

Research Article

Time Series Analysis and Forecasting for Wind Speeds Using Support Vector Regression Coupled with Artificial Intelligent Algorithms

Ping Jiang,¹ Shanshan Qin,^{2,3} Jie Wu,³ and Beibei Sun⁴

¹School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China

²MOE Key Laboratory of Western China's Environmental Systems, Research School of Arid Environment & Climate Change, Lanzhou University, Lanzhou 730000, China

³School of Mathematics and Statistics, Lanzhou University, Lanzhou 730000, China

⁴China Water Resources Beifang Investigation Design and Research Co. Ltd., Tianjin 300222, China

Correspondence should be addressed to Shanshan Qin; qinsslzu@gmail.com

Received 8 October 2014; Accepted 16 December 2014

Academic Editor: Erol Egrioglu

Copyright © 2015 Ping Jiang 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.

Wind speed/power has received increasing attention around the earth due to its renewable nature as well as environmental friendliness. With the global installed wind power capacity rapidly increasing, wind industry is growing into a large-scale business. Reliable short-term wind speed forecasts play a practical and crucial role in wind energy conversion systems, such as the dynamic control of wind turbines and power system scheduling. In this paper, an intelligent hybrid model for short-term wind speed prediction is examined; the model is based on cross correlation (CC) analysis and a support vector regression (SVR) model that is coupled with brainstorm optimization (BSO) and cuckoo search (CS) algorithms, which are successfully utilized for parameter determination. The proposed hybrid models were used to forecast short-term wind speeds collected from four wind turbines located on a wind farm in China. The forecasting results demonstrate that the intelligent hybrid models outperform single models for short-term wind speed forecasting, which mainly results from the superiority of BSO and CS for parameter optimization.

1. Introduction

In recent decades, increasing attention has been paid to renewable energy around the world due to the limited reserves of nonrenewable resources and the emerging crisis of global climate warming resulting from large amounts of greenhouse gases emission generated by fossil fuel combustion [1, 2]. The cumulative installed wind capacity in China was reported to be 91412.89 MW in 2013, up 21.4% over the previous year. Adoption will not slow because increased wind power capacity goals have been set as promising solutions to the energy crises in many countries [3]. With the installed capacity of wind energy increasing, resulting in the large-scale integration of wind power into electrical power systems, additional problems and challenges have appeared, including power stability, quality, and, especially, power dispatching [4].

Uncertainties related to wind power can put power quality and system reliability at risk as the penetration of wind power increases, and major grid integration issues such as reserve capacities and balance management also arise [5]. Because wind power is proportional to the cube of wind speeds and a 10% deviation of the expected wind speed results in an approximately 30% deviation in the expected wind power production [6], the prediction error of wind energy largely depends on the accuracy of wind speed forecasts. Accurate wind speed predictions for each of the farm's turbines are critical for the management of wind farms, which is usually the basis of wind power prediction and effective wind power utilization [7] and can increase the reliability of the power grid and reduce operating costs [8].

However, accurate and reliable wind speed forecasts are a significant challenge due to its stochastic nature with high

rates of change, highly nonlinear behavior with no typical patterns [9], and dependency on elevation, terrain, atmospheric pressure, and temperature, which results in large uncertainties of wind speeds. Extensive efforts have been devoted to develop efficient wind speed forecasting models. To date, many forecasting models have been examined and proposed by applying different predictive methods and techniques performed on different forecasting horizons. According to the length of the prediction horizon, wind speed forecasting can be classified into long-term forecasting and short-term forecasting. The former can provide critical information for site location, windmill planning, and proper wind turbine sections for specific wind farms [10]. Precise short-term wind speed forecasts can minimize scheduling errors that can exert a large impact on grid reliability and market-based ancillary costs [11]. According to these approaches, wind speed prediction can be clustered into two main categories, that is, physical methods and statistical methods [12].

Physical methods, which take into account physical factors, that is, temperature, pressure, wind farm layout, and local terrain, are based on numerical weather prediction (NWP) tools that provide weather forecasts by utilizing mathematical models of the atmosphere [2, 13–15]; these models require long operation times and large amounts of computational resources. Landberg initially proposed the concept of applying NWP tools as an input; tools such as the wind atlas analysis and application program (WASP) and PARK are now used for wind prediction correction [13].

Statistical methods that are used to determine the relation between historical wind speeds by generally recursive techniques can be utilized for short-term wind speed forecasting. Many models have been developed to improve wind speed forecasting accuracy, including autoregression (AR), autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), artificial neural networks (ANN), fuzzy logic (FL), support vector machine (SVM), and spatial-temporal models [16]. Torres et al. [17] used ARMA and persistence models to forecast the hourly wind speed up to 10 h ahead. The ARIMA and ANN approaches have been used for wind speed time series forecasting on the south coast of the state of Oaxaca, Mexico [18]. Three types of ANN models, namely, adaptive linear element, back propagation, and radial basis function, were investigated for hourly mean wind speed forecasting at two observational sites in North Dakota [19]. A fuzzy model was proposed for wind speed prediction and provided wind speed forecasts from 30 min to 2 h ahead [20]. Zhou et al. [21] suggested a systematic study on fine-tuning least-squares support vector machines (LS-SVM) model parameters for one-step-ahead wind speed forecasting for the first time. A methodology to characterize the stochastic processes applied for wind speed at different geographical locations via scenarios was provided [22].

Moreover, hybrid models that hybridize multiple features of different predictive models are usually adopted for wind speed forecasting because this type of model can comprehensively capture the intricate characteristics of wind speed series. Combining several forecasting methodologies is another strategy that can significantly improve predictive

performance by taking advantage of each method's performance with respect to data sets, capability of describing nonlinearity and linearity, as well as prediction horizons; these combined models can be superior to individual models [12]. Li et al. proposed a hybrid model consisting of the ANN and Bayesian approaches, and the results indicated that the hybrid approaches produced forecasting errors that were always smaller than those produced by ANN [23]. Monfared et al. [24] developed an ANN and FL hybrid model to predict actual wind speed time series sampled in Rostamabad from 2002 to 2005, which demonstrated that this approach requires less computational time and provides better prediction performance. Salcedo-Sanz et al. [25] combined a hybridized ANN with a mesoscale model, and this combined strategy produced superior forecasting results. Additionally, Cadenas and Rivera investigated hybrid models that consisted of ANN and ARIMA and concluded that the hybrid models outperformed the individual ANN and ARIMA approaches [26].

Based on the aforementioned research, predictive models with different relative weaknesses and strengths have been widely studied and developed. Among these models, the potential for applying the combined approaches over a much wider application area has a special significance because individual models perform well only under specific and corresponding conditions and may therefore require the use of different models [12]. In this paper, an intelligent hybrid forecasting model based on support vector regression (SVR), brainstorm optimization (BSO), and the Cuckoo search (CS) algorithm is proposed. Firstly, to determine the spatial and temporal relations of wind speed series collected from different wind turbines, cross correlation (CC) analysis was performed to discover information about the auto- and cross correlations of the wind speed time series. Then, the proposed hybrid models were applied for wind speeds prediction 1- or 3-step ahead based on averaged hourly and 10 min wind speed series, respectively. The applicative case studies show that the proposed approach has far better performance for short-term wind speeds forecasting, which can assist wind power scheduling and wind farm management.

The remainder of the paper is organized as follows. Section 2 describes the related methodology. Section 3 presents the case study analysis. The conclusions are summarized in Section 4. Finally, acknowledgments and references are presented.

2. Methodology

2.1. Cross Correlation Analysis. In time series analysis, cross correlation refers to the correlations between two time series X_1 and X_2 , while the autocorrelation of a random vector X_1 is the correlation of the series X_1 with itself. Cross correlation is a measure of the similarity of two time series as a function of a time-lag applied to one of them, with the correlation ranging from -1 to $+1$. Moreover, cross correlation is useful to determine the time delay between two time series; the maximum value indicates the point in time where the signals are best aligned [27]. Detailed information about cross correlation can be found in [28].

