IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received September 26, 2019, accepted January 7, 2020, date of publication January 14, 2020, date of current version January 23, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2966641

Electric Load Forecasting Use a Novelty Hybrid Model on the Basic of Data Preprocessing Technique and Multi-Objective Optimization Algorithm

HE BO[1], YING NIE[2], AND JIANZHOU WANG[2]
[1]Postdoctoral Research Station, Dongbei University of Finance and Economics, Dalian 116025, China
[2]School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China

Corresponding author: Ying Nie (nieyingaa@163.com)

This work was supported by the Social Science Foundation of Liaoning Province in China under Grant L18CTJ003.

ABSTRACT Power load forecasting has an influence of great signification on improving the operational efficiency and economic benefits of the power grid system. Aiming at improving forecast performance, a substantial number of load forecasting models are proposed. However, these models have disregarded the limits of individual prediction models and the necessity of data preprocessing, resulting in poor prediction accuracy. In this article, a novelty hybrid model which combines data preprocessing technology, individual forecasting algorithm and weight determination theory is presented for obtaining higher accuracy and forecasting ability. In this model, an effective data preprocessing method named SSA is adopted to extract the load data characteristics and further improve the prediction performance. In addition, a combined forecasting mechanism composed of BP, SVM, GRNN and ARIMA is successfully established using the weight determination theory, which exceeds the limits of individual prediction models and comparatively improves prediction accuracy. And the thought of combine linear and nonlinear model together can further take the advantage of two kinds of models to forecast power load more effectively. To assess the validity of the combined model, four datasets of 30-minutes power load from Australia are selected for research. The experimental results show that the established model not only has obvious advantages over other individual models, but also can be applied as an available technology for electrical system programming.

INDEX TERMS Power load forecasting, hybrid model, data preprocessing technology, weight determination theory, MOEA/D optimization algorithm.

I. INTRODUCTION

The electricity supply system is a kind of production and consumption system consisting of electricity generation, substation, transmission and distribution [1]. Currently, the power system has become one of the important signs for national economic development [2], so that it is very necessary and important to predict the relevant indicators of the power system [3]. The electric load forecast plays an important role for the economic and safe operation of the power system [4], as several decisions of great importance for the economic operation is based on the prediction of the electric load, such as fuel distribution, short-term maintenance, and emergency [5]. Therefore, all countries of the world are looking for effective methods of predicting the electric load [6]. With the rapid development of Australian economy and the increasing demand for electricity, the improvement of power load forecasting should be highly valued [7], [8]. In addition, reliable power load forecasting can reduce energy consumption and environmental pollution [9]. Therefore, accurate and reliable forecasting for the planning of electricity generation and the adaptation of energy policy and structure is conducive [10].

Electric load forecasting, which combines Very Short-Term Load Forecast (VSTLF), Short-Term Load Forecast (STLF), and Long-Term Load Forecast (LTLF), is significant for the safety of power system, particularly in power market [11]. The VSTLF as well as STLF are necessary basis for electrical grid dispatching. VSTLF primarily pay close attention to load prediction within one hour [12]. One of the most significant goals of VSTLF is to formulate

The associate editor coordinating the review of this manuscript and approving it for publication was Ziyan Wu.

the daily power plan [13]. Moreover, VSTLF can also be applied to cold standby and rotary standby [14]. At the same time, VSTLF and STLF can be applied to draw up the power grid overhaul program [15]. A number of articles have achieved well consequences for VSTLF [16], yet power grid dispatching and management is still a difficult task [17]. Compared with VSTLF, STLF is more widely used and more difficult. As a result, diminishing the error of STLF is one of the advisable methods to enhance the operation level of electronic system [18]. Accurate prediction of electric load can save a lot of time to manage smart grid ahead of time and avoid major changes [19].

From the literature review, it can be clearly seen that the research of power prediction still occupies a significant place in the power system, which can impact the management and operation of the smart grid greatly. Consequently, it is essential to develop an electric load prediction model with good performance. For the sake of achieving the precise and steady STLF, a great quantity of methods are adopted, where statistical method, machine learning method and hybrid method are the main methods.

Compared with other models, statistical methods such as ARIMA and ARMA have good real-time performance [20], [21]. Feinberg and Eugene [5] predict the load at multiple point of time in a 1-24h time scope by fractional ARIMA model. The experiment consequences show the hybrid model can enhance the prediction ability. Steinherz *et al.* [22] evaluate the situation of power market by the method which combined ARMA and GARCH (Generalized AutoRegressive Conditional Heteroskedasticity). Models mentioned above are the first choice for evaluating electric load in power market. Pappas *et al.* [23] proposed a novel approach for load prediction of power demand based on Multi-model partition theory. The experimental results show that the model has good load forecasting effect. However, this method cannot well reflect the complex non-linear relationship between load and various random factors, such as hours, days, weeks, months and social events, which may lead to unpredictable changes in electricity demand.

Considering the outstanding ability of nonlinear systems, neural networks have been widely used in STLF. In previous studies, different artificial neural network (ANN) models have been verified, regardless of load profiling. Chaturvedi *et al.* [24] used generalized neural network (GNN) to predict electric load, which overcomes the shortcomings of artificial neural network. Combining GNN with fuzzy system and adaptive genetic algorithm, a most effective prediction model is proposed. Lou and Dong [25] uses random fuzzy variables to model the fuzzy random uncertainty of power load prediction. On this basis, a new stochastic fuzzy neural network (RFNN) is proposed, which has good application prospects in micro grid and small power system.

In order to improve predictability and stability, neural networks are often combined with other technologies to establish a hybrid model base on the systems characteristics. A two-stage forecasting system is performed in [26] by Zhang *et al.* based on the data preprocessing approach, improved multi-objective optimization algorithm named IMODA, error correction and nonlinear ensemble strategy, which observably prompts wind speed forecasting capacity. Mohammadi *et al.* [27] proposed a new prediction model based on a new feature selection algorithm using Pearson's correlation and hybrid forecast engine based on combined improved Elman neural network (IENN) and novel shark smell optimization (NSSO) algorithm. The proposed forecast engine is combined with novel shark smell optimization to increase the prediction accuracy. A multi-layer bidirectional recurrent neural network model based on LSTM and GRU is proposed by Tang *et al.* [28] to forecast short-term power load and the experimental result shows that the proposed method is superior to the competition winner. Ghadimi *et al.* [29] proposes a hybrid forecast strategy including novel feature selection technique (MIT-MIT), and a complex forecast engine based on RNN and ENN optimized by chaotic binary shark smell optimization (CBSSO) algorithm, and the results validate the effectiveness of the proposed method.

An overview of the previous work shows the forecasting approaches mentioned above possess certain inherent flaws. The weaknesses of these approaches are summing up below:

(1) Traditional statistical algorithms cannot deal with the prediction of high noise and fluctuations, irregular and nonlinear trends, nor can they deal with the characteristics of the power load series data, which are mostly restricted by the priori hypothesis of linear forms among time series. Moreover, these methods in practice require a lot of historical data for the load forecast, and they depend heavily on data. If the primary data suddenly change as a result of environmental or social factors, the forecast error will expend sharply.

(2) Unlike other methods, artificial intelligence algorithm can effectively find the implicit nonlinear relationship between historic data, which has been extensively investigated and used in resolving complex relations and precisely predict. Nevertheless, a number of shortcomings still exist in artificial intelligence methods, such as a slight fall into the local optimal, over-adapted and comparatively low rate of convergence.

(3) Conventional prediction models don't attach importance to the inevitability and cannot always attain higher prediction ability and fulfill the requests of the time series prediction. As a result, the forecasting methods mentioned above cannot always capture load trends due to the inevitable deficiencies of individual models and cannot always be applied in all cases.

Based on this, a new combinatorial model is proposed. It associates data preprocessing technology, multi-objective optimization algorithm, weight determined theory [30], and linear prediction models called BPNN [31], SVM [32], and GRNN [33], as well as nonlinear model ARIMA [34]. It effectively utilized the advantage of prediction algorithm, and has been further improved. To be more specific, according to the decomposition and integration policy, the primary load sequence is decomposed into the time series after filtering to effective reconstruct the high frequency noise signal. Afterwards some algorithms are applied to predict the power load after processing. After that, the new decision making method was developed based on the evolution calculation technology of the collective intelligence and the one-time exit strategy, and the final prediction result was obtained by integrating the single model. As is known to all, this advanced computational technology is successfully applied to several mixed models, first used to improve prediction accuracy and calculate the weights of every model, applying to the combination prediction model.

The leading contributions and innovations of the research are as follows:

(1) ***Base on decomposition and integration policy, a data preprocessing technology named SSA is adopted to diminish the adverse impact of high frequency noise and select the major features from data.*** By filtering the original power load series, the uncertainty and irregularity of load data can be reduced and the electric load prediction performance can be effectively improved.

(2) ***A new evolutionary computation method based on swarm intelligence MOEA/D and a decision-making right method is proposed.*** In order to calculate the weight of selected model, the "leave-one-out" method was applied to preserve the data set, and the weight of each model are selected based on the multi-objective optimization algorithm.

(3) ***Based on four kinds of individual methods, including nonlinear models BPNN, SVM, GRNN and linear model ARIMA, the combined model has immensely enhanced the ability of electric load forecasting.*** The established model effectively takes not only the advantages of linear model but also the advantages of nonlinear model, and conquers the limitations of low accuracy and instability of traditional single model.

(4) ***A more scientific and comprehensive prediction and assessment of the combined model developed in this work.*** The evaluation system uses three multi-step prediction experiments, five performance indicators, and three discussions, which provides an effective evaluation for the prediction accuracy of the model.

(5) ***The new combined model developed provides strong technical support for smart grid dispatching management.*** According to the load data of several sites in large power system, the model is simulated and tested. The results show that the model can effectually enhance the ability of load forecasting base on the traditional prediction model.

The construction of this article is introduced below. Section II expounds methods used in the model, which include data preprocessing technology, multi-objective optimization algorithm, and established combination model. To verify the forecasting ability of the established model, three certain experiments were implemented and display in Section III. Particularly, sections IV introduce the dataset, performance indicators, and test methods used in this study, as well as the experimental results of the established model and other models. In order to further prove the accuracy and validity of this new combination model, section V discusses it in detail. Finally, section VI gives important results and conclusions.

NOMENCLATURE

GWO	Gray Wolf algorithm
MOGWO	Multi-objective Gray Wolf algorithm
RBF	Radial basis function neural network
GRNN	Generalized Regression Neural Network
WNN	Wavelet neural network
EMD	Empirical Mode Decomposition
EEMD	Ensemble Empirical Mode Decomposition
CEEMD	Complementary Ensemble Empirical Mode Decomposition
SSA	Singular Spectrum Analysis
WAVEN	Wavelet transform Preprocessing Method
IMFs	Intrinsic Mode Functions
ZDT	Test functions for multi-objective algorithm
BP	Back Propagation Neural Network
ENN	Evolutionary neural network
ELM	Extreme Learning Machine Neural Network
SVR	Support Vactor Regerssion
SSE	Sum of square error
RMSE	Root mean square error
MAE	Mean absolute error
MSE	Mean squared error
MAPE	Mean absolute percentage error
DC	Directional change
MOPSO	Multi-objective Particle Swarm Optimization
MODA	Multi-objective Dragonfly algorithm

II. METHODS

In this section, an effective data preprocessing method called singular spectrum analysis (SSA) used in the composition model is introduced, and then an optimized algorithm named the multi-objective evolutionary algorithm based on decomposition (MOEA/D) is discussed in detail.

A. SINGULAR SPECTRUM ANALYSIS

Singular spectrum analysis (SSA) is a classic technology for researching time series [35]. It is mostly used to resolve problems of trend extracting or quasi-periodic component extracting and noise suppressing [36]. Its analytical performance is very impressive [37]. The standard content of SSA is divided into four parts [38]: (1) embedding, (2) singular value decomposition, (3) grouping and (4) diagonal averaging [39].

1) EMBEDDING

The goal of SSA is to decompose a sequence of time series $X = (x_1, x_2, \cdots, x_N)$ into the aggregation of a kind of time sequences to discern the primary sequence factor, which include the trend, noise and period. SSA consists of two complementary phases, named decomposing and restructuring respectively [40].

Sequence $X = (x_1, x_2, \cdots, x_N)$ is transformed into a sequence $Y = (y_1, y_2, \cdots, y_K)$, which can be specified as

$$X = (x_1, x_2, \cdots, x_N) \rightarrow Y = (y_1, y_2, \cdots, y_K) \tag{1}$$

where $Y_i = (x_i, x_{i+1}, \cdots, x_{i+L-1})^T \in R^L$ and $L \in [2, N]$ is the window length. Besides, K is the length of Y and it calculated as K = N-L + 1. As a result, the consequences of the mapping operation are the trajectory matrix $Y = [Y_1, Y_2, \cdots, Y_K] = (y_{ij})_{i,j=1}^{L,K}$, which could be indicated as

$$Y = \begin{bmatrix} x_1 & x_2 & \cdots & x_K \\ x_2 & x_3 & \cdots & x_{K+1} \\ \cdots & \cdots & & \cdots \\ x_L & x_{L+1} & \cdots & x_N \end{bmatrix} \tag{2}$$

The Y matrix is the Hankel matrix [41], and the matrix elements are equivalent to the diagonal.

