

## Research Article

# Hybrid Wind Speed Forecasting Model Study Based on SSA and Intelligent Optimized Algorithm

Wenyu Zhang, Zhongyue Su, Hongli Zhang, Yanru Zhao, and Zhiyuan Zhao

*Key Laboratory for Semi-Arid Climate Change of the Ministry of Education, College of Atmospheric Sciences, Lanzhou University, Key Laboratory of Arid Climatic Change and Reducing Disaster of Gansu Province, Lanzhou 730000, China*

Correspondence should be addressed to Wenyu Zhang; [zhangwy@lzu.edu.cn](mailto:zhangwy@lzu.edu.cn)

Received 9 January 2014; Accepted 13 February 2014; Published 3 April 2014

Academic Editor: Suohai Fan

Copyright © 2014 Wenyu Zhang 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.

Accurate wind speed forecasting is important for the reliable and efficient operation of the wind power system. The present study investigated singular spectrum analysis (SSA) with a reduced parameter algorithm in three time series models, the autoregressive integrated moving average (ARIMA) model, the support vector machine (SVM) model, and the artificial neural network (ANN) model, to forecast the wind speed in Shandong province, China. In the proposed model, the weather research and forecasting model (WRF) is first employed as a physical background to provide the elements of weather data. To reduce these noises, SSA is used to develop a self-adapting parameter selection algorithm that is fully data-driven. After optimization, the SSA-based forecasting models are applied to forecasting the immediate short-term wind speed and are adopted at ten wind farms in China. Finally, the performance of the proposed approach is evaluated using observed data according to three error calculation methods. The simulation results from ten cases show that the proposed method has better forecasting performance than the traditional methods.

## 1. Introduction

Entering the 21st century, countries worldwide face the dual pressures of environmental protection and economic growth. To reduce energy-related toxic emissions in the current energy infrastructure, renewable energy should be utilized with the goal of maintaining sustainable development and creating a better ecological environment. In its “Special Report on Renewable Energy Sources” the IPCC 2011 states that renewable energies are affordable and economically viable options for meeting the electricity needs of people in developing countries [1]. The extensive use of principal energy sources has raised concerns about future security and the impact they have had on climate. To overcome the challenges of improving or maintaining energy security and mitigating climate change, low-carbon emission technologies have attracted considerable interest worldwide, resulting in increased electricity generation from renewable sources. Unquestionably, renewable energies have huge potential, but how quickly their benefits can meet the growth of global energy demand hinges on government support to make

renewable energy cost-competitive in energy markets. As fossil fuel prices increase and renewable technologies mature, renewable energies are becoming increasingly competitive; a global effort will be required to construct a low-carbon society [2]. For instance, the International Energy Agency anticipates a rapid expansion of renewables such as hydroelectricity, wind, solar, and geothermal energy; production will rise from 840 million tons of oil equivalent in 2008 to nearly 3250 million tons of oil equivalent in 2035 [3]. The “Global Wind Energy Outlook 2010” projects that world wind energy production will increase more than tenfold by 2030, emphasizing the importance of wind as a key contributor to improving energy security and reducing greenhouse gas emissions [4].

Among the various renewable energies, wind energy is the most promising. Wind energy has been the fastest growing renewable energy technology in the last ten years. Furthermore, it has been playing a crucial role in everyday life for people in developing countries, who account for one-third of the world's total population. Wind energy also supports developed countries; as one source of clean energy, it helps them meet the 21st century energy demands [5].

Because of its economic and ecological advantages, wind power has recently become one of the most popular alternative energies, accounting for approximately 10% of the national power usage in European countries, and this figure exceeds 15% in Spain, Germany, and the USA [6]. According to the Global Wind Energy Council Report (2011), the world's wind power capacity grew by 22.5% in 2010, adding 35,802 MW to bring total installations to 194,390 MW. Almost half of these additions were made in China, which experienced an annual growth rate of approximately 65% [7]. The world's total installed capacity of wind power reached 254 GW in June 2012 [8]. However, the main obstacle facing wind industry development is that wind is an intermittent energy source, which means that there is large variability in the production of energy due to various factors such as wind speed, air density, and turbine characteristics. Specifically, horizontal air motion is defined as wind [9], and wind is driven by uneven cooling and heating on the earth's surface, as well as the rotation of the earth. Thus, the occurrence of wind has strong uncertainty, in both space and time. The nondeterminacy seriously limits wind power penetration and threatens grid security. Although wind power drives the turbines that generate electricity, the complex fluctuations of wind make it difficult to predict the power output. The theoretical amount of energy that might be generated from wind is proportional to the cube of the wind speed, and slight changes in the wind speed might cause significant changes in the total amount of electricity generated from the wind; this also causes obstacles for power transportation [10]. To guarantee the security of the grid system, the dispatching department must balance the grid's production and consumption within very small time intervals [11]. Moreover, due to the lack of accurate information about wind occurrence, the efficiency of wind turbines can also be limited [12]. Several academic works addressing wind energy attest to the increasing interest in this important theme.

Generally, there are two ways to solve this problem in wind power generation. One is the large-scale transformation of the existing electrical power system; the most popular method is smart grid transformation, which consists of a digitally enabled power system [13]. The core of a smart grid is the integration of secure and high-speed data communication—based on advanced computers, electronic equipment, intelligent components, and more—to operate the mixture system intelligently and effectively [14, 15]. Using the real-time information, a smart grid can facilitate the development of wind power. When wind-generated power is insufficient, other forms of energies can be used to supplement power for a short time through a smart grid. However, currently, there is no widely accepted standardized communication/network infrastructure that could be applied to power transformation through a smart grid [16]. Another possible solution is to improve wind power techniques, including wind-prediction techniques, wind turbine techniques, wind energy storage techniques, and combined dispatching. Accurate wind information is important for estimating the wind power output. Accurate estimations can benefit not only increasing wind power penetration and maintaining a stable grid system but also the combined

dispatching in a mixture power grid. Under current electrical systems, developing wind-prediction techniques can be an effective way to guarantee the security and stability of the electrical system without increasing the running cost.

In the past decades, many approaches have been developed for short-term load forecasting. These methods can be categorized into different groups. Some of these methods assume a time series model structure and then try to identify its parameters. In actual power generation, wind predictions—especially the short-term forecasts—are important for scheduling, controlling, and dispatching the energy conversion systems [17]. The most important characteristic of wind—speed—can be easily influenced by other meteorological factors, such as air temperature and air pressure, as well as obstacles and terrain [18]. Thus, wind speed prediction is not easy to address; moreover, wind speed modeling has become one of the most difficult problems [19, 20]. The forecasting approach can be determined based on the available information and the time scale in question, which will affect its application. Short-term wind speed forecasting is a subclass of wind speed forecasting. The time scales for short-term forecasting range from a few seconds to minutes, hours, or several days [21].

Many methods of forecasting wind speed have been proposed. In general, they can be classified into two categories: physical methods and statistical methods.

Physical methods are often referred to as meteorological predictions of wind speed; they involve the numerical approximation of models that describe the state of the atmosphere [22]. Numerical weather prediction (NWP) techniques, one type of physical model-based approach, rely on a class of physical models with numerical parameters characterizing local meteorological and geographical properties, such as temperature, atmospheric pressure, surface roughness, and obstacles [23]. NWP techniques include the weather research and forecasting (WRF) model and the fifth generation mesoscale model (MM5). These models always use physical data such as topography information, pressure, and temperature to forecast wind speed in the future [24, 25]. Typically, prediction methods using NWP forecasts outperform statistical approaches after a 3–6 h look-ahead time, whereas statistical approaches turn out to be quite reliable for very short-term forecasts, that is, less than 6 h. The optimal model most likely consists of a mixed approach, which is very often adopted by utilities to combine high accuracy for very short horizons with longer forecasts of up to 48–72 h [26]. Unlike physical models, statistical methods make forecasts by finding relationships using historical wind speed data and, sometimes, other variables (e.g., wind direction or temperature). The data used are recorded at the observation site or at other nearby locations where data are available. In the literature, many statistical methods have been applied to this topic, such as the autoregressive integrated moving average (ARIMA) model, Kalman filters, and the generalized autoregressive conditional heteroscedasticity (GARCH) model. The statistical models can be used at any stage in the modeling process, and they often combine various methods into one. Physical and statistical models each have their own advantages for wind speed prediction, but few forecasts use

only one model. Often, the results of the physical prediction merely represent the first step towards forecasting the wind; thus, the physically predicted wind speed can be regarded as an auxiliary input to other statistical models [27]. Currently, grey models (GM) [28, 29] and models based on artificial intelligence (AI) techniques have been developed for this area, including the artificial neural networks (ANNs) of multilayer perceptrons (MLP) [30, 31], radial basis function (RBF) [32], recurrent neural networks [33, 34], and fuzzy logic [35, 36]. Neural networks can learn from past data and recognize hidden patterns or relationships in historical observations and use them to forecast future values. The results indicate that prediction errors resulting from the Bayesian combination approach always become smaller, which is in contrast to artificial neural networks, whose performance is not consistent when the site or evaluation criterion changes [3]. An annealing clustering ANN [37], a BP (back propagation) neural network with rough set [38], a neural network combining wavelet transform and adaptive mutation particle swarm optimization [39], an ANN based on fuzzy logic methods [40], and Bayesian neural networks integrating the Monte Carlo algorithm [41] have been proposed for short-term wind speed forecasting. Combination models include the adaptive particle swarm optimization-based combined method [42], wavelet transform combination model based on a neural network and an evolutionary algorithm [43], wavelet transforms and adaptive models [44], and an adaptive fuzzy combination model based on the self-organizing map and support vector regression [45]. In fact, conventionally, the forecasting method is not classified as either physical or statistical, because most forecasting models include both techniques.