2.2. Elman Recurrent Neural Network (ERNN). The Elman recurrent neural network (ERNN), first proposed by Elman in 1990 [29], is a partial recurrent network model and lies somewhere between a classic feed-forward perception and a pure recurrent network. Recurrent neural networks have superior temporal and spatial behaviors, such as stable and unstable fixed points and limits cycles, and chaotic behaviors. These behaviors can be utilized to model certain cognitive functions, such as associative memory, unsupervised learning, self-organizing maps and temporal reasoning [30]. Detailed information about ERNN can be found in [31].

The recurrent Elman architecture was chosen for this comparative work due to its nonlinear mapping ability, which can also be used to describe wind speed nonlinear patterns. A simple Elman artificial neural network structure is described in Figure 1. Because the dynamic characteristics of an Elman network are provided by internal connections, it does not need to use the state as an input or training signal, which makes ERNN superior to static feed-forward networks and explains why it is widely used in dynamic system identification.

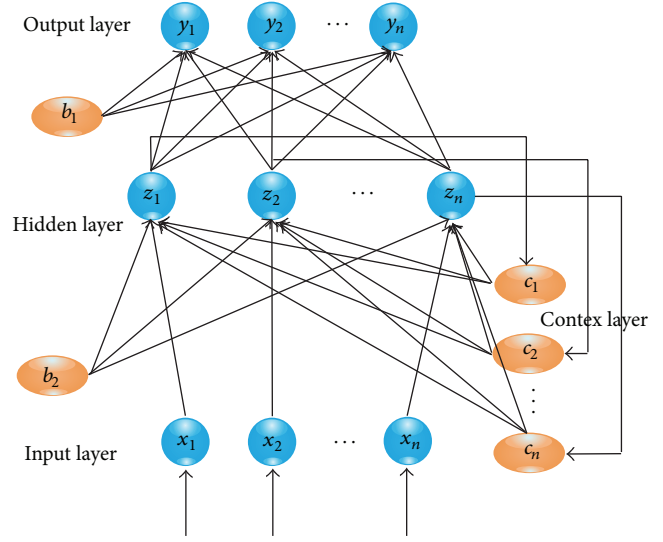


FIGURE 1: Elman recurrent neural network structure.

2.3. Support Vector Regression (SVR). SVR is an adaptation of a recently developed machine learning theory (MLT) known as support vector machines (SVM) proposed by Vapnik et al. [32]. In the SVR model, a regression function $y = f(x)$ is fit, and it is then applied to predict the outputs based on a new input set. A brief review of SVR is introduced as follows [33–35].

Step 1. A nonlinear mapping $\varphi(\cdot) : \mathfrak{R}^n \rightarrow \mathfrak{R}^{n_h}$ is defined to solve a nonlinear regression problem by mapping the training sets $\{(\mathbf{x}_i, y_i)\}_{i=1}^N$ into a high dimensional feature space \mathfrak{R}^{n_h} .

Step 2. In the high-dimensional feature space, the nonlinear regression problem in the lower dimension space is transformed into a linear one by a linear function, namely, the SVR function

$$f(\mathbf{x}) = \mathbf{w}^T \cdot \varphi(\mathbf{x}) + b, \quad (1)$$

where $f(\mathbf{x})$ denotes the forecasting values and the coefficients \mathbf{w} ($\mathbf{w} \in \mathfrak{R}^{n_h}$) and b ($b \in \mathfrak{R}$) are adjustable.

Step 3. Define the empirical risk, $R_{\text{emp}}(f)$,

$$R_{\text{emp}}(f) = \frac{1}{N} \sum_{i=1}^N \Theta_{\varepsilon}(y_i, \mathbf{w}^T \varphi(\mathbf{x}_i) + b), \quad (2)$$

where $\Theta_{\varepsilon}(\mathbf{y}, f(\mathbf{x}))$ is the ε -intensive loss function given by

$$\Theta_{\varepsilon}(\mathbf{y}, f(\mathbf{x})) = \begin{cases} |f(\mathbf{x}) - \mathbf{y}| - \varepsilon, & \text{if } |f(\mathbf{x}) - \mathbf{y}| \geq \varepsilon \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

The ε -intensive loss function is utilized to control the sparsity of the solutions and the generalization of the models.

Step 4. Determine the overall training errors between the training data and the ε -insensitive loss function, which can be

defined as a quadratic optimization problem with inequality constraints

$$\text{Min}_{\mathbf{w}, b, \xi^*, \xi} R_{\varepsilon}(\mathbf{w}, \xi^*, \xi) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^N (\xi_i + \xi_i^*), \quad (4)$$

$$\text{subject to } y_i - \mathbf{w}^T \varphi(\mathbf{x}_i) - b \leq \varepsilon + \xi_i^*, \quad i = 1, 2, \dots, N,$$

$$\mathbf{w}^T \varphi(\mathbf{x}_i) - y_i + b \leq \varepsilon + \xi_i, \quad i = 1, 2, \dots, N,$$

$$\xi_i^*, \xi_i \geq 0, \quad i = 1, 2, \dots, N. \quad (5)$$

The first term in (4) is employed to regularize weight sizes, penalize large weights, and maintain regression function flatness. The second term in (4) penalizes training errors of $f(\mathbf{x})$ and \mathbf{y} by exploiting the ε -intensive loss function. Herein, C is a parameter to balance those two terms. The training errors below $-\varepsilon$ are denoted as ξ_i^* ; otherwise they are denoted as ξ_i^* .

Step 5. Obtain the parameter vector \mathbf{w} by solving the quadratic optimization problem defined in Step 4

$$\mathbf{w} = \sum_{i=1}^N (\beta_i^* - \beta_i) \varphi(\mathbf{x}_i), \quad (6)$$

where β_i^* and β_i are the Lagrangian multipliers.

Step 6. Establish the SVR regression function by the following equation:

$$f(\mathbf{x}) = \sum_{i=1}^N (\beta_i^* - \beta_i) K(\mathbf{x}_i, \mathbf{x}_j) + b, \quad (7)$$

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\lambda \|\mathbf{x}_i - \mathbf{x}_j\|^2), \quad \lambda > 0,$$

where $K(\mathbf{x}_i, \mathbf{x}_j)$ is the kernel function and $K(\mathbf{x}_i, \mathbf{x}_j) = \varphi(\mathbf{x}_i) \cdot \varphi(\mathbf{x}_j)$. In this paper, the radial basis function (RBF) was

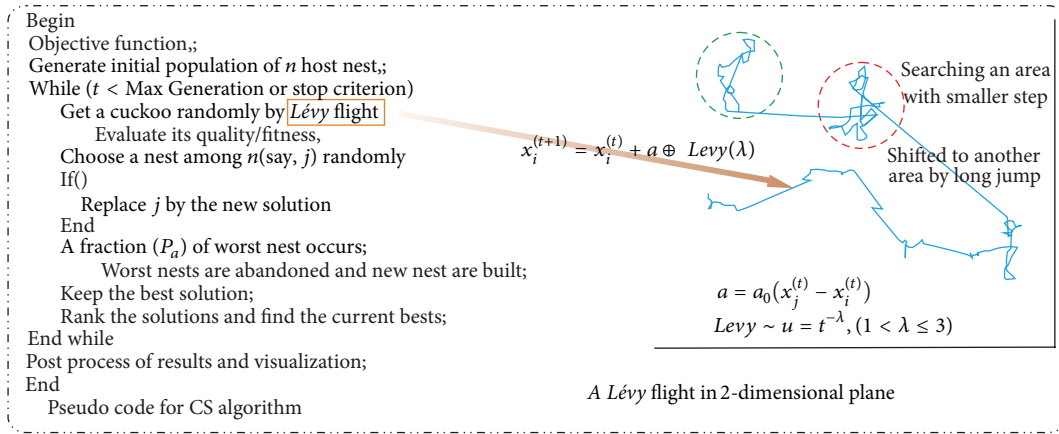


FIGURE 2: The basic procedure of the CS algorithm [38, 41].

selected as the kernel function due to its strong capability for nonlinearly mapping the training sets into an infinite-dimensional space, which is suitable to handle nonlinear relationship problems.

