IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 21 May 2024, accepted 15 July 2024, date of publication 18 July 2024, date of current version 26 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3430830

RESEARCH ARTICLE

# A Novel Ensemble Wind Speed Forecasting System Based on Artificial Neural Network for Intelligent Energy Management

**MERVE ERKINAY OZDEMIR**
Department of Electrical and Electronic Engineering, Iskenderun Technical University, 31200 Iskenderun, Turkey
e-mail: merve.erkinayozdemir@iste.edu.tr

**ABSTRACT** Accurate and consistent wind speed forecasting is vital for efficient energy management and the market economy. Wind speed is non-linear, non-stationary, and irregular, so it is very difficult to forecast. There are many forecasting methods currently in use; however, selecting and developing the most appropriate method for a particular region in wind speed forecasting is still a hot topic. This study presents a new and unique neural network-based ensemble system for forecasting wind speed, which is very difficult to predict but is directly related to the power generated by wind farms for individual and different sites. With the developed ensemble model, average mean absolute error, mean absolute percentage error and root mean square error values are obtained as 0.1269, 3.074%, 0.1596 respectively. Test results demonstrate significant contributions of the proposed system compared to existing statistical, heuristic and ensemble models, indicating that the developed model is a promising alternative for wind speed forecasting models. The obtained results show that this system is an effective and useful intelligent tool that can be used by various companies and government facilities that invest and operate in intelligent wind energy technologies.

**INDEX TERMS** Artificial neural networks, ensemble forecasting, intelligent energy management, wind energy, wind speed.

## ACRONYMS

| | |
|---|---|
| ANN | Artificial Neural Network. |
| MAE | Mean Absolute Error. |
| MAPE | Mean Absolute Percentage Error. |
| RMSE | Root Mean Square Error. |
| NWF | Numerical Weather Forecast. |
| LM | Levenberg Marquardt. |
| SCG | Scaled Conjugate Gradient. |
| TanSig | Tangent Sigmoid. |

## I. INTRODUCTION

With technological developments in many industries and the world's growing population, the need for affordable and sustainable energy sources has increased. Existing energy production and consumption systems cause air, water, and soil pollution on a local, regional, and global scale. Pollutants, especially greenhouse gases of fossil fuels, cause global warming and climate change. The most important method of reducing pollutants includes environment-friendly, sustainable, and renewable energy sources [1]. Among the other renewable options, wind energy is one of the fastest developing energy resources in the world and is the most financed type of energy [2]. It is the most advanced and commercially available energy type among renewable energies and also one of the cheapest energy sources [3].

Wind power plants need constant and convenient wind speed for adequate energy generation [4]. Energy production in wind turbines is a direct function of wind speed and unlike conventional generation systems, the production consumption balance cannot be easily adjusted [5]. Forecasting wind speed is becoming progressively crucial for supply-demand balancing and intelligent wind farm management systems [6]. Therefore, for the stability and efficiency of the power systems, it is necessary to advance immensely accurate wind speed forecasting methods. Wind speed forecasting is also

The associate editor coordinating the review of this manuscript and approving it for publication was Chandan Kumar.

essential for wind energy conversion systems as it significantly impacts production rate, including the dynamic control of a wind turbine and the programming of a power system [7]. Wind speed forecasting is very important for estimating the energy expected to be produced from wind turbines [8] and increasing the efficiency of wind power generating systems [9], [10]. However, due to its unsteady and intermittent nature, wind speed is often difficult to accurately forecast during the operation. Wind speed is affected by meteorological parameters such as humidity, air temperature and pressure, so these parameters are critical to determine wind speed [11].

Various methods have been developed in the literature to increase wind speed forecasting accuracy at different conditions. These methods can be grouped as persistence, physical, statistical, heuristic and ensemble models [12], [13]. The persistence method, which is accepted as the simplest time series model, is based on the assumption that the measured value of the wind speed will not change for the time to be estimated [14]. When trying to predict wind speed with physical methods such as Mesoscale models and Numerical Weather Forecast (NWF), data such as orography, roughness, obstacles, pressure, and temperature are primarily used. In the statistical method, the relationship between wind power data and wind speed is tested to determine historical data. Wind speed forecasting using persistence and physical models performs well in the time series forecasting task and has fast computation times. It performs particularly well in long-term wind speed forecasting. In statistical methods such as AR, ARMA, and ARIMA, wind speed forecasting is made based on the time series method [15]. Statistical models are easily trained and implemented. Heuristic models include artificial neural networks (ANNs) [16], neural fuzzy logic [17], genetic algorithm [18], and support vector machine [19], [20]. These methods do not require assumptions about the data and can be easily applied to linear or non-linear data. Neural networks can effectively solve complex and non-linear problems by learning complex non-linear relationships with its architecture that uses weight parameters to transform input data in hidden layers [33], [34]. Fuzzy logic provides an effective solution by modelling complex non-linear problems through local linear models such as Takagi Sugeno [35]. The previous studies showed that using ANN methods instead of complex rules and mathematical routines in forecasting wind speed can provide faster solutions [21], [22], [23]. In particular, physical methods such as numerical weather prediction uses complex mathematical models (Predictor, High-Resolution Limited Area Model, Global Forecasting System) [36]. As these models are computationally intensive, they are run on supercomputers. These models are not used for short and medium term wind speed forecasts as they require additional cost and processing [37]. Ensemble models use various types of forecasting methods to predict wind speed. To improve the forecast performance, ensemble models combine different methods, such as statistical models, ANN, support vector machine and other machine learning methods [24], [25].

These methods mentioned above have several disadvantages. Persistence and physical methods need large data, which slows down and complicates the operation of the system. The forecasting performance of statistical models is poor for non-linear data. Heuristic models and ensemble models, on the other hand, have disadvantages such as needing too much training data requiring pre-processing for a large number of samples [44], and being stuck on local solutions as the number of samples increases, overlearning, and slowing down the system. As a result, the forecasting performance of these systems decreases as the data increases. In these models, too many parameters affect performance and it is tough to determine them [27]. Additional methods used to increase performance both slow down the system and require the addition of other solutions. This is both time and resource consuming. Despite all these requirements, performance cannot be increased for large sites [26]. In ensemble models, finding and optimizing the appropriate parameters of the methods is very difficult, which considerably reduces the forecasting performance [27].

However, almost all wind speed forecasting studies using ensemble models have considered time series data and common frequency data or mixed frequency data [28]. These data require a lot of pre-processing, which can lead to loss of information and poor forecasting results. Although researchers have achieved robust wind speed forecasting performance, there is still some research and study gaps in wind speed forecast using meteorological data in ensemble models. Meteorological data directly affect wind speed, contain useful information, are easy to obtain and process, and can provide significantly improved modeling performance. In previous studies in the literature, Neural Network based forecasting systems have been preferred due to their performance in wind speed forecasting for a limited number of sites. Prediction performance decreased significantly as the number of regions and the number of data increased. Therefore, introducing meteorological data into the field of wind speed prediction in ensemble models and creating a practical prediction method whose performance does not decrease as the number of sites and data increases is an important task to overcome the lack of existing studies on wind speed forecast.

In this paper, a novel ensemble wind speed forecasting model based on an artificial neural network is presented to forecast wind speed for five different sites in the Eastern Mediterranean Region of Turkey. In the proposed model, the prediction performance doesn't decrease with the increasing number of data, it uses meteorological data, works very fast thanks to its parallelism, takes advantage of the methods used in similar studies in the literature and eliminates their disadvantages. The model can accurately forecast average daily wind speed using temperature, humidity, pressure, and wind speed. Statistical error analyses methods such as mean absolute percentage error (MAPE), mean absolute error

(MAE) and, root mean square error (RMSE) are used to evaluate the forecasting performance of the network and to indicate the error levels.

