



Short-term wind power prediction based on ICEEMDAN-Correlation reconstruction and BWO-BiLSTM

Jingxia Liu¹ · Yanqi Wu¹ · Xuchu Cheng¹ · Baoli Li² · Peihong Yang¹

Received: 25 November 2023 / Accepted: 26 June 2024

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

Abstract

To solve the problems of high volatility and low prediction accuracy of wind farm output power, this paper proposes a short-term wind power prediction model with improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN), dispersive entropy combined with zero crossing rate (DE-ZCR) correlation reconstruction and beluga whale optimization (BWO) to optimize bidirectional long short-term memory (BiLSTM) neural network. Firstly, the original wind power is decomposed into multiple modal components by ICEEMDAN; secondly, the DE-ZCR method is used to evaluate the complexity and correlation of each component, and each modal component is reconstructed into a high frequency oscillation component, a medium frequency regular component, and a low frequency stable component; then the BWO-BiLSTM is used to predict each reconstructed power component separately, and finally the predicted values are superimposed to obtain the final results. The prediction model constructed in this paper is compared with four other models under different wind seasons, the results show that the model of this paper is superior to other models, validating the effectiveness of the combined prediction model.

Keywords Short-term wind power prediction · Improved complete ensemble empirical mode decomposition with adaptive noise · Correlation reconstruction · Beluga whale optimization · Bidirectional long short-term memory

1 Introduction

1.1 Background

In the context of the energy revolution and the strategic goal of “double carbon” the use and research of renewable energy has been developed by leaps and bounds, among which the wind power industry has been accelerated and the installed capacity has been growing rapidly. In the report released by the International renewable energy agency (IRENA) “World Energy Transition Outlook 2022”, it is pointed out that by 2030, wind power installed capacity will exceed 3.38 TW [1]. Wind power is gradually becoming one of the mainstream power generation technologies.

Due to the large random fluctuation of wind power output and large-scale wind power connected to the grid, it poses a great challenge to the stability of the power system and affects the operation and scheduling plan [2, 3]. Therefore, high-precision prediction of wind power output in the short-term effectively formulates reasonable power generation plans and reduces the occurrence of wind abandonment and power restriction [4–6], which is of great significance for the stable operation of the power grid.

1.2 Literature review

At present, the commonly used wind power forecasting methods mainly include physical methods [7], statistical methods [8], and artificial intelligence methods and combined forecasting methods [9]. The physical method uses the meteorological data provided by numerical weather prediction (NWP) and combines the geographic and physical information of wind farm to build a mathematical model to realize power prediction. This method does not have high technical requirements for wind farm, but relies on a large number of scene information when building the model. The

✉ Yanqi Wu
2083582564@qq.com

¹ The College of Information Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China

² China Academy of Aerospace Science and Industry Power Technology, Hohhot 010010, China

statistical method analyzes the historical meteorological data and extracts the characteristic law to predict the power. The modeling is complicated, and it is difficult to reflect the influence of nonlinear factors.

In recent years, with the development of big data application technology, a series of shallow machine learning methods have been widely applied in the field of power prediction [6, 10–12]. In order to further explore the correlation between historical data and extract deep nonlinear features [13], artificial intelligence and deep learning have become the hot and cutting-edge methods for wind power forecasting. Compared with traditional shallow machine learning methods, this method can effectively deal with huge and cumbersome historical data and has strong learning and generalization capabilities [14, 15]. Xiong et al. [16] proposed to use convolutional neural networks (CNN) for short-term feature extraction, and then use long-term memory (LSTM) to extract the long-term trend of local high-dimensional features. The prediction inaccuracy caused by mixing of original data is effectively reduced. Zhang et al. [17] proposed an interpretable short-term wind power prediction model based on integrated depth map neural network. Firstly, the graph network model (GNN) with attention mechanism is applied to aggregate and extract the spatio-temporal characteristics of wind power data. Then, LSTM method is used to process the extracted features, and the wind power prediction model is established, and good prediction accuracy is obtained. However, in view of the problems such as too many parameters and difficult parameter selection of deep learning models, some researchers propose to use optimization algorithms to adapt to the hyperparameters of neural networks, so as to significantly improve the training efficiency and performance of neural networks [18–20]. Tian et al. [21] proposed an improved whale optimization algorithm to optimize these parameters, and the optimized neural network improved the prediction effect. Pavlov et al. [22] proposed a method combining signal decomposition technology and LSTM neural network for wind power generation prediction. The improved crawling search algorithm (RSA) was used to further adjust hyperparameters, and the proposed tuned LSTM model was tested on wind power production data sets with a two-hour resolution, which improved the prediction performance.

Combined prediction method combines the advantages of different models to achieve better prediction accuracy and broaden application scenarios [23]. Due to the strong volatility and instability of wind power [24], the data decomposition and integration method has become a popular direction in the field of power forecasting [25, 26]. Commonly used decomposition techniques include wavelet transform (WT) [27], variational mode decomposition (VMD) [28], empirical mode decomposition (EMD) [30] and other decomposition methods, which can decompose time series into smooth, regular and accurately identified components, eliminate noise

redundancy to a certain extent, and improve the robustness of the prediction model. Cui et al. [30] proposed a point-interval forecasting framework for ultra-short-term wind power. In the constructed point prediction model, the VMD algorithm is used to decompose the wind power sequence, and a more regular sequence is obtained. Then, combining ISSA, bidirectional gated cycle unit (BiGRU) and Attention, the ISA-BiGRU-Attention model is constructed for wind power subsequence prediction. Li et al. [31] proposed a method to decompose the wind speed series by using CEEMDAN, and then input the components into the PSO-ELM model, effectively reducing the random fluctuations of the original data. In the process of signal decomposition, CEEMDAN still produces too many components and residual noise of different degrees. Bommidi et al. [32] used ICEEMDAN for noise reduction of wind speed data and Transformer model for training of wind speed subsequence, achieving good results. In order to improve the computational efficiency, Literature [33–35] adopted a series of entropy theories to reconstruct the components generated by modal decomposition, effectively analyzed the complexity of each modal component of wind power, analyzed the characteristic relations of each modal component, and improved the computational efficiency of the algorithm. However, due to the choice of subjective experience, the correlation between modal components is not fully considered.

1.3 The contributions

To solve the above problems, this paper proposes a short-term wind power prediction model based on ICEEMDAN decomposition, DE-ZCR correlation reconstruction, and BWO-BiLSTM.

- (1) This paper presents a method of wind power timing decomposition and correlation reconstruction. First, ICEEMDAN is used to decompose the raw wind power, smoothing out the data fluctuations to extract the internal hidden information. then the wind power components are simplified and reconstructed into three components with different frequency characteristics by DE-ZCR correlation analysis method, which are high frequency oscillation component, medium frequency regular component, and low frequency stable component to strengthen the correlation degree among the modal components and reduce the computational complexity of the model.
- (2) BWO was used to optimize the hyperparameters of BiLSTM, and the optimization effect was tested from the perspective of the whole and components, and the optimal parameters were identified; Then the reconstructed

subsequences were input into BWO-BiLSTM for prediction.

- (3) Finally, the predicted values are superimposed to obtain the final result. Using wind farm data for multiple experimental comparisons and analysis, the method in this paper achieves higher prediction accuracy.

1.4 The structure of the paper

The chapters of this paper are arranged as follows: Sect. 2 introduces the relevant principles of ICEEMDAN, DE and ZCR; Sect. 3 introduces the principle and method of BWO-BiLSTM method. In Sect. 4, the implementation steps of ICEEMDAN-DE-ZCR-BWO-BiLSTM are introduced in detail. In Sect. 5, wind speed data sets and performance indicators are introduced and the validity of BWO is verified. Finally, the concrete experimental results are given and the error analysis is carried out. Part 7 summarizes the conclusions of this study and the future work.

2 Data decomposition and selection

2.1 ICEEMDAN theory

Improved complete ensemble empirical mode decomposition with adaptive noise (ICEEMDAN) is an improved modal decomposition method based on EMD [36], compared with CEEMDAN The algorithm further weakens the residual noise signal and solves the problem of residual noise and pseudo-modality in the CEEMDAN algorithm, ensuring that the decomposed signal retains all the information of the original signal and improving the stability and accuracy of the decomposition. The following is the decomposition process of ICEEMDAN for historical power time series:

- (1) Define $X(t)$ as the sequence of historical wind power to be decomposed, add group I Gaussian white noise to $X(t)$, and construct the sequence $X^i(t)$ can be expressed as Eq. (1):

$$X^i(t) = X(t) + \beta_0 E_1(\omega^i(t)), \quad i = 1, 2, \dots, I \quad (1)$$

$E_k(\cdot)$ is the k th order modal component operator generated by EMD, $k = (1, 2, \dots, k)$, ω^i represents Gaussian white noise.