2.4. Cuckoo Search (CS) Algorithm. The CS algorithm, which was inspired by the breeding behavior of cuckoos, is a metaheuristic algorithm recently developed by Yang and Deb [36]. For an optimization problem, the quality or fitness of a solution can simply be proportional to the value of the objective function. In the present study, the CS algorithm was used for parameter optimization. The CS is described briefly as follows.

When generating new solutions $x_i^{(t+1)}$ for, say, a cuckoo i , a Lévy flight is performed

$$x_i^{(t+1)} = x_i^{(t)} + \alpha \oplus \text{Levy}(\lambda), \quad (8)$$

where $\alpha > 0$ is a step size that should be related to the scales of the problem of interest. Based on the constraints imposed by the optimization problem, α must be tuned to the desired step size. Usually, α is utilized, as in the following equation (9):

$$\alpha = \alpha_0 (x_j^{(t)} - x_i^{(t)}), \quad (9)$$

where α_0 is the initial step change. The above equation is essentially the stochastic equation for a random walk. In general, a random walk is a Markov chain whose next status/location only depends on the current location (the first term in the above equation) and the transition probability (the second term). The product \oplus means entrywise multiplication. This entrywise product is similar to that used in PSO, but here the random walk via a Lévy flight is more efficient in exploring the search space, as its step length is ultimately much longer.

The Lévy flight essentially provides a random walk while the random step length is drawn from a Lévy distribution

$$\text{Levy} \sim u = t^{-\lambda}, \quad (1 < \lambda \leq 3), \quad (10)$$

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 should be generated by a Lévy walk around the best solution obtained so far, which will speed up the local search. However, a substantial fraction of the new solutions should be generated by far-field randomization, 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 (Figure 2).

2.5. The Brainstorm Optimization (BSO) Algorithm. We can describe the brainstorm optimization (BSO) algorithm as follows [37].

Step 1. Randomly select a population of n individuals that are feasible for the problem. An individual can be represented by a vector $x_i = (x_{i1}, x_{i2}, \dots, x_{id})$ $i = 1, 2, \dots, n$, where d is the number of dimensions for the individuals. Generally, we initialize the individuals by using $x_{ij} = b_l + (b_r - b_l) \cdot \text{rand}$, where b_l and b_r represent the left boundary and the right boundary of the variable and rand is a random value selected from 0 to 1.

Step 2. Divide the n individuals into m groups by using the k -means cluster method.

Step 3. Calculate the fitness function value for each individual and take the individual that has the best value in the group as the center for the group.

Step 4. Randomly select a number value r_1 between 0 and 1 and compare it with a predetermined probability p_1 . If $r_1 < p_1$, randomly select a group center and then randomly generate an individual to replace it.

Step 5. To update the individuals, first randomly select a number value r_2 from 0 to 1 and then continue according to the following description. If r_2 is less than a predetermined probability p_2 , choose a random value r_3 between 0 and 1.

(a1) If its value is less than a stationary probability p_3 , select the group center and add random values to it for updating the individuals.

- (a2) Otherwise, choose an individual randomly from this group and update it by adding a random value.

Otherwise, generate a random value r_4 between 0 and 1.

- (b1) If its value is less than a stationary probability p_4 , select two centers of two groups randomly and combine them; then add random values to update the individual.
- (b2) Otherwise, choose two individuals from two selected groups randomly and combine them, then add random values to update the individual.
- (b3) Compare the fitness function values obtained by the new updated individuals with the corresponding original individuals, then save the better one and consider it as the new individual.

Step 6. Repeat Step 5 until the all of the individuals have been updated.

Step 7. When the termination criterion has been reached, end the process; otherwise, go to Step 2 and repeat.

It is clear that the individual update in Step 5 is vital to the BSO process. The replacement of the individual in this step is carried out by

$$P_{\text{updated}} = P_{\text{selected}} + \phi * n(\mu, \sigma), \quad (11)$$

where P_{selected} and P_{updated} represent the selected and updated individual, respectively, $n(\mu, \sigma^2)$ is the normal random function of which μ is the mean and σ^2 is the variance, and the coefficient ϕ is determined by

$$\phi = \frac{1}{1 + \exp [(0.5 * \text{max_iteration} - \text{current_iteration}) / k]} * \text{rand}, \quad (12)$$

where max_iteration and current_iteration represent the maximum iteration number and the current iteration number, respectively, rand is a value randomly chosen between 0 and 1, and k is a predetermined value.

2.6. The Proposed BSO- and CS-Based Hybrid Model. In this study, an intelligent hybrid model is proposed for short-term wind speed forecasting by integrating intelligent optimization algorithms with a support vector regression model with inputs determined by spatial-temporal correlation analysis. It is well known that the forecasting accuracy of a SVR model largely relies on the reasonable assignment of the kernel parameter σ and hyper parameter C . Therefore, the determination of these parameters is a significant issue [38]. However, there is neither a structural approach nor any shortage of opinions on how to determine efficient parameters. A grid search, which is a noninteractive traversal algorithm, is the basic method for determining both the kernel parameter σ and hyper parameter C . The parameters largely depend on the size of a grid. On the one hand, a confined grid can be well

suiting for determining accurate parameters for SVR, but the large workload and computational time results in poor model efficiency and practicability. On the other hand, use of a wider grid rarely chooses precise parameters, which can then lead to poor model performance. Lately, the problem of determining a suitable kernel parameter and hyper parameter is often solved using artificial intelligent optimization algorithms due to their superiority in addressing optimization problems when the objective function is nondifferentiable or has a great number of local minima. Artificial intelligence optimization algorithms have been proposed and rapidly developed. Thus, to improve the forecasting efficiency and performance of the SVR model, two recently developed artificial intelligence algorithms (BSO and CS) were adopted to identify the SVR parameters.

To determine accurate short-term wind speed forecasts using an intelligent hybrid model ensemble, two different SVR models were considered: BSO-SVR and CS-SVR. BSO-SVR was utilized with an averaged hourly wind speed series at sites S_1 and S_2 . BSO-SVR was used for 10 min wind speed prediction with 3-steps ahead. The main structures of the proposed models are shown in Figure 3.

3. Case Study and Analysis

3.1. Performance Evaluation Criteria. To evaluate the performance of the different models, three error evaluation criteria, mean absolute error (MAE), root mean squared error (MSE), and the mean absolute percentage error (MAPE), were adopted; the corresponding definitions are given by

$$\begin{aligned} \text{MAE} &= \frac{1}{T} \sum_{i=1}^T |y(i) - \hat{y}(i)|, \\ \text{MSE} &= \frac{1}{T} \sum_{i=1}^T (y(i) - \hat{y}(i))^2, \end{aligned} \quad (13)$$

$$\text{MAPE} = \frac{1}{T} \sum_{i=1}^T \left| \frac{y(i) - \hat{y}(i)}{y(i)} \right| \times 100\%,$$

where $y(i)$ and $\hat{y}(i)$ are the actual value and the forecasted value at time i , respectively.

3.2. Wind Speed Data Description. In our case study, a wind park located in a reference region in China was selected, and 10 min wind speed series V_1, V_2, V_3, V_4 over a period of time were collected from four wind turbines installed at S_1, S_2, S_3, S_4 in the park, respectively. Because the time units of the wind speed time series are 10 min, short-term forecasts at future times that are multiple time units ahead are calculated. As shown in Figure 4, the observed data reveal that the wind speeds at the study sites S_1, S_2, S_3, S_4 exhibited similar fluctuations and trends.

