Short-Term Wind Speed Forecasting Based on Fuzzy C-Means Clustering and Improved MEA-BP

Gonggui Chen, *Member, IAENG*, Jing Chen, Zhizhong Zhang, and Zhi Sun

***Abstract*— Overuse of non-renewable energy has seriously affected the natural environment. Wind energy is a kind of clean energy with great potential. The higher accuracy of wind speed forecasting, the higher utilization efficiency of wind energy will be. BP neural network can solve nonlinear problems, but it has different generalization ability to different data. Therefore, this paper proposed a forecasting model based on fuzzy c-means clustering (FCM) and improved mind evolutionary algorithm-BP (IMEA-BP). Firstly, the input data set of BP is divided into several classes by FCM, and the number of class is obtained by multiple experiments. After classification, the coefficient of variation in each input vector is used for outlier detection, and outliers are removed from the input data set. Then different IMEA-BP models are built for each class of input data set. Finally, the class of forecasting input is determined and the corresponding IMEA-BP is used for forecasting. The experimental results of two cases showed that the proposed model is not only suitable for one-step forecasting, but also improves the accuracy of multi-step forecasting.**

***Index Terms*—Wind Speed Forecasting, Fuzzy C-means Clustering (FCM), Outlier Detection (OD), Improved Mind Evolutionary Algorithm-BP (IMEA-BP).**

I. Introduction

UNCERTAINTY of wind power output affects the stability of power system. Wind energy is a clean and renewable energy, but its utilization rate is low due to the intermittent and volatility of wind energy [1-4].

Wind speed forecasting models can be divided into physical models and statistical models [5-7]. The physical models establishes equations through information such as air pressure, temperature, altitude, and then solves the equations to forecast wind speed. Numerical weather forecasting is a typical physical forecasting model. Statistical models are divided into linear and nonlinear forecasting models. Linear forecasting models include auto regressive model (AR), moving average model (MA) and auto regressive integrated moving average model (ARIMA) [8-10]. The nonlinear forecasting models mainly includes the artificial neural network (ANN) models and the support vector machine (SVM) models [11-14].

Wang et al. [15] used the ensemble empirical mode decomposition to divide the original data into signals of different frequencies, and used the decomposed signals as the input of GA-BP. Sun et al. [16] preprocessed the wind speed data by fast ensemble empirical mode decomposition (FEEMD) and sample entropy, and then used the improved BP neural network to forecast the wind speed. Song et al. [17] proposed a forecasting model based on improved complete ensemble empirical mode decomposition adaptive noise (ICEEMDAN) and gray wolf algorithm. Yu et al. [18] used the wavelet packet to decompose the original data, and then applied the gradient boosted regression trees (GBRT) to determine the Elman structure. Finally, the density-based spatial clustering of applications with noise (DBSCAN) method was used to select the data training forecasting for Elman. They also proposed a forecasting model of wavelet decomposition based on singular spectrum analysis [19]. Liu et al. [20] used the empirical wavelet transform (EWT) to divide the original wind speed data into several sub-sequences, the long short term memory neural network used to forecast the low-frequency part, and the Elman neural network used to forecast the high-frequency part. Ren et al. [21] proposed a PSO-BP forecasting model with parameter selection. Sun et al. [22] used phase space reconstruction (PSR) to select the input vector for the core vector regression (CVR) models, and then used kernel principal component analysis (KPCA) to extract the nonlinear features of PSR, and finally used the CVR models on forecasting. Shao et al. [23] combined infinite feature selection and recurrent neural network (RNN) to solve the problem of short-term wind power forecasting. Reference [24] used wavelet transform (WT) and the sparsity of the correlation between meteorological stations to establish forecasting model, which greatly improved the forecasting accuracy. Reference [25] used WT to filter wind speed data, and then used radial basis function neural network (RBF) to make preliminary forecasting. Then Levenberg–Marquardt (LM), Broyden-Fletcher-Goldfarb-Shanno (BFGS), and Bayesian Regularization (BR) were combined into three multi-layer perceptron (MLP). Meta-heuristic technique imperialist competitive algorithm (ICA) was used to optimize neural networks.

Manuscript received June 20, 2019; revised September 20, 2019. This work was supported by the National Natural Science Foundation of China (Nos. 51207064 and 61463014).

Gonggui Chen is with the Key Laboratory of Industrial Internet of Things & Networked Control, Ministry of Education, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; Key Laboratory of Complex Systems and Bionic Control, Chongqing University of Posts and Telecommunications, Chongqing 400065, China(e-mail:chenggpower@126.com).

Jing Chen is with the Key Laboratory of Industrial Internet of Things & Networked Control, Ministry of Education, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: chenjing_0627@163.com).

Zhizhong Zhang is with the Key Laboratory of Communication Network and Testing Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (corresponding author, Tel: 023-62461681; e-mail: zhangzztx@163.com).

Zhi Sun is with the Chn Energy Enshi Hydropower Co., Ltd, Enshi 445000, China (e-mail: sunzhi24@126.com).

Reference [26] had established PCA-MEA-MLP hybrid model by combining MEA, PCA and MLP, and then applied the hybrid model to strip shape forecasting. Reference [27] divided wind speed sequence into several sub-layers by FEEMD, and then established MEA-MLP, GA-MLP, FEEMD-MLP models. The results showed that MEA/GA algorithm can improve the performance of MLP. Reference [28] used MEA to find the optimal initial weights and wavelet parameters of wavelet neural network, and then used the improved wavelet neural network to forecast short-term traffic flow.

A BP neural network has different forecasting accuracy for different input data. If a BP neural network has a good forecasting result of input data with low volatility, then it has a worse forecasting result of input data with high volatility. Therefore, we used FCM to divide the input data of BP neural network with similar characteristics into one class, and establish different BP neural networks for each class.

The wind speed data of the two cases are the measured wind speed of a wind farm in China. The sampling interval of wind speed is 15 minutes.

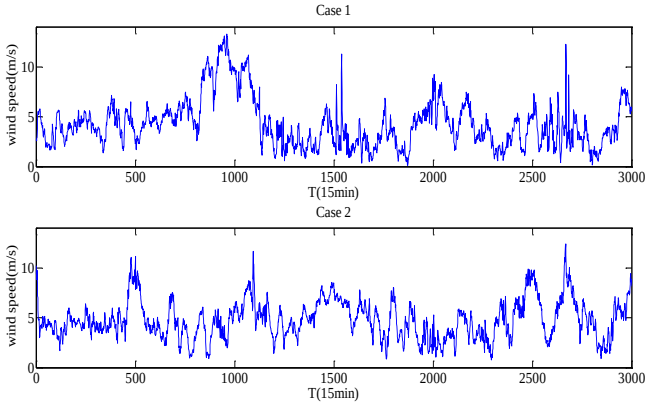

Fig. 1.Wind speed data

II. Methodology

A. Fuzzy C-Means Clustering

The Fuzzy C-Means Clustering (FCM) algorithm proposed by Bezdek used degree of membership to judge the class of data [29-31]. The data matrix X consists of p variables of n samples.

$$X=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \tag{1}$$

FCM divides n samples into c classes, and $V=(v_1,v_2,\cdots,v_c)$ is the cluster center, where $v_i=(v_{i1}, v_{i2}, \cdots, v_{ip})$, $i=(1,2,\cdots c)$.