Several methods are being used to diagnose the dynamical characteristics of observational wind speed time series. The singular spectrum analysis (SSA) method, which is a powerful technique for time series analysis, has been employed elegantly and effectively in several areas, such as hydrology, geophysics, climatology, and economics [46–49]. Traditional approaches are based on statistical models, including linear regression methods [50], exponential smoothing [51], Box-Jenkins approaches [52], and Kalman filters [53]. Essentially, most of the traditional approaches are based on linear analysis. However, the wind speed series are usually nonlinear. As discussed earlier, noise signals caused by unstable factors increase the difficulties associated with convergence and forecasting accuracy. Guo et al. [54] and Dong et al. [55] considered the first IMF (intrinsic mode function) obtained by EMD (empirical mode decomposition) as noise and demonstrated that eliminating the first IMF improved the forecasting accuracy in their experiments. In addition, the EMD-based signal filtering proposed in [56] indicates that noise potentially contains the first IMF or the first several IMFs. Here, SSA [57, 58] has been employed to characterize the properties of wind speed, and the hybrid model with SSA is used for short-time forecasting. The SSA technique incorporates the elements of classical time series analysis, multivariate geometry, multivariate statistics, signal processing, and dynamical systems. The aim of SSA is to obtain a decomposition of the original series into a set of

independent and interpretable components, which include a slowly varying trend, oscillatory components, and random noise [57]. One of the differences between traditional time series analysis methods and SSA is that SSA and SSA-related methods could be applied to both classical time series analysis problems and various other situations, such as exploratory analysis for data-mining and parameter estimation in signal processing [59]. The main reason for using nonparametric or data-driven techniques is that no previous assumptions are required to analyze and perform forecasts, such as the normality of residuals, the stationarity of the time series, or a predefined model structure. Although SSA does not involve any forecasting algorithms, the analysis reveals the natural characteristics of time series. One of the main advantages of SSA compared to other nonparametric methods is that only two parameters are used to simulate the time series in many implementations [60], and no model is assumed before the SSA method is adopted; the subspace-based model is built adaptively. The other essential difference between SSA and the majority of methods is that SSA does not require *a priori* model of a trend, such as *a priori* knowledge of the number of periodicities and period values, to analyze time series with a trend or periodicity.

Briefly, the SSA method decomposes a time series into a number of components with simpler structures, such as a slowly varying trend, oscillations, and noise. SSA belongs to the general category of principal component analysis (PCA) methods, which apply a linear transformation of the original data space into a feature space, where the data set may be represented by effective features while retaining most of the information content of the data [61]. Wavelet analysis is a powerful tool for PCA and has been used for feature extraction and denoising wind speed for a long time. To optimally fit the forecasting model, the data must be stationary and have a normal distribution. In SSA, these problems do not exist, as the technique does not depend on any parameters, as does the other model for the trend. The SSA technique has been used in a variety of fields, such as signal processing, nonlinear dynamics, climate, medicine, and mathematical statistics [62]. Additionally, some denoising methods are based on singular value decomposition (SVD), similar to the SSA method.

The basic SSA method and forecasting models are presented to address the short-term wind speed forecasting problem. The employed models meet two goals: (1) using the real data for SSA to eliminate cumulative error and (2) estimating the data trend using the history of the data to forecast short-term data in wind speed. Simulation results present the effectiveness of the proposed method in characterizing and predicting time series.

In this paper, a hybrid prediction algorithm is proposed for short-term wind speed forecasting. The proposed algorithm, which integrates the advantages of the time series and SSA, is adopted to develop a prediction model. The rest of this paper is organized as follows. In Section 2, the data description and the research background are outlined. The NWP model WRF is also introduced in Section 2; in Section 3, the SSA model for feature extraction is introduced and the main steps of the model are given, which include

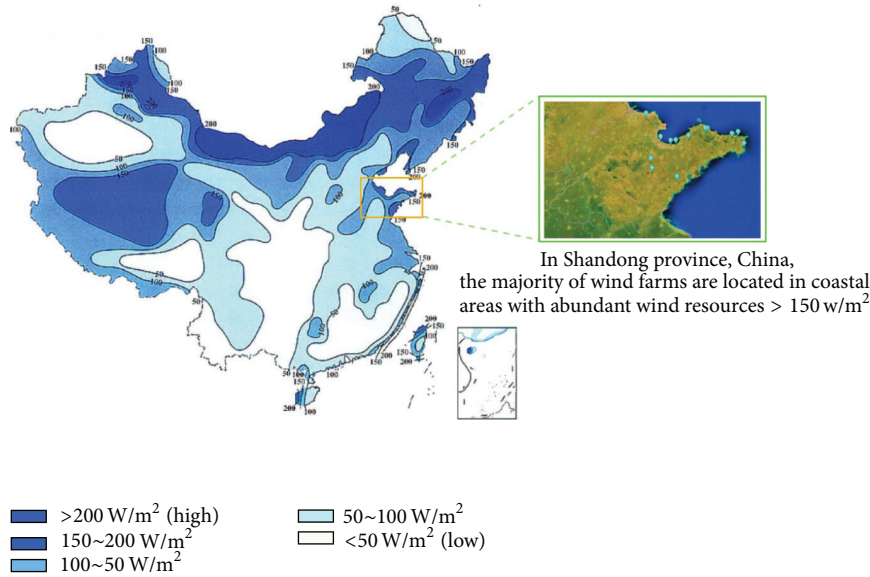


FIGURE 1: Chinese wind energy resource distribution and wind farms in Shandong province.

decomposition and reconstruction. The methods contain the basic processing of physical methods. The performance of wind speed data is described. The main forecasting methods and the results are presented and discussed in Section 5, with comparisons to other methods. A brief review of this paper and the future research are in Section 6.

## 2. Background and Data Collection

Wind farms at 10 sites in Shandong, China, applied our methodologies. The Shandong province, located between  $114^{\circ}47.5' E$  and  $122^{\circ}42.3' E$  and between  $34^{\circ}22.9' N$  and  $38^{\circ}24.01' N$  in the eastern coastal area of China on the lower reaches of the Yellow River, covers an area of  $1.57 \times 10^5 \text{ km}^2$ , which accounts for 1.6% of the total land area of China. As an economically powerful province with a large population ( $9.4 \times 10^7$  inhabitants in 2007), it has experienced rapid and sustained economic development, with enormous energy consumption. With the rapid increase in wind energy in China, the total installed capacity of wind power in Shandong province reached 4562.3 MW in 2012. Some of wind farms are near the Pacific Ocean, and their nominal power can reach up to 49.5 MW. The prevailing wind speed in Shandong in April averages 10.27 m/s. No additional information is provided in the database about these wind farms (location, nominal power, and so on) for confidentiality reasons. This database contains information about 10 wind farms. We denote the selected wind farms numbered 1 to 10 by wind farms A to J, respectively. These wind farms have variable situations and land surfaces, which allowed the model's effectiveness to be sufficiently tested. Thus, our proposed model could be demonstrated in very different settings.

As shown in Figure 1, some wind farms in this study are located near the sea; because of the different heat capacity values of the sea and the land, there will be a sea breeze, which

will blow in opposite directions during the day and night. In this region, frequent strong winds make it quite difficult to accurately forecast synoptic processes.

