

# Research of Short-Term Wind Speed Forecasting Based on the Hybrid Model of Optimized Quadratic Decomposition and Improved Monarch Butterfly

Gonggui Chen, Pan Qiu, Xiaorui Hu, Fangjia Long and Hongyu Long\*

**Abstract**—With the rapid consumption of fossil fuels, traditional power generation methods not only cannot issue future energy needs, but also bring serious environmental problems. As a clean and renewable energy, wind energy plays an increasingly important role in energy supply structure. However, the wind speed itself is intermittent, unstable and random, which brings severe challenges to wind power generation. Aimed at improving the accuracy and reliability of short-term wind speed forecasting, this paper proposes a new hybrid model. The model includes time-varying filter, modal decomposition, permutation entropy, adaptive noise modal decomposition, adaptive neuro-fuzzy inference system (ANFIS), packet data processing method neural network (GMDH), and improved monarch butterfly optimization algorithm (IMBO). First, the original wind speed sequence is significantly decomposed twice to obtain the sub-sequence to be predicted. Then, the reconstructed data uses ANFIS and GMDH neural network models to predict sequences in different frequency domains to get prediction results. To further improve the performance of the model, the improved monarch butterfly optimization algorithm is used to modify the model parameters. Finally, the final prediction result is obtained by summing the prediction results of each component. In addition, for verifying the performance of the model, this paper designs six sets of comparative experiments from two dimensions to verify the model on three data sets. The results show that the model proposed in this paper has high prediction accuracy and good stability.

**Index Terms**— wind speed prediction, neural network, secondary decomposition, data mining, IMBO, hybrid predictor

Manuscript received June 10, 2021; revised November 24, 2021. This work was supported by the National Natural Science Foundation of China under Grant 51207064.

Gonggui Chen is a professor of Key Laboratory of Industrial Internet of Things and Networked Control, Ministry of Education, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: chenggpw@126.com).

Pan Qiu is a master degree candidate of Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: 15084418828@163.com).

Xiaorui Hu is a professor level senior engineer of Marketing Service Center, State Grid Chongqing Electric Power Company, Chongqing 401123, China (e-mail: xiaorui4832@sina.com).

Fangjia Long is a senior engineer of State Grid Chongqing Electric Power Company, Chongqing 400015, China (e-mail: 157990759@qq.com).

Hongyu Long is a professor level senior engineer of Chongqing Key Laboratory of Complex Systems and Bionic Control, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (corresponding author to provide phone: +8613996108500; e-mail: longhongyu20@163.com).

## I. INTRODUCTION

WITH the rapid development of human society, the consumption of energy is also increasing day by day. According to the "2017 World Energy Outlook" issued by the IEA, total global energy consumption will increase by 30% by the end of 2040.

Due to the limited and non-renewable fuel, the energy supply is far from meeting the needs of future social development. In order to solve the problem of energy depletion, human beings have found that the emergence of renewable energy can cause tremendous changes in the energy consumption structure, and may become the pillar of future energy [1]. The Paris Agreement was adopted at the Paris Climate Change Conference on December 12, 2015. One of the goals of this agreement is to fully universalize the supply of decarbonized electricity by 2050, and wind power generation plays an indispensable role in achieving this goal. Fig. 1 shows the future global wind energy installation plan.

As a typical renewable energy, wind power generation has the advantages of clean, pollution-free, wide distribution, and mature technology. However, there are some unavoidable problems in the process of wind power generation, which is mainly reflected in the random, non-stationary, and non-continuous output of wind speed [2]. These problems have seriously affected the safety and operating efficiency of power grid installations. To solve these problems as much as possible, accurate short-term wind speed prediction is an effective way. Improving the accuracy of short-term wind speed forecasting can not only help the grid to carry out effective power dispatching, but also avoid a huge loss of grid supply, thereby achieving the purpose of improving economic benefits [3, 4].

From the perspective of forecasting methods, wind speed forecasting can be divided into three categories [5]: physical methods, statistical methods and hybrid methods. Shukul and Li integrated the ARIMA model into the Kalman filtering method and artificial neural network to improve the performance of wind speed prediction [6]. The support vector machine (SVM) method based on statistical theory is commonly used to solve nonlinear problems. Some experiments have proved that this method is superior to general artificial neural networks in short-term wind speed prediction [7].

Regional on shore wind and offshore  
wind outlook(GW)

The data source:  
Global Wind Energy Council

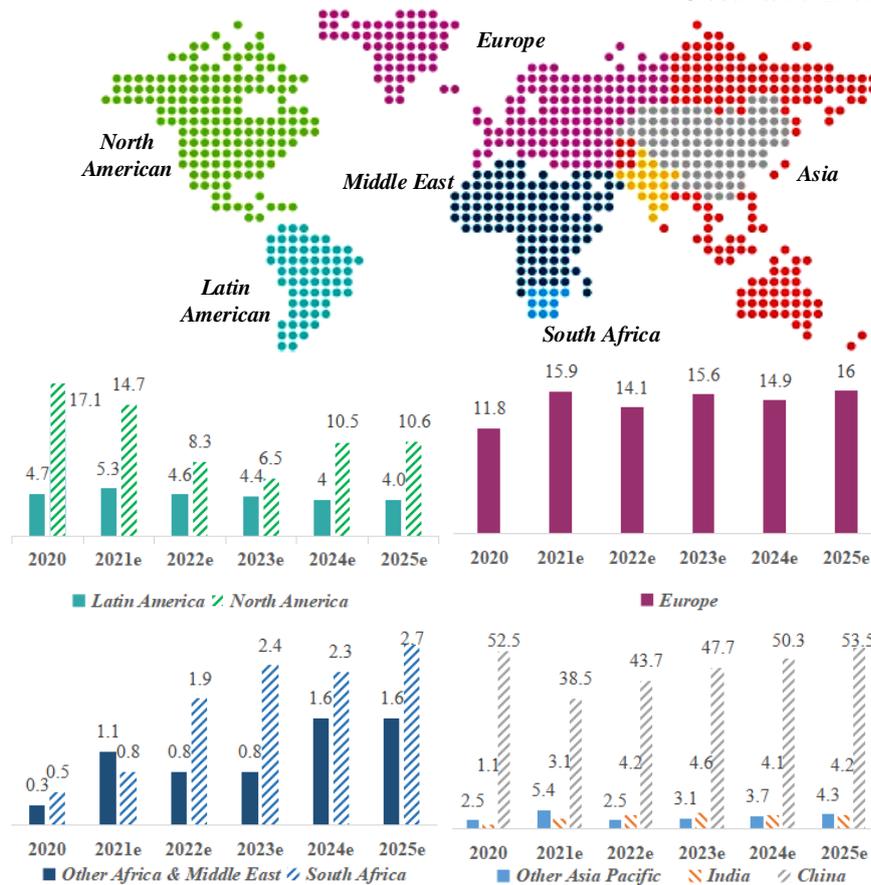


Fig. 1. Global wind energy installation plan

For example, in 2017, Jiang *et al.* proposed a hybrid model for short-term wind speed prediction based on the support vector machine (SVM) method, and achieved good prediction results [8]. Meanwhile, Zhang *et al.* used the genetic algorithm to optimize the parameters of least squares to support the vector of the machine. However, genetic algorithms are complicated to operate and are not suitable for overall modeling [9].

As the research continues to deepen, researchers have found that it is difficult to obtain forecast data with higher accuracy and precision only by relying on a single model. In the follow-up research process, a large number of experimental results show that the prediction results of the mixed model are far better than the single model. For example, Zhang *et al.* propose that the wind speed is predicted by different models, and the final prediction result is the product of the prediction results of each model and its weight coefficient [10]. Zhang's team combined EEMD (Ensemble Empirical Mode Decomposition)-CS (Cuckoo Search)-WNN (Wavelet Neural Network) to propose a new hybrid method of short-term wind speed prediction based on the cuckoo search optimization algorithm [11]. Zhang *et al.* proposed a wind speed prediction method based on the quadratic decomposition algorithm and Elman neural network [12]. The data decomposition strategy in the proposed model uses WPD (Wavelet Packet Decomposition) and FEEMD (Fast Integrated Empirical Mode Decomposition).

In the process of determining the weight coefficients of the mixed model, it is found that it is easy to fall into a local

optimal situation, which leads to deviations in the prediction results [13]. In order to solve this problem, the researchers found through experiments that intelligent optimization algorithms help to achieve higher accuracy in the hybrid forecasting model [14, 15]. In 2019, Jiang *et al.* proposed a hybrid wind speed prediction system based on fuzzy time series and intelligent optimization algorithms. Comparative experiment results show that the prediction results are better than the traditional hybrid model [16].

Other examples include the multi-step wind power prediction model based on the error factor of the single objective gray wolf algorithm and the integrated method proposed by Hao *et al.* [17]. To obtain a stable input sequence and make the prediction result better, the preprocessing of the data is also particularly important [18]. For example, the decomposition of wind speed series as a common data preprocessing method can make the input data more stable. The predictive model uses the pre-processed data to more easily capture the non-linear characteristics, which is beneficial to improve the precision and accuracy of the predictive results [19, 20].

The data pre-processing strategies adopted in this paper combine the empirical mode decomposition (TVF-EMD) based on time-varying filtering technology [21], permutation entropy improvement and the modal decomposition with adaptive noise (P-CEEMAND) [18, 22, 23]. Through the deep secondary decomposition of the original wind speed series, the time series obtained can better meet the

requirements of the model for the typical characteristics of the input data.

In this paper, the main contributions are as follows:

(1) Unlike most forecasting models that use a single raw data processing method. This paper adopts a new master-slave decomposition algorithm (TPD), which is based on time-varying filtering, permutation entropy to improve modal decomposition and adaptive noise decomposition to achieve secondary decomposition, thereby eliminating the interference caused by non-stationary factors in the original data. Make full use of the instantaneous amplitude and frequency information to adaptively design the local cut-off frequency.

(2) Based on effective learning strategies, a hybrid prediction model is established. More specifically, for improving model training efficiency and prediction accuracy, the model combines two learning methods of supervised learning and adaptive fuzzy neural network as a training model for data samples. The improved adaptive fuzzy inference system combined with GMDH neural network is mainly used for the prediction of complex systems and can deal with linear and nonlinear time series signals well.

(3) In order to avoid the proposed hybrid model from falling into the local optimum, the monarch butterfly optimization algorithm based on greedy strategy optimization is introduced on the basis of the model in this paper.

(4) For three sets of mean wind speed data collected with 15-minutes time stamps. In this paper, six different models are tested independently to verify the working effect of each model, and the prediction results of each model are compared and corresponding conclusions are drawn.

The original data set in this article originated from a wind farm in Jiangsu Province, China. The data contains real-time wind speed data at different times and places in different months for easy comparison. In the research of this article, the first 1500 samples of each site sample are used to train the proposed prediction model, 1201-1500 are used as the verification data, and the total of 500 samples 1501-2000 are as for test set. The three sets of original wind speed data are shown in Fig. 2.

## II. METHODOLOGY

In this part, it specifically incorporates the selection of wind speed sequence sites, the principle of data preprocessing related methods, the principle and structure of hybrid models and the principle of intelligent optimization algorithm. Finally, the realization processes of experimental prediction are given.

### A. The Running Flow of the Model

(1) First of all, on the basis of obtaining the original wind speed data, use the time-varying filter (TVF) optimized EMD for data decomposition. After discarding the high-frequency noise components, two main frequency component sequences are obtained.

(2) Secondly, after the two main frequency components are obtained by the previous decomposition, the CEEMDA grouping method based on permutation entropy is used to decompose the data for the second time. The data is decomposed into sub-sequences of different frequencies

according to the relationship of frequency and domain distribution.

