International Journal of Power and Energy Systems, Vol. 44, No. 10, 2024
An Open Access Paper

A NEW DESIGN OF VMD-BAIPSO-GRU POWER FORECASTING ALGORITHM

Feng Huang,* Rui Wang,* Yuelin Yu,* Feiyu Hu,* Xin Xie,* and Lingxiang Huang**

Abstract

Wind power forecasting (WPF) has important practical value for grid-connected systems. To address the difficulty of predicting wind power, an improved particle swarm optimisation (IPSO) algorithm for gated recurrent unit (GRU) power forecasting was proposed. Firstly, to deal with problems, such as the instability and large fluctuations of wind power data, the variational mode decomposition (VMD) algorithm was applied to preprocess historical wind power data. Then, a GRU neural network model was established, and the Bat algorithm (BA) and IPSO were used in the PSO process to obtain the hyperparameters of the GRU neural network, to determine the parameters of the forecasting model. Finally, the new VMD-BAIPSO-GRU neural network wind power algorithm was proposed to divide the VMD decomposed data into training and testing sets, and the model was then trained and tested. Compared with other similar algorithms, the VMD-BAIPSO-GRU algorithm has a lowers root mean square error (RMSE) of 0.37 and mean absolute percentage error (MAPE) of 12.09%, indicating higher forecasting accuracy.

Key Words

WPF, VMD, BA, IPSO, GRU neural network

List of Abbreviations

Particle swarm optimisation (PSO)
Improved particle swarm optimisation (IPSO)
Gated recurrent unit (GRU)
Variational mode decomposition (VMD)
Bat algorithm (BA)
Coati optimisation algorithm (COA)
Cuckoo search algorithm (CSA)
Sparrow search algorithm (SSA)
Long short term memory (LSTM)
Recurrent neural network (RNN)
Mean absolute error (MAE)
Root mean square error (RMSE)
Mean absolute percentage error (MAPE)
Wind power forecasting (WPF)
Numerical weather forecasting (NWF)
Improved variational mode decomposition (IVMD)
Intrinsic mode function (IMF)

1. Introduction

The instability and unsafe of wind power cause the fluctuation of wind power output, resulting in unstable power generation [1]. This is because wind power is affected by many factors, such as wind speed, meteorological conditions, and equipment status. If the grid is unable to accurately predict wind power generation, it can lead to an imbalance between supply and demand on the grid, triggering voltage instability, frequency bias, or even a breakdown of the power system. Through accurate wind power forecasting (WPF) the grid operation strategy can be adjusted accordingly. So in order to realise the economical and safe of the power grid, accurate, and rapid forecasting of wind power is required [2]–[5].

In the early stages, WPF technology used wind speed from numerical weather forecasting (NWF) as input for forecasting. However, the forecasting accuracy depended solely on the accuracy of NWP, and the impact of terrain and weather conditions in areas where the terrain was complex and the climate was changeable led to significant errors in NWF, resulting in large forecasting errors.

Models based on time series, such as those developed by reference [6]–[8], have alleviated the issue of decreasing WPF accuracy over time. Reference [9] used artificial neural networks to establish forecasting models that improve the information exchange capabilities between forecasting units and, consequently, increase single forecasting accuracy. References [10], [11] established forecasting models based on the Kalman filtering method that showed some efficacy in handling noise in wind power data. However, predicting wind power is challenging because it is influenced by numerous factors, such as time, space,

* School of Electrical and Information Engineering, Hunan Institute of Engineering, Xiangtan, China 411101; e-mail: {hf7825, 247570654, 374402091, 1691455049, 598460233}@qq.com
** Harbin Electric Corporation Wind Power Co., LTD, Xiangtan, China 411101; e-mail: lingxiang.huang@hewp.com
Corresponding Author: Feng Huang

(DOI: 10.2316/J.2024.203-0530)

and wind turbine conditions. Furthermore, wind turbine factors have many subordinate factors [12], which makes it difficult to consider all factors in one model. As a result, current WPF models predominantly emphasise improving individual factors to enhance forecasting accuracy, and there remains a need for research to identify and establish comprehensive forecasting models that integrate multiple factors to improve overall forecasting accuracy.

At present, WPF consists of four predominant research domains: forecasting time scale, forecasting spatial scale, forecasting object, and forecasting model. In the realm of forecasting time scale, researchers have proposed a system of classification, as suggested by reference [13], which arranges WPF into four distinct time increments: long-term forecasting, medium-term forecasting, short-term forecasting, and ultra-short-term forecasting. In the arena of forecasting spatial scale, reference [14] has proffered four specific modes of forecasting: single turbine power forecasting, wind farm power forecasting, single wind field power forecasting, and wind power cluster power forecasting. Regarding the forecasting object domain, reference [15] has advocated an approach in which researchers conduct indirect forecasting of wind speed and indirect forecasting. Finally, the domain of forecasting models can be divided into several different categories, including the physical forecasting model elucidated in references [16], [17], the statistical learning forecasting model described in references [18], [19], and the nascent development of the combination forecasting model in the last decade. Based on contextual constraints, a combination model has been chosen to coordinate ultra-short-term forecasting for a single turbine. Reference [20] designs a new ultra-short-term forecasting method considering some characteristics of the field group, act as the unit state, the wind turbine wake flow. But unfortunately, it loses the change of power series or wind speed series. Reference [21] proposes a new dynamic combination forecasting model of the new observation information and self-adaptive index on accounted of the weight ratio of historical wind data. The new model reduce the error of forecasting to a certain extent.