In all cases, we collect observational speed data from the wind farm and meteorological weather forecasts from a NWP model. In this database, historical wind speed data are obtained every 15 min and then averaged for each hour. We have quarter data from 00:00, April 1, 2013, to 23:45, April 30, 2013, which amounts to 2650 data points. Meteorological wind simulation results include the temperature, pressure, humidity, wind speed, and direction provided by the WRF model. The forecasts of the models are hourly and are given in terms of the coordinated universal time (UTC), and the maximum prediction horizon is 96 h. Generally, in these models, historical wind speed is used as an input at one of the points. However, in the NN model, the temperature, pressure, humidity, and direction results from the WRF are also used as input data. In all of the wind farms analyzed, data from the two groups (historical measurement and model forecasting) are divided in two sets: the first portion of data is used to train the models and the remaining data points are used to validate the models.

The WRF mesoscale numerical model is now the current generation "community" physics-based atmospheric model, serving the needs of both atmospheric research and operational forecasting. Recently, the WRF model has become one of the most popular and widely used tools for numeric weather prediction. In this paper, the WRF model is selected as a representative of the physical models. The main forecasting data are used to provide forecasting factors to NN model.

WRF is a fully compressible, nonhydrostatic model with a large number of physics options regarding cumulus parameterization, cloud microphysics, radiation, PBL parameterization, and land-surface model. In the WRF model, a grid is defined as an integration of three-dimensional points.



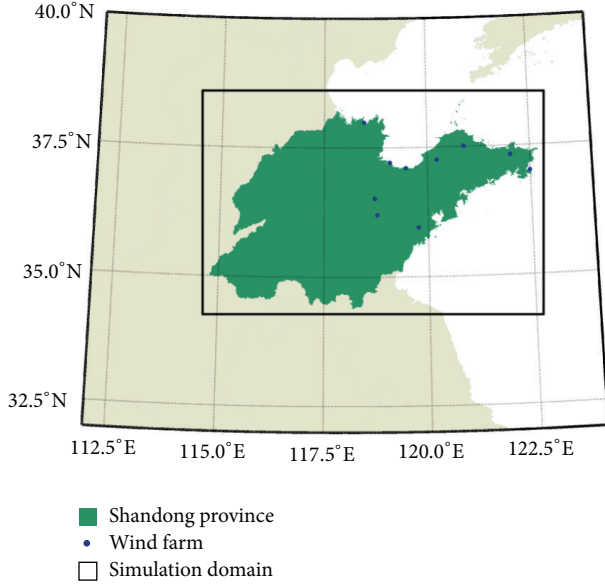


FIGURE 2: The simulation domains in WRF.

TABLE 1: Model configuration of the WRF simulation.

	Physical options
Cumulus parameterization	Grell 3D ensemble cumulus scheme
Short-wave radiation	RRTM scheme
Long-wave radiation	Dudhia scheme
Surface layer physics	Eta similarity
Land-surface processes	Fractional sea-ice
Planetary boundary layer	Mellor-Yamada-Janjic scheme

It contains a set of weather data (wind speed, atmospheric pressure, etc.). For each grid, there are a current time and an associated stop time. The atmospheric status is simulated by calculating a series of physical equations; this is based not only on the on-grid data but also on a specific physical model. Then, the current time of the grid can be advanced by a time-step, a unit of time [63]. The WRF simulation in this paper is performed for the case of April 2013 in Shandong province, China. The domains used in this simulation are shown in Figure 2.

The National Centers for Environmental Prediction Final Analysis ( $1^\circ \times 1^\circ$ , 6-hourly) data (NCEP FNL), which include 24 levels from surface to 10 hPa, are used as the initial and lateral boundary conditions. Details about the NCEP FNL are available at <http://rda.ucar.edu/datasets/ds083.2/>. The physics options selected in this simulation are also shown in Table 1.

### 3. SSA of Hourly Wind Speed of Wind Farms in Shandong Province

**3.1. SSA of a Time Series.** SSA is defined as a method to obtain detailed information from a noisy time series [64]. Consider the real-valued nonzero time series  $Y_T = (y_1, y_2, \dots, y_T)$  of sufficient length  $T$ . The primary reconstruction process of

SSA is to decompose the original series into a set of subseries  $X = X_1 + \dots + X_d$  and then reconstruct the original series [57, 58]. The SSA technique consists of two complementary stages: decomposition and reconstruction; both stages include two separate steps.

**3.2. Decomposition.** This stage is subdivided into two steps: embedding and singular value decomposition (SVD).

**3.2.1. Embedding.** Embedding can be regarded as a mapping that converts a one-dimensional time series  $Y_T = (y_1, y_2, \dots, y_T)$  into the multidimensional series  $X_1, X_2, \dots, X_K$  with vectors  $X_i = (y_i, y_{i+1}, \dots, y_{i+L-1})' \in \mathbf{R}^L$ , where  $K = T - L + 1$ . Vectors  $X_i$  are called  $L$ -lagged vectors, where  $L$  is an integer such that  $2 \leq L \leq T$ . The main result of this step is the definition of a trajectory matrix, defined as

$$X = [X_1, X_2, \dots, X_K] = \begin{pmatrix} y_1 & y_2 & \cdots & y_K \\ y_2 & y_3 & \cdots & y_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & y_T \end{pmatrix}. \quad (1)$$

The trajectory matrix is a Hankel matrix, where all of the elements along the diagonal  $i + j = \text{const}$  are equal [57].

**3.2.2. Singular Value Decomposition (SVD).** The SVD of  $XX'$  produces a set of  $L$  eigenvalues  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$  and the corresponding eigenvectors  $U_1, U_2, \dots, U_L$  (often denoted by empirical orthogonal functions (EOF)). Then, the SVD of the trajectory matrix can be written as  $X = E_1 + E_2 + \dots + E_d$ , where  $E_i = \sqrt{\lambda_i} U_i V_i'$ ,  $d$  is the rank of  $XX'$  (i.e., the number of nonzero eigenvalues), and  $V_1, V_2, \dots, V_d$  are the principal components (PC), defined as  $V_i = X' U_i / \sqrt{\lambda_i}$ . The collection  $(\sqrt{\lambda_i}, U_i, V_i)$  is referred to as the  $i$ th Eigen triple of the matrix  $X$ . If  $\Gamma = \sum_{i=1}^d \lambda_i$ , then  $\lambda_i / \Gamma$  is the proportion of variance of  $X$  explained by  $E_i : E_1$  which has the highest contribution [49], while  $E_d$  has the lowest contribution. SVD could be time-consuming if the length of the time series is large ( $T > 1000$ ).

**3.3. Reconstruction.** After decomposing the time series, the results include subseries  $X = X_1 + \dots + X_d$ . As in the decomposition stage, the reconstruction stage consists of two steps: grouping and averaging.

**3.3.1. Grouping.** In this step,  $r$  out of  $d$  Eigen triples are selected by the user. Let  $I = \{i_1, i_2, \dots, i_r\}$  be a group of  $r$ -selected Eigen triples  $X_i = X_{i_1} + X_{i_2} + \dots + X_{i_r}$ , where  $X_i$  is related to the "signal" of  $y$ , while the rest of the  $(d - r)$  Eigen triples denote the error term  $\varepsilon$ .

**3.3.2. Averaging.** The group of  $r$  components selected in the previous stage is then used to reconstruct the deterministic components of the time series. The basic idea is to transform each of the terms  $X_{i_1}, X_{i_2}, \dots, X_{i_r}$  into reconstructed time series  $y_{i_1}, y_{i_2}, \dots, y_{i_r}$  through the Hankelization process  $H()$  or diagonal averaging: if  $z_{ij}$ , then  $i + j = k + 1$ . Thus,  $H(Z)$  is a time series of length  $T$  reconstructed from the matrix  $Z$ .

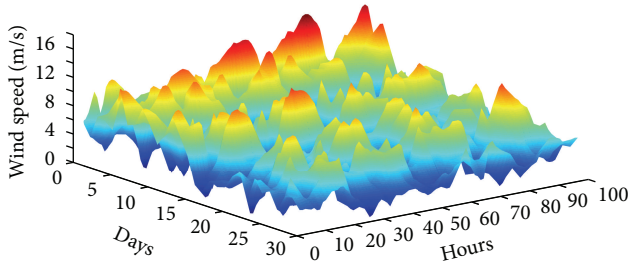


FIGURE 3: Wind speed diagram of wind farm A in April.

At the end of the averaging step, the reconstructed time series is an approximation of  $y$ :

$$y = H(X_{i_1}) + H(X_{i_2}) + \cdots + H(X_{i_r}) + e. \quad (2)$$

As noted by Alexdradov and Golyandina, the reconstruction of a single Eigen triple is based on the whole time series. This means that SSA is not a local method and, hence, is robust to outliers.

**3.4. Wind Speed and Its Properties.** The experimental data are the wind speed time series of Shandong province in m/s for 30 days from April 1 to April 30, 2013. For each series, we take 96 time series of different hourly wind speeds as a forecasting unit during forecasting processing. An example of one such wind farm's wind speed is given in Figure 3.

