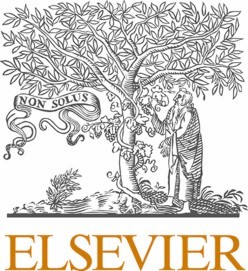
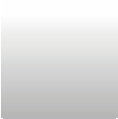
[Applied Energy](https://doi.org/10.1016/j.apenergy.2021.117446) [301](https://doi.org/10.1016/j.apenergy.2021.117446) [(2021) 117446](https://doi.org/10.1016/j.apenergy.2021.117446)

Review of meta-heuristic algorithms for wind power prediction: Methodologies, applications and challenges



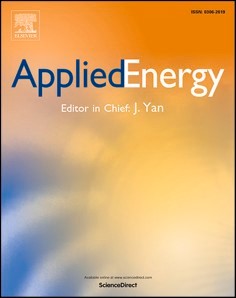
Contents lists available at

ScienceDirect

Applied Energy

journal homepage:

www.elsevier.com/locate/apenergy



Peng Lu a, Lin Ye a,\*, Yongning Zhao a, Binhua Dai a, Ming Pei a, Yong Tang b

a *College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, PR China* b *State Key Laboratory of Power Gird Safety and Energy Conservation (China Electric Power Research Institute), Haidian District, Beijing 100192, PR China*

H I G H L I G H T S

* A comprehensive literature review of meta-heuristic optimization algorithms for wind power forecasting is conducted.
* A well-designed taxonomy of meta-heuristic algorithms for optimizing wind power forecasting model parameters is developed.
* A deeply comprehensive and scientific multiple error evaluation metric is proposed for wind power error analysis.
* Further research topics for future works in wind power forecasting are presented.

A R T I C L E I N F O A B S T R A C T

|  |  |
| --- | --- |
| *Keywords:*  Meta-heuristic algorithms  Wind power forecasting Combined approach  Multiple time horizons  Multiple error evaluation metrics | The integration of large-scale wind power introduces issues in modern power systems operations due to its strong randomness and volatility. These issues can be resolved via wind power forecasting that can provide comprehensive future information about wind power uncertainties. This paper presents a timely and comprehensive review of meta-heuristic algorithms in the framework of wind power forecasting. The framework is based on the auxiliary layer, forecasting base layer, and core layer. The auxiliary layer, such as the data-decomposition layer, decomposes the wind power time series into many relatively stationary subseries, and uses prediction models, including artificial neural networks (ANNs) and machine learning (ML). The core layer is based on meta-heuristic algorithms, which include evolutionary-based algorithms, physics-based algorithms, human-based algorithms, swarm-based algorithms, hybrid algorithms, and multi-objective optimization algorithms. These algorithms aim to search for the optimal solutions under constraints, which is highly significant for optimizing the key parameters of the prediction models. Besides, multiple error evaluation metrics, e.g., deterministic, uncertainty, and testing methods used in the field of wind power prediction are described. A quantitative analysis focusing on their advantages, disadvantages, forecasting accuracy, and computational costs are also provided. Finally, a few open research issues and trends related to the topic are discussed, which can contribute to improving the understanding of each wind power forecasting method. In general, this review paper provides valuable insights to wind power engineers. |

# 1. Introduction

## 1.1. Motivation

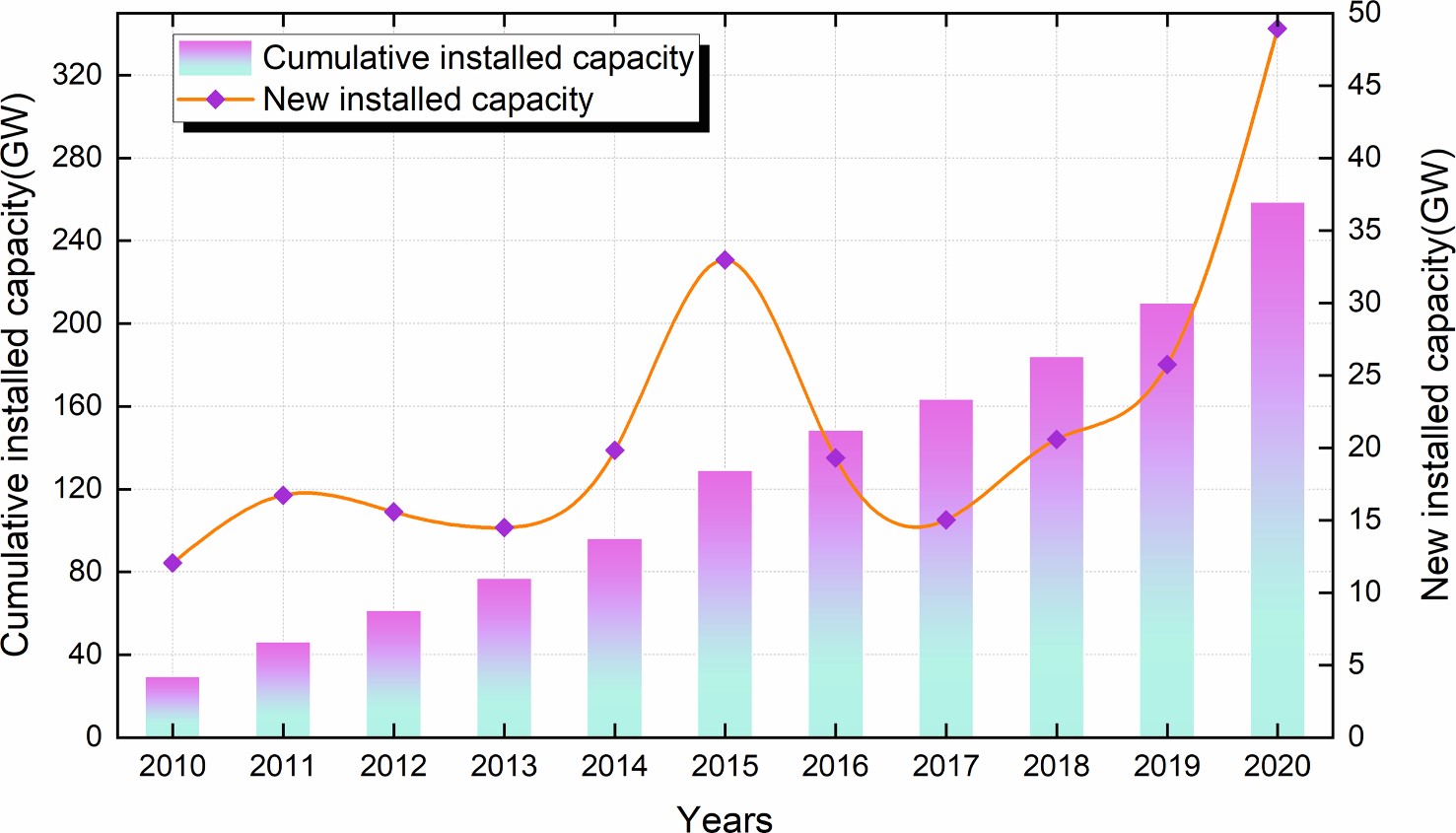
Wind energy is an alternative energy source that can effectively alleviate environmental issues. Therefore, several countries in the world have devoted their attention to this form of energy and the development of facilities for wind energy production. The report by the National Energy Administration (NEA) shows that the total wind capacity of China was more than 240 GW at the end of 2020 [1], the total wind capacity of China during 2010–2020 is shown in Fig. 1. It can be observed that the development and promotion of wind energy have been supported by China, and it has become one of the most rapidly growing renewable energy sources over the last decades. However, the integration of large-scale wind power in power systems will face significant challenges due to the uncertainty and intermittence of wind power. One of the challenges is related to operational problems, e.g., power and voltage control, and the other challenges are related to the planning and

|  |
| --- |
| \* Corresponding author.  *E-mail addresses:* lupeng@cau.edu.cn (P. Lu), yelin@cau.edu.cn (L. Ye).  <https://doi.org/10.1016/j.apenergy.2021.117446> |

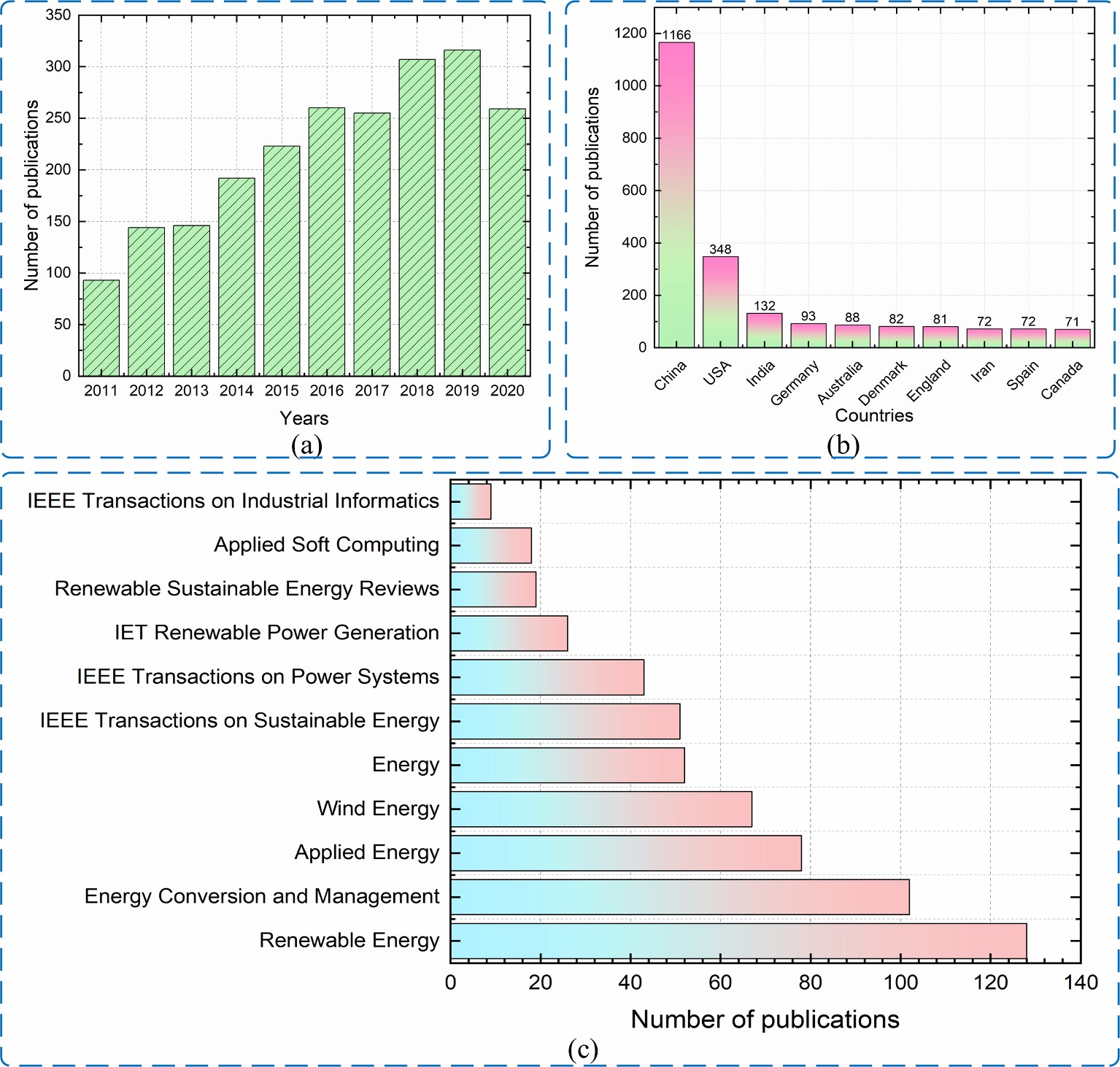
Received 9 April 2021; Received in revised form 18 June 2021; Accepted 11 July 2021

Available online 6 August 2021

0306-2619/© 2021 Elsevier Ltd. All rights reserved.



**Fig. 1.** New and cumulative wind power installed capacity from 2010 to 2020 in China.



**Fig. 2.** Number of publications in years, countries, and most popular journals.

dispatch problems, e.g., wind power dispatch, load dispatch, and spinning reserve [2,3]. A possible solution to overcome these challenges is to improve wind power forecasting accuracy and reliability, where an efficacy wind power forecasting model can help power system operators to make better decisions and increase their profit [4,5].

## 1.2. Literature review

A bibliometric analysis was conducted on December 31, 2020 to provide an overview of the existing research on wind power prediction using the well-established and acknowledged databases, e.g., Web of Science (WoS). The query used for WoS is as follows: TS = (“wind power forecasting” OR “prediction” OR “wind speed forecasting” OR “wind prediction”).

Fig. 2(a) shows the number of publications indexed by the WoS from 2011 to 2020. A total of 2195 publications are found in the WoS. The number of publications increased rapidly from 2011 to 2019. However, the number in 2020 is relatively smaller compared to that in 2019. Fig. 2 (b) depicts the countries ranked according to the number of relevant papers published in the WoS. Out of all the countries, the most relevant papers are published in China and totaled 1166. This number of publications is not surprising, given the significant amount of funding allocated to the wind power industry by the Chinese government in the past decade. The Energy Administration (EA) and the State Grid Corporation of China (SGCC) have also issued guidelines for wind power dispatch and control, which consider wind power prediction as a key technology. Furthermore, from 2010 to 2020, the National Natural Science Foundation of China (NSFC) sponsored more than 30 projects related to wind prediction. Fig. 2(c) shows the journals ranked by the number of relevant papers published according to the WoS. A total of 128 papers are published on the topic of wind prediction in Renewable Energy, a well- known international journal.

Since the publication of the first article on wind power prediction [6], several forecasting approaches have been proposed that use the historical wind speed and wind power to predict the output wind power in wind farms [7,8]. These approaches can be divided into four groups according to the characteristics of a forecasting model: physical forecasting methods [9], statistical forecasting methods [10], artificial intelligence-based forecasting methods [11], and combined forecasting methods [12]. The main aim of all the methods is to increase the wind power prediction accuracy while simultaneously minimizing the computational cost.

The physical forecasting methods utilize a complex mathematical model to calculate the actual wind power output using the numerical weather prediction (NWP) data [13], which is well suited for short-term wind power forecasting. However, the disadvantage of the physical forecasting methods lies in their high computational cost. Moreover, in most cases, it is difficult to obtain the NWP data from weather stations, which is one of the reasons for the lack of popularity of the physical forecasting methods among academic researchers.

The statistical forecasting methods are used for ultra short-term prediction [14] or short-term prediction [15]. These methods include the auto-regressive model (AR) [16], vector autoregression (VAR) [17], auto-regressive moving average model (ARMA) [18], autoregressive integrated moving average model (ARIMA) [19], fractional autoregressive integrated moving average model (FARIMA) [20], and Kalman filter-based model [21]. These methods are simple and easy to implement. However, these methods can not accurately process noisy wind power time series data with high fluctuations, due to which the wind power forecasting results include large errors [22]. Furthermore, it is difficult for these methods to achieve satisfactory prediction results for input data with nonlinear characteristics.

The artificial neural networks (ANNs) forecasting methods can give more information on the uncertainty of wind power forecasts [23], which the results are provided by the output layer(s) [24]. The ANNs are trained by using historical wind power values to learn the relationship between wind power input and output [25]. The extreme learning machine (ELM) [26] and backpropagation neural network (BPNN) [27] are the most popular ANNs for wind power forecasting. However, it is difficult to obtain accurate and reliable forecasting results from the ANNs due to the inherent randomness of the initial network parameters, such as weights and thresholds, smoothing factors, etc.

The combined forecasting methods can compensate for the shortcomings of an individual forecasting method. These methods can be further divided into two groups [28,29]: the first group that combines physical and statistical forecasting methods [30]. The methods in this group have a computational complexity that is significantly high for a personal computer. The second group consists of using meta-heuristic algorithms to tune the key parameters of the prediction model [31], which can be summarized as the following paradigm: “data processing + forecasting model + optimization algorithm”. In these methods, a proper selection of forecasting model parameters is essential for predicting the output wind power.

In recent years, meta-heuristic algorithms have become popular among many researchers due to their effectiveness, robustness, and flexibility. Many articles about these algorithms have been published in top journals. These algorithms have been applied to a variety of problems, such as cogeneration scheduling [32], architectural design optimization [33], maximum power point tracking of PV systems [33], optimal allocation of distributed generation [34], wind energy integration to the grid [35], photovoltaic-based power forecasting, and short- term scheduling of hydro-based power plants [36,37]. The low wind power prediction accuracy could be improved by using meta-heuristic algorithms.

The goal of the meta-heuristic algorithms is to find the optimal parameters of forecasting models, e.g., the initial parameters of the ANNs, the kernel parameters of the support vector machine (SVM), or the weights of a combined model, subject to different constraints [38]. In [39], a combined prediction model that consists of ANNs and SVM is proposed for the real-time wind-electric power generation prediction, which obtains a promising result. A similar wind prediction framework is applied in [40], in this case, a new wind pattern recognition and reference wind mast (RWM) data correlations with numerical weather predictions (NWP) is first proposed, which can improve the accuracy of short-term wind-electric power prediction. In [41], the optimal prediction intervals (PIs) of wind power based on the ELM were obtained using particle swarm optimization (PSO). Furthermore, the weights of the neural networks were tuned using the weight-improved PSO [42]. The application of a genetic algorithm (GA) to select the weights and thresholds of the backpropagation neural network (BPNN) is presented in [43]. Besides, the initial weights and thresholds of a BPNN and the SVM model’s parameters are also optimized by the GA in [44].

Different kernel functions have different performances. In [45], the kernel function parameters of the least squares support vector machine (LSSVM) are optimized using the gravitational search algorithm (GSA), and subsequently, the best parameters are selected for the LSSVM model. In addition, a few other meta-heuristic algorithms, such as the bat algorithm (BA), grey wolf optimizer (GWO) [46], whale optimization algorithm (WOA) [47], artificial bee colony (ABC) [48], and firefly algorithm (FA) [49] have been successfully applied for predicting the wind power output.

To solve multiple objective functions with high prediction accuracy, stability, and reliability, multi-objective optimization algorithms inspired by nature can be used. These algorithms include the multi- objective whale optimization algorithm (MOWOA) [50], non- dominated sorting genetic algorithm II (NSGA-II) [51], multi-objective chaotic water cycle algorithm (MOCWCA) [52], multi-objective imperialist competitive algorithm (MOICA) [53], multi-objective particle swarm optimization algorithm (MOPSO) [54], and multi-objective bat algorithm (MOBA) [55]. The multi-objective optimization algorithms may be preferable due to the non-stationary and nonlinear characteristics of the wind power values.

There have been several comprehensive reviews on the topics of uncertainty analysis [56], probabilistic forecasting [56], combined approaches forecasting [57], physical approaches forecasting [8], statistical approaches forecasting [15], and artificial intelligence approaches forecasting [58]. Furthermore, a larger number of publications have focused on the modeling of meta-heuristic optimization algorithms to predict the output wind power. However, as there is no review paper on this topic of wind power prediction, the main target of this paper is to review wind power forecasting approaches based on meta-heuristic algorithms. An additional key contribution of this paper is the study of using the diversity of meta-heuristic algorithms to predict the wind power output with the optimal parameters.

|  |
| --- |
| **Fig. 3.** The framework of wind power forecasting based on the meta-heuristic algorithms. |

## 1.3. Contribution

Initial wind power prediction usually focused on a model formulation. Related literature contained reviews of physical methods, statistical methods, probability forecasting, and combined forecasting methods, ignoring meta-heuristic optimization algorithms. However, the meta-heuristic optimization algorithms are essential for improving the prediction accuracy and stability of the models. Therefore, it is also necessary to summarize the meta-heuristic optimization algorithms, which is essential for obtaining a reliable prediction result.

The contributions of this paper are as follows:

1. Conducting a comprehensive literature review of meta-heuristic optimization algorithms for wind power forecasting.
2. Providing a well-designed taxonomy for wind power forecasting.
3. The diversity of meta-heuristic algorithms for optimizing the parameters of the wind power forecasting model is analyzed.
4. Open research issues and future works are discussed.

The rest of this paper is organized as follows: Section 2 provides the framework of wind power prediction based on meta-heuristic algorithms. Section 3 illustrates the advantages and disadvantages of the auxiliary layer. The forecasting base layer and Key core layer are presented in Sections 4 and 5, respectively. Section 6 describes multiple error evaluation metrics for the developed forecasting models. The performance of different meta-heuristic algorithms is compared in Section 7. Open research issues and future work are discussed in Section 8. The conclusions of this review paper are given in Section 9.

# 2. The framework of wind power prediction based on meta- heuristic algorithms *2.1. Basic mathematical statement*

For a trained prediction model *f*(⋅), the typical wind power prediction problem is translated in mathematical terms as follows:

⎛ ⎞

*pret* ⏟⏞⏞⏟ ⎜ *t* ⎟ (1) *p* = *f* ⎝ ⏟⏞⏞⏟*θ ,* ⏟⏞⏞⏟*x* ⎠

prediction model unknown parameters inputs

where ***θ*** and ***x****t* are the unknown parameters and inputs of the prediction model, respectively, and *f*(⋅) represents any prediction model, such as artificial neural networks (ANNs) and machine learning (ML).

The unknown parameters in formula 1 are very important for the prediction model. Generally, to obtain the unknown parameters of the model, a single objective and multi-objective optimization objective function can be established.

In addition to the above-mentioned prediction model based on parameter optimization, the weight-based combination model is also a widely accepted prediction method. Assume that *pprek,t* is the prediction value of the *k* prediction model at time *t*, the combination models’ output can be expressed as:

*pprec,t* ⏟⏞⏞⏟*k,t pprek,t* (2)

*weights*

where *k* is the number of the individual prediction model, *ωk,t* is the weight of the model *k*, *pprek,t* is the prediction value of model *k* at time *t*.

## 2.2. Overall framework description

Based on state-of-the-art references from well-known international energy journals, this paper proposes a wind power forecasting framework based on the meta-heuristic algorithm, which aims to help researchers to study wind power forecasts in-depth and improve forecasting accuracy and reliability. Fig. 3 shows the framework of wind power forecasting based on meta-heuristic algorithms. It consists of three layers: an auxiliary layer (data-decomposition layer), a forecasting base layer (the forecasting model), and a key core layer (optimization- algorithm layer). The goal of the auxiliary layer is to cope with undesired data obtained from actual running wind farms to ensure data quality. The goal of the forecasting base layer is to predict the output wind power over multi-time horizons. A detailed description of this layer can be found in [9,59]. In the key core layer, the optimal parameters, e.g., weights and thresholds of a prediction model such as ANNs are selected to meet the prediction requirements.

|  |
| --- |
| **Fig. 4.** The diagram of the auxiliary layer. |

# 3. Auxiliary layer (Data Decomposition)

## 3.1. Basic mathematical statement

It is difficult to develop a prediction model due to the intermittency and non-stationarity of wind power time series. The data decomposition method can extract the inherent characteristics of the time series and improve the quality of the data. A standard data decomposition method can be expressed as follows:

̃*pt* = ⏟⏞⏞⏟*F* ( ⏟⏞⏞⏟*ϑ ,* ⏟⏞⏞⏟̃***x****t* ) (3)

Decomposition Model Original model parameter signal

where ̃*pt* is the decomposed wind power signal at time horizon *t*, *F*(⋅) is the decomposition methods, *ϑ* is the parameters of decomposition model, and ̃*xt* is the original signal that loaded to the decomposition model.

## 3.2. Decomposition method description

Data decomposition can cope with undesired artifacts in the data acquired from actual running wind farms, which plays an important role in ensuring data quality [9]. Fig. 4 shows the popular decomposition algorithms. Different decomposition methods used in the data- decomposition layer include the empirical mode decomposition (EMD) and its improved versions, the wavelet transform (WT), the variational mode decomposition (VMD), the wavelet packet decomposition (WPD), and the singular spectrum analysis (SSA). Generally, the original wind

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Table 1**  The advantage and disadvantage of decomposition methods.   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | Methods | Results | Advantage | Disadvantage | Preprocessing | References | | EMD | Decomposing wind power time series into several IMFs and one residue | • Decomposition of complex wind power time series | • Endpoint effect and mode mixing | ✓ | [60,61] | | EEMD | Adding white noise to wind power time series | • Wind power of different scales is associated with white noise | • Each IMF has more noise | ✓ | [62] | | EEMD | Using opposite polarity | • Greatly reduce the computational efficiency of EEMD | • Increase the number of EMD operations | ✓ | [63,64] | | CEEMDAN | Using adaptive noise | • Reconstruction is noise-free. | • Containing a certain amount of residual noise | ✓ | [65,66] | | WT | Decomposing wind power into components in low frequency and high frequency | • Having a good localization property | • Have a limitation on the size of the input wind power | ✓ | [67-69] | | WPD | Decomposing approximation component and detailed component | • Obtaining a complete decomposition binary tree | • Difficult to choose wavelet packet base | ✓ | [70] | | VMD | Decomposing the non-stationary Wind power into multiple band-limited intrinsic mode functions | • Avoiding errors caused during the recursive calculations | • Easy to cause decomposition or under-decomposition | ✓ | [71,72] | | SSA | Decomposing wind power into a set of independent principal components | • Removing wind power time series noise | • Data conversion can be difficult to understand | ✓ | [73] |     **Fig. 5.** The parameters optimization of different wind power prediction models. |

power values are decomposed into subsets of different frequencies and then fed into the forecasting models. This reduces the prediction complexity and improves the prediction accuracy. The advantages and disadvantages of different types of data decomposition methods are given in Table 1.