The innovations and main contributions of this study can be summarized as follows;

(1) A new system is presented which has never been tried before in the literature and which utilizes the advantages of using ANN and eliminates the disadvantages of using ensemble methods for wind speed forecasting.

(2) Compared to ANN-based systems used in the literature, ANNs in the ensemble model are designed individually for each region in parallel, so a system that responds quickly and accurately despite the large amount of data and site is designed.

(3) Unlike the data used in other ensemble models in the literature, a system that performs with high accuracy using meteorological nonlinear data has been developed and a gap in this field has been filled.

(4) Since a different ANN model is designed for each site, the learning rate of each ANN is different, thus eliminating problems such as getting stuck in local solutions and overlearning, which are encountered in other studies using artificial neural networks and directly affecting the performance of the system.

(5) This system uses more practical data processing technology because of parallelized structure rather than other studies in the literature.

(6) The developed system can be used for any type of wind farm and it is a very useful and more practical system for smart energy management. This system will likely meet the needs of companies and large government facilities that invest and operate in renewable energy technologies.

This paper is organized as follows. Section I is the introduction of this paper. Section II presents the materials and methods used in the ensemble forecasting model provides details about the framework of the developed ensemble model. Section III introduces the statistical evaluations and related analyses used, and also includes some in-depth discussions. Section IV concludes of this study.

## II. MATERIALS AND METHODS

### A. THE STUDY REGION AND DESCRIPTION OF THE DATASET

Turkey is considered to be with a high wind energy potential country. There are some cities with high wind energy potential in many regions of Turkey and wind power plant investments have been going on in those regions. There are primarily seven different regions in Turkey. Aegean, Marmara, and Eastern Mediterranean regions in Turkey are regions with high wind energy potential [29]. The region located in the east of the Mediterranean region is called the Eastern Mediterranean Region. In this study, the wind speed was forecasted in 5 different sites in the Eastern Mediterranean Region which are Adana as Site 1, Antakya as Site 2, Kahramanmaras as Site 3, Iskenderun as Site 4 and Mersin as Site 5, as shown in Figure 1.


**FIGURE 1. Turkey map and selected sites of the eastern mediterranean region.**

The Mediterranean climate is dominant in the eastern Mediterranean region. Summers are generally hot and dry, and winters are warm and rainy. However, climatic features vary depending on altitude above sea level. An increase in terrestrial effects on the climate is observed in mountainous regions, but the intense continental climate is not observed in these regions due to the Mediterranean climate effect. Factors such as humidity, temperature, pressure and elevation affect wind speed [30]. In this study, the data set is created consisting of daily average humidity, temperature, pressure, and wind speed data obtained from Iskenderun Meteorology Directorate for the cities such as Adana, Antakya, Kahramanmaras, İskenderun, Mersin between 1998-2022. These data were used for the forecasting system, which was developed to forecast the daily average wind speed of the sites. In the methods used for forecasting in this study, daily average humidity, temperature, pressure data were used for input and daily average wind speed data were used for output.

Fig.2 shows boxplot figures of each site's data. As seen as in Fig. 2, boxplot figures show the minimum, median, maximum values and data quartiles of the data in each site used in this study.

### B. FORECASTING

In this study, wind speed forecasting was performed with two different methods. Firstly, wind speed forecasting was performed for 5 different sites with a multilayer feed-forward ANN model. Forecasting performances for different parameters of ANNs are compared. Secondly, the ensemble forecasting model developed for this study was used to forecast wind speed for 5 different sites. In order to determine the networks that found the best results in both methods, it was examined how accurately they responded to the verification data separated from the database. The networks with the highest accuracy rate were used.

#### 1) ARTIFICIAL NEURAL NETWORK

ANNs are an information processing technology used to solve complex problems in various applications such as control systems, pattern recognition [31], diagnosis, classification,

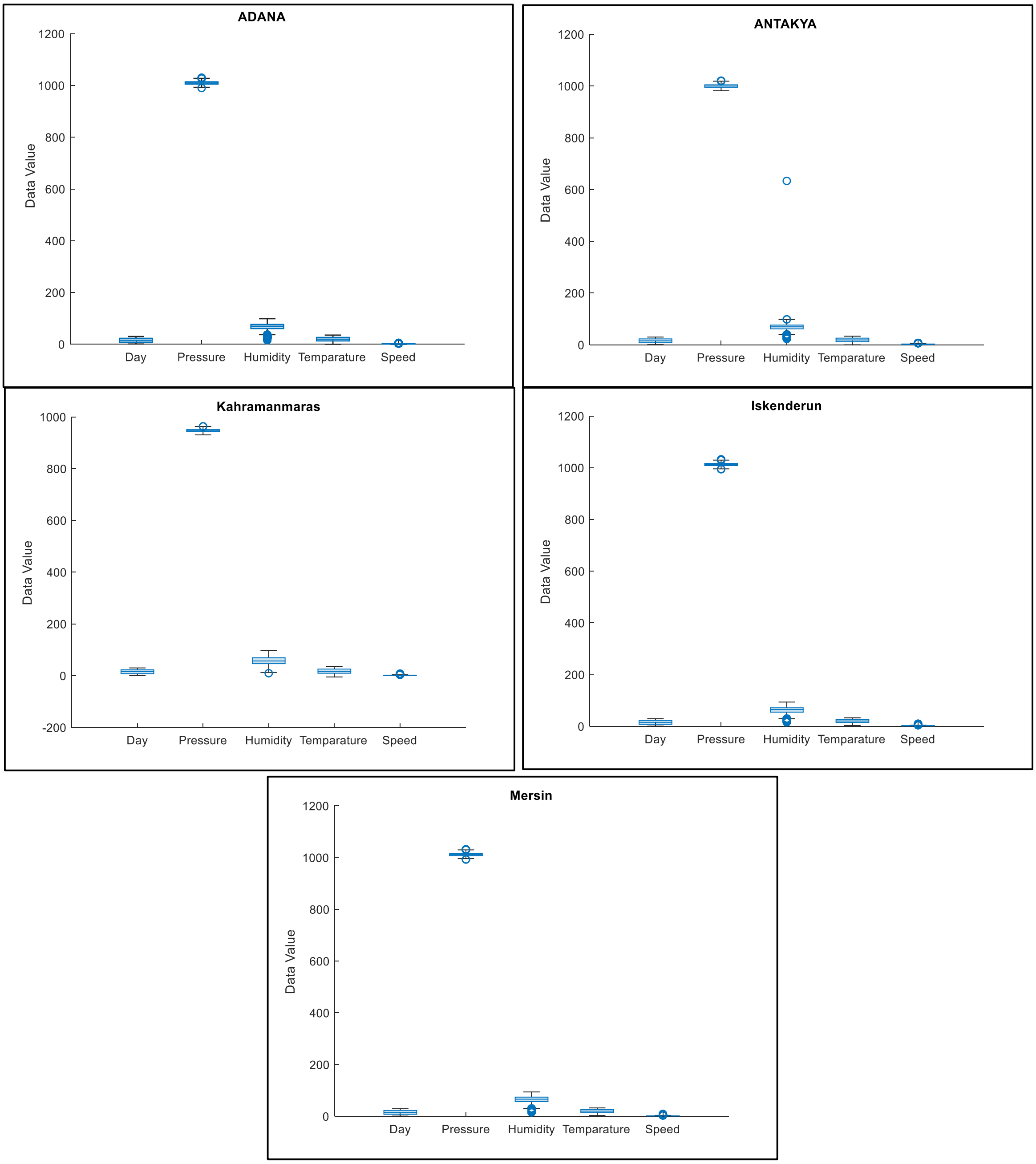



**FIGURE 2.** **Boxplot figures of each site.**

prediction, data association, data filtering, and interpretation [32]. The structure of ANN is based on the biological nervous system of the human brain. Neurons, the simplest and smallest unit of ANN, can manage complex behaviors between neurons and weight parameters. Neurons form networks by connecting the system to each other in various ways. These networks can learn, memorize and reveal the relationship between data for the predictions. ANN models consist of interconnected neurons. Artificial neurons have main components such as input, weight, threshold, summation function, activation function, and output [38].