- (2) Defining $M(\cdot)$ as the local mean of the generated signal, the first residual equation is expressed as Eq. (2):

$$r_1 = \frac{1}{I} \sum_{i=1}^I M(X^i(t)) \quad (2)$$

- (3) Calculate the first modal component expressed as Eq. (3):

$$d_1 = X(t) - r_1 \quad (3)$$

- (4) Continue to add Gaussian white noise to $X(t)$, calculate the second residual component as r_2 , define the second modal component as d_2, r_2 and d_2 are expressed as Eqs. (4) and (5), respectively:

$$r_2 = r_1 + \beta_1 E_2(\omega^i(t)) \quad (4)$$

$$d_2 = r_1 - r_2 = r_1 - \frac{1}{I} \sum_{i=1}^I M(r_1 + \beta_1 E_2(\omega^i(t))) \quad (5)$$

- (5) Using the same method to calculate the k th modal component and the k th residual expressed as Eqs. (6) and (7), respectively:

$$r_k = \frac{1}{I} \sum_{i=1}^I M(r_{k-1} + \beta_{k-1} E_k(\omega^i(t))) \quad (6)$$

$$d_k = r_{k-1} - \frac{1}{I} \sum_{i=1}^I M(r_{k-1} + \beta_{k-1} E_k(\omega^i(t))) \quad (7)$$

- (6) The above method is iterated until all modal components and residuals are obtained, and the algorithm ends.

2.2 Dispersion entropy theory

Dispersion entropy (DE) is a kinetic index to analyze the degree of signal irregularity, which can be used as a measure of time series complexity in the field of forecasting, and can effectively reflect the characteristic mutational behavior of complex nonstationary data [37] and accurately depict the internal complexity of wind power series data. Compared to traditional sample entropy, approximate entropy, and permutation entropy, it is faster to calculate, more resistant to interference and can take into account the magnitude relationship between magnitudes. It is more suitable for online real-time applications and has more stable results in measuring signal complexity. The basic process is as follows:

Given historical wind power: $x = \{x_1, x_2, \dots, x_n\}$, the x is mapped to $y = \{y_1, y_2, \dots, y_n\}$ by the normal cumulative distribution function, $y_i \in (0, 1)$ where y_i can be expressed as:

$$y_i = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x_j} e^{-\frac{(x-rms)^2}{2\sigma^2}} dt \quad (8)$$

In the above equation, σ denotes the standard deviation and rms denotes the root mean square of the original signal.

Assign y_i to the range of $[1,c]$ integers, as indicated by:

$$z_i^c = \text{round}(c \cdot y_i + 0.5) \tag{9}$$

Setting the delay time as d and the number of embedding dimensions as m , get a set of embedding vectors:

$$z_i^{m,c} = \{z_i^c, z_{i+d}^c, \dots, z_{i+(m-1)d}^c\} \tag{10}$$

, where $i = 1, 2, \dots, n - (m - 1)d$, the embedding vector $z_i^{m,c}$ maps to dispersion patterns $\pi_{v_0, v_1, \dots, v_{m-1}}$, $v_0 = z_i^c$, $v_1 = z_{i+d}^c, \dots, v_{m-1} = z_{i+(m-1)d}^c$. The number of dispersion patterns assigned to each $z_i^{m,c}$ is c^m .

The relative frequencies of the dispersive modes are expressed as:

$$p(\pi_{v_0, v_1, \dots, v_{m-1}}) = \frac{\text{Num}(\pi_{v_0, v_1, \dots, v_{m-1}})}{n - (m - 1)d} \tag{11}$$

The dispersion entropy value of the original sequence signal is defined as:

$$E_{DE}(x, m, c, d) = - \sum_{\pi=1}^{c^m} p(\pi_{v_0, v_1, \dots, v_{m-1}}) \cdot \ln(p(\pi_{v_0, v_1, \dots, v_{m-1}})) \tag{12}$$

3 BWO-BiLSTM model theory

3.1 Beluga whale optimization algorithm

Beluga whale optimization [38] (BWO) is a new swarm intelligence optimization algorithm proposed in 2022. BWO establishes three phases of exploration, exploitation, and whale fall. The equilibrium factor and whale fall probability in BWO are adaptive and have the characteristics of quick convergence and robustness. In addition, the algorithm introduces Lévy flight to enhance the global convergence of the development phase.

BWO Based on the population mechanism, beluga whales are considered as search agents, while each beluga whale is a candidate solution, which is updated during optimization. The BWO algorithm can be gradually transformed from exploration to development. Depending on the balance factor B_f , which is modeled as:

$$B_f = B_0 \left(1 - \frac{T}{2T_{\max}}\right) \tag{13}$$

, where T is the current iteration, T_{\max} is the maximum iterative number, B_0 and randomly changes between $(0, 1)$ at each

iteration. The exploration phase happens when the balance factor $B_f > 0.5$, while the exploitation phase happens when $B_f < 0.5$.

(1) Exploration phase: The mathematical model for simulating the swimming behavior of beluga whales is as follows:

$$X_{i,j}^{T+1} = X_{i,p_j}^T + (X_{r,p_1}^T - X_{i,p_j}^T)(1 + r_1) \sin(2\pi r_2), \quad j = \text{even} \tag{14}$$

$$X_{i,j}^{T+1} = X_{i,p_j}^T + (X_{r,p_1}^T - X_{i,p_j}^T)(1 + r_1) \cos(2\pi r_2), \quad j = \text{odd}, \tag{15}$$

, where T is the current iteration, $X_{i,j}^{T+1}$ is the original position for the i -th beluga whale on the j -th dimension, $p_j (j = 1, 2, \dots, d)$ is a random number chosen from the d -dimension, X_{i,p_j}^T is the position of the i -th beluga whale in p_j dimension, X_{i,p_j}^T and X_{r,p_1}^T are the current positions for i -th and r -th beluga whale, r is a randomly selected beluga whale, r_1 and r_2 are random numbers between $(0, 1)$, to enhance random operations during the exploration phase.

(2) Development phase: simulating the foraging behavior of beluga whales, which can forage and move cooperatively based on their nearby locations. Lévy's flight strategy is introduced into the development phase of BWO to enhance convergence. Assuming that the prey is captured using the Lévy flight strategy, the mathematical model is represented as follows:

$$X_i^{T+1} = r_3 X_{best}^T - r_4 X_i^T + C_1 \cdot L_F \cdot (X_r^T - X_i^T) \tag{16}$$

$$C_1 = 2r_4(1 - T/T_{\max}) \tag{17}$$

X_{best}^T is the best location for beluga whale, r_3 and r_4 are random numbers between $(0, 1)$, C_1 is a measure of the random jump intensity of Levy's flight. L_F is the Lévy flight function, which is calculated as follows:

$$L_F = 0.05 \times \frac{u \cdot \sigma}{|v|^{1/\beta}} \tag{18}$$

$$\sigma = \left(\frac{\Gamma(1 + \beta) \cdot \sin(\pi\beta/2)}{\Gamma((1 + \beta)/2) \cdot \beta \cdot 2(\beta - 1)/2} \right)^{1/\beta} \tag{19}$$

, where u and v are normally distributed random numbers, $\beta = 1.5$.

(3) Whale fall: To simulate the behavior of whale fall, the probability of whale fall is selected by the individuals of the population as a subjective assumption during each iteration and used to simulate small changes in the population. The beluga whale population is assumed to move randomly, and to ensure that the population size is constant, the beluga whale’s position and whale fall step size are used to determine the updated position. The mathematical model is expressed as follows:

$$X_i^{T+1} = r_5 X_i^T - r_6 X_r^T + r_7 X_{step} \tag{20}$$

, where r_5 and r_6 are random numbers between (0, 1), X_{step} is the step length of the whale fall, as follows:

$$X_{step} = (u_b - l_b) \exp(-C_2 T / T_{max}) \tag{21}$$

C_2 is the step factor associated with whale fall probability and population size, u_b and l_b are the upper and lower bounds of the variables, respectively. From the above equation, it can be seen that the step size is influenced by the variable bound, the number of iterations and the maximum number of iterations.

3.2 BiLSTM neural network

Long short-term memory neural network [39] (LSTM) is a deformation of recurrent neural network(RNN), and the internal structure of LSTM neural network replaces the simpler neurons in the hidden layer with memory cells compared to RNN, and adds concepts such as cell states and the structure of gates. The internal structure includes input gates, output gates, forget gates, and cellular cell states [40], which work together with the three gating units to maintain, update, and transmit state information. The formula is as follows:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + \beta_f) \tag{22}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + \beta_i) \tag{23}$$

$$\tilde{C} = \tanh(W_c[h_{t-1}, x_t] + \beta_c) \tag{24}$$

$$C_t = f_t C_{t-1} + i_t \tilde{C} \tag{25}$$

$$y_t = \sigma(W_o \cdot [h_{t-1}, x_t] + \beta_o) \tag{26}$$

$$h_t = y_t \tanh(C_t) \tag{27}$$

In the above equation, f_t, i_t, y_t are the outputs of the oblivion gate, the input gate and the output gate, respectively at

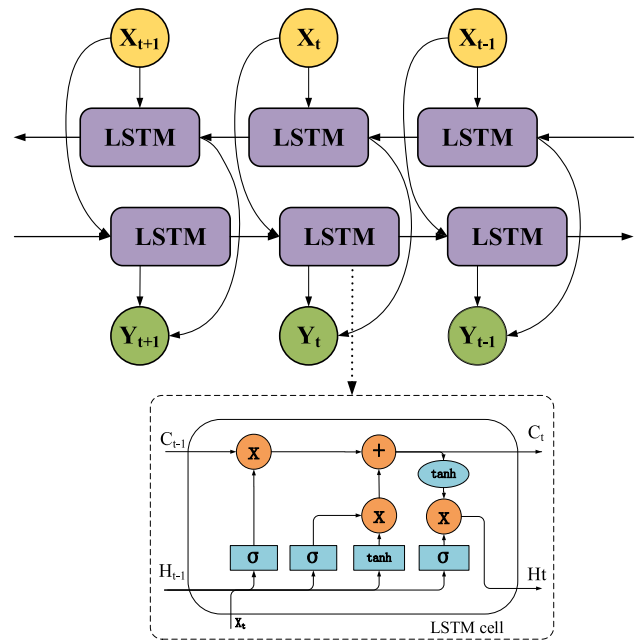


Fig. 1 Structure of BiLSTM

time t ; \tilde{C} and C_t are cell and intermediate cell states, respectively; W_f, W_i, W_c, W_o are the weight matrices of each gating unit structure; $\beta_f, \beta_i, \beta_c, \beta_o$ are bias terms; σ is the activation function; x_t is the input to the hidden layer at time t ; h_t is the output of the hidden layer at time t ; h_{t-1} is the output of the hidden layer at time $t - 1$.