# 4. Forecasting base layer (Prediction Model)

## 4.1. Basic mathematical statement

The forecasting base layer contains two types of prediction models, one is based on ANNs, and the other is based on ML. Based on equations 1–2, the basic idea is as follows: in the training model stage, the input variables of the model at time *t* are given, the optimal parameter ***θ*** is determined by the meta-heuristic optimization method, and then the optimal parameters are set in the model testing stage for wind power forecasting, which can be defined as:

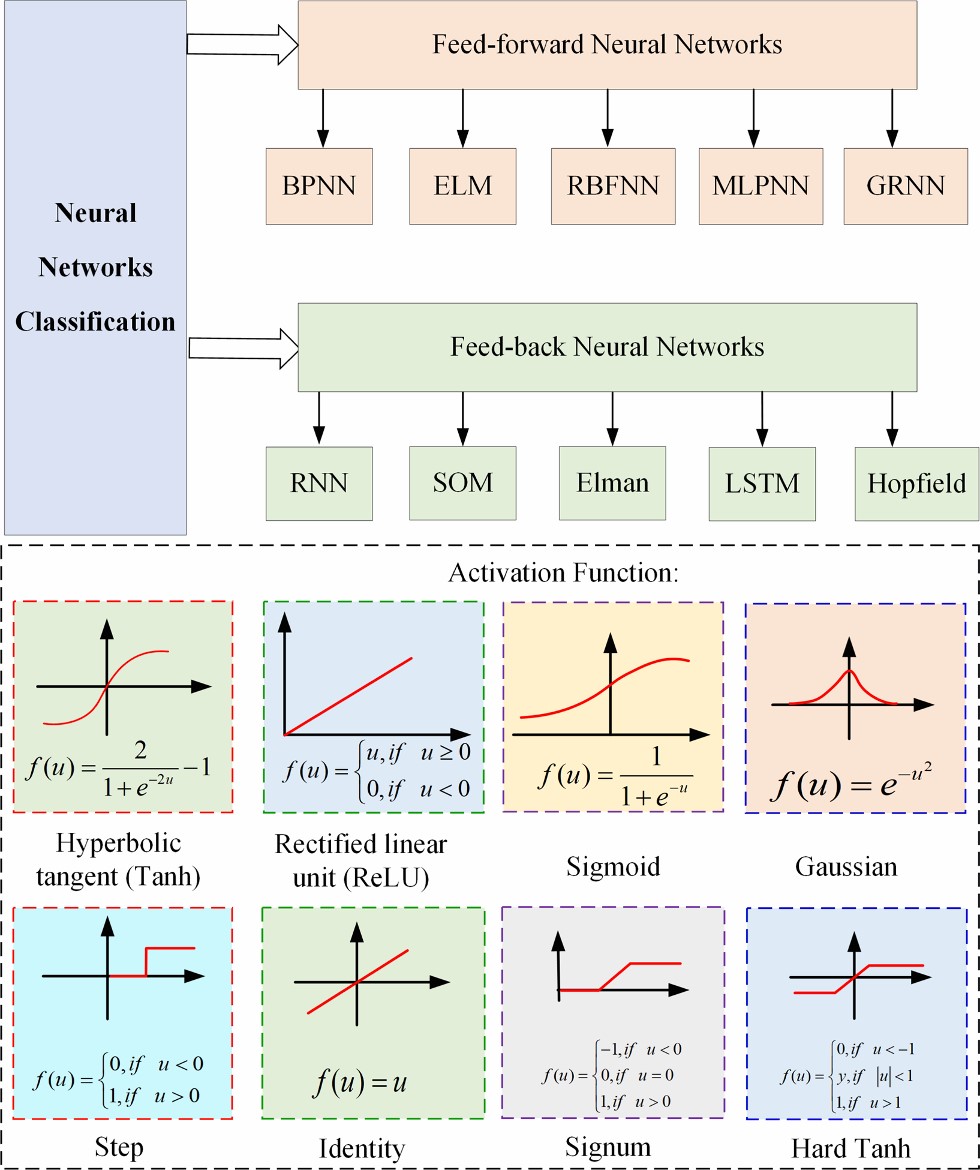
ANNs

⏞⏟⏟⏞ *ptpre*+*l*|*t* = ⏟̅̅̅̅̅̅̅̅⏞⏞̅̅̅̅̅̅̅̅⏟*f* or ML prediction (⏟⏞⏞⏟***θ*** key *,* ⏟⏞⏞⏟***ω*** weight *,* ⏟⏞⏞⏟***x****t* inputs)

model parameters coefficient

(4)

where *ppret*+*l*|*t* is the prediction value at time horizon *t* + *l*, ***x****t* is the inputs of the prediction model. *f*(⋅) is the prediction model, which includes ANNs and ML. ***θ***is the key parameters of a prediction model that affect the prediction accuracy. The parameter of different models as follows: the initial weights and thresholds of the back propagation neural network (BPNN) or radial basis function neural networks (RBFNN), the weights and biases of the ELM or multi-layer perception (MLP), the smoothing



**Fig. 6.** Classification of neural networks and activation functions [78].

factors of generalized regression neural network (GRNN), the learning rates of Elman neural networks (Elman), the biases of Long-short term memory (LSTM) networks, and the regularization and kernel parameters of SVM/LSSVM [74,75]. ***ω***is the weight coefficient of the combined model. Fig. 5. Shows the parameters optimization of different wind power prediction models. The principles of ANNs and ML for wind power prediction are detailed in [29,58]. In the following, we describe the principles of each prediction model.

## 4.2. Artificial neural networks (ANNs)

The ANNs have been a research hotspot in the field of artificial intelligence since the 1980s [76]. They abstract the neural network of the human brain from the perspective of information processing, establish a few simple models, and form different networks based on different connection modes. In engineering and academic circles, they are often referred to as neural networks or quasi-neutral networks [77]. The ANNs are composed of a large number of nodes, called neurons, each of which represents a specific output function. The signal passing through the connection between each node is weighed by a certain value, which is called the weight. The output of an ANN depends on the connection mode of the network, the weight values, and the activation functions. The ANNs have been widely used for wind speed and power forecasting. They can be grouped into two main types depending on the network structure, as shown in Fig. 6: feed-forward neural networks (FFNNs) and feedback-forward neural networks (FBNNs). The former includes the MLP, ELM, RBFNN, BPNN and GRNN. The latter include the recursive neural network (RNN), LSTM, Hopfield neural network (Hopfield), self- organizing mapping network (SOM) and Elman.

### 4.2.1. Feed-forward neural networks (FFNNs)

The FFNNs model is the simple kind of neural network, where each neuron is arranged in layers. Each neuron is only connected with the neuron of the preceding layer. It receives the output of the previous layer, which is multiplied by the weight and the result is then output to the next layer. A nonlinear mapping relation is implemented between the input and output layers. There is no feedback between the layers. The FFNNs models include the BPNN, ELM, RBFNN, MLP and GRNN. The structure of five types of feed-forward neural networks is shown in Fig. 7.

*4.2.1.1. General regression neural network (GRNN).* A GRNN model based on mathematical statistics was proposed in [79]. It is based on nonlinear regression analysis. Its structure is composed of four layers, namely the input layer, pattern layer, sum layer, and output layer. The GRNN has strong nonlinear mapping capability and learning speed. It can also obtain high forecast accuracy for small wind power/speed time series sample data. It has been successfully applied for wind power/wind speed prediction.

In[80], GRNN and MLP were used to predict wind speeds in 67 cities of India. The results showed that the former performed better than the latter. Moreover, the GRNN was used to predict the cooling load profiles of each building. The results showed the effectiveness of the GRNN model for predicting cooling load [81].

*4.2.1.2. Radial basis function neural network (RBFNN).* An RBFNN model [82] uses radial basis functions as activation functions. It is a three-layer network, consisting of the input layer, hidden layer, and output layer. In the hidden layer, Gaussian function is used as the activation function of the local response. In the forward network, the activation function is generally the global response function. The RBFNN network can approximate any continuous function with high precision within a short training time.

The structure of RBF was modified in [83] for wind power forecasting, which could work with non-stationary series exhibiting homogeneous no-stationary behavior. The optimal number of neurons in the RBFNN hidden layer was estimated by a cross-validation subspace method [84], followed by the use of the RBFNN to predict chaotic noise time series. Furthermore, the RBFNN model was used to forecast a set of quantiles with predefined nominal wind power probability [82]. The RBF was also developed for short-term system load prediction. This case demonstrates that the RBF model can predict and estimate the short- term prediction confidence interval, and gives the index of calculation reliability [85].

*4.2.1.3. Back propagation neural network (BPNN).* The BPNN is a popular ANN model [86]. It is a multi-layer feed-forward network trained by an inverse error propagation algorithm, which is the most outstanding part of an ANN. A BPNN was established to forecast wind speeds by the nonlinear wind speed sequences [87]. The feasibility and validity of this method were proved by simulation experiments. In [88], a BPNN model was used to forecast the wind power generation of wind energy conversion system (WECS). However, it is difficult to choose the optimal number of neurons in the hidden layer during the complex mapping process. Furthermore, in [89], a novel criterion is proposed to select the number of neurons in the hidden layer, which can improve the accuracy of the BPNN model.

*4.2.1.4. Extreme learning machine (ELM).* To overcome the problem of low learning efficiency and tedious parameter settings in a traditional neural network algorithm, the ELM is proposed in [90], which has good generalization performance. Several studies have demonstrated in recent years that the ELM model is effective for wind speed/power forecasting.

|  |
| --- |
| **Fig. 7.** The structure of five kinds of feed-forward neural networks. |

An ELM model was employed in [91] to predict short-term wind speed. A short-term combined prediction model based on the ELM and masking signal-based VMD (VMD-EMD) was proposed in [92] for wind power forecasting. As the quantile regression can more comprehensively describe the wind power time series, a composite interval forecasting ELM was used for multi-step wind speed prediction. In [93], a novel wind speed prediction framework is developed, which composites quantile regression outlier-robust ELM with feature selection and parameter optimization using a hybrid population-based algorithm. In addition, different models have been proposed for wind speed/power forecasting, such as the outlier robust ELM for probabilistic wind speed forecasting [94], AdaBoost ELM [95], ridge ELM (R-ELM) [96], and evolutionary ELM [97] for short-term wind power forecasting and electricity price forecasting. All these studies demonstrate that the ELM has higher accuracy than other benchmark forecasting models.

|  |
| --- |
| **Fig. 8.** Structures of five kinds of FBNNs. |

*4.2.1.5. Multi-layer perceptron neural network (MLP).* An MLP model belongs to the class of FFNNs, which can be seen as a directed graph consisting of multiple node layers [98], where each layer is connected to the next layer. Except for the input nodes, the neurons in each node have nonlinear activation functions [99].

An MLP model was used for short-term load prediction, using data from three years for training. The simulation results demonstrated that this method performed well for load prediction [100]. A neural network- based hybrid computing model was proposed in [101] for forecasting wind speed in renewable energy systems. The results showed that the proposed model performed better in terms of error minimization compared to other methods. Moreover, it was shown that the number of neurons in the hidden layer affected the wind speed prediction accuracy.

A method to estimate the optimal number of neurons in the hidden layer of ANN was proposed in [102]. The simulation results showed that the method could reduce the calculation error.

**Remark:**. *In summary*, *the FFNN models*, *such as the BPNN*, *ELM*, *RBFNN*, *MLP and GRNN are based on classical nonlinear statistical models*. *These models have the advantages of simple and easy implementation*, *which are widely used in wind power*/*wind speed prediction*. *However*, *a few parameters of the models affect the prediction accuracy*, *such as the weight and threshold*, *and smoothing factor*. *The prediction accuracy can be further improved by selecting optimal parameters of a selected model*.

Each model can be used for ultra-short and short-term wind speed/ wind power prediction. Unfortunately, it is impossible to single out one model that is the most suitable for wind speed/wind power prediction. An increasing number of researchers use ELM for wind power forecasting, due to its good generalization performance and faster computation speed compared to the networks trained using the back- propagation algorithms.

### 4.2.2. Feedback neural networks (FBNNs)

The FBNNs consider the time delay between input and output and should be described by dynamic equations. Each neuron in the network sends its output signal as input to other neurons, which requires a certain amount of time to achieve stability. The FBNN models include the SOMNN, Elman, LSTM, RNN and Hopfield, and the structures of these five kinds of feedback neural networks are shown in Fig. 8.

*4.2.2.1. Self-organizing mapping network (SOMNN).* The SOMNN is an important type of neural network based on unsupervised learning [103]. It stimulates the division of labor between nerve cells in different regions of the brain, i.e., different regions have different response characteristics, and the process is automatic. The SOMNN classifies the input pattern set by finding the optimal reference vector set, where each reference vector is a connecting weight vector corresponding to an output unit.

A phase space reconstruction method and the self-organizing map (SOM) were used for wind speed prediction in [104], Correlation analysis was used to determine the number of clusters and subsequently, the input of the prediction model was determined. A combination of SOM and ELM (SOM-ELM) was used for short-term wind power prediction in [105]. The SOM method was also applied for short-term load forecasting and solar power forecast (SPF) with weather classification [106,107].

*4.2.2.2. Elman neural network (Elman).* The Elman was proposed by J. L. Elman in 1990 for speech processing [108], which is a typical global feedforward local regression. The Elman can be considered as an RNN with local memory units and local feedback connections. Its prediction accuracy can be improved using data preprocessing. Several scholars have exploited the remarkable ability of EDM, EEMD, WPD and other decomposition technologies to process non-linear data by combining them with Elman to predict wind power. In [109], a hybrid method was proposed for short-term wind speed forecasting, which was based on wavelet packet decomposition, density-based spatial clustering of applications with noise, and the Elman (WPD-DBSCAN-ENN). Similarly, a novel hybrid deep-learning method with empirical wavelet transformation and two kinds of RNN was proposed for wind speed prediction [110]. Besides, several data preprocessing techniques for wind speed forecasting were proposed, including SSA-ENN [111], hour-ahead wind speed forecasting [112], rolling Elman [113], forecasting wind speed using FEEMD-Elman [114].

*4.2.2.3. Long-short term memory (LSTM).* The LSTM was first proposed in 1997. The LSTM is suitable for processing and predicting very long intervals and delays in time series, thanks to its unique design structure [115]. It is often used as a combined model for wind power prediction, and is rarely used as an individual prediction model. A novel hybrid model based on WPD, convolutional neural network (CNN) and CNNLSTM was developed for wind speed prediction in [116]. As both LSTM and ESN showed good performance for time series prediction, a hybrid model combining the advantages of these two networks was proposed for wind power forecasting [117,118]. In [119], the LSTM and deep learning neural network (DLNN) was first used to wind prediction. The results showed that the prediction precision of the model was higher than other benchmark models.

#### Table 2

Explanation of five kinds of kernel functions.

Function Function expression Function explanation/applications name

|  |  |  |
| --- | --- | --- |
| Liner | *k*(*xi,xj*) = *xTi xj* | Lack of training data |
| Polynomial | *k*(*xi,xj*) = (*γxTi xj* + *r*)*d,γ >*  0 | Denotes the similarity of training samples in feature space |
| Gaussian RBF | *k*(*xi,xj*) = *exp*(−  ⃦ ⃦2 0  *γ*⃦*xi* − *xj*⃦ )*,γ >* | Used in infinite dimensional feature space |
| Exponential RBF | (  *k*(*xi,xj*) = *exp* −  ⃦ ⃦)  ⃦*xi* − *xj*⃦ | It can implement nonlinear mapping |

Sigmoid *k*(*x*2*i,σx*2*j*) = *tanh*(*γxTi xj* + *r*) *r*is a shifting parameter that controls the threshold of aping.

*4.2.2.4. Recursive neural network (RNN).* The RNN is a general term used for two types of ANNs [120]: 1) recurrent neural network, and 2) recursive neural network. The time recursive neural network’s interneuron connection constitutes a matrix. The structural recursive neural network uses a similar neural network structure to recursively construct a more complex deep network.

In [121], the spatial information from a remote measurement station was analyzed. An online learning algorithm based on recursive prediction error (RPE) strategy was used to complete the training task. Subsequently, the trained RNN was used to predict wind speed. The RNN was also used to predict wind speed and wind power based on meteorological information, and showed better predictive performance than the ARIMA model [122,123].

*4.2.2.5. Hopfield neural network (Hopfield).* The Hopfield is an FBNNs invented by John Hopfield in 1982 [124]. Its output is feedback to its input, and under the excitation of the input, the output produces constant state changes. This feedback process is repeated. Once a stable equilibrium state is reached, the Hopfield outputs a stable constant value. The Hopfield model is trained using weather information including temperature, wind speed, and relative humidity. Compared with the MLP, ERNN, and RBFNN, high accuracy in 24 h-ahead weather forecasting can be achieved using the Hopfield [125].

## 4.3. Machine learning for wind power forecasting

Based on our current study, we find that the most common and popular models used for wind speed/wind power prediction are: machine learning (SVM and LSSVM) and ANN family models. The SVM was proposed in 1995 [126], which has become a hotspot in intelligent technology research. It can solve small sample, nonlinear and high- dimensional pattern recognition problems, exhibits many unique advantages, and can be applied to other machine learning problems such as function fitting. The SVM applied to power prediction falls in the category of regression problems. In this method, the wind power time series samples are mapped to the high-dimensional feature space via nonlinear mapping, and subsequently regressed linearly in the feature space. In recent years, the SVM has been used for wind prediction, which has improved the prediction accuracy to some extent [127].

However, the SVM has the drawback of requiring relatively complex calculations. To overcome this drawback, the least square support vector machine (LSSVM) was proposed in [128] The main difference between SVM and LSSVM is that the latter uses squared errors instead of non- negative errors in the cost function, as well as uses equality constraints rather than inequality constraints. Therefore, the essence of LSSVM is to transform the quadratic programming problem into a problem involving linear equations, which significantly shortens the computation time required for model learning.

In the SVM and LSSVM models, the choice of kernel functions, penalty factor (*γ*) and kernel function parameters (*σ*, *r*, *d*) directly influences their generalization capabilities. The use of LSSVM to solve the prediction problem is significantly dependent on the selection of kernel function and its parameter. Many kernel functions with different properties can be used in the SVM/LSSVM, such as Linear, Polynomial, Gaussian RBF, Sigmoid, and Exponential RBF. The features of different kernel functions are described in Table 2 [129,130].

In the SVM model, choosing different kernel functions can affect the prediction performance [131]. In [132], three SVM kernels, such as linear, Gaussian, and polynomial kernels are implemented for wind speed prediction in a wind farm. In [133], a hybrid wind speed prediction method based on the SVR was proposed. In this method, an autoregressive model of time-delay coordinates was used to carry out phase space reconstruction for feature selection. The SVR model was then trained using a single variable wind speed time series. The parameters of the SVR were adjusted using the GA. The results showed that the prediction performance of this model was more accurate for short- term WSF and WPF compared to the persistence and autoregressive models. In [132], the LSSVM model was used to carry out one-step ahead wind speed prediction. The effects of linear, Gaussian and polynomial kernel functions on prediction accuracy were discussed. The results showed that the performance of the linear kernel function was worse than the other two kernel functions for a small training sample size.

**Remark:**. *The aforementioned two types of individual prediction models including machine learning and artificial neural networks are widely used for wind speed*/*power forecasting*. *Their features can be summarized as follows*:

* The training of the FFNN is mainly based on the BP algorithm. When the sample size of wind power/wind speed time series is large, the calculation process and convergence speed are slow. For the FBNN, the training is mainly based on Hebb’s rule. The convergence speed of the calculation is very fast that can meet the requirements of having a short simulation time.
* The purpose of FFNN training is to quickly converge, and the degree of convergence is generally determined by the error function. The learning purpose of the FBNN is to quickly find a stable point, and the

⎧

{

}

⎪

*Obj*

(

1

)

=

1

*N*

∑

*T*

# 5. Key core layer (Meta-heuristic algorithms)

## 5.1. Basic mathematical statement

The main task of the key core layer is to optimize the parameters of the forecasting model by using various meta-heuristic algorithms in the model training stage, which can improve the accuracy and stability of the prediction model.

**Definition 1.**. ***Single*-*objective optimization problem*:**

The general single-objective optimization problem can be rewritten as:

( )2

*Objmse N ppret* − *prealt* (5) where *prealt* and *ppret* are the actual and predicted wind power value at time *t*, respectively. *N* is the number of the training set, *T* is the total prediction time.

=

1

∑

*T*

*t*

=

1

**Definition 2.**. ***Multi*-*objective optimization problem*:**

Unlike the single-objective optimization problem, multi-objective optimization with minimization problem can be expressed as:

*Minimize*: *F*(*x*) = {*f*1(*x*)*,f*2(*x*)*,*…*,fn*(*x*)}

*Subjectto*: *gj*(*x*)⩽0*, j* = 1*,*2*,*…*,m* (6) *hj*(*x*) = 0*, j* = 1*,*2*,*…*,p*

where *n* is the number of objective function*s*, *m* and *p* represent the number of inequality constraints and equality constraints, respectively.

|  |
| --- |
| ⏞̅̅̅̅̅⏟⏟̅̅̅̅̅⏞ *accuracy* ⏞̅̅̅̅̅⏟⏟̅̅̅̅̅⏞ *stability* ⎪⎪⎨ *mse* ( *ppret* − *prealt* )2  *Minimize Obj* = *Objmse*(1) *,Objstd*(2) = *t*=1 (7)  ⎪⎪⎪⎩ ( )  *Objstd*(2) = *std ppret* − *prealt* |

In the field of wind power forecasting, considering both accuracy and stability is a significant factor in power grid security. Firstly, a widely applied single-objective function *Objmse*(1) is developed to achieve high prediction accuracy. Next, the other objective function *Objstd*(2) that denotes the stability of the prediction errors is carried out, which has a significant impact on the stability of the prediction results. The multi- objective function *Obj* can be described as [51,134]:

stability is generally determined by the energy function.

• In general, the aforementioned individual prediction model is influenced by the parameter values. The accuracy of wind power prediction is low because the optimal parameters are usually selected manually by trial-and-error. Various statistical error tests show that individual forecasting models do not exhibit high wind power forecasting accuracy in most of the locations. The solution is to establish a combined forecasting model.

The core strategy for improving the low wind power prediction accuracy of the individual prediction models due to improper parameter selection is to use meta-heuristic algorithms. The next section will focus on the development of heuristic algorithms and parameter optimization of the aforementioned prediction models.

where *Objmse*(1)and *Objstd*(2) denote the accuracy and stability of the prediction model, *std*(⋅) is the standard deviation between predicted and actual values.

**Definition 2.**. *Weighting*-*based optimization problem*:

In addition to the above-mentioned prediction model based on parameter optimization, the weight-based combination model is also a widely accepted prediction method and has been widely used in the field of wind power prediction. Assuming that *pprek,t* is the prediction value of the *k* prediction model at time *t*, the combination models’ output can be expressed as:

*pprec,t* ⏟⏞⏞⏟*k,t pprek,t* (8)

*weights*

where *K* is the number of the individual prediction model, *ωk,t* is the weight of the model *k*, *pprek,t* is the prediction value of model *k* at time *t*.

|  |
| --- |
| **Fig. 9.** Classification of meta-heuristic algorithms. |

Usually, minimizing the sum of squares errors (SSE) of the combined prediction models’ output is used as the objective function, which can be expressed as:

⎧

⎪⎪⎪⎪⎪⎪⎪⎪⎨ *Minimize ObjSSEK*= ***ω****T****Eω***



⎪⎪⎪⎪⎪⎪⎪⎪⎩ *e* = *ptreal* − *pprek,t* *j*=1*ωk*=*ω*1 *jekke,tkej,t* (9)

{

*s.t* ***R****T****ωω***⩾0= 1 (10)

where *ek,t* is the error of the *k*-th method at time *t*, *pprek,t* is the prediction value of the *k*-th method at time *t*. ***ω***=(*ω*1, *ω*2,…, *ωk*)*T* is the weight vector, ***R***=(1,1,…,1)*T* is a column vector, and ***E***=(***E****ij*)*K*×*K* is the combination forecasting wind power error matrix.