Recent years have seen research that demonstrates the efficacy of fusing multiple forecasting models, thus combining their respective benefits to enhance WPF accuracy. Some machine learning algorithms have strong effectiveness in power forecasting [22]. Reference [23] proposes a new forecasting model based on LSTM, which uses correlation in time series for long-term, and solve some long term learning problems of classical neural networks. Reference [24] proposes a hybrid wavelet LSTM model for fault forecasting in electrical power grids that has a higher predictive capacity. Deal with short-term wind and solar power forecasting. Reference [25] designs a new COA-CNN-LSTM model to deal with short-term WPF. Reference [26] proposes a new LSTM neural networks forecasting model combining the optimised Cuckoo search algorithm (CSA). Reference [27] uses a Sparrow search algorithm (SSA) to determine number of neurons and the learning rate in LSTM and CNN-LSTM networks. However, intelligent optimisation algorithms, such as PSO algorithm, Cuckoo algorithm, and Bat algorithm (BA), is often difficult to obtain global optimal solutions and need to set niche parameters for solving problems.

Wind power data is characterised by fluctuations and non-stationarity. The accuracy of predicting oscillation points is insufficient. Constructing a hybrid prediction model can achieve better results. A new combination forecasting algorithm is designed here, which integrates variational mode decomposition (VMD), Bat algorithm improved particle swarm optimisation (BAIPSO) and gated recurrent unit (GRU) neural network model (VMD-BAIPSO-GRU). Firstly, the VMD is used to decompose the power data into a series of relatively stable sub-signals, highlighting the local feature information of the data. Then, a GRU neural network model is established, and the BA is introduced to improve the PSO process and optimise the hyper parameters of the GRU. The new combination forecasting model demonstrates higher forecasting accuracy in comparison to other algorithms of similar nature with the data of 23rd turbine at the Dabangling Power Plant of Harbin Electric Corporation Wind Power Co., Ltd. The proposed algorithm is particularly effective when forecasting the real wind power output of a single turbine.

The paper Section 1 introduces the preprocessing of wind power, and Section 2 presents the principles of VMD, BAIPSO, and GRU, as well as the configuration of the combination model. Finally, Section 3 validates the effectiveness of the new VMD-BAIPSO-GRU model. Finally, summarise the entire paper.

2. Preprocessing of Wind Power Data

Due to equipment failures and other factors, the original dataset often contains exceptions and missing values. The method employed for missing wind power data is nearest neighbour interpolation. Additionally, the missing value is replaced with the value that is closest to the mean of the attribute column by comparing the preceding and succeeding data points.

The Isolation Forest algorithm is used for detecting power anomalies. The power data used is from the 23rd turbine at the Dabangling Power Plant of Harbin Electric Corporation Wind Power Co., Ltd. The mean of the consecutive two observed values is used for correction, and the resulting outliers are shown in Fig. 1.

The method used for wind power data transformation is minimum–maximum normalisation.

$$P = (p - p_{\min}) \,/\, (p_{\max} - p_{\min}) \tag{1}$$

where, p = the power at the current time; $p_{\min}$ = the minimum value of wind power; $p_{\max}$ = the maximum value of power, and P = the normalised power result.

After data normalisation, it is necessary to restore it to the original data. The standardisation function for minimum and maximum normalisation is shown in (2).

$$p_{\text{pre}} = {p'}_{\text{pre}} \,(p_{\max} - p_{\min}) + p_{\min} \tag{2}$$

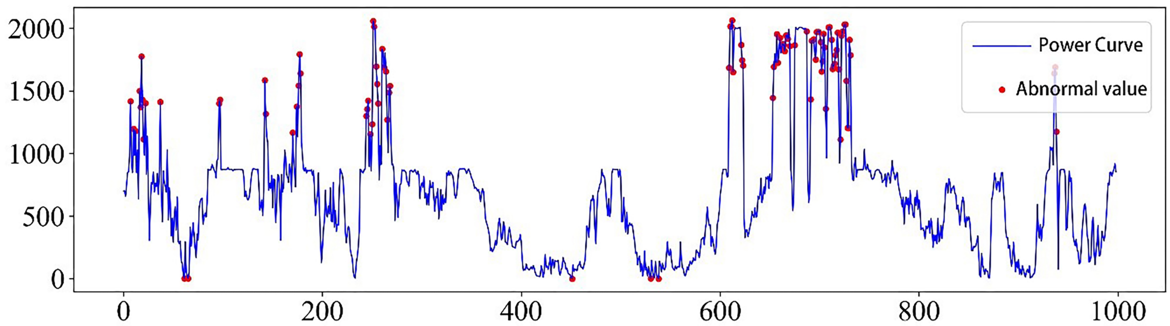

Figure 1. Outliers of wind power data.

where p_{pre} = the restored WPF value; p'_{pre} = the unprocessed WPF value.

3. Components of the Combination Model

3.1 Improved Variational Mode Decomposition (IVMD)

In 2014, the VMD algorithm was proposed [28] VMD transforms the original wind data into a number of intrinsic mode function (IMF). It can solve the optimal solution and the variational problem with its central frequency of each IMF. The operational steps of VMD are as follows.

3.1.1 Establish the Variational Problem

(1) To calculate the marginal spectrum of each mode function $u_k(t)$ by the Hilbert transform and obtain a one-sided frequency spectrum:

$$\left[\delta(t)+\frac{j}{\pi \mathrm{t}}\right] * u_k(t) \tag{3}$$

(2) Mix the exponential terms corresponding to the central frequency ω_k of each mode function to condition the frequency spectrum of the mode function into the fundamental modulation band:

$$\left[\left[\delta(t)+\frac{j}{\pi \mathrm{t}}\right] * u_k(t)\right] e^{-j\omega_k t} \tag{4}$$

(3) Find the bandwidth corresponding to each mode component and transform the target problem into a variational one with constraints:

$$\frac{\min_{\{u_k w_k\}}\left\{\sum_k \left\| \ \delta_t\left[\left(\frac{j}{\pi \mathrm{t}}+\delta(t)\right)u_k(t)\right]e^{-jw_k t}\right\|^2\right\}}{s.t. \sum_k u_k = f} \tag{5}$$

In above (5), $\{u_k\} = \{u_1, \ldots, u_k\}$ represents the decomposed mode component, and $\{w_k\} = \{w_1, \ldots, w_k\}$ is the central frequency of the mode function.

