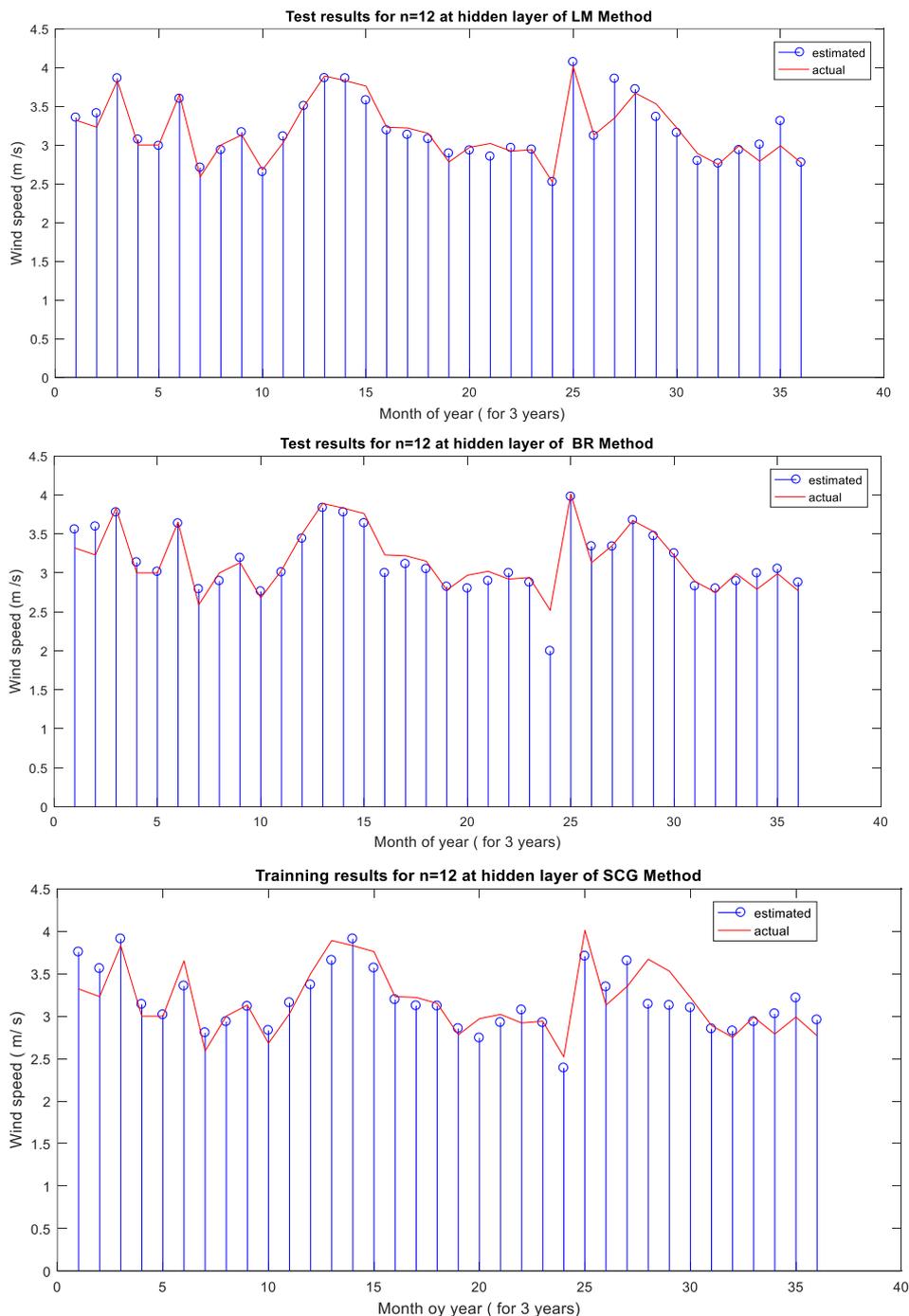
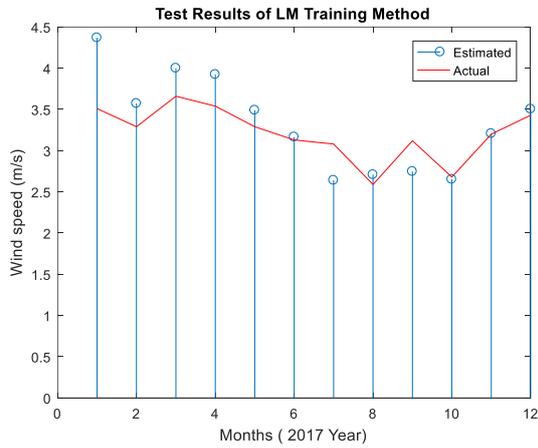
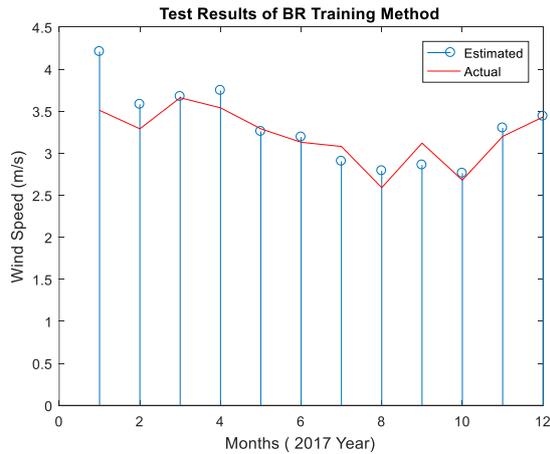


