



One-hour ahead wind speed forecasting using deep learning approach

Arif Ozbek¹ · Akin Ilhan² · Mehmet Bilgili¹ · Besir Sahin³

Accepted: 7 June 2022 / Published online: 8 July 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Accurate forecasting of wind speed (WS) data plays a crucial role in planning and operating wind power generation. Nowadays, the importance of WS predictions overgrows with the increased integration of wind energy into the electricity market. This work proposes machine learning algorithms to forecast a one-hour ahead short-term WS. Forecasting models were developed based on past time-series wind speeds to estimate the future values. Adaptive Neuro-Fuzzy Inference System (ANFIS) with Fuzzy c-means, ANFIS with Grid Partition, ANFIS with Subtractive Clustering and Long Short-Term Memory (LSTM) neural network were developed for this purpose. Three measurement stations in the Marmara and Mediterranean Regions of Turkey were selected as the study locations. According to the hourly WS prediction, the LSTM neural network based on the deep learning approach gave the best result in all stations and among all models applied. Mean Absolute Error values in the testing process were obtained to be 0.8638, 0.9603 and 0.5977 m/s, and Root Mean Square Error values were found to be 1.2193, 1.2573 and 0.7531 m/s from the LSTM neural network model for measuring stations MS1, MS2, and MS3, respectively. In addition, the analyzes revealed that the best correlation coefficient (R) results among the algorithms in the test processes were obtained to be 0.9498, 0.9147, and 0.8897 for the MS1, MS2, and MS3 measurement stations, respectively. In this regard, it is shown that the LSTM method gave high sensitive results and mainly provided greater performance than the ANFIS models for one hour-ahead WS estimations.

Keywords Deep learning · Wind speed · One-hour ahead prediction · ANFIS-FCM · ANFIS-GP · ANFIS-SC · LSTM

1 Introduction

In recent years, renewable energy sources (RESs) have grown excessively around the world. RESs have significant spreading potential among the other energy sources, and they are concentrated in many geographical regions (Bilgili and Sahin 2009). All countries have at least one large number of RES power plants, and many countries have a resource portfolio in this sector. Among whole RESs, wind

energy has become widely used in the World and makes an essential contribution to renewable energy generation (Bilgili et al. 2015; Bulut and Muratoglu 2018; Wang and Wang 2015; Jones and Eiser 2010; Zheng et al. 2016). Besides, wind power has become a reliable and competitive electric-producing technology. The progress regarding the wind power installations continues to be more robust in the World, and every coming year, this sector involves more active countries and manufacturers and enhances yearly installed wind capacity with additional investments (Mendecka and Lombardi 2019; Kazimierczuk 2019; Köktürk and Tokuç 2017; Gualtieri 2019; Ahmed 2018; Tagliapietra et al. 2019; Kılıç 2019). While the cumulative installed wind power capacity in 2008 was 120.7 GW globally, the total worldwide wind power capacity at the end of 2018 increased to 600.3 GW. A 61.16 GW of global wind energy capacity has been installed by 2018 (GWEC 2020; EWEA 2020). On the other hand, global wind energy capacity is expected to increase substantially in the coming years. However, because of the natural irregular characteristic of

✉ Arif Ozbek
arozbek@cu.edu.tr

¹ Department of Mechanical Engineering, Ceyhan Engineering Faculty, Cukurova University, 01950 Ceyhan, Adana, Turkey

² Department of Energy Systems Engineering, Faculty of Engineering and Natural Sciences, Yıldırım Beyazıt University, 06010 Ankara, Turkey

³ Department of Mechanical Engineering, Engineering Faculty, Aydin University, 34295 Kucukcekmece, Istanbul, Turkey

wind speed, it can be quite challenging to fulfill reliable and seasonal wind energy management in power generation systems (Chen et al. 2018). With the high integration of wind power into electric power, the power system becomes more unreliable due to wind speed fluctuations' intermittent and stochastic nature. Therefore, accurate estimation of wind speed is undoubtedly inevitable to increase wind energy usage efficiency and reduce the operating cost of wind speed (Wang and Li 2018).

The literature studies show that deterministic wind speed and wind power forecasting can be divided into three parts, including physical, statistical, and artificial intelligence tools (Zhang et al. 2019a, b, c, d; Jung and Broadwater 2014). Physical information such as atmospheric temperature, atmospheric pressure, roughness, and obstacles are used in physical models, which are basic methods (Tascikaraoglu and Uzunoglu 2014). For instance, numerical weather prediction (NWP) processes exploit a series of mathematical equations and physical information. Physical approaches are usually by medium to long term wind power forecasting. On the other hand, they are unfortunately complex in computations and require significant computing resources. Recently, statistical models have also been used for wind speed anticipation. In statistical methods, historical wind speed data is often used in a suitable time frame to predict future wind speed. The statistical approaches mainly include moving average models (MA), autoregressive models (AR), autoregressive moving average models (ARMA), an autoregressive integrated moving average models (ARIMA), as well as seasonal autoregressive integrated moving average models (SARIMA). The statistical models are initially fitted to the historical data and then employed to generate the predictions. Statistical models are well applied in the predictions of ultra-short-term and short-term wind power forecasting. However, their prediction accuracies are somewhat limited. A recent model regarding the recursive ARIMA and EMD (empirical mode decomposition) was proposed by Liu et al. (2015). In this study, short-term wind speed prediction applied in the railway strong wind warning system was implemented. Kavasseri (2009) developed a fraction-ARIMA and Seetharaman (2009) to predict one day and two-day ahead of wind speed in the region of North Dakota.

Finally, wind speed anticipations can be achieved based on artificial neural networks (ANN). The networks can be structured using back-propagation, radial basis functions, multi-layer perceptron neural networks, extreme learning machines, and Bayesian neural networks. Wang and Li (2018) suggested a hybrid model depending on optimal feature extraction, deep learning algorithm, and error correction strategy for multi-step wind speed forecasting. Braun et al. (2020) suggested and developed a stochastic

reduced-form model of power time series to prosper the modeling perspective of complex and large wind power systems. Zhang et al. (2019a, b, c, d) predicted the wind turbine power, handling the algorithm of a long short-term memory network (LSTM). Besides analyzing the error distribution characteristics considering short-term wind turbine power predictions, the Gaussian mixture model (GMM) was utilized in this study. An innovative approach named Ensem-LSTM is used by Chen et al. (2018). In their model, the nonlinear-learning ensemble of deep learning time series forecasting depending on the long short-term memory neural networks (LSTMs), support vector regression machine (SVRM), and extremal optimization algorithm (EO) were used. An innovative forecasting outline involving an RNN (recurrent neural network) structure model and LSTM (long short-term memory) is provided by Dong et al. (2018). In the study, an operative prediction map was adapted to different forecasting horizons. LSTM method, taking the experimental results into account, was demonstrated to obtain more sensitive predictions than conventional neural network methods in the study of Shi et al. (2018). The multi-task training method is used by Qin et al. (2019a, b) to predict wind turbine energy. Deep neural network (DNN) models such as long short-term memory (LSTM) recurrent neural networks are shown by Wu et al. (2016) to generate better prediction results when compared to the conventional methods. Yu et al. (2018) demonstrated that innovative proposed hybrid models, including support vector machine (SVM), standard recurrent neural network (RNN) and LSTM, and gated recurrent unit neural network (GRU), generated more accurate predictions compared to the conventional methods. Han et al. (2019) suggested a solution to these situations considering copula function and LSTM. They proposed an effective origination of the critical meteorological factors affecting power generation depending on nonlinear influences and tendencies. K-nearest neighbours regressor (KNNR), random forest regressor (RFR), decision tree regressor and multiple-layer perceptron regressor (MLPR) are four machine learning regression methods presented by Mogos et al. (2022) (DTR). For wind speed forecasts, twenty minutes of real data of wind speed collected at one-minute intervals were used. Their proposed approaches were designed to provide a low-cost and effective solution for predicting short-term wind speeds.

A hybrid type of model, including the decomposition-virtual nodes-pruning and LSTM depending on the neighbourhood gates, is recommended by Zhang et al. (2019a, b, c, d). In the study, they attained wind speed forecasting of high-precisions. Lopez et al. (2018) demonstrated a combination of architectural structures, including long short-term memory (LSTM) and Echo State Network (ESN), which contained the characteristics of

both together. Zhang et al. (2019a, b, c, d) proffered a shared weight long short-term memory network (SWLSTM) to reduce the number of optimized variables and the necessary time for the training. In their study, while high forecasting precision is obtained in the utilization of LSTM, the training time of the computing process was observed to be reduced. Liang et al. (2018) used a model of multi-variable stacked LSTMs (MSLSTM). In the model, the wind speed values were predicted based on the meteorological data, including wind speed, temperature, pressure, humidity, dew point, and solar radiation. Their results indicated that the suggested MSLSTM structure could seize and learn the uncertainties and generate competitive output performance. Liu et al. (2018a, b) presented a multi-step forecasting model through the combination of variational mode decomposition (VMD), singular spectrum analysis (SSA), LSTM, and extreme learning machine (ELM). The study demonstrated that the combined model yielded superior performance in predictions compared to the other tested methods. Xiaoyun et al. (2016) demonstrated that the LSTM forecasting model has higher forecasting precision when compared to the back-propagation (BP) algorithm and the support vector machine (SVM) model. They also indicated that its application in wind energy forecasting is obtained more suitable. So, it was concluded that the method might find more applications for wind engineering. Balluff et al. (2015) conducted the prediction of wind speed and pressure using recurrent neural networks (RNN). Zaytar and Amrani (2016) used the deep neural network structure to perform weather prediction by time series. They handled multi-stacked LSTMs for the map sequences of weather values having the same length in this regard. Their results based on LSTM computations were shown to be competitive with the conventional methods, and they presented a good alternative in the predictions of general weather conditions.

Liu et al. (2018a, b) suggested a model which included empirical wavelet transform (EWT), LSTM, and Elman neural network (ENN) had a good performance in the forecasting of wind speed with high precision. Yuan et al. (2019) performed a prediction of wind power by a hybrid model of LSTM neural network and beta distribution function-based particle swarm optimization (Beta-PSO). The study demonstrated that the reliability of the hybrid model could be applied safely and stably for the operation of power systems. Ding et al. (2022) introduced a new hybrid prediction system that includes a fuzzy entropy-based double decomposition strategy, piecewise error correction, Elman neural network, and ARIMA model. The proposed hybrid model was used to solve issues with both linear and nonlinear wind speed series. Huang et al. (2018) utilized a hybrid model of the combination, including ensemble empirical mode decomposition (EEMD) and

Gaussian process regression (GPR), as well as the LSTM neural network. According to the experimental measurements obtained from China, it was revealed that these cited prediction methods gave better results compared to the other prediction methods. Hu and Chen (2018) conducted a nonlinear hybrid model to obtain forecasting using LSTM neural network, differential evolution algorithm (DE), and hysteretic extreme learning machine (HELM), where the advantage of this suggested model compared based on the others in terms of performance was shown in their study. Chen et al. (2019a, b) performed wind speed prediction by proposing a convolutional neural network and LSTM neural network. They showed that their proposed model gave better results than the other conventional methods of high forecasting precision. Chen et al. (2019a, b) proposed the EEL-ELM method involving an extreme learning machine (ELM), Elman neural network (ENN), and long short term memory neural network (LSTM), in which they demonstrated this hybrid method generated successful forecasting performance. Ozen et al. (2019) studied wind speed predictions of İstanbul, İzmir, Muğla, Tekirdağ, and Eskişehir provinces of Turkey. The wind speed data of these provinces were used in performing the forecasting based on long short-term memory (LSTM) neural network. Ehsan et al. (2020) indicated the LSTM approach exceeded the other twelve models in terms of the accuracy of predictions by 97.8%. Three local recurrent neural network types were used by Barbounis et al. (2006), including the infinite impulse response multi-layer perceptron (IIR-MLP), and the diagonal recurrent neural network (RNN), and the local activation feedback multi-layer network (LAF-MLN). In this study, simulation outcomes indicated that the RNN models outperform the other models. The predictions were performed by Zhang et al. (2019a, b, c, d) utilizing LSTM, auto-regressive moving average (ARMA), and kernel density estimation (KDE). This study contributed to the construction of the smart grid. Prabha et al. (2019) proposed LSTM wind speed forecasting using one hour ahead of clustering. In their study, they considered four different sites for wind speed predictions.

Qin et al. (2019a, b) used ensemble empirical mode decomposition (EEMD), fuzzy entropy (FuzzyEn), and long short-term memory neural network (LSTMNN) for the decomposition of the original wind speed series into a series of components. Their computations indicated that the forecasting accuracy of the model referred to as EEMD-FuzzyEn-LSTMNN was shown to be better than the other three models applied alone, including back-propagation neural network (BPNN), support vector machines (SVM), and long short-term memory neural network (LSTMNN). Based on the selection of essential input variables, Yaxue et al. (2022) suggested an interval type-2 (IT2) wind speed forecasting model. They also used the recursive least

squares (RLS) approach to get the conclusion parameters for wind speed prediction in real time. Using their proposed algorithm, they were able to improve forecast accuracy. Liu et al. (2019) suggested a hybrid model composed of a discrete wavelet transform (DWT) and LSTM networks. The LSTM and DWT networks were studied to exhibit the wind power time series dynamic behaviour and decompose wind power time series into non-stationary components. In their study, it was presented that this proposed method applied to three different wind farms improved the prediction accuracy. Cali and Sharma (2019) presented the numerical prediction of wind power and weather using the long short-term memory based on a recurrent neural network (LSTM-RNN). They used wind power data and data of numerical weather prediction, which both were obtained from a wind farm in Spain. Qian et al. (2019) proposed a recent technique for wind turbines depending on the algorithms of the LSTM neural network, in which this suggested method was shown to increase the reliability and the economic benefits of the wind farms. Marndi et al. (2020) suggested a prediction methodology for analyzing the wind time series data of New Delhi in North India and Bengaluru in South India meteorological stations. In their study, the LSTM model was utilized. The forecasting results of the LSTM were compared with the predictions obtained from support vector machine (SVM), and extreme learning machine (ELM) analyses and the LSTM model in their study was demonstrated to improve the short-term wind speed forecasting ability at the station level. Bilgili and Sahin (2010) came up with a model for determining wind speed using some meteorological data. These data included wind speed, atmospheric pressure, ambient temperature, rainfall, and relative humidity. They considered the methods of linear regression (LR), nonlinear regression (NLR), and artificial neural network (ANN). Besides, it was shown that the provided artificial neural network (ANN) method is superior to linear regression (LR) and nonlinear regression (NLR) methods. Yang et al. (2018) proposed an ultra-short-term multi-step wind power forecasting model considering the usual unit method. The least-squares support vector machine (LSSVM) model was utilized as a prediction model based on the measured power data of wind farms. They demonstrated that the suggested model could effectively improve forecasting accuracy. Yang et al. (2022) proposed a data-decomposition-based ultra-short-term/short-term wind speed prediction approach based on improved singular spectrum analysis (ISSA). They compared the outcomes of numerous data preprocessing decomposition methods, including EMD, EEMD, and CEEMD, against the results of the prediction without data preprocessing. Their findings revealed that ISSA can significantly increase prediction accuracy. Table 1 gives a

summary of the typical studies on wind forecasting available in the literature.

