WEP Warwick Evans Publishing

International Journal of Mechanical and Electrical Engineering

ISSN: 3005-9615 (Print), ISSN: 3005-7132 (Online) | Volume 2, Number 3, Year 2024
DOI: https://doi.org/10.62051/ijmee.v2n3.04
Journal homepage: https://wepub.org/index.php/IJMEE/index

# Short-term Photovoltaic Power Prediction based on ICEEMDAN and Optimized Deep Hybrid Kernel Extreme Learning Machine

Hao Yan

School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, China

## ABSTRACT

Aiming at the problem of low prediction accuracy for photovoltaic power generation due to the strong randomness and volatility, a prediction model based on Improved Complete Ensemble Empirical Mode Decomposition With Adaptive Noise (ICEEMDAN) and Beluga Whale Optimization Deep Hybrid Kernel Extreme Learning Machine (BWO-DHKELM) is proposed. Firstly, the historical data are analyzed by Pearson Correlation Coefficient, and the meteorological data with high correlation are obtained as the input features of the prediction model. Secondly, the PV power is decomposed by ICEEMDAN to reduce its volatility. Then, DHKELM is constructed for each subsequence and several parameters of the model are optimized by BWO. Finally, the predicted values of each subsequence are summed to obtain the final prediction results. The effectiveness and superiority of the proposed model are verified by using real data from a PV plant in Jiangsu, China as an example.

## KEYWORDS

Photovoltaic Power Generation; ICEEMDAN; Beluga Whale Optimization; Deep Hybrid Kernel Extreme Learning Machine.

## 1. INTRODUCTION

In recent years, with the continuous promotion of the "dual-carbon" goal, the development of renewable energy in China has been rapid. As one of the most popular renewable energy sources, photovoltaic (PV) power generation is developing rapidly[0-2]. However, PV power generation is volatile and stochastic, and a high percentage of PV access in the grid will affect the scheduling and operation of the grid[3]. Improving the accuracy of PV power prediction is important for operation control in the grid[4].

There are three main methods for short-term PV power prediction: the physical method, the statistical method and the combined method[5]. The physical method mainly predicts the power generation through the principle of solar radiation, which requires accurate geographic information and equipment information of the PV power plant and has the weak ability of anti-interference. In the statistical method, it is mainly based on PV power generation prediction model by machine learning algorithm and deep learning algorithm, which realizes the prediction by mining the nonlinear mapping relationship between meteorological elements and PV power, and has strict requirements for historical data[6]. Currently, prediction models such as Artificial Neural Networks and Support Vector Regression (SVR) are well known[7]. However, these early shallow neural networks with a single hidden layer have limited ability to mine complex nonlinear mappings containing rich features. Extreme Learning Machine (ELM), as a computationally efficient single hidden layer forward feedback neural network with strong generalization ability, also suffers from the above problem[8]. DHKELM improves ELM by adding hybrid kernel function and autoencoder, respectively, which effectively improves the problem of randomly generating initial weights and thresholds of ELM and strengthens the depth of the model structure to effectively solve the limitations of. machine learning. However, the model also has the problem to determine the parameters[9]. Heuristic algorithms are the research hotspots for solving the parameter selection problem, and common algorithms include particle swarm algorithm[10], differential evolution algorithm[11], arithmetic optimization algorithm[12], etc. BWO is a new intelligent optimization algorithm which was firstly proposed in 2022[13], and this algorithm has a greater advantage in the aspects of global search ability and local exploitation ability, and is able to optimize and solve the complex problems. In this paper, we propose to use BWO to optimize the parameters of DHKELM.

A single prediction model has its own limitations, so the combination of prediction models based on multiple methods has become a hot research topic for scholars, and the combination of prediction models can give full play to the advantages of different prediction methods to achieve better prediction results. Among them, the decomposition method is widely used because of its effectiveness[14]. The researchers used Empirical Mode Decomposition (EMD) to decompose the original PV power and then establish LSSVM prediction model for each sub-sequence, and finally reconstructed to get the final predicted power, which has a higher prediction accuracy than wavelet decomposition, but EMD decomposition will cause modal mixing phenomenon[15]. Ensemble Empirical Mode Decomposition (EEMD) effectively solves the modal mixing phenomenon of EMD, but its white noise is not easy to eliminate[16]. The researchers uses the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) to decompose the photovoltaic power, and then utilizes the error-compensated LSSVM model to predict the components, this decomposition effectively solves the problem of modal mixing phenomenon after adding noise to the signal, but there are still residual noise and pseudo-modal problems[17]. Compared with the above decomposition methods, the ICEEMDAN algorithm adopts local mean estimation, and selects the white noise components decomposed by EMD to determine the final decomposition subsequence, which can effectively reduce the influence of white noise on the decomposition[18].

In this paper, a combined ICEEMDAN-BWO-DHKELM prediction model is proposed. First, the correlation coefficient method is utilized to obtain the input features by correlating the meteorological factors and PV power. Secondly, the original PV power sequence is decomposed by ICEEMDAN to obtain subsequences with different frequencies. Then, the prediction model of DHKELM is constructed for each subsequence, and BWO is utilized to optimize the parameters which are difficult to determine in the DHKELM model. Finally, the predicted values of each subsequence are reconstructed to obtain the final power prediction.

## 2. FEATURE SELECTION

Pearson Correlation Coefficient (PCC) is an indicator to show the degree of correlation[19]. In order to analyze the correlation between different meteorological factors and PV power, this paper chooses to use the Pearson correlation coefficient to get the meteorological factors with high correlation, and the coefficient can be obtained using the equation:

$$P = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{1}$$

where $P$ is the Pearson correlation coefficient, $x_i$ and $y_i$ are the number $i$ point in the two series, $\bar{x}$ and $\bar{y}$ are the average value of the respective series, and $n$ is the number of data points in the series.

The experimental data in this paper is from a PV power station in Jiangsu, China, and the correlation analysis of meteorological factors and PV power generation in the data is shown in Table 1.

**Table 1.** PCC between power and factors

| Factor | Total radiation | Direct radiation | Scattered radiation | Component temperature | Ambient temperature | Barometric pressure | Relative humidity |
|---|---|---|---|---|---|---|---|
| Correlation Coefficient | 0.95 | 0.77 | 0.65 | 0.61 | 0.42 | 0.17 | -0.21 |

As can be illustrated in Table 1, the correlation of total radiation, direct radiation, scattered radiation and component temperature is higher; followed by ambient temperature, however, barometric pressure and relative humidity have less influence. Therefore, in this paper, total radiation, direct radiation, scattered radiation, component temperature and ambient temperature are selected as input features.