In a basic ANN model, input $x_j$ is multiplied by a weight ($w_{jk}$) and a threshold value ($\emptyset_j$) is then added. For $n$ input values, the net input $u_k$ is calculated using (1);

$$u_k = \sum\nolimits_{k=1}^{n} w_{jk}x_k - \emptyset_j \tag{1}$$

Apply $u_k$ to a linear or non-linear activation function $f(x)$ to obtain the output $y_j$. The output of the network is calculated mathematically as given in (2)

$$y_j = f(\sum\nolimits_{k=1}^{n} w_{jk}x_k - \emptyset_j) \tag{2}$$

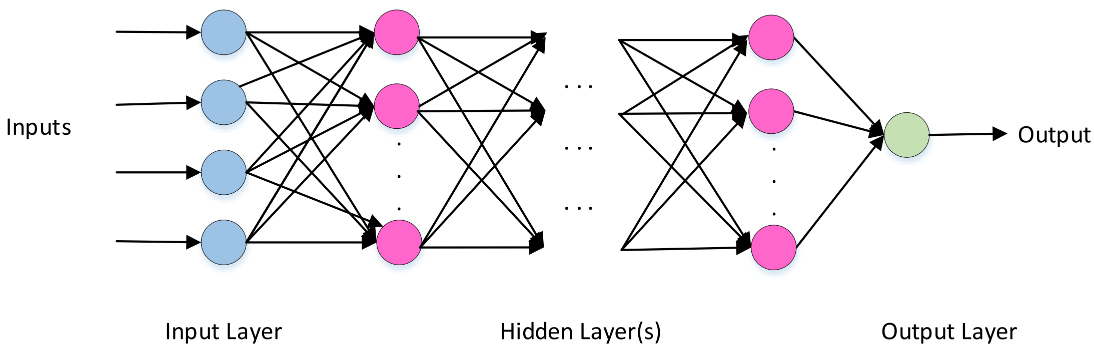



**FIGURE 3.** Schematic diagram of a multilayer feedforward neural network.

**TABLE 1.** The verification results for different ANN parameters.

| ANN Models | Training Algorithm | Transfer Function | The numbers of hidden layers/ numbers of neurons in hidden layers | Accuracy (%) Site 1/Site 2/Site 3/Site 4/ Site 5 | Average Accuracy (%) |
|---|---|---|---|---|---|
| ANN 1 | LM | Tangent sigmoid | 3/12,13,9 | 50.6/66.3/70.2/56.3/71.4 | 62.9 |
| ANN 2 | LM | Tangent sigmoid | 4/8,14,6,7 | 67.3/70.6/73.2/80.3/76.3 | 73.5 |
| ANN 3 | LM | Hyperbolic Tangent Sigmoid | 2/22,18 | 50.3/64.4/69.2/70.4/63.4 | 63.5 |
| ANN 4 | LM | **Hyperbolic tangent sigmoid** | 5/8,12,18,10,20 | 82.3/85.2/88.3/77.5/86.1 | 83.8 |
| ANN 5 | SCG | Tangent Sigmoid | 2/15,23 | 68.2/75.5/67.3/58.7/65.6 | 67.1 |
| ANN 6 | SCG | Tangent Sigmoid | 3/7,13,14 | 63.2/70.3/74.8/65.7/82.3 | 71.3 |
| ANN 7 | SCG | Hyperbolic tangent Sigmoid | 4/19,12,20,13 | 67.3/82.2/69.7/58.7/83.5 | 72.3 |
| ANN 8 | SCG | Hyperbolic tangent Sigmoid | 5/11,8,14,6,23 | 73.6/85.7/68.4/60.2/74.3 | 72.4 |

In Eq. 2, $n$ is the number of entries, $w_{jk}$ is the weight value of the $k$ input of neuron $j$, $\emptyset_j$ is the applied threshold. Output $y_j$ is compared with the targeted value $e_j$ using the error shown in (3);

$$\delta_j = y_j(1 - y_j)(e_j - y_j) \tag{3}$$

Training the ANN is one of the most important processes. Training the ANN is described properly since the process of modification of the connection weights is necessary in a specific order using a learning algorithm. The purpose of training in artificial neurons is to compile the values of the weights of the inputs of each neuron. The weights are mainly random before the network is trained and do not contain any meaning, but after training, they contain meaningful and satisfactory information [39].

One of the most used algorithms for training is the back-propagation algorithm. In this algorithm, the inputs and desired outputs are implemented to the network and the output is obtained. Then, the outputs of the network and the desired outputs are compared to each other and the error is found. This process is implemented to all data in the training dataset and this training proceeding, called an epoch, is iterated until the error level decreases to an admissible value [40]. The error is reflected back to the neurons in the hidden layers and the performance of the model is increased. Briefly, the error is disseminated back from the output layer neurons to the input layers, neurons, as is given in (4);

$$\mathrm{w_{jk}}\ (\mathrm{t}+1) = \mathrm{w_{jk}}\ (\mathrm{t}) + \delta_{\mathrm{j}}\mathrm{y_j}\alpha(\mathrm{w_{jk}}\ (\mathrm{t}) - \mathrm{w_{jk}}\ (\mathrm{t}-1)) \tag{4}$$

where,

| | |
|---|---|
| $w_{jk}\ (t-1)$: | previous weight of the $k_{th}$ neuron |
| $w_{jk}\ (t+1)$ : | updated weight of the $k_{th}$ neuron |
| $w_{jk}\ (t)$: | current weight of the $k_{th}$ neuron |
| $\alpha$: | learning rate of the network |
| $\delta_j$: | error |
| $y_j$: | output |

In this study, a multilayer feedforward ANN model is used. The basic topology of multilayer feedforward ANN is the input layer, hidden layer(s), and output layer [43] as seen in Figure 3. The ANN was trained and tested on a