The objective function:

$$J(U,V)=\sum_{k=1}^{n}\sum_{i=1}^{c}u_{ik}^{m}d_{ik}^{2} \tag{2}$$

where u_{ik} is degree of membership, $U=(u_{ik})_{c\times n}$ is the membership matrix, and d_{ik} is the distance between data point and cluster center $d_{ik}=\|x_k-v_i\|$.

The clustering criterion of FCM is to find U and V to minimize $J(U,V)$. The specific steps for FCM are as follows:

(i) Determine the number of class c and the initial membership matrix $U^{(0)}=(u_{ik}^{(0)})$. l is the number of iterations.

(ii) Calculate the cluster center $V^{(l)}$ of the lth iteration by the following formula:

$$v_i^{(l)}=\frac{\sum_{k=1}^{n}(u_{ik}^{(l-1)m}x_k)}{\sum_{k=1}^{n}(u_{ik}^{l-1})^m},i=1,2,\cdots c \tag{3}$$

(iii) Correct the membership matrix $U^{(l)}$ and calculate the objective function value $J^{(l)}$.

$$u_{ik}^{(l)}=\frac{1}{\sum_{j=1}^{n}\left(\frac{d_{ik}^{(l)}}{d_{jk}^{(l)}}\right)^{\frac{2}{m-1}}},i=1,2,\cdots c;k=1,2\cdots n \tag{4}$$

$$J^{(l)}(U^{(l)},V^{(l)})=\sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik}^{(l)})^m(d_{ik}^{(l)})^2 \tag{5}$$

where $d_{ik}^{(l)}=\left\|x_k-v_i^{(l)}\right\|$.

(iv) When the condition $\max\left|u_{ik}^{(l)}-u_{ik}^{(l-1)}\right|<\varepsilon_u$ is met, the iteration will end. Otherwise, $l=l+1$, and step (ii) is performed. Where ε_n is a constant greater than zero.

Through the above steps, the final membership matrix U and the cluster center V can be obtained to minimize J. The sample class can be determined according to the value of U.

The number of classes c has a great influence on the forecasting result. For each case, we need to figure out the best number of classes. For the two cases in this paper, we set the c as 1, 2, 3, 4, 5, 6 respectively, and then conducted 20 experiments using BP neural network. The average error criteria of the experiments are shown in TABLE I.

For case 1 and case 2, when c is 4 and 6 respectively, the forecasting result is the best. The FCM iteration process for both cases is shown in Fig.2.

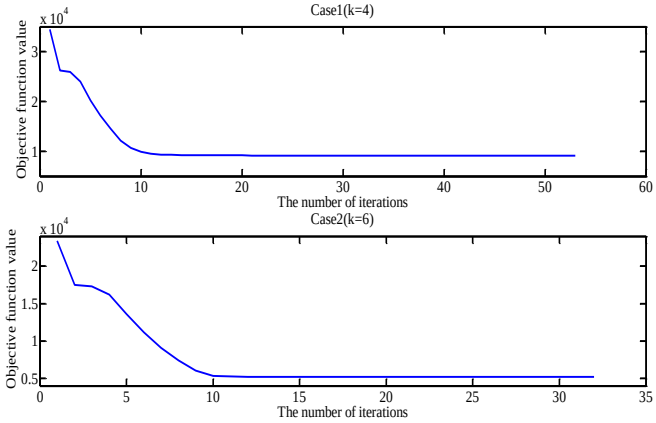

Fig. 2.FCM iteration process

TABLE I
Error Criteria Of Different Classes

	Case 1						Case 2					
	1	2	3	4	5	6	1	2	3	4	5	6
MAE	0.264	0.271	0.311	0.261	0.279	0.275	0.313	0.305	0.319	0.320	0.322	0.306
RMSE	0.329	0.337	0.389	0.325	0.351	0.343	0.421	0.411	0.414	0.425	0.424	0.405
MAPE(%)	5.236	5.357	5.881	5.135	5.560	5.416	5.902	5.729	5.872	5.960	6.029	5.582

B. Outlier Detection

Outlier detection (OD) is to find the maximum or minimum value in the data set. This paper used the outlier detection method based on statistics. The coefficient of variation can reflect the dispersion degree of data, so the variation coefficient of the input data is used for outlier detection. The outlier detection steps are as follows:

(i) Calculate the coefficient of variation *CV* of the forecasting matrix *X*.

$$X_{n\times p}=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \tag{6}$$

$$CV_{1\times n}=[\frac{\sigma_{X_1}}{\mu_{X_1}},\frac{\sigma_{X_2}}{\mu_{X_2}},\cdots,\frac{\sigma_{X_n}}{\mu_{X_n}}] \tag{7}$$

(ii) Calculate the mean μ_{CV} and standard deviation σ_{CV} of $CV_{1\times n}$.

(iii) When $CV_i > \mu_{CV}+3\sigma_{CV}$ or $CV_i < \mu_{CV}-3\sigma_{CV}$ is met, X_{ip} will be removed from $X_{n\times p}$.

C. MEA

The Mind Evolutionary Algorithm (MEA) proposed by Sun et al. can converge rapidly [32, 33]. The similartaxis and dissimilation operation of the MEA can avoid the problem which the crossover and mutation process of genetic algorithms may destroy existing genes.

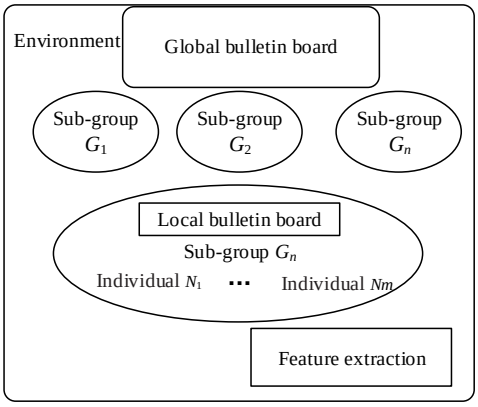

Fig. 3.Structure of MEA

Sub-groups are divided into superior groups and temporary groups. The billboard records the serial number, the action and the score of the group in order to communicate information between individuals or subgroups. The serial number of the group can distinguish between different individuals or subgroups. Local billboard and global billboard post internal information of sub-groups and information of sub-groups respectively. The operation in which individuals compete to become superior individual in a sub-group is called similartaxis. The sub-group is matured when a sub-group does not produce new superior individual in the operation of similartaxis. The operation in which each sub-group evolves in order to be the superior individual of the whole group is called dissimilation. The superior sub-group will be replaced by the temporary sub-group in the operation of dissimilation when the score of a temporary sub-group is higher than that of a superior sub-group. If all the scores of the temporary sub-group are lower than those of the superior sub-group, the iterative operation can be ended. The steps of MEA are as follows:

(i) Firstly, several individuals are randomly generated, and n individuals with the highest scores are divided into N_s superior individuals and N_t temporary individuals. $N_s+N_t=n$.