3.1.2 Solve the Variational Problem

(1) Use the Lagrange multiplier operator λ to transform (5) into an unconstrained variational problem. The expression is as follows:

$$\begin{aligned} &L\left(\{u_k\},\{w_k\},\lambda\right) \\ &\quad= \alpha\sum_k \left\|\partial_t\left[u_k(t)\left(\frac{j}{\pi \mathrm{t}}+\delta(t)\right)\right]e^{-jw_k t}\right\|^2 \\ &\qquad+\left\|f(t)-\sum_k u_k(t)\right\|_2^2 \\ &\qquad+\left\langle \lambda(t), f(t)-\sum_k u_k(t)\right\rangle \end{aligned} \tag{6}$$

(2) Take the solution of the previous minimisation problem in (5) as the solution of the augmented Lagrange expression (6). By putting to use the alternating direction multiplier method, update u_k^{N+1}, w_k^{N+1}, and λ^{N+1} iteratively. The updating of the mode u_k becomes an equivalent minimisation problem:

$$\begin{aligned} \widehat{u}_k^{N+1} = \underset{\widehat{u}_k, u_k\in X}{\operatorname{argmin}} \Bigg\{ \alpha \Bigg\| \Bigg\| \widehat{f}(w) - \sum_i \widehat{u}_i(w) + \frac{\widehat{\lambda}(w)}{2}\Bigg\|_2^2 \\ +jw\left[(sgn(w+w_k))\widehat{u}_k(w+w_k)+1\right]\Bigg\|_2^2\Bigg\} \end{aligned} \tag{7}$$

Replacing w with $w - w_k$ in the above equation, the non-negative frequency interval integration form becomes:

$$\begin{aligned} \widehat{u}_k^{N+1} = \underset{\widehat{u}_k, u_k\in X}{\operatorname{argmin}} \Bigg\{ \int_0^\infty 4\alpha\left|\widehat{u}_k(w)\right|^2(w-w_k)^2 \\ +2\left|f(w)+\frac{\widehat{\lambda}(w)}{2}-\sum_i \widehat{u}_i(w)\right|^2 \mathrm{dw}\Bigg\} \end{aligned} \tag{8}$$

The solution to this is:

$$\widehat{u}_k^{N+1}(w) = \frac{\widehat{f}(w)-\sum_i \widehat{u}_i(w)+\frac{\widehat{\lambda}(w)}{2}}{1+2\alpha(w-w_k)^2} \tag{9}$$

The central frequency w_k does not appear in the reconstructed function fidelity term but only in the previous bandwidth term. The equivalent minimisation

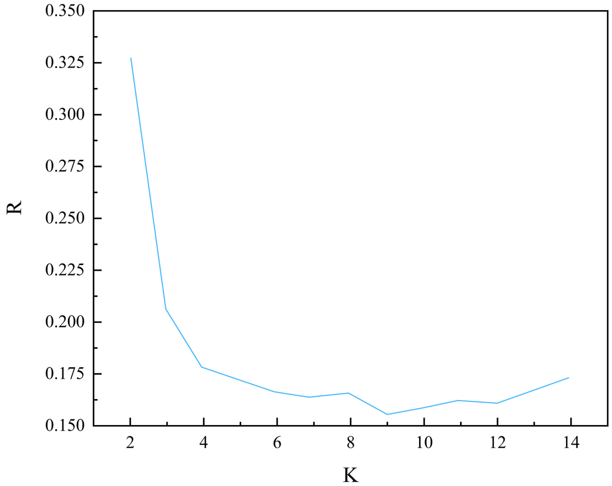

Figure 2. Trend chart of residual correlation coefficient of VMD.

expression for w_k is:

$$w_k^{N+1} = \underset{w_k}{\operatorname{argmin}} \left\{ \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi \mathrm{t}} * u_k(t) \right) \right] e^{-jw_k t} \right\|_2^2 \right\} \tag{10}$$

The central frequency:

$$w_k^{N+1} = \underset{w_k}{\operatorname{argmin}} \left\{ \int_0^\infty \|\widehat{u}_k(w)\|^2 (w - w_k)^2 \, \mathrm{dw} \right\} \tag{11}$$

The updated expression for the central frequency is obtained as follows:

$$w_k^{N+1} = \frac{\int_0^\infty w \left|\widehat{u}_k(w)\right|^2 \mathrm{dw}}{\int_0^\infty \left|\widehat{u}_k(w)\right|^2 \mathrm{dw}} \tag{12}$$

where, $\widehat{u}_k^{N+1}(w)$ is the Wiener filtering of the remaining components at the current stage; w_k^{N+1} is the centroid of the power spectrum at the current stage; The real part of $\widehat{u}_k(w)$ is $\{u_k(t)\}$ by inverse Fourier transform.

To ensure sufficient decomposition of the wind power sequence, the number of modes K can be determined by the Pearson coefficient R_k between the residuals obtained by VMD decomposition and the wind power sequence.

For VMD decomposition of historical power generation data, the residual correlation coefficients for different numbers of modes K are shown in Fig. 2, where the horizontal axis K represents the number of decomposition modes, and the vertical axis R represents the abbreviation for the residual correlation coefficient R_K. When $K < 9$, the R_K shows a decreasing trend as a whole, indicating insufficient decomposition of the power time series. When $K > 9$, the R_K starts to rise, indicating over-decomposition of the wind power time series. The value of K is determined to be 9.

3.2 Bat Algorithm Improved Particle Swarm Optimisation (BAIPSO)

3.2.1 IPSO Optimisation Algorithm

Eberhart and Kennedy proposed PSO algorithm in 1995, which was based stochastic optimisation technique. The PSO algorithm mimics the swarm behaviour of herds of animals, insects, fish, and birds. The groups search for food in a cooperative manner, and each member in the group continuously changes its search mode by learning its own experience and the experiences of other members.