2 Research significance and novelty of the work

In this study, machine learning algorithms were applied to forecast one-hour ahead of WS. Forecasting models were developed based on past time-series WS values to estimate the future values. For this purpose, ANFIS with Fuzzy c-means (FCM), ANFIS with Grid Partitioning (GP), ANFIS with Subtractive Clustering (SC) and long short-term memory (LSTM) neural network models were used. The novelty of the study is to present the superiority of the LSTM neural network algorithm compared to the conventional ANFIS approaches. In this manner, the LSTM algorithm has been compared with the ANFIS-FCM, ANFIS-GP and ANFIS-SC. So, these four algorithms were initially trained according to the observed real wind speed data of three wind farms operating in Turkey. As soon as the training of the algorithms was completed, the testing processes were conducted to demonstrate the supremacy of the LSTM algorithm over three conventional ANFIS approaches.

3 Material and methods

3.1 Adaptive network fuzzy inference system (ANFIS)

ANFIS is known as a kind of ANN method based on the Takagi–Sugeno fuzzy inference system. It was developed by Jang in the early 1990s and has been used for modeling nonlinear functions and estimating chaotic time series. The generated ANFIS model commonly utilizes a hybrid learning algorithm. ANFIS takes advantage of both constructs as it integrates both neural networks and fuzzy logic inference methods. To use the ANFIS technique, a dataset based on input–output is generally needed. The model, which is established depending on the type and number of membership functions (MFs) selected, is created using a learning algorithm. The method uses the set of fuzzy if–then rules it creates. The ANFIS structure is constructed by parameterizing the difference between the output of the entire network and the target value that is, minimizing the error. Theoretically, ANFIS can satisfactorily approximate any continuous function (Abyaneh et al. 2011). General details concerning the ANFIS architecture are available in the literature (Jang 1993; Karakuş et al. 2017; Tabari et al. 2012; Mathworks 2020).

Table 1 Some studies on wind estimation available in the literature

| Refs | Approach | Estimation type | Input variables | Wind data |
|----------------------------------|------------------------------|-----------------|--|--|
| Chen et al. (2018) | EnsemLSTM | Wind speed | Past data | Ten-minute data 23.11.2012–28.11.2012 |
| Wang and Li (2018) | LSTMN | Wind speed | Past data | Hourly data of 5 years |
| Liu et al. (2015) | EMD and RARIMA | Wind speed | Past data | Minutely measured 400 data |
| Kavasseri and Seetharaman (2009) | Fractional-ARIMA and f-ARIMA | Wind speed | Past data | 10-min average data to predict one day or two days ahead |
| Dong et al. (2018) | RNN and LSTM | Wind power | Past data | Hourly wind data was measured between July 2009 to June 2012 |
| Shi et al. (2018) | R (recursive), VMD, and LSTM | Wind power | Past data | Hourly data was measured between 1–31.08.2017 and 1–31.03.2017 |
| Wu et al. (2016) | DNN, LSTM, and RNN | Wind power | Past data | 15-min data for 1 the year 2014 |
| Lopez et al. (2018) | LSTM and ESN | Wind power | Past data | Hourly data between 14.01.2002 to 25.08.2002 |
| Liang et al. (2018) | MSLSTM | Wind speed | Temperature, solar radiation, dew point, wind speed, humidity and pressure | 5-min data of February 2016 |
| Xiaoyun et al. (2016) | LSTM | Wind power | Past data | 5-min data of 01.05.2010 to 31.05.2011 |
| Huang et al. (2018) | EEMD, GPR, and LSTM | Wind speed | Past data | 5 min data between 01.01.2014 to 04.01.2014 and 60 min data between 01.07.2014 to 25.07.2014 |
| Chen et al. (2019a, b) | ELM, ENN, and LSTM | Wind speed | Past data | 10 min and 1-h data |
| Prabha et al. (2019) | LSTM | Wind speed | Past data | 1-h data of 2009–2013, 2011–2015, 2013–2016, and 2009–2014 |
| Mamdi et al. (2020) | LSTM, SVM, and ELM | Wind speed | wind speed, temperature, pressure and humidity | 30 min data measured between 2010–2014 |

Fuzzy c-means (FCM) is a clustering method of ANFIS, allowing each data point to have multiple clusters and belong to different membership degrees. The basis of the FCM algorithm involves minimizing the objective function. Similarly, the concept of FCM is available in the literature (Mathworks 2020). Subtractive Clustering Algorithm (SC) is also an algorithm of ANFIS that considers each data point as a candidate cluster center, and the potential of each data point is calculated by measuring the density of the data point surrounding the cluster center. The algorithm uses an iterative process, assuming each point is potentially a cluster center considering its location for other data points. Studies regarding the SC model can also be found in the literature (Benmouiza and Cheknane 2019). The grid partitioning (GP) algorithm divides the input data space into a rectangular subspace with the help of an axis-parallel partition. Each input is partitioned into identically shaped membership functions. The number of the

fuzzy if–then rules is the same as M^p . Here, the input dimension is shown by n , and the number of partitioned fuzzy subsets for each input variable is denoted by M (Abyaneh et al. 2011). This approach to solving a problem is simply referred to as functional decomposition. This phenomenon is as well extensively explained in the literature (Chandy 1992).

3.2 Long short term memory (LSTM) neural network

LSTM neural network, presented by Hochreiter and Schmidhuber (1997), is an RNN architecture utilized in deep learning approaches. Unlike traditional feed-forward neural networks, LSTM has memory cells with constant errors or feedback connections. An ordinary LSTM unit consists of a cell and three different gates. The cell remembers values in variable-length time intervals, and these three gates

regulate the flow of information into and out of the cell. An input gate teaches to protect the fixed error stream in the memory cell from irrelevant inputs, while an output gate teaches to protect other units from unrelated memory content reserved in the memory cell. On the other hand, the forget gate learns how long the value is in the memory cell (Zahroh et al. 2019; Salman et al. 2018).

An LSTM layer architecture is presented in Fig. 1. In the graph, h_t and c_t are the output and the cell state at time step t , respectively. To compute the first output and the updated cell state, the first LSTM element is used for the first state of the neural network and the early time step of the series. At time step t , this block utilizes the current state of the network (c_{t-1} , h_{t-1}) and the next time step of the sequence to calculate the output and the updated cell state c_t .

The flow of data at a time step, t is presented in Fig. 2. In this flow, input gate (i) and output gate (o) check the level of cell state update and the level of cell state added to the hidden state, respectively. Forget gate (f) controls the level of cell state reset. Moreover, cell candidate (g) adds the information to the cell state.

The learnable weights of an LSTM layer are the bias b , the recurrent weights R , and the input weights W . They are expressed as follows:

$$W = \begin{bmatrix} W_i \\ W_f \\ W_g \\ W_o \end{bmatrix}, R = \begin{bmatrix} R_i \\ R_f \\ R_g \\ R_o \end{bmatrix}, b = \begin{bmatrix} b_i \\ b_f \\ b_g \\ b_o \end{bmatrix} \quad (1)$$

where i, f, g and o are the input gate, forget gate, the cell candidate and output gate, respectively. The cell state at a time step of t is indicated as,

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (2)$$

where \odot is the Hadamard product (element-wise

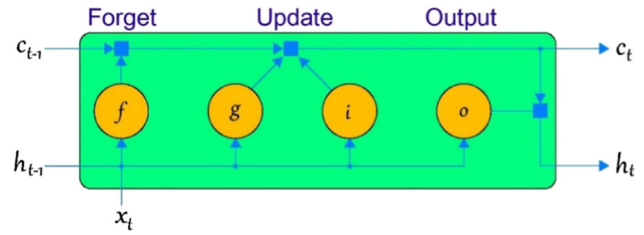


Fig. 2 The flow of data at a time step, t

multiplication of vectors). The hidden state at a time step of t is defined as,

$$h_t = o_t \odot \sigma_c(c_t) \quad (3)$$

where σ_c presents the state activation function. The hyperbolic tangent function is utilized to calculate the state activation function for the LSTM layer function by default. The components at time step t are defined as follows (Liu and Liu, 2019):

$$i_t = \sigma_g(W_i x_t + R_i h_{t-1} + b_i) \quad (4)$$

$$f_t = \sigma_g(W_f x_t + R_f h_{t-1} + b_f) \quad (5)$$

$$g_t = \sigma_c(W_g x_t + R_g h_{t-1} + b_g) \quad (6)$$

$$o_t = \sigma_g(W_o x_t + R_o h_{t-1} + b_o) \quad (7)$$

where, σ_g presents the gate activation function. An LSTM layer function utilizes the sigmoid function to calculate the gate activation function, and it is defined as,

$$\sigma(x) = (1 + e^{-x})^{-1} \quad (8)$$

Figure 3 illustrates the basic steps of the proposed LSTM neural network model for WS prediction. After supplying WS data from the wind turbine, simulation parameters were set. The original data were trained, validated, tested, and then normalized, respectively.

Fig. 1 LSTM layer architecture

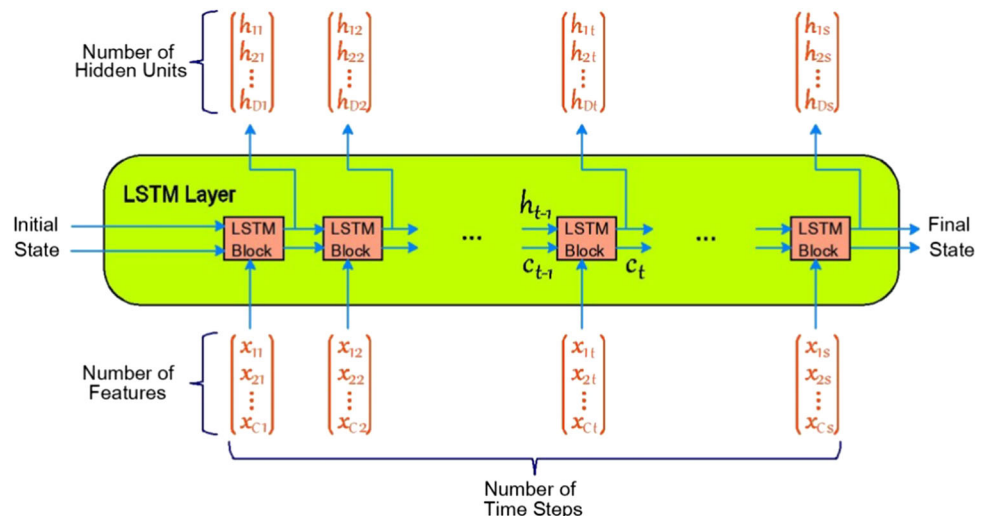
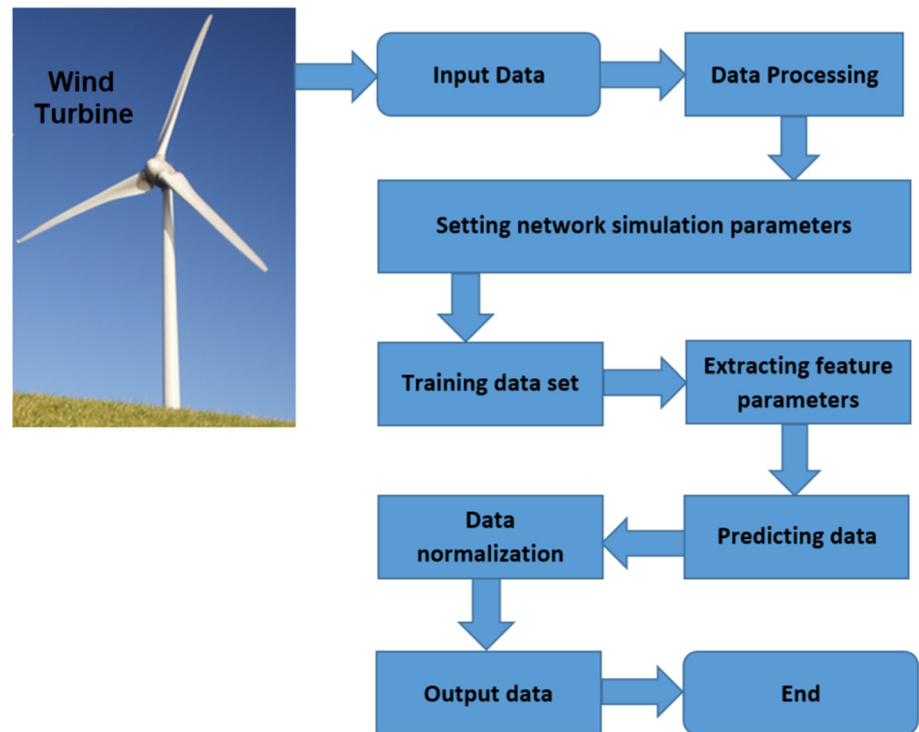


Fig. 3 Wind speed prediction flow chart based on the LSTM model



3.3 Study area

Interested area locations for the current study are plotted on the map of Turkey, as shown in Fig. 4. In this regard, three wind power plants were chosen from the south and the northwest of Turkey. Namely, two wind power plants were chosen from the Hatay province of Turkey. It is a city located in the south of Turkey, and the measurement stations for wind power plants are named MS1 and MS3, as shown in Fig. 4. On the other hand, one wind power plant was chosen from the Tekirdag province of Turkey. It is a city neighbour to Istanbul province, and the measurement station is named MS2 for this province. The data of these three measurement stations in the wind power plants were used for the current study.

