

RESEARCH

Open Access



# A design of new wind power forecasting approach based on IVMD-WSA-IC-LSTM model

Zhenhui Li<sup>1\*</sup> and Shuchen Xiang<sup>1</sup>

\*Correspondence:  
940341533@qq.com

<sup>1</sup> Hunan Institute of Engineering,  
Xiangtan 411104, China

## Abstract

The wind power forecasting (WPF) technology can reduce the adverse impact of wind power grid connection. Based on the characteristics of wind power data, an algorithm based on improved variational mode decomposition (IVMD) and long short-term memory (LSTM) Network is proposed to predict the wind power, and hyper parameter optimization search of LSTM using Whale Swarm Algorithm with Iterative Counter (WSA-IC). Firstly, through correlation analysis, the characteristics of 10 different wind power data are screened, and two kinds of data with large correlation with wind power are determined as input of the mode. Secondly, IVMD is used to calculate the maximum envelope kurtosis, determine the best decomposition parameters of the variational mode decomposition (VMD), and the original wind power and wind speed sequences are decomposed to obtain the IMF with different time scales. Finally, to address the problems of difficult optimization of hyper parameter and difficulty in obtaining optimal solutions for LSTM neural network modes, the WSA-IC algorithm is proposed to optimize its key hyper parameter, and the IVMD-WSA-IC-LSTM forecasting mode is established to obtain the short-term forecasting results of wind power. The algorithm is tested with the data of China Longyuan Power Group Corporation Limited. Compared with other common forecasting approaches using same data, the mean absolute error (MAE) of the forecasting approach is reduced to 0.007859, the mean square error (MSE) is reduced to 0.00011, and the determination coefficient is improved to 0.998828, which has higher forecasting accuracy.

**Keywords:** WPF, LSTM neural network, WSA, EMD

## Introduction

Wind power, as a new energy, has developed rapidly in recent years [1]. Wind power plays an important role in our country and is widely used. However, randomness, intermittent and uncertainty of wind power will adversely affect the stable and safe of power grid. Accurate wind power forecasting is beneficial to efficient operation, safety analysis, and energy trading of power grid [2].

In response to the problems of low accuracy and poor anti-interference ability of traditional power forecasting algorithms, researchers propose some artificial intelligence-based WPF algorithms. Xu et al. [3] proposes a new WPF approach by mining the data of numerical weather prediction (NWP). The WPF approach based on LSTM neural

network often has better accuracy than some artificial intelligence approaches [4, 5]. LSTM can effectively solve the problem of RNN's inability to capture long-term dependence, but it has high complexity, long training time and multiple hyperparameters difficult to select. Yang et al. [6] proposed a bidirectional long short-term memory (Bi-LSTM) to predict multiple attributes of a product. Delgado et al. [7] presents a recurrent neural network-based variant LSTM-based power forecasting. A new prediction approach for wind power is proposed in Sun et al. [8], in which LSTM network, wavelet decomposition (WT), and principal component analysis (PCA) are combined together. In Pu et al. [9], an ultra-short-term WPF mode based on PSO and LSTM neural network combination is proposed, which improves the precision of ultra-short-term WPF. However, particle swarm optimization (PSO) has some problems such as slow convergence rate and local optimal solution. The optimization algorithms such as Drosophila Optimization (DO), Whale Swarm Optimization (WSO), and Bat Swarm Optimization (BSO) have emerged. In the Peng et al. [10], an improved DO algorithm is used to propose a wind speed forecasting mode to obtain wind power data with higher forecasting accuracy by expanding the parameter search range and improving the convergence speed of the algorithm. However, the original data is characterized by the rapid variational of wave crests and troughs, which leads to some defects in the forecasting mode. For data problems, Shi et al. [11] uses wavelet decomposition (WD) and artificial neural network (ANN) for forecasting, which improves forecasting performance, but wavelet packet decomposition in the Zheng et al. [12] requires artificial wavelet function setting, and noise interference may occur if improper decomposition result is set. In the Cheng et al. [13], a short-term WPF mode based on the combination of empirical mode decomposition (EMD) [14] and radial basis function neural network (RBFNN) [15] was proposed to make the decomposed wind power component series with strong regularity and improve the forecasting accuracy. Wang et al. [16, 17] adopts empirical mode decomposition empirical model decomposition (EMD) method to better solve the problem of modal aliasing phenomenon and susceptibility to noise interference, but the total number of modes  $K$  is difficult to determine, affecting the accuracy the mode. These methods solve the nonlinear problem of wind power data forecasting to some extent, but there are some problems such as poor generalization ability of mode, long training time and easy to over-fit or under-fit. Konstantin et al. [18] proposes RNN mode based on LSTM to predict wind power, which can make more accurate forecasting by using long-term correlation in time series, and solve the defects that other neural networks cannot learn for a long time. The Deng et al. [19] proposed a method to optimize the design of key hyper parameter of LSTM neural networks by combining the cuckoo algorithm. In the Chang et al. [20], a sparrow search algorithm is proposed to optimize the learning rate and number of neurons in LSTM and CNN-LSTM networks. However, it is difficult to select the optimal hyper parameter of LSTM neural network mode, and intelligent optimization algorithms such as cuckoo bird algorithm and sparrow algorithm need to set niche parameters for different problems, so it is difficult to obtain global optimal solutions.

