

Research Article

Short-Term Wind Speed Forecasting Study and Its Application Using a Hybrid Model Optimized by Cuckoo Search

Xuejun Chen,^{1,2} Shiqiang Jin,³ Shanshan Qin,^{4,5} and Laping Li⁶

¹Key Laboratory of Arid Climatic Change and Reducing Disaster of Gansu Province/Key Open Laboratory of Arid Climatic Change and Disaster Reduction of Chinese Meteorological Administration, Institute of Arid Meteorology, Chinese Meteorological Administration, Lanzhou 730020, China

²Gansu Meteorological Service Center, Lanzhou, Gansu 730020, China

³School of Mathematics & Statistics, Lanzhou University, Lanzhou, Gansu 730000, China

⁴MOE Key Laboratory of Western China's Environmental Systems, Research School of Arid Environment & Climate Change, Lanzhou University, Lanzhou 730000, China

⁵Scientific Information Center for Resources and Environment, Lanzhou Branch of the National Science Library, Chinese Academy of Sciences, Lanzhou 730000, China

⁶Datong Meteorological Bureau of Shanxi Province, Datong 037010, China

Correspondence should be addressed to Shiqiang Jin; jinsq1@gmail.com

Received 15 October 2014; Accepted 16 January 2015

Academic Editor: Erol Egrioglu

Copyright © 2015 Xuejun Chen 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.

The support vector regression (SVR) and neural network (NN) are both new tools from the artificial intelligence field, which have been successfully exploited to solve various problems especially for time series forecasting. However, traditional SVR and NN cannot accurately describe intricate time series with the characteristics of high volatility, nonstationarity, and nonlinearity, such as wind speed and electricity price time series. This study proposes an ensemble approach on the basis of 5-3 Hanning filter (5-3H) and wavelet denoising (WD) techniques, in conjunction with artificial intelligence optimization based SVR and NN model. So as to confirm the validity of the proposed model, two applicative case studies are conducted in terms of wind speed series from Gansu Province in China and electricity price from New South Wales in Australia. The computational results reveal that cuckoo search (CS) outperforms both PSO and GA with respect to convergence and global searching capacity, and the proposed CS-based hybrid model is effective and feasible in generating more reliable and skillful forecasts.

1. Introduction

In the contemporary energy market, the demand for electricity soars intensely due to the development of economy and society, while reserves of fossil fuel for power generation are becoming exhaustive and various ecosystem problems are increasing. Under this serious condition, renewable, clean, and nonpolluting energy becomes alternative energy for substituting fossil fuel. So wind energy becomes the one satisfying the above requirements. Meanwhile, as the increasing generation of wind power and the growth of integration of wind power into grid system, electricity generation based on wind energy resource has been playing an increasing role in China. The installed wind power capacity has been increased

by approximately 200% between 2005 and 2009 [1]. Despite the high cost of wind power plant, wind power has its unique advantages especially at remote locations which are rich in wind energy resource in China.

Wind series from the southwest of China, Wuwei City and Jinchang City in Gansu province, appear to have complex characteristics, such as high volatility, nonstationarity, and nonlinearity. In order to work efficiently on the market of the wind power, it is apparent that forecasting the wind power production is essential for farm owners and assists producers in making decisions for the sale of energy, thus increasing production and profits. If an accurate prediction of the wind speed for the following time can be evaluated, the total amount of active power that can be produced by each

generator on a wind farm can be determined. So wind speed prediction is getting more and more attention [2].

However, as the result of the complicated characteristics of wind speed, such as chaotic fluctuation, nonstationarity, and nonlinearity, forecasting has been the most challenging task. In order to predict wind speed efficiently, research in the field of forecasting the wind power or wind speed has been devoted to the development of reliable and effective tools and many different methods have been reviewed and proposed in [3–5].

As wind speed appears to be of high volatility and nonstationary, some additional techniques as preprocessing procedures are proposed to remove the irregular wind speed, such as empirical mode decomposition (EMD) [6–8] method and wavelet transform (WT) or wavelet denoising (WD) method [9–14]. WT has been widely applied to present a signal in both frequency and time domains. Wavelet transform method has been extensively applied recently in analyzing a nonstationary and high fluctuant series. It decomposes the original complicated data into several components of wavelet transform, one of which is smooth and reflects the inherent and real information. Because of the complexity of factors of wind speed fluctuation, wavelet transform as a preprocessing procedure is used so as to obtain a further excellent performance. In [13], WT is used to decompose the original wind speed series into detail signal and approximation signal to remove the abnormal fluctuation of wind speed series for further modeling. Catalão et al. [9] propose artificial neural networks combined with WT for short-term wind power forecasting in Portugal. The proposed model is both effective and novel, outperforming persistence, ARIMA, and NN approaches. EMD is based upon the local characteristic time scale of signal and could decompose the complicated signal function into a number of intrinsic model functions (IMFs) for further modeling [15].

Zhang et al. [16] said that the frequently used statistical approaches of wind speed series forecasting can be classified into statistical models and artificial intelligent algorithms (AI). The former establish time series models to predict the future speed by mining information contained in the historical signals. Time series method includes Autoregressive Integrated Moving Average (ARIMA), which is used for forecasting wind power in US wind farms in [17, 18]. A part of the models outperform the persistence model. The Autoregressive Conditional Heteroscedastic (ARCH) model is conjoined with ARIMA model to take the heteroscedasticity influence of the residual series [19] into account. An ARMA-GARCH-M framework is employed to examine the 26 regional wind power energy markets in the US using daily average wind speed [20]. It revealed that wind speed displays a characteristic of time-varying volatility and there is different relationship between the mean and volatility of wind speed series across the different locations. In [21], the ARIMA-ARCH model is employed to predict wind speed series itself. It is demonstrated that the ARIMA-ARCH model offers better performance than single ARIMA model.

Another approach is the intelligent algorithm models building a nonlinear model to fit the historical wind speed series by minimizing the training error, such as Artificial

Neural Networks (ANN). It is a widely used statistical method for many fields, such as stock price [22], electricity price [23], load forecasting [24, 25], gas consumption [26], and wind speed [27, 28]. A typical artificial neural network, Backpropagation Neural Network (BPNN) [29], is actually a mapping function relation from the vector(s) of input to output with unknowing the correlation between the data. It has been proven in mathematical theories that BPNN can implement any complicated nonlinear mapping function and approximate an arbitrary nonlinear function with satisfactory accuracy [30]. By learning the historical data pattern, BPNN can be effectively utilized to predict series in new horizon. Similarly, support vector regression (SVR) is also designed to capture the nonlinear patterns from time series [1, 31, 32]. Also, it has been observed that it can model nonlinear wind speed with an excellent performance. Nevertheless, one of disadvantages of the method is dilemma of selection of values of parameters in support vector machine because the way of selecting values for the parameters will affect the generalization performance remarkably. In this paper, chaos optimization is applied to accomplish selection of values of parameters.

As chaotic fluctuation, nonstationarity, and nonlinearity of wind speed series, hybrid models based on linear and artificial intelligence are popularly proposed in the research of wind speed series forecasting. Liu et al. [33] proposed two hybrid methods: ARIMA-ANN and ARIMA-Kalman models. ARIMA model is utilized to determine the structure of ANN and initialize the Kalman measurement and the state equations for Kalman. Su et al. [34] proposed ARIMA and Kalman filter to predict the daily mean wind speed in the west of China. To develop a novel hybrid model which is adapted to the data set and increase the fitting accuracy, this approach used Particle Swarm Optimization (PSO) to optimize the parameters of the ARIMA model. Both of them obtain good performance and are applied to the wind speed forecasting. A hybrid of ARIMA-ANN is employed in [35]. The ARIMA models were used to forecast the linear pattern and then with the obtained errors ANN were built to forecast the nonlinear tendencies that the ARIMA could not identify. It reveals that these hybrid models have a higher forecasting accuracy than the single ARIMA and ANN.

There is a large amount of research directed to the development of reliable and accurate wind speed and power prediction models. However, it is difficult to draw a conclusion of which model is the best because a model could perform well at its site, but not at other sites. In other words, a potential best forecasting model at one site does not guarantee the model to work well at another site. This paper discusses forecasting accuracy in different sites and months based on a preprocessing method and comparison between a new optimal algorithm and some conventional optimal algorithms that are used in the forecasting models. In most of the cases, the statistical tools can provide accurate results in the short-term, medium-term, and long-term prediction. However, as to the very short-term and short-term horizon, the effect of atmospheric dynamics on the wind speed becomes more important, so in these cases the use of physical approaches becomes important. This paper will explore the accuracy of very

short-term (10 minutes) of 3-step forecasting by the use of statistical approaches.

The main contributions of this paper are as follows. Several standard forecasting models (SVR, BP, and Elman) are used to forecast wind series. These models make an excellent performance, respectively. In order to improve accuracy further, another two kinds of techniques are proposed in this paper. The first kind is to use 5-3 Hanning filter and Wavelet denoising as a preprocessing procedure. The second kind is a new mate-heuristic algorithm, cuckoo search, which is introduced to optimize the parameters of SVR and compare with grid search (GS) and two conventional optimal algorithms (GA and PSO). To demonstrate that our proposed method is effective, electricity price in New South Wales is utilized to build proposed models and get satisfying results.

This paper is organized as follows. The explicit theories of the proposed approach are described in Section 3, including 5-3 Hanning filter and WD, SVR, BP, Elman, and optimal algorithms. Section 3 provides the proposed methods in this paper. In Section 4, numerical results and evaluation of forecasting performance in the case study are shown. Section 5 provides some conclusions and suggestions.

2. The Related Methodology

2.1. The Data Preprocessing Method. The proper data preprocessing can effectively remove the useless information, such as outliers and noises, in a time series. As wind speed appears to be of high volatility and nonstationary, some preprocessing procedures are introduced to remove the irregular wind speed and outliers of electricity price.

2.1.1. The Proposed 5-3 Hanning Filter (5-3H) Technique. 5-3H method is short for the medians of five-three-Hanning smoothing method (“five” is a method for a median-of-five smoothing, “three” for a median-of-three smoothing, and “H” for Hanning smoothing). This method, presented by Tukey, adopts weighted smoothing by three times to the original data to generate the ultimate smoothed estimates. Tukey introduces three steps for the signal preprocessing: five-point moving average smoothing, three-point moving average smoothing, and Hanning moving average smoothing, respectively. Flowchart of this method is illustrated in Figure 1.

Let the original data be $\{x(i), i = 1, 2, \dots, n\}$, where n is the length of time series x . And three steps are illustrated and expressed as follows.

