



Multi-objective interval prediction of wind power based on conditional copula function

Gang ZHANG¹ , Zhixuan LI¹, Kaoshe ZHANG¹, Lei ZHANG², Xia HUA³, Yongqing WANG⁴



Abstract Interval prediction of wind power, which features the upper and lower limits of wind power at a given confidence level, plays a significant role in accurate prediction and stability of the power grid integrated with wind power. However, the conventional methods of interval prediction are commonly based on a hypothetical probability distribution function, which neglects the correlations among various variables, leading to decreased prediction accuracy. Therefore, in this paper, we improve the multi-objective interval prediction based on the conditional

copula function, through which we can fully utilize the correlations among variables to improve prediction accuracy without an assumed probability distribution function. We use the multi-objective optimization method of non-dominated sorting genetic algorithm-II (NSGA-II) to obtain the optimal solution set. The particular best solution is weighted by the prediction interval average width (PIAW) and prediction interval coverage probability (PICP) to pick the optimized solution in practical examples. Finally, we apply the proposed method to three wind power plants in different Chinese cities as examples for validation and obtain higher prediction accuracy compared with other methods, i.e., relevance vector machine (RVM), artificial neural network (ANN), and particle swarm optimization kernel extreme learning machine (PSO-KELM). These results demonstrate the superiority and practicability of this method in interval prediction of wind power.

CrossCheck date: 27 November 2018

Received: 18 May 2018 / Accepted: 27 November 2018 / Published online: 12 February 2019
© The Author(s) 2019

✉ Gang ZHANG
zhanggang3463003@163.com

Zhixuan LI
15594818128@163.com

Kaoshe ZHANG
zhangks@263.net

Lei ZHANG
zhanglzei@163.com

Xia HUA
kevinxhua@163.com

Yongqing WANG
3087984821@qq.com

Keywords Wind power prediction, Interval prediction, Conditional copula function, Empirical distribution function, Multi-objective optimization algorithm

1 Introduction

China is rich in wind resources [1, 2] and has adopted a national policy to vigorously develop wind power. However, wind power has strong randomness and volatility and is difficult to accurately predict [3–8], thus making wind power prediction a current hotspot in research. Most current studies of wind power prediction focus more heavily on point prediction [9–14] and less on interval prediction [15]. Because the point prediction approach causes the wind power prediction results to greatly deviate from the actual values, it is difficult to implement real-time

¹ Institute of Water Resources and Hydro-Electric Engineering, Xi'an University of Technology, Xi'an 710048, China

² State Grid Shaanxi Baoji Electric Power Company, Baoji 721000, China

³ State Grid Gansu Electric Power Company, Gansu Electric Power Research Institute, Lanzhou 730050, China

⁴ State Grid Shaanxi Electric Power Company, Shaanxi Electric Power Research Institute, Xi'an 710054, China

scheduling [16]. Interval prediction can supply more accurate information for dispatching operations [17], and consequently, the interval prediction method has become increasingly important. The beta distribution calculation model [18] and its improved model [19] obtain interval prediction results under different confidence levels from the point prediction error distribution. The non-parametric kernel density estimation method fits the probability distribution curve of wind power through interpolation and finally obtains the prediction interval that satisfies the preset confidence level [20]. Based on the standardized Gaussian distribution, Kou Peng et al used a modified Gaussian model to realize interval prediction of wind power [21]. The interval prediction results can also be obtained via establishment of the probability distribution function of wind power prediction based on the empirical distribution model and a non-parametric regression technique [22]. Additionally, to improve the interval coverage and reduce the interval width, Fan Lei et al [23] used the method of ensemble empirical mode decomposition and correlation vector machine to realize short-term wind power interval prediction.

The development of heuristic algorithms has injected new vitality into solution of nonlinear problems. Khosravi A et al [24] proposed a method for predicting the upper and lower bounds of wind power based on the neural network method. Quan H et al [25] proposed a lower upper bound estimation (LUBE) model based on the single-layer forward neural network structure. This model uses particle swarm optimization (PSO) to obtain a satisfactory wind power prediction interval, but it is slow. Yang Xiyun et al [26] proposed the particle swarm optimization kernel extreme learning machine (PSO-KELM) model, which has higher prediction accuracy and a faster speed of operation and can supply decision supports for wind power interval prediction and secure and stable operation of wind power on the grid.

Although the above research has shown useful exploration of the interval prediction of wind power, two types of problems still remain:

- 1) Certain current prediction methods must assume that wind power obeys a certain distribution function [18–21] and fail to consider the correlation between adjacent periods of wind power series, which reduces the prediction accuracy of wind power to a certain extent;
- 2) Each method has different parameters, and varying prediction results can be obtained if the parameters are different. Prediction interval coverage probability (PICP) and prediction interval average width (PIAW) are usually selected for accuracy evaluation [16, 26], but the evaluation results of these two indicators have

often been inconsistent (when the effect of PICP is preferable, PIAW has poorer evaluation results and vice versa), leading to an inability to obtain the optimal parameters and prediction results.

Because the copula function can more comprehensively and accurately describe the correlation relationship among multiple variables, prediction research conducted based on this observation has become the focus of many scholars. Selected scholars have attempted to solve the wind power prediction problem using the conditional copula function [27], and although they reported certain achievements, they did not find the boundary problem of the model parameters, could not obtain a group of optimal model parameter values and could not guarantee the prediction accuracy. All of the above problems have made the copula prediction method difficult to apply in reality.

To address these problems, this paper perfected the wind power prediction method based on the conditional copula function and achieved a combination of this method with multi-objective optimization algorithms. This method takes advantage of the copula function to make full use of the correlation existing in the adjacent wind power sequences and improve the prediction accuracy. To this end, we use the non-dominated sorting genetic algorithm-II (NSGA-II) multi-objective optimization algorithm to seek the optimal solution set of the parameters for the copula prediction model, and the optimal parameters and corresponding prediction results are obtained via evaluation, which has a strong practical value. To verify the effectiveness of the method, we take three wind farms in China as research objects and apply the proposed method to ultra-short term wind power prediction. After comparing with the common interval prediction methods, the predicting accuracy and effect of the proposed method are verified.