WSA (Whale Swarm Algorithm) is a meta heuristic algorithm, which belongs to Swarm intelligence algorithm [21]. WSA-IC(Whale Swarm Algorithm with Iterative Counter) does not need to introduce small habitat parameters and can effectively jump out of local

optimal solutions by identifying extreme value points based on stability thresholds and fitness thresholds during the iterative process. An IVMD-WSA-IC-LSTM WPF algorithm is proposed. The maximum envelope kurtosis is calculated by IVMD to determine the optimal decomposition parameter  $K$  of VMD, and the WSA-IC is used to improve the multi-modal optimization iteration rule to obtain the optimal hyper parameter and global optimal solution of LSTM to improve the WPF accuracy. Compared with other forecasting algorithms, the forecasting error is smaller, the average absolute error is lower, the coefficient of determination and the forecasting accuracy are higher.

The first part of the paper introduces the VMD algorithm, and improves in the second part. In the third part, the improved whale swarm optimization algorithm is introduced to optimize the hyper parameter of LSTM neural network, and the fourth part tests the algorithm.

### Improve methods

#### VMD Algorithm

VMD takes the original signal data into a number of Eigen mode components IMF with different center frequencies with finite bandwidth [22]. By constructing and solving the variational problem, the optimal solution and its central frequency of each Eigen mode function are solved by iteration [23].

Firstly, the variational problem [24] is constructed, assuming that the original signal  $f$  is decomposed into  $k$  components, and the decomposition signal is guaranteed to be a modal component of finite bandwidth with central frequency. The constraint variational expression is shown in Eq. (1):

$$\begin{cases} \min \left\{ \sum_{k=1}^K \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-jw_k t} \right\|_2^2 \right\} \\ s.t. \sum_{k=1}^K u_k = f \end{cases} \tag{1}$$

where  $u_k$  is mode signal;  $f$  is time series;  $k$  is mode number;  $*$  refers to convolution operator;  $w_k$  refers to center frequency of  $k$ th mode component;  $\delta(t)$  refers to Dirac function, indicating density; and  $e^{-jw_k t}$  is the phase of signal rotation with time.

Secondly, the variational problem is transformed into a non-constraint problem, to facilitate the optimal solution of the Variational constraint mode and obtain an augmented Lagrange expression [25].

$$L(\{u_k\}, \{w_k\}, \lambda) = \alpha \sum_{k=1}^K \left\| \partial_t \left[ \left( \delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-jw_k t} \right\|_2^2 + \|f(t) - u_k(t)\|_2^2 + \left[ \lambda(t), f(t), f(t) - \sum_{k=1}^K u_k(t) \right] \tag{2}$$

where,  $\alpha$  is the secondary penalty factor,  $\lambda$  is the Lagrangian multiplier.

Finally, by using alternate direction multiplier method, the sum is optimized and the ‘‘saddle point’’ of Eq. (2) is searched, which is the optimal solution of Eq. (1). Expressions of  $u_k^{n+1}, w_k^{n+1}, \lambda_k^{n+1}$  are obtained by alternately updating:

$$\hat{u}_k^{n+1}(w) = \frac{\hat{f}(w) - \sum_{i \neq k} \hat{u}_i(w) + \frac{\hat{u}(w)}{2}}{1 + 2\alpha(w - w_k)^2} \tag{3}$$

$$w_k^{n+1} = \frac{\int_0^\infty w |\hat{u}_k(w)|^2 dw}{\int_0^\infty |\hat{u}_k(w)|^2 dw} \tag{4}$$

$$\hat{\lambda}^{n+1}(w) = \hat{\lambda}^n(w) + \gamma \left( \hat{f}(w) - \sum_k \hat{u}_k^{n+1}(w) \right) \tag{5}$$

$$\sum_k \frac{\|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2}{\|\hat{u}_k^n\|_2^2} < \varepsilon, n < N? \tag{6}$$

where,  $\varepsilon$  is the convergence criterion and  $N$  is the maximum number of iterations.

**IVMD ALGORITHM**

Since VMD decomposition is to decompose the original data sequence from low frequency to high frequency one by one, the low frequency IMF sequence is more likely to reflect the big trend of data, when envelope kurtosis is maximum, the frequency of IMF sequence is the highest, at this time, no further VMD decomposition is necessary. Therefore, finding the maximum envelope kurtosis means finding the optimum number of layers  $K$ . The maximum envelope kurtosis solution process is as follows:

Assuming that the number of decomposition layers of VMD is  $K$ , the envelope of each IMF can be calculated from the following equation, i.e.:

$$xt_k^i = \left| \frac{1}{\pi t} \cdot xt_k^i(t) \right| \quad i = 1, 2, \dots, k \tag{7}$$

where,  $xt_k^i$  is the absolute value obtained by Hilbert transformation, and  $xt_k^i(t)$  is the  $i$ th IMF when the number of decomposition layers  $K$ .

The  $i$ th IMF envelope kurtosis calculation formula is

$$ek(i) = \frac{E(xt_k^i - \mu(xt_k^i))^4}{\sigma(xt_k^i)^4} \tag{8}$$

where,  $\mu(xt_k^i)$  is the mean value of  $xt_k^i$ ;  $\sigma(xt_k^i)$  is the standard deviation of  $xt_k^i$ ,  $E(xt_k^i - \mu(xt_k^i))^4$  is the fourth order central moment of  $xt_k^i$ .