Step 1. Five-point moving median average smoothing. The original data $x(i-2), x(i-1), x(i), x(i+1), x(i+2)$ sorted in ascending order is displayed as $x_{(i-2)}, x_{(i-1)}, x_{(i)}, x_{(i+1)}, x_{(i+2)}$. The five-point moving median average smoothing signal $\{y(i)\}$ can be presented as follows:

$$y(i) = x_{(i)} \quad (1)$$

for $i = 3, 4, \dots, n-2$, where $y(i)$ is the i th point and $x(i)$ is the i th point in the original signal. Then, the four items missing in the series $y(i)$ can be estimated as follows:

$$y(1) = y(2) = y(n-1) = y(n) = 0. \quad (2)$$

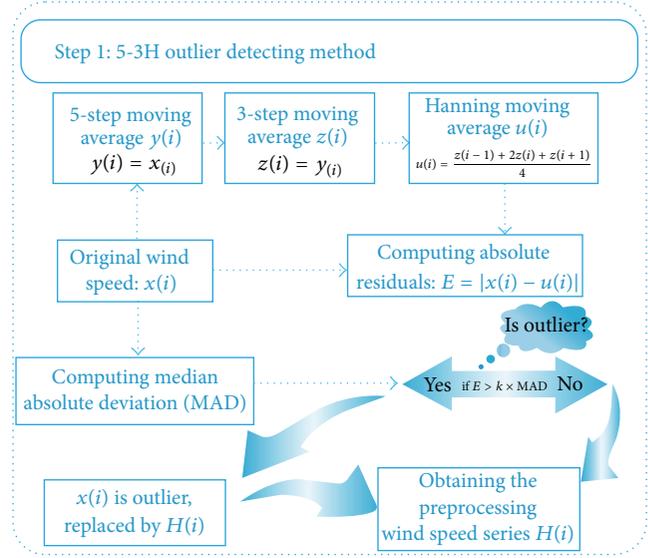


FIGURE 1: The flowchart of 5-4 Hanning filter.

Step 2. Three-point moving average smoothing. For the smoothed signal in the first step, we use three-point moving average smoothing method to form the second smoothed estimates. The series $y(i-1), y(i), y(i+1)$ sorted from small to large is expressed as $y_{(i-1)}, y_{(i)}, y_{(i+1)}$. So three-point moving median average smoothing signal $\{z(i)\}$ can be presented as follows:

$$z(i) = y_{(i)} \quad (3)$$

for $i = 4, \dots, n-3$, where $z(i)$ is the i th point in the second smoothed time series. Then, the six items missing in the series $z(i)$ can be estimated as follows:

$$z(1) = z(2) = z(3) = z(n-2) = z(n-1) = z(n) = 0. \quad (4)$$

Step 3. Hanning moving average smoothing. As for the second smoothed signal, we use Hanning filter to produce final smoothed signal. For a Hanning smooth,

$$u(i) = \frac{z(i-1) + 2z(i) + z(i+1)}{4} \quad (5)$$

for $i = 2$ to $n-1$, where $u(i)$ is the i th point in the final smoothed signal. Then, the six items missing in the series $u(i)$ can be estimated as

$$u(1) = u(2) = u(3) = u(n-2) = u(n-1) = u(n) = 0. \quad (6)$$

Step 4. Compute median absolute deviation (MAD). MAD reflects the degree of absolute dispersion of every original data. The median M_e of $\{x(i)\}$ can be presented as

$$M_e = \begin{cases} x_{(n/2)}, & \text{if } n \text{ is odd,} \\ \frac{1}{2}x_{(n/2)} + \frac{1}{2}x_{(n/2+1)}, & \text{if } n \text{ is even.} \end{cases} \quad (7)$$

MAD can be expressed as

$$\text{MAD} = \frac{|x(i) - M_e|}{n}. \quad (8)$$

Step 5. Set threshold to remove outliers and smooth data. In this paper, we set threshold value as 0.3. $p(i)$ is series to replace original data needed to be replaced by this following formula:

$$p(i) = \{|x(i) - u(i)| > k \times \text{MAD}\}, \quad (9)$$

where p is a logical variable valued either 0 or 1. So the preliminary 5-3H values can be expressed as

$$H(i) = \begin{cases} u(i), & p(i) = 1 \\ x(i), & p(i) = 0. \end{cases} \quad (10)$$

And by replacing the eight values in the beginning and end of preliminary 5-3H values we could obtain the final 5-3H values:

$$H(i) = x(i) \quad (11)$$

for $i = 1, 2, 3, 4, n-3, n-2, n-1, n$.

2.1.2. Wavelet Transform (WT). The WT method is an effective mathematical method used to analyze signal by decomposition into various frequencies. WTs can be categorized into two kinds: Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT). DWT is for wavelets discretely sampled. As for WTs, a key advantage over Fourier transforms is their temporal resolution, which captures both location information and its frequency. In this work, DWT is used to decompose the original wind speed data.

WT decomposes a signal into many detail components and an approximation component, where approximation component contains low-frequency information, the most essential part to identify its signal, and where the detail components reveal the noise of signal. Figure 2 is a tree of wavelet decomposition displaying the decomposition procedure. Firstly, the original data is decomposed into an approximation component A_1 and a detail component D_1 ; and then the A_1 is continued to be decomposed into another approximation component A_2 and detail component D_2 if it is necessary to analyze the signal with higher level resolution. Continue this process until it reaches a suitable number of levels.

The original wind speed data is decomposed into several components, one approximation component and multiple detail components, to reflect the characteristics of the wind speed data on different levels. The approximation is designed to present the main trend of the original wind speed and the details are designed to present the stochastic volatilities on different levels. A suitable number of levels can be determined by comparing the similarity between the approximation and the original wind speed.

2.2. Artificial Intelligence Algorithm. Zhang et al. [16] considered that statistical models are not perfect in forecasting. As

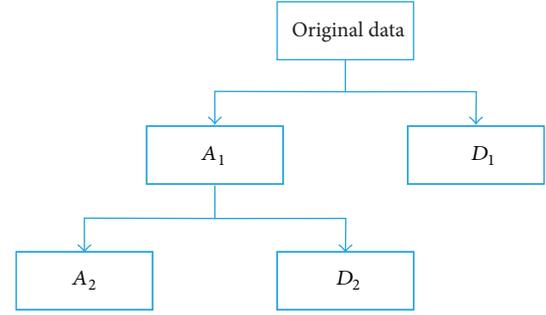


FIGURE 2: Wavelet decomposition.

most of statistical models assume that the data is normally distributed, however, wind speed series is not normally distributed [36]. Additionally, the stochastic and intermittent characteristics of wind speed series require more complex models and functions for capturing the nonlinear trends and relations, whereas these models are built based on a hypothesis: a linear correlation structure exists among time series values [37]. Consequently, the wind speed series is difficult to be forecasted accurately by statistical models. To address these problems of statistical approaches, the AI models, mainly including ANN and SVR, have got more and more concerns for accurate short-term wind speed prediction.

2.2.1. Artificial Neural Network (ANN). ANN consists of interconnected artificial neurons which are programmed to imitate the natural properties of biological neurons. It has been widely used in forecasting time series, especially the data nonnormally distributed, such as wind speed.

2.2.2. Backpropagation Neural Network (BPNN). In this work, a backpropagation (BP) is adopted as one of the comparative approaches for short-term wind speed forecasting. In Figure 3 the BP contains an input layer, at least one hidden layer, and an output layer, which implement the map of an input vector to output scalar via activation function in different neurons. With n inputs and m hidden neurons, the output of the j th hidden node Z_j can be calculated as

$$Z_j = f_h \left(\sum_{i=1}^n w_{ij} y_{k-i} \right), \quad (12)$$

where w_{ij} denotes the connection weight from the i th input node to the j th hidden node, y_{k-i} is i -step behind past wind speed y_k , and $f_h(\cdot)$ denotes a sigmoid activation function in the hidden layer. Then, the wind speed prediction can be estimated by

$$\hat{y}_k = f_o \left(\sum_{j=1}^m w_j Z_j \right), \quad (13)$$

where w_j denotes the connection weight from the j th hidden node to the output node, \hat{y}_k denotes the forecasted wind speed at the k th sampling instant, and $f_o(\cdot)$ is a linear activation function for the output layer. The nonlinear mapping

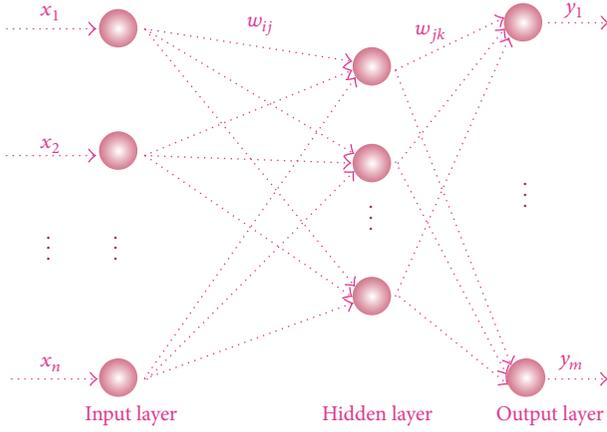


FIGURE 3: BP topology chart.

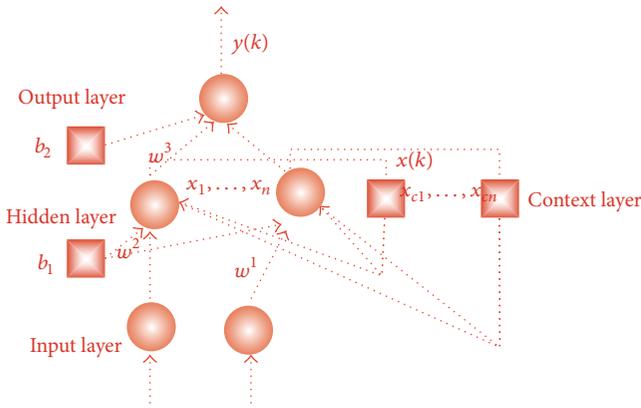


FIGURE 4: ERNN topology chart.

capability of ANN is achieved by minimizing the overall error between the actual wind speed y_k and the predicted wind speed \hat{y}_k through Levenberg-Marquardt (LM) algorithm [38].

2.2.3. Elman Recurrent Neural Network (ERNN). Elman recurrent neural network (ERNN) is a famous recurrent topology, developed by Tong et al. [39]. In a typical ERNN, the hidden layer neurons are fed by the outputs of the context neurons and the input neurons (Figure 4). Context neurons are known as previous states (memory units) of output of hidden neurons [40]. This recurrent topology makes the ERNN more sensitive to the historical data, increasing its capacity of dealing with the dynamic information. In addition, it is not necessary to use state variable as the input or the training data. Its dynamic characteristics are provided by its internal connections, which make network more suitable for time-varying system modeling. This is also an important factor making ERNN superior to the feed-forward neural network, such as multilayer perceptions (MLP) and radial basis function networks (RBF).

Because of ERNN's training algorithm which is mainly based on the gradient descent method, this may cause

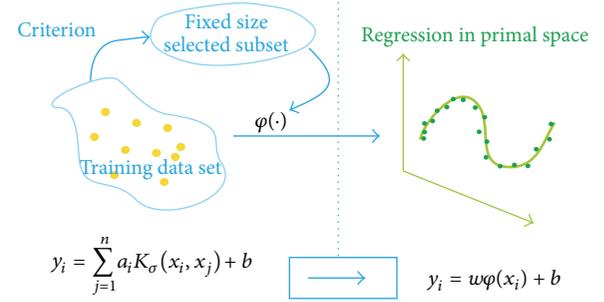
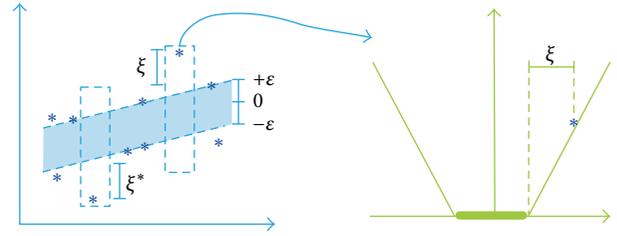


FIGURE 5: Input space and feature space.