3.4 Data analysis and model structure

In the current study, the WS forecasting was actualized according to the measured data of three installed wind farms located in Turkey. In this regard, the WS predictions have been carried out using a time-series analysis-based, ANFIS-FCM, ANFIS-GP, ANFIS-SC and LSTM neural network. In the whole of these proposed simulations, the measurement data were split down into two datasets composed of training and testing clusters. While the training cluster was used in model training for the whole of the suggested methods, the testing dataset was utilized for over-fitting the model validation. Besides, the power of

these proposed forecasting methods was evaluated in terms of the mean absolute error (MAE), root means square error (RMSE), and the correlation coefficient (R). In LSTM applications, the total number of hidden layers was considered between 5 and 150. The epoch number was kept constant at 300. In ANFIS structures, Sugeno's fuzzy approach was used to derive the output values obtained from the input variables. In this way, a variety of ANFIS structures were tested, and the optimal model structures, the number of membership functions (MFs), and the total number of iterations (i.e., epoch number) were adjusted by the trial and error method. Then, the utilized ANFIS models were analyzed based on the comparisons of the given statistical parameters. Ultimately, in the ANFIS-FCM model, the amount of MFs was set in a range of 2 to 10 increased one by one. However, the input number and the max epoch number were adjusted to correspond to values 5 and 100, respectively.

Similarly, in the ANFIS-SC model, the input number and the max epoch number were set at 5 and 100, respectively. The influence of the radius in this method was analyzed in a range of 0.2 to 0.9. In the ANFIS-GP model, the input number and the max epoch number were tried, as in ANFIS-SC and ANFIS-FCM models. The number of the MFs was set to 2 and 3 in this model.

Figure 5 presents the WS data used for MS1, MS2, and MS3, respectively. Wind speed values for these wind farms include the measured dataset considering the year ranges between 01.01.2015 and 30.04.2017, 05.05.2018 and

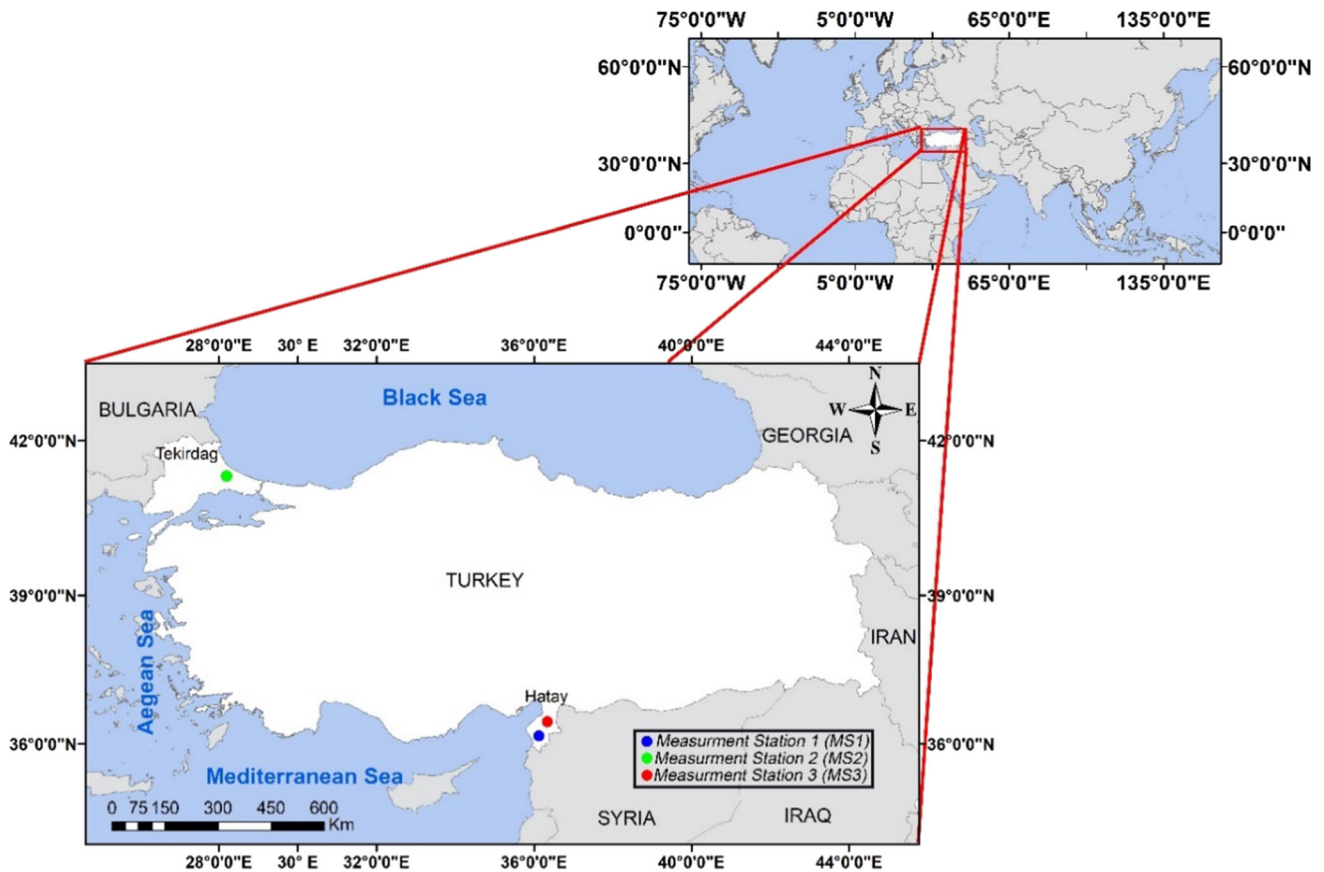


Fig. 4 The map locations of MS1, MS2, and MS3

08.07.2018 01.06.2016 and 31.07.2016 for the wind farms of MS1, MS2, and MS3, respectively. Wind speed data in these wind farms were measured mechanically with a cup anemometer with a vane. Besides, the wind blowing speeds in the cited wind farms were observed to be in the ranges of $0.01 \leq WS \leq 23.80$ m/s, $0.47 \leq WS \leq 17.23$ m/s, and $4.23 \leq WS \leq 14.20$ m/s, for MS1, MS2, and MS3, respectively. On the other hand, 20,424, 1,531, and 743 samples of hours were respectively used considering MS1, MS2, and MS3 measuring stations. In this regard, these samples were discretized into two pieces for 70% and 30%, respectively, corresponding to the training and testing data clusters. The prediction models were initially trained to utilize the training data cluster, then secondly validated using the testing data cluster. The other features of the measured data for these wind power plants are summarized in Table 2, including the arithmetic mean values of the measured data, standard deviation values, and the cumulative values for the training and testing data clusters.

3.5 Error analysis for the proposed methods

In our study, three statistical error criteria comprising mean absolute error (MAE), root mean square error (RMSE), and

correlation coefficient (R) are utilized for the evaluation of the goodness of a model. These error criteria are applied to control the accuracy of the estimations according to the observed variables (Bilgili et al. 2013). The mathematical expressions of these error criteria are provided in Eqs. (9), (10), and (11), respectively.

$$MAE = \frac{1}{N} \sum_{i=1}^N |p(i) - o(i)| \quad (9)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N [p(i) - o(i)]^2} \quad (10)$$

$$R = \left(\frac{\sum_{i=1}^N [p(i) - \bar{p}][o(i) - \bar{o}]}{\left(\sqrt{\sum_{i=1}^N [p(i) - \bar{p}]^2} \sqrt{\sum_{i=1}^N [o(i) - \bar{o}]^2} \right)} \right) \quad (11)$$

where $p(i)$ and $o(i)$ are the predicted value and observed value at the time i , respectively. Also, \bar{p} and \bar{o} present mean values of the predicted data and the actual data, respectively. The total number of data is represented by N .

Fig. 5 Time series of the hourly observed WS data for MS1, MS2 and MS3

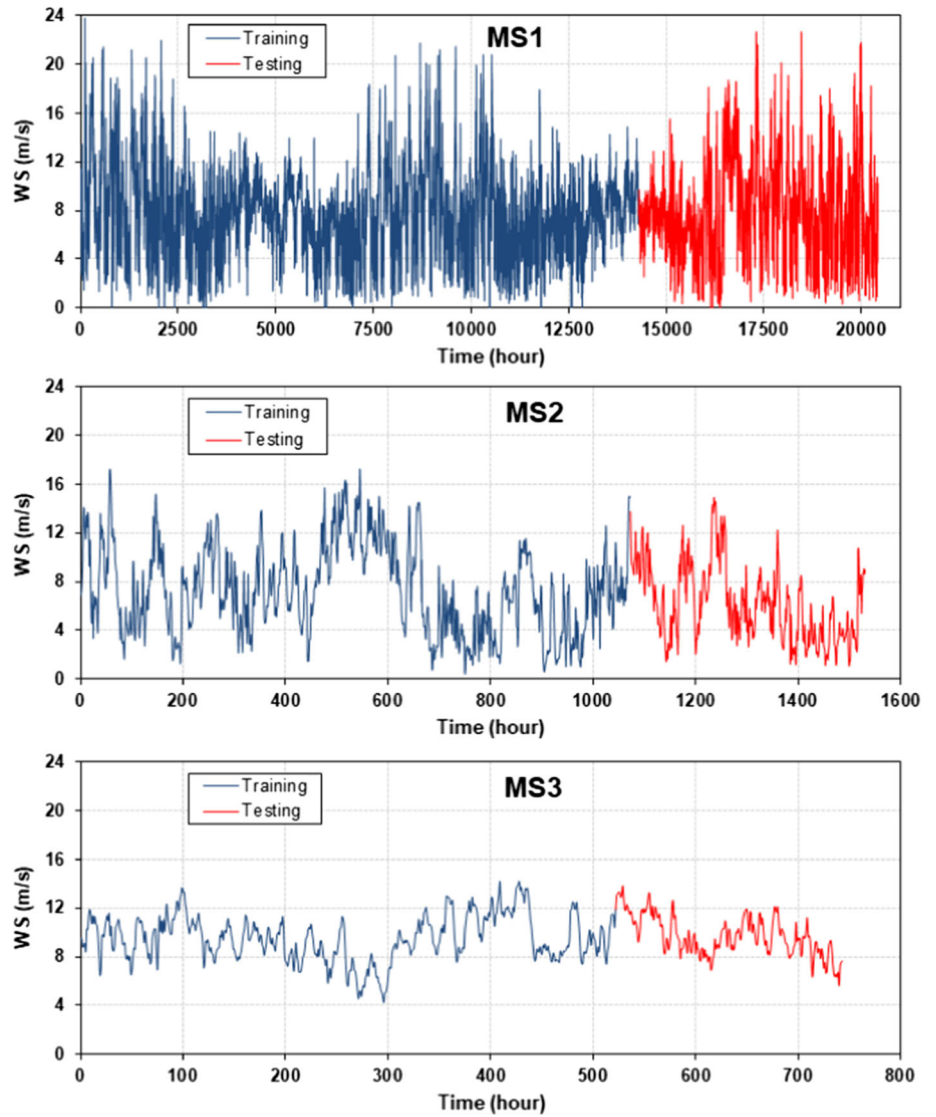


Table 2 Information on the original observed data of the measurement stations

| Data | MS1 | MS2 | MS3 |
|----------------------------------|---------------------------|---------------------------|---------------------------|
| Minimum value (m/s) | 0.01 | 0.47 | 4.23 |
| Maximum value (m/s) | 23.80 | 17.23 | 14.20 |
| Arithmetic mean value (m/s) | 7.89 | 6.99 | 9.58 |
| Standard deviation (m/s) | 3.6682 | 3.5207 | 1.8247 |
| Data date range | 01.01.2015 and 30.04.2017 | 05.05.2018 and 08.07.2018 | 01.06.2016 and 31.07.2016 |
| Number of cumulative data (100%) | 20,424 | 1531 | 743 |
| Number of training data (70%) | 14,299 | 1074 | 521 |
| Number of testing data (30%) | 6125 | 457 | 222 |

Table 3 Optimal parameter determination for WS prediction using the LSTM model

| Station | Hidden layer number | Epoch number | MAE (m/s) | RMSE (m/s) | R |
|---------|---------------------|--------------|---------------|---------------|---------------|
| MS1 | 5 | 300 | 0.8691 | 1.2289 | 0.9491 |
| | 10 | 300 | 0.8638 | 1.2193 | 0.9498 |
| | 25 | 300 | 0.8636 | 1.2227 | 0.9495 |
| | 50 | 300 | 0.8662 | 1.2263 | 0.9492 |
| | 75 | 300 | 0.8656 | 1.2278 | 0.9490 |
| | 100 | 300 | 0.8674 | 1.2305 | 0.9488 |
| | 125 | 300 | 0.8736 | 1.2348 | 0.9483 |
| | 150 | 300 | 0.8822 | 1.2469 | 0.9474 |
| MS2 | 5 | 300 | 0.9603 | 1.2573 | 0.9147 |
| | 10 | 300 | 0.9727 | 1.2658 | 0.9133 |
| | 25 | 300 | 1.0493 | 1.3593 | 0.9016 |
| | 50 | 300 | 0.9753 | 1.2742 | 0.9126 |
| | 75 | 300 | 1.2996 | 1.5912 | 0.8774 |
| | 100 | 300 | 1.7896 | 2.2146 | 0.7870 |
| | 125 | 300 | 1.2861 | 1.6145 | 0.8583 |
| | 150 | 300 | 1.4604 | 1.8173 | 0.8352 |
| MS3 | 5 | 300 | 0.5977 | 0.7531 | 0.8897 |
| | 10 | 300 | 0.6355 | 0.8048 | 0.8750 |
| | 25 | 300 | 0.6163 | 0.7876 | 0.8801 |
| | 50 | 300 | 0.6947 | 0.8789 | 0.8356 |
| | 75 | 300 | 0.6616 | 0.8343 | 0.8648 |
| | 100 | 300 | 0.8747 | 1.1568 | 0.7224 |
| | 125 | 300 | 0.9022 | 1.1467 | 0.7384 |
| | 150 | 300 | 1.3313 | 1.7018 | 0.4883 |

The best results are shown in bold

4 Results and discussions

4.1 Results of the LSTM neural network

Table 3 provides special accuracy measures of the LSTM results. The best results are shown in bold. The assessment criteria in Table 3 were conducted based on the prediction values obtained from the test process results. A cumulative of 8 LSTM neural network models, including a range of hidden layer numbers between 5 to 150 for three measuring stations, were tried and then tested. As seen in this table, the best results were obtained in the hidden layer numbers corresponding to 10, 5, and 5 for MS1, MS2, and MS3, respectively.

