IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 26, 2019, accepted September 10, 2019, date of publication September 17, 2019, date of current version October 1, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2942012

Wind Power Prediction Based on PSO-SVR and Grey Combination Model

YI ZHANG, HEXU SUN, (Senior Member, IEEE), AND YINGJUN GUO
[1]School of Artificial Intelligence, Hebei University of Technology, Tianjin 300130, China
[2]School of Information Science and Engineering, Hebei University of Science and Technology, Shijiazhuang 050018, China

Corresponding author: Hexu Sun (sun_prof@163.com)

This work was supported in part by the China National Natural Science Foundation under Grant 61472273, and in part by the Scientific Research Foundation of Hebei Education Department under Grant QN2016109.

ABSTRACT As a kind of green, clean and renewable energy, wind power generation has been widely utilized in various countries in the world. With the rapid development of wind energy, it is also facing prominent problems. Because wind power generation is intermittent, unstable and stochastic, it has caused serious difficulties for power grid dispatch. At present, the important method to solve this problem is to predict wind speed and wind power. Grey model is suitable for uncertain systems with poor information and needs less operation data, so it can be used for wind speed and wind power prediction. However, the traditional grey system model has the disadvantage of low prediction accuracy. Therefore, firstly the GM (1,1) for wind speed prediction is improved by background value optimization in this paper. In order to comprehensively reveal the inherent uncertainty of wind speed random series, the fractional order grey system models with different orders are constructed. Secondly, in order to overcome the shortcoming of single grey model, each grey model is effectively united, and a combination prediction model based on neural network is presented. The two NWP outputs, i.e. ECMWF and GRAPES-MESO, have been added to the prediction model for reducing the uncertainty. The structure parameters of the neural network are optimized by trial and error. Thirdly, the support vector regression model is established to fit the scatter operation data of wind speed-power, and the parameters of the model are optimized by the particle swarm algorithm. Then the power prediction value is obtained by the fitted wind speed-power relationship and the corresponding result of the grey combination model for wind speed prediction. Finally, wind speed and wind power are predicted based on the actual operation data. In addition, the prediction model based on ARIMA is also constructed as a benchmark model. The results show that the proposed grey combination model improves the prediction accuracy.

INDEX TERMS Grey model, background value, support vector regression, fractional order, particle swarm algorithm, wind speed prediction, wind power prediction, combination model.

I. INTRODUCTION

With the rapid development of global economy, energy demand is rapidly growing. The traditional fossil energy has been depleted, and the problems of climate warming and environmental pollution are becoming increasingly serious. As a clean and renewable resource, wind energy has been paid more and more attention. Wind power generation has been developing rapidly all over the world [1]–[3]. In 2018, the new installed capacity of wind power generation in the world was 53.9 GW, and the cumulative installed capacity exceeded 600 GW for the first time. In China, the new installed capacity was 21 GW, and the cumulative installed capacity was 221 GW.

Due to the rising scale of grid-connected wind power generation, the randomness, fluctuation and uncertainty of wind energy have a profound impact on the security, stability and economy of power system [4]–[7]. At present, the important method to solve this problem is to carry out wind power prediction research. According to the historical data and current information of wind speed and wind turbine, the changing trend of wind power generation is predicted in order to enhance the safety, reliability and controllability of the system [8]–[10]. Thus, the grid dispatch plan is

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

optimized. The ability of peak regulation is improved and the spinning reserve capacity is also reduced. With the decrease of wind abandonment, the wind power generation can meet the requirement of electricity market transaction and actively participate in electricity market bidding.

At present, there are some methods for wind power prediction. The prediction method based on the physical model needs to establish thermodynamic and kinetic equations describing the atmosphere layer evolution. Considering the boundary conditions and the actual topography, the trend of wind speed and wind power is predicted. The prediction model constructed by this method is complex with huge computation complexity, and it is very sensitive to false initial information [11]–[13].

The wind power prediction models based on machine learning are to establish the calculation procedure for solving forecasting problem by means of bio-intelligence [14]–[18]. These methods include the fuzzy logic algorithm for simulating fuzzy concept in the human brain [19]–[22], the long-term and short-term memory network (LSTM) , the expert experience method [23]–[26] ,etc. These models can be applied to the problem with unclear intrinsic mechanism and the non-linear relationship between input variable and output variable is established. However, the robustness of the algorithm needs further proof, and a lot of data and computation are needed for model training and parameter estimation.

The prediction method based on the Weather Research and Forecasting ensembles has been developed in order to overcome the disadvantage of the single NWP prediction [27]. The recent initial condition and the high-resolution model do not always contribute to the performance enhancement. The novel fuzzy method has been adopted to measure the effectiveness of the each prediction and select the best three members to establish the final prediction result. The cuckoo research method is applied to optimize the model.

The hybridization of the numerical weather prediction and machine learning algorithm has been presented for wind speed forecasting [28]–[31]. The diversity of the data is produced by the system model with the different initial parameters. The final forecasting output of the model is processed by means of the intelligent no-linear method in order to exploit the diversity of the system.

Machine learning approach is increasingly adopted to extract the inner patterns from the data, which is published in the latest journal Nature in 2019 [32]. Physical model (theory driven) and machine learning model (data driven) have different paradigms. But the two methods can complement each other. The former has strong extrapolation ability. The latter is more flexible, and new laws can be found. The combination of these two methods can improve parameterization and replace the physical sub-model with machine learning approach [33].

The grey prediction model is suitable for the uncertain system with only known partial information. The model needs less historical data and computation, and can adapt to the dynamic changes of data. Therefore, it is widely used in agriculture, electricity, finance and other fields [34]–[36]. However, the traditional single grey model has the disadvantage of low prediction accuracy.

The contributions of the prediction model proposed in this paper are summarized as follows.

(1) Compared with the single grey prediction model, an improved grey model based on background value optimization is proposed. In order to better reveal the inherent law of wind speed and wind power series, the fractional order grey models with different orders are designed, and the complementary characteristic of each model is analyzed.

(2) In order to integrate single prediction model well, a combination prediction model based on multi-hidden layers neural network is established. The NWP output is applied to the combination model to reduce the uncertainty of the system. In order to evaluate the prediction performance of the proposed model, the ARIMA prediction model is also constructed as the baseline method.

(3) According to the actual wind speed-wind power operation data, a wind speed-power curve-fitting model is constructed by support vector regression, and the parameters of the model are optimized by particle swarm algorithm.

II. PREDICTION METHOD

A. PSO-SVR FITTING ALGORITHM

Support vector machine method maps input space to high-dimensional feature space using kernel function, and obtains the non-linear relationship between input and output variables. By minimizing the structure risk, the generalization ability of the model can be improved, so as to obtain good statistical law in the case of fewer input samples. Support vector regression has advantage in solving non-linear problem and can obtain the global optimal solution [37]–[39].

Support vector machine adopts ε-insensitivity function, which assumes that all training data are fitted under accuracy ε as follows:

$$\begin{cases} y_i - f(x_i) \le \varepsilon + \xi_i \\ f(x_i) - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases} \qquad i = 1, 2, \ldots, n \tag{1}$$

In the formula (1), ξ_i, ξ_i^* are the relaxation factors. The problem is to minimize the objective function:

$$R(\omega, \xi, \xi^*) = \frac{1}{2}\omega \cdot \omega + C \sum_{i=1}^{n} (\xi_i + \xi_i^*) \tag{2}$$

The constant C represents the penalty degree for the samples exceeding the accuracy range. The dual form of Lagrange function is to maximize the following function:

$$\begin{aligned} W(\alpha, \alpha^*) = &-\frac{1}{2} \sum_{i=1,j=1}^{n} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) \cdot K(x \cdot x_i) \\ &+ \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) y_i - \sum_{i=1}^{n} (\alpha_i + \alpha_i^*)\varepsilon \end{aligned} \tag{3}$$

The constraints are as follows:

$$\begin{cases} \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) = 0 \\ 0 \le \alpha_i,\ \alpha_i^* \le C \end{cases} \tag{4}$$

The radial basis kernel function is as follows:

$$K(x, y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \tag{5}$$

The expression of the obtained non-linear fitting function is as follows:

$$f(x) = \omega \cdot \phi(x) + b = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x, x_i) + b \tag{6}$$

Thus, the key problem of support vector regression algorithm is how to solve parameters α_i, α_i^*. These parameters α_i, α_i^* are determined by the sequential minimal optimization method.

However, the performance of the support vector regression model is greatly influenced by the hyper-parameters, such as penalty factor C, accuracy ε and kernel function variance σ. How to scientifically and effectively identify these hyper-parameters is a issue to be tackled. Therefore, it is necessary to introduce optimization algorithm into hyper-parameters search process.

Particle swarm optimization algorithm is a parallel global search strategy based on population. It has faster convergence speed. The optimal particle is found in each iteration, and the rest of the particles can follow it. Particles are updated according to two extremums: one is its own extremum, which represents the cognitive level of the particle itself in finding the best position; the other is the global extremum, which represents the ability of all particles to find the best solution at present [40].

The particle swarm optimization uses the velocity-position search model. The iteration formula adjusting the position and speed of the particle is as follows:

$$\begin{aligned} V_i(k+1) = w * V_i(k) + c_1 * r_1 * (p_{ibest} - X_i(k)) \\ + c_2 * r_2 * (g_{best} - X_i(k)) \end{aligned} \tag{7}$$

c_1 is a cognitive learning factor that regulates the step size of the particle to its best position. c_2 is a social learning factor, which is used to adjust the step size of the particle to the global best position. w is the weight factor. The value automatically decreases with the iteration, which is defined as:

$$w = w_{\min} + (N_{\max} - n) * (w_{\max} - w_{\min}) / N_{\max} \tag{8}$$

Thus w decreases linearly from 0.9 to 0.4 with the iteration .