Within a month of each day, the wind speed time series has an inconspicuous periodicity of 24 hours. The general rule is to select  $L = T/2$ . However, in this case, the time series of the forecasting period could define  $L$ , and that will be more appropriate. In addition to the  $L$  parameter for SSA, the number of components  $r$  defines how the components can be separated. To locate the parameter  $r$ , the detailed properties of the wind speed time series should be displayed. A method for selecting  $r$  out of  $d$  components is requiring that the sum of their contributions be at least a predefined threshold, such as 90%. Generally, noise components will have a low contribution.

To evaluate the contributions of the different components, Figure 4 shows the principal components of wind speed in the SSA. Each PCx represents the ordinal components separated from the original time series. Note that PC1 and PC2 are dominating in all components, while PC3 and PC4 are subordinate. Components PC5–PC7 and PC8–PC11 could be classified as small parts which could be grouped. However, their contributions are still substantial. The rest of components diminish very slowly, but their contributions are low. The construction of the time series will be based on these Eigen triples.

**3.5. Reconstruction and Preparation for Forecasting Approaches.** As previously mentioned, the window length  $L$  is decided using a characteristic forecasting time series in the decomposition stage. Therefore,  $L = 96$  h is assumed,

here, which corresponds to a time series of 4 days of wind speed. The reconstruction of a single Eigen triple is based on the entire time series. In this case, a useful Eigen triple is defined as one that improves the accuracy of the final forecast. Figure 5 depicts the contributions of different Eigen triples in forecasting. The “positive” means that the contribution of these components is positive, which reduces the MAPE and vice versa.

The useful Eigen triples set is the group of  $r$  components selected in the previous stage; they are deterministic components of the time series. Because the simulated data consist of the wind speed, temperature, humidity, and pressure, they are reconstructed in the same way. The detailed process is shown in Figure 6. First, the wind speed and the other weather data are rearranged as an  $L$ -length time series, using the  $L$  determined in last section. Second, according to the calculated accuracy for each component, the principal components are divided into a positive components set and a negative components set. The positive component set includes the trend and the harmonic components, and the negative set mostly contains the noise. Finally, the reconstructed time series is provided as an  $r \times 4$  matrix.

By selecting a set of useful components and considering other negative components as noise, some frequencies may be filtered out completely. In Figure 6, original predictions represent the forecasting result without SSA, which includes the noise in simulation processing. In other words, this part could be regarded as a forecast including all of the Eigen triples.

#### 4. Performance Metrics of Forecast Accuracy

To date, a number of performance measures have been proposed and employed to evaluate the forecast accuracy, but no single performance measure has been recognized as the universal standard. This actually complicates the performance comparison of different forecasting models. As a result, we need to assess the performance using multiple metrics, and it is interesting to see if different metrics will give the same performance ranking for the tested models. The metrics included in this study are mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) [21]:

$$\begin{aligned} \text{MAE} &= \frac{1}{T} \sum_{t=1}^T |y_t - f_t|, \\ \text{RMSE} &= \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - f_t)^2}, \\ \text{MAPE} &= \frac{1}{T} \sum_{t=1}^T \left| \frac{y_t - f_t}{y_t} \right|, \end{aligned} \quad (3)$$

where  $y_t$  and  $f_t$  denote the observations and the forecast value from model  $t$ , respectively, and  $T$  is the number of data used for performance evaluation and comparison.

MAE measures the average magnitude of the errors of the forecasting sets. More specifically, these involve the average of

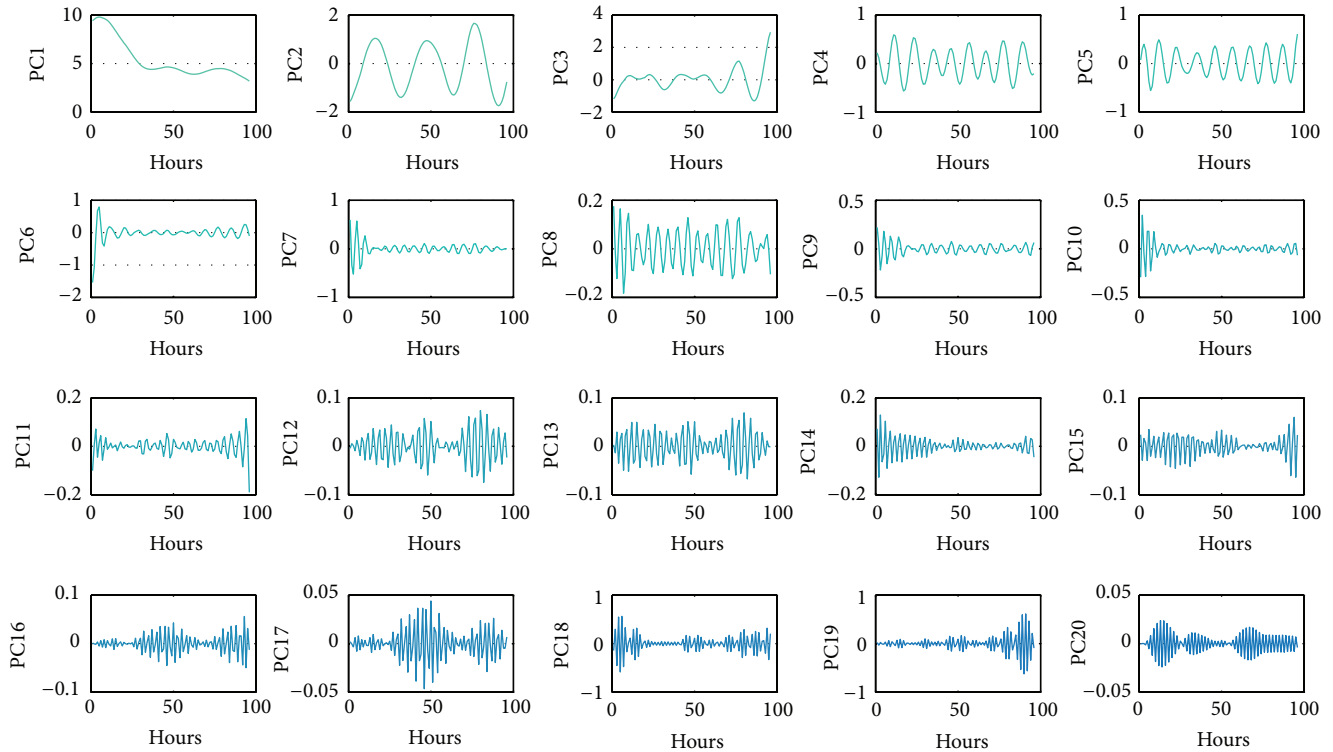


FIGURE 4: Principal components related to the 20 Eigen triples.

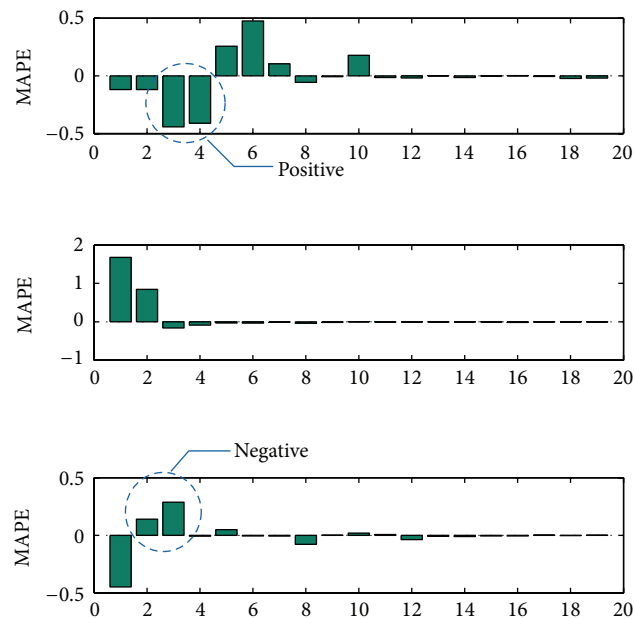


FIGURE 5: The contributions of different Eigen triples to forecasting accuracy.

the verification sample and the absolute values of the differences between the forecasted results and the corresponding observations. MAE is a linear measure, which means that all of the individual differences are equally weighted in the average. In contrast, RMSE is a quadratic scoring rule that measures the average magnitude of the error. Because the

errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors. This means that the RMSE is most useful when large errors are particularly undesirable. MAPE is a measure of accuracy in a fitted time series value in statistics, specifically, a trending value. The difference between the actual value and the forecasted value