The MAE, the RMSE, and the R results were respectively obtained as 0.8638 m/s, 1.2193 m/s, and 0.9498 for MS1. Similarly, for MS2, the same statistical parameters were obtained as 0.9603 m/s, 1.2573 m/s, and 0.9147, respectively. Finally, for MS3, the values of 0.5977 m/s MAE, 0.7531 m/s RMSE, and 0.8897 R were computed after the simulations of the testing cluster. These obtained results indicated that the proposed LSTM model got a satisfactory performance in WS forecasting of installed wind farms in Turkey.

On the other hand, Fig. 6 demonstrates the LSTM network testing time series, including the observed and forecasted wind speed data clusters, respectively, for MS1, MS2, and MS3. The number of the hourly samples and the wind speed pile were respectively located on X-axis and Y-axis. The significant fluctuation in the hourly measured samples can be seen in the WS time series plot. The WS forecasting technique reveals that the predictions of the tested WS time series functions almost overlap with the actual values when three measuring stations are considered. These overlapping incidents were also proved regarding the statistical calculations, as also demonstrated in Table 2. It is glad to report that the utilization of the LSTM neural network generated satisfactory results in such a sinusoidal data cluster. The forecasting outcomes can be analyzed in more detail to bring forward that Fig. 6 presented initially close looks to the hourly testing samples provided in Fig. 5, respectively, for measuring stations of MS1, MS2, and MS3, as well as provided LSTM prediction results for these wind farms. Intercalary to the LSTM neural network predictions shown in Fig. 6. Figure 7 demonstrates the LSTM histogram and regression plots of the observed actual values for the WS's forecasted values. In this regard,

Fig. 6 The hourly forecasted time series for WS data cluster with proposed LSTM; the predictions for MS1, MS2 and MS3

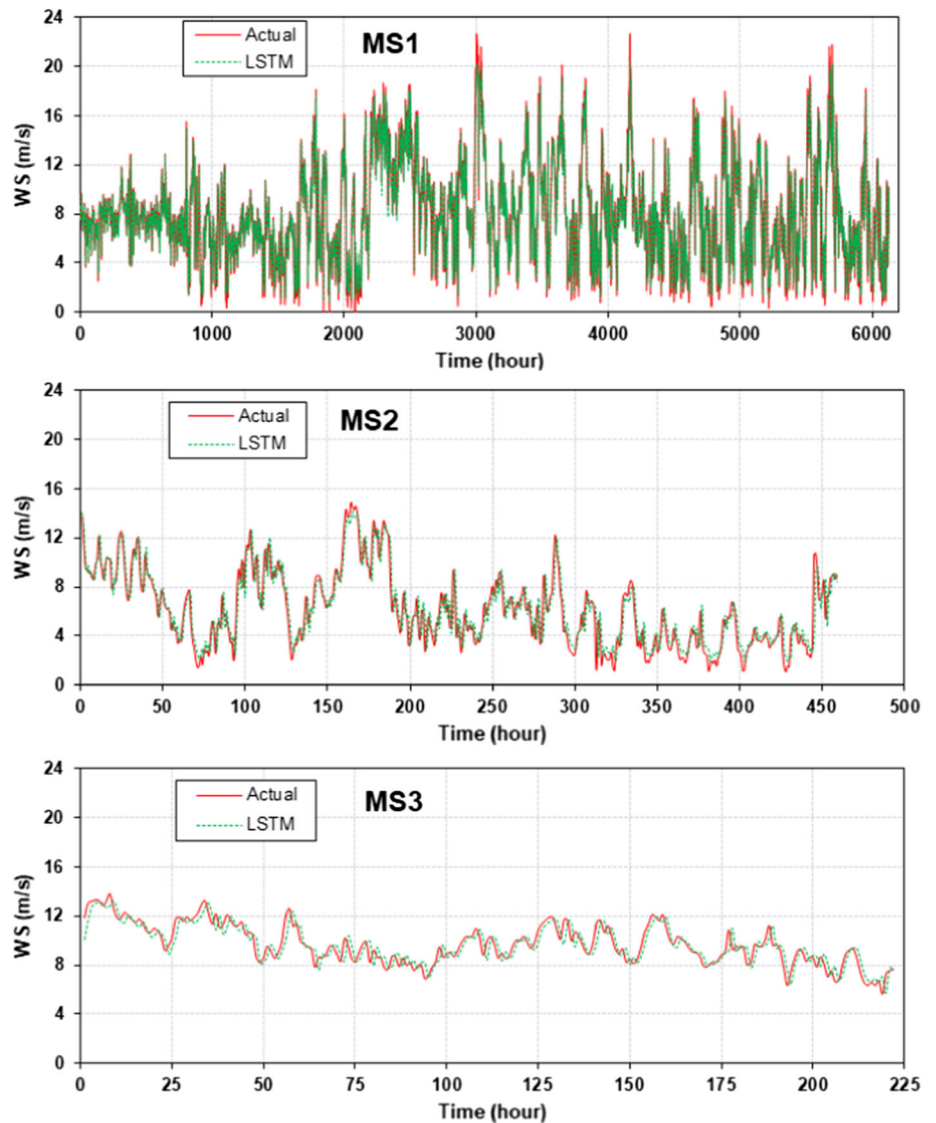


Fig. 7 stands for the histogram and regression plots of MS1, MS2, and MS3. The histogram part of this figure demonstrates the frequency distribution of differences between actual and predicted WS values for the LSTM method. The results in all stations obtained from this model appear to be skewed to the positive. As can be seen from this figure, while more than 93% of the model predictions have an error $\leq \pm 2$ m/s for MS1 and MS2; 100% of the model predictions have the same error for MS3. Regression part of this figure, the X-axis and Y-axis respectively denote the observed real and forecasted values of the WS data in the m/s unit. The correlation coefficient results of 0.9498, 0.9147, and 0.8897 for three measuring stations also accompany the actual and predicted value distributions in Fig. 7. So, it gives rise to understanding how precisely the applied model’s forecasting outcomes fit the actual observed data.

4.2 Results of the ANFIS-FCM model

Table 4 gives different evaluation criteria of the ANFIS-FCM model. A cumulative of 9 different models were tried and tested for the measuring stations in this suggested model. These models involved the number of MFs in the range between 2 and 10 with an increment of 1. In this model, the best forecasting results were obtained at the number of MFs corresponding to 6, 2, and 3 for MS1, MS2, and MS3 measuring stations. As observed in this table, the number of membership functions less than six did not generate good results for MS1, since the low number of membership functions in amount causes not-well partitioning of inputs. Additionally, membership functions greater than six did not also give good outcomes for this station. In this case, many nodes and fuzzy rules were used, which enhance the computation time, consequently

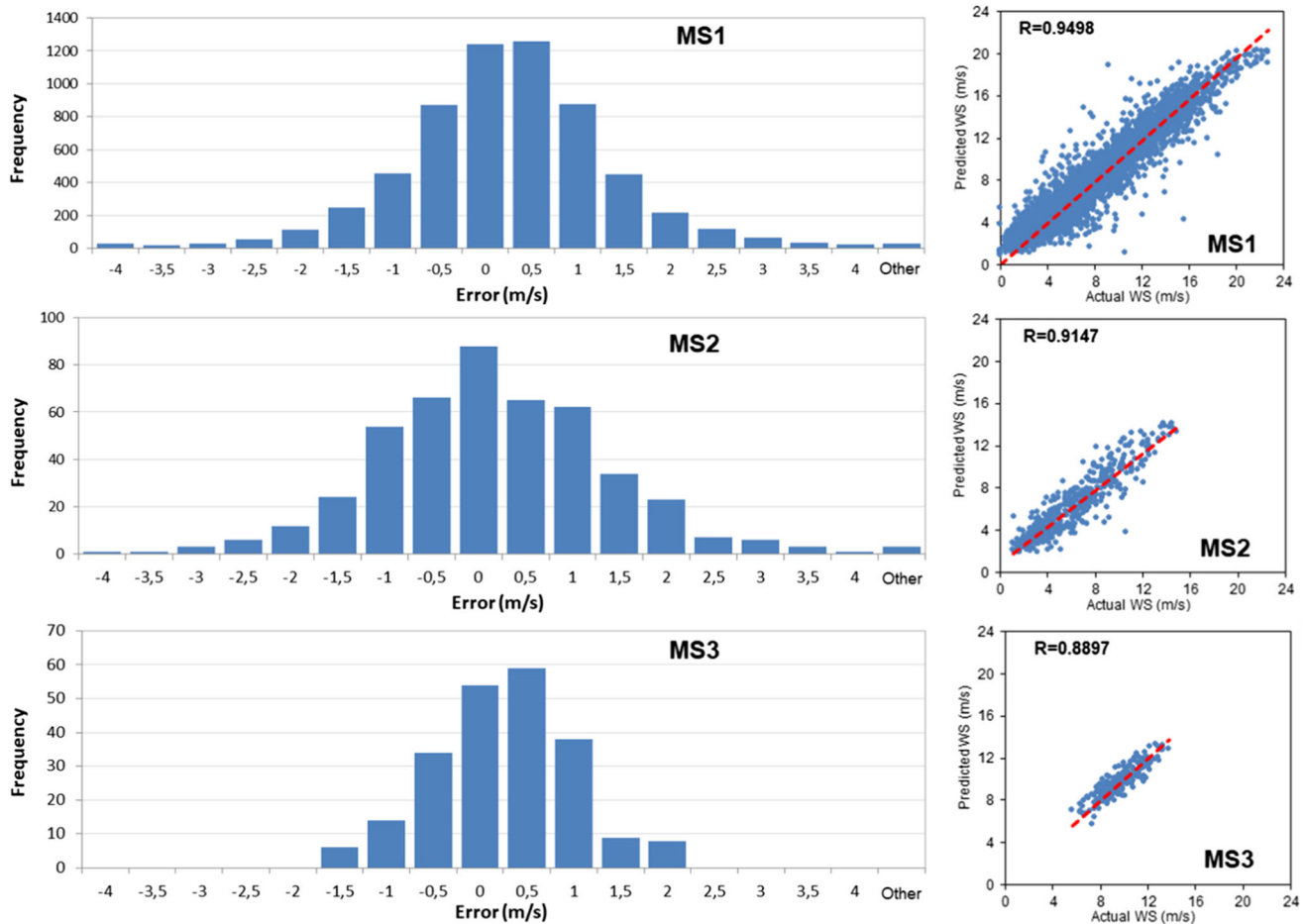


Fig. 7 Histogram and statistical regression graphs regarding the predicted vs the actual values of the WS data cluster concerning the suggested LSTM neural network model; the results presented for MS1, MS2 and MS3

resulting in relatively poor estimations. These situations are similar for MS2 and MS3 measuring stations. On the other hand, based on the testing process results in three measuring stations, at MFs numbers corresponding to 6, 2, and 3, respectively; the best performance statistical results were reported to be 0.8669 m/s MAE, 1.2260 m/s RMSE, and 0.9492 R for MS1, 0.9634 m/s MAE, 1.2606 m/s RMSE, and 0.9129 R for MS2, and 0.6182 m/s MAE, 0.7709 m/s RMSE, and 0.8848 R for MS3.

The testing time series containing the actual observed and predicted WS data clusters obtained by the suggested ANFIS-FCM model are shown in Fig. 8 for three measuring stations. As revealed in this figure, considering the testing part, it is reported that the forecasting of the wind speed time series matches well with the actual observed data cluster values. Besides, at this stage, the forecasting results should be analyzed in more detail; i.e., Fig. 8 initially provides a close look at the testing results of hourly WS samples for the measuring stations, and it demonstrated ANFIS-FCM forecasting outcomes. Histogram and regression plots are indicated in Fig. 9, considering the

observed actual values demonstrated according to the predictions obtained by the ANFIS-FCM on the WS values. The histogram results in MS1 and MS3 appear to be skewed to the positive, while the result in MS2 appears to be skewed to the negative. The histogram results in MS1 and MS3 appear to be skewed to the positive, while the result in MS2 appears to be skewed to the negative. While more than 92% of the model predictions have an error $\leq \pm 2$ m/s for MS1 and MS2; more than 99.5% of the model predictions have the same error for MS3. The R results of 0.9492, 0.9129, and 0.8848 for MS1, MS2, and MS3 indicate a correlation between the actual and forecasted value distributions. Also, the results lead to understanding how accurately the prediction values of the applied model match with the data of the actual observed values.