(ii) Then, superior sub-groups and temporary sub-groups take superior individuals and temporary individuals as the center respectively to produce the new individuals.

(iii) After the operation of similartaxis, the score of the optimal individual is taken as the score of the sub-group.

(iv) With the maturity of sub-groups, the scores of each sub-group are published on the global billboard, and the dissimilation operation is performed.

The score of the MEA is the same as the fitness function of the genetic algorithm, and it is the performance of the individual's ability to adapt to the environment. This paper uses the coefficient of determination (R^2) as the individual's score.

The total sum of squares：

$$SS_{tot}=\sum(y_i-\bar{y})^2 \tag{8}$$

where $\bar{y}$ is the mean of data set.

The sum of squares of residuals：

$$SS_{res}=\sum(y_i-y_i^*)^2 \tag{9}$$

where y_i^* is the forecasted wind speed.

Coefficient of determination：

$$R^2=1-\frac{SS_{res}}{SS_{tot}} \tag{10}$$

In the similartaxis operation, the sub-group of MEA is randomly generated by the center of the group, and the movement direction of the individual relative to the center of the group is also random. The randomness of the algorithm may skip the optimal solution. Therefore, the following improvements are proposed for the similartaxis operation of MEA.

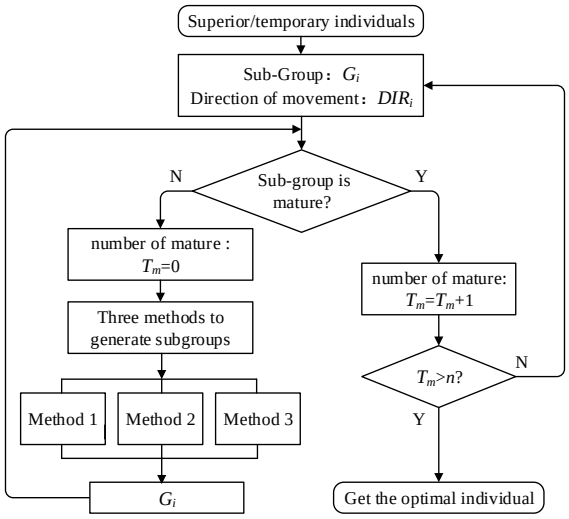

Fig. 4.Improved similartaxis operation

DIR_i is the movement direction of each individual in the sub-group relative to the center of the group, and T_m is the number of maturity. We also use three methods to generate new sub-groups. *n* is the number of consecutive matured.

(i) The superior/temporary individuals are used as the center of the sub-group to randomly generate the sub-group G_i, and DIR_i is recorded.

(ii) Determine whether the sub-group is matured.

(iii) If the sub-group is matured, T_m=T_m+1. If T_m is greater than n, the optimal individual is obtained and the iteration ends, otherwise step (i) is performed.

(iv) If the sub-group is not matured, T_m =0, and a new sub-group is generated. Three methods to generate new sub-groups are as follows:

a) The first method is to take the current superior individual as the center and move forward in the direction of the current superior individual movement.

b) The second method is to randomly generate new sub-groups centered on the current superior individuals.

c) The third method is completely random generation of new individuals.

After generating a new sub-group, step (ii) is performed.

The first method of generating sub-groups can make individuals move to the optimal solution, the second method can enhance the local search ability, and the third method can enhance the global search ability. The algorithm can avoid falling into local optimal solution by specifying the number of maturity. The similartaxis operation of IMEA and MEA are shown in Fig.5 and Fig.6.