2 Conditional copula function

The copula function, also known as the connection function, is a type of function that connects multiple random variable joint distribution functions with their respective marginal distribution functions [28–31]. In using the copula function for wind power interval prediction, the copula function in its general form (Gaussian copula function, *t*-copula function and others) always reveals differences from the actual situation. At the same time, use of the continuous copula function form in calculation is relatively difficult because it requires the operation of obtaining the inverse function in calculation of the confidence interval. Therefore, this paper uses the conditional copula function in discrete form to make wind power predictions.



The construction method of the discrete condition copula function [32–35] is given as follows. It is assumed that N independent samples for $t + 1$ variables are known, as shown in (1), where each row represents one sample, and each column represents the value of the marginal distribution function of one variable under different samples; $X_1, X_2, \dots, X_t, X_{t+1}$ is a value in the real number domain; and $F_1, F_2, \dots, F_t, F_{t+1}$ are respectively the marginal distribution functions of variable 1 to variable $t + 1$.

$$[X_1, X_2, \dots, X_t, X_{t+1}] = \begin{bmatrix} F_1(X_1^1) & \dots & F_t(X_t^1) & F_{t+1}(X_{t+1}^1) \\ F_1(X_1^2) & \dots & F_t(X_t^2) & F_{t+1}(X_{t+1}^2) \\ \vdots & & \vdots & \vdots \\ F_1(X_1^N) & \dots & F_t(X_t^N) & F_{t+1}(X_{t+1}^N) \end{bmatrix} \quad (1)$$

The value of each element in (1) is between 0 and 1, and we divide the interval $[0, 1]$ into K equal sub-intervals. Thus, we have K subintervals for each variable, and the $t + 1$ variables together form the K^{t+1} subspaces.

The conditional copula function refers to the probability distribution function of the $(t + 1)$ th variable under the condition that the former t variable values of a certain sample F are known, i.e., the probability distribution function corresponding to $F_{t+1}(x_{t+1}^F)$ under the condition that (2) is known.

$$F_1(x_1^F), F_2(x_2^F), \dots, F_t(x_t^F) \quad (2)$$

In the N samples of (1), the samples in the former t variables that fall in the same subspace as (2) are removed to form a conditional matrix. Suppose there are N_1 samples in the matrix. The samples for which all of the elements in the $(t + 1)$ th column fall in the same subinterval are classified into the same class, and it is assumed that a J class is obtained. The numbers of samples included in each class are recorded respectively as M_1, M_2, \dots, M_J . Finally, an identical value $F_{t+1}^j (j = 1, 2, \dots, J)$ is calculated using (3) to represent the $(t + 1)$ th element in each class, and the probability of occurrence of each class is given.

$$F_{t+1}^j = \frac{1}{M_j} \sum_{i=1}^{M_j} F_{t+1}(x_{t+1}^i) \quad p^j = M_j/N_1 \quad (3)$$

Therefore, $[F_{t+1}^j, p^j]_{j=1, 2, \dots, J}$ is the discrete probability distribution function of $F_{t+1}(x_{t+1}^F)$ based on (2), which is the discrete conditional copula function.

3 Multi-objective prediction method based on conditional copula function

The wind power prediction model proposed in this paper includes two main components. One component is the conditional copula interval prediction method based on the discrete form, which uses wind power historical data after giving a set of prediction model parameters (K, t), to establish the probability distribution function of the wind power to be predicted by calculating the marginal distribution function such that the upper and lower bounds of the prediction interval under the given confidence level are obtained, and the interval prediction effect is measured by two evaluation indices containing PICP and PIAW. The second component is the NSGA-II multi-objective optimization method that uses PICP and PIAW as the objective functions within the value range of the copula interval prediction model parameters K and t and finds the non-inferior solution set for predicting the parameters of K and t . It is assumed that each evaluation index is equally important such that each index is endowed with 50% weights, and a set of optimal prediction model parameters are obtained by weighting calculation. The copula interval prediction is performed with the recently obtained parameter, and the results are compared with the prediction results of the currently used interval prediction methods.

To understand the detailed process of this prediction method better, we give the specific implementation steps of the prediction method in Fig. 1.

Step 1: Using the historical wind power sequence of the wind farm, in the process of establishing conditional copula distribution function in discrete form, we first determine the value range of K (interval division number) and t (condition number) that is applicable to the current historical data.

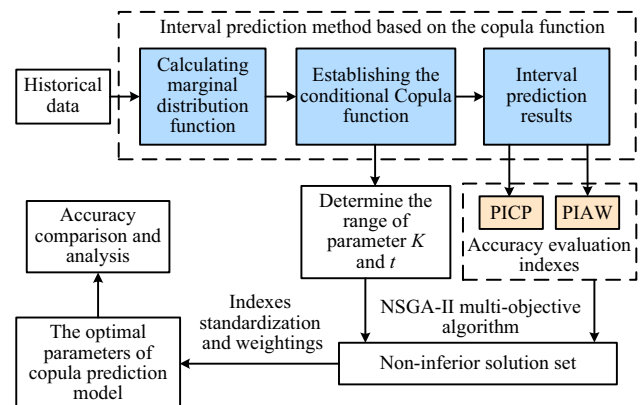


Fig. 1 Process diagram of multi-objective prediction method based on conditional copula function

Step 2: We use two indicators, PIAW and PICP, to evaluate the accuracy of the conditional copula interval prediction results under different K and t values.

Step 3: Taking PICP and PIAW as objective functions, we use the NSGA-II multi-objective optimization method to find the non-inferior solution set of the conditional copula prediction model parameters within the range of K, t .

Step 4: We obtain a set of optimal parameters for the prediction model using the average weight method.

Step 5: We use the optimal group of parameters of K, t for conditional copula interval prediction and compare the prediction results with those of relevance vector machine (RVM), artificial neural network (ANN) and PSO-KELM (three commonly used interval prediction algorithm results) to verify the effectiveness of the proposed method.