The available wind speed data were divided into training sets and test sets, which were utilized to determine the model's structure and to evaluate its predictive performance, respectively. In this paper, we were able to assume that

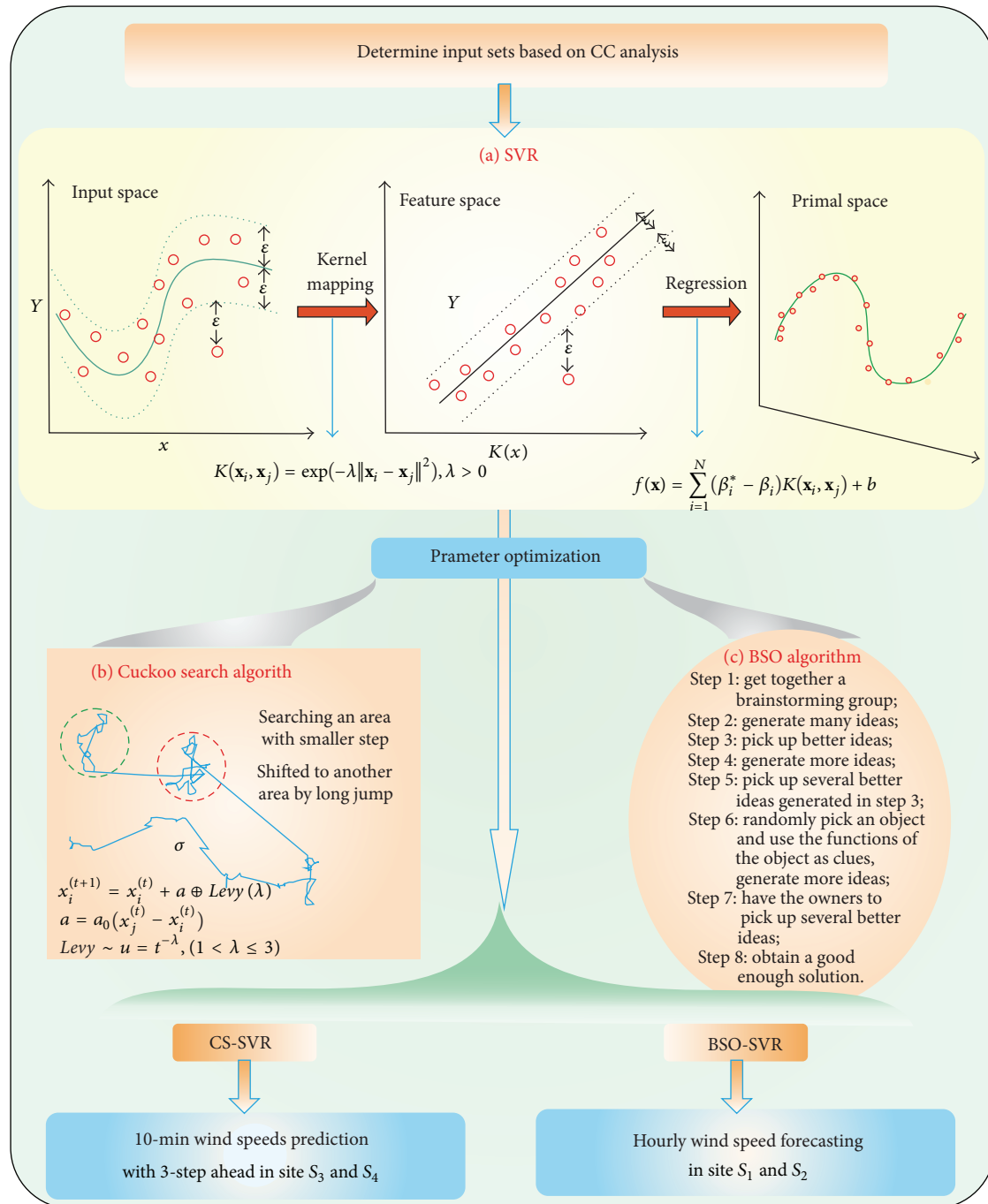


FIGURE 3: The structure of the proposed intelligent hybrid models.

the distances between sites S_i and S_j ($i \neq j$) are short, which indicates that a wind speed value occurring at S_i at time t may occur in S_j at the same time or at a delayed time. As a result, it is reasonable to assume that there are strong and weaker spatial-temporal relations between two wind speed series. The native predictor suggests that as the forecasting time lag increases, the correlation with past measurements becomes negligible [39]. Before constructing the models, we first examined the effects of spatial-temporal correlation between two sites.

3.3. Spatial-Temporal Correlation (STC) Analysis. A cross correlation-based STC analysis was performed to explore the time delays between the different wind speeds (V_i and V_j), which were referenced for the model inputs; this is critical for accurate forecasting model building.

From Figure 5, the coefficients between S_1 and S_2 , S_1 and S_3 , S_1 and S_4 , S_2 and S_3 , S_2 and S_4 , and S_3 and S_4 , are 0.9931, 0.9520, 0.9342, 0.8937, 0.8973 and 0.9577, respectively. This indicates that the spatial relationships between two sites are highly correlated, as the corresponding coefficients are

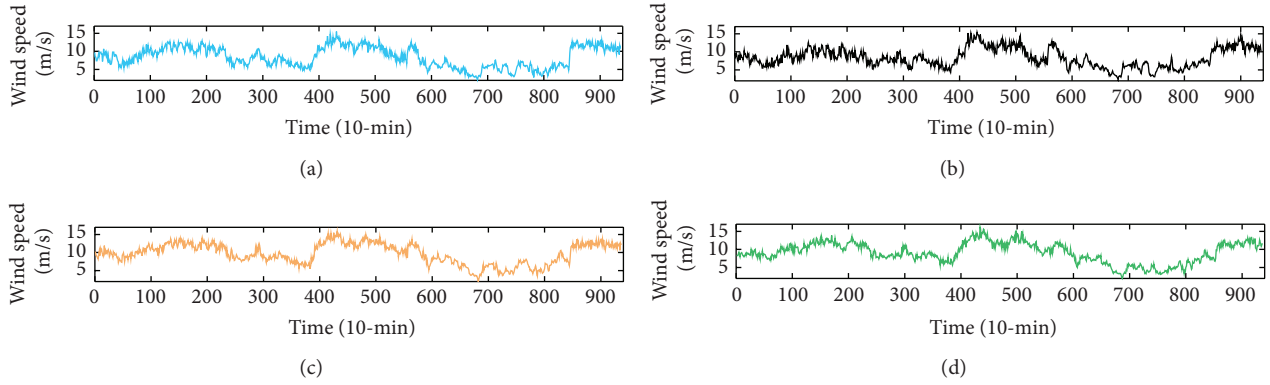


FIGURE 4: Observed wind speed series V_1, V_2, V_3, V_4 at the corresponding sites S_1, S_2, S_3, S_4 , which are illustrated in subplot (a), (b), (c), (d), respectively.

TABLE 1: The quantified cross correlation values.

Lags	Quantified cross correlation (QCC)									
	(V_1, V_1)	(V_1, V_2)	(V_1, V_3)	(V_1, V_4)	(V_1, V_5)	(V_2, V_2)	(V_2, V_3)	(V_2, V_4)	(V_3, V_3)	(V_3, V_4)
-10	0.77	0.71	0.77	0.76	0.72	0.74	0.74	0.80	0.79	0.80
-9	0.78	0.72	0.78	0.78	0.73	0.75	0.75	0.82	0.81	0.82
-8	0.80	0.73	0.80	0.79	0.75	0.76	0.76	0.83	0.83	0.83
-7	0.82	0.76	0.82	0.81	0.77	0.78	0.78	0.86	0.84	0.85
-6	0.84	0.78	0.84	0.83	0.80	0.79	0.80	0.87	0.86	0.87
-5	0.86	0.82	0.85	0.85	0.81	0.80	0.81	0.88	0.88	0.88
-4	0.88	0.82	0.87	0.87	0.84	0.82	0.83	0.90	0.89	0.90
-3	0.90	0.84	0.88	0.88	0.86	0.83	0.85	0.91	0.91	0.91
-2	0.92	0.86	0.90	0.90	0.89	0.85	0.87	0.93	0.92	0.93
-1	0.95	0.89	0.92	0.92	0.93	0.87	0.89	0.95	0.94	0.96
0	1.00	0.93	0.95	0.93	1.00	0.89	0.90	1.00	0.96	1.00
1	0.95	0.90	0.92	0.92	0.93	0.86	0.88	0.95	0.94	0.96
2	0.92	0.87	0.90	0.90	0.89	0.84	0.86	0.93	0.92	0.93
3	0.90	0.85	0.88	0.89	0.86	0.83	0.85	0.91	0.91	0.91
4	0.88	0.83	0.87	0.87	0.86	0.81	0.83	0.90	0.89	0.90
5	0.86	0.81	0.86	0.86	0.81	0.79	0.81	0.88	0.88	0.88
6	0.84	0.80	0.84	0.84	0.80	0.78	0.79	0.87	0.86	0.87
7	0.82	0.78	0.83	0.83	0.77	0.76	0.77	0.86	0.85	0.85
8	0.80	0.76	0.81	0.81	0.75	0.74	0.76	0.83	0.83	0.83
9	0.78	0.74	0.79	0.80	0.73	0.73	0.75	0.82	0.82	0.82
10	0.77	0.73	0.78	0.79	0.72	0.71	0.73	0.80	0.80	0.80