Meta-heuristic techniques attempt to optimize the above objective function to obtain the best parameters of the prediction model. The key task of wind power optimization is to select alternative solutions for finding the best optimal solution(s) for predicting wind power output. This is a complex multidisciplinary task involving meteorological factors, geographic information, and artificial intelligence. Numerous meta-heuristic algorithms for wind power optimization have been developed over the past few decades. Fig. 9 shows that there are five main categories of meta-heuristic algorithms: physics-based algorithms, human-based algorithms, swarm-based algorithms, and hybrid algorithms.

## 5.2. Evolutionary-based algorithms

The biological evolutionary algorithm (EA) can be defined as a stochastic search method that simulates biological evolution processes or species’ social behavior, beyond the other species through learning, adaptation, and evolution. There are four main EA methods, including genetic algorithms (GA), evolutionary programming (EP), genetic programming (GP), and differential evolution (DE), which can be used to predict the wind power output.

|  |
| --- |
| **Fig. 10.** The principle and pseudo code of GSA. |

### 5.2.1. Genetic algorithm (GA)

The GA is an algorithm based on modeling the biological evolution process to solve the optimization and searching problems [44]. The GA contains several chromosomes, each of which represents a solution or a feasible solution to a specific problem.

The data decomposition technique can improve the accuracy of ANNs in wind power prediction. The fast ensemble empirical mode decomposition (FEEMD) technique is employed to decompose the time series of wind speed [135]. The GA, the thinking evolution algorithm, and a neural network were used to form a hybrid model (FEEMD-GA- MLP) for wind speed forecasting, which improved the prediction accuracy in comparison to the other benchmark models. In the other ANNs forecasting models, the wind speed data are decomposed into several sub-sequences by using EMD, the weights and thresholds of ANNs are tuned by the GA [136]. In recent years, machine learning approaches have been introduced to predict wind power output. Among them, the SVM/LSSVM is one of the most widely used. The GA was also used to ensure the generalization of the SVM, and the best parameters were loaded to the forecasting model [137]. Consequently, the kernel function parameters of an SVM were tuned by the GA, and a GA-SVM model was developed to predict the short-term wind power output [44]. Similarly, the GA was also used to optimize the weights of the combined model with an adaptive neuro-fuzzy inference system, SVM, and ELM [138].

### 5.2.2. Evolutionary programming (EP)

The EP is a finite state machine evolution model proposed by L. J. Fogel in the study of artificial intelligence in [139]. In this model, the state of the machine is compiled according to the distribution law. D. B. Fogel expanded the EP idea in [140], making it able to deal with the optimization problems of real conditions, and introduced a normal distribution mutation operator into the mutation operation. The EP has become a widely-used optimization algorithm and thus has been applied in many practical problems.

A neural network with hyperbolic tangential activation functions (HTU) mixed with the evolution programming was proposed [141], which can achieve high accuracy for predicting the short-term speed prediction. The HTU was used in the feed-forward neural network. Meanwhile, the hybrid evolution programming algorithm was used to select the best number of neurons in the network layers. For instance, two different algorithms, the EP and PSO, were discussed in solving the problem of hyper-parameters estimation in the SVM [142]. Results show that the EP-SVM was superior to the multi-layer perceptron in wind speed prediction of Spanish wind farms.

### 5.2.3. Genetic programming (GP)

The GP is an evolutionary computation method derived from Darwin’s theory of evolution [143]. Unlike genetic algorithms in which each individual represents a piece of chromosome coding, in the GP, the individual denotes computer programs, represented by a tree structure in the computer memory.

The hybrid model can compensate for the shortcomings of the individual forecasting model. An effective short-term wind forecasting method is proposed [144], which combined the ANN and GP. Compared with the existing combination models, this combination method avoids the defect of error accumulation in individual prediction models by introducing a semi-random neural network based on the Gaussian regression (GP) model. Results showed that the average RMSE for the five wind farms was 0.1175.

### 5.2.4. Differential evolution (DE)

The DE algorithm is an evolutionary computing technology. Similar to the other evolutionary algorithms, the DE is a stochastic model that simulates biological evolution, allowing individuals to be preserved through repeated iterations [145]. The DE reduces the complexity of genetic manipulation by using real coding, differential-based mutation operations, and one-to-one competitive survival strategies.

A combination model based on the adaptive neural fuzzy reasoning system (ANFIS), evolutionary PSO algorithm, WT, and mutual information, was established and used for predicting short-term wind power [146]. The implementation steps were as follows: The WT was used to decompose wind power into some sequences that values fluctuate slightly, and the newly obtained sequences were input to the ANFIS model for prediction, the EPSO adjusted its member function to enhance the ANFIS performance to obtain smaller errors.

## 5.3. Physics-based algorithms

The physics-based optimization algorithm is an inanimate algorithm of meta-heuristic algorithms, a type of algorithm developed according to the physical properties of certain physical behaviors or similar physical laws. The most popular algorithms are the gravitational search algorithm (GSA), simulated annealing (SA), sine cosine algorithm (SCA), and chaotic optimization algorithm (COA).

### 5.3.1. Gravitational search algorithm (GSA)

The principle of GSA is the law of gravity and newton’s second law [147], which finds the optimal solution by moving the particle position in the search space. The principle and pseudo code of GSA is shown in Fig. 10. The particles move in the space according to the gravities between them, when one of the particles moves to the optimal position, the optimal solution is found.

In [148], a hybrid model LSSVM-GSA is proposed for short-term wind power prediction, and the forecasting model’s parameters are tuned by GSA. Results showed that the hybrid LSSVM-GSA model can obtain a better forecasting result compare to BPNN and SVM. Further, an improved GSA model is used to analyze the uncertainty of wind power systems [148]. In [149], the application of seasonal SVR and chaotic gravity search algorithm in power prediction is used. Similarly, the SVR parameters were optimized by chaotic GSA [149]. In [150], three decomposition techniques: feature selection (FS), mutual information (MI) and wavelet transform (WT) were applied to decompose electricity- pricing data. An LSSVM model is optimized by chaos theory-based GSA to predict electricity price. A similar concept was also used in [151], a new hybrid modeling method that combines time series decomposition, feature selection and a basic prediction model in the synchronous optimization framework is proposed, and then, the ELM forecasting model is tuned by GSA for short-term wind speed forecasting.

### 5.3.2. Simulated annealing (SA)

SA algorithms can get an approximate solution to optimization problems based on the Monte Carlo idea [152]. In the random search in solution space, the better solution is accepted according to a certain probability. The algorithm includes initialization, perturbation, cooling schedule, and acceptance probability program, in which temperature initialization and cooling play a key role in achieving good results.

In terms of wind power prediction, some authors divided the prediction process into data preprocessing-prediction model-optimization model parameter process, and significantly reduced the error compared with other benchmark prediction models. The non-stationary wind speed data were processed by a singular spectrum, and weights and thresholds of BPNN were optimized by combining SA and PSO [153]. In [154], LSSVM is used as the prediction model and the kernel function parameters are tuned by a coupled SA, and then a rolling prediction strategy is adopted for wind speed prediction.

### 5.3.3. Sine cosine algorithm (SCA)

The idea of SCA is derived from the mathematical model of the sine cosine function [155], which is a new optimization algorithm. Inspired by the sine and cosine functions, the algorithm simulates outward or toward the optimal solution wave direction and creates multiple initial random candidate solutions. Besides, the algorithm integrates some random variables and adaptive variables to emphasize the exploration and use of search space in different optimization problems.

Apart from the meteorological factor, the parameters of the prediction model have a great influence on prediction performance. In [156], the SCA is developed to tune the parameters of the BiLSTM model, the autocorrelation function (ACF) and the partial autocorrelation function (PACF) are used to determine the input variables. To the same end, a multi-objective sine–cosine-algorithm (MOSCA) is developed to optimize the WNN neural network to achieve high accuracy and reduce volatility [78]. Experimental results show that the algorithm is stable and reliable.

### 5.3.4. Chaotic optimization algorithm (COA)

Chaos is a general nonlinear phenomenon, whose behavior is complex but has inherent regularity. A fundamental feature of chaotic systems is that small variations in parameters can lead to differences in future behavior, such as stable fixed points, periodic oscillations, and bifurcation.

In [157], the chaotic characteristics of wind speed and the Apriori algorithm are used for wind speed prediction. In the model, four environmental factors were considered. The implementation steps are as follows: 1) use the k-means method for clustering, 2) establish association rules by the Apriori algorithm, 3) the chaotic time wind speed series is forecasted, and 4) correct the predicted wind speed data. A prediction technique of chaotic data is proposed [158], which embedded a time series in state space using delay coordinates, and the information of past values are used to make a short-term prediction of the future behavior of the time series.

## 5.4. Human-based algorithms

The human-based algorithms derived from simulating certain behaviors in human behaviors, for example, teaching and learning algorithm (TLBO) inspired by the teacher and student learning activities, cultural algorithms (CA) inspired by human sociology, and harmony search (HS) inspired by the musicians to adjust the tone of the instrument in the band.

### 5.4.1. Teaching and learning algorithm (TLBO)

The TLBO is a new swarm intelligence algorithm proposed by Rao in [159]. It simulates the teaching process of teachers and the learning process of students and aims to improve students’ learning performance through “teaching” and “learning” between students.

The authors in [160] used Gaussian process regression to capture the randomness of wind energy, and a variant Gaussian process was used for time-series prediction, and then the prediction model was trained based on TLBO to improve learning efficiency. A hybrid model based on TLBO and DE was proposed [161]. DE is used in updating the previous best positions of an individual in TLBO, which avoided stagnation. The hybrid algorithm is used to predict three typical chaotic time series. In [162], an improved teaching–learning-based optimization algorithm with neighborhood searching (NSTLBO) is presented, then, the parameters of FFNNs are tuned. Compared with other evolution-based training techniques, this method makes use of both ANNs’ ability of nonlinear mapping between different complex data and TLBO’s strong ability in global searching and exploration [163].

### 5.4.2. Cultural algorithm (CA)

CA is an optimization algorithm of two-layer evolutionary mechanism [164]. CA acquires and integrates the knowledge of problem- solving through a belief space independent of the population space, which enables the evolution speed of the population to go beyond the evolution speed solely relying on biological genes. CA has good global optimization performance.

A novel RBFNN model that combines orthogonal least-squares (OLS) and CA is proposed for wind power prediction [165], and the hidden number of RBFNN is obtained by OLS and the CA is used to adjust the parameters in the network. Results showed a better forecasting result can be got compare to the individual forecasting model.

### 5.4.3. Harmony search (HS)

The harmony search (HS) is a global search algorithm, which has been applied in many optimization problems [166]. According to the principle, the musicians adjust the tones of each instrument in the band repeatedly according to their memory, and a perfect harmony state is finally reached.

Salcedo-Sanz et al [167] combined elements of coral reef optimization (CRO) algorithm with operators of HS method to form a hybrid CRO-HS optimization technique. The parameters of an ELM network are optimized by the hybrid technique for wind power prediction. Further, an HS-LSSVM forecasting model where the LSSVM is used to predict the output wind power, and the HS is used to select the regularized parameter gamma and the kernel function parameter sigma [168].

### 5.4.4. Group search optimization (GSO)

The inspiration of the group search optimization (GSO) method comes from animal search behavior [169]. Group members are divided into three categories: producers, scroungers and dispersed members.

|  |
| --- |
| **Fig. 11.** The principle and pseudo code of PSO. |

In each iteration, the members with the best positions are producers and they keep their positions unchanged, while the other members are randomly selected as scroungers or dispersed members. The scroungers move towards the producers by a certain distance and the dispersed members perform random walks. A nonlinear interval optimization (NIO) model is proposed [170], which took into account both the mean and deviation values of the scheduling targets. A Pareto solution is obtained using a multi-process group search algorithm (GSOMP), which displayed a trade-off relationship between the mean and deviation of the scheduling target. Then, the decision method, evidence reasoning (ER), is applied to determine the final scheduling solution.

## 5.5. Swarm-based algorithms

As an emerging evolutionary computing technique, the swarm intelligence algorithm is used to simulate the collective behavior of animals in social nature. Although these agents (insects or groups of individuals) are relatively simple and have limited capabilities, they can get the necessary tasks through collaborative interaction. For their survival, keep a distance from other agents. The swarm-based algorithm is detailed below.

### 5.5.1. Particle swarm optimization PSO

Particle swarm optimization (PSO) inspired by artificial life research results, by simulating the flock foraging behavior in the process of migration and cluster [171]. Fig. 11 shows the principle and pseudo code of PSO. All the organisms in nature have certain swarm behaviors, and one of the main research areas of artificial life is to explore the biological nature of group behavior, to build the group on the computer model. In PSO, the solution of each optimization is called “particle”. Each particle is given memory to remember the best location to search.

Data preprocessing and input parameter selection can improve prediction accuracy [172]. EEMD technique is used to decompose wind speed data into several independent IMFs [173], and the first IMF can be removed, and then a new dataset is aggregated. The WNN parameters were optimized by GA-APSO. The GA-APSO-WNN model is used for forecasting wind speed. Then, the weights and thresholds of BP were tuned by PSO. Besides, some combined prediction methods with MI, WT, evolutionary PSO, are successfully applied to predict the multi-time scale wind speed [174]. Similarly, the selection parameters can provide considerable contributions to the prediction performance [70], a combining forecasting model combined WT, NNs, and PSO to predict the output wind cluster power with several wind farms, the best forecasting model’s parameters can be got by PSO [175]. Furthermore, a chaotic PSO algorithm was applied to choose the suitable parameter combination for an SVR model for load forecasting [176].

### 5.5.2. Whale optimization algorithm(WOA)

Whale optimization algorithm (WOA) is proposed by Mirjalili in 2016 [47]. The WOA imitates the hunting behavior of humpback whales in which the “spiral bubble net” strategy is used, and captures food through contraction and surrounding, spiral position updating, and random hunting technique[177]. The algorithm has a simple structure, few adjustment parameters. It has been applied in areas of engineering and has a good prospect.

The machine learning (e.g., LSSVM) parameters such as kernel parameter, penalty factor (*c*) have a great effect on the complexity and reliability of the forecasting model. In [48], LSSVM is selected as the prediction model and WOA is used to optimize model parameters. The model is successfully applied for wind speed forecasting. Compared with individual prediction, this model has a wide application range and a good prediction effect. Moreover, reducing input redundancy can improve prediction accuracy. In [178], DWT- IR is adopted to select the optimal features, and the wavelet kernel function is used to replace the kernel function of LSSVM to improve the nonlinear mapping capability of LSSVM. The parameters of W-LSSVM were optimized by WOA in wind power prediction. In order to improve the accuracy of wind speed prediction and model stability, a new multi-objective whale optimization algorithm (MOWOA) [50] is developed to optimize the weights and thresholds of the Elman neural network.

### 5.5.3. Bat algorithm (BA)

Bat algorithm (BA), a heuristic search algorithm proposed by Yang in 2010, is an effective method to search for global optimal solutions [179]. This algorithm uses an iterative optimization technique, in which random solutions are initialized, and then the optimal solution is obtained through iteration. It generates local new solutions via random flight around the optimal solution, which strengthens the local search.

Data mining technique has received more and more attention in the area of wind speed prediction, but there is still a lack of research on multi-step prediction, which hinders further development in this area. A hybrid method based on optimal feature selection and ANNs for multi- time scale wind speed prediction is proposed [180]. The prediction model is optimized by the improved bat algorithm and cognitive strategy. The proposed hybrid model compensates for the shortcomings of neural networks in multi-step prediction and can be verified in different prediction time scales. A similar forecast framework can be seen in [181], in the combining model, the ELM is firstly used to predict the wind speed, and the BA, improved by the conjugate gradient method, optimizes GRNN parameters to get the final predicted results. Results showed that the hybrid prediction model has higher accuracy than the benchmark model.

### 5.5.4. Cuckoo search (CS)

Cuckoo search (CS) is a new heuristic algorithm [182], simulating the parasitic breeding behavior of cuckoo birds. The algorithm can be enhanced by so-called Levy flight, rather than a simple isotropic random flying.

Parameter optimization can make a model more reliable. In [68], an improved CS algorithm is proposed to obtain the optimal weights of WNN, and then, the hybrid model is used for short-term wind speed forecasting. and make. A similar prediction framework can be seen in [183], in this case, the prediction is closer to the actual. Moreover, solar radiation is predicted using an optimized hybrid model by CS [184]. Forecasting accuracy and forecasting stability are also considered in [185], a non-dominated sorting based multi-objective cuckoo search algorithm (NSMOCS) is used to tune the parameters of GRNN, then the best parameters are load to prediction model to forecast short-term load.

### 5.5.5. Grey wolf optimizer (GWO)

GWO is an algorithm proposed by Mirjalili in 2014 [186]. It originated from the simulation of the hierarchical mechanism and hunting behavior of grey Wolf and realized the optimization of search purposes through the process of wolves’ tracking, surrounding, hunting and attacking prey. GWO has the advantages of simple principle, easy implementation and few parameters to be adjusted.

In order to achieve satisfactory forecasting results, a novel hybrid modeling strategy is proposed [46], which combines EWT, GWO, RELM, and IEWT reconstruction. The hybrid method effectively improves prediction accuracy. In [187], the parameters of the intrinsic mode function (IMF) were optimized by GWO, and the LSTM was used as a predictor. While, the optimal value of fractional order can be found by GWO [74], and then the optimal order is loaded to the forecasting model to forecast the coal and natural gas consumption. The procedure is similar in [188], and the multi-objective GWO to optimize ELM parameters for wind speed forecasting. The case study compared single and traditional models, and the results showed that the model can obtain high accuracy and strong stability, simultaneously.

### 5.5.6. Artificial bee colony (ABC)

The artificial bee colony algorithm (ABC) is proposed in [189], which imitates bee behaviors. Its main feature is that it does not need to know the special information about the problem, but only needs to compare the advantages and disadvantages of the problem. Through the local optimization behavior of each worker bee, the global optimal value is finally highlighted in the population.

In [190], the weights of each group method of data handling (GMDH) models, were optimized by ABC, and then, the results were summed to generate a new forecast. A similar wind forecast framework was applied in [191], in this case using ABC to select the optimal kernel parameters of relevance vector machine (RVM), WD technique decomposes wind signals into different groups according to frequency ranges. Results showed that the combined model can improve the accuracy of RVM prediction.

### 5.5.7. Ant lion optimizer (ALO)

Ant lion optimizer (ALO) is first designed by Mirjalili in 2015 [192], named after the ant lion foraging behavior in nature. Similar to other intelligent algorithms, it is a meta-heuristic algorithm. This algorithm can explore the search space through random walks of ants around the ant lions, and learn from the selected ant lions to ensure the diversity of the population and the optimization performance of the algorithm. ALO has the advantages of a few adjustment parameters.

A new hybrid wind power forecasting method is proposed in [51], CEEMD technique is used to decompose the original wind speed series into several sub-layers. The Elman neural network model optimized by the multi-objective ant lion optimization algorithm (MOALO) was used as the prediction model. The experimental results indicated that the MAPE of the developed model utilizing 10-min, 30-min and 60-min interval data were 2.8220%, 5.0216%, and 7.7205%, respectively, which were much lower than those of the benchmark models.

### 5.5.8. Firefly algorithm (FA)

Firefly algorithm (FA) is proposed by Yang in [193]. It mimics the behavior of firefly flickering. The main purpose of a flickering firefly is to attract other fireflies. The algorithm assumes that a firefly, regardless of its sex, will be attracted to other fireflies according to their higher brightness. The brightness is associated with the target function.

A novel approach, the parameters of the MLP optimized by FFA, is proposed to predict wind speed [194]. And monthly data from 2004 to 2014 is used to train. Besides, an SVR/SVM and FA are combined to predict the output wind power [195,196]. The procedure is similar in [49], the FEEMD, VMD and BP models are combined for electricity price prediction, in which the weights and thresholds were optimized by the FA. Results showed that satisfactory forecasting results with multiple time scales can be obtained. In [197], a model combining a variety of seasonal patterns and improved FA for predicting power load is presented. The combining model’s weight coefficients are optimized by non-positive constraint combination theory.

## 5.6. Hybrid optimization algorithms

The purpose of biological meta-heuristic optimization algorithms is to quickly get the optimal value of the evaluation function. Different optimization algorithms have different global and local searching abilities, when the search capability is complementary, a hybrid optimization algorithm can be established by combining multiple individual optimization algorithms, which can be superior to individual algorithms.

*5.6.1. Particle swarm optimization with gravitational search algorithm*

# *(PSO* + *GSA)*

The PSO has a strong global capability, while GSA has a local searching ability. The hybrid algorithm, named PSOGSA [198], which combines the advantages of two algorithms, can outperform the PSO and GSA in calculating capability and convergence speed.

An SVM and Markov are combined for wind speed forecasting [151], and the C-C method and the phase space reconstruction method are used to determine the forecasting model’s input, then, the parameters of LSSVM were optimized by the PSOGSA. The authors in [99] used longitudinal data selection (LDS) and singular spectral analysis (SSA) techniques to decompose the original data, gravity search algorithm, and adaptive PSOGSA to optimize the nonlinear MLP structure. The LDS technology ensures that the input and output data have the same properties. The SSA technique extracts the trend and seasonality of power load values. Finally, the combined model is used to predict the output short-term power. In [93], the original wind speed sequence is decomposed into several IMFs using the EMD technique of time adaptive filter and the PSOGSA algorithm is used to tune the ORELM network structure’s weight and deviation, and the model with the optimal value is obtained for wind speed forecasting.

## 5.6.2. Particle swarm optimization with ant colony optimization (PSO + ACO)

|  |
| --- |
| **Fig. 12.** Multi-objective optimization algorithms for wind speed/power forecasting [204]. |

The ideal solution is to get the global optimal solution of the system, and the solution parameters are found by PSO and ACO. Therefore, a novel hybrid optimization method integrating PSO and ACO is developed, which has been proven to solve complex problems compared to gradient search methods [199]. The hybrid optimization method can achieve global optimization. In ACO, it only updates the pheromone path with the best solution and allows each ant to jump, and the PSO allows more particles to fly freely in space. Therefore, PSO provides a solution in each iteration, and ACO can get the expected solution. The hybrid technology of ACO and PSO has the advantages of both algorithms. the hybrid model based on PSO and ACO is proposed [200], and the target is to find the constant values in a way that we achieve the lowest possible error value between the observed and estimated values. In a similar way, a hybrid model based on ACO and PSO to predict the energy output is adopted [201]. The hybrid optimization prediction model can get satisfactory results and a faster convergence effect, compared to the BP model.

## 5.6.3. Flower pollination with gravitational search algorithm (FPA + GSA)

The principle of FPA comes from a simulation of self-pollination and cross-pollination. It is a new meta-heuristic intelligent algorithm that features simple programming and strong global searching ability. GSA locates the best solution by moving the particles of the population around and it has the superior local searching ability. The best solution is to avoid the problems like local minima trap and slow convergence. Due to the GSA can evaluate global optimum but is slow convergence, and the FPA is fast convergence but falls into local optimum. Therefore, a new hybrid optimization method integrating FPA and GSA is developed. In [202], a hybrid algorithm FPA-GSA is developed integrating the global searching ability of FPA and the local searching ability of GSA. Conventional back-propagation algorithm for training FFNN suffers from problems like slow convergence and local minima trap.