Figure 7. Results of LM, BR, and SCG for training set of the ANN

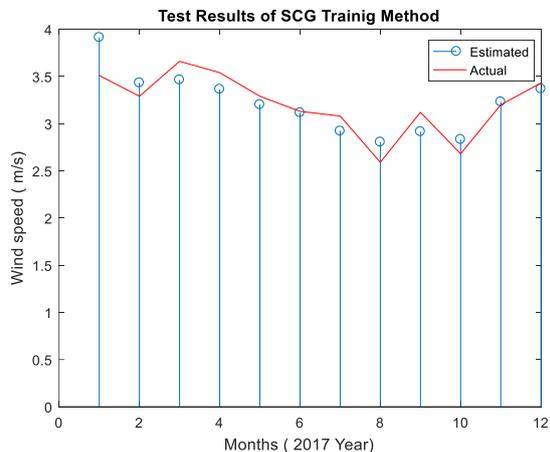
The wind speed values are predicted by ANN using multiple local meteorological data in order to investigate the effects of three different NN training algorithms. Figures 8, 9, and 10 depict the estimated results for Levenberg-Marquardt, Bayesian Regularization, and Scaled conjugate gradient algorithms, respectively, in test set. As can be seen clearly, predicted values follows the envelope of the true ones. Most of the estimated values are near to the measured values. However, the deviation from the actual value in January is quite noticeable for the three approaches.



**Figure 8.** Measured and predicted speed values in 2017 using Levenberg-Marquardt



**Figure 9.** Measured and predicted speed values in 2017 using Bayesian Regularization



**Figure 10.** Measured and predicted speed values in 2017 using Scaled Conjugate Gradient

Figure 11 illustrates the estimated values resulted from the three algorithms together. It can be observed that all predicted values in June are nearly close to the actual value. All of the training algorithms show the lowest estimation accuracy in January. In spite of that similar observation also occurs in April, the estimated values are not much away from the true value. Although the data set used for training is limited, only 3-year measurements, the predicted wind speed values for months in 2017 have a high degree of estimation.