approximately equal to or larger than 0.9. Moreover, certain winds measured at S_i are expected to have arrived at S_j after a time delay and, to some extent, with decreasing wind speed. A traditional method of identifying the time delays between a local site and other sites, is to consider the cross correlation derived from the wind speeds. Figure 6 intuitively illustrates the CC between V_i and V_j for different time delays, and the quantified values are listed in Table 1. In Figure 6, the degree of correlation is indicated using distinctive colors. The red or hot color indicates a stronger correlation, while the blue color indicates a weaker correlation. When the lag is equal to zero, it implies a spatial correlation of (V_i, V_j) with no time delay. Moreover, if the lag k is larger or smaller than zero, a temporal

correlation exists between V_i and V_j with one time delayed or prior to the other. As can be visualized from each subplot of Figure 6, if the winds at site S_i are strong at S_j , there is a high correlation between V_i and V_j for a time delay within 5 units with the color of the region ranging from red to color. Additionally, the majority of the quantitative correlations of (V_i, V_j) for different time delays are 0.8, which demonstrates a high spatial-temporal correlation (STC) between S_i and S_j .

In this study the principal objective is to forecast wind speed at site S_i multiple time steps ahead. Then, we assume that the wind speed forecast series V_i should be composed of the historical data of the wind speeds at S_i and the wind speeds at previous times measured at the other sites

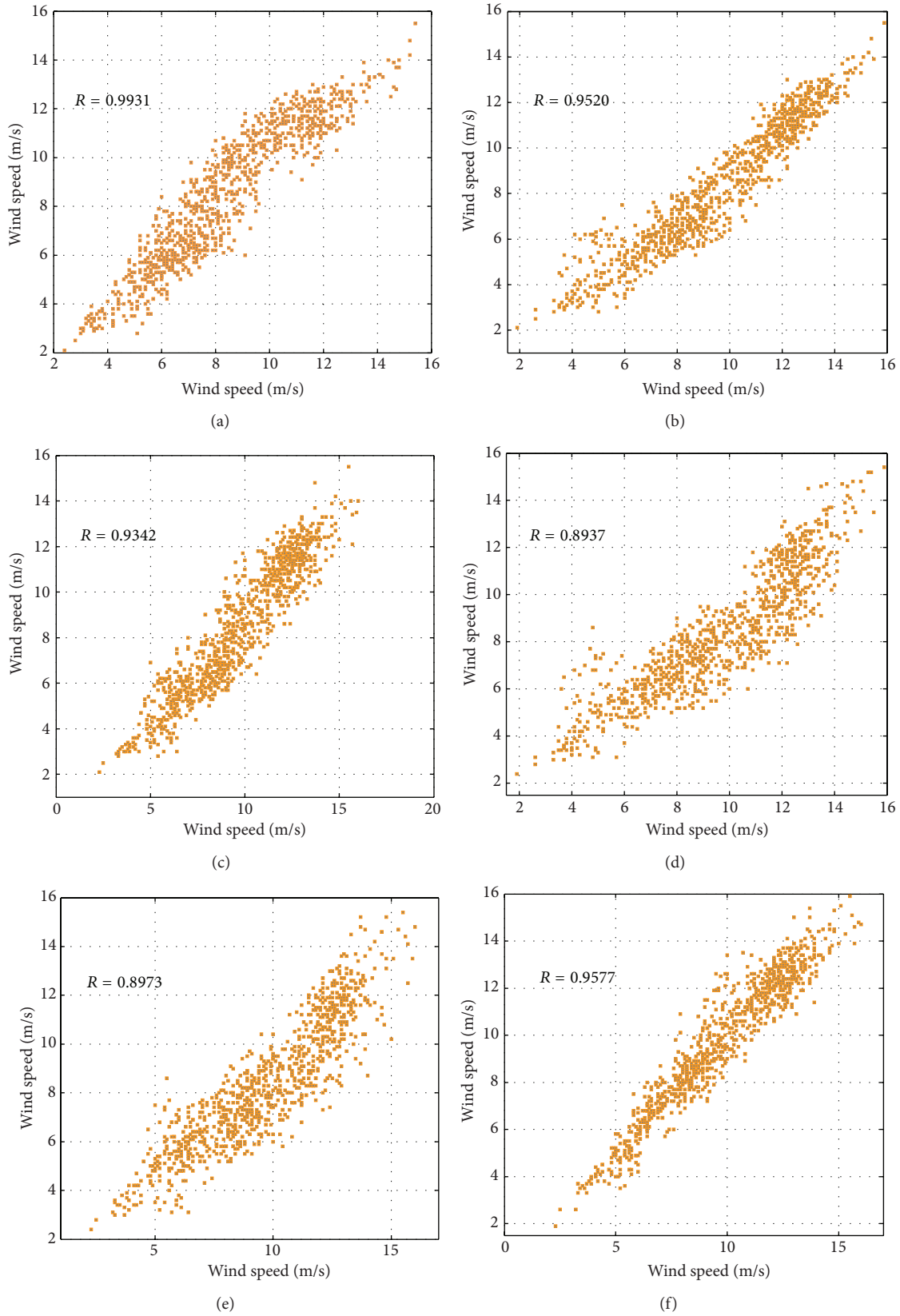


FIGURE 5: Scatter mapping of the measured wind speeds V_i and V_j at the corresponding sites S_i and S_j ($i \neq j$). (a)–(f) denotes the relationship between V_1 and V_2 , V_1 and V_3 , V_1 and V_4 , V_2 and V_3 , V_2 and V_4 , and V_3 and V_4 , respectively.

