Journal of Engineering, Project, and Production Management
2025, 15(4), 2024-292

Comparative Analysis of ANN Algorithms for Wind Speed Forecasting in Renewable Energy Management

Yousef Altork[1], Duaa Salem[2], and Nabeel Abu Shaban[3]

[1] Assistant Professor, Department of Alternative Energy Technology, Faculty of Engineering and Technology, Al-Zaytoonah University of Jordan, P.O. Box 130, Amman 11733, Jordan, E-mail: y.altork@zuj.edu.jo (corresponding author).

[2] Master Student, Department of Civil and Infrastructure Engineering, Faculty of Engineering and Technology, Al-Zaytoonah University of Jordan, P.O. Box 130, Amman 11733, Jordan, E-mail: duaa.qadoumi@gmail.com

[3] Associate Professor, Department of Mechanical Engineering, Faculty of Engineering and Technology, Al-Zaytoonah University of Jordan, P.O. Box 130, Amman 11733, Jordan, E-mail: n.shaban@zuj.edu.jo

Emgineering Management
Received December 19, 2024; revised February 1, 2025; accepted April 21, 2025
Available online October 6, 2025

Abstract: Wind energy is critical to meeting the power requirement of the global population. Accurate wind speed forecasting is vital in energy trading, power system operations, and enhanced market balance. This study examines three Artificial Neural Network (ANN) algorithms using meteorological parameters as inputs. The three methods of training that have been discussed in the current paper are the Scaled Conjugate Gradient Backpropagation technique, the Bayesian regularization algorithm, and the Levenberg-Marquardt (LM) training algorithm. In this work, 42 datasets of the total 60 datasets procured from the National Renewable Energy Laboratory (NREL) over five years were used for training, 9 for testing, and the remaining 9 for validation. Wind speed will be the study's dependent variable while surrounding temperature, barometrical pressure, wind orientation, relative humidity, and rainfall amount are the independent variables. In the present study, the performance of the ANN algorithms is assessed using measures such as Mean Squared Error (MSE) and correlation coefficient (R). The results indicate that all three ANN algorithms exhibit excellent performance, enabling precise wind speed predictions. This reliability supports their potential application in optimizing wind energy operations. Among these, the LM algorithm provided the most precise predictions, exhibiting the lowest error rates. Therefore, it is concluded that the LM algorithm was the most accurate in predicting wind speed for the NWTC, providing valuable insights for renewable energy planning and management.

Keywords: Wind speed forecasting, Artificial Neural Network (ANN) algorithms, renewable energy management, comparative analysis, NREL datasets.

DOI 10.32738/JEPPM-2024-292

---

1. Introduction

Wind energy is a renewable energy source that is considered environmentally friendly, cost-competitive, and socially beneficial. However, precise information on wind dynamics at wind farm sites is crucial for managing and operating wind energy conversion systems because wind speed varies periodically and from site to site (Murthy and Rahi, 2017; Pinson et al., 2009). Short-term wind speed forecasting is essential for advancing the effectiveness of wind power generation systems and integrating wind energy into the power system (Chen et al., 2022; Luickx et al., 2008; Sfetsos, 2000), while long-duration wind speed forecasting is necessary for placement and appraisingwind energy applications (Alsamamra et al., 2024; Ucar and Balo, 2009).

The forecasting approach and its application can vary based on the information available and the relevant timeframe. Short-term wind speed forecasting is a subgroup of wind speed forecasting that covers periods ranging from seconds to days. Short-term forecasts are particularly useful for system operators, power producers, and wind farm developers who require forecasts for daily and intraday spot markets, system operations, and maintenance planning. Wind speed forecasts spanning a few hours attempt to tackle scheduling issues within a power system, while forecasts spanning several days are intended for maintenance and resource planning. For periods ranging from a few seconds to several minutes, the forecasting objective is to control wind energy conversion systems (Kaldellis et al., 2009; Costa et al., 2008; Masoud et al., 2022).

Over the years, several studies have been conducted to develop accurate wind forecasting techniques. Researchers have employed various soft computing approaches, including Artificial Neural Network (ANN), in a variety of fields (Manasrah et al., 2023; De Pauli et al., 2020; Thabtah et al., 2019).

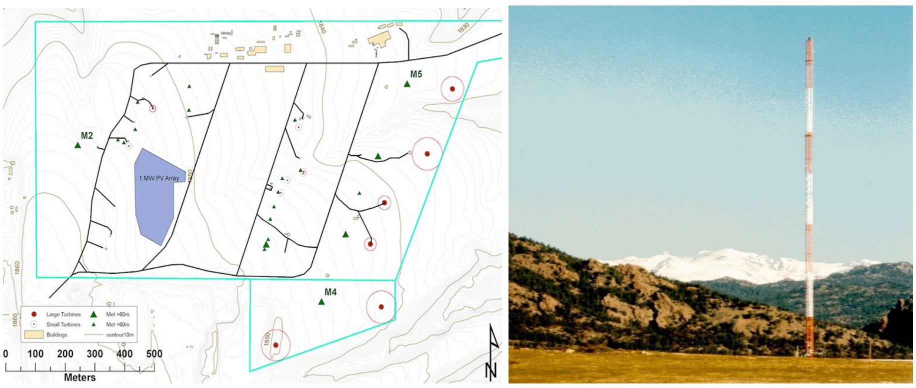
Huang and Kuo (2018) used the ANN model to estimate interim wind velocity by collecting a year's worth of data, using two months for training and one month for testing. The study compared WindNet, Random Forest (RF), Support Vector Machine (SVM), and Decision Tree (DT), revealing that WindNet technology outperformed the other models. Meanwhile, Barhmi et al., (2019) employed hourly meteorological satellite data gathered between 2011 and 2013 to predict wind speed using the ANN technique. The study compared the ANN approach with the Multiple Linear Regression (MLR) model, and the authors found that the ANN model yielded better results. Similarly, Dupré et al. (2020) used an ANN model to estimate wind speed and energy hourly for the years 2015 and 2016 using 10-minute observations. They compared the findings with the ARMA model and concluded that the ANN model performed better.