The SVR fitting process optimized by particle swarm algorithm is shown in figure 1.

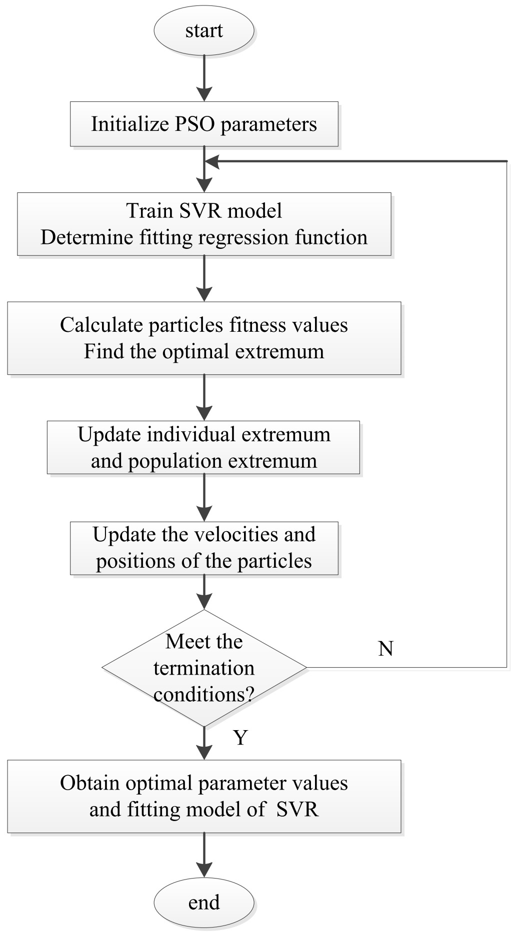


FIGURE 1. PSO-SVR method flow.

B. GREY SYSTEM PREDICTION MODEL

Grey system theory extends the viewpoints and methods of information theory, system theory and cybernetics. The grey system theory is applicable to the study of systems with small samples and poor information. By means of the development and analysis of partial information, useful information can be extracted, and the behavior and evolution of the system can be predicted [41].

1) GM(1,1)

The original wind speed series is as follows: $V^{(0)}(i)$ $i = 1, \cdots, n$. In order to weaken the inherent randomness of the original series and enhance its regularity, the one-order accumulated generation series are calculated as follows:

$$V^{(1)}(k) = \sum_{i=1}^{k} V^{(0)}(i) \quad \forall k = 1, \cdots, n \tag{9}$$

The background value is constructed according to the trapezoid method as follows:

$$Z^{(1)}(k) = 0.5V^{(1)}(k) + 0.5V^{(1)}(k-1) \tag{10}$$

The difference equation of GM (1,1) is described by the following formula.

$$V^{(0)}(k) + aZ^{(1)}(k) = b \quad (k = 2, \cdots, n) \tag{11}$$

where a is the development coefficient, b is the grey action coefficient.

The whitening equation of GM (1,1) is established as follows:

$$\frac{dV^{(1)}}{dt} + aV^{(1)} = b \tag{12}$$

The parameters a and b are calculated by the least square method .

$$\begin{pmatrix} a \\ b \end{pmatrix} = \left(\gamma^T \gamma\right)^{-1} \gamma^T Y \tag{13}$$

where

$$\gamma = \begin{pmatrix} -Z^{(1)}(2) & 1 \\ -Z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -Z^{(1)}(n) & 1 \end{pmatrix} Y = \begin{pmatrix} V^{(0)}(2) \\ V^{(0)}(3) \\ \vdots \\ V^{(0)}(n) \end{pmatrix} \tag{14}$$

Assuming the initial condition $V^{(1)^\wedge}(1) = V^{(1)}(1)$, the grey prediction $V^{(1)^\wedge}(k)$ for the one-order accumulated generation series $V^{(1)}(k)$ is established as follows:

$$V^{(1)^\wedge}(k) = \left(V^{(1)}(1) - \frac{b}{a}\right) e^{-a(k-1)} + \frac{b}{a} \quad k = 1, 2, \cdots \tag{15}$$

Prediction value of original series is obtained by the inverse accumulated generation operation.

$$V^{(0)^\wedge}(k) = V^{(1)^\wedge}(k) - V^{(1)^\wedge}(k-1) \tag{16}$$

Considering that the new data can reflect the change law of the series better than the old data, a rolling mechanism is proposed, which replaces the old data with the new data.

2) DISCRETE GREY MODEL

The discrete grey model , i.e. DGM (1,1), is represented as follows:

$$V^{(1)}(k+1) = \alpha_1 V^{(1)}(k) + \alpha_2 \tag{17}$$

These parameters α_1 and α_2 are solved by the least square method as follows:

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \left(\gamma^T \gamma\right)^{-1} \gamma^T Y \tag{18}$$

where

$$\gamma = \begin{pmatrix} V^{(1)}(1) & 1 \\ V^{(1)}(2) & 1 \\ \vdots & \vdots \\ V^{(1)}(n-1) & 1 \end{pmatrix} Y = \begin{pmatrix} V^{(1)}(2) \\ V^{(1)}(3) \\ \vdots \\ V^{(1)}(n) \end{pmatrix} \tag{19}$$

Assuming the initial condition $V^{(1)^\wedge}(1) = V^{(1)}(1)$, the prediction value of the discrete grey model is calculated as follows:

$$V^{(1)^\wedge}(k) = \alpha_1^{k-1} V^{(0)}(1) + \frac{1-\alpha_1^{k-1}}{1-\alpha_1}\alpha_2 \quad k = 1, 2, \cdots \tag{20}$$

3) FRACTIONAL ORDER GREY MODEL

Fractional order model implies an "in-between" idea, which decreases the perturbation bound of grey prediction model.

Definition 1: $\frac{p}{q}(0 < \frac{p}{q} < 1)$-order accumulation operator is expressed as follows:

$$v^{(\frac{p}{q})}(k) = \sum_{i=1}^{k} C_{k-i+\frac{p}{q}-1}^{k-i} v^{(0)}(i) \tag{21}$$

where $C_{\frac{p}{q}-1}^{0} = 1$, $C_{k-1}^{k} = 0, k = 1, 2, \cdots, n$,

$$C_{k-i+\frac{p}{q}-1}^{k-i} = \frac{(k-i+\frac{p}{q}-1)(k-i+\frac{p}{q}-2)\cdots(\frac{p}{q}+1)\frac{p}{q}}{(k-i)!} \tag{22}$$

Definition 2: $\frac{p}{q}(0 < \frac{p}{q} < 1)$-order inverse accumulation operator is expressed as follows:

$$\varsigma^{(1)} v^{(1-\frac{p}{q})}(k) = v^{(1-\frac{p}{q})}(k) - v^{(1-\frac{p}{q})}(k-1) \tag{23}$$

Definition 3: $\frac{p}{q}(0 < \frac{p}{q} < 1)$-order accumulation discrete grey model ,i.e. $DGM^{(\frac{p}{q})}(1,1)$, is established as follows:

$$v^{(\frac{p}{q})}(k+1) = \alpha_1 v^{(\frac{p}{q})}(k) + \alpha_2, k = 1, 2, \cdots, n-1 \tag{24}$$

α_1 and α_2 are solved by the least square method as follows:

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = \left(\gamma^T \gamma\right)^{-1} \gamma^T Y$$

where

$$\gamma = \begin{pmatrix} V^{(\frac{p}{q})}(1) & 1 \\ V^{(\frac{p}{q})}(2) & 1 \\ \vdots & \vdots \\ V^{(\frac{p}{q})}(n-1) & 1 \end{pmatrix} Y = \begin{pmatrix} V^{(\frac{p}{q})}(2) \\ V^{(\frac{p}{q})}(3) \\ \vdots \\ V^{(\frac{p}{q})}(n) \end{pmatrix} \tag{25}$$

The modeling procedure of $DGM^{(\frac{p}{q})}(1,1)$ is described as follows:

Step 1: $\frac{p}{q}(0 < \frac{p}{q} < 1)$-order accumulation series is calculated as follows: $V^{(\frac{p}{q})} = (v^{(\frac{p}{q})}(1), v^{(\frac{p}{q})}(2), \cdots, v^{(\frac{p}{q})}(n))$.

Step 2: α_1 and α_2 are solved according to the formula (25).

Step 3: The prediction value of $\frac{p}{q}(0 < \frac{p}{q} < 1)$-order accumulation series is obtained.

$$V^{(\frac{p}{q})^\wedge}(k) = \left(V^{(0)}(1) - \frac{\alpha_2}{1-\alpha_1}\right)\alpha_1^{k-1} + \frac{\alpha_2}{1-\alpha_1} \quad k = 1, 2, \cdots$$

Step 4: $\frac{p}{q}(0 < \frac{p}{q} < 1)$-order inverse accumulation operation of $V^{(\frac{p}{q})^\wedge}(k)$ is carried out to obtain the prediction value of the raw data series, i.e. $\hat{V}^{(0)}(k) = \varsigma^{(\frac{p}{q})} V^{(\frac{p}{q})^\wedge}(k)$.