To improve the global search ability of the PSO algorithm in the early stages of iteration and the local search ability in the later stages, a nonlinearly decreasing inertia weight is introduced for improvement, as follows [29]:

$$w = w_{\min} + (w_{\max} - w_{\min}) \times 2exp\left(-\alpha \left(\frac{t_{\max}}{t}\right)^4\right) \tag{13}$$

where, α is the constant coefficient; t is the current number; $t_{\max}$ is the maximum one.

3.2.2 Bat Algorithm Improved Particle Swarm Optimisation (BAIPSO)

The principle of the BA is to consider each individual in the population of bats as a feasible solution in space, and to view the process of bats searching for prey and flying behaviour as a process of random search and optimisation. The fitness function to be solved is considered as a criterion for judging the feasibility of the bat in space, and the algorithm's retention or elimination of feasible solutions corresponds to the laws of natural selection of individual bats in competitive survival.

Set the flight speed of each bat at position X_i to V_i, the emission frequency of the pulse to f_i, the loudness of the pulse to A_i, and the emissivity of the pulse to r_i, η is a random number in [0,1], $f_{\max}$ and $f_{\min}$ are the maximum frequency and the minimum frequency, respectively:

$$\begin{aligned} V_{\mathrm{id}}^{t+1} &= w_i V_{\mathrm{id}}^t + f_i (P_{\mathrm{best}} - X_{\mathrm{nd}}^t) \\ &\quad + c_2 r_2 (G_{\mathrm{best}} - X_{\mathrm{nd}}^t) && (14) \\ X_{\mathrm{id}}^{t+1} &= f_i X_{\mathrm{id}}^t + V_{\mathrm{id}}^{t+1} && (15) \\ f_i &= f_{\min} + \eta (f_{\max} - f_{\min}) && (16) \end{aligned}$$

Assuming [30]: 1) each bat uses echolocation for localisation analysis, and can also accurately identify obstacles through flight; 2) when approaching prey, bats emit pulse sound waves and adjust themselves; 3) the main way to adjust the volume of sound waves is to gradually adjust from the maximum volume $A_{\max}$ to the minimum volume $A_{\min}$, based on these three assumptions and using some approximations for simplification, the steps of BAIPSO can be summarised as follows:

(1) Initialise parameters, determine the objective function, set the iteration times of the algorithm, the population size of the bats, the position X_i^0, velocity V_i^0 , and pulse emission frequency of each bat r_i^0 $(i = 1, 2, \ldots, n)$.

(2) Calculate the fitness value, find the bat individual with the best position in the population, and update the position and velocity.
(3) Generate a random number rand$_1$ within the range of $[0, 1]$. If rand$_1 > r_i^t$, select the best bat individual and generate a local solution near the current best individual. Otherwise, go to step (2).
(4) Generate a random number rand$_2$ within the range of $[0, 1]$. If rand$_2 < A_i^t$, and the fitness value of the objective function at the time is better than the local solution found in step (3), then accept the new solution.
(5) Update the global best solution, determine whether the maximum iteration times have been reached, if the maximum iteration times have been reached, output the global best solution, otherwise go to step (2).

During the operation of BAIPSO algorithm, each particle is able to use echolocation to accurately identify its own position. As a result, all particles are able to complete the replacement of their positional information within the particle swarm, and record the optimal positional information. This facilitates the precise determination of a particle's position as it arrives at a new location, thereby enhancing its local search ability.

3.3 GRU Recurrent Neural Network

LSTM is a special type of RNN. By using a gating mechanism, it solves the problem of gradient vanishing and avoids the long-range dependency problem when dealing with long-term sequences. The GRU is a variant of LSTM proposed by Cho *et al.* in 2014. Structurally, GRU has the same input and output structures as a regular recurrent neural network. Each unit receives the current raw input value and the output value from the previous unit, and after internal processing, the output information vector serves as the input value for the next neural unit. The difference lies in the optimisation of the gate apparatus based on LSTM, which simplifies the internal processing of the neural unit.

In LSTM, forget gate, output gate and input gate, control memory, output, and input. The GRU model simplifies the structure with only two gating structures: update gate and reset gate. The update gate can be used to control the size of information from the previous moment entering the current moment. For the update gate, the larger the value, the more information from the previous moment will be saved. The reset gate controls how much information will be included in the current candidate set at the previous time. For the reset gate the smaller the value, the less information will be included in the previous time.

3.4 The Steps of VMD-BAIPSO-GRU Forecasting Model

The paper proposes a new VMD-BAIPSO-GRU forecasting model. The steps of VMD-BAIPSO-GRU model:
(1) Step one of the model: The raw power time series data is cleaned by completing missing values and removing outliers. The data is then normalised using MinMaxscale method to improve data quality and model convergence speed.
(2) Step two of the model: The wind power time series data is decomposed using VMD, resulting in K modalities with different central frequencies. The VMD decomposed sub-sequences are then divided into training sets and testing sets.
(3) Step three of the model: The BAIPSO-GRU model parameters are initialised and optimised using training data. The optimal parameters obtained through BAIPSO optimisation are then assigned to the GRU neural network. The optimised GRU is then used to train and forecast all sub-datasets. The forecastings of all sub-datasets are then reconstructed and reverse-normalised to obtain the final forecasting values.
(4) Step four of the model: The forecasting performance is evaluated using the MAE, MAPE, and RMSE coefficients, using the following formulas:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y})_i^{\,2} \tag{17}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \widehat{y}_i}{y_i}\right| \tag{18}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - y_t')^2} \tag{19}$$

where n is the number, the forecasting value is $\widehat{y} = \{\widehat{y}_1, \widehat{y}_2, \widehat{y}_3, \ldots, \widehat{y}_n\}$ and the test value is $y = \{y_1, y_2, y_3, \ldots, y_n\}$.

4. Validation of Combination Model

By using VMD to decompose wind power data, the instability and nonlinearity of the data can be reduced. GRU has advantages in time series prediction, especially in power prediction. The use of IPSO can improve convergence speed and accuracy but it is prone to getting trapped in local optima and cannot find the global optimal solution. The BA, with its resistance to interference and high computational rate, can be used to improve the efficiency of GRU parameter optimisation and find the global optimal parameters.