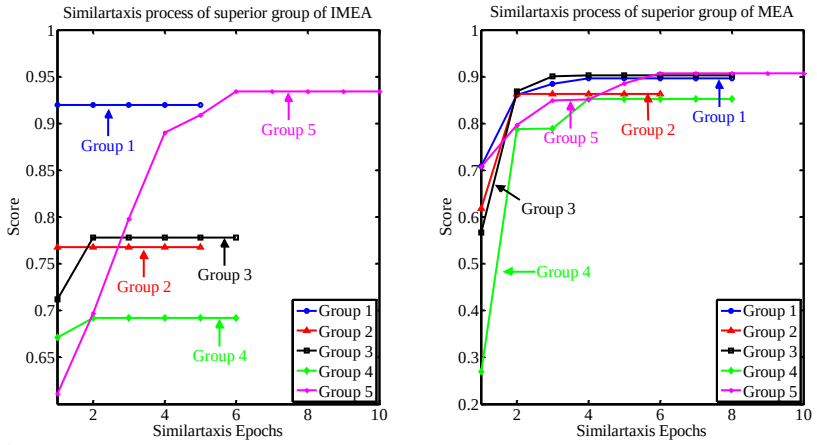

Fig. 5.Similartaxis operation of superior group

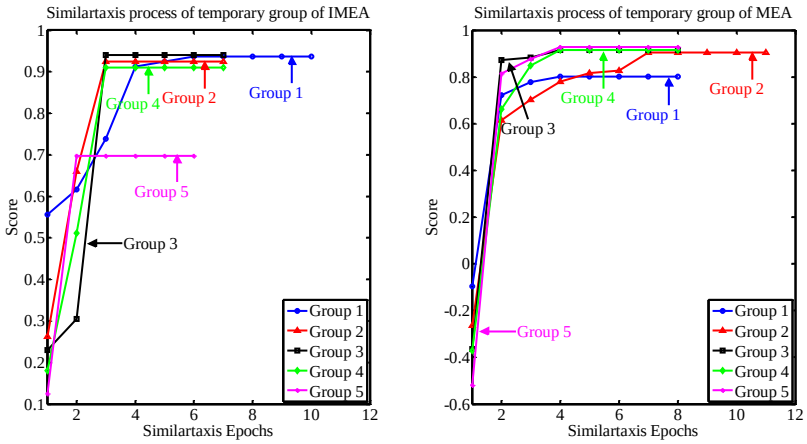

Fig. 6.Similartaxis operation of temporary group

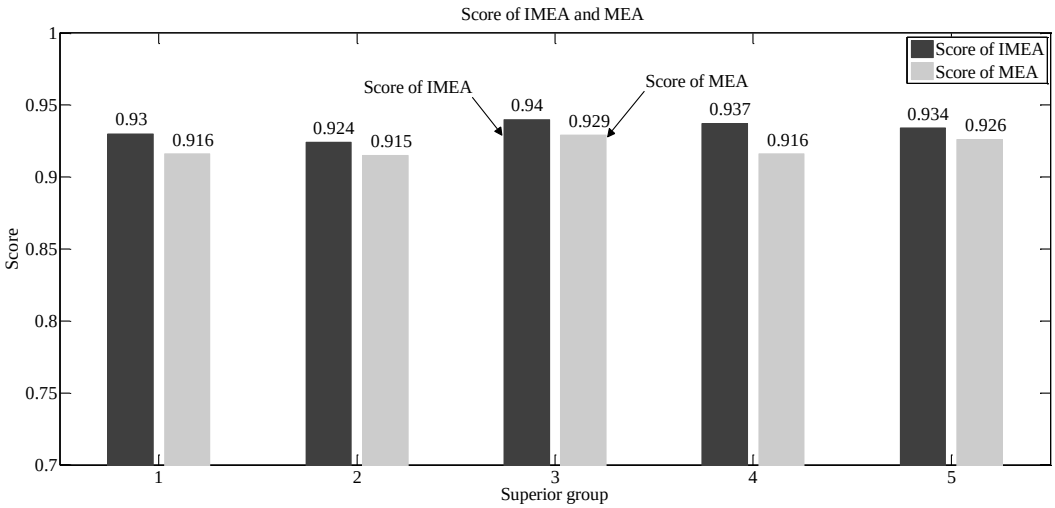

Fig. 7.Score of the superior sub-group

As can be seen from Fig.5 and Fig.6, The number of iterations of IMEA is less than 4, while the number of iterations of MEA is less than 6. The average convergence time of the sub-groups of the IMEA is less than the average convergence time of MEA, so the convergence rate of the IMEA is faster than that of MEA. Fig.7 shows the final score of IMEA and MEA sub-group. The scores of the two algorithms are above 0.9, but the highest score of IMEA is higher than MEA, and the scores of each group of IMEA are also higher than MEA. The optimization result of IMEA is better than that of MEA.

D. IMEA-BP

The initial weights and thresholds of the BP neural network are given randomly, so the forecasting accuracy of the model depends on the choice of initial weights and thresholds. In order to improve this limitation of BP neural network, we use IMEA to optimize the initial weights and thresholds of BP neural network, and then use the optimization result of IMEA as the initial weights and thresholds of BP neural network. The structure of IMEA-BP is shown in Fig.8.

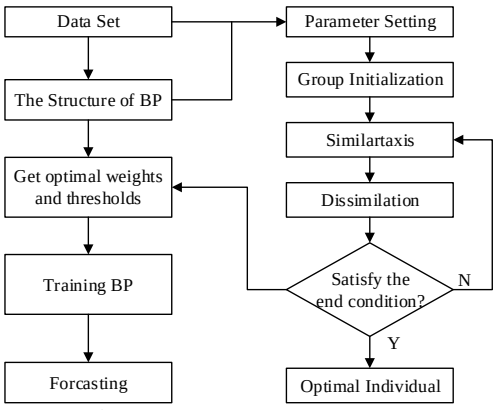

Fig. 8.Structure of IMEA-BP

III. Framework of Hybrid Model

The framework of the hybrid model is shown in Fig.9. For case 1, FCM classifies the input data set into 4 classes. For case 2, FCM classifies the input data into 6 classes. Then outlier detection method is used to eliminate outliers in each class of input data set. Finally, different IMEA-BPs are used for training and forecasting.

The initial group of IMEA contains 100 individuals. The number of superior and temporary sub-groups is 5. The number of individuals in the sub-group is 300. The number of nodes in input layer, hidden layer and output layer of BP neural network is 6, 13 and 1 respectively.

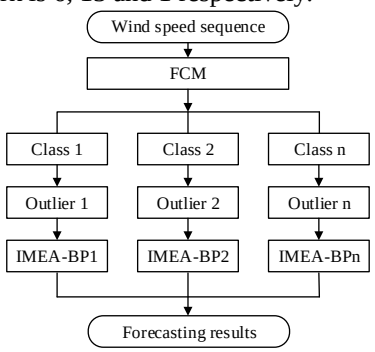

Fig. 9.Framework of hybrid model

IV. Case Study

The following three error criteria are selected to measure the accuracy of wind speed forecasting:

(i) Mean Absolute Error:

$$\mathrm{MAE} = \frac{1}{p}\left|y_i - y_i^*\right| \tag{11}$$

(ii) Root Mean Squared Error:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{p}(y_i - y_i^*)^2}{P}} \tag{12}$$

(iii) Mean Absolute Percentage Error (%):

$$\mathrm{MAPE} = \frac{1}{p}\sum_{i=1}^{p}\left|\frac{y_i - y_i^*}{y_i}\right| \times 100\% \tag{13}$$

where y_i is the true wind speed, y_i^* is the forecasted wind speed, and p is the number of wind speed forecasting.

In this paper, one-step forecasting, two-step forecasting and three-step forecasting are used to verify the validity of the model. The multi-step forecasting diagram is shown in Fig.10.

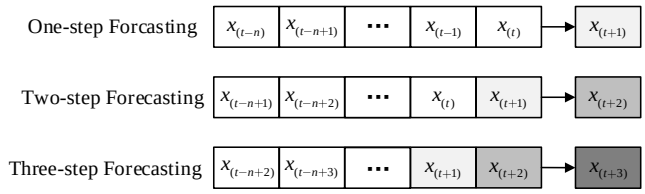

Fig. 10. Multi-step forecasting diagram

A. Case 1

The forecasting error criteria of each model in case 1 are shown in TABLE II. The results of one-step forecasting, two-step forecasting and three-step forecasting are shown in Fig.11, Fig.12 and Fig.13 respectively.

By analyzing the error criteria in TABLE II, for one-step forecasting, the MAE, RMSE and MAPE of the proposed method are 0.25, 0.316 and 4.899% respectively. MAE and RMSE of the proposed method are slightly larger than FCM-BP model, while MAPE is slightly smaller. This shows that the method proposed in is close to the forecasting result of FCM-BP in one-step forecasting. For two-step forecasting and three-step forecasting, each error criteria of the proposed method is the smallest, and its forecasting accuracy is the highest. Comparing proposed method with FCM-MEA-BP, the IMEA algorithm can find more suitable initial weights and thresholds for BP, so the forecasting results are better. The comparison between the forecasting results of ARIMA model and BP shows that BP is more suitable to deal with nonlinear problems than ARIMA. The comparison between BP and Elman model shows that BP is better than Elman in this case. By comparing the multi-step forecasting results of each model, the forecasting accuracy decreases with the increase of forecasting step size.

B. Case 2

The forecasting error criteria of each model in case 2 are shown in TABLE III. The results of one-step forecasting, two-step forecasting and three-step forecasting are shown in Fig.14, Fig.15 and Fig.16 respectively.