(3) The obtained sub-sequences are divided into training sets and test sets. For each site sample, the first 1200 samples are used to train the prediction model, 1201-1500 are used as the verification data, and the 1501-2000 samples are used as the test set.

(4) The GMDH neural network and the ANFIS adaptive fuzzy neural network are used to process the wind speed sequence of the corresponding frequency. The first prediction results that need to be improved are obtained, and the two models are described and explained in detail in section 2.

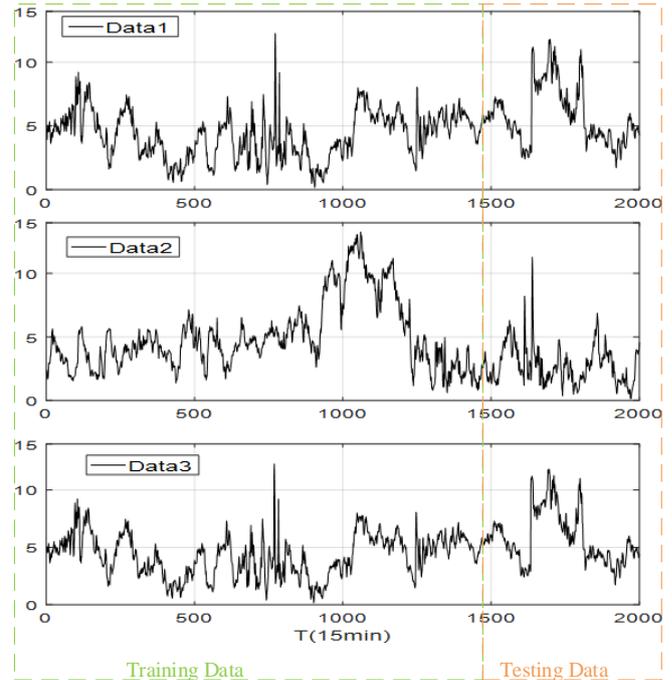


Fig. 2. Original wind speed datasets

(5) We find that the prediction result is found to be unsatisfactory after finishing step four. The specific manifestation is that it is easy to fall into the local optimum during the model prediction process, which leads to a decrease in overall prediction accuracy. To solve this problem, the Monarch Butterfly algorithm is improved in this study to optimize the prediction model, and finally the final prediction result is completed. The algorithm improvement method is given in the second section of the third part.

(6) Finally, in order to verify the prediction ability and generalization ability of the proposed model, three different classical prediction models and four different optimization algorithms are used for comparison with a total of seven different prediction models. The performance evaluation indicators to verify the proposed model are given in the second part of the third section.

The structure of the proposed model and execution process of the hybrid model proposed in this article are given in Fig. 3 and Fig. 4.

### B. Principles of Data Preprocessing

The basic idea of the data preprocessing method adopted in this paper is based on decomposing the sequence  $a(t)$  into a collection of multiple single-component IMFs plus a non-zero

average residual component  $r_Q(t)$  [24]. The processes of the second decomposition algorithm are shown in Fig. 5.

$$a(t) = \sum_{i=1}^Q \overline{IMF} + r_Q(t) \quad (1)$$

where  $Q$  is the number of *IMFs*,  $r_Q(t)$  is the final residual signal, and the first-generation residual signal  $r_1 = \langle m(x(i)) \rangle$ .

#### a: The first decomposition method

The whole process of the algorithm is divided into three steps: estimating the local cut-off frequency, estimating the local mean, and checking whether the residual margin meets the stopping criterion [21].

(1) Estimated cut off frequency

$$f'_{bis}(t) = \frac{f'_1(t) + f'_2(t)}{2} = \frac{\eta_2(t) - \eta_1(t)}{4I_1(t)I_2(t)} \quad (2)$$

(2) Estimated local mean

Different from the convention methods of calculating the local meaning, this paper uses the time-varying filter technique to estimate the local meaning. Furthermore, it uses the non-uniform *b-spline* approximation as the time-varying filter[25]. The advantage is that it can improve the stability and robustness of traditional modal decomposition and reprocessing data under the condition of satisfying the low sampling rate. Let  $\beta^n_m$  be the  $n$ -order *b-spline* function,  $m$  is the step length of the sequence, and define the time sequence in the *b-spline* space [26, 27].

$$h_m^n(t) = \sum_{k=-\infty}^{\infty} c(k) \beta^n(t/m - k) \quad (3)$$

where  $c(k)$  is the *b-spline* coefficient, and  $t$  represents the order and knot of a given *b-spline*.

$$b_m^n(t) := \beta^n(t/m) \quad (4)$$

The asterisk indicates the convolution operator. For a given time series  $a(t)$ , the *b-spline* coefficients are uniquely determined by minimizing the approximation error  $\rho_m^2$ .

$$\rho_m^2 = \sum_{t=-\infty}^{+\infty} (a(t) - [c]_{\uparrow m} * b_m^n(t))^2 \quad (5)$$

where  $[.]_{\uparrow m}$  is the up-sampling operation (adding zeros between each sample) by  $m$ .

(3) Check whether the residual margin meets the stopping criterion.

Calculation stop criterion  $\theta(t)$ :

$$\theta(t) = \frac{B_{Loughlin}(t)}{f_{avg}(t)} \quad (6)$$

where  $B_{Loughlin}(t)$  is the instantaneous bandwidth, and  $f_{avg}(t)$  is the average cutoff frequency. When  $\theta(t)$  becomes smaller instantaneously, the bandwidth also becomes smaller. Given a threshold  $\zeta$ , if  $\theta(t) \leq \zeta$ , stop the decomposition.

#### b: The second decomposition method

After the modal decomposition of the time-varying filter, the sub-sequences retained after removing high-frequency noise have good smoothness and good linear characteristics. However, due to the wind speed sequence itself, there are still low-frequency signals and other unstable factors. In order to meet the data requirements of the predictive model, it is required to accurately capture the linear characteristics of wind speed [3]. The proposed permutation entropy grouping

method based on adaptive noise modal decomposition performs a deep secondary decomposition of the primary data, and then separates the basic sequence and the remaining low-frequency signals. At the same time, the introduction of permutation entropy can help us detect the complexity of the decomposed sequence as an indicator of whether the demand is met. The adjacent entropy values are used to reconstitute *IMFs* into new sub-sequences. The secondary decomposition wind speed sequence can be expressed as [28]:

(1) Calculate the first-order modal component *IMF1*:

$$IMF1 = a - r_1 \quad (7)$$

(2) Calculate the  $n$ -order residual  $r_Q$ :

$$r_Q = \left\langle m \left( r_{Q-1} + \beta_{q-1} b_q(w^i) \right) \right\rangle \quad (8)$$

where  $b_q$  is the  $k$ -th mode obtained from EMD.

(3) Calculate the  $n$ -th order modal component *IMFn*:

$$IMFQ = r_Q - r_{Q-1} \quad (9)$$

The phase space reconstruction of the second decomposition signal, for the time series  $a(t)$ ,  $t=1,2,\dots, n$ , embedding dimension is defined as  $u$ , the extension time is  $t$ , and the reconstruction component  $k=n-(\psi-1)*t$ .

(4) Get the reconstruction space as [22, 23]:

$$\begin{pmatrix} \alpha(1) & \alpha(1+\tau) & \cdots & \alpha(1+(\mu-1)\tau) \\ \vdots & \vdots & & \vdots \\ \alpha(\varphi) & \alpha(\varphi+\tau) & \cdots & \alpha(\varphi+(\mu-1)\tau) \\ \vdots & \vdots & & \vdots \\ \alpha(K) & \alpha(K+\tau) & \cdots & \alpha(K+(\mu-1)\tau) \end{pmatrix} \varphi=1,2,\dots,K \quad (10)$$

(5) Extract symbol sequence

Arrange the  $j$ -th reconstructed component in the reconstruction matrix in ascending numerical order, and the resulting sorted index value will form a set of completely arranged symbol sequence  $S$  numbered  $u$ .

(6) Calculate position information from reconstructed components

Count the  $k$  reconstruction components corresponding to each arrangement in  $u!$ . The number of occurrences in the full arrangement of the number of times corresponding to each column is divided by the total number of times in the column to get the probability of each corresponding position. Therefore, there are  $m!$  column corresponding to the full array, and the probability of each column corresponding to the position is  $P_1, P_2, \dots, P_k$ .

(7) Calculate permutation entropy

$$H_p(m) = - \sum_{j=1}^k P_j \ln P_j \quad (11)$$

(8) Normalization processing

$$0 \leq H_p = H_p / \ln(m!) \leq 1 \quad (12)$$

when  $P_j=1/m!$ ,  $H_p(m)$  reaches the maximum value  $\ln(m!)$ . The eigenvalues  $\lambda_i$  ( $\lambda_i \geq 0, i=1, 2, \dots, K$ ), Compute the left singular vector ( $U_i$ ) and the right singular vector ( $R_i$ ) of  $X$ , and the matrix  $X$  and  $x_i$  are described as follows:

$$x_i = \sqrt{\lambda_i} U_i R_i^T \quad (13)$$

where,  $d = \max(i, \lambda_i > 0) = \text{Rank}(XX^T)$ ,  $U_i$  and  $R_i$  represent the left and right singular vectors of  $XX^T$  respectively.