4.1 Analysis of VMD Decomposition Results

The historical wind power data was decomposed using VMD into nine sub-sequences with different frequencies, and the nine IMF components obtained from VMD are shown in Fig. 3.

It can be observed that the average amplitude of sub-sequence IMF1 is relatively large and changes smoothly. Sub-sequences IMF 2-IMF 4 exhibit good regularity and significant periodicity. Sub-sequences IMF 5-IMF 9 exhibit large fluctuations in amplitude, but still exhibit obvious regularity and periodicity. This suggests that the choice of $K = 9$ for the number of sub-sequences obtained from VMD decomposition is ideal.

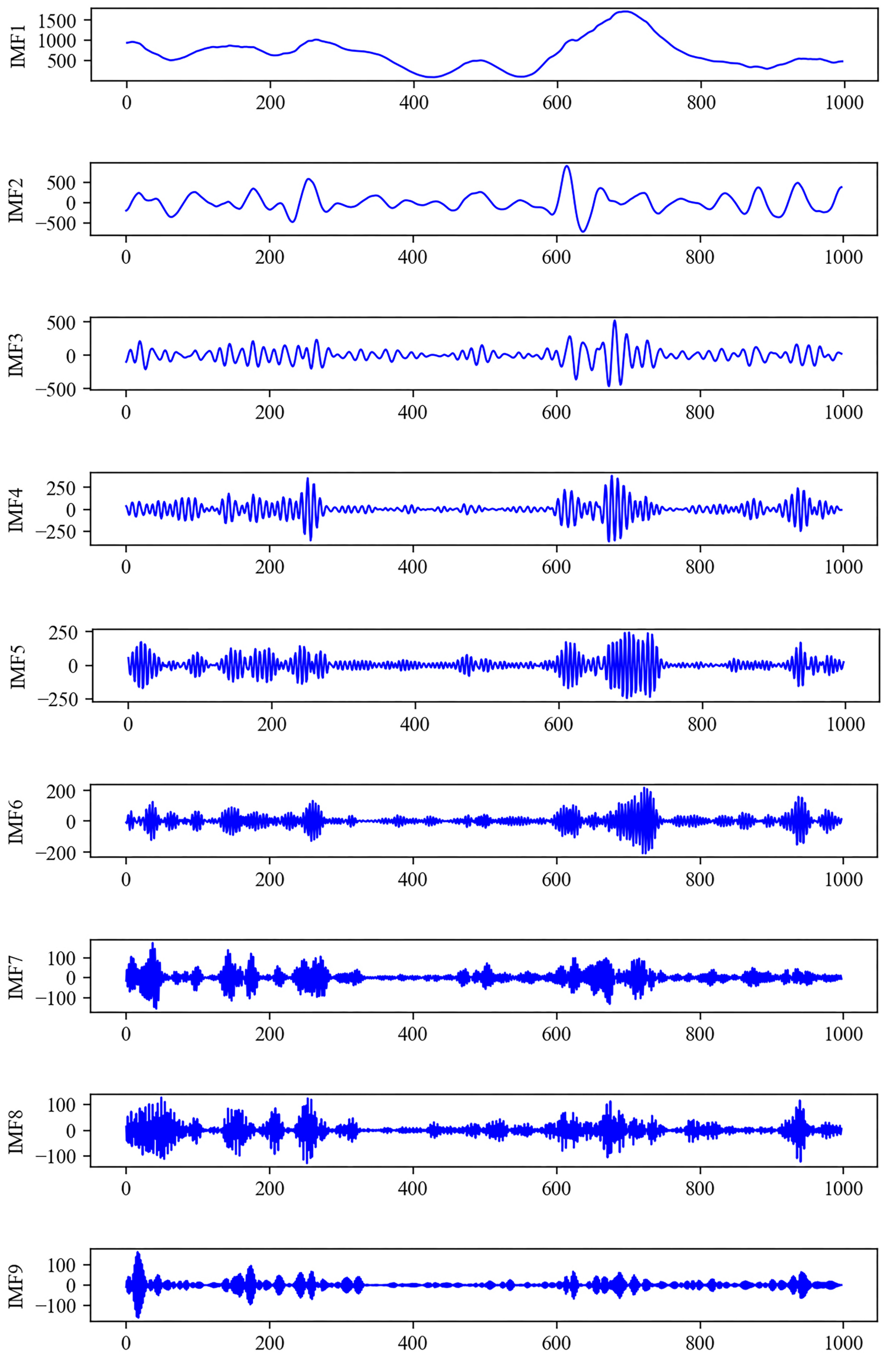

Figure 3. VMD decomposition vectors IMF 1–IMF 9.

4.2 BAIPSO Algorithm Testing

To test the effect of the algorithm, six test functions were used and the BAIPSO algorithm was compared with the basic PSO. The results of algorithm operation are shown in Table 1:

The test functions include single peak functions $f_1 \sim f_3$, which are relatively simple, and multi peak functions $f_4 \sim f_6$, which are relatively complex. Except for the function f_3, the optimal position for all other test functions is $[0]^{\mathrm{n}}$, and the optimal value for all six test functions is 0. This indicates that as long as the optimal position is reached, the function will converge to 0. The parameters for both BAIPSO and original PSO algorithms were set as follows: population size of $m = 30$, maximum iteration number of 1,000, and learning factor $c_1 = c_2 = 1.98$.

Both algorithms were tested with a dimension of $D = 10$ and $D = 20$ using the test functions for 20 iterations. The optimisation statistics are shown in Table 2. Through the data we can compare and analyse the performance of the BAIPSO and the original PSO.

