Energy Reports 8 (2022) 638–648

Contents lists available at ScienceDirect

Energy Reports

journal homepage: www.elsevier.com/locate/egyr

Research paper

Feasibility of soft computing techniques for estimating the long-term mean monthly wind speed

Shahab S. Band [a,*], Sina Ardabili [a], Amir Mosavi [b,c,**], Changhyun Jun [d,***], Helaleh Khoshkam [e], Massoud Moslehpour [f,g,****]

[a] Future Technology Research Center, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan
[b] Institute of Information Society, University of Public Service, 1083 Budapest, Hungary
[c] John von Neumann Faculty of Informatics, Obuda University, Budapest, Hungary
[d] Department of Civil and Environmental Engineering, College of Engineering, Chung-Ang University, Seoul 06974, Republic of Korea
[e] Department of Civil and Environmental Engineering and Water Resources Research Center, University of Hawaii at Manoa, Honolulu, HI, 96822, USA
[f] Department of Business Administration, Asia University, 500, Lioufeng Rd., Wufeng, Taichung 41354, Taiwan
[g] Department of Management, California State University, San Bernardino, 5500 University Parkway, San Bernardino, CA 92407, USA

ARTICLE INFO

Article history:
Received 14 August 2021
Received in revised form 12 November 2021
Accepted 22 November 2021
Available online 22 December 2021

Keywords:
Long-term mean monthly wind speed
Energy
Soft computing techniques

ABSTRACT

Estimating wind energy plays an important role in energy science as it can be considered a crucial source of renewable and sustainable energy. In this study, five types of soft computing approaches were implemented to estimate the long-term mean monthly wind speed (W) at 50 weather stations in Iran. The applied models were artificial neural networks (ANN), gene expression programming (GEP), multivariate adaptive regression spline (MARS), adaptive neuro-fuzzy inference system (ANFIS), and random forest (R.F.). In addition, the geographical information (i.e., latitude, longitude, and altitude) and periodicity term (i.e., the number of months in a year) were used to input the models. Results demonstrated that the R.F. technique was the best model for estimating W, utilizing the geographical information and number of the month. Hence, it can be concluded that the applied soft computing techniques can employ the aforementioned inputs for estimating W.

1. Introduction

In recent decades, wind energy has received growing attention because it is clean, renewable, sustainable, and cost-effective (Chen and Yu, 2014; Wang et al., 2018; Zheng et al., 2016, 2019). Knowledge of wind energy is of vital importance to understand its potential in a certain area (Celik and Kolhe, 2013; Zheng et al., 2013; Hur, 2021), protect the safety of conversion systems related to wind energy (Liu et al., 2015; Maroufpoor et al., 2019), and generate electricity via wind turbines (Noorollahi et al., 2016).

Wind speed has the most influence on wind energy. Therefore, its accurate estimation can enhance the performance of wind turbines and power systems (Liu et al., 2014). Wind speed includes inherent stochastic and intermittency characteristics (Chen and Yu, 2014; Jiang and Huang, 2017; Wang et al., 2018). On the other hand, climate change can affect the wind speed time series at a specific region. Hence, an accurate estimation of wind speed is a challenging task. Recently, many studies have been carried out to estimate wind speed by soft computing approaches. Some of the previous works on wind speed forecasting are briefly presented below.

The efficiency of the fractional-autoregressive integrated moving average (f-ARIMA) and persistence approaches in forecasting 24h- and 48h-ahead wind speed was examined by Kavasseri and Seetharaman (2009). The results showed the better performance of f-ARIMA at four sites in North Dakota (USA). Monfared et al. (2009) evaluated the accuracy of fuzzy logic (F.L.), and artificial neural networks (ANN) approaches in modeling the 30-min wind speed at Rostamabad in the north of Iran. F.L. was found to be

* Corresponding author at: Future Technology Research Center, National Yunlin University of Science and Technology, 123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan.
** Corresponding author at: John von Neumann Faculty of Informatics, Obuda University, Budapest, Hungary.
*** Corresponding author at: Department of Civil and Environmental Engineering, College of Engineering, Chung-Ang University, Seoul 06974, Republic of Korea.
**** Corresponding author at: Department of Business Administration, Asia University, 500, Lioufeng Rd., Wufeng, Taichung 41354, Taiwan.
E-mail addresses: shamshirbands@yuntech.edu.tw (S.S. Band), amir.mosavi@nik.uni-obuda.hu (A. Mosavi), cjun@cau.ac.kr (C. Jun), writetodrm@gmail.com (M. Moslehpour).

https://doi.org/10.1016/j.egyr.2021.11.247

more accurate than ANN. Cadenas and Rivera (2009) successfully employed ANN for predicting the short-term wind speed at Oaxaca in Mexico. Fadare (2010) modeled wind speeds at 28 stations in Nigeria, utilizing the ANN approach. Erdem and Shi (2011) applied four types of autoregressive moving average (ARMA) techniques for the short-term forecast of wind speed at two sites in North Dakota (USA). The vector autoregressive (VAR) model outperformed the other types of ARMA methods. Mean daily, and monthly wind speed in Zhangye (China) were modeled by Guo et al. (2012) using the classical feed-forward neural networks (FFNN) and hybrid empirical mode decomposition-FFNN (EMD-FFNN). They concluded that the hybrid model yielded better results than single model. Velo et al. (2014) estimated the wind speed time series at 20 weather stations in Spain, using the multi-layer perceptron (MLP) ANN. They stated that the MLP ANN method could reliably estimate wind speed. Su et al. (2014) recommended the coupled ARIMA-Kalman filter technique for estimating mean daily wind speed at wind farms in the west of China. Wang et al. (2015) developed a coupled model by integrating the least square support vector machine (LSSVM), particle swarm optimization, and gravitational search algorithm to model the wind speed at wind farms in the northwest of China. They compared the performance of their hybrid model with those of ARIMA; radial basis function neural networks (RBFNN), generalized regression neural networks (GRNN), back-propagation neural networks (BPNN), and Elman models. Wang et al. (2016) hybridized the ensemble empirical mode decomposition (EEMD) with genetic algorithm back propagation-neural networks (GABP-NN) to estimate the short-term wind speed at a wind farm in Inner Mongolia (China). Their model outperformed the conventional GA-BPNN approach. Chang et al. (2017) introduced an improved RBFNN with an error feedback scheme (IRBFNN-EF) for forecasting the short-term wind speed at a wind farm in Taiwan. To show the superiority of their model, its outcomes were compared with those of adaptive neuro-fuzzy inference system (ANFIS), BPNN, and RBFNN. Han et al. (2017) coupled the ARMA time series model with soft computing techniques such as the BPNN, SVM, and random forest (R.F.) to model the wind speed. Their results denoted the higher accuracy of hybrid models than those of single models. Sun et al. (2019) forecasted the wind speed at the Sotavento Galicia wind farm in Spain by a hybrid technique. They coupled fast ensemble empirical mode decomposition (FEEMD) with regularized extreme learning machine (RELM) optimized by a backtracking search algorithm (BSA). The hybrid FEEMD-BSA-RELM was found to present superior performance than the conventional RELM. Zhang et al. (2019) utilized a new hybrid approach to estimate the short-term wind speed at two wind farms in Spain and China. They coupled variational mode decomposition-wavelet transformation (VMD-WT) with principal component analysis-back propagation-radial basis function (PCA-BP-RBF). The developed model provided significantly better results compared to those of other models. Qu et al. (2019) modeled the multi-step ahead wind speed at two wind farms in Shandong (China) by coupling the decomposition approach with an improved flower-pollination algorithm-back propagation neural network (FPA-BPNN). They showed the superiority of their model over other approaches. Samadianfard et al. (2020) utilized the hybrid MLP-whale optimization algorithm (WOA) to estimate the wind speed at several sites in Iran. They found out that the coupled MLP-WOA method outperforms MLP.

This study aims to examine the feasibility of five soft computing techniques, namely the multivariate adaptive regression splines (MARS), gene expression programming (GEP), ANN, ANFIS, and R.F., to estimate the long-term mean monthly wind speed (W). The models were tested at 50 stations in Iran. Previous studies estimated wind speed by using climatic variables such as relative humidity, air temperature, air pressure, lagged wind speeds as inputs, and other factors (López and Arboleya, 2021; Li et al., 2022). However, to the best of our knowledge, no studies estimated the wind speed by using only the geographical information (i.e., latitude, longitude, altitude) and the number of months in a year. The organization of this paper is as follows. The methodology is explained in Section 2. In Section 3, the study sites and data are presented. In Section 4, the results are given. Finally, Section 5 reports the conclusions.