## 5.6.4. Whale optimization algorithm with simulated annealing (WOA + SA)

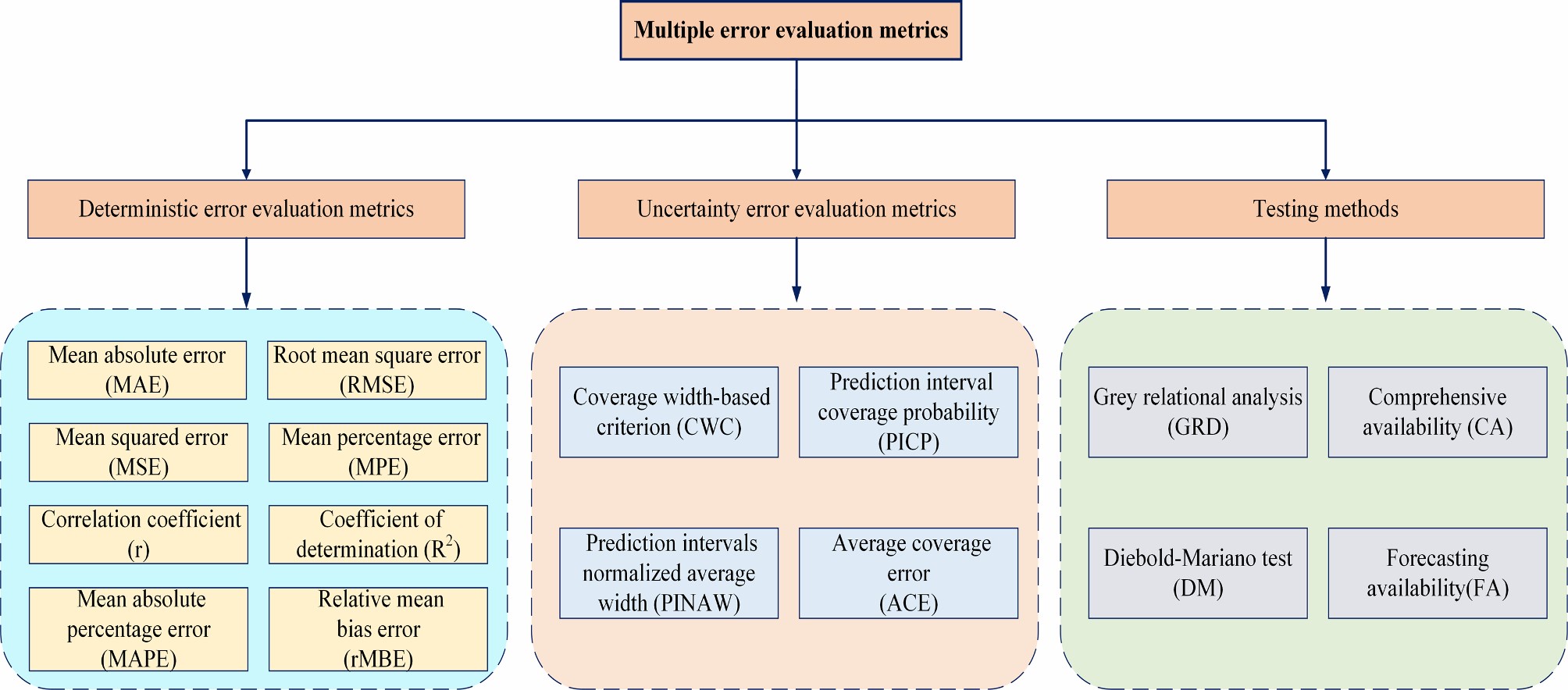
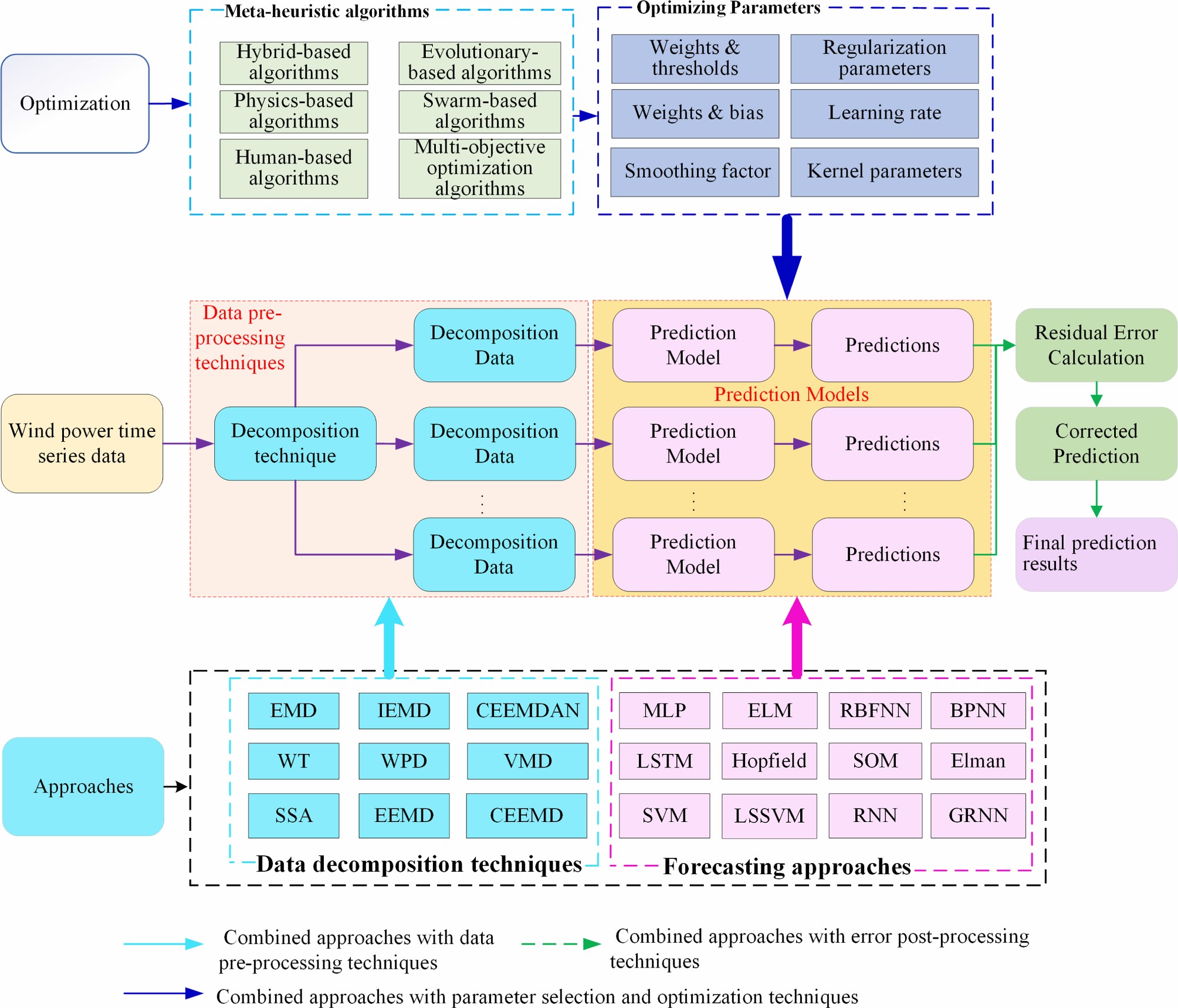
The advantage of the WOA is that it takes less computation time in simple planning, and can achieve the local optimum. While the SA can obtain the global optimum in large-scale problems, and not depend on the initial solution. In other words, a WOA can be hybridized with SA, and SA will be enhancing the exploitation in WOA. In [203], a hybrid model is developed combining WOA and SA. SA is embedded into WOA to improve its searching ability for optimal solutions in each iteration. The hybrid model is tested using benchmark data sets and compared with three well-known wrapper feature selection methods.

## 5.7. Multi-objective optimization algorithms

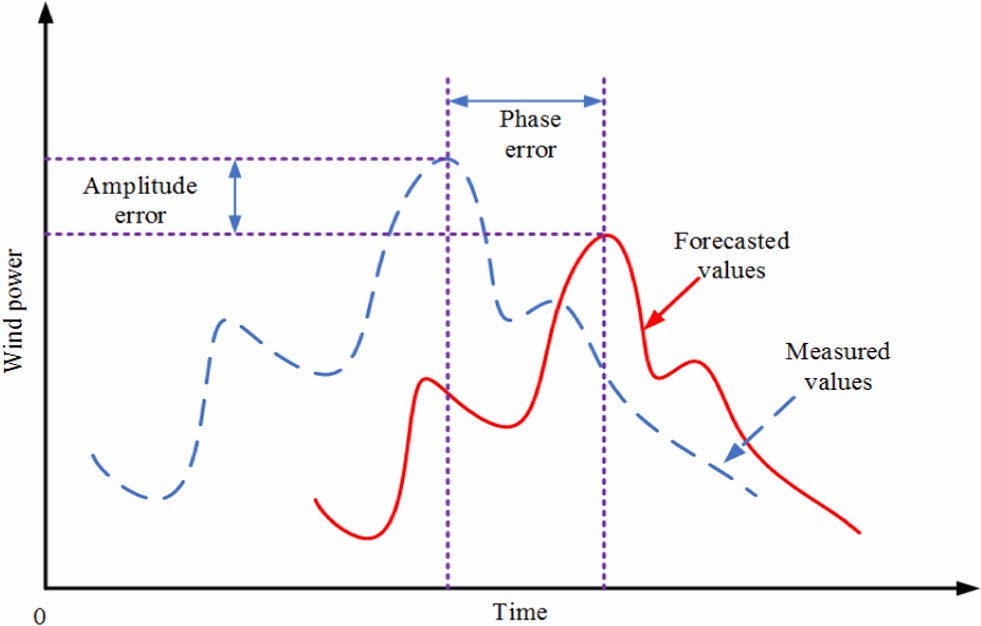
At present, multi-objective optimization algorithms have been progressively receiving more attention. These algorithms can simultaneously meet the requirements of prediction model accuracy and reliability. In this section, multi-objective optimization algorithms are categorized into five categories for wind power prediction, which are shown in Fig. 12. These categories include swarm intelligence-based algorithms, evolution-based algorithms, geography-based algorithms, hybrid algorithms, physics-based algorithms.

**Remark:**. *To summarize*, *there are several models and applications based on the meta*-*heuristic optimization algorithms*, *as shown in* Fig. 13. *These can be categorized into three prediction modes*: 1) *combined approaches with data pre*-*processing techniques*, 2) *combined approaches with parameter selection and optimization techniques*, 3) *combined approaches with error processing techniques*. *In these three modes*, *the meta*-*heuristic algorithm optimization plays a significant role in improving the wind power prediction accuracy*.

(1) Data preprocessing stage

**Fig. 13.** The framework of the mainstream wind power forecasting method.

**Fig. 14.** Multiple error evaluation metrics for wind power forecasting error.



**Fig. 15.** Two types of wind power forecasting error.

### **Table 3**

Performance evaluation metrics.

Index Definition

Equation

*MAE* Mean absolute error *MAE* ⃒⃒⃒*pacti,t* − *pprei,t* ⃒⃒⃒

*RMSE* Root mean square error *RMSE* = √̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅1∑*N*= (*pacti,t* − *pprei,t* )2̅

*N i*⃒1 ⃒ *MAPE* Mean absolute percentage error 1∑*N*= ⃒⃒*pacti,t* − *pprei,t* ⃒⃒ × 100%

|  |  |
| --- | --- |
| *r* Correlation coefficient | *N* 1(*pi*=*i,*1*t* − *p~~p~~actiacti,,t t* )(*pipre,t* − *pprei,t* ) *act*  *r* = ~~√̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅̅~~ |

*MSE* Mean squared error *MAPEMSE* ==1∑*N i p*1*acti,* − *Wptrueprei,t*

*N t*

∑*N act* 2 *pre pre* 2

*R*2 Coefficient of determination *i*

*R act*~~∑~~*, N*= *p*(*p, acti,t* − *pacti,t* )~~2~~

Mean percentage error *pi t i* 1 *pre*

*MPE*

*MPE*

=

*p*

*i*

*,*

*t*

*rMBE*

*rMBE*

=

∑

*N*

*i*

=

1

*p*

*act*

*i*

*,*

*t*

−

*p*

*pre*

*i*

*,*

*t*

*act*

*i*

*,*

*t*

−*act i t* × 100%

Relative mean bias error 1

*N p*

**\*Note:** The actual and predicted values are represented by *pacti,t* and*pprei,t* , respectively, *~~p~~acti,t* and*pprei,t* stand for the average of the actual and predicted values, respectively, *N* is the length of the data set, and *Wtrue* is the capacity of wind farms.

In this stage, data decomposition techniques are used to convert the nonlinear wind power time series into stable subseries following particular rules. Subsequently, the redundant data can be filtered out. As a result, stable subsets and highly informative training data can be obtained, which improves the data quality used for forecasting.

1. Parameter selection and optimization stage

In this stage, using meta-heuristic algorithms for parameter selection of the forecasting base layer can contribute considerably to improving the prediction performance in the training stage. Selecting the best parameters from a large number of candidate parameters can reduce the simulation time and provide satisfactory prediction performance. A few popular swarm-based algorithms, e.g., GA and PSO have been successfully applied to predict the output wind power.

1. Error post-processing stage

In this stage, the predicted value is mostly either overestimated or underestimated. The prediction accuracy can be further improved by combining the residual error value with the prediction sequence.

## **6. Multiple error evaluation metrics**

Three evaluation metrics are commonly used to quantitatively evaluate the performance of wind prediction methods. These metrics include deterministic error evaluation metrics, uncertainty forecasting metrics and testing methods, which are shown in Fig. 14. In the next sub- section, we describe the deterministic error evaluation metrics.

### 6.1. Deterministic error evaluation metrics

In this section, we provide details about a scientific and reasonable evaluation metric. Fig. 15 shows the two typical evaluation metrics, longitudinal (amplitude) error and lateral (phase) error, which can be used for wind power prediction [205]. The former metric denotes the difference between the real and predicted wind power values in the vertical direction at a certain time period. On the other hand, the latter metric represents the difference between the real and predicted wind power values in the horizontal direction at a certain time period, which can generally be described by the lead or lag of the predicted sequence peak. The longitudinal error can be improved using systematic error correction or error time series statistics.

Various deterministic error evaluation metrics are used for wind power/speed prediction [206]. These metrics are summarized in Table 3 and include: mean absolute error (*MAE*), root mean square error (*RMSE*), mean absolute percentage error (*MAPE*), mean squared error (*MSE*), correlation coefficient (*r*), coefficient of determination (*R*2), mean percentage error (*MPE*) and relative mean bias error (*rMBE*).

### 6.2. Uncertainty error evaluation metrics

Uncertainty error evaluation metrics include prediction interval coverage probability (*PICP*), coverage width-based criterion (*CWC*), average coverage error (*ACE*), and prediction intervals normalized average width (*PINAW*). Their expressions are given as follows [207]:

⎧

⎨ *PINAW*[1 + I(*PICP*{)e− *ρ*(*PICP*− *υ*)]*,ρ* ∈ [50*,*100]

⎩ I(*PICP*) = 1*, if PICP < υ* (11)

0*, otherwise*

|  |  |  |  |
| --- | --- | --- | --- |
| 1 *N*  *PICP* = 1 -  *ci,ci* =  *N i*=1 | 1*,*  0*,* | *pi* ∈ [*Li,Ui*] *pprei* ∈∕[*Li,Ui*] | (12) |
| *N*  1  (  *maxp*  −  *minp*  )  ∑  *n*  *i*  =  1  *U*  *i*  − | *Li* |  | (13) |

{ ∑ *pre*

where *υ* is the nominal confidence level, *ρ* is penalize coefficient of the invalid prediction intervals, *Li* land *Ui* are the lower and upper bounds of *i*-th prediction intervals, *N* is the total number of prediction samples, *maxp* and *minp* are the maximum and minimum of the autual wind power, respectively. *6.3. Testing methods*

Although the performance evaluation metric is highly significant for forecasting performance evaluation, testing is needed to confirm the effectiveness of a forecasting method from a statistical perspective. Different testing methods include grey relation degree (*GRD*), comprehensive availability (*CA*), diebold-mariano (*DM*) test and forecasting availability (*FA*). Each of these methods is described in the following.

#### *6.3.1. Grey relational analysis (GRD)*

The GRD is an index that can be used to describe the correlation level of different forecasting model results with observed wind power values. Denoting the actual and forecasted wind power time sequences by

{*pact*(*i*)}*ni*=1 and {*pact*(*j*)}*mj*=1, respectively. The GRD between *pact*0 and *pprej* (*k*), which can be expressed as follows [99]:

|  |
| --- |
| **Fig. 16.** Development of meta-heuristic algorithms. |

*ri* *n* *i*(*k*) (14) where the relational coefficient of *pacti* (*k*) and *pprej* (*k*) in point *k* is given as ⃒ ⃒ ⃒ ⃒ *minmin*⃒*pacti* (*k*) − *pjpre*(*k*)⃒ + *ρmaxmax*⃒*pacti* (*k*) − *pprej* (*k*)⃒ *ξj*(*k*) = *i* ⃒⃒*pjacti* (*k*) − *pjpre*(*k*)⃒⃒ + *ρmaxmaxi* ⃒⃒*pjacti* (*k*) − *pjpre* (*k*)⃒⃒ (15)

*i j*

where *ρ* is the distinguishing coefficient, its value is in the range *ρ* ∈ [0*,* 1]. If *r*0 *< ri*, the predicted and actual values are very similar, which indicates that the forecasting models have good predictive performance.

#### *6.3.2. Forecasting availability (FA)*

Forecasting availability is a metric measured by the square of the prediction error and the MSE of the prediction accuracy. The skewness and kurtosis of the *FA* distribution should be considered in the field of wind power forecasting [208]. The *FA* of the *i-*th forecasting method at time *t* is given by *Ait* = 1− |*εit*|, where the relative forecasting error of the *i-*th forecasting method at time *t*, which can be calculated as follows:

|  |  |  |  |
| --- | --- | --- | --- |
| ⎧  ⎪⎪⎨ (  *εi,t* = *pacti,t* −  ⎪⎪⎩ | (  − 1 *pacti,t* −  )/  *pprei,t pacti,t* −  (  1 *pacti,t* − | )/  *pprei,t pacti,t >* 1  ( )/  1⩽ *pacti,t* − *pprei,t pacti,t >* 1  )/  *pprei,t pacti,t* ⩾ − 1 | (16) |

#### *6.3.3. Comprehensive availability (CA)*

The prediction model comprehensive availability (*CA*) based on *GRD* and *FA* considers the trend and accuracy of the prediction sequence and fully reflects the prediction model performance [209]. The *CA* statistical evaluation of the prediction model is expressed as follows:

*CAi* = *ψri* +(1− *ψ*)*Hi* (17) where *ψ* is the control parameter, *ψ* ∈ [0*,*1], *ri* is the GRD of the *i-*th prediction model, and *Hi* is the FA of the *i-*th prediction model. The higher the value of *ψ*, the higher the emphasis on GRD evaluation.

#### *6.3.4. Diebold-Mariano (DM) test*

The *DM* test focuses on prediction accuracy [209,210], which is used to compare the performance of the developed prediction model with those of other benchmark prediction models. Denoting the actual wind { } power values as *pacti* ;*i* = 1*,*2*,*3*...,t* + *k* , and the wind power values { } predicted by two methods as *pprei ,*1;*i* = 1*,*2*,*3*...,t* + *k* and

{ }

*pprei ,*2;*n* = 1*,*2*,*3*...,t* + *k* , respectively, the wind power forecasting

errors are given by

|  |  |  |
| --- | --- | --- |
| *ε*(*t*+1)*h* = *pactt*+*h* − | *ppret*+*h*(1) | (18) |
| *ε*(*t*+2)*h* = *pactt*+*h* − | *ppret*+*h*(2) | (19) |

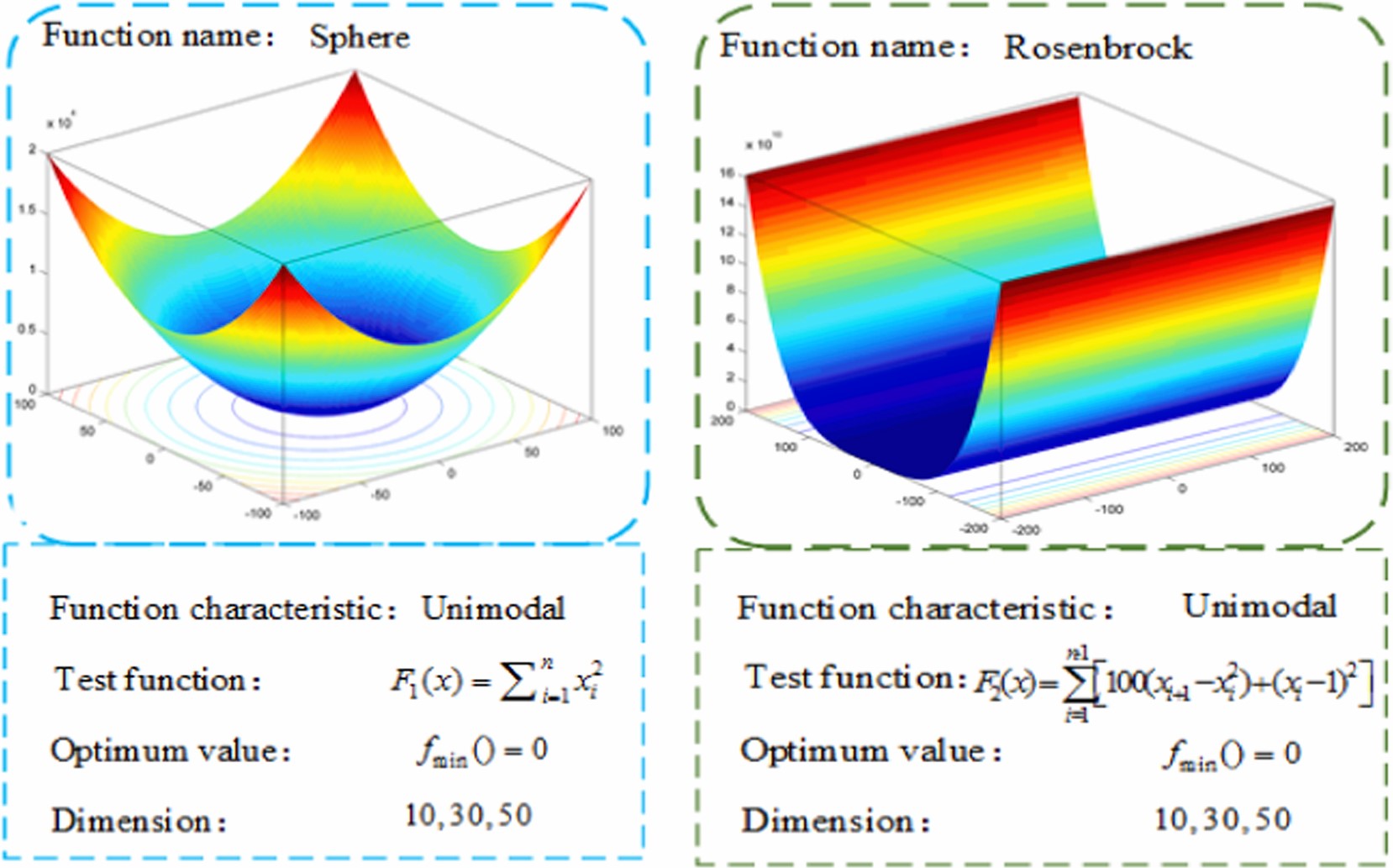
where *εt*(+1)*h* and *εt*(+2)*h* represent the forecasting errors of the two models being compared. The accuracy of each forecasting model is measured by a suitable loss function given by*L*(*εt*(+*i*)*h*) *i* = 1*,*2*.*, which is then used to calculate the DM statistics as follows:

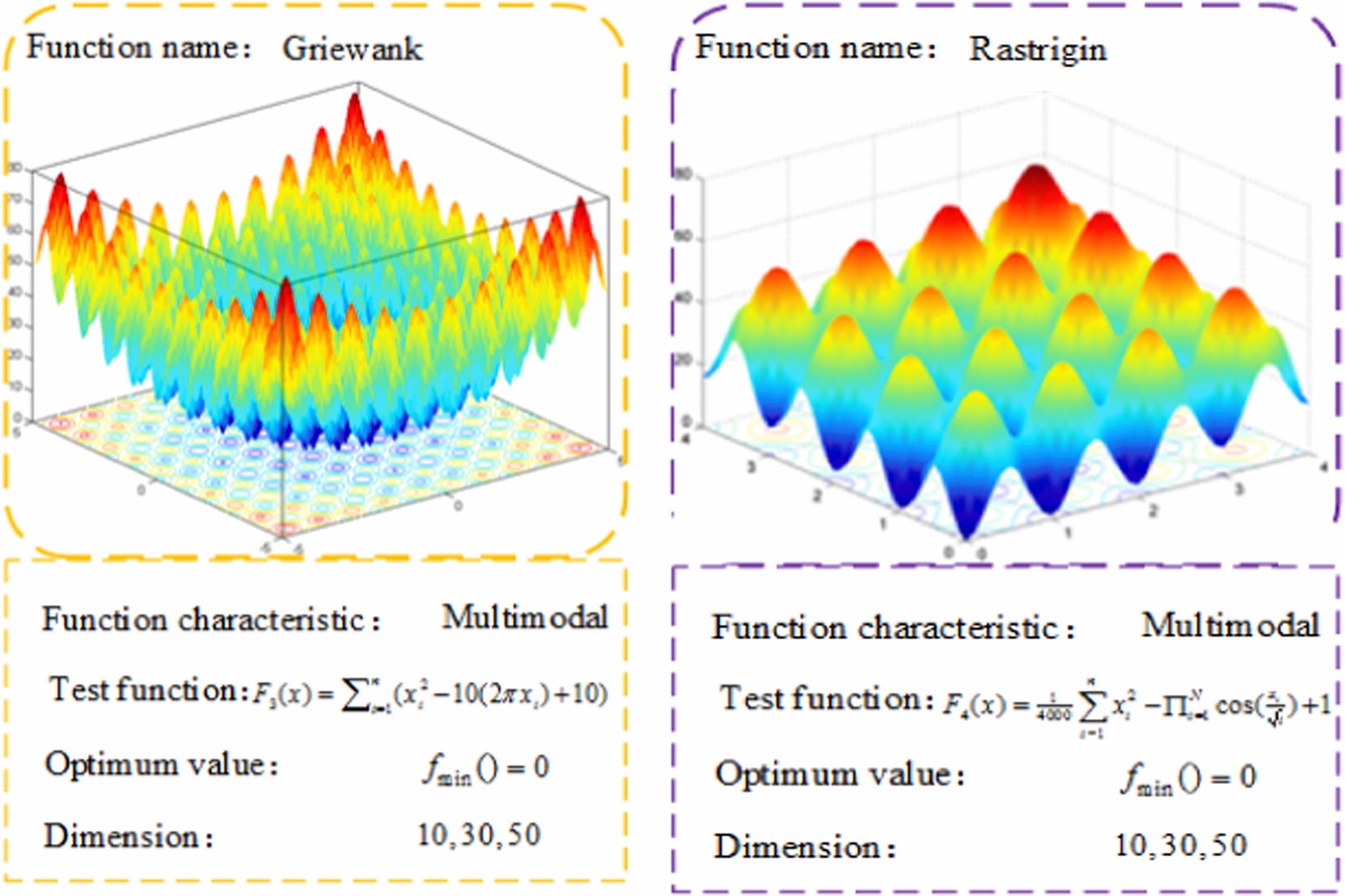
∑*k* ( ( (1) ) ( (2) ))/ *h*=1 *L εt*+*h* − *L εt*+*h k* 2

*DM* *s* (20)

where *s*2is a variance estimator. The null hypothesis is *H*0 : *E*(*dh*) = 0 ∀*n*and the alternative hypothesis is *H*1 : *E*(*dh*) =∕ 0. The value of DM is compared with *Zα/*2, where *α* is the level of significance. If DM is within the interval [− *Zα/*2*, Zα/*2], the null hypothesis *H*0. will be rejected; otherwise, it will be accepted.