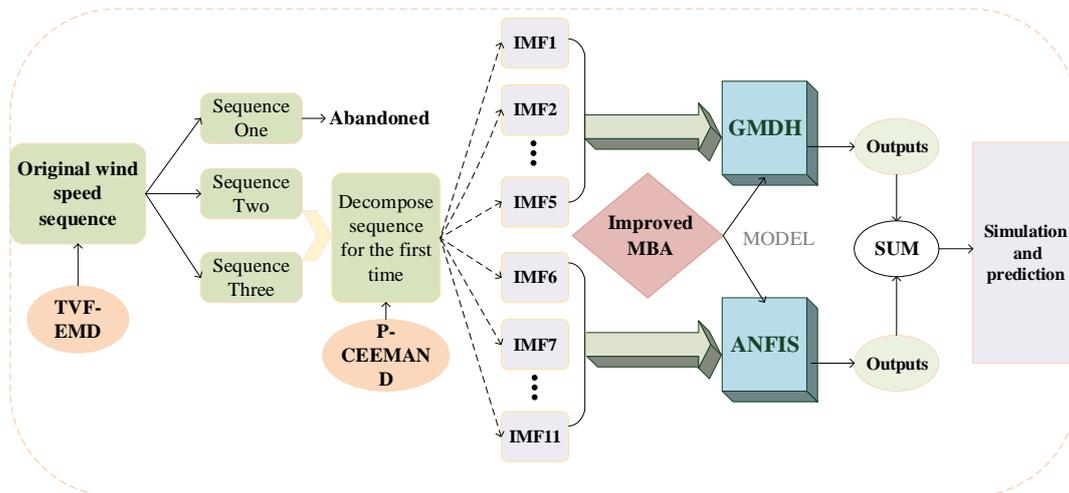


Fig. 3. The structure of the proposed model

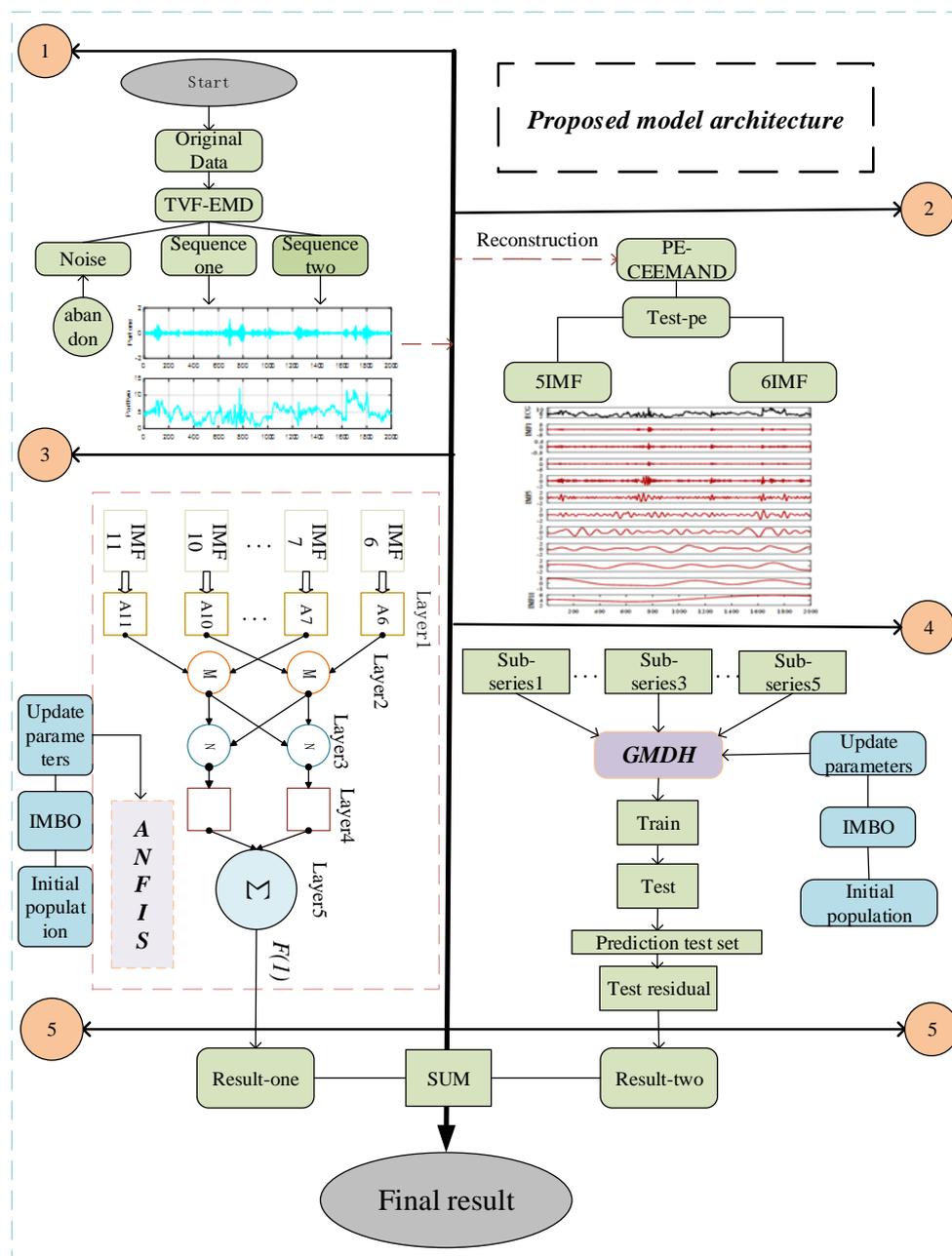


Fig. 4. Model implementation path

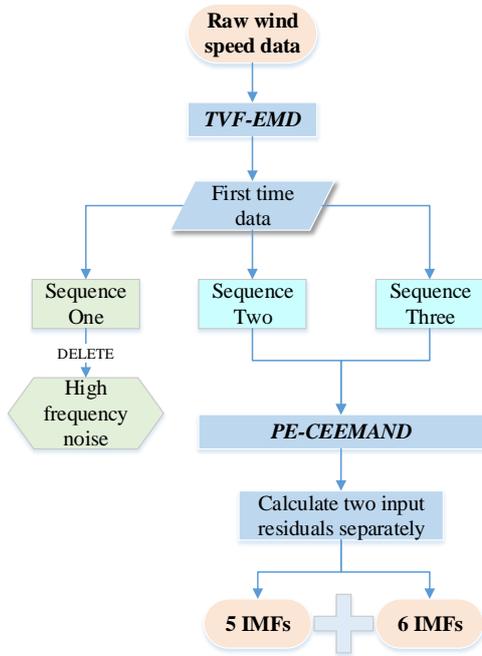


Fig. 5. Data preprocessing process

### C. Low-frequency model

The system is composed of a radical basis function neural network combined with a fuzzy system [29]. Combined with the data characteristics after the second decomposition of this article, Fig. 9 shows a basic ANFIS model structure. In general, basic ANFIS architecture has five layers [30].

The first layer: as a fuzzy layer, the function is to pass the input signal obtained by each node to the next layer, and the output of each unit can be expressed as:

$$O_1^i = \mu(x) = \exp \left\{ - \left[ \left( \frac{x - c_i}{a_i} \right)^2 \right]^{b_i} \right\} \quad (14)$$

where  $x$  is the input signal, and  $a_i$ ,  $b_i$ , and  $c_i$  are the premise parameters of the model.

The second layer: as a rule layer, the degree of membership output by each node represents the trigger strength of the fuzzy rule.

$$O_2^i = v_i = \mu(x) \cdot \mu(y), i = 1, 2 \quad (15)$$

The third layer: the normalization layer, each node of this layer is a fixed node, marked as  $n$ . The ratio of the trigger intensity of the  $i$ th node to the sum of the trigger intensity of all rules is:

$$O_3^i = \bar{v}_i = \frac{v_i}{v_1 + v_2}, i = 1, 2 \quad (16)$$

The fourth layer: the de-blurring layer, to get the output value for each rule.

$$O_4^i = \bar{v}_i f_i = \bar{v}_i (p_i x + q_i y + r_i), i = 1, 2 \quad (17)$$

where  $v_i$  is the normalized combustion intensity of the third layer, and  $p_i$ ,  $q_i$ ,  $r_i$  are independent variable parameters.

The fifth layer: the summation layer, which realizes the output of the ANFIS model by summing the output value of each rule obtained in the previous layer.

$$O_5^i = \sum \bar{v}_i f_i, i = 1, 2 \quad (18)$$

### D. High-frequency nonlinear model

GMDH neural network is a way of using high-order binomial iteration to obtain the nonlinear relationship between input and output data. For a certain actual time series, the goal of the model is to find a multiple-input single-output mathematical relationship that can replace the actual function. This mathematical relationship can be used to indirectly predict the actual value. A better way to establish the connection between input and output is to use the Volterra functional series. The basic expression is as follows [31]:

$$y = a_0 + \sum_{i=1}^n a_i x_i + \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \dots \quad (19)$$

For each pair of  $(x_i, x_j)$  as input variables, the regression method is used to calculate the coefficient  $a_i$ . After calculating the results of parameters, the prediction accuracy should be as small as possible, so in this model let the coefficient take the minimum value.

The mathematical relationship is expressed as:

$$a_i = \frac{\sum_{i=1}^M (y_i - G_i)^2}{M} \rightarrow \min \quad (20)$$

where  $M$  is the total number of samples.

Using the second-order expression of Eq. (21), the matrix equation can be determined as  $Aa=Y$ , where  $Y$  is the vector of observed output values.

$$A = \begin{bmatrix} 1 & x_{1p} & x_{1q} & x_{1p}x_{1q} & x_{1p}^2 & x_{1q}^2 \\ 1 & x_{2p} & x_{2q} & x_{2p}x_{2q} & x_{2p}^2 & x_{2q}^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_{Mp} & x_{Mq} & x_{Mp}x_{Mq} & x_{Mp}^2 & x_{Mq}^2 \end{bmatrix} \quad (21)$$

In order to solve the above standard equation, the least square method of multiple regression analysis can be used to achieve.

$$a = (A^T A)^{-1} A^T Y \quad (22)$$

For the total amount of data  $M$ , this equation will determine the vector of the best coefficients of the quadratic equation Eq. (22) [32]. Fig. 10 shows the basic model architecture of GMDH in combination with the research objects of this article.

The GMDH neural network has the ability to self-update the next-level neural unit to ensure prediction accuracy. The input layer neuron of GMDH neural network is only responsible for transmitting the input signal to the middle hidden layer neurons. Each node and output node of the hidden layer has only two inputs, so the GMDH neural network has only one output. Fig. 6 shows the processing unit of the GMDH neural network.

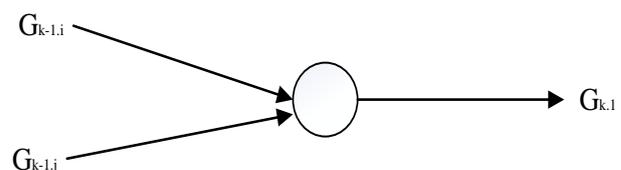


Fig. 6. Processing unit of GMDH neural network

E. Improved Monarch Butterfly optimization algorithm

The algorithm includes two processes of individual migration and adaptation to the environment. The monarch butterfly optimization algorithm defines each individual butterfly as  $X=x_1, x_2, \dots, x_D$ , and the entire group is distributed on two continents, denoted as "land 1" and "land 2" respectively. For a random individual located in "land 1", the probability is  $p$ , then the probability of being located in "land 2" is  $1-p$ , and the whole group can be divided into two groups. The behavior of the group on "land 1" is migration, and the behavior of the group on "land 2" is adaptation to the environment. The details of the algorithm implementation are as follows:

(1) Migration

The calculation formula for the migration behavior of the individual located in "land1" is expressed as:

$$x_{new}^d = \begin{cases} x_{r1}^d, rand * pre \leq p \\ x_{r2}^d, rand * pre > p \end{cases} \quad (23)$$

where  $x_{new}^d$  is the  $d$ th dimension of the new individual,  $r1$  is a random individual in land1,  $r2$  is a random individual in land 2,  $rand$  is a uniform random number between 0-1, and  $pre$  is a constant. The new individual belongs to the land 1 group. If the individual is better than the corresponding parent individual, then replace the parent position, otherwise the individual will be discarded.

(2) Adapt to the environment

The formula for calculating the adaptive behavior of an individual located in "land 2" is expressed as:

$$x_{i,new}^d = \begin{cases} x_{best}^d, rand1 \leq p \\ x_{r3}^d, rand1 > p \ \& \ rand2 \leq BAR \\ x_{r3}^d + \frac{S_{max}}{t^2} \cdot Levy(x_i^d), rand1 > p \ \& \ rand2 > BAR \end{cases} \quad (24)$$

where  $x_{i,new}^d$  best is the  $d$ th dimension of the optimal individual,  $r3$  is the random individual in land 2,  $S_{max}$  is the maximum step size of the monarch butterfly, generally taken as 1,  $t$  is the current iteration number,  $levy$  is the Levy flying random, and  $BAR$  is a constant.

The main disadvantage of MBO is that it has poor standard deviations and poor average fitness on some benchmarks. In order to improve the overall performance of the algorithm, this paper introduces a greedy strategy in the basic MBO method. The improved algorithm is called IMBO.

In the traditional MBO algorithm, every new individual butterfly has the same probability of being passed on to the next generation [33, 34]. The IMBO algorithm after introducing the greedy strategy only accepts butterfly individuals with better fitness. Specifically, the use of greedy strategies can continuously update and optimize the generation of the next generation of individuals. The details of this greedy strategy are as follows:

$$x_{i,new}^{d+1} = \begin{cases} x_i^{d+1}, f(x_i^{d+1}) < f(x_i^d) \\ x_i^d, else \end{cases} \quad (25)$$

where  $x_{i,new}^{d+1}$  is the new butterfly individual produced by the next generation,  $f(x_i^{d+1})$  and  $f(x_i^d)$  are the fitness of the corresponding individuals  $x_i^{d+1}$  and  $x_i^d$ , respectively. Based

on the above analysis, the introduction process of the entire greedy strategy can be expressed as follow.