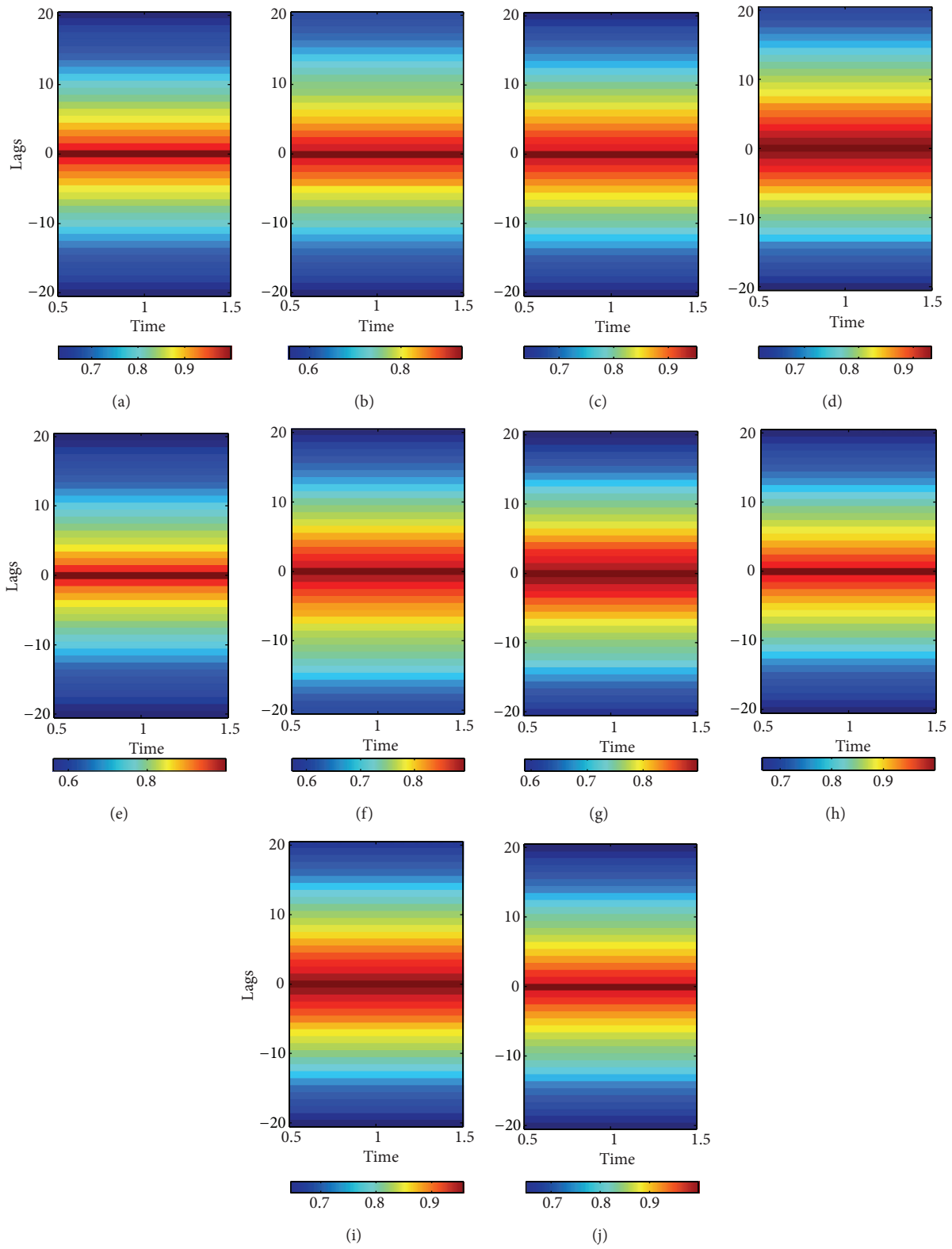


FIGURE 6: Cross correlation (CC). (a)–(j) denote cross correlation between two wind series V_i and V_j . V_i, V_j denote the measured wind speeds at sites S_i and S_j ($i, j = 1, 2, 3, 4$). (a) $RC(V_1, V_1)$; (b) $RC(V_1, V_2)$; (c) $RC(V_1, V_3)$; (d) $RC(V_1, V_4)$; (e) $RC(V_2, V_2)$; (f) $RC(V_2, V_3)$; (g) $RC(V_2, V_4)$; (h) $RC(V_3, V_3)$; (i) $RC(V_3, V_4)$; (j) $RC(V_4, V_4)$.

S_j ($j \neq i, j = 1, 2, 3, 4$). Thereby, an advanced predictive model providing multiple-steps wind forecasts, can be formulated as follows:

$$\widehat{V}_i(t+h) = f_i(V_1(t), V_2(t), V_3(t), V_4(t)), \quad (i = 1, 2, 3, 4), \quad (14)$$

where $\widehat{V}_i(t+h)$ denotes the h -step ahead forecasting wind speed, and $V_1(t), V_2(t), V_3(t)$ and $V_4(t)$ are the past values of the measured wind speeds.

3.4. Analysis of Forecasting Results. In this section, four cases were used to validate the performance of the proposed BSO-based and CS-based SVR models. Firstly, the examined 10 min wind speed data at sites S_1 and S_2 were averaged over 1 h intervals, with 258 pairs of data for model training and 30 pairs for model testing. Furthermore, owing to the cross correlation, the input sets were lagged time steps of the current values of the wind speeds at the local and remote sites.

For SVR-based models, to avoid excessive errors during the training stages, the cross validation (CV) technique was utilized. It is well known that the predictive accuracy of a SVR model largely relies on reasonable values of the kernel parameter σ and hyper parameter C . Then, for the purpose of improving the forecasting capacity of the SVR, the BSO algorithm and the CS method were used to optimize the parameter selection of the SVR model. The radial basis function (RBF) was selected as the kernel function because of its ability to nonlinearly map the training sets onto an infinite dimensional space, which is suitable to handle linear and nonlinear relationship problems.

For the sake of comparison, in addition to the proposed model another model was tested, namely the Elman recurrence neural network (ERNN). The ERNN, a modified form of the artificial neural networks with an internal layer of self-recurrent neurons, is a nonlinear mapping function that can capture the nonlinearity of the studied wind speed series. Then, for fair comparison, both models with their corresponding mapping capacity were chosen for comparison with the proposed model. Based on the STC analysis discussed in the above subsection, the inputs to the models were the wind values from each one of the four measurement sites at several lag times. Thereby, both the local and the remote past information were introduced into these models for adequate periods of 1 h and 30 min behind.

3.4.1. BSO-Based SVR Forecasting Model: Two Cases for Averaged over 1 h Interval. Herein, we established the proposed and comparison models based on the same wind speed dataset and obtained simulation and forecasting results to compare the models. Referring to Section 3.1, the evaluation criteria adopted in this paper are MAPE, MAE and MSE. The forecasted values and errors for the proposed BSO-SVR, basic SVR and ERNN models are shown in Figure 7 and Table 2.

As shown in Figure 7, over the predictive horizon the timing and range of variation of the forecasted series approximate the measured series, but some forecasted values slightly over- or underestimate the actual values. To evaluate the models' forecasting performance quantitatively, the evaluation errors

TABLE 2: Forecasting errors over the prediction horizon: averaged over 1-h interval for Case 1 and Case 2.

Criteria	Case 1			Case 2		
	BSO-SVR	SVR	ERNN	BSO-SVR	SVR	ERNN
MAE	0.37	0.79	0.89	0.55	1.03	1.09
MSE	0.52	1.03	1.05	1.23	1.75	1.74
MAPE	4.56%	9.59%	10.61%	7.35%	13.49%	14.23%

TABLE 3: Forecasting errors over the prediction horizon: averaged over 10-min interval for Case 3 and Case 4.

Criteria	Case 3			Case 4		
	CS-SVR	SVR	ERNN	CS-SVR	SVR	ERNN
MAE	0.63	0.72	0.63	0.47	0.83	0.78
MSE	0.83	0.89	0.73	0.61	0.98	1.02
MAPE	7.13%	8.20%	7.20%	6.01%	8.15%	7.73%

over the prediction horizon are listed in Table 2. For both case studies, the smallest values of MAPE, MAE, and MSE were obtained by the proposed BSO-SVR model, which demonstrates that BSO-SVR outperforms the other two models. As shown in Figures 7(c)–7(e) and 7(f)–7(h), for both cases the forecasting errors obtained by the BSO-SVR model were much lower than that of single SVR model; MAPE, MAE and MSE decreased by 5.03%, 0.42 (m/s) and 0.51 (m/s) for Case 1, and 6.14%, 0.58 (m/s) and 0.52 (m/s) for case 2, respectively. Lower values of the error criteria indicate better forecasting results. Thus, it can be concluded that the forecasting performance of the BSO-based SVR model was significantly improved because of the strong ability of BSO algorithm to optimize parameters, and the proposed method can be well utilized for short-term wind speed forecasting over 1 h ahead.

3.4.2. CS-Based SVR Forecasting Model: Two Cases for Averaged over 10 min Interval. The Cuckoo Search (CS) SVR model, which is based on a recently proposed artificial intelligence optimization algorithm, was studied using the 10 min wind speeds at sites S_3 and S_4 . Forecasts using the ERNN, single SVR and the proposed CS-SVR are compared in Figure 8 and Table 3.