TABLE II
FORECASTING ERROR CRITERIA OF DIFFERENT MODELS FOR CASE 1

	MAE			RMSE			MAPE(%)		
	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Proposed Method	0.250	0.293	0.321	0.316	0.379	0.410	4.899	5.494	6.277
FCM-BP	0.249	0.304	0.324	0.312	0.396	0.431	4.973	5.660	6.355
FCM-MEA-BP	0.248	0.301	0.329	0.312	0.393	0.428	4.956	5.687	6.411
ARIMA	0.278	0.340	0.389	0.331	0.425	0.491	5.414	6.355	7.493
BP	0.257	0.323	0.337	0.322	0.414	0.459	5.130	6.116	6.589
Elman	0.280	0.366	0.370	0.347	0.462	0.497	5.594	7.128	7.139

TABLE III
FORECASTING ERROR CRITERIA OF DIFFERENT MODELS FOR CASE 2

	MAE			RMSE			MAPE(%)		
	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Proposed Method	0.276	0.343	0.447	0.354	0.449	0.608	4.903	6.074	7.886
FCM-BP	0.322	0.362	0.460	0.366	0.475	0.615	4.971	6.446	8.214
FCM-MEA-BP	0.284	0.354	0.469	0.366	0.482	0.672	5.137	6.106	8.536
ARIMA	0.310	0.398	0.500	0.412	0.526	0.655	5.857	7.419	9.057
BP	0.294	0.380	0.476	0.405	0.505	0.628	5.554	7.026	8.785
Elman	0.357	0.440	0.497	0.463	0.579	0.647	6.536	8.211	9.053