2) SINGULAR VALUE DECOMPOSITION (SVD)

The task of this part is decomposing the trajectory matrix Y. If an orthogonal matrix $U = [u_1, u_2, \cdots, u_{S_r}] \in R^{S_r \times n_c}$ and $V = [v_1, v_2, \cdots, v_{S_r}] \in R^{S_r \times n_c}$ are given, and a matrix $B \in R^{S_r \times n_c}$, the following formula can be defined:

$$\begin{aligned} U^T BV &= \Lambda = diag(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \cdots, \sqrt{\lambda_{pm}}) \\ &= \begin{bmatrix} \sqrt{\lambda_1} & & & \\ & \sqrt{\lambda_2} & & \\ & & \cdots & \\ & & & \sqrt{\lambda_2} \end{bmatrix} \end{aligned} \tag{3}$$

As a result, $B = U \Lambda V^T$ where $p_m = \min(s_r, n_c)$ and $\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \cdots \geq \sqrt{\lambda_{pm}}$ represents the singular value of B. v_i represent the singular vector of the singular value $\sqrt{\lambda_i}$ which satisfy $Bv_i = \sqrt{\lambda_i}$ and $B^T u_i = \sqrt{\lambda_i} v_i (i = 1, 2, \ldots, p_m)$. The SVD of B is the factorization of B.

$\lambda_1, \lambda_2, \cdots, \lambda_L (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L \geq 0)$ is assumed to be the eigenvalues. The orthogonal systems of eigenvector are $U_1, U_2, \cdots, U_L$. Where d represents the subscript of the maximum eigenvalue, that is, the rank of the matrix Y, and it can be indicated as

$$d = \max(i, \lambda_i > 0) = rank Y \tag{4}$$

Supposing that $V_i = Y^T U_i / \sqrt{\lambda_i} (i = 1, 2, \ldots, d)$, the SVD of trajectory matrix Z is expressed below

$$Y = Y_1 + Y_2 + \cdots + Y_d \tag{5}$$

In which $Y_i = \sqrt{\lambda_i} U_i V_i^T$ and the rank of Y_i is 1, as a result the Y_i is an elementary matrix. Where U_i is the left eigenvector and V_i represent the right eigenvector, respectively. $\sqrt{\lambda_i} (i = 1, 2, \ldots, d)$ denote the singular value of matrix Y. The $\left\{\sqrt{\lambda_i}\right\}$ set represents the spectrum of Y. U_i, V_i and $\left\{\sqrt{\lambda_i}\right\}$ are the characteristic loop $(\left\{\sqrt{\lambda_i}\right\}, U_i, V_i)$.

If matrix $\sum_{i=1}^{r} Y_i$ fits the trajectory $Y(||Y - Y^{(r)}||)$, the matrix has the minimum value. And SVD is the optimal.

3) GROUPING

The elementary matrices X_i into different groups and calculate summation of matrices in each group. If $I = \{i_1, i_2, \ldots, i_p\}$, the matrix X_I and group I can be described as $Y_I = Y_{i1} + Y_{i2} + \cdots + Y_{ip}$. The spilled of the set of indices $I = 1, 2, \ldots, d$ is partitioned into the disjoint subsets $I_1, I_2, \cdots, I_m$. The practice is called grouping and it can be represented as:

$$Y_I = Y_{i1} + Y_{i2} + \cdots + Y_{im} \tag{6}$$

4) DIAGONAL AVERAGING

The task of this part is designed to convert the guaranteed groups of last stage into a series of length N. Supposing that Y represents an matrix with the dimension of L × K, $L^* = \min(L, K)$ and $K^* = \max(L, K)$. If $L < K$, $y_{ij}^* = y_{ij}$, or else $y_{ij}^* = y_{ji}$, where y_{ij}^* is the element of matrix Y. The sequence $RC = (rc_1, rc_2, \ldots, rc_N)$ restructured can be expressed as

$$rc_{n_c} = \begin{cases} \dfrac{1}{n_c + 1} \sum_{s_g=1}^{n_c+1} y_{s_g, n_c - s_g + 2}^*, & 1 \leq n_c \leq L^* \\ \dfrac{1}{L^*} \sum_{s_g=1}^{L^*} y_{s_g, n_c - s_g + 2}^*, & L^* \leq n_c \leq K^* \\ \dfrac{1}{N - n_c} \sum_{s_g=n_c-s_g+2}^{n_c - K^* + 1} y_{s_g, n_c - s_g + 2}^*, & K^* \leq n_c \leq N \end{cases} \tag{7}$$

B. MULTI-OBJECTIVE EVOLUTIONARY ALGORITHM BASED ON DECOMPOSITION

Recently, decomposition-based multi-objective evolutionary algorithm (MOEA/D) proposed by Zhang and Li [42] has attracted more and more researchers' interest because of its concise and effective characteristics, and many theoretical and practical achievements have emerged. The MOEA/D algorithm is introduced below.

A multi-objective optimization problem (MOP) with M objectives and N decision variables can be expressed as follows:

$$\begin{aligned} &Minimize\ F(x) = (f_1(x), f_2(x), \cdots, f_m(x)) \\ &Subject\ to\ x \in \Omega \end{aligned} \tag{8}$$

where $\Omega \in R^n$ is feasible region of decision space, and the decision vector $x = \{x_1, x_2, \cdots, x_n\} \in \Omega$ is a candidate solution of MOP for optimization problem. Here, the objective function $F(x) : x \rightarrow R^m$ includes M conflicting object functions with continuous real values $f_1(x), f_2(x), \cdots, f_m(x)$, which R^m represents the target space.

The pareto dominance relation of individuals is as follows: if there are decision vectors U and V, and satisfy the following two conditions at the same time, we call U dominance V:

(1) For all objectives, U is no worse than V, that is $f_i(u) \leq f_i(v)$, $i \in \{1, \ldots, m\}$.

(2) There is at least one goal in which U is better than V, that is $\exists j \in \{1, \ldots, m\}$, make $f_i(u) \leq f_i(v)$. In this case V is said to be dominated by U, which can denoted by $u \succ v$, and among $\succ$ it are dominant relations.

For the multi-objective optimization problem MOP which optimizes several conflicting objectives at the same time,

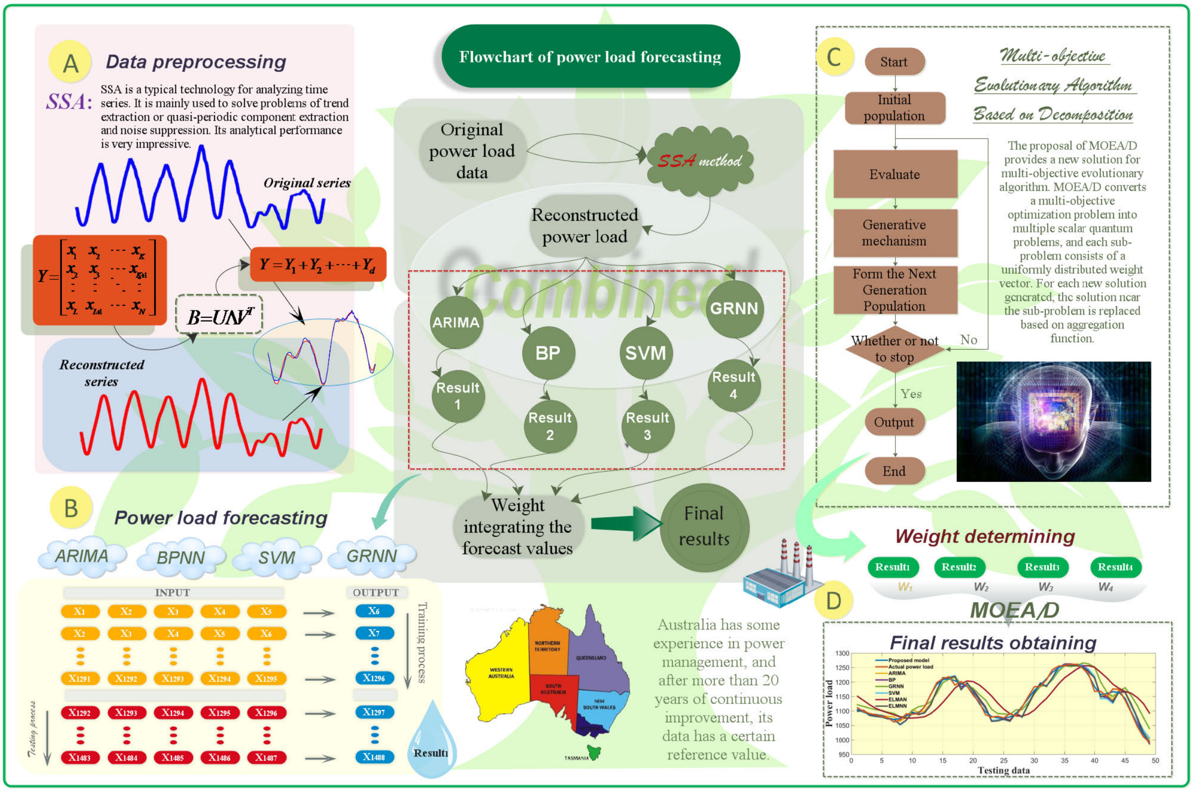


FIGURE 1. The flowchart of power load forecasting using hybrid model.

there is no unique solution to achieve the optimization of all objectives at the same time. There is only one best set of compromise solutions, that is, the non-dominated (not dominated by all other solutions). The optimal compromise solution set of multiple objectives is called the Pareto optimization solution set of MOP. The value of Pareto optimization solution in decided space and target space is defined as Pareto solution set (PS) and Pareto frontier (PF) [43], respectively.

MOEA/D algorithm decomposes the multi-objective optimization problem MOP into a series of sub-problems represented by weight vectors, and then uses evolutionary algorithm to optimize these sub-problems simultaneously. The algorithm has strong search ability for continuous optimization, combinatorial optimization and PS complex problems.

If a multi-objective optimal problem similar to (8) and a weight vector $\lambda = (\lambda_1, \cdots, \lambda_m)$ are given, and the given weight vector satisfies $\sum_{i=1}^{m} \lambda_i = 1, \lambda_i \geq 0, i = 1, 2, \ldots, m$, MOEA/D based on Tchebycheff decomposition uses this weight vector to optimize a MOP into several sub-problems by the following methods.

$$\min_{x\in\Omega} g^{tc}(x|\lambda, z^*) = \min_{x\in\Omega} \max_{1\leq i\leq m} \{\lambda_i|f_i(x) - z_i^*\} \tag{9}$$

where $z^* = (z_1^*, z_2^*, \cdots, z_m^*)$ (*i.e.*, $z_i^* < \min\{f_i(x)|x \in \Omega\}, 1, 2, \ldots, m$) is the ideal point. By solving multiple sub-problems with different weight vectors in (9), the Pareto optimal solution set [44] with good diversity can be obtained. The complete PF of the MOP consists of the optimal solutions of all the sub-problems, in which the sub-problems with different weight vectors are searched in different regions of the complete PF. Therefore, the MOEA/D can change its search direction by changing the weight vector of the sub-problem.

In the algorithm MOEA/D, the population is made up of the optimal solution of the sub-problem currently found. Each sub-problem maintains a list of neighbors, which preserves sub-problems with weight vectors similar to the sub-problem. Therefore, under the assumption of continuity, two neighbor sub-problems should have similar optimal solutions. In each generation of MOEA/D, each sub-problem is optimized using only the information of its neighboring sub-problems.

The pseudo code of MOEA/D is described in Appendix A.

III. ESTABLISHED HYBRID FORECASTING MODEL

According to the mythologies discussed above, the established hybrid forecasting model is mainly established by four steps, including data preprocessing step, individual models forecasting step, combined model building step, and electric load forecasting step. The method of data preprocessing technique, combined with weight decided method, and combination single model, can effectively enhance the accuracy of load forecasting. The proposed step of our established model shown in **Fig. 1** is statement as follows:

❶ Data preprocessing module.

The SSA preprocessing technique is selected in the established model to obtain a reconstructed sequence by refining and identifying the period and vibration parts of the original signal. Through this method, a time series with less noise signal and random volatility can receive to apply in the following forecasting steps.

❷ Optimization module.

The first step is individual model forecasting. Four individual forecasting models—BPNN, SVM, and GRNN, as well as ARIMA—are selected to carry on multi-step forecasting, respectively. And four forecasting results are obtained by this operation. There are there neural network for nonlinear prediction and one linear prediction, and the models of both the linear and nonlinear prediction is excellent in predicting power load.