After reviewing existing literature, it is evident that various soft computing methods, particularly ANN techniques, are widely used in different fields including wind energy. However, there is a lack of comparative analysis with other approaches in wind energy forecasting. Therefore, this study focuses on employing different ANN algorithms to accurately estimate wind speed. Utilizing five years of time series data (2018-2022) from NREL, key meteorological variables such as barometrical pressure, surrounding temperature, wind orientation, relative humidity, and rainfall amount are considered independent variables. The main objectives of this study include developing and comparing different ANN algorithms based on statistical analysis to determine the most effective approach for wind speed prediction, which is crucial for wind farm operators.

2. Site Details

Located in the United States of America, the NWTC is a famous research facility established by the NREL for the US Department of Energy; the facility is located at the base of the Rocky Mountains near Boulder, Colorado. As one of its objectives, NWTC aims at collaborating with the wind energy industry to develop wind power technology so that the cost of wind energy is made cheaper. This is attained by innovation through research and development projects that focus on design innovations of wind turbines.

Fig. 1 shows the location of the towers at NREL's Flatirons Campaus and NWTC's M2 tower The geospatial location of the site includes 39.9106º North, Longitude: 105.2347º West, Elevation: 1855 AMSL, and Time Zone: MST and is accessible through NREL's measurement and instrument data center. The data is collected using devices installed on or near an 82-m meteorological tower located at the western end of the NWTC site at approximately 7.5 km south of Boulder, Colorado, and 11.5 km west of Broomfield, Colorado (Jager and Andreas, 2014). Three heights are employed to measure the air temperature. On the other hand, the tower uses six different heights to measure wind direction and speed. There is also other available information on the temperature regarding the dew point, the relative humidity, the barometric pressure, the total liquid precipitation, and sun radiation.



(a) (b)

Fig. 1. (a) The towers at NREL's Flatirons Campus. (b) M2 tower (Jager and Andreas, 2014)

3. Methodology

This study investigates the effectiveness of three ANN algorithms — Levenberg-Marquardt, Bayesian regularization, and Scaled Conjugate Gradient Backpropagation — for wind speed prediction. Utilizing meteorological data and past wind speed measurements, these models aim to provide accurate forecasts crucial for efficient wind energy management. Comparative investigation of these algorithms offers insights into their correctness and performance in renewable energy applications, contributing to sustainable energy practices and technological advancements.

ANN algorithms can provide accurate wind speed predictions to enhance the overall performance of wind farms by providing real-time adjustments to the turbine settings which in turn increases the energy output as well as decreasing mechanical stress. This results in improved power quality, improved energy trading strategies, and precise predictive maintenance, which in turn translates to higher operational efficiency and wind farm revenues.

The wind speed data obtained from ANN algorithms provide accurate information about wind speed, which in turn enables operators to produce energy in response to market needs and, therefore, realize better returns on their energy trading.

ANN-driven predictions can help in planning for maintenance in advance, thus reducing time that a turbine has to be out of service and also helping in increasing the lifespan of the turbines by resolving issues before they reach a point where they are catastrophic.

An accurate wind speed forecast increases the efficiency of wind energy integration into the electric grid, thereby minimizing the use of spinning reserves and guaranteeing a stable and firm power supply. Predictions assist in the determination of when to increase or decrease production, hence controlling the cost of human resources and operations while maximizing the generation of energy.

The prediction process, in this study, involves three stages: data pre-processing, model building, and model assessment. First, the raw meteorological data is normalized to the range between -1 and 1 to improve the model's effectiveness. After this, the ANN algorithms are trained with 70% of the data, validated with 15%, and tested on the remaining 15%. Last of all, to compare the predictive accuracy of the model, statistical coefficients like R and MSE are computed in the next section.

3.1. The Levenberg-Marquardt Training Algorithm

The training algorithm for artificial neural networks is an approximation of Newton's method known as the Levenberg-Marquardt (LM) technique (Aliouane, 2022). It is widely acknowledged as the most effective optimization algorithm for artificial neural networks.

When dealing with performance functions represented as a summation of squares, the Hessian matrix as outlined by Lampton can be approximated as Eq. (1) (Lampton, 1997):

$$H = J * J^T \tag{1}$$

The gradient can then be calculated as Eq. (2):

$$g = J^T * e \tag{2}$$

Here, J denotes the Jacobian matrix, which consists of the first-order partial derivatives of the network errors concerning the weights and biases, while e represents the vector of network errors.Determining the Jacobian matrix is computationally less intensive compared to the Hessian matrix. Therefore, the updated formula can be adjusted as in Eq. (3):

$$X_{k+1} = X_k - \frac{J^T * e}{J^T * J + \mu * I} \tag{3}$$

Here, μ is a scalar parameter that controls the algorithm's behavior. When μ = 0, the algorithm behaves like Newton's method using the approximate Hessian matrix. Higher values of μ transform it into gradient descent with a reduced step size.

In practice, this algorithm, particularly in neural networks, converges with fewer iterations. However, each iteration demands more calculations, especially for matrix inversion. Hence, its usage is limited to cases with a manageable number of parameters to optimize. The computational complexity of matrix inversion is directly proportional to the matrix and vector sizes, as noted by Hagan and Menhaj (1994).