(a) Sphere function (b) Rosenbrock function

**Fig. 17.** Two typical unimodal benchmark functions.

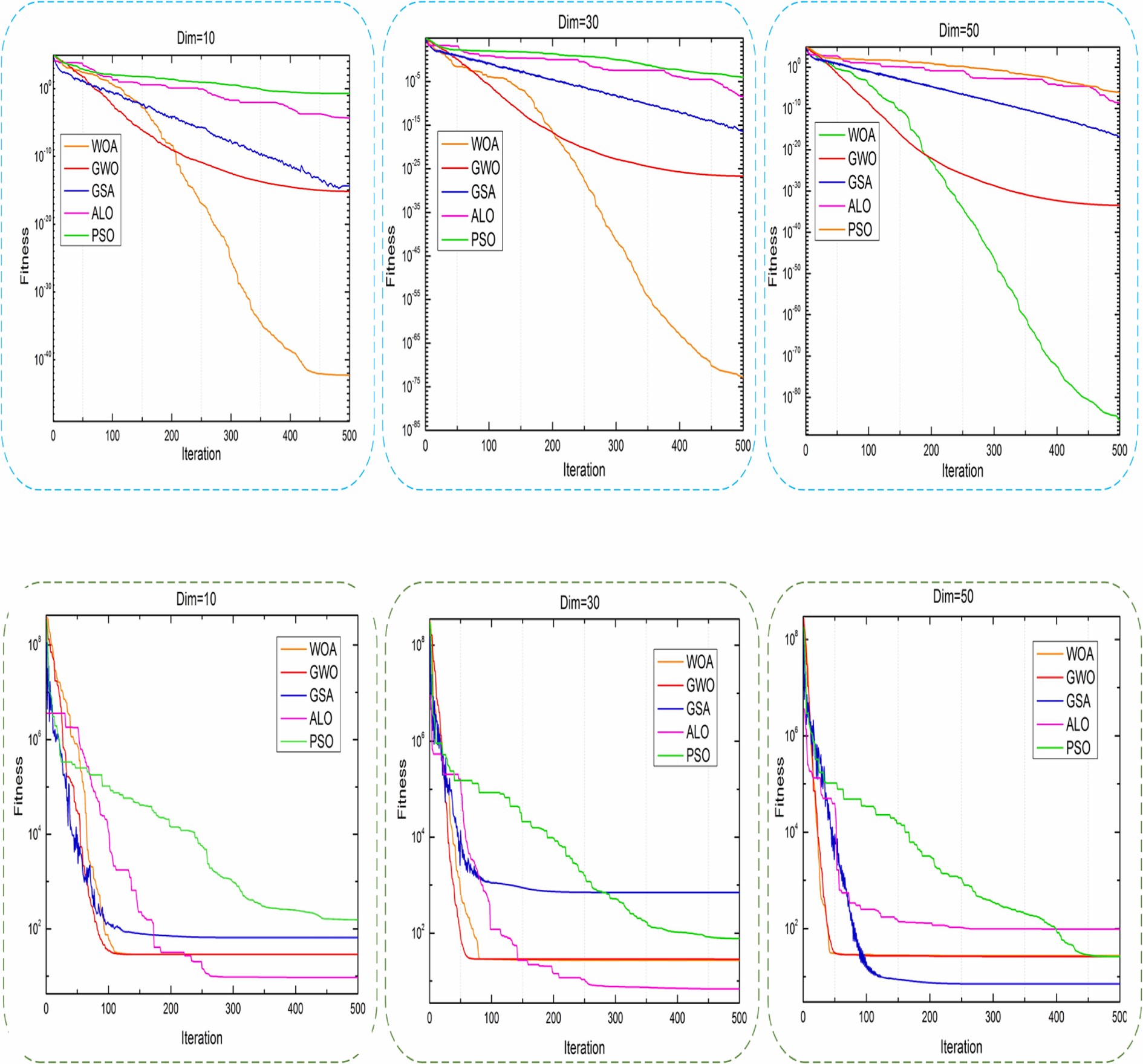


(a) Griewank function (b) Rastrigin function

**Fig. 18.** Two typical unimodal benchmark functions.

1. **Comparison of performance of meta-heuristic algorithms** the 1970s. The first algorithms of this type are evolutionary-based algorithms. Subsequently, with the development of other disciplines,

*7.1. Detailed description of the development process* other algorithms such as physics-based algorithms, human-based algorithms, swarm-based algorithms, and hybrid algorithms have been The development of meta-heuristic algorithms can be traced back to rapidly developed. Fig. 16 shows the development of the five types of



(

a) The results of Sphere function with

different dim

(b) The results of Rosenbrock function with different dim

**Fig. 19.** Unimodal function test results of WOA, GWO, GSA, ALO, and PSO.

meta-heuristic algorithms. Among all these algorithms, the GA and PSO are the most popular due to their powerful prediction ability. These algorithms are often used as the benchmark for comparing newly developed algorithms. The literature contains a large number of publications related to swarm-based algorithms.

### 7.2. Comparison of convergence performance of meta-heuristic algorithms

#### *7.2.1. Evaluation for single-objective optimization*

This study employs two typical unimodal benchmark functions, e.g., Sphere and Rosenbrock, and two typical multimodal benchmark functions, e.g., Griewank and Rastrigin, to measure and validate the performance of the selection algorithm. These functions are shown in Fig. 17-18. Furthermore, the classic optimization algorithms such as the PSO, and relatively new optimization algorithms, such as WOA, GSA, ALO and GWO, are employed to reduce the probability of falling into local optima.

Figs. 19-20 show the convergence performance of different meta- heuristic algorithms. It can be concluded that the newly developed algorithms, including the WOA and GWO, perform better than the classic PSO. The WOA shows the best performance for the Sphere benchmark function.

#### *7.2.2. Evaluation for multi-objective optimization*

Standard test functions are used to evaluate the convergence, accuracy and robustness of different multi-objective optimization algorithms. The reason for this choice is the difficulty in the theoretical evaluation of the performance of different multi-objective optimization algorithms. In the field of wind speed/power prediction, ZDT1, ZDT2, ZDT3, ZDT4, ZDT1 with linear PF, Kursawe and Viennet3 are commonly used standard test functions. These test functions are chosen due to their different shapes and sizes, which can help in the performance evaluation of the performance of the multi-objective optimization algorithms from multiple angles. Fig. 21 shows the Pareto frontiers of test functions.

|  |
| --- |
| (  a) The results of  Griewank function  with different dim  (b) The results of Rastrigin function with different dim  **Fig. 20.** Multimodal function test results of WOA, GWO, GSA, ALO, and PSO. |

Different researchers have recently used standard test functions to evaluate the performance of multi-objective optimization algorithms for wind speed prediction. The Pareto front of the test function was analyzed, and the proposed multi-objective optimization algorithms were compared with benchmark algorithms to evaluate the superiority of the algorithms. Table 4 compares the performance of different multi- objective algorithms. It can be observed that the multi-objective optimization algorithms proposed in the literature can better approximate the real Pareto front compared with the benchmark algorithms, which also demonstrates the superiority of the proposed algorithm.

### 7.3. Advantages and disadvantages of different algorithms

The advantages and disadvantages of different meta-heuristic algorithms including the swarm-based algorithm, evolutionary-based algorithm, physics-based algorithm, and human-based algorithm are listed in Tables 5–8. Table 5 shows the advantages and disadvantages of evolutionary-based algorithms. The GA is widely used for wind power prediction due to its significant advantages and ease of implementation via the MATLAB optimization toolbox. Specifically, the final solution given by the GA does not depend on the initial solution. Table 6 displays the advantages and disadvantages of the physics-based algorithm. Due to the robustness of GSA and SA against initialization, simplicity and ease of implementation, and capability to obtain global hearing in large- scale problems, these two algorithms have become the most commonly used physics-based algorithm. Table 7 shows that the HS is more efficient than the TLOA, TS and GSO out of all the human-based algorithms.

Table 8 shows the advantages and disadvantages of swarm-based algorithms. The PSO is widely applied to wind power prediction due to its advantages such as its relative maturity and ease of implementation. It can also be used as a benchmark optimization model. In recent years, new algorithms, e.g., CS, WOA, ALO, etc. have been attracting more and more attention from researchers due to their good convergence accuracy, robust performance and global optimization capabilities.

Table 9 shows the advantages and disadvantages of the meta- heuristic algorithms. The pros and cons of each algorithm are summarized, which include optimization performance in simple problems, tuning complexity, independence with respect to the initial solution, computational complexity and popularity. The individual algorithms, e.

g., GA, ES, GSA, etc. show excellent performance in simple problems. On the contrary, the hybrid and multi-objective optimization algorithms show worse performance in simple problems, but perform better in complex optimization problems. Moreover, the hybrid and multi- objective optimization algorithms have high tuning complexity and computational complexity, which indicates that high-performance computers are required to implement these algorithms.

This paper innovatively summarizes the hierarchical forecasting framework of “auxiliary layer-forecasting base layer layer-key core

|  |
| --- |
| **Fig. 21.** Obtained Pareto optimal solutions by different multi-objective optimization algorithms [50]. |

#### Table 4

The performance comparison of different multi-objective algorithms.

IGD (%)

Ref

Algorithms

Benchmark

Improvement

Test functions

Metrics (Average)

SP (%)

–

[50] MOWOA MOALO ZDT1 81.8

ZDT2 89.4 –

ZDT3 1.4 –

[78] MOSCA MOPSO ZDT1 68.8 – ZDT2 68.8 –

ZDT3 − 209.8 –

1. MOSBO MOALO ZDT1 78.3 – ZDT2 81.6 –

ZDT3 1.2 –

1. MOGWO MOALO ZDT1 90.0 – ZDT2 96.5 –

ZDT3 96.4 –

1. MSSA NSGA-II ZDT1 7.8 99.3 ZDT2 51.2 87.1

ZDT3 10.9 75.2

[53] MOICA MOPSO ZDT1 16.1 86.7 ZDT2 30.8 84.2

ZDT3 31.9 87.6

ZDT4 29.6 90.5

[214] IMOWCA MOWCA ZDT1 58.5 (GD) − 7.6 ZDT3 61.7 (GD) − 20.7 Kursawe 31.0 (GD) 29.0 Viennet3 55.8 (GD) − 3.4

Note: *E*1 is the evaluation metric of the references algorithm, *E*2 of the benchmark algorithms. Improvement percentage = (*E*2 − *E*1)/*E*2 × 100%.

layer”, and the methods in each layer have been proven to be effective. The benefits and limitations of the proposed framework for wind power forecasting are summarized in Table 10, which can be summarized as follows: 1) The linear and nonlinear characteristics of power time series can be characterized; 2) Both the accuracy and stability of the

#### Table 5

Summary of the evolutionary-based algorithm.

|  |  |  |  |
| --- | --- | --- | --- |
| **Optimization methods** | **Refs** | **Advantages** | **Disadvantages** |
| Evolutionary  Programming  (EP) | [215] | The function is simpler;  Easy to implement; Suitable for various solutions. | There is no guarantee that the best solution will be found in a limited time. |
| Genetic  Programming  (GP) | [144] | Suitable  convergence speed; Easy to use. | Search operator space is too large;  Subprogram reusability. |
| Genetic Algorithm (GA) | [43] | Not depend on the initial solution. | Convergence speed is slower  than other stochastic algorithms;  Not reliable for large-scale networks. |
| Differential Evolution (DE) | [216] | Cannot be easily trapped in local minima; Few numbers of control parameters; Easy to use. | Parameter tuning mostly by trial-and-error;  May drop in local best. |

forecasting can be improved; 3) Compared with an individual forecasting model, the forecasting performance of the proposed framework is better. The limitations of the proposed framework are summarized as follows: 1) Serious multicollinearity will occur when the forecasting results of an individual model are similar, which will lead to the decline of the forecasting model accuracy; 2) Decomposition technique is adopted to model the optimal subsequence, and the results depend on experience without objective basis.

Tables 11-12 show a brief description of recent studies on wind speed/power forecasting based on the meta-heuristic algorithms, including prediction horizon, prediction model, error evaluation, CPU time, benefits and limitations, main contributions. Particularly, Table 11

#### Table 6

Summary of the physics-based algorithm.

|  |  |  |  |
| --- | --- | --- | --- |
| **Optimization methods** | **Refs** | **Advantages** | **Disadvantages** |
| Gravitational  Search Algorithm  (GSA) | [147] | Robust against initialization;  Ability to explore local solutions;  Simple and easy to implement. | The local search ability of GSA is weak. |
| Simulated Annealing (SA) | [152] | Not depend on the initial solution; Obtain the global optimum in large-scale problems. | Need more computation time; Need to set the controller parameters appropriately. |
| Intelligent Water  Drops algorithm  (IWD) | [217] | Get a better solution; Can effectively jump out of the local optimal solution;  Approximate optimal solution. | Update object is single; High probability of global search capability. |
| Chaotic  Optimization  Algorithm (COA) | [218] | Avoid getting into the local best and get the global optimum. | Aperiodicity and local instability;  Blind repeat search within the search space. |
| **Table 7**  Summary of Human-ba | sed algo | rithm. |  |
| **Optimization methods** | **Refs** | **Advantages** | **Disadvantages** |
| Teaching and learning algorithm (TALA) | [159] | Balanced global search ability and convergence rate; A good capability for global and local searching. | Need more support for exploration features; Easy to be trapped locally and has a poor global search ability; Random search. |
| Tabu search (TS) | [218] | Fast convergences; Easy to tune the controller parameters; Suitable for simple and uncritical problems. | Need an appropriate initial solution. |
| Harmony Search (HS) | [166] | Simple in concept;  Easy to be implemented; Suitable convergence speed;  Easy to understand. | More computational burden and complexity; Dependent on harmony memory considering the rate. |
| Group search optimization (GSO) | [169] | Faster convergence; Better global search ability and robustness. | Easy to fall into local optimum  “premature” phenomenon. |

#### Table 8

Summary of the swarm-based algorithm.

|  |  |  |  |
| --- | --- | --- | --- |
| **Optimization methods** | **Refs** | **Advantages** | **Disadvantages** |
| Ant Colony  Optimization  (ACO) | [219] | Need less computation time insimple planning; | Can’t achieve the global optimum; Tuning the controller parameters is a challenge. |
| Whale optimization  algorithm  (WOA) | [47,48] | Easy to solve every optimization problem; Easy to solve multidimensional and multidimensional solution problems. | The variable problem cannot be solved;  No guarantee that a global optimum will be found. |
| Particle Swarm  Optimization  (PSO) | [173,220] | Fewer controller parameters than IPSO;  Easy to understand; Very effective for global search optimization. | The ability to  search locally is weak; Tuning the controller parameters is a challenge. |
| Fruit Fly  Optimization  Algorithm (FFOA) | [221,222] | Less parameter setting and faster convergence. | Not high global optimization accuracy;  Slow convergence speed;  Easy to fall into local optimum. |
| Bat Algorithm (BA) | [180,181] | Random optimization; New features of echolocation. | Trapped into a local optimal solution; Slow solution. |
| Cuckoo Search (CS) | [68] | Strong global search ability; Less choice of parameters;  Search path. | Slow convergence in late evolution; Defects that are easy to fall into local optimal solutions. |
| Grey Wolf  Optimizer (GWO) | [223,224] | Simple structure;  Clear concept and easy to implement; Good global performance. | Easy to fall into local optimum poor stability; Accuracy is not high. |
| Artificial bee colony (ABC) | [225] | Robust against initialization;  Ability to explore local solutions;  Easy to implement and simple. | Poor exploitation characteristics; May get stuck in local optimum. |
| Ant Lion  Optimizer (ALO) | [226,227] | Good convergence accuracy, robust performance and global optimization capabilities. | Poor iteration performance;  Easy to fall into the extreme value. |
| Firefly Algorithm (FA) | [193,194] | Easy to implement and simple;  A good at exploration. | Traps into local optima; Single goal optimization. |

summarizes the results of the optimal parameters of the ANNs using the meta-heuristic algorithm. It can be seen that the ANN-based method has a wide range of applications and is mainly applied to ultra-short-term prediction. Table 12 summarizes the results of the SVM and LSSVM using the meta-heuristic algorithm. It can be seen that the prediction results of SVM/LSSVM are highly dependent on the selected parameters of kernel function and are mainly used for ultra-short-term prediction. Therefore, to obtain a stable and reliable prediction result based on ANNS and SVM/LSSVM, it is necessary to combine key parameter optimization, data preprocessing method, and optimization algorithm.

### 7.4. Performance comparison of multi-objective optimization

The combined models can make use of the diversity of wind power data and obtain the final prediction results from different individual prediction models using linear or non-linear weights. As shown in Fig. 22, a multi-objective optimization algorithm is used to find the optimal weights for weighing these individual prediction results and then adding them. Table 13 summarizes the optimization results with the variable weight prediction model. The table shows that the accuracy and stability of the combined forecasting model based on multi- objective optimization are better than those of other models.

Moreover, a multi-objective algorithm can be used to optimize the optimal parameters of the prediction model to obtain high-precision and highly reliable prediction results. Fig. 23 shows the process of optimizing the key parameters in the prediction model. The optimization objectives include the weights and thresholds of the neural network, and the penalty coefficient and kernel parameter of SVM and LSSVM. In the optimization process, the wind power data are usually divided into three sets: training, validation, and testing sets. The parameters with the smallest squared error are selected as the optimal parameters and used

#### Table 9

The comparison of characteristics of different algorithms.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Algorithms** | **Optimization capability in simple problem** | **Optimization capability in complex problem** | **Tuning complexity** | **Independency on initial solution** | **Computation cost** | **Popularity** |
| GA  ES  ACO  DE  GSA  SCA  TS  HS  SA  PSO  BA  CS  FA  GWO  ABC  PSO + GSA  PSO + ACO  FPA + GSA  WOA + SA  MODA  MOALO  MOBA  MOMFO  MOGWO  MOGOA  MOGWA  MOPSO  MOSBO  MSSA  MOWOA  MOFEPSO  NSGA-II  NSGA-III  MODE  ISMODA  MOBSFPA  MOCWCA  IMOWCA  MOICA  MOMVO  MOSCA  ISMODA |  |  |  |  |  |  |

Notes: Elcellent: Good: Poor: .



#### Table 10

Summary of the proposed framework of wind power forecasting.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Auxiliary layer (Data Decomposition) | Forecasting base layer (Prediction Model) | Core optimization layer (Meta- heuristic algorithms) | Benefits | Limitations |
| EMD and improved version | * Artificial neural networks * Machine learning | Evolutionary-based algorithms | Simple and quite efficient; Easy to implement. | Need large computational time;  Slow convergence speed. |
| VMD |  | Physics-based algorithms | Get the global optimum;  Get reliable and accurate results. | Need more computation time; Not often used. |
|  |  | Human-based algorithms | Higher variability;  Simple and quite efficient. | More computational burden and complexity;  Not often used;  Poor global search ability. |
| WT |  | Swarm-based algorithms | Very popular;  Good convergence accuracy, robust performance and global optimization capabilities. | Cannot achieve the desired performance;  The result is affected by the initial parameters. |
| WPD |  | Hybrid optimization algorithms | Have strong global searching ability; Obtain the best global solution. | Need large computational time;  Complex structure;  Hard to implement. |
| SSA |  | Multi-objective optimization algorithms | Easy to implement and strong robustness;  Preserve the best non-dominated solutions; Both prediction accuracy and stability can be improved, simultaneously. | Need more computation time for iterations; Sensitivity to random initialvalues. |

with the prediction model to obtain the final prediction result on the test data.

A multi-objective ant lion optimization (MOALO) model was used in [242] to search the optimal parameters of the LSSVM. Besides, the parameters of VMD and multi-kernel robust ridge (MKRR) were optimized using the multi-objective chaotic water cycle algorithm (MOCWCA) in [52].

The results showed the effectiveness and superiority of the proposed method compared to other individual models. In [134], the weights and thresholds of the Elman neural network were optimized using the MOWOA. Similarly, the Elman was optimized using the multi-objective satin bowerbird optimizer algorithm (MOSBOA) [211]. The results showed its superiority over other benchmark forecasting models in terms of accuracy and stability. The weights of ELM were optimized using the MOICA [53]. In [243], the weights were optimized using a multi-objective differential evolution algorithm (MODE). A similar prediction framework was applied in [78], where the initial weight and threshold of the WNN were optimized using a multi-objective sine cosine algorithm (MOSCA). A few other cases of individual prediction model parameter optimization are summarized and shown in Table 14.

**Remark:**. *The task of the above*-*mentioned multi*-*objective optimization algorithm is to optimize the weight coefficients of the combined model and the parameters of an individual prediction model*. *In terms of weight coefficients optimization*, *the optimal Pareto solution can be obtained by the multi*- *objective algorithm*, *and the accuracy of the combined prediction model can be improved*. *In terms of parameter optimization of an individual prediction model*, *such as weights and bias of ELM*, *the weights and thresholds of RBF and the kernel function parameters of SVM*, *etc*., *multi*-*objective optimization algorithms are used to improve the accuracy and reliability of the prediction model*, *simultaneously*, *and show good computational performance and robustness*. *Based on the above analysis*, *the multi*-*objective optimization algorithm has great potential in practical engineering applications*.

### 7.5. Comparison of computation time

Computation time is one of the important factors for evaluating the performance of prediction models. Computational time analysis of several optimization algorithms and prediction models can visually demonstrate the performance of the algorithms. The computation times of different optimization algorithms are shown in Fig. 24. Fig. 24(a) shows the computation times of multi- and single-objective algorithms.

It can be found that the former usually takes more time than the latter. Fig. 24 (b) shows the calculation times of different prediction models and multi-objective algorithms. It can be observed that the computation times increase significantly when multi-objective algorithms are used to optimize the weights of the prediction models. Although the implementation of a multi-objective algorithm increases the computation time, it is able to meet the practical needs from the perspective of accuracy and reliability of the prediction model. Moreover, the shortcoming of long computation time can be overcome by using high- performance computers in the future. **8. Open research issues and future works**

At present, meta-heuristic optimization algorithms are being widely applied in wind power/speed forecasting. However, based on an extensive literature review, it can be gathered that there are still a few shortcomings and challenges. Next, we describe a few possible directions for the future development of meta-heuristic optimization algorithms from the following aspects:

* Big data issue

The current applications of meta-heuristic algorithms for predicting the output wind power are limited to individual wind farm power prediction. However, the size of a wind farm dataset can hardly be called big data. With the increasing installed wind power capacity, the size of the dataset also significantly increases, which provides the potential for realizing the accumulation of big data. However, it is difficult to improve the accuracy of wind power prediction with computational efficiency constraints. Namely, for the high-dimensional dataset, e.g., data acquired from hundreds of weather stations and wind farms, the applicability and computational efficiency of the meta-heuristic algorithms are questionable. The effective integration of a sizeable amount of data available from large wind farms and proposing new novel meta- heuristic algorithms is an emerging problem.

* Offline and online modes

Based on an extensive literature survey, it is found that the meta- heuristic optimization algorithms optimize the prediction model parameters that belong to the offline optimization mode. This means that once the best parameters are selected for the prediction model, they will

#### Table 11

Recent works based on ANNs for wind power forecasting.