 FIGURE 6: ϵ -insensitive loss function.

a number of problems [41]: (1) the speed of network convergence is slow, and the training may give rise to a lower learning efficiency; (2) as the network structure and weights are not trained concurrently, a good performance of dynamic approximation cannot be guaranteed; (3) lack of the global search capacity easily makes it fall into a local best.

2.2.4. Support Vector Regression (SVR). Developed by Vapnik [42], support vector machines (SVMs) are one of the most widely used models based on statistical learning theory. A nonlinear mapping $\varphi(\cdot) : \mathfrak{R} \rightarrow \mathfrak{R}$ is defined to map the training data set (input data) $\{(\mathbf{x}, y)\}$ into a high dimensional feature space (which has infinite dimensions), \mathfrak{R} (Figure 5). Then, in this high dimensional feature space, theoretically there exists a linear function, f , to formulate the nonlinear relationship between input data and output data. Such a linear function, namely, SVR function, is given as follows:

$$f(x) = \mathbf{w}\varphi(\mathbf{x}) + b, \quad (14)$$

where $f(x)$ is the forecasting values; the coefficients \mathbf{w} ($\mathbf{w} \in \mathfrak{R}$) and b ($b \in \mathfrak{R}$) are adjustable. As mentioned above, through SVM method one aims to minimize the empirical risk,

$$R_{\text{emp}}(f) = \frac{1}{N} \sum \Theta_{\epsilon}(y, \mathbf{w}(\varphi(\mathbf{x}) + b)), \quad (15)$$

where $\Theta_{\epsilon}(y, f(\mathbf{x}))$ is the ϵ -insensitive loss function (Figure 6) and is defined as follows:

$$\Theta_{\epsilon}(y, f(\mathbf{x})) = \begin{cases} |f(\mathbf{x}) - y| - \epsilon, & \text{if } |f(\mathbf{x}) - y| \geq \epsilon \\ 0, & \text{otherwise.} \end{cases} \quad (16)$$

In addition, $\Theta_{\epsilon}(y, f(\mathbf{x}))$ is employed to find out an optimal hyper plane on the high dimensional feature space

(Figure 5) to maximize the distance separating the training data. So, the SVR focuses on finding the optimal hyper plane and minimizing the ε -insensitive loss function and the training error between the training data.

Then, the SVR minimizes the overall errors

$$\text{Min } R_\varepsilon(\mathbf{w}, \xi, \xi^*) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_i^n (\xi_i + \xi_i^*) \quad (17)$$

with the constraints

$$\begin{aligned} \mathbf{y}_i - \mathbf{w}^T \varphi(\mathbf{x}_i) - b &\leq \varepsilon + \xi_i^*, \quad i = 1, 2, \dots, N, \\ -\mathbf{y}_i + \mathbf{w}^T \varphi(\mathbf{x}_i) + b &\leq \varepsilon + \xi_i \quad i = 1, 2, \dots, N, \\ \xi_i^* &\geq 0 \quad i = 1, 2, \dots, N, \\ \xi_i &\geq 0 \quad i = 1, 2, \dots, N. \end{aligned} \quad (18)$$

The first term in (17) employed the concept of maximizing the distance between two separated training data sets and is used to regularize weight sizes, to penalize large weights, and to maintain regression functional flatness. And the second term $C \sum_i^n (\xi_i + \xi_i^*)$ penalizes training errors of $f(\mathbf{x})$ and \mathbf{y} by using the ε -insensitive loss function. C is a parameter to trade off two terms. Training errors above ε are denoted as ξ_i^* , and training errors below $-\varepsilon$ are denoted as ξ_i (Figure 6).

After solving the problem of quadratic optimization with inequality constraints, the parameter vector \mathbf{w} of (14) is obtained:

$$\mathbf{w} = \sum_{i=1}^N (\beta_i^* - \beta_i) \varphi(\mathbf{x}_i), \quad (19)$$

where β_i^* , β_i are obtained by solving a quadratic problem and are the Lagrangian multipliers. Finally, the SVR regression function is obtained as (21) in the dual space,

$$f(\mathbf{x}) = \sum_{i=1}^N (\beta_i^* - \beta_i) K(\mathbf{x}_i, \mathbf{x}_j) + b, \quad (20)$$

where $K(\mathbf{x}_i, \mathbf{x}_j)$ is called the kernel function, and the value of the kernel equals the inner product of two vectors, \mathbf{x}_i and \mathbf{x}_j , in the feature space $\varphi(\mathbf{x}_i)$, $\varphi(\mathbf{x}_j)$, respectively; that is, $K(\mathbf{x}_i, \mathbf{x}_j) = \varphi(\mathbf{x}_i) \circ \varphi(\mathbf{x}_j)$. Any function that meets Mercer's condition [42] can be used as the kernel function.

2.3. Artificial Intelligent Optimization Algorithm. The empirical results show that the selection of the two parameters c and g (the parameter of Gaussian kernel function) in SVR influences the forecasting accuracy significantly. In order to further improve forecasting accuracy of wind speed, we have employed different evolutionary algorithms (GA, PSO, and CS) for parameters determination, to identify which algorithm is suited for specified data patterns.

2.3.1. Genetic Algorithm (GA). GA was firstly developed by John Holland et al. in the 1960s. It is an effective algorithm for nonlinear global optimization that was inspired by

the biological evolution process. It is especially suitable for solving complicated optimization problems for simplicity and robustness, and it has been in use extensively in various forecasting and optimization fields. The GA approach is listed as follows.

- (i) Select a group of random candidate solutions.
- (ii) Iterate the following steps until reaching stop criterions:
 - (1) computing the fitness values of the candidate solutions in accordance with the adaptive condition,
 - (2) producing the next generation according to the proportionate principle (the one with higher fitness is more inclined to be chosen),
 - (3) performing a crossover and mutation operation to the candidate solutions and generating new ones.
- (iii) Return the solutions.

2.3.2. Particle Swarm Optimization (PSO). The PSO algorithm was first proposed by Kennedy and Eberhart [43], inspired by the social swarming behavior of animals moving in large groups (birds and insects in particular). Like other swarm-based techniques, the algorithm contains a number of individuals refining their knowledge of the given search space. In this search space, the individuals, called as particles, have a position and a velocity. The PSO algorithm works via attracting the particles of the given search space positions of high fitness. A memory function in each particle adjusts its trajectory according to two pieces of information: the best position that it has so far visited and the global best position attained by the whole swarm. The whole swarm can be considered as a society, and the first piece of information can be thought of as a result from the particle's memory about its past states, and the second piece of information is resulting from the collective experience of all individuals of the society. A fitness evaluation function in PSO computes each particle's position and assigns it a fitness value. Each particle can remember the global best, which can be identified when the position of highest fitness value is visited by the swarm. The position of the highest fitness value that has been personally visited is called the local best.

2.3.3. Cuckoo Search (CS). The cuckoo search (CS) algorithm is a new optimization metaheuristic algorithm (Yang and Deb in 2009 [44]), based on a stochastic global search and the obligate brood-parasitic behavior of cuckoos by laying their eggs in the nests of host birds. In this optimization algorithm, each nest represents a potential solution. They choose the recently spawned nests so that they can be sure that eggs could hatch first for the reason that a cuckoo egg usually hatches earlier than its host bird. In addition, by mimicking the host chicks, a cuckoo chick can deceive the host bird to grab more food resources. If the host birds discover that an alien cuckoo egg has been laid in (with the probability P_a), they

either propel the egg or abandon the nest and completely build a new nest in a new location. New eggs (solutions) laid by cuckoo choose the nest by Levy flights around the current best solutions. And with the levy flight behavior, the cuckoo speeds up the local search efficiency.

In sum, two search capabilities have been used in cuckoo search: global search (diversification) and local search (intensification), controlled by a switching/discovery probability (P_a). Yang and Deb simplified cuckoo parasitic breeding process by the following three idealized rules.

- (i) Each cuckoo lays only one egg at a time and randomly searches a nest to lay it.
- (ii) The egg of high quality will be considered to survive to the next generation.
- (iii) The host bird of the nest, where a cuckoo lays its egg, can discover an alien egg with a possibility, $P_a \in [0, 1]$. And the host bird either propels the egg out of the nest or abandons its nest to build a new one in a new location. The number of available nests is fixed during these rules.

To better understand these rules, they can be transformed into the following steps.

Step 1. A cuckoo randomly chooses a nest to hatch only one egg. An egg represents a potential best solution.

Step 2. To maximize the probability of their eggs survival, the cuckoo birds search the most suitable nests by law of Levy flight. According to the elitist selection principle, the best egg (minimum solution) will survive to the next generation and will have the opportunity to grow into a mature cuckoo bird. In this step, the aim of cuckoo algorithm is to obtain the ability of intensification.

Step 3. The number of available nests (population) is fixed during these rules. The alien egg laid by a cuckoo bird is discovered by the host with a probability $P_a \in [0, 1]$, and this egg is thrown away or the host abandons the nest to completely build a new one in a new location (with a new random solution). In this step, the aim of cuckoo algorithm is to obtain the ability of diversification.

For minimization problems the quality or fitness function value may be the reciprocal of the objective function. Each egg in a nest represents a solution and the cuckoo egg represents a new solution. Therefore, there is no difference between an egg, a nest, and a solution.

When generating new solutions for, say, a cuckoo i , a Levy flight is performed as follows:

$$x_i^{t+1} = x_i^t + \alpha \oplus \text{Levy}(\lambda), \quad (21)$$

where $\alpha > 0$ is the step size and should be related to the scales of the problem of interest. In most cases, $\alpha = 1$ is proposed. Equation (21) is essentially the stochastic process for a random walk. In addition, a random walk is a Markov chain process where next status only depends on the current status (the first term in (21)) and the transition probability

(the second term in (21)). The product \oplus means entrywise multiplications, which is similar to those used in PSO, but the random walk process via Levy flight here is more efficient in exploring the search space, for its step length is much longer in the long run.

The Levy flight provides a random walk process while the random step length is drawn from a Levy distribution:

$$\text{Levy} \sim u = t^{-\lambda} \quad (22)$$

which has an infinite variance with an infinite mean. Here, the steps essentially form a random walk process with a power-law step-length distribution with a heavy tail. Some of the new solutions are generated by Levy flight around the current best solution obtained so far, which will intensify the local search (intensification). However, a substantial fraction of the new solutions should be generated by far-field randomization (diversification), whose locations should be far enough from the current best solution; this will ensure that the system will not be trapped in a local optimum.

The simple flowchart of the cuckoo search algorithm is presented in Figure 7.

3. The Proposed Hybrid Model

As the high volatility, nonstationarity, and nonlinearity of wind speed series, many useful tools are introduced to pre-dispose so as to make an accurate forecasting.

The procedure for applying the proposed method to predict the 10 min wind speed is illustrated in Figure 8 and described as follows.

Step 1. Conduct the 5-3H method to test and discover the outliers and then replace by 5-3H values.

In this step, after a large number of experiments, we set threshold parameter $k = 0.3$. The result shows that not only could 5-3H detect the outliers effectively, but it can also smooth the original data to some extent, in which it captures the majority of the trend of wind speed.