3.1 Wind power interval prediction

A relatively strong correlation relationship exists between wind power sequences in short time intervals (1 hour) [29]. Therefore, we can take the continuous wind power sequences collected from certain wind farm as historical data and use the copula function to describe this correlation between wind power sequences. We assume that the continuous wind power sequence of a certain wind farm is $[y_1, y_2, \dots, y_{n-1}, y_n]$, and we need to predict the interval of wind power y_{n+1} at the next moment. In using the discrete conditional copula function to make wind power prediction, assuming that the condition number is t , the $t + 1$ variables and $n - t$ samples can be transformed from the historical data:

$$\begin{bmatrix} y_1 & \cdots & y_t & y_{t+1} \\ y_2 & \cdots & y_{t+1} & y_{t+2} \\ \vdots & & \vdots & \vdots \\ y_{n-t+1} & \cdots & y_{n-1} & y_n \end{bmatrix} \quad (4)$$

Each column in the matrix is a wind power sequence at certain time intervals, which is viewed as $[X_1, X_2, \dots, X_t, X_{t+1}]$. Thus, we establish the conditional copula distribution function of wind power to be predicted. In other words, we use the marginal distribution function value $F_1(x_1^F), F_2(x_2^F), \dots, F_t(x_t^F)$ of the wind power $[y_{n-t+1}, \dots, y_{n-1}, y_n]$ as a known condition and obtain the probability distribution function of y_{n+1} .

We assume that the conditional copula function obtained is $[F_{t+1}^j, P^j]_{j=1, 2, \dots, J}$, where p^1 is the largest and p^J is the smallest. Because all values of the marginal distribution functions fall into the interval $[0, 1]$, we divide the interval into K subintervals, which are denoted as S_1, S_2, \dots, S_K in turn. Thus p^j is accumulated from $j = 1$ until the summation is greater than or equal to the preset confidence level β .

We assume that when accumulating to $j = q, \sum_{j=1}^{j=q} p^j \geq \beta$, and additionally, the subinterval corresponding to $F_{t+1}^1, F_{t+1}^2, \dots, F_{t+1}^q$ is given as follows:

$$\begin{cases} F_{t+1}^1 \in (S_1^L, S_1^u] \\ F_{t+1}^2 \in (S_2^L, S_2^u] \\ \vdots \\ F_{t+1}^q \in (S_q^L, S_q^u] \end{cases} \quad (5)$$

where superscript L and u refer to the lower and upper bound of the interval, respectively. Under the confidence level β , the prediction interval of the wind power at the $(t + 1)$ th moment is the union of (5), namely:

$$(S_1^L, S_1^u] \cup (S_2^L, S_2^u] \cup \dots \cup (S_q^L, S_q^u] = (S^L, S^u] \quad (6)$$

Using the inverse operation of the marginal distribution function, we obtain the prediction interval $F_{t+1}^{-1}(S^L) \leq W_{t+1} \leq F_{t+1}^{-1}(S^u)$ of the wind power at the $(t + 1)$ th moment, and the wind power interval prediction at $(t + 1)$ th moment is realized.

3.2 Interval prediction evaluation index

We adopt PICP and PIAW as the evaluation indices to validate the prediction accuracy of the proposed algorithm. PICP indicates the number of actual values of wind power that fall within the prediction interval. Based on satisfaction of the confidence level, the larger the PICP value, the larger the number of actual wind power values that will fall into the prediction interval and the better the prediction effect will be. PIAW represents the average width between the upper and lower boundaries of the interval. If the same confidence level is satisfied, the smaller the value, the narrower the prediction interval will be, which means that the prediction interval is closer to the actual value. The calculation formulas of PICP and PIAW are written as follows:

1) PICP

$$PICP = \frac{1}{U} \sum_{u=1}^U A_u \times 100\% \quad (7)$$

where U is the total amount of wind power to be predicted; $u = 1, 2, \dots, U$; and A_u is the characteristic function defined as follows:

$$A_u = \begin{cases} 1 & V_u \in [V_u, \overline{V}_u] \\ 0 & V_u \notin [V_u, \overline{V}_u] \end{cases} \quad (8)$$



where \overline{V}_u and \underline{V}_u are respectively the upper and lower bounds of the prediction interval; V_u is the actual value of wind power at the moment to be predicted.

If the actual value of wind power to be predicted falls into the prediction interval, the value of A_u is 1. Otherwise, A_u is 0.

2) PIAW

$$PIAW = \frac{1}{U} \sum_{u=1}^U (\overline{V}_u - \underline{V}_u) \quad (9)$$

3.3 NSGA-II multi-objective optimization algorithm

The NSGA algorithm is derived from the genetic algorithm (GA) but is different from the traditional GA. NSGA has a strong advantage in the field of multi-objective optimization but also contains many defects. Aimed at addressing the defects of the NSGA method, Deb et al proposed the NSGA-II method in 2002. This version is better than the NSGA, and it can improve the operation speed and system robustness [36].

In this paper, according to [37], two target functions of PIAW and PICP are used, and NSGA-II is applied for multi-objective optimization to obtain a non-inferior solution set of the conditional copula prediction model.

4 Example analysis

4.1 Example data source

To verify the effectiveness of the proposed method in this paper, we apply the multi-objective prediction method of the conditional copula function proposed in Section 3 to three wind farms, which are recorded as wind farms 1–3.

Wind farm 1 is the Huanglongjiecumiao wind power plant, which is located in Yanan, Shaanxi Province of China, and the total installed capacity is 49.5 MW. In establishing the interval prediction model based on the copula function, we use the wind power data from the Huanglongjiecumiao wind power plant from June to November in 2016 as the modelling data, and the time interval of the data is 1 hour. We take the data from 72 points (three days) in December as the validation data. Wind farm 2 is the wind farm referred to as Liuhe under the jurisdiction of the China Huaneng Group Co., Ltd. in Nanjing, Jiangsu Province. The total installed capacity of the project is 49.5 MW, and a total of 25 sets of wind turbines are installed. We use the wind power data from the Liuhe wind power plant from March to October in 2017 as the modelling data, and the time interval of the data is 1 hour. We take the data from 72 points (three days) in

November as the validation data and analyse the interval prediction effectiveness of the proposed method. Wind farm 3 is the wind farm Baitiancifeng located in Jingbian, Yulin, Shaanxi Province. The total installed capacity is 49.5 MW. We use the wind power data of the Baitiancifeng wind power plant from April to August in 2016 as the modelling data, and the time interval of the data is 1 hour. We take the data from 72 points (three days) in September as the validation data and analyse the interval prediction effectiveness of the proposed method.

4.2 NSGA-II multi-objective optimization

In this section, we use only wind farm 1 as an example to specifically describe the application process for the method proposed in this paper. The steps of wind farms 2–3 are the same as those of wind farm 1, and only the final result is given.