The second step is individual result weighted combining. For the sake of obtaining the weight coefficients of each model, a kind of decision-making weight method on the basic of MOEA/D algorithm and ''left-one-out'' policy is proposed to gain the optimal consequences. Say concretely, the last three days of the training set are retained to obtain the weights of selected models. It is remarkable that when the algorithm reaches the maximum iteration number or the minimum fitness function value, it will stop. In addition, the iteration times and dimensions of MOEA/D are set to 200 and 4, and the weight coefficients are between -1 and 1. Then, according to the weight coefficients of each model, the prediction consequences of each model are combined for getting the final wind speed prediction results.

❸ Forecasting module.

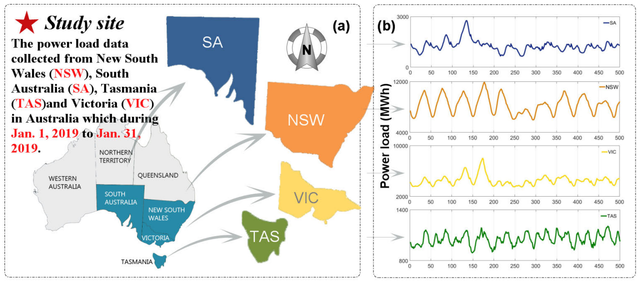
In this study, a combination model has been developed to predict electric load on the basic of historic data. The multi-step prediction method is used to assess the forecast ability of the combination model. In our established model, three-step forecasting is adopted to verify the predictive effect of the model.

❹ Evaluation module.

Six assessment indices, named MAE (Mean Absolute Error), the RMSE (Root Mean Square Error), the MAPE (Mean Absolute Percentage Error), the MSE (Mean Squared Error), and the DC (Directional Change), as well as the R^2 (Pearson's correlation coefficient), were adopted to implement the evaluation of the designed prediction architecture quantitatively.

IV. EXPERIMENTS AND ANALYSIS

Applications of the proposed hybrid model and several comparison models are shown in this subsection, and the comparation is divided into three experimental presentations. The operation environment of all experiments is: 2.40GHz CPU, 4.00GB RAM, Windows7 and Matlab R2018A. Considering the random factors, in order to ensure that the final results are reliable, 10 experiments are conducted each time, and then take the average value respectively.



(c)

Dataset	Numbers	Power load of four datasets from Jan. 1, 2019 to Jan. 31, 2019			
		Max	Min	Mean	Std.
Site 1	1488	13701.0	5973.3	8984.0	1731.9
Site 2	1488	2994.3	647.6	1483.9	474.5
Site 3	1488	1338.3	716.2	1076.3	81.4
Site 4	1488	9281.2	3320.3	5075.3	1206.1

FIGURE 2. Description of observations in four datasets. (a) Location of the study sites. (b) The original power load series from Jan. 1 to Jan. 31, 2019. (c) The statistical measures for the power load.

TABLE 1. Data structure of four selected datasets in Australia.

Dataset	Samples	Numbers	Statistical Indicator			
			Max	Min	Mean	Std.
Site 1	All samples	1488	13701.0	5973.3	8984.0	1731.9
	Training	1296	13545.0	5973.3	8897.6	1694.3
	Testing	192	13701.0	6719.8	9567.6	1869.4
Site 2	All samples	1488	2994.3	647.6	1483.9	474.5
	Training	1296	2994.3	647.6	1500.5	493.4
	Testing	192	2166.3	768.4	1371.6	295.7
Site 3	All samples	1488	1338.3	716.2	1076.3	81.4
	Training	1296	1338.3	716.2	1072.5	81.8
	Testing	192	1261.9	967.4	1101.5	73.9
Site 4	All samples	1488	9281.2	3320.3	5075.3	1206.1
	Training	1296	9281.2	3320.3	5050.7	1212.2
	Testing	192	8292.4	3640.2	5241.2	1153.2

Power load of New South Wales, South Australia, Tasmania and Victoria from January 1, 2019 to January 31, 2019.

A. DATASET DESCRIPTION

This study collects four datasets based on Australian power load data, which are from the 30-minute load data of National Electric Market (NEM) of AEMO Company. Australia has some experience in power management, and after more than 20 years of continuous improvement, its data has a certain reference value. We selected four typical areas, New South Wales (NSW), South Australia (SA), Tasmania (TAS), and Victoria (VIC). These four regions are typical areas of the National Electricity Market (NEM), whose Scheduled Demand are 8,821.74 MW, 1,053.28 MW, 1,263.03 MW and 5,250.73 MW, respectively. A brief diagram of the select set is presented in **Fig. 2**. In addition, a kind of evaluate indices of load data sampling for four datasets, which include minimum, maximum, average, and standard deviation, are shown in **Table 1**.

In order to evaluate the prediction effect of the new established model, 30-minute load data were selected from the National Electricity Market (NEM) for 31 days from January 1, 2019 to January 31, 2019. We divide each dataset into two groups: training set and testing set. More concretely, from January 1, 2019 to January 27, 2019, the first 28 days

TABLE 2. Six error metrics.

Metric	Definition	Equation
MAPE	Mean Absolute Percentage Error	$\mathbf{MAPE}=\frac{1}{N}\sum_{i=1}^{N}\lvert\frac{x_i-\hat{x}_i}{x_i}\rvert\cdot 100\%$
MAE	Mean Absolute Error	$\mathbf{MAE}=\frac{1}{N}\sum_{i=1}^{N}\lvert x_i-\hat{x}_i\rvert$
RMSE	Root Mean Square Error	$\mathbf{RMSE}=\sqrt{\frac{1}{N}\cdot\sum_{i=1}^{N}(x_i-\hat{x}_i)^2}$
MSE	Mean Squared Error	$\mathbf{MSE}=\frac{1}{N}\sum_{i=1}^{N}(x_i-\hat{x}_i)^2$
DC	Directional Change	$\mathbf{DC}=\frac{100}{N-1}\sum_{t=1}^{N-1}a_t\ ,\ a_t=\begin{cases}0, & otherwise\\ 1, & if((x(t+1)-x(t))(\hat{x}(t+1)-x(t)))>0\end{cases}$
$\boldsymbol{R^2}$	Pearson's correlation coefficient	$\boldsymbol{R^2}=\frac{\sum_{i=1}^{N}(y_i-\bar{y})(\hat{y}_i-\bar{\hat{y}})}{\sqrt{\sum_{i=1}^{N}(y_i-\bar{y})^2(\hat{y}_i-\bar{\hat{y}})^2}}$

Note: $\hat{x}_i$ is the i-th forecasting value, N is the total number, $\bar{\hat{y}}$ and $\bar{y}$ are the average of the forecasting and observed values, respectively.

included 1296 data points as training samples for power load prediction. We retain the last four days of the training set to decide the weights of each model. The remaining four days from January 28, 2019 to January 31, 2019 included 192 data points as testing set. The same rolling forecasting mechanism is adopted in training set and the testing set. The results of one, two and three-step prediction are output, and the rolling forecasting mechanism is applied in four-day load forecasting. **Fig. 2** presents the detail of data construction of the established composite model.

B. EVALUATION CRITERIA: MAPE, MAE, RMSE, MSE, DC, R^2

Many performance indicators have been studied and applied to assess the predictive effect of different models. More diversity of evaluation indicators can better evaluate the quality of the proposed model. Therefore, as shown in **Table 2**, the predictive capacity of the established model proposed in this study is evaluated using multiple error indicators, which including the Mean Absolute Error (MAE) [45], the Root Mean Square Error (RMSE) [46], the Mean Absolute Percentage Error (MAPE) [47], the Mean Squared Error (MSE) [48], and the Directional Change (DC), as well as the Pearson's correlation coefficient (R^2) [49].

In detail, RMSE and MAE are adopted to estimate the mean value of the expected value and the actual value, so as to prevent the mutual cancelling of positive and negative errors in prediction happening. MAPE is the mean value of the absolute error, which is the most common indicator to reflect the reliability of model. MSE is used to supplement the estimation of the average size between the expected value and the actual value. On the other hand, DC reflects the predicted direction of motion or turning point. With regard to the above five indices, the smaller the values of MAPE, MAE, RMSE and MSE refer to the greater the model performance. DC is the contrary. Specially, Pearson's correlation coefficient judges the correlation degree between the forecasting series and the original series and it is applied in the discussion to further judge the model effectiveness.

To better explain the performance indicators, definitions and formula of five indicators are listed in **Table 2**.

C. DIEBOLD-MARIANO TEST

To verify that the developed hybrid model has better prediction ability than other compared models, an effective verification method called DM test proposed by Diebold FX and Mariano RS [50] is adopted. The theory of DM test is introduced first.

Considering the significance level α, zero hypothesis H_0 shows that the predictive performance of the developed hybrid model and the compared model are not significantly different. The meaning of H_1 is contrast with H_0. The relevant formulas can be expressed as follows:

$$H_0 : E[L(err_i^1)] = E[L(err_i^2)] \tag{10}$$

$$H_1 : E[L(err_i^1)] \neq E[L(err_i^2)] \tag{11}$$

In the formula, L represents the loss function of prediction error. Error_i^1 and Error_i^2 are the error sequence predicted by selected model.

In addition, the statistics of DM test can be defined in the following ways:

$$DM = \frac{\sum_{i=1}^{n}(L(err_i^1) - L(err_i^2))}{\sqrt{S^2/n}} s^2 \tag{12}$$

In which, S^2 is the estimate of the variance of $d_i = L(err_i^1) - L(err_i^2)$. Assuming a certain significance level α, the obtained value DM is in comparison with that of $z_{\alpha/2}$. Once DM statistics exceed the interval [-$z_{\alpha/2}$, $z_{\alpha/2}$], H_0 can be rejected. This shows the predictive performance of the established model and that of the comparative model are significantly different, which means that H_1 will be accepted.

TABLE 3. Prediction abilities of the established model and other SSA-based models.

Dataset	Model	MAPE (%)			MAE			RMSE			MSE			DC (%)		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site1 NSW	SSA-ARIMA	0.4526	2.3464	4.6742	42.9450	224.7733	444.3365	54.3809	280.6294	518.9202	2957.2832	78752.8346	269278.1383	97.3822	95.7895	95.2381
	SSA-BPNN	0.5045	1.1539	1.5910	48.2283	107.4331	148.0065	64.9853	143.0682	195.8262	4223.0928	20468.5234	38347.8820	96.8586	88.4211	79.8942
	SSA-GRNN	1.4966	2.1497	2.9563	147.6534	210.5202	290.4661	203.1603	275.4202	366.7030	41274.0882	75856.2695	134471.1043	91.0995	91.5789	88.8889
	SSA-SVM	0.7255	1.5065	2.3659	69.7167	144.6228	225.4907	92.2153	197.4440	305.3867	8503.6591	38984.1354	93261.0551	94.7644	89.4737	87.8307
	Proposed Model	**0.4514**	**1.0903**	**1.2577**	**42.8534**	**103.1178**	**121.6497**	**54.3128**	**140.9177**	**170.5751**	**2949.8836**	**19857.7990**	**29095.8541**	**96.3351**	**92.1053**	**88.8889**
Site2 SA	SSA-ARIMA	1.0163	3.5520	6.3665	13.6469	48.7132	86.4533	18.9296	59.8857	105.1676	358.3303	3586.2986	11060.2195	93.1937	86.3158	76.7196
	SSA-BPNN	1.0306	2.1388	3.6067	13.8767	28.9330	48.5429	19.5287	36.7757	62.0926	381.3717	1352.4502	3855.4861	92.6702	85.7895	84.1270
	SSA-GRNN	3.8782	5.6826	7.6613	52.6740	78.4061	106.0495	64.4060	98.1571	135.8120	4148.1377	9634.8096	18444.8972	89.5288	85.7895	80.9524
	SSA-SVM	1.4163	3.0062	6.3699	18.7442	40.8630	85.6046	25.1162	52.5920	106.4923	630.8240	2765.9220	11340.6201	91.6230	84.7368	71.9577
	Proposed Model	**0.9970**	**2.1160**	**3.5597**	**13.4636**	**28.8656**	**48.0672**	**19.0909**	**36.9727**	**61.9447**	**364.4606**	**1366.9831**	**3837.1501**	**91.0995**	**81.0526**	**62.9630**
Site3 TAS	SSA-ARIMA	0.5871	1.9259	2.7610	6.4807	21.2882	30.5548	8.5117	27.6736	38.1816	72.4496	765.8302	1457.8354	86.9110	82.1053	83.0688
	SSA-BPNN	0.5758	1.5917	1.5991	6.3555	17.4251	17.5133	8.2954	22.0543	22.4346	68.8135	486.3917	503.3096	88.4817	75.2632	75.1323
	SSA-GRNN	1.5299	2.3386	2.9360	16.8945	25.7630	32.2374	21.5945	32.1621	40.2896	466.3213	1034.4012	1623.2534	86.9110	77.8947	80.4233
	SSA-SVM	0.6332	1.8502	2.8902	6.9871	20.2353	31.5518	9.1682	26.0206	39.2098	84.0568	677.0735	1537.4051	91.0995	70.5263	60.8466
	Proposed Model	**0.5701**	**1.5897**	**1.5969**	**6.2878**	**17.3789**	**17.4961**	**8.3044**	**22.0649**	**22.4198**	**68.9636**	**486.8606**	**502.6490**	**91.6230**	**71.0526**	**71.4286**
Site4 VIC	SSA-ARIMA	0.6030	2.6833	4.7946	30.2420	145.7823	259.7349	38.8561	192.0960	322.5816	1509.7980	36900.8818	104058.8941	94.7644	90.5263	91.0053
	SSA-BPNN	0.5945	1.4461	1.6066	29.9897	77.8820	86.6667	38.5383	113.5722	120.8811	1485.2003	12898.6504	14612.2445	93.7173	82.6316	82.5397
	SSA-GRNN	2.3859	3.5421	4.6783	118.8554	181.0846	239.9875	146.2953	231.4429	313.8030	21402.3097	53565.8179	98472.3103	87.4346	87.3684	86.7725
	SSA-SVM	0.9749	2.0819	3.4279	49.4949	111.5463	185.0911	66.8503	147.8164	245.9356	4468.9590	21849.6841	60484.3216	90.0524	79.4737	74.0741
	Proposed Model	**0.5841**	**1.4433**	**1.5776**	**29.2273**	**77.7097**	**85.4954**	**37.8023**	**114.2885**	**122.6129**	**1429.0176**	**13061.8713**	**15033.9125**	**94.2408**	**81.0526**	**79.3651**