However, some slight white noises still exist in the series after 5-3H. Hence, it is necessary to further smooth via wavelet in Step 2.

Step 2. Decompose 5-3H values by wavelet denoise by db3 wavelet basis function and reconstruct the series.

In this approach, we adopt db3 as the wavelet basis function in only one layer to decompose the data. As the result of respective smooth preprocessing data after 5-3H, and as making many an experiment, we discover that decomposing the data to one layer has the best effectiveness of denoising, which otherwise could denoise excessively to get rid of useful information of original data. In relation to threshold selection, we use the popular method of threshold selecting, Birge-Massart method. After being filtered, the wind speed of high frequency, that is, white noises, could be smoother so as to be better used in forecasting.

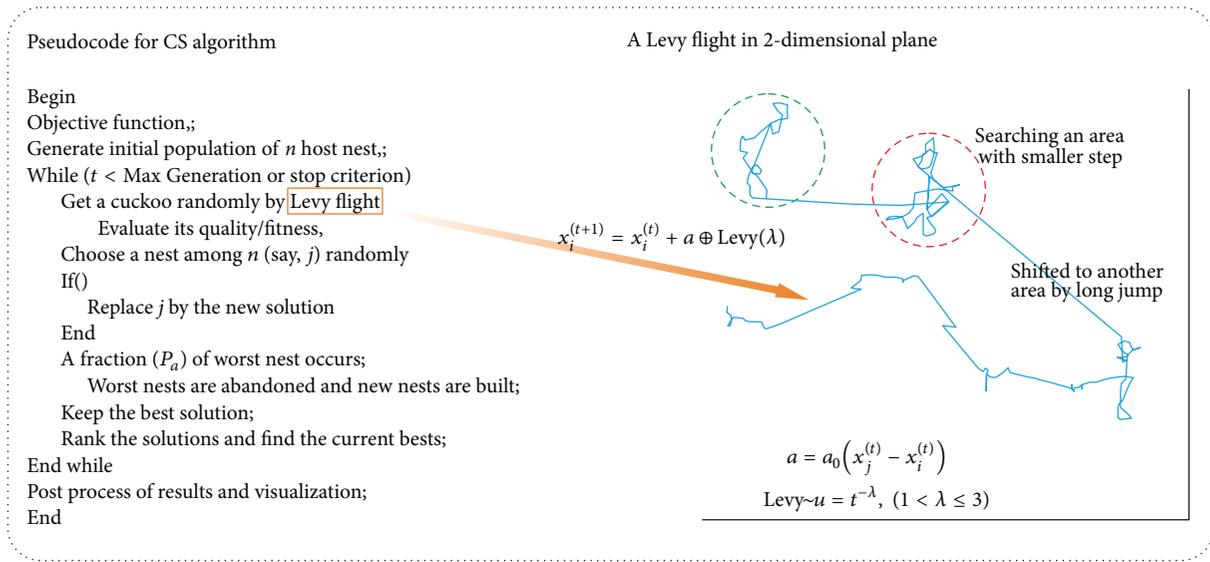


FIGURE 7: The flowchart of cuckoo search.

Step 3. Use three popular artificial intelligent algorithms, BP, Elman, and SVR, to fit the models and predict the future values of one day.

We discover that the SVR functions are the best among these two models. To further improve the performance of SVR, we propose another two steps at the same time, which are, respectively, another three artificial intelligent optimization algorithms in Step 4.

Step 4. Conduct the GA, PSO, and CS to optimize the two main parameters of SVR and make a comparison with the conventional approach of grid search.

A nonheuristic algorithm of searching parameters of SVR is grid search in this paper to search the best parameters c and g . Although, in the sense of grid search, it could find the best accuracy (the global optimum), employing the metaheuristic algorithm can find the global optimum more efficiently if considering the search in a larger field of c and g by grid search is time-wasting. Therefore, under this consideration and in order to further improve accuracy of forecasting, GA, PSO, and CS are employed to search the two main parameters of SVR.

4. Analysis and Discussion of the Applicative Case Studies

4.1. Data Presentation. To validate the proposed forecasting method, three cases are introduced. The first two are 10 min average wind speed series from wind towers of 70 meters in two sites in four seasonal months (January, April, July, and October in 2011, which are the representative months for each quarter of the year). The first site locates in the Jiling Shoal, Jinchang City, with longitude of 101.7999, latitude of 38.5248, and altitude of 2195.000. The second wind tower is in Qingtu

Lake, Wuwei City, with longitude of 103.6201, latitude of 39.1031, and altitude of 1298.000. Of each wind tower in each month, we draw 744 samples and make a 3-step forecasting. The previous 600 samples are used to build a model and then predict the remaining 144 ($48 * 3$) points ($144 * 10$ minutes, which amounts to a whole day). To further validate the universe of approach, the data of electricity price from New South Wales (NSW) in January 2012 is also used as the third case. In Figures 9–11, the raw data in three cases are illustrated.

4.2. Error Evaluation. Table 1 shows the results of proposed intelligent algorithms to forecasting the wind speed by 3 steps in Jiling Shoal and Qingtu Lake and electricity price by 3 steps in NSW in 2011. We refer to PBP, PELman, and PSVR as prediction method after preprocessing. To validate the proposed approach, we mainly contrast the results of PBP and BP, PELman and Elman, and SVR and PSVR. The mean absolute error (MAE), mean square error (MSE), mean absolute percentage error (MAPE), and symmetric mean absolute percentage error (SMAPE) are utilized to scale the prediction accuracy of these three models [45]. The MAE values can be calculated by

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |x_i(t) - \hat{x}_i(t)| \quad (23)$$

and the values of MSE can be computed by

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (x_i(t) - \hat{x}_i(t))^2 \quad (24)$$

and the MAPE values can be calculated by

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i(t) - \hat{x}_i(t)}{x_{i(t)}} \right| \times 100\%. \quad (25)$$

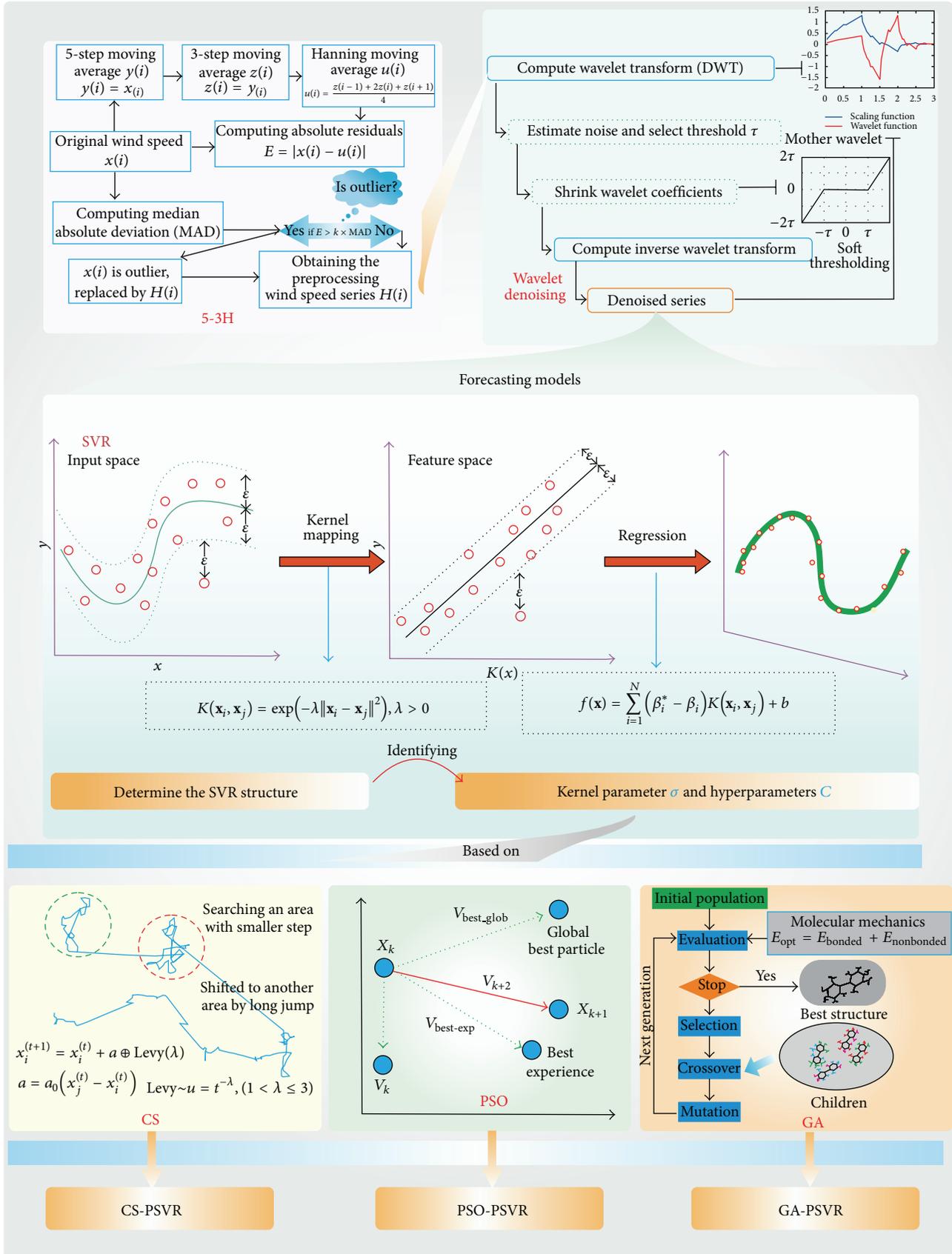


FIGURE 8: The flowchart of the proposed method.

TABLE 1: Forecasted results in three cases.