As shown in Figure 8, the timing and range of variation of the forecasted series mostly agree with the actual series, but some forecasted values were significantly underestimated and overestimated. Each model's forecasting performance can be quantitatively evaluated based on the errors over the performance horizon. Table 3 gives the predictive error values for both case studies; the smallest value of MAPE, MAE, and MSE was obtained by the proposed CS-SVR model, which demonstrates that the CS-SVR had better forecasting performance. As listed in Table 3, the forecasting errors obtained by CS-SVR model were much smaller than those of the single SVR model; MAPE, MAE and MSE decreased by 1.07%, 0.09 (m/s) and 0.06 (m/s) for Case 1, and 2.14%, 0.36 (m/s) and 0.37 (m/s) for case 2, respectively. However, the forecasting errors of the ERNN model are approximately the same as the CS-SVR errors, which means that ERNN

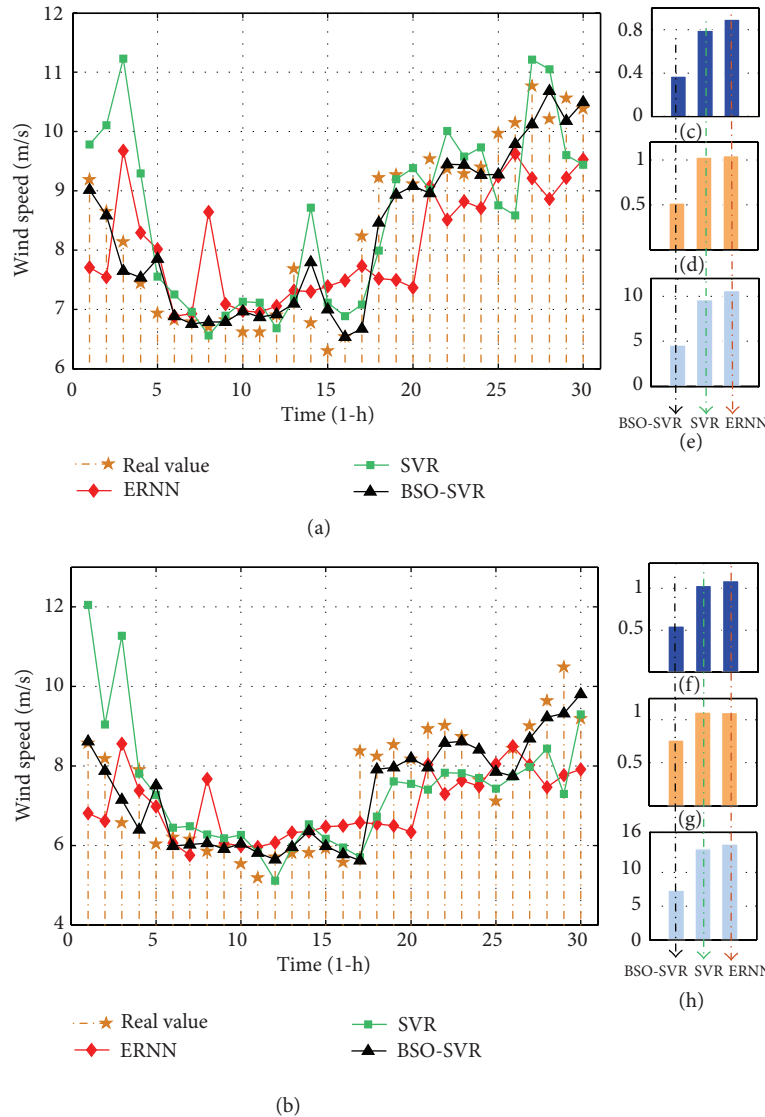


FIGURE 7: Forecasting values over the prediction horizon: averaged over 1 h interval for Case 1 (a) and Case 2 (b). (c)–(h): the corresponding histogram of the forecasting errors. (a) and (f): MAE; (d) and (g): MSE; (e) and (h): MAPE %.

performed well in Case 3. These results do not indicate that a particular method was determined to be superior to the others; model performance under specific conditions should be analyzed and understood, and incremental improvements should be made based on the knowledge gained [40]. However, on the whole, the CS-based SVR model was significantly improved, and the proposed method can be utilized for short-term wind speed forecasting over 45 min ahead.

The proposed BSO-SVR forecasting model outperformed the other two models when using average hourly wind speed data. The CS-SVR model performed well in 10 min wind speed forecasting. Based on the STC analysis results, we believe that the developed two models can be applied to the other two data sets because of their similar distributions and statistical characteristics.

4. Conclusions

As one of the promising forms of renewable energy, wind power has been growing rapidly over the world. Given the significant influence of wind speed on wind power, considerable efforts have been devoted to wind speed prediction.

In this study, two intelligent hybrid models are proposed for short-term wind speed forecasting by integrating intelligent optimization into a SVR model with inputs determined by a CC-based analysis of spatial-temporal relationships. To determine the spatial and temporal relations between wind speed series collected from different wind turbines, CC analysis provided adequate information about the auto- and cross correlations of the wind speed time series. For accurate

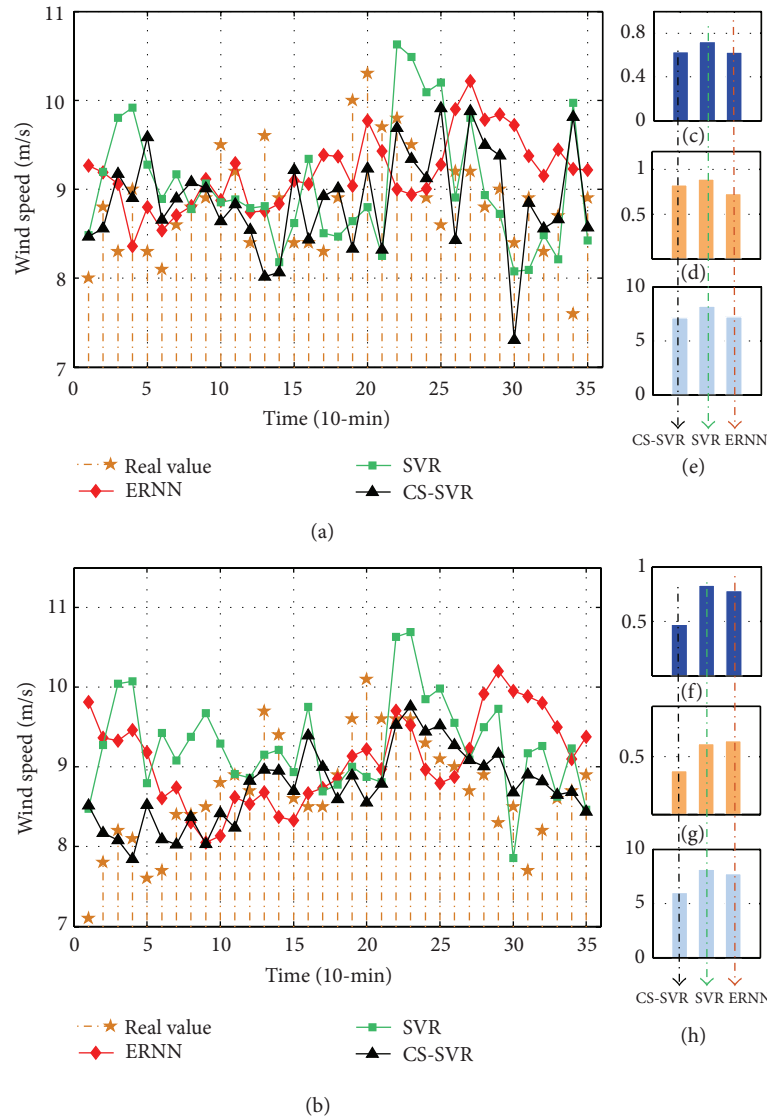


FIGURE 8: Forecasting values over the prediction horizon over multiple-steps: averaged over 10 min interval for Case 3 (a) and Case 4 (b). (c)–(h): the corresponding forecasting error histogram of (a) and (b): MAE; (d) and (g): MSE; and (e) and (h): MAPE %.

short-term forecasts of wind speeds using an intelligent hybrid model ensemble, two different SVR models were considered: BSO-SVR and CS-SVR. BSO-SVR was used with averaged hourly wind speed series at sites S_1 and S_2 . BSO-SVR was used for 10 min wind speed prediction for 3-steps ahead. The results of this study demonstrated that the proposed hybrid models were efficient when conducting short-term wind prediction compared to the single SVR and ERNN models, which indicates that the hybrid models can improve model prediction and enhance forecasting efficiency.