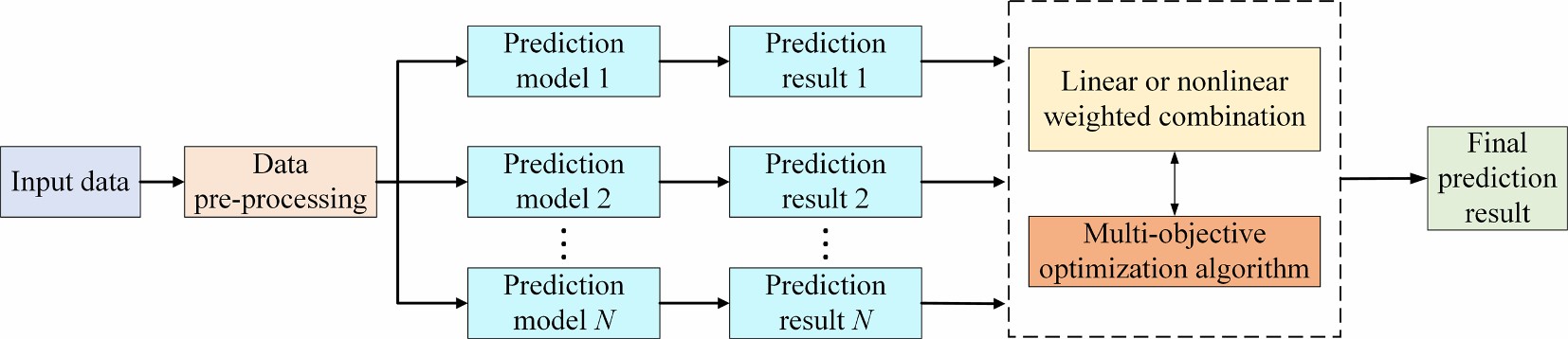
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Refs | Prediction horizon | Year | Prediction model | Error | CPU time | Benefits | Limitations | Main contributions |
| [134] | Ahead-10 min | 2018 | CEEMD-MOWOA- ENN | MAE =  0.2375  NMSE =  0.0039  RMSE =  0.3187  MAPE =  4.2247 | – | Mode-mixing the problem can be solved;  Prediction accuracy be improved. | Need large  computational time; Spurious mode problem. | • A hybrid multi-objective optimization model based on MOWOA is proposed; • The weights and thresholds of the Elman are optimized by MOWOA then loaded to the forecasting model. |
| [46] | Ahead-10 min  (1-step) | 2018 | EWT-GWO-RELM | MAE =  0.0361 m/s RMSE =  0.0447 m/s MAPE =  0.3631% |  | Adaptive;  Original signals are decomposed. | Hard to distinguish time-varying elements. | • The wind speed series decomposed by EWT then are forecasted by RELM optimized by GWO. |
| [194] | Ahead-15 min | 2018 | MLPNN-FA | RRMSE =  9.51%  RMAE =  11.03% | – | Simple structure of the model;  Easy to implement. | Traps into local optima. | * A novel approach is presented based on   MLPNN-FA for wind speed forecasting;   * The parameters of MLPNN are optimized by FA. |
| [51] | Ahead-10 min | 2017 | ENN-MOALO | MAE =  0.185 m/s RMSE =  0.2628 m/s MAPE =  2.4823% |  | Hard to implement; Need few parameters. | Poor iteration performance; Accuracy is affected by initial parameters. | * The original wind speed series are decomposed by CEEMD; * The parameters of the Elman model are optimized by the MOALO. |
| [228] | Ahead-10 min  (1-step) | 2017 | EEMD-CS-WNN | MAE =  0.0975 m/s RMSE =  0.1158 m/s MAPE =  1.6235% |  | Mode-mixing the problem can be solved. | Lack of strict mathematical theory. | * The wind speed time series are decomposed into several subsequences; * The parameters of the WNN are optimized by CS to avoid the over- fitting problem and improve fitting accuracy. |
| [93] | Ahead-10 min | 2017 | TVF-EMD- PSOGSA- CQR-ORELM | MAPE =  0.2373  MAE =  0.0370  RMSE =  0.0720 |  | Adaptive. | Mode-mixing problem; Need large computational time;  Hard to implement. | * A novel model based on TVF-EMD- PSOGSA-CQR-ORELM is developed for wind speed forecasting; * The weights and deviation of ORELM are optimized by PSOGSA. |
| [229] | Ahead-15 min | 2016 | EEMD-FOA-GRNN | MAE =  0.5345  RMSE =  0.7064 | – | Mode-mixing problem can be solved;  Easy to implement. | Lack of strict mathematical theory. | * A hybrid model is proposed based on EEMD-FOA-GRNN for wind speed   forecasting;   * The parameters of GRNN are optimized by FOA. |
| [43] | Ahead-10 min | 2016 | EEMD-GA-BP | MAPE =  6.82  RMSE =  0.59 | 280.84 | Mode-mixing problem can be solved;  Easy to implement. | Lack of strict mathematical theory. | * A hybrid wind speed forecasting model based on the improved EMD and GA-BP neural network method is proposed; * The weights and thresholds of BPNN were optimized by GA. |
| [135]. | Ahead-15 min | 2015 | FEEMD-GA-MLP | MAE =  0.3020  MAPE =  2.67%  RMSE =  4.019 | 45.23 | Adaptive;  Easy to implement. | Added noises. | * The raw wind speed data are decomposed by FEEMD; * The weights and thresholds of MLPNN were optimized by GA. |
| [230] | Ahead-15 min | 2011 | TS-BP | RMSE =  16.05% MAE =  13.09% | – | Easy to implement. | Need large computational time; Dependent on initial parameters. | • A novel hybrid model based on BPNN optimized by TS is proposed for ultra- short-term wind power prediction. |
| [157] | Ahead-45 min | 2016 | SSA–APSOSA–BP | AE =  0.0040  MAE =  0.4285  RMSE =  0.5620  MAPE =  5.9362 |  | Need few parameters. | Time lag  phenomenon; Hard to implement. | * A hybrid algorithm and data processing technologies are combined for wind speed prediction; * The parameters of the BPNN are optimized by APSOSA. |
| [174] | Ahead-24H | 2014 | IS-PSO-BP | MAE =  0.16  MSE = 0.17  MAPE =  15.51% | – | Easy to implement;  Reduced redundant  Information | Need large computational time. | * A hybrid forecasting model is proposed to enhance the forecasting accuracy of wind speed; * The parameters of BPNN are optimized by PSO. |

#### Table 12

The SVM/LSSVM model for wind power forecasting.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Refs | Prediction horizon | Year | Key parameters | Kernel function | Prediction model | Error | Benefits | Limitations | Main contributions |
| [44] | Ahead − 0.5 h | 2014 | *γ* = 3*.*9932  *σ* =  0*.*6863*ε* =  0*.*0004 | Gaussian | WT-SVM-GA | RMSE =  1.2234  MAE =  0.6169  MAPE =  14.79 | Adaptive;  Original signals are decomposed. | Rely on wavelet function;  Decomposition levels determines the prediction accuracy. | * A hybrid model combining WT, GA and SVM is proposed for wind power forecasting; * The parameters (Gaussian) of SVM are optimized by GA. |
| [45] | Ahead − 24 h | 2015 | *γ* = 5*.*4691 *σ* = 0*.*3214 | ERBF | LSSVM-GSA | RMSE =  4.5524  MAE =  3.4305 | Easy to implement; Simple structure. | Rely on parameters of prediction model; Need large historical data. | • A hybrid strategy based on LSSVM–GSA is proposed to forecast the short-term wind power. |
| [231] | Ahead − 48 h | 2020 | – | RBF | SVM-IDA | NRMSE  = 3.25% NMAE =  2.75% | Easy to implement; Simple structure. | The initial parameters depend on experience. | • The key parameters of SVM are tuned by IDA. |
| [128] | Ahead-15 min | 2021 | – | Kernel function switch | CEEMDAN- WPE-LSSVM-  ICS | NRMSE  = 2.74% NMAE =  1.63% | The prediction accuracy has been significantly improved. | Hard to implement; Complex structure. | • A kernel function switch mechanism is firstly presented. |
| [232] | Ahead-50 min | 2021 | – | Gaussian | MWD-LSSVM- NSGA-II | RMSE =  8.78%  MAE =  6.59% | Easy to observe the decomposed signal; Kernel function capability is improved. | Complex structure; Need large computational time. | • A novel multi-step prediction method based on the decomposition, classification, and reconstruction is developed. |
| [233] | Ahead-15 min | 2020 | – | Gaussian | ST-GWO- MSVM | NRMSE  = 4.65%; NMAE =  3.32% | Several wind farms outputs prediction is realized. | Hard to model; Need large computational time. | • The spatio-temporal correlation of wind farms is considered.  Multiple wind farms output is predicted. |
| [234] | Ahead-1 day | 2011 | *σ* =  50*.*265*γ* = 7681*.*3 *ε* = 19*.*375 | – | seasonal recurrent SVR-CABC | MAPE =  2.960% | The modeling stage is detailed. | Complex structure; Hard to implement. | * A concept of seasonal recurrent is firstly proposed; * The parameters of SRSVR are selected by the chaotic behavior of ABC. |
| [235] | Ahead-15 min | 2019 | – | – | LSSVM-DBN- SSA-LS | NRMSE  = 2.81%; NMAE =  1.79% | Different kernel functions are considered;  Prediction accuracy is improved. | No kernel type is defined; Need large  computational time; Phase shift phenomenon. | * Data decomposition and clustering methods are proposed; * Different kernel functions is investigated. |

**Notes:** The regularization parameter (γ) controls the penalty imposed on data points that deviate from the regression function. Meanwhile, the kernel parameter (σ) affects the smoothness of the regression function.



**Fig. 22.** Multi-objective optimization algorithm used for optimizing combination weights.

remain fixed during the wind power prediction process. Consequently, the reliability of the results may be unconvincing, and the optimal parameters of the prediction model may no longer apply to a new dataset. Therefore, the establishment of a meta-heuristic optimization online strategy is a future research direction. This strategy should enable the recognition of new data changes and quick convergence.

* Computation time and accuracy

The meta-heuristic algorithms have been widely used for wind power prediction, and have become a popular research topic in wind power prediction. However, an important challenge is the large amount of time taken by the optimization process. As stated in the no-free lunch (NFL) theory, no single meta-heuristic algorithm can always produce the best results in multi-time scale wind power prediction. Therefore, a promising research direction in the future is to combine different optimization algorithms. The resulting algorithm should have fast convergence speed and global search capability, which can further improve the accuracy and reliability of power prediction models.

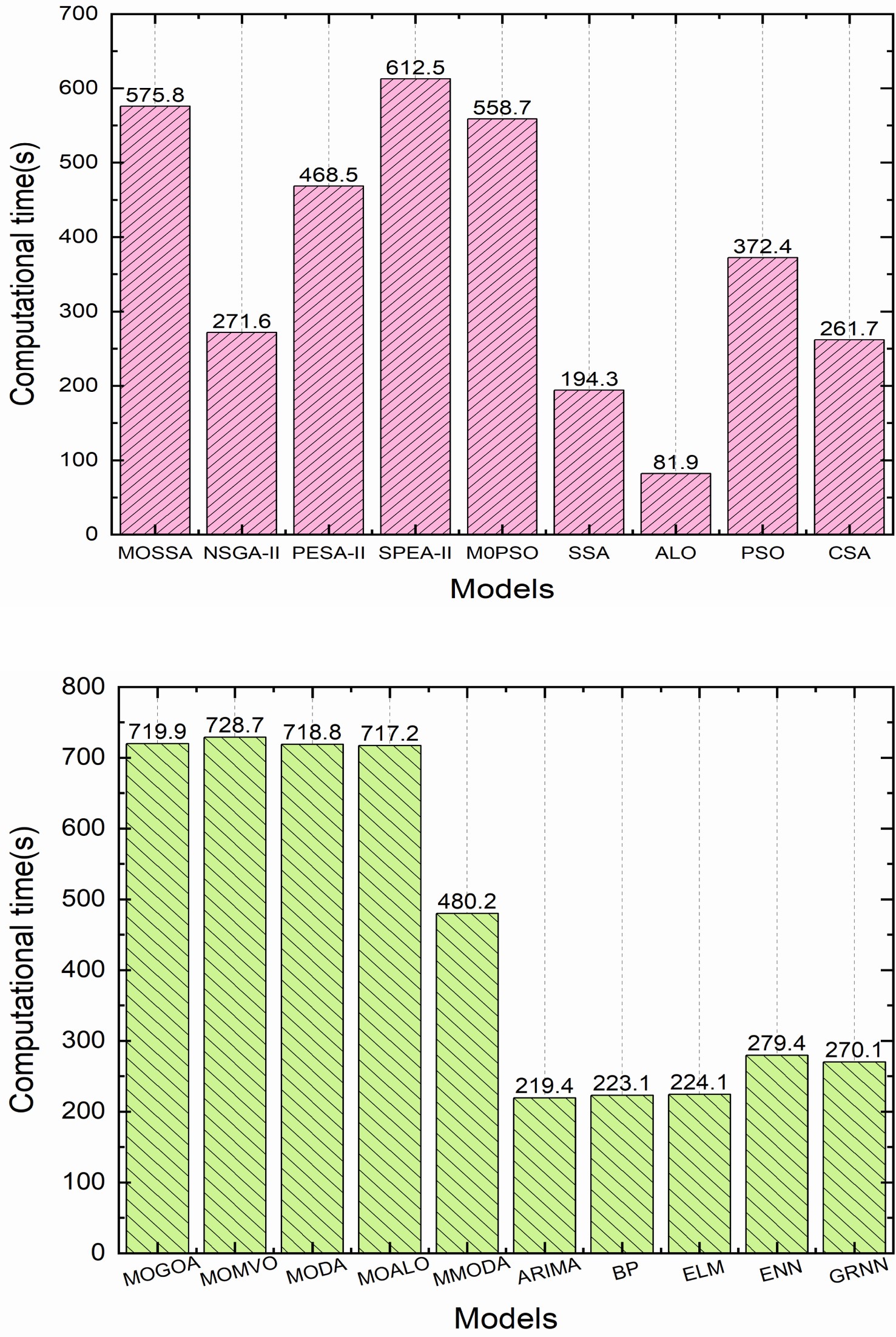
* Synchronization prediction

A complete combined prediction framework includes three stages: the data pre-processing stage, prediction model selection stage, and data post-processing stage. However, meta-heuristic algorithms are used to optimize only the key parameters of the prediction model stage, and the other stages are rarely studied. This makes it impossible to identify the stage in which a large prediction error is caused by the combined

#### Table 13

Variable weight prediction model is optimized by multi-objective optimization algorithms.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Refs | Models | | Optimization algorithms | | Optimized parameters | | **Benefits** | | **Limitations** | | Time horizons |
| [236] | BPNN, WNN, RBFNN, and GRNN | | MOBA | | The weight coefficients for these individual models. | | Obtain pareto optimal solutions;  Obtain nonlinear results.  Clear framework. | | Trapped into a local optimal solution; Slow solution. | | Very short- term |
| [237] | BPNN, RBFNN, GRNN, and WNN | | MOBSFPA | | The weight coefficients for ensemble forecasting. | | Obtain pareto optimal solutions;  Obtain nonlinear results. | | Complex structure; Need large computational time. | | Very short- term |
| [238] | WPD-AdaBoost.MRT- ORELMs | | MOGWO | | The coefficients of all base forecasting data. | | Easy to implement; Preserve the best non- dominated solutions;  More effective and efficient. | | Hard to model;  Need large computational time. | | Very short- term |
| [31] | ARIMA, BPNN, ELM, ENN, and GRNN | | MODA | | The weight coefficients of the multiple prediction models. | | Both linear and nonlinear results are obtained. | | Have a requirement for data;  Need more computation time for iterations. | | Very short- term |
| [239] | Two Bi-LSTM networks | | MOMVO | | The weight coefficients of two forecasting models. | | Obtain high accuracy and strong stability. | | Complex structure; Hard to implement. | | Very short- term |
| [240] | ARIMA, HW, CS-BPNN, and DE-OSELM | | NSGA-III | | The weights of the four branch models. | | Easy to implement and strong robustness;  Obtain pareto optimal solutions. | | Complicated calculation; Uneven solution distribution. | | Very short- term |
| [54] | GRNN, ENN, and CBPNN | | MOPSO | | The weight coefficients. | | Easy to implement; Simple structure. | | Performance in general. | | Very short- term |
| [241] | BPNN, ENN, and ARIMA | | MOGOA | | The weight coefficients of three prediction models. | | Both linear and nonlinear results are obtained. | | Sensitivity to random initial values;  Have a requirement for data stability. | | Very short- term |
| **Table 14** | Optimization the parameter | | **Fig. 23.** Multi-objective optimization algorithm used for optimizing model parameters.  s of prediction models by multi-objective optimization algorithms. | | | | | | | |  |
| Refs Prediction models | | Optimization algorithms | | Optimized parameters | | **Benefits** | | **Limitations** | | Time horizons | |
| [244] RBFNN | | NSGA-II | | The solution with the smallest PINAW. | | Easy to implement and strong robustness;  Obtain pareto optimal solutions. | | Complicated calculation;  Uneven solution distribution. | | Short-term | |
| [51] ENN | | MOALO | | The initial weights and thresholds of the ENN. | | High convergence and coverage; Robustness. | | Poor iteration performance; | | Very short-term and short-term | |
| [78] WNN | | MOSCA | | The initial weight and threshold of the WNN. | | Obtain high accuracy and strong stability. | | Suffer from computational burden. | | Very short-term | |
| [245] MLPNN | | NSGA-II | | The parameters of the MLPNN. | | Easy to implement and strong robustness;  Obtain pareto optimal solutions. | | Complicated calculation;  Uneven solution distribution. | | Short-term | |
| [53] ELM | | MOICA | | The weight of the ELM. | | Both high accuracy and stability. Good fit ability. | | Complex structure; | | Very short-term | |
| [246] SVM | | MODE | | The parameters of the SVM. | | Adaptive search rules;More accurate optimization results. | | May drop in local best. | | Short-term | |
| [213] LSSVM | | MSSA | | The key parameters of the LSSVM. | | Obtain pareto optimal solutions. | | Complicated calculations and not often used. | | Very short-term and short-term | |
| [247] MIMOLSSVM | | MOGWO | | The parameters in the MIMOLSSVM. | | Easy to implement;  Preserve the best non-dominated solutions;  More effective and efficient. | | Easy to fall into local optimum; Hard to model;  Need large computational time. | | Very short-term | |



(

a) multi-objective and individual-objective algorithms

(b) optimization of ensemble models **Fig. 24.** Computation times of different optimization algorithms [204].

forecasting framework. A few key parameters of data pre-processing and post-processing techniques can also affect the prediction accuracy. Therefore, in future research, synchronous optimization of the entire prediction framework can be achieved by using meta-heuristic algorithms to optimize each stage of the prediction framework.

* Artificial neural network and deep learning

Nowadays, the focus of research in wind power prediction is the application of meta-heuristic algorithms to optimize the key parameters of ANNs, e.g., initial weights and thresholds, smoothing factors, etc. As the neural network used is a shallow model, the learning ability of sample features is relatively weak. On the other hand, the deep learning networks have a large number of hidden layers, and excellent training and learning abilities, which can help improve the wind power prediction accuracy. Therefore, some researchers have applied deep learning algorithms to wind power prediction, obtaining promising prediction results. To develop a multi-time scale prediction model, a meta-heuristic algorithm combined with deep learning networks could be used to mitigate the problem of poor prediction results.

* Testing platform

Most of the current wind power prediction programs are implemented in MATLAB. Although some research results in the literature show that wind power prediction can be implemented in Python software, only a selected group of meta-heuristic algorithms have been implemented in this software to optimize the key parameters of the prediction model. Highly efficient algorithms and tools should be further investigated to deal with the high computational complexity of meta-heuristic algorithms. One approach is to implement distributed computing on the Hadoop platform, and another high-performance computation approach is based on using the graphics processing unit (GPU) for performing the computations. Both approaches can be implemented in parallel and reduce the burden of an individual computer.

## **9. Conclusions**

In this paper, different meta-heuristic algorithms that can be used to improve wind power prediction accuracy and stability are systematically summarized. The pros and cons of data decomposition methods to improve the quality of wind power sample data are provided. Then, the impact of key parameters on the prediction accuracy of different predictors is summarized. Furthermore, optimization of the predictor parameters is introduced, which is divided into six categories: evolutionary-based algorithms, physics-based algorithms, human- based algorithms, swarm-based algorithms, hybrid algorithms, and multi-objective optimization algorithms. A comparative study showed that different optimization methods had different characteristics in terms of convergence speed, accuracy, performance, efficiency, and calculation speed. Therefore, the effectiveness of an optimization algorithm can change with the dataset size, parameter settings, application type, etc. Multiple error evaluation metrics (e.g., deterministic, uncertainty, and testing methods) for identifying the performance of each prediction model are presented. A brief discussion on the limitations and future trends of the meta-heuristic algorithms for wind power forecasting problems is also provided.

Based on the literature research, an interesting conclusion can be drawn: By selecting optimal parameters of the optimization algorithm, almost every meta-heuristic algorithm can achieve better and faster convergence compared to traditional optimization algorithms, such as the gradient descent algorithm. However, the rationality of this conclusion is worthy of discussion by relevant researchers. Namely, it is difficult to compare different meta-heuristic algorithms because the parameters of these algorithms, such as population size, quantity, and iteration number differ for multi-time scale wind power prediction. Moreover, it can be gathered based on the no free lunch (NFL) theorem that there is no individual meta-heuristic algorithm that can achieve the best prediction accuracy at all time scales. Therefore, the review presented in this paper suggests that wind power prediction requires the selection of the most suitable meta-heuristic algorithm at different time scales.

Most of the above studies focus on an individual wind farm, in future work, it is expected that the proposed forecasting framework will be extended to the wind power cluster that including hundreds of wind farms. As the number of wind farms increases, the size of data sets increases significantly (e.g., hundreds of wind farms), and more factors need to be considered in the objective function and constraints established, such as spatial correlation constraints among wind farms, upstream and downstream wind speed and direction constraints, and time series delay constraints, etc. It is very increasingly important to improve the computational efficiency and accuracy for the development of power system regulation strategy. Therefore, the future research work is to combine meta-heuristic algorithms and efficient computing processing methods (e.g., based on distributed computing on the Hadoop platform and high-performance computation based on the graphics processing unit (GPU)) to develop a high-accuracy and high-efficiency forecasting model.

## **Declaration of Competing Interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## **Acknowledgments**

This work is supported in part by the Science and Technology Project of State Grid Corporation of China (5108-202155037A-0-0-00) and the National Natural Science Foundation of China (51977213).

## **References**

1. http://www.nea.gov.cn; 2021 [accessed 10 June 2021].
2. [Gass V, Strauss F, Schmidt J, Schmid E. Assessing the effect of wind power](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0010) [uncertainty on profitability. Renew Sustain Energy Rev 2011;15:2677–83.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0010)
3. [Georgilakis PS. Technical challenges associated with the integration of wind power into power systems. Renew Sustain Energy Rev 2008;12:852–63.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0015)
4. [Vargas SA, Esteves GRT, Maçaira PM, Bastos BQ, Cyrino Oliveira FL, Souza RC. Wind power generation: a review and a research agenda. J Cleaner Prod 2019;](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0020)

[218:850–70](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0020).

1. [Pinson P, Nielsen HA, Madsen H, Kariniotakis G. Skill forecasting from ensemble](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0025) [predictions of wind power. Appl Energy 2009;86:1326–34.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0025)
2. Brown BG, Katz RW, Murphy AHJJoAM, Climatology. Time series models to simulate and forecast wind speed and wind power. Journal of Applied Meteorology and Climatology. 1984;23:1184-95.
3. [Nowotarski J, Weron R. Recent advances in electricity price forecasting: a review](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0035) [of probabilistic forecasting. Renew Sustain Energy Rev 2018;81:1548–68.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0035)
4. [Ma L, Luan SY, Jiang CW, Liu HL, Zhang Y. A review on the forecasting of wind](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0040) [speed and generated power. Renew Sustain Energy Rev 2009;13:915–20](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0040).
5. [Qian Z, Pei Y, Zareipour H, Chen N. A review and discussion of decomposition-](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0045) [based hybrid models for wind energy forecasting applications. Appl Energy 2019;](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0045)

[235:939–53](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0045).

1. [Gonzalez-Sop´](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0050) [ena JM, Pakrashi V, Ghosh B. An overview of performance˜](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0050)  [evaluation metrics for short-term statistical wind power forecasting. Renew Sustain Energy Rev 2021;138:110515](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0050).
2. [Raza MQ, Khosravi A. A review on artificial intelligence based load demand](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0055) [forecasting techniques for smart grid and buildings. Renew Sustain Energy Rev](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0055)

[2015;50:1352–72](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0055).