Site	Month	Forecasted results						
		BP	PBP	Elman	PElman	SVR	PSVR	
Jiling Shoal	January	MAE	0.6719	0.6585	0.7360	0.7316	0.6620	0.6551
		MSE	0.9760	0.9480	1.0770	1.0610	0.9792	0.9613
		MAPE (%)	15.28	13.39	17.40	16.33	13.01	12.88
		SMPAE (%)	12.45	12.07	13.54	13.39	12.35	12.15
	April	MAE	0.6467	0.5242	0.5849	0.5345	0.5209	0.4947
		MSE	0.8420	0.7016	0.7522	0.7187	0.6922	0.6705
		MAPE (%)	9.96	8.30	9.19	8.46	7.93	7.79
		SMPAE (%)	9.96	8.41	9.18	8.61	8.20	7.96
	July	MAE	0.8534	0.8007	0.8133	0.792	0.8108	0.7653
		MSE	1.1730	1.1700	1.1280	1.147	1.1640	1.1430
		MAPE (%)	13.42	12.29	12.9	12.38	12.53	11.58
		SMPAE (%)	13.24	12.39	12.69	12.37	12.56	11.86
	October	MAE	0.5814	0.5499	0.5812	0.5614	0.5432	0.5182
		MSE	0.7665	0.7583	0.7745	0.7681	0.7940	0.6811
		MAPE (%)	11.73	10.48	11.70	10.67	10.92	10.16
		SMPAE (%)	11.25	10.25	11.34	10.51	10.70	9.97
Qingtu Lake	January	MAE	0.6657	0.6275	0.6522	0.6420	0.6664	0.6271
		MSE	0.9204	0.8949	0.9096	0.8919	0.9303	0.8969
		MAPE (%)	17.23	15.66	16.82	16.33	17.28	15.68
		SMPAE (%)	15.90	14.67	15.56	15.15	15.98	14.64
	April	MAE	0.7838	0.7752	0.8357	0.8249	0.7533	0.7455
		MSE	1.0920	1.0850	1.1570	1.1530	1.0620	1.0830
		MAPE (%)	22.00	21.07	23.76	23.33	20.99	20.16
		SMPAE (%)	20.05	19.79	21.95	21.49	19.31	19.11
	July	MAE	0.7822	0.7643	0.7422	0.7271	0.7560	0.7418
		MSE	1.1640	1.2080	1.1050	1.0630	1.1420	1.1140
		MAPE (%)	19.76	18.80	18.62	17.24	18.83	17.61
		SMPAE (%)	17.02	16.16	16.23	15.66	16.21	15.69
	October	MAE	0.5724	0.5593	0.5469	0.5347	0.5477	0.5336
		MSE	0.7772	0.7789	0.7616	0.7458	0.7574	0.7436
		MAPE (%)	12.81	12.42	12.18	11.97	12.33	12.02
		SMPAE (%)	12.34	12.18	11.76	11.52	11.77	11.47
NSW	January	MAE	1.6519	1.5083	2.5160	1.9229	1.5308	1.3921
		MSE	2.3697	2.0388	3.4833	2.6229	2.1335	2.0616
		MAPE (%)	6.99	6.39	10.57	8.24	6.43	5.63
		SMPAE (%)	6.83	6.48	10.36	7.98	6.40	5.71

And the values of SMAPE can be computed by

$$\text{SMPAE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i(t) - \hat{x}_i(t)}{(x_{i(t)} + \hat{x}_i(t))/2} \right| \times 100\%, \quad (26)$$

where $x_i(t)$ and $\hat{x}_i(t)$ signify the i th actual and predicted values at time t , respectively. In Table 1 listed are comparisons of the MAE, MSE, MAPE, and SMAPE values for the PBP and BP, PEIman and Elman, and SVR and PSVR models.

As shown in Table 1, plenty of MAE, MSE, MAPE, and SMPAE values obtained through the proposed methods are displayed. To further facilitate understanding of the performance of improved approaches, the 4 kinds of decreased

relative error (RE) obtained by forecasting results of these three models are calculated by

$$\text{RE}_1(t, i) = \frac{\text{MAE}_{\text{reference model}}(t) - \text{MAE}_{\text{model } i}(t)}{\text{MAE}_{\text{reference model}}(t)} \times 100\%,$$

$$\text{RE}_2(t, i) = \frac{\text{MSE}_{\text{reference model}}(t) - \text{MSE}_{\text{model } i}(t)}{\text{MSE}_{\text{reference model}}(t)} \times 100\%,$$

$$\text{RE}_3(t, i) = \frac{\text{MAPE}_{\text{reference model}}(t) - \text{MAPE}_{\text{model } i}(t)}{\text{MAPE}_{\text{reference model}}(t)} \times 100\%,$$

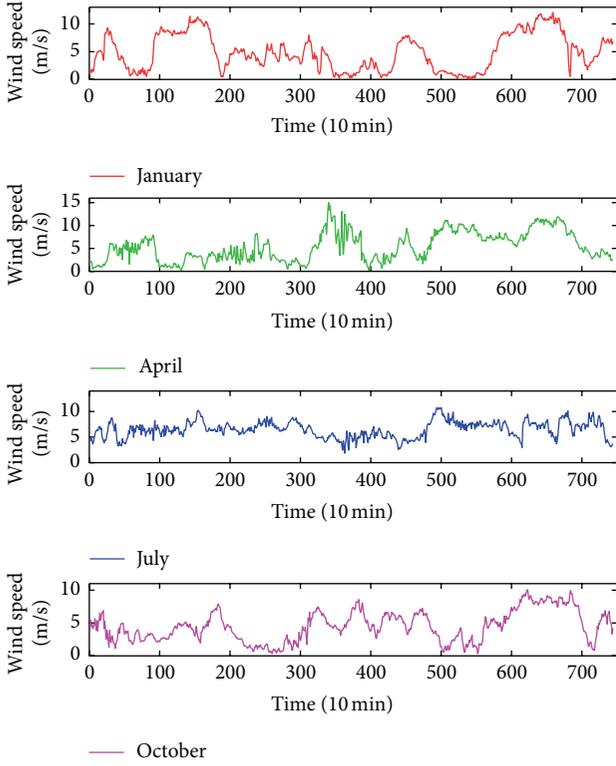


FIGURE 9: The raw data in Jiling Shoal.

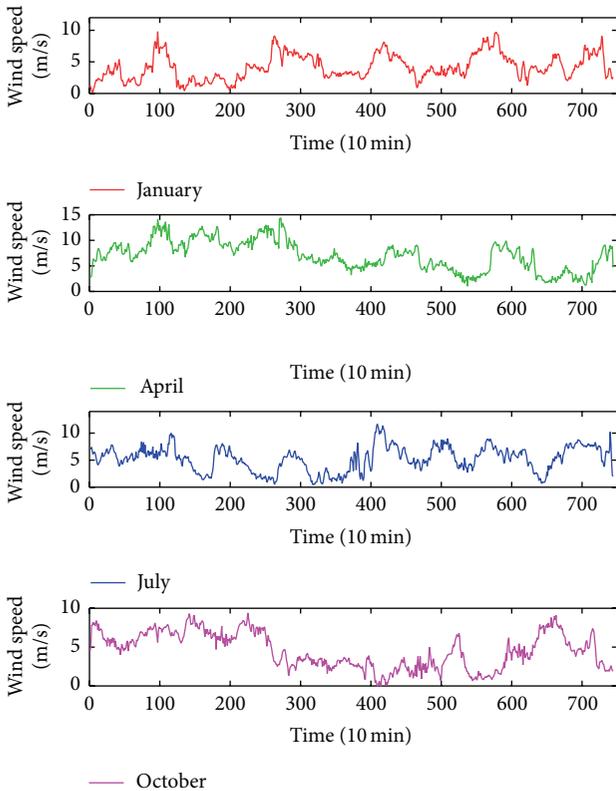


FIGURE 10: The raw data in Qingtu Lake.

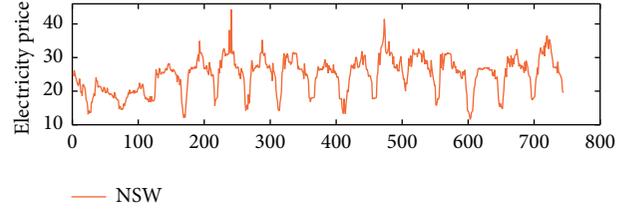


FIGURE 11: The raw data in NSW.

$$RE_4(t, i) = \frac{SMAPE_{reference\ model}(t) - SMAPE_{model\ i}(t)}{SMAPE_{reference\ model}(t)} \times 100\%, \quad (27)$$

where the reference model in our case is the model without preprocessing, model i ($i = 1, 2, 3$) represents one of three models, and t ($t = 1, 2, 3, 4$) stands for one of four seasonal months. The results of RE values are listed in Table 2 and illustrated in Figure 14. Additionally, to provide a comprehensive evaluation of performances of proposed methods, the average (Ave.) error criterion is introduced, which is computed according to

$$Ave.(i) = \sum_{t=1}^4 RE_j(t, i). \quad (28)$$

4.3. Simulating Results Analysis

Steps 1 and 2. Conduct 5-3H and wavelet denoise method to predispose wind speed in Jiling Shoal in four seasonal months in 2011 by use of proposed method.

As we can see from Figures 12-13, the first subplot to fourth subplot are pictures of original and preprocessing wind speed series in January, April, July, and October, respectively. Figure 14 illustrates the raw electricity price series in NSW. The preprocessing data are apparently smoother as the result of 5-3H outliers filtering and wavelet denoising. According to the algorithm of 5-3H, not only could it detect the outliers, but, more importantly, it has also the characteristics of smoothing data. Through these proposed predisposing approaches, the tendency of wind speed series and electricity price become clear and are more adaptive to be forecasted; that will be illustrated in next step.

Step 3. Use three popular artificial intelligent algorithms, BP, Elman, and SVR, to fit the models and predict future values of one day.

As is listed in Table 1, PBP, PELman, and PSVR represent models after data preprocessing. We can easily discover that almost all the accuracies of PBP, PELman, and PSVR outperform those of BP, Elman, and SVR. The more explicit results, RE, are shown in Table 2 and Figure 15. Table 2 reveals the improved percentage of accuracy by 3 models evaluated by four error measures (MAE, MSE, MAPE, and SPMAE). Figures 16-17 are forecasting results of wind speed in a certain month and electricity price in NSW. And Figure 15,

TABLE 2: The decreased RE values of each site.

Site	Month	PBP				PElman				PSVR			
		MAE	MSE	MAPE	SMPAE	MAE	MSE	MAPE	SMPAE	MAE	MSE	MAPE	SMPAE
Jiling Shoal	January	1.99	2.87	12.37	3.05	0.60	1.49	6.15	1.11	1.04	1.83	1.00	1.62
	April	18.94	16.67	16.67	15.56	8.62	4.45	7.94	6.21	5.03	3.13	1.77	2.93
	July	6.18	0.26	8.42	6.42	2.62	-1.68	4.03	2.52	5.61	1.80	7.58	5.57
	October	5.42	1.07	10.66	8.89	3.41	0.83	8.80	7.32	4.60	14.22	6.96	6.82
	Ave.	8.13	5.22	12.03	8.48	3.81	1.27	6.73	4.29	4.07	5.25	4.33	4.24
Qingtu Lake	January	5.74	2.77	9.11	7.74	1.56	1.95	2.91	2.63	5.90	3.59	9.26	8.39
	April	1.10	0.64	4.23	1.30	1.29	0.35	1.81	2.10	1.04	-1.98	3.95	1.04
	July	2.29	-3.78	4.86	5.05	2.03	3.80	7.41	3.51	1.88	2.45	6.48	3.21
	October	2.29	-0.22	3.04	1.30	2.23	2.07	1.72	2.04	2.57	1.82	2.51	2.55
	Ave.	2.85	-0.15	5.31	3.85	1.78	2.04	3.46	2.57	2.85	1.47	5.55	3.79
NSW	January	8.69	13.96	8.58	5.12	23.57	24.70	22.04	22.97	9.06	3.37	12.44	10.78

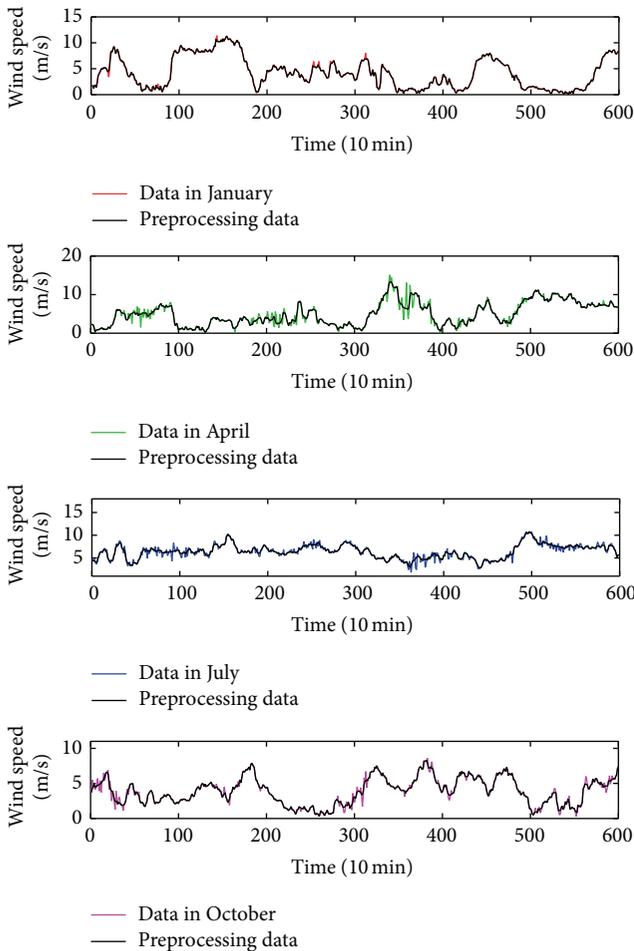


FIGURE 12: The data preprocessing in Jiling Shoal.