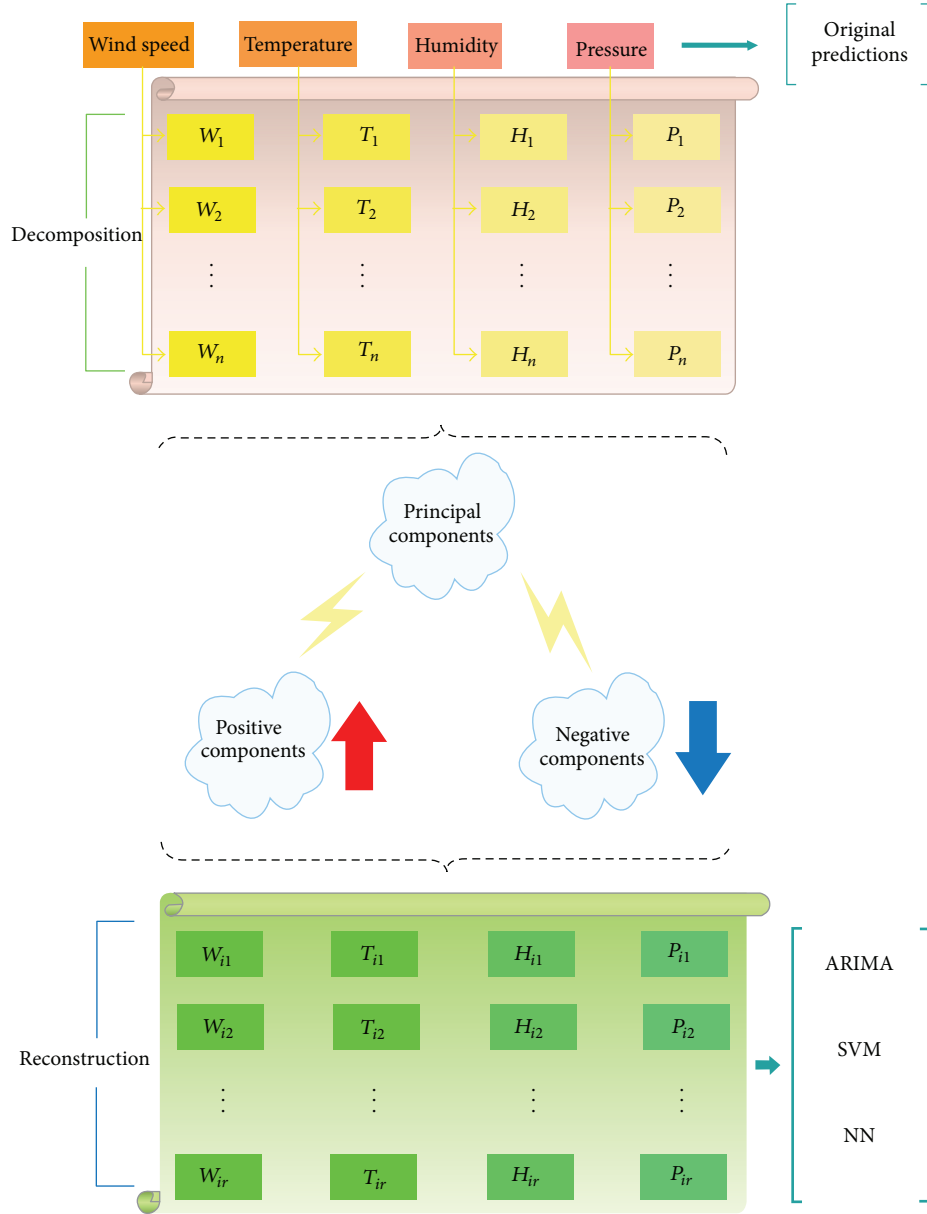


FIGURE 6: Decomposition and reconstruction diagram of simulation data.

is divided by the actual value. The absolute value of this calculation is summed for every forecast point in time and divided again by the total number of forecast points.

## 5. Forecasting Models

The forecasting model is used to measure the performance of this hybrid algorithm. It consists of the following algorithms.

(1) *Autoregressive Integrated Moving Average (ARIMA) Algorithm.* Introduced by Box and Jenkins [65], the ARIMA linear models have dominated many areas as a popular time series forecasting approach. As the application of these models is very common, a brief description is provided here. The linear function is based upon three parametric

linear components: autoregression (AR), integration (I), and moving average (MA). The autoregressive or ARIMA  $(p, 0, 0)$  model is represented as follows:

$$y_t = \theta_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + e_t, \quad (4)$$

where  $p$  is the number of the autoregressive terms,  $y_t$  is the forecasted output,  $y_{t-p}$  is the observation at time  $t - p$ , and  $\varphi_1, \varphi_2, \dots, \varphi_p$  are a finite set of parameters. The  $\varphi$  terms are determined by linear regression. The  $\theta_0$  term is the intercept, and  $e_t$  is the error associated with the regression. This time series depends only on the  $p$  past values of itself and a random term  $e_t$ . The moving average or ARIMA  $(0, 0, q)$  method is represented as

$$y_t = \mu - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} + e_t, \quad (5)$$



where  $q$  is the number of moving average terms,  $\theta_1, \theta_2, \dots, \theta_q$  are the finite weights or parameters set, and  $\mu$  is the mean of the series. This time series depends only on  $q$  past random terms and a present random term  $e_t$ . As a particular case, an ARIMA( $p, 0, q$ ) or ARMA( $p, q$ ) is a model for a time series that depends on  $p$  past values of itself and on  $q$  past random terms  $e_t$ . The equation is shown as follows:

$$y_t = \theta_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \mu - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q} + e_t. \quad (6)$$

Finally, an ARIMA( $p, d, q$ ) is an ARIMA( $p, 0, q$ ) model for a time series that has been different  $d$  times. The ARIMA models have the capability to include external independent or predictor variables.

(2) *Support Vector Machine (SVM) Algorithm.* The support vector machine (SVM) was proposed by Vapnik [66]. Based on the structured risk minimization (SRM) principle, the SVM minimizes an upper bound of the generalization error instead of the empirical error, as in other neural networks. Additionally, the SVM model generates the regression function by applying a set of high-dimensional linear functions. The SVM regression function is formulated as follows:

$$y = w\varphi(x) + b, \quad (7)$$

where  $\varphi(x)$  is called the feature, which is nonlinearly mapped from the input space  $x$ . The coefficients  $w$  and  $b$  are estimated by minimizing

$$R(C) = C \frac{1}{N} \sum_{i=1}^N L_\varepsilon(d_i, y_i) + \frac{1}{2} \|w\|^2 [0, 1], \quad (8)$$

$$L_\varepsilon(d, y) = \begin{cases} |d - y| - \varepsilon, & |d - y| \geq \varepsilon, \\ 0, & \text{others,} \end{cases}$$

where both  $C$  and  $\varepsilon$  are prescribed parameters. The first term  $L_\varepsilon(d, y)$  is called the  $\varepsilon$ -intensive loss function. The  $d_i$  is the actual wind speed in the  $i$ th period. This function indicates that errors below  $\varepsilon$  are not penalized. The term  $C(1/N) \sum_{i=1}^N L_\varepsilon(d_i, y_i)$  is the empirical error. The second term,  $(1/2)\|w\|^2$ , measures the flatness of the function.  $C$  evaluates the model. The positive slack variables  $\zeta$  and  $\zeta^*$  represent the distance from the actual values to the corresponding boundary values of the  $\varepsilon$ -tube. The equation is transformed into the following constrained formation:

$$R(w, \zeta, \zeta^*) = \frac{1}{2} ww^T + C^* \left( \sum_{i=1}^N (\zeta_i + \zeta_i^*) \right), \quad (9)$$

subject to

$$\begin{aligned} w\varphi(x_i) + b_i - d_i &\leq \varepsilon + \zeta_i^*, \\ d_i - w\varphi(x_i) - b_i &\leq \varepsilon + \zeta_i, \\ \zeta_i, \zeta_i^* &\geq 0, \quad i = 1, 2, K, N. \end{aligned} \quad (10)$$

Finally, they satisfy the equality

$$f(x, \alpha, \alpha^*) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x, x_i) + b. \quad (11)$$

Here,  $K(x, x_i)$  is called the kernel function. The value of the kernel is equal to the inner product of two vectors  $x_i$  and  $x_j$  in the feature space  $\varphi(x_i)$  and  $\varphi(x_j)$ , such that  $K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j)$ . Any function that satisfies Mercer's condition can be used as the kernel function [67]. The Gaussian kernel function is used in this paper.