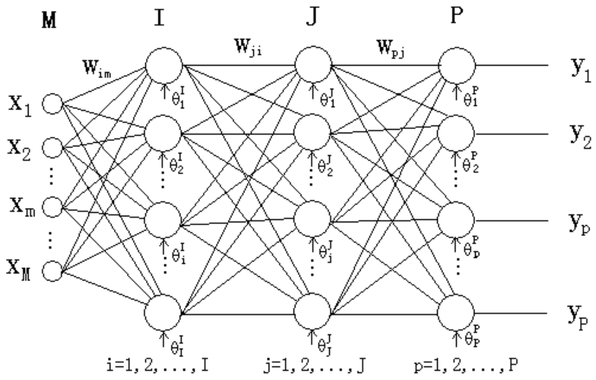


FIGURE 2. Neural network structure.

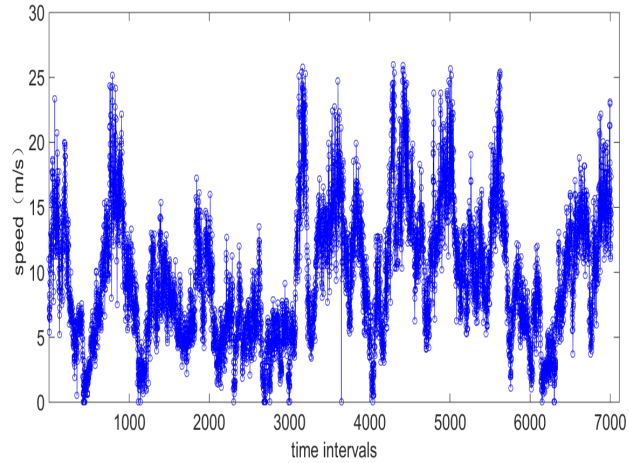


FIGURE 3. Wind speed series.

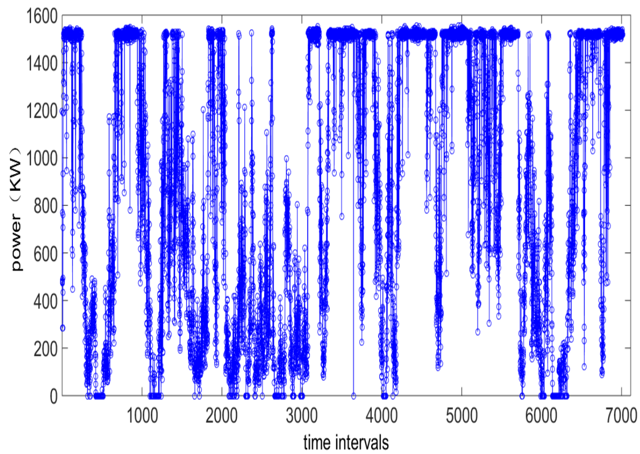


FIGURE 4. Wind turbine power series.

C. NEURAL NETWORK COMBINATION MODEL

Neural network has the advantages of distributed storage and processing, self-adaptation and self-learning. It is especially suitable for dealing with grey information system and can sufficiently fit complex non-linear relation [42]–[44].

The structure of the neural network with multi-hidden layers is shown in Figure 2.The back-propagation algorithm can be described in two steps as follows: (1) The input signal is transmitted from the input layer to the output layer. (2) Error signal is propagated from the output layer to the input layer.

The corresponding updating expression can be derived as follows.

$$\Delta w_{pj}^{PJ}(n) = -\eta \frac{\partial E(n)}{\partial w_{pj}^{PJ}(n)} = \eta \delta_p^P(n) v_j^J(n) \quad (26)$$

$$\Delta w_{ji}^{JI}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}^{JI}(n)} = \eta \delta_j^J(n) v_i^I(n) \quad (27)$$

$$\Delta w_{im}^{IM}(n) = -\eta \frac{\partial E(n)}{\partial w_{im}^{IM}(n)} = \eta \delta_i^I(n) x_{km}(n) \quad (28)$$

where

$$\delta^P(n) = 2 * (d_k - Y_k). * Y_k. * (1 - Y_k) \quad (29)$$

$$\delta^J(n) = v^J(n). * (1 - v^J(n)). * \{(\delta^P(n)^T * W^{PJ}(n)\}^T \quad (30)$$

$$\delta^I(n) = v^I(n). * (1 - v^I(n)). * \{(\delta^J(n)^T * W^{JI}(n)\}^T \quad (31)$$

The learning rate η of neural network is set between zero and one. In order to speed up the learning convergence process, the additional momentum method is adopted.

III. MATERIAL AND SCHEME

A. WIND SPEED AND WIND POWER DATA

Wind speed and wind turbine power data are derived from the actual operation system of some wind farm in China. The sampling interval is ten minutes. The variation of wind speed with time is shown in figure 3, which shows obvious randomness and uncertainty. The range of wind speed is wide. The maximum wind speed can reach 25.93 m/s, and the minimum wind speed is 0 m/s, i.e. no-wind period. Within ten minutes interval, the maximum change rate of wind speed is 10.38 m/s, when the wind speed violently fluctuates. In some period, the wind speed is stable, and the change rate of wind speed is nearly zero.

The variation of wind turbine power with time is shown in figure 4. When the wind speed is higher than the rated wind speed of wind turbine, the wind turbine power reaches 1.5 MW, i.e. the rated power. Therefore, compared with the wind speed series in figure 3, the power output is obviously limited. When the wind speed is lower than the cut-in wind speed, the power is close to 0 W. Thus, the period of power approaching zero is obviously increased. When the actual wind speed is between the cut-in and the rated wind speed, the wind turbine power is increased with the enhancement of the wind speed, and vice versa.

B. PREDICTION MODEL DEVELOPMENT

The wind power prediction model is shown in figure 5. Firstly, the grey system model is suitable for uncertain system with partly known information. The random fluctuation of wind speed shows inherent uncertainty. So the grey model is selected to predict wind speed. However, the prediction accuracy of the traditional grey model is low,

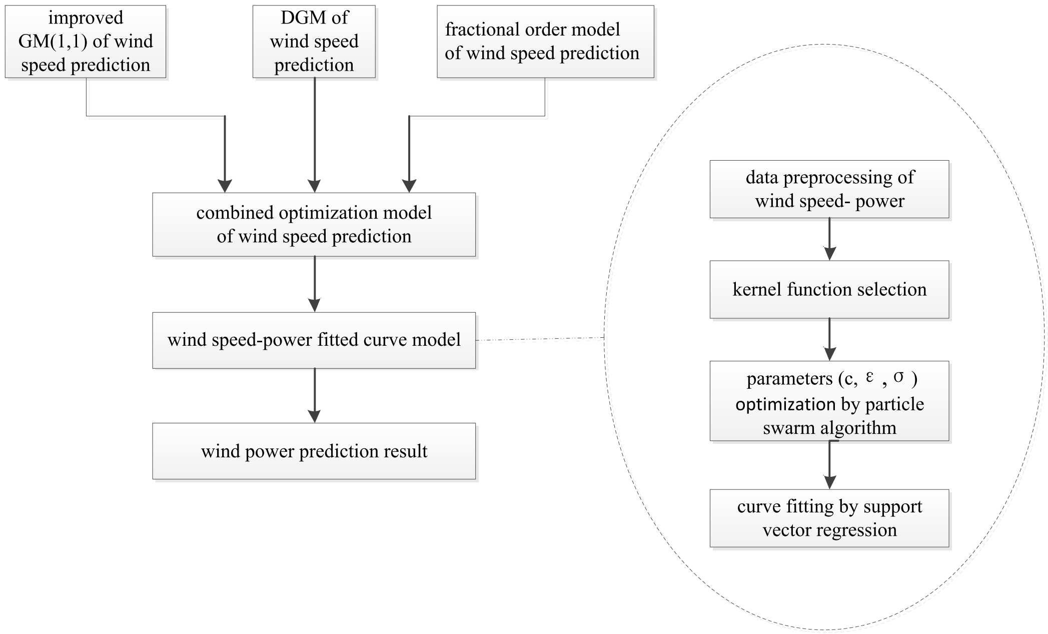


FIGURE 5. Wind power prediction model.

and part of the intrinsic information in the original time series is lost. Therefore, the improved grey model, discrete grey model and fractional order grey model are also designed. Secondly, in order to effectively integrate each single prediction model, a combined optimization model of wind speed prediction is constructed by means of neural network with multi-hidden layers. The original intrinsic information of wind speed series is fully utilized and the prediction accuracy and universality are improved. Finally, a support vector regression model based on particle swarm optimization is established. According to the actual wind speed-power scatter data, the wind speed-power relationship curve is fitted, and then the wind power prediction value is obtained.

IV. RESULTS AND ANALYSIS

A. WIND SPEED-WIND POWER FITTED CURVE

The power characteristic curve of wind turbine can reflect the power generation performance under different wind speed conditions. At present, there are two kinds of power characteristic curves: theoretical power curve and actual operation curve. The theoretical wind speed-power characteristic can be expressed as $P=\frac{1}{2}C_pA\rho v^3$, where P is wind turbine output power, C_p is wind energy utilization coefficient, $A = \pi R^2$ is the area swept by the wind turbine, R is hub radius, ρ is air density, v is wind speed. The power curve is obtained by simulation under ideal condition. However, the wind turbine power is often affected by turbulence intensity, wind shear and other factors in the environment.

The measured power curve is obtained from wind speed-power data recorded in SCADA. However, in practical application, the part of the measured scatter points are disorderly and highly dispersed.