---

**Title** The structure of optimize Monarch Butterfly optimization algorithm.

---

**Input:**

$X$ : individual butterfly

$P$ : the random probability

$x_{new}^d$ : the dimension of the new individual

$R$ : a random individual

$x_{i,new}^d$ : the dimension of the optimal individual

**Output:**  $x_{i,new}^{d+1}$

**Step 1: Initialization.**

Initialize the population  $P$  of  $NP$  butterflies; set the maximum generation

$Max(g)$ , butterfly number  $NP1$  in Land 1 and butterfly number  $NP2$  in Land 2.

**Step 2: Fitness evaluation.**

**Step 3: While**  $t < Max(g)$  **do**

To classify the offspring population.

Divide all population into two sub-populations.

for  $i=1$  to  $NP1$  do

for  $k=1$  to  $D$  do

Generate by Eq. (30) and Eq. (24).

end for  $k$

Generate by greedy strategy as Eq. (25).

---

F. IMBO-Hybrid model

Monarch Butterfly Optimization Algorithm (MBO) was proposed in 2016, which was mainly inspired by Monarch Butterfly migration and reproduction behavior. In order to obtain the optimal individuals in the population, the main strategy is to update the location and migration distance of breeding individuals. However, this algorithm has some limitations in global search. Considering this problem, corresponding improvement measures are proposed.

At the same time, ANFIS and GMDH neural networks are very sensitive to the input layer and hidden layer, the initial connection weight between hidden layer and output layer neurons, and the initial threshold between the hidden layer and the output layer. The initial weight of neural network and the threshold of network layer are two important parameters which determine the accuracy and precision of the model. However, these two parameters of conventional ANFIS and GMDH neural network are artificially assigned according to engineering experience, which is obviously unreasonable. In order to solve this problem, this paper uses the improved monarch optimization algorithm to determine these two important parameters. The structure of the IMBO-Hybrid model is shown in Fig. 7.

Based on the above improved strategies, parameters used in the mixed model were determined after repeated independent tests, as shown in TABLE I.

III. DATA PREPROCESSING RESULTS

Due to the repeatability of data decomposition, so as to ease the unnecessary workload of this article, this article does not show the preprocessing results of the data sets of the other two sites. In the hybrid model advocated in this work, TVF-EMD and P-CEEMAND are used to process the original wind speed, including de-noise, decomposition, and localization. Fig. 11 shows the primary data of the site after

being decomposed by TVF-EMD. On the basis of a decomposition, 11 intrinsic mode functions (*IMFs*) are obtained after P-CEEMAND decomposition. The secondary data is shown in Fig. 8.

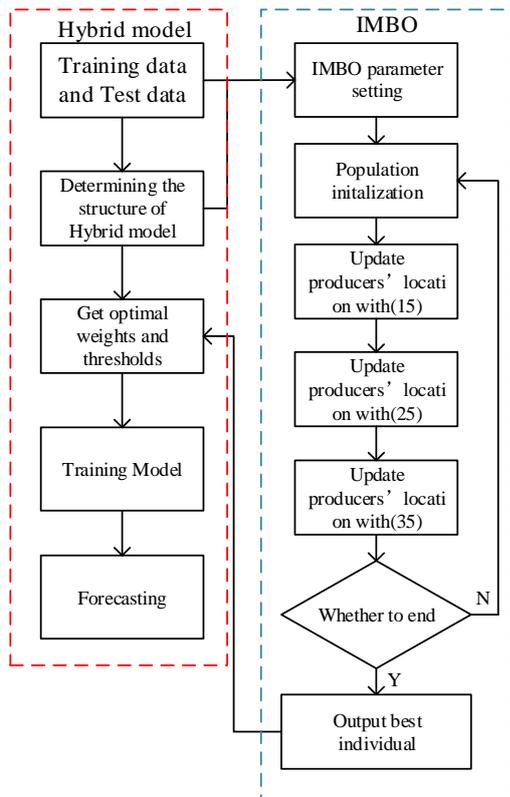


Fig. 7. The structure of IMBO-Hybrid model

TABLE I  
PARAMETER INITIALIZATION

Parameter name	Value
Number of iterations	300
Population size	50
Evolution times	20
Proportion of producers	0.2
Security threshold	0.8
Number of input layer nodes	5
Number of hidden layer nodes	20
Number of output layer nodes	1

In order to further verify the data characteristics of the captured *IMF* components, this paper uses component linear characteristics and frequency domain amplitude analysis on the basis of the data in Fig. 8. Only by satisfying the data characteristic requirements of these two aspects can it be enough to demonstrate the effectiveness of the secondary decomposition strategy proposed in this paper.

It can be seen from the Fig. 8 that the abnormal value of the component data is far smaller than the standard value, accounting for less than five thousandths of the total data. In addition, from the margin assignment data in the figure, it can be seen that the first five component data have good linear characteristics. This also provides conditions and basis for us

to use the ANFIS model to predict sub-frequency domain components.

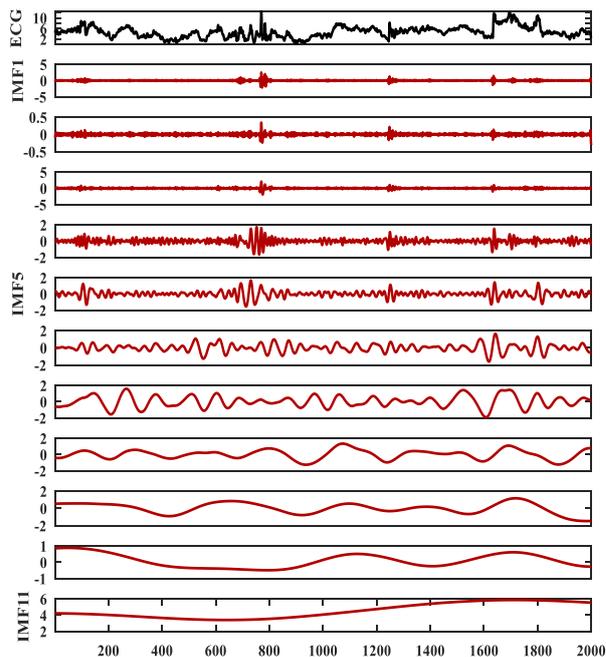


Fig. 8. Secondary data decomposition

#### IV. RESULT ANALYSIS

##### A. Performance evaluation index

In the basic experiment of this paper, four classic error indicators, MAE, MSE, RMSE and MAPE, were used to measure and evaluate the prediction accuracy of wind speed by the hybrid model. Their significance and calculation formulas are shown in TABLE II.

TABLE II  
THE PERFORMANCE EVALUATION INDEX

Index	Significance	Formula
MAE	Mean Absolute Error	$MAE = \frac{1}{n} \sum_{i=1}^n  y_i - \hat{y}_i $
MSE	Mean Square Error	$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$
RMSE	Root Mean Squared Error	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$
MAPE	Mean Absolute Percentage Error	$MAPE = \frac{1}{n} \sum_{i=1}^n \left  \frac{\hat{y}_i - y_i}{y_i} \right  \times 100\%$

where  $y_i$  is the true wind speed;  $\hat{y}_i$  is the predicted wind speed, and  $n$  is the number of wind speeds forecasted.

Firstly, MSE can measure the difference between the actual value and the predicted value of wind speed data. Intuitively speaking, the smaller the MSE value is, the higher the prediction accuracy of the model will be, and the smaller the error will be. Then MAE and RMSE, a pair of indicators, can be verified with each other to ensure that the sign problem of prediction error will bring bad influence to the result. Finally, MAPE shows that the error between the predicted value and the real value can reflect the stability of the model to a certain extent

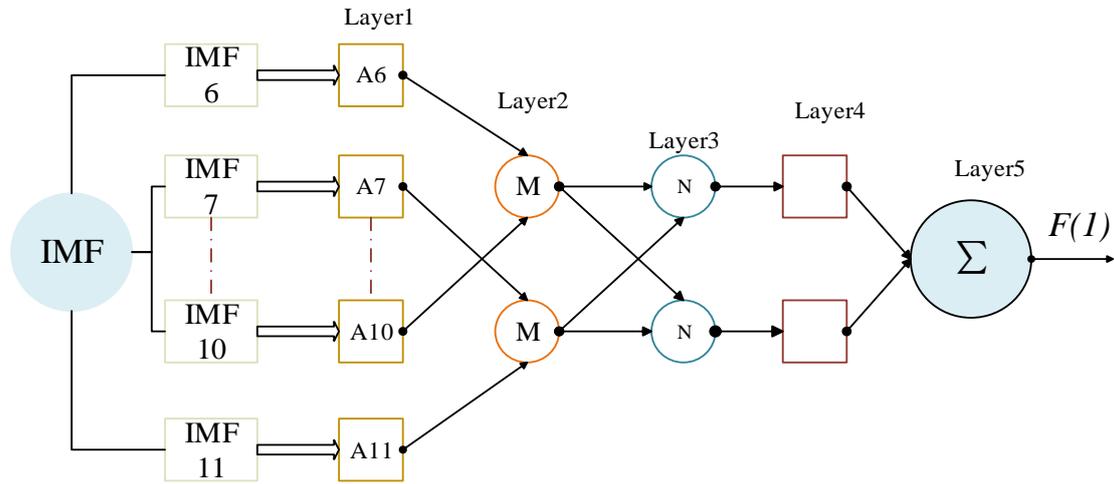


Fig. 9. ANFIS neural network model structure

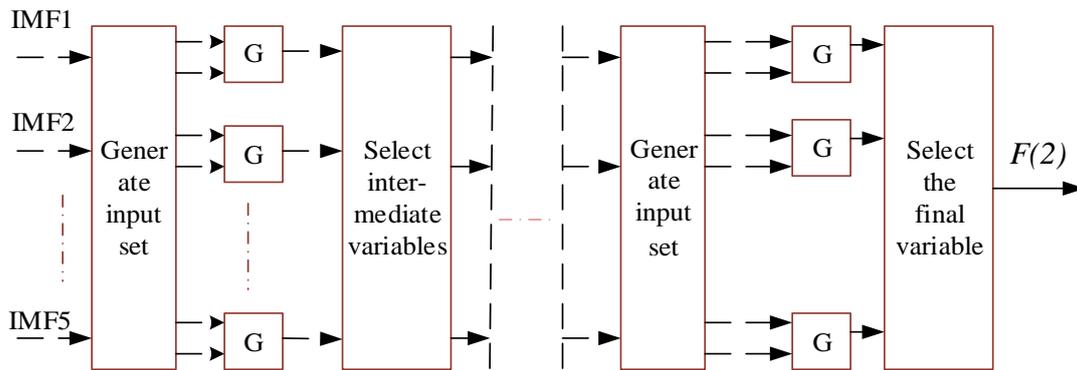


Fig. 10. GMDH neural network structure

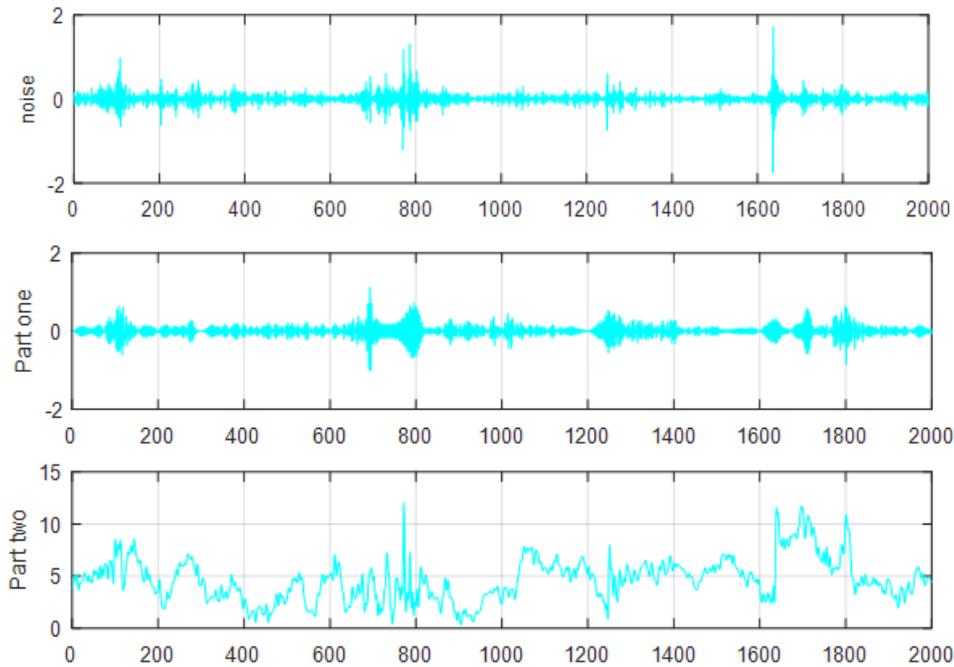


Fig. 11. First data decomposition sequence diagram

The purpose of this paper is to verify the performance of the proposed model in a more comprehensive way. The following is the specific process in the experiment. First, 11 stationary wind velocity sub-sequences are obtained based on the secondary decomposition (TPD decomposition) technique. Then, on the basis of the proposed model, the optimized monarch butterfly algorithm is used to optimize the