In establishing the discrete conditional copula function, t and K must be determined. The confidence level β is a free parameter that can reflect the prediction accuracy, and it is unified as $\beta = 0.9$ in this paper. The values of parameter K and t are limited by the sample number of the historical wind power sequence. Under the condition that K and t are too large because of too many subspaces composed of wind power sequences, there is no sample, which is the same as t wind power subspaces in t consecutive moments before the point to be predicted in (4). Thus, the corresponding conditional matrix cannot be formed, which leads to failure of the conditional copula function construction. In this example, we take the historical sample as the condition and take the existence of conditional matrix of each points in the prediction period as the principle. We programme in MATLAB and cyclically take the value of the model parameters to determine the value range of K and t . Finally, we obtain the following range: while $t = 1$, the range of K is from 2 to 408, while $t = 2$, the range of K is from 2 to 11, while $t = 3$, the range of K is from 2 to 5.

In the value range of model parameters K and t , the two evaluation indices of PICP and PIAW are taken as the objective function for calculation using the NSGA-II multi-objective optimization method to find the Pareto optimal solution of the conditional copula prediction model parameters. The obtained Pareto optimal solution set is shown in Fig. 2.

It can be observed that there are 13 optimal points in the Pareto optimal solution set. All of the points in the solution set have relatively better prediction performance. The PICP values of points 1, 2, and 3 are large, and their PIAW is also wide. In the remaining points, the PIAW changes are not much different. Point 4 can be viewed as the sudden point of the entire Pareto optimal solution set curve, which

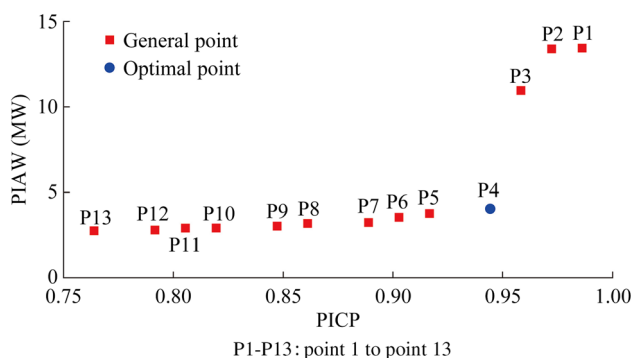


Fig. 2 Pareto optimal solution set

can also be said to be the inflection point. In the figure, we mark the general Pareto optimal solution with red dots, and the inflection point is marked with a blue circle.

In the Pareto optimal solution set, the larger the PICP point, the larger the PIAW will be. The PIAW and PICP values are contradictory to each other, and thus the solution sets cannot be compared with each other to seek the better choice. Therefore, all solution sets are feasible solutions. To obtain the optimal parameter settings, we first use the extreme value method to standardize the PIAW and PICP values of each group of parameters. According to the definition in Section 3.2, the smaller the PIAW value, the better the prediction performance will be, and therefore, we use (10) to standardize the index PIAW. Additionally, the larger the PICP value, the better the prediction performance will be, and we adopt (11) to standardize the index PICP, where x_i is the value of point i to be normalized, y_i is the normalized value of point i .

$$y_i^{PIAW} = \frac{x_1^{PIAW} - x_i^{PIAW}}{x_1^{PIAW} - x_{13}^{PIAW}} \quad 0 \leq i \leq 13 \quad (10)$$

$$y_i^{PICP} = \frac{x_i^{PICP} - x_{13}^{PICP}}{x_1^{PICP} - x_{13}^{PICP}} \quad 0 \leq i \leq 13 \quad (11)$$

Equations (10) and (11) are used to standardize the PIAW and PICP value of the 13 points. We consider that the two precision indices are equally important, and the weights of PICP and PIAW can all be considered as 50%. Finally, we conduct the weight calculation for the two evaluation indices of all points in the Pareto optimal solutions set, and the results are shown in Table 1.

From Table 1, we note that after standardization of the two interval prediction indices of each point, from point 1 to 13, the PIAW value gradually increases, and the PICP value gradually decreases. In other words, the PICP value of point 1 is the best, and the PIAW value is the worst. Otherwise, the PICP value of point 13 is the worst, and the PIAW value is the best. Additionally, after taking 50% as the weight of the two indices and summing them, the result of point 4 is 0.8465, which is the maximum value in the 13

Table 1 Weight calculation results of the points in Pareto optimal solutions set

Point	PICP	PIAW	Standardized PIAW	Standardized PICP	Weight calculation
P1($t = 3, K = 4$)	0.986	13.442	0	1	0.5000
P2($t = 1, K = 4$)	0.972	13.396	0.004	0.937	0.4705
P3($t = 1, K = 5$)	0.958	10.957	0.233	0.875	0.554
P4($t = 1, K = 51$)	0.944	4.032	0.881	0.812	0.8465
P5($t = 1, K = 140$)	0.917	3.766	0.906	0.688	0.797
P6($t = 1, K = 124$)	0.903	3.545	0.926	0.625	0.7755
P7($t = 1, K = 230$)	0.889	3.239	0.955	0.563	0.759
P8($t = 1, K = 198$)	0.861	3.185	0.960	0.437	0.6985
P9($t = 1, K = 282$)	0.847	3.028	0.975	0.375	0.675
P10($t = 1, K = 362$)	0.819	2.918	0.985	0.250	0.6175
P11($t = 1, K = 396$)	0.806	2.913	0.985	0.188	0.5865
P12($t = 1, K = 397$)	0.792	2.801	0.996	0.125	0.5605
P13($t = 1, K = 399$)	0.764	2.758	1	0	0.5000

points. This observation shows that the interval prediction result is optimal in this parameter setting. In Fig. 2, it can be observed that the point is located at the inflection point of the approximate curve, which is marked with a blue circle, and the specific prediction model parameters are set to $K = 51, t = 1$.

Similarly, we apply this method to the other two wind farms. The optimal prediction model parameters for wind farm 2 are $K = 244, t = 1$, and the optimal prediction model parameters for wind farm 3 are $K = 200, t = 1$.

4.3 Result comparison for different algorithms

To validate the effectiveness of the proposed method, RVM [23], ANN [24], and PSO-KELM [26] are introduced to compare the results of interval prediction with those of the proposed method.