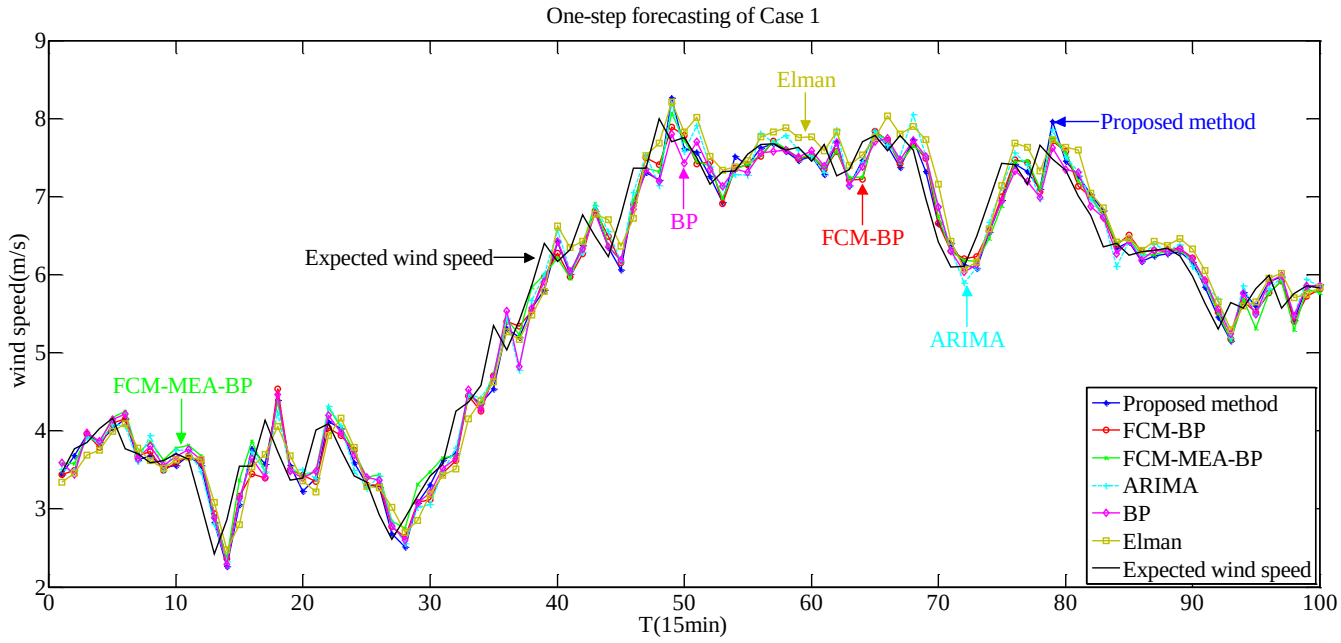

Fig. 11. One-step forecasting of Case 1

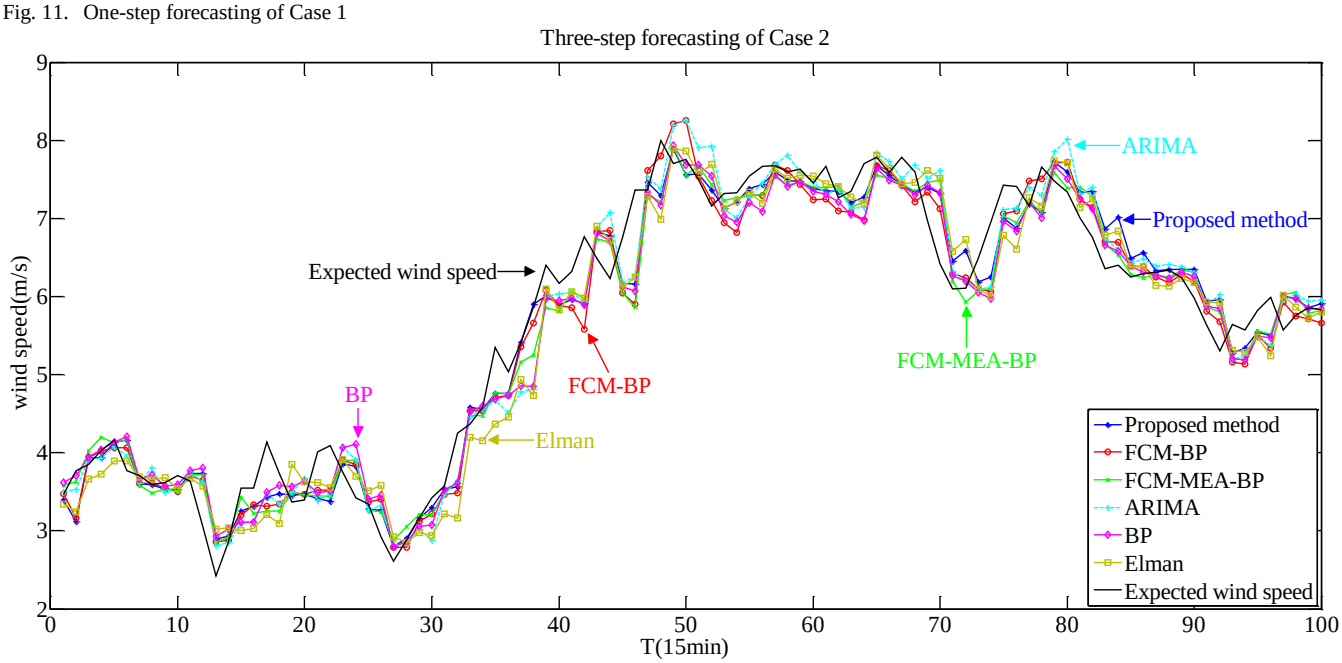

Fig. 12. Two-step forecasting of Case 1

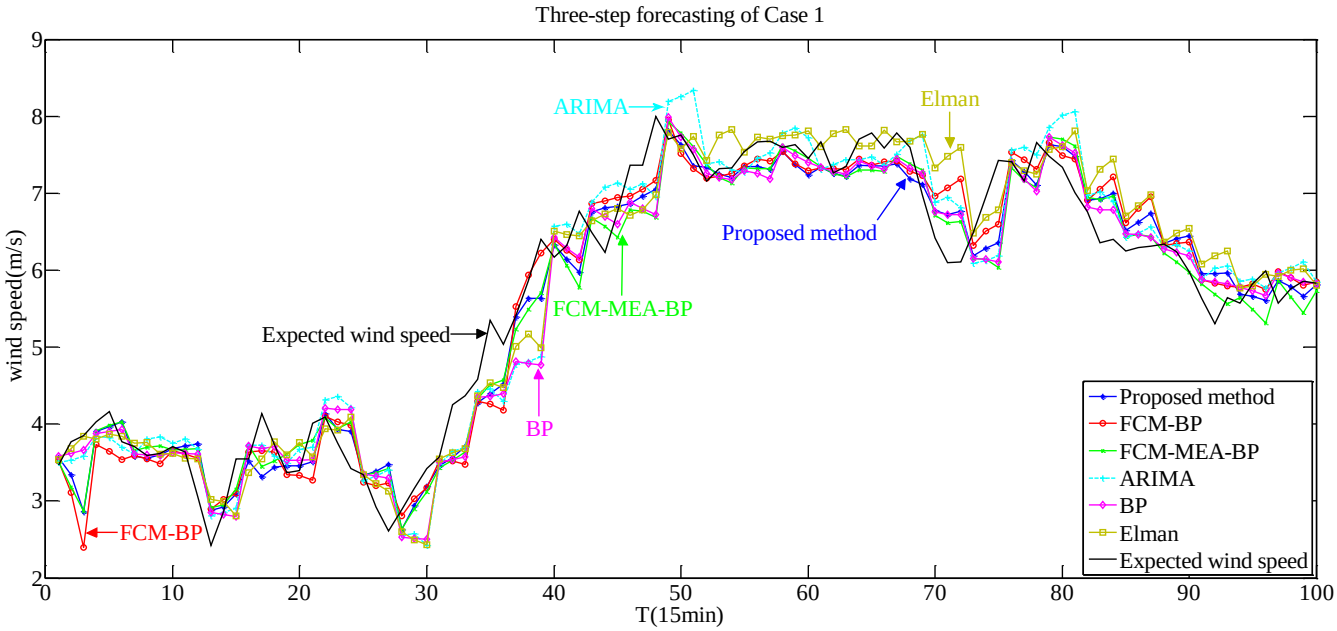

Fig. 13. Three-step forecasting of Case 1

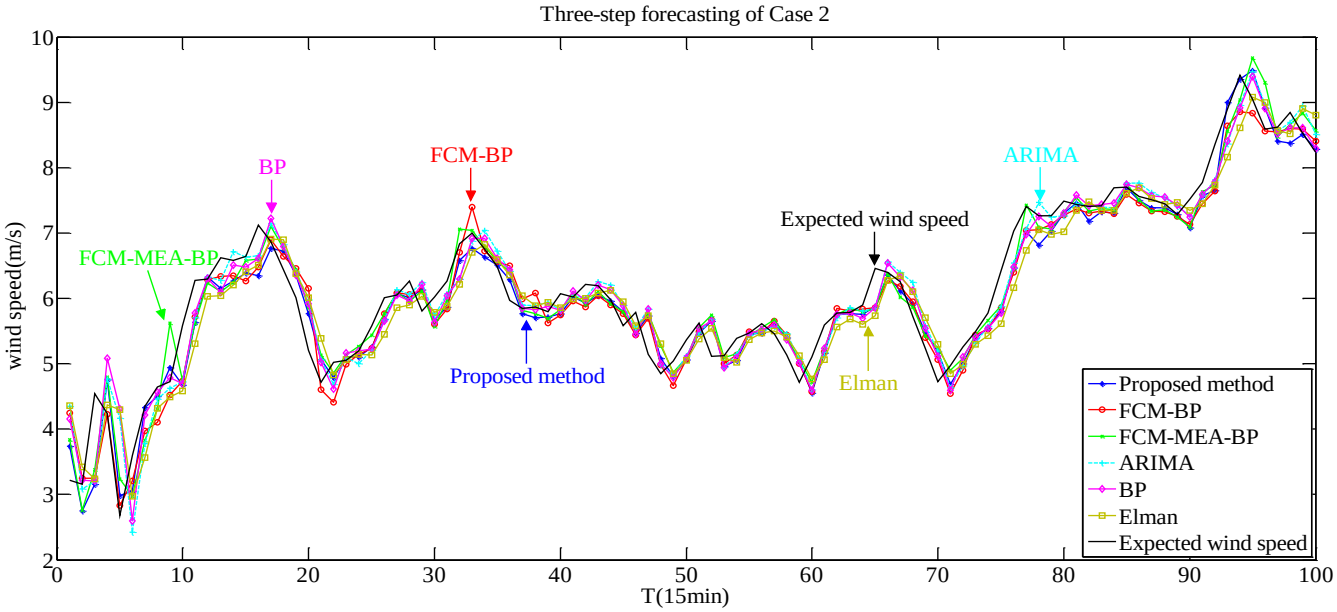

Fig. 14. One-step forecasting of Case 2

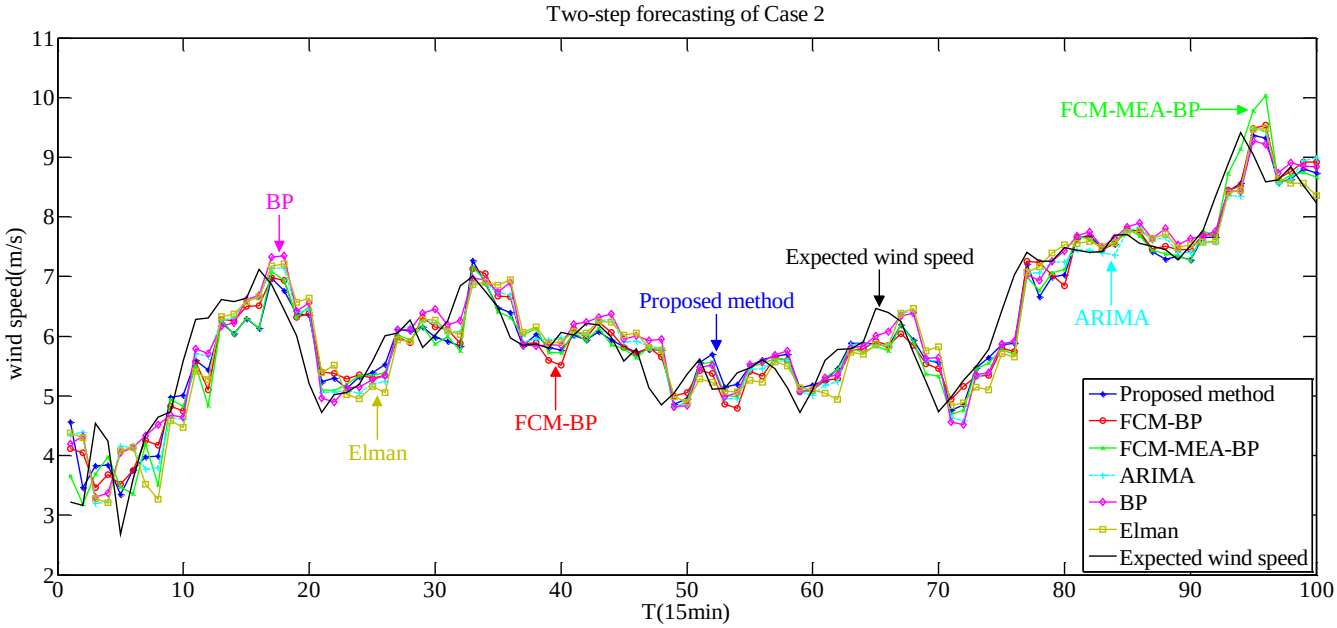

Fig. 15. Two-step forecasting of Case 2

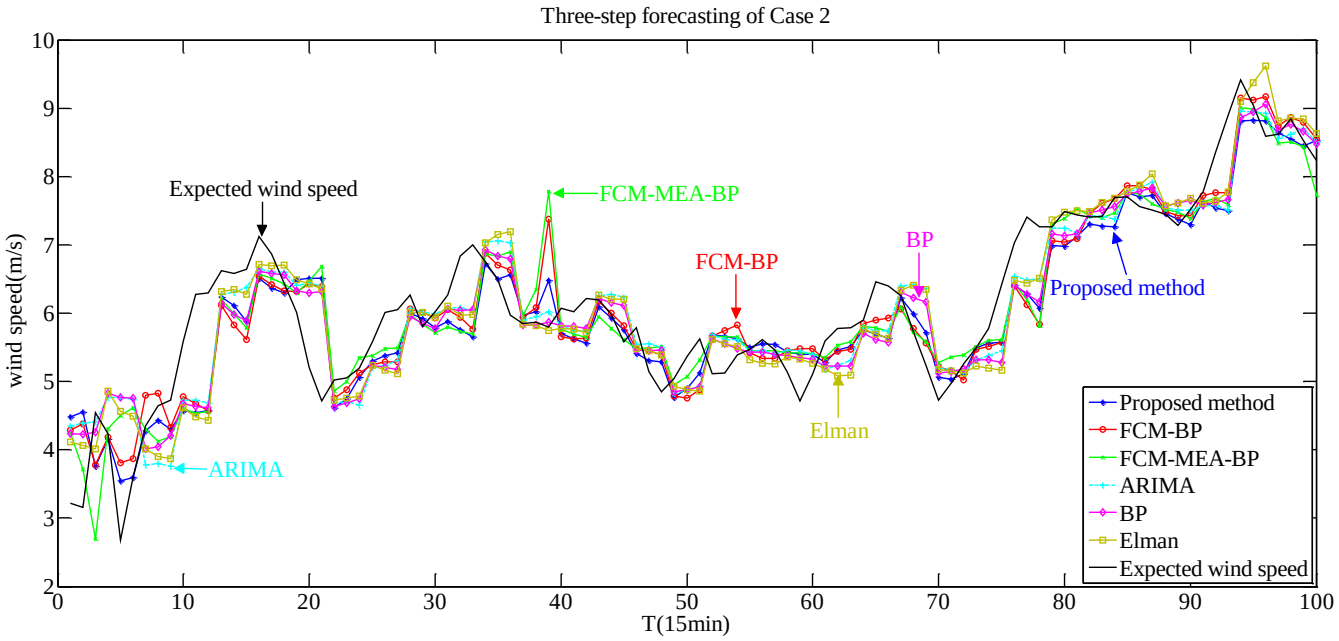

Fig. 16. Three-step forecasting of Case 2

By analyzing the error criteria of the forecasting results in case 2, MAE, RMSE and MAPE of proposed method are the smallest in one-step forecasting, two-step forecasting and three-step forecasting. As the number of forecasting steps increases, proposed method has more obvious advantages compared with other models. Compared with case one, no matter which prediction model is selected, the error criterion is larger than case one. For different wind speed sequences, BP neural network has different adaptability. For the wind speed series with high volatility, the forecasting result is poor, while for the wind speed series with low volatility, the forecasting result is better.

V. Conclusion

This paper combines FCM, outlier detection and IMEA-BP to propose a new forecasting model. Compared with k-means clustering, FCM does not have repeated iterative operation, so FCM can save computing time. FCM is suitable for high dimensional data processing and has good scalability. But for different data sets, the number of class c should be different. The