Research Article

A Hybrid Forecasting Model Based on EMD-GASVM-RBFNN for Power Grid Investment Demand

Jinchao Li,^{1,2} Shaowen Zhu ,¹ Qianqian Wu,¹ and Pengfei Zhang³

¹School of Economics and Management, North China Electric Power University, Changping, Beijing, China

²Beijing Key Laboratory of New Energy and Low-Carbon Development, North China Electric Power University, Changping, Beijing, China

³State Grid Corporation of China, Xicheng, Beijing, China

Correspondence should be addressed to Shaowen Zhu; zhusw@ncepu.edu.cn

Received 28 July 2018; Accepted 5 September 2018; Published 26 September 2018

Academic Editor: Vincenzo Bianco

Copyright © 2018 Jinchao Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Power grid as an important infrastructure which ensures the healthy development of economy and society and accurate and reasonable prediction of the power grid investment demand has always been the focus problem of the power planning department and the power grid enterprises. In view of the complex nonlinear and nonstationary characteristics of the power grid investment demand sequence, a novel hybrid EMD-GASVM-RBFNN forecasting model based on empirical mode decomposition (EMD) method, support vector machines optimized by genetic algorithm (GA-SVM) model, and radial basis function neural network (RBFNN) model is proposed. Firstly, the EMD method is used to decompose the original power grid investment data sequence into a series of IMF components and a residual component which have stronger regularity compared with the original data. Then, according to the different characteristics of each subsequence, the GA-SVM and RBFNN model will be used to forecast different subsequences, respectively. Next, the prediction results of different subsequences are aggregated to obtain the final prediction results of the power grid investment. Finally, this paper dynamically simulates China's power grid investment from 2018 to 2020 based on the EMD-GASVM-RBFNN hybrid forecasting model and Monte Carlo method.

1. Introduction

Power grid is a basic carrier of electricity collection, transmission, and allocation; it is an important guarantee for the healthy development of the economy and society. Moderate investment and construction on power grid in advance is the prerequisite for ensuring the power supply. In general, the huge investment amount of power grid investment with long payback period has greatly increased the difficulty and risk of power grid investment decision. So, accurate and effective prediction of power grid investment can not only help pool funds and rationally arrange investment in power grid construction, but also reduce capital costs and economic risks, which plays a crucial role in promoting power grid investment planning and construction process [1]. Power grid investment is affected by various influencing factors such as economy, society, and load level. The development trend of power grid investment in different countries has significant

diversities. The power grid investment graphs of US, Japan, Germany, and France are presented in Figure 1 while China's power grid investment is depicted in Figure 2.

As can be seen from Figure 1, from the mid-1970s to the end of the 1990s, the overall power grid investment in United States has shown a downward trend, and the annual decline is about 50 million dollars. But after entering the 21st century, the US power grid investment began to rebound. At the same time, the development trend of the power grid investment in Japan was in the opposite direction to United States; the power grid investment in Japan fell from the peak level of more than 1200 billion yen in 1993 to only 300 billion yen in 2003. In only 10 years, the Japanese power grid investment dropped by nearly 3 times. In Europe, the overall power grid investment in Germany basically retained a slight growth trend in the latest 10 years except for the sharp decline in 2008 caused by international financial crisis. In the same period, the RTE power grid investment in France increased from 600

FIGURE 1: Power grid investment sequence diagram of United States, Japan, Germany, and France.

FIGURE 2: China's power grid investment demand data sequence from 1990 to 2017.

million euros in 2006 to more than 1500 million euros in 2016, with an average annual growth rate of more than 8%. It can be seen from Figure 2 that China's power grid investment basically keeps a sharp growth trend with an annual growth rate of 20.5%, and the power grid investment has increased from 5.4 billion yuan in 1990 to 531.5 billion yuan in 2017.

According to the analysis of the power grid investment trend in the above countries, we can conclude that the development trend of power grid investment in different countries is obviously distinguished; such distinction mainly originates from the differences in the national conditions of different countries. The power grid investment is influenced by many influencing factors such as economy, society, politics, and

environment, and different countries have remarkable difference in these factors. These factors can be divided into short-term impact factors, medium-term impact factors, and long-term impact factors. The division and characteristics of three types of factors are listed in Table 1.

The short-term impact factors, medium-term impact factors, and long-term impact factors work together making the power grid investment have complicated nonlinear and nonstationary characteristics. Therefore, accurate and effective forecasting of the power grid investment has become a key and difficult problem for power grid enterprises and the power planning department.

A reasonable forecasting method is an important prerequisite for improving the effectiveness and accuracy of the power grid investment prediction in the future. The exploration of forecasting methods has always been a hot spot for scholars, and a lot of forecasting methods have been put forward over the years [2].

The traditional forecasting models based on parameter estimation include regression analysis, time series prediction, and grey model. After continuous development for decades in the last century, these traditional methods have mature theoretical foundations, and a large number of scholars at home and abroad have conducted extensive researches and applications of them, which are summarized in Table 2.

In recent years, with the continuous development of the methods such as artificial neural networks, machine learning, and intelligent computing, these methods have been widely used in prediction field. As a result, the single or combined prediction model based on the above-mentioned intelligent

TABLE 1: Partitioning of grid investment demand influencing factors.

Type	Acting factors	Features
Short-term impact factors	Major policy changes, the international financial crisis and other uncertainties.	Most of them are emergent events, with strong uncertainty and randomness.
Medium-term impact factors	National five-year development plan, phased planning of power planning department, etc.	These factors have obvious periodic characteristics
Long-term impact factors	Safeguarding the country's current level of economic and social development and serving the goal of long-term national planning	Such factors have the characteristics of long-term, tendency and stability.

algorithms have also become the mainstream and trend in prediction field. The summary of domestic and foreign research on these methods is shown in Table 3.

Although intelligent algorithms are adept in coping with the nonlinear problems, they would fail to identify and extract internal features and characteristics effectively from complex nonlinear nonstationary data. In response to this problem, wavelet analysis, EMD method, and other decomposition algorithms were developed. The decomposition algorithms are used to decompose the nonlinear nonstationary original sequence into a series of subsequences with stationary characteristics of different characteristic scales, and the decomposed subsequences contain different physical characteristics of the original sequence; then different forecasting method will be used to predict subsequences with different physical characteristics; finally the final forecasting result of the original sequence will be obtained by aggregating the forecasting results of the subsequences. The decomposition algorithms can effectively identify and extract the inherent features and trends of complex nonlinear nonstationary signals and significantly improve the prediction accuracy of nonlinear nonstationary time series, so it has been widely used in the prediction field. The summary of domestic and foreign research on decomposition methods for prediction is shown in Table 4.