2. Methodology

As mentioned above, wind speed has inherent stochastic and intermittency characteristics. Therefore, powerful soft computing approaches are needed to estimate wind speed accurately. Also, the inputs introduced to the soft computing methods play an important role in their performance. The geographical information, including the latitude, longitude, and altitude of sites as well as the number of months, were employed as the predictors of MARS, GEP, ANN, ANFIS, and R.F. to estimate W. These input predictors are readily available even in remote areas where weather data are often inaccessible to feed the soft computing models.

2.1. MARS

MARS was introduced by Friedman (1991). It is a non-parametric and non-linear regression technique. MARS does not need any prior information on the basic functional relationship among the independent and dependent variables. A set of basic functions (B.F.s) is constructed during the learning phase of MARS. The B.F.s define the relationship between a target variable and a series of input data. A general type of a B.F. can be shown by,

$$\begin{aligned} BF_m(x) &= \max(0; x - c) \\ BF_m(x) &= \max(0; c - x) \end{aligned} \tag{1}$$

where $BF_m(x)$, c, and x denote the mth basis function, a threshold value, and an input predictor, respectively.

The MARS model estimates the response variable via an iterative two-step procedure consisting of forwarding and backward phases. First, all of the possible B.F.s are added to the model in the forward stage. This phase creates an over-fitted model. Then, the backward stage is pruned by B.F.s with the least contribution in building the optimal MARS model. Thus, the backward phase enhances the efficiency of this model. Finally, the MARS model with the lowest generalized cross-validation (GCV) is chosen as the optimal one.

A typical form of MARS to estimate a response variable (y) can be shown as follows:

$$y = \alpha_0 + \sum_{m=1}^{M} \alpha_m BF_m(x) \tag{2}$$

where a_o is the intercept, BF_m is the mth B.F., a_m is its coefficient, and M is the number of B.F.s.

2.2. GEP

Ferreira (2001) proposed the GEP as an evolutionary algorithm. The features of genetic programming and genetic algorithm were merged to build the GEP. First, we create a random population in the GEP. Then some chromosomes are generated from the initial population. Next, these chromosomes are encoded with the shape of expression trees. A fitness function finally evaluates the outcomes to determine the suitability of the solution. If the solution is found to be the best one, the evolution process will be stopped. Otherwise, the solution from the prior generation

Table 1
The magnitude of various parameters of the GEP model.

Head size	7
Number of chromosomes	30
Number of genes per chromosome	3
Mutation rate	0.044
Inversion rate	0.1
IS transposition rate	0.1
RIS transposition rate	0.1
One point recombination rate	0.3
Two point recombination rate	0.3
Gene recombination rate	0.1
Gene transposition rate	0.1

will be transferred to the next generation. In this context, the chromosomes are modified by applying some genetic operators. The evolution procedure will be continued until the best solution is achieved.

The modeling process via the GEP is comprised of the following five steps:

• Defining a fitness function for assessing the performance of GEP. Root mean square error (RMSE) was utilized in this study as the fitness function.

• Introducing the terminal and function sets. The terminal set includes inputs (i.e., geographical information and periodicity term) and output (i.e., mean monthly wind speed) variables. Additionally, $(+, -, \times, \div, \ln x, e^x, x^2, x^3, \sqrt{x}, \sqrt[3]{x}, \sin x, \cos x, and \arctan x)$ were applied as the function set.

• Assigning the chromosomal architecture consisting of the head size, the number of chromosomes, and the number of genes per chromosome.

• Determining a linking function. Herein, an addition linking function was applied.

• Assigning the rates of various genetic operators.

Table 1 shows the magnitude of parameters of GEP.

2.3. ANN

Recently, ANN has been increasingly used to solve engineering problems. It consists of input, hidden, and output layers and the layers act as the input receiver layer, the information processing layer, and the output generator. Each layer consists of neurons. Every neuron receives the inputs, processes them, and produces an output signal.

Here, the feed-forward back propagation (FFBP) ANN approach was applied to estimate W. The performance of FFBP ANN mainly depends on the number of neurons in the hidden layer. The optimum number of neurons in the hidden layer is often found by trial-and-error (Bateni et al., 2007). The tangent-sigmoid and linear transfer functions are used in the hidden and output layers, respectively, due to their widespread use. The Levenberg–Marquardt training algorithm trained the ANN model because of its fast convergence and common use (Zanetti et al., 2007).

2.4. ANFIS

Originally, Jang (1993) proposed ANFIS and its combination of the fuzzy logic and ANN approaches. The fuzzy logic system establishes a relationship among the input and output variables. At the same time, ANN determines the characteristics of the membership functions of the fuzzy logic approach. This has overcome the major challenge with designing the fuzzy system (i.e., obtaining its if-then rules). The ANFIS uses a combination of the least square errors and back-propagation gradient descent algorithms for the learning process.

Two methods, namely grid partitioning and subtractive clustering, have been used in ANFIS classify data (Jang, 1993). In the grid partitioning method, input data space is categorized in some local fuzzy areas. While, in subtractive clustering, each data point is used as a potential cluster center. In this study, the subtractive clustering approach is utilized.

2.5. RF

RF is an ensemble-learning algorithm, which was proposed by Breiman (2001). R.F. falls in the category of tree-based approaches in which all trees are dependent on a set of random variables. Regression trees together cause the forest to grow. As its name implies, a collection or a forest of trees is used in this algorithm. In the R.F. algorithm, a large number of decision trees are initially created. Next, all trees are combined to ensure the proper modeling.

The R.F. approach overcomes the over-fitting and low generalizability of the decision tree method. Also, a small variation in learning patterns of the decision tree approach changes its structure (Quinlan, 1986). A specific category of existing patterns is selected to create each tree, considering the replacement of each selected pattern. This kind of sampling usually lists several patterns outside of the category, the so-called out-of-bag patterns. Out-of-bags that are not used in the formation of the corresponding tree can be used to test the generalizability of that tree.

The number of trees (*ntrees*) affects the performance of the R.F. approach. Based upon the user-controlled parameter, a decision tree is grown to the largest possible size.

3. Study sites

The feasibility of MARS, GEP, ANN, ANFIS, and R.F. to estimate W was tested at 50 stations across Iran. The locations of study stations are shown in Fig. 1. The long-term mean monthly wind speed data over 1951–2018 were obtained from the Iran Meteorological Organization (IMO). Table 2 summarizes the geographical information, the mean, standard deviation, and coefficient of variation of wind speed measurements at the study sites. The mean annual wind speed varies between 0.76 m/s at the Kashan station and 5.38 m/s at the Zabol station. The lowest and highest standard deviations are seen at Zabol (2.74 m/s) and Bandar Anzali (0.17 m/s), respectively (Table 2). The coefficient of variation ranges from 0.08 at Bandar Anzali to 0.51 at Zabol. Fig. 2 shows the mean monthly variability of wind speed for all of the 50 study stations. The mean monthly wind speed varies from 1.7 m/s in November/December to 2.9 m/s in July.

4. Results

The long-term mean monthly wind speed over 1951–2018 was estimated at 50 stations in Iran by MARS, GEP, ANN, ANFIS, and R.F. The latitude, longitude, and altitude of each station, as well as the number of months in a year (1 for January and 12 for December), were used as inputs to models. 40 and 10 of the study stations were chosen randomly as the training and testing stations, respectively. Testing stations are Bandar Abbas, Ramsar, Saqez, Shahrekord, Shiraz, Tabriz, Torbat-e Heydarieh, Yasuj, Zahedan, and Zanjan. The performance of all methods was compared with each other using the root mean square error (RMSE), mean absolute error (MAE), Bias, correlation coefficient (CC), and scatter index (S.I.). The expressions for the abovementioned statistical metrics are given by,

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(WS_{o,i} - WS_{e,i}\right)^2}{N}} \tag{3}$$

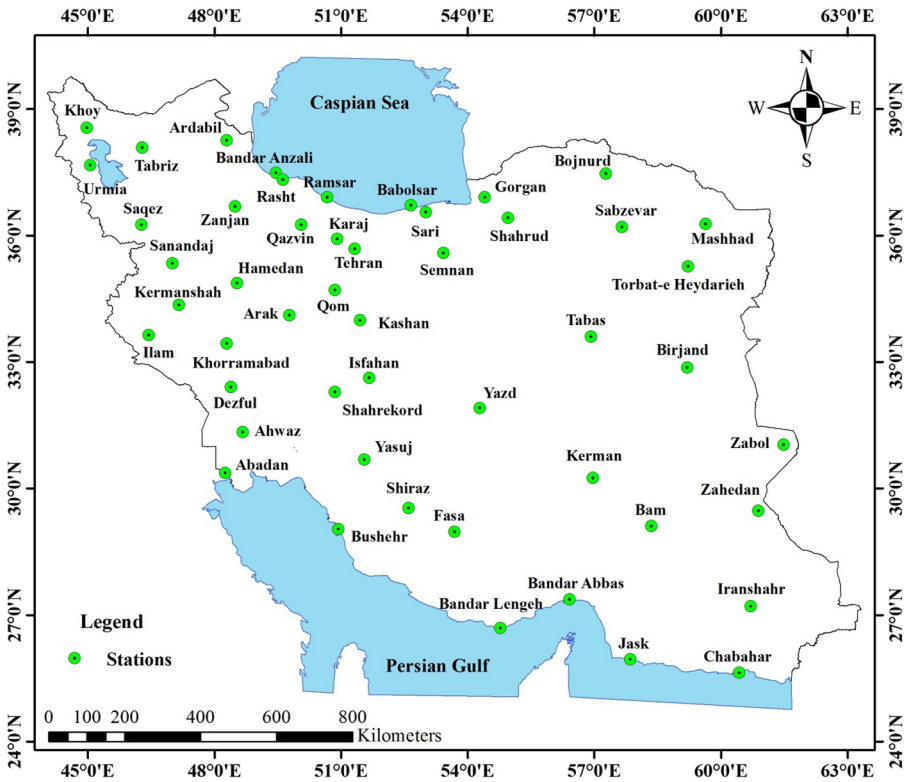

Fig. 1. Geographical locations of the study stations in Iran.