model so as not to fall into the local optimum during the prediction process. Finally, on this basis, the ANFIS model is used to process the first five sequences with low-frequency timing while using the GMDH model to process. In addition, by combining the frequency domains of different frequency domains, the subsequences of each group are obtained. The results are shown in TABLE III.

TABLE III  
NUMBER OF *IMF* AFTER FREQUENCY DOMAIN

Data	<i>IMFS</i>	sub-series
Data1	11	5
Data2	11	5
Data3	11	6

Finally, the final prediction result is achieved by summing the prediction results of each component. The prediction result is the performance index value of each sub-series and the index corresponding to the original wind speed sequence of each station. The sites and subsequences are shown in TABLE III. At the same time, Fig. 12, Fig. 13 and Fig. 14 are the comparison between the actual wind speed curves. And the predicted wind speed curve after each sub-sequence is predicted by the model in step experiment is shown. It can be concluded from TABLE IV.

(1) MAE and RMSE values of sub-sequence decomposed in frequency domain are lower than those obtained by direct prediction of original wind speed data. The results show that the proposed data preprocessing strategy achieves good results and can effectively improve the prediction accuracy of the prediction model.

(2) It is not difficult to find that the MAPE value of subsequence increases with the increase of sequence frequency. This is because the variation of MAPE values depends not only on the error between the actual value and the predicted value, but also on the degree of concentration in the sample size.

(3) For the prediction of different sites, the corresponding indicators are quite different. The enlightenment to us is that in the subsequent model comparison experiment, the error between the predicted value and the real value of each site should be evaluated separately.

### B. Prediction results analysis

Based on existing real wind speed data and predicted wind speed data, the performance of the proposed hybrid model is verified. In this paper, we design a comparison experiment between transverse and other classical prediction models and longitudinal intelligent optimization algorithm.

Experiment 1 compares the proposed model with several classical time series models. In this paper, two classic neural network prediction models (ANFIS and GMDH) are used as comparison models to test the prediction ability of the mixed model. The main purpose of choosing ANFIS and GMDH as comparison models is to comprehensively consider linear and non-linear models and neural network models to test the performance of the models.

Experiment 2 compares the prediction effects of using different optimization algorithms to optimize parameters based on the proposed hybrid model. The classic monarch butterfly optimization algorithm (MBO), particle swarm optimization algorithm (PSO), and genetic algorithm (GA) are selected to compare with the updated monarch butterfly optimization algorithm (IMBO) based on the greedy strategy in this paper. The overall structure of the input and output of each group of models remains unaffected, and only the corresponding changes are made in the optimization algorithm to find the best parameters for model optimization.

### a: Comparative experiment of different classic prediction models

The experimental results show that the proposed model has a strong predictive ability for short-term wind speed prediction. After setting the prediction parameters of the model, experiments are carried out on three stations, and the experimental results gained are shown in TABLE V. At the same time, Fig. 15 shows prediction curves.

For site 1, the four model evaluation indicators of the established model are better than the other two models, and the most satisfactory prediction accuracy is obtained.

For site 2, according to the values of MAE, MAPE, RMSE and MAPE, a single model has the ability to predict by analogy. In the prediction laboratory, the MAPE values of ANFIS and GMDH are 15.990% and 12.206%, respectively. In contrast, the MAPE value of the proposed model is 9.871%, which is 6.199% and 2.335% percent higher than the above model respectively.

For site 3, according to the four evaluation criteria adopted, the hybrid model established in this paper is still superior to other models. Among them, MSE, MAE, MAPE and RMSE are 0.186%, 0.234%, 4.426% and 0.432%, respectively. In the remaining model, the methods of predicting accuracy from good to bad are ANFIS and GMDH, and their MAPE values are 15.550% and 15.593%, respectively.

The prediction results of the built model are significantly different from those of other independent models. Specifically, regardless of the prediction step, the evaluation index value obtained by the established model is significantly lower than the evaluation index value obtained by the comparison model. Therefore, we can conclude from the experiment that the hybrid prediction model proposed in this paper is superior to the traditional single model in terms of short-term wind speed prediction.

### b: Model experiment of different optimization algorithms

The purpose of this experiment is to compare the IMBO-based hybrid model proposed in this paper with other hybrid models based on different algorithms. Specifically, this paper adopts particle swarm optimization, genetic algorithm, monarch butterfly optimization algorithm as the comparison algorithm. So as to reasonably verify the performance of IMBO, this paper selects two classical optimization algorithms and the unimproved Monarch butterfly optimization algorithm. Experiments are carried out on the premise that the common parameters are maintained and other relevant parameters are reasonable. TABLE XIV shows the calculated evaluation index values of these three sites. It can be seen from the table that the prediction performance of the proposed hybrid model is better than the combined model optimized by other comparison optimization algorithms.

For example, the MAPE value of the hybrid model optimized by the improved monarch butterfly optimization algorithm is 1.5%~4.5%, while the MAPE value of the hybrid model improved based on the legacy algorithm is 3.6%~5.9%. Fig. 17 shows the prediction curves of various multi-objective optimization algorithms, and it is not difficult to draw the conclusions given above.

TABLE IV  
THE ERROR RESULT OF THE PREDICTED RESULT

	Dataset1			Dataset2			Dataset3		
	MAE	RMSE	MAPE	MAE	RMSE	MAPE	MAE	RMSE	MAPE
sub-series1	0.103	0.157	3.215	0.101	0.222	4.215	0.124	0.105	2.154
sub-series2	0.053	0.089	2.984	0.054	0.084	1.145	0.042	0.084	0.983
sub-series3	0.045	0.065	1.364	0.045	0.061	0.878	0.013	0.053	0.624
sub-series4	0.031	0.051	0.542	0.007	0.021	0.688	0.021	0.051	0.314
sub-series5	0.013	0.014	0.064	None	None	None	0.016	0.042	0.226
sub-series6	None	None	None	None	None	None	0.008	0.034	0.185
<b>Our-model</b>	<b>0.245</b>	<b>0.376</b>	<b>8.151</b>	<b>0.270</b>	<b>0.411</b>	<b>6.935</b>	<b>0.224</b>	<b>0.369</b>	<b>4.486</b>