From the above review of the literature, we can conclude that the existing forecasting methods have developed relatively maturely. However, the study on power grid investment forecasting model is still lacking. HU Bai-chu et al. [3] use the analytic hierarchy process and grey model to forecast the power grid infrastructure investment, but the forecasting model is too simple and subjective, so the accuracy of the prediction result still remained to be improved. ZHAO Huiru et al. [4] propose a cointegration theory and error correction model to forecast the power grid investment demand; this model finds out the logarithmic linear relation between the power grid investment and influencing factors of the power grid investment. However, the power grid investment has complicated nonlinear nonstationary characteristics; the proposed forecasting model in this paper hardly excavates the internal laws and characteristics of the power grid investment. Shuyu Dai et al. [1] propose a hybrid support vector machine optimized by differential evolution algorithm and grey wolf optimization algorithm (DE-GWO-SVM) to forecast China's power grid investment based on the historical data from 1990 to 2016. The support vector machine can effectively process the complex nonlinear data and brilliant

parameters optimization ability of the intelligent algorithm; the DE-GWO-SVM model has an outstanding prediction performance for China's power grid investment.

In the light of the defects and advantages of existing power grid investment forecasting models, this paper proposes a novel EMD-GASVM-RBFNN hybrid model to forecast China's power grid investment based on the historical data from 1990 to 2017. The structure diagram of the EMD-GASVM-RBFNN hybrid forecasting model is shown in Figure 3. The innovations of this article are as follows.

(1) The empirical mode decomposition (EMD) is used to decompose the original data of China's power grid investment which has complex nonlinear, nonstationary characteristics from 1990 to 2017 into several intrinsic mode function (IMF) components and a residual component which represent the random characteristics, periodic characteristics, and trend characteristics of the original power grid investment data, respectively. These subsequences decomposed from the original power grid investment data by the EMD method have simpler frequency feature and stronger correlations [5] which are easier for building the prediction model than the original data series and which improve the prediction accuracy prominently.

(2) The support vector machine optimized by genetic algorithm (GA-SVM) model and radial basis function neural network (RBFNN) model are used to forecast the subsequences of the original power grid investment data decomposed by the EMD method. Twenty-eight power grid investment data samples from 1990 to 2017 will be used to establish the prediction model in this paper; it is difficult to build an accurate and robust forecasting model with just 28 samples. The SVM and RBFNN model are both adept in processing nonlinear problems with small sample numbers. The periodic subsequence with low frequency and the trend subsequence will be predicted by GA-SVM model while the random subsequence with high frequency will be forecasted by RBFNN model which has strong nonlinear fitting ability and parameter learning ability.

(3) Monte Carlo dynamic simulation method will be used to simulate China's power grid investment in 2018-2020 based on the EMD-GASVM-RBFNN hybrid prediction model proposed in this paper. The power grid investment is affected by many influencing factors, and the change of these influencing factors in the future will have a huge influence on power grid investment; but the exact value of these influencing factors in the future are difficult to determine. In this paper, we use the Monte Carlo method to simulate the

TABLE 2: Summary of traditional forecasting methods and applied documents.

Author	Forecasting object	Forecasting model	Characteristics
Antonio J. Conejo et al. [6]	Electricity prices	ARIMA	<p>In the traditional method, the basic function relation of the model must be determined in advance, and the parameter estimation method is used to estimate the model parameters. Moreover, the traditional method is difficult to effectively identify the internal characteristics of the signal to be predicted, and the prediction accuracy is low.</p>
Volkan S. Ediger et al. [7]	Energy demand	ARIMA	
Diyar Akay et al. [8]	Electricity demand	Grey Model	
Rajesh G. Kavasseri et al. [9]	Wind speed	Fractional-ARIMA	
Vincenzo Bianco et al. [10]	Electricity consumption	Linear regression models	
Li Jun-fang et al. [11]	Wind speed and wind power	Grey Model	
Hsiao-Tien Pao et al. [12]	CO2 emissions, energy consumption and economic growth in China	Improved grey model	

TABLE 3: Summary of prediction methods and application literature based on intelligent algorithms.

Author	Forecasting object	Forecasting model	Characteristics
Pedro A. González et al. [13]	Energy consumption in buildings	Feedback Artificial Neural Network	It has a powerful function approximation ability and can fit the function expression of unknown system. It is an effective method to deal with complex nonlinear systems. However, it is difficult to fully identify and extract the internal characteristics of complex nonlinear and non-stationary time series.
Ping-Feng Pai et al. [14]	Stock price	ARIMA-SVM	
ZOU Zheng-da et al. [15]	Short-Term Load Forecasting	Recurrent Neural Network-Ant Colony Optimization Algorithm	
Ping-Feng Pai et al. [16]	Electricity load	SVM-GA	
Thanasis G. Barbounis et al. [17]	Long-Term Wind Speed and Power Forecasting	Local Recurrent Neural Network Models	
Nima Amjadi [18]	Electricity Prices	Fuzzy Neural Network	
Ping-Feng Pai et al. [19]	Rainfall Forecasting	Recurrent Support Vector Regression	
J.P.S. Catalao et al. [20]	Electricity prices	Neural network approach	
Nicholas I. Sapankevych et al. [21]	Time series prediction	SVM	
Dongxiao Niu et al. [22]	Power load	SVM and ant colony optimization Algorithm	
LI Jin et al. [23]	Mid-long Term Load Forecasting	Simulated Annealing and SVM Algorithm	Generalized regression neural network with fruit fly optimization algorithm
Hong-ze Li et al. [24]	Power load	SVR with Chaotic Immune Algorithm	
Wei-Chiang Hong [25]	Traffic Flow Forecasting		
Geng, J et al. [26]	Load Forecasting	SVR	

TABLE 4: Summary of prediction methods and application literature based on decomposition algorithm.