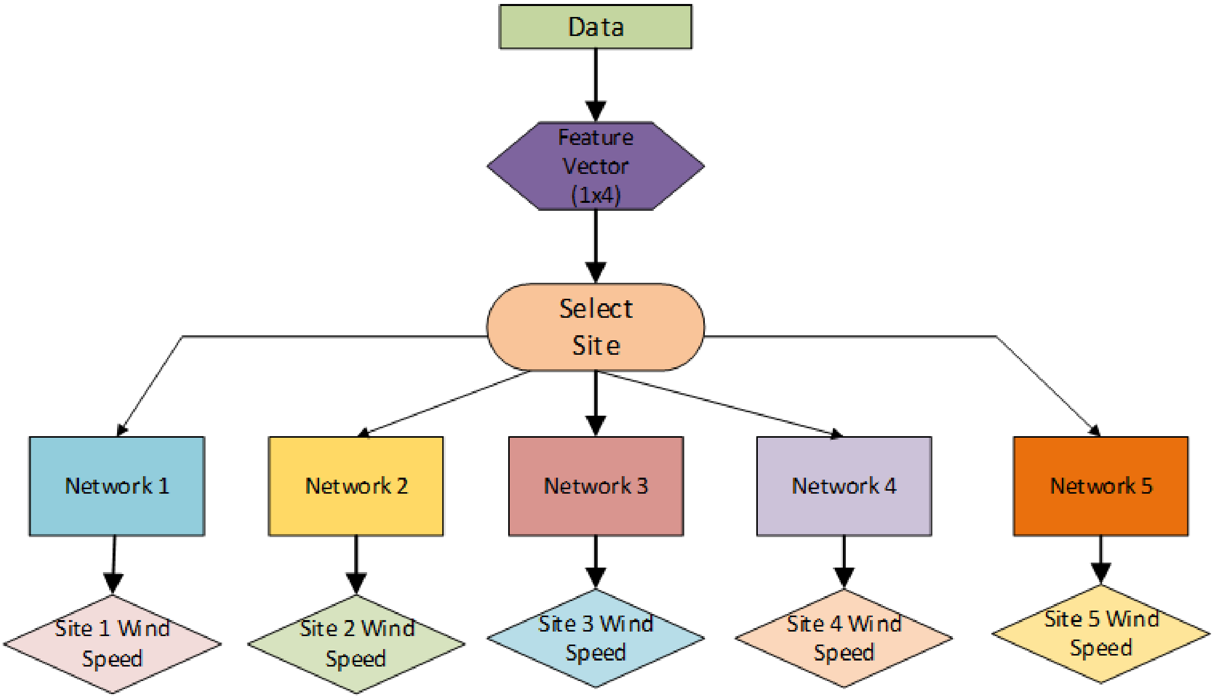



**FIGURE 4.** **Overall framework of the developed model.**

total of 6048 out of 8640 data in the database. 70% of the 6048 data in the database were used for training ANN and 30% for verification. Different backpropagation algorithms are used to train the ANN. In the developed ANN model, supervised learning technique with multilayer feed forward network with Scaled Conjugate Gradient (SCG), Levenberg-Marquard (LM) learning algorithms are employed. Tangent sigmoid (tansig) and hyperbolic tangent sigmoid transfer functions used respectively. The accuracy of the model depends on the number of hidden layer, the number of neurons in the hidden layer, the transfer function and the learning algorithm [41].

Many studies have shown that the training process of neural networks slows down when a large number of hidden layers are used [45]. Researchers have concluded that if the criterion of the problem is to achieve better accuracy, then a large number of hidden layers can be used, but if the main factor of a problem is time, then the use of a large number of hidden layers should be considered in such problems [46]. On the other hand, adding unnecessary neurons or layers leads to overfitting. Therefore, before designing the neural network, it is very important to analyze the training database samples to correctly determine the approximate number of neurons and hidden layers.

MATLAB R2023A software is employed to implement the presented model.

The verification results are given in Table 1 for different ANN parameters. Best accuracy results were developed with ANN 4 model. In this ANN is used LM learning algorithm and hyperbolic tangent sigmoid transfer function. There are 5 hidden layers, the number of neurons in the 1st, 2nd, 3rd, 4th and 5th hidden layers are 8, 12, 18, 10 and 20 respectively. The learning rate is 0.4, the number of epochs for training the network is 169. After several trial-and-error experiments, the training algorithm, the transfer function, the learning rate, the number of hidden layers, and the number of neurons in the hidden layers were selected. The configuration that gave the highest accuracy was selected. As seen in Table 1 wind speed forecasting accuracy results are low. In order to increase the forecasting rate, forecasting model is changed and developed a new ensemble forecasting model.

### 2) THE ENSEMBLE FORECASTING MODEL

Figure 4 shows the ensemble forecasting model developed in this study. Each network represents a different ANN which forecasts one site. By combining these networks, ensemble forecasting model is established. This system has five networks to forecast wind speed: Site 1 Adana, Site 2 Antakya, Site 3 Kahramanmaras, Site 4 Iskenderun, and Site 5 Mersin. Each network of the model represents a specific ANN model which is trained to forecast daily wind speed for one site.

To test any data, 4 feature information such as the day information and daily average humidity, temperature and pressure obtained from the test data are entered into the system. Then the desired site is selected. Test data is applied as network input to the selected site. After these processes, the test data of the selected network in Figure 4 is tested and the wind speed for the selected day in the relevant site is forecasted.

### a: TRAINING AND TESTING FOR EACH ANN IN THE ENSEMBLE MODEL

For each site there are 8640 data collected between 1998 and 2022. For 5 sites there is a total of 43200 data in the database. The ANN to predict wind speed for each site was trained and tested using a total of 6048 data from the 8640 data in the database. For each site, 70% of the 6048 data were used for training and 30% for validation. Each site was trained and tested with its own training and test data.

MATLAB R2023A software is used for the ensemble model. For all network structure in this system, feedforward ANN model is used. Levenberg – Marquardt (LM) algorithm was used for training, and the transfer function, the hyperbolic tangent sigmoid was used.

After several trial-and-error experiments, the training algorithm, the transfer function, the learning rate, the number of hidden layers were selected. Also, the number of neurons in the hidden layers were selected based on the trial-and-error approach. In this approach, each network in the model was trained with different training algorithms, transfer functions, learning rates, different number of hidden layers and different numbers of neurons [42]. Then, the performances of the networks were obtained by using different training algorithms, transfer functions, learning rates, different numbers of hidden layers and different numbers of neurons in hidden layers. For each network in the ensemble model, the configuration that gave the highest accuracy was selected.

Figure 5 shows how each ANN in the ensemble model was designed. Each network in the ensemble model has different configurations. Each site represented by the networks has different geographical characteristics, even if they are in the same region, and each has different meteorological data accordingly. With this model developed, it is possible to design different network configurations according to the specific data and characteristics of each site. The aim is to improve ANN prediction performance and thus the accuracy of wind speed forecasting. Structures are carried out step by step for each network, as described below.

*Network 1:* In this network of the ensemble model, an ANN forecasts wind speed for the input data of Site 1. In the ANN, there are 4 hidden layers, the number of neurons in the 1st, 2nd, 3rd, and 4th hidden layers are 13,10, 9 and 15 respectively. The learning rate is 0.3, the number of epochs for training the network is 128. The verification accuracy rate for this network as given in Table 2 shows that the performance of this network for forecasting wind speed for Site 1 is high.

*Network 2:* In this network, an ANN is established to forecasts wind speed for Site 2. For this ANN, there are 3 hidden layers, the number of neurons in the 1st, 2nd, and 3rd hidden layers are 18,12, and 12 respectively. The learning rate is 0.2, the number of epochs for training the network is 105.The verification accuracy rate for this step as given in Table 2 shows that the performance of this network for forecasting wind speed for Site 2 is high.

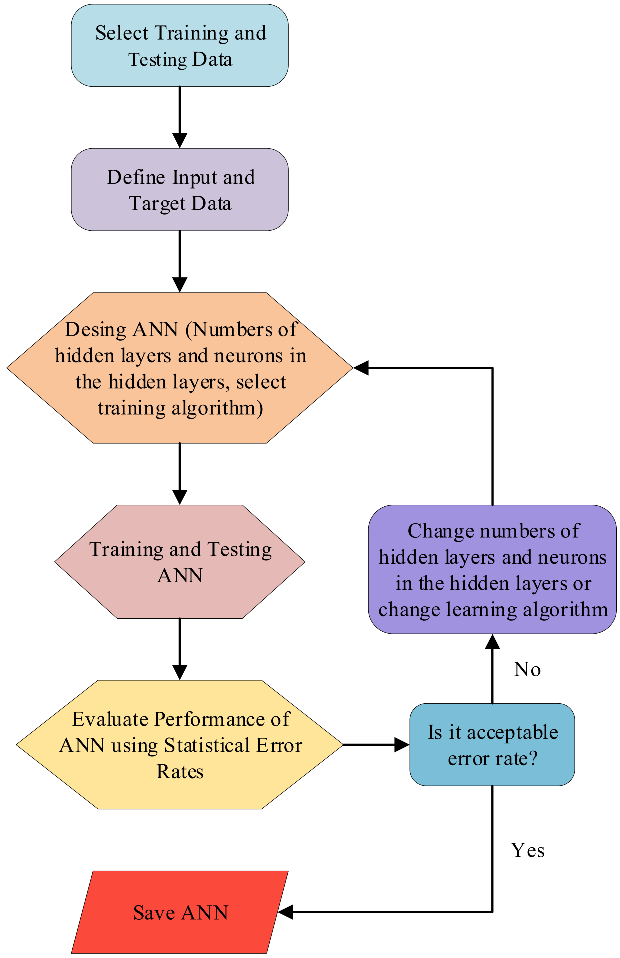



**FIGURE 5. Flowchart of the designed each ANN in the ensemble model.**

*Network 3:* The ANN in this step, forecasts wind speed for Site 3. For the ANN structure, there are 4 hidden layers, the number of neurons in the 1st, 2nd, 3rd, and 4th hidden layers are 22,18, 19 and 18 respectively. The learning rate is 0.2, the number of epochs for training the network is 97. The verification results for this step as given in Table 2 demonstrates that this ANN model has a high forecasting performance for Site 3.