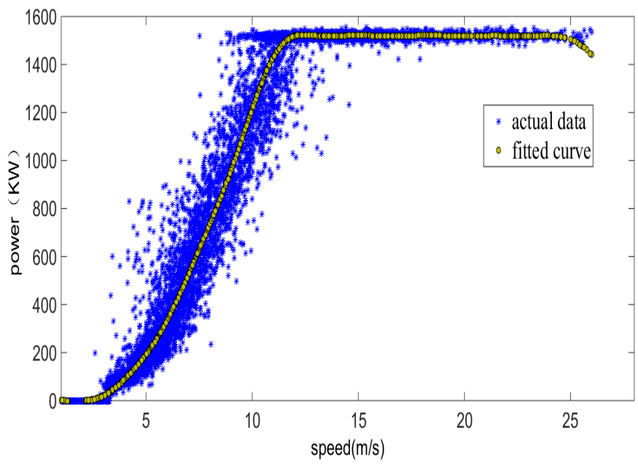


FIGURE 6. Scatter points and fitted curve of wind speed-power.

These abnormal data has a great impact on the power characteristic analysis. Therefore, the scatter data need to be preprocessed and optimally fitted in order to eliminate the impact of these outlier data.

The support vector regression model optimized by particle swarm algorithm is used to fit the wind speed-power curve as shown in figure 6. The optimized parameters of the model are as follows: penalty factor $C = 32$, precision $\varepsilon = 0.1$ and variance $\sigma = 2$ of RBF kernel function.

The blue scatter points in figure 6 reflect the actual operation status of wind turbine, and the green curve is the fitted result of the proposed model. The following conclusion can be obtained: the cut-in wind speed of wind turbine is 3 m/s, the rated wind speed is 12.5 m/s and the cut-out wind speed is 25 m/s. The fitted curve can perfectly represent the wind speed-power operation characteristic.

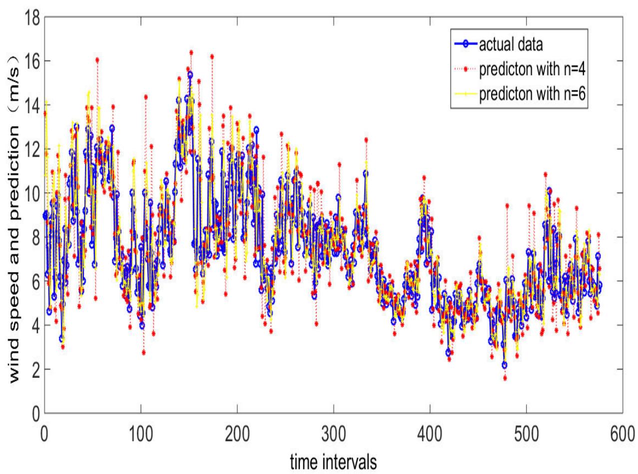


FIGURE 7. Wind speed prediction result of GM(1,1) with various sample number.

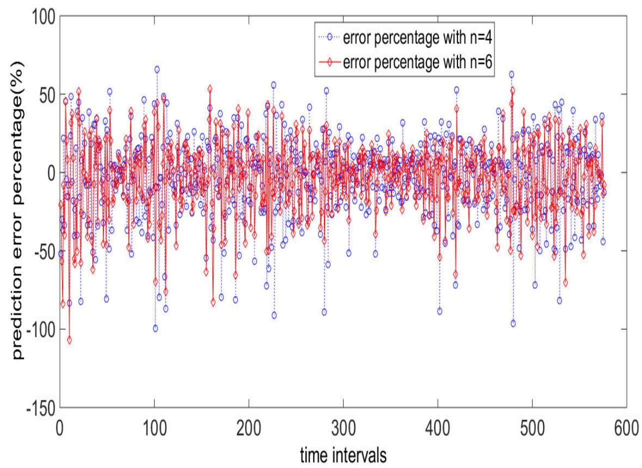


FIGURE 8. Wind speed prediction error of GM(1,1) with various sample number.

B. WIND SPEED PREDICTION BASED ON GREY MODEL

The evaluation indexes of model prediction performance are as follows: mean absolute error(MAE):

$$MAE = \frac{1}{n}\left(\sum_{i=1}^{n}\left(\left|P^{(0)}(i) - P^{(0)^{\wedge}}(i)\right|\right)\right) \tag{32}$$

mean absolute percentage error(MAPE):

$$MAPE = \frac{1}{n}\left(\sum_{i=1}^{n}\left(\left|\frac{P^{(0)}(i) - P^{(0)^{\wedge}}(i)}{P^{(0)}(i)}\right|\right)\right) \tag{33}$$

root mean square error(RMSE):

$$RMSE = \sqrt{\frac{1}{n}\left(\sum_{i=1}^{n}\left(P^{(0)}(i) - P^{(0)^{\wedge}}(i)\right)^2\right)} \tag{34}$$

The prediction result of GM (1,1) is shown in figure 7. When the number of points used to construct the grey model is equal to 4, the prediction error is obvious. The prediction overshoot occurs at the violent fluctuation of the original data. The prediction performance is remarkably improved when the number of data points is set as six.

The prediction error percentage of GM (1,1) is shown in figure 8. When the value of n is 6, the situation of large prediction error is obviously reduced compared with the case where n equals 4. However, the prediction accuracy of grey system model is not high, and large error still occurs at some prediction points.

GM (1,1) is suitable for small samples and poor information. The less data is required to build the model, usually no less than four data points. Experiments show that different sample number has an impact on the prediction performance. When n is set from 4 to 8, the prediction error is presented in table 1. The GM(1,1) with n as 6 is optimal due to 17.5% for the MAPE, 1.27m/s for the MAE and 1.69m/s for the RMSE.

TABLE 1. wind speed prediction error of GM(1,1) with various sample number.

	n=4	n=5	n=6	n=7	n=8
MAE(m/s)	1.46	1.36	1.27	1.28	1.30
MAPE	20.1%	18.7%	17.5%	17.8%	18.1%
RMSE(m/s)	1.93	1.78	1.69	1.71	1.73

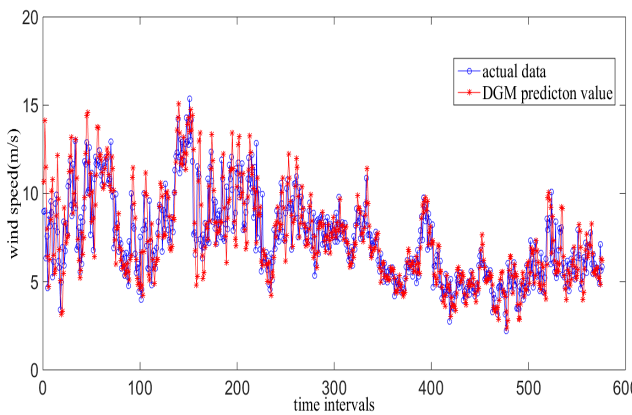


FIGURE 9. Wind speed prediction result of DGM.

C. WIND SPEED PREDICTION BASED ON DGM

The prediction result of DGM is shown in figure 9. It can be concluded from the prediction curve that the prediction value of this model is very close to that of GM (1,1), with 17.2% for the MAPE.

The prediction results of DGM and GM (1,1) are basically consistent, because the two models are equivalent under the condition of small development coefficient.

From formulas (10) and (11), the expression can be obtained as follows:

$$V^{(1)}(k) = \frac{1-0.5a}{1+0.5a}V^{(1)}(k-1) + \frac{b}{1+0.5a} \quad k = 2, \cdots, n \tag{35}$$

Considering DGM expression (17), it can be concluded that:

$$\alpha_1 = \frac{1-0.5a}{1+0.5a}\alpha_2 = \frac{b}{1+0.5a}\frac{b}{a} = \frac{\alpha_2}{1-\alpha_1} \tag{36}$$

The prediction expression of GM (1,1) is transformed as follows:

$$V^{(1)^\wedge}(k) = V^{(1)}(1)e^{-a(k-1)} + \frac{b}{a}\left(1-e^{-a(k-1)}\right) \quad k=1,2,\cdots \tag{37}$$

Considering formula (20), e^{-a} and $\alpha_1 = \frac{1-0.5a}{1+0.5a}$ are expanded by power series:

$$\begin{aligned} e^{-a} &= 1-a+\frac{a^2}{2!}-\frac{a^3}{3!}+\cdots+(-1)^n\frac{a^n}{n!} \\ \alpha_1 &= \frac{1-0.5a}{1+0.5a} = 1-a+\frac{a^2}{2}-\frac{a^3}{2^2}+\cdots+(-1)^{n+1}\frac{a^{n+1}}{2^n} \end{aligned} \tag{38}$$

If only the first three items are retained, the expression is established as follows: $e^{-a} = \alpha_1$. When the value of a is very small, the following expression is established: $e^{-a} \approx \alpha_1$. In this case, the DGM is equivalent to the GM (1,1).

D. WIND SPEED PREDICTION BASED ON IMPROVED GREY MODEL

1) EFFECT OF INITIAL CONDITION ON GREY MODEL

It is necessary to determine an initial condition for grey system model. The initial condition is generally assumed as: $V^{(1)^\wedge}(1) = V^{(1)}(1)$. Thus, the fitted curve passes through the first data point. However, according to the principle of least square method, the fitted curve does not necessarily pass through the first data point.