# 3. PRINCIPLES RELATED TO PREDICTIVE MODEL

## 3.1. ICEEMDAN

Due to the strong volatility of PV power, its prediction after decomposition can improve the accuracy. CEEMDAN is based on EMD by adding adaptive Gaussian white noise at each stage to reduce the reconstruction error. The proposal of ICEEMDAN effectively solves the pseudo-modal problem of CEEMDAN. The specific steps of ICEEMDAN are as follows:

1) Add a set of noise $W_i(t)$, and add the modal components obtained from its EMD to the original signal $X(t)$, construct the sequence $X_i^1(t) = X(t) + \gamma_0 E_1\left(W_i(t)\right)$, then obtain the first residual component as:

$$r_1(t) = \left\langle N(X_i^1(t)) \right\rangle \tag{2}$$

where $E_j(\cdot)$ is the jth Intrinsic Mode Function (IMF) component generated by EMD, $\gamma_0$ is the ratio of the signal-to-noise ratio to the standard deviation of this noise component, $N(\cdot)$ is the local mean of the signal, and $\langle \cdot \rangle$ is the mean of the signal.

2) Calculate the first modal component IMF1:

$$IMF_1(t) = X(t) - r_1(t) \tag{3}$$

3) construct a sequence $X_i^2(t) = r_1(t) + \gamma_1 E_2\left(W_i(t)\right)$, the second modal component $\text{IMF}_2$ is calculated:

$$IMF_2(t) = r_1(t) - r_2(t) = r_1(t) - \left\langle N\left(X_i^2(t)\right) \right\rangle \tag{4}$$

4) Repeat step 3) to obtain the *k*th residual component and the *k*th modal component $\text{IMF}_k$ as:

$$r_k(t) = \left\langle N\left(r_{k-1}(t) + \gamma_{k-1} E_k W_i(t)\right) \right\rangle \tag{5}$$

$$IMF_k(t) = r_{k-1}(t) - r_k(t) \tag{6}$$

## 3.2. Deep Hybrid Kernel Extreme Learning Machine (DHKELM)

### 3.2.1. Hybrid Kernel Extreme Learning Machine

Kernel Extreme Learning Machine (KELM) is a single hidden layer feedforward neural network based on kernel function[20], the feedforward neural network model outputs is achieved by the following equation:

$$f(\boldsymbol{x}) = h(\boldsymbol{x})\boldsymbol{\beta} = \boldsymbol{H\beta} \tag{7}$$

where $y = f(\boldsymbol{x})$ is the output of the network, $h(\boldsymbol{x})$ is the input function of the hidden layer, $\boldsymbol{x}$ is the input vector, $\boldsymbol{H}$ is the feature mapping matrix, $\boldsymbol{\beta}$ is the output weight vector connecting the hidden layer and the output layer. The powerful nonlinear mapping ability of the kernel function overcomes the problem of random selection on parameter for ELM, and can better improve the ability to extend of the model. The output function of the KELM can be represented as:

$$y(x) = h(x)\beta = h(x)H^T\left(\frac{I}{C} + HH^T\right)^{-1} T \tag{8}$$

$$\begin{cases} HH^T = \Omega_{KELM} \\ \Omega_{KELM_{i,j}} = h(x_i)h(x_j) = \mathrm{K}(x_i, x_j) \end{cases} \tag{9}$$

where $I$ is the unit matrix, $C$ is the penalty parameter, $\Omega_{KELM}$ is the kernel function.

Therefore, the model function of KELM is:

$$y(x) = \begin{bmatrix} \mathrm{K}(x, x_1) \\ \vdots \\ \mathrm{K}(x, x_N) \end{bmatrix} \left(\frac{I}{C} + \Omega_{KELM}\right)^{-1} T \tag{10}$$

In order to effectively improve the ability that local exploitation and global search of the kernel function, a method of constructing a hybrid kernel function by the linear weight is proposed to obtain the Hybrid Kernel Extreme Learning Machine (HKELM)[21], whose kernel function has the following equations:

$$\begin{cases} \mathrm{K}(x, x_i) = \alpha \cdot \exp\left(-\dfrac{\|x, x_i\|^2}{2\sigma^2}\right) + \beta \cdot \left((x \cdot x_i) + q\right)^p \\ \beta = 1 - \alpha, \alpha \in [0,1] \end{cases} \tag{11}$$

where $2\sigma^2$ is the parameter used to control the radial range, $q$ and $p$ are the constant and exponential parameters of the polynomial kernel function, respectively, $\alpha$ is the weight coefficient of the hybrid function.

### 3.2.2. Autoencoder-improved Deep Hybrid Kernel Extreme Learning Machine

Compared to other deep learning, multi-layer ELMs do not need to be fine-tuned several times and have the advantage of fast training[22]. Therefore, the deep hybrid kernel extreme learning machine (DHKELM) based on autoencoder is proposed. The model has an input layer, an hidden layer, a kernel mapping layer, a fully connected layer and an output layer. The autoencoder needs to satisfy that the target input is equal to the output for better extraction of nonlinear features.

In this model, when the number of training samples is greater than the hidden layer nodes, the following formula is used to calculate $\beta$:

$$\beta = \left(\frac{I}{\lambda} + H^T H\right)^{-1} H^T T \tag{12}$$

DHKELM contains several implicit layer nodes[23].The DHKELM model is shown in Fig. 1.