4.3 Results of the ANFIS-SC model

The values of different evaluation criteria regarding the ANFIS-SC model are given in Table 5. In this regard, the

Table 4 Optimal parameter determination for WS prediction using ANFIS-FCM model

| Station | Number of MFs | Input number | Max epoch | MAE (m/s) | RMSE (m/s) | R |
|---------|---------------|--------------|------------|---------------|---------------|---------------|
| MS1 | 2 | 5 | 100 | 0.8682 | 1.2294 | 0.9489 |
| | 3 | 5 | 100 | 0.8697 | 1.2324 | 0.9486 |
| | 4 | 5 | 100 | 0.8673 | 1.2291 | 0.9489 |
| | 5 | 5 | 100 | 0.8669 | 1.2273 | 0.9490 |
| | 6 | 5 | 100 | 0.8669 | 1.2260 | 0.9492 |
| | 7 | 5 | 100 | 0.8719 | 1.2340 | 0.9486 |
| | 8 | 5 | 100 | 0.8701 | 1.2341 | 0.9486 |
| | 9 | 5 | 100 | 0.8697 | 1.2620 | 0.9484 |
| | 10 | 5 | 100 | 0.8680 | 1.2318 | 0.9488 |
| | MS2 | 2 | 5 | 100 | 0.9634 | 1.2606 |
| 3 | | 5 | 100 | 0.9762 | 1.2692 | 0.9118 |
| 4 | | 5 | 100 | 0.9962 | 1.3054 | 0.9066 |
| 5 | | 5 | 100 | 0.9699 | 1.2895 | 0.9091 |
| 6 | | 5 | 100 | 0.9906 | 1.2841 | 0.9101 |
| 7 | | 5 | 100 | 0.9799 | 1.2997 | 0.9079 |
| 8 | | 5 | 100 | 1.0104 | 1.3428 | 0.9019 |
| 9 | | 5 | 100 | 0.9986 | 1.3312 | 0.9032 |
| 10 | | 5 | 100 | 1.0219 | 1.3607 | 0.8994 |
| MS3 | | 2 | 5 | 100 | 0.6176 | 0.7741 |
| | 3 | 5 | 100 | 0.6182 | 0.7709 | 0.8848 |
| | 4 | 5 | 100 | 0.6328 | 0.7950 | 0.8785 |
| | 5 | 5 | 100 | 0.6365 | 0.8085 | 0.8747 |
| | 6 | 5 | 100 | 0.6315 | 0.7968 | 0.8782 |
| | 7 | 5 | 100 | 0.6785 | 0.8507 | 0.8628 |
| | 8 | 5 | 100 | 0.6764 | 0.8426 | 0.8635 |
| | 9 | 5 | 100 | 0.6801 | 0.8491 | 0.8660 |
| | 10 | 5 | 100 | 0.6748 | 0.8420 | 0.8635 |

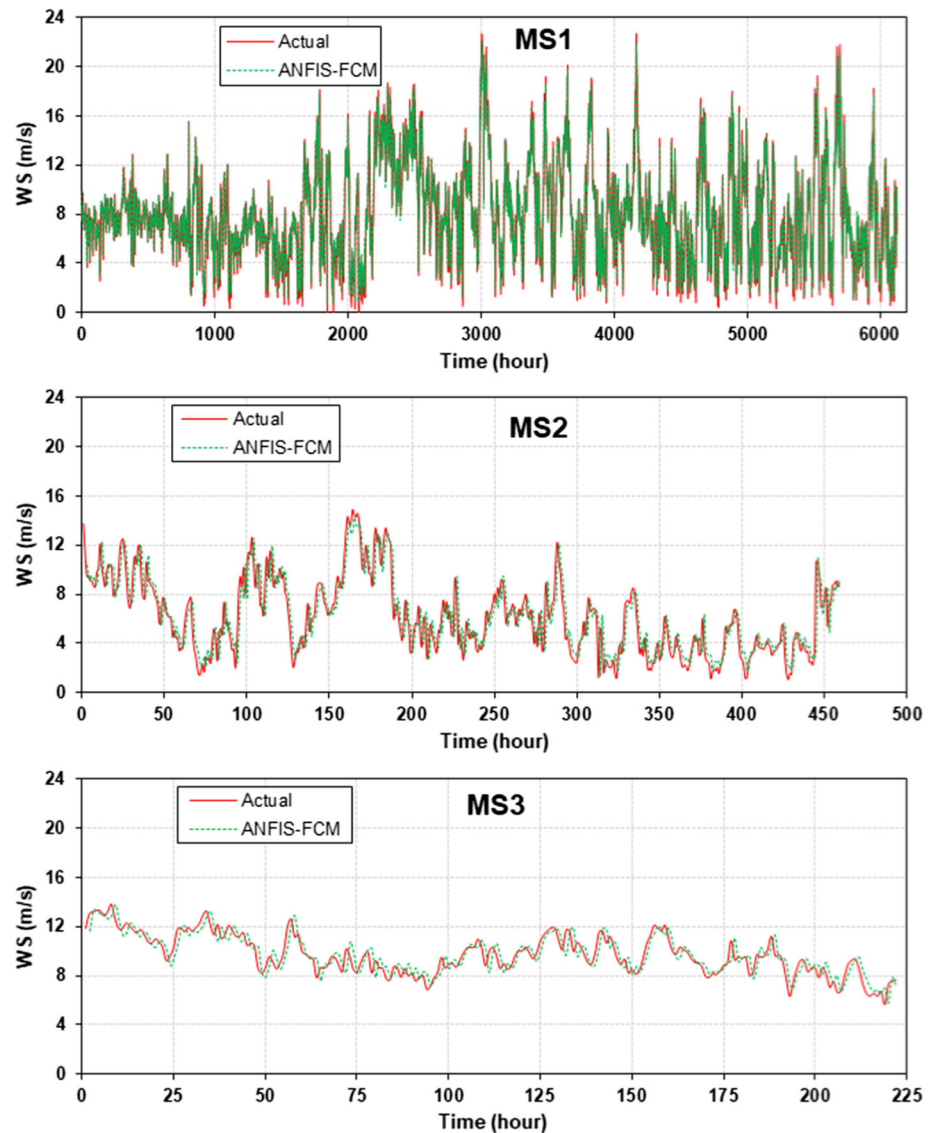
The best results are shown in bold

applied forecasting methodology using this method was studied under variant cluster radius sets, including ranges of $0.2 \leq R \leq 0.8$ for MS1 and $0.2 \leq R \leq 0.9$ for MS2 and MS3. Table 4 presented the forecasting outcomes obtained from the testing process of the ANFIS-SC model. As seen from the table, all the tried ANFIS-SC models presented more or less similar results in respect of the precision measures. Nevertheless, the radius of 0.3 for MS1 demonstrated a little better MAE, RMSE, and R statistics. Similarly, considering the measuring station MS2, radius 0.9 provided better results again in MAE, RMSE, and R statistics. Finally, for MS3, again, the radius of 0.9 came off with better results in terms of all the statistical parameters. For the whole measuring stations, at these radii values, the MAE, RMSE, and R statistics were reported as 0.8644 m/s, 1.2250 m/s, and 0.9494, respectively for MS1, whereas, 0.9925 m/s, 1.2838 m/s, and 0.9098, respectively for MS2, and finally, 0.6123 m/s, 0.7683 m/s, and 0.8848 for MS3. These results presented that low values of the

cluster radii for MS2 and MS3 did not permit the excellent mapping of the model. On the other hand, high values of the cluster radii for MS1 caused some difficulties in the training and caused over-fitting or undesirable input memorization.

Figure 10 exhibits the tested time series regarding forecasted WS hourly data obtained by the ANFIS-SC model for the observed actual values. The proposed ANFIS-SC model performed the prediction results for three measuring stations are shown in this figure. As seen in this figure, the ANFIS-SC estimations of the WS time series for three wind farms coincide well with the values of the actual observed data cluster. Additionally, to obtain a detailed analysis of the three measuring stations, it is important to mention that Fig. 10 shows a close look at the hourly wind speed samples used in the testing stage and demonstrated in Fig. 5. So, Fig. 10 indicates the ANFIS-SC prediction results of these test data for these wind farms. Besides, Fig. 11 displays the histogram and regressions of the

Fig. 8 The hourly time series wind speed data cluster with actual values and forecasted values for the suggested ANFIS-FCM model, at the testing stage of the WS data with the usage of the testing data cluster; the predicting operation conducted for the stations



ANFIS-SC model according to the WS forecasted values shown considering the measured values. The histogram results in MS1 and MS3 appear to be skewed to the positive and, results obtained for MS2 tend to show the closest thing to a normal distribution, with an almost symmetrical distribution of about zero. In the MS2 station, the error distribution follows a Gaussian curve. The histogram part of this figure states that, while more than 93% of the model predictions have an error $\leq \pm 2$ m/s for MS1 and MS2; more than 99.5% of the model predictions have the same error for MS3. The R results of the ANFIS model, including the values of 0.9494, 0.9098, and 0.8848 for three stations, have been also presented in Fig. 11. These R results represent the actual and corresponding forecasted data distributions and indicate the high precision of the forecasting model concerning the real measured data.

4.4 Results of the ANFIS-GP model

The forecasting methodology was also applied to the model of ANFIS-GP. In this ANFIS model, Gaussian membership and linear membership functions were used regarding input and output data. The trials were executed at the number of MFs corresponding to 2 for MS1, 2, and 3 for MS2 and 2 and 3 for MS3. The input number and the maximum epoch number for those trials were set to 5 and 100, respectively, as shown in Table 6.

The testing time series, including the measured actual and forecasted hourly WS data obtained by the ANFIS-GP model, are shown in Fig. 12 for three measuring stations. As observed in subparts of the figure, in the testing part of the ANFIS-GP model, it is observed that the guesses of the WS time series coincide well with the actual measured values. To analyze the ANFIS-GP forecasting results in

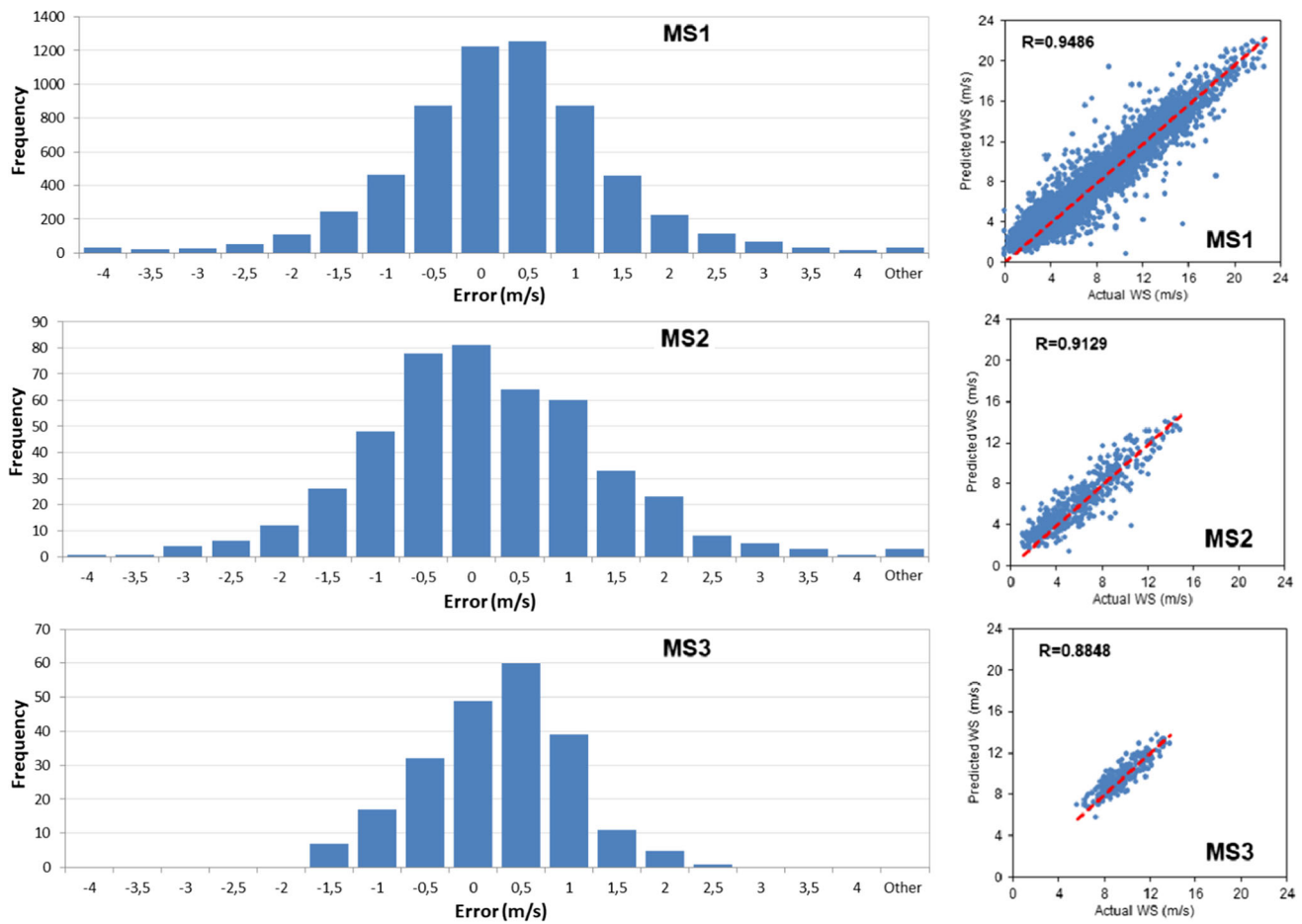


Fig. 9 Histogram and statistical regression graphs regarding the predicted vs the actual values of the WS data cluster concerning the proposed ANFIS-FCM model; the results presented for the stations

more detail, Fig. 12 is presented to display a closer look at the testing part of Fig. 5. Figure 12 also involves the predicted samples of the method concerning the observed time series considering three measuring stations at the testing stage. The histogram and regression plots on actual and predicted data of WS with the utilization of the ANFIS-GP model are exhibited in Fig. 13. The results in all stations obtained from this model appear to be skewed to the positive. More than 93% of the model predictions have an error $\leq \pm 2$ m/s for MS1 and MS2 and more than 99.5% of the model predictions have the same error for MS3 as can be seen from this figure. As demonstrated in Table 6, the MAE, RMSE, and R statistical values were computed as 0.8821 m/s, 1.3186 m/s, and 0.9411, respectively, at MS1. The results were obtained much better when the number of MFs was set to 2 instead of 3 at MS2 and MS3. The comparisons based on the statistical results of MAE, RMSE, and R demonstrate this situation. The statistical results indicated that the values of 1.1097 m/s MAE, 1.5461 m/s RMSE, and 0.8758 R were obtained when the number of MFs was set 2 for MS2. Similarly, at MS3,

0.8518 m/s MAE, 1.1419 m/s RMSE, and 0.7748 R were obtained when the number of MFs was set 2.

4.5 Comparison of the results

Table 7 represents the statistical accuracy results of the WS predictions for each method and different stations. This table presents that the best WS predictions have been obtained with the ANFIS-FCM method. This method can be a helpful tool in the forecasting of WS with high accuracy.

To evaluate and analyze the accuracy and performance of the results of all models, statistical data were compared with test data of WS models in the literature. All models were found to be highly reliable and predictive in the study. However, the proposed LSTM model provided the best results in predictive performance. Figure 14 depicts a comparison of actual and predicted daily WS values for 2-day time periods for this purpose.