The envelope kurtosis  $ek_i (i = 1, 2, \dots, k)$  of each IMF can be calculated from Eq. (8), therefore, the local maximum envelope kurtosis  $ek_k^{\max}$  is

$$ek_k^{\max} = \max\{ek_1, ek_2, \dots, ek_k\} \tag{9}$$

The local maximum value  $ek_k^{\max}$  when  $K=(2, \dots, k)$  is calculated respectively from Eq. (9), and the global maximum value of maximum envelope kurtosis  $ek_g^{\max}$  is calculated by Eq. (10), and the trend diagram of maximum envelope kurtosis is drawn. The value of  $K$  corresponding to global maximum value is the optimal decomposition layer number of maximum VMD.

$$ek_g^{\max} = \max\{ek_1^{\max}, ek_2^{\max}, \dots, ek_k^{\max}\} \tag{10}$$

### Improve Whale Swarm Algorithm

WSA is a new meta-heuristic algorithm, belonging to group intelligence algorithm [26–30]. It is used to solve the merit-seeking problem by imitating behaviors such as searching and hunting among whale populations using ultrasonic waves as information carrier. The position iterative formula of the whale swarm algorithm is:

$$\begin{cases} x_i^{t+1} = x_i^t + \beta(y_i^t - x_i^t) \\ \beta = rand(0, \rho_0 e^{-\eta d_{X,Y}}) \end{cases} \quad (11)$$

where,  $\eta$  is the ultrasonic attenuation factor;  $d_{xy}$  is the distance between  $X$  and  $Y$ ;  $\rho_0$  represents the initial ultrasonic intensity;  $y_i^t$  indicates the position of whale  $y_i$  at step  $t$ ;  $x_i^t$  and  $x_i^{t+1}$  is the position of whale  $x_i$  at step  $t$  and  $t+1$  respectively; and  $\beta$  is a random number of  $[0, \rho_0 e^{-\eta d_{X,Y}}]$ .

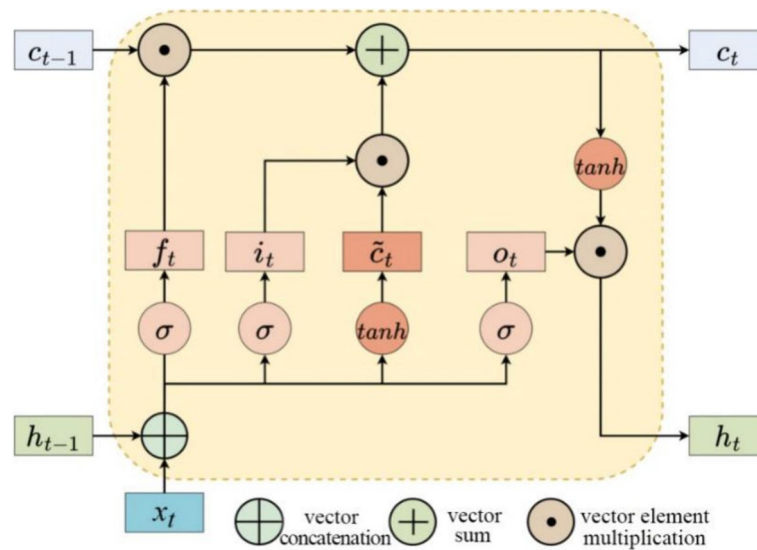
WSA-IC improves the iteration rule of WSA and solves the problem of getting into local optimality easily without specifying different attenuation coefficient values for different problems, and without introducing any niche parameters. WSA-IC can identify and jump out the found extremum effectively during iterative process, which is helpful to find global optimal solution.

### Improvement of WSA iteration rules

Improve WSA location update rule: generate a replica of that current whale, and if the replica whale does not move to a better position than the original whale under the guidance of its “better and nearest” whale, then the original whale remains in its original position, otherwise the original whale moves to the copy whale position. Without introducing arbitrary niche parameters [31], this strategy effectively generates multiple subgroups, which enhances the local optimization ability of the algorithm and helps to find multiple optimal solutions.

### Identify and jump out of found extreme points

Two new parameters, namely stability threshold  $T_s$  and fitness threshold  $T_f$  are introduced for identifying and jumping out of the found extreme points. Firstly, the preset  $T_s$  is used to determine whether a whale has reached a steady state (i.e., an extreme point is found), and an iteration counter  $c$  is set for each whale to record the number of iterations for the un-updated position of the whale. Secondly, the global optimal solution is identified by  $T_f$ . By checking the iterative counters of the whales that have not made position updates in each iteration, if the whale reaches the steady state ( $X.c = T_s$ ), the size of  $T_f$  and  $|f_{g_{best}} - f(X)|$  determine whether to update to the current global optimal solution set, then initialize the whale to jump out of the found extreme value point. Otherwise, the iteration counter is incremented by 1.