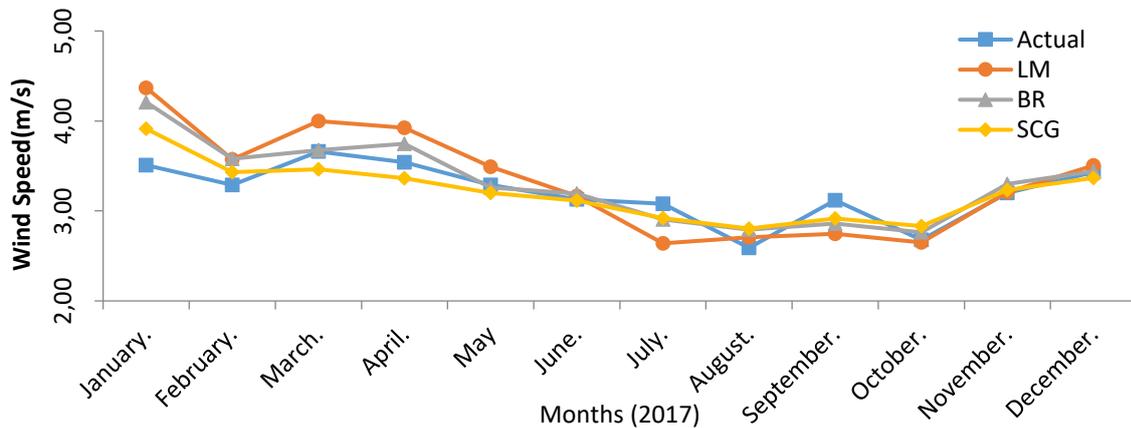


Figure 11. Comparison of the results obtained from LM, BR and SCG methods for 2017 test data set

Moreover, Figure 12 shows the regression scores (R) for each method in training phase. These scores are viewed as a measure of the generalization ability of the NN model. The score near to 1 implies that the NN estimates each example in the training set accurately and it loses its generalization, i.e., the estimation error of the NN becomes larger, for the test set. This is verified empirically in this study. The SCG method which has the lowest score leads to more accurate estimates than the LM and BR, which discussed below.

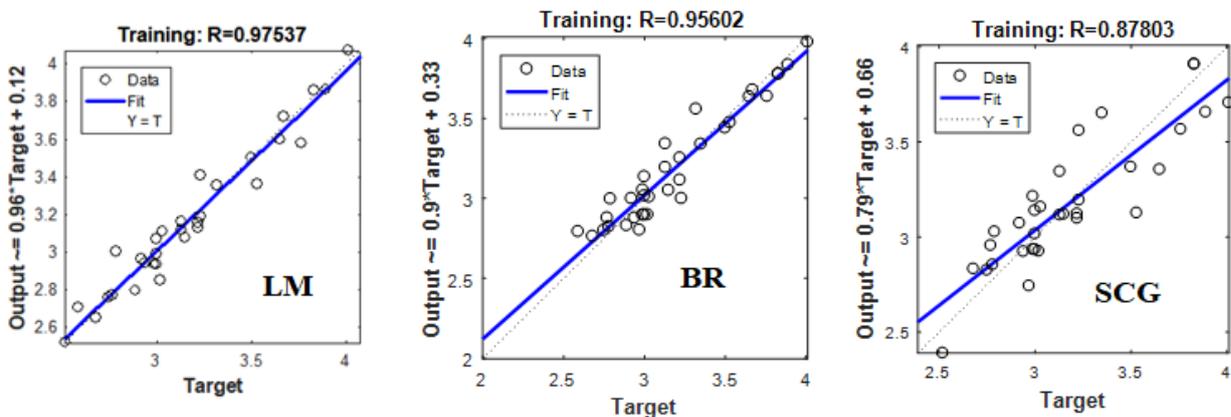


Figure 12. Regression values for LM, BR, and SCG in training phase

Apart from all the discussions above, we employed the well known evaluation metrics in the literature, which are Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) to measure the performance of the NN models. RMSE is a second-order scoring rule that also measures the average magnitude of the error and is defined as follows.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (v_{estimated} - v_{actual})^2}. \tag{10}$$

The mean absolute percentage error (MAPE) is a commonly used statistical method that measures how close the prediction result made by a prediction system is to the truth. It measures accuracy as a percentage.