*Network 4:* The ANN forecasts wind speed of Site 4. For this ANN, there are 4 hidden layers, the number of neurons in the 1st, 2nd, 3rd, and 4th hidden layers are 25, 15, 23, 7 respectively. The learning rate is 0.4, the number of epochs for training the network is 159. The accuracy rate given in Table 2 indicates that this ANN model forecasts wind speed successfully for Site 4.

*Network 5:* The ANN structure is used to forecast wind speed of Site 5. The ANN structure has 3 hidden layers. The first, the second and the third hidden layers include 12, 16 and 19 neurons, respectively. The learning rate is 0.3, the number of epochs for train the network is 102. The forecasting

**TABLE 2.** The verification results for ANN models in the ensemble model.

| Network | Training Algorithm | Transfer Function | The numbers of hidden layers/ numbers of neurons in hidden layers | Accuracy (%) |
|---|---|---|---|---|
| Network 1 | LM | Hyperbolic tangent sigmoid | 4/13,10,9,15 | 98.3 |
| Network 2 | LM | Hyperbolic tangent sigmoid | 3/18,15,12 | 98.2 |
| Network 3 | LM | Hyperbolic tangent sigmoid | 4/22,18,19,18 | 98.5 |
| Network 4 | LM | Hyperbolic tangent sigmoid | 4/25,15,23,7 | 97.6 |
| Network 5 | LM | Hyperbolic tangent sigmoid | 3/12,16,19 | 98.7 |
| | | | Average | 98.3 |

performance results given in Table 2 shows that this ANN model forecasts wind speed successfully for Site 5.

### C. EVALUATION METRICS

Statistical error analysis methods were used to evaluate the performance of the network and to indicate the error. *Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and, Root Mean Square Error (RMSE)* analyses are used to evaluate the results of ANN and the ensemble model. These analysis methods are calculated using in 5-7;

$$\text{MAE} = \frac{\sum_{t=1}^{n} \left|(x_t - x_t')\right|^2}{n} \tag{5}$$

$$\text{MAPE}\,(\%) = \frac{\sum_{t=1}^{n} \left|(x_t - x_t')/x_t\right|}{n} \times 100 \tag{6}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{n} (x_t - x_t')^2}{n}} \tag{7}$$

In the above equations, $x_t$ represents the output value and $x_t'$ represents the target output value.

## III. RESULTS AND DISCUSSION

This section introduces the statistical analysis of ANN and the ensemble model at five sites. The evaluation criteria for forecasting *MAE, MAPE* and *RMSE* are used. The *MAE, MAPE, RMSE* values calculated to determine the performance of the developed ensemble model represent the average absolute difference between the predicted and measured data, the average absolute percentage difference between the predicted and measured data, and the difference between the predicted and measured data, respectively. From the database, 2592 data were selected for each site from 8640 data to test the ANN and the ensemble model. These 2592 data were randomly selected from the 8640 data in the database. The 2592 data were not used for training any forecasting structure.

### A. EXPERIMENT 1: ARTIFICIAL NEURAL NETWORK

The developed different ANN models which details were presented in section II/1, were tested. *MAPE, MAE* and *RMSE* statistical analysis methods presented in Section II-C were used to evaluate the accuracy of ANN models. The *MAE, MAPE* (%) and *RMSE* values obtained from different ANN models developed for 5 different sites are illustrated in Table 3.

As can be seen from Table 3, the best results were obtained for ANN4. ANN 4 model also gave the highest accuracy rate in the validation result. The obtained test results proved that this model is the ANN model with the best prediction performance among the other models. *MAE* values ranged between 0.2426 and 0.3982. *MAPE (%)* values vary between 20.93 and 28.23, while *RMSE* values vary between 0.2774 and 0.3432. The *MAE, MAPE* and *RMSE* results obtained show that the forecasting performance of this model is not at the desired performance. These test results show that the forecasting performance of ANN for wind forecasting is low.

The results obtained with ANN revealed that a different method should be developed to increase the prediction performance as the number of sites and the number of data increases.

### B. EXPERIMENT II: THE ENSEMBLE SYSTEM

Detailed test results of the developed ensemble system are presented in Table 4. As can be seen from Table 4, Site 1 has the best MAE value and the *MAE* value is 0.1131. In general, it can be seen that the *MAE* values of all the sites are less than 0.2 and that the values are very close to each other. *MAE* results are within the desired range and very good for all sites. The average *MAE* value is 0.1269 and this result confirms that the *MAE* values of the other sites are quite good.

The site with the best *MAPE* value is Site 5 with a *MAPE* value of 2.5318%. When analyzing the *MAPE* values of the other sites, it can be seen that they all have *MAPE* values of

**TABLE 3.** Statistical metrics of test results for different ANN Models.

| Model | Evaluation Metrics | Site 1 | Site 2 | Site 3 | Site 4 | Site 5 |
|---|---|---|---|---|---|---|
| ANN1 | MAE | 0.4812 | 0.6845 | 0.4924 | 0.5927 | 0.4337 |
| | MAPE (%) | 50.83 | 30.01 | 53.71 | 36.15 | 22.85 |
| | RMSE | 0.6413 | 0.49 | 0.69 | 0.6123 | 0.5951 |
| ANN 2 | MAE | 0.5141 | 0.6014 | 0.4718 | 0.6825 | 0.4431 |
| | MAPE (%) | 60.23 | 30.93 | 50.53 | 34.58 | 23.15 |
| | RMSE | 0.6178 | 0.5142 | 0.6781 | 0.6247 | 0.6014 |
| ANN 3 | MAE | 0.5013 | 0.6224 | 0.4819 | 0.4985 | 0.4654 |
| | MAPE (%) | 59.85 | 30.84 | 56.83 | 35.11 | 23.78 |
| | RMSE | 0.6517 | 0.5434 | 0.6913 | 0.5142 | 0.6214 |
| ANN 4 | MAE | 0.3982 | 0.3756 | 0.2873 | 0.3017 | 0.2426 |
| | MAPE (%) | 28.23 | 27.98 | 25.53 | 25.85 | 20.93 |
| | RMSE | 0.3432 | 0.3016 | 0.2917 | **0.2924** | 0.2774 |
| ANN 5 | MAE | 0.4837 | 0.6924 | 0.4771 | 0.6013 | 0.4557 |
| | MAPE (%) | 60.15 | 30.38 | 56.33 | 36.27 | 24.25 |
| | RMSE | 0.6312 | 0.4874 | 0.6856 | 0.6134 | 0.4028 |
| ANN 6 | MAE | 0.5054 | 0.6814 | 0.5112 | 0.6123 | 0.4856 |
| | MAPE (%) | 61.32 | 30.65 | 60.38 | 37.24 | 23.97 |
| | RMSE | 0.6553 | 0.5822 | 0.7014 | 0.6334 | 0.5478 |
| ANN 7 | MAE | 0.4921 | 0.6423 | 0.4874 | 0.6762 | 0.4687 |
| | MAPE (%) | 59.74 | 31.38 | 55.82 | 33.67 | 23.63 |
| | RMSE | 0.6254 | 0.5787 | 0.6925 | 0.6716 | 0.6127 |
| ANN 8 | MAE | 0.5123 | 0.4716 | 0.5028 | 0.6425 | 0.4557 |
| | MAPE (%) | 60.53 | 29.71 | 50.47 | 33.98 | 23.62 |
| | RMSE | 0.6612 | 0.5837 | 0.7071 | 0.6256 | 0.5327 |