**Fig. 1** The structure of the LSTM cycle unit

### LSTM forecasting mode

LSTM network is an improvement of recurrent neural network (RNN), which effectively solves the gradient disappearance and explosion problems existing in the original RNN by adding a gating structure. LSTM is known as a gate-based cyclic neural network. The LSTM network is able to obtain important information for a specific period of time and can maintain a certain time interval.  $c$  Stores information longer than short-term memory, but much shorter than long-term memory. The structure of the LSTM cycle unit is shown in Fig. 1, and the calculation formula is as follows:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{12}$$

$$h_t = o_t \odot \tanh(c_t) \tag{13}$$

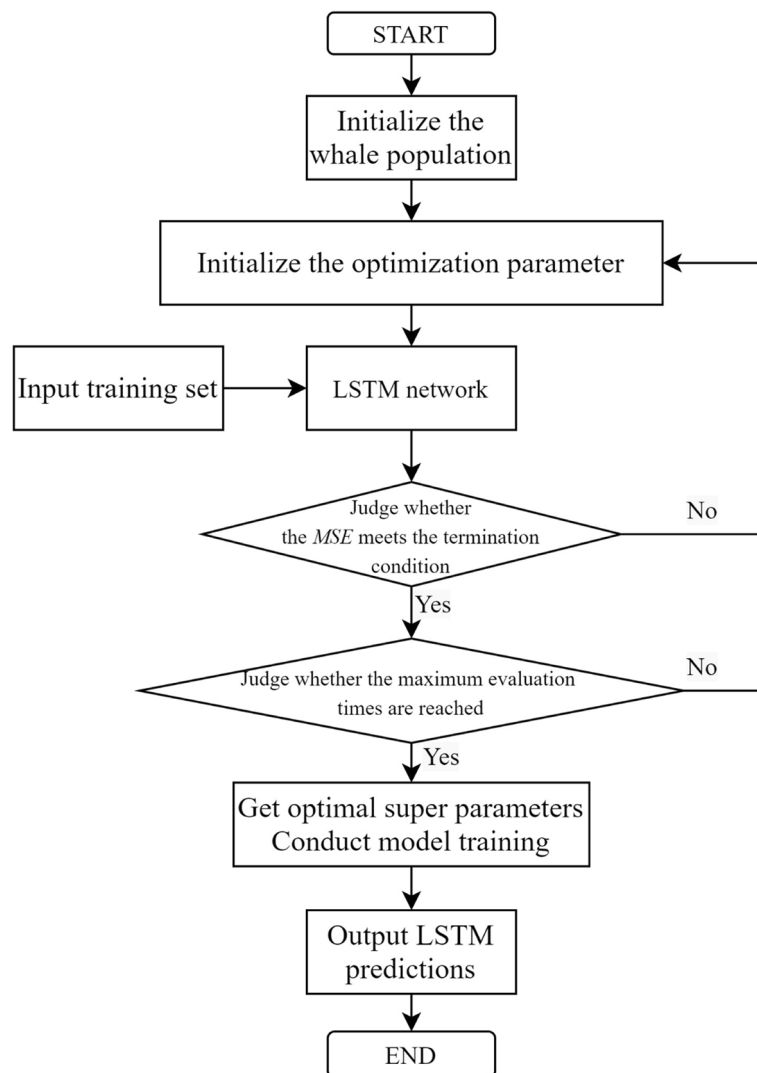
$$\tilde{c}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c) \tag{14}$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{15}$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{16}$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{17}$$

where  $\odot$  is the product of vector element;  $c_{t-1}$  is the memory cell at the previous time;  $\tilde{c}_t$  is the candidate state,  $\tanh$  is the activation function;  $W$ ,  $U$ , and  $b$  are network parameters;  $\sigma()$  is the Sigmoid activation function, whose output interval is (0, 1);  $x_t$  is the input at the current time;  $h_{t-1}$  is the external state at the previous time;  $W$ ,  $U$  and  $b$  are network parameters.



**Fig. 2** The flow of component forecasting based on WSA-IC-LSTM

**IVMD-WSA-IC-LSTM forecasting mode**

The accuracy of WPF will be affected by the setting of LSTM hyper parameter. First, a WSA-IC-LSTM forecasting algorithm mode is designed based on the learning rate  $l_p$ , the learning rate planning  $r_p$ , the number of hidden layer neurons  $U_h$  and the number of samples  $b_s$  of the LSTM mode as the hyper parameter of the WSA-IC algorithm for finding the optimal object, and the fitness value of  $MSE$  is taken as the target function. During the training process, the corresponding optimal hyper parameter with the smallest  $MSE$  is searched. Using WSA-IC for hyper parameter search can avoid the problem of randomness caused by artificial parameter adjustment, and the optimization effect is stable.

The flow of component forecasting based on WSA-IC-LSTM is shown in Fig. 2. The steps of WAS-IC-LSTM are as follows:

- Step 1: Initialize the whale population. Including whale position initialization, iteration counter set to zero;
- Step 2: Initialize the optimization parameter. initialize superparameters such as  $l_p$ ,  $r_p$ ,  $U_{h1}$ ,  $U_{h2}$ ,  $bs$ , etc.;
- Step 3: Input training set, train LSTM network and calculate training error  $MSE$ ;
- Step 4: Judge whether the  $MSE$  meets the termination condition. If the termination condition is satisfied, save the optimal hyper parameter at this time and execute step 5. Otherwise, repeat step 2;
- Step 5: The optimal hyper parameter of LSTM is set for mode training;
- Step 6: Output LSTM forecasting in preparation for subsequent work.