The MAPE method has a high sensitivity to the scale size of system and is not recommended for use with low volume data. The fact that the MAPE values are at the desired level with a low volume data set shows the robustness of the estimation system.

Furthermore, when the actual value is not zero, but quite small, the MAPE will often take on extreme values. It is defined as follows.

$$MAPE = \left( \frac{1}{N} \sum \frac{|v_{actual} - v_{estimated}|}{|v_{actual}|} \right) * 100. \quad (11)$$

In present study, we used the MAPE and the RMSE values to choose the training algorithm leading to the smallest estimation error. It can be seen from Table 1 that SCG algorithm has obvious preference in the view of prediction accuracy with MAPE of 4.8164% as annually, followed by LM and BR algorithms with MAPE of 7.8860% and 5.5226% accordingly. Therefore, SCG method is suggested to build a wind speed forecasting model within the scope of this study. In addition, the SCG results in RMSE of 0.0444, which verifies its MAPE.

When the results obtained from this study are evaluated, it is seen that the SCG algorithm produces more minimal and consistent RMSE than LM and BR. As shown in Figure 13, the LM algorithm gives maximum RMSE for this study and shows poor performance as it fluctuates widely.

**Table 1.** Accuracies for predicted wind speed

Month (2017)	Levenberg-Marquardt		Bayesian Regularization		Scaled Conjugate Gradient	
	RMSE	MAPE (%)	RMSE	MAPE (%)	RMSE	MAPE (%)
1-January	0.2480	2.0393	0.2018	1.6593	0.1164	0.9571
2-February	0.0815	0.7151	0.0841	0.7380	0.0410	0.3598
3-March	0.0982	0.7745	0.0042	0.0332	0.0568	0.4481
4-April	0.1108	0.9033	0.0598	0.4875	0.0510	0.4160
5-May	0.0579	0.5077	0.009	0.0786	0.0257	0.2251
6-June	0.0102	0.0937	0.0181	0.1667	0.0042	0.0391
7-July	0.1270	1.1903	0.0506	0.4747	0.0460	0.4311
8-August	0.0342	0.3810	0.0582	0.6486	0.0618	0.6892
9-September	0.1077	0.9964	0.0751	0.6947	0.0592	0.5482
10-October	0.0081	0.0874	0.0237	0.2556	0.0439	0.4734
11-November	0.0019	0.0174	0.029	0.2616	0.0088	0.0796
12-December	0.0214	0.1799	0.0029	0.0241	0.0178	0.1497
<b>Annual</b>	<b>0.0756</b>	<b>7.8860</b>	<b>0.0514</b>	<b>5.5226</b>	<b>0.0444</b>	<b>4.8164</b>