The LSTM neural network can only learn and transmit information in one direction when training, without taking into account the information of future moments, which is inherent in this learning and training method, and therefore cannot obtain all the information contained before and after the time series, and the data utilization is low and the training time is prolonged.

The BiLSTM neural network uses a two-layer LSTM structure [41], which is a combination of RNN and LSTM neural networks. Its network structure is shown in Fig. 1. Compared with the traditional LSTM network, its prominent feature is to build a bidirectional recurrent neural network, including forward and backward layers, where the forward layer learns to train historical information and the backward layer learns to train future information. BiLSTM neural networks can be used to obtain past and future state information to train on both forward and backward data, thus giving it stronger nonlinear processing power to handle uncertain relationships in the data and improve accuracy of prediction.

4 ICEEMDAN-DE-ZCR-BWO-BiLSTM Wind power prediction model

4.1 Process of the prediction model

Based on the above modeling methods, this study proposes a combined wind power prediction model based on ICEEMDAN decomposition, DE-ZCR Relevance reconstruction, and BWO-BiLSTM. The overall structure of the model and the flow of the BWO algorithm are shown in Fig. 2, and the specific prediction process is as follows:

- (1) The original wind power data is pre-processed to construct a high-quality dataset;
- (2) After normalization of the wind power series, the IMF components and one residual component are obtained by ICEEMDAN decomposition, and the IMF components and residuals are reconstructed by dispersive entropy algorithm and combined with Zero Crossing Rate to optimize the frequency division process, and the components are divided into high frequency oscillation sequences (IMF₁-IMF_m), medium frequency regular sequences and low frequency stable sequences according to DE-ZCR correlation reconstruction;
- (3) Construct the model input, in which the high frequency power sequence is used as the reconstructed input component alone, and the mid-frequency and low frequency power sequences selected by the DE-ZCR method are combined as one power component each as the training set input to the BWO-BiLSTM model;
- (4) Initialize the parameters of the BiLSTM network and optimize BiLSTM network hyperparameters using BWO. Set the number of BWO populations, the maximum number of iterations, and the fitness function, with the goal of finding the set of hyperparameter combinations that make BiLSTM training samples the most accurate, calculating fitness values and updating positions, and using the corresponding parameter values to build a prediction model to train samples until the optimal solution is reached;
- (5) Transform the optimal solution after BWO optimization into the optimal parameters of the prediction model, and assign the optimal parameters to the BiLSTM model;
- (6) Use the BWO-BiLSTM network to predict the optimized reconstructed power components separately, and reconstruct the prediction structure of each power component by superposition to obtain the final wind power prediction results;

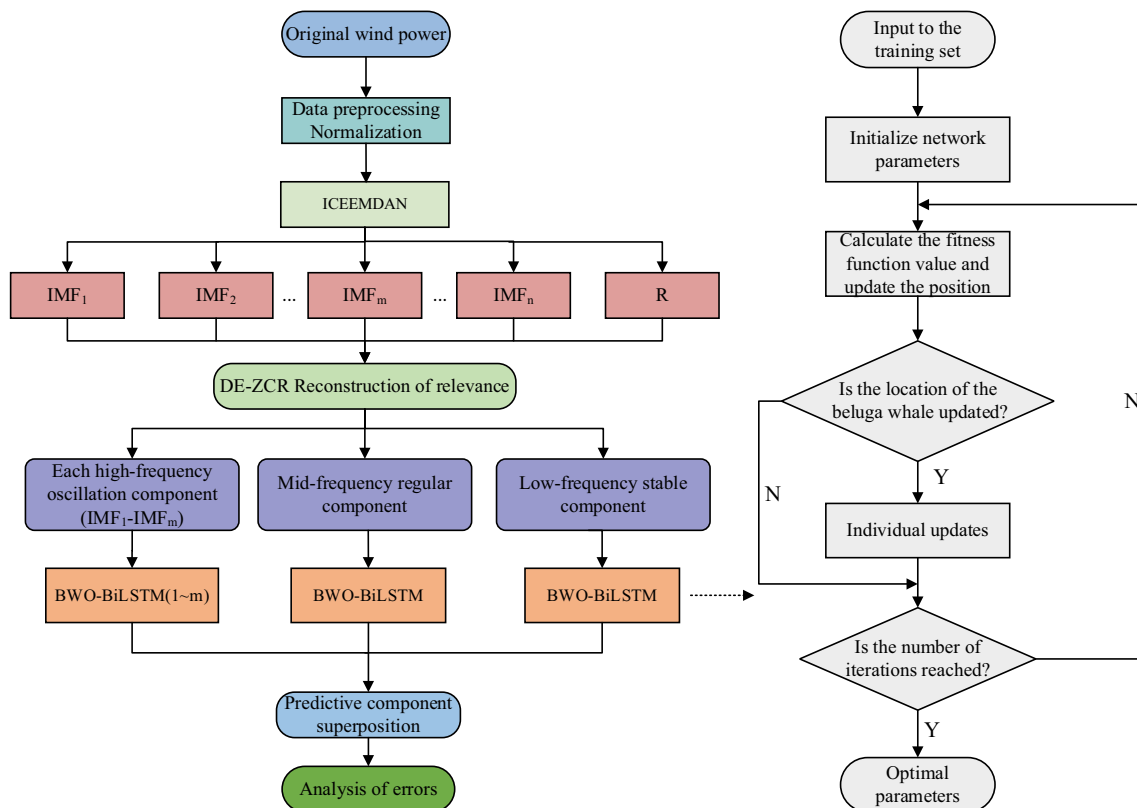


Fig. 2 Flow chart of combined prediction model

- (7) The actual wind power data are compared with the wind power forecasts obtained and error analysis.

4.2 Evaluation of the model

In order to test the effectiveness of the short-term wind power portfolio forecasting model proposed in this paper, three different evaluation criteria are used to evaluate the forecasting accuracy of the model. They are: mean absolute error (MAE), root mean square error (RMSE) and fitting coefficient R-Square (R^2) [42], normalized root mean square error (NRMSE) and normalized mean absolute error (NMAE) [43].

$$MAE = \frac{1}{n} \sum_{i=1}^n |f(x_i) - a(x_i)| \tag{28}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n |f(x_i) - a(x_i)|^2}{n}} \tag{29}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (a(x_i) - f(x_i))^2}{\sum_{i=1}^n (a(x_i) - \bar{y})^2} \tag{30}$$

$$NRMSE = \frac{1}{C} \sqrt{\frac{1}{n} \sum_{i=1}^n |f(x_i) - a(x_i)|^2} * 100 \tag{31}$$

$$NMAE = \frac{1}{n \cdot C} \sum_{i=1}^n |f(x_i) - a(x_i)| \tag{32}$$

In the above equation, $a(x_i)$ is the true value of the training sample at time i , $f(x_i)$ is the predicted value of the training sample at time i , \bar{y} is the average of the wind power values at time i , n is all data sample points, and C is the installed capacity of the wind farm.

5 Example verification analysis

5.1 Data analysis and selection

This study uses the measured data of a wind farm in Inner Mongolia in China to conduct short-term wind power forecasting experiments. According to the characteristics of the distribution pattern of high and minor wind seasons in Inner Mongolia, typical wind speed scale scenarios throughout the year are mostly divided into high wind seasons (January—May and September–December) and minor wind seasons (June–August). In order to verify the validity of the model prediction, the experiment in this paper distinguishes the high wind season and minor wind season as the typical wind season scenarios of wind farms for analysis, and the typical wind

speed seasonal wind power output scenarios for short-term prediction.