Conflict of Interests

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

Acknowledgment

This work was supported by the National Natural Science Foundation of China (Grant no. 70973015).

References

- [1] J. Wu, J. Wang, and D. Chi, "Wind energy potential assessment for the site of Inner Mongolia in China," *Renewable and Sustainable Energy Reviews*, vol. 21, pp. 215–228, 2013.
- [2] J. Wang, S. Qin, Q. Zhou, and H. Jiang, "Medium-term wind speeds forecasting utilizing hybrid models for three different sites in Xinjiang, China," *Renewable Energy*, vol. 76, pp. 91–101, 2015.
- [3] 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.
- [4] P. Kou, F. Gao, and X. Guan, “Sparse online warped Gaussian process for wind power probabilistic forecasting,” *Applied Energy*, vol. 108, pp. 410–428, 2013.
 - [5] P. S. Georgilakis, “Technical challenges associated with the integration of wind power into power systems,” *Renewable and Sustainable Energy Reviews*, vol. 12, no. 3, pp. 852–863, 2008.
 - [6] J. Hu, J. Wang, and G. Zeng, “A hybrid forecasting approach applied to wind speed time series,” *Renewable Energy*, vol. 60, pp. 185–194, 2013.
 - [7] J. Chen, X. Xue, M. Ha, D. Yu, and L. Ma, “Support vector regression method for wind speed prediction incorporating probability prior knowledge,” *Mathematical Problems in Engineering*, vol. 2014, Article ID 410489, 10 pages, 2014.
 - [8] Z. Guo, D. Chi, J. Wu, and W. Zhang, “A new wind speed forecasting strategy based on the chaotic time series modelling technique and the Apriori algorithm,” *Energy Conversion and Management*, vol. 84, pp. 140–151, 2014.
 - [9] Z. Guo, Y. Dong, J. Wang, and H. Lu, “The forecasting procedure for long-term wind speed in the Zhangye area,” *Mathematical Problems in Engineering*, vol. 2010, Article ID 684742, 17 pages, 2010.
 - [10] M. A. Mohandes, T. O. Halawani, S. Rehman, and A. A. Hussain, “Support vector machines for wind speed prediction,” *Renewable Energy*, vol. 29, no. 6, pp. 939–947, 2004.
 - [11] R. G. Kavasseri and K. Seetharaman, “Day-ahead wind speed forecasting using f-ARIMA models,” *Renewable Energy*, vol. 34, no. 5, pp. 1388–1393, 2009.
 - [12] A. Tascikaraoglu and M. Uzunoglu, “A review of combined approaches for prediction of short-term wind speed and power,” *Renewable and Sustainable Energy Reviews*, vol. 34, pp. 243–254, 2014.
 - [13] L. Landberg, “Short-term prediction of the power production from wind farms,” *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 80, no. 1–2, pp. 207–220, 1999.
 - [14] M. C. Alexiadis, P. S. Dokopoulos, H. S. Sahsamanoglou, and I. M. Manousaridis, “Short-term forecasting of wind speed and related electrical power,” *Solar Energy*, vol. 63, no. 1, pp. 61–68, 1998.
 - [15] M. Negnevitsky and C. W. Potter, “Innovative short-term wind generation prediction techniques,” in *Proceedings of the IEEE PES Power Systems Conference and Exposition (PSCE '06)*, pp. 60–65, IEEE, November 2006.
 - [16] W. Zhang, J. Wu, J. Wang, W. Zhao, and L. Shen, “Performance analysis of four modified approaches for wind speed forecasting,” *Applied Energy*, vol. 99, pp. 324–333, 2012.
 - [17] J. L. Torres, A. García, M. De Blas, and A. De Francisco, “Forecast of hourly average wind speed with ARMA models in Navarre (Spain),” *Solar Energy*, vol. 79, no. 1, pp. 65–77, 2005.
 - [18] E. Cadenas and W. Rivera, “Wind speed forecasting in the South Coast of Oaxaca, México,” *Renewable Energy*, vol. 32, no. 12, pp. 2116–2128, 2007.
 - [19] 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.
 - [20] I. G. Damousis, M. C. Alexiadis, J. B. Theocharis, and P. S. Dokopoulos, “A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation,” *IEEE Transactions on Energy Conversion*, vol. 19, no. 2, pp. 352–361, 2004.
 - [21] 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.
 - [22] J. M. Morales, R. Mínguez, and A. J. Conejo, “A methodology to generate statistically dependent wind speed scenarios,” *Applied Energy*, vol. 87, no. 3, pp. 843–855, 2010.
 - [23] G. Li, J. Shi, and J. Zhou, “Bayesian adaptive combination of short-term wind speed forecasts from neural network models,” *Renewable Energy*, vol. 36, no. 1, pp. 352–359, 2011.
 - [24] M. Monfared, H. Rastegar, and H. M. Kojabadi, “A new strategy for wind speed forecasting using artificial intelligent methods,” *Renewable Energy*, vol. 34, no. 3, pp. 845–848, 2009.
 - [25] S. Salcedo-Sanz, E. G. Ortiz-García, A. Portilla-Figueras, L. Prieto, and D. Paredes, “Hybridizing the fifth generation mesoscale model with artificial neural networks for short-term wind speed prediction,” *Renewable Energy*, vol. 34, no. 6, pp. 1451–1457, 2009.
 - [26] 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.
 - [27] http://en.wikipedia.org/wiki/Cross-correlation#Time_delay_analysis.
 - [28] L. Liu and Y. Wang, “Cross-correlations between spot and futures markets of nonferrous metals,” *Physica A: Statistical Mechanics and its Applications*, vol. 400, pp. 20–30, 2014.
 - [29] J. L. Elman, “Finding structure in time,” *Cognitive Science*, vol. 14, no. 2, pp. 179–211, 1990.
 - [30] Q. He, *Neural Network and Its Application in IR*, Graduate School of Library and Information Science, University of Illinois at Urbana, Champaign, Ill, USA, 1999.
 - [31] J. Hertz, *Introduction to the Theory of Neural Computation. Vol. 1. Basic Books*, 1991.
 - [32] V. Vapnik, S. E. Golowich, and A. Smola, “Support vector method for function approximation, regression estimation, and signal processing,” in *Advances in Neural Information Processing Systems*, pp. 281–287, 1997.
 - [33] W.-C. Hong, Y. Dong, W. Y. Zhang, L.-Y. Chen, and B. K. Panigrahi, “Cyclic electric load forecasting by seasonal SVR with chaotic genetic algorithm,” *International Journal of Electrical Power and Energy Systems*, vol. 44, no. 1, pp. 604–614, 2013.
 - [34] 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.
 - [35] 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.
 - [36] X.-S. Yang and S. Deb, “Engineering optimisation by cuckoo search,” *International Journal of Mathematical Modelling and Numerical Optimisation*, vol. 1, no. 4, pp. 330–343, 2010.
 - [37] Y. Shi, “An optimization algorithm based on brainstorming process,” *International Journal of Swarm Intelligence Research*, vol. 2, no. 4, pp. 35–62, 2011.
 - [38] J. Ahmed and Z. Salam, “A Maximum Power Point Tracking (MPPT) for PV system using Cuckoo Search with partial shading capability,” *Applied Energy*, vol. 119, pp. 118–130, 2014.
 - [39] T. G. Barbounis and J. B. Theocharis, “Locally recurrent neural networks for wind speed prediction using spatial correlation,” *Information Sciences*, vol. 177, no. 24, pp. 5775–5797, 2007.
 - [40] Z. Guo, J. Zhao, W. Zhang, and J. Wang, “A corrected hybrid approach for wind speed prediction in Hexi Corridor of China,” *Energy*, vol. 36, no. 3, pp. 1668–1679, 2011.

- [41] J. Wang, S. Qin, S. Jin, and J. Wu, "Estimation methods review and analysis of offshore extreme wind speeds and wind energy resources," *Renewable and Sustainable Energy Reviews*, vol. 42, pp. 26–42, 2015.



Hindawi

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