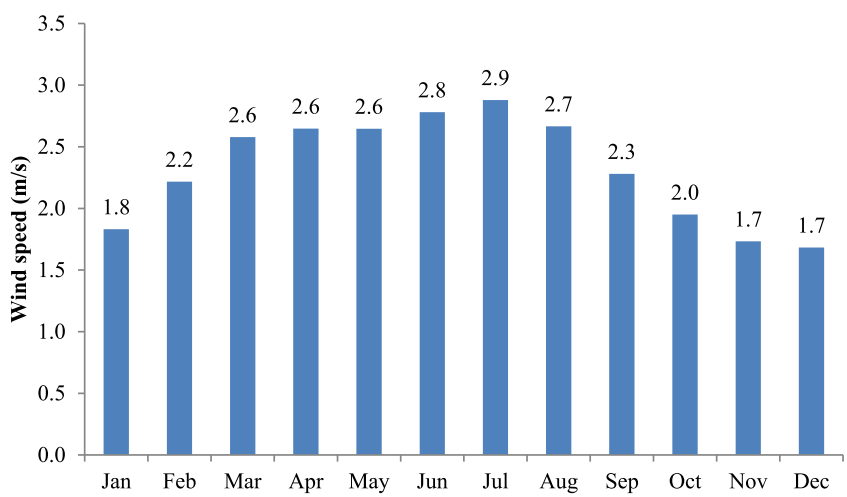

Fig. 2. Long-term mean monthly variations of wind speed at the study stations in Iran.

$$MAE = \frac{\sum_{i=1}^{N} \left| WS_{o,i} - WS_{e,i} \right|}{N} \tag{4}$$

$$Bias = \overline{WS_e} - \overline{WS_o} \tag{5}$$

$$CC = \frac{\sum_{i=1}^{N} \left(WS_{o,i} - \overline{WS_o} \right) \cdot \left(WS_{e,i} - \overline{WS_e} \right)}{\sqrt{\sum_{i=1}^{N} \left(WS_{o,i} - \overline{WS_o} \right)^2 \cdot \sum_{i=1}^{N} \left(WS_{e,i} - \overline{WS_e} \right)^2}} \tag{6}$$

$$SI = \sqrt{\frac{\sum_{i=1}^{N} ((WS_{o,i} - \overline{WS_o}) - (WS_{e,i} - \overline{WS_e}))^2}{\sum_{i=1}^{N} WS_{o,i}^2}} \tag{7}$$

where $WS_{o,i}$ and $WS_{e,i}$ are the ith observed and estimated long-term mean monthly wind speed, respectively. $\overline{WS_o}$ and $\overline{WS_e}$ denote the average of observed and estimated long-term mean monthly wind speed, respectively. N represents the number of long-term mean monthly wind speed observations for the training and testing phases. A lower value of RMSE, MAE, Bias, and S.I. and a higher value of CC illustrate a better results in terms of prediction.

The ANN, ANFIS, and R.F. models must be calibrated to obtain the best wind speed estimates. Hence, the optimal number of neurons in the hidden layer for FFBP ANN, the optimal radii value for the subtractive clustering in ANFIS, and the optimal number of trees for R.F. were found by trial-and-error. The number of neurons in the hidden layer was varied from 1 to 20. The optimal number of neurons was found to be 4 for ANN. We changed the radii value for the subtractive clustering from 0 to 1, and an optimum value of 0.67 is obtained for ANFIS. Ntree was set to 500 to establish the R.F. model. Adding more trees to R.F. did not improve its results.

Table 2
Geographical information of the study stations and the statistical metrics of wind speed data at each station.

Station name	Latitude (°N)	Longitude (°E)	Altitude (m)	Mean (m/s)	Standard deviation (m/s)	Coefficient of variation
Abadan	30.37	48.25	6.6	3.23	0.80	0.25
Ahwaz	31.33	48.67	22.5	2.49	0.59	0.24
Arak	34.10	49.77	1708	1.57	0.45	0.29
Ardabil	38.25	48.28	1332	3.77	0.36	0.10
Babolsar	36.72	52.65	−21	1.49	0.23	0.16
Bam	29.10	58.35	1066.9	3.00	0.48	0.16
Bandar Abbas	27.22	56.37	9.8	3.05	0.40	0.13
Bandar Anzali	37.48	49.45	−23.6	2.23	0.17	0.08
Bandar Lengeh	26.53	54.83	22.7	3.55	0.38	0.11
Birjand	32.87	59.20	1491	2.76	0.75	0.27
Bojnurd	37.47	57.27	1112	2.24	0.72	0.32
Bushehr	28.97	50.82	9	3.09	0.47	0.15
Chabahar	25.28	60.62	8.0	2.90	0.46	0.16
Dezful	32.40	48.38	143	1.74	0.48	0.28
Fasa	28.97	53.68	1288.3	1.68	0.36	0.21
Gorgan	36.90	54.40	0	1.42	0.32	0.23
Hamedan	34.87	48.53	1741.5	1.81	0.49	0.27
Ilam	33.63	46.43	1337	2.30	0.39	0.17
Iranshahr	27.20	60.70	591.1	2.07	0.49	0.24
Isfahan	32.62	51.67	1550.4	2.07	0.65	0.31
Jask	25.63	57.77	5.2	4.05	0.64	0.16
Karaj	35.92	50.90	1312.5	2.42	0.47	0.19
Kashan	33.98	51.45	982.3	0.76	0.36	0.48
Kerman	30.25	56.97	1753.8	3.14	0.72	0.23
Kermanshah	34.35	47.15	1318.6	2.53	0.38	0.15
Khorramabad	33.43	48.28	1147.8	1.79	0.24	0.14
Khoy	38.55	44.97	1103	1.21	0.30	0.24
Mashhad	36.27	59.63	999.2	2.23	0.54	0.24
Qazvin	36.25	50.05	1279.2	1.97	0.41	0.21
Qom	34.70	50.85	877.4	2.26	0.57	0.25
Ramsar	36.90	50.67	−20	1.71	0.18	0.10
Rasht	37.32	49.62	−8.6	1.33	0.18	0.13
Sabzevar	36.20	57.65	972	3.39	0.83	0.24
Sanandaj	35.33	47.00	1373.4	2.02	0.36	0.18
Saqez	36.25	46.27	1522.8	2.05	0.43	0.21
Sari	36.55	53.00	23	1.55	0.22	0.14
Semnan	35.58	53.42	1127	1.52	0.52	0.35
Shahrekord	32.28	50.85	2048.9	1.43	0.39	0.27
Shahrud	36.42	54.95	1349.1	2.27	0.60	0.27
Shiraz	29.53	52.60	1484	2.18	0.45	0.21
Tabas	33.60	56.92	711	1.50	0.52	0.34
Tabriz	38.08	46.28	1361	3.16	0.95	0.30
Tehran	35.68	51.32	1190.8	2.73	0.64	0.24
Torbat-e Heydarieh	35.27	59.22	1450.8	2.22	0.88	0.40
Urmia	37.67	45.05	1328	1.62	0.34	0.21
Yasuj	30.68	51.55	1816.3	1.43	0.28	0.19
Yazd	31.90	54.28	1237.2	2.59	0.46	0.18
Zabol	31.03	61.48	489.2	5.38	2.74	0.51
Zahedan	29.47	60.88	1370	3.32	0.66	0.20
Zanjan	36.68	48.48	1663	1.98	0.27	0.14

The overall statistical indices of utilized models are shown in Table 3. ANFIS (RMSE = 0.67 m/s, MAE = 0.51 m/s, Bias = 0.00 m/s, CC = 0.79, and SI = 0.26) and RF (RMSE = 0.71 m/s, MAE = 0.47 m/s, Bias = −0.02 m/s, CC = 0.83, and SI = 0.27) for the training stations, as well as MARS (RMSE = 0.56 m/s, MAE = 0.42 m/s, Bias = 0.02 m/s, CC = 0.75, and SI = 0.23) and RF (RMSE = 0.56 m/s, MAE = 0.41 m/s, Bias = 0.07 m/s, CC = 0.82, and SI = 0.23) for the testing stations are the best-performing models to estimate *W*.