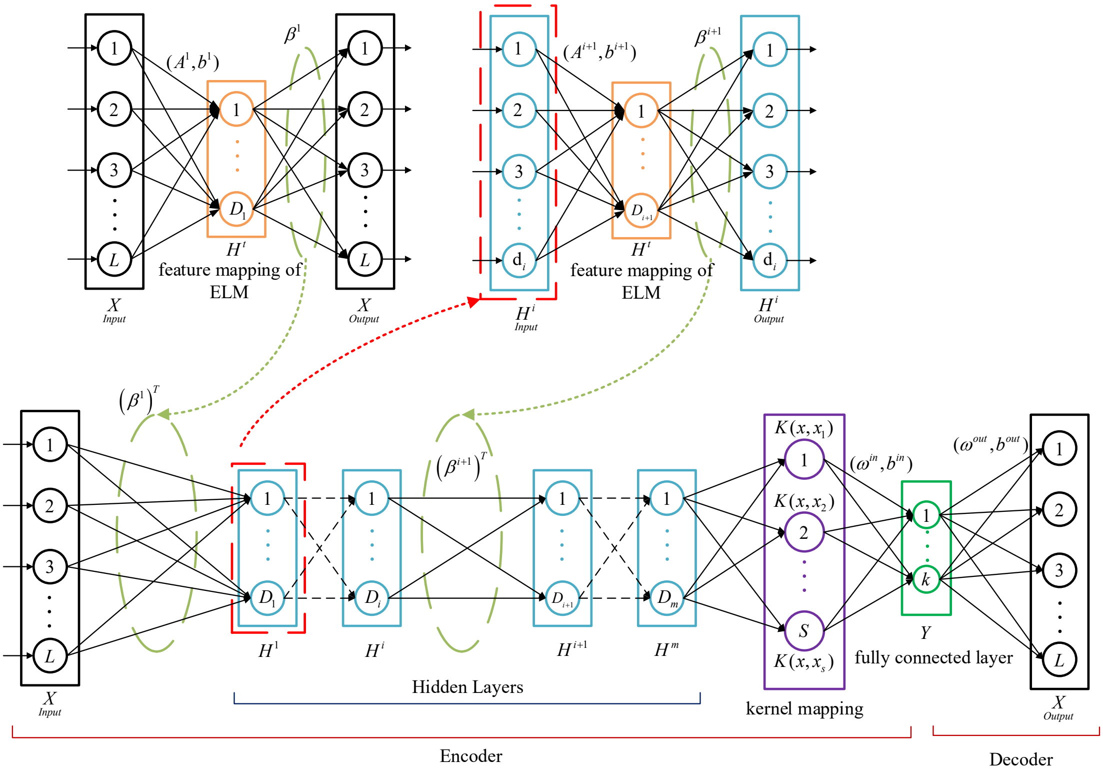


**Fig 1.** Model of DHKELM

The DHKELM model is composed of an encoder and a decoder. The weights of hidden layer $\{\beta^i\}_{i=1}^{m}$ are determined as follows:

1) The orthogonal parameters between the input layer and the hidden layer that $A^i$ and $b^i$ are arbitrarily selected.

2) Find the state matrix of hidden layer $H^t$ and select the activation function $g(x)$:

$$H^t = \left(\frac{I}{\lambda} + H^T H\right)^{-1} T \tag{13}$$

$$g(x) = \frac{1}{1+e^{-x}} \tag{14}$$

3) Find the weight of the $i+1$th hidden layer $\beta^{i+1}$:

$$\beta^{i+1} = \left(\frac{I}{\lambda} + \left(H^t\right)^T H^t\right)^{-1} (Ht)^T H^i \tag{15}$$

The stacked hidden layers form a hierarchy:

$$\begin{cases} H^0 = X \\ H^{i+1} = g\left(\left(\beta^{i+1}\right)^T H^i\right), \quad i = 0,1,\cdots,m-1 \end{cases} \tag{16}$$

The output of the fully connected layer $Y$ is:

$$Y = g\left(\omega^{in} H^{m} + b^{in}\right) \tag{17}$$

Thus , the parameters between $H^m$ and $Y$ are obtained. The fully connected layer of network is taken up by the decoder to get the reconstructed data $\underset{Output}{X}$ as:

$$\begin{aligned} \underset{Output}{X} &= g\left(\omega^{out} Y + b^{out}\right) \\ &= g\left(\omega^{out} g\left(\omega^{in} H^{m} + b^{in}\right) + b^{out}\right) \end{aligned} \tag{18}$$

The objective function of the model is to minimize the error as follows:

$$\underset{\omega^{in}, b^{in}, \omega^{out}, b^{out}}{argmin} \parallel \underset{Output}{X} - X \parallel_2^2 \tag{19}$$

where $\omega^{in}$、$b^{in}$、$\omega^{out}$、$b^{out}$ are the parameters before and after the fully connected layer.

Both the hidden layer weights and the state matrix of DHKELM are computationally accessible, and after initialization only the weights between the last two layers need to be trained to take values. This drastically reduces the parameters in the deep encoder and achieves an increase in efficiency.

### 3.3. Beluga Whale Optimization (BWO)

BWO simulates three behaviours of beluga whales such as swimming, preying and whale fall. and proposes the corresponding three position updating processes. Similar to other swarm intelligence optimization algorithms, BWO contains an exploration phase of global search and a exploitation phase of local search. The position vector of each beluga whale represents the alternate solution for parameter search,

In BWO the position of each beluga whale changes as the optimization algorithm proceeds, and the population position of the beluga whales as search agents can be represented by the following matrix:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d} \end{bmatrix} \tag{20}$$

where $n$ is the population size of beluga whales, $d$ is the dimension of the problem variable. For all beluga whales, the corresponding fitness values are given by the following equation:

$$F_X = \begin{bmatrix} f\left(x_{1,1}, x_{1,2}, \ldots, x_{1,d}\right) \\ f\left(x_{2,1}, x_{2,2}, \ldots, x_{2,d}\right) \\ \vdots \\ f\left(x_{n,1}, x_{n,2}, \ldots, x_{n,d}\right) \end{bmatrix} \tag{21}$$

The balance factor $B_f$ that determines the transformation between the exploration and exploitation phases of BWO is given by the following equation:

$$B_f = B_0\left(1 - T / 2T_{\max}\right) \tag{22}$$

where $T$ is the current iteration, $T_{\max}$ is the maximum iteration number, $B_0$ is a random number between (0,1). When the balance factor $B_f > 0.5$ the exploration phase happens; while the exploitation phase happens when $B_f \le 0.5$ . It can be seen from the above equation, with the increase in the number of iterations, the algorithm gradually changes from the exploration phase to the exploitation phase.