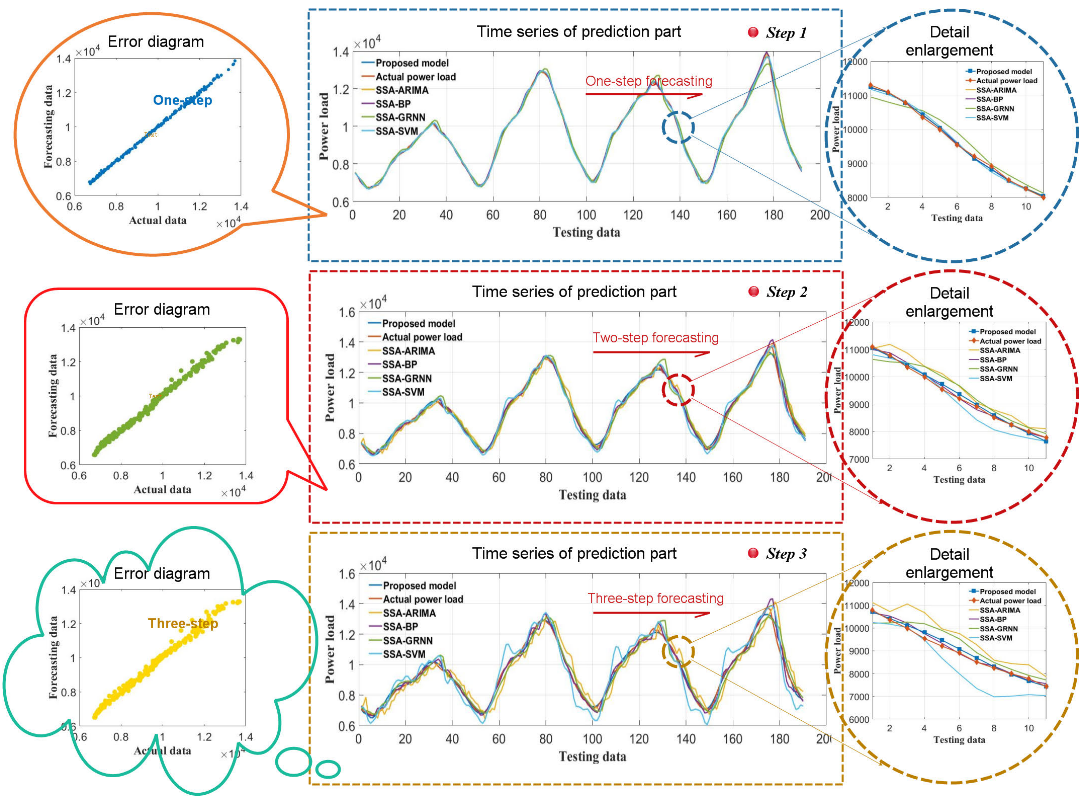


FIGURE 3. The multi-step prediction ability in Experiment I for Site1.

D. EXPERIMENT I: TESTS OF CEEMD-SSA-BASED MODELS

To evaluate the superiority of weighted combination model to single model, four single model which constituting the established composite model are used to compare with the established model. The four single models are SSA-ARIMA, SSA-BPNN, SSA-GRNN, and SSA-SVM, and the power load data of four datasets are used to carry on this experiment. The results of the experiment in **Table 3** and **Fig. 3** show that:

For *dataset NSW*, the SSA-based combination model achieves the best forecast performance when it is in one-step forecast, for the MAPE value obtained 0.4514%. On the contrary, the MAPE values of SSA-ARIMA, SSA-BPNN, SSA-GRNN, and SSA-SVM were relatively low. In predicting two and three-steps, the developed SSA-based combination model is more effective than other methods in predicting power load. According to the MAE, RMSE, MSE, and DC value obtained, ARIMA, as a representative of linear prediction, is less stable than the nonlinear neural network model in multi-step prediction.

TABLE 4. Performances of models with different data preprocessing techniques.

Dataset	Model	MAPE (%)			MAE			RMSE			MSE			DC (%)		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site1 NSW	EMD-ABGS.	0.9167	1.3529	1.7991	86.3480	127.2199	169.6478	155.0231	184.9586	233.7170	24032.1667	34209.6787	54623.6355	92.1466	85.2632	82.0106
	EEMD-ABGS.	0.9928	1.1991	1.6839	93.9331	97.3530	166.4404	128.8898	138.3064	247.3451	16612.5792	19128.6465	61179.6120	94.2408	91.5789	80.9524
	CEEMD-ABGS.	0.7587	1.4414	2.1425	72.7748	141.2789	211.1692	98.5799	206.6094	310.9776	9718.0026	42687.4548	96707.0820	94.7644	86.8421	73.5450
	Proposed Model	**0.4514**	**1.0903**	**1.2577**	**42.8534**	**103.1178**	**121.6497**	**54.3128**	**140.9177**	**170.5751**	**2949.8836**	**19857.7990**	**29095.8541**	**96.3351**	**92.1053**	**88.8889**
Site2 SA	EMD-ABGS.	3.0451	3.3371	4.0864	43.2054	46.8973	56.1778	64.0362	66.0727	71.9806	4100.6393	4365.5969	5181.2104	77.4869	71.5789	61.3757
	EEMD-ABGS.	1.6018	1.9325	3.6262	22.0816	25.8481	49.5616	31.3309	34.8460	63.2510	981.6245	1214.2423	4000.6842	84.8168	80.0000	58.2011
	CEEMD-ABGS.	1.8376	3.4948	5.3700	24.8476	47.9767	73.7901	31.3431	62.9823	96.4134	982.3903	3966.7670	9295.5489	83.2461	60.0000	44.4444
	Proposed Model	**0.9970**	**2.1160**	**3.5597**	**13.4636**	**28.8656**	**48.0672**	**19.0909**	**36.9727**	**61.9447**	**364.4606**	**1366.9831**	**3837.1501**	**91.0995**	**81.0526**	**62.9630**
Site3 TAS	EMD-ABGS.	0.8320	1.2238	2.1546	9.1775	13.5333	23.7981	12.4520	18.1714	30.2941	155.0534	330.1984	917.7301	87.4346	77.3684	58.7302
	EEMD-ABGS.	1.5799	1.5693	1.9320	17.3089	9.5446	21.1974	22.1874	12.4179	26.8388	492.2794	154.2050	720.3224	71.7277	85.2632	57.6720
	CEEMD-ABGS.	0.7467	1.6841	2.6155	8.1768	18.5173	28.8124	10.8473	23.4754	36.7185	117.6631	551.0945	1348.2509	86.3874	71.0526	47.6190
	Proposed Model	**0.5701**	**1.5897**	**1.5969**	**6.2878**	**17.3789**	**17.4961**	**8.3044**	**22.0649**	**22.4198**	**68.9636**	**486.8606**	**502.6490**	**91.6230**	**71.0526**	**71.4286**
Site4 VIC	EMD-ABGS.	0.8915	1.5048	2.1733	47.0685	88.6744	114.8432	75.9919	124.8487	153.4390	5774.7672	14996.2779	23543.5197	90.5759	85.7895	71.4286
	EEMD-ABGS.	1.4544	1.6678	2.3090	77.7436	89.9231	118.5727	113.7022	139.3561	156.1096	12928.1938	14297.3932	24370.2074	81.1518	77.3684	74.0741
	CEEMD-ABGS.	1.1249	2.1846	3.4101	57.5246	112.6209	175.5059	76.0399	147.3175	222.6383	5782.0596	21702.4389	49567.8297	87.9581	74.7368	59.2593
	Proposed Model	**0.5841**	**1.4433**	**1.5776**	**29.2273**	**77.7097**	**85.4954**	**37.8023**	**114.2885**	**122.6129**	**1429.0176**	**13061.8713**	**15033.9125**	**94.2408**	**81.0526**	**79.3651**

ABGS. represents the combined model contains ARIMA, BPNN, GRNN and SVM models.

For *dataset SA*, according to the five performance indicators, the established model is more outstanding than other single models at all stages of the forecast. According to the MAPE values in **Table 3**, SSA-GRNN model was the least effective model in the one, two and three-step forecast among the other single models, but when it is selected into the combination model, the accuracy of the combination model is still higher than that of the other three models. So it can be concluded that the combination model can effectively improve the prediction accuracy.

For *dataset TAS*, when the forecast is a one-step forecast, the MAPE, MAE, RMSE, and MSE of the developed combined model are the lowest, and the DC is 91.6230% which is the highest. This indicates the forecast accuracy of the established model is excellent. As for the two and three-step prediction, the index values of established model are obviously better than those of other SSA-based models, so it can draw the conclusion that the predictive ability of the established model is even better than that of other models.

For *dataset VIC*, the situation is similar to that of *NSW*, *SA* and *VIC*, which the values of five indicators shows that the established model in this paper has better forecast ability to other SSA-based models.

Remark: The error metrics obtained by the established model is obviously better than the value of single SSA-based prediction model for three forecasting steps. Experimental results therefore show that the developed SSA-MOEA/D combination model is better than individual SSA-based models in multi-stage prediction.

E. EXPERIMENT II: TESTS OF DATA PREPROCESSING METHODS MAPE, MAE, RMSE, MSE, DC

This section designs a comparison on the basis of different preprocessing methods. The preprocessing methods are EMD, EEMD, and CEMD as well as the SSA applying in this proposed model. These models are built to emphasize the advantages and disadvantages of the data processing technology. By comparing the forecasting model with the above methods, the excellent performance of the hybrid model proposed in this article is further clarified. The prediction errors of the other preprocessing methods based model and the proposed model are shown in **Tables 4** and **Fig.4**. More details of the experiment are given below:

First, the proposed model has the best MAPE, MAE, RMSE, MSE, and DC for one-step forecasting for dataset NSW, respectively. Second, the models with predictive accuracy of high to low are SSA model, CEEMD model, EEMD model, and EMD model and the MAPE values are decreasing step by step. There are still a few EMDs, EEMD, and CEEMD does not follow this rank but the SSA is always the best preprocessing method. In addition, the SSA-MOEA/D combination model has the best prediction power when the model is predicted with two-step and the MAPE value is lower than 2%. As for the three-step prediction, the established model is still superior among the other preprocessing based models. **Fig. 4** presents a comparison of the predictive performance of one, two and three-step at the second study site. It shows that SSA-MOEA/D combination model is the most accurate prediction model among other models.

The same situation happens in the other three sites for the MAPE of the established model are no more than 1% for step one, respectively. They are the lowest value among other compared preprocessing methods based model, and the same as MAE, RMSE, and MSE. For DC, the proposed model has the largest value among all the trial.

Remark: On this basis, the SSA-MOEA/D model has the highest predictive accuracy and the average MAPE values from one step to three steps are 0.6507%, 1.5598%, and 1.9980%. Furthermore, it indicates that the established model is better than the other models on the basic of different preprocessing methods, which can verify the validity of the model.