To assess the accuracy of the utilized soft computing techniques in more details, the statistical metrics of *W* estimates at all the training and testing stations are shown in Tables 4 and 5, respectively. For the training stations, the best *W* estimates were obtained by MARS at Bam (RMSE = 0.13 m/s, MAE = 0.10 m/s, Bias = −0.05 m/s, CC = 0.98, SI = 0.04), GEP at Sanandaj (RMSE = 0.25 m/s, MAE = 0.22 m/s, Bias = 0.22 m/s, CC = 0.95, SI = 0.06), ANN at Sanandaj (RMSE = 0.15 m/s, MAE = 0.11 m/s, Bias = 0.08 m/s, CC = 0.95, SI = 0.06), ANFIS at Sanandaj (RMSE = 0.15 m/s, MAE = 0.13 m/s, Bias = −0.01 m/s, CC = 0.93, SI = 0.07), and RF at Khorramabad (RMSE = 0.21 m/s, MAE = 0.18 m/s, Bias = 0.17 m/s, CC = 0.89, SI = 0.07). Also, all the models had the worst performance at Zabol (Table 4).

Table 3
Statistical metrics of mean monthly wind speed estimates from the utilized soft computing approaches at training and testing stations.

Models	Training stations					Testing stations				
	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI
MARS	0.83	0.59	0.00	0.64	0.32	0.56	0.42	0.02	0.75	0.23
GEP	0.96	0.66	0.01	0.50	0.37	0.76	0.60	0.11	0.45	0.31
ANN	0.84	0.58	−0.07	0.63	0.33	0.68	0.49	0.16	0.64	0.28
ANFIS	0.67	0.51	0.00	0.79	0.26	0.81	0.61	0.16	0.56	0.33
RF	0.71	0.47	−0.02	0.83	0.27	0.56	0.41	0.07	0.82	0.23

For the testing stations (Table 5), MARS at Ramsar (RMSE = 0.25 m/s, MAE = 0.20 m/s, Bias = −0.14 m/s, CC = 0.75, SI = 0.12), GEP at Saqez (RMSE = 0.24 m/s, MAE = 0.21 m/s, Bias = 0.15 m/s, CC = 0.92, SI = 0.09), ANN at Zanjan (RMSE = 0.17 m/s, MAE = 0.14 m/s, Bias = 0.05 m/s, CC = 0.88, SI = 0.08), ANFIS at Ramsar (RMSE = 0.20 m/s, MAE = 0.17 m/s, Bias = 0.17 m/s,

Table 4
Statistical metric of mean monthly wind speed estimates from the utilized soft computing approaches at each training station.

Stations	MARS					GEP					ANN				
	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI
Abadan	0.77	0.61	−0.55	0.75	0.16	0.93	0.71	−0.71	0.70	0.18	1.01	0.88	−0.88	0.79	0.15
Ahwaz	0.37	0.29	−0.01	0.76	0.14	0.40	0.32	−0.04	0.76	0.15	0.41	0.29	−0.23	0.81	0.13
Arak	0.50	0.44	0.38	0.66	0.20	0.75	0.73	0.73	0.96	0.12	0.39	0.37	0.35	0.91	0.12
Ardabil	1.42	1.37	−1.37	0.20	0.10	1.70	1.66	−1.66	0.22	0.10	1.85	1.83	−1.83	0.48	0.08
Babolsar	0.41	0.36	0.33	0.79	0.16	0.69	0.68	0.68	0.96	0.05	0.41	0.38	0.38	0.83	0.10
Bam	0.13	0.10	−0.05	0.98	0.04	0.53	0.40	−0.40	0.67	0.11	0.28	0.25	0.00	0.89	0.09
Bandar Anzali	1.00	0.92	−0.92	−0.72	0.17	0.41	0.35	−0.09	−0.80	0.18	0.41	0.34	−0.33	−0.15	0.11
Bandar Lengeh	0.28	0.21	−0.16	0.88	0.06	0.80	0.77	−0.77	0.81	0.06	0.39	0.29	−0.22	0.89	0.09
Birjand	0.33	0.25	0.00	0.90	0.12	0.71	0.45	−0.39	0.63	0.21	0.33	0.23	−0.13	0.91	0.11
Bojnurd	0.34	0.29	0.15	0.90	0.13	0.54	0.42	−0.10	0.74	0.23	0.50	0.40	−0.31	0.87	0.17
Bushehr	0.49	0.44	−0.08	0.39	0.16	0.59	0.50	−0.49	0.70	0.10	0.65	0.53	−0.48	0.57	0.14
Chabahar	0.65	0.53	0.52	0.66	0.13	0.40	0.32	−0.03	0.47	0.13	0.53	0.36	0.30	0.67	0.15
Dezful	0.55	0.49	0.46	0.75	0.17	0.70	0.65	0.65	0.87	0.15	0.50	0.45	0.45	0.90	0.12
Fasa	0.52	0.46	0.46	0.84	0.14	0.94	0.92	0.92	0.87	0.10	1.00	0.94	0.94	0.94	0.21
Gorgan	0.62	0.58	0.58	0.88	0.15	0.77	0.75	0.75	0.81	0.13	0.54	0.50	0.50	0.76	0.14
Hamedan	0.46	0.41	0.22	0.51	0.22	0.52	0.48	0.45	0.88	0.14	0.30	0.27	0.16	0.84	0.14
Ilam	0.71	0.66	−0.66	0.77	0.10	0.24	0.22	0.03	0.77	0.10	0.24	0.21	−0.12	0.90	0.09
Iranshahr	1.34	1.32	1.32	0.87	0.12	0.73	0.66	0.66	0.76	0.15	1.07	1.06	1.06	0.96	0.09
Isfahan	0.51	0.41	−0.13	0.62	0.23	0.48	0.43	0.31	0.97	0.17	0.34	0.27	−0.11	0.86	0.15
Jask	0.68	0.60	−0.58	0.83	0.08	1.31	1.21	−1.21	0.58	0.12	0.78	0.70	−0.68	0.80	0.09
Karaj	0.37	0.31	−0.23	0.77	0.12	0.32	0.27	−0.21	0.91	0.10	0.63	0.58	−0.58	0.87	0.10
Kashan	1.37	1.35	1.35	0.80	0.27	1.55	1.55	1.55	0.92	0.17	1.21	1.20	1.20	0.88	0.20
Kerman	0.78	0.62	−0.62	0.72	0.15	0.75	0.65	−0.62	0.95	0.14	0.59	0.51	−0.51	0.90	0.09
Kermanshah	0.76	0.73	−0.73	0.77	0.09	0.28	0.25	−0.24	0.95	0.05	0.45	0.42	−0.42	0.94	0.06
Khorramabad	0.20	0.17	0.01	0.77	0.11	0.55	0.54	0.54	0.96	0.04	0.36	0.31	0.28	0.96	0.12
Khoy	0.94	0.91	0.91	0.71	0.19	0.90	0.89	0.89	0.91	0.10	0.97	0.96	0.96	0.91	0.12
Mashhad	0.42	0.36	0.34	0.91	0.11	0.34	0.26	−0.03	0.82	0.15	0.26	0.21	0.20	0.95	0.07
Qazvin	0.30	0.24	0.20	0.83	0.11	0.30	0.24	0.22	0.91	0.10	0.25	0.21	−0.10	0.82	0.12
Qom	0.38	0.32	−0.20	0.83	0.14	0.31	0.28	0.01	0.93	0.13	0.45	0.36	−0.33	0.89	0.13
Rasht	0.35	0.30	0.05	−0.33	0.26	0.88	0.82	0.82	−0.17	0.24	0.59	0.58	0.58	0.59	0.11
Sabzevar	0.99	0.93	−0.93	0.94	0.10	1.36	1.19	−1.19	0.67	0.19	1.32	1.26	−1.26	0.94	0.12
Sanandaj	0.24	0.18	−0.06	0.74	0.11	0.25	0.22	0.22	0.95	0.06	0.15	0.11	0.08	0.95	0.06
Sari	0.46	0.38	0.35	0.66	0.19	0.64	0.63	0.63	0.95	0.06	0.37	0.34	0.34	0.73	0.09
Semnan	0.79	0.74	0.74	0.86	0.16	0.78	0.71	0.71	0.81	0.20	0.41	0.33	0.32	0.92	0.16
Shahrud	0.28	0.23	0.11	0.90	0.11	0.44	0.34	−0.08	0.71	0.19	0.58	0.48	−0.48	0.86	0.14
Tabas	1.09	1.07	1.07	0.92	0.13	0.89	0.83	0.83	0.77	0.21	0.94	0.93	0.93	0.95	0.10
Tehran	0.74	0.59	−0.57	0.65	0.17	0.63	0.51	−0.51	0.96	0.13	0.97	0.88	−0.88	0.89	0.14
Urmia	0.71	0.66	0.66	0.64	0.16	0.53	0.51	0.51	0.92	0.08	0.61	0.57	0.57	0.83	0.12
Yazd	0.42	0.36	−0.29	0.76	0.12	0.28	0.23	−0.17	0.91	0.08	0.43	0.36	−0.34	0.87	0.10
Zabol	3.06	2.24	−2.24	0.93	0.35	3.86	2.91	−2.91	0.35	0.42	3.07	2.18	−2.18	0.78	0.36