Algorithm in the exploration phase, from the swimming behaviour of two mirror image or the same direction of the beluga whales can be obtained from the beluga whale's position update formula is:

$$\begin{cases} X_{i,j}^{T+1} = X_{i,p_j}^{T} + \left(X_{r,p_1}^{T} - X_{i,p_j}^{T}\right)(1+r_1)\sin(2\pi r_2), j = 2k \\ X_{i,j}^{T+1} = X_{i,p_j}^{T} + \left(X_{r,p_1}^{T} - X_{i,p_j}^{T}\right)(1+r_1)\cos(2\pi r_2), j = 2k+1 \end{cases} \tag{23}$$

where $X_{i,j}^{T+1}$ is the new position of the *i*th beluga whale in the *j*th dimension, $p_j\,(j = 1,2,...,d)$ is a random number selected from *d* dimensions, $X_{i,p_j}^{T}$ is the current position of the *i*th beluga whale in the $p_j$th dimension, $X_{r,p_1}^{T}$ is the current position of the *r*th beluga whale (where *r* is a random number), $r_1$ and $r_2$ are random numbers between (0,1).

The formula for updating the beluga whales position of BWO in the exploitation phase is:

$$X_i^{T+1} = r_3 X_{Best}^{T} - r_4 X_i^{T} + C_1 \cdot L_F \cdot \left(X_r^{T} - X_i^{T}\right) \tag{24}$$

where $X_{Best}^{T}$ is the optimal position in the beluga whale population, $C_1 = 2r_4(1 - T / T_{\max})$ is the random jump strength used to control the intensity of Levy flight. $L_F$ is the Levy flight function:

$$L_F = 0.05 \times \frac{u \times \sigma}{|v|^{1/\beta}} \tag{25}$$

$$\sigma = \left( \frac{\Gamma(1+\beta) \times \sin(\pi\beta / 2)}{\Gamma((1+\beta)/2) \times \beta \times 2^{(\beta-1)/2}} \right)^{1/\beta} \tag{26}$$

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

In the process of migration and preying, beluga whales will be threatened by killer whales, polar bears and humans, etc. Some beluga whales will be attacked and fall into the deep sea or lost from the group and separated from the population is called whale fall. In the algorithm, whale fall are added in each iteration, in order to ensure the stability of the population size, the position update formula is affected by the whale fall step size as:

$$X_i^{T+1} = r_5 X_i^{T} - r_6 X_r^{T} + r_7 X_{step} \tag{27}$$

where $r_5$, $r_6$ and $r_7$ are random numbers between (0,1), $X_{step}$ is the step size of whale fall which is obtained by the following equation:

$$X_{step} = (u_b - l_b)\exp(-C_2 T / T_{\max}) \tag{28}$$

where $C_2 = 2W_f \times n$ is the step factor which is related to the probability of whale fall and the population size, $u_b$ and $l_b$ are the upper and lower bounds of the variables respectively. In the model, the probability of whale fall $W_f$ varies linearly:

$$W_f = 0.1 - 0.05T / T_{\max} \tag{29}$$

The above equation represents a gradual decrease from 0.1 to 0.05, which means that the closer the beluga whale is to the food, the fewer the whale fall happens.

### 3.4. Beluga Whale Optimization Deep Hybrid Kernel Extreme Learning Machine (BWO-DHKELM)

In the DHKELM model, the input weights and bias quadrature parameters between the input and hidden layers are random, and the kernel parameters and weights of the kernel function also affect the accuracy of the model. Since DHKELM has more hidden layers than ELM, a large number of parameter values will affect the generalization ability of the model, mainly including the kernel parameters, kernel function weights, the number of hidden layers, the number of neurons in the hidden layers, the regularization coefficients, the penalty coefficients, and the weights of the last two layers.

In this paper, the prediction model of BWO-DHKELM is built, and the overall process of optimization is shown in Fig. 2:

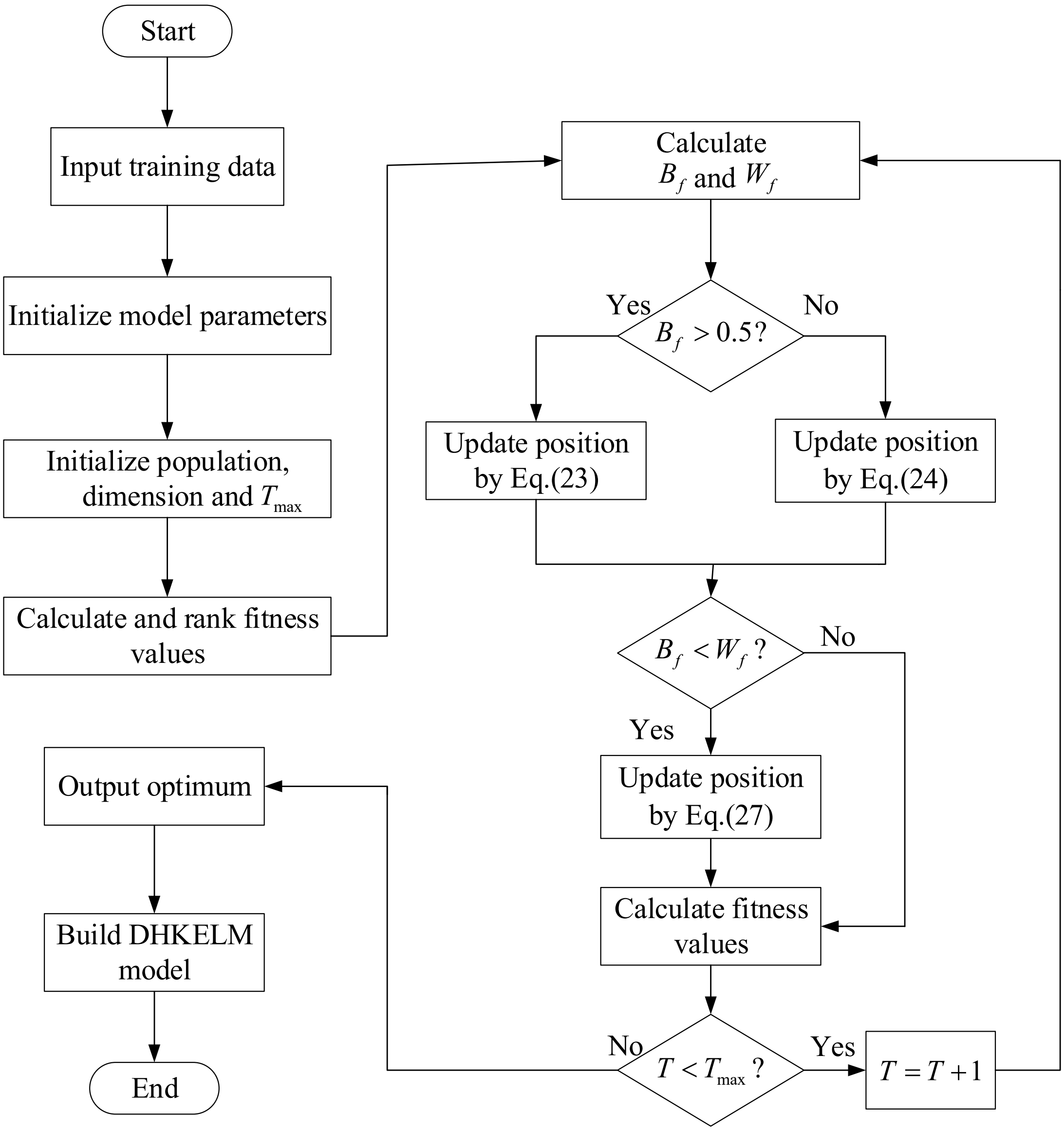


**Fig 2.** Flow chart of BWO-DHKELM

1) Initialize the parameters and population, determine parameters such as population size and maximum number of iterations.

2) Calculate the individual beluga whale fitness value, and the fitness function of the training process is the root mean square error.

3) Update the position of beluga whales according to the value of $B_f$. The value $B_f$ is calculated first, if $B_f > 0.5$, the beluga whale updates its position according to equation (23); if $B_f \leq 0.5$, the beluga whale updates its position according to equation (24).

4) Perform the whale fall to update the position and fitness value. Calculate the value of $W_f$, and compare it with the value of $B_f$. If $B_f < W_f$, the beluga whale update position by equation (27) and calculates the fitness value.

5) Arrange the fitness value and judge the iteration condition, if the condition is satisfied then output the optimal solution; otherwise return to 3).

6) The optimal results are input into the DHKELM model for short-term PV power prediction.

### 3.5. Short-term Photovoltaic Power Prediction based on ICEEMDAN-BWO-DHKELM

The combined prediction model proposed in this paper is the ICEEMDAN-BWO-DHKELM PV power prediction model. Fig.3 is the flowchart of the combined prediction model with the following steps:

1) Data preprocess, divide the data into training set and test set;

2) Finish correlation analysis of meteorological factors and photovoltaic power using the Pearson correlation coefficient and get the input features;

3) Decomposition of PV power using ICEEMDAN to obtain subsequences of different frequencies;

4) Establish the DHKELM model to train and predict the sequences obtained from the above decomposition;

(5) Optimize the parameters of the above model using BWO to build the BWO-DHKELM model;

6) Superimpose the predicted values obtained from the above prediction models separately to obtain the final power prediction.

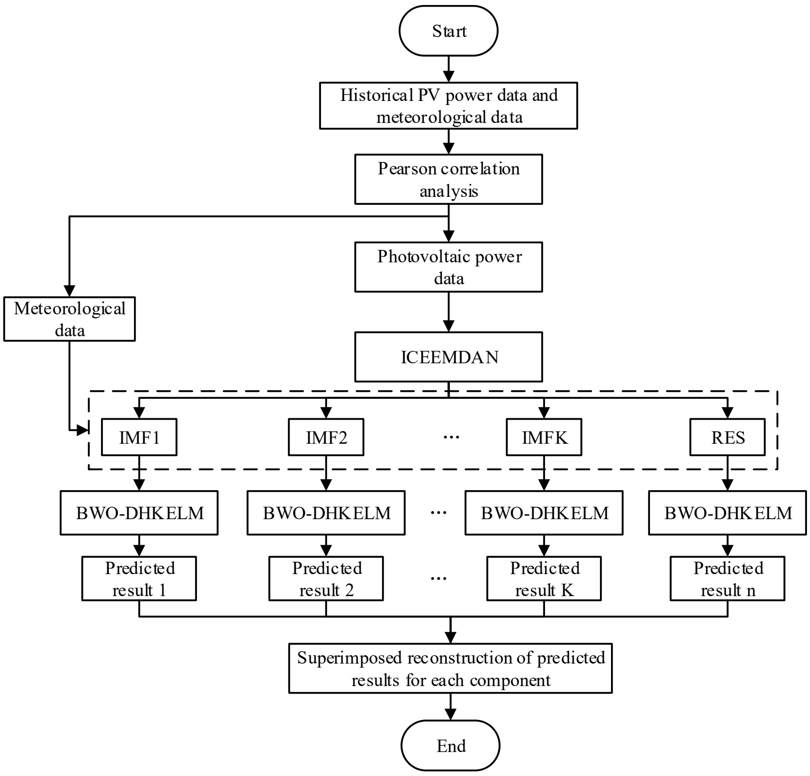


**Fig 3.** Flow chart of ICEEMDAN-BWO-DHKELM

# 4. EXAMPLE ANALYSIS

## 4.1. Evaluation Methodology

In this paper, root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) are chosen as the evaluation indexes, and the expressions are:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{p}_i - p_i\right)^2} \tag{30}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{p}_i - p_i\right| \tag{31}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|\hat{p}_i - p_i\right|}{p_i} \times 100\% \tag{32}$$

where $p_i$ is the actual power, $\hat{p}_i$ is the predicted power.

## 4.2. Data Selection and Process

Take the measured data of a PV power station in Gaoyou, Jiangsu Province, China as an example, the power generation and meteorological data from August 1st 2019 to September 30th 2019 are selected. The PV power is generated in the light moment, so the data is selected in the time period from 07:00 to 19:00, and the sampling interval is 15 min, with a total of 48 sampling points in a day. The data are divided according to the ratio of 4:1 between the training set and the test set.

Using ICEEMDAN to decompose the historical PV power data to reduce its volatility, after decomposition to obtain a number of sub-sequences with a fixed frequency, the results are shown in Figure 4. From Fig. 4, it can be seen that the PV power series is decomposed into IMF1~IMF9 and residual (RES), IMF1 is a high-frequency component, which reflects the local characteristics of the PV power, and with the gradual reduction on the frequency of the sub-sequence, the curve is gradually smoothed and the fluctuation is gradually reduced, and these low-frequency components characterize the overall trend of the PV power. In this paper, different types of subsequences are predicted separately to improve the prediction accuracy.