coefficient of variation is used to reflect the dispersion of data, and statistical knowledge is used to remove the input data with large dispersion from the original data set. In addition, this paper proposed improvements to the similartaxis operation of MEA. The number of consecutive matured n can prevent subgroups from falling into local optimal solution. DIR_i reduces randomness in subgroup optimization. Three ways of generating subgroups also ensure the global optimization ability of the algorithm. The IMEA not only enhances the local search ability of the algorithm, but also guarantees a certain global search ability, which can quickly and effectively find the optimal initial weights and thresholds of BP. Through the analysis of case 1 and case 2, the forecasting accuracy of the proposed method is improved. However, the accuracy of multi-step forecasting needs to be further improved.

References

[1] S. Harbola and V. Coors, "One dimensional convolutional neural network architectures for wind prediction", *Energy Conversion and Management*, vol. 195, pp. 70-75, 2019.

[2] M. Yang, X. Chen, J. Du, and Y. Cui, "Ultra-short-term multistep wind power prediction based on improved EMD and reconstruction method using run-length analysis", *IEEE Access*, vol. 6, pp. 31908-31917, 2018.

[3] B. Khorramdel, C.Y. Chung, N. Safari, and G.C.D. Price, "A fuzzy adaptive probabilistic wind power prediction framework using diffusion kernel density estimators", *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 7109-7121, 2018.