Stations	ANFIS					RF				
	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI
Abadan	0.75	0.57	−0.21	0.34	0.22	0.73	0.56	−0.32	0.78	0.20
Ahwaz	0.55	0.48	0.24	0.47	0.19	0.44	0.36	−0.10	0.81	0.17
Arak	0.64	0.60	0.60	0.87	0.14	0.45	0.40	0.34	0.85	0.18
Ardabil	1.23	1.19	−1.19	0.52	0.09	1.10	1.05	−1.05	0.37	0.09
Babolsar	0.34	0.30	0.30	0.77	0.11	0.23	0.19	0.19	0.94	0.09
Bam	0.44	0.38	−0.10	0.93	0.14	0.38	0.33	−0.13	0.68	0.12
Bandar Anzali	0.42	0.34	−0.34	−0.34	0.11	0.43	0.37	−0.37	−0.69	0.10
Bandar Lengeh	0.25	0.18	0.08	0.79	0.07	0.59	0.53	−0.53	0.77	0.07
Birjand	0.47	0.38	0.14	0.83	0.16	0.58	0.40	−0.20	0.79	0.19
Bojnurd	0.61	0.41	−0.38	0.73	0.20	0.55	0.45	0.07	0.84	0.23
Bushehr	0.43	0.37	0.26	0.67	0.11	0.47	0.41	−0.29	0.72	0.12
Chabahar	0.55	0.43	0.35	0.51	0.15	0.44	0.41	0.30	0.76	0.11
Dezful	0.38	0.29	−0.18	0.87	0.19	0.38	0.34	0.18	0.85	0.19
Fasa	0.57	0.50	0.19	0.77	0.31	0.63	0.58	0.58	0.83	0.15
Gorgan	0.28	0.24	−0.01	0.71	0.19	0.36	0.30	0.28	0.87	0.15
Hamedan	0.42	0.39	0.26	0.86	0.17	0.39	0.36	0.14	0.73	0.20
Ilam	0.62	0.58	−0.58	0.82	0.10	0.25	0.21	−0.11	0.89	0.09
Iranshahr	0.53	0.42	0.32	0.78	0.20	0.74	0.65	0.65	0.81	0.16
Isfahan	0.38	0.30	0.04	0.80	0.17	0.47	0.39	−0.10	0.86	0.21
Jask	0.70	0.60	−0.55	0.71	0.10	0.86	0.72	−0.72	0.78	0.12
Karaj	0.21	0.16	0.03	0.94	0.08	0.26	0.21	−0.16	0.93	0.09
Kashan	0.99	0.95	0.95	0.94	0.35	0.81	0.77	0.77	0.94	0.30
Kerman	0.42	0.32	−0.16	0.82	0.12	0.55	0.50	−0.23	0.92	0.16
Kermanshah	0.64	0.62	−0.62	0.93	0.05	0.37	0.32	−0.32	0.90	0.07
Khorramabad	0.31	0.28	0.03	0.87	0.17	0.21	0.18	0.17	0.89	0.07
Khoy	0.85	0.81	0.81	0.56	0.19	0.74	0.71	0.71	0.91	0.19
Mashhad	0.98	0.97	0.97	0.93	0.09	0.42	0.34	0.26	0.87	0.15

(*continued on next page*)

Table 4 (*continued*).

Stations	ANFIS					RF				
	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI
Qazvin	0.56	0.50	0.50	0.91	0.12	0.28	0.21	0.18	0.90	0.11
Qom	0.80	0.77	−0.77	0.90	0.10	0.41	0.34	−0.23	0.93	0.15
Rasht	0.41	0.37	0.37	0.34	0.12	0.48	0.45	0.45	0.14	0.13
Sabzevar	1.02	0.96	−0.96	0.92	0.10	0.87	0.71	−0.63	0.74	0.17
Sanandaj	0.15	0.13	−0.01	0.93	0.07	0.23	0.20	0.13	0.88	0.09
Sari	0.36	0.33	−0.33	0.83	0.11	0.23	0.19	0.19	0.90	0.07
Semnan	0.63	0.60	0.60	0.96	0.12	0.50	0.38	0.32	0.81	0.24
Shahrud	0.40	0.31	−0.15	0.78	0.16	0.41	0.34	0.02	0.86	0.18
Tabas	1.10	0.94	0.94	0.80	0.36	0.71	0.61	0.61	0.83	0.23
Tehran	0.46	0.40	−0.40	0.93	0.08	0.58	0.45	−0.42	0.82	0.15
Urmia	0.44	0.32	0.25	−0.23	0.22	0.53	0.49	0.49	0.87	0.13
Yazd	0.68	0.59	−0.59	0.82	0.13	0.34	0.26	−0.24	0.95	0.09
Zabol	1.76	1.22	−0.73	0.90	0.27	2.95	2.20	−1.51	0.49	0.42

Table 5
Statistical metric of mean monthly wind speed estimates from the utilized soft computing approaches at each testing station.

Stations	MARS					GEP					ANN				
	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI
Bandar Abbas	0.31	0.29	0.26	0.95	0.05	0.42	0.33	−0.32	0.71	0.09	0.50	0.44	0.38	0.89	0.11
Ramsar	0.25	0.20	−0.14	0.75	0.12	0.48	0.46	0.46	0.92	0.07	0.22	0.19	0.19	0.74	0.07
Saqez	0.35	0.29	0.13	0.64	0.15	0.24	0.21	0.15	0.92	0.09	0.26	0.21	0.18	0.89	0.09
Sharekord	0.45	0.39	0.18	0.41	0.28	0.99	0.97	0.97	0.87	0.13	0.75	0.65	0.59	0.67	0.32
Shiraz	0.46	0.38	−0.34	0.76	0.14	0.43	0.39	0.39	0.98	0.08	0.55	0.49	0.43	0.93	0.16
Tabriz	1.12	0.83	−0.76	0.78	0.25	1.30	1.04	−1.04	0.61	0.24	1.31	0.99	−0.98	0.31	0.26
Torbat-e Heydarieh	0.54	0.47	0.37	0.91	0.17	0.70	0.59	0.02	0.63	0.30	0.49	0.46	0.28	0.90	0.17
Yasuj	0.29	0.25	0.11	0.76	0.19	1.07	1.06	1.06	0.96	0.05	0.96	0.85	0.82	0.87	0.34
Zahedan	0.81	0.72	−0.03	0.08	0.24	0.90	0.76	−0.75	0.67	0.15	0.73	0.51	−0.35	0.43	0.19
Zanjan	0.46	0.42	0.42	0.69	0.10	0.24	0.21	0.20	0.85	0.07	0.17	0.14	0.05	0.88	0.08

Stations	ANFIS					RF				
	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI	RMSE (m/s)	MAE (m/s)	Bias (m/s)	CC	SI
Bandar Abbas	0.69	0.62	0.61	0.77	0.10	0.29	0.24	0.01	0.75	0.09
Ramsar	0.20	0.17	0.17	0.78	0.06	0.13	0.11	0.09	0.89	0.05
Saqez	0.41	0.37	−0.17	0.49	0.18	0.30	0.25	0.06	0.79	0.14
Sharekord	0.91	0.83	0.83	0.60	0.25	0.56	0.49	0.48	0.68	0.19
Shiraz	0.54	0.41	−0.24	0.81	0.22	0.35	0.30	0.13	0.90	0.15
Tabriz	1.37	1.02	−1.00	−0.11	0.29	1.01	0.74	−0.65	0.81	0.23
Torbat-e Heydarieh	0.76	0.58	−0.18	0.56	0.31	0.74	0.66	0.31	0.76	0.28
Yasuj	0.90	0.76	0.76	0.74	0.34	0.65	0.63	0.63	0.92	0.13
Zahedan	1.16	1.07	0.58	0.11	0.30	0.68	0.56	−0.46	0.76	0.15
Zanjan	0.32	0.29	0.29	0.90	0.07	0.15	0.13	0.08	0.89	0.06

CC = 0.78, SI = 0.06), and RF at Ramsar (RMSE = 0.13 m/s, MAE = 0.11 m/s, Bias = 0.09 m/s, CC = 0.89, SI = 0.05) had the best results. Also, all the models indicated the worst performance at Tabriz.