(3) *Artificial Neural Networks (ANN) Algorithm.* Artificial neural network (ANN) models are generally used for time series forecasting. A neural network is a mathematical representation that is inspired by the way the brain processes information. The model consists of an input layer, an output layer, and one or more intervening layers, also referred to as hidden layers. The hidden layers can capture the nonlinear relationship between variables. Each layer consists of multiple neurons that are connected to the neurons in adjacent layers. Because these networks contain many interacting nonlinear neurons in multiple layers, the networks can capture relatively complex performance. ANN is already one of the models that is able to approximate various nonlinearities in the data series. Many types of ANN models have been suggested in related research, with the most popular one used for classification being the multilayer perceptron (MLP) with back propagation. The output of the  $i$ th hidden neuron is then computed by processing the weighted inputs and its bias term  $h_i$  as follows:

$$h_i = f \left( b_i + \sum_{j=1}^n w_{ij} x_j \right), \quad (12)$$

where  $w_{ij}$  denotes the weight connecting input  $w_{ij}$  to hidden unit  $h_i$ .

Similarly, the output of the output layer is computed as follows:

$$y = f^{\text{output}} \left( b + \sum_{j=1}^n w_j x_j \right), \quad (13)$$

with  $n$  presenting the number of hidden neurons and  $w_j$  representing the weight connecting hidden unit  $j$  to the output neuron. A threshold function is then applied to map the network output  $y$  to a classification label. The transfer functions  $f^h$  and  $f^{\text{output}}$  allow the network to model nonlinear relationships in the data.

## 6. Forecast Results and Comparative Analysis

The last section consists of forecasting data calculation, error comparison, and results analysis. The error calculation module provides different methods to compare the SSA technique for different forecasting methods in the 10 wind farms. The algorithms are chosen based on the different theoretical principles.

TABLE 2: Comparison of RMSE for 10 wind farms.

	Wind farm A	Wind farm B	Wind farm C	Wind farm D	Wind farm E	Wind farm F	Wind farm G	Wind farm H	Wind farm I	Wind farm J
ARIMA										
Ori-RMSE	1.38	3.54	7.52	3.27	3.00	5.98	3.81	1.34	1.61	9.75
SSA-RMSE	1.34	3.49	1.33	3.04	2.28	5.59	4.81	1.57	1.43	8.96
SVM										
Ori-RMSE	5.01	6.85	4.27	7.11	8.42	7.19	7.92	6.89	5.43	4.95
SSA-RMSE	4.43	5.91	3.66	6.38	7.54	6.85	7.58	6.51	5.73	4.52
ANN										
Ori-RMSE	1.45	4.34	1.88	6.17	1.22	3.63	3.10	2.53	1.64	8.98
SSA-RMSE	1.31	3.98	1.37	2.92	0.64	6.76	5.14	1.63	0.97	9.36

TABLE 3: Comparison of MAE for 10 wind farms.

	Wind farm A	Wind farm B	Wind farm C	Wind farm D	Wind farm E	Wind farm F	Wind farm G	Wind farm H	Wind farm I	Wind farm J
ARIMA										
Ori-MAE	1.01	3.25	7.13	2.99	2.63	5.84	3.53	0.98	1.11	9.67
SSA-MAE	0.94	3.12	0.97	2.78	1.45	5.44	4.44	1.03	0.95	8.88
SVM										
Ori-MAE	4.82	6.76	4.14	7.04	8.36	7.07	7.81	6.70	5.24	4.82
SSA-MAE	4.31	5.89	3.59	6.29	7.49	6.76	7.45	6.41	5.61	4.39
ANN										
Ori-MAE	0.98	4.14	1.76	5.86	1.05	3.41	2.77	2.37	1.37	8.90
SSA-MAE	0.94	3.58	1.19	2.74	0.50	6.56	4.61	1.04	0.58	9.26

TABLE 4: Comparison of MAPE for 10 wind farms.

	Wind farm A	Wind farm B	Wind farm C	Wind farm D	Wind farm E	Wind farm F	Wind farm G	Wind farm H	Wind farm I	Wind farm J
ARIMA										
Ori-MAPE	7.68%	24.33%	50.62%	20.38%	19.91%	40.81%	26.38%	7.46%	8.85%	68.29%
SSA-MAPE	7.28%	23.50%	7.72%	18.97%	11.64%	37.90%	33.10%	8.15%	7.54%	62.55%
SVM										
Ori-MAPE	33.76%	47.49%	28.85%	49.74%	59.13%	49.86%	55.17%	47.22%	36.38%	33.70%
SSA-MAPE	30.16%	41.81%	25.23%	44.34%	52.98%	47.97%	52.47%	45.20%	39.22%	30.68%
ANN										
Ori-MAPE	7.77%	30.61%	12.44%	40.88%	7.64%	23.66%	20.92%	16.32%	10.13%	62.77%
SSA-MAPE	7.26%	26.83%	8.46%	18.85%	3.90%	45.67%	34.56%	8.30%	4.63%	65.21%

In the following, the combined methodology is applied to predict the future wind speed of the furthestmost hours to each forecasting period; the results are shown in Tables 2, 3, and 4. Specifically, the prediction performance is evaluated using the RMSE, MAPE, and MSE. For these forecasting cases in different wind farms, the only actual requirement from the wind farms is reducing the prediction error as much as possible. In our recent research into wind speed forecasting, the time series method relied on the historical data, which are primarily used in immediate-short-term forecasting. The forecast period is usually limited to 4 hours, divided into quarter-hour periods. To expand the forecast

period and observe the forecasting performance using the SSA technique, the wind data in this paper are merged into an hour, as we mentioned earlier, and the forecast period is extended from 4 hours to 7 hours. The schematic diagram of the forecasting process is shown in Figure 7.

As seen in Table 2, the forecasting performance values using the SSA technique based on the ARIMA and the NN are much better than those of the SVM algorithm for most of the wind farms. For instance, the values of RMSE for the ARIMA and ANN at wind farm C are 1.33 and 1.37, less than the value of 3.66 obtained using SVM. As we mentioned, whether the SSA technique reduces forecasting error is the

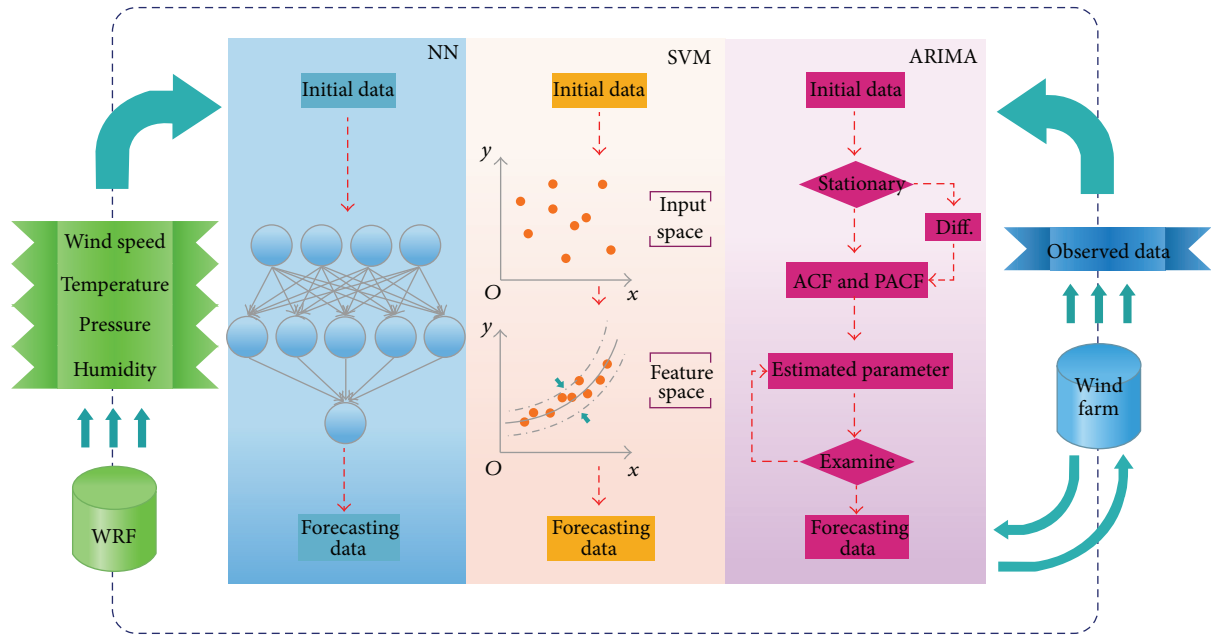


FIGURE 7: Procedure chart of the forecasting process.