Considering that the first data point is the oldest data, it is not closely related to the future change. It is not generated by accumulation operation, and its regularity is not high. If the newer data point is used as initial condition to solve the problem, i.e. $V^{(1)^\wedge}(m) = V^{(1)}(m)$ $(m = 2, 3, \cdots, n)$, prediction expression of GM (1,1) is changed as follows:

$$V^{(1)^\wedge}(k) = \left(V^{(1)}(m) - \frac{b}{a}\right)e^{-a(k-m)} + \frac{b}{a} \quad k=1,2,\cdots \tag{39}$$

For some specific applications, GM (1,1) is very sensitive to the selection of initial condition. Different initial conditions have a great impact on the prediction result and accuracy. The reason is that the original data sequence is very close to the exponential growth. The values of data points vary greatly, which can reach two orders of magnitude.

However, due to the randomness and uncertainty of wind speed, the original data is not an ideal exponential growth sequence. Therefore, this experiment demonstrates that with the change of initial condition, the prediction results of GM (1,1) are basically the same. In the case of predicting the wind speed, the GM (1,1) is not sensitive to the initial condition, and the initial condition has very weak influence on the prediction result.

2) GREY MODEL WITH BACKGROUND VALUE OPTIMIZATION

The development coefficient a and grey action b are important parameters of GM (1,1). Background value is needed to solve these parameters by least square method. Optimizing the construction method of background value is an important means to improve the accuracy and applicability of grey model.

The definite integral of the expression (12) in [k-1,k] interval is as follows:

$$\begin{aligned} &\int_{k-1}^{k} \frac{dV^{(1)}}{dt}dt + a\int_{k-1}^{k} V^{(1)}dt = b \\ &V^{(1)}(k) - V^{(1)}(k-1) + a\int_{k-1}^{k} V^{(1)}dt = b \\ &V^{(0)}(k) + a\int_{k-1}^{k} V^{(1)}dt = b \end{aligned} \tag{40}$$

The background value obtained by comparing expression (11) with (40) is essentially as follows:

$$Z^{(1)}(k) = \int_{k-1}^{k} V^{(1)}dt \tag{41}$$

The expression (10) for background value is actually constructed according to formula (41) by means of trapezoidal method. The general expression for constructing background value is as follows:

$$Z^{(1)}(k) = \alpha V^{(1)}(k) + (1-\alpha)V^{(1)}(k-1) \quad (0 \le \alpha \le 1) \tag{42}$$

Thus, five methods for constructing background value are proposed as follows:

$$Z^{(1)}(k) = \left\{\begin{array}{l} V^{(1)}(k-1); \\ \frac{3}{4}V^{(1)}(k-1) + \frac{1}{4}V^{(1)}(k); \\ \frac{1}{2}V^{(1)}(k-1) + \frac{1}{2}V^{(1)}(k); \\ \frac{1}{4}V^{(1)}(k-1) + \frac{3}{4}V^{(1)}(k); \\ V^{(1)}(k) \end{array}\right\} \tag{43}$$

According to the above expressions, the constructed models are labeled as $GM(1,1)^a$, $GM(1,1)^b$, $GM(1,1)^c$, $GM(1,1)^d$, $GM(1,1)^e$, respectively. The corresponding prediction results are shown in figure 10.

From figure 10, $GM(1,1)^e$ shows the best prediction effect and accuracy among the five models. Correspondingly, the prediction error percentage is shown in figure 11. The number of large prediction error for $GM(1,1)^e$ model is significantly reduced.

The background value of $GM(1,1)^e$ is set as the right endpoint value, i.e. $Z^{(1)}(k) = V^{(1)}(k)$, which is introduced

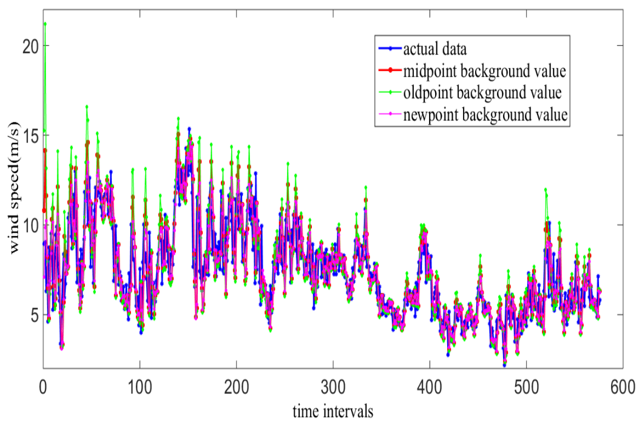


FIGURE 10. Wind speed prediction result of GM(1,1) with various background values.

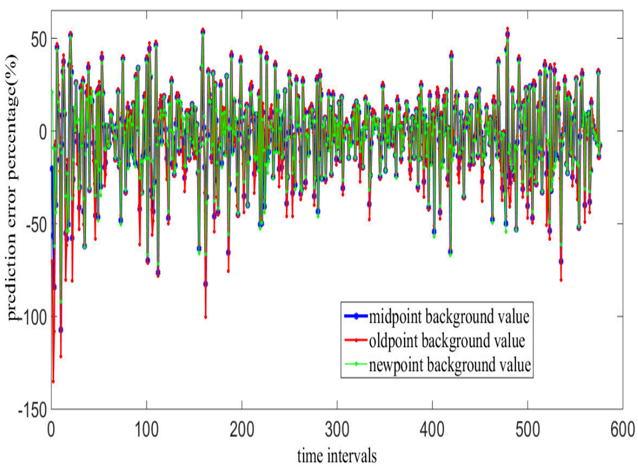


FIGURE 11. Wind speed prediction error of GM(1,1) with various background values.

into formula (11):

$$V^{(1)}(k) = \frac{1}{1+a}V^{(1)}(k-1) + \frac{b}{1+a} \quad k = 2, \cdots, n \tag{44}$$

Considering the formulas (17) and (44), the following expression is obtained:

$$\alpha_1 = \frac{1}{1+a}\alpha_2 = \frac{b}{1+a} \tag{45}$$

α_1 can be expanded by power series:

$$\alpha_1 = \frac{1}{1+a} = 1 - a + a^2 - a^3 + a^4 + \cdots \tag{46}$$

Comparing it with formula (38), it is concluded that the prediction result greatly varies with the different methods for constructing background value.

The prediction performance of model based on background value optimization is shown in Table 2. GM(1,1)^a (taking oldest value as background value) shows the worst performance, while GM(1,1)^e (taking newest value as background value) shows the best performance. Compared with GM(1,1)^c , i.e. the traditional grey model, the prediction performance

TABLE 2. Wind speed prediction error of GM(1,1) with various background values.

	GM(1,1)^a	GM(1,1)^b	GM(1,1)^c	GM(1,1)^d	GM(1,1)^e
MAE(m/s)	1.43	1.35	1.27	1.21	1.16
MAPE	19.4%	18.5%	17.5%	16.8%	16.2%
RMSE(m/s)	1.96	1.81	1.69	1.61	1.53

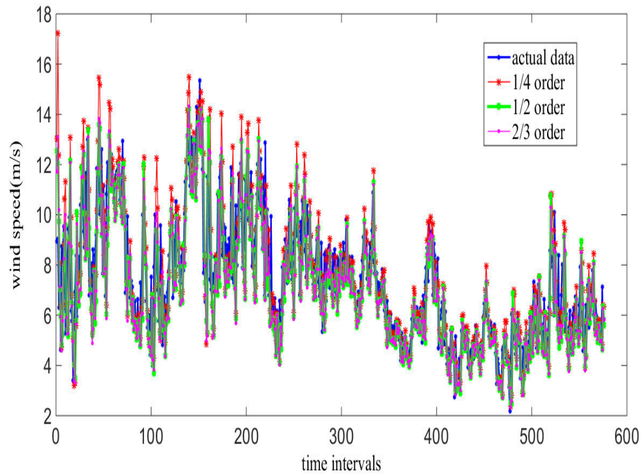


FIGURE 12. Wind speed prediction result of fractional order DGM.

of GM(1,1)^e is improved by 7.4% (from 17.5% to 16.2%)for MAPE.

E. WIND SPEED PREDICTION BASED ON FRACTIONAL ORDER GREY MODEL

The wind speed has strong fluctuation and inherent uncertainty. In order to extract the inherent law and information of wind speed change, the fractional order grey models with different orders are established to improve the model adaptability and reduce the prediction disturbance bound. The result of fractional order grey model for wind speed prediction is shown in figure 12. The rule and trend of wind speed change can be reflected by these different models. However, for some prediction points, these models present complementary characteristic.

The error of fractional order model for wind speed prediction is shown in figure 13. The overall performance of the three fractional order models is basically the same, but the error distribution is not uniform. For some local prediction points, the single fractional order model shows higher prediction accuracy and smaller prediction error. However, for other local prediction points, the model presents lower prediction accuracy and larger prediction error.

F. WIND SPEED AND POWER PREDICTION BASED ON COMBINATION MODEL

When different grey models work independently, the optimal prediction ratio of each model is shown in figure 14. Obviously, no model has a great advantage in prediction performance. These models are very close to each other in

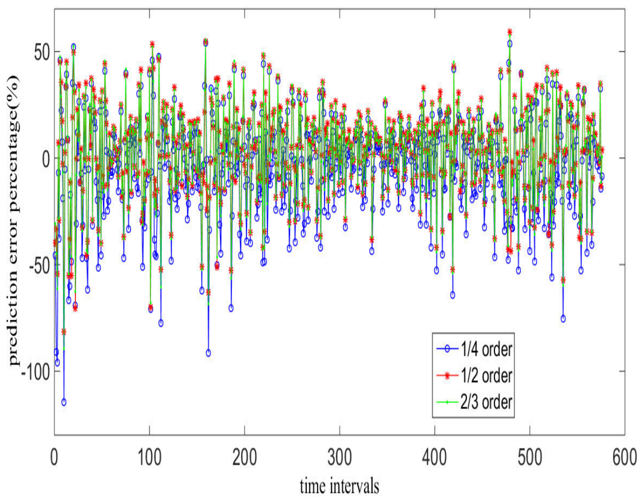


FIGURE 13. Wind speed prediction error of fractional order DGM.