The utilized models were ranked for the training and testing stations based on their RMSE values in Tables 6 and 7. MARS, GEP, ANN, ANFIS, and R.F. performed the best at 8, 4, 10, 10, and 8 stations for the training stations, respectively. The R.F. and ANFIS techniques are the best models at the training stations according to their lowest total ranks. Additionally, 6, 2, 1, and 1 stations from the test stations had the first rank in the R.F., MARS, GEP, and ANN. The lowest total rank value from R.F. indicates its outperformance compared to the other models at the testing stations.

Fig. 3 shows the histogram of observed and estimated W values at the testing stations. As can be seen, all the models underestimate W for some months and overestimate it for other months. For example, the highest W underestimations are seen at Tabriz in June, July, and August.

The performance of soft computing models depends mainly on their inputs. In addition, air temperature, air pressure, relative humidity and other weather data could be used as inputs. However, these climatic variables may be unavailable in some regions. Hence, it is advantageous to utilize easily available predictors such as geographic information (latitude, longitude, and altitude) and the number of months.

Other studies also used geographical information and the number of months in soft computing techniques to estimate different variables of interest. For example, Sanikhani et al. (2018) and Mehdizadeh (2018) showed that soft computing methods could estimate long-term mean monthly air temperature using geographical properties and periodicity terms. Kisi and Sanikhani (2015) employed these predictors to predict the long-term mean monthly precipitation in machine learning methods. In addition, Kisi et al. (2015) estimated long-term mean monthly reference evapotranspiration. The abovementioned studies proved that geographic information and periodicity components could estimate long-term mean monthly hydro-climatological variables.

According to the rank of the models in Tables 6 and 7, R.F. is generally the best-performing approach for the training and test stations. Unlike many soft computing techniques, R.F. does not need to be tuned (Lahouar and Ben Hadj Slama, 2017). For example, a time-consuming trial and error procedure should obtain the optimal number of neurons of the hidden layer in ANN and the optimal radii value for the ANFIS. In contrast, the R.F. development requires less time (Jia et al., 2021).

Other studies (e.g., Lahouar and Ben Hadj Slama (2017), Jia et al. (2021), Wang et al. (2021), and Sun et al. (2021)) also illustrated the robustness of R.F. in estimating wind characteristics

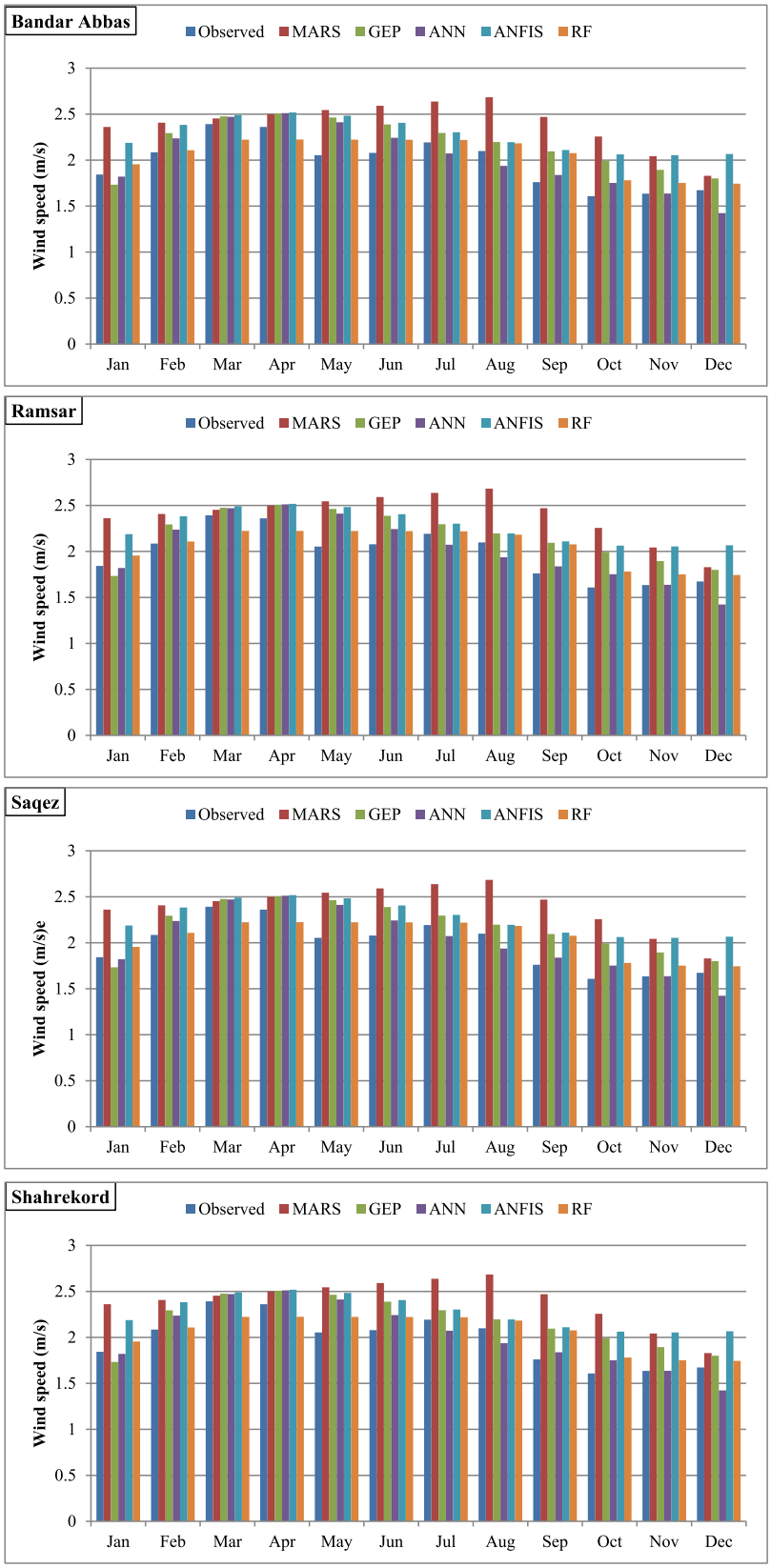

Fig. 3. Histogram of observed and estimated long-term monthly mean wind speed values from the utilized soft computing approaches at the testing stations.

such as wind power. In addition, Feng et al. (2017), Jiajun et al. (2020) and Neupane et al. (2021) reported the capability of R.F. for estimating the wind speed time series.

5. Conclusions

This study investigated the performance of five soft computing techniques, namely multivariate adaptive regression splines