Author	Forecasting object	Forecasting model	Characteristics
Antonio J. Conejo et al. [27]	Electricity Price	Wavelet-ARIMA Models	It can effectively identify and extract the internal features and laws of nonlinear non-stationary time series, and significantly improve the prediction accuracy of nonlinear non-stationary time series.
Z. A. Bashir and M. E. El-Hawary [28]	Short-Term Load Forecasting	Wavelets-PSO-Based Neural Networks	
Rahmat-Allah Hooshmand et al. [29]	Short-term load forecasting	Wavelet transform and artificial neural network model	
Lean Yu et al. [30]	Crude oil price	EMD-based neural network model	
Chun-Fu Chen et al. [5]	Tourism demand	EMD-based neural network model	
YE Lin and LIU Peng [31]	Short-term Wind Power Prediction	EMD-SVM model	
Ning An et al. [32]	Electricity demand	EMD-FNN model	
L. Karthikeyan and D. Nagesh Kumar [33]	Non-stationary time series	EMD-ARMA models	
W.Y. Duan et al. [34]	Short-Term Wave Height	EMD-SVR model	
Shouxiang Wang et al. [35]	Wind Speed	EMD-GA-BP neural network	
Fan, G.-F. et al. [36].	Load Forecasting	EMD-PSO-GA-SVR

FIGURE 3: Structure diagram of the EMD-GASVM-RBFNN hybrid forecasting model.

distributed data of these influencing factors in the next three years; then the distributed data of these influencing factors will be used to predict China's power grid investment in 2018-2020 based on the EMD-GASVM-RBFNN hybrid prediction model. Finally, the distributed data of the China's power grid investment from 2018 to 2020 will be obtained.

The main structure of this article is arranged as follows: Section 2 introduces the methodology; Section 3 carries out empirical analysis to verify the validity of the proposed model for the power grid investment prediction in China; Section 4 uses the Monte Carlo dynamic simulation method to simulate the power grid investment in China from 2018 to 2020; Section 5 summarizes the whole paper.

2. Methodology

2.1. Empirical Mode Decomposition. Empirical mode decomposition (EMD) was first proposed by Huang et al. in 1998 [37]. This method can effectively process nonlinear and nonstationary data signals. It is an adaptive signal processing method. This method decomposes the original signal into a series of intrinsic mode functions (IMFs) and a residue which represent different time scales. IMF is a function that satisfies the following two conditions: (1) in the entire data set, the number of extrema and the number of zero crossings must either be equal or differ at most by one; (2) at any point, the mean value of the envelopes defined by the local maxima and the local minima must be zero.

The specific steps of the empirical mode decomposition are as follows:

- (1) Determine all the local extreme points of the original sequence $x(t)$
- (2) Connect all local extrema by a cubic spline line to generate its upper and lower envelopes $m_1(t)$ and $m_2(t)$

- (3) Compute the point-by-point envelope mean $m(t)$ from upper and lower envelopes; i.e., $m(t) = 1/2[m_1(t) + m_2(t)]$
- (4) Then extract the details, $h_1^1(t) = x(t) - m(t)$. If $h_1^1(t)$ meets the properties of IMF, then $c_1(t) = h_1^1(t)$ will be extracted as the first IMF signal. If $h_1^1(t)$ is not an IMF, replace $x(t)$ with $h_1^1(t)$, and then repeat the above steps until the extracted $h_1^k(t)$ is an IMF, and $c_1(t) = h_1^k(t)$ will be extracted as the first IMF signal

- (5) If an IMF is derived and meantime replace $x(t)$ with the residue $r_1(t) = x(t) - c_1(t)$, repeat steps (1)-(4) until the stop criterion is satisfied

The EMD method decomposes the original data sequence into several IMF components and a residue. The initial data sequence can be expressed by

$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (1)$$

$c_i(t)$ ($i=1,2,3..n$) are the IMF components of the initial signal sequence, n is the number of IMF components, and $r_n(t)$ is the residue of the initial sequence.

The EMD method has several distinct advantages. First, it is relatively easy to understand and implement. Second, the fluctuations within a time series are automatically and adaptively selected from the time series, and it is robust for nonlinear and nonstationary time series decomposition. Third, it lets the data speak for themselves. EMD method can adaptively decompose a time series into several independent IMF components and one residual component. The IMFs and the residual component displaying linear and nonlinear behavior depend only on the nature of the time series being studied.

2.2. GASVM Model. The support vector machines (SVMs) were proposed by Vapnik. It is a statistical learning algorithm

based on VC (Vapnik-Chervonenkis) dimension theory and structural risk minimization (SRM) principle [21, 23]. SVM has a strong capacity for processing nonlinear data. The basic principle of support vector machines is to find a nonlinear mapping function $\phi(x)$ to make the linear inseparable data x in the low-dimensional feature space projected into the high-dimensional feature space F to make it a linear separable problem. Then the following function $f(x)$ is used to carry out linear regression in the high-dimensional feature space F .

$$f(x) = [\omega \cdot \phi(x)] + b, \quad \phi: \mathbf{R}^m \rightarrow F, \quad \omega \in F \quad (2)$$

In the formula, ω is a weight vector; b is a bias. The problem of the function approximate is equivalent with minimizing the following problem:

$$R = \frac{1}{2} \|\omega\|^2 + C \frac{1}{l} \sum_{i=1}^l |y_i - f(x_i)|_\varepsilon \quad (3)$$

y_i is the expected output while $f(x_i)$ is the actual output of the sample i . l is the number of samples. $\|\omega\|^2$ is the weights vector norm, which is used to constrain the model structure capacity in order to obtain better generalization performance. C is the regularized constant determining the trade-off between the empirical error and the regularization term.

With the introduction of Vapnik's ε -insensitive loss function [38], we adopted Vapnik's linear loss function with ε -insensitive zone as a measure for empirical error which is shown below:

$$|y_i - f(x_i)|_\varepsilon = \begin{cases} 0 & \text{if } |y_i - f(x_i)| \leq \varepsilon \\ |y_i - f(x_i)| - \varepsilon & \text{otherwise} \end{cases} \quad (4)$$

In the formula, ε is a specified parameter. $|y_i - f(x_i)|_\varepsilon$ is called ε -insensitive loss function. Then the support vector machine (SVM) determines the regression function by minimizing the objective function, and the objective function and constraints are as follows:

$$\min \left\{ \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i^* + \xi_i) \right\} \quad (5)$$

subject to

$$\begin{aligned} y_i - \omega \cdot \phi(x) - b &\leq \varepsilon + \xi_i^* \\ (\omega \cdot \phi(x)) + b - y_i &\leq \varepsilon + \xi_i \\ \xi_i^*, \xi_i &\geq 0 \end{aligned} \quad (6)$$

$i = 1, 2, \dots, l$. ξ_i^* and ξ_i are slack variables. With the Lagrange multipliers introduced, the support vector regression function can be obtained by solving the dual problem [22, 23, 31]:

$$f(x) = \sum_{i=1}^l (a_i - a_i^*) K(X_i, X) + b \quad (7)$$

a_i, a_i^* are Lagrange multipliers. $K(X_i, X)$ is kernel function, then linear and RBF kernel function would be adopted in this paper which are shown as follows, respectively:

$$K(x_i, x_j) = x_i^T x_j \quad (8)$$

$$K(x_i, x_j) = \exp\left(-\frac{1}{2\sigma^2} \|x_i - x_j\|^2\right) \quad (9)$$

σ is RBF kernel function parameter while the linear kernel function has no parameter. In summary, penalty parameter C , parameter ε , and RBF kernel function parameter σ are parameters that need to be determined. Therefore, genetic algorithm (GA) is introduced to optimize the parameters C , ε , and σ .