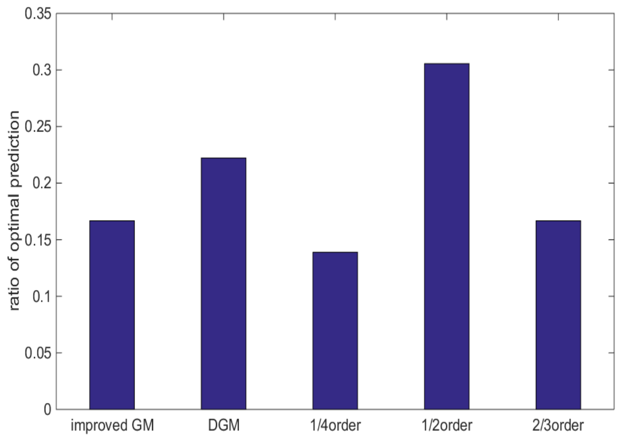


FIGURE 14. Ratio of optimal prediction points for different models.

terms of the optimal prediction points. The single grey model is not comprehensive in revealing the intrinsic characteristic of wind speed change. Therefore, in order to effectively integrate these models, the combined optimization prediction model based on neural network is constructed. The combination model can improve the prediction performance, and fully exploit the advantages of each model [45]–[47]. In addition, the NWP outputs are introduced into the combination model in order to further reduce the prediction fuzziness. The resolution of ECMWF(European Centre for Medium-Range Weather Forecasts) mode is 15km. The resolution of GRAPES-Meso(mesoscale of the Global and Regional of Assimilation and Prediction System) mode is 10km.

The structure of the neural network is determined as three hidden layers by trial and error method. The numbers of neurons in the hidden layers are respectively (6, 9, 7).

In order to objectively show the prediction performance of the proposed model, ARIMA(Auto-regressive Integrated Moving Average) model is also constructed as a baseline method. ARIMA model is composed of auto-regressive model and moving average model. It is a typical method

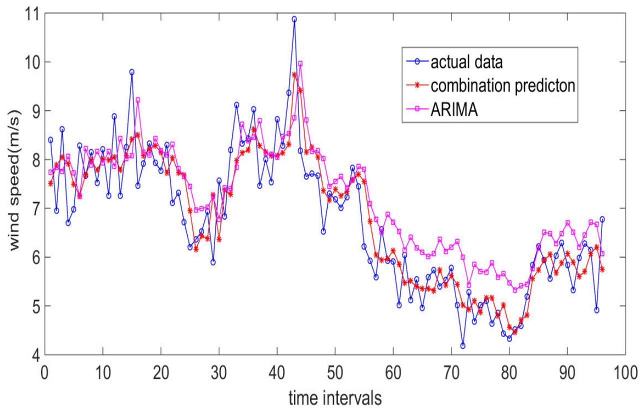


FIGURE 15. Wind speed prediction result of combination model.

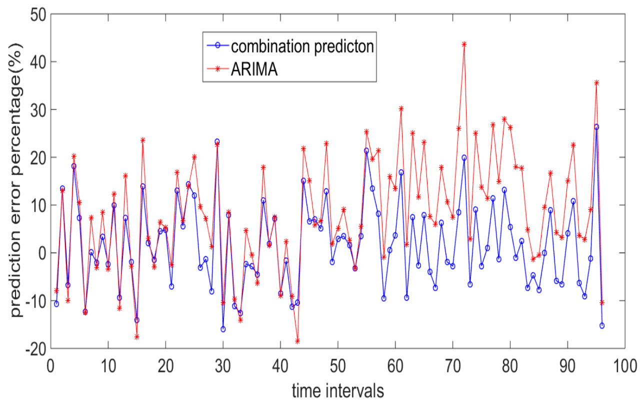


FIGURE 16. Wind speed prediction error of combination model.

in time series analysis. The time-varying data sequence is regarded as a random sequence, which is represented by a certain mathematical model. The model can predict future values from past and present values of time series [48]–[51].

The order and parameters of ARIMA model are determined by partial auto-correlation function method and long auto-regression method. The identified model, i.e. ARIMA(3,1,2), is described as follows:

$$\hat{v}(t) = 0.436v(t-1) + 0.636v(t-2) - 0.203v(t-3) \\ +0.0631\varepsilon(t-1) - 0.412\varepsilon(t-2)$$

The result of the grey combination model for wind speed prediction is shown in figure 15. The prediction value can well track the change of wind speed. When the wind speed dramatically changes, compared with ARIMA model, there is no serious prediction overshoot by means of the combination model. In terms of tracking ability, the combination model is superior to ARIMA.

The error of the grey combination model for wind speed prediction is shown in figure 16. Compared with the ARIMA model, the prediction error of the combination model is significantly reduced. The proportion of relative error less than 10% is significantly increased, and the proportion of relative error over 20% is greatly reduced.

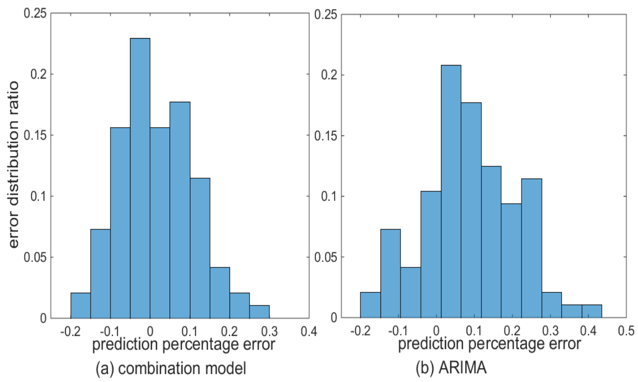


FIGURE 17. Error distribution histogram for wind speed prediction.

The error distribution characteristic of the combination model and ARIMA for wind speed prediction is shown in the figure 17. The prediction error of the combination model is mostly concentrated in the low error interval, and the corresponding ratio of the low error is higher than that of the ARIMA. The error distribution of ARIMA is relatively wide, and there are still some prediction points with the error higher than 30%.

The error of different models for wind speed prediction is shown in table 3. The combination model presents the optimal prediction performance. Compared with the grey model optimized by background value, the combination model results in the improvement by 53.1% for MAPE (from 16.2% to 7.6%). The prediction performance of ARIMA is better than that of any single grey model. However, the combination model is superior to ARIMA, with the improvement by 37.2% for MAPE (from 12.1% to 7.6%).

The result of the grey combination model for wind power prediction is shown in figure 18. Compared with the ARIMA model, the proposed combination model can well present the law and trend of actual power fluctuation without prediction overshoot. ARIMA model cannot well reflect the fluctuation of power in different intervals of wind speed.

The error of grey combination model for wind power prediction is shown in figure 19.

Compared with the ARIMA model, the prediction error of the combination model is significantly reduced. Relative prediction errors are mostly within 10%. However, ARIMA presents the large prediction error, and the error of some points is higher than 30%.

The error distribution characteristic of the combination model and ARIMA for power prediction is shown in the figure 20. The prediction error of the combination model is mostly concentrated in the low error interval, and the corresponding ratio of the low error is higher than that of the ARIMA. The error distribution of ARIMA is relatively wide, and the ratio of high error is slightly higher than that of the low error. There are still some prediction points with the error approaching 40%.

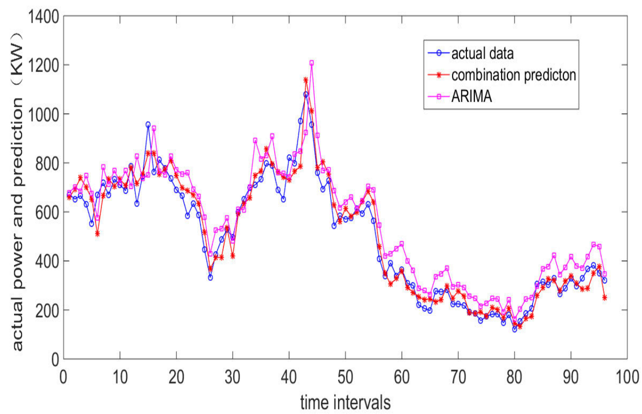


FIGURE 18. Wind power prediction result of combination model.

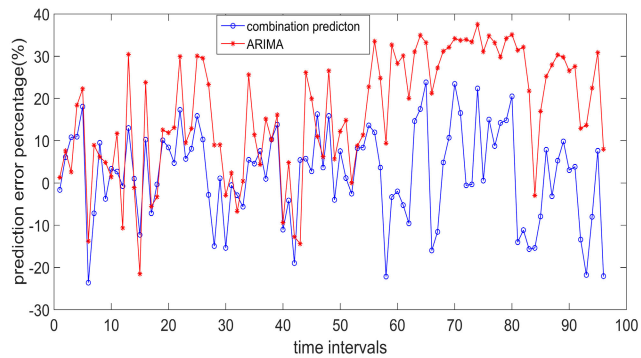


FIGURE 19. Wind power prediction error of combination model.