F. EXPERIMENT III: COMPARE COMBINED MODEL WITH CLASSIC INDIVIDUAL MODELS

To further compare and discuss the availability and superiority of the hybrid model system, six other individual forecasting models (ARIMA, BPNN, ELMAN, GRNN, ELM, and

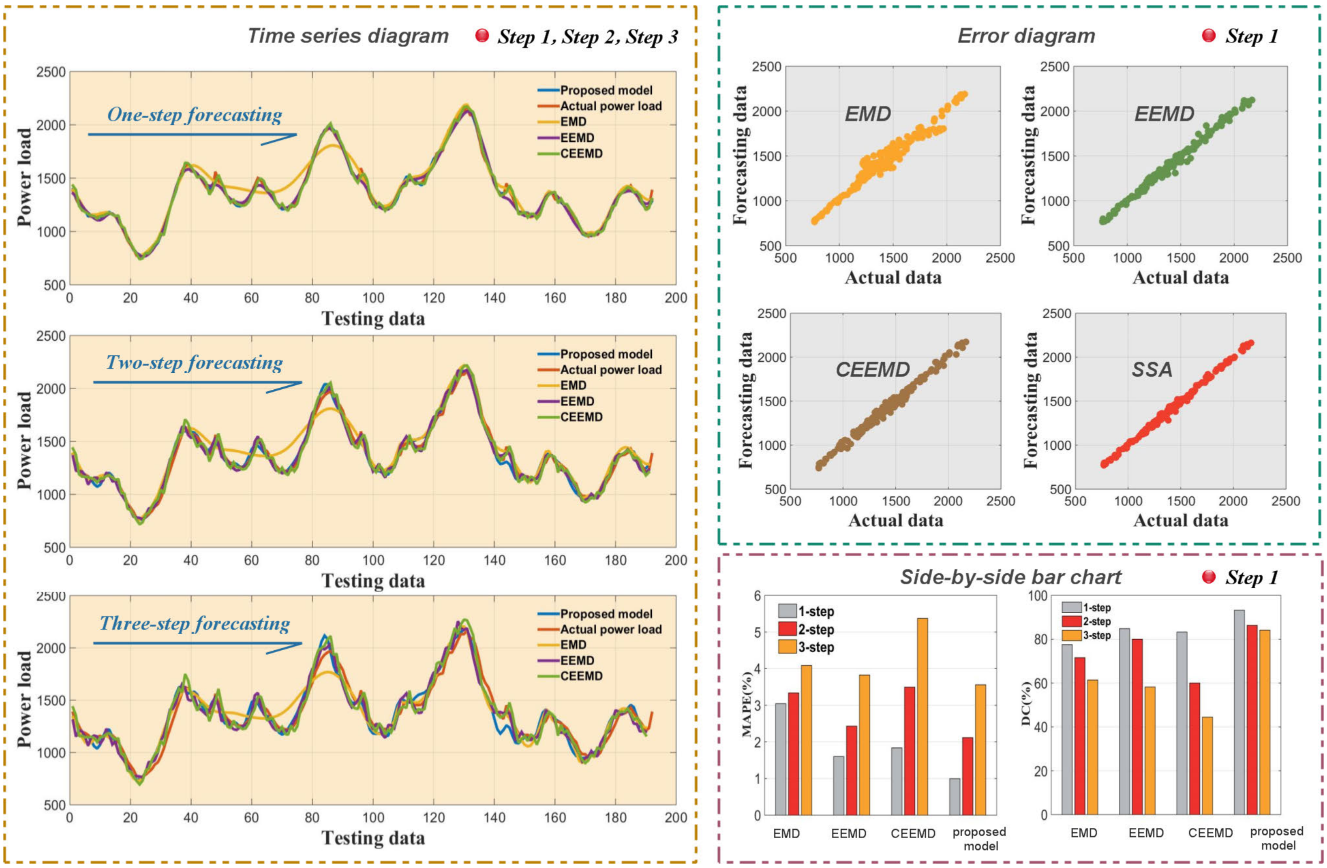


FIGURE 4. The multi-step prediction ability in Experiment II for Site2.

SVM) were used to carry out additional experiments with calculating their respective index values. All experimental prediction errors are listed in **Tables 5** and **Fig.5**. The conclusions of Experiment 1 and Experiment 2 are very similar. Therefore, the conclusions of this experiment are discussed in detail.

For four datasets, The SSA-MOEA/D hybrid model has achieved the highest accuracy for the three-step power load prediction, where its error metrics is better than other models. It can also be found that all MAPE values of proposed model have very visible improvement compared with these individual models.

In addition, the BP model ranks second among the other six individual models in site 1. For other sites besides site 1, the ELM model has the relatively better prediction level, and it takes turns with BP to occupy the second rank. ARLMA achieves high precision in one-step prediction, but in two-step and three-step, the precision decreases rapidly. This shows that ARIMA is not suitable for multi-step prediction. For ELMAN model, the accuracy is always lowest among other individual models.

For the metrics MAE, RMSE, and MSE, the error values of proposed model are lowest. And DC value of proposed model is highest for three steps. These indicators are all consistent with MAPE, shows that the proposed hybrid model has the best forecasting ability no matter in one-step or multi-step prediction.

Remark: Given the evaluation metrics in **Table 5**, SSA-MOEA/D hybrid model is still superior to those individual models, namely ARIMA, BPNN, ELMAN, GRNN, ELM, and SVM. In addition, the BP model and ELM model ranked second or three among other models but all the predictive performance of individual models is far inferior to the combined model. Accordingly, the proposed SSA-MOEA/D hybrid model is a good proposal for forecasting power load.

V. DISCUSSIONS

In order to discuss the experiment results in more detail and reduce the error of load prediction, several discussions are put forward, including the validity of the established model, the optimize ability of the optimization algorithm, and the practical application in power system.

A. DM TEST

Firstly, the validity of the model is verified by DM test. All other models are compared with SSA-MOEA/D models. Based on the theory of the DM test, the zero hypothesis is there is no significant difference between the prediction performance of the two models, and the other hypothesis is there is a significant difference between the prediction

TABLE 5. Performances of the hybrid model comparing with other single models.

Dataset	Model	MAPE (%)			MAE			RMSE			MSE			DC (%)		
		1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
Site1 NSW	ARIMA	0.8008	3.1289	5.9940	77.5742	299.1464	561.5114	107.7794	388.9793	692.1589	11616.3993	151304.8890	479083.9896	93.7173	85.7895	74.6032
	BPNN	0.7851	1.4306	2.1054	76.7429	142.4956	209.2378	109.6494	222.3156	317.6569	12022.9802	49424.2265	100905.9022	93.1937	86.8421	80.4233
	ELMAN	6.8802	6.8398	10.7411	650.4953	647.0488	1016.5607	734.2595	730.8930	1146.3169	539137.0449	534204.5105	1314042.4988	94.2408	93.1579	91.0053
	GRNN	1.4902	2.2803	3.1626	147.5198	223.0965	306.4669	209.0493	296.5749	397.0889	43701.5941	87956.6991	157679.5590	89.5288	87.8947	86.2434
	ELMNN	0.8077	1.4493	2.1728	79.4467	141.9924	212.4226	126.4863	204.4684	311.5069	15998.7741	41807.3440	97036.5494	92.1466	90.0000	84.6561
	SVM	0.8033	1.5648	2.3981	77.3678	150.0580	226.2159	109.8636	202.7321	300.6984	12070.0034	41100.3118	90419.5558	93.1937	89.4737	83.0688
	Proposed Model	**0.4514**	**1.0903**	**1.2577**	**42.8534**	**103.1178**	**121.6497**	**54.3128**	**140.9177**	**170.5751**	**2949.8836**	**19857.7990**	**29095.8541**	**96.3351**	**92.1053**	**88.8889**
Site2 SA	ARIMA	2.3266	6.5635	10.6315	31.7404	89.3403	142.8283	44.2329	108.7593	171.3388	1956.5511	11828.5901	29356.9999	86.3874	76.8421	70.8995
	BPNN	2.1839	4.1669	5.9071	30.1575	58.3607	84.8544	41.2517	71.1800	105.6235	1701.7017	5066.5936	11156.3268	89.5288	79.4737	72.4868
	ELMAN	7.8439	7.8501	15.4750	106.2673	106.3629	179.8859	127.6232	127.0956	229.8498	16287.6762	16153.2998	52830.9456	86.9110	85.7895	69.8413
	GRNN	4.0765	6.1195	8.0473	55.5051	84.4488	111.2389	67.6195	101.9606	138.3091	4572.3950	10395.9573	19129.3983	89.0052	82.6316	80.4233
	ELMNN	2.1566	4.0896	6.0191	30.0038	57.2873	84.4657	40.9540	72.5912	104.2474	1677.2262	5269.4875	10867.5223	86.3874	83.1579	78.3069
	SVM	2.1914	4.1844	6.6188	30.3003	58.1321	92.3741	44.1851	77.1300	120.2370	1952.3261	5949.0344	14456.9456	87.4346	81.0526	75.6614
	Proposed Model	**0.9970**	**2.1160**	**3.5597**	**13.4636**	**28.8656**	**48.0672**	**19.0909**	**36.9727**	**61.9447**	**364.4606**	**1366.9831**	**3837.1501**	**91.0995**	**81.0526**	**62.9630**
Site3 TAS	ARIMA	1.3709	3.0914	4.4497	15.0824	34.0448	48.1381	20.1702	43.5093	62.3954	406.8377	1893.0588	3893.1892	82.1990	71.5789	68.7831
	BPNN	1.4408	2.4307	3.3564	15.8846	26.6111	36.5835	21.2982	35.3131	46.1934	453.6154	1247.0142	2133.8280	79.0576	72.6316	69.3122
	ELMAN	3.4310	4.0748	4.7186	37.8454	44.9061	51.9607	47.2058	55.6548	64.0333	2228.3860	3097.4580	4100.2610	84.2932	76.8421	75.6614
	GRNN	2.2459	2.7048	3.9029	24.9257	30.0179	43.9623	30.7680	38.1684	54.3537	946.6724	1456.8267	2954.3275	78.0105	77.8947	68.7831
	ELMNN	1.4675	2.4422	3.4582	16.1488	26.7772	37.7623	21.3395	35.1657	47.2355	455.3754	1236.6252	2231.1939	81.1518	70.5263	67.7249
	SVM	1.4515	2.4676	3.6970	16.0050	27.0305	40.2733	21.3866	35.6514	50.7697	457.3883	1271.0258	2577.5621	79.5812	71.0526	68.2540
	Proposed Model	**0.5701**	**1.5897**	**1.5969**	**6.2878**	**17.3789**	**17.4961**	**8.3044**	**22.0649**	**22.4198**	**68.9636**	**486.8606**	**502.6490**	**91.6230**	**71.0526**	**71.4286**
Site4 VIC	ARIMA	1.3558	4.3191	9.2166	69.9244	228.2493	462.2151	94.6869	295.3592	666.3611	8965.6084	87237.0759	444037.1635	87.4346	76.8421	76.8421
	BPNN	1.3011	2.5817	3.8702	66.7873	133.6466	202.1717	91.3115	173.0312	263.4848	8337.7837	29939.7931	69424.2492	86.3874	79.4737	79.4737
	ELMAN	6.7978	8.1496	9.8226	363.4089	435.2846	524.4578	434.6551	519.6187	623.5999	188925.0909	270003.5719	388876.7806	92.1466	90.5263	90.5263
	GRNN	2.5187	3.6485	4.9355	127.3236	185.0871	250.3580	158.6006	236.9293	323.8019	25154.1563	56135.5163	104847.6411	90.5759	83.6842	83.6842
	ELMNN	1.2812	2.5996	3.7373	66.2920	137.0779	198.9694	90.8736	177.4982	261.9393	8258.0157	31505.6009	68612.1765	86.9110	79.4737	79.4737
	SVM	1.3224	2.6016	4.1027	68.9995	135.8771	213.5477	95.4889	182.9121	279.6161	9118.1385	33456.8292	78185.1753	85.8639	77.3684	77.3684
	Proposed Model	**0.5841**	**1.4433**	**1.5776**	**29.2273**	**77.7097**	**85.4954**	**37.8023**	**114.2885**	**122.6129**	**1429.0176**	**13061.8713**	**15033.9125**	**94.2408**	**81.0526**	**79.3651**

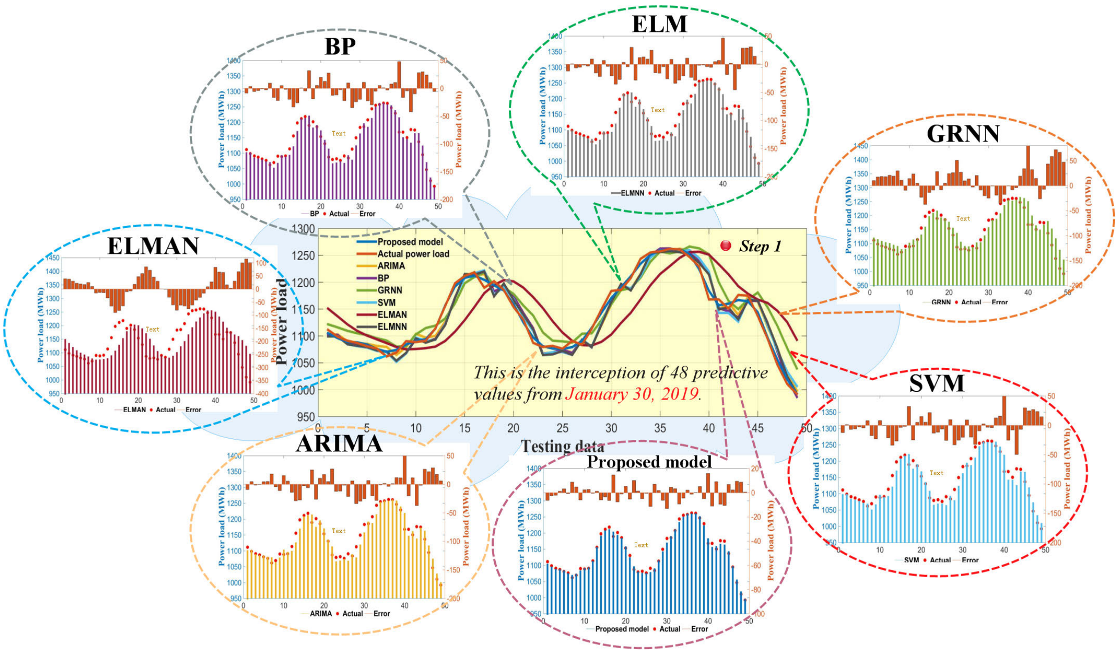


FIGURE 5. The multi-step prediction ability in Experiment III for Site3.

performance of the two models. **Table 6** gives the DM statistics and average values of the four sites.