1. [Hu S, Xiang Y, Zhang H, Xie S, Li J, Gu C, et al. Hybrid forecasting method for wind power integrating spatial correlation and corrected numerical weather prediction. Appl Energy 2021;293:116951](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0060).
2. [Andrade JR, Bessa RJ. Improving renewable energy forecasting with a grid of numerical weather predictions. IEEE Trans Sustain Energy 2017;8:1571–80.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0065)
3. [Pawlikowski M, Chorowska A. Weighted ensemble of statistical models. Int J Forecast 2020;36:93–7](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0070).
4. [Sideratos G, Hatziargyriou ND. An advanced statistical method for wind power](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0075) [forecasting. IEEE Trans Power Syst 2007;22:258–65.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0075)
5. [Hill DC, McMillan D, Bell KR. Infield DJIToSE. Application of auto-regressive models to UK wind speed data for power system impact studies. IEEE Trans](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0080) [Sustain Energy 2011;3:134–41](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0080).
6. [Dowell J, Pinson P. Very-short-term probabilistic wind power forecasts by sparse](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0085) [vector autoregression. IEEE Trans Smart Grid 2016;7:763–70](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0085).
7. [Flores JJ, Graff M, Rodriguez HJRE. Evolutive design of ARMA and ANN models](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0090) [for time series forecasting. Renew Energy 2012;44:225–30.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0090)
8. [Yuan C, Liu S, Fang Z. Comparison of China’s primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average)](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0095) [model and GM (1, 1) model. Energy. 2016;100:384–90.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0095)
9. [Sadaei HJ, Enayatifar R, Guimaraes FG, Mahmud M, Alzamil ZA. Combining˜](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0100) [ARFIMA models and fuzzy time series for the forecast of long memory time series.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0100)

[Neurocomputing. 2016;175:782–96.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0100)

1. [Hua K, Mishra Y, Ledwich G. Fast unscented transformation-based transient](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0105) [stability margin estimation incorporating uncertainty of wind generation. IEEE Trans Sustain Energy 2015;6:1254–62.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0105)
2. [Bludszuweit H, Dominguez-Navarro JA, Llombart A. Statistical analysis of wind](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0110) [power forecast error. IEEE Trans Power Syst 2008;23:983–91](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0110).
3. [Jiang P, Liu Z, Niu X, Zhang L. A combined forecasting system based on statistical](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0115) [method, artificial neural networks, and deep learning methods for short-term](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0115) [wind speed forecasting. Energy. 2021;217:119361](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0115).
4. [Chitsazan MA, Fadali MS, Trzynadlowski AM. Wind speed and wind direction](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0120) [forecasting using echo state network with nonlinear functions. Renew Energy](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0120)

[2019;131:879–89](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0120).

1. [Karabacak K, Cetin N. Artificial neural networks for controlling wind–PV power systems: a review. Renew Sustain Energy Rev 2014;29:804–27](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0125).
2. [Wan C, Xu Z, Pinson P, Dong ZY, Wong KP. Probabilistic forecasting of wind power generation using extreme learning machine. IEEE Trans Power Syst 2014;](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0130) [29:1033–44](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0130).
3. [Sun W, Wang Y. Short-term wind speed forecasting based on fast ensemble empirical mode decomposition, phase space reconstruction, sample entropy and](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0135) [improved back-propagation neural network. Energy Convers Manage 2018;157:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0135)

[1–12.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0135)

1. [Tascikaraoglu A, Uzunoglu M. A review of combined approaches for prediction of short-term wind speed and power. Renew Sustain Energy Rev 2014;34:243–54](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0140).
2. [Jung J, Broadwater RP. Current status and future advances for wind speed and](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0145) [power forecasting. Renew Sustain Energy Rev 2014;31:762–77.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0145)
3. [Shi J, Ding Z, Lee W, Yang Y, Liu Y, Zhang M. Hybrid forecasting model for very-](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0150) [short term wind power forecasting based on grey relational analysis and wind](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0150) [speed distribution features. IEEE Trans Smart Grid 2014;5:521–6](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0150).
4. [Liu Z, Jiang P, Zhang L, Niu X. A combined forecasting model for time series:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0155)

[application to short-term wind speed forecasting. Appl Energy 2020;259:114137](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0155).

1. [Wang J, Song Y, Liu F, Hou R. Analysis and application of forecasting models in wind power integration: A review of multi-step-ahead wind speed forecasting](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0160) [models. Renew Sustain Energy Rev 2016;60:960–81.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0160)
2. [Machairas V, Tsangrassoulis A, Axarli K. Algorithms for optimization of building](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0165) [design: a review. Renew Sustain Energy Rev 2014;31:101–12.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0165)
3. [Pesaran HAM, Huy PD, Ramachandaramurthy VK. A review of the optimal allocation of distributed generation: objectives, constraints, methods, and](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0170) [algorithms. Renew Sustain Energy Rev 2017;75:293–312.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0170)
4. [Mp HA, Huy PD, Ramachandaramurthy VK. A review of the optimal allocation of distributed generation: objectives, constraints, methods, and algorithms. Renew Sustain Energy Rev 2017;75:293–312.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0175)
5. [Behera S, Sahoo S, Pati B. A review on optimization algorithms and application to wind energy integration to grid. Renew Sustain Energy Rev 2015;48:214–27.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0180)
6. [Das UK, Tey KS, Seyedmahmoudian M, Mekhilef S, Idris MYI, Van Deventer W,](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0185) [et al. Forecasting of photovoltaic power generation and model optimization: a](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0185) [review. Renew Sustain Energy Rev 2018;81:912–28](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0185).
7. [Wang J, Niu T, Lu H, Yang W, Du P. A novel framework of reservoir computing](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0190) [for deterministic and probabilistic wind power forecasting. IEEE Trans Sustain](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0190) [Energy 2020;11:337–49.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0190)
8. [Buhan S, Çadırcı I. Multistage wind-electric power forecast by using a combination of advanced statistical methods. IEEE Trans Ind Inf 2015;11:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0195)

[1231–42.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0195)

1. [Buhan S, Ozkazanç Y, Çadırcı I. Wind pattern recognition and reference wind¨](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0200) [mast data correlations with NWP for improved wind-electric power forecasts.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0200)

[IEEE Trans Ind Inf 2016;12:991–1004](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0200).

1. [Nazari-Heris M, Mohammadi-Ivatloo B, Gharehpetian G. Short-term scheduling of hydro-based power plants considering application of heuristic algorithms: a](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0205) [comprehensive review. Renew Sustain Energy Rev 2017;74:116–29.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0205)
2. [Wan C, Xu Z, Pinson P. Direct interval forecasting of wind power. IEEE Trans](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0210) [Power Syst 2013;28:4877–8.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0210)
3. [Wang S, Zhang N, Wu L, Wang Y. Wind speed forecasting based on the hybrid](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0215) [ensemble empirical mode decomposition and GA-BP neural network method.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0215)

[Renew Energy 2016;94:629–36](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0215).

1. [Liu D, Niu D, Wang H, Fan L. Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. Renew Energy 2014;62:592–7](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0220).
2. [Yuan X, Chen C, Yuan Y, Huang Y, Tan Q. Short-term wind power prediction based on LSSVM–GSA model. Energy Convers Manage 2015;101:393–401](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0225).
3. [Liu H, Wu H, Li Y. Smart wind speed forecasting using EWT decomposition, GWO](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0230) [evolutionary optimization, RELM learning and IEWT reconstruction. Energy](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0230) [Convers Manage 2018;161:266–83.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0230)
4. [Mirjalili S, Lewis A. The whale optimization algorithm. Adv Eng Softw. 2016;95:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0235)

[51–67.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0235)

1. [Du P, Wang J, Yang W, Niu T. Multi-step ahead forecasting in electrical power system using a hybrid forecasting system. Renew Energy 2018;122:533–50.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0240)
2. [Wang D, Luo H, Grunder O, Lin Y, Guo H. Multi-step ahead electricity price forecasting using a hybrid model based on two-layer decomposition technique and BP neural network optimized by firefly algorithm. Appl Energy 2017;190:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0245)

[390–407.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0245)

1. [Wang J, Du P, Niu T, Yang W. A novel hybrid system based on a new proposed algorithm—multi-objective whale optimization algorithm for wind speed](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0250) [forecasting. Appl Energy 2017;208:344–60](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0250).
2. [Du P, Wang J, Guo Z, Yang W. Research and application of a novel hybrid](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0255) [forecasting system based on multi-objective optimization for wind speed](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0255) [forecasting. Energy Convers Manage 2017;150:90–107.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0255)
3. [Naik J, Dash PK, Dhar S. A multi-objective wind speed and wind power prediction](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0260) [interval forecasting using variational modes decomposition based Multi-kernel robust ridge regression. Renew Energy 2019;136:701–31](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0260).
4. [Li C, Zhu Z, Yang H, Li R. An innovative hybrid system for wind speed forecasting based on fuzzy preprocessing scheme and multi-objective optimization. Energy](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0265)

[2019;174:1219–37.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0265)

1. [He Z, Chen Y, Shang Z, Li C, Li L, Xu M. A novel wind speed forecasting model based on moving window and multi-objective particle swarm optimization](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0270) [algorithm. Appl Math Model 2019;76:717–40.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0270)
2. [Zhou Q, Wang C, Zhang G. A combined forecasting system based on modified multi-objective optimization and sub-model selection strategy for short-term](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0275) [wind speed. Appl Soft Comput 2020;94:106463.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0275)
3. [Yan J, Liu Y, Han S, Wang Y, Feng S. Reviews on uncertainty analysis of wind power forecasting. Renew Sustain Energy Rev 2015;52:1322–30.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0280)
4. [Xiao L, Wang J, Dong Y, Wu J. Combined forecasting models for wind energy forecasting: a case study in China. Renew Sustain Energy Rev 2015;44:271–88](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0285).
5. [Foley AM, Leahy PG, Marvuglia A, McKeogh EJ. Current methods and advances in forecasting of wind power generation. Renew Energy 2012;37:1–8.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0290)
6. [Liu H, Chen C. Data processing strategies in wind energy forecasting models and](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0295) [applications: a comprehensive review. Appl Energy 2019;249:392–408.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0295)
7. [Mandal P, Zareipour H, Rosehart WD. Forecasting aggregated wind power](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0300) [production of multiple wind farms using hybrid wavelet-PSO-NNs. Int J Energy Res. 2014;38:1654–66.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0300)
8. [Heng J, Wang C, Zhao X, Wang J. A hybrid forecasting model based on empirical](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0305) [mode decomposition and the cuckoo search algorithm: a case study for power](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0305) [load. Math Probl Eng 2016;2016:1–28](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0305).
9. [Wang C, Zhang H, Fan W, Ma P. A new chaotic time series hybrid prediction](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0310) [method of wind power based on EEMD-SE and full-parameters continued fraction. Energy 2017;138:977–90.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0310)
10. [Ren Y, Suganthan PN, Srikanth N. A comparative study of empirical mode](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0315) [decomposition-based short-term wind speed forecasting methods. IEEE Trans](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0315) [Sustain Energy 2015;6:236–44](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0315).
11. [Zhang J, Li D, Hao Y, Tan Z. A hybrid model using signal processing technology,](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0320) [econometric models and neural network for carbon spot price forecasting.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0320) 
    1. [Cleaner Prod 2018;204:958–64.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0320)
12. [Liu H, Mi X, Li Y. Comparison of two new intelligent wind speed forecasting](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0325) [approaches based on wavelet packet decomposition, complete ensemble](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0325) [empirical mode decomposition with adaptive noise and artificial neural networks.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0325)

[Energy Convers Manage 2018;155:188–200.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0325)

1. [Zhang W, Qu Z, Zhang K, Mao W, Ma Y, Fan X. A combined model based on CEEMDAN and modified flower pollination algorithm for wind speed forecasting.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0330)

[Energy Convers Manage 2017;136:439–51.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0330)

1. [Mi X-w, Liu H, Li Y-f. Wind speed forecasting method using wavelet, extreme learning machine and outlier correction algorithm. Energy Convers Manage 2017;](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0335) [151:709–22](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0335).
2. [Xiao L, Shao W, Yu M, Ma J, Jin C. Research and application of a hybrid wavelet](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0340) [neural network model with the improved cuckoo search algorithm for electrical](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0340) [power system forecasting. Appl Energy 2017;198:203–22.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0340)
3. [Catalao JPS, Pousinho HMI, Mendes VMF. Hybrid wavelet-PSO-ANFIS approach˜](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0345) [for short-term wind power forecasting in Portugal. IEEE Trans Sustain Energy](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0345)

[2011;2:50–9.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0345)

1. [Meng A, Ge J, Yin H, Chen S. Wind speed forecasting based on wavelet packet](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0350) [decomposition and artificial neural networks trained by crisscross optimization algorithm. Energy Convers Manage 2016;114:75–88](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0350).
2. [Liu H, Mi X, Li Y. Smart multi-step deep learning model for wind speed](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0355) [forecasting based on variational mode decomposition, singular spectrum analysis,](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0355) [LSTM network and ELM. Energy Convers Manage 2018;159:54–64](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0355).
3. [Li Z, Jiang Y, Guo Q, Hu C, Peng Z. Multi-dimensional variational mode decomposition for bearing-crack detection in wind turbines with large driving- speed variations. Renewable Energy 2018;116:55–73.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0360)
4. [Safari N, Chung CY, Price GCD. Novel multi-step short-term wind power](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0365) [prediction framework based on chaotic time series analysis and singular spectrum analysis. IEEE Trans Power Syst 2018;33:590–601.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0365)
5. [Song J, Wang J, Lu H. A novel combined model based on advanced optimization](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0370) [algorithm for short-term wind speed forecasting. Appl Energy 2018;215:643–58.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0370)
6. [Jw A, Hj B, Yw C, Yd C. Forecasting solar radiation using an optimized hybrid model by Cuckoo Search algorithm. Energy. 2015;81:627–44.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0375)
7. [Kaviani S, Sohn I. Application of complex systems topologies in artificial neural](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0380) [networks optimization: an overview. Expert Syst Appl 2021;180:115073](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0380).
8. [Marugan AP, ´](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0385) [Marquez FPG, Perez JMP, Ruiz-Her´](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0385) [nandez D. A survey of artificial´](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0385) [neural network in wind energy systems. Appl Energy 2018;228:1822–36](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0385).
9. [Wang J, Yang W, Du P, Niu T. A novel hybrid forecasting system of wind speed](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0390) [based on a newly developed multi-objective sine cosine algorithm. Energy Convers Manage 2018;163:134–50.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0390)
10. [Specht DF. A general regression neural network. IEEE Trans Neural Networks](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0395)

[2002;2:568–76.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0395)

1. [Kumar G, Malik H. Generalized regression neural network based wind speed](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0400) [prediction model for western region of India. Procedia Comput Sci 2016;93:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0400)

[26–32.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0400)

1. [Ben-Nakhi AE, Mahmoud MA. Cooling load prediction for buildings using general](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0405) [regression neural networks. Energy Convers Manage 2004;45:2127–41.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0405)
2. [Sideratos G, Hatziargyriou ND. Probabilistic wind power forecasting using radial](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0410) [basis function neural networks. IEEE Trans Power Syst 2012;27:1788–96.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0410)
3. [Chng ES, Chen S, Mulgrew B. Gradient radial basis function networks for nonlinear and nonstationary time series prediction. IEEE Trans Neural Networks](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0415) [1996;7:190–4.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0415)
4. [Leung H, Lo T, Wang S. Prediction of noisy chaotic time series using an optimal](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0420) [radial basis function neural network. IEEE Trans Neural Networks 2001;12:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0420)

[1163–72](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0420).

1. [Ranaweera DK, Hubele NF, Papalexopoulos AD. Application of radial basis](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0425) [function neural network model for short-term load forecasting. IET Proc - Gener](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0425) [Trans Distribut 1995;142:45–50.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0425)
2. [Wang S, Qin C, Feng Q, Javadpour F, Rui Z. A framework for predicting the production performance of unconventional resources using deep learning. Appl Energy 2021;295:117016](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0430).
3. [Ti Z, Deng XW, Zhang M. Artificial Neural Networks based wake model for power](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0435) [prediction of wind farm. Renewable Energy 2021;172:618–31.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0435)
4. [Wen-Yeau C. Application of back propagation neural network for wind power](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0440) [generation forecasting. Int J Digital Content Technol Appl 2013;7:502–9.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0440)
5. [Madhiarasan M, Deepa SN. A novel criterion to select hidden neuron numbers in](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0445) [improved back propagation networks for wind speed forecasting. Appl Intell](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0445)

[2016;44:878–93.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0445)

1. [Tang J, Deng C, Huang G-B. Extreme learning machine for multilayer perceptron. IEEE Trans Neural Networks Learn Syst 2016;27:809–21.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0450)
2. [Wang L, Li X, Bai Y. Short-term wind speed prediction using an extreme learning machine model with error correction. Energy Convers Manage 2018;162:239–50.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0455)
3. [Qi M, Gao H, Wang L, Xiang Y, Lv L, Liu J. Wind power interval forecasting based on adaptive decomposition and probabilistic regularised extreme learning](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0460) [machine. IET Renew Power Gener 2020;14:3181–91.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0460)
4. [Zheng W, Peng X, Lu D, Zhang D, Liu Y, Lin Z, et al. Composite quantile regression extreme learning machine with feature selection for short-term wind speed](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0465) [forecasting: A new approach. Energy Convers Manage 2017;151:737–52.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0465)
5. [Peng X, Zheng W, Zhang D, Liu Y, Lu D, Lin L. A novel probabilistic wind speed forecasting based on combination of the adaptive ensemble of on-line sequential ORELM (Outlier Robust Extreme Learning Machine) and TVMCF (time-varying mixture copula function). Energy Convers Manage 2017;138:587–602.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0470)
6. [Peng T, Zhou J, Zhang C, Zheng Y. Multi-step ahead wind speed forecasting using a hybrid model based on two-stage decomposition technique and AdaBoost- extreme learning machine. Energy Convers Manage 2017;153:589–602.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0475)
7. [Mishra SP. Short-term forecasting of wind power generation using extreme](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0480) [learning machine and its variants. Int J Power Energy Convers 2017;8:68–89](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0480).
8. [Xiao C, Dong Z, Xu Y, Meng K, Zhou X, Zhang X. Rational and self-adaptive](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0485) [evolutionary extreme learning machine for electricity price forecast. Memetic](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0485) [Computing 2016;8:223–33.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0485)
9. [Liu H, Tian H-q, Li Y-f. Comparison of new hybrid FEEMD-MLP, FEEMD-ANFIS,](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0490) [wavelet packet-MLP and wavelet packet-ANFIS for wind speed predictions.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0490)

[Energy Convers Manage 2015;89:1–11](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0490).

1. [Niu M, Sun S, Wu J, Yu L, Wang J. An innovative integrated model using the singular spectrum analysis and nonlinear multi-layer perceptron network](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0495) [optimized by hybrid intelligent algorithm for short-term load forecasting. Appl Math Model 2016;40:4079–93.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0495)
2. [Yeh W-C, Yeh Y-M, Chang P-C, Ke Y-C, Chung V. Forecasting wind power in the Mai Liao Wind Farm based on the multi-layer perceptron artificial neural network model with improved simplified swarm optimization. Int J Electr Power Energy](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0500) [Syst 2014;55:741–8.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0500)
3. [Sheela KG, Deepa SN. Neural network based hybrid computing model for wind speed prediction. Neurocomputing 2013;122:425–9](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0505).
4. [Madhiarasan M, Deepa SN. Comparative analysis on hidden neurons estimation in multi layer perceptron neural networks for wind speed forecasting. Artif Intell Rev 2017;48:449–71](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0510).
5. [Kohonen T. Self-organized formation of topologically correct feature maps. Biol Cybern 1982;43:59–69](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0515).
6. [Hu R, Hu W, Gokmen N, Li P, Huang Q, Chen Z. High resolution wind speed¨](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0520) [forecasting based on wavelet decomposed phase space reconstruction and self- organizing map. Renew Energy 2019;140:17–31.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0520)
7. Tan K, Logenthiran T, Woo WL. Forecasting of wind energy generation using self- organizing maps and extreme learning machines. 2016 IEEE Region 10 Conference (TENCON): IEEE; 2016. p. 451-4.
8. [Bohari ZH, Azemy HS, Nasir MNM, Baharom MF, Sulaima MF, Jali MH. Reliable short term load forecasting using self organizing map (SOM) in deregulated](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0530) [electricity market. J Theoret Appl Inform Technol 2015;79:389–94](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0530).
9. [Lan H, Yin H, Hong Y-Y, Wen S, Yu DC, Cheng P. Day-ahead spatio-temporal forecasting of solar irradiation along a navigation route. Appl Energy 2018;211:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0535)

[15–27](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0535).

1. [Elman JL. Finding structure in time. Cognit Sci 1990;14:179–211.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0540)
2. [Yu C, Li Y, Xiang H, Zhang M. Data mining-assisted short-term wind speed](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0545) [forecasting by wavelet packet decomposition and Elman neural network. J Wind Eng Ind Aerodyn 2018;175:136–43](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0545).
3. [Liu H, Mi X-W, Li Y-F. Wind speed forecasting method based on deep learning](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0550) [strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. Energy Convers Manage 2018;156:498–514](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0550).
4. [Yu C, Li Y, Zhang M. An improved wavelet transform using singular spectrum analysis for wind speed forecasting based on elman neural network. Energy](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0555) [Convers Manage 2017;148:895–904.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0555)
5. [Yu C, Li Y, Zhang M. Comparative study on three new hybrid models using Elman neural network and empirical mode decomposition based technologies improved by singular spectrum analysis for hour-ahead wind speed forecasting. Energy](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0560) [Convers Manage 2017;147:75–85.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0560)
6. [Yang Z, Wang J. Multistep wind speed forecasting using a novel model hybridizing singular spectrum analysis, modified intelligent optimization, and rolling elman neural network. Math Probl Eng 2016;2016:1–21.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0565)
7. [Liu H, Tian H-q. Liang X-f, Li Y-f. Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks. Appl Energy](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0570)

[2015;157:183–94](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0570).

1. [Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0575)

[1735–80](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0575).

1. [Liu H, Mi X, Li Y. Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0580) [long short term memory network. Energy Convers Manage 2018;166:120–31.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0580)
2. [Nguyen H-P, Baraldi P, Zio E. Ensemble empirical mode decomposition and long short-term memory neural network for multi-step predictions of time series signals in nuclear power plants. Appl Energy 2021;283:116346](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0585).
3. [Hu H, Wang L, Tao R. Wind speed forecasting based on variational mode](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0590) [decomposition and improved echo state network. Renew Energy 2021;164:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0590)

[729–51.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0590)

1. [Qin Y, Li K, Liang Z, Lee B, Zhang F, Gu Y, et al. Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal. Appl Energy 2019;236:262–72.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0595)
2. [Barbounis TG, Theocharis JB, Alexiadis MC, Dokopoulos PS. Long-term wind](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0600) [speed and power forecasting using local recurrent neural network models. IEEE Trans Energy Convers 2006;21:273–84.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0600)
3. [Barbounis TG, Theocharis JB. Locally recurrent neural networks for wind speed prediction using spatial correlation. Inf Sci 2007;177:5775–97](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0605).
4. [Barbounis TG, Theocharis JB. Locally recurrent neural networks for long-term wind speed and power prediction. Neurocomputing 2006;69:466–96](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0610).
5. [Cao Q, Ewing BT, Thompson MA. Forecasting wind speed with recurrent neural](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0615) [networks. Eur J Oper Res 2012;221:148–54.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0615)
6. [Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. Proc Natl Acad Sci 1982;79:2554–8.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0620)
7. Ahmadi A, Zargaran Z, Mohebi A, Taghavi F. Hybrid model for weather forecasting using ensemble of neural networks and mutual information. 2014 IEEE geoscience and remote sensing symposium: IEEE; 2014. p. 3774–7.
8. [Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273–97](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0630).
9. [Zendehboudi A, Baseer MA, Saidur R. Application of support vector machine models for forecasting solar and wind energy resources: a review. J Cleaner Prod](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0635)

[2018;199:272–85](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0635).

1. [Lu P, Ye L, Tang Y, Zhao Y, Zhong W, Qu Y, et al. Ultra-short-term combined prediction approach based on kernel function switch mechanism. Renew Energy 2021;164:842–66](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0640).
2. [Guo XC, Yang JH, Wu CG, Wang CY, Liang YC. A novel LS-SVMs hyper-parameter](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0645) [selection based on particle swarm optimization. Neurocomputing 2008;71:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0645)

[3211–5](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0645).