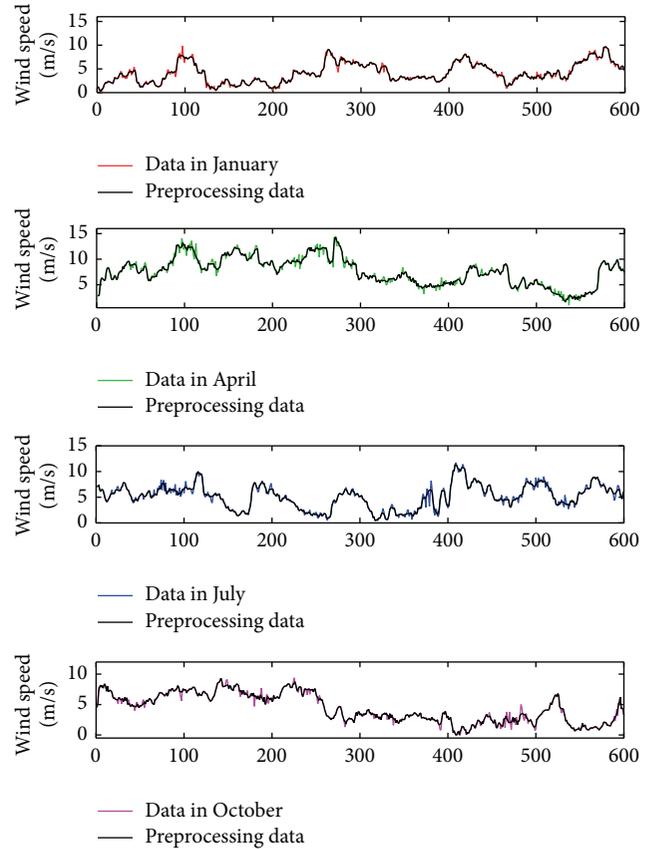


FIGURE 13: The data preprocessing in Qingtu Lake.

made through the result of Ave. in Table 2, shows the average improved percentage of RE values. As is shown in the PBP column from Table 2, almost all the RE values of MAE are positive, which implies that all of the MAE values attained by PBP are smaller than those obtained by BP. In addition,

almost all the RE values in these 3 models are positive, implying smaller errors than original approaches. Figure 15 shows a comprehensive result, average percentage of RE values, in three cases. In Jiling Shoal, PBP has the greater improvement than the other two evidently, while PBP and PSVR perform better than PELman in Qingtu Lake, and in NSW the PELman outperforms the other two intensively. As is shown in Figure 14, the data preprocessing in NSW is more effective because it removes more outliers of electricity price

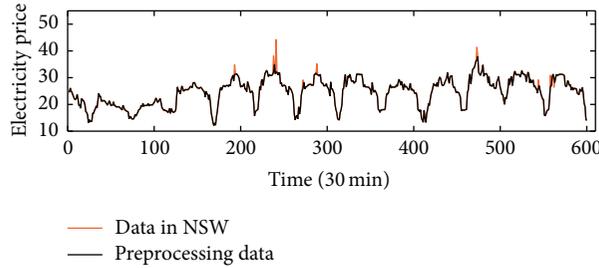


FIGURE 14: The data preprocessing in NSW.

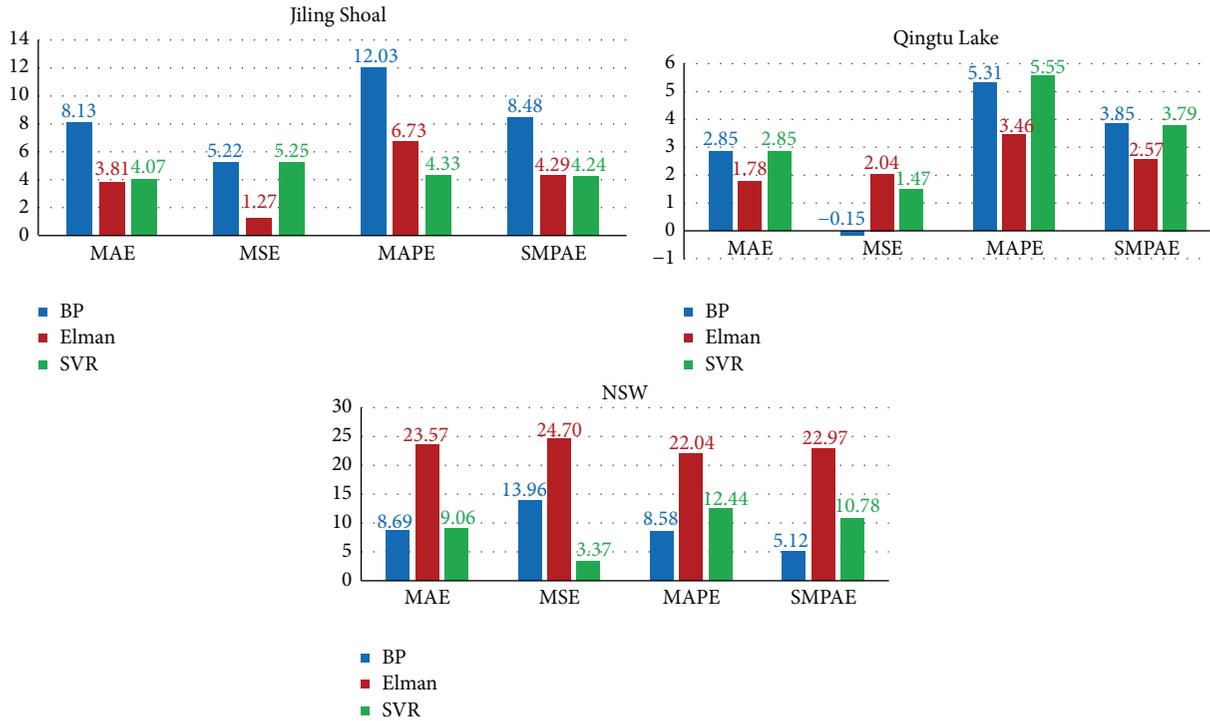


FIGURE 15: Average percentage of RE values of BP, Elman, and SVR in three cases.

than that of wind speed. In conclusion, both Table 2 and Figure 15 demonstrate excellent performance of the proposed preprocessing methods. As the fact that the whole installed electricity capacity in China in 2011 is 62364.2 MW, this slightly improved accuracy could even economize a large amount of money.

In particular, from Table 1, the finding shows that the PSVR functions best among these other two models. To further improve the accuracy of forecasting of wind speed, we propose another three artificial intelligent optimization algorithms to improve performance of PSVR in Step 4.

Step 4. Conduct the CS to optimize the two main parameters of PSVR and make a comparison with the conventional approach, GS, GA, and PSO.

Using the metaheuristic algorithms, GA and PSO, to optimize the hyperparameters of SVR could generally attain a better accuracy than using a nonheuristic conventional

method, such as grid search (GS). However, as Moghram and Rahman [46] said, no certain model or algorithm that forecasts effectively in a wind farm could be applied to any wind farms as a result of difference of wind speed between wind farms and various location-specific factors influencing the wind speed patterns. To explore the potential best algorithm forecasting wind speed in Jiling Shoal and Qingtu Lake, it is necessary to make a comparison between different algorithms. In next part, we choose the most commonly used algorithms (including artificial intelligent and non-artificial-intelligent algorithms, GS, GA, and PSO) to optimize the hyperparameters of SVR and then make a contrast with a new metaheuristic algorithm (CS) optimizing SVR. The final results are shown in Tables 3-4 and Figures 18-21.

Figures 20-21 are forecasting results of wind speed in a certain month and electricity price. Table 3 displays four kinds of forecasting error indexes of PSVR through optimization of GS, GA, PSO, and CS in three cases. The final four models are marked as PSVR, GA-PSVR, PSO-PSVR,

TABLE 3: Optimized results of PSVR in three cases.

Site	Month		Forecasted results			
			PSVR	GA-PSVR	PSO-PSVR	CS-PSVR
Jiling Shoal	January	MAE	0.6551	0.6716	0.6504	0.6597
		MSE	0.9613	0.9578	0.9618	0.9597
		MAPE (%)	12.88	14.79	13.23	13.21
		SMPAE (%)	12.15	12.49	12.16	12.18
	April	MAE	0.4947	0.4963	0.4961	0.4967
		MSE	0.6705	0.6740	0.6742	0.6737
		MAPE (%)	7.79	7.85	7.86	7.84
		SMPAE (%)	7.96	8.02	8.03	8.02
	July	MAE	0.7653	0.7648	0.7653	0.7619
		MSE	1.1431	1.1432	1.1424	1.1430
		MAPE (%)	11.58	11.58	11.57	11.53
		SMPAE (%)	11.86	11.86	11.85	11.82
	October	MAE	0.5128	0.5095	0.5104	0.5062
		MSE	0.6811	0.6791	0.6806	0.6743
		MAPE (%)	10.16	10.12	10.15	10.06
SMPAE (%)		9.97	9.94	9.99	9.86	
Qingtu Lake	January	MAE	0.6271	0.6271	0.6315	0.6259
		MSE	0.8969	0.8962	0.9013	0.8967
		MAPE (%)	15.68	15.68	15.79	15.67
		SMPAE (%)	14.64	14.64	14.76	14.59
	April	MAE	0.7455	0.7544	0.7410	0.7447
		MSE	1.0827	1.0893	1.0717	1.0812
		MAPE (%)	20.16	20.18	20.11	20.14
		SMPAE (%)	19.11	19.39	18.82	19.06
	July	MAE	0.7418	0.7431	0.7467	0.7392
		MSE	1.1144	1.1266	1.1291	1.1004
		MAPE (%)	17.60	17.72	17.93	17.50
		SMPAE (%)	15.69	15.71	15.84	15.70
	October	MAE	0.5336	0.5344	0.5299	0.5295
		MSE	0.7436	0.7425	0.7379	0.7360
		MAPE (%)	12.02	12.07	11.99	11.93
SMPAE (%)		11.47	11.49	11.43	11.39	
NSW	January	MAE	1.3921	1.4439	1.4459	1.3001
		MSE	2.0616	2.0178	1.9080	1.8034
		MAPE (%)	5.63	6.05	6.04	5.34
		SMPAE (%)	5.71	6.09	6.06	5.38
Total			7	2	6	21

and CS-PSVR. And PSVR uses the GS method to search parameters in this paper. The boldfaced numbers denote the best accuracy in each line. The bottom line of Table 3, "Total line," sums up the times of best accuracy in each algorithm, which is illustrated in Figure 18. Obviously, CS-PSVR performs best and is 21 times better than the other three algorithms. However, the performance of GA-PSVR and PSO-PSVR falls behind the GS slightly.