It can be seen from Table 2 that the result of the BAIPSO algorithm is relatively poor when optimising the function f_3, but it also converges near the optimal value, and the optimisation accuracy is better than the accuracy of original PSO algorithm. The BAIPSO algorithm has higher accuracy than the original PSO algorithm for all

Table 1
The Results of Two Algorithms Operation

Test functions	Function name	Value range	Optimal value position	Theoretical value
f_1	Phere	$[-100{,}100]^n$	$[0]^n$	0
f_2	Schwefel2	$[-10{,}10]^n$	$[0]^n$	0
f_3	Rosenbrock	$[-30{,}30]^n$	$[0]^n$	0
f_4	Griewank	$[-600{,}600]^n$	$[0]^n$	0
f_5	Ackley	$[-32{,}32]^n$	$[0]^n$	0
f_6	Rastrigin	$[-5.12{,}5.12]^n$	$[0]^n$	0

Table 2
Results Comparison

Functions	Dimension	PSO		BAIPSO	
第1章	第1章	Mean value	Standard deviation	Mean value	Standard deviation
f_1	10	3.19E+001	7.72E-0002	5.22E-006	1.43E-006
	20	1.37E+001	5.21E-002	3.27E-007	1.19E-007
f_2	10	1.54E-001	2.27E-002	4.63E-015	1.28E-012
	20	1.36E-001	2.11E-002	3.63E-015	2.48E-014
f_3	10	3.68E+002	4.68E+001	8.91E+000	3.21E+000
	20	2.49E+001	2.35E+001	1.35E+000	6.41E+001
f_4	10	2.90E+002	6.51E+001	1.47E-013	3.71E-013
	20	1.39E+002	5.14E+001	2.81E-015	4.32E-015
f_5	10	4.48E-001	6.27E-001	3.68E-009	1.67E-009
	20	2.83E-001	2.36E-001	2.65E-011	1.88E-011
f_6	10	3.68E-001	2.50E-001	1.32E-013	1.61E-013
	20	5.74E-001	3.89E-001	2.54E-015	6.28E-015

test functions, and also has higher accuracy in terms of function optimisation.

In conclusion, the BAIPSO algorithm has a stronger search ability, faster speed, higher precision than the original PSO algorithm.

4.3 Testing of VMD-BAIPSO-GRU Wind Power Forecasting Model

The testing was conducted with a sampling time interval of 10 min, for a total of 200 points after preprocessing. The first 160 points were selected as the training set for the forecasting model, while the last 40 points were used as the test set for forecasting, with a forecasting time window of 20 min.

The evaluation metrics for each model are shown in Table 3. It can be observed from the Fig. 4 that both the BA-GRU and IPSO-GRU models can predict the general trend of wind power, but their accuracy is not very good. However, after decomposing the original data with VMD, the forecasting accuracy improved and is closer to the actual values. In the last little figure, the forecasting values are closer to the true values.

Table 3
Evaluation Indicators for Each Forecasting Model

Models	MAE	MAPE	RMSE
BA-GRU	0.52	34.53%	0.72
IPSO-GRU	0.73	42.69%	0.97
BAIPSO-GRU	0.44	22.73%	0.64
VMD-BAIPSO-GRU	0.37	12.09%	0.22

It can be seen from Table 3 that the new VMD-BAIPSO-GRU combination WPF model has the closest forecasting results to the actual values, with lower error indicators than the other models. The MAE of new forecasting model is 0.37 smaller than MAEs of other

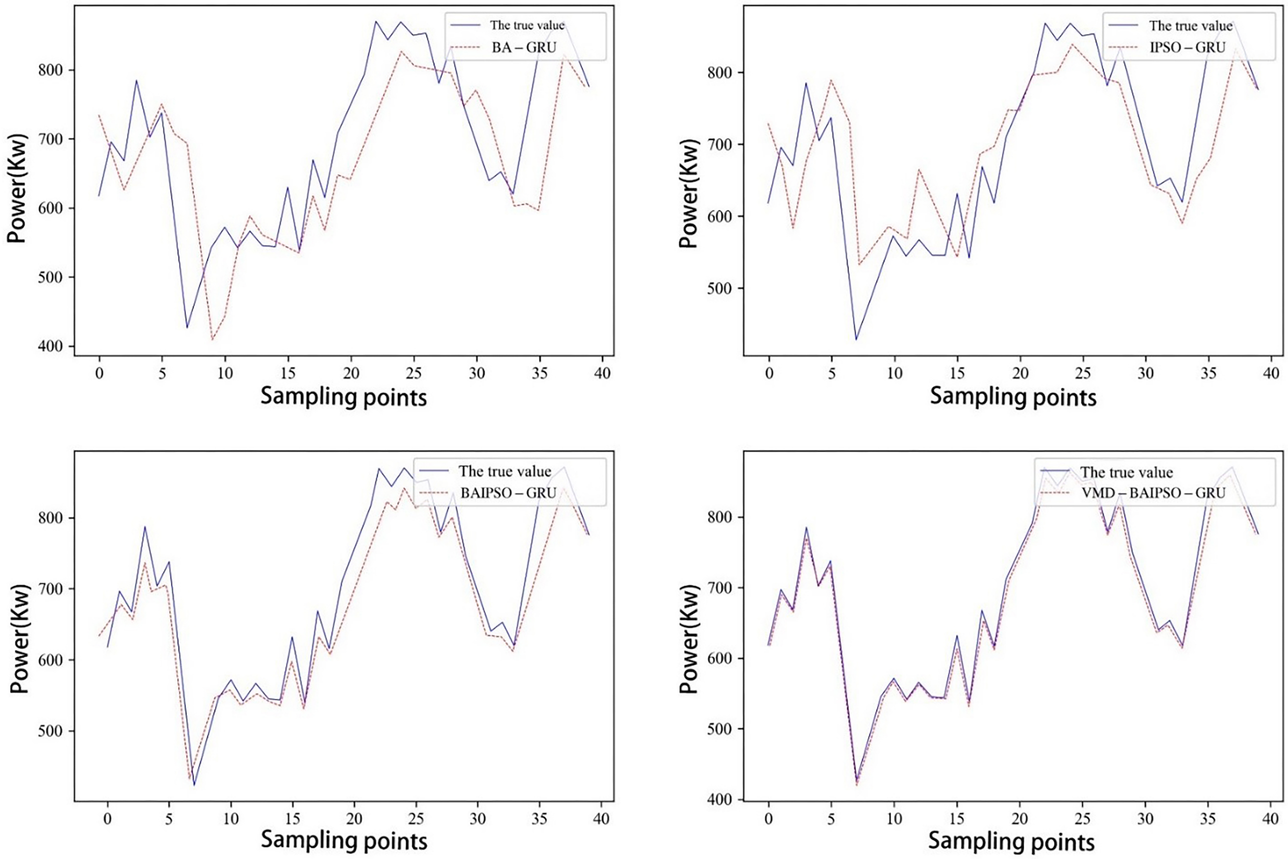

Figure 4. The forecasting and testing values of the model.