3.2. Bayesian Regularization Algorithm

Bayesian regularization was established to transform nonlinear classifications into properly formulated problems, as discussed by Burden and Winkler (2008) and MacKay (1992). Typically, during training, the goal is to minimize the sum of squared errors between the model output and the goal value. Bayesian regularization introduces an extra term into this objective, given by Eq. (4) (Ticknor, 2013):

$$F = \beta E_D + \alpha E_w \tag{4}$$

Here, F represents the objective function, E_D is the sum of squared errors, E_w is the sum of squares of network weights, and α and β are parameters of the objective function, as explained by MacKay (1992). In Bayesian networks, weights are treated as random variables, and their density function is derived using Bayes' rules, it stated by Forsee and Hagan (1997) as in Eq. (5):

$$P(w|D, \alpha, \beta, M) = \frac{P(D|w, \beta, M)P(w|\alpha, M)}{P(D|\alpha, \beta, M)} \tag{5}$$

Here, w stands for the weight vector, D for the data vector, and M for the neural network model. Forsee and Hagan (1997) modeled the noise in the data as Gaussian to arrive at the probability density function of the weights. The choice of the regularization parameters α and β is the solution of the Hessian matrix of F(w) in the minimum point w_{MP}. Forsee and Hagan (1997) proposed the Gauss-Newton approximation of the Hessian matrix which is possible while using the Levenberg-Marquardt training to seek the minimum. This approach minimizes the chances of getting stuck at local optima, which improves the network's ability to generalize.

This technique is a new approach in handling the weights of the networks with reference to the data given and the architecture of the model. Growing neural networks with additional hidden layer neurons greatly increase the chances of overfitting; a validation set is used to decide upon an optimal point to stop. In the case of Bayesian regularized networks, model complexity is reduced as extra connection weights are considered unnecessary hence they approach zero. The network concentrates on the training of the large weights which are also known as the effective number of parameters that become fixed as the size of the network grows (Burden and Winkler, 2008). The noise and fluctuation characteristics of stock markets contribute to overfitting and overtraining problems in the conventional backpropagation networks. These even more spare networks minimize overfitting dangers and exclude the necessity of having a validation subset, which gives more training data.

3.3. Scaled Conjugate Gradient Backpropagation Algorithm

The fundamental backpropagation algorithm corrects weights in the direction of precipitous descent, which is the descending trajectory leading to the fastest decrease in the performance function. However, this approach may not always result in the fastest convergence, as noted by Hagan et al. (2002). In contrast, conjugate gradient (CG) algorithms perform searches along directions that typically lead to faster convergence compared to the precipitous descent, although still maintaining error minimization attained in prior steps, as highlighted by Ali and Mahdi (2023). These directions are referred to as conjugate directions.

Most conjugate gradient techniques find the value of the step size which optimizes the performance function laterally the conjugate gradient direction and the step size is changed in every iteration. During the first iteration of the CG algorithms, the search vectors move in the direction of the steepest descent. To estimate step size, the CG algorithms often involve line search techniques. This does away with the need to calculate the Hessian matrix to determine the optimal step size in the current direction of the search. It is then seen that the next search direction is conjugate to the current search direction. Hagan et al. (2002) describe the typical approach to obtain the new search direction from the previous search direction and the new steepest descent direction as in Eqs. (6-8).

$$p_o = -g_o \tag{6}$$

$$x_{k+1} = x_k + \alpha_k g_k \tag{7}$$

$$p_k = -g_k + \beta_k p_{k-1} \tag{8}$$

Different versions of conjugate gradient (CG) algorithms are distinguished based on how the factor β is computed, as discussed by Ali and Mahdi (2023).

Instead, another method that provides an estimation of the step size is the CG approach combined with the trust region model which is part of the LM algorithm rather than the line search technique. The Scaled Conjugate Gradient (SCG) is another combination method that was developed by Møller (1993). Scaling factors λ and σ are presented in this method to estimate the Hessian matrix, where s is the Hessian matrix approximation, E is the entire error function, and E' is the gradient of E. At the beginning of the procedure, the user initializes these factors so that $0 < \sigma < 10^{10}$ and $0 < \lambda < 10^{10}$. For SCG, the following expression can be used to calculate the β factor and determine the new search direction as in Eqs. (9-11) (Møller, 1993):

$$s_k = \frac{\grave{E}(w_k + \sigma_k p_k) - \grave{E}(w_k)}{\sigma_k} + \lambda_k p_k \tag{9}$$

$$\beta_k = \frac{(|g_{k+1}|^2 + {g_{k+1}}^T g_k)}{{g_k}^T g_k} \tag{10}$$

$$p_{k+1} = -g_{k+1} + \beta_k p_k \tag{11}$$

The design parameters are updated independently by the user at each iteration, which is essential for the algorithm's success. This autonomy is a significant advantage compared to algorithms based on line search.

3.4. Performance Evaluation Indices

The precision of the wind speed prediction model is evaluated using metrics such as mean square error (MSE), and the correlation coefficient (R), which are derived from comparing actual and predicted wind speed values. The formulas for these parameters are referenced as Eqs. (12, 13)(Barhmi et al., 2019):

$$MSE = \frac{1}{n}\sum_{i=1}^{M}(X_{A,i} - X_{p,i})^2 \tag{12}$$

$$R = \frac{\sum_{i=1}^{n}\left(X_{P,i} - \bar{X}_P\right).\left(X_{A,i} - \bar{X}_A\right)}{\sqrt{\sum_{i=1}^{n}\left(X_{P,i} - \bar{X}_P\right)^2.\sum_{i=1}^{n}\left(X_{A,i} - \bar{X}_A\right)^2}} \tag{13}$$

4. Results and Discussion

This study utilizes MATLAB software and arithmetical models to accurately estimate the wind speed in NWTC, Colorado, USA. To achieve this, three different ANN algorithms prediction techniques are employed; Levenberg-Marquardt, Bayesian regularization, and Scaled Conjugate Gradient Backpropagation. The development of these models incorporates meteorological and operational parameters, as well as evaluated wind speed values.

All models utilize five parameters as independent variablesas in Table 1. Monthly average data spanning five years (from 2018 to 2022) is analysed, resulting in a total of 60 datasets (12 datasets per year), each containing values for the five input variables and one output variable (wind speed). Statistical parameters for the raw data are summarized in Table 1, indicating an average wind speed at the site of 4.25 m/s, with fluctuations ranging from 3.33 to 6.55 m/s. These five inputs were chosen based on prior literature on meteorological conditions impacting wind energy (Hodge et al., 2011; Chen et al., 2019), with wind speed being most sensitive to the changes in these parameters.