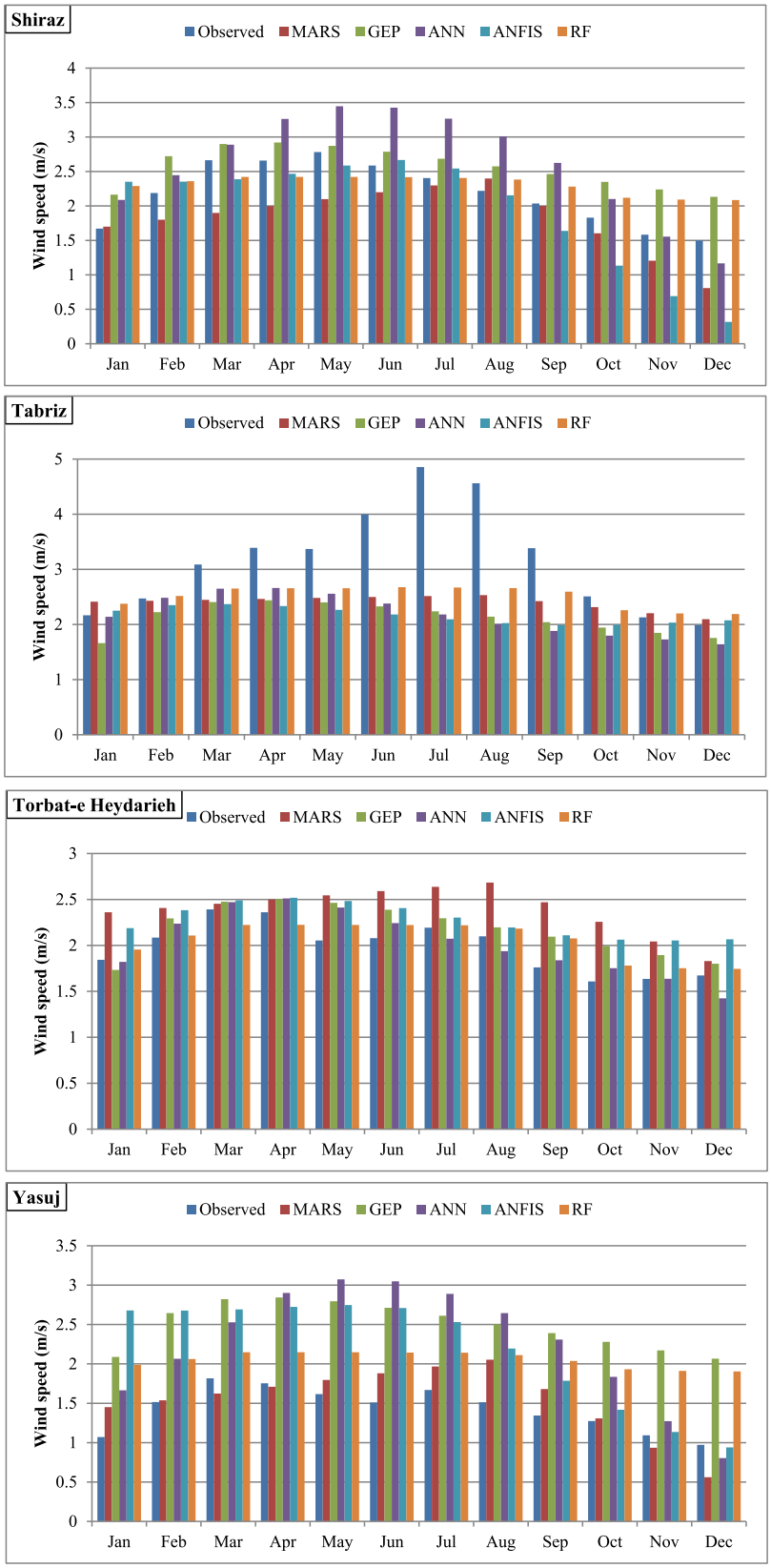

Fig. 3. (*continued*).

(MARS), gene expression programming (GEP), artificial neural networks (ANN), adaptive neuro-fuzzy inference system (ANFIS), and random forest (R.F.) for the estimation of long-term mean monthly wind speed. Fifty stations in Iran were chosen as the study sites. The long-term mean monthly wind speed data were used to train and test the models at 40 (i.e., 40 stations × 12 months = 480 data points) and 10 (i.e., 10 stations × 12 months = 120 data points) stations, respectively. Latitude, longitude and

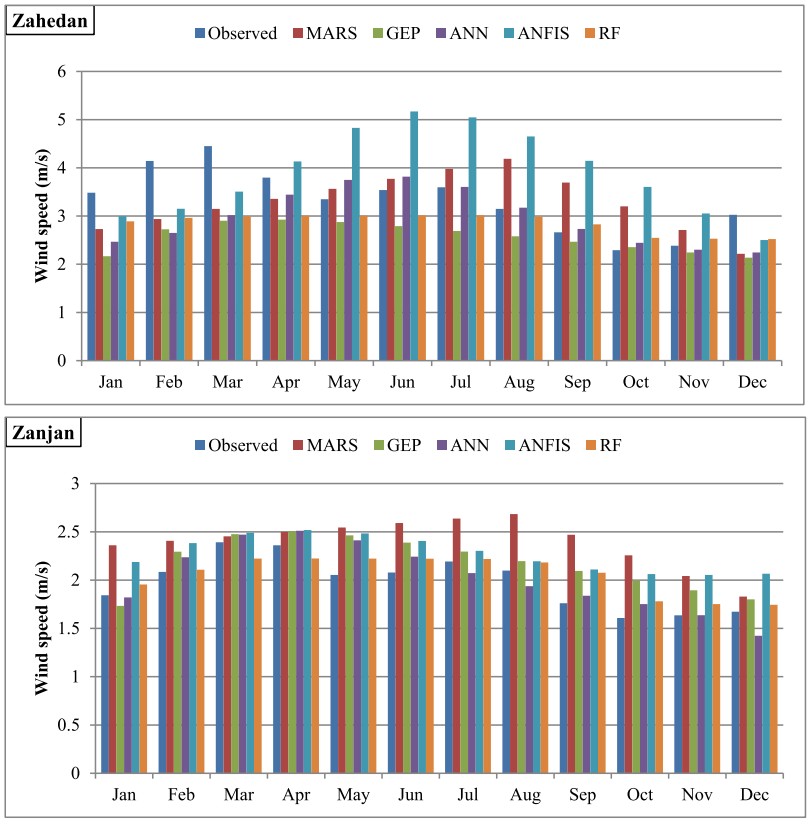

Fig. 3. (*continued*).

Table 6
The rank of utilized models at the training stations based on their accuracy.

Stations	MARS	GEP	ANN	ANFIS	RF
Abadan	3	4	5	2	1
Ahwaz	1	2	3	5	4
Arak	3	5	1	4	2
Ardabil	3	4	5	2	1
Babolsar	3	5	4	2	1
Bam	1	5	2	4	3
Bandar Anzali	5	2	1	3	4
Bandar Lengeh	2	5	3	1	4
Birjand	2	5	1	3	4
Bojnurd	1	3	2	5	4
Bushehr	3	4	5	1	2
Chabahar	5	1	3	4	2
Dezful	4	5	3	1	2
Fasa	1	4	5	2	3
Gorgan	4	5	3	1	2
Hamedan	4	5	1	3	2
Ilam	5	2	1	4	3
Iranshahr	5	2	4	1	3
Isfahan	5	4	1	2	3
Jask	1	5	3	2	4
Karaj	4	3	5	1	2
Kashan	4	5	3	2	1
Kerman	5	4	3	1	2
Kermanshah	5	1	3	4	2
Khorramabad	1	5	4	3	2
Khoy	4	3	5	2	1
Mashhad	4	2	1	5	3
Qazvin	4	5	1	5	2
Qom	2	1	4	5	3

(*continued on next page*)

Table 6 (*continued*).

Stations	MARS	GEP	ANN	ANFIS	RF
Rasht	1	5	4	2	3
Sabzevar	2	5	4	3	1
Sanandaj	4	5	1	2	3
Sari	4	5	3	2	1
Semnan	5	4	1	3	2
Shahrud	1	4	5	2	3
Tabas	4	2	3	5	1
Tehran	4	3	5	1	2
Urmia	5	3	4	1	2
Yazd	3	1	4	5	2
Zabol	3	5	4	1	2
Total ranks	130	148	123	107	94

Table 7
The rank of utilized models at the testing stations based on their accuracy.

Stations	MARS	GEP	ANN	ANFIS	RF
Bandar Abbas	2	3	4	5	1
Ramsar	4	5	3	2	1
Saqez	4	1	2	5	3
Sharekord	1	5	3	4	2
Shiraz	3	2	5	4	1
Tabriz	2	3	4	5	1
Torbat-e Heydarieh	2	3	1	5	4
Yasuj	1	5	4	3	2
Zahedan	3	4	2	5	1
Zanjan	5	3	2	4	1
Total ranks	27	34	30	42	17

altitude (i.e., the geographical attributes) and the number of months (i.e., the periodicity term) were employed as inputs to the models. The aforementioned soft computing techniques successfully estimated long-term mean monthly wind speed using the geographical information and periodicity term. It was found that the R.F. method generally yielded superior results than the other four methods at the training and testing stations. ANFIS and MARS were ranked the second and third best models at the training

and test stations. GEP and ANFIS provided the lowest accuracy at the training and testing stations, respectively. The most accurate estimates of long-term mean monthly wind speed were obtained by MARS at Bam (RMSE = 0.13 m/s, MAE = 0.10 m/s), GEP at Saqez and Zanjan (RMSE = 0.24 m/s, MAE = 0.21 m/s), ANN at Sanandaj (RMSE = 0.15 m/s, MAE = 0.11 m/s), ANFIS at Sanandaj (RMSE = 0.15 m/s, MAE = 0.13 m/s), and RF at Ramsar (RMSE = 0.13 m/s, MAE = 0.11 m/s).

The present study examined the potential of the abovementioned soft computing techniques for estimating the long-term mean monthly wind speed. Further studies should be directed towards evaluating the performance of other soft computing techniques. Furthermore, the feasibility of soft computing techniques can also be investigated to estimate long-term mean monthly solar radiation as another source of clean energy utilizing geographical information and periodicity term.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Bateni, S.M., Borghei, S.M., Jeng, D.S., 2007. Neural network and neuro-fuzzy assessments for scour depth around bridge piers. Eng. Appl. Artif. Intell. 20 (3), 401–414.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Cadenas, E., Rivera, W., 2009. Short term wind speed forecasting in La Venta, Oaxaca, Mexico, using artificial neural networks. Renew. Energy 34 (1), 274–278.