Then a combined forecasting method based on IVMD-WSA-IC-LSTM is proposed:

- Step 1: Using IVMD to decompose the preprocessed data into  $K$  IMF components;
- Step 2: Normalize each component;
- Step 3: Initialize the boundary threshold, fitness threshold, stability threshold, distance threshold and maximum evaluation times of WSA-IC algorithm, and initialize the number of whale population of WSA-IC algorithm;
- Step 4: The forecasting mode of LSTM wind power is constructed, the parameter optimization range is determined, and a whale population is generated. The learning rate  $l_p$ , learning rate planning  $r_p$ , the number of neurons in hidden layer  $U_{h1}$  and  $U_{h2}$ , the number of batch samples  $bs$  are taken as hyper parameter variables, and the parameter optimization range is determined, and each IMF component is divided into training set and test set;
- Step 5: Defining the fitness, using the mean square error  $MSE$  of the LSTM mode output values as the fitness value;
- Step 6: Each IMF component is sent to LSTM network training, superparameter optimization using WSA-IC algorithm is used to obtain the optimal hyper parameter minimum time of  $MSE$ . The optimal superparameter of LSTM is set for training, and the predicted value of each component is output.
- Step 7: The forecasting results of each IMF component are superimposed to obtain a final power forecasting value.

## Results and discussion

### Algorithm implementation and testing

Hardware platform configuration: Windows 10 operating system, 16 G memory, GTX1650 super graphics card, using Python 3.8 for programming. Select the 245-day data of Baidu KDD CUP 2022 No. 1 blower, taking 10 min as an adoption point, with 35,280 sampling points in total. The forecasting mode adopts rolling multi-step forecasting mode. Since 10 min is taken as the sampling point, the time window (input sequence length)  $N$  of forecasting of wind power is set to 24 and the measurement step length (output sequence length)  $l$  is set as 10.



**Table 1** The local maximum envelope kurtosis under different  $K$ 

$K$	2	3	4	5	6	7	8	9
Kurtosis of wind speed series	1.3	2.9	4.1	6.9	6.9	6.9	6.8	6.8
Kurtosis of power series	0.1	1.1	5.7	8.4	8.6	8.5	8.4	8.4

**Table 2** The number of layers of the VMD

Parameters	$\alpha$	$\tau$	$K$	$DC$	$init$	$tol$
Value	7000	0	6	0	1	$10^{-7}$

### IVMD decomposition

The wind power active power and wind speed series decomposed by IVMD are the data of 245 days with 10 min as a sampling point of the above data set, and the maximum envelope kurtosis method is adopted for analysis.  $K$  is taken as 2–9. By calculating the local maximum envelope kurtosis under different values of  $K$ , the  $K$  value corresponding to the global maximum envelope kurtosis is found, as shown in Table 1.

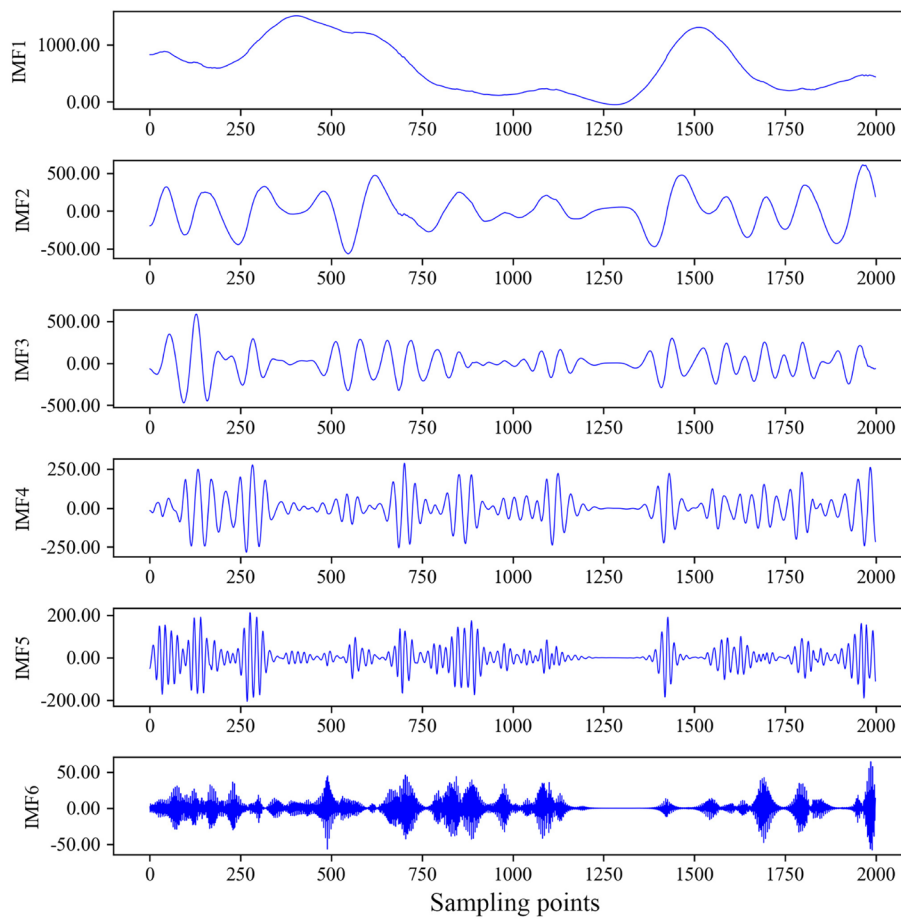
According to Table 1, when  $K=6$ , the maximum envelope kurtosis is the global maximum envelope kurtosis, so the optimum number of VMD layers  $K=6$  is determined.