Figures 3, 4 and 5 respectively show the prediction results of applying the four methods in wind farms 1–3, and the red line and the green line form 72 little intervals, which the result of the interval prediction. If the black “×”



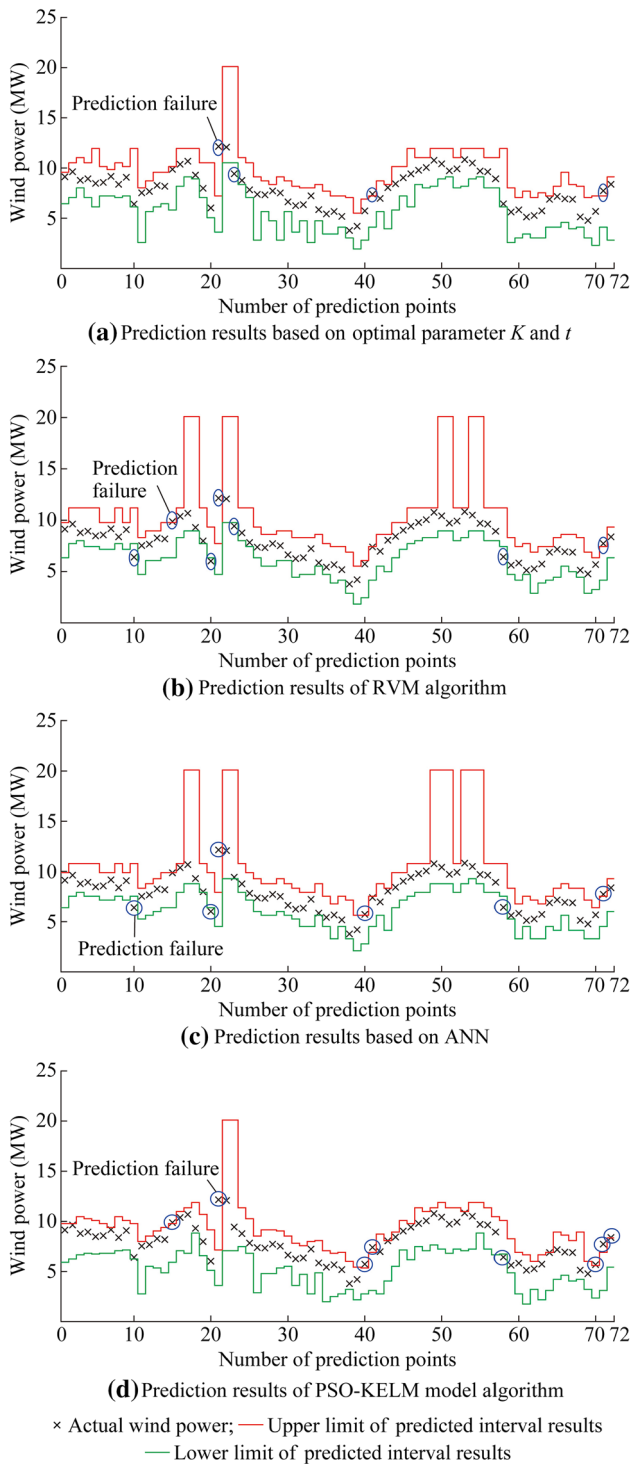


Fig. 3 Prediction results of applying four methods in wind farm 1

mark is within the red line and the green line, the point prediction can be considered successful. If the black “×” mark is beyond the red line, then the measured value is considered beyond the upper limit of the prediction. If the black “×” mark is lower than the green line, then the measured value is lower than the lower limit of the

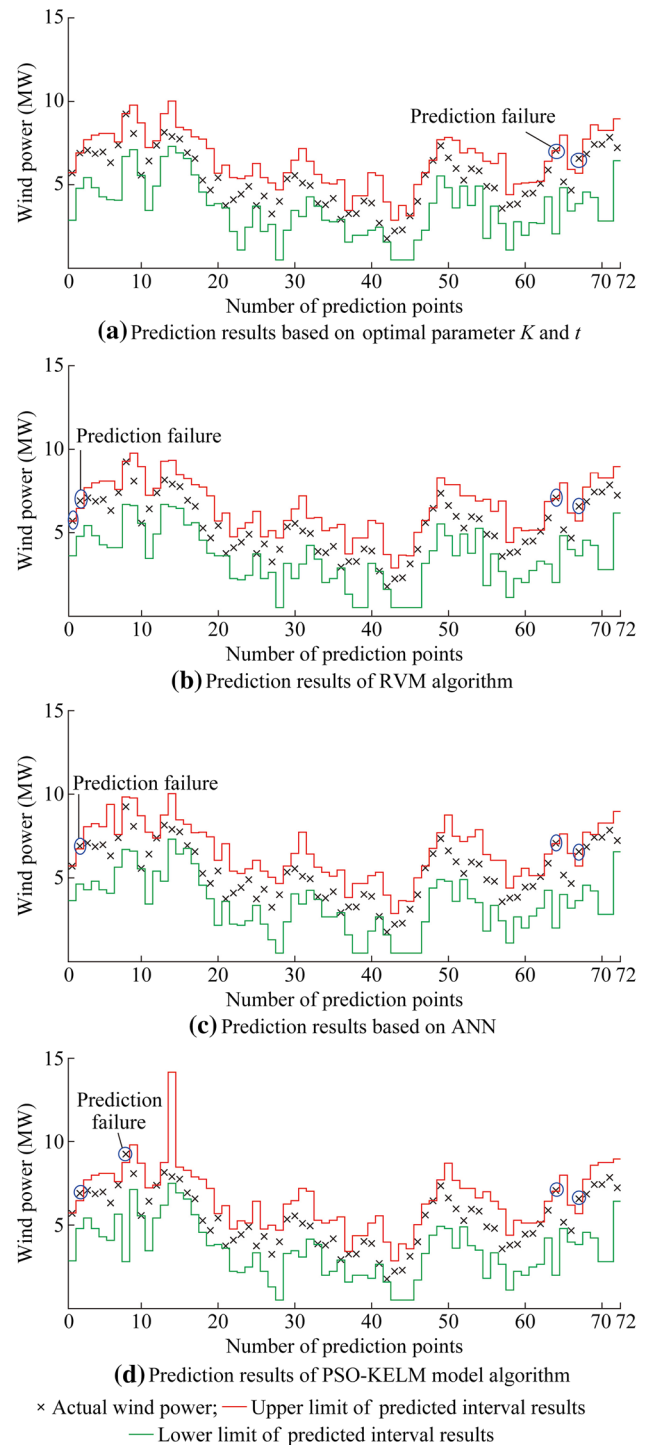


Fig. 4 Prediction results of applying four methods in wind farm 2

prediction. When the measured value exceeds the upper limit or lower limit, the point prediction is considered to be a failure.