**TABLE 4.** Statistical metrics of test results for proposed ensemble model.

| Site | MAE | MAPE(%) | RMSE |
|---|---|---|---|
| Site 1 | 0.1131 | 3.2143 | 0.1488 |
| Site 2 | 0.1269 | 2.9817 | 0.1333 |
| Site 3 | 0.1429 | 3.4211 | 0.1815 |
| Site 4 | 0.1318 | 3.2217 | 0.1930 |
| Site 5 | 0.1199 | 2.5318 | 0.1418 |
| Avarage | 0.1269 | 3.0741 | 0.1596 |

less than 4%. The average MAPE value was 3.0741%. *MAPE* results confirm that the model forecasts very well.

The site with the best *RMSE* value is Site 2 and the *RMSE* value is 0.1333. The *RMSE* values of the other sites vary between 0.1418 and 0.1930. The average *RMSE* value is 0.1596. The *RMSE* values confirm that the developed model forecasts quite accurately for all sites.

The *MAE*, *MAPE* and *RMSE* values calculated to evaluate the forecasting performance of the developed model were obtained at desired values for all sites. The values show that the developed model forecasts the wind speed with very high accuracy.

The comparison of measured and predicted values for each Site is shown in Figure 6. In Fig. 6 the superiority of the proposed system is clear. The forecasting results of the ensemble model and the actual wind speed values are very close to each other. The results show that the use of the ensemble model for wind speed forecasting is an appropriate choice.

The proposed ensemble model forecasts wind speed accurately and practically under different geographical and climatic conditions. This is because a different ANN model has been developed for each site. The ANN model

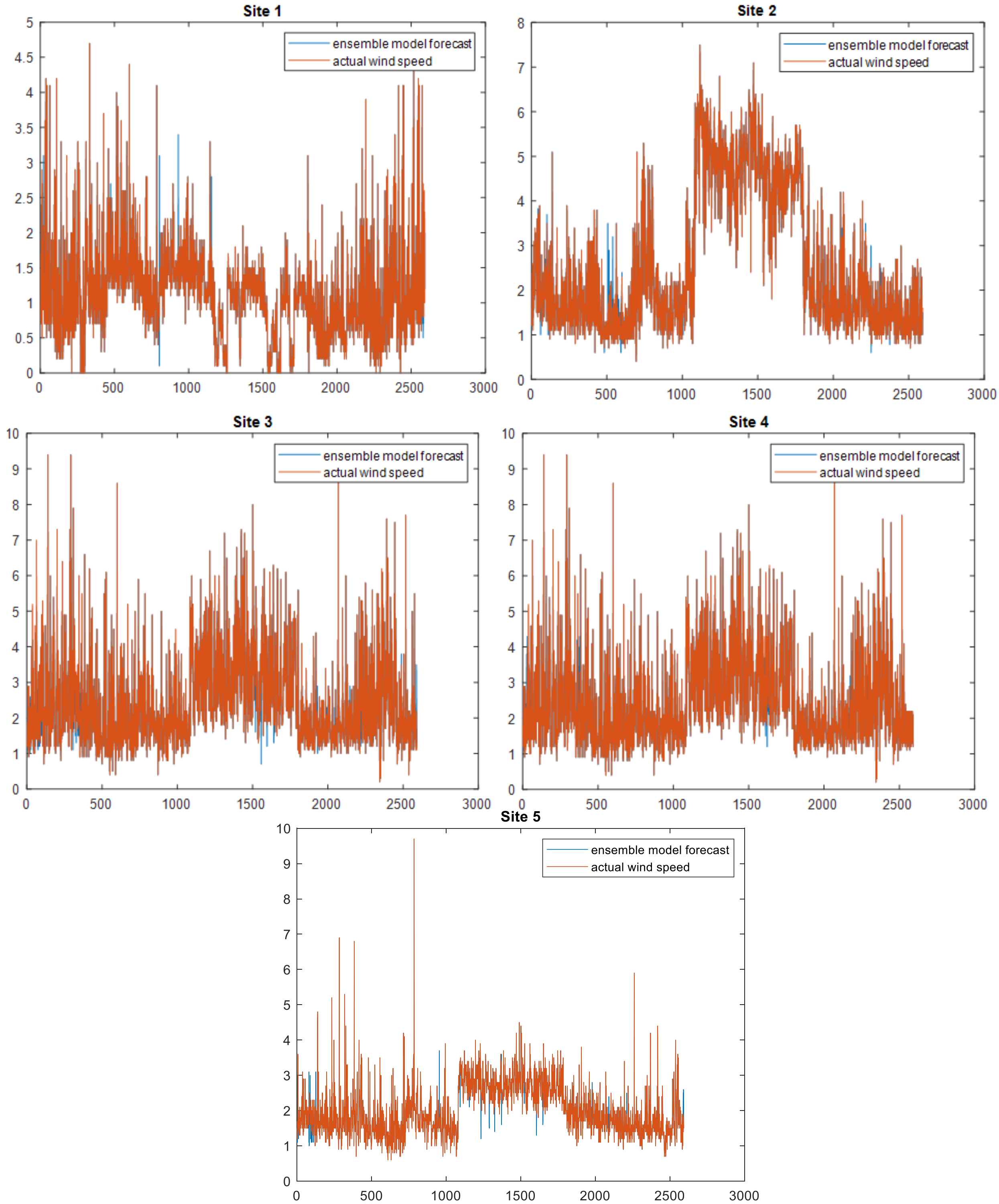



**FIGURE 6.** **Forecasting results of the developed system.**

with the highest accuracy rate was obtained for each site and this strengthened the reliability of the model. The developed model overcomes a lot of non-linear and complex data and forecasts the wind speed with high accuracy.

## IV. CONCLUSION

Due to technological advances in many sectors and population growth around the world, the need for affordable and sustainable energy sources is increasing significantly. Among other renewable energy options, wind energy is the most advanced and commercially available type of energy. Accurate and stable wind speed forecasting is vital for efficient energy management and the market economy.

Many existing wind speed forecasting models use persistence and physical methods. These methods need large data, which slows down and complicates the functioning of the system. As a result, the accuracy of the forecasting system decreases. The forecasting performance of statistical models is poor for nonlinear data. Heuristic and ensemble models, on the other hand, have the disadvantages of needing too many training examples, requiring preprocessing for a large number of examples, getting stuck in local solutions as the number of examples increases, overlearning, and slowing down the system. In previous studies, neural network based forecasting systems have been preferred due to their performance in forecasting wind speed for a limited number of sites and data. In this study, it was developed a unique ensemble forecasting model to increase the accuracy of wind speed forecast. Raw, random and non-linear meteorological data such as wind speed, humidity, temperature, and pressure were used in this model. In contrast, the developed model doesn't need preprocessing the data, this increases the speed of the system considerably. Unlike the methods mentioned above, the model developed in this study does not get stuck in local solutions since different ANN models are combined in parallel, although it works with a large number of non-linear data and more sites. Contrary to other studies using ANN and also the results obtained in this study, the performance of the system does not decrease as the number of sites and the number of data increases. In ensemble models, it is very difficult to find and optimize the appropriate parameters of the methods, which considerably reduces the forecast performance. In this study, since the parameter optimization is performed separately for each site, the prediction performance is considerably improved.