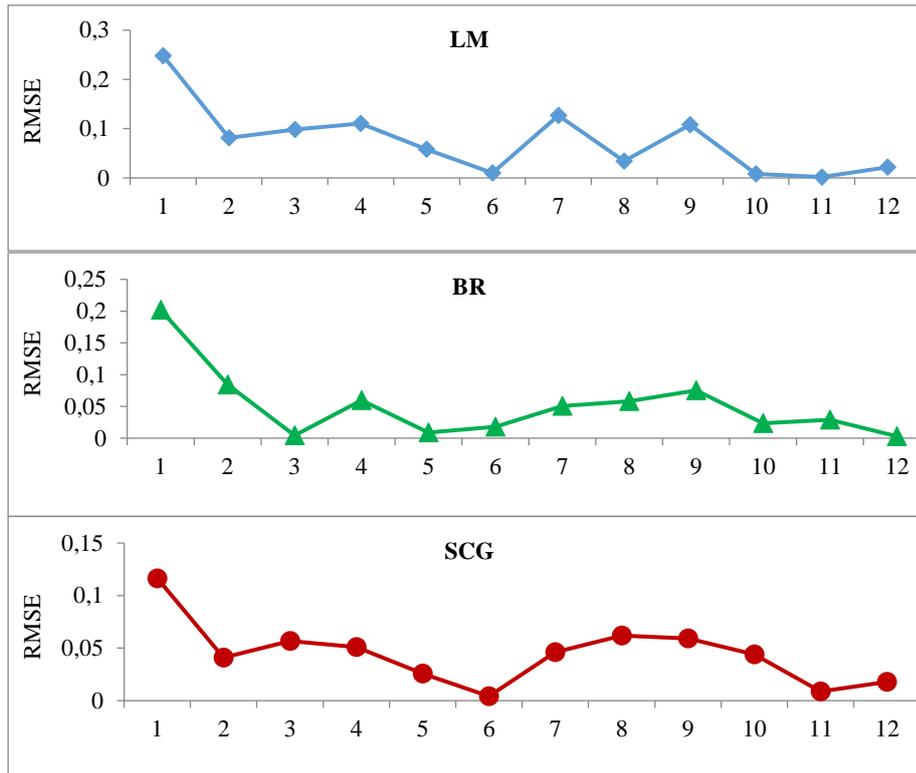


Figure 13. RMSE comparisons of LM, BR and SCG

**Test Results of Three Training Methods**

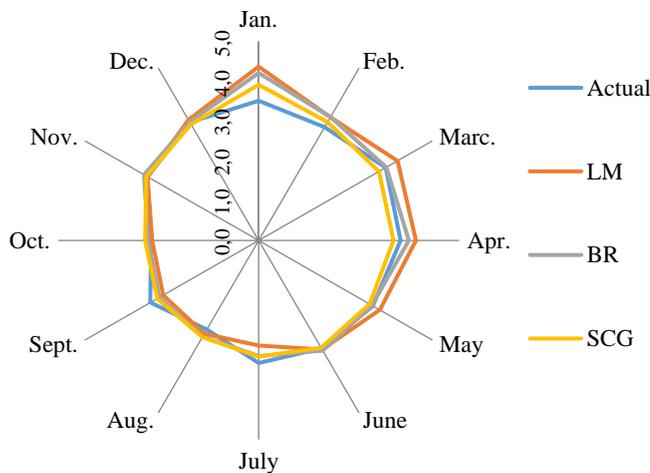


Figure 14. Actual and predicted wind speed values for three training methods ( LM, BR, SCG)

Looking at the low degree of estimation for January and April months, the lowest temperature in the data set between 2014-2017 is January 2017, since this value is not represented in the data set used for training of the network, it was estimated with low MAPE value. The estimated value of April has a relatively low percentage compared to the other months. The 3-year data set used for training is not sufficient to represent all meteorological conditions. However, it is seen that the degree of estimation is highly accurate according to the RMSE and MAPE values used to measure the estimation performance of the proposed network. All of the training methods showed lower performance in January compared to other months. Similarly, in April, the methods produced a lower accuracy in estimating the real value than in the other months. Although the data set used for training was only 3, the 2017 wind speed values were estimated to be close to the actual values in the test study (Figure 14).

### 3. CONCLUSIONS

In the electricity sector, it is important to make accurate forecasts of demand, to ensure the smooth supply of electricity and to utilize the resources correctly. Estimating the short-term demands can be difficult due to the presence of random elements in the estimation values obtained. In the general sense, it can be stated that no other attempts have been made so far to apply or improve the quality of forecasting wind velocity across Libya using an ANN model. For this reason, the model in the present work implies a major development in this field and a contribution to the available literature.