Table 1. The numerical overview of the documented data from NREL employed in the current investigation

Parameters	Ta [°C] at 50m height	RH [%]	Pressure [mBar]	Precipitation [mm]	Wind Direction [°]	Wind speed [m/s]
Minimum	-13.5940	32.5706	806.6900	0	179.3376	3.3309
Maximum	24.0808	69.1236	819.8150	2.2604	231.4231	6.5523
Standard deviation	9.3133	7.3287	2.9626	1.2701	10.1287	0.7039
Variance	86.7378	53.7110	8.7770	1.6132	102.5914	0.4955
Median	8.4660	45.5812	813.6645	0.3850	201.0284	4.2526
Mean	9.9496	46.4526	813.5569	0.3316	202.3264	4.3969

The common approach for prediction is contained within the framework of Artificial Neural Network – the Multilayer Perceptron (MLP) model. This was done to develop an MLP structure that was to be used specifically for wind speed prediction at the research site. As shown in Fig. 2, five parameters are input layers of the MLP model and one parameter is specified as an output layer. Some examples of input parameters include meteorological and operating parameters such as air pressure, relative humidity, wind direction, temperature, and precipitation. The output layer contains the output parameter which is the wind speed as identified during the training of the model.

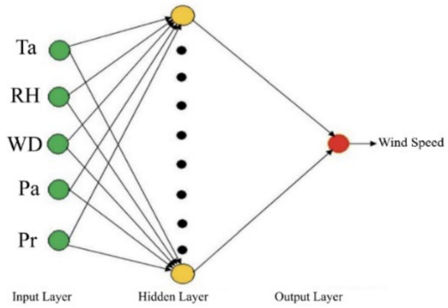


Fig. 2. Illustrates the MLP architecture comprising 5 input parameters and 1 output parameter

A dataset comprising 60 data points obtained from NREL was organized for use in the MLP model. Out of these datasets, 70% were allocated for training, and the remaining 30% were divided equally for validation and testing the model. Prior to constructing the neural model, the experimental datasets were normalized to a range between -1 and 1 using a prescribed Eq. (14) (Ghritlahre and Prasad, 2018):

$$Z = 2 * \frac{Z_i - Z_{min}}{Z_{max} - Z_{min}} - 1 \tag{14}$$

where Z is the experimental data.

Numerous researchers have relied on trial approaches to determine the optimal number of neurons, while others have devised formulas to calculate the desired number of hidden neurons. One of the most widely used formulas is that proposed by Ghritlahre and Prasad (2018) and presented in Eq. (15):

$$H_a = \frac{A + B}{2} + \sqrt{T_n} \tag{15}$$

where A represents the input parameters. B represents the number of output parameters. Tn presents the training datasets. By applying the formula provided in Eq. (15), a total of 10 neurons were identified. Consequently, this specific quantity of neurons was selected to establish the comprehensive construction of the neural model.

4.1. The Levenberg-Marquardt Training Algorithm Results

The 5-10-1 neuronic model was ultimately trained using the LM learning algorithm. To assess the effectiveness of the MLP model, arithmetical parameters such as R and MSE were employed, as computed using Eqs. (12) and (13). The statistical analysis of the 5-10-1 neuronic model is presented in Table 2.

From the arithmetical analysis, the R values for train, validate, and test were determined to be 0.9872, 0.9755, and 0.9965, respectively. These values represent the highest levels observed compared to other neuron-based models. Additionally, the lowest MSE values were 0.0703, 0.0513, and 0.0245, respectively.

Table 2. The arithmetical investigation of 5-10-1 neuronic model using LM algorithm

	Samples	MSE	R
Training	42	0.0703	0.9872
Validation	9	0.0513	0.9755
Testing	9	0.0245	0.9965

Fig. 3 below depicts the 5-10-1 model's performance curve where the y-axis represents the MSE values and the x-axis represents the training epoch. The training process was terminated at epoch 14, which corresponds to the minimum values of MSEs for the validation sets. In the case of validation, MSE of 0.051315 was achieved during epoch 14 and hence it can be said that it had the best validation performance.

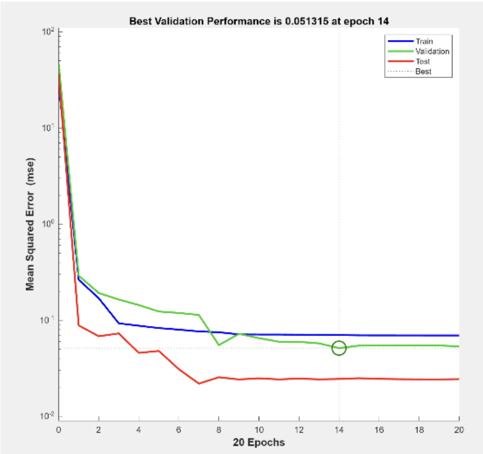


Fig. 3. LM algorithm performance curve

Furthermore, the regression plot of the optimal 5-10-1 model is shown in Fig. 4, indicating R values of 0.98723, 0.97552, 0.99653, and 0.98861 for model train, test, validate, and all processes, respectively. This alignment underscores the consistency between the actual values of recorded data from NREL and the output data predicted by the LM ANN algorithm.