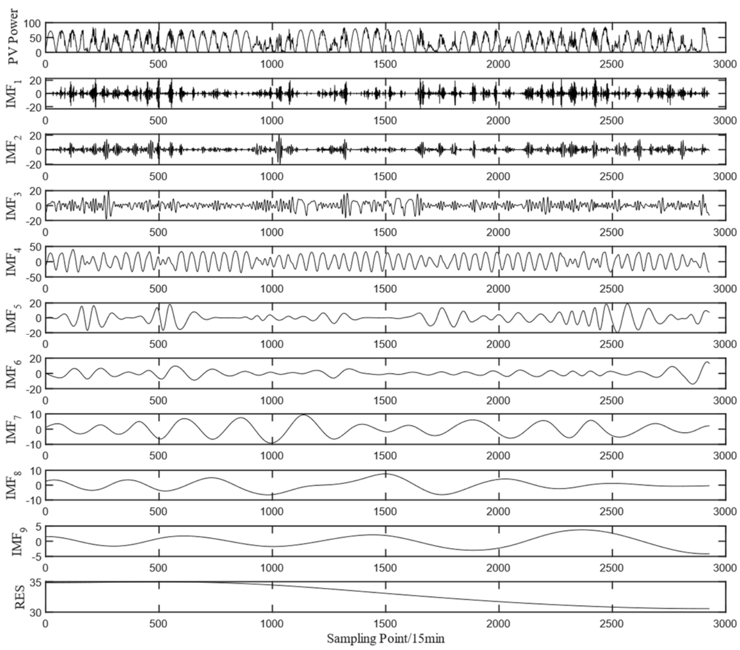


**Fig 4.** Decomposition results of PV power

## 4.3. Analysis of Prediction Results

The validity of the decomposition in the model of this paper and the superiority of the proposed model are verified by comparing the experiment with the models including SVR, ELM, HKELM, DHKELM, CEEMDAN-DHKELM, and ICEEMDAN-DHKELM. The parameters of each model are set as follows: the penalty factor $\gamma$ and kernel parameters $\eta$ in SVR model are obtained by grid-search method between $[2^{-9}, 2^{9}]$; the number of hidden layers of ELM is 1 and the number of neurons is 100; HKELM has a hybrid kernel function of RBF and Poly, the weight coefficient is 0.5, and the regularization coefficient $C = 100$, parameters of hybrid kernel are all 100; for the unoptimized DHKELM, the kernel function is a hybrid kernel function of RBF and Poly, the weight coefficient is 0.5, the kernel parameter of RBF is 100, the kernel parameter of Poly is 100 and 0.5, the number of hidden layers is 5, the number of neurons is 50, the regularization coefficient is 100, and the penalty coefficient is 100; for the BWO, the number of population $n = 20$, the maximum number of iterations $T_{\max} = 100$, and the whale fall coefficients are 0.1 and 0.05.

The dataset is normalized before prediction. The prediction results under three different weather types are compared and analyzed. Figures 5-7 show the results of the different prediction models for each weather type.