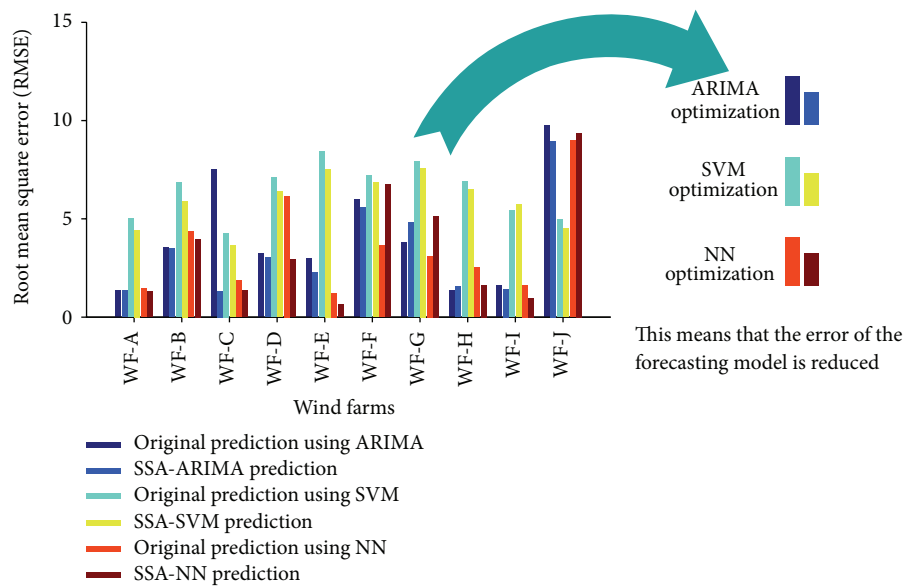


FIGURE 8: Final error reduction of wind farms in Shandong province.

critical problem addressed in this paper. So, comparing the Ori-group with the SSA-group, it is apparent that the error of the SSA-group is lower as measured by the RMSE. For example, the value of SSA-RMSE in wind farm E using ANN decreases from 1.22 to 0.64. The value of SSA-RMSE in wind farm C using ARIMA decreases from 7.52 to 1.33. However, this is not the case across all 10 wind farms. In cases such as wind farm F in ANN, wind farm G in ARIMA, and wind farm I in SVM, the RMSEs are relatively high. However, the number of these groups is small, and there is no site where the RMSE of all three methods increases. This conclusion

is presented in Table 2, and a detailed explanation will be provided next.

Similar to Tables 2 and 3 displays the values for MAE. MAE is expressed as the absolute value of the error, which can also be used to evaluate forecasting errors. The MAE exhibits the same pattern as the RMSE. The MAE of the Ori-group is lower than the error of the SSA-group, which means that the SSA technique significantly improved the prediction accuracy. As observed from the table, the SSA-group is different from the Ori-group, which has a substantially higher forecasting accuracy. However, the forecasting

errors associated with different methods make a substantial difference.

Details of the MAPE of forecasting results are given in Table 4; they are similar to the results from Tables 2 and 3. The results reported in Table 4 are the MAPE results obtained using three different methods. As we can see from the table, the SSA-group has higher forecasting accuracy than the Ori-group. The MAPE results in Table 4 show that the combined method improves accuracy from 0.4% to 42%. For example, the ARIMA forecasting of wind farm C reduces the error from 50.62% to 7.72%, whereas the ANN forecasting of wind farm D reduces the error from 40.88% to 18.85%. Most results indicate that the SSA technique significantly improves the prediction accuracy.

The RMSE is widely used in wind farms and is the primary error measure used in this paper. The RMSE results are displayed in Figure 8. These results suggest that the SSA technique is an excellent method for time series forecasting.

## 7. Discussion and Conclusion

This paper provided a new wind speed forecasting method by introducing an SSA algorithm. In wind speed forecasting cases, we proposed an effective method for defining the parameters of SSA and applied the new method to 10 wind farms in Shandong province. Here, SSA-based filtering consists of a data-separation technology with fewer parameters, which is more efficient than the traditional noise reduction methods used in recent research. The proposed algorithm significantly outperforms the basic forecasting algorithm in a variety of different situations using ARIMA, SVM, and NN; this conclusion is especially true for forecasting with a time series algorithm. Secondly, the method employs SSA to eliminate the noise series with an algorithm based on the data itself, which overcomes the limitations imposed by the complexity of the algorithms. Further, the goal of the proposed model is not only to present an exact representation of the forecasting method itself but also to set up a series of methods for the process of decomposition and reconstruction that will be generally capable of receiving new inputs.

The interest in employing three different forecasting methods based on different theoretical structures confirmed that the accuracy and applicability of the forecast result are improved; however, the forecasting capacities of the methods themselves were significantly different. On the one hand, the forecasting horizon was expanded from 4 hours to 7 hours in these immediate-short-term winds speed forecasts. This method, which consists of the decomposition and reconstruction of SSA, will result in a better evaluation of forecasting method performance in a time series. On the other hand, there is not a universal method for wind speed forecasting; different methods can be applied under different conditions. Although the behavior of SVM is unsatisfactory, the results of the other two forecasting methods are still suitable. In fact, this result is exactly what we expect in wind speed forecasting. Because of the fluctuating nature of the wind series, it is difficult to find an efficient and versatile optimization method. Even to this day, wind speed

forecasting remains a very laborious problem, and the MAPE of such wind farms usually ranges from 25% to 40% [68]. Although the versatility of this approach remains to be tested, this experiment, which included 10 wind farms, demonstrates that the method is useful for wind speed forecasting.

## Conflict of Interests

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

## Acknowledgments

This research was supported by the National Natural Science Foundation of China (41225018) and IAM (IAM201305).

## References

- [1] "IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation," 2011, <http://www.ipcc-wg3.de/publications/special-reports>.
- [2] K. Y. Oh, J. Y. Kim, J. K. Lee, M. S. Ryu, and J. S. Lee, "An assessment of wind energy potential at the demonstration offshore wind farm in Korea," *Energy*, vol. 46, pp. 555–563, 2012.
- [3] N. Chaudhry and L. Hughes, "Forecasting the reliability of wind-energy systems: a new approach using the RL technique," *Applied Energy*, vol. 96, pp. 422–430, 2012.
- [4] GWEC, "Global Wind Energy Council," 2010, <http://www.gwec.net/fileadmin/documents/Publications/GWEO%202010%20final.pdf>.
- [5] 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.
- [6] M. Carolin Mabel and E. Fernandez, "Analysis of wind power generation and prediction using ANN: a case study," *Renewable Energy*, vol. 33, no. 5, pp. 986–992, 2008.
- [7] S. Rehman, A. M. Mahbub Alam, J. P. Meyer, and L. M. Al-Hadhrami, "Wind speed characteristics and resource assessment using weibull parameters," *International Journal of Green Energy*, vol. 9, no. 8, pp. 800–814, 2012.
- [8] 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.
- [9] A. Öztöpal, "Artificial neural network approach to spatial estimation of wind velocity data," *Energy Conversion and Management*, vol. 47, no. 4, pp. 395–406, 2006.
- [10] X. Li, K. Hubacek, and Y. L. Siu, "Wind power in China—dream or reality?" *Energy*, vol. 37, no. 1, pp. 51–60, 2012.
- [11] L. Lazić, G. Pejanović, and M. Živković, "Wind forecasts for wind power generation using the Eta model," *Renewable Energy*, vol. 35, no. 6, pp. 1236–1243, 2010.
- [12] 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.
- [13] Smart grid, "Wikipedia," 2012, [http://en.wikipedia.org/wiki/Smart\\_grid](http://en.wikipedia.org/wiki/Smart_grid).
- [14] W. Wang, Y. Xu, and M. Khanna, "A survey on the communication architectures in smart grid," *Computer Networks*, vol. 55, no. 15, pp. 3604–3629, 2011.