Next, the input parameters such as penalty factor of VMD  $\alpha$ , noise tolerance  $\tau$ , direct current component parameter  $DC$ , convergence tolerance criterion  $tol$ , initialized center frequency parameter  $init$ , and the number of layers  $K$  of the VMD are set as shown in Table 2.

The VMD decomposition was performed on the power and wind speed series respectively, and for the convenience of observation, 2000 sampling points were intercepted to do the decomposition results schematically. The power IVMD decomposition results are shown in Fig. 3. Adopting the power sub-sequence (IMF1-IMF6) and wind speed sequence (IMF1-IMF6) after IVMD decomposition, it can be seen from the figure that there are obvious differences in the frequency of modes among each sub-sequence, which effectively avoids the problem of mode aliasing and end effect. The new sub-sequence is constructed by reconstituting the corresponding position, using the reconstructed sequence to predict the components, and finally each forecasting results are superimposed to obtain the optimized forecasting result.

### Analysis of new WPF mode

In order to verify the forecasting effect of forecasting algorithm, BP neural network forecasting algorithm, RNN forecasting algorithm, LSTM network forecasting algorithm, LSTM network forecasting algorithm based on WSA-IC optimization and IVMD-LSTM forecasting algorithm are introduced. Among them, BP forecasting algorithm, RNN forecasting algorithm, LSTM forecasting algorithm and IVMD-LSTM forecasting algorithm set the same hyper parameter. The IVMD-LSTM algorithm and IVMD-WSA-IC-LSTM algorithm set the same VMD decomposition times, and the parameter settings are shown in the following table:



**Fig. 3** The power IVMD results

**Table 3** The optimal hyper parameter

Model	$l_r$	$r_f$	$U_{h1}$	$U_{h2}$	$bs$
WSA-IC-LSTM	0.00474	0.2538	23	88	249
WSA-IC-LSTM1	0.00019	0.3542	69	89	142
WSA-IC-LSTM2	0.01004	0.4485	66	93	205
WSA-IC-LSTM3	0.01199	0.4682	69	56	68
WSA-IC-LSTM4	0.02532	0.3988	72	38	36
WSA-IC-LSTM5	0.02485	0.6570	23	48	142
WSA-IC-LSTM6	0.00361	0.5341	21	43	22

The IVMD-WSA-IC-LSTM algorithm mode consists of 6 sub-modes. Hyperparameters optimization is performed through WSA-IC algorithm. The optimal hyper parameter found is shown in Table 3.

**Comparison and analysis of single forecasting algorithms**

The forecasting effects of BP, RNN, LSTM, and WSA-IC-LSTM are compared. The hyper parameter of BP, RNN and LSTM forecasting algorithms are set artificially, and

**Table 4** The optimal hyper parameters

Mode	$l_r$	$r_f$	$U_{h1}$	$U_{h2}$	$bs$
BP	0.01	0.2	52	52	128
RNN	0.01	0.2	52	52	128
LSTM	0.01	0.2	52	52	128
WSA-IC-LSTM	0.0047	0.2538	23	88	249

**Table 5** The comparison diagrams of the forecasting results

Mode	MSE	MAE	$R^2$
BP	0.02503	0.126586	0.904189
RNN	0.02091	0.110173	0.935742
LSTM	0.01823	0.108281	0.938887
WSA-IC-LSTM	0.01415	0.074738	0.963176
IVMD-LSTM	0.001481	0.028092	0.984204
IVMD-WSA-IC-LSTM	0.000110	0.007862	0.998824

the WSA-IC-LSTM forecasting algorithm finds the optimal hyper parameter through WSA-IC. The hyper parameter settings for the BP, RNN, LSTM, and WSA-IC-LSTM forecasting algorithm modes are shown in Table 4:

The error analysis indicators for the forecasting effects of the four algorithm modes are shown in Table 5. Here, mean absolute error (MAE) refers to the average value of the absolute error between the real value and the predicted value, mean squared error (MSE) is the expected value of the residual sum of squares,  $R^2$  represents the numerical feature of the relationship between a variable and other random variables.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{18}$$

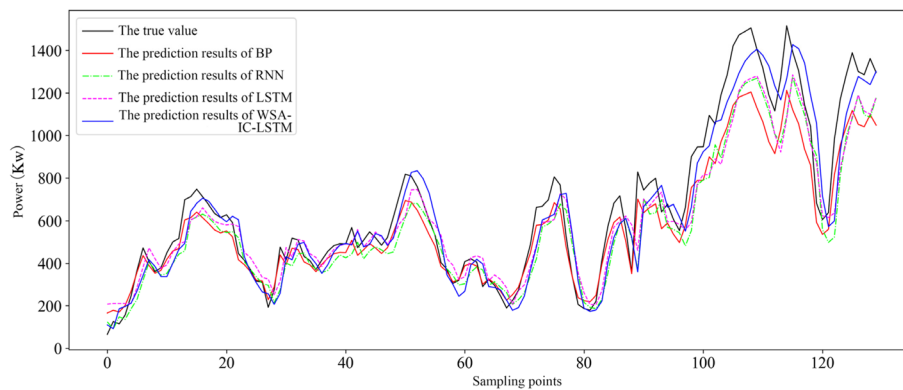
$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{19}$$

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y}_i)^2 \tag{20}$$

$$SST = \sum_{i=1}^n (y_i - \bar{y}_i)^2 \tag{21}$$

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \tag{22}$$

where  $n$  is the number of samples,  $y_i$  is the test value of wind power, and  $\hat{y}_i$  is the predicted value of wind power.