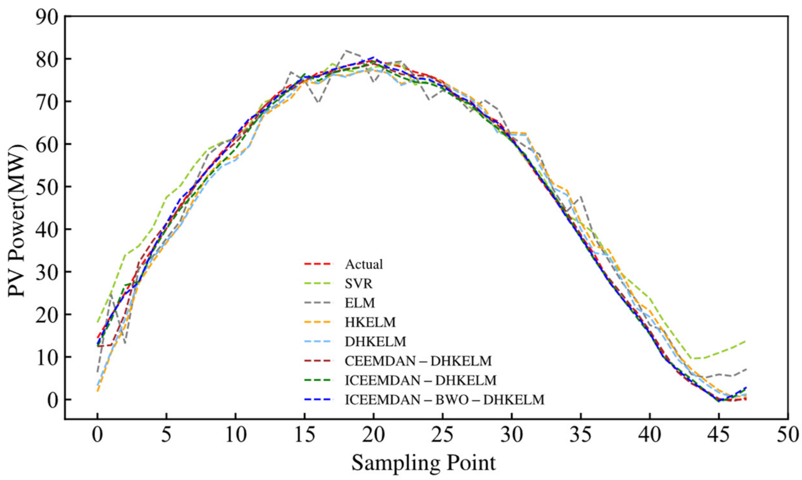


**Fig 5.** Prediction results on sunny day

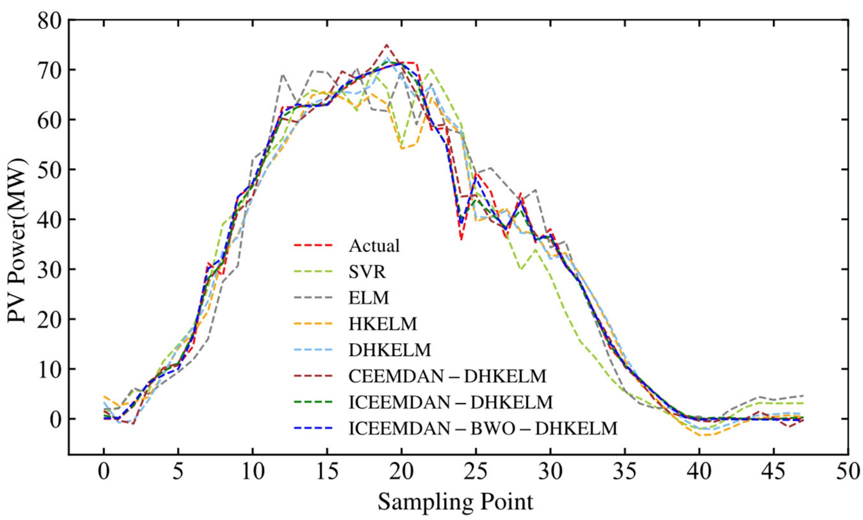


**Fig 6.** Prediction results on cloudy day

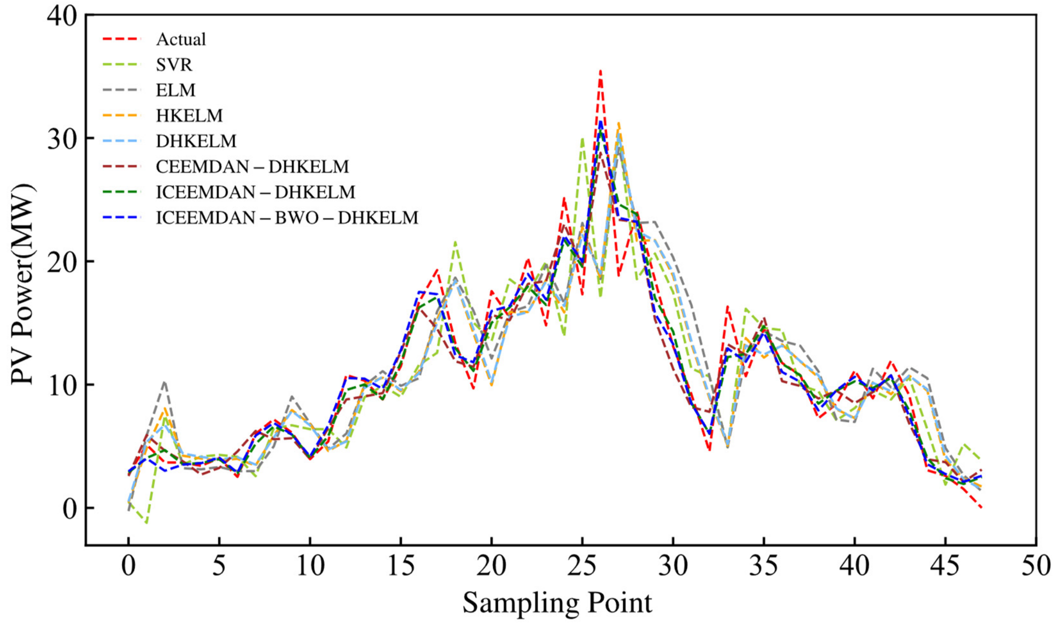


**Fig 7.** Prediction results on rainy day

**Table 2.** Comparison of predictive errors of each model in different weathers

| weather | Prediction model | RMSE | MAE | MAPE |
|---|---|---|---|---|
| Clear Weather | SVR | 4.9894 | 3.6643 | 11.1432 |
| | ELM | 4.1161 | 3.2839 | 10.0382 |
| | HKELM | 4.0620 | 3.2353 | 9.4405 |
| | DHKELM | 3.6140 | 2.9007 | 9.1440 |
| | CEEMDAN-DHKELM | 1.4587 | 1.1687 | 2.9835 |
| | ICEEMDAN-DHKELM | 1.3997 | 0.8683 | 2.7804 |
| | ICEEMDAN-BWO-DHKELM | 0.9128 | 0.6660 | 1.6521 |
| cloudy | SVR | 6.6696 | 4.7435 | 18.5632 |
| | ELM | 6.3012 | 4.4576 | 14.8269 |
| | HKELM | 6.1275 | 4.2643 | 14.1965 |
| | DHKELM | 5.0384 | 3.4576 | 13.1015 |
| | CEEMDAN-DHKELM | 2.5397 | 1.7837 | 5.7453 |
| | ICEEMDAN-DHKELM | 1.7428 | 1.0366 | 4.0412 |
| | ICEEMDAN-BWO-DHKELM | 1.2773 | 0.8134 | 3.2761 |
| Rainy day | SVR | 5.4884 | 4.0119 | 33.8696 |
| | ELM | 5.0429 | 3.7968 | 29.2467 |
| | HKELM | 4.8601 | 3.5467 | 28.3150 |
| | DHKELM | 4.7316 | 3.4532 | 28.1592 |
| | CEEMDAN-DHKELM | 2.1074 | 1.6440 | 13.3965 |
| | ICEEMDAN-DHKELM | 1.7152 | 1.1824 | 9.5095 |
| | ICEEMDAN-BWO-DHKELM | 1.5646 | 1.1503 | 8.9802 |

As can be seen from Fig. 5, the PV generation power is more stable on sunny days, and the prediction models can predict better results except for the SVR and ELM, which fluctuate, the SVR predicts poorly at moments of the day when the light is small, and the prediction effect of the ELM is affected by the randomly generated weights and biases. This shows that for more stable generation power, the models have better prediction results.

It can be observed from Fig. 6 that the prediction of each model has a large deviation from the actual value, and the larger volatility and randomness come from the fact that the PV power is affected by cloud movement and irradiance in cloudy weather. The SVR and ELM models are slightly poorly fitted under this influence, and the predictions of the single models all have hysteresis, and the combined models using CEEMDAN, ICEEMDAN decomposition method and DHKELM can overcome this problem to a certain extent, and obtain better prediction results than the single models.

As can be seen from Fig. 7, the amplitude of PV power curve in rainy days is the lowest, and there are the strong volatility and randomness. The prediction values of each model has the largest deviation from the actual values, and under this weather type, the prediction value of the prediction model proposed in this paper is the closest to the actual value of PV power, which indicates the effectiveness of the model proposed in this paper.

Table 2 shows the comparison of the prediction errors of different prediction models under various weather types. From the table, it can be seen that all the prediction error values for sunny days are smaller than those for cloudy and rainy days, which is due to the fact that the volatility of the PV power in sunny days is less, and it is more adaptable for each model. Comparing the single models, it can be seen that the RMSE of DHKELM prediction under sunny, cloudy, and rainy days reduces the RMSE of SVR, ELM, and HKELM by 27.57%, 12.20%, and 11.03%; 24.46%, 20.04%, and 17.77%; 13.79%, 6.17%, and 2.64%, respectively. From the above, it can be seen that the DHKELM model is optimal in the single model, which verifies the superiority of choosing DHKELM in the single model. The combined models that CEEMDAN-DHKELM and ICEEMDAN-DHKELM models have better predictive performance compared to the single model. For the above two combined models, ICEEMDAN-DHKELM decreased RMSE by 4.04%, 31.38%, and 18.61%, MAE by 25.70%, 41.88%, and 28.08%, MAPE by 6.81%, 29.66%, and 29.02%, for sunny, cloudy, and rainy day types, compared to CEEMDAN-DHKELM, respectively. It shows that the decomposition effect of ICEEMDAN is better than that of CEEMDAN, and better prediction effect can be obtained when dealing with PV power volatility. Compared to the ICEEMDAN-DHKELM model, the ICEEMDAN-BWO-DHKELM model reduces the RMSE by 34.79%, 26.71%, and 8.78%, the MAE by 23.30%, 21.53%, and 2.71%, and the MAPE by 40.58%, 18.93%, and 5.57% for sunny, cloudy, and rainy weather types, respectively. This shows that the optimization of DHKELM using BWO after decomposition by ICEEMDAN has better adaptability for each weather type.