Figure 15 illustrates the Taylor diagrams (Taylor 2001) of the proposed model’s errors for the test data. A Taylor

Table 5 Optimal parameter determination for WS prediction using ANFIS-SC model

| Station | Influence radius | Input number | Max epoch | MAE (m/s) | RMSE (m/s) | R |
|---------|------------------|--------------|------------|---------------|---------------|---------------|
| MS1 | 0.2 | 5 | 100 | 0.8674 | 1.2310 | 0.9488 |
| | 0.3 | 5 | 100 | 0.8644 | 1.2250 | 0.9494 |
| | 0.4 | 5 | 100 | 0.8647 | 1.2270 | 0.9492 |
| | 0.5 | 5 | 100 | 0.8656 | 1.2279 | 0.9491 |
| | 0.6 | 5 | 100 | 0.8656 | 1.2277 | 0.9491 |
| | 0.7 | 5 | 100 | 0.8688 | 1.2309 | 0.9489 |
| | 0.8 | 5 | 100 | 0.8711 | 1.2341 | 0.9486 |
| MS2 | 0.2 | 5 | 100 | 1.0351 | 1.3546 | 0.9007 |
| | 0.3 | 5 | 100 | 1.0235 | 1.3385 | 0.9027 |
| | 0.4 | 5 | 100 | 1.0007 | 1.2920 | 0.9086 |
| | 0.5 | 5 | 100 | 1.0036 | 1.3013 | 0.9073 |
| | 0.6 | 5 | 100 | 1.0026 | 1.2983 | 0.9077 |
| | 0.7 | 5 | 100 | 0.9942 | 1.2862 | 0.9095 |
| | 0.9 | 5 | 100 | 0.9925 | 1.2838 | 0.9098 |
| MS3 | 0.2 | 5 | 100 | 0.7847 | 1.0278 | 0.8133 |
| | 0.3 | 5 | 100 | 0.6805 | 0.8572 | 0.8611 |
| | 0.4 | 5 | 100 | 0.6486 | 0.8141 | 0.8729 |
| | 0.5 | 5 | 100 | 0.6363 | 0.7910 | 0.8793 |
| | 0.6 | 5 | 100 | 0.6296 | 0.7821 | 0.8817 |
| | 0.7 | 5 | 100 | 0.6213 | 0.7753 | 0.8838 |
| | 0.9 | 5 | 100 | 0.6123 | 0.7683 | 0.8848 |

The best results are shown in bold

diagram, as is well known, can be used to assess the accuracy of suggested models in a variety of ways. The overall similarity of the models to the measured values is inversely proportional to the distance between the points representing the models and the point representing the observed (actual) value. In comparison to the models for all stations, Fig. 15 clearly shows that the LSTM model predictions are closer to the observations. The lowest RMSE values were obtained as 1.2193 m/s, 1.2573 m/s and 0.7531 m/s in MS1, MS2 and MS3, respectively when the LSTM method was used as can be seen from Table 7.

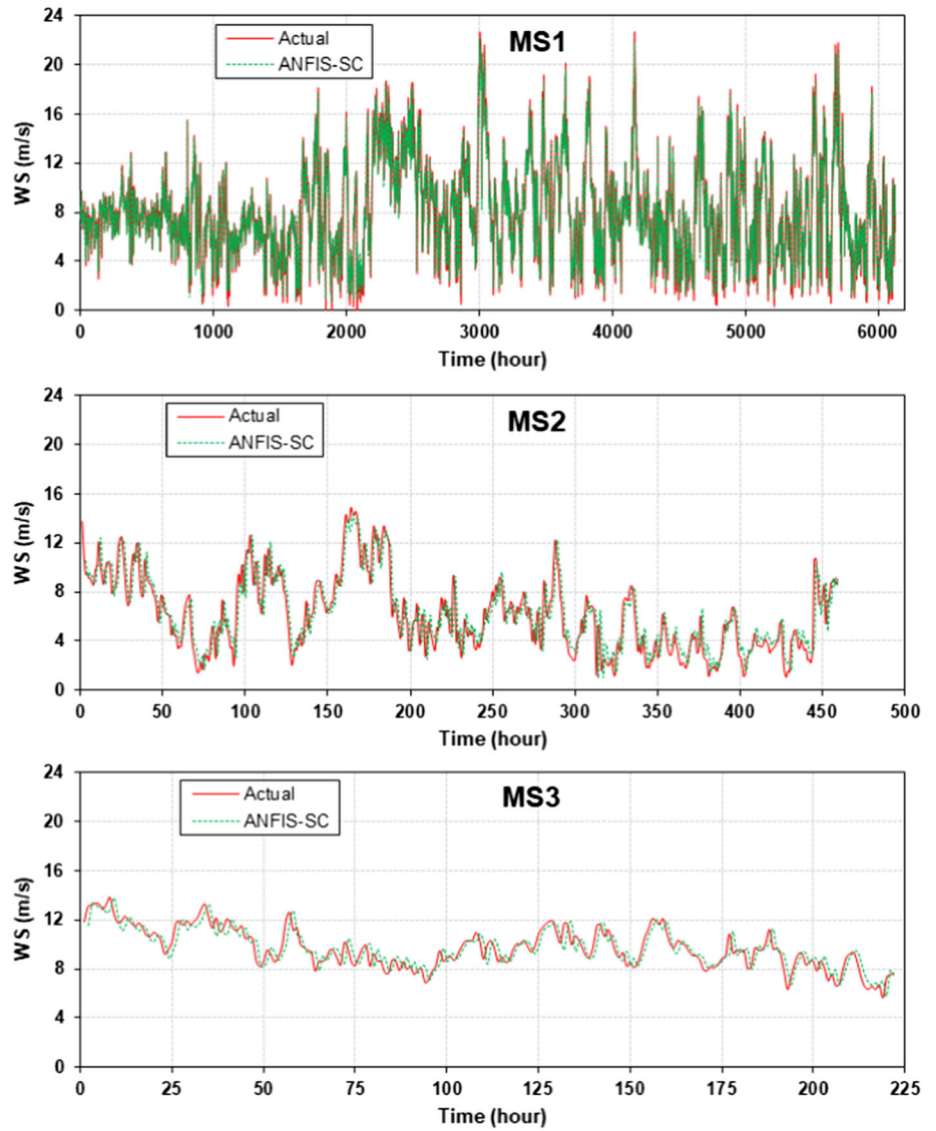
5 Conclusions

In this study, WS forecasting was performed using ANFIS-(FCM, GP, SC) and long short-term memory (LSTM) methods. A total of 20,424, 1531, and 743 hourly data were used in MS1, MS2, and MS3, respectively. The testing process of the methods was compared according to the statistical parameters, including the MAE (m/s), RMSE (m/s), and R parameters. Generally, high WS prediction accuracies were obtained in the used methods. Considering the measuring stations MS1, MS2, and MS3, the error

analyses indicated that the LSTM structure provided the best results in three statistical error results. If to summarize statistical data regarding LSTM, this method provided 0.8638 m/s MAE, 1.2193 m/s RMSE, and 0.9498 R for MS1. Whereas considering MS2, the values of 0.9603 m/s, 1.2573 m/s, and 0.9147 were obtained for MAE, RMSE, and R, respectively. Finally, for MS3, the best results regarding the statistical parameters were computed to be 0.5977 m/s for MAE, 0.7531 m/s for RMSE, and 0.8897 for R, respectively.

The time-series approach based on the LSTM neural network estimates is made by considering the hidden periodicity in the data. This approach has the significant advantage of using univariate modeling as an independent variable of the data obtained in the past. Future work can be focused on using hybrid deep learning functions and architectures to improve the precision and accuracy of predictive results. More accurate estimation of the WS can reduce costs and risks and increase the power system's security can help managers develop a program of action and optimal management of the electricity grid. Thus, its economic and social benefits can be increased.

Fig. 10 The hourly time series WS data cluster with actual values and forecasted values for the proposed ANFIS-SC model, at the testing stage of the WS data with the testing data cluster utilization; the estimation operation conducted for the stations



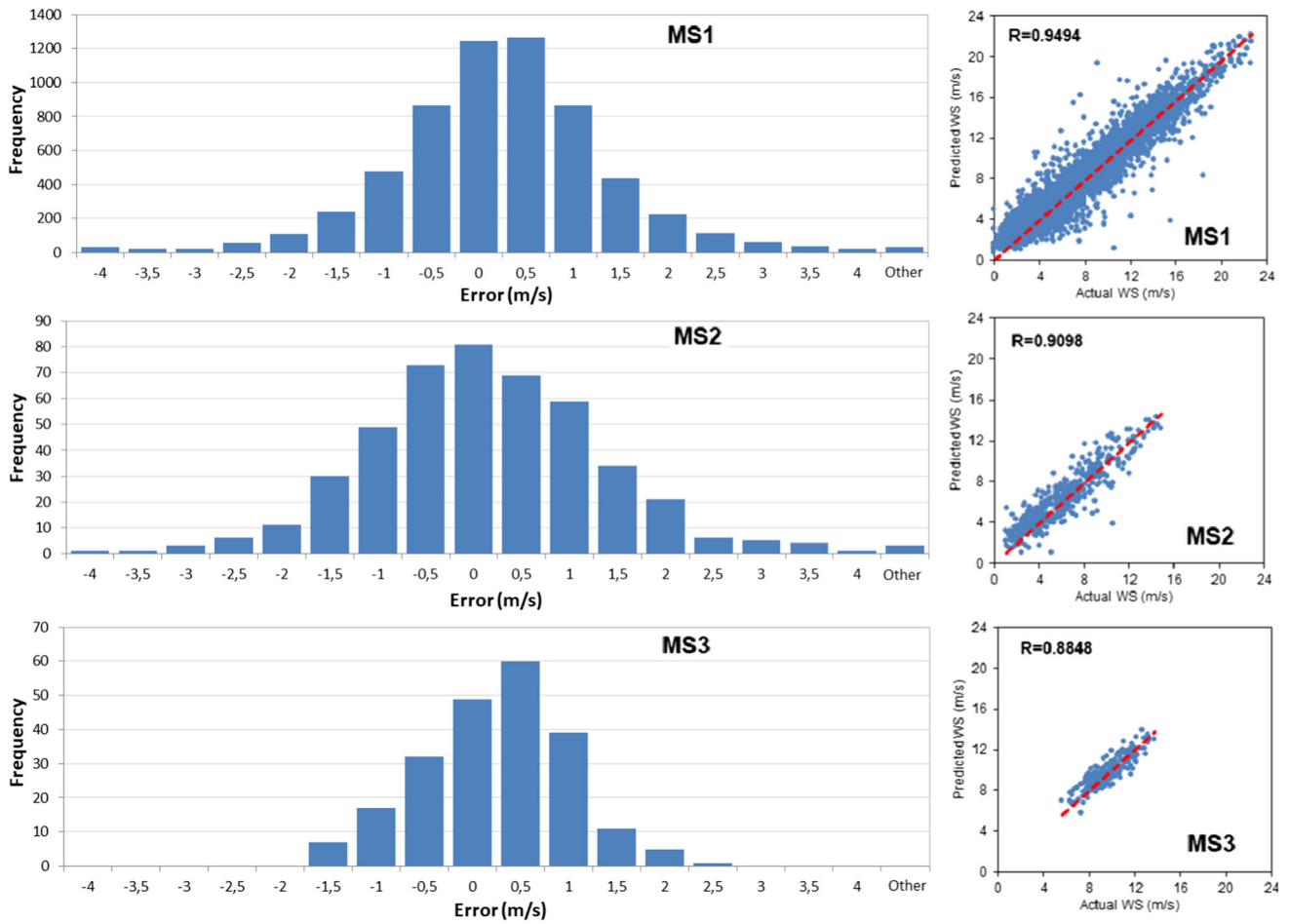


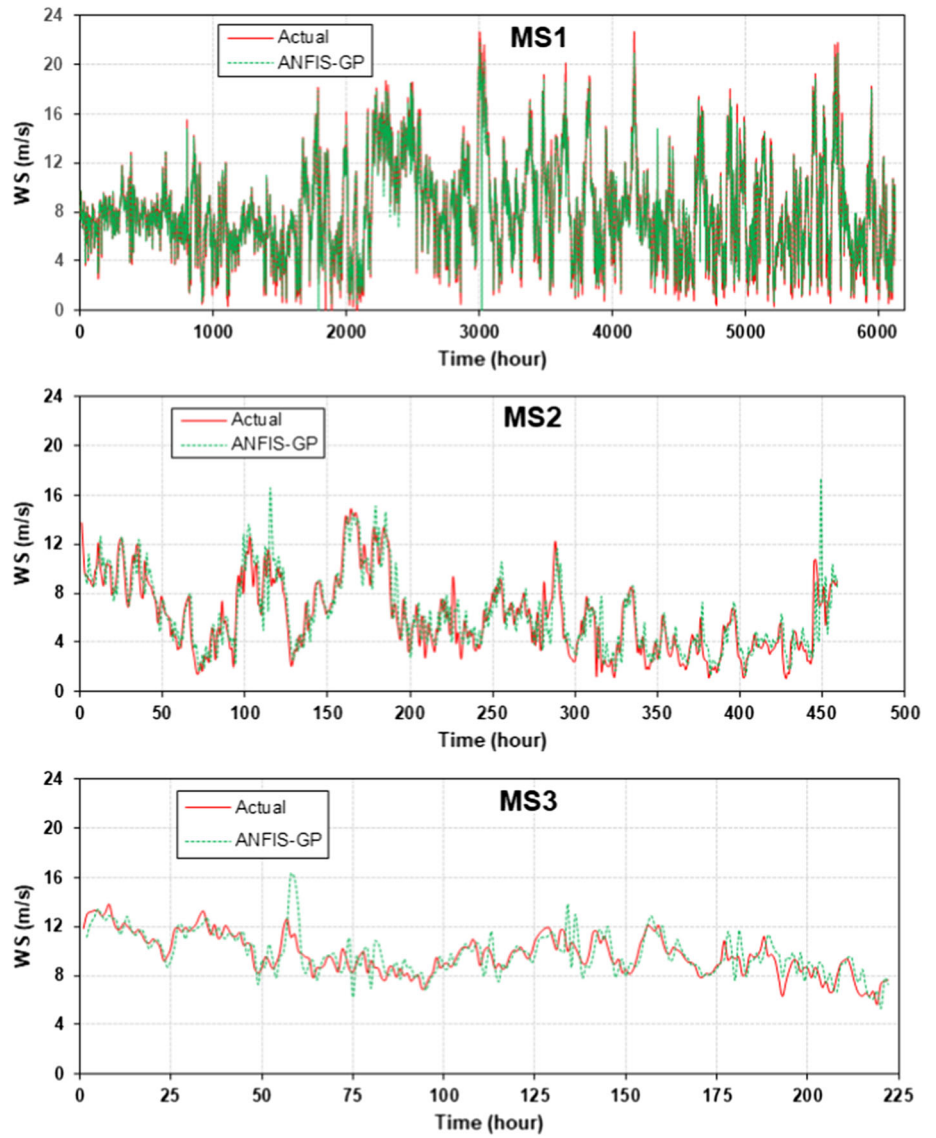
Fig. 11 Histogram and statistical regression graphs regarding the predicted vs the actual values of the WS data cluster into account and using the recommended ANFIS-SC model; the model outputs presented for the stations