Genetic algorithm (GA) was first proposed by Professor Holland in the 1970s. It is an adaptive global optimization algorithm based on the experience of biological natural genetic mechanism. The genetic algorithm improves the individual fitness by imitating the selection, crossover, and mutation mechanism of biological genetics, through continuous evolution and iteration; it adaptively controls the search process to get the optimal solution. The genetic algorithm (GA) is a method of parallel stochastic searching optimization, which has been widely used in combinatorial optimization, machine learning, signal processing, and adaptive control [39]. The operation steps of GA-SVM are described in detail in some literatures [16, 40]. The GA-SVM model structure is illustrated in Figure 4.

2.3. RBF Neural Network. Radial basis function (RBF) neural network is a three-layer forward neural network model with good performance and global approximation and is free from the local minima problems [41]. It is a multi-input, single-output system consisting of an input layer, a hidden layer, and an output layer. During the data processing, the hidden layer performs nonlinear transforms for feature extraction and the output layer gives a linear combination of output weights.

The transformation of the RBF neural network from the input space to the hidden layer space is nonlinear, and the transformation from the hidden layer space to the output layer space is linear. RBF neural network has a simple structure, simple training, and fast learning convergence, which can fit any nonlinear function, so it is widely used in time series prediction [41–44].

The RBF neural network commonly uses the Gauss radial basis function as the activation function of the hidden layer neurons. The Gauss radial basis function can be expressed as follows:

$$R(x_p - c_i) = \exp\left(-\frac{1}{2\sigma^2} \|x_p - c_i\|^2\right) \quad (10)$$

in which x_p is the input to the hidden layer, c_i is the center of the Gaussian function, $\|x_p - c_i\|$ is the distance between the input vector and the center of the Gaussian function, and σ is the variance of the Gaussian function. The output of the network can be expressed as follows:

$$y_j = \sum_{i=1}^h \omega_{ij} \exp\left(-\frac{1}{2\sigma^2} \|x_p - c_i\|^2\right) \quad j = 1, 2, \dots, n \quad (11)$$

FIGURE 4: GA-SVM model structure.

In the formula, $x_p = (x_1^p, x_2^p, \dots, x_m^p)^T$ is the p th input sample of the network; $p=1, 2, \dots, P$; P is the total number of samples; $i=1, 2, \dots, h$ is the number of hidden layer nodes; y_i is the actual output of the j th output node of the network corresponding to the input sample.

There are three parameters that need to be solved in the RBF neural network model: the center c_i of the radial basis function, the variance σ of the radial basis function, and the weight ω_{ij} from the hidden layer to the output layer.

According to the different selection methods of radial basis function centers, there are many learning methods for parameters c_i and σ , such as random selection center method, self-organization selection method, supervised selection center method, and orthogonal least squares method [45]. The least square algorithm is also applied to train the output weight ω_{ij} [41].

3. Case Study

This section uses the data of China's power grid investment and influencing factors of the investment from 1990 to 2017 to verify the effectiveness and accuracy of the EMD-GASVM-RBF hybrid forecasting model proposed in this paper. And the primary industry added value, the secondary industry added value, the tertiary industry added value, urbanization rate, the electricity installed capacity, total electricity consumption, clean energy power generation ratio, and population are selected as the influencing factors of the power grid investment.

Firstly, the EMD method is used to decompose the original sequence of the power grid investment demand, and two IMF components and a residual component are decomposed. As we can see from Figure 5, the IMF1

component has significant random characteristics while the IMF2 component has significant periodic characteristics; they reflect the different internal characteristics of China's power grid investment demand, respectively. The residual component reflects the development trend of the power grid investment in China. In view of the different characteristics of each subsequence, we will use appropriate models to predict different subsequences separately. In this paper, the IMF1 subsequence will be forecasted by RBF neural network while IMF2 and residual subsequences will be forecasted by GA-SVM model. Finally, the prediction results of different subsequences will be aggregated to get the final power grid investment forecasting result.

In this paper, we use the "newrb" function from MATLAB platform to create the RBF neural network for the prediction of IMF1 component; we use the historical data of the past 8 years of the IMF1 component as the inputs of the RBF network to predict the next year's IMF1 data, so only 20 data samples are used to build the prediction model for IMF1. The activation functions of the hidden layer are all set as Gauss radial basis functions, the error margin is set to $1e-8$, the spreading factors are set to 2, and the maximum number of the hidden layer neurons is set to 100.

The GA-SVM models are built to forecast IMF2 component and residual component. The parameters of each model are set as shown in Table 5.

As we can see from the table, the main difference between the GA-SVM models for IMF2 prediction and residue prediction is the type of kernel function. The eigenmode component IMF2 with high wave frequency and high complexity will be predicted by the radial basis function with better generalization ability while the residual component with significant linear and trend characteristics will be predicted by linear

TABLE 5: Parameters of GA-SVM model for IMF2 and residue prediction.

Model	G	NP	TOS	TKF
GA-SVM for IMF2	200	20	Epsilon-SVR	Radial Basis Function
GA-SVM for residue	200	20	Epsilon-SVR	Linear

*G and NP represent the maximum evolutionary algebra and population size of genetic algorithm, respectively. TOS and TKF represent the type of SVM and the type of kernel function of the SVM model, respectively.

FIGURE 5: EMD decomposition diagram (unit/100 million yuan).

kernel function [10]. Moreover, we use the historical data of the past 10 years of the IMF2 component as the inputs of the GA-SVM model to predict the next year's IMF2 data, so only 18 data samples are used to build the prediction model for IMF2. The residual component reflects a trend characteristic of the power grid investment; it is directly affected by the level of economic and social development as well as electricity production and consumption, so the primary industry added value, the secondary industry added value, the tertiary industry added value, urbanization rate, the electricity installed capacity, total electricity consumption, clean energy power generation ratio, and population are set

as inputs of the GA-SVM model for the forecasting of the residual component.