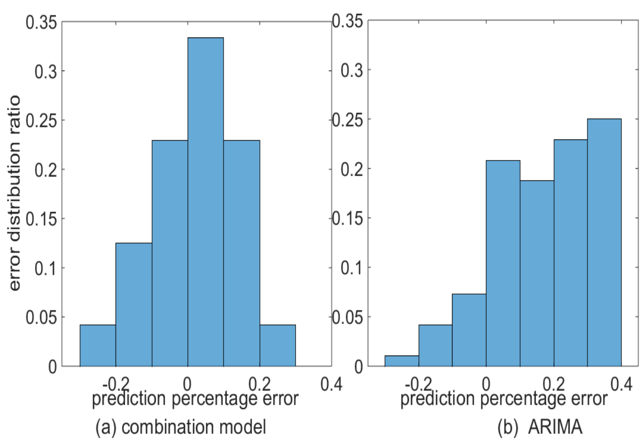


FIGURE 20. Error distribution histogram for power prediction.

TABLE 3. Wind speed prediction error of different models.

	improved GM	DGM	1/2order	1/4order	2/3order	combination	ARIMA
MAE(m/s)	1.16	1.26	1.27	1.25	1.28	0.51	0.76
MAPE	16.2%	17.3%	17.1%	16.7%	17.2%	7.6%	12.1%
RMSE(m/s)	1.53	1.69	1.67	1.70	1.68	0.64	0.91

TABLE 4. Wind power prediction error of different models.

	improved GM	DGM	1/2order	1/4order	2/3order	combination	ARIMA
MAE(kw)	76.9	89.4	106.6	101.2	104.6	43.4	69.7
MAPE	15.7%	18.2%	21.1%	19.7%	20.8%	9.7%	14.9%
RMSE(kw)	109.4	129.6	143.7	142.3	139.7	54.2	82.6

The power prediction error of different models is shown in table 4. The combination model shows the best prediction performance. Compared with the grey model optimized by background value, the combination model results in the improvement by 38.2% for MAPE (from 15.7% to 9.7%). The prediction performance of ARIMA is better than that of any single grey model. However, the combination model is superior to ARIMA, with the improvement by 34.9% for MAPE (from 14.9% to 9.7%).

V. FURTHER DISCUSSIONS

In order to understand how the proposed method reduces model uncertainty and prediction error, the following section will further discuss the advantage of the combination model and how the prediction model is effectively integrated.

Although the traditional grey model is suitable for small sample and poor information system, it is discovered that the number of data points adopted in the modeling process has an impact on the model performance. The intuitive idea is that using the more data points could lead to the better prediction performance. However, the experiment results show that along with the increasing number of selected data points, the prediction error decreases at the beginning, and then changes little with the slight increase. The reason may be that the prediction performance is improved by the increase of known information. The change of wind speed is sometimes strong and sudden, which causes the decrease of the correlation among the large sample data. Meanwhile, the construction method of background value has great influence on the grey model. Compared with the traditional construction method, the model using the latest data as background value shows better performance. This is because the latest data in time series can more accurately reflect the law and trend of system change. But the same idea does not apply to the selection of initial condition. The performance of the grey model is insensitive to the initial condition, including the latest data as initial condition. In fact, this just illustrates the uncertainty and fuzziness of the wind speed series. The sequence generated by one order accumulation is not ideally close to the exponential law.

The prediction performance indexes of different fractional order models are generally close to each other. However, by observing and comparing these prediction sequences, it is revealed that there is complementary characteristic in terms of abrupt change points and prediction delay. A single model can not comprehensively indicate inherent law of the sequence. The combination model can well solve this problem. The number of hidden layers and neurons of the neural network demonstrates the representation ability of the system. But too many layers and neurons may lead to overfitting. These parameters need to be determined by experiments.

The wind speed-power curve shows obvious non-linear characteristic and greatly varies in different range of wind speed. Support vector regression (SVR) model adopts the kernel function mapping mechanism, and the allowable error limit and penalty term are all adjustable. Therefore the SVR model can well identify the power characteristic in the whole range of wind speed and eliminate the abnormal data points. How to select these parameters directly affects the accuracy of curve fitting. Hence, it is necessary to introduce optimization algorithm for parameter search and model correction. Particle swarm optimization (PSO) can search the global optimal value in parallel and quickly converge with simple computation. Thus, the output result is further modified, along with the enhancement of the model applicability, by integrating PSO into the SVR model.

VI. CONCLUSION

The stochastic fluctuation of wind speed and wind power is inherently uncertain. The traditional grey model has the disadvantage of large prediction error. The experiment result shows that the number of points used to construct the grey model directly affects the prediction performance. When the number, i.e. n, equals 6, the GM (1,1) is optimal. GM (1,1) is not sensitive to the selection of initial condition, but the construction method of background value has a significant impact on the prediction result. Compared with the traditional method of selecting midpoint value as background value, the improved GM (1,1) with right endpoint value as background value has significantly enhanced the wind speed prediction performance. For MAE, MAPE and RMSE, the prediction performance is increased by 8.7%, 7.4% and 9.5% ,respectively.

The prediction performance of discrete grey model and GM (1,1) is basically same. In order to reveal the inherent law and change trend of wind speed more comprehensively, the fractional order grey models with different orders are designed for wind speed prediction. The prediction accuracy of each fractional order model is similar, but the prediction result presents obvious complementary characteristic among these models.

The support vector regression model is constructed to fit the actual scatter points of wind speed and wind power. Particle swarm algorithm is used to optimize the model parameters, such as penalty factor C, precision ε and variance σ of kernel function. The fitted curve can perfectly reveal the relationship between wind speed and power.

In order to overcome the shortcoming of single prediction model, the combination prediction model based on neural network is established. For further reducing uncertainty, the NWP outputs of ECMWF and GRAPES-MESO are introduced into the combination prediction model. The structure parameters of the neural network are optimized by trial and error method. Compared with the single grey system model with the lowest prediction error, the prediction performance of proposed combination model is significantly improved. For MAE, MAPE and RMSE, the wind speed prediction performance is increased by 56.0%, 53.1% and 58.2% respectively, and the wind power prediction performance is increased by 43.6%, 38.2% and 50.5% respectively.

In order to fairly compare the performance of the proposed model , ARIMA(3,1,2) prediction model is established as a baseline method. Compared with the ARIMA(3,1,2) method, the MAE, MAPE and RMSE of the proposed model for wind speed prediction are improved by 32.9%, 37.2% and 29.7% respectively, and the MAE, MAPE and RMSE of the proposed model for wind power prediction are improved by 37.7%, 34.9% and 34.4% respectively.

REFERENCES

[1] Z. Qian, Y. Pei, H. Zareipour, and N. Chen, "A review and discussion of decomposition-based hybrid models for wind energy forecasting applications," *Appl. Energy*, vol. 235, pp. 939–953, Feb. 2019.

[2] P. Du, J. Z. Wang, W. D. Yang, and T. Niu, "A novel hybrid model for short-term wind power forecasting," *Appl. Soft. Comput.*, vol. 80, pp. 93–106, Jul. 2019.

[3] N. Safari, S. M. Mazhari, and C. Y. Chung, "Very short-term wind power prediction interval framework via bi-level optimization and novel convex cost function," *IEEE Trans. Power Syst.*, vol. 34, no. 2, pp. 1289–1300, Mar. 2019.

[4] C.-M. Huang, C.-J. Kuo, and Y.-C. Huang, "Short-term wind power forecasting and uncertainty analysis using a hybrid intelligent method," *IET Renew. Power Gener.*, vol. 11, no. 5, pp. 678–687, 2017.

[5] R. Ak, V. Vitelli, and E. Zio, "An interval-valued neural network approach for uncertainty quantification in short-term wind speed prediction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 11, pp. 2787–2800, Nov. 2015.

[6] D. Lee and R. Baldick, "Short-term wind power ensemble prediction based on Gaussian processes and neural networks," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 501–510, Jan. 2014.

[7] I. Okumus and A. Dinler, "Current status of wind energy forecasting and a hybrid method for hourly predictions," *Energy Convers. Manage.*, vol. 123, pp. 362–371, Sep. 2016.

[8] A. U. Haque, M. H. Nehrir, and P. Mandal, "A hybrid intelligent model for deterministic and quantile regression approach for probabilistic wind power forecasting," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1663–1672, Jul. 2014.

[9] E. T. Renani, M. F. M. Elias, and N. A. Rahim, "Using data-driven approach for wind power prediction: A comparative study," *Energy Convers. Manage.*, vol. 118, pp. 193–203, Jun. 2016.

[10] A. Khosravi, L. Machado, and R. O. Nunes, "Time-series prediction of wind speed using machine learning algorithms: A case study Osorio wind farm, Brazil," *Appl. Energy*, vol. 224, pp. 550–566, Aug. 2018.

[11] A. Tascikaraoglu and M. Uzunoglu, "A review of combined approaches for prediction of short-term wind speed and power," *Renew. Sustain. Energy Rev.*, vol. 34, no. 6, pp. 243–254, 2014.

[12] J. Jung and R. P. Broadwater, "Current status and future advances for wind speed and power forecasting," *Renew. Sustain. Energy Rev.*, vol. 31, no. 2, pp. 762–777, 2014.

[13] J. Wang, J. Hu, K. Ma, and Y. Zhang, "A self-adaptive Hybrid approach for wind speed forecasting," *Renew. Energy*, vol. 78, pp. 374–385, Jun. 2015.

[14] R. Ak, O. Fink, and E. Zio, "Two machine learning approaches for short-term wind speed time-series prediction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 8, pp. 1734–1747, Aug. 2016.

[15] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1033–1044, May 2014.