Table 6 Optimal parameter determination for WS prediction using ANFIS-GP model

| Station | Number of MFs | Input number | Max epoch | MAE (m/s) | RMSE (m/s) | R |
|---------|---------------|--------------|------------|---------------|---------------|---------------|
| MS1 | 2 | 5 | 100 | 0.8821 | 1.3186 | 0.9411 |
| MS2 | 2 | 5 | 100 | 1.1097 | 1.5461 | 0.8758 |
| | 3 | 5 | 100 | 2.4987 | 8.6798 | 0.2730 |
| MS3 | 2 | 5 | 100 | 0.8518 | 1.1419 | 0.7748 |
| | 3 | 5 | 100 | 1.9331 | 3.1643 | 0.4700 |

The best results are shown in bold

Fig. 12 The hourly time series WS data cluster with observed values and guessed values for the proffered ANFIS-GP model, at the testing step of the WS data with the employment of the testing data cluster; the approximation operation conducted for the stations



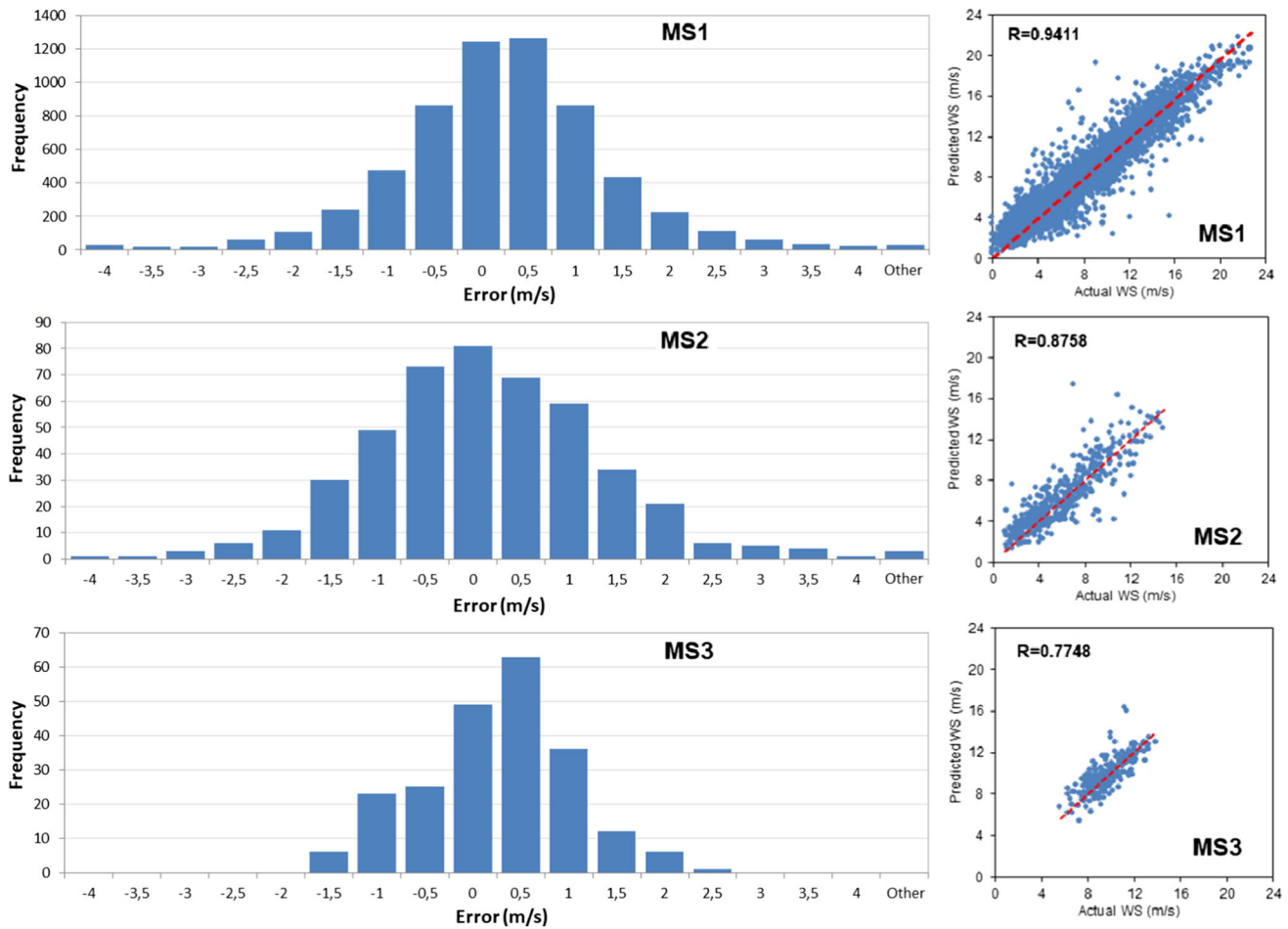


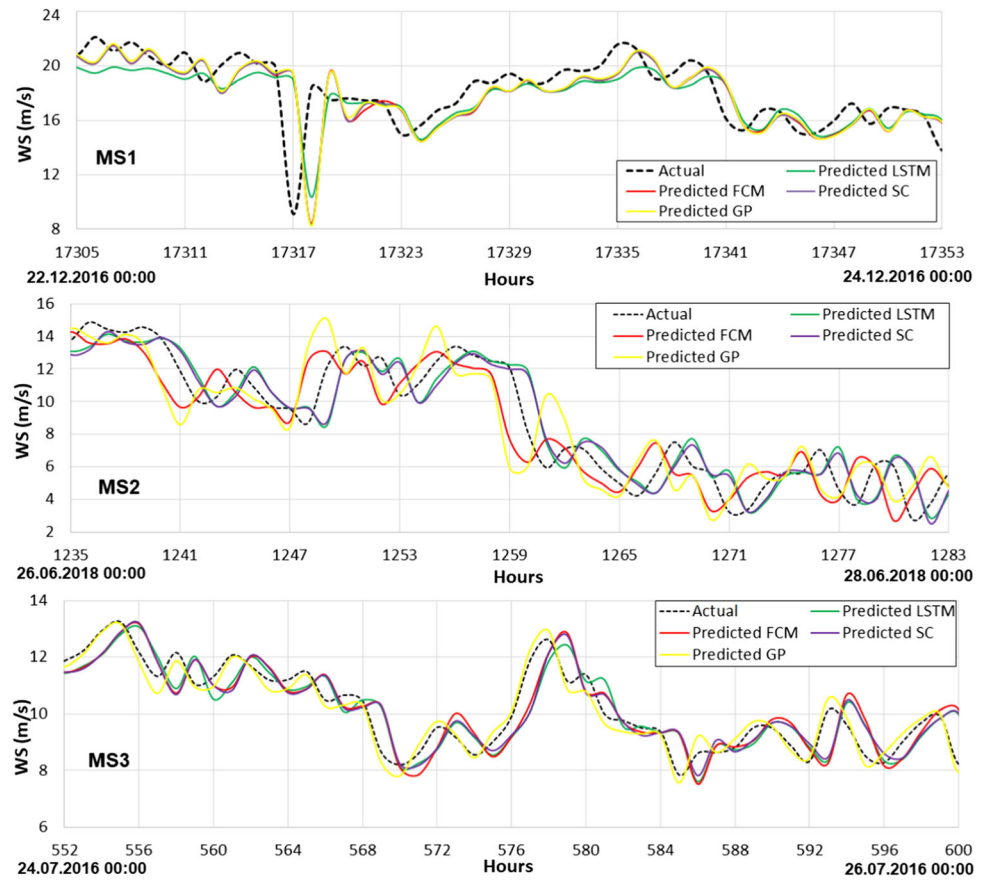
Fig. 13 Histogram and statistical regression graphs regarding the predicted vs the actual values of the WS data cluster and using the recommended ANFIS-GP model; the model outcomes presented for the stations

Table 7 Statistical accuracy results of the WS prediction

| Station | Forecasting Method | MAE (m/s) | RMSE (m/s) | R |
|---------|--------------------|---------------|---------------|---------------|
| MS1 | LSTM | 0.8638 | 1.2193 | 0.9498 |
| | ANFIS FCM | 0.8669 | 1.2260 | 0.9486 |
| | ANFIS SC | 0.8644 | 1.2250 | 0.9494 |
| | ANFIS GP | 0.8821 | 1.3186 | 0.9411 |
| MS2 | LSTM | 0.9603 | 1.2573 | 0.9147 |
| | ANFIS FCM | 0.9634 | 1.2606 | 0.9129 |
| | ANFIS SC | 0.9925 | 1.2838 | 0.9098 |
| | ANFIS GP | 1.1097 | 1.5461 | 0.8758 |
| MS3 | LSTM | 0.5977 | 0.7531 | 0.8897 |
| | ANFIS FCM | 0.6182 | 0.7709 | 0.8848 |
| | ANFIS SC | 0.6123 | 0.7683 | 0.8848 |
| | ANFIS GP | 0.8518 | 1.1419 | 0.7748 |

The best results are shown in bold

Fig. 14 Comparison of the models for the stations



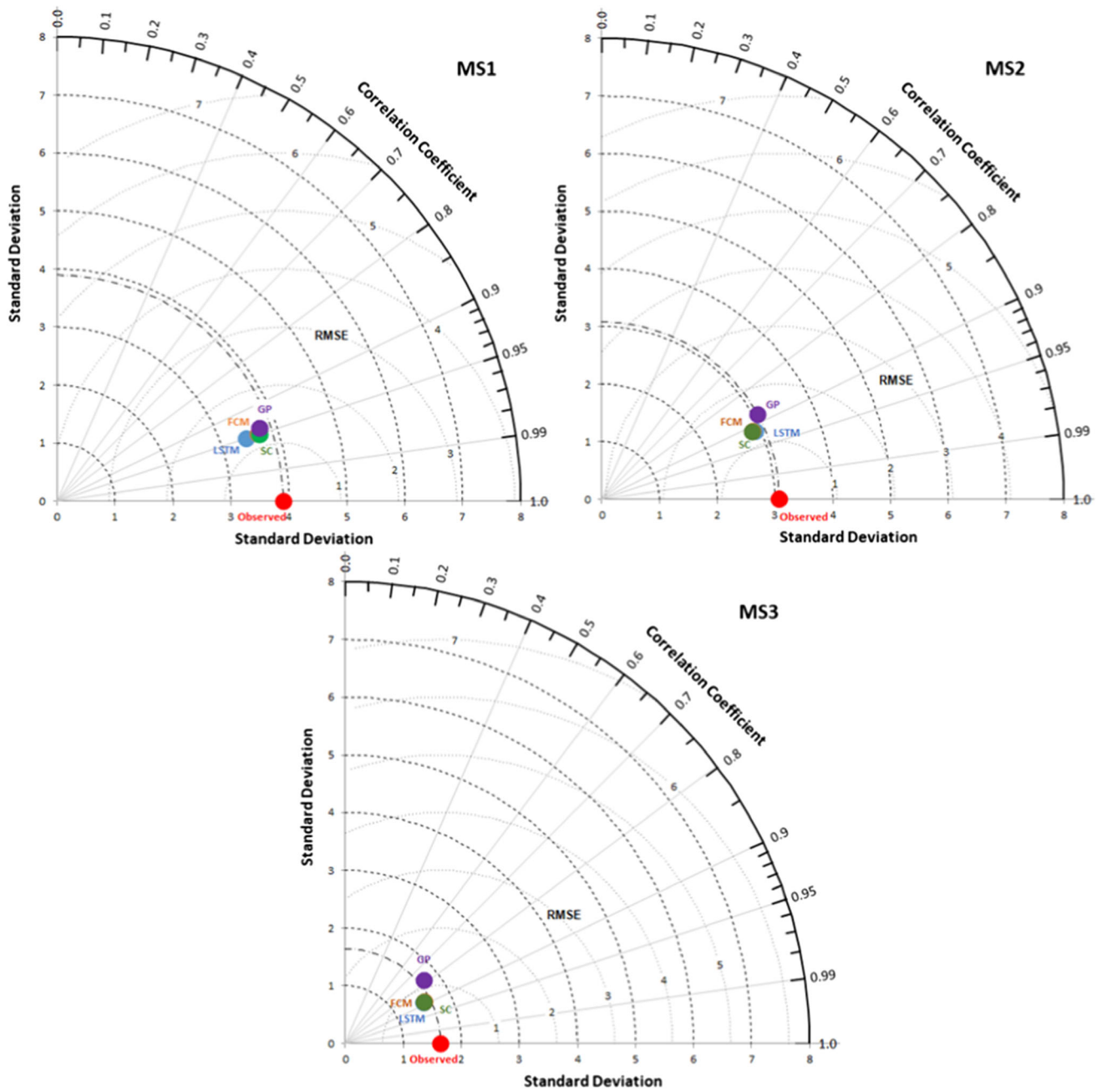


Fig. 15 Taylor diagrams of the proposed model’s errors for the test data

Appendix

Some parts of the software/code used for implementing the machine learning algorithm:

```

dataTrain = data(1:numTimeStepsTrain+1);
dataTest = data(numTimeStepsTrain+1:end);
mu = mean(dataTrain);
sig = std(dataTrain);
dataTrainStandardized = (dataTrain - mu) / sig;
XTrain = dataTrainStandardized(1:end-1);
YTrain = dataTrainStandardized(2:end);
numFeatures = 1;
numResponses = 1;
numHiddenUnits = 10;
layers = [ ...
    sequenceInputLayer(numFeatures)
    lstmLayer(numHiddenUnits)
    fullyConnectedLayer(numResponses)
    regressionLayer];
options = trainingOptions('adam', ...
    'MaxEpochs',300, ...
    'GradientThreshold',1, ...
    'InitialLearnRate',0.005, ...
    'LearnRateSchedule','piecewise', ...
    'LearnRateDropPeriod',125, ...
    'LearnRateDropFactor',0.2, ...
    'Verbose',0, ...
    'Plots','training-progress');

net = trainNetwork(XTrain,YTrain,layers,options);
dataTestStandardized = (dataTest - mu) / sig;
XTest = dataTestStandardized(1:end-1);
net = predictAndUpdateState(net,XTrain);
[net,YPred] = predictAndUpdateState(net,YTrain(end));
numTimeStepsTest = numel(XTest);
for i = 2:numTimeStepsTest
    [net,YPred(:,i)] = predictAndUpdateState(net,YPred(:,i-1),'ExecutionEnvironment','cpu');
end
YPred = sig*YPred + mu;
YTest = dataTest(2:end);
rmse = sqrt(mean((YPred-YTest).^2))

```

Declarations

Conflict of interest The authors do not have any conflicts of interest.