**Fig. 4** The error analysis indicators for the forecasting effects of the four algorithm modes

And Fig. 4 shows the comparison diagrams of the forecasting results of the BP, RNN, LSTM and WSA-IC algorithm modes with 130 sampling points.

The effect of LSTM algorithm forecasting mode is significantly better than BP and RNN algorithm forecasting effect. Second, the WSA-IC-LSTM forecasting algorithm mode outperforms with the LSTM forecasting algorithm mode.

## Conclusions

In view of the higher precision of WPF in power system, an ultrashort-term wind power forecasting method based on IVMD-WSA-IC-LSTM is proposed in this paper. And the following conclusions are drawn through comparative experiments:

- (1) IVMD determines the number of decomposition layers of VMD, and can accurately acquire the intrinsic characteristics between wind power data. Compared with LSTM, IVMD-LSTM forecasting mode has better forecasting effect, which verifies the feasibility of original wind power sequence based on IVMD decomposition processing.
- (2) The WSA-IC-LSTM forecasting mode is based on the LSTM forecasting, and the WSA-IC algorithm is added for hyperparameter search to find the smallest loss function MSE, corresponding to the learning rate  $l_r$ , the learning rate attenuation  $r_f$ , the number of hidden layer units  $U_{h1}$  and  $U_{h2}$  and the number of batch samples  $bs$ . By comparing the forecasting results of the two modes, it is found that the forecasting effect of WSA-IC-LSTM is better than LSTM. The validity of hyper parameter optimization based on intelligent optimization algorithm (WSA-IC) is verified.
- (3) Further in-depth comparison analysis of the forecasting effects of the IVMD-WSA-IC-LSTM forecasting mode and the IVMD-LSTM forecasting mode, combined with the above findings, led to the conclusion that the mode presented in this paper has the highest forecasting accuracy. The predicted  $MAE$  was reduced to 0.007859,  $MSE$  decreased to 0.00011, and  $R^2$  increased to 0.998828. The feasibility is verified.

### Abbreviations

WPF	Wind power forecasting
IVMD	Improved variational mode decomposition
LSTM	Long short-term memory
WSA-IC	Whale Swarm Algorithm with Iterative Counter
VMD	Variational mode decomposition
MAE	Mean absolute error
MSE	Mean square error
Bi-LSTM	Bidirectional long short-term memory
WT	Wavelet transform
PCA	Principal component analysis
PSO	Particle swarm optimization
DO	Drosophila optimization
WSO	Whale swarm optimization
BSO	Bat swarm optimization
WD	Wavelet decomposition
ANN	Artificial neural network
RBFNN	Radial basis function neural network
EMD	Empirical mode decomposition
RNN	Recurrent neural network

### Acknowledgements

This work was supported by National Natural Science Foundation of China (grant number 62006075), National Natural Science Foundation of Hunan province of China (grant number 2022JJ50116 and 2022JJ50014), Key Supported Project of the 14th Five Year Plan for Education and Science in Hunan Province (XJK21AGD005).

### Authors' contributions

ZHL: supervision; SCX: writing. Both authors have read and approved the final manuscript.

### Funding

National Natural Science Foundation of China (grant number 62006075), National Natural Science Foundation of Hunan province of China (2022JJ50116 and 2022JJ50014), Key Supported Project of the 14th Five Year Plan for Education and Science in Hunan Province (XJK21AGD005).

### Availability of data and materials

All data, models, and code generated or used during the study appear in the submitted article.

### Declarations

#### Competing interests

The authors declare that they have no competing interests.