In order to ensure the reliability of the experimental results of the prediction model under the case of a small sample size, leave-one-out cross validation is selected to do the cross validation of the sample. The basic idea of the leave-one-out cross validation is that one of the N samples is taken as the validation sample, and the remaining $N-1$ samples are used as training sets for the GA-SVM model and repeat N times until all N samples are taken as validation sets in turn. In this paper, the leave-one-out cross validation method will repeat 10 times for each sample; then the average value of the leave-one-out cross validation results for each sample will be regarded as the final forecasting result of the forecasting model for each sample.

The forecasting results of the IMF1, IMF2, and residual components by different forecasting model and the final power grid investment forecasting result are showed in Figure 6. From Figure 6 we can intuitively find out that the forecasting results of the IMF2 component and the residual component are basically consistent with the actual value. Although the RBF neural network has a poor prediction accuracy of the IMF1 component compared with the GA-SVM model for IMF2 and residual component prediction because of the strong random and nonlinear characteristics of the IMF1 component, it still effectively reflects the development trend of the IMF1 component. Besides, the proportion of the IMF1 component is relatively low compared with the IMF2 component and the residual component in the original data, so the prediction accuracy of the IMF1 component has lower impact on the forecasting result of the final power grid investment than the IMF2 component and residual component. The MAPE and RMSE of the final forecasting result of the power grid investment from 2000 to 2017 are only 6.69% and 167.08 hundred million yuan, respectively. The forecasting result is basically consistent with the actual value which confirms the accuracy and effectiveness of the EMD hybrid power grid investment forecasting model.

To further verify the validity and accuracy of the EMD-GASVM-RBFNN hybrid power grid investment forecasting model, the forecasting results of GA-SVM model and BP neural network model will be comparisons of the EMD-GASVM-RBFNN hybrid power grid investment forecasting model. A BP neural network with two hidden layers is built for the power grid investment forecasting. After repeated tests, the numbers of neurons in the hidden layers are both set to 10 to get a relatively good forecasting performance of BP neural network model; the hidden layer neuron activation functions are both set to "logsig" functions; the learning algorithm of the network is set to Levenberg-Marquardt (LM)

FIGURE 6: Subsequences forecasting results and the final power grid investment forecasting result by EMD-GASVM-RBFNN hybrid model. ((a), (b), (c), and (d) represent the IMF1, IMF2, residue, and final power grid investment forecasting results, respectively. The unit is 100 million yuan.)

algorithm; the error margin of the network is set to $1e-8$; the learning rate of the network is set to 0.01. Primary industry added value, secondary industry added value, tertiary industry added value, urbanization rate, electricity installed capacity, total electricity consumption, clean energy power generation ratio, and population are set as the inputs of the BP neural network. The inputs of the GA-SVM power grid investment forecasting model are the same as the inputs of the BP neural network, and the type of the kernel function is set to radial basis function. The leave-one-out cross validation is also used to obtain the power grid investment forecasting results for both models from 1990 to 2017. The power grid investment prediction results of the GA-SVM model and BP neural network model are shown in Figure 7.

It can be intuitively seen from Figure 7 that the forecasting results based on the GA-SVM model and BP neural network model are not quite good enough compared with the EMD hybrid power grid investment forecasting model. The mean absolute percentage error (MAPE) and the root mean square error (RMSE) of the forecasting results from 2000 to 2017 will be the evaluation criteria for quantitative analysis of different forecasting models. The formulas for the calculation of MAPE and RMSE are as follows:

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y'_i - y_i}{y_i} \right| \quad (12)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y'_i - y_i)^2} \quad (13)$$

N is the number of the samples and y'_i is the prediction result of the power grid investment model while y_i is the actual value of the power grid investment.

The normalized power grid investment prediction results and the MAPE and RMSE results of different forecasting model are shown in Table 6.

Based on the calculation results of the MAPE and RMSE, we can conclude that the EMD hybrid forecasting model has a better prediction performance than the GA-SVM model and BP neural network. Although the MAPE of the EMD hybrid model is just slightly lower than the GA-SVM model, the RMSE of the EMD hybrid model is considerably lower than the GA-SVM and the BP neural network model; as we can see from Table 6, the RMSE of the EMD hybrid model is just about half of the GA-SVM and the BP neural network

FIGURE 7: (a) and (b) represent the power grid investment forecasting results by GA-SVM model and RBFNN model, respectively. The unit is 100 million yuan.

TABLE 6: Comparison of forecasting results of different forecasting models from 2000 to 2017.

Year	AV	EMD Model		GA-SVM		BPNN	
		FV	APE/%	FV	APE/%	FV	APE/%
2000	-0.50	-0.57	12.78	-0.64	25.70	-0.61	21.27
2001	-0.56	-0.47	20.14	-0.55	3.05	-0.56	0.23
2002	-0.46	-0.58	22.00	-0.55	15.92	-0.54	14.46
2003	-0.55	-0.57	5.24	-0.46	19.51	-0.39	34.78
2004	-0.54	-0.43	24.54	-0.52	5.65	-0.38	33.72
2005	-0.45	-0.44	2.42	-0.42	5.17	-0.36	16.55
2006	-0.24	-0.29	5.82	-0.28	4.72	-0.26	2.48
2007	-0.11	-0.10	0.64	-0.09	1.66	-0.10	0.79
2008	0.06	0.11	5.03	0.06	0.38	0.10	4.06
2009	0.43	0.32	7.79	0.14	20.36	0.15	19.31
2010	0.26	0.30	2.76	0.43	13.31	0.40	10.38
2011	0.35	0.32	2.25	0.26	7.01	0.26	6.82
2012	0.34	0.33	0.60	0.31	2.40	0.33	0.61
2013	0.42	0.42	0.39	0.40	1.07	0.44	1.55
2014	0.51	0.50	0.93	0.59	5.09	0.69	11.43
2015	0.71	0.66	2.52	0.83	6.99	0.73	1.27
2016	1.00	1.03	1.47	0.83	8.54	0.74	12.69
2017	0.96	0.90	3.12	1.12	8.33	1.07	5.79
MAPE/%		6.69		8.60		11.01	
RMSE		167.08		311.71		345.77	

*AV and FV represent the actual value and the forecasting value of the power grid investment, respectively. APE, MPAE, and RMSE represent the absolute percentage error, mean absolute percentage error, and root mean square error, respectively.

model. Figure 8 gives the absolute percentage error (APE) of the power grid investment forecasting results by different prediction models from 2000 to 2017.