The wind farm has an installed capacity of 96 MW and consists of 48 wind turbines rated at 2 MW. A total of 2880 wind power sample points were selected in the wind farm during the light wind season from 0:00 on June 1, 2020 to 0:00 on July 1, 2020, and 2873 wind power sample points during the strong wind season from 0:00 on November 1, 2020 to 0:00 on December 1, 2020, for prediction experiment, with a time sampling interval of 15 min. There were 96 data points measured in one day. After data preprocessing, the last 280 sampling points (3d) were selected as the test set of the prediction model in the two prediction experiments of the small and small wind seasons, and the remaining data were used as the training set. The training set: test set was 9:1.

Due to the influence of natural wind power, wind power output shows a certain volatility, which is reflected in a large seasonal change rate, showing strong seasonality and intermittency. It can be seen from the Figs. 3 and 4 that during the small wind season, the wind power fluctuates greatly and is not stable. The overall fluctuation of wind power during the gale season is relatively small, and due to the high wind speed, wind turbines are able to capture more wind energy and convert it into electricity. Therefore, high wind seasons usually correspond to higher wind power output.

5.2 Data preprocessing and normalization

Normalization process: Different evaluation indexes have different scale units, so the data normalization process effectively solves the comparability between the error evaluation indexes, the method is min–max standardization, which is a linear transformation of the original data, so that the result value is mapped to [0,1], the specific formula is as follows:

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{33}$$

, where x is the wind power sample point, x_{\max} is the maximum value among the sample points, x_{\min} is the minimum value among the sample points, and x^* is the normalized is the sample point.

5.3 Wind power sequence decomposition

The sample points of wind power from 0:00 on June 1, 2020 to 0:00 on July 1, 2020, and from 00:00 on November 1, 2020 to 0:00 on December 1, 2020, in the wind farm light wind season, were taken for decomposition experiment, and the sampling interval was 15 min. The collected and processed wind power series are decomposed by ICEEMDAN, and 9 IMF components (IMF₁ ~ IMF₉) and 1 residual component R are obtained for each of the large and small wind season data

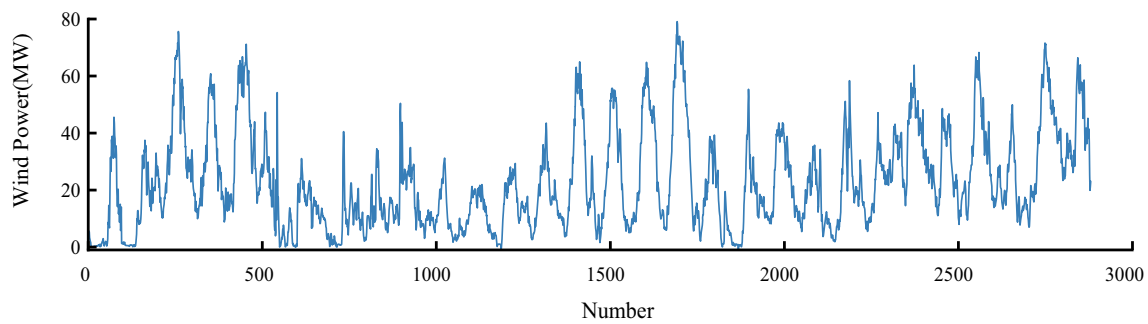


Fig. 3 Historical wind power (June)

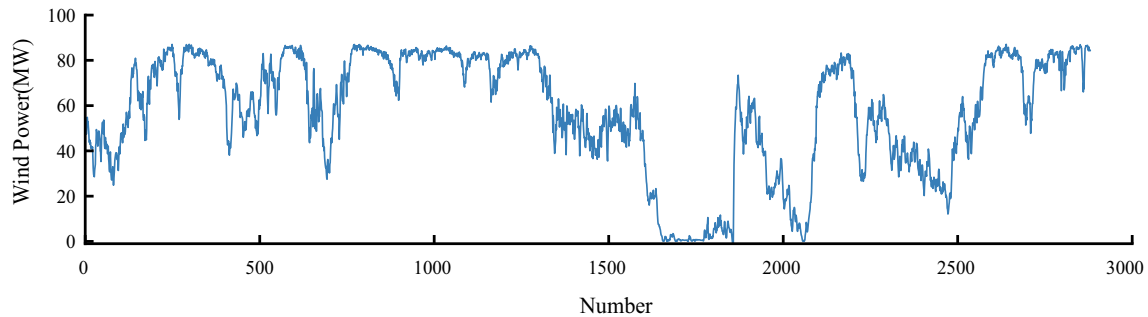


Fig. 4 Historical wind power (November)

samples, and the subseries are distributed from high to low frequency. The standard deviation of the white noise is 0.2, the number of additions is 100, and the number of iterations is 1000. The curves of each wind power component are shown in Fig. 5.

From the analysis of Fig. 5, the IMF component of the original wind power series decomposed by ICEEMDAN has a small fluctuation range and good smoothness compared with the wind power curve without decomposition, which can also well reflect the internal fluctuation situation and characteristic information. When the frequency decreases, the signal curve of the decomposed sequence components gradually stabilizes and has a certain periodicity.

5.4 DE-ZCR sequence component relevance reconstruction

In this study, the complexity of each component is evaluated by calculating the dispersion entropy value of each power component, and the IMF components are regrouped, and the larger the dispersion entropy value, the higher the complexity. The dispersion entropy value is calculated for each parameter setting m to take the value of 5, d to take the value of 1, and c to take the value of 4. Meanwhile, zero crossing rate (ZCR) of the power components is calculated to analyze the amplitude frequency variation and distinguish the frequency characteristics of the components, which better reflects the overall trend of variation.

ZCR calculated by the following formula:

$$Z = \frac{n_{\text{zero}}}{N} \quad (34)$$

In the above equation, n_{zero} is the number of transitions, N is the number of sample points, Z is the transitions rate, and the results are kept in four decimal places. When the transitions rate is lower than 0.01, it is defined as the low frequency component, and higher than 0.01, it is defined as the high frequency component. The dispersion entropy values and zero crossing rate of each sample are calculated as shown in Table 1.

As can be seen from Table 1, the dispersion entropy values of the original power components $\text{IMF}_1 \sim \text{R}$ are ranked from high to low, which better reflects the complexity variation of the power components. If the power components are selected based on the variation in DE size with subjective experience, without considering the frequency correlation between the power components, they are not easily trained adequately by the prediction model and yield better prediction results. Therefore, in this study, dispersion entropy combined with over-zero rate (DE-ZCR) is used as the correlation determination method for reconstructing the components, and the calculation of ZCR can divide the high and low frequency bands, and simultaneously combine DE to evaluate the complexity of the power components, which eliminates

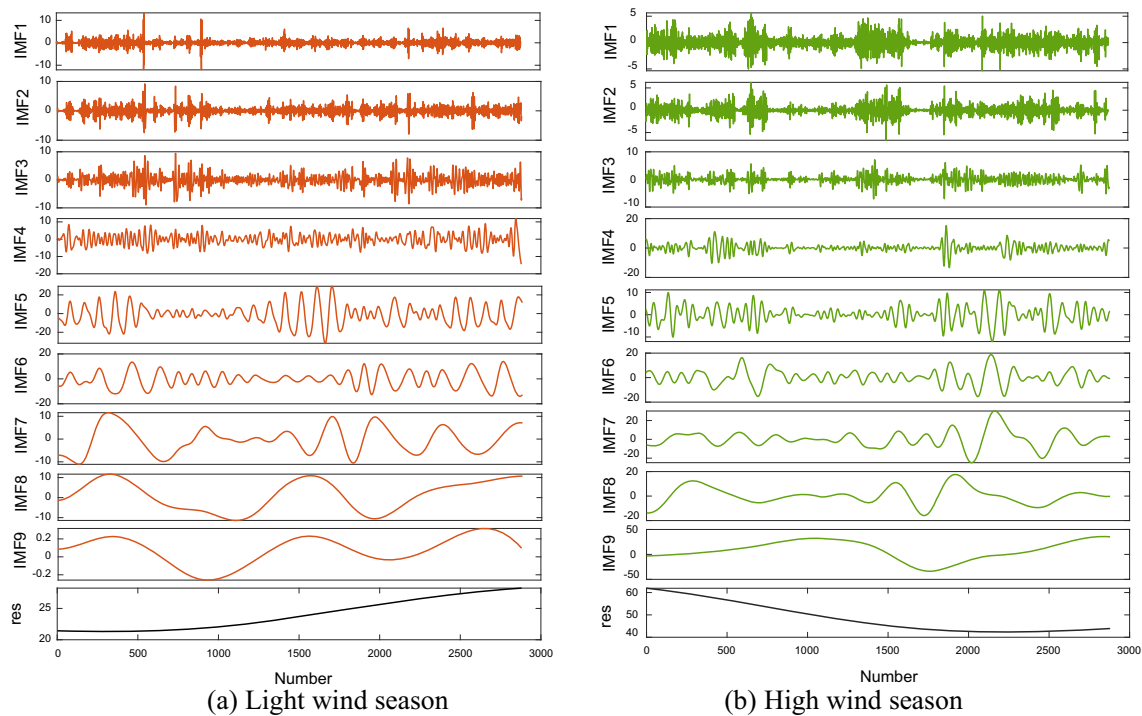


Fig. 5 Wind power series decomposition results

Table 1 Wind power modal component DE values and ZCR values

Power component	Light wind season (DE)	High wind season (DE)	Light wind season (ZCR)	High wind season (ZCR)
IMF ₁	6.9403	7.1100	1.5200	1.5558
IMF ₂	6.2332	6.1184	0.7042	0.7350
IMF ₃	5.1062	5.0678	0.3508	0.3733
IMF ₄	3.6968	3.9575	0.1342	0.1833
IMF ₅	3.1192	3.3742	0.0592	0.0900
IMF ₆	2.7616	2.8042	0.0350	0.0392
IMF ₇	2.4175	2.3510	0.0133	0.0192
IMF ₈	1.9215	2.0099	0.0042	0.0092
IMF ₉	1.9207	1.8950	0.0025	0.0025
R	1.7457	1.7503	0.0000	0.0000

the limitations of the single determination method in reconstructing the power components and improves the robustness and consistent correlation of the power component division.