- [15] Y. N. Wu, J. Chen, and L. R. Liu, "Construction of China's smart grid information system analysis," *Renewable & Sustainable Energy Reviews*, vol. 15, pp. 4236–4241, 2011.
- [16] J. Gao, Y. Xiao, J. Liu, W. Liang, and C. L. P. Chen, "A survey of communication/networking in Smart Grids," *Future Generation Computer Systems*, vol. 28, no. 2, pp. 391–404, 2012.
- [17] M. G. de Giorgi, A. Ficarella, and M. Tarantino, "Assessment of the benefits of numerical weather predictions in wind power forecasting based on statistical methods," *Energy*, vol. 36, no. 7, pp. 3968–3978, 2011.
- [18] 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.
- [19] I. J. Ramírez-Rosado and L. A. Fernández-Jiménez, "An advanced model for short-term forecasting of mean wind speed and wind electric power," *Control and Intelligent Systems*, vol. 32, no. 1, pp. 21–26, 2004.
- [20] A. Sfetos, "A novel approach for the forecasting of mean hourly wind speed time series," *Renewable Energy*, vol. 27, no. 2, pp. 163–174, 2002.
- [21] 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.
- [22] H. Liu, J. Shi, and E. Erdem, "Prediction of wind speed time series using modified Taylor Kriging method," *Energy*, vol. 35, no. 12, pp. 4870–4879, 2010.
- [23] K. Chen and J. Yu, "Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach," *Applied Energy*, vol. 113, pp. 690–705, 2014.
- [24] 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.
- [25] 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, November 2006.
- [26] F. Cassola and M. Burlando, "Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output," *Applied Energy*, vol. 99, pp. 154–166, 2012.
- [27] 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.
- [28] S. J. Wu and S. L. Lin, "Intelligent web-based fuzzy and grey models for hourly wind speed forecast," *International Journal of Computers*, vol. 4, pp. 235–242, 2010.
- [29] J.-F. Li, B.-H. Zhang, G.-L. Xie, Y. Li, and C.-X. Mao, "Grey predictor models for wind speed-wind power prediction," *Power System Protection and Control*, vol. 38, no. 19, pp. 151–159, 2010.
- [30] L. Lin, J. T. Eriksson, H. Vihriala, and L. Soderlund, "Predicting wind behavior with neural networks," in *Proceedings the European Wind Energy Conference*, pp. 655–658.
- [31] 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.
- [32] H. G. Beyer, T. Degner, J. Haussmann, M. Hoffman, and P. Rujan, "Short term forecast of wind speed and power output of a wind turbine with neural networks," in *Proceedings of the 2nd European Congress on Intelligent Techniques and Soft Computing*, pp. 349–352, 1994.
- [33] G. N. Kariniotakis, G. S. Stavrakakis, and E. F. Nogaret, "Wind power forecasting using advanced neural networks models," *IEEE Transactions on Energy Conversion*, vol. 11, no. 4, pp. 762–767, 1996.
- [34] A. More and M. C. Deo, "Forecasting wind with neural networks," *Marine Structures*, vol. 16, no. 1, pp. 35–49, 2003.
- [35] G. Kariniotakis, G. S. Stavrakakis, and E. F. Nogaret, "A fuzzy logic and neural network based wind power model," in *Proceedings of the European Wind Energy Conference*, pp. 596–599, 1996.
- [36] X. Wang, G. Sideratos, N. Hatzigargyriou, and L. H. Tsoukalas, "Wind speed forecasting for power system operational planning," in *Proceedings of the International Conference on Probabilistic Methods Applied to Power Systems*, pp. 470–474, September 2004.
- [37] H. Mori and A. Yuihara, "Deterministic annealing clustering for ANN-based short-term load forecasting," *IEEE Transactions on Power Systems*, vol. 16, no. 3, pp. 545–551, 2001.
- [38] 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.
- [39] X. Zhang, W. Yan, and Z. Bao, "Short-term load forecasting of power systems by combination of wavelet transform and AMPSO based neural network," *Energy Procedia*, vol. 13, pp. 6006–6016, 2011.
- [40] A. Badri, Z. Ameli, and A. Birjandi, "Application of artificial neural networks and fuzzy logic methods for short termload forecasting," *Energy Procedia*, vol. 14, pp. 1883–1888, 2012.
- [41] D.-X. Niu, H.-F. Shi, and D. D. Wu, "Short-term load forecasting using bayesian neural networks learned by Hybrid Monte Carlo algorithm," *Applied Soft Computing Journal*, vol. 12, no. 6, pp. 1822–1827, 2012.
- [42] H. Y. Yamin, Q. El-Dwairi, and S. M. Shahidehpour, "A new approach for GenCos profit based unit commitment in day-ahead competitive electricity markets considering reserve uncertainty," *International Journal of Electrical Power & Energy Systems*, vol. 29, no. 8, pp. 609–616, 2007.
- [43] G. Mokryani and P. Siano, "Evaluating the integration of wind power into distribution networks by using monte carlo simulation," *International Journal of Electrical Power & Energy Systems*, vol. 53, pp. 244–255, 2013.
- [44] G. Mokryani and P. Siano, "Combined monte carlo simulation and OPF for wind turbines integration into distribution networks," *Electric Power Systems Research*, vol. 103, pp. 37–48, 2013.
- [45] G. Mokryani and P. Siano, "Optimal wind turbines placement within a distribution market environment," *Applied Soft Computing*, vol. 13, no. 10, pp. 4038–4046, 2013.
- [46] R. Vautard and M. Ghil, "Singular spectrum analysis in non-linear dynamics, with applications to paleoclimatic time series," *Physica D: Nonlinear Phenomena*, vol. 35, no. 3, pp. 395–424, 1989.
- [47] R. Vautard, P. Yiou, and M. Ghil, "Singular-spectrum analysis: a toolkit for short, noisy chaotic signals," *Physica D: Nonlinear Phenomena*, vol. 58, no. 1–4, pp. 95–126, 1992.
- [48] M. de Carvalho, P. C. Rodrigues, and A. Rua, "Tracking the US business cycle with a singular spectrum analysis," *Economics Letters*, vol. 114, no. 1, pp. 32–35, 2012.



- [49] H. Hassani, S. Heravi, and A. Zhigljavsky, "Forecasting European industrial production with singular spectrum analysis," *International Journal of Forecasting*, vol. 25, no. 1, pp. 103–118, 2009.
- [50] A. D. Papalexopoulos and T. C. Hesterberg, "A regression-based approach to short-term system load forecasting," *IEEE Transactions on Power Systems*, vol. 5, no. 4, pp. 1535–1547, 1990.
- [51] W. R. Christiaanse, "Short term load forecasting using general exponential smoothing," *IEEE Transactions on Power Apparatus Systems*, vol. 90, no. 2, pp. 900–911, 1971.
- [52] S.-J. Huang and K.-R. Shih, "Short-term load forecasting via ARMA model identification including non-Gaussian process considerations," *IEEE Transactions on Power Systems*, vol. 18, no. 2, pp. 673–679, 2003.
- [53] H. M. Al-Hamadi and S. A. Soliman, "Short-term electric load forecasting based on Kalman filtering algorithm with moving window weather and load model," *Electric Power Systems Research*, vol. 68, no. 1, pp. 47–59, 2004.
- [54] 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.
- [55] Y. Dong, J. Wang, H. Jiang, and J. Wu, "Short-term electricity price forecast based on the improved hybrid model," *Energy Conversion and Management*, vol. 52, no. 8-9, pp. 2987–2995, 2011.
- [56] A.-O. Boudraa and J.-C. Cexus, "EMD-based signal filtering," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, no. 6, pp. 2196–2202, 2007.
- [57] H. Hassani, "Singular spectrum analysis: methodology and comparison," *Journal of Data Science*, vol. 5, pp. 239–257, 2007.
- [58] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of Time Series Structure: SSA and Related Techniques*, Chapman & Hall/CRC, 2001.
- [59] N. Golyandina and A. Korobeynikov, "Basic singular spectrum analysis and forecasting with R," *Computational Statistics & Data Analysis*, vol. 71, pp. 934–954, 2014.
- [60] M. Claudio and S. Rocco, "Singular spectrum analysis and forecasting of failure time series," *Reliability Engineering & System Safety*, vol. 114, pp. 126–136, 2013.
- [61] G. Tzagkarakis, M. Papadopoulou, and P. Tsakalides, "Trend forecasting based on Singular Spectrum Analysis of traffic workload in a large-scale wireless LAN," *Performance Evaluation*, vol. 66, no. 3-5, pp. 173–190, 2009.
- [62] C. A. F. Marques, J. A. Ferreira, A. Rocha et al., "Singular spectrum analysis and forecasting of hydrological time series," *Physics and Chemistry of the Earth*, vol. 31, no. 18, pp. 1172–1179, 2006.
- [63] S. M. Sadjadi, *Weather Research and Forecasting Model 2. 2 Documentation: A Step-by-step guide of a Model Run*, 2007.
- [64] K. Afshar and N. Bigdeli, "Data analysis and short term load forecasting in Iran electricity market using singular spectral analysis (SSA)," *Energy*, vol. 36, no. 5, pp. 2620–2627, 2011.
- [65] G. E. P. Box and G. Jenkins, *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco, Calif, USA, 1970.
- [66] V. Vapnik, *The Nature of Statistic Learning Theory*, Springer, New York, NY, USA, 1995.
- [67] P.-F. Pai and C.-S. Lin, "A hybrid ARIMA and support vector machines model in stock price forecasting," *Omega*, vol. 33, no. 6, pp. 497–505, 2005.
- [68] Z.-H. Guo, J. Wu, H.-Y. Lu, and J.-Z. Wang, "A case study on a hybrid wind speed forecasting method using BP neural network," *Knowledge-Based Systems*, vol. 24, no. 7, pp. 1048–1056, 2011.