4. Scenario Analysis

This section dynamically simulates China's power grid investment demand from 2018 to 2020 based on the EMD hybrid

forecasting model and Monte Carlo method. First, the RBF neural network model and GA-SVM are used to predict the IMF1 and IMF2 components separately in advance for three years. The prediction results of each IMF component in the next three years are shown in Table 7. Then, the residual component from 2018 to 2020 is dynamically simulated based on the GA-SVM model and the Monte Carlo method. Finally, the simulation results of the power grid investment demand

FIGURE 8: The absolute percentage error from 2000 to 2017 of different forecasting models.

TABLE 7: Prediction results of the IMF components from 2018 to 2020.

	2018	2019	2020
IMF1	-105.4	176.0	-167.4
IMF2	-69.7	-549.6	-433.9

*The unit is 100 million yuan.

TABLE 8: Setting results of the urbanization ratio, clean energy installed ratio, and growth rate of population from 2018 to 2020.

Influencing Factors	2018	2019	2020
Urbanization Ratio	60%	61%	62%
Clean Energy Installed Ratio	38%	39%	40%
Growth Rate of Population	5%	5%	5%

in the next three years can be obtained by aggregating the results of the residue simulation result and the predicted IMF components.

In Monte Carlo simulation, the distribution of the growth rate of each variable of the model inputs should be assumed, and different distribution assumptions will have different effects on the residue prediction result. This paper first assumes that the growth rate of the input variables is normal distribution; then the K-S (Kolmogorov-Smirnov) method will be used to verify whether the growth rate of the input variables is in accordance with normal distribution.

When the K-S method is used to test the normality of the data distribution, it is necessary to test the difference between the detection sequence and the standard normal distribution sequence. If the result of the significance test is more than 0.05, the detection sequence is considered to be normal distribution. In this paper, the urbanization ratio and the clean energy installed ratio in the next three years are set as a fixed value, respectively, while the growth rate of the population in the next three years is set to 5%. The specific setting results are shown in Table 8.

The K-S test method is used to test whether the growth rates of primary industry, secondary industry, tertiary

industry, electricity installed capacity, and total electricity consumption meet the normal distribution based on the historical data. The results of the K-S test are shown in Table 9. It can be seen from the table that the results of the exact significance test (2-tailed) of each input variable are all greater than 0.05, indicating that the growth rates of the input variables are all in accordance with normal distribution. Figure 9 also gives the frequency distribution map of the growth rate of each variable. It can be seen from the graph that the input variables' growth rates have approximately normal distribution. Because the historical data of the variables are limited, the mean and standard deviation of the sample data of these input variables will be approximated as the mean and standard deviation of the corresponding normal distribution.

After verifying the normal distribution of the growth rate of each input variable by K-S test, the Monte Carlo method will be used to create 100 thousand groups' next year's input variables. Then the distribution data of the next year's input variables are calculated on the basis of the growth rate results of each input variable simulated by the Monte Carlo method and the data of the input variables in this year, and the distribution data of the next year's residue will be predicted based on the GA-SVM model and the next year's variables data. Finally, the simulation results of the power grid investment in the next year can be obtained by aggregating the results of the residual simulation result and the predicted IMF components. Figure 10 shows the distribution data of power grid investment obtained by Monte Carlo simulation from 2018 to 2020. The average value of each input variable's distribution data in 2018 obtained by Monte Carlo simulation will be the basis for the calculation of the distribution data of the variables in 2019, and the average value of the distribution data of each variable in 2019 will be the basis for the calculation of the distribution data of each variable in 2020.

From Figure 10, we can intuitively learn that the power grid investments in the next three years are all basically in line with normal distribution. Table 10 shows the statistical test results of the Monte Carlo dynamic simulation results of the power grid investment distribution data in the next three years. It can be seen from the table that the exact significance test results of the K-S method from 2018 to 2020 are all far more than 0.05; the absolute value of skewness and kurtosis coefficients of the frequency distribution curves are all less than 0.012; it shows that the distribution data of the power grid investment demand are highly in accordance with the normal distribution and the fitting precision is satisfactory.

The Monte Carlo simulation results show that the average value of the power grid investment simulated by the Monte Carlo method in the next three years will reach 4837.6, 4980.2, and 5109.2 hundred million yuan, respectively. Table 11 and Figure 11 give the interval probability values and the development trends of the power grid investment for the next three years, respectively, based on the property of normal distribution, in which the parameter μ and the parameter σ represent the mean value and the standard deviation of the normal distribution, respectively.

As can be seen from Table 11 and Figure 11, the scenario forecasting results based on the EMD hybrid model and the

FIGURE 9: Frequency distribution map of the variables' growth rate.

TABLE 9: Statistical characteristics and the K-S exact significance test (2-tailed) results of the variables' growth rate.

Variable	Number of Samples	Mean	Median	Std. Deviation	Exact Sig. (2-tailed)
PI	65	0.0875	0.0767	0.0882	0.4140
SI	65	0.1362	0.1296	0.1402	0.3235
TI	65	0.1310	0.1329	0.1062	0.3982
EIC	39	0.0925	0.0907	0.0325	0.1066
TEC	38	0.0870	0.0881	0.0367	0.8865

*PI, SI, TI, EIC, and TEC represent primary industry, secondary industry, tertiary industry, electricity installed capacity, and total electricity consumption, respectively.

Monte Carlo method show that the distribution interval of the power grid investment in years 2018-2020 will remain stable, which has no significant increase or decrease trend in the next three years; China's power grid investment in the next three years will remain between 4300 and 5700 hundred million yuan as a whole. In the light of the mean value of the simulation results, the power grid investment in 2018 will reduce for about 10% from 2017 to 5000.6 hundred million yuan; then there will be a slight increase of the power grid investment in 2019 and 2020, reaching 5160.6 and 5303.0 hundred million yuan, respectively.

Compared with the rapid growth of the power grid investment from 2012 to 2016, the simulation result indicates that the power grid investment will not substantially increase in the next three years. At the present stage, China's economic development has entered in "new normal" state, in which the economic growth is slowing down and the economic structure is constantly adjusting. The continuous decline of the proportion of the secondary industry has also curbed

TABLE 10: Statistical test of the Monte Carlo simulation results from 2018 to 2020.

Year	2018	2019	2020
Number of Samples	100000	100000	100000
Mean	4837.6	4980.2	5109.2
Std. Deviation	163	180.5	193.8
Exact Sig. (2-tailed)	0.9076	0.7515	0.5115
Skewness	-0.003	0.008	-0.002
Kurtosis	0.008	-0.012	-0.01

*The unit is 100 million yuan.

the rapid growth of the total electricity consumption. In addition, the oversupply of electricity and the low utilization rate of clean energy power generation need to be solved urgently, and a new round of electricity market reform has further promoted the process of marketization. Under the above backgrounds, the focus of power grid investment in

FIGURE 10: Frequency distribution graph of the power grid investment demand based on the Monte Carlo method in the next three years (the unit is 100 million yuan).