According to the data results shown from Table 1, further correlation analysis is performed on the power components, and the DE-ZCR correlation is optimally reconstructed as shown in Fig. 6.

Using DE-ZCR method to reconstruct the correlation of power components, it is known from the definition of over-zero rate that the power components IMF₈-R are divided into low frequency part, and their IMF₈-R distribution points are very close to each other in the correlation analysis graph, and their distances are basically the same, so they

are reconstructed and superimposed as low frequency stable components (C₅); the power components IMF₁-IMF₇ are divided into high frequency part according to the over-zero rate, but The input components are still too much, in order to tap the coupling characteristic information between the components and make further division, it can be seen that the dispersion entropy value of IMF₃-IMF₄ has obvious abruptness compared with other components, in the above figure IMF₄-IMF₇ has obvious regular arrangement, the frequency is more centered, so it will be reconstructed and superimposed as medium frequency regular component (C₄), while the IMF₁-IMF₃ component has its IMF₁-IMF₃ components have the highest complexity and contain the most

Fig. 6 DE-ZCR relevance reconstruction analysis

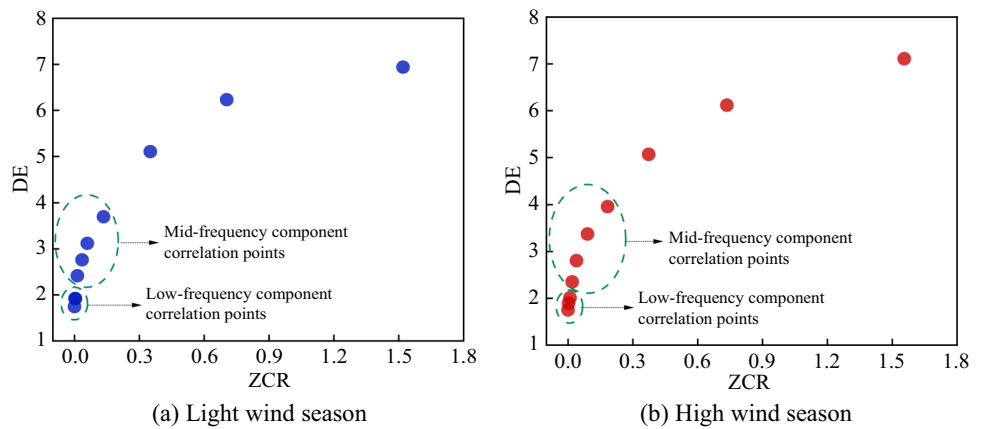


Table 2 Different types of reconstructed power components

Original power component	Reconstructing the fraction	Type
IMF ₁	C ₁	High frequency oscillation component
IMF ₂	C ₂	
IMF ₃	C ₃	
IMF ₄ ~ IMF ₇	C ₄	Medium frequency regular component
IMF ₈ ~ R	C ₅	Low frequency stable component

wind power hidden information, so they are reconstructed as high frequency oscillation components for prediction (C₁, C₂, C₃), and the reconstructed sequences are C₁, C₂, C₃, C₄, C₅ sequences, the reconstruction results are shown in Table 2. After the reconstruction of each component, the eigenvector was constructed by DE-ZCR optimization, and the input parameter optimization prediction model was used for power prediction.

In order to further validate the effectiveness of the decomposition reconstruction method in this paper, three different control models were selected for comparison tests in the test set using BiLSTM neural network, which are Method 1: IMF₁ ~ R components are modeled and predicted separately; Method 2: C₁ + C₂ + C₃, C₄, C₅ are modeled and predicted separately; Method 3 (the method in this paper): C₁, C₂, C₃, C₄, C₅ are modeled and predicted separately, modeling prediction. The judgment criteria of the above prediction models are shown in Fig. 7. Vertical coordinates of Fig. 7 shows the error rate, and it can be seen from the figure that the RMSE and MAE of this method decrease by 0.22 MW and 0.11 MW, respectively, compared with method 1, and by 1.14 MW and 1.00 MW, respectively, compared with method 2, which verifies the reliability of correlation reconstruction

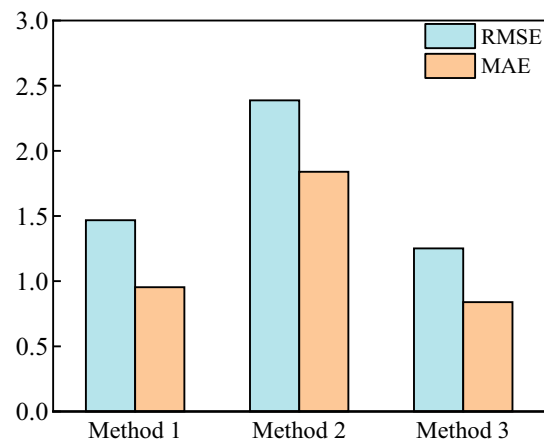


Fig. 7 Comparison of power prediction performance of different decomposition methods

Table 3 Relative error of reconstructed power by different methods

Method	RMSE/MW	MAE/MW
1	1.47	0.95
2	2.39	1.84
3	1.25	0.84

The experimental results of method 3 are highlighted in boldface

in optimizing the prediction model in this paper. Boldface denotes Method 3 (Table 3).

5.5 BWO-BiLSTM model validation

The BiLSTM neural network model has subjective and random parameter adjustment during the training process, which can easily lead to problems of instability and poor prediction accuracy during model convergence. Therefore, this paper verifies the effectiveness of BWO algorithm on BiLSTM parameter optimization from two aspects: (1) Taking IMF1

Table 4 Parameter interval setting

	Parameter name	Parameter value range
BWO-BiLSTM	Learning rate	[0.0001, 0.1]
	Iterations	[0, 200]
	Hidden layer 1 nodes	[0, 100]
	Hidden layer 2 nodes	[0, 100]

Table 5 Basic parameter settings of the algorithm

Optimization algorithm	Parameter
CS	$P = 0.2, \alpha = 0.01, \beta = 1.5$
GWO	$\alpha, \beta, \delta [0 \ 0 \ 0]$
WOA	$P_i = 0.5, l \in [-1, 1]$
BWO	$W_f \in (0.05, 0.1), \beta = 1.5$

power component as an example, three hyperparameters, namely hidden layer unit, training cycle and learning rate, are set to verify the effectiveness of BWO from the component perspective; (2) taking the data of light wind season in Sect. 4.1 as an example, four hyperparameters, namely learning rate Lr, number of iterations K, number of nodes in hidden layer 1 and number of nodes in hidden layer 2, were set to verify the effectiveness of BWO from an overall perspective (Table 4).

(1) In order to verify the superiority of the BWO algorithm and the effectiveness of regulating the BiLSTM hyperparameters, the experiments are conducted with the IMF1 power component as an example using four groups of algorithms to optimize the BiLSTM model, namely CS-BiLSTM, GWO-BiLSTM, WOA-BiLSTM, and BWO-BiLSTM models for effectiveness comparison experiments, with the root mean square error function as the target fitness function to solve for the set of hyperparameter combinations with the smallest RMSE in the training sample, and the RMSE indicates the degree of difference between the actual power and the predicted power. The optimization intervals for the three parameters of BiLSTM are set as follows: the number of implicit layer cells is [50,200], the training period is [100,300], the learning rate is [0.005,0.1], the Adam algorithm is used as the network solver, and the tanh function is used for the activation function. The population size for each algorithm is set to 20 and the maximum number of iterations is 100, and the additional parameters for each specific algorithm are set as follows (Table 5).