## 5. CONCLUSION

In this paper, a PV power prediction method based on ICEEMDAN-BWO-DHKELM is proposed for the characteristics of PV power with stochasticity and volatility, and the following conclusions are obtained through the analysis of the arithmetic examples:

For the PV power with strong randomness and strong volatility, the decomposition method using ICEEMDAN is chosen to be processed to obtain several sub-sequences, and the ICEEMDAN effectively solves the problem of modal mixing and effectively overcomes the hysteresis problem occurring in the prediction with better decomposition effect.

For the problem that the parameters of DHKELM model are more and not easy to determine, the BWO-DHKELM is proposed, and experiments show that the BWO has a better ability to find the optimal and can effectively improve the prediction effect of DHKELM.

Simulation results show that the combined model that ICEEMDAN-BWO-DHKELM has a better prediction effect on short-term PV power under different weather types, and its prediction accuracy is better than that of some typical single models and other combined models, which has a better application prospect.

## ACKNOWLEDGMENTS

The work has been supported by the National Natural Science Foundation of China (61973105); Supported by the Key Scientific and Technological Projects in Henan (222102210120). Thanks are due to the Henan Polytechnic University Library for providing a detailed reference for the research.

## REFERENCES

[1] LAI Changwei, LI Jinghua, CHEN Bo, et al. Review of Photovoltaic Power Output Prediction Technology[J]. Transactions of China Electrotechnical Society,2019,34(06):1201-1217.

[2] ZHANG Xiangying, YANG Yongbiao, XU Qingshan, et al. Photovoltaic Power Combination Prediction Method Based on Multi-Temporal Similarity Day Theory[J/OL]. Southern Power System Technology: 1-9[2023-03-16].

[3] SHANG Liqun, LI Hongbo, HOU Yadong, et al. Short-term photovoltaic power generation prediction based on VMD-ISSA-KELM[J]. Power System Protection and Control, 2022,50(21):138-148.

[4] LI Yi, YANG Mao, SU Xin. Short-Term Prediction of Photovoltaic Power Based on Integrated Clustering and Improved Markov Chain Model[J/OL]. Southern Power System Technology: 1-10[2023-03-16].

[5] GONG Yingfei, LU Zongxiang, QIAO Ying, et al. An Overview of Photovoltaic Energy System Output Forecasting Technology[J]. Automation of Electric Power Systems, 2016, 40(4): 140-151.

[6] ZHU Qiongfeng, LI Jiateng, QIAO Ji, et al. Application and Prospect of AI Technology in Renewable Energy Forecasting[J/OL]. Proceedings of the CSEE: 1-23[2023-03-16].

[7] WANG Xin, HUANG Ke, ZHENG Yihui, et al. Short-term Forecasting Method of Photovoltaic Output Power Based on PNN/PCA/SS-SVR[J]. Automation of Electric Power Systems, 2016,40(17):156-162.

[8] HUANG G B, WANG D H, LAN Y. Extreme learning machines: a survey[J]. International Journal of Machine Learning and Cybernetics, 2011, 2(2): 107-122.

[9] Chu Y, Feng C, Guo C, et al. Network embedding based on deep extreme learning machine[J]. International Journal of Machine Learning and Cybernetics, 2019, 10(10): 2709-2724.

[10] WU Chunhua, DONG Along, LIZhihua, et al. Photovoltaic Power Prediction Based on Graph Similarity Day and PSO-XGBoost[J]. High Voltage Engineering, 2022,48(08):3250-3259.

[11] QIAO Weibiao, CHEN Baodong, WU Shijuan, et al. A forecasting model of natural gas daily load based on wavelet transform and LSSVM-DE[J]. Natural Gas Industry,2014,34(09):118-124.

[12] YANG Haizhu, TIAN Fuming, ZHANG Peng, et al. Short-term load forecasting based on CEEMD-FE-AOA-LSSVM[J]. Power System Protection and Control, 2022,50(13):126-133.

[13] Changting Zhong, Gang Li, Zeng Meng. Beluga whale optimization: A novel nature-inspired metaheuristic algorithm[J]. Knowledge-Based Systems,2022,251:109215.

[14] Zhang Na, Ren Qiang, Liu Guangchen. et al. Short-term PV Output Power Forecasting Based on CEEMDAN-AE-GRU[J]. Journal of Electrical Engineering & Technology. 17, 1183–1194 (2022).

[15] YANG Shuang, LUO Diansheng, HE Hongying, et al. Output power forecast of PV power system based on EMD-LSSVM model[J]. Acta Energiae Solaris Sinica, 2016,37(06):1387-1395.

[16] WANG Zhenhao, WANG Chong, CHENG Long, et al. Photovoltaic Power Combined Prediction Based on Ensemble Empirical Mode Decomposition and Deep Learning[J]. High Voltage Engineering, 2022, 48(10):4133-4142.

[17] WANG Rui, GAO Qiang, LU Jing. Short-term photovoltaic power prediction based on CEEMDAN-LSSVM-ARIMA model[J]. Transducer and Microsystem Technologies, 2022, 41(5): 118-122.

[18] COLOMINAS M A, SCHLOTTHAUER G, TORRES M E. Improved complete ensemble EMD: a suitable tool for biomedical signal processing[J]. Biomed Signal Process Control, 2014, 14(1): 19-29.

[19] Li Hongxia, Li Jianlin, Mi Yang. Research on Photovoltaic power prediction technology Based on Machine Learning[J]. Journal of Physics: Conference Series, 2021, 2087(1).

[20] XIANG Ling, DENG Zeqi, ZHAO Yue. Multi-step Wind Speed Prediction Model Based on LPF-VMD and KELM [J]. Power System Technology, 2019,43(12):4461-4467.

[21] WANG Rui, XU Xinchao, LU Jing. Short-term Wind Power Prediction Based on SSA Optimized Variational Mode Decomposition and Hybrid Kernel Extreme Learning Machine[J/OL]. Information and Control: 1-11[2023-03-17].

[22] OUYANG Sen, SHI Yili. A new improved entropy method and its application in power quality evaluation[J]. Automation of Electric Power Systems, 2013, 37(21):156-159.

[23] Khan J, Fayaz M, Hussain A, et al. An improved alpha beta filter using a deep extreme learning machine[J]. IEEE Access, 2021, 9: 61548-61564.