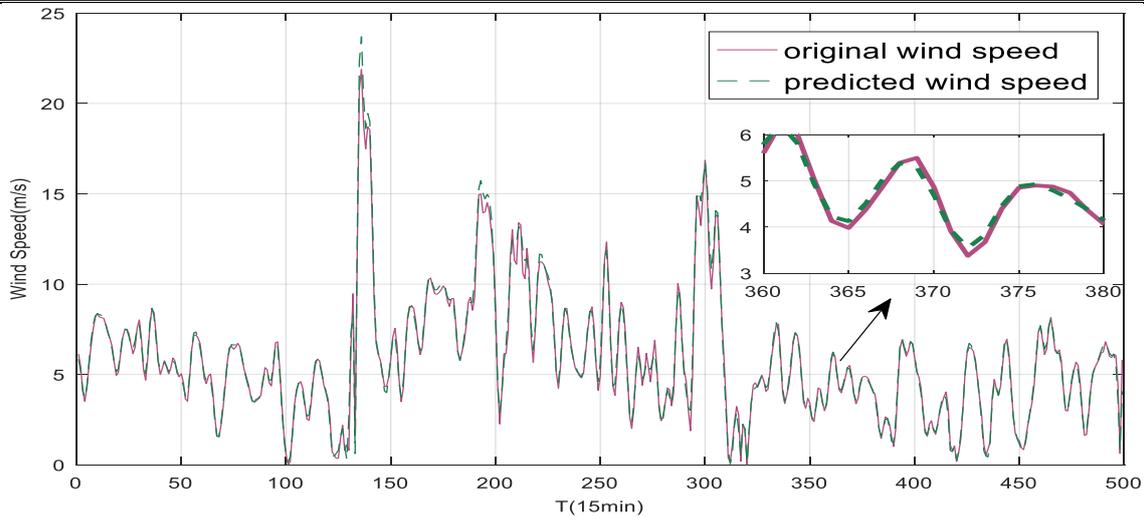


Fig. 12. The prediction result curve of the verification data set 1

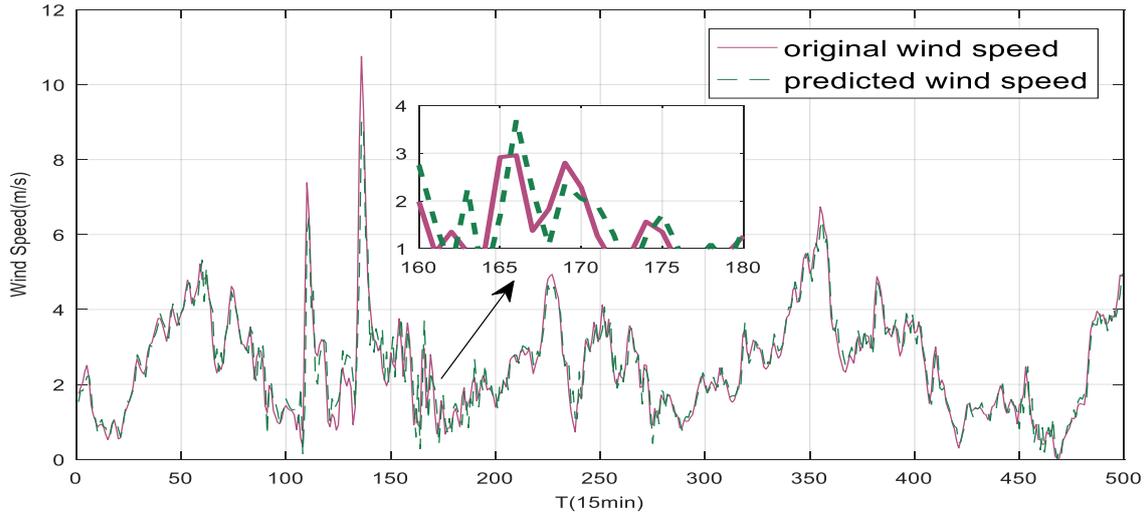


Fig. 13. The prediction result curve of the verification data set 2

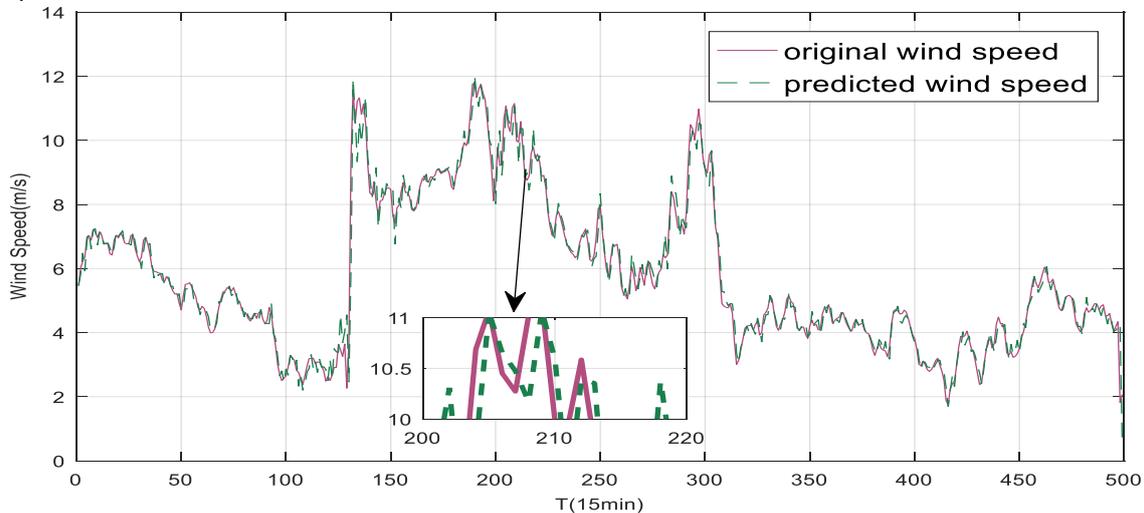


Fig. 14. The prediction result curve of the verification data set 3

TABLE V  
THE PREDICTION ERROR INDEX VALUE OF EXPERIMENT 1

		MSE(m/s)	MAE(m/s)	MAPE(m/s)	RMSE(m/s)
Dataset1	Proposed-model	<b>0.304</b>	<b>0.217</b>	<b>8.33</b>	<b>0.551</b>
	ANFIS	1.314	1.325	10.21	0.556
	GMDH	1.351	1.451	9.77	0.631
Dataset2	Proposed-model	<b>0.326</b>	<b>0.372</b>	<b>9.871</b>	<b>0.571</b>
	ANFIS	1.456	1.414	15.990	1.859
	GMDH	1.256	1.069	12.206	1.804
Dataset3	Proposed-model	<b>0.186</b>	<b>0.234</b>	<b>4.426</b>	<b>0.432</b>
	ANFIS	0.907	0.805	15.550	1.440
	GMDH	0.720	0.798	15.593	1.456

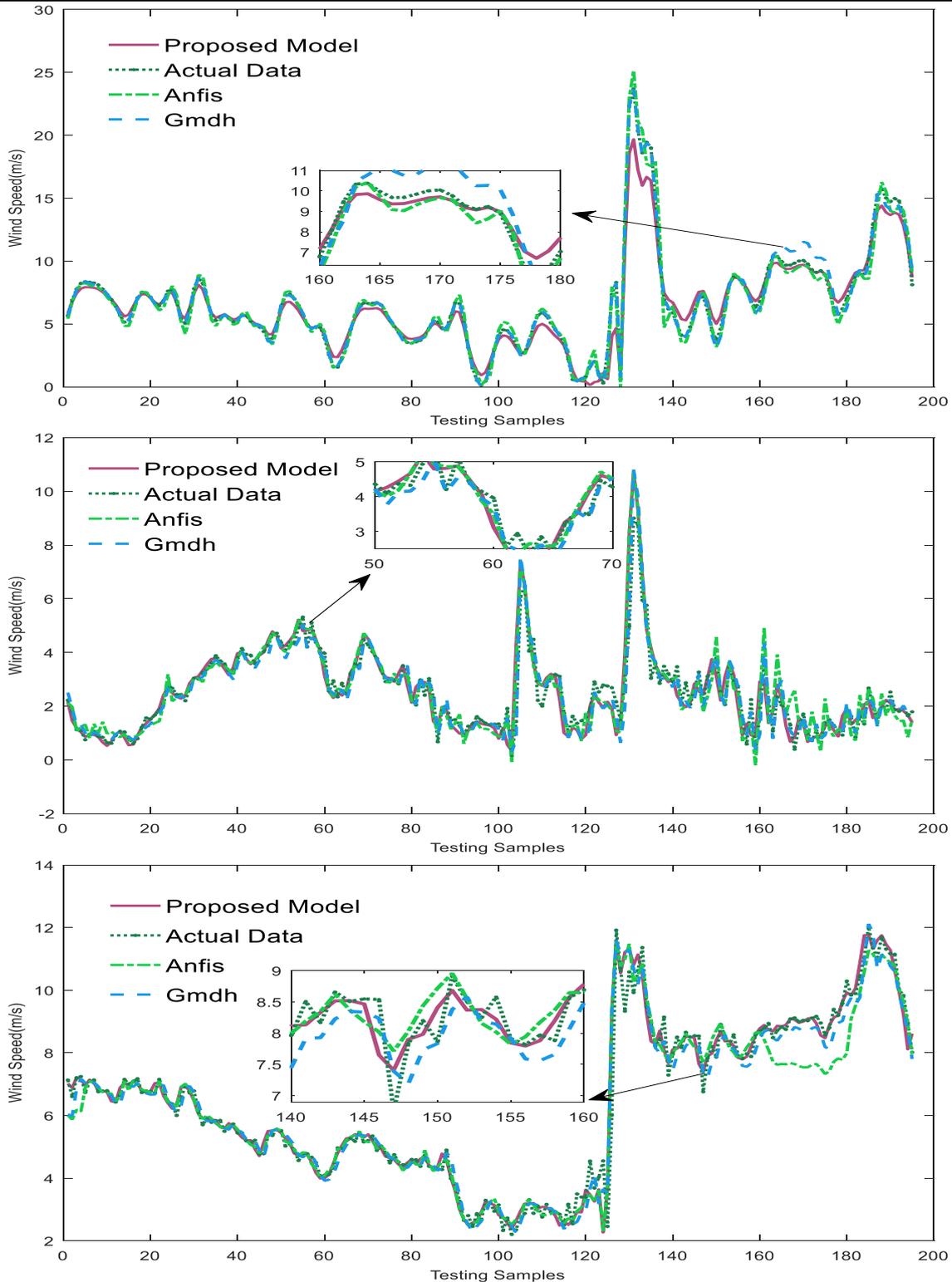


Fig. 15. The result curve of experiment 1

**C. Computational complexity comparison**

Based on the previous experimental results, the superiority of the new hybrid model and other models and algorithms can be well verified, but some performance of the model itself cannot be well verified.

To further illustrate the superiority of IMBO in running time, in this experiment, the running time and convergence of 15 independent experiments are used to measure the performance of IMBO in time and space. Other parameters of the selected algorithm are shown in TABLE VI.

TABLE VI  
SELECTED ALGORITHM PARAMETERS

Algorithms	The definition parameters	Value
IMBO,MBO	Max iteration	200
	Particles	50
	Position	[-2,2]
	Max velocity	2
GA	Min velocity	-2
	Particles	50
	Position	[-5,5]
	Max velocity	-5
	Min velocity	5

The running time of 15 groups of independent experiments and the maximum, minimum, and average values generated by convergence are summarized in TABLE VII.

TABLE VII  
THE OPERATION OF THE FOUR ALGORITHMS

Indexes		IMBO	MBO	GA	PSO
Running times	Max	<b>78.66</b>	79.54	80.14	85.14
	Min	<b>71.21</b>	72.14	73.45	76.14
	Average	<b>75.93</b>	75.84	76.79	80.64
Convergent generation	Max	<b>58</b>	69	65	84
	Min	<b>42</b>	53	64	65
	Average	<b>48</b>	61	69	74

The operation of the four algorithms is summarized in TABLE VII, from which it can be seen that the average running time of IMBO is slightly higher than that of MBO, but smaller than that of GA and PSO. At the same time, it is not difficult to find that the average number of convergence iterations of IMBO is less than that of MBO, GA, and PSO. The average number of convergence iterations of the latter three are 61, 69, and 74, respectively, while the average number of convergence iterations of IMBO is 48 and the corresponding average running time is 75.93s. Fig. 16 shows the iterative convergence curve of the algorithm. The experimental results prove that the improvement of MBO in this paper is effective.

**D. Stability and directionality**

In order to further verify the stability and sensitivity of the mixed model, this paper uses prediction error variance (*Var*) to evaluate the models' prediction stability. The definition and calculation formula of *Var* and *DA* are expressed as:

$$Var = E(\hat{y} - y) - E((\hat{y} - y))^2 \quad (26)$$

$$DA = \frac{100}{n} \sum_{t=1}^{n-1} a_t, a_t = \begin{cases} 1, & \text{if } (\hat{y}_{t+1} - y_t)(y_{t+1} - y_t) > 0 \\ 0, & \text{otherwise} \end{cases} \quad (27)$$

where  $\hat{y}$  is the predicted value, and  $y$  is the observed value.  $y(t)$  and  $y(t+1)$  represent wind speed observations at time  $t$  and  $t+1$ ,  $\hat{y}(t+1)$  is the predicted value of wind speed at  $t+1$ . Based on these two indicators, the stability and directivity are compared and evaluated respectively.

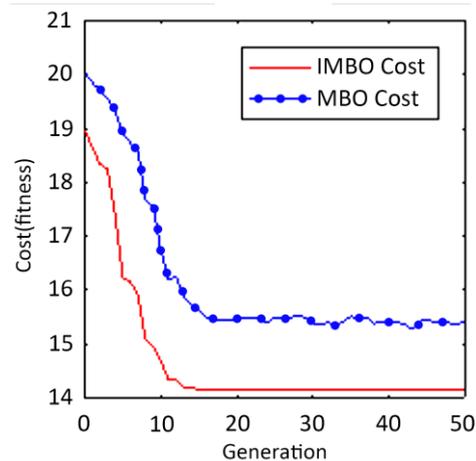
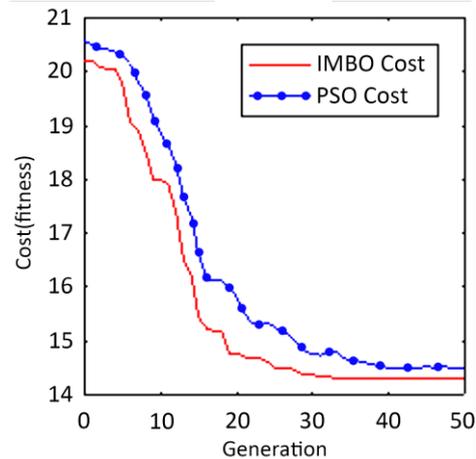
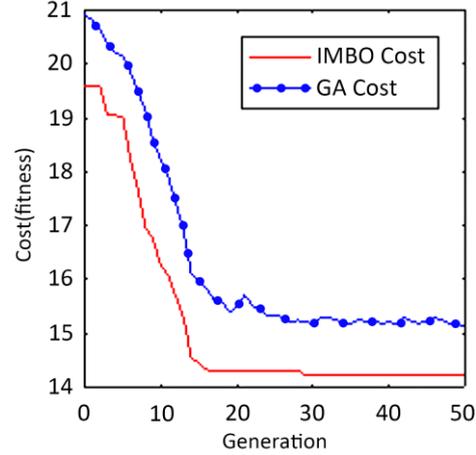