Table 6 shows that most of the values of DM test calculate by the established model and the compared models of the above three experiments are greater than the upper limit of 1% significance level. Specially, the DM test value between SSA-BPNN and the proposed model is greater than the significant level upper limit of 5%. It indicated in **Table 6** that the minimum value of DM in the table is 1.99, which is larger than $Z_{0.05/2} = 1.96$, except for a few points which are not significant. Therefore, we can conclude that H_0 can be declined and the alternative hypothesis could be accepted at 5% significance level, and this also implies the probability accepted alternative hypothesis is 95%. This proves the performance of the established model is significant superior to that of the 13 compared combined models and the individual model, with a significant level of 95%. In other words, compared with other models, the accuracy of the proposed model has been significantly improved.

B. PERFORMANCE OF THE OPTIMIZATION ALGORITHM

In order to research on the excellent performance of the MOEA/D algorithm, two aspects are considered, which are parameter settings and convergence testing of metaheuristic algorithms.

TABLE 6. The statistics value of DM test for experimented models.

Model	NSW			SA			TAS			VIC			Average		
	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step	1-step	2-step	3-step
SSA-ARIMA	4.86*	8.95*	11.16*	6.94*	6.3*	6.33*	6.24*	5.02*	7.47*	4.97*	7.3*	10.21*	4.6*	5.52*	7.03*
SSA-BPNN	2.92*	1.99**	2.56*	2.9*	2.14**	1.71	2.01**	2.03**	2**	1.98**	1.99**	1.95**	2.45**	2.04**	2.05**
SSA-GRNN	5.18*	7.85*	8.57*	10.29*	7.45*	5.56*	8.73*	6.79*	7.77*	10.3*	8.76*	8.58*	6.9*	6.17*	6.1*
SSA-SVM	4.02*	4.62*	6.06*	5.24*	5.98*	7.71*	2.65*	4.55*	8.02*	4.55*	5.26*	7.16*	3.29*	4.08*	5.79*
EMD-ABGS.	4.35*	3.92*	4.85*	7.07*	6.17*	1.95*	5.18*	2.19**	3.76*	3.91*	1.97*	2.95*	4.1*	2.85*	2.7*
EEMD-ABGS.	4.85*	1.85	2.5**	4.23*	1.76	2.25**	8.13*	1.42	2.65*	4.86*	2.61*	2.18**	4.41*	1.53	1.92*
CEEMD-ABGS.	4.52*	3.13*	4.29*	6.7*	6.45*	5.39*	4.12*	0.84	6.59*	6.12*	2.54**	6.77*	4.29*	2.59*	4.61*
ARIMA	3.86*	5.99*	9.31*	5.38*	8.81*	10.67*	6.76*	7.98*	8.62*	5.82*	7.39*	3.52*	4.36*	6.03*	6.42*
BPNN	3.36*	3.33*	4.87*	5.63*	7.44*	5.7*	6.79*	5.35*	7.94*	5.25*	3.72*	5.78*	4.21*	3.97*	4.86*
ELMAN	5.66*	7.52*	7.35*	10.51*	9.18*	6.52*	8.96*	6.57*	8.61*	9.87*	7.8*	8.67*	7.00*	6.21*	6.23*
GRNN	4.14*	3.15*	5.72*	4.99*	5.45*	5.36*	7.00*	5.49*	8.25*	5.22*	3.99*	6.55*	4.27*	3.62*	5.18*
ELMNN	13.45*	13.86*	14.11*	11.98*	11.15*	9.74*	9.93*	9.13*	9.49*	13.05*	13.3*	13.61*	9.68*	9.49*	9.39*
SVM	3.08*	3.2*	4.99*	5.93*	7.57*	6.54*	7.24*	5.52*	8.08*	5.32*	4.25*	6.43*	4.31*	4.11*	5.21*

* 1% significance level; ** 5% significance level.
ABGS. represents the combined model contains ARIMA, BPNN, GRNN and SVM models.

TABLE 7. Parameter default settings of MOEA/D.

Experimental parameters	Default value
iterations	250
Population size	200
Optimizing parameter dimension	4
Neighbor size	20
Crossover rate	2
Simulated binary crossover parameter	2
Polynomial variation parameter	5

1) PARAMETER SETTINGS

Almost all metaheuristic algorithms need to set a series of parameters, which may lead to different results. It is not easy to find out the rule of parameter allocation, or to find out that it follows this principle in general application. Therefore, a discussion is carried out on MOEA/D, which uses various parameter configurations to study the impact of changes in parameters on forecasting performance, and to find out the optimal parameter settings of the algorithm. **Table 7** represents the parameter settings of MOEA/D algorithm.

The experiment for every values of each parameter is running 100 times, and the mean value and variance of MAE and R^2 are taken as the reference indexes. The specific experimental value are listed in **Table 8**.

a: EXTERNAL PARAMETERS

External parameters are commonly used in all metaheuristic algorithms. In this paper, three important parameters are selected, which are iteration number, population size and neighbor size. The effects of different values on the optimization results are discussed respectively.

b: ITERATIONS NUMBER

Since the number of iterations greatly affects the running time and calculation results of the algorithm, it is the first issue to be discussed. We select 150, 250, and 350 iterations respectively to discuss the performance of the algorithm. It can be seen that when the number of iterations is 250, the mean value and variance of MAE are the smallest, which are 42.8766 and 0.0005 respectively, and the mean value of R^2 is the smallest, which is 0.9991. In terms of running time, the smaller the iterations, the smaller the operation time. We take the number of iterations that the error is small and the time is not big as the parameter value.

c: POPULATION SIZE

A certain amount of population size can ensure the accuracy of the algorithm search, but too many will produce redundancy. It can be observed that when the population size is set to 100, 200, and 300 respectively, there is little difference in mean value of MAE and R^2, but there is a big difference in variance. When the population size is 200, the MAE variance is the smallest, which is 0.0005. When the population size is 100, the MAE variance is 1.1583. The variance of R^2 has the same situation. For the running time, the larger the population size, the longer the running time. This is because the larger the population size, the more the population to loop through, the more time spent.

d: NEIGHBOR SIZE

Neighbor size is the particle number contained in each population. The default neighbor size of the algorithm is 30. We discuss 20, 30, and 40 cases for comparison. It can be seen that different neighbor sizes have little effect on the MOEA/D algorithm. So neighbor size with relatively small running time can be selected as the parameter value of the algorithm.

e: INTERNAL PARAMETERS

Internal parameters are relative to external parameters, which are only used for specific algorithms. In this experiment, three internal parameters of MOEA/D algorithm, crossover rate, simulated binary crossover parameter, and polynomial variation parameter are selected for comparison and discussion.

f: CROSSOVER RATE

Crossover rate is used to judge the probability that two individuals need to cross in MOEA/D. The default crossover rate of this MOEA/D algorithm is 2. We discuss the crossover rate of 1, 2, and 3 respectively. It can be seen that when the crossover probability is 1, 2, and 3, MAE and R^2 are not very different. But the variance is significantly smaller when the

TABLE 8. Comparison of different parameter settings using NSW dataset.

External parameters	Iterations number			Population size			Neighbor size		
	150	250	350	100	200	300	20	30	40
Mean (MAE)	43.8828	42.8766	43.8697	43.5525	42.8766	42.9334	42.8770	42.8766	42.8876
Var (MAE)	0.0009	0.0005	0.0007	1.1583	0.0005	0.0433	0.0005	0.0005	0.0007
Mean (R^2)	0.9990	0.9991	0.9990	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991
Var (R^2)	1.36E-11	2.151E-11	1.5E-11	5.02E-09	2.15E-11	2.76E-10	1.66E-11	2.15E-11	1.45E-11
Time	3.0980	4.9576	6.9417	2.4992	4.9576	7.5438	3.9259	4.9576	6.3723

Internal parameters	Crossover rate			Analog binary crossover parameters			Polynomial Variation Parameters		
	1	2	3	1	2	3	3	5	7
Mean (MAE)	43.0896	42.8766	42.8661	43.1007	42.8766	42.9268	42.8859	42.8766	43.1812
Var (MAE)	0.2794	0.0005	0.0005	0.2134	0.0005	0.0206	0.0009	0.0005	0.6781
Mean (R^2)	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991	0.9991
Var (R^2)	1E-09	2.15E-12	6.53E-12	1.04E-09	2.15E-11	2.04E-10	1.61E-11	2.15E-11	2.79E-09
Time	4.9617	4.9576	5.0064	4.9831	4.9576	4.9133	4.9277	4.9576	4.9998

crossover rate is 2 and 3. From the point of view of running time, the elapsed time is not very different.

g: SIMULATED BINARY CROSSOVER PARAMETER

Simulated binary crossover (SBX) is a real-parameter combination operator which is commonly used in MOEA/D to optimization problems. SBX is actually designed with respect to the one-point crossover properties. The specific form of the operator is as follows

$$\begin{aligned} \tilde{x}_{1j}(t) &= 0.5 \times [(1+\gamma_j)x_{1j}(t) + (1-\gamma_j)x_{2j}(t)] \\ \tilde{x}_{2j}(t) &= 0.5 \times [(1-\gamma_j)x_{1j}(t) + (1+\gamma_j)x_{2j}(t)] \end{aligned} \tag{13}$$

$$\gamma_j = \begin{cases} (2u_j)^{\frac{1}{\eta+1}}, & if\ u_j \leq 0.5 \\ (\dfrac{1}{2(1-u_j)})^{\frac{1}{\eta+1}}, & others \end{cases} \tag{14}$$

where $u_j \in \mathrm{U}(0, 1)$ and $\eta > 0$ is the Simulated binary crossover parameter discussed below. It can be seen from the results that when the simulated binary crossover parameter are 1, 2, and 3 respectively, the mean and time of MAE and R^2 do not change much, but the variance is obviously smaller when the parameter is 2.

h: POLYNOMIAL VARIATION PARAMETER

The mutation operation of this algorithm is polynomial mutation. Polynomial variation parameter 3, 5, and 7 are selected for discussion, of which 5 are default polynomial variation parameter. It can be seen that the mean and time of MAE and R^2 at these points are not much different, but the variance is very small when the polynomial variation parameter is 5.

2) CONVERGENCE TESTING OF METAHEURISTIC ALGORITHMS

To verify the effectiveness and efficiency of MOEA/D compared with other multi-objective optimization algorithms, four test problems are carried out respectively. **Table 9** shows the contents of the four test functions. By comparing different optimization models, it is proved that the forecast ability of the MOEA/D algorithm is superior to other multi-objective algorithms. We select three other well-known multi-objective algorithms as comparison algorithms, which are NSGA2, MOALO, and MODA, and conduct 100 experiments on

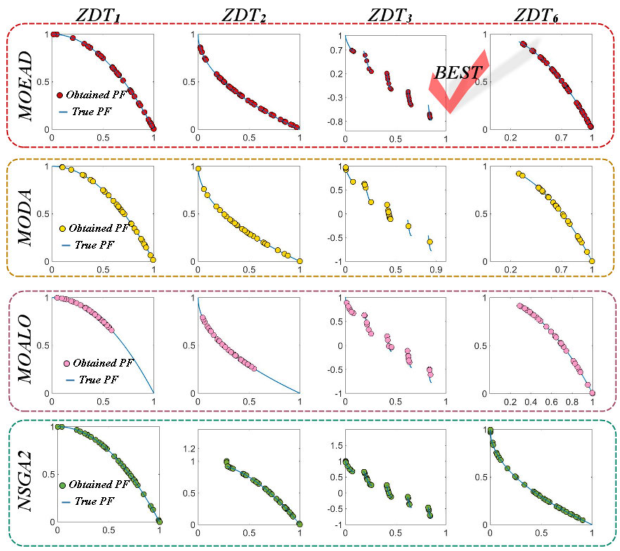


FIGURE 6. Pareto optimal solutions of MOEA/D, MODA, MOALO and NSGA2 for ZDT1, ZDT2, ZDT3 and ZDT4.

different population numbers [51]. The average values of the indicators are listed in **Table 10**.