References

- Abyaneh HZ, Nia AM, Varkeshi MB, Marofi S, Kisi O (2011) Performance evaluation of ANN and ANFIS models for estimating garlic crop evapotranspiration. *J Irrig Drain Eng* 137:280–286
- Ahmed AS (2018) Wind energy characteristics and wind park installation in Shark El-Ouinat. *Egypt Renew Sust Energ Rev* 82:734–742
- Balluff S, Bendfeld J, Krauter S (2015) Short term wind and energy prediction for offshore wind farms using neural networks. In: 4th international conference on renewable energy research and applications, Palermo, Italy, 22–25 Nov. 2015
- Barbounis TG, Theocharis JB, Alexiadis MC, Dokopoulos PS (2006) Long-term wind speed and power forecasting using local recurrent neural network models. *IEEE Trans Energy Convers* 21(1):273–284
- Benmouiza K, Cheknane A (2019) Clustered ANFIS network using fuzzy c-means, subtractive clustering, and grid partitioning for hourly solar radiation forecasting. *Theor Appl Climatol* 137:31–43
- Bilgili M, Sahin B (2009) Investigation of wind energy density in the southern and southwestern region of Turkey. *J Energy Eng* 135(1):12–20

- Bilgili M, Sahin B (2010) Comparative analysis of regression and artificial neural network models for wind speed prediction. *Meteorol Atmos Phys* 109:61–72
- Bilgili M, Sahin B, Sangun L (2013) Estimating soil temperature using neighboring station data via multi-nonlinear regression and artificial neural network models. *Environ Monit Assess* 185(1):347–358
- Bilgili M, Ozbek A, Sahin B, Kahraman A (2015) An overview of renewable electric power capacity and progress in new technologies in the World. *Renew Sust Energ Rev* 49:323–334
- Braun T, Waechter M, Peinke J, Guhr T (2020) Correlated power time series of individual wind turbines: A data driven model approach. *J Renew Sustain Energy* 12(023301):1–13
- Bulut U, Muratoglu G (2018) Renewable energy in Turkey: great potential, low but increasing utilization, and an empirical analysis on renewable energy-growth nexus. *Energy Policy* 123:240–250
- Cali U, Sharma V (2019) Short-term wind power forecasting using long-short term memory based recurrent neural network model and variable selection. *Int J Smart Grid Clean Energy* 8(2):103–110
- Chandy KM, Taylor S (1992) An introduction to parallel programming. Jones and Bartlett, Boston
- Chen J, Zeng GQ, Zhou W, Du W, Lu KD (2018) Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization. *Energy Convers Manag* 165:681–695
- Chen MR, Zeng GQ, Lu KD, Weng J (2019a) A two-layer nonlinear combination method for short-term wind speed prediction based on ELM, ENN, and LSTM. *IEEE Internet Things J* 6(4):6997–7010
- Chen Y, Zhang S, Zhang W, Peng J, Cai Y (2019b) Multifactor spatio-temporal correlation model based on a combination of convolutional neural network and long short-term memory neural network for wind speed forecasting. *Energy Convers Manag* 185:783–799
- Ding L, Bai Y, Liu M-D, Fan M-H, Yang J (2022) Predicting short wind speed with a hybrid model based on a piecewise error correction method and Elman neural network. *Energy* 244:122630. <https://doi.org/10.1016/j.energy.2021.122630>
- Dong D, Sheng Z, Yang T (2018) Wind power prediction based on recurrent neural network with long short-term memory units. In: 2018 IEEE international conference on renewable energy and power engineering, Toronto, ON, Canada, pp. 34–38, 24–26 Nov. 2018
- Ehsan MA, Shahirinia A, Zhang N, Oladunni (2020) T Wind speed prediction and visualization using long short-term memory networks (LSTM). In: 10th International Conference on Information Science and Technology (ICIST 2020), London, England, 9–15 Sep. 2020
- EWEA (2020) Wind energy in Europe in 2018. Wind Europe Business Intelligence. <https://windeurope.org/about-wind/statistics/european/wind-energy-in-europe-in-2018/> (2018). Accessed 14 November 2020
- Gualtieri G (2019) A comprehensive review on wind resource extrapolation models applied in wind energy. *Renew Sust Energ Rev* 102:215–233
- GWEC (2020) Global wind report, 2018. Global Wind Energy Council. <http://www.gwec.net> (2018). Accessed 14 Nov 2020
- Han S, Qiao YH, Yan J, Liu YQ, Li L, Wang Z (2019) Mid-to-long term wind and photovoltaic power generation prediction based on copula function and long short term memory network. *Appl Energy* 239:181–191
- Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9:1735–1780
- Hu YL, Chen LA (2018) Nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and differential evolution algorithm. *Energy Convers Manag* 173:123–142
- Huang Y, Liu S, Yang L (2018) Wind speed forecasting method using EEMD and the combination forecasting method based on GPR and LSTM. *Sustainability* 10(3693):1–15
- Jang JR (1993) ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans Syst Man Cybern* 23(3):665–685
- Jones CR, Eiser JR (2010) Understanding “local” opposition to wind development in the UK: How big is a backyard? *Energy Policy* 38:3106–3117
- Jung J, Broadwater RP (2014) Current status and future advances for wind speed and power forecasting. *Renew Sust Energ Rev* 31:762–777
- Karakuş O, Kuruoğlu EE, Altinkaya MA (2017) One-day ahead wind speed/power prediction based on polynomial autoregressive model. *IET Renew Power Gener* 11(11):1430–1439
- Kavasseri RG, Seetharaman K (2009) Day-ahead wind speed forecasting using f-ARIMA models. *Renew Energy* 34(5):1388–1393
- Kazimierczuk AH (2019) Wind energy in Kenya: A status and policy framework review. *Renew Sust Energ Rev* 107:434–445
- Kılıç B (2019) Determination of wind dissipation maps and wind energy potential in Burdur province of Turkey using geographic information system (GIS). *Sustain Energy Technol Assess* 36(100555)
- Köktürk G, Tokuç A (2017) Vision for wind energy with a smart grid in Izmir. *Renew Sust Energ Rev* 73:332–345
- Liang S, Nguyen L, Jin F (2018) A multi-variable stacked long-short term memory network for wind speed forecasting. In: 2018 IEEE international conference on Big Data (Big Data), Seattle, WA, USA, pp. 4561–4564, 10–13 Dec. 2018
- Liu H, Mi X, Li Y (2018a) Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis, LSTM network and ELM. *Energy Convers Manag* 159:54–64
- Liu H, Mi XW, Li YF (2018b) Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. *Energy Convers Manag* 156:498–514
- Liu H, Tian H, Li Y (2015) An EMD-recursive ARIMA method to predict wind speed for railway strong wind warning system. *J Wind Eng Ind Aerodyn* 141:27–38
- Liu R, Liu L (2019) Predicting housing price in China based on long short-term memory incorporating modified genetic algorithm. *Soft Comput: Methodol Appl* 23:11829–11838
- Liu Y, Guan L, Hou C, Han H, Liu Z, Sun Y, Zheng M (2019) Wind power short-term prediction based on LSTM and discrete wavelet transform. *Appl Sci* 9(1108):1–17
- Lopez E, Valle C, Allende H, Gil E, Madsen H (2018) Wind power forecasting based on echo state networks and long short-term memory. *Energies* 11(526):1–22
- Marndi A, Patra GK, Gouda KC (2020) Short-term forecasting of wind speed using time division ensemble of hierarchical deep neural networks. *Bull Atmos Sci Technol* 1:91–108
- Mathworks: Long Short-Term Memory Networks (2020) <https://www.mathworks.com/help/deeplearning/ug/long-short-term-memory-networks.html>. Accessed 01 May 2020
- Mendecka B, Lombardi L (2019) Life cycle environmental impacts of wind energy technologies: a review of simplified models and harmonization of the results. *Renew Sust Energ Rev* 111:462–480
- Mogos AS, Salauddin M, Liang X, Chung C (2022) An effective very short-term wind speed prediction approach using multiple regression models. *IEEE Can J Electr Comput Eng*. <https://doi.org/10.1109/ICJECE.2022.3152524>

- Özen C, Kaplan O, Özcan C, Dinç U (2019) Short term wind speed forecast by using long short term memory. In: 9th international symposium on atmospheric sciences (ATMOS 2019), İstanbul, Turkey, 23–26 Oct. 2019
- Prabha PP, Vanitha V, Resmi R (2019) Wind speed forecasting using long short term memory networks. In: 2nd international conference on intelligent computing, instrumentation and control technologies (ICICICT), Kannur, Kerala, India, pp 1310–1314, 5–6 July 2019
- Qian P, Tian X, Kanfoud J, Lee JLY, Gan TH (2019) A novel condition monitoring method of wind turbines based on long short-term memory neural network. *Energies* 12(3411):1–15
- Qin Q, Lai X, Zo J (2019a) Direct multi-step wind speed forecasting using LSTM neural network combining EEMD and fuzzy entropy. *Appl Sci* 9(126):1–19
- Qin Y, Li K, Liang Z, Lee B, Zhang F, Gu Y, Zhang L, Wu F, Rodriguez D (2019b) Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal. *Appl Energy* 236:262–272
- Salman AG, Heryadi Y, Abdurahman E, Suparta W (2018) Single layer & multi-layer long short-term memory (LSTM) model with intermediate variables for weather forecasting. *Procedia Comput Sci* 135:89–98
- Shi X, Lei X, Huang Q, Huan S, Ren K, Hu Y (2018) Hourly day-ahead wind power prediction using the hybrid model of variational model decomposition and long short-term memory. *Energies* 11(3227):1–20
- Tabari H, Kisi O, Ezani A, Talae PH (2012) SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. *J Hydrol* 444–445:78–89
- Tagliapietra S, Zachmann G, Fredriksson G (2019) Estimating the cost of capital for wind energy investments in Turkey. *Energy Policy* 131:295–301
- Tascikaraoglu A, Uzunoglu M (2014) A review of combined approaches for prediction of short-term wind speed and power. *Renew Sust Energ Rev* 34:243–254
- Wang J, Li Y (2018) Multi-step ahead wind speed prediction based on optimal feature extraction, long short term memory neural network and error correction strategy. *Appl Energy* 230:429–443
- Wang S, Wang S (2015) Impacts of wind energy on environment: a review. *Renew Sust Energ Rev* 49:437–443
- Wu W, Chen K, Qiao Y, Lu Z (2016) Probabilistic short-term wind power forecasting based on deep neural networks. In: 2016 international conference on probabilistic methods applied to power systems (PMAPS 2016), Beijing, China, 05 Dec. 2016
- Xiaoyun Q, Xiaoning K, Chao Z, Shuai J, Xiuda M (2016) Short-term prediction of wind power based on deep long short-term memory. In: 2016 IEEE PES Asia-Pacific power and energy conference, Xi'an, China, pp 1148–1152, 25–28 Oct. 2016
- Yang M, Liu L, Cui Y, Su X (2018) Ultra-short-term multi-step prediction of wind power based on representative unit method. *Hindawi Math Probl Eng*. Article ID: 1936565
- Yang Q, Deng C, Chang X (2022) Ultra-short-term/short-term wind speed prediction based on improved singular spectrum analysis. *Renewable Energy* 184:36–44. <https://doi.org/10.1016/j.renene.2021.11.044>
- Yaxue R, Yintang W, Fucal L, Yuyan Z (2022) A short-term wind speed prediction method based on interval type 2 fuzzy model considering the selection of important input variables. *J Wind Eng Ind Aerodyn* 225:104990. <https://doi.org/10.1016/j.jweia.2022.104990>
- Yu C, Li Y, Bao Y, Tang H, Zhai G (2018) A novel framework for wind speed prediction based on recurrent neural networks and support vector machine. *Energy Convers Manag* 178:137–145
- Yuan X, Chen C, Jiang M, Yuan Y (2019) Prediction interval of wind power using parameter optimized Beta distribution based LSTM model. *Appl Soft Comput* 82(105550):1–10
- Zahroh S, Hidayat Y, Pontoh RS, Santoso A (2019) Sukono Bon (2019) AT Modeling and forecasting daily temperature in Bandung. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, Riyadh, Saudi Arabia 406–412:26–28
- Zaytar MA, Amrani CE (2016) Sequence to sequence weather forecasting with long short-term memory recurrent neural networks. *Int J Comput Appl* 143(11):7–11
- Zhang J, Yan J, Infiel D, Liu Y, Lien F (2019a) Short-term forecasting and uncertainty analysis of wind turbine power based on long short-term memory network and Gaussian mixture model. *Appl Energy* 241:229–244
- Zhang Y, Gao S, Han J, Ban M (2019b) Wind speed prediction research considering wind speed ramp and residual distribution. *IEEE Access* 7:131873–131887
- Zhang Z, Qin H, Liu Y, Wang Y, Yao L, Li Q, Li J, Pei S (2019c) Long short-term memory network based on neighborhood gates for processing complex causality in wind speed prediction. *Energy Convers Manag* 192:37–51
- Zhang Z, Ye L, Qin H, Liu Y, Wang C, Yu X, Yin X, Li J (2019d) Wind speed prediction method using shared weight long short-term memory network and Gaussian process regression. *Appl Energy* 247:270–284
- Zheng CW, Li CY, Pan J, Liu MY, Xia L (2016) An overview of global ocean wind energy resource evaluations. *Renew Sust Energ Rev* 53:1240–1251

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.