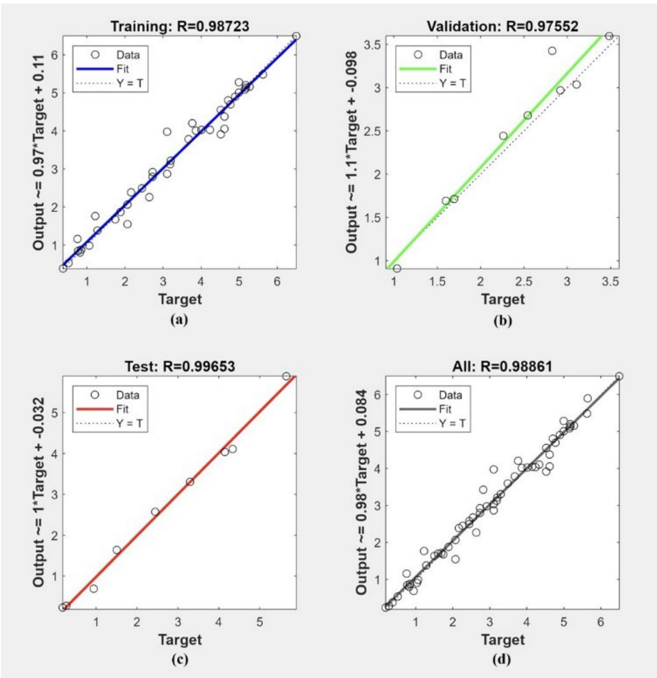


Fig. 4. LM algorithm regression plot of 5-10-1 and R-values of the model (a) Train, (b) Validate, (c) Test, and (d) Overall

4.2. Bayesian Regularization Algorithm Results

By applying the Bayesian regularization algorithm, it was determined that the values of R for training, and testing stood at 0.98969, and 0.97733, respectively. These figures signify the highest levels observed in comparison to other neuron-based models.

Fig. 5 illustrates the performance curve of the 5–10-1 model, demonstrating a gradual decrease in MSE values with each epoch. The training process concluded at epoch 36, coinciding with the lowermost MSE of the validating groups. The optimal validation execution was achieved at epoch 36, with an MSE throughout validation of 0.053713.

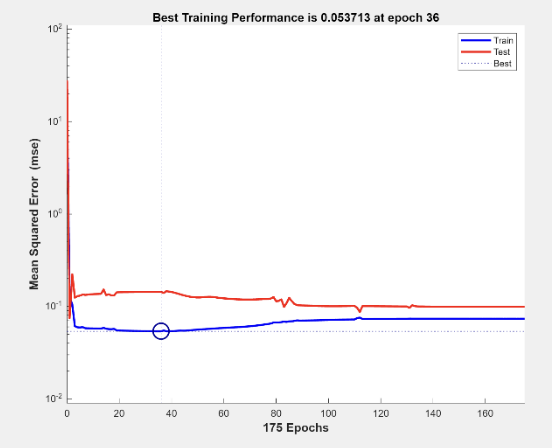


Fig. 5. Bayesian regularization algorithm performance curve

Additionally, Fig. 6 presents the regression plot of the optimal 5-10-1 model, highlighting R values of 0.98969, 0.97733, and 0.98748 for training, testing, and all processes, respectively. This coherence underscores the consistency between the real values of documented data from NREL and the output data predicted by the Bayesian regularization ANN algorithm.

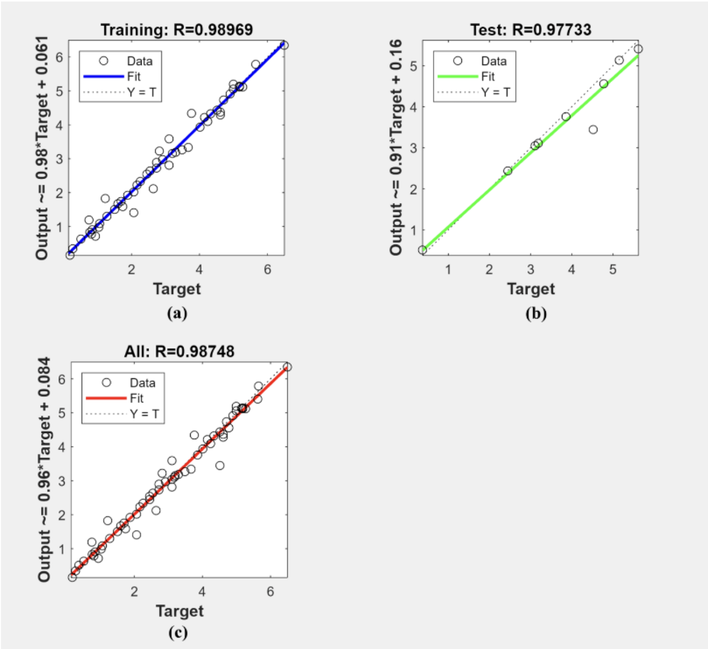


Fig. 6. Bayesian regularization algorithm regression plot of 5-10-1 and R-values of the model (a) Train, (b) Test, and (c) Overall

4.3. SCG Backpropagation Algorithm Results

Based on the scaled conjugate gradient backpropagation algorithm, the R values for training, validation, and testing were determined as 0.98325, 0.97282, and 0.98670, respectively, marking the highest levels observed compared to other neuron-based models.

Fig. 7 portrays the performance curve of the 5–10-1 model, showcasing a progressive decline in MSE values with each epoch. The training process concluded at epoch 19, coinciding with the lowermost MSE of the validating groups. Optimal validation performance was achieved at epoch 19, with an MSE during validation of 0.11103.

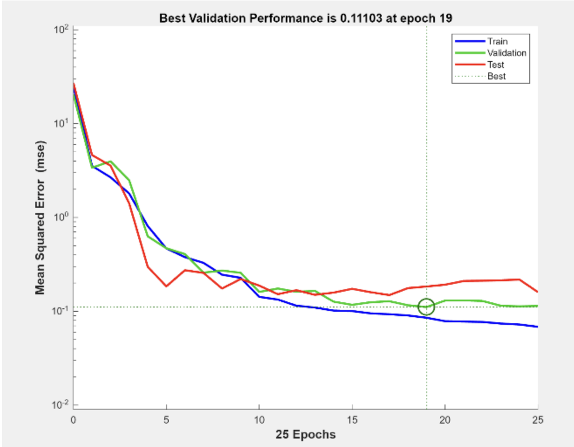


Fig. 7. SCG algorithm performance curve.