In this section, the selected performance indicators of optimization algorithm include Inverted Generation Distance (IGD) indicators (Mirjalili et al., 2016) and Spread (SP). Besides, the operation time is considered as an indicator to judge the efficiency. Specially, IGD is an index to show the convergent condition of the algorithm. The results of IGD could be applied to judge the robustness and stability of used algorithm (Mirjalili et al., 2016; Mirjalili et al., 2017; Kusacci and Cann, 2013). The smaller the IGD value, the better the performance of the algorithm. In Pareto sets, SP is usually used to evaluate the distribution of solutions. If SP equals 0, all non-dominant solutions are equidistant.

The final simulation consequences are presented in **Table 10** and **Fig. 6**. The results show that for ZDT1, ZDT2, ZDT3, and ZDT6, MOEA/D has the best performance. The IDG of MOEA/D is much smaller than other algorithms, which shows that the MOEA/D has the best convergence performance. In particular, NSGA2 and MODA are suboptimal algorithms. The convergence effect of MOALO and MOGWO is much worse than other algorithms. For SP,

TABLE 9. Details of Multi-objective test functions.

Function name	The expression of objective function	Constraint scope	Characteristic	
			Continuity	Convexity
ZDT1	$f_1(x)=x_1; f_2(x)=g\left(1-\sqrt{f_1/g}\right);$ $g(x)=1+9\sum_{i=2}^{n}x_i/(n-1)$	[0,1]	✓	✓
ZDT2	$f_1(x)=x_1; f_2(x)=g\left(1-(f_1/g)^2\right);$ $g(x)=1+9\sum_{i=2}^{n}x_i/(n-1)$	[0,1]	✓	✗
ZDT3	$f_1(x)=x_1;$ $f_2(x)=g\left(1-\sqrt{f_1/g}-(f_1/g)\sin(10\pi f_1)\right);$ $g(x)=1+9\sum_{i=2}^{n}x_i/(n-1)$	[0,1]	✗	✓
ZDT6	$f_1(x)=x_1; f_2(x)=g(x)\left(1-(f_1/g)^2\right);$ $g(x)=1+10(n-1)+\sum_{i=2}^{n}x_i^2-10\cos(4\pi x_i)$	[0,1]	✓	✗

TABLE 10. Assessment results of MOEA/D and compared algorithms NSGA2, MOALO and MODA.

	Population	MOEA/D			NSGA2			MOALO			MODA			MOGWO		
		IDG	Spread	time	IDG	Spread	time	IDG	Spread	time	IDG	Spread	time	IDG	Spread	time
ZDT1	100	0.0013	0.2939	3.6488	0.0059	0.4072	9.0092	0.0059	1.4280	8.3820	0.0041	1.5406	32.0982	0.0115	0.9850	59.9657
	150	0.0005	0.2994	5.3414	0.0053	0.4716	17.9865	0.0304	1.3203	11.5836	0.0069	1.4513	59.7313	0.0067	0.8426	87.7022
	200	0.0004	0.2959	7.0971	0.0037	0.4712	30.0166	0.0120	1.4065	15.2488	0.0075	1.4525	89.3144	0.0118	0.8895	121.6576
	300	0.0001	0.3069	10.0975	0.0025	0.4540	62.7633	0.0167	1.4765	24.9062	0.0039	1.5637	170.8933	0.0090	0.9123	211.3944
ZDT2	100	0.0758	0.1922	4.5328	0.0072	0.4388	11.1098	0.0286	1.2572	7.4870	0.0080	1.2839	37.6540	0.0052	0.8398	67.9923
	150	0.0004	0.1822	6.9841	0.0044	0.4129	24.7541	0.0058	1.5592	10.5166	0.0085	1.4867	65.1603	0.0114	0.8728	101.3634
	200	0.0004	0.1571	7.2786	0.0037	0.4012	36.8125	0.0052	1.5376	19.2959	0.0038	1.7190	92.4846	0.0095	0.9641	135.6078
	300	0.0003	0.1672	10.1822	0.0025	0.4657	80.6811	0.0073	1.5137	21.1422	0.0013	1.3605	175.7402	0.0116	0.8576	209.7601
ZDT3	100	0.0244	0.9890	3.7555	0.0243	0.8015	15.1603	0.0317	1.4228	6.7319	0.0242	1.3365	36.2297	0.0258	1.1287	41.0739
	150	0.0243	1.0038	5.4889	0.0245	0.7886	31.1654	0.0253	1.4435	9.0774	0.0250	1.4101	63.5233	0.0280	1.1578	56.6127
	200	0.0243	1.0116	6.9201	0.0245	0.8245	51.5123	0.0252	1.4096	11.1368	0.0245	1.2588	90.8669	0.0266	1.0586	69.9916
	300	0.0243	1.0480	10.6065	0.0245	0.8214	107.6241	0.0254	1.4544	17.1206	0.0245	1.5307	168.5860	0.0260	1.1898	107.4549
ZDT6	100	0.0001	0.1468	3.2826	0.0505	0.6934	18.1816	0.0267	1.3053	7.8851	0.0088	1.2535	33.3456	0.0115	0.9135	56.7804
	150	0.0000	0.1534	4.6395	0.0587	0.5645	33.9415	0.0179	1.4581	11.0791	0.0066	1.4848	67.0615	0.0142	0.8681	86.3532
	200	0.0000	0.1559	5.8932	0.0587	0.9345	59.5900	0.0165	1.4111	14.7687	0.0037	1.5385	95.8804	0.0120	0.8778	115.4161
	300	0.0000	0.1658	8.7469	0.0365	1.2987	126.5335	0.0158	1.4556	23.0035	0.0026	1.5457	196.0288	0.0104	0.9421	201.7482

MOEA/D still has the best distribution performance. The elapsed time of MOEA/D algorithm is significantly lower than the time of the other three algorithms. In terms of work efficiency, MOEA/D algorithm is undoubtedly the fastest running and the best performance algorithm. When considering different population size, the population size is 100, 150, 200, and 300, respectively. Generally speaking, the larger the population size, the better the convergence effect.

C. PRACTICAL APPLICATION IN POWER SYSTEM

Electricity has become an indispensable part of national economic construction and people's lives. The role of power grid is to provide reliable, continuous and good quality power to all kinds of users as economically as possible. In terms of power system department terminology, it is necessary to supply power reliably, safely and economically to meet the requirements of load. Interruption, reduction, and inferiority of power supply will affect all sectors of the national economy and people's lives, and even cause serious political impact. The size and characteristics of load are very important factors for power grid planning and operation management. Therefore, prior estimation of load variation and characteristics is an important part of power system planning and operation research. The load forecasting theory of power system has been developed accordingly. The research of load forecasting has more important significance especially in the process of forming power trading market.

1) LOAD FORECASTING IS THE BASIS OF ECONOMIC OPERATION OF POWER GRID

The economic operation of power grid is to achieve obvious power saving effect without material investment. On the basis

Algorithm 1 MOEA/D

Input:

- MOP - multi-objective optimization problem
- N - the number of the MOEA/D subproblems
- $\lambda_1, \ldots, \lambda_N$ - a uniform distribution of N weight vectors
- T - the number of the weight vectors in the neighborhood of each weight vector
- *max_gen* - the maximum number of generations

Output:

- EP - external population

Setup:

- Set EP $= \emptyset$
- gen $= 0$

Step 1: Initialization

- /*Initialize an primary internal population uniformly randomly.*/
  $P_0 = \{x_1, \cdots, x_N\}$ and $FV_i = F(x_i)$
- /*Initialize $z = (z_1, \cdots, z_n)^{\mathrm{T}}$ by a specific problem method. */
- /* Calculate the Euclidean distance between any two weight vectors, and then calculate the closest T weight vectors to each weight vector.*/
- $\forall i = 1, \ldots, N$, set B(i) $= \{i_1, \cdots, i_{\mathrm{T}}\}$
  $\lambda_{i1}, \ldots, \lambda_{iT}$ represent the T closest weight vectors to λ_i

Step 2: Updating

- **WHILE** (t <*max_gen*) **DO**
- **FOR EACH** $i = 1, \ldots, N$ **DO**
  /* Genetic operators */ /* Randomly select two indexes k, l from $B(i)$, and then generate a new solution y from x_k and x_l by using genetic operators. */
- **FOR EACH** $j = 1, \ldots, n$ **DO**
  /*Update of z.*/ if $z_j < f_j(y)$, then set $z_j = f_j(y)$ **END FOR**
- **FOR EACH** index $j \in B(i)$ **DO**
  /*Update of neighboring solutions.*/
  if $g_{te}(y|\lambda_j, z) \leq g_{te}(x_j|\lambda_j, z^*)$, then set $x_j = y$ and $FV_j = F(y_j)$.

END FOR

- /*Update of EP.*/
  /*Remove from EP all the vectors dominated by F(y).Add F(y)to EP if no vector in EP dominate F(y). */
  END FOR
- $t = t+1$
  END WHILE
- **RETURN EP**

of the safety, stability, economic operation of power grid, and the guarantee of power supply, the existing equipment and original funds are fully utilized to forecast the normal load and the maximum load through the actual operation data of the power network. Computation of various operation modes allowed by the power grid and comparison with actual operation modes are made to rationally adjust the annual operation mode, quarterly operation mode, and monthly operation mode of transformers so as to make the main transformers operate economically. By changing the connection mode of transmission line, adjusting load reasonably, and cutting peak and filling valley, as well as reducing line loss, the maximum economic benefit of network operation can be achieved.

2) LOAD FORECASTING IS THE GUARANTEE OF SAFE AND RELIABLE OPERATION OF POWER GRID

The stable and safe operation of power system has become an important prerequisite for the operation of national economy. Because it is difficult to store large quantities of electric energy, the production and consumption of electric energy must be balanced at all times in the process of power supply. Excessive electric energy will lead to the decrease of security and stability of power grid, and the lack of electric energy will affect the normal production activities of society. Accurately load prediction can: A. sparingly manage the operations of generators in electric system, B. retain the security and steady of smart grid running, C. decrease rotating reservation capacity which are not necessary, D. rationally carry out the unit maintenance project, E. ensure the normal production activities of society, F. effectually decrease the expenses of generating electricity, enhance social and economic benefits.

3) POWER LOAD FORECASTING IS AN IMPORTANT PART OF POWER SYSTEM PLANNING

The load level of power system determines the scale and speed of its development. Therefore, the results of load forecasting determine the development of power system in the future planning period to a certain extent. The accuracy of load forecasting will directly affect the rationality of investment, network layout and operation. Accurate prediction of local future power demand and grid capacity is of great guiding significance to the determination of local power supply points and grid planning. Predicting the location, time and quantity of load distribution is the basis of location and capacity selection of high voltage substation, and its accuracy also determines the operability and adaptability of power grid planning.

VI. CONCLUSION

Credible and effectual electric load forecasting not only occupies a significant place in the electric control and the securty dispatch of the electricity grid, but also be beneficial to environment and economic. However, the complex fluctuation of power load makes it difficult to predict accurately. Consequently, this study proposes a new hybrid SSA-MOEA/D power load combination forecasting model.

This new model uses the advantages of each individual prediction model effectively, and finally gets further improvements. In particular, the original sequence is decomposed by the data pre-processing technology, and the filtered time series is reconstructed to eliminate the noise signals. Then, the load data are predicted using the non-linear neural network and the linear method ARIMA respectively. In addition, In addition, a new algorithm based on MOEA/D is proposed successfully, which integrates the weighted models to get the final prediction results. Taking the 30-minute load data set of Australian power system as an example, the ability of the hybrid model are evaluated. The experiment concequences show that the accurate and stabable ablility of the established model are superior to other individual models for comparison. Overall, the hybrid model developed in this article effectively improves the efficiency of the power load forecast and adds a new feasible scheme for smart network planning.

APPENDIXES

APPENDIX A

The pseudo code of MOEA/D.

APPENDIX B

See Tables 2–10.

CONFLICT OF INTEREST

The authors declare that there are no conflicts of interest regarding the publication of this paper.

REFERENCES

[1] J. Harmsen, A. Roes, and M. Patel, "The impact of copper scarcity on the efficiency of 2050 global renewable energy scenarios," *Energy*, vol. 50, pp. 62–73, Feb. 2013.

[2] I. Koprinska, M. Rana, and V. G. Agelidis, "Correlation and instance based feature selection for electricity load forecasting," *Knowl.-Based Syst.*, vol. 82, pp. 29–40, Jul. 2015.