The four groups of experiments in the optimization process are shown in Fig. 8, and the analysis shows that CS and GWO converge slowly in the optimization process of

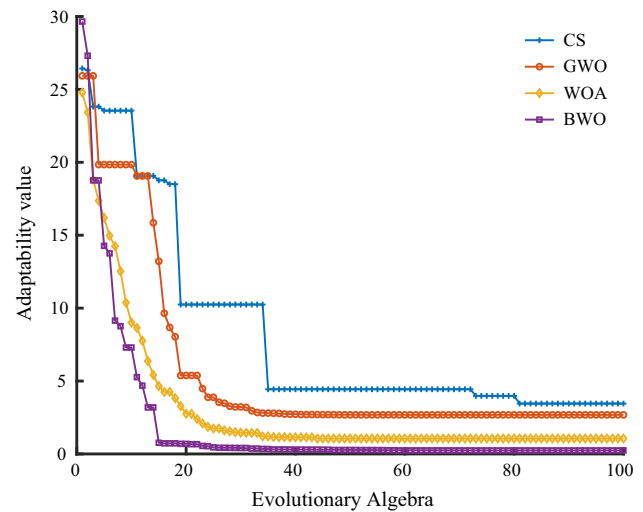


Fig. 8 Iterative process of adaptation value change

BiLSTM, WOA converges faster in the early stages and gradually tends to converge, but the root mean square error of the above three methods is relatively large and the results are not satisfactory. The BWO is smooth and converges rapidly, reducing the RMSE to a minimum. Compared to the remaining three algorithms, BWO is more robust, faster and has the best results.

(2) In the experiment, data from the light wind season shown in Sect. 4.1 were utilized. To rigorously evaluate the optimization performance of the BWO algorithm, it was employed to further enhance the BiLSTM model. With RMSE serving as the target fitness function, a BiLSTM neural network comprising two hidden layers was devised. The BWO algorithm was applied to optimize several parameters of the BiLSTM model, including the learning rate (Lr), iteration number (k), the number of nodes in hidden layer 1 (L1), and the number of nodes in hidden layer 2 (L2). The parameter settings for the integrated BWO-BiLSTM approach are presented in the accompanying table (Table 4).

After experimental iteration, the optimization process of Lr, K, L1, L2 and other hyperparameters by BWO algorithm is shown in Fig. 9.

According to the analysis in Fig. 9, after 20 optimization iterations, all four hyperparameters tend to converge after the 27th iteration. After convergence, the learning rate of the model is 0.0095, the number of iterations is 98, the number of nodes in hidden layer 1 is 63, and the number of nodes in hidden layer 2 is 59. In addition, the fitness function (RMSE) optimization process is shown in Fig. 10. As can be seen from the figure, fitness function values also reach final convergence after 27 iterations.

In order to further verify the effectiveness of BWO algorithm, comparison tests were carried out by CS-BiLSTM model, GWO-BiLSTM model, WOA-BiLSTM model and

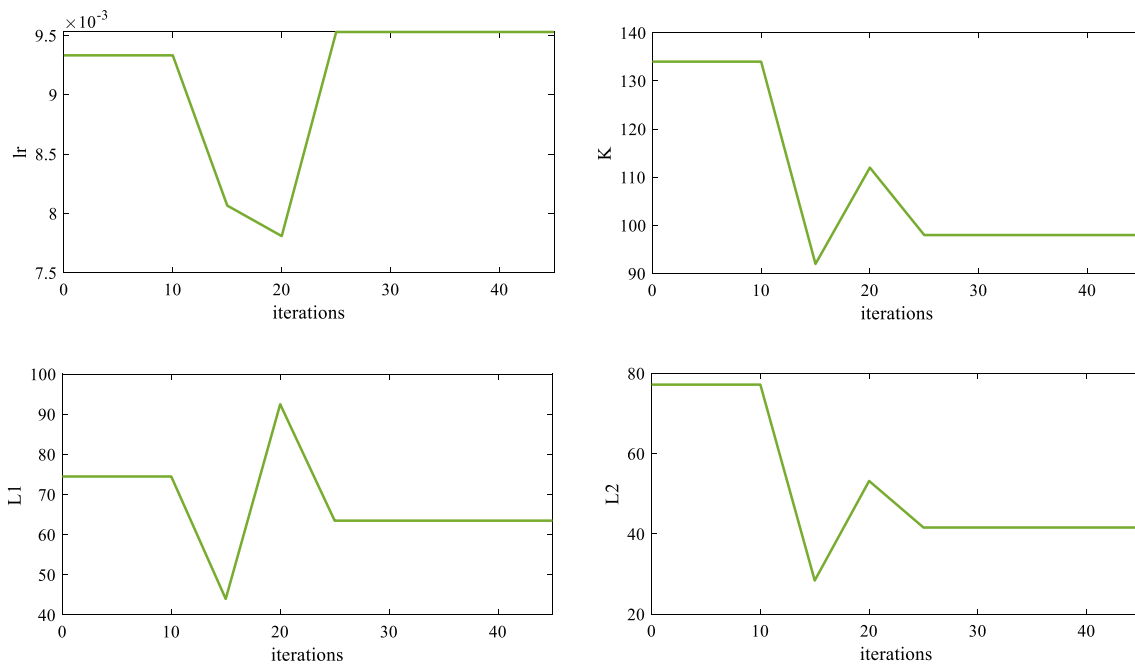
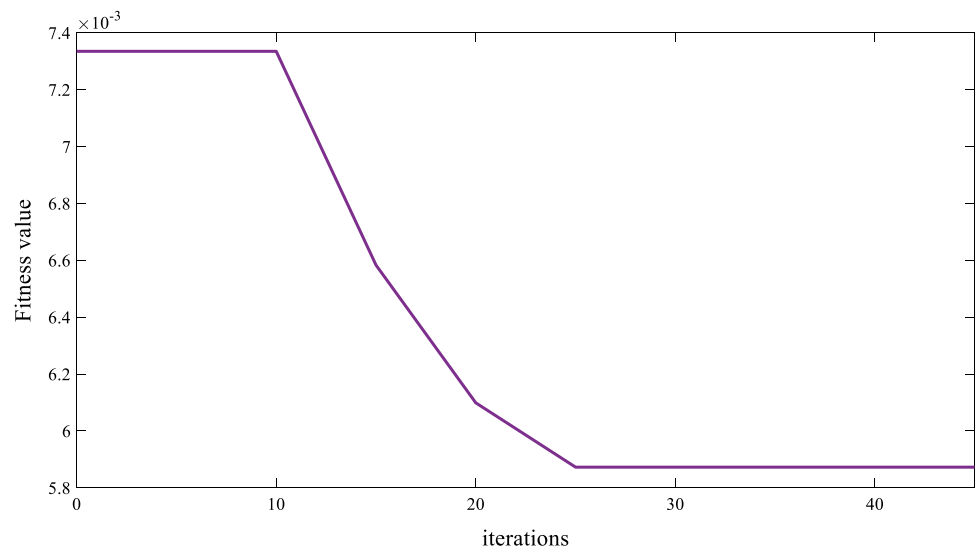


Fig. 9 Hyperparameter optimization procedure

Fig. 10 Process of fitness value change



BWO-BiLSTM under the same experimental environment. The forecast results are shown in Table 6. Boldface denotes BWO-BiLSTM.

As evident from the analysis shown in Table 6, the BWO-BiLSTM exhibits distinct advantages. In comparison to CS-BiLSTM, GWO-BiLSTM, and WOA-LSTM, RMSE demonstrates reductions of 2.11 MW, 1.94 MW, 8.60%, and 1.14 MW, respectively. Similarly, MAE shows decreases of 1.63 MW, 1.38 MW, 0.55 MW, and a significant improvement of 13.56%. Furthermore, the R^2 value attains 92.51%, indicating a strong correlation between predicted and actual

values. These results suggest that BWO is effective in hyperparameter optimization for the BiLSTM neural network, thereby enhancing prediction accuracy.

As can be seen from Fig. 11, BWO-BiLSTM has the highest accuracy in wind power prediction compared with other models. The overall prediction trend of the model is relatively close to the real value, and it still has a good degree of fitting at the peak stage of wind power, which verifies the effectiveness of the BWO-BiLSTM model in wind power prediction.

Fig. 11 Model prediction results

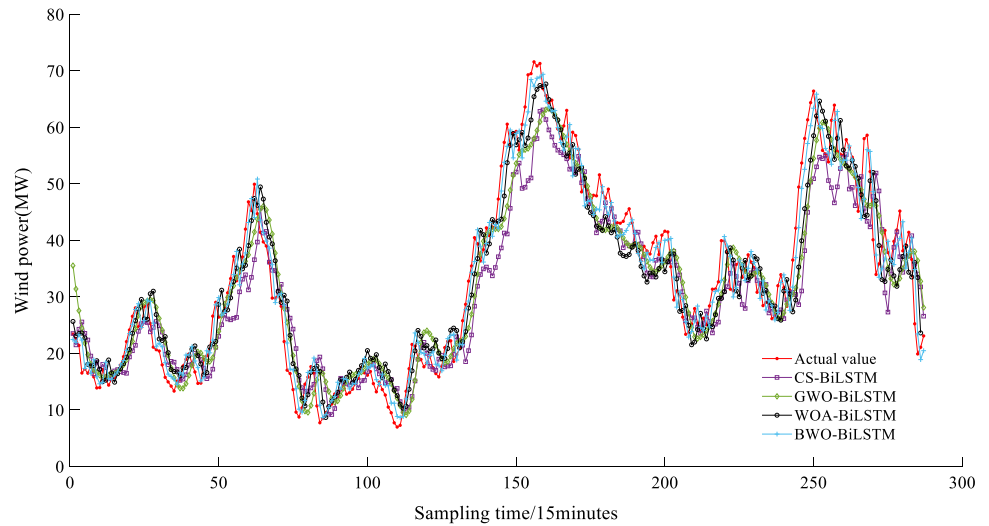


Table 6 Comparison of predicted performances

Model	MAE/MW	RMSE/MW	NMAE/%	NRMSE/%	R ² /%
CS-BiLSTM	5.06	6.12	5.27	6.38	85.98
GWO-BiLSTM	4.81	5.95	5.01	6.20	86.66
WOA-BiLSTM	3.98	5.15	4.15	5.36	89.85
BWO-BiLSTM	3.43	4.01	3.57	4.18	92.51