Table 4 reveals the explicit RE values of GA-PSVR, PSO-PSVR, and CS-PSVR relative to PSVR and demonstrates specific respective accuracy. Figure 19 is drawn according to the Ave. in Table 4. Although, from the analysis of Table 4 and

Figure 19, we may figure out that CS-PSVR is the best while GA-PSVR and PSO-PSVR are worse than PSVR, from the column CS in Table 4, CS-PSVR only improves Ave. accuracy less than 0.5% in Jiling Shoal and about 0.3% to 0.6% in Qingtu Lake. This proves that limited is the capability of artificial intelligent optimization algorithm forecasting wind speed patterns in Jiling Shoal and Qingtu Lake. However, we should notice that CS-PSVR forecasting electricity price in NSW performs very well due to the result that it augments more than an average of 5% accuracy and even approaches to that of 12.52% in MSE. This is because electricity price has more apparently seasonal trends, while the wind speed series is

TABLE 4: RE values of PSVR optimized by GA, PSO, and CS in three cases.

Site	Month	GA-PSVR				PSO-PSVR				CS-PSVR			
		MAE	MSE	MAPE	SMPAE	MAE	MSE	MAPE	SMPAE	MAE	MSE	MAPE	SMPAE
Jiling Shoal	January	-2.52	0.36	-14.81	-2.87	0.72	-0.06	-2.67	-0.09	-0.71	0.17	-2.56	-0.30
	April	-0.33	-0.53	-0.69	-0.69	-0.28	-0.55	-0.79	-0.77	-0.41	-0.48	-0.61	-0.68
	July	0.06	-0.01	-0.02	-0.02	-0.01	0.06	0.08	0.09	0.44	0.00	0.43	0.29
	October	0.65	0.29	0.41	0.34	0.46	0.07	0.09	-0.17	1.28	0.99	1.02	1.07
	Ave.	-0.54	0.03	-3.78	-0.81	0.22	-0.12	-0.82	-0.24	0.15	0.17	-0.43	0.10
Qingtu Lake	January	0.01	0.08	-0.05	-0.04	-0.69	-0.49	-0.75	-0.84	0.20	0.02	0.05	0.33
	April	-1.20	-0.60	-0.09	-1.46	0.61	1.02	0.23	1.51	0.11	0.14	0.11	0.25
	July	-0.18	-1.09	-0.72	-0.10	-0.66	-1.32	-1.87	-0.95	0.35	1.26	0.53	-0.07
	October	-0.16	0.14	-0.35	-0.14	0.69	0.76	0.30	0.35	0.76	1.01	0.81	0.74
	Ave.	-0.38	-0.37	-0.30	-0.43	-0.01	-0.01	-0.52	0.02	0.36	0.61	0.38	0.31
NSW	January	-3.73	2.12	-7.35	-6.51	-3.87	7.45	-7.14	-6.02	6.61	12.52	5.23	5.80

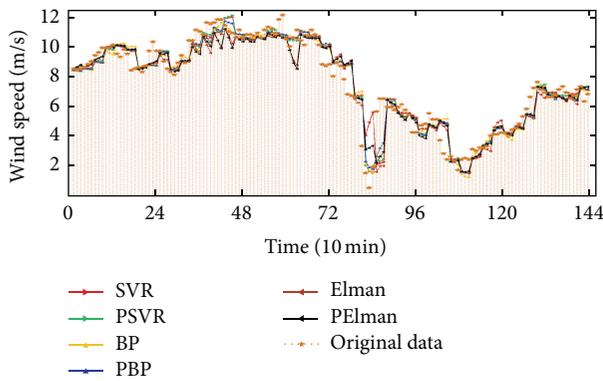


FIGURE 16: The forecasting results of Jiling Shoal after preprocessing in October.

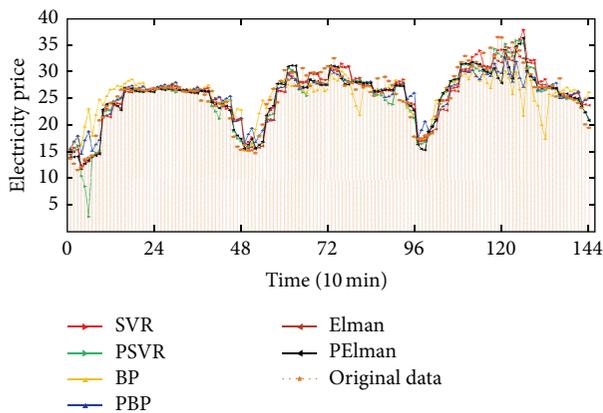


FIGURE 17: The forecasting results of NSW after preprocessing.

fluctuated more intensively randomly. In conclusion, through comparing average RE values of GA-PSVR, PSO-PSVR, and CS-PSVR in NSW, the proposed method CS-PSVR is more excellent than the other three algorithms.

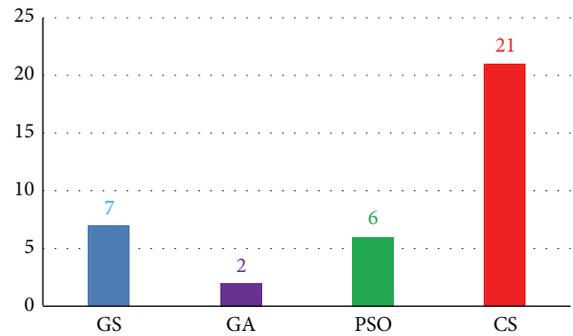


FIGURE 18: Times of best accuracy in three cases.

5. Conclusion

Wind power has been rapidly growing in the world. The forecasting of wind speed plays an important role in the wind energy. Accurate wind speed prediction is becoming increasingly important to improve and optimize renewable wind power generation. Particularly, reliable short-term wind speed prediction can enable model predictive control of wind turbines and real-time optimization of wind farm operation. In this paper we utilize 5-3H and wavelet denoising method to preprocess the original data and then conduct BP, Elman, and SVR models to make a 3-step prediction every 10 minutes. Finally, we adopt GA, PSO, and CS to optimize the PSVR. It is discovered that 5-3H combined with wavelet denoising can significantly improve accuracy of BP network, Elman network, and SVR forecasting wind speed in two sites and electricity price in NSW. These results reveal that excellent ability of removing outliers and denoising of 5-3H combined with wavelet denoising can be applied into the wind speed forecasting in the Jiling Shoal and Qingtu Lake and the electricity prediction in NSW. Relating to the optimization of the two main hyperparameters of SVR, the capacity of a new metaheuristic intelligent optimization algorithm, cuckoo search, outperforms that of traditional methods that are GS, GA, and PSO.

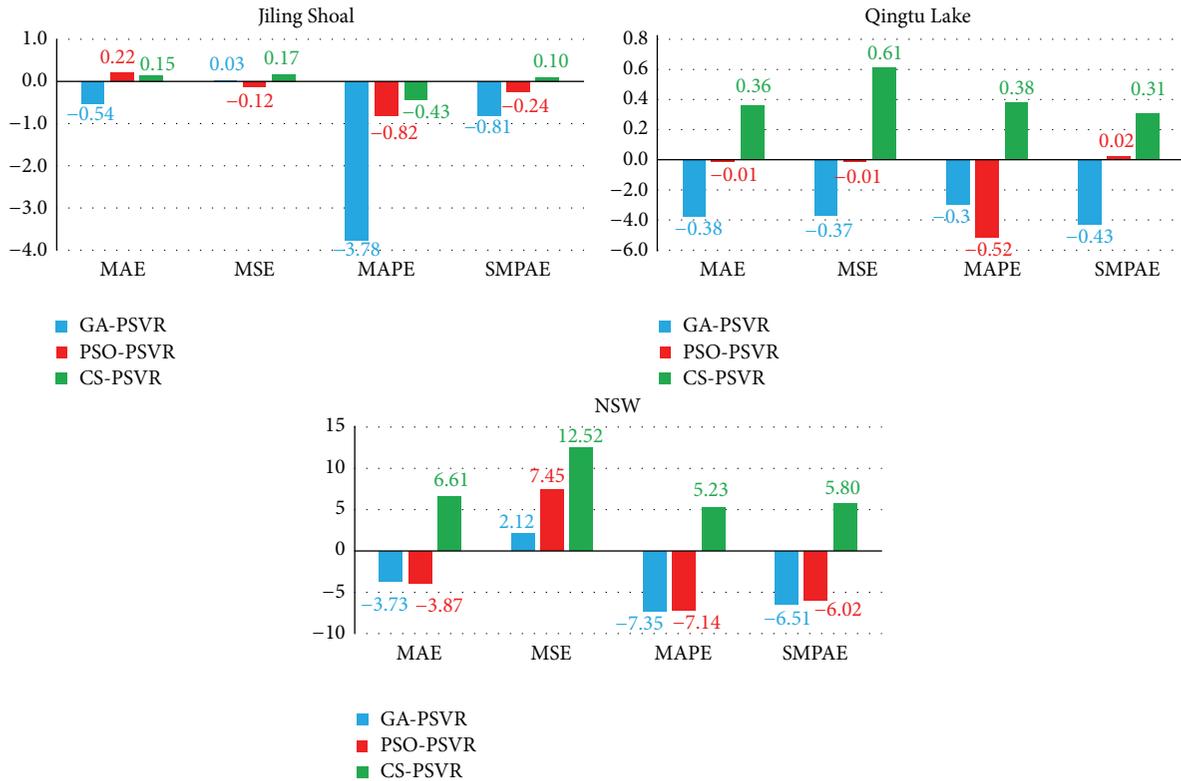


FIGURE 19: Average RE values of GA-PSVR, PSO-PSVR, and CS-PSVR in 3 cases.

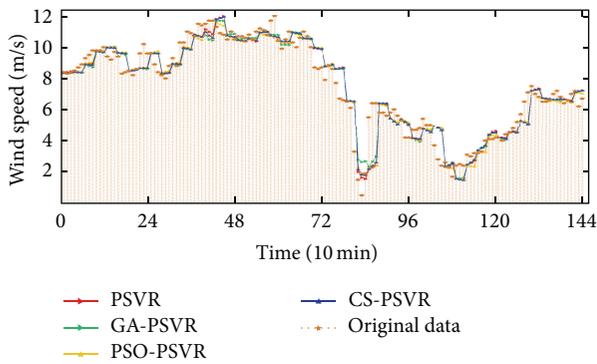


FIGURE 20: The forecasting results of SVR by optimization in Jiling Shoal in October.