TABLE 11: Interval probability statistics results of the forecasting power grid investment from 2018 to 2020.

Year	Interval					
	$\mu - \sigma$	$\mu + \sigma$	$\mu - 2\sigma$	$\mu + 2\sigma$	$\mu - 3\sigma$	$\mu + 3\sigma$
2018	4674.5	5000.6	4511.5	5163.7	4348.4	5326.7
2019	4799.7	5160.6	4619.2	5341.1	4438.7	5521.6
2020	4915.3	5303.0	4721.5	5496.8	4527.7	5690.7
Interval Probability	68.27%		95.45%		99.73%	

*The unit is 100 million yuan.

FIGURE 11: Interval statistics graph of the forecasting power grid investment demand from 2018 to 2020.

the next few years will be improving the overall efficiency and safety of the power grid rather than continuing to expand the existing enormous scale of the power grid. On the other hand, in order to support the development of China's total economic volume and meeting the demand for electricity consumption, the power grid investment demand is unlikely to show a significant downward trend. Therefore, the stable

development trend of the power grid investment demand in the next three years would be reasonable.

5. Conclusion

In the light of the complicated nonlinear and nonstationary characteristics, this paper proposed an EMD-GASVM-RBFNN hybrid forecasting model which can effectively identify and extract the internal characteristics of China's power grid investment. Firstly, the EMD method is used to decompose the original power grid investment into two IMFs and a residual component. Then the IMF1 component with high frequency will be forecasted by RBF neural network while the IMF2 component with low frequency and the residual component will be forecasted by GA-SVM model. Finally, the prediction results of different components will be aggregated to get the final power grid investment forecasting result. Next, the forecasting results of GA-SVM model and BPNN model will be comparisons of the EMD-GASVM-RBFNN hybrid forecasting model. The RMSE and MAPE calculation results of different forecasting models indicate that the EMD-GASVM-RBFNN hybrid model has a better prediction performance than GA-SVM model and BPNN model. Finally, the Monte Carlo method is used to simulate China's power grid investment in 2018-2020. The simulation

result indicates that the development trend of China's power grid investment will remain stable in the next three years.

Data Availability

The data used in this paper are all from «China Electric Power Statistics Yearbook» and National Bureau of Statistics: <http://www.stats.gov.cn>.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

Jinchao Li contributed to the conception and design. Shaowen Zhu contributed to the computation. Qianqian Wu and Pengfei Zhang collected and interpreted the data. All of the authors drafted and revised the manuscript together and approved its final publication.

Acknowledgments

This work has been supported by “Ministry of Education, Humanities and Social Science Fund, no. 15YJC630058”, “Beijing Social Science Fund, no. 18GLB023”, “the Fundamental Research Funds for the Central Universities, no. 2017MS083”, and “the Science and Technology Project of SGCC”.

References

- [1] S. Dai, D. Niu, and Y. Han, “Forecasting of power grid investment in china based on support vector machine optimized by differential evolution algorithm and grey wolf optimization algorithm,” *Applied Sciences*, vol. 8, no. 4, p. 636, 2018.
- [2] C. Kang, Q. Xia, and B. Zhang, “Review of power system load forecasting and its development,” *Automation of Electric Power Systems*, vol. 28, no. 17, pp. 1–11, 2004.
- [3] B.-C. Hu, G. Hu, C.-H. Hu, S. Qing, M.-W. Li, and C. Peng, “Grid infrastructure investment calculation model based on gray prediction,” *Journal of University of Electronic Science and Technology of China*, vol. 42, no. 6, pp. 890–894, 2013.
- [4] H. Zhao, L. Yang, C. Li, and X. Ma, “Research on prediction to investment demand of power grid based on co-integration theory and error correction model,” *Power System Technology*, vol. 35, no. 9, pp. 193–198, 2011.
- [5] C.-F. Chen, M.-C. Lai, and C.-C. Yeh, “Forecasting tourism demand based on empirical mode decomposition and neural network,” *Knowledge-Based Systems*, vol. 26, pp. 281–287, 2012.
- [6] A. J. Conejo, J. Contreras, R. Espínola, and M. A. Plazas, “Forecasting electricity prices for a day-ahead pool-based electric energy market,” *International Journal of Forecasting*, vol. 21, no. 3, pp. 435–462, 2005.
- [7] V. Ş. Ediger and S. Akar, “ARIMA forecasting of primary energy demand by fuel in Turkey,” *Energy Policy*, vol. 35, no. 3, pp. 1701–1708, 2007.
- [8] D. Akay and M. Atak, “Grey prediction with rolling mechanism for electricity demand forecasting of Turkey,” *Energy*, vol. 32, no. 9, pp. 1670–1675, 2007.
- [9] R. G. Kavasseri and K. Seetharaman, “Day-ahead wind speed forecasting using f-ARIMA models,” *Journal of Renewable Energy*, vol. 34, no. 5, pp. 1388–1393, 2009.
- [10] V. Bianco, O. Manca, and S. Nardini, “Electricity consumption forecasting in Italy using linear regression models,” *Energy*, vol. 34, no. 9, pp. 1413–1421, 2009.
- [11] J.-F. Li, B.-H. Zhang, G.-L. Xie, Y. Li, and C.-X. Mao, “Grey predictor models for wind speed-wind power prediction,” *Power System Protection and Control*, vol. 38, no. 19, pp. 151–160, 2010.
- [12] H. T. Pao, H. C. Fu, and C. L. Tseng, “Forecasting of CO₂ emissions, energy consumption and economic growth in China using an improved grey model,” *Energy*, vol. 40, no. 1, pp. 400–409, 2012.
- [13] P. A. González and J. M. Zamarreño, “Prediction of hourly energy consumption in buildings based on a feedback artificial neural network,” *Energy and Buildings*, vol. 37, no. 6, pp. 595–601, 2005.
- [14] P. F. Pai and C. S. Lin, “A hybrid ARIMA and support vector machines model in stock price forecasting,” *Omega*, vol. 33, no. 6, pp. 497–505, 2005.
- [15] Z.-D. Zou, Y.-M. Sun, and Z.-S. Zhang, “Short-term load forecasting based on recurrent neural network using ant colony optimization algorithm,” *Power System Technology*, vol. 29, no. 3, pp. 59–63, 2005.
- [16] P.-F. Pai and W.-C. Hong, “Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms,” *Electric Power Systems Research*, vol. 74, no. 3, pp. 417–425, 2005.
- [17] T. G. Barbounis, J. B. Theocharis, M. C. Alexiadis, and P. S. Dokopoulos, “Long-term wind speed and power forecasting using local recurrent neural network models,” *IEEE Transactions on Energy Conversion*, vol. 21, no. 1, pp. 273–284, 2006.
- [18] N. Amjady, “Day-ahead price forecasting of electricity markets by a new fuzzy neural network,” *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 887–896, 2006.
- [19] P.-F. Pai and W.-C. Hong, “A recurrent support vector regression model in rainfall forecasting,” *Hydrological Processes*, vol. 21, no. 6, pp. 819–827, 2007.
- [20] J. P. S. Catalão, S. J. P. S. Mariano, V. M. F. Mendes, and L. A. F. M. Ferreira, “Short-term electricity prices forecasting in a competitive market: a neural network approach,” *Electric Power Systems Research*, vol. 77, no. 10, pp. 1297–1304, 2007.
- [21] N. Sapankevych and R. Sankar, “Time series prediction using support vector machines: a survey,” *IEEE Computational Intelligence Magazine*, vol. 4, no. 2, pp. 24–38, 2009.
- [22] D. Niu, Y. Wang, and D. D. Wu, “Power load forecasting using support vector machine and ant colony optimization,” *Expert Systems with Applications*, vol. 37, no. 3, pp. 2531–2539, 2010.
- [23] J. Li, J. Liu, and J. Wang, “Mid-long term load forecasting based on simulated annealing and SVM algorithm,” *Proceedings of the Chinese Society of Electrical Engineering*, vol. 31, no. 16, pp. 63–66, 2011.
- [24] W.-C. Hong, “Application of seasonal SVR with chaotic immune algorithm in traffic flow forecasting,” *Neural Computing and Applications*, vol. 21, no. 3, pp. 583–593, 2012.
- [25] H. Li, S. Guo, C. Li, and J. Sun, “A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm,” *Knowledge-Based Systems*, vol. 37, pp. 378–387, 2013.