[4] C. Zhang, H. Wei, L. Xie, Y. Shen, and K. Zhang, "Direct interval forecasting of wind speed using radial basis function neural networks in a multi-objective optimization framework", *Neurocomputing*, vol. 205, pp. 53-63, 2016.

[5] X. Yang, X. Ma, N. Kang, and M. Maihemuti, "Probability Interval Prediction of Wind Power Based on KDE Method with Rough Sets and Weighted Markov Chain", *IEEE Access*, vol. 6, pp. 51556-51565, 2018.

[6] J. Wang, W. Zhang, J. Wang, T. Han, and L. Kong, "A novel hybrid approach for wind speed prediction", *Information Sciences*, vol. 273, pp. 304-318, 2014.

[7] J. Jung and R.P. Broadwater, "Current status and future advances for wind speed and power forecasting", *Renewable and Sustainable Energy Reviews*, vol. 31, pp. 762-777, 2014.

[8] Y. Zhou and M. Huang, "Lithium-ion batteries remaining useful life prediction based on a mixture of empirical mode decomposition and ARIMA model", *Microelectronics Reliability*, vol. 65, pp. 265-273, 2016.

[9] D.C. Kiplangat, K. Asokan and K.S. Kumar, "Improved week-ahead predictions of wind speed using simple linear models with wavelet decomposition", *Renewable Energy*, vol. 93, pp. 38-44, 2016.

[10] S. Sim, P. Maass and P.G. Lind, "Wind speed modeling by nested ARIMA processes", *Energies*, vol. 12, no. 1, 2019.

[11] A. Zameer, J. Arshad, A. Khan, and M.A.Z. Raja, "Intelligent and robust prediction of short term wind power using genetic programming based ensemble of neural networks", *Energy Conversion and Management*, vol. 134, pp. 361-372, 2017.

[12] Y. Zhang, B. Chen, Y. Zhao, and G. Pan, "Wind Speed Prediction of IPSO-BP Neural Network Based on Lorenz Disturbance", *IEEE Access*, vol. 6, pp. 53168-53179, 2018.

[13] Z. Shi, H. Liang and V. Dinavahi, "Wavelet neural network based multiobjective interval prediction for short-term wind speed", *IEEE Access*, vol. 6, pp. 63352-63365, 2018.

[14] H.S. Dhiman, D. Deb and J.M. Guerrero, "Hybrid machine intelligent SVR variants for wind forecasting and ramp events", *Renewable and Sustainable Energy Reviews*, vol. 108, pp. 369-379, 2019.

[15] S. Wang, N. Zhang, L. Wu, and Y. Wang, "Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and

GA-BP neural network method", *Renewable Energy*, vol. 94, pp. 629-636, 2016.
[16] W. Sun and Y. Wang, "Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and improved back-propagation neural network", *Energy Conversion and Management*, vol. 157, pp. 1-12, 2018.
[17] J. Song, J. Wang and H. Lu, "A novel combined model based on advanced optimization algorithm for short-term wind speed forecasting", *Applied Energy*, vol. 215, pp. 643-658, 2018.
[18] C. Yu, Y. Li, H. Xiang, and M. Zhang, "Data mining-assisted short-term wind speed forecasting by wavelet packet decomposition and Elman neural network", *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 175, pp. 136-143, 2018.
[19] C. Yu, Y. Li and M. Zhang, "An improved Wavelet Transform using Singular Spectrum Analysis for wind speed forecasting based on Elman Neural Network", *Energy Conversion and Management*, vol. 148, pp. 895-904, 2017.
[20] H. Liu, X. Mi and Y. Li, "Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network", *Energy Conversion and Management*, vol. 156, pp. 498-514, 2018.
[21] C. Ren, N. An, J. Wang, L. Li, B. Hu, and D. Shang, "Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting", *Knowledge-Based Systems*, vol. 56, pp. 226-239, 2014.
[22] S. Sun, H. Qiao, Y. Wei, and S. Wang, "A new dynamic integrated approach for wind speed forecasting", *Applied Energy*, vol. 197, pp. 151-162, 2017.
[23] H. Shao, X. Deng and Y. Jiang, "A novel deep learning approach for short-term wind power forecasting based on infinite feature selection and recurrent neural network", *Journal of Renewable and Sustainable Energy*, vol. 10, no. 4, 2018.
[24] A. Tascikaraoglu, B.M. Sanandaji, K. Poolla, and P. Varaiya, "Exploiting sparsity of interconnections in spatio-temporal wind speed forecasting using Wavelet Transform", *Applied Energy*, vol. 165, pp. 735-747, 2016.
[25] A. Aghajani, R. Kazemzadeh and A. Ebrahimi, "A novel hybrid approach for predicting wind farm power production based on wavelet transform, hybrid neural networks and imperialist competitive algorithm", *Energy Conversion and Management*, vol. 121, pp. 232-240, 2016.
[26] Z. Wang, G. Ma, D. Gong, J. Sun, and D. Zhang, "Application of Mind Evolutionary Algorithm and Artificial Neural Networks for Prediction of Profile and Flatness in Hot Strip Rolling Process", *Neural Processing Letters*, 2019.
[27] H. Liu, H. Tian, X. Liang, and Y. Li, "New wind speed forecasting approaches using fast ensemble empirical model decomposition, genetic algorithm, Mind Evolutionary Algorithm and Artificial Neural Networks", *Renewable Energy*, vol. 83, pp. 1066-1075, 2015.
[28] L. Xu, X. Du and B. Wang, "Short-Term Traffic Flow Prediction Model of Wavelet Neural Network Based on Mind Evolutionary Algorithm", *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 32, no. 12, 2018.
[29] J.C. Bezdek, R. Ehrlich and W. Full, "FCM: The fuzzy c-means clustering algorithm", *Computers and Geosciences*, vol. 10, pp. 191-203, 1984.
[30] J.C. Bezdek, R.J. Hathaway, M.J. Sabin, and W.T. Tucker, "Convergence theory for fuzzy c-means: counterexamples and repairs", *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-17, pp. 873-877, 1987.
[31] A.K. Abd-Elaal, H.A. Hefny and A.H. Abd-Elwahab, "Forecasting of Egypt wheat imports using multivariate fuzzy time series model based on fuzzy clustering", *IAENG International Journal of Computer Science*, vol. 40, no. 4, pp. 230-237, 2013.
[32] N. Cui, J. Wei, L. Zhao, Q. Zhang, D. Gong, and M. Wang, "Reference Crop Evapotranspiration Prediction Model of Arid Areas of Northwest China Based on MEA-BPNN", *Transactions of the Chinese Society for Agricultural Machinery*, vol. 49, no. 8, pp. 228-236 and 307, 2018.
[33] W. Wang, R. Tang, C. Li, P. Liu, and L. Luo, "A BP neural network model optimized by Mind Evolutionary Algorithm for predicting the ocean wave heights", *Ocean Engineering*, vol. 162, pp. 98-107, 2018.