The comparative test results of CS-BiLSTM model, GWO-BiLSTM model, WOA-BiLSTM model and BWO-BiLSTM model, which are shown in bolditalics

Table 7 Comparative analysis of multi-model prediction results

Windy season	Model	MAE/MW	RMSE/MW	NMAE/%	NRMSE/%	R ² /%
Light wind season	SVM	5.29	6.46	5.51	6.73	84.35
	BiLSTM	4.48	5.91	4.67	6.16	87.01
	EMD-BiLSTM	3.18	3.95	3.31	4.11	93.82
	VMD-BWO-BiLSTM	2.69	3.78	2.81	3.94	94.63
	ICEEMDAN-BWO-BiLSTM	2.19	2.75	2.28	2.87	97.03
	<i>Method of this paper</i>	1.80	2.33	1.88	2.43	97.78
High wind season	SVM	2.86	3.69	2.98	3.86	78.94
	BiLSTM	2.79	3.51	2.91	3.66	81.63
	EMD-BiLSTM	1.93	2.55	2.01	2.66	90.67
	VMD-BWO-BiLSTM	1.62	2.18	1.70	2.27	93.25
	ICEEMDAN-BWO-BiLSTM	1.23	1.60	1.28	1.67	96.73
	<i>Method of this paper</i>	0.71	0.91	0.74	0.95	98.86

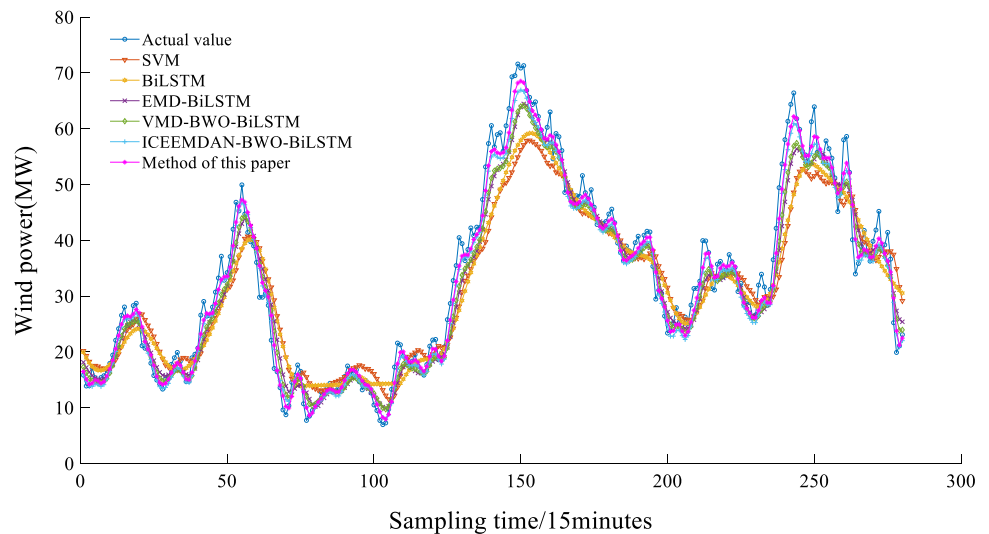
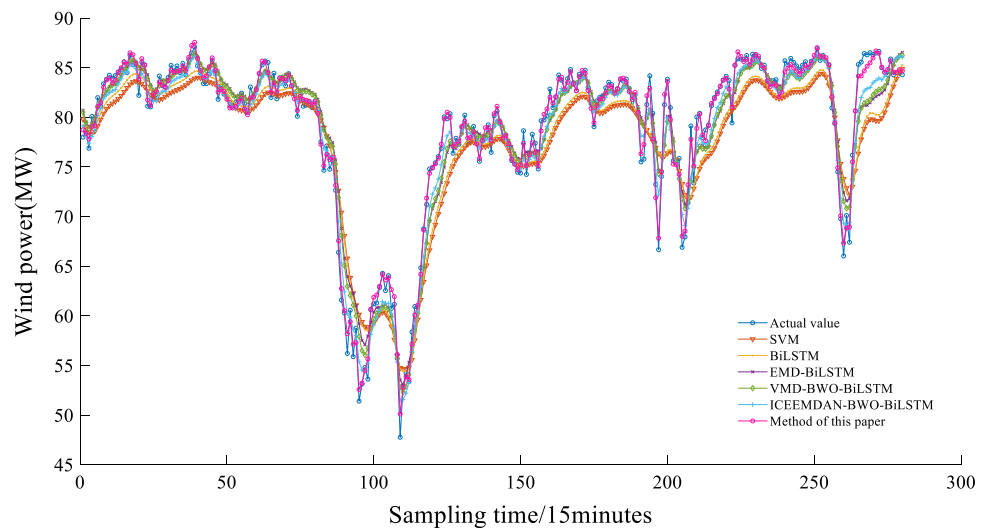
The experimental results of the proposed method are highlighted in bolditalics

5.6 Analysis of experimental results

In order to further verify the effectiveness of the proposed model, SVM, BiLSTM, EMD-BiLSTM, VMD-BWO-BiLSTM, ICEEMDAN-BWO-BiLSTM and the proposed method (ICEEMDAN-DE-ZCR-BWO-BiLSTM) are selected for comparison. The error evaluation index proposed

in 3.3 was used for evaluation, and each group of experiments was independently conducted 20 times to ensure the objectivity of the experiment and the stability of the results. The experimental results are shown in Table 7.

As can be seen from Table 7, compared with the single prediction model, the combined prediction model effectively

Fig. 12 Wind power prediction results (Light wind season)**Fig. 13** Wind power prediction results (High wind season)

improves the prediction performance. Compared with BiLSTM, RMSE and MAE decreased by 196% and 130%, respectively, and R^2 increased by 6.81% in EMD-BiLSTM during light wind season. During the gale season, RMSE and MAE decreased by 96% and 86% respectively, and R^2 increased by 9.04%. This shows that the EMD decomposition power series has better performance in improving the prediction performance, and verifies the superiority of the combined prediction model. Compared with VMD-BWO-BiLSTM, ICEEMDAN-BWO-BiLSTM reduced RMSE and MAE by 103% and 50%, and increased R^2 by 2.4%, respectively, in breeze season, ICEEMDAN-BWO-BiLSTM reduced RMSE and MAE by 39% and 58%, and increased R^2 by 3.48%, respectively, in gale season. It shows that ICEEMDAN decomposition algorithm has better performance than VMD decomposition algorithm in improving prediction performance, and has more advantages in processing time series data.

When DE-ZCR method is further used to reconstruct each power and integrate BWO-BiLSTM depth network for prediction, compared with ICEEMDAN-BWO-BiLSTM, RMSE and MAE of this method in the breeze season are reduced by 37% and 39%, respectively, and R^2 is increased by 0.75%. During the gale season, RMSE and MAE decreased by 69% and 52%, respectively, and R^2 increased by 2.13%. Compared with other models, the model proposed in this paper is superior to other methods in all evaluation indexes. In summary, the combined prediction model constructed in this paper has high prediction accuracy. Bold type indicates the method of this article.

The prediction results of each model are shown in Figs. 12 and 13, and prediction curves are used to compare the fitting effects of the five models. In the prediction of wind farm

season scenario, the fitting curves based on ICEEMDAN-DE-ZCR-BWO-BiLSTM model are fundamentally consistent with the curves of real values with the smallest error, indicating that the model has excellent prediction results.

6 Conclusion

In order to improve the accuracy of wind farm output power prediction, this paper proposes a short-term wind power prediction method based on ICEEMDAN decomposition, DE-ZCR correlation reconstruction, and BWO-BiLSTM model, and the following conclusions are obtained through experimental verification:

- (1) ICEEMDAN is used to decompose the wind power series, and DE-ZCR correlation reconstruction is used to restructure the subseries, which effectively solves the problem of high output power volatility. DE-ZCR correlation method has obvious advantages of reconstruction, eliminating the disadvantages of subjective selection, and the effectiveness of the method is proved by experiments. The time-dependent characteristics of long-time wind power are captured by BiLSTM, and the hyperparameters of the BiLSTM model are optimized by BWO, and the evaluation indexes of RMSE, MAE, and R^2 are verified to be better than other models by comparison experiments, which further illustrates the high accuracy of the model in this paper.
- (2) The ICEEMDAN-DE-ZCR-BWO-BiLSTM combined prediction model proposed in this paper has excellent prediction ability in short-term prediction, The shortcoming of this model is that it does not consider the influence of multi-factor variables and temporal and spatial characteristics on wind power prediction. Next research should increase the feature dimension of the model input for multivariate data fusion, and data from multiple wind farms can be considered to further validate the model described in this paper and improve its practical application value.