Fig. 16. Algorithm iteration convergence curve

The experimental results are shown in TABLE VIII. The experimental result data can prove that the mixed model proposed in this paper has the smallest *Var* value of 0.0117, 0.0012, 0.0025 and the largest *DA* value of 85.4557, 82.6464, 79.1461 on the three data sets, respectively. It can be proved that the hybrid model is the most stable. At the same time, the hybrid model has the best generalization ability for wind speed data of different stations

TABLE VIII  
FORECAST ERROR VARIANCE AND DIRECTION ACCURACY VALUE

Indexes	Models	Dataset		
		Dataset1	Dataset2	Dataset3
Var	The proposed model	<b>0.0117</b>	<b>0.0012</b>	<b>0.0025</b>
	TPD-ANFIS	0.1430	0.1256	0.2653
	TPD-GMDH	0.1610	0.1132	0.1295
	MBO-ANFIS	0.2794	0.1398	0.1354
	GA-Combination model	0.2662	0.0663	0.0366
	PSO-Combination model	0.1256	0.0465	0.0368
	MBO-Combination model	0.2344	0.0821	0.0363
DA(%)	The proposed model	<b>85.4557</b>	<b>82.6464</b>	<b>79.1461</b>
	TPD-ANFIS	71.2136	61.3216	67.2665
	TPD-GMDH	80.8966	45.2133	71.3293
	MBO- ANFIS	79.2324	67.3431	73.3218
	GA-Combination model	82.3154	82.8953	80.3129
	PSO-Combination model	83.5463	82.8613	82.3743
	MBO-Combination model	84.2498	84.9243	81.1665

**E. Diebold-Mariano (DM) test**

The error between the predicted value and the actual value is defined as a variable, so the confidence level of this variable under a certain confidence level can be reflected by DM test. In order to further verify the difference in prediction accuracy between the hybrid model proposed in this paper and other models on the basis of previous experiments, the experiment conducted DM index test to further verify, and its theoretical description is as follows.

When the confidence level is given as  $\alpha$ , the hypothesis test can be expressed as:

$$\begin{cases} H_0 : E\left(L\left(e_t^{(1)}\right)\right) - E\left(L\left(e_t^{(2)}\right)\right) = 0 \\ H_1 : E\left(L\left(e_t^{(1)}\right)\right) - E\left(L\left(e_t^{(2)}\right)\right) \neq 0 \end{cases} \quad (28)$$

where  $e^t$  represents the prediction error value at time  $t$ .

In order to accurately calculate the value of  $L(\bullet)$ , the best way is to use a loss function, and this paper takes the error rate as the loss function.  $H_0$  is defined as the null hypothesis, meaning that there is no difference in the performance of the two models. Conversely,  $H_1$  indicates that there are differences between the models. The *DM* test can be described as:

$$DM = \frac{s^2 \sum_t \left( L\left(e_t^{(1)}\right) - L\left(e_t^{(2)}\right) \right) / N}{\sqrt{s^2 / N}} \quad (29)$$

where  $s^2$  is the variance estimate of  $d_t = L(e_t^{(1)}) - L(e_t^{(2)})$ .

The value of confidence in this paper is 0.01. TABLE IX records the prediction error DM of the proposed model and the model performance comparison of the two models. The DM statistics are compared with the critical value  $Z/2$ . If the value of the DM statistics is greater or less than the confidence curve, the null hypothesis is rejected. From the DM value of the prediction error of the three stations in the experiment, it is not difficult to find that the confidence level of the comprehensive wind speed prediction strategy adopted in this paper is higher than that of other prediction methods.

Therefore, it is not difficult to prove that the differences and advantages between the hybrid model proposed in this paper and other combination models with different strategies are obvious. The results show that compared with the comparative wind speed prediction model, the prediction of this model is better.

TABLE IX  
FORECAST ERROR DM VALUE

Model	Site-One	Site-Two	Site-three
TPD-ANFIS	5.6610	5.3283	6.3212
TPD-GMDH	5.5612	5.3464	6.2166
Proposed Model	5.8413	5.8461	5.2184
PSO based ANFIS	3.4131	3.9845	4.3431
GA based ANFIS	4.4166	4.4655	3.8495
PSO based GMDH	4.3498	4.2164	4.3189
GA based GMDH	3.9721	4.4986	3.6894
MBO based Propose Model	4.3152	3.9213	3.9462
TPD-MBO based GMDH	4.9499	4.1328	4.8942
TPD-MBO based Propose Model	4.4213	4.6187	5.3791
TPD-IMBO based Propose Model	<b>3.9846</b>	<b>3.6548</b>	<b>3.9943</b>

The experimental results obtained based on the three sets of horizontal experiments carried out in this research can be seen. The hybrid model based on the improved monarch butterfly algorithm proposed in this paper has obvious advantages in model stability, generalization ability, robustness, and prediction accuracy and prediction efficiency.

**F. Parametric analysis**

When a parameter of the model changes, it will affect the prediction ability of the hybrid model. This paper takes the standard deviation of predicted value and error value as the corresponding index to conduct sensitivity analysis from two aspects of data pretreatment technology and optimization algorithm to discuss the sensitivity of model state and output results to parameter changes. Index definition is shown in TABLE X.

The three parameters considered in TPD data preprocessing are the ratio of the standard deviation of noise to the standard deviation of the original sequence, The Times of implementation and the maximum number of screening

iterations. In the IMBO optimization algorithm, the number of monarch butterflies, the number of iterations and the archive size are the other three parameters. The detailed comparison results are shown in TABLE XI, TABLE XII and TABLE XIII. During the first analysis, only one parameter is changed at a time, leaving the rest unchanged. The ratios of the standard deviation of the noise to the standard deviation of the sequence are 0.01, 0.05, 0.1, 0.15, and the achieved numbers

are 50, 100, 150, 200, and the maximum screening iterations are 200, 300, 400, and 500, respectively. Similarly, in the second analysis program relative to IMBO, the parameter values of the monarch butterfly are set to 20, 40, 60, and 80. The number of iterations is 50, 100, 150, which means the archive size is 200, 300,400,500.

TABLE X  
PREDICTIVE SENSITIVITY OF THE FOUR MEASUREMENT INDICATORS TO EACH RELATED MODE.

Metric	Definition	Equation
PMAE	STD value of MAE of n times	$PMAE = \text{Std}(MAE1, MAE2, \dots MAEn)$
PMape	STD value of MAPE of n time	$PMape = \text{Std}(MAPE1, MAPE2, \dots MAPEn)$
PRMSE	STD value of RMSE of n time	$PRMSE = \text{Std}(RMSE1, RMSE2, \dots RMSEn)$
PMSE	STD value of MSE of n time.	$PMSE = \text{Std}(MSE1, MSE2, \dots MSEn)$

TABLE XI  
STATION 1, THE SENSITIVITY ANALYSIS TABLE OF THE PARAMETERS INVOLVED IN THE MIXED MODEL.

Parameter	Site-One			
	PMAE	PMape	PRMSE	PMSE
Ratio	0.0033	0.0313	0.0032	0.4642
Number of realizations	0.0012	0.0137	0.0013	0.7942
Maximum number of filtering iterations	0.0046	0.0891	0.0012	0.4316
Number of Monarch Butterflies	0.0032	0.0135	0.0021	0.9546
Number of iterations	0.0092	0.0321	0.0089	0.4649
Archive size	0.0075	0.0121	0.0036	0.6434

TABLE XII  
STATION 2, THE SENSITIVITY ANALYSIS TABLE OF THE PARAMETERS INVOLVED IN THE MIXED MODEL.

Parameter	Site-Two			
	PMAE	PMape	PRMSE	PMSE
Ratio	0.0033	0.0361	0.0033	0.4492
Number of realizations	0.0014	0.0236	0.0024	0.8421
Maximum number of filtering iterations	0.0043	0.0491	0.0013	0.5461
Number of Monarch Butterflies	0.0038	0.0124	0.0025	0.9456
Number of iterations	0.0084	0.0245	0.0046	0.5162
Archive size	0.0065	0.0521	0.0064	0.3315

TABLE XIII  
STATION 3, THE SENSITIVITY ANALYSIS TABLE OF THE PARAMETERS INVOLVED IN THE MIXED MODEL.

Parameter	Site-Three			
	PMAE	PMape	PRMSE	PMSE
Ratio	0.0043	0.0565	0.0046	0.5561
Number of realizations	0.0026	0.0322	0.0036	0.6231
Maximum number of filtering iterations	0.0045	0.0434	0.0011	0.4315
Number of Monarch Butterflies	0.0034	0.0164	0.0026	0.9231
Number of iterations	0.0086	0.0313	0.0053	0.5242
Archive size	0.0014	0.5034	0.0034	0.3312

TABLE XIV  
THE PREDICTION ERROR INDEX VALUE OF EXPERIMENT 2

	Indexes	MSE(m/s)	MAE(m/s)	MAPE(m/s)	RMSE(m/s)
Dataset1	Proposed Mode	<b>0.072</b>	<b>0.552</b>	<b>1.610</b>	<b>0.269</b>
	MBO-Model	0.512	0.144	4.461	0.716
	PSO- Model	0.540	0.150	4.747	0.735
	GA- Model	0.503	0.149	5.161	0.709
Dataset1	Proposed Mode	<b>0.031</b>	<b>0.036</b>	<b>1.933</b>	<b>0.178</b>
	MBO-Model	0.135	0.083	4.844	0.366
	PSO- Model	0.131	0.082	4.895	0.131
	GA- Model	0.127	0.082	4.827	0.357
Dataset1	Proposed Mode	<b>0.186</b>	<b>0.234</b>	<b>4.553</b>	<b>0.432</b>
	MBO-Model	0.219	0.354	6.451	0.486
	PSO- Model	0.251	0.370	6.738	0.501
	GA- Model	0.235	0.352	6.366	0.485