[3] H. Liu, H.-Q. Tian, X.-F. Liang, and Y.-F. Li, "Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks," *Appl. Energy*, vol. 157, pp. 183–194, Nov. 2015.

[4] L. Liu, Q. Wang, J. Wang, and M. Liu, "A rolling grey model optimized by particle swarm optimization in economic prediction," *Comput. Intell.*, vol. 32, no. 3, pp. 391–419, Aug. 2016.

[5] E. A. Feinberg and D. Genethliou, "Load forecasting," *Power Electron. Power Syst.*, vol. 2005, pp. 269–285.

[6] N. Bowden and J. E. Payne, "Short term forecasting of electricity prices for MISO hubs: Evidence from ARIMA-EGARCH models," *Energy Econ.*, vol. 30, no. 6, pp. 3186–3197, Nov. 2008.

[7] W. K. Wong and Z. X. Guo, *Intelligent Sales Forecasting for Fashion Retailing Using Harmony Search Algorithms and Extreme Learning Machines Optimizing Decision Making in the Apparel Supply Chain Using Artificial Intelligence (AI)*. Amsterdam, The Netherlands: Elsevier, 2013, pp. 170–195.

[8] C. Pilling, V. Dodds, M. Cranston, D. Price, T. Harrison, and A. How, "Flood forecasting—A national overview for great britain," *Flood Forecasting*, pp. 201–247, 2016.

[9] W. El-Baz and P. Tzscheutschler, "Short-term smart learning electrical load prediction algorithm for home energy management systems," *Appl. Energy*, vol. 147, pp. 10–19, Jun. 2015.

[10] J. Zhang, C. Draxl, T. Hopson, L. D. Monache, E. Vanvyve, and B.-M. Hodge, "Comparison of numerical weather prediction based deterministic and probabilistic wind resource assessment methods," *Appl. Energy*, vol. 156, pp. 528–541, Oct. 2015.

[11] J. Wang, Y. Gao, and X. Chen, "A novel hybrid interval prediction approach based on modified lower upper bound estimation in combination with multi-objective salp swarm algorithm for short-term load forecasting," *Energies*, vol. 11, no. 6, p. 1561, Jun. 2018.

[12] S. Sepasi, E. Reihani, A. M. Howlader, L. R. Roose, and M. M. Matsuura, "Very short term load forecasting of a distribution system with high PV penetration," *Renew. Energy*, vol. 106, pp. 142–148, Jun. 2017.

[13] X. Qiu, Y. Ren, P. N. Suganthan, and G. A. Amaratunga, "Empirical mode decomposition based ensemble deep learning for load demand time series forecasting," *Appl. Soft Comput.*, vol. 54, pp. 246–255, May 2017.

[14] C. Tienxin and P. Zhiyong, "Load-frequency control with short-term load prediction," *Power Syst. Power Plant Control*, vol. 1, pp. 205–208, Jan. 1987.

[15] B. Punantapong, P. Punantapong, and I. Punantapong, "Improving a grid-based energy efficiency by using service sharing strategies," *Energy Procedia*, vol. 79, pp. 910–916, Nov. 2015.

[16] C. Sandels, J. Widén, L. Nordström, and E. Andersson, "Day-ahead predictions of electricity consumption in a Swedish office building from weather, occupancy, and temporal data," *Energy Buildings*, vol. 108, pp. 279–290, Dec. 2015.

[17] K. W. Kow, Y. W. Wong, R. Rajkumar, and R. K. Rajkumar, "A review on performance of artificial intelligence and conventional method in mitigating PV grid-tied related power quality events," *Renew. Sustain. Energy Rev.*, vol. 56, pp. 334–346, Apr. 2016.

[18] Z. Falomir and A.-M. Olteţeanu, "Logics based on qualitative descriptors for scene understanding," *Neurocomputing*, vol. 161, pp. 3–16, Aug. 2015.

[19] R. Zimroz, W. Bartelmus, T. Barszcz, and J. Urbanek, "Diagnostics of bearings in presence of strong operating conditions non-stationarity—A procedure of load-dependent features processing with application to wind turbine bearings," *Mech. Syst. Signal Process.*, vol. 46, no. 1, pp. 16–27, May 2014.

[20] A. Lasek, N. Cercone, and J. Saunders, "Smart restaurants: Survey on customer demand and sales forecasting," in *Smart Cities Homes*. Amsterdam, The Netherlands: Elsevier, 2016, pp. 361–386.

[21] H. Kwon, B. Lyu, K. Tak, J. Lee, and I. Moon, "Optimization of petrochemical process planning using naphtha price forecasting and process modeling," in *Proc. 12th Int. Symp. Process Syst. Eng.*, 2015, pp. 2039–2044.
[22] H. Hippert, C. Pedreira, and R. Souza, "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Trans. Power Syst.*, vol. 16, no. 1, pp. 44–55, Feb. 2001.
[23] S. Pappas, L. Ekonomou, P. Karampelas, D. Karamousantas, S. Katsikas, G. Chatzarakis, and P. Skafidas, "Electricity demand load forecasting of the Hellenic power system using an ARMA model," *Electr. Power Syst. Res.*, vol. 80, no. 3, pp. 256–264, Mar. 2010.
[24] D. Chaturvedi, A. Sinha, and O. Malik, "Short term load forecast using fuzzy logic and wavelet transform integrated generalized neural network," *Int. J. Electr. Power Energy Syst.*, vol. 67, pp. 230–237, May 2015.
[25] C. W. Lou and M. C. Dong, "A novel random fuzzy neural networks for tackling uncertainties of electric load forecasting," *Int. J. Electr. Power Energy Syst.*, vol. 73, pp. 34–44, Dec. 2015.
[26] L. Zhang, Y. Dong, and J. Wang, "Wind speed forecasting using a two-stage forecasting system with an error correcting and nonlinear ensemble strategy," *IEEE Access*, vol. 7, pp. 176000–176023, 2019, doi: 10.1109/access.2019.2957174.
[27] M. Mohammadi, F. Talebpour, E. Safaee, N. Ghadimi, and O. Abedinia, "Small-scale building load forecast based on hybrid forecast engine," *Neural Process. Lett.*, vol. 48, no. 1, pp. 329–351, Aug. 2018.
[28] X. Tang, Y. Dai, T. Wang, and Y. Chen, "Short-term power load forecasting based on multi-layer bidirectional recurrent neural network," *IET Gener., Transmiss. Distrib.*, vol. 13, no. 17, pp. 3847–3854, Sep. 2019.
[29] N. Ghadimi, A. Akbarimajd, H. Shayeghi, and O. Abedinia, "Two stage forecast engine with feature selection technique and improved meta-heuristic algorithm for electricity load forecasting," *Energy*, vol. 161, pp. 130–142, Oct. 2018, doi: 10.1016/j.energy.2018.07.088.
[30] L. Xiao, J. Wang, Y. Dong, and J. Wu, "Combined forecasting models for wind energy forecasting: A case study in China," *Renew. Sustain. Energy Rev.*, vol. 44, pp. 271–288, Apr. 2015, doi: 10.1016/j.rser.2014.12.012.
[31] C. Ren, N. An, J. Wang, L. Li, B. Hu, and D. Shang, "Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting," *Knowl.-Based Syst.*, vol. 56, pp. 226–239, Jan. 2014, doi: 10.1016/j.knosys.2013.11.015.
[32] Y. Xu, W. Yang, and J. Wang, "Air quality early-warning system for cities in China," *Atmos. Environ.*, vol. 148, pp. 239–257, Jan. 2017.
[33] Z. Liu, P. Jiang, L. Zhang, and X. Niu, "A combined forecasting model for time series: Application to short-term wind speed forecasting," *Appl. Energy*, vol. 259, Feb. 2020, Art. no. 114137, doi: 10.1016/j.apenergy.2019.114137.
[34] P. Jiang and Z. Liu, "Variable weights combined model based on multi-objective optimization for short-term wind speed forecasting," *Appl. Soft Comput.*, vol. 82, Sep. 2019, Art. no. 105587, doi: 10.1016/j.asoc.2019.105587.
[35] X. Zhang, J. Wang, and Y. Gao, "A hybrid short-term electricity price forecasting framework: Cuckoo search-based feature selection with singular spectrum analysis and SVM," *Energy Econ.*, vol. 81, pp. 899–913, Jun. 2019.
[36] S. Soltani, "On the use of the wavelet decomposition for time series prediction," *Neurocomputing*, vol. 48, pp. 267–277, Oct. 2002.
[37] D. S. Broomhead and G. P. King, "Extracting qualitative dynamics from-experimental data," *Phys. D, Nonlinear Phenomena*, vol. 20, pp. 217–236, 1986.
[38] J. B. Elsner and A. A. Tsonis, *Singular Spectrum Analysis a New Tool in Time Series Analysis*. New York, NY, USA: Plenum Press, 1996.
[39] Y. Wang, P. Shang, L. He, Y. Zhang, and D. Liu, "Can China achieve 2020 2030 carbon intensity targets through energy structure adjustment," *Energies*, vol. 11, no. 10, p. 2721, 2018, doi: 10.3390/en11102721.
[40] P. Du, Y. Jin, and K. Zhang, "A hybrid multi-step rolling forecasting model based on SSA and simulated annealing—Adaptive particle swarm optimization for wind speed," *Sustainability*, vol. 8, no. 8, p. 754, 2016, doi: 10.3390/su8080754.
[41] H. Hassani, N. Alharbi, M. Ghodsi, "A study on the empirical distribution of thescaled Hankel matrix eigenvalues," *J. Adv. Res.*, vol. 6, pp. 925–929, Nov. 2015.
[42] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Trans. Evol. Comput.*, vol. 11, no. 6, pp. 712–731, Dec. 2007.
[43] K. Deb and D. Kalyanmoy, *Multi-Objective Optimization Using Evolutionary Algorithms*, vol. 2. Hoboken, NJ, USA: Wiley, 2001, p. 509.
[44] H. Ishibuchi, Y. Sakane, N. Tsukamoto, and Y. Nojima, "Simultaneous use of different scalarizing functions in MOEA/D," in *Proc. 12th Annu. Conf. Genetic Evol. Comput. (GECCO)*, 2010, pp. 519–526.
[45] P. Du, J. Wang, W. Yang, and T. Niu, "A novel hybrid model for short-term wind power forecasting," *Appl. Soft Comput.*, vol. 80, pp. 93–106, Jul. 2019, doi: 10.1016/j.asoc.2019.03.035.
[46] J. Wang, P. Du, H. Lu, W. Yang, and T. Niu, "An improved grey model optimized by multi-objective ant lion optimization algorithm for annual electricity consumption forecasting," *Appl. Soft Comput.*, vol. 72, pp. 321–337, Nov. 2018, doi: 10.1016/j.asoc.2018.07.022.
[47] P. Du, J. Wang, and W. Yang, "Container throughput forecasting using a novel hybrid learning method with error correction strategy," *Knowl. Based Syst.*, vol. 182, Oct. 2019, Art. no. 104853, doi: 10.1016/j.knosys.2019.07.024.
[48] H. Li, J. Wang, R. Li, and H. Lu, "Novel analysis–forecast system based on multi-objective optimization for air quality index," *J. Cleaner Prod.*, vol. 208, pp. 1365–1383, Jan. 2019, doi: 10.1016/j.jclepro.2018.10.129.
[49] P. Jiang, H. Yang, and X. Ma, "Coal production and consumption analysis, and forecasting of related carbon emission: Evidence from China," *Carbon Manage.*, vol. 10, no. 2, pp. 189–208, Mar. 2019, doi: 10.1080/17583004.2019.1577177.
[50] F. Diebold and R. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Statist.*, vol. 20, no. 1, pp. 134–144, Jan. 2002, doi: 10.1080/07350015.1995.10524599.
[51] R. Li and Y. Jin, "A wind speed interval prediction system based on multi-objective optimization for machine learning method," *Appl. Energy*, vol. 228, pp. 2207–2220, Oct. 2018.

HE BO received the Ph.D. degree from the Tianjin University of Finance and Economics, in 2011. She is currently pursuing the Ph.D. degree with the School of Economics, Dongbei University of Finance and Economics. She has published some refereed journal articles. Her research interests include computational intelligence, macroeconomic analysis, and wind speed forecasting.

YING NIE received the B.S. degree from Northeast Petroleum University, in 2013. She is currently pursuing the master's degree with the Dongbei University of Finance and Economics. Her research interests include power load forecasting, wind energy forecasting, artificial intelligence, optimization algorithms, and machine learning.

JIANZHOU WANG received the B.S. degree from Northwest Normal University, in 1988, and the M.Sc. and Ph.D. degrees from Lanzhou University, in 1998 and 2004, respectively. He is currently a Professor with the School of Statistics, Dongbei University of Finance and Economics. He has published over 150 refereed journal articles. His research interests include wind energy forecasting, data mining, machine learning, statistics learning, and forecast theory.