Each of the components of the developed ensemble system has significantly improved the forecast performance. The contribution made by each component is verified using test data. The developed system performed very well, demonstrating excellent wind speed forecasting capability due to the superiority of each network. The obtained results show that this system is an effective and useful intelligent tool that can be used by various companies and government facilities that invest and operate in intelligent wind energy technologies. Moreover, the developed ensemble system can be used in other fields such as electricity price forecasting, load forecasting and power plant scheduling.

## REFERENCES

[1] F. G. Montoya, F. Manzano-Agugliaro, S. López-Márquez, Q. Hernández-Escobedo, and C. Gil, "Wind turbine selection for wind farm layout using multi-objective evolutionary algorithms," *Expert Syst. Appl.*, vol. 41, no. 15, pp. 6585–6595, Nov. 2014, doi: 10.1016/j.eswa.2014.04.044.

[2] F. Manzano-Agugliaro, A. Alcayde, F. G. Montoya, A. Zapata-Sierra, and C. Gil, "Scientific production of renewable energies worldwide: An overview," *Renew. Sustain. Energy Rev.*, vol. 18, pp. 134–143, Feb. 2013, doi: 10.1016/j.rser.2012.10.020.

[3] N. Ullah, I. Sami, Md. S. Chowdhury, K. Techato, and H. I. Alkhammash, "Artificial intelligence integrated fractional order control of doubly fed induction generator-based wind energy system," *IEEE Access*, vol. 9, pp. 5734–5748, 2021, doi: 10.1109/ACCESS.2020.3048420.

[4] E. Espina, J. Llanos, C. Burgos-Mellado, R. Cárdenas-Dobson, M. Martínez-Gómez, and D. Sáez, "Distributed control strategies for microgrids: An overview," *IEEE Access*, vol. 8, pp. 193412–193448, 2020, doi: 10.1109/ACCESS.2020.3032378.

[5] J. Ahmad, M. Tahir, and S. K. Mazumder, "Improved dynamic performance and hierarchical energy management of microgrids with energy routing," *IEEE Trans. Ind. Informat.*, vol. 15, no. 6, pp. 3218–3229, Jun. 2019, doi: 10.1109/TII.2018.2877739.

[6] A. A. Alsakati, C. A. Vaithilingam, J. Alnasseir, K. Naidu, and G. Rajendran, "Transient stability enhancement of grid integrated wind energy using particle swarm optimization based multi-band PSS4C," *IEEE Access*, vol. 10, pp. 20860–20874, 2022, doi: 10.1109/ACCESS.2022.3151425.

[7] R. Carli, M. Dotoli, R. Pellegrino, and L. Ranieri, "A decision making technique to optimize a buildings' stock energy efficiency," *IEEE Trans. Syst. Man, Cybern. Syst.*, vol. 47, no. 5, pp. 794–807, May 2017.

[8] C.-J. Huang and P.-H. Kuo, "A short-term wind speed forecasting model by using artificial neural networks with stochastic optimization for renewable energy systems," *Energies*, vol. 11, no. 10, p. 2777, Oct. 2018, doi: 10.3390/en11102777.

[9] A. Ucar and F. Balo, "Investigation of wind characteristics and assessment of wind-generation potentiality in uludağ-bursa, Turkey," *Appl. Energy*, vol. 86, no. 3, pp. 333–339, Mar. 2009, doi: 10.1016/j.apenergy.2008.05.001.

[10] A. Khosravi, R. N. N. Koury, L. Machado, and J. J. G. Pabon, "Prediction of wind speed and wind direction using artificial neural network, support vector regression and adaptive neuro-fuzzy inference system," *Sustain. Energy Technol. Assessments*, vol. 25, pp. 146–160, Feb. 2018, doi: 10.1016/j.seta.2018.01.001.

[11] P. Pinson, H. A. Nielsen, H. Madsen, and G. Kariniotakis, "Skill forecasting from ensemble predictions of wind power," *Appl. Energy*, vol. 86, nos. 7–8, pp. 1326–1334, Jul. 2009, doi: 10.1016/j.apenergy.2008.10.009.

[12] A. Costa, A. Crespo, J. Navarro, G. Lizcano, H. Madsen, and E. Feitosa, "A review on the young history of the wind power short-term prediction," *Renew. Sustain. Energy Rev.*, vol. 12, no. 6, pp. 1725–1744, Aug. 2008, doi: 10.1016/j.rser.2007.01.015.

[13] D. A Fadare, "The application of artificial neural networks to mapping of wind speed profile for energy application in Nigeria," *Appl. Energy*, vol. 87, no. 3, pp. 934–942, 2010, doi: 10.1016/j.apenergy.2009.09.005.

[14] J. M. Morales, R. Mínguez, and A. J. Conejo, "A methodology to generate statistically dependent wind speed scenarios," *Appl. Energy*, vol. 87, no. 3, pp. 843–855, Mar. 2010, doi: 10.1016/j.apenergy.2009.09.022.

[15] Shivani, K. S. Sandhu, and A. Ramachandran Nair, "A comparative study of ARIMA and RNN for short term wind speed forecasting," in *Proc. 10th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jul. 2019, pp. 1–7.

[16] J. Wang, J. Hu, K. Ma, and Y. Zhang, "A self-adaptive hybrid approach for wind speed forecasting," *Renew. Energy*, vol. 78, pp. 374–385, Jun. 2015, doi: 10.1016/j.renene.2014.12.074.

[17] W. Yang, Z. Tian, and Y. Hao, "A novel ensemble model based on artificial intelligence and mixed-frequency techniques for wind speed forecasting," *Energy Convers. Manage.*, vol. 252, Jan. 2022, Art. no. 115086, doi: 10.1016/j.enconman.2021.115086.

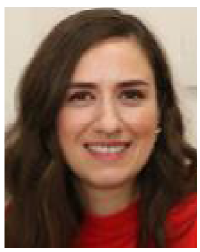
[18] J.-Z. Wang and Y. Wang, "A novel wind speed forecasting model for wind farms of Northwest China," *Int. J. Green Energy*, vol. 14, no. 5, pp. 463–478, Apr. 2017, doi: 10.1080/15435075.2016.1278373.

[19] H. Zhang, X. Wang, and Y. Wu, "Application of an improved SVM algorithm for wind speed forecasting," in *Future Intelligent Information Systems* (Lecture Notes in Electrical Engineering), vol. 86, D. Zeng, Ed., Berlin, Germany, 2011, pp. 333–340, doi: 10.1007/978-3-642-19706-2_43.

[20] M. A. Mohandes, T. O. Halawani, S. Rehman, and A. A. Hussain, "Support vector machines for wind speed prediction," *Renew. Energy*, vol. 29, no. 6, pp. 939–947, May 2004, doi: 10.1016/j.renene.2003.11.009.

[21] E. Çam, E. Arcaklıoğlu, A. Çavuşoğlu, and B. Akbıyık, "A classification mechanism for determining average wind speed and power in several regions of Turkey using artificial neural networks," *Renew. Energy*, vol. 30, no. 2, pp. 227–239, Feb. 2005.