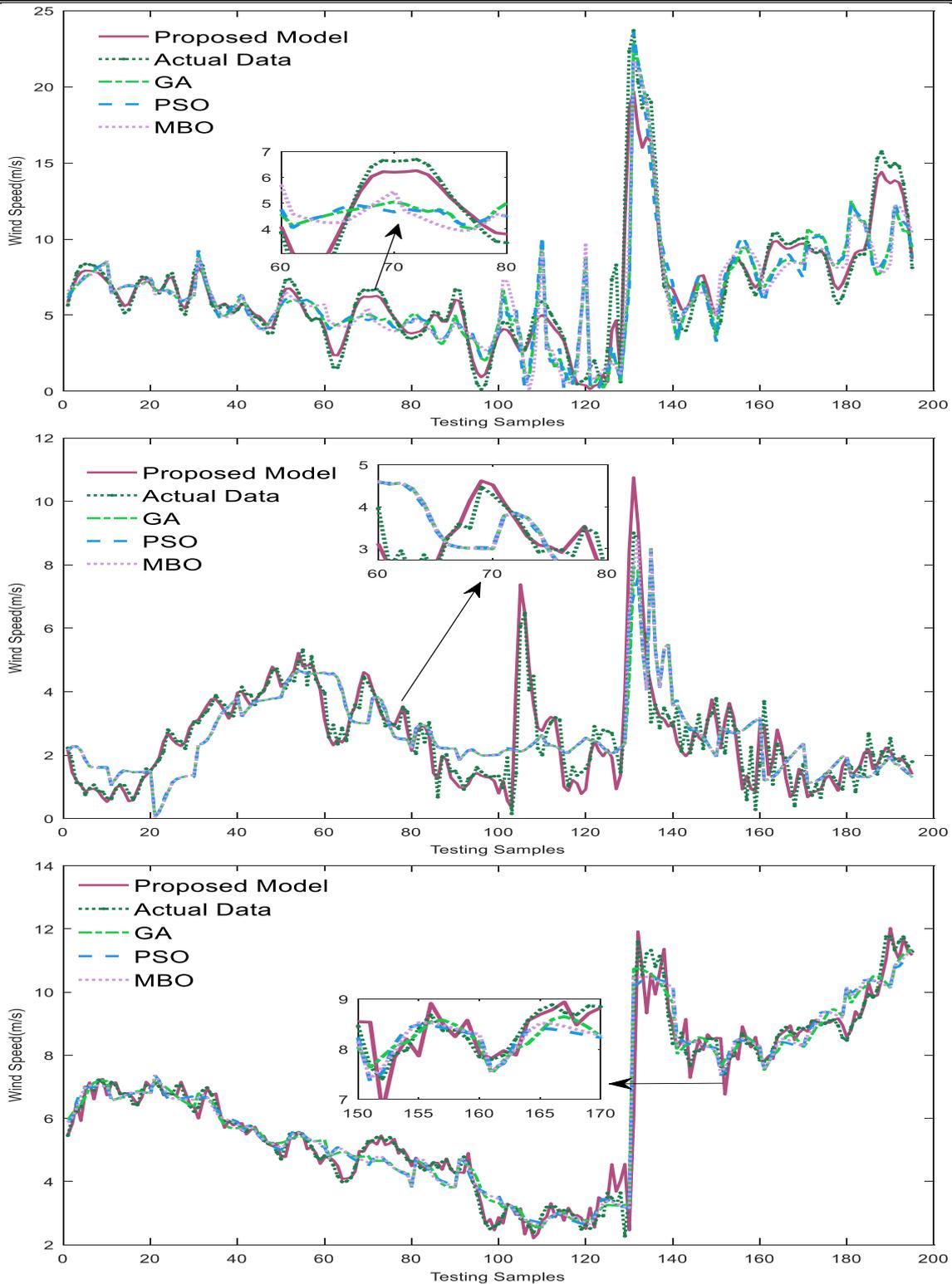


Fig. 17. The result curve of experiment 2

From the experimental results, it can be found that with the change of the parameters in the TPD decomposition model, the transformation degree of the four parameters PMAE, PMPE, PRMSE and PMSE are relatively low. For example, in the site 1 prediction, the ratio of PMAPE values is 0.0313, the number of realization is 0.0137, and the maximum screening iteration is 0.0891. These values are higher than most SSE values, but still at a low level. This indicates that the TPD decomposition model has a stronger tolerance to parameter changes. At the same time, when the parameters in the IMBO optimization algorithm change, it can be concluded that the sensitivity of the measurement index in the IMBO algorithm is not as good as that of the parameters in the TPD decomposition technology. This shows that the variation of parameters in the optimization algorithm has little influence on the prediction accuracy.

#### V. CONCLUSION

The short-term wind speed prediction model proposed in this paper combines TVF, EMD, PE, CEEMDAN, GMDH, ANFIS, and IMOB. In the model, the deep secondary decomposition method implemented on TVF, EMD, PE, CEEMDAN, is used as a data preprocessing strategy to effectively decompose the original wind speed sequence to generate 11 sub-sequences. Then the proposed IMOB optimizes the GMDH and ANFIS models to obtain the optimal parameters of the model, and at the same time avoids the model from falling into the local optimum during the prediction process. After that, GMDH is used to process the first five sequences, while ANFIS is used to process the remaining sequences. Finally, the components of the obtained prediction results are summed to obtain the predicted value of the original wind speed series.

The model extracts the advantages of meta-heuristic algorithms and artificial neural networks to obtain a hybrid predictor, and uses data mining technology to process the original wind speed data to improve data quality. The TPD-IBMO-ANFIS model, TPD-IBMO-GMDH model, PSO-PROPOSED model, GA-PROPOSED model and MOB-PROPOSED model are compared to three wind speed data sets. After obtaining the final wind speed prediction of different models, we can get that the secondary decomposition strategy advocated in this paper can make the model additionally use the data features to achieve better prediction results. For example, in the predictive index of data set three in experiment one, the MAPE and RMSE are 4.426% and 0.423 respectively. The values of MAPE and RMSE of the GMDH model are 15.593% and 1.456.

#### REFERENCES

- [1] Y. Cui, C. Huang and Y. Cui, "A novel compound wind speed forecasting model based on the back propagation neural network optimized by bat algorithm," *Environmental Science and Pollution Research*, vol. 27, no. 7, pp. 7353-7365, 2020.
- [2] M. U. Yousuf, I. Al-Bahadly and E. Avci, "Current perspective on the accuracy of deterministic wind speed and power forecasting," *IEEE Access*, vol. 7, pp. 159547-159564, 2019.
- [3] Z. Liu, P. Jiang, L. Zhang and X. Niu, "A combined forecasting model for time series: Application to short-term wind speed forecasting," *Applied Energy*, vol. 259, pp. article. 114137 (1-25), 2020.
- [4] J. Zhao, J. Wang, Z. Guo and Y. Guo, et al., "Multi-step wind speed forecasting based on numerical simulations and an optimized stochastic ensemble method," *Applied Energy*, vol. 255, pp. article. 113833(1-16), 2019.
- [5] G. Chen, L. Li, Z. Zhang and S. Li, "Short-Term wind speed forecasting with principle-subordinate predictor based on Conv-LSTM and improved BPNN," *IEEE Access*, vol. 8, pp. 67955-67973, 2020.
- [6] O. B. Shukur and M. H. Lee, "Daily wind speed forecasting through hybrid KF-ANN model based on ARIMA," *Renewable Energy*, vol. 76, pp. 637-647, 2015.
- [7] P. Zhang, Y. Liu, F. Wu and S. Liu, "Low-Overhead and High-Precision prediction model for content-based sensor search in the internet of things," *IEEE Communications Letters*, vol. 20, no. 4, pp. 720-723, 2016.
- [8] P. Jiang, Y. Wang and J. Wang, "Short-term wind speed forecasting using a hybrid model," *Energy*, vol. 119, pp. 561-577, 2017.
- [9] K. Rouzbehi, A. Miranian, J. I. Candela and A. Luna, et al, "A control strategy for DC-link voltage control containing PV generation and energy storage," *An intelligent approach*, vol. 01, no. 1, pp.268-271, 2014.
- [10] W. Zhang, Z. Qu, K. Zhang and W. Mao, "A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting," *Energy Conversion and Management*, vol. 136, pp. 439-451, 2017.
- [11] Q. Z. W. J. Zhang K, "A novel hybrid approach based on cuckoo search optimization algorithm for short-term wind speed forecasting," *Environmental Progress & Sustainable Energy*, vol. 36, pp. 943-952, 2016.
- [12] H. Liu, C. Chen, X. Lv, X. Wu and M. Liu, "Deterministic wind energy forecasting: A review of intelligent predictors and auxiliary methods," *Energy Conversion and Management*, vol. 195, pp. 328-345, 2019.
- [13] Y. Hu and L. Chen, "A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and differential evolution algorithm," *Energy Conversion and Management*, vol. 173, pp. 123-142, 2018.
- [14] W. Jian-zhou, H. Jia-ni, X. Liye and W. Chen, "Research and application of a combined model based on multi-objective optimization for multi-step ahead wind speed forecasting," *Energy*, vol. 125, pp. 591-613, 2017.
- [15] X. Yuan, C. Chen, M. Jiang and Y. Yuan, "Prediction interval of wind power using parameter optimized beta distribution based LSTM model," *Applied Soft Computing*, vol. 82, pp. article. 105550 (1-10), 2019.
- [16] P. Jiang, H. Yang and J. Heng, "A hybrid forecasting system based on fuzzy time series and multi-objective optimization for wind speed forecasting," *Applied Energy*, vol. 235, pp. 786-801, 2019.
- [17] Y. Hao and C. Tian, "A novel two-stage forecasting model based on error factor and ensemble method for multi-step wind power forecasting," *Applied Energy*, vol. 238, pp. 368-383, 2019.
- [18] X. Luo, J. Sun, L. Wang and W. Wang, "Short-term wind speed forecasting via stacked extreme learning machine with generalized correntropy," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4963-4971, 2018.
- [19] S. Nan, Z. Su-quan, Z. Xian-hui, S. Xun-wen and Z. Xiao-yan, "Wind speed forecasting based on grey predictor and genetic neural network models," *IEEE Access*, vol. 14, pp. 1479-1482, 2013.
- [20] H. S. Dhiman, D. Deb and J. M. Guerrero, "Hybrid machine intelligent SVR variants for wind forecasting and ramp events," *Renewable and Sustainable Energy Reviews*, vol. 108, pp. 369-379, 2019.
- [21] Z. L. W. M. Heng Li, "A time varying filter approach for empirical mode decomposition," *Accepted Manuscript*, vol. 138, pp. 146-158, 2017.
- [22] Y. Cao, W. Tung, J. B. Gao, V. A. Protopopescu and L. M. Hively, "Detecting dynamical changes in time series using the permutation entropy," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 70, no. 4, pp. 46217, 2004.
- [23] L. Li, Z. Liu, T. Ming, Z. Sheng, L. Ming, "Improved tunicate swarm algorithm: Solving the dynamic economic emission dispatch problems," *Applied Soft Computing*, vol. 108, pp. 1754-1767, 2021.
- [24] S. Phoong and P. P. Vaidyanathan, "Time-varying filters and filter banks: some basic principles," *IEEE Transactions on Signal Processing*, vol. 44, no. 12, pp. 2971-2987, 1996.
- [25] W. Abd.Fatah, Z. Mohammad Izat Emir, Z. Rozaimi, "B-Spline curve interpolation model by using intuitionistic fuzzy approach," *IAENG International Journal of Applied Mathematics*, vol. 50, no. 4, pp. 760-766, 2020.
- [26] A. A. M. E. M. Unser, "B-Spline signal processing: Part I-theory," *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 821-833, 1993.
- [27] Y. Yang, "Empirical mode decomposition as a time-varying multirate signal processing system," *Mechanical Systems and Signal Processing*, vol. 76-77, pp. 759-770, 2016.

- [28] P. B. C. W. E. Tsironi, "An analysis of convolutional long short-term memory recurrent neural networks for gesture recognition," *Neurocomputing*, vol. 268, pp. 76-86, Dec. 2017.
- [29] H. Putriaji, Subanar, Abdurakhman and Tarno, "ANFIS performance evaluation for predicting time series with calendar effects," *IAENG International Journal of Applied Mathematics*, vol. 50, no. 3, pp. 587-598, 2021.
- [30] A. Khosravi, L. Machado and R. O. Nunes, "Time-series prediction of wind speed using machine learning algorithms: A case study Osorio wind farm, Brazil," *Applied Energy*, vol. 224, pp. 550-566, 2018.
- [31] Y. Jiang and G. Huang, "Short-term wind speed prediction: Hybrid of ensemble empirical mode decomposition, feature selection and error correction," *Energy Conversion and Management*, vol. 144, pp. 340-350, 2017.
- [32] B. Bahadir, Y. Yusuf, "Empirical mode decomposition based denoising method with support vector regression for time series prediction: a case study for electricity load forecasting," *Measurement*, vol. 103, pp. 52-61, 2017.
- [33] G. Wang, X. Zhao and S. Deb, "A novel monarch butterfly optimization with greedy strategy and self-adaptive crossover operator," *2015 Second International Conference on Soft Computing and Machine Intelligence*, vol. 2015, pp. 23-24, 2015.
- [34] A. Ahmed and M. Khalid, "An intelligent framework for short-term multi-step wind speed forecasting based on Functional Networks," *Applied Energy*, vol. 225, pp. 902-911, 2018.