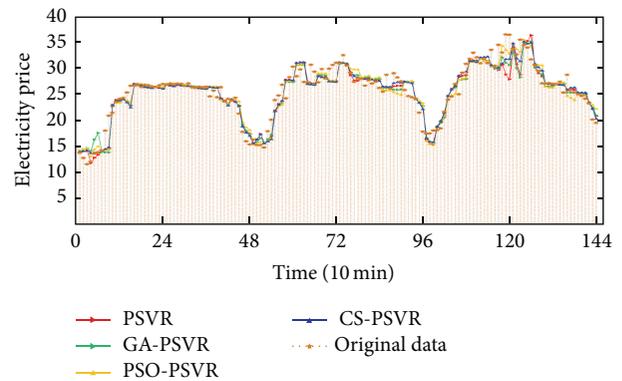


FIGURE 21: The forecasting results of SVR by optimization in NSW.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This paper was supported by the R&D Special Fund for Public Welfare Industry (meteorology) (GYHY201206013), the Key Technology Integration and Application of the Chinese Meteorological Administration Projects (CMAGJ2013M35): the popularization and application of low frequency oscillation

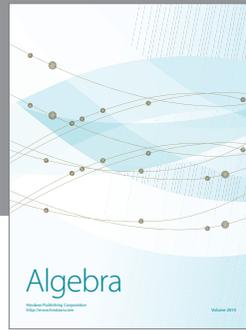
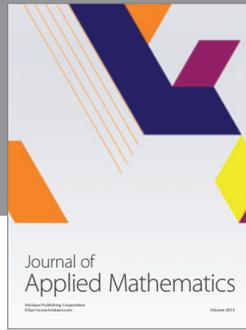
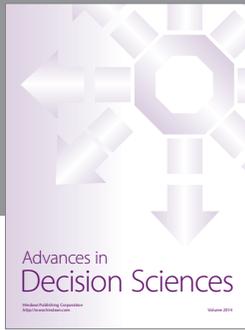
on the monthly scale of north drought forecast in China, and the Gansu Provincial Meteorological Service Center Innovation Fund (2013-10).

References

- [1] J. Zhou, J. Shi, and G. Li, "Fine tuning support vector machines for short-term wind speed forecasting," *Energy Conversion and Management*, vol. 52, no. 4, pp. 1990-1998, 2011.
- [2] X. Zheng and Y. Fu, "Research of wind speed and wind power forecasting," *Advanced Materials Research*, vol. 347-353, pp. 611-614, 2012.

- [3] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, "Current methods and advances in forecasting of wind power generation," *Renewable Energy*, vol. 37, no. 1, pp. 1–8, 2012.
- [4] M. Lei, L. Shiyang, J. Chuanwen, L. Hongling, and Z. Yan, "A review on the forecasting of wind speed and generated power," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 4, pp. 915–920, 2009.
- [5] S. S. Soman, H. Zareipour, O. Malik, and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons," in *Proceedings of the North American Power Symposium (NAPS '10)*, pp. 1–8, September 2010.
- [6] Z. Guo, W. Zhao, H. Lu, and J. Wang, "Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model," *Renewable Energy*, vol. 37, no. 1, pp. 241–249, 2012.
- [7] H. Liu, C. Chen, H.-Q. Tian, and Y.-F. Li, "A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks," *Renewable Energy*, vol. 48, pp. 545–556, 2012.
- [8] X. An, D. Jiang, M. Zhao, and C. Liu, "Short-term prediction of wind power using EMD and chaotic theory," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 2, pp. 1036–1042, 2012.
- [9] J. P. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "Short-term wind power forecasting in Portugal by neural networks and wavelet transform," *Renewable Energy*, vol. 36, no. 4, pp. 1245–1251, 2011.
- [10] J. P. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "Hybrid intelligent approach for short-term wind power forecasting in Portugal," *IET Renewable Power Generation*, vol. 5, no. 3, pp. 251–257, 2011.
- [11] J. Wang, J. Wang, Y. Li, S. Zhu, and J. Zhao, "Techniques of applying wavelet de-noising into a combined model for short-term load forecasting," *International Journal of Electrical Power & Energy Systems*, vol. 62, pp. 816–824, 2014.
- [12] M. M. Alam and H. Sakamoto, "Investigation of Strouhal frequencies of two staggered bluff bodies and detection of multistable flow by wavelets," *Journal of Fluids and Structures*, vol. 20, no. 3, pp. 425–449, 2005.
- [13] M. M. Alam, M. Moriya, and H. Sakamoto, "Aerodynamic characteristics of two side-by-side circular cylinders and application of wavelet analysis on the switching phenomenon," *Journal of Fluids and Structures*, vol. 18, no. 3–4, pp. 325–346, 2003.
- [14] D. Liu, D. Niu, H. Wang, and L. Fan, "Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm," *Renewable Energy*, vol. 62, pp. 592–597, 2014.
- [15] Z. Guo, W. Zhao, H. Lu, and J. Wang, "Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model," *Renewable Energy*, vol. 37, no. 1, pp. 241–249, 2012.
- [16] W. Zhang, J. Wang, J. Wang, Z. Zhao, and M. Tian, "Short-term wind speed forecasting based on a hybrid model," *Applied Soft Computing Journal*, vol. 13, no. 7, pp. 3225–3233, 2013.
- [17] M. Milligan, M. Schwartz, and Y. Wan, "Statistical wind power forecasting models: results for U.S. wind farms," in *Proceedings of the American Wind Energy Association (AWEA) Windpower Conference (WINDPOWER '03)*, Austin, Tex, USA, May 2003.
- [18] M. Milligan, M. N. Schwartz, and Y. Wan, "Statistical wind power forecasting for U.S. wind farms," in *Proceedings of the 17th Conference on Probability and Statistics in the Atmospheric Sciences*, American Meteorological Society Annual Meeting, 2004.
- [19] P. Lv and L. Yue, "Short-term wind speed forecasting based on non-stationary time series analysis and ARCH model," in *Proceedings of the 2nd International Conference on Multimedia Technology (ICMT '11)*, pp. 2549–2553, July 2011.
- [20] B. T. Ewing, J. B. Kruse, and M. A. Thompson, "Analysis of time-varying turbulence in geographically-dispersed wind energy markets," *Energy Sources, Part B: Economics, Planning and Policy*, vol. 3, no. 4, pp. 340–347, 2008.
- [21] M.-D. Wang, Q.-R. Qiu, and B.-W. Cui, "Short-term wind speed forecasting combined time series method and arch model," in *Proceedings of the International Conference on Machine Learning and Cybernetics (ICMLC '12)*, pp. 924–927, IEEE, Xian, China, July 2012.
- [22] E. Hadavandi, H. Shavandi, and A. Ghanbari, "Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting," *Knowledge-Based Systems*, vol. 23, no. 8, pp. 800–808, 2010.
- [23] A. Khosravi, S. Nahavandi, and D. Creighton, "Quantifying uncertainties of neural network-based electricity price forecasts," *Applied Energy*, vol. 112, pp. 120–129, 2013.
- [24] J. Wang, S. Zhu, W. Zhang, and H. Lu, "Combined modeling for electric load forecasting with adaptive particle swarm optimization," *Energy*, vol. 35, no. 4, pp. 1671–1678, 2010.
- [25] Z. Xiao, S.-J. Ye, B. Zhong, and C.-X. Sun, "BP neural network with rough set for short term load forecasting," *Expert Systems with Applications*, vol. 36, no. 1, pp. 273–279, 2009.
- [26] A. Khotanzad, H. Elragal, and T.-L. Lu, "Combination of artificial neural-network forecasters for prediction of natural gas consumption," *IEEE Transactions on Neural Networks*, vol. 11, no. 2, pp. 464–473, 2000.
- [27] T. G. Barbounis and J. B. Theocharis, "A locally recurrent fuzzy neural network with application to the wind speed prediction using spatial correlation," *Neurocomputing*, vol. 70, no. 7–9, pp. 1525–1542, 2007.
- [28] 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," *Knowledge-Based Systems*, vol. 56, pp. 226–239, 2014.
- [29] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.
- [30] L. Zhang, K. Wu, Y. Zhong, and P. Li, "A new sub-pixel mapping algorithm based on a BP neural network with an observation model," *Neurocomputing*, vol. 71, no. 10–12, pp. 2046–2054, 2008.
- [31] E. G. Ortiz-García, S. Salcedo-Sanz, Á. M. Pérez-Bellido, J. Gascón-Moreno, J. A. Portilla-Figueras, and L. Prieto, "Short-term wind speed prediction in wind farms based on banks of support vector machines," *Wind Energy*, vol. 14, no. 2, pp. 193–207, 2011.
- [32] J. Wang, W. Zhu, W. Zhang, and D. Sun, "A trend fixed on firstly and seasonal adjustment model combined with the ϵ -SVR for short-term forecasting of electricity demand," *Energy Policy*, vol. 37, no. 11, pp. 4901–4909, 2009.
- [33] H. Liu, H.-Q. Tian, and Y.-F. Li, "Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction," *Applied Energy*, vol. 98, pp. 415–424, 2012.
- [34] Z. Su, J. Wang, H. Lu, and G. Zhao, "A new hybrid model optimized by an intelligent optimization algorithm for wind speed forecasting," *Energy Conversion and Management*, vol. 85, pp. 443–452, 2014.

- [35] E. Cadenas and W. Rivera, "Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model," *Renewable Energy*, vol. 35, no. 12, pp. 2732–2738, 2010.
- [36] M. S. Miranda and R. W. Dunn, "One-hour-ahead wind speed prediction using a Bayesian methodology," in *Proceedings of the IEEE Power Engineering Society General Meeting (PES '06)*, June 2006.
- [37] S. A. P. Kani and M. M. Ardehali, "Very short-term wind speed prediction: a new artificial neural network-Markov chain model," *Energy Conversion and Management*, vol. 52, no. 1, pp. 738–745, 2011.
- [38] G. Li and J. Shi, "On comparing three artificial neural networks for wind speed forecasting," *Applied Energy*, vol. 87, no. 7, pp. 2313–2320, 2010.
- [39] X. Tong, Z. Wang, and H. Yu, "A research using hybrid RBF/Elman neural networks for intrusion detection system secure model," *Computer Physics Communications*, vol. 180, no. 10, pp. 1795–1801, 2009.
- [40] S.-Y. Chen, F.-J. Lin, and K.-K. Shyu, "Direct decentralized neural control for nonlinear MIMO magnetic levitation system," *Neurocomputing*, vol. 72, no. 13–15, pp. 3220–3230, 2009.
- [41] H.-W. Ge, F. Qian, Y.-C. Liang, W.-l. Du, and L. Wang, "Identification and control of nonlinear systems by a dissimilation particle swarm optimization-based Elman neural network," *Nonlinear Analysis: Real World Applications*, vol. 9, no. 4, pp. 1345–1360, 2008.
- [42] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.
- [43] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the IEEE International Conference on Neural Networks*, pp. 1942–1948, December 1995.
- [44] X.-S. Yang and S. Deb, "Cuckoo search via Lévy flights," in *Proceedings of the World Congress on Nature and Biologically Inspired Computing (NABIC '09)*, pp. 210–214, IEEE, December 2009.
- [45] W. Zhao, J. Wang, and H. Lu, "Combining forecasts of electricity consumption in China with time-varying weights updated by a high-order Markov chain model," *Omega (United Kingdom)*, vol. 45, pp. 80–91, 2014.
- [46] I. Moghram and S. Rahman, "Analysis and evaluation of five short-term load forecasting techniques," *IEEE Transactions on Power Systems*, vol. 4, no. 4, pp. 1484–1491, 1989.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