Celik, A.N., Kolhe, M., 2013. Generalized feed-forward based method for wind energy prediction. Appl. Energy 101, 582–588.

Chang, G.W., Lu, H.J., Chang, Y.R., Lee, Y.D., 2017. An improved neural network-based approach for short-term wind speed and power forecast. Renew. Energy 105, 301–311.

Chen, K., Yu, J., 2014. Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach. Appl. Energy 113, 690–705.

Erdem, E., Shi, J., 2011. ARMA Based approaches for forecasting the tuple of wind speed and direction. Appl. Energy 88 (4), 1405–1414.

Fadare, D.A., 2010. The application of artificial neural networks to mapping of wind speed profile for energy application in Nigeria. Appl. Energy 87 (3), 934–942.

Feng, C., Cui, M., Hodge, B.M., Zhang, J., 2017. A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. Apll. Energy 190, 1245–1257.

Ferreira, C., 2001. Gene expression programming: a new adaptive algorithm for solving problems. Complex Syst. 13 (2), 87–129.

Friedman, J.H., 1991. Multivariate adaptive regression splines. Ann. Stat. 19 (1), 1–67.

Guo, Z., Zhao, W., Lu, H., Wang, J., 2012. Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model. Renew. Energy 37 (1), 241–249.

Han, Q., Meng, F., Hu, T., Chu, F., 2017. Non-parametric hybrid models for wind speed forecasting. Energy Convers. Manage. 148, 554–568.

Hur, S-h., 2021. Short-term wind speed prediction using extended Kalman filter and machine learning. Energy Rep. 7, 1046–1054.

Jang, J.S.R., 1993. ANFIS: Adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man. Cybern. 23, 665–685.

Jia, X., Han, Y., Li, Y., Sang, W., Zhang, G., 2021. Condition monitoring and performance forecasting of wind turbines based on denoising autoencoder and novel convolutional neural networks. Energy Rep. 7, 6354–6365.

Jiajun, H., Chuanjin, Y., Yongle, L., Huoyue, X., 2020. Ultra-short term wind prediction with wavelet transform, deep belief network and ensemble learning. Energy Convers. Manage. 205, 112418.

Jiang, Y., Huang, G., 2017. Short-term wind speed prediction: hybrid of ensemble empirical mode decomposition, feature selection and error correction. Energy Convers. Manage. 144, 340–350.

Kavasseri, R.G., Seetharaman, K., 2009. Day-ahead wind speed forecasting using f-ARIMA models. Renew. Energy 34 (5), 1388–1393.

Kisi, O., Sanikhani, H., 2015. Prediction of long-term monthly precipitation using several soft computing methods without climatic data. Int. J. Climatol. 35 (14), 4139–4150.

Kisi, O., Sanikhani, H., Zounemat-Kermani, M., Niazi, F., 2015. Long-term monthly evapotranspiration modeling by several data-driven methods without climatic data. Comput. Electron. Agric. 115, 66–77.

Lahouar, A., Ben Hadj Slama, J., 2017. Hour-ahead wind power forecast based on random forests. Renew. Energy 109, 529–541.

Li, D., Jiang, F., Chen, M., Qian, T., 2022. Multi-step-ahead wind speed forecasting based on a hybrid decomposition method and temporal convolutional networks. Energy 238 (C), 121981.

Liu, D., Niu, D., Wang, H., Fan, L., 2014. Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. Renew. Energy 62, 592–597.

Liu, H., Tian, H.Q., Li, Y.F., 2015. Comparison of new hybrid FEEMD-MLP, FEEMDANFIS, Wavelet Packet-MLP and Wavelet Packet-ANFIS for wind speed predictions. Energy Convers. Manage. 89, 1–11.

López, G., Arboleya, P., 2021. Short-term wind speed forecasting over complex terrain using linear regression models and multivariable LSTM and NARX networks in the Andes Range. Ecuador. Renew. Energy http://dx.doi.org/10.1016/j.renene.2021.10.070.

Maroufpoor, S., Sanikhani, H., Kisi, O., Deo, R.C., Yaseen, Z.M., 2019. Long-term modelling of wind speeds using six different heuristic artificial intelligence approaches. Int. J. Climatol. 39 (8), 3543–3557.

Mehdizadeh, S., 2018. Assessing the potential of data-driven models for estimation of long-term monthly temperatures. Comput. Electron. Agric. 144, 114–125.

Monfared, M., Rastegar, H., Kojabadi, H.M., 2009. A new strategy for wind speed forecasting using artificial intelligent methods. Renew. Energy 34 (3), 845–848.

Neupane, A., Raj, N., Deo, R.C., Ali, M., 2021. Chapter 6 - development of data-driven models for wind speed forecasting in Australia. Predic. Model. Energy Manage. Power Syst. Eng. 143–190.

Noorollahi, Y., Jokar, M.A., Kalhor, A., 2016. Using artificial neural networks for temporal and spatial wind speed forecasting in Iran. Energy Convers. Manage. 115, 17–25.

Qu, Z., Mao, W., Zhang, K., Zhang, W., Li, Z., 2019. Multi-step wind speed forecasting based on a hybrid decomposition technique and an improved back-propagation neural network. Renew. Energy 133, 919–929.

Quinlan, J.R., 1986. Induction of decision trees. J. Mach. Learn. 1 (1), 81–106.

Samadianfard, S., Hashemi, S., Kargar, K., Izadyar, M., Mostafaeipour, A., Mosavi, A., Nabipour, N., Shamshirband, S., 2020. Wind speed prediction using a hybrid model of the multi-layer perceptron and whale optimization algorithm. Energy Rep. 6, 1147–1159.

Sanikhani, H., Deo, R.C., Samui, P., Kisi, O., Mert, C., Mirabbasi, R., Gavili, S., Yaseen, Z.M., 2018. Survey of different data-intelligent modeling strategies for forecasting air temperature using geographic information as model predictors. Comput. Electron. Agric. 152, 242–260.

Su, Z., Wang, J., Lu, H., Zhao, G., 2014. A new hybrid model optimized by an intelligent optimization algorithm for wind speed forecasting. Energy Convers. Manage. 85, 443–452.

Sun, Z., Zhao, M., Dong, Y., Cao, X., Sun, H., 2021. Hybrid model with secondary decomposition, randomforest algorithm, clustering analysis and long short memory network principal computing for short-term wind power forecasting on multiple scales. Energy 221, 119848.

Sun, N., Zhou, J., Liu, G., He, Z., 2019. A hybrid wind speed forecasting model based on a decomposition method and an improved regularized extreme learning machine. Energy Proc. 158, 217–222.

Velo, R., Lopez, P., Maseda, F., 2014. Wind speed estimation using multi-layer perceptron. Energy Convers. Manage. 81, 1–9.

Wang, J., Niu, T., Lu, H., Guo, Z., Yang, W., Du, P., 2018. An analysis-forecast system for uncertainty modeling of wind speed: a case study of large-scale wind farms. Appl. Energy 211, 492–512.

Wang, Y., Wang, J., Wei, X., 2015. A hybrid wind speed forecasting model based on phase space reconstruction theory and Markov model: a case study of wind farms in northwest China. Energy 91, 556–572.

Wang, A., Xu, L., Xing, J., Chen, X., Liu, K., Liang, Y., Zhou, Z., 2021. Random-forest based adjusting method for wind forecast of WRF model. Comput. Geosci. 155, 104842.

Wang, S., Zhang, N., Wu, L., Wang, Y., 2016. Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method. Renew. Energy 94, 629–636.

Zanetti, S.S., Sousa, E.F., Oliveira, V.P., Almeida, F.T., Bernardo, S., 2007. Estimating evapotranspiration using artificial neural network and minimum climatological data. J. Irrig. Drain. Eng. 133 (2), 83–89.

Zhang, Y., Chen, B., Pan, G., Zhao, Y., 2019. A novel hybrid model based on VMD-WT and PCA-BP-RBF neural network for short-term wind speed forecasting. Energy Convers. Manage. 195, 180–197.

Zheng, C.W., Li, X.Y., Luo, X., Chen, X., Qian, Y.H., Zhang, Z.H., Gao, Z.S., Du, Z.B., Gao, Y.B., Chen, Y.G., 2019. Projection of future global offshore wind energy resources using CMIP data. Atmos. Ocean 57 (2), 134–148.

Zheng, C.W., Pan, J., Li, J.X., 2013. Assessing the China sea wind energy and wave energy resources from 1988 to 2009. Ocean Eng. 65, 39–48.

Zheng, C.W., Pan, J., Li, C.Y., 2016. Global oceanic wind speed trends. Ocean Coast. Manage. 129, 15–24.