It can be observed from Fig. 3a that the measured values of three points (21, 41, 71) exceed the upper limit of the prediction, the measured values of one point (23) exceed

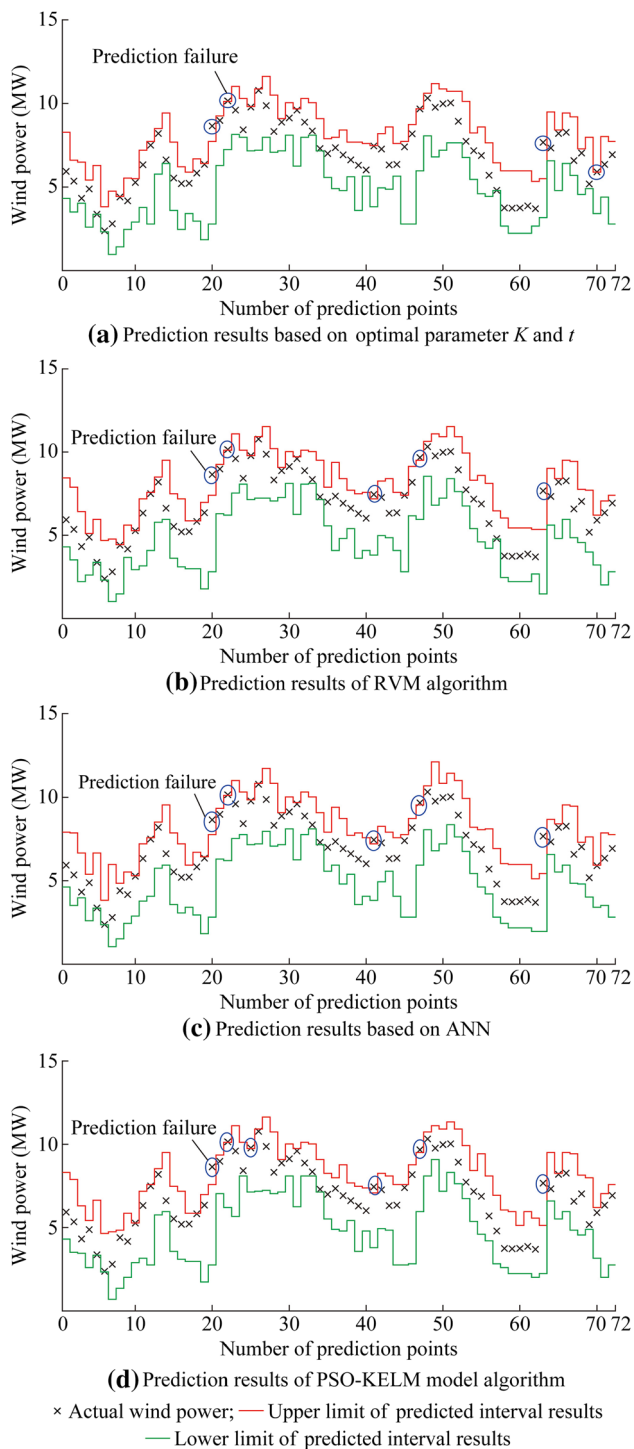


Fig. 5 Prediction results of applying four methods in wind farm 3

the lower limit of the prediction, and thus a total of four points do not meet the prediction requirements. In Fig. 3b, the measured values of three points (15, 21, 71) exceed the upper limit of the prediction, and the measured values of four points (10, 20, 23, 58) exceed the lower limit of the prediction, and thus a total of seven points do not meet the

prediction requirements. In Fig. 3c, the measured values of three points (21, 40, 71) exceed the upper limit of the prediction, and the measured values of three points (10, 20, 58) exceed the lower limit of the prediction, and thus a total of six points do not meet the prediction requirements. In Fig. 3d, the measured values of seven points (15, 21, 40, 41, 70, 71, 72) exceed the upper limit of the prediction, the measured values at the 58th point exceed the lower limit of the prediction, and thus a total of 8 points do not meet the prediction requirements. The prediction results obtained using the method of this paper has the highest number of points that meet the requirement. In Fig. 3, we note that all four prediction methods are failures at point 21, which might be due to a large sudden change of wind power for this time slot. However, the prediction range of all four methods at point 22 is too wide, which might be due to the continuous increase of wind power in the adjacent periods, thereby lifting the upper bound of the prediction interval. However, the prediction model reduces the error as soon as possible according to the actual situation, and it shows an upward small and protruding rectangle in the predicted image. However, from Fig. 3, when using the method of this paper to make predictions on wind farm 1, the figure shows only a small upward and protruding rectangle, whereas Fig. 3b and c has more than three, indicating that the method in this paper fully uses the correlation between wind power sequences and has a better prediction effect.

In Fig. 4a, only two points (64, 67) in the predicted segment do not meet the prediction requirements. In Fig. 4b, only four points (1, 2, 64, 67) in the predicted segment do not meet the prediction requirements. In Fig. 4c, only three points (2, 64, 67) in the predicted segment do not meet the prediction requirements. In Fig. 4d, only four points (2, 8, 64, 67) in the predicted segment do not meet the prediction requirements. It can be observed that the method of this paper is still the most reliable. However, the overall prediction effect of each method under wind farm 2 is better than that of wind farm 1, which might be due to the smooth wind power of wind farm 2 during the prediction period, which only fluctuates between 3 and 10 MW, and the rate of change is not large. However, at points 43–45, the lower bounds of the interval predictions given by each method are quite low, and the predictions are relatively conservative. This result might be due to the decrease in wind power at points 40–42, such that the prediction model expects that this downward trend in wind power will continue to exist, thereby reducing the predicted lower bound and increasing the success rate. This process shows a small downward and protruding rectangle in the predicted image. It can be observed from Fig. 4 that for prediction in wind farm 2 using the method of this paper, the number of small rectangles protruding downward is the least, indicating that the prediction effect of

method in this paper is better than that of other three methods.