The aim of this study was to investigate the ability of BR, LM, and SCG neural network training algorithms to predict wind speed over a limited set of meteorological data. The performance of a system for forecasting is not only dependent on the network configuration, but also the correct selection of training algorithms is an important parameter in developing the optimum ANN structure. The objective of this work is to demonstrate the predictive abilities of BR, LM, and SCG neural network training algorithms.

This study evaluates the performance of the three different backpropagation learning algorithms by estimating the wind speed. The 3-year data out of the four-year meteorological data was used to train the network and the trained network was tested with one-year data. The statistical metrics root-mean-square error (RMSE) and mean absolute percentage error (MAPE) are employed to analyze the performance of the three learning algorithms. A neural network trained with the scaled Conjugate Gradient (SCG) algorithm has been shown to yield more accurate results than NN trained with the Levenberg-Marquardt (LM) and Bayesian Arrangement (BR). It is well known that the more data is available, the more accurate results are obtained in estimation. Therefore, the 3-year data set used for training in the present work is not sufficient to represent all meteorological conditions. However, it is seen that the degree of estimation is highly accurate according to the RMSE and MAPE values.

Finally, the study showed that the accuracy of predictability to wind speed using different training algorithm, and their performance was in the order of that SCG is best and LM is worst for this data set. Furthermore, when the exact ANN model is available, it has shown that the limited data used in the training phase is capable of estimating the wind speed within acceptable limits. Thus, choosing the right training algorithm for a limited dataset enables maximizing the predictive capability of ANN models. However, doing a similar study over a wider training set that covers long years may result in different results in the performances of the algorithms used for ANN.

### CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

### REFERENCES

- [1] Pinson, P., Nielsen, H.A., Madsen, H., Kariniotakis, G., "Skill forecasting from ensemble predictions of wind power", *Applied Energy*, 86 (7–8): 1326–34, (2009).
- [2] Watson, S.J., Landberg, L., Halliday, J.A., "Application of wind speed forecasting to the integration of wind energy into a large scale power system", *IEE Proc. Gen. Transm. Distrib.*, 141(4): 357–62, (1994).
- [3] Torres, J., Garcia, A., Deblas, M., Defrancisco, A., "Forecast of hourly average wind speed with ARMA models in Navarre (Spain)", *Sol. Energ.*, 79(1): 65–77, (2005).
- [4] Lin, L., Eriksson, J.T., Vihriala, H., Soderlund, L., "Predicting wind behavior with neural Networks", In *Proceeding the 1996 European wind energy conference*, Sweden, 655–8, (1996).