forecasting 0.52, 0.73, 0.44, 0.37, MAPE is 12.09% smaller than MAPEs of other forecasting 34.53%, 42.69%, 22.73%, and RMSE is 0.22 smaller than other forecasting models RMSE of 0.72, 0.97, 0.64.

In summary, the combination of BA and IPSO algorithm greatly improves the optimisation performance of the origin WPF algorithm, and is not easily trapped in a local optimal solution. Compared to the other three modes this new model has obvious advantages. From the forecasting results, the new VMD-BAIPSO-GRU is a well-performing model.

5. Conclusion

In this study, it proposed an IPSO algorithm for GRU WPF. Several analyses and tests were conducted. Firstly, forecasting analysis was carried out on each component of the VMD decomposition with the optimal K-value. Secondly, tests and comparisons were conducted on the BAIPSO algorithm, and its optimisation performance was verified by comparing its convergence effect with that of the original PSO algorithm by six test functions. Finally, the performance of the VMD-BAIPSO-GRU forecasting algorithm was compared and analysed with three other forecasting algorithms, to verify its feasibility. The VMD-BAIPSO-GRU algorithm has a lower error and higher forecasting accuracy.

(1) The VMD decomposition can accurately explore hidden characteristics in wind power data. The forecasting model has better forecasting performance compared with BAIPSO-GRU, which proves the effectiveness of original wind power data based on VMD decomposition processing.
(2) The VMD-BAIPSO-GRU forecasting model is based on the GRU forecasting, and the BAIPSO algorithm is added for optimising hyperparameter to get the smallest error. By comparing the forecasting results of the other three modes, it is proved that the forecasting accuracy of VMD-BAIPSO-GRU is better than BAIPSO-GRU. It does verify the validity of hyper parameter optimisation based on BAIPSO.
(3) By further in-depth and objective comparison analysis of the scientific and rational effects of the VMD-BAIPSO-GRU forecasting model and the BAIPSO-GRU forecasting model, combined with the previous findings, it can get the conclusion that the new forecasting model presented in the paper has the best forecasting results with highest forecasting accuracy. The MAE of forecasting results was reduced to 0.37, MAPE decreased to 12.09%, and RMSE decreased to 0.22.

Currently, the forecasting accuracy of this model is greatly affected by historical data information. In the next step, various factors, such as weather and temperature in the actual operation of wind turbines can be used as independent variables input into the model, making the forecasting model more perfect and having greater practicality.

Acknowledgement

This work was supported by National Natural Science Foundation of China (grant number 62006075), Hunan Provincial Natural Science Foundation of China (grant numbers 2022JJ50014 and 2022JJ50116), and Special Project for Construction of Changzhutan National Independent Innovation Demonstration Zone (grant numbers CG-YB20211030 and CG-ZDGS20221004).

Data Availability Statement

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

References

[1] Y. Wang, Analysis of the current situation and development trend of wind power generation, *Electric Power Equipment Management, 11*, 2020, 21–22.

[2] Xu Mengtian, Wang Hongzhe, Zhao Chengping, Yan Hua, Research on data preprocessing policy based on short-term wind power prediction, *Renewable Energy Resources, 37*(01), 2019, 119–125.

[3] C. Ma, Research on development status of wind power energy and its grid-connected technology, *Hunan Hydro & Power, 2*, 2020, 65–69.

[4] S. Wang and D. Xie, A simple analysis of the impact of grid-connected wind power on power quality, *Electronic Technology & Software Engineering, 23*, 2018, 220–221.

[5] M. Moradi, Y. Weng, and Y.C. Lai, Defending smart electrical power grids against cyberattacks with deep Q-learning, *P R X Energy, 1*, 2022, 033005.

[6] Maheswari Chenniappan, Divya Gnanavel, Kavi Priya Gunasekaran, R.R. Rajalakshmi, A.S Ramya, Albert Alexander Stonier, Geno Peter and Vivekananda Ganji, Prediction of fault occurrences in smart city water distribution system using time-series forecasting algorithm, *Mathematical Problems in Engineering*, *2022*, 2022, 9678769.

[7] L. Changro and P. Keith Key-Ho, Forecasting trading volume in local housing markets through a time-series model and a deep learning algorithm, *Engineering Construction and Architectural Management, 29*(1), 2022, 165–178.

[8] Zhang Xiaojie, Cai Wenchuan, Gan Zhongxue, Coordinated control strategy for active power of wind power cluster based on model, predictive control, *Proceedings of the CSEE*, *41*(17), 2021, 5887–5899.

[9] M.O. Moreira, P.P. Balestrassi, A.P. Paiva, P.F. Ribeiro, B.D. Bonatto, Design of experiments using artificial neural network ensemble for photovoltaic generation forecasting, *Renewable and Sustainable Energy Reviews*, *135*, 2021, 110450.

[10] Q. Zhu, Y. Wang, and Y. Luo, Improvement of multi-layer soil moisture prediction using support vector machines and ensemble Kalman filter coupled with remote sensing soil moisture datasets over an agriculture dominant basin in China, *Hydrological Processes, 35*(4), 2021, e14154.

[11] M. Sompop and C. Nawinda, Short-term forecasting of renewable energy consumption: Augmentation of a modified grey model with a Kalman filter, *Applied Soft Computing Journal, 87*, 2020, 105994.