In wind farm 3, the accuracy of the method proposed in this paper is still the highest, and four points (20, 22, 63, 70) in Fig. 5a do not meet the requirements. Five points (20, 22, 41, 47, 63) in Fig. 5b do not meet the requirements. Five points (20, 22, 41, 47, 63) in Fig. 5c do not meet the requirements. Six points (20, 22, 25, 41, 47, 63) in Fig. 5d do not meet the requirements. It can be observed that in addition to the method in this article, the remaining three methods have points of failed prediction in each day, especially for each night. We speculate that this result might be related to local meteorological factors during the night. At the 41st point, the wind power shows a downward trend at points 31–40, which results in prediction of failure for Fig. 5b, c and d. However, the method in this paper can make full use of the correlation between wind power and finds the “trend” of such a continuous decline is actually weakened. The prediction contains a certain margin for the prediction interval at the next moment, and that makes the prediction successful.

5 Discussion

After obtaining the prediction results of each method under three wind farms, for further analysis and discussion, we use (7) and (9) to separately calculate the prediction accuracy indicators of each method. The results are shown in Table 2.

It can be observed from Table 2 that in the three wind farms, the PICP values obtained by each method are all greater than 0.85, indicating that the prediction effects of these methods are relatively excellent. However, the PICP of the interval prediction method (M1) proposed in this paper is the largest, and at the same time, the corresponding PIAW is smallest. This observation illustrates that the proposed method in this paper is superior to the commonly

used interval prediction methods for which these three methods are representative.

The comparison shows that the prediction interval coverage of M3 is better than that of M2 and M4, but its interval width is too large. The interval average width of M2 is narrower, but the success rate of interval prediction is not high. In general, each of these three methods has advantages and disadvantages, and it is impossible to choose an optimal method. However, the optimal parameters selected in this paper are based on the same importance of PICP and PIAW, but at times it is necessary to emphasize security in the scheduling process, and thus the requirements for PICP are higher in the two evaluation indicators. Applying the method in this paper endows the PICP with a higher weight, but because of the need to consider high security, the prediction process tends to be conservative, resulting in an increase in the interval width. In contrast, at times we require a narrower prediction interval in the scheduling process. In this case, the requirements for PIAW are higher. The smaller the PIAW, the easier it is to meet the requirements such that the PIAW is endowed with a higher weight when applying this method, but this increases the probability of prediction failure and decreases PICP. It can be observed that the selection of the optimal parameters for the prediction model actually changes with different situations, and the actual scheduling must be considered.

6 Conclusion

Due to the strong randomness of wind power, making point predictions accurately is relatively difficult. This paper improved the wind power prediction method based on the conditional copula function and proposed a combination of this method with multi-objective optimization algorithms, thus supplying reference information for short-term or real-time dispatch of the power system.

First, we take full advantage of the correlation relationship between wind power sequences in adjacent time periods by establishing the discrete condition copula function of the point to be predicted. Based on the problem of contradiction between multiple evaluation indicators when the values of the prediction model parameters change, we use the NSGA-II multi-objective algorithm to seek the optimal model parameter within its value range such that the non-inferior solution set of the conditional copula prediction model parameters is obtained. Second, to validate the effectiveness and applicability of the method proposed in this paper, the method is applied to three wind farms in China. We compare the results obtained using the proposed method with those from three more mature methods. The comparative analysis shows that the method

Table 2 Accuracy comparison of three methods

Method	W1		W2		W3	
	PICP	PIAW	PICP	PIAW	PICP	PIAW
M1	0.944	4.032	0.972	2.965	0.944	3.198
M2	0.903	4.154	0.944	3.068	0.931	3.342
M3	0.917	4.337	0.958	3.169	0.931	3.400
M4	0.889	4.085	0.944	3.089	0.917	3.455

Note: M1 means the method proposed by this paper; M2 means RVM; M3 means ANN; M4 means PSO-KELM; W1–W3 mean the wind farms 1–3 respectively

proposed in this paper has the smallest PIAW, the largest PICP, and the best effect. Finally, in the discussion, we believe that when PICP and PIAW are endowed with their respective weights, we must combine the actual requirements of scheduling. The results obtained using this method are not exactly the same in different practical situations.

However, certain shortcomings remain in our work. If the correlation relationship among wind power sequences is weak, the prediction results obtained using this method are usually poor. In the prediction, we only consider the correlation relationship between the wind power in the adjacent time periods but ignore the correlation relationship between the wind power sequence and influencing factors such as meteorology and unit maintenance. Therefore, in future work, we will further study the copula function wind power interval prediction based on the combination of meteorological factors and wind power.