Additionally, Fig. 8 presents the regression plot of the optimal 5-10-1 model, revealing R values of 0.98325, 0.97282, 0.98670, and 0.98037 for the model's train, test, validate, and all processes, respectively. This alignment underscores the consistency between the actual values of recorded data from NREL and the output data predicted by the SCG ANN algorithm.

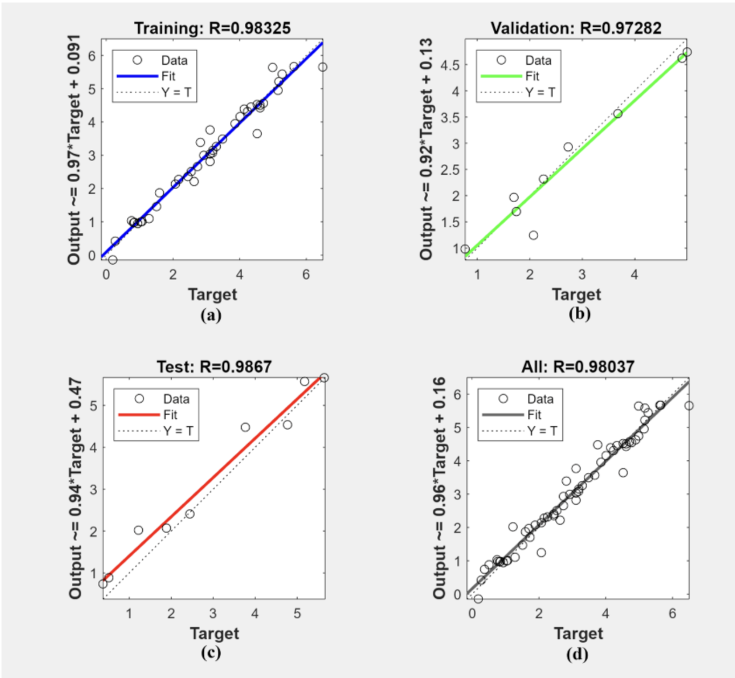
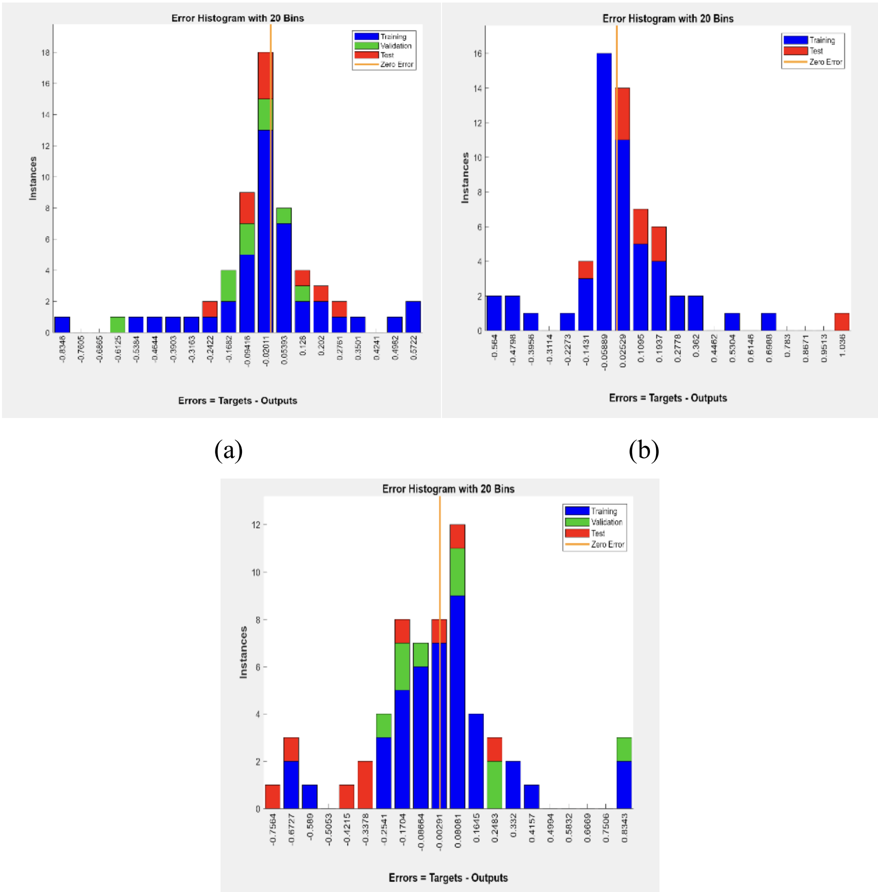


Fig. 8. SCG algorithm regression plot of 5-10-1 and R-values of the model (a) Train, (b) Validate, (c) Test, and (d) Overall.

4.4. Comparison of Different Algorithms

Fig. 9 displays individual error samples, revealing that the errors of the LM algorithm are notably close to zero and significantly less in comparison to those of the Bayesian regularization, and SCG algorithms. Upon analyzing the errors, it was observed that in the LM algorithm, 35% of errors fell between −0.09416 and +0.05393. Conversely, in the Bayesian regularization, and SCG algorithms, 47% and 43% of errors respectively accumulated within the ranges of −0.1431 to +0.1937 and −0.2541 to +0.1645 as seen in the figure below. Additionally, Fig. 9 illustrates that the errors in the LM algorithm are fewer than those in the Bayesian regularization, and SCG algorithms.

Based on the above observations, it can be concluded that ANN algorithms such as LM, Bayesian regularization, and SCG accurately predicted the wind speed at the research site. Furthermore, among these models, the LM algorithm demonstrates the least error when compared to the Bayesian regularization, and SCG algorithms.



(a) (b)

(c)

Fig. 9. Histogram error graph of (a)LM, (b) Bayesian regularization, and (c) SCG algorithms

These differences in performance levels of the three can be explained by the difference in the way the three algorithms are optimized. LM algorithm which is a mixture of the gradient descent and Gauss-Newton algorithm is more suitable for the problem involving a small number of data points and yields faster convergence along with better precision in error minimization. This characteristic makes it more accurate than the other algorithms with less error range as shown above.