Acknowledgements The authors are grateful for the 2020LH05019 grant from Inner Mongolia Natural Science Foundation.

Author contributions All authors contributed to the conception and design of the study. The first draft of the manuscript was written by Jingxia Liu, with all authors commenting on previous versions of the manuscript, all authors read and approved the final manuscript. Jingxia Liu wrote the main manuscript text, Yanqi Wu was responsible for data collection and analysis, and Xuchu Cheng was responsible for data visualization and graph drawing. Baoli Li conducted preliminary experimental investigation, Peihong Yang prepared relevant literature research, and all authors reviewed the manuscript.

Data availability The authors will supply the relevant data in response to reasonable requests.

Declarations

Conflict of interest The authors declared that we have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

References

1. Liu L, Wang J, Li J, Wei L (2023) An online transfer learning model for wind turbine power prediction based on spatial feature construction and system-wide update. *Appl Energy* 15(340):121049
2. Liao X, Liu Z, Zheng X, Ping Z (2023) He X Wind power prediction based on periodic characteristic decomposition and multi-layer attention network. *Neurocomputing* 534:119
3. Meka R, Alaeddini A, Bhaganagar K (2021) A robust deep learning framework for short-term wind power forecast of a full-scale wind farm using atmospheric variables. *Energy* 221:119759
4. Castorrini A, Gentile S, Galdi E (2021) Bonfiglioli a increasing spatial resolution of wind resource prediction using NWP and RANS simulation. *J Wind Eng Ind Aerodyn* 210:104499
5. Yin X, Zhao X (2019) Big data driven multi-objective predictions for offshore wind farm based on machine learning algorithms. *Energy* 186:115704
6. Xue H, Jia Y, Wen P, Farkoush SG (2020) Using of improved models of gaussian processes in order to regional wind power forecasting. *J Clean Prod* 262:121391
7. Liu H, Chen C (2019) Data processing strategies in wind energy forecasting models and applications: a comprehensive review. *Appl Energy* 249:392
8. Yunus K, Thiringer T, Chen P (2016) ARIMA-based frequency-decomposed modeling of wind speed time series. *IEEE Trans Power Syst A Publ Power Eng Soc* 31(4):2546
9. Ogliairi E, Guizzoni M, Giglio A (2021) Pretto s wind power 24-h ahead forecast by an artificial neural network and an hybrid model: comparison of the predictive performance. *Renew Energy* 178:1466
10. Lipu MSH, Miah MS, Hannan MA (2021) Artificial intelligence based hybrid forecasting approaches for wind power generation: progress, challenges and prospects. *IEEE Access* 9:102460–102489
11. Mi X, Liu H, Li Y (2019) Wind speed prediction model using singular spectrum analysis, empirical mode decomposition and convolutional support vector machine. *Energy Convers Manag* 180:196
12. Jalali SMJ, Ahmadian S, Khodayar M, Khosravi A, Shafie-khah M, Nahavandi S, Catalão JPS (2022) An advanced short-term wind power forecasting framework based on the optimized deep neural network models. *Int J Elect Power Energy Syst* 141:108143
13. Yaghoubirad M, Azizi N, Farajollahi M, Ahmadi A (2023) Deep learning-based multistep ahead wind speed and power generation forecasting using direct method. *Energy Convers Manag* 281:116760
14. Chu Y, Li M, Coimbra CFM, Feng D, Wang H (2021) Intra-hour irradiance forecasting techniques for solar power integration: a review. *IScience* 24(10). <https://doi.org/10.1016/j.isci.2021.103136>
15. Hossain MA, Chakraborty RK, Elsawah S, Ryan MJ (2021) Very short-term forecasting of wind power generation using hybrid deep learning model. *J Clean Prod* 296:126564

16. Xiong B, Lou L, Meng X, Wang X, Ma H, Wang Z (2022) Short-term wind power forecasting based on attention mechanism and deep learning. *Electr Power Syst Res* 206:107776
17. Zhang J, Li H, Cheng P et al (2024) Interpretable wind power short-term power prediction model using deep graph attention network. *Energies* 17(2):384
18. Zhongda T (2021) Approach for short-term wind power prediction via kernel principal component analysis and echo state network optimized by improved particle swarm optimization algorithm. *Trans Inst Meas Control* 43(16):3647–3662
19. Shahid F, Zameer A, Muneeb M (2021) A novel genetic LSTM model for wind power forecast[J]. *Energy* 223:120069
20. Zhongda T, Hao C (2021) A novel decomposition-ensemble prediction model for ultra-short-term wind speed. *Energy Convers Manag* 248:114775
21. Zhongda T, Hao L, Feihong L (2021) A combination forecasting model of wind speed based on decomposition. *Energy Rep* 7:1217–1233
22. Marijana KP, Luka J, Nebojsa B et al (2024) Optimizing long-short-term memory models via metaheuristics for decomposition aided wind energy generation forecasting. *Artif Intell Rev* 57(3):45
23. Duan J, Wang P, Ma W, Fang S, Hou Z (2022) A novel hybrid model based on nonlinear weighted combination for short-term wind power forecasting. *Int J Electr Power Energy Syst* 134:107452
24. Tian Z (2021) Analysis and research on chaotic dynamics behaviour of wind power time series at different time scales. *J Ambient Intell Humaniz Comput* 14(2):1–25
25. Sajjad M, Khan ZA, Ullah A, Hussain T, Ullah W, Lee M, Baik SW (2020) A Novel CNN-GRU based hybrid approach for short-term residential load forecasting. *IEEE Access* 278:127799
26. Sareen K, Panigrahi BK, Shikhola T, Sharma R (2023) An imputation and decomposition algorithms based integrated approach with bidirectional LSTM neural network for wind speed prediction. *Energy* 278:800
27. Shan J, Wang H, Pei G, Zhang S, Zhou W (2022) Research on short-term power prediction of wind power generation based on WT-CABC-KELM[J]. *Energy Rep* 8:800
28. Rayi VK, Mishra SP, Naik J, Dash PK (2022) Adaptive VMD based optimized deep learning mixed kernel ELM autoencoder for single and multistep wind power forecasting. *Energy* 244:122585
29. Zhongda T, Hao C (2021) A novel decomposition-ensemble prediction model for ultra-short-term wind speed. *Energy Convers Manag* 248:114775
30. Xiwen C, Xiaoyu Y, Dongxiao N (2024) The ultra-short-term wind power point-interval forecasting model based on improved variational mode decomposition and bidirectional gated recurrent unit improved by improved sparrow search algorithm and attention mechanism. *Energy* 288:129714
31. Li G, Pan Z, Qi Z, Wang H, Wang T, Zhao Y, Zhang Y, Li G, Wang P (2023) Hybrid forecasting system considering the influence of seasonal factors under energy sustainable development goals. *Measurement* 211:112607
32. Bommidi BS, Teeparthi K (2023) Kosana V hybrid wind speed forecasting using ICEEMDAN and transformer model with novel loss function. *Energy* 265:126383
33. Zhang Y, Li C, Jiang Y, Zhao R, Yan K, Wang W (2023) A hybrid model combining mode decomposition and deep learning algorithms for detecting TP in urban sewer networks. *Appl Energy* 333:120600
34. Yue W, Liu Q, Ruan Y, Qian F, Meng H (2022) A prediction approach with mode decomposition-recombination technique for short-term load forecasting. *Sustain Cit Soc* 85:104034
35. Tian Z, Li S, Wang Y (2020) A prediction approach using ensemble empirical mode decomposition-permutation entropy and regularized extreme learning machine for short-term wind speed. *Wind Energy* 23(2):177–206
36. Ghimire S, Deo RC, Casillas-Pérez D (2022) Salcedo-sanz s improved complete ensemble empirical mode decomposition with adaptive noise deep residual model for short-term multi-step solar radiation prediction. *Renew Energy* 190:408
37. Mostafa R, Mohammad RA (2019) Hamed A application of dispersion entropy to status characterization of rotary machines. *J Sound Vib* 438:291
38. Zhong C, Li G (2022) Meng Z Beluga whale optimization: a novel nature-inspired metaheuristic algorithm[J]. *Knowl Based Syst* 251:109215
39. Wu X, Jiang S, Lai CS, Zhao Z, Lai LL (2022) Short-term wind power prediction based on data decomposition and combined deep neural network. *Energies* 15(18):6734
40. Li Z, Luo X, Liu M, Cao X, Du S (2022) Sun H Short-term prediction of the power of a new wind turbine based on IAO-LSTM. *Energy Rep* 8:9025
41. Boubaker S, Benganem M, Mellit A, Lefza A, Kahouli O, Kolsi L (2021) Deep neural networks for predicting solar radiation at hail region, Saudi Arabia. *IEEE Access* 9:36719
42. Zhongda T, Hao C (2021) Multi-step short-term wind speed prediction based on integrated multi-model fusion. *Appl Energy* 298:117248
43. Qingcheng L, Huiling C, Hanwei L et al (2024) A novel ultra-short-term wind power prediction model jointly driven by multiple algorithm optimization and adaptive selection. *Energy* 288:129724

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.