- [5] Beyer, H.G., Degner, T., Haussmann, J., Homan, M., Rujan P., “Short term forecast of wind speed and power output of a wind turbine with neural Networks”, In: Proceeding the second European congress on intelligent techniques and soft computing. Germany, (1994).
- [6] Kariniotakis, G., Stavrakakis, G.S., Nogaret, E.F., “Wind power forecasting using advanced neural network model”, IEEE Trans. Energy Convers., 11(4): 762–7, (1996).
- [7] Kariniotakis, G., Stavrakakis G.S., Nogaret, E.F., “A fuzzy logic and neural network-based wind power model”, In: Proceeding the 1996 European wind energy conference, Sweden, 596–9, (1996)
- [8] Celik, A.N., “Energy output estimation for small-scale wind power generation using Weibull representative wind data”, J. Wind Eng. Ind. Aerodyn., 91(5): 693 -707, (2003).
- [9] Oztopal, A., Kahya, C., Sahin, A.D., “Wind speed modelling with artificial neural network”, 3. National Clean Energy Symposium, Istanbul, Turkey, 415-422, (2000).
- [10] Kalogirou, S.A., “Artificial neural networks in renewable energy systems applications: a review”, Renewable Sustain. Energy Rev., 5: 373–401, (2001).
- [11] Sahin, A.D., Sen, Z., “Spatial normal distribution graphics methodology in western part of Anatolia”, II. National Clean Energy Symposium, Istanbul, Turkey, (1997).
- [12] Sirdas, S., “Daily wind speed harmonic analysis for Marmara region in Turkey”, Energy Conversion Manage., 46 (7- 8): 1267 – 127, (2005).
- [13] Sen, Z., “Terrain topograph classification for wind energy generation”, Renew. Energy, 16: 904 – 907, (1999).
- [14] Matheron, G., “Principles of geostatistics”, Econ Geol., 58: 1246 -1266, (1963).
- [15] Krige, D.G., “A statistical approach to some basic mine evaluation problems on the Witwateround”, J. Chimie. Min. Soc. South Africa, 119 – 139, (1951).
- [16] Sen, Z., Sahin, A.D., “Regional wind energy evaluation in some parts of Turkey”, J. Wind Eng. Ind. Aerodyn., 37(7): 740 – 741, (1998).
- [17] Bechrakis, D.A., Sparis, P.D., “Correlation of wind speed between neighboring measuring stations”, IEEE Trans. Energy Convers.,19: 400–406, (2004).
- [18] Abohedma, M. B., Alshebani, M. M., “Wind load characteristics in Libya”, Int. Journal of Civil, Env., Structural, Construction, and Architectural Engineering, 4(3): 88-91, (2010).
- [19] Saleh, I. M., “Prospects Of Renewable Energy In Libya”, International Symposium On Solar Physics And Solar Eclipses (SPSE), (2006).
- [20] Moharued, A. A., Ehnabrouk, A. M., “Assessment of the wind energy potential on the coast of Tripoli”, Hydrology, (2007).
- [21] El-Osta, W., Belhag, M., Klat, M., Fallah, I., Khalifa, Y., “Wind Farm Pilot Project In Libya”, Renewable Energy, 6(5-6): 639~42, (1995).
- [22] Embirsh, H. S. A., Ikshadah, Y.A., “Future prospects of The Wind Energy in Libya”, International Journal of Scientific and Research Publications, 7(10), (2017).

- [23] Wang, X., Zhao, Y., Pourpanah F., “Recent advances in deep learning”, *Int. Journal of Machine Learning and Cybernetics*, 11: 747–750, (2020).
- [24] Abiodun, O. I., Jantan, A., Omolara, A.E., Dada, K.V., Mohamed, N.A., Arshad, H., “State-of-the-art in artificial neural network applications: A survey.”, *Heliyon* 4 (2018). doi: 10.1016/j.heliyon.2018.e00938
- [25] Azizi, E., Kharrati-shishavan, H., Mohammadi-ivatloo, B., Mohammadpour Shotorbani, A., “Wind Speed Clustering Using Linkage-Ward Method: A Case Study of Khaaf, Iran”, *Gazi University Journal of Science*, 32(3), 945-954, (2019). DOI: 10.35378/gujs.459840
- [26] Kumar, K., Prabhu, K.R., Ramesh Babu, N., “Design and Analysis of Modified Single P&O MPPT Control Algorithm for a Standalone Hybrid Solar and Wind Energy Conversion System”. *Gazi University Journal of Science*, 30(4), 296-312, (2017).
- [27] Demuth, H., Beale, M., “Neural network toolbox user’s guide”, The Math Works, Inc., Natick, MA, 01760-2098, (2003).
- [28] Marquardt, D. W., “An Algorithm for Least- Squares Estimation of Nonlinear Parameters”, *Journal of the Society for Industrial and Applied Mathematics*, 11(2): 431–441, (1963).
- [29] MacKay, D. J. C., “A Practical Bayesian Framework for Backpropagation Networks”, *Neural Comput.*, 4(3): 448–472, (1992).
- [30] Abbo, K. K., Mohamed, H. H., “New Scaled Conjugate Gradient Algorithm for Training Artificial Neural Networks Based on Pure Conjugacy Condition, *Kirkuk University Journal for Scientific Studies*, 10(3): 230-241, (2015).