However, the Bayesian Regularization algorithm also contains a term for regularization to reduce overfitting which helps in generalization. But with a few more errors, the algorithm provides equal importance to fitting the model on the training data and the model's predictive strength.

Although relying on SCG yields computational efficiency and scalability, second-order optimization is an approximation method. This makes it ideal for use with large amounts of data, however, it appears to have a larger error margin because of low precision convergence in small datasets.

Hence, the differences described above reveal what was mentioned before: the trade-off between the performance in terms of accuracy, samples' coverage, and the time the algorithms take to make their prediction. The better performance achieved by the LM algorithm for this dataset suggests that the optimization of the given set of parameters is more important for applications used where accuracy is important compared to the amount of data available or the complexity of the problem. The Bayesian Regularization and SCG algorithms may therefore be more suitable for larger, noisy data sets and problems with high degree of generalization.

4.5. Limitations and Future Research Directions

This study has some limitations which are discussed below, although it offers useful information about the performance of various ANN algorithms for wind speed prediction. First, the study only employs three different ANN algorithms while it does not compare them with other machine learning or even classical statistical techniques. This hampers the generalization of the findings and may miss the possible benefits of other methods. The study might be improved in future by the comparison of more machine learning techniques as well as traditional methods to give a broader analysis.

Second, the evaluation of how important each input variable is relative to the others was not carried out in this study. This raises the possibility that detailed knowledge about the importance of individual input variables, including barometrical pressure, air temperature, wind orientation, relative humidity, and rainfall, might help to yield better insights into how they affect the prediction of the wind speed and hence enhance the performance of the model. Future research should take into account the sensitivity analysis to establish the significance of each factor to the model.

Furthermore, the study uses historical data and therefore may not necessarily reflect real-life wind energy operations. Real-world implementation and testing of the models on real wind farm scenarios are also required to evaluate their application in real-life conditions.

In the same way, subsequent research studies can extend from the current work, remove the aforementioned limitations, increase the predictive accuracy, and contribute to the development of more efficient wind energy forecasting models.

5. Conclusions

In this study, three distinct approaches were employed to forecast wind speed for the NWTC, Colorado, USA, namely ANN algorithms: the LM training algorithm, Bayesian regularization algorithm, and SCG Backpropagation algorithm. The recorded data span various parameters, ranging from 32.57 % to 69.12 % for relative humidity, -13.59 °C to 24.08 °C for temperature, 179.34° to 231.42° for wind direction, 806.69 mbar to 819.82 mbar for pressure, and 0 mm to 2.26 mm for precipitation.

Surrounding temperature, barometrical pressure, wind orientation, relative humidity, and rainfall amount were predictor variables in entire models, while documented wind speed was the dependent variable. An overall of 60 datasets, representing monthly average data from 2018 to 2022, were employed across all models, with 42 datasets allocated for model train, 9 for testing, and 9 for validating. Comparative analyses of the models were conducted by means of performance indices such as MSE, and R.

The results revealed that all algorithms, LM, Bayesian regularization, and SCG, exhibited satisfactory performance, with correlation coefficients (R) of 0.98861, 0.98748, and 0.98037, respectively. Notably, the LM algorithm revealed the maximum accurate predictions, boasting the lowest errors. Consequently, it is concluded that the LM algorithm most precisely forecasts the wind speed for the NWTC. Furthermore, the authors suggest exploring relevant input parameters to forecast wind speed at the anticipated site or other research locations using ANN techniques and determining the optimal model. Additionally, techniques such as genetic algorithms (GA) could be considered for comparative studies in future research endeavors.

Author Contributions

Yousef Altork contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, and project administration. Duaa Salem contributes to conceptualization, methodology, draft preparation, manuscript editing, and visualization. Nabeel Abu Shaban contributes to analysis, investigation, data collection, draft preparation, and manuscript editing. All authors have read and agreed with the manuscript before its submission and publication.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

References

Ali, E. and Mahdi, S. (2023). A family of developed hybrid four-term conjugate gradient algorithms for unconstrained optimization with applications in image restoration. *Symmetry*, 15(6), 1203.

Aliouane, L. (2022). Levenberg-Marquardt algorithm neural network for clay volume estimation from well-log data in an unconventional tight sand gas reservoir of Ahnet basin (Algerian Sahara). *Bulletin of Geophysics and Oceanography*, 63, 443-454.

Alsamamra, H. R., Salah, S., and Shoqeir, J. H. (2024). Performance analysis of ARIMA Model for wind speed forecasting in Jerusalem, Palestine. *Energy Exploration & Exploitation*, 01445987241248201.

Barhmi, S., Elfatni, O., and Belhaj, I. (2019). Forecasting of wind speed using multiple linear regression and artificial neural networks. *Energy Systems*. doi: 10.1007/s12667-019-00338-y.

Burden, F. and Winkler, D. (2008). Bayesian regularization of neural networks. *Methods in Molecular Biology*, 458, 25-44.

Chen, G., Tang, B., Zeng, X., Zhou, P., Kang, P., and Long, H. (2022). Short-term wind speed forecasting based on long short-term memory and improved BP neural network. *International Journal of Electrical Power & Energy Systems*, 134, 107365.

Chen, Y., He, Z., Shang, Z., Li, C., Li, L., and Xu, M. (2019). A novel combined model based on echo state network for multi-step ahead wind speed forecasting: A case study of NREL. *Energy conversion and management*, 179, 13-29.

Costa, A., Crespo, A., Navarro, J., Lizcano, G., Madsen, H., and Feitosa, E. (2008). A review on the young history of wind power short-term prediction. *Renewable and Sustainable Energy Reviews*, 12(6), 1725-1744. doi: 10.1016/j.rser.2007.01.015.