[22] X. Liu, H. Zhang, X. Kong, and K. Y. Lee, "Wind speed forecasting using deep neural network with feature selection," *Neurocomputing*, vol. 397, pp. 393–403, Jul. 2020, doi: 10.1016/j.neucom.2019.08.108.
[23] T. Blanchard and B. Samanta, "Wind speed forecasting using neural networks," *Wind Eng.*, vol. 44, no. 1, pp. 33–48, Feb. 2020, doi: 10.1177/0309524x19849846.
[24] S. Wang, J. Wang, H. Lu, and W. Zhao, "A novel combined model for wind speed prediction—Combination of linear model, shallow neural networks, and deep learning approaches," *Energy*, vol. 234, Nov. 2021, Art. no. 121275, doi: 10.1016/j.energy.2021.121275.
[25] X. Niu and J. Wang, "A combined model based on data preprocessing strategy and multi-objective optimization algorithm for short-term wind speed forecasting," *Appl. Energy*, vol. 241, pp. 519–539, May 2019, doi: 10.1016/j.apenergy.2019.03.097.
[26] L. Luo, H. Li, J. Wang, and J. Hu, "Design of a combined wind speed forecasting system based on decomposition-ensemble and multi-objective optimization approach," *Appl. Math. Model.*, vol. 89, pp. 49–72, Jan. 2021.
[27] Z. Zhang, J. Wang, D. Wei, T. Luo, and Y. Xia, "A novel ensemble system for short-term wind speed forecasting based on two-stage attention-based recurrent neural network," *Renew. Energy*, vol. 204, pp. 11–23, Mar. 2023, doi: 10.1016/j.renene.2022.12.120.
[28] W. Yang, M. Hao, and Y. Hao, "Innovative ensemble system based on mixed frequency modeling for wind speed point and interval forecasting," *Inf. Sci.*, vol. 622, pp. 560–586, Apr. 2023, doi: 10.1016/j.ins.2022.11.145.
[29] B. Sahin and M. Bilgili, "Wind characteristics and energy potential in Belen-Hatay, Turkey," *Int. J. Green Energy*, vol. 6, no. 2, pp. 157–172, Apr. 2009, doi: 10.1080/15435070902784947.
[30] Z. Liu, J. Ping, L. Zhang, and X. Niu, "A combined forecasting model for time series: Application to short-term wind speed forecasting," *Appl. Energy*, vol. 259, Feb. 2020, Art. no. 114137, doi: 10.1016/j.apenergy.2019.114137.
[31] M. E. Özdemir, Z. Telatar, O. Eroğul, and Y. Tunca, "Classifying dysmorphic syndromes by using artificial neural network based hierarchical decision tree," *Australas. Phys. Eng. Sci. Med.*, vol. 41, no. 2, pp. 451–461, Jun. 2018, doi: 10.1007/s13246-018-0643-x.
[32] M. E. Ozdemir, Z. Ali, B. Subeshan, and E. Asmatulu, "Applying machine learning approach in recycling," *J. Mater. Cycles Waste Manage.*, vol. 23, no. 3, pp. 855–871, May 2021, doi: 10.1007/s10163-021-01182-y.
[33] Y. Liu, L. Wang, K. Gu, and M. Li, "Artificial neural network (ANN)—Bayesian probability framework (BPF) based method of dynamic force reconstruction under multi-source uncertainties," *Knowledge-Based Syst.*, vol. 237, Feb. 2022, Art. no. 107796, doi: 10.1016/j.knosys.2021.107796.
[34] J. Aguila-Leon, C. Vargas-Salgado, C. Chiñas-Palacios, and D. Díaz-Bello, "Energy management model for a standalone hybrid microgrid through a particle swarm optimization and artificial neural networks approach," *Energy Convers. Manage.*, vol. 267, Sep. 2022, Art. no. 115920, doi: 10.1016/j.enconman.2022.115920.
[35] W.-J. Chang, C.-L. Su, and C.-C. Ku, "Passive decentralized fuzzy control for Takagi–Sugeno fuzzy model based large-scale descriptor systems," *IEEE Access*, vol. 10, pp. 28656–28669, 2022, doi: 10.1109/ACCESS.2022.3158671.
[36] J. Zhao, Z. Guo, Y. Guo, W. Lin, and W. Zhu, "A self-organizing forecast of day-ahead wind speed: Selective ensemble strategy based on numerical weather predictions," *Energy*, vol. 218, Mar. 2021, Art. no. 119509, doi: 10.1016/j.energy.2020.119509.
[37] M. U. Yousuf, I. Al-Bahadly, and E. Avci, "Current perspective on the accuracy of deterministic wind speed and power forecasting," *IEEE Access*, vol. 7, pp. 159547–159564, 2019, doi: 10.1109/ACCESS.2019.2951153.
[38] M. G. M. Abdolrasol, S. M. S. Hussain, T. S. Ustun, M. R. Sarker, M. A. Hannan, R. Mohamed, J. A. Ali, S. Mekhilef, and A. Milad, "Artificial neural networks based optimization techniques: A review," *Electronics*, vol. 10, no. 21, p. 2689, Nov. 2021, doi: 10.3390/electronics10212689.
[39] A. D. Rasamoelina, F. Adjailia, and P. Sincák, "A review of activation function for artificial neural network," in *Proc. IEEE 18th World Symp. Appl. Mach. Intell. Informat. (SAMI)*, Herlany, Slovakia, Jan. 2020, pp. 281–286, doi: 10.1109/SAMI48414.2020.9108717.
[40] Z. Pang, F. Niu, and Z. O'Neill, "Solar radiation prediction using recurrent neural network and artificial neural network: A case study with comparisons," *Renew. Energy*, vol. 156, pp. 279–289, Aug. 2020, doi: 10.1016/j.renene.2020.04.042.
[41] C. R. Nagarathna, M. Kusuma, and K. Seemanthini, "Classifying the stages of Alzheimer's disease by using multi layer feed forward neural network," *Proc. Comput. Sci.*, vol. 218, pp. 1845–1856, Jan. 2023, doi: 10.1016/j.procs.2023.01.162.
[42] A. Ashraf, W. H. Bangyal, H. T. Rauf, S. Pervaiz and J. Ahmad, "Training of artificial neural network using new initialization approach of particle swarm optimization for data classification," in *Proc. Int. Conf. Emerg. Trends Smart Technol. (ICETST)*, Karachi, Pakistan, 2020, pp. 1–6, doi: 10.1109/ICETST49965.2020.9080707.
[43] J. Khan, E. Lee, and K. Kim, "A higher prediction accuracy–based alpha–beta filter algorithm using the feedforward artificial neural network," *CAAI Trans. Intell. Technol.*, vol. 8, no. 4, pp. 1124–1139, Dec. 2023, doi: 10.1049/cit2.12148.
[44] F. Rodríguez, S. Alonso-Pérez, I. Sánchez-Guardamino, and A. Galarza, "Ensemble forecaster based on the combination of time-frequency analysis and machine learning strategies for very short-term wind speed prediction," *Electric Power Syst. Res.*, vol. 214, Jan. 2023, Art. no. 108863, doi: 10.1016/j.epsr.2022.108863.
[45] O. A. M. López, A. M. López, and J. Crossa, "Fundamentals of artificial neural networks and deep learning," in *Multivariate Statistical Machine Learning Methods for Genomic Prediction*. Cham, Switzerland: Springer, 2022, doi: 10.1007/978-3-030-89010-0_10.
[46] M. I. C. Rachmatullah, J. Santoso, and K. Surendro, "A novel approach in determining neural networks architecture to classify data with large number of attributes," *IEEE Access*, vol. 8, pp. 204728–204743, 2020, doi: 10.1109/ACCESS.2020.3036853.



**MERVE ERKINAY OZDEMIR** received the Ph.D. degree in electrical and electronics engineering from Ankara University, Ankara, Turkey, in 2012.

Since 2016, she has been an Assistant Professor with the Department of Electrical Electronics Engineering, Iskenderun Technical University, Hatay, Turkey. Her research interests include signal processing, machine learning, artificial intelligence, and pattern recognition.