- [26] J. Geng, M.-L. Huang, M.-W. Li, and W.-C. Hong, "Hybridization of seasonal chaotic cloud simulated annealing algorithm in a SVR-based load forecasting model," *Neurocomputing*, vol. 151, no. 3, pp. 1362–1373, 2015.
- [27] A. J. Conejo, M. A. Plazas, R. Espínola, and A. B. Molina, "Day-ahead electricity price forecasting using the wavelet transform and ARIMA models," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 1035–1042, 2005.
- [28] Z. A. Bashir and M. E. El-Hawary, "Applying wavelets to short-term load forecasting using PSO-based neural networks," *IEEE Transactions on Power Systems*, vol. 24, no. 1, pp. 20–27, 2009.
- [29] R.-A. Hooshmand, H. Amooshahi, and M. Parastegari, "A hybrid intelligent algorithm based short-term load forecasting approach," *International Journal of Electrical Power & Energy Systems*, vol. 45, no. 1, pp. 313–324, 2013.
- [30] L. Yu, S. Wang, and K. K. Lai, "Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm," *Energy Economics*, vol. 30, no. 5, pp. 2623–2635, 2008.
- [31] L. Ye and P. Liu, "Combined Model Based on EMD-SVM for Short-term Wind Power Prediction," *Proceedings of the CSEE*, vol. 31, pp. 102–108, 2011.
- [32] N. An, W. Zhao, J. Wang, D. Shang, and E. Zhao, "Using multi-output feedforward neural network with empirical mode decomposition based signal filtering for electricity demand forecasting," *Energy*, vol. 49, no. 1, pp. 279–288, 2013.
- [33] G. Fan, L. Peng, X. Zhao, and W. Hong, "Applications of Hybrid EMD with PSO and GA for an SVR-Based Load Forecasting Model," *Energies*, vol. 10, no. 11, p. 1713, 2017.
- [34] L. Karthikeyan and D. N. Kumar, "Predictability of nonstationary time series using wavelet and EMD based ARMA models," *Journal of Hydrology*, vol. 502, pp. 103–119, 2013.
- [35] W. Y. Duan, Y. Han, L. M. Huang, B. B. Zhao, and M. H. Wang, "A hybrid EMD-SVR model for the short-term prediction of significant wave height," *Ocean Engineering*, vol. 124, pp. 54–73, 2016.
- [36] S. Wang, N. Zhang, L. Wu, and Y. Wang, "Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method," *Journal of Renewable Energy*, vol. 94, pp. 629–636, 2016.
- [37] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings A*, vol. 454, pp. 903–995, 1998.
- [38] V. Vapnik, S. Golowich, and A. Smola, "Support vector machine for function approximation, regression estimation, and signal processing," *Advances in Neural Information Processing Systems*, vol. 9, pp. 281–287, 1996.
- [39] K. S. Tang, K. F. Man, S. Kwong, and Q. He, "Genetic algorithms and their applications," *IEEE Signal Processing Magazine*, vol. 13, no. 6, pp. 22–37, 1996.
- [40] C.-H. Wu, G.-H. Tzeng, Y.-J. Goo, and W.-C. Fang, "A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting bankruptcy," *Expert Systems with Applications*, vol. 32, no. 2, pp. 397–408, 2007.
- [41] Z. Yun, Z. Quan, S. Caixin, L. Shaolan, L. Yuming, and S. Yang, "RBF neural network and ANFIS-based short-term load forecasting approach in real-time price environment," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 853–858, 2008.
- [42] L. Yu, K. K. Lai, and S. Y. Wang, "Multistage RBF neural network ensemble learning for exchange rates forecasting," *Neurocomputing*, vol. 71, no. 16–18, pp. 3295–3302, 2008.
- [43] H. Jiang, Y. Dong, J. Wang, and Y. Li, "Intelligent optimization models based on hard-ridge penalty and RBF for forecasting global solar radiation," *Energy Conversion and Management*, vol. 95, pp. 42–58, 2015.
- [44] C.-M. Lee and C.-N. Ko, "Time series prediction using RBF neural networks with a nonlinear time-varying evolution PSO algorithm," *Neurocomputing*, vol. 73, no. 1–3, pp. 449–460, 2009.
- [45] S. Chen, C. F. N. Cowan, and P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 2, no. 3, pp. 302–309, 1991.