De Pauli, S. T. Z., Kleina, M., and Bonat, W. H. (2020). Comparing artificial neural network architectures for Brazilian stock market prediction. *Annals of Data Science*, 7(4), 613-628. doi: 10.1007/s40745-020-00305-w.

Dupré, A., Drobinski, P., Alonzo, B., Badosa, J., Briard, C., and Plougonven, R. (2020). Sub-hourly forecasting of wind speed and wind energy. *Renewable Energy*, 145, 2373-2379. doi: 10.1016/j.renene.2019.07.161.

Forsee, F. D. and Hagan, M. T. (1997). Gauss-Newton approximation to Bayesian learning. *1997 IEEE International Conference on Neural Networks*, 1930-1935. doi: 10.1109/ICNN.1997.614194.

Ghritlahre, H. K. and Prasad, R. K. (2018). Exergetic performance prediction of solar air heater using MLP, GRNN, and RBF models of artificial neural network technique. *Journal of Environmental Management*, 223, 566-575. doi: 10.1016/j.jenvman.2018.06.

Hagan, M. T. and Menhaj, M. B. (1994). Training feedforward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks*, 5, 989-993.

Hagan, M. T., Demuth, H. B., and Jesús, O. D. (2002). An introduction to the use of neural networks in control systems. *International Journal of Robust and Nonlinear Control*, 12(11), 959-985.

Hodge, B. M., Zeiler, A., Brooks, D., Blau, G., Pekny, J., and Reklatis, G. (2011). Improved wind power forecasting with ARIMA models. *Computer aided chemical engineering,* 29, 1789-1793.

Huang, C. J. and Kuo, P. H. (2018). A short-term wind speed forecasting model using artificial neural networks with stochastic optimization for renewable energy systems. *Energies*, 11(10), 2777. doi: 10.3390/en11102777.

Jager, D. and Andreas, A. (2014). NREL National Wind Technology Center (NWTC): M2 Tower; Boulder, Colorado (Data). *NREL Data Catalog*. doi: 10.7799/1052222.

Kaldellis, J. K., Kavadias, K. A., and Filios, A. E. (2009). A new computational algorithm for the calculation of maximum wind energy penetration in autonomous electrical generation systems. *Applied Energy*, 86(7-8), 1011-1023. doi: 10.1016/j.apenergy.2008.08.014.

Lampton, M. (1997). Damping-undamping strategies for the Levenberg-Marquardt nonlinear least-squares method. *Computational Physics Journal*, 11, 110-115.

Luickx, P. J., Delarue, E. D., and D'haeseleer, W. D. (2008). Considerations on the backup of wind power: Operational backup. *Applied Energy*, 85(9), 787-799. doi: 10.1016/j.apenergy.2008.01.004.

MacKay, D. J. C. (1992). A practical Bayesian framework for backpropagation networks. *Neural Computation*, 4, 448-472.

Manasrah, A., Masoud, M., Jaradat, Y., Alia, M., Almanasra, S., and Suwais, K. (2023). Shading anomaly detection framework for bi-facial photovoltaic modules. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 45(4), 12955-12972.

Masoud, M., Jaradat, Y., and Alsakarnah, R. (2022). A non-content multilayers hybrid machine learning web phishing detection model. *International Review on Modelling and Simulations*, 15(2), 108-115.

Møller, M. F. (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 6(4), 525-533.

Murthy, K. S. R. and Rahi, O. P. (2017). A comprehensive review of wind resource assessment. *Renewable and Sustainable Energy Reviews*, 72, 1320-1342.

Pinson, P., Nielsen, H. A., Madsen, H., and Kariniotakis, G. (2009). Skill forecasting from ensemble predictions of wind power. *Applied Energy*, 86(7-8), 1326-1334. doi: 10.1016/j.apenergy.2008.10.005.

Sfetsos, A. (2000). A comparison of various forecasting techniques applied to mean hourly wind speed time series. *Renewable Energy*, 21(1), 23-35. doi: 10.1016/S0960-1481(99)00075-2.

Thabtah, F., Zhang, L., and Abdelhamid, N. (2019). NBA game result prediction using feature analysis and machine learning. *Annals of Data Science*, 6(1), 103-116. doi: 10.1007/s40745-018-00189-x.

Ticknor, J. L. (2013). A Bayesian-regularized artificial neural network for stock market forecasting. *Expert Systems with Applications*, 40(14), 5501-5506.

Ucar, A. and Balo, F. (2009). Investigation of wind characteristics and assessment of wind-generation potentiality in Uludag-Bursa, Turkey. *Applied Energy*, 86(3), 333-339. doi: 10.1016/j.apenergy.2008.05.008.

Dr. Yousef Altork has been the head of the Alternative Energy Technology Department at Al-Zaytoonah University of Jordan (ZUJ) since 2022. He holds a PhD in Mechanical Engineering from the University of Southern Queensland, Australia, and a Master's in Renewable Energy from Jordan University. Dr. Altork is renowned for his work in solar thermal energy and the integration of AI in renewable energy systems. His innovations in solar-powered refrigeration and AI-optimized renewable energy systems have significantly advanced efficiency and sustainability in the field. His research interests include solar thermal energy, AI in renewable energy, and sustainable energy systems.

Eng. Duaa Salem holds a BSc in Mechatronics Engineering from Al-Balqa Applied University and is currently a Master's student in Engineering Management at Al-Zaytoonah University of Jordan. Her research areas focus on engineering management, sustainability, artificial intelligence, and renewable energy. Eng. Salem combines her technical background in mechatronics with her advanced studies in management to explore innovative solutions that enhance efficiency and promote sustainable practices.

Dr. Nabeel Abu Shaban Ph.D. in Mechanical Engineering from University of Jordan (2014). He is Dean of the Faculty of Engineering and Technology, Al-Zaytoonah University of Jordan, (Oct. 2022 – present) and Associate professor in the Department of Mechanical Engineering (Oct. 2016 – present). His research interests are Energy, Fluid mechanics, Manufacturing, and sustainability.