

# Wind Power Generation Forecasting Using Least Squares Support Vector Machine Combined with Ensemble Empirical Mode Decomposition, Principal Component Analysis and a Bat Algorithm

## **Authors:**

Qunli Wu, Chenyang Peng

*Date Submitted:* 2018-11-27

*Keywords:* grey relational analysis, bat algorithm (BA), principal component analysis (PCA), least squares support vector machine (LSSVM), ensemble empirical mode decomposition (EEMD)

## **Abstract:**

Regarding the non-stationary and stochastic nature of wind power, wind power generation forecasting plays an essential role in improving the stability and security of the power system when large-scale wind farms are integrated into the whole power grid. Accurate wind power forecasting can make an enormous contribution to the alleviation of the negative impacts on the power system. This study proposes a hybrid wind power generation forecasting model to enhance prediction performance. Ensemble empirical mode decomposition (EEMD) was applied to decompose the original wind power generation series into different sub-series with various frequencies. Principal component analysis (PCA) was employed to reduce the number of inputs without lowering the forecasting accuracy through identifying the variables deemed as significant that maintain most of the comprehensive variability present in the data set. A least squares support vector machine (LSSVM) model with the pertinent parameters being optimized by bat algorithm (BA) was established to forecast those sub-series extracted from EEMD. The forecasting performances of diverse models were compared, and the findings indicated that there was no accuracy loss when only PCA-selected inputs were utilized. Moreover, the simulation results and grey relational analysis reveal, overall, that the proposed model outperforms the other single or hybrid models.

*Record Type:* Published Article

*Submitted To:* LAPSE (Living Archive for Process Systems Engineering)

*Citation (overall record, always the latest version):*

LAPSE:2018.0993

*Citation (this specific file, latest version):*

LAPSE:2018.0993-1

*Citation (this specific file, this version):*

LAPSE:2018.0993-1v1

*DOI of Published Version:* <https://doi.org/10.3390/en9040261>

*License:* Creative Commons Attribution 4.0 International (CC BY 4.0)

Article

# Wind Power Generation Forecasting Using Least Squares Support Vector Machine Combined with Ensemble Empirical Mode Decomposition, Principal Component Analysis and a Bat Algorithm

Qunli Wu and Chenyang Peng \*

Department of Economics and Management, North China Electric Power University, Baoding 071003, China; wuqunli2002@aliyun.com

\* Correspondence: ncepupengchenyang@163.com; Tel.: +86-150-3128-0893

Academic Editor: José C. Riquelme

Received: 10 January 2016; Accepted: 29 March 2016; Published: 1 April 2016

**Abstract:** Regarding the non-stationary and stochastic nature of wind power, wind power generation forecasting plays an essential role in improving the stability and security of the power system when large-scale wind farms are integrated into the whole power grid. Accurate wind power forecasting can make an enormous contribution to the alleviation of the negative impacts on the power system. This study proposes a hybrid wind power generation forecasting model to enhance prediction performance. Ensemble empirical mode decomposition (EEMD) was applied to decompose the original wind power generation series into different sub-series with various frequencies. Principal component analysis (PCA) was employed to reduce the number of inputs without lowering the forecasting accuracy through identifying the variables deemed as significant that maintain most of the comprehensive variability present in the data set. A least squares support vector machine (LSSVM) model with the pertinent parameters being optimized by bat algorithm (BA) was established to forecast those sub-series extracted from EEMD. The forecasting performances of diverse models were compared, and the findings indicated that there was no accuracy loss when only PCA-selected inputs were utilized. Moreover, the simulation results and grey relational analysis reveal, overall, that the proposed model outperforms the other single