[16] J. Heinermann and O. Kramer., "Machine learning ensembles for wind power prediction," *Renew. Energy*, vol. 89, pp. 671–679, Apr. 2016.

[17] J. Hu, J. Heng, J. Tang, and M. Guo, "Research and application of a hybrid model based on Meta learning strategy for wind power deterministic and probabilistic forecasting," *Energy Convers. Manage.*, vol. 173, pp. 197–209, Oct. 2018.

[18] O. Karakuş, E. E. Kuruoğlu, and M. A. Altinkaya, "One-day ahead wind speed/power prediction based on polynomial autoregressive model," *IET Renew. Power Gener.*, vol. 11, no. 11, pp. 1430–1439, 2017.

[19] A. Kavousi-Fard, A. Khosravi, and S. Nahavandi, "A new fuzzy-based combined prediction interval for wind power forecasting," *IEEE Trans. Power Syst.*, vol. 31, no. 1, pp. 18–26, Jan. 2016.

[20] G. Zhang, H.-X. Li, and M. Gan, "Design a wind speed prediction model using probabilistic fuzzy system," *IEEE Trans. Ind. Informat.*, vol. 8, no. 4, pp. 819–827, Nov. 2012.

[21] A. E. Saleh, M. S. Moustafa, K. M. Abo-Al-Ez, and A. A. Abdullah, "A hybrid neuro-fuzzy power prediction system for wind energy generation," *Int. J. Elect. Power Energy Syst.*, vol. 74, pp. 384–395, Jan. 2016.

[22] M. Morshedizadeh, M. Kordestani, R. Carriveau, D. S.-K. Ting, and M. Saif, "Application of imputation techniques and adaptive neuro-fuzzy inference system to predict wind turbine power production," *Energy*, vol. 138, pp. 394–404, Nov. 2017.

[23] K. Wang, X. Qi, H. Liu, and J. Song, "Deep belief network based *K*-means cluster approach for short-term wind power forecasting," *Energy*, vol. 165, pp. 840–852, Dec. 2018.

[24] 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 Convers. Manage.*, vol. 134, pp. 361–372, Feb. 2017.

[25] Z. Shi, H. Liang, and V. Dinavahi, "Direct interval forecast of uncertain wind power based on recurrent neural networks," *IEEE Trans. Sustain. Energy*, vol. 9, no. 3, pp. 1177–1187, Jul. 2018.

[26] E. E. Elattar, "Prediction of wind power based on evolutionary optimised local general regression neural network," *IET Gener., Transmiss. Distrib.*, vol. 8, no. 5, pp. 916–923, Nov. 2014.

[27] J. Zhao, Z.-H. Guo, Z.-Y. Su, Z.-Y. Zhao, X. Xiao, and F. Liu, "An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed," *Appl. Energy*, vol. 162, pp. 808–826, Jan. 2016.

[28] S. Salcedo-Sanz, Á. M. Pérez-Bellido, E. G. Ortiz-García, A. Portilla-Figueras, L. Prieto, and D. Paredes, "Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term wind speed prediction," *Renew. Energy*, vol. 34, pp. 1451–1457, Jun. 2009.

[29] S. Salcedo-Sanz, Á. M. Pérez-Bellido, E. G. Ortiz-García, A. Portilla-Figueras, L. Prieto, and F. Correoso, "Accurate short-term wind speed prediction by exploiting diversity in input data using banks of artificial neural networks," *Neurocomputing*, vol. 72, pp. 1336–1341, Jan. 2009.

[30] A. J. Deppe, W. A. Gallus, Jr., and E. S. Takle, "A WRF ensemble for improved wind speed forecasts at turbine height," *Weather Forecasting*, vol. 28, pp. 212–228, Feb. 2013.

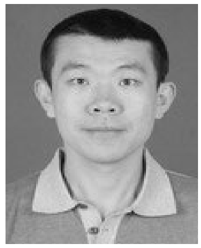
[31] F. Cassola and M. Burlando, "Wind speed and wind energy forecast through Kalman filtering of numerical weather prediction model output," *Appl. Energy*, vol. 99, pp. 154–166, Nov. 2012.

[32] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and M. Prabhat, "Deep learning and process understanding for data-driven Earth system science," *Nature*, vol. 566, pp. 195–204, Feb. 2019.

[33] W. T. Lin, J. Z. Wang, and W. Y. Zhang, "Program of wind speed prediction based on numerical simulation with intelligent optimization algorithm," *Climatic Environ. Res.*, vol. 17, no. 5, pp. 646–658, 2012.

[34] J. Shi, Z. Ding, W.-J. Lee, Y. Yang, Y. Liu, and M. Zhang, "Hybrid forecasting model for very-short term wind power forecasting based on grey relational analysis and wind speed distribution features," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 521–525, Jan. 2014.

[35] Y. Li, J. Zhang, J. Xiao, and Y. Tan, "Short-term prediction of the output power of PV system based on improved grey prediction model," in *Proc. Int. Conf. Adv. Mechatronic Syst.*, Kumamoto, Japan, Aug. 2014, pp. 547–551.
[36] Z. Z. Zhang, J. Zou, and G. Zheng, "Ultra-short term wind power prediction model based on modified grey model method for power control in wind farm," *Wind Eng.*, vol. 35, no. 1, pp. 55–68, 2011.
[37] F. Tagliaferri, I. M. Viola, and R. G. J. Flay, "Wind direction forecasting with artificial neural networks and support vector machines," *Ocean Eng.*, vol. 97, no. 15, pp. 65–73, 2015.
[38] X. Yuan, Q. Tan, X. Lei, Y. Yuan, and X. Wu, "Wind power prediction using hybrid autoregressive fractionally integrated moving average and least square support vector machine," *Energy*, vol. 129, pp. 122–137, Jun. 2017.
[39] J. Zeng and W. Qiao, "Short-term wind power prediction using a wavelet support vector machine," *IEEE Trans. Sustain. Energy*, vol. 3, no. 2, pp. 255–264, Apr. 2012.
[40] N. Amjady, F. Keynia, and H. Zareipour, "Wind power prediction by a new forecast engine composed of modified hybrid neural network and enhanced particle swarm optimization," *IEEE Trans. Sustain. Energy*, vol. 2, no. 3, pp. 265–276, Jul. 2011.
[41] X. Yang, Z. Fang, Y. Yang, D. Mba, and L. Xiaochuan, "A novel multi-information fusion grey model and its application in wear trend prediction of wind turbines," *Appl. Math. Model.*, vol. 71, pp. 543–557, Jul. 2019.
[42] H. Quan, D. Srinivasan, and A. Khosravi, "Short-term load and wind power forecasting using neural network-based prediction intervals," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 2, pp. 303–315, Feb. 2014.
[43] A. Sharifian, M. J. Ghadi, S. Ghavidel, L. Li, and J. F. Zhang, "A new method based on type-2 fuzzy neural network for accurate wind power forecasting under uncertain data," *Renew. Energy*, vol. 120, pp. 220–230, May 2018.
[44] G. Sideratos and N. D. Hatziargyriou, "Probabilistic wind power forecasting using radial basis function neural networks," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 1788–1796, Nov. 2012.
[45] M. C. Mabel and E. Fernandez, "Analysis of wind power generation and prediction using ANN: A case study.," *Renew. Energy*, vol. 33, pp. 986–992, May 2008.
[46] L. Zjavka, "Wind speed forecast correction models using polynomial neural networks," *Renew. Energy*, vol. 83, pp. 998–1006, Nov. 2015.
[47] Z.-H. Guo, J. Wu, H.-Y. Lu, and J.-Z. Wang, "A case study on a hybrid wind speed forecasting method using BP neural network," *Knowl.-Based Syst.*, vol. 24, no. 7, pp. 1048–1056, 2011.
[48] O. A. Maatallah, A. Achuthan, K. Janoyan, and P. Marzocca, "Recursive wind speed forecasting based on Hammerstein auto-regressive model," *Appl. Energy*, vol. 145, pp. 191–197, May 2015.
[49] B.-M. Hodge, A. Zeiler, D. Brooks, G. Blau, J. Pekny, and G. Reklatis, "Improved wind power forecasting with ARIMA models," *Comput. Aided Chem. Eng.*, vol. 29, pp. 1789–1793, Jan. 2011.
[50] P. Chen, T. Pedersen, B. Bak-Jensen, and Z. Chen, "ARIMA-based time series model of stochastic wind power generation," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 667–676, May 2010.
[51] M. B. Shams, S. Haji, A. Salman, H. Abdali, and A. Alsaffar, "Time series analysis of Bahrain's first hybrid renewable energy system," *Energy*, vol. 103, pp. 1–15, May 2016.



YI ZHANG received the master's degree from the Hebei University of Technology, China, in 2005, where he is currently pursuing the Ph.D. degree in control theory and control engineering. He is also a Lecturer with the School of Information Science and Engineering, Hebei University of Science and Technology. His research interests include wind power prediction and the application of machine learning in the wind energy.

HEXU SUN (M'90–SM'96) received the Ph.D. degree in automation from Northeastern University, China, in 1993. He is currently a Professor with the School of Artificial Intelligence, Hebei University of Technology, and also a Professor with the Hebei University of Science and Technology. His current research interests include wind energy and complex engineering systems.

YINGJUN GUO received the master's degree from the Beijing Institute of Technology, China, in 2004. He is currently pursuing the Ph.D. degree in control theory and control engineering with the Hebei University of Technology. He is also an Associate Professor with the Hebei University of Science and Technology. His research interests include wind power control and power electronic devices.