[12] Yang Xiaofeng, Fang Yihang, Zhao Pengzhen, Wang Chenming, Xie Ning, Non-mechanism modeling method for state recognition of wind turbine, *Electric Power*, Aug. 2022.

[13] N. Korprasertsak and T. Leephakpreeda, Robust short-term prediction of wind power generation under uncertainty via statistical interpretation of multiple forecasting models, *Energy, 180*, 2019, 387–397.

[14] L. Zi and L. Xiaolei, Wind power forecasting of an offshore wind turbine based on high-frequency SCADA data and deep learning neural network, *Energy, 201*, 2020, 117693.

[15] Z. Gan, C. Li, J. Zhou, and G. Tang, Temporal convolutional networks interval prediction model for wind speed forecasting, electric power systems research, *Electric Power Systems Research, 191*(4), 2021, 106865.

[16] Wu Dingan, Zhong JianWei, Wang Xinlei, Xiang Jiaguo, Zeng Fanwei, Hu kai, Chen Chen, Principal component analysis and long and short-term memory networks for power load forecasting, *Intelligent Processing and Application, 11*(08), 2021, 4751.

[17] Zhang Yunqin, Cheng Qize, Jiang Wenjie, Liu Xiaofeng, Shen Liang, Chen zehua, Photovoltaic power prediction model based on EMD-PCA-LSTM, *Acta Energiae Solaris Sinica*, *09*, 2021, 62–69.

[18] S. Sheng, H. Jin, and C. Liu, Short-term and mid-short-term wind power forecasting based on VMD-WSGRU, *Power System Technology, 46*(03), 2022, 897–904.

[19] W.S. Rosenthal, A.M. Tartakovsky, and Z. Huang, Ensemble Kalman filter for dynamic state estimation of power grids stochastically driven by time-correlated mechanical input power, *IEEE Transactions on Power Systems, 99*, 2017, 1–10.

[20] Fu Yang, Ren Zixu, Wei Shurong, Wang Yang, Huang Lingling, Jia Feng, Ultra-short-term power prediction of offshore wind power based on improved LSTM-TCN model, *Proceedings of the CSEE, 42*(12), 2022,4292–4303.

[21] L. Ye, Q. Zhu, and Y. Zhao, Dynamic optimal combination model considering adaptive exponential for ultra-short term wind power prediction, *Automation of Electric Power Systems, 39*(20), 2015, 12–18.

[22] Z.M. Zhai, M. Moradi, L.W. Kong, and Y.C. Lai, Model-free tracking control of complex dynamical trajectories with machine learning, *Nature Communications, 14*(1), 2023, 5698.

[23] K. Dragomiretskiy and Z. Dominique, Variational mode decomposition, *IEEE Transactions on Signal Processing, 62*(3), 2013, 531–544.

[24] N.W. Branco, M.S.M. Cavalca, S.F. Stefenon, and V.R.Q. Leithardt, Wavelet LSTM for fault forecasting in electrical power grids, *Sensors, 22*(21), 2022, 8323.

[25] M. Abou Houran, S.M.S. Bukhari, M.H. Zafar, M. Mansoor, and W. Chen, COA-CNN-LSTM: Coati optimization algorithm-based hybrid deep learning model for PV/wind power forecasting in smart grid applications, *Applied Energy, 349*, 2023, 121638.

[26] Y. Deng, J. Duan, and H. Jia, Ultra-short-term wind power prediction based on deep learning with independent recurrent neural network via cuckoo algorithm optimized, *Power System and Clean Energy, 37*(09), 2021, 18–26.

[27] D. Chang and X. Nan, Short-term photovoltaic power prediction based on back propagation neural network improved by hybrid sparrow algorithm, *Modern Electric Power, 39*(3), 2022, 12.

[28] K. Dragomiretskiy and D. Zosso, Variational mode decomposition, *IEEE Transactions on Signal Processing, 62*(3), 2014, 531–544.

[29] C. Li, *The research on wind power prediction based on the IPSO-BP neural network model* (Wuhan: Wuhan University of Science and Technology, 2016).

[30] X.S. Yang and X. He, Bat algorithm: Literature review and applications, *International Journal of Bio-Inspired Computation, 5*(3), 2013, 141–149.

Biographies

Feng Huang was born in 1978 in Shaoyang, Hunan, China. He received the Ph.D. degree in power electronics and power drives from Harbin Institute of Technology, P. R. China, in 2007. He is a Professor with the College of Electrical & Information Engineering, Hunan Institute of Engineering, Xiangtan, China. His research interests include image encryption and Automated Test System. He had over ten years' experience in teaching, research and development in projects and published over 30 scientific papers.

Rui Wang was born in 1997 in Liu Yang, Hunan, China. He received the bachelor's degree in electrical engineering and automation major from Hunan Institute of engineering in 2019. Now he is major in power engineering of Hunan Institute of engineering. His main research direction is short-term wind power prediction.

Yuelin Yu was born in 1998 in Xiangtan, Hunan, China. He received the bachelor's and master's degrees in electrical engineering from Hunan Institute of Engineering. He is currently working in Hunan Electrical College of Technology. His research direction is intelligent algorithms.

Feiyu Hu was born in 1997 in Shaodong, Hunan, China. She received the B.S. degree in electrical engineering from Changsha University of Science and Technology, China. She is currently pursuing a master's degree with Hunan Institute of Engineering. Her research is big data in power systems.

Xin Xie was born in 1996 in Guanxi, China. He received the B.S. degree in electrical engineering from Zhengzhou University, China. He is currently pursuing the master's degree with Hunan Institute of Engineering. His research is wind power prediction.

Lingxiang Huang was born in 1981 in Xiangtan, Hunan, China. He is the Vice President and a Senior Engineer of the Wind Energy Research Institute, Harbin Electric Corporation Wind Power Co., Ltd. As a head of the electrical control system, he mainly participates in the development and design of the electrical control system for 2 MW, 2.5 MW, and 5 MW series direct drive permanent magnet wind turbines.