Received: 2 June 2023 Accepted: 31 July 2023

Published online: 09 August 2023

### References

1. Qia W (2021) Forecast of China's wind power installed capacity and corresponding CO2 reduction from 2020 to 2060. *Ecol Econ* 37(7):13–21
2. Xiyun Y, Yanfeng Z, Tianze Ye et al (2020) Prediction of combination probability interval of wind power based on naive Bayes. *High Voltage Eng* 46(3):1096–1104
3. Qianyao Xu, He D, Zhang N et al (2015) A short-term wind power forecasting approach with adjustment of numerical weather prediction input by data mining. *IEEE Trans Sustainable Energy* 6(4):1283–1291
4. Huang F, Li Z, Xiang S et al (2021) A new wind power forecasting algorithm based on long short-term memory neural network. *Int Trans Electrical Energy Syst* 31(12):e13233
5. Haritha N V, Anand J (2022) Solar power forecasting using long short-term memory algorithm in Tamil Nadu State. *Emerging Technologies for Sustainable and Smart Energy*. CRC Press, 73–96
6. Zaoli Yang, Yue Gao, and Xiangling Fu (2023) A Decision-making algorithm combining the aspect-based sentiment analysis and intuitionistic fuzzy-VIKOR for online hotel reservation[J]. *Ann Oper Res* 326(Suppl 1):49. <https://doi.org/10.1007/s10479-021-04339-y>
7. Delgado I, Fahim M (2020) Wind turbine data analysis and LSTM-based prediction in SCADA system. *Energies* 14(1):125
8. Sun Y, Wang P, Zhai S et al (2020) Ultra short-term probability prediction of wind power based on LSTM network and condition normal distribution. *Wind Energy* 23(1):63–76
9. Xianyi Pu, Guihong Bi, Kai W et al (2021) Ultra-short-term wind power prediction based on two-layer decomposition technique and PSO-LSTM Model. *Electr Machines Control Appl* 48(05):86–92
10. Chenyu P, Ning C, Bingtuan G (2020) Ultra-short-term wind power forecasting method combining multiple clustering and hierarchical clustering. *Automation Electr Power Syst* 44(2):172–179
11. Jun S, Yanan Z, Wei X et al (2021) Deep scattering network with fractional wavelet transform. *IEEE Trans Signal Process* 69:4740–4757

12. Zheng J, Xu C, Zhang Z, et al (2017) Electric load forecasting in smart grid using long-short-term-memory based recurrent neural network. *The 51st Annual Conference on Information Systems and Sciences Proceedings* 107:191–208
13. Qiming C, Chen Lu, Yinman C et al (2018) Short-term wind power forecasting method based on EEMD and LS-SVM Model. *Electr Power Automation Equip* 38(5):27–35
14. Wang JZ, Zhou GH, Zhao XS, et al (2007) Gearbox fault diagnosis and prediction based on empirical mode decomposition scheme. *2007 International Conference on Machine Learning and Cybernetics* 2:1072–1075
15. Shuai K, Li Z, Zhou L et al (2022) Multi-objective optimization design of PMASynRM based on RBF neural network. *J Physics Conference Series*. 2183(1):012013 (IOP Publishing)
16. Jixuan W, Bin D, Jiang W (2020) Short-term wind power prediction based on empirical mode decomposition and RBF neural network. *Proc CSU-EPSA* 32(11):109–115
17. Liu H, Tian H-q, Pan D-f et al (2013) Forecasting models for wind speed using wavelet, wavelet packet, time series and artificial neural networks. *Appl Energy* 107:191–208
18. Konstantin D, Zosso D (2013) Variational mode decomposition. *IEEE Trans Signal Process* 62(3):531–544
19. Yaping D, Jiandong D, Hao J et al (2021) Ultra-short-term wind power prediction based on deep learning with independent recurrent neural network via Cuckoo Algorithm Optimized. *Power Syst Clean Energy* 37(09):18–26
20. Dongfeng C, Xinyuan N (2022) Short-term photovoltaic power prediction based on back propagation neural network improved by hybrid sparrow algorithm. *Modern Electr Power* 39(3):12
21. Zeng B, Li X, Gao L et al (2020) Whale swarm algorithm with iterative counter for multimodal function optimization. *Neural Comput Appl* 32(9):5071–5091
22. Chaudhari H, Nalbalwar S L, Sheth R (2016) A review on intrinsic mode function of EMD. *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)*. 2349–235
23. Han Li Wu, Xiangyang LL et al (2021) Acoustic diagnosis method of rail corrugation in high speed railway based on component energy ratio of intrinsic mode function. *Railw Eng* 61(09):117–120
24. Sirsat AV, Padhee S (2022) A variational principle based approach for general solution to transverse isotropic axisymmetric cylinder problem. *J Appl Mech*. 107:191–208
25. Kovtunen VA, Itou H, Khludnev AM et al (2022) Non-smooth variational problems and applications. *Philos Trans R Soc A* 380(2236):20210364
26. Bing Zeng, Xinyu Li, Liang Gao, et al (2020) Whale swarm algorithm with the mechanism of identifying and escaping from extreme points for multimodal function optimization[J]. *Neural Comput Appl* 32(9):5071–91.
27. Bing Zeng, Lingxiang Huang, Langjun Peng et al (2020) Parameter tuning method of PID individual pitch control based on Whale Swarm Algorithm. *Ship Eng* 42(S1):550–553+557
28. Zeng B, Dong Y, Li X et al (2017) IHSCR: energy-efficient clustering and routing for wireless sensor networks based on harmony search algorithm. *Int J Distrib Sens Netw* 13(11):1550147717741103
29. Li X, Qin K, Zeng B et al (2016) Assembly sequence planning based on an improved harmony search algorithm. *Int J Adv Manuf Technol* 84(9–12):2367–2380
30. Li M, Zhang Y, Zeng B et al (2016) The modified firefly algorithm considering fireflies' visual range and its application in assembly sequences planning. *Int J Adv Manuf Technol* 82(5–8):1381–1403
31. Gu Qinghua, Luo Jiale, Li Xuexian (2023) Evolutionary algorithm Based on Niche for many-objective optimization[J]. *Comput Engr Appl* 59(1):126–139

### Publisher's Note

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

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---