Acknowledgements This work was supported by the National Natural Science Foundation of China (No. 51507141), Key research and development plan of Shaanxi Province (No. 2018ZDCXL-GY-10-04), the National Key Research and Development Program of China (No. 2016YFC0401409), and the Shaanxi provincial education office fund (No. 17JK0547).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- [1] Liu Y, Xiao LY, Wang HF et al (2013) Analysis on the hourly spatiotemporal complementarities between China's solar and wind energy resources spreading in a wide area. *Sci China (Technol Sci)* 56(3):683–692
- [2] Liu L, Ji TY, Li MS et al (2018) Short-term local prediction of wind speed and wind power based on singular spectrum analysis and locality-sensitive hashing. *J Modern Power Syst Clean Energy* 6(2):317–329
- [3] Lobo MG, Sanchez I (2012) Regional wind power forecasting based on smoothing techniques, with application to the Spanish peninsular system. *IEEE Trans Power Syst* 27(4):1990–1997
- [4] Bouffard F, Galiana FD (2008) Stochastic security for operations planning with significant wind power generation. *IEEE Trans Power Syst* 23(2):306–316
- [5] Wang ZM (2009) Study on cultivating and developing the industry chain of wind power in China. Dissertation, Jiangsu University
- [6] Radziukynas V, Klementavičius A (2016) Short-term forecasting of loads and wind power for Latvian power system: accuracy and capacity of the developed tools. *Latv J Phys Tech Sci* 53(2):3–13
- [7] Kazuki O, Kamiyama S, Michael P et al (2013) Optimal operation planning of wind farm installed BESS using wind power forecast data of wind turbine generators considering forecast error. *Int J Emerg Electric Power Syst* 14(3):207–218
- [8] Ranaboldo M, Giebel G, Codinaet B (2013) Implementation of a model output statistics based on meteorological variable screening for short-term wind power forecast. *Wind Energy* 16(6):811–826
- [9] Alexiadis M, Dokopoulos P, Sahsamanoglou H et al (1998) Short-term forecasting of wind speed and related electrical power. *Sol Energy* 63(1):61–68
- [10] Yuan K, Zhang KF, Zhang YX et al (2018) Irregular distribution of wind power prediction. *J Modern Power Syst Clean Energy* 6(6):1172–1180
- [11] Dong L, Wang LJ, Liao XZ et al (2009) Prediction of wind power generation based on time series wavelet transform for large wind farm. In: 2009 3rd international conference on power electronics systems and applications (PESA), Hong Kong, China, 20–22 May 2009, 4 pp
- [12] Taylor JW, McSharry PE, Buizza R (2009) Wind power density forecasting using ensemble predictions and time series models. *IEEE Trans Energy Convers* 24(3):775–782
- [13] Chun YL, Yan LH (2012) Wind prediction based on general regression neural network. In: 2012 second international conference on intelligent system design and engineering application, Sanya, China, 6–7 January 2012, 5 pp
- [14] Zhang G, Liu HC, Zhang JB et al (2018) Wind power forecasting based on variational mode decomposition multi-frequency combinations. *J Modern Power Syst Clean Energy*. <https://doi.org/10.1007/s40565-018-0471-8>
- [15] Wang CQ, Wei H, Wu SY (2017) Stochastic security constrained unit commitment considering uncertainty of wind power. *Power Syst Technol* 41(5):1419–1427
- [16] Chen J, Shen YX, Lu X et al (2016) An intelligent multi-objective optimized method for wind power prediction intervals. *Power Syst Technol* 40(8):2281–2287
- [17] Jiang YY, Lin XY (2015) Wind power short-term point forecast confidence interval estimation based on parametric and non-parametric calculation. *J Fuzhou Univ (Nat Sci Ed)* 43(1):75–82
- [18] Fabbri A, Roman T, Abbad J et al (2005) Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market. *IEEE Trans Power Syst* 20(3):1440–1446
- [19] Liu XJ, Xie CY (2014) Wind power fluctuation interval estimation based on beta distribution. *Electric Power Autom Equip* 34(12):26–30
- [20] Jeon J, Taylor JW (2012) Using conditional kernel density estimation for wind power density forecasting. *J Am Stat Assoc* 107(497):66–79
- [21] Kou P, Gao F, Guan XH et al (2012) Prediction intervals for wind power forecasting: using sparse warped Gaussian process. In: 2012 IEEE power and energy society general meeting, San Diego, USA, 22–26 July 2012, 8 pp
- [22] Wang CX, Lu ZX, Qiao Y et al (2010) Short-term wind power forecast based on non-parametric regression model. *Autom Electric Power Syst* 34(16):78–82
- [23] Fan L, Wang Y, Liang Z et al (2016) Prediction on short-term wind power interval based on EEMD-RVM. *Guangdong Electric Power* 29(2):14–20
- [24] Khosravi A, Nahavandi S, Creighton D (2013) Prediction intervals for short-term wind farm power generation forecasts. *IEEE Trans Sustain Energy* 4(3):602–610
- [25] Quan H, Srinivasan D, Khosravi A (2014) Short-term load and wind power forecasting using neural network-based prediction intervals. *IEEE Trans Neural Netw Learn Syst* 25(2):303–315



- [26] Yang XY, Guan WY, Liu YQ et al (2015) Prediction intervals forecasts of wind power based on PSO-KELM. *Proc CSEE* 35(S1):146–153
- [27] Lan F, Sang CC, Liang JJ et al (2016) Interval prediction for wind power based on conditional Copula function. *Proc CSEE* 36(S1):79–86
- [28] Aas K, Czado C, Frigessi A et al (2009) Pair-copula constructions of multiple dependence. *Insurance Math Econ* 44(02):182–198
- [29] Lan F, Nong ZG, Li JH (2016) Research on spatial and temporal correlation of wind power sequence. *Proc CSU-EPSA* 28(1):24–31
- [30] Genest C, Remillard B, Beaudoin D (2009) Goodness-of-fit tests for copulas: a review and a power study. *Insurance Math Econ* 44(2):199–213
- [31] Bhat CR, Eluru N (2009) A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. *Transp Res Part B Methodol* 43(7):749–765
- [32] Yu ZP (2011) The properties of conditional copula. *J Jiamusi Univ (Nat Sci Ed)* 29(5):763–765
- [33] Wang L (2009) A study on financial contagion effect among stock markets based on the conditional Copula model. Dissertation, Xiamen University
- [34] Cao YD (2014) The simulation of Garch model by multiple conditions Copula via vine structure. Dissertation, Lanzhou University
- [35] Li YD (2012) Wind speed models considering dependence and their applications in reliability evaluation of generating systems. Dissertation, Chongqing University
- [36] Bandyopadhyay S, Bhattacharya R (2013) Solving multi-objective parallel machine scheduling problem by a modified NSGA-II. *Appl Math Model* 37(10–11):6718–6729
- [37] He XH, He K (2017) Dynamic economic dispatch considering wind power penetration based on NSGA-II. *Electron Des Eng* 25(11):170–175

Gang ZHANG received his Ph.D. degree from Xi'an University of Technology in 2013. He is currently working in the Institute of Water Resources and Hydro-electric Engineering, Xi'an University of

Technology. His research interests focus on the renewable energy prediction and power system dispatching.

Zhixuan LI is currently studying in the Institute of Water Resources and Hydro-electric Engineering, Xi'an University of Technology. His research interests focus on the renewable energy prediction and power system dispatching.

Kaoshe ZHANG received his Ph.D. degree from Xi'an Jiaotong University. He is currently working in the Institute of Water Resources and Hydro-electric Engineering, Xi'an University of Technology. His research interests focus on the renewable energy prediction and power system dispatching.

Lei ZHANG received her master degree from Xi'an University of Technology in 2018. She is currently working in the Shaanxi Baoji electric power company. Her research interests focus on the renewable energy prediction and power system dispatching.

Xia HUA received his Ph.D. degree in semiconductor physics from the Department of Physics and Astronomy, Shanghai Jiao Tong University in 2014. He is currently working in the Gansu Electric Power Research Institute, Lanzhou, China. His research interest focuses on the renewable energy forecasting model.

Yongqing WANG is a professorate senior engineer, and currently working in Shaanxi Electric Power Research Institute, State Grid Shaanxi Electric Power Company. His research interest focuses on the renewable energy dispatching.