1. [Ouyang T, Zha X, Qin L, Xiong Y, Xia T. Wind power prediction method based on](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0650) [regime of switching kernel functions. J Wind Eng Ind Aerodyn 2016;153:26–33.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0650)
2. [Li C, Lin S, Xu F, Liu D, Liu J. Short-term wind power prediction based on data](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0655) [mining technology and improved support vector machine method: a case study in](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0655) [Northwest China. J Cleaner Prod 2018;205:909–22.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0655)
3. [Zhou J, Shi J, Li G. Fine tuning support vector machines for short-term wind](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0660) [speed forecasting. Energy Convers Manage 2011;52:1990–8.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0660)
4. [Santamaria-Bonfil G, Reyes-Ballesteros A, Gershenson C. Wind speed forecasting](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0665) [for wind farms: a method based on support vector regression. Renew Energy](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0665)

[2016;85:790–809](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0665).

1. [Wang J, Du P, Niu T, Yang W. A novel hybrid system based on a new proposed algorithm-multi-objective whale optimization algorithm for wind speed forecasting. Appl Energy 2017;208:344–60.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0670)
2. [Liu H, Tian H, Liang X, Li Y. New wind speed forecasting approaches using fast ensemble empirical model decomposition, genetic algorithm, mind evolutionary](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0675) [algorithm and artificial neural networks. Renew Energy 2015;83:1066–75.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0675)
3. Wang Y, Wang S, Zhang N. A novel wind speed forecasting method based on ensemble empirical mode decomposition and GA-BP neural network. 2013 IEEE Power & Energy Society General Meeting: IEEE; 2013. p. 1–5.
4. [Jawad M, Ali SM, Khan B, Mehmood CA, Farid U, Ullah Z, et al. Genetic algorithm-based non-linear auto-regressive with exogenous inputs neural network](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0685) [short-term and medium-term uncertainty modelling and prediction for electrical](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0685) [load and wind speed. J Eng 2018;2018:721–9.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0685)
5. Prado F, Minutolo MC, Kristjanpoller W. Forecasting based on an ensemble

Autoregressive Moving Average - Adaptive neuro - Fuzzy inference system – Neural network - Genetic Algorithm Framework. Energy. 2020;197:117159.

1. [Wong H-S, Guan L. Application of evolutionary programming to adaptive regularization in image restoration. IEEE Trans Evol Comput 2000;4:309–26](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0695).
2. [Yao X, Liu Y, Lin G. Evolutionary programming made faster. IEEE Trans Evol Comput 1999;3:82–102.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0700)
3. Hervas-Martínez C, Guti´ ´errez PA, Fernandez JC, Salcedo-Sanz S, Portilla-Figueras ´ A, Perez-Bellido A, et al. Hyperbolic tangent basis function neural networks ´ training by hybrid evolutionary programming for accurate short-term wind speed prediction. 2009 Ninth International Conference on Intelligent Systems Design and Applications: IEEE; 2009. p. 193-8.
4. [Salcedo-Sanz S, Ortiz-Garcia EG, Perez-Bellido AM, Portilla-Figueras A, Prieto L.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0710) [Short term wind speed prediction based on evolutionary support vector regression](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0710) [algorithms. Expert Syst Appl 2011;38:4052–7.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0710)
5. [Koza JR. Genetic programming as a means for programming computers by natural](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0715) [selection. Statist Comput 1994;4:87–112](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0715).
6. [Zameer A, Arshad J, Khan A, Raja MAZ. Intelligent and robust prediction of short](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0720) [term wind power using genetic programming based ensemble of neural networks.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0720)

[Energy Convers Manage 2017;134:361–72.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0720)

1. [Hu Y-L, Chen L. A nonlinear hybrid wind speed forecasting model using LSTM](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0725) [network, hysteretic ELM and differential evolution algorithm. Energy Convers](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0725) [Manage 2018;173:123–42.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0725)
2. [Osorio GJ, Matias JCO, Catalao JPS. Short-term wind power forecasting using](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0730) [adaptive neuro-fuzzy inference system combined with evolutionary particle](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0730) [swarm optimization, wavelet transform and mutual information. Renew Energy](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0730)

[2015;75:301–7](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0730).

1. [Rashedi E, Nezamabadi-Pour H, Saryazdi SGSA. A gravitational search algorithm.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0735)

[Inf Sci 2009;179:2232–48](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0735).

1. [Yuan XH, Chen C, Yuan YB, Huang YH, Tan QX. Short-term wind power](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0740) [prediction based on LSSVM-GSA model. Energy Convers Manage 2015;101:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0740)

[393–401](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0740).

1. [Ju FY, Hong WC. Application of seasonal SVR with chaotic gravitational search algorithm in electricity forecasting. Appl Math Model 2013;37:9643–51.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0745)
2. [Shayeghi H, Ghasemi A. Day-ahead electricity prices forecasting by a modified CGSA technique and hybrid WT in LSSVM based scheme. Energy Convers Manage](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0750)

[2013;74:482–91](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0750).

1. [Wang Y, Wang JZ, Wei X. A hybrid wind speed forecasting model based on phase space reconstruction theory and Markov model: A case study of wind farms in northwest China. Energy. 2015;91:556–72.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0755)
2. [Kirkpatrick S, Gelatt CD, Vecchi MP. Optimization by simulated annealing.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0760)

[Science 1983;220:671–80.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0760)

1. [Jiang P, Ge YJ, Wang C. Research and application of a hybrid forecasting model based on simulated annealing algorithm: a case study of wind speed forecasting. J Renew Sustain Energy. 2016;8:17.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0765)
2. [Du P, Jin Y, Zhang K. A hybrid multi-step rolling forecasting model based on ssa and simulated annealing—adaptive particle swarm optimization for wind speed.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0770)

[Sustainability. 2016;8:754–79.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0770)

1. [Mirjalili S. SCA: a sine cosine algorithm for solving optimization problems.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0775)

[Knowl-Based Syst 2016;96:120–33](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0775).

1. [Peng T, Zhang C, Zhou J, Nazir MS. An integrated framework of Bi-directional long-short term memory (BiLSTM) based on sine cosine algorithm for hourly solar radiation forecasting. Energy. 2021;221:119887](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0780).
2. [Guo ZH, Chi DZ, Wu J, Zhang WY. A new wind speed forecasting strategy based on the chaotic time series modelling technique and the Apriori algorithm. Energy](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0785) [Convers Manage 2014;84:140–51.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0785)
3. [Farmer JD, Sidorowich JJ. Predicting chaotic time series. Phys Rev Lett. 1987;59: 845.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0790)
4. [Rao RV, Savsani VJ, Vakharia D. Teaching–learning-based optimization: a novel method for constrained mechanical design optimization problems. Comput-Aided](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0795) [Des. 2011;43:303–15.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0795)
5. [Yan J, Li K, Bai E, Yang Z, Foley A. Time series wind power forecasting based on variant Gaussian Process and TLBO. Neurocomputing. 2016;189:135–44.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0800)
6. [Wang L, Zou F, Hei X, Yang D, Chen D, Jiang Q, et al. A hybridization of teaching- learning-based optimization and differential evolution for chaotic time series prediction. Neural Comput Appl 2014;25:1407–22.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0805)
7. [Wang L, Zou F, Hei XH, Yang DD, Chen DB, Jiang QY. An improved teaching-](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0810) [learning-based optimization with neighborhood search for applications of ANN.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0810)

[Neurocomputing. 2014;143:231–47.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0810)

1. [Satapathy SC, Naik A, Parvathi K. Teaching learning based optimization for neural networks learning enhancement. Int Conf Swarm Evolut Memetic Comput: Springer 2012:761–9](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0815).
2. Reynolds RG. An introduction to cultural algorithms. Proceedings of the third annual conference on evolutionary programming: World Scientific; 1994. p. 131- 9.
3. Chen B, Zhao L, Lu JH. Wind power forecast using RBF network and culture algorithm. 2009 International Conference on Sustainable Power Generation and Supply: IEEE; 2009. p. 1-6.
4. [Geem ZW, Kim JH, Loganathan GV. A new heuristic optimization algorithm: harmony search. Simulation 2001;76:60–8](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0830).
5. [Salcedo-Sanz S, Pastor-Sanchez A, Del Ser J, Prieto L, Geem ZW. A Coral reefs](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0835) [optimization algorithm with harmony search operators for accurate wind speed](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0835) [prediction. Renew Energy 2015;75:93–101.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0835)
6. [Song Z-y, Li J-j. Prediction model based on least squares support vector machine with harmony search and its application. J Harbin Inst Technol 2009;41:207–10](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0840).
7. [He S, Wu QH, Saunders JR. Group search optimizer: an optimization algorithm](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0845) [inspired by animal searching behavior. IEEE Trans Evol Comput 2009;13:973–90](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0845).
8. [Li Y, Wu Q, Jiang L, Yang J, Xu D. Optimal power system dispatch with wind](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0850) [power integrated using nonlinear interval optimization and evidential reasoning approach. IEEE Trans Power Syst 2015;31:2246–54.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0850)
9. [Poli R, Kennedy J, Blackwell T. Particle swarm optimization. Swarm Intell 2007;](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0855)

[1:33–57](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0855).

1. [Wang J, Zhang F, Liu F, Ma J. Hybrid forecasting model-based data mining and](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0860) [genetic algorithm-adaptive particle swarm optimisation: a case study of wind](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0860) [speed time series. IET Renew Power Gener 2016;10:287–98](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0860).
2. [Amjady N, Keynia F, Zareipour H. Wind power prediction by a new forecast engine composed of modified hybrid neural network and enhanced particle](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0865) [swarm optimization. IEEE Trans Sustain Energy 2011;2:265–76](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0865).
3. [Ren C, An N, Wang JZ, Li L, Hu B, Shang D. Optimal parameters selection for BP](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0870) [neural network based on particle swarm optimization: a case study of wind speed forecasting. Knowl-Based Syst 2014;56:226–39.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0870)
4. [Pousinho HMI, Mendes VMF, Catalao JPS. A hybrid PSO-ANFIS approach for](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0875) [short-term wind power prediction in Portugal. Energy Convers Manage 2011;52:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0875)

[397–402](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0875).

1. [Hong WC. Chaotic particle swarm optimization algorithm in a support vector](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0880) [regression electric load forecasting model. Energy Convers Manage 2009;50:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0880)

[105–17.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0880)

1. [Gharehchopogh FS, Gholizadeh H. A comprehensive survey: whale optimization algorithm and its applications. Swarm Evol Comput 2019;48:1–24](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0885).
2. [Liu JP, Li CL. The short-term power load forecasting based on sperm whale](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0890) [algorithm and wavelet least square support vector machine with DWT-IR for feature selection. Sustainability 2017;9:20](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0890).
3. [Kar AK. Bio inspired computing –a review of algorithms and scope of applications. Expert Syst Appl 2016;59:20–32.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0895)
4. [Niu T, Wang JZ, Zhang KQ, Du P. Multi-step-ahead wind speed forecasting based on optimal feature selection and a modified bat algorithm with the cognition strategy. Renew Energy 2018;118:213–29.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0900)
5. [Xiao LY, Qian F, Shao W. Multi-step wind speed forecasting based on a hybrid](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0905) [forecasting architecture and an improved bat algorithm. Energy Convers Manage](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0905)

[2017;143:410–30](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0905).

1. [Yang X-S, Deb S. Cuckoo search: recent advances and applications. Neural Comput Appl 2014;24:169–74.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0910)
2. [Ong P, Zainuddin Z. Optimizing wavelet neural networks using modified cuckoo](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0915) [search for multi-step ahead chaotic time series prediction. Appl Soft Comput](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0915)

[2019;80:374–86](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0915).

1. [Wang J, Jiang H, Wu Y, Dong Y. Forecasting solar radiation using an optimized](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0920) [hybrid model by Cuckoo Search algorithm. Energy 2015;81:627–44](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0920).
2. [Wu Z, Zhao X, Ma Y, Zhao X. A hybrid model based on modified multi-objective](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0925) [cuckoo search algorithm for short-term load forecasting. Appl Energy 2019;237:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0925)

[896–909](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0925).

1. [Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimizer. Adv Eng Softw 2014;69:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0930)

[46–61](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0930).

1. [Altan A, Karasu S, Zio E. A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0935) [grey wolf optimizer. Appl Soft Comput 2021;100:106996](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0935).
2. [Wu C, Wang J, Chen X, Du P, Yang W. A novel hybrid system based on multi- objective optimization for wind speed forecasting. Renew Energy 2020;146:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0940)

[149–65](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0940).

1. [Karaboga D, Ozturk C. A novel clustering approach: artificial Bee Colony (ABC) algorithm. Appl Soft Comput 2011;11:652–7](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0945).
2. [Yahya NA, Samsudin R, Shabri A, Saeed F. Combined group method of data](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0950) [handling models using artificial bee colony algorithm in time series forecasting.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0950)

[Procedia Comput Sci 2019;163:319–29.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0950)

1. [Fei S-w, He Y. Wind speed prediction using the hybrid model of wavelet](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0955) [decomposition and artificial bee colony algorithm-based relevance vector](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0955) [machine. Int J Electr Power Energy Syst 2015;73:625–31.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0955)
2. [Mirjalili S. The ant lion optimizer. Adv Eng Softw 2015;83:80–98](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0960).
3. [Yang X-S, He X. Firefly algorithm: recent advances and applications. Int J Swarm](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0965) [Intell 2013;1:36–50.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0965)
4. [Deo RC, Ghorbani MA, Samadianfard S, Maraseni T, Bilgili M, Biazar M. Multi-](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0970) [layer perceptron hybrid model integrated with the firefly optimizer algorithm for](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0970) [windspeed prediction of target site using a limited set of neighboring reference station data. Renew Energy 2018;116:309–23.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0970)
5. [Kavousi-Fard A, Samet H, Marzbani F. A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0975)

[Expert Syst Appl 2014;41:6047–56.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0975)

1. [Gani A, Mohammadi K, Shamshirband S, Altameem TA, Petkovic D, Ch S.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0980)

[A combined method to estimate wind speed distribution based on integrating the](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0980) [support vector machine with firefly algorithm. Environ Prog Sustain Energy 2016;](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0980) [35:867–75](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0980).

1. [Xiao L, Shao W, Liang T, Wang C. A combined model based on multiple seasonal](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0985) [patterns and modified firefly algorithm for electrical load forecasting. Appl Energy 2016;167:135–53.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0985)
2. [Kar D, Ghosh M, Guha R, Sarkar R, Garcia-Hernandez L, Abraham A. Fuzzy mutation embedded hybrids of gravitational search and Particle Swarm Optimization methods for engineering design problems. Eng Appl Artif Intell 2020;95](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0990).
3. [Kiran MS, Ozceylan E, Gunduz M, Paksoy T. A novel hybrid approach based on particle swarm optimization and ant colony algorithm to forecast energy demand](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0995) [of Turkey. Energy Convers Manage 2012;53:75–83](http://refhub.elsevier.com/S0306-2619(21)00836-9/h0995).
4. [Kıran MS, Ozceylan E, Gündüz M, Paksoy T. A novel hybrid approach based on¨](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1000) [particle swarm optimization and ant colony algorithm to forecast energy demand](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1000) [of Turkey. Energy Convers Manage 2012;53:75–83](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1000).
5. [Rahmani R, Yusof R, Seyedmahmoudian M, Mekhilef S. Hybrid technique of ant](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1005) [colony and particle swarm optimization for short term wind energy forecasting.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1005) 
   1. [Wind Eng Ind Aerodyn 2013;123:163–70.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1005)
6. Chakraborty D, Saha S, Maity S. Training feedforward neural networks using hybrid flower pollination-gravitational search algorithm. 2015 International Conference on Futuristic Trends on Computational Analysis and Knowledge Management (ABLAZE): IEEE; 2015. p. 261-6.
7. [Mafarja MM, Mirjalili S. Hybrid Whale Optimization Algorithm with simulated annealing for feature selection. Neurocomputing 2017;260:302–12.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1015)
8. [Liu H, Li Y, Duan Z, Chen C. A review on multi-objective optimization framework in wind energy forecasting techniques and applications. Energy Convers Manage](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1020) [2020;224:113324](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1020).
9. [De Giorgi MG, Congedo PM, Malvoni M, Laforgia D. Error analysis of hybrid photovoltaic power forecasting models: a case study of mediterranean climate.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1025)

[Energy Convers Manage 2015;100:117–30.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1025)

1. [Wang J-Z, Wang Y, Jiang P. The study and application of a novel hybrid forecasting model –a case study of wind speed forecasting in China. Appl Energy](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1030)

[2015;143:472–88](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1030).

1. [Wang Y, Hu Q, Meng D, Zhu P. Deterministic and probabilistic wind power](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1035) [forecasting using a variational Bayesian-based adaptive robust multi-kernel regression model. Appl Energy 2017;208:1097–112.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1035)
2. [Chen H, Hou D. Research on superior combination forecasting model based on forecasting effective measure. J Univ Sci Technol China 2002;32:172–80.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1040)
3. [Heng JN, Wang JZ, Xiao LY, Lu HY. Research and application of a combined model based on frequent pattern growth algorithm and multi-objective](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1045) [optimization for solar radiation forecasting. Appl Energy 2017;208:845–66.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1045)
4. [Wang J-Z, Wang Y, Jiang P. The study and application of a novel hybrid forecasting model –a case study of wind speed forecasting in China. Appl Energy 2015;143:472–88](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1050).
5. [Tian C, Hao Y, Hu J. A novel wind speed forecasting system based on hybrid data](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1055) [preprocessing and multi-objective optimization. Appl Energy 2018;231:301–19.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1055)
6. [Wang J, Wang S, Yang W. A novel non-linear combination system for short-term](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1060) [wind speed forecast. Renew Energy 2019;143:1172–92](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1060).
7. [Jiang P, Li R, Li H. Multi-objective algorithm for the design of prediction intervals for wind power forecasting model. Appl Math Model 2019;67:101–22](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1065).
8. [Wang J, Niu T, Lu H, Guo Z, Yang W, Du P. An analysis-forecast system for uncertainty modeling of wind speed: a case study of large-scale wind farms. Appl Energy 2018;211:492–512.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1070)
9. [Hu Y, Qu B, Wang J, Liang J, Wang Y, Yu K, et al. Short-term load forecasting using multimodal evolutionary algorithm and random vector functional link network based ensemble learning. Appl Energy 2021;285:116415](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1075).
10. [Li Z, Lin X, Zhang Q, Liu H. Evolution strategies for continuous optimization: a survey of the state-of-the-art. Swarm Evol Comput 2020;56:100694.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1080)
11. [Mokhtari H. A nature inspired intelligent water drops evolutionary algorithm for parallel processor scheduling with rejection. Appl Soft Comput 2015;26:166–79](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1085).
12. [An X, Jiang D, Liu C, Zhao M. Wind farm power prediction based on wavelet decomposition and chaotic time series. Expert Syst Appl 2011;38:11280–5.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1090)
13. [Niu D, Wang Y, Wu DD. Power load forecasting using support vector machine and ant colony optimization. Expert Syst Appl 2010;37:2531–9.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1095)
14. [Huang CM, Huang CJ, Wang ML. A particle swarm optimization to identifying the ARMAX model for short-term load forecasting. IEEE Trans Power Syst 2005;20:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1100)

[1126–33](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1100).

1. [Pan W-T. A new fruit fly optimization algorithm: taking the financial distress](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1105) [model as an example. Knowl-Based Syst 2012;26:69–74.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1105)
2. [Cao G, Wu L. Support vector regression with fruit fly optimization algorithm for seasonal electricity consumption forecasting. Energy 2016;115:734–45.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1110)
3. [Xu D, Ding S. Research on improved GWO-optimized SVM-based short-term load](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1115) [prediction for cloud computing. Comput Eng Appl 2017;53:68–73](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1115).
4. [Khalilpourazari S, Khalilpourazary S. Optimization of production time in the multi-pass milling process via a Robust Grey Wolf Optimizer. Neural Comput Appl](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1120)

[2018;29:1321–36](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1120).

1. [Akay B, Karaboga D, Gorkemli B, Kaya E. A survey on the Artificial Bee Colony algorithm variants for binary, integer and mixed integer programming problems.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1125)

[Appl Soft Comput 2021;106:107351](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1125).

1. [Ali ES, Abd Elazim SM, Abdelaziz AY. Ant lion optimization algorithm for renewable distributed generations. Energy 2016;116:445–58.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1130)
2. [Wang J, Du P, Lu H, Yang W, Niu T. An improved grey model optimized by multi- objective ant lion optimization algorithm for annual electricity consumption forecasting. Appl Soft Comput 2018;72:321–37](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1135).
3. [Zhang K, Qu Z, Wang J, Zhang W, Yang F. A novel hybrid approach based on cuckoo search optimization algorithm for short-term wind speed forecasting.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1140)

[Environ Prog Sustainable Energy 2017;36:943–52](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1140).

1. [Qu Z, Zhang K, Wang J, Zhang W, Leng W. A hybrid model based on ensemble empirical mode decomposition and fruit fly optimization algorithm for wind](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1145) [speed forecasting. Adv Meteorol 2016;2016](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1145).
2. [Han S, Li J, Liu Y. Tabu search algorithm optimized ANN model for wind power](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1150) [prediction with NWP. Energy Procedia 2011;12:733–40.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1150)
3. [Li L-L, Zhao X, Tseng M-L, Tan RR. Short-term wind power forecasting based on support vector machine with improved dragonfly algorithm. J Cleaner Prod 2020; 242:118447](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1155).
4. [Ding M, Zhou H, Xie H, Wu M, Liu K-Z, Nakanishi Y, et al. A time series model based on hybrid-kernel least-squares support vector machine for short-term wind](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1160) [power forecasting. ISA Trans 2021;108:58–68.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1160)
5. [Lu P, Ye L, Zhong W, Qu Y, Zhai B, Tang Y, et al. A novel spatio-temporal wind power forecasting framework based on multi-output support vector machine and](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1165) [optimization strategy. J Cleaner Prod 2020;254:119993](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1165).
6. [Hong WC. Electric load forecasting by seasonal recurrent SVR (support vector](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1170) [regression) with chaotic artificial bee colony algorithm. Energy 2011;36:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1170)

[5568–78](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1170).

1. [Zhang Y, Le J, Liao X, Zheng F, Li Y. A novel combination forecasting model for](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1175) [wind power integrating least square support vector machine, deep belief network,](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1175) [singular spectrum analysis and locality-sensitive hashing. Energy 2019;168:](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1175)

[558–72](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1175).

1. [Wang J, Heng J, Xiao L, Wang C. Research and application of a combined model](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1180) [based on multi-objective optimization for multi-step ahead wind speed](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1180) [forecasting. Energy 2017;125:591–613.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1180)
2. [Qu Z, Zhang K, Mao W, Wang J, Liu C, Zhang W. Research and application of ensemble forecasting based on a novel multi-objective optimization algorithm for](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1185) [wind-speed forecasting. Energy Convers Manage 2017;154:440–54.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1185)
3. [Liu H, Duan Z, Li Y, Lu H. A novel ensemble model of different mother wavelets](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1190) [for wind speed multi-step forecasting. Appl Energy 2018;228:1783–800](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1190).
4. [Liu H, Chen C. Multi-objective data-ensemble wind speed forecasting model with stacked sparse autoencoder and adaptive decomposition-based error correction.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1195) [Appl Energy 2019;254:113686](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1195).
5. [Zhang S, Liu Y, Wang J, Wang C. Research on combined model based on multi- objective optimization and application in wind speed forecast. Appl Sci 2019;9: 423](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1200).
6. [Jiang P, Liu Z. Variable weights combined model based on multi-objective optimization for short-term wind speed forecasting. Appl Soft Comput 2019;82: 105587](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1205).
7. [Li R, Jin Y. A wind speed interval prediction system based on multi-objective](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1210) [optimization for machine learning method. Appl Energy 2018;228:2207–20](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1210).
8. [Jiang P, Yang H, Heng J. A hybrid forecasting system based on fuzzy time series](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1215) [and multi-objective optimization for wind speed forecasting. Appl Energy 2019;](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1215)

[235:786–801](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1215).

1. [Zhang C, Wei H, Xie L, Shen Y, Zhang K. Direct interval forecasting of wind speed](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1220) [using radial basis function neural networks in a multi-objective optimization](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1220) [framework. Neurocomputing 2016;205:53–63.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1220)
2. [Ak R, Li Y-F, Vitelli V, Zio E. Adequacy assessment of a wind-integrated system using neural network-based interval predictions of wind power generation and load. Int J Electr Power Energy Syst 2018;95:213–26](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1225).
3. [Shrivastava NA, Lohia K, Panigrahi BK. A multiobjective framework for wind](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1230) [speed prediction interval forecasts. Renew Energy 2016;87:903–10.](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1230)
4. [Niu T, Wang J, Lu H, Du P. Uncertainty modeling for chaotic time series based on](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1235) [optimal multi-input multi-output architecture: application to offshore wind](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1235) [speed. Energy Convers Manage 2018;156:597–617](http://refhub.elsevier.com/S0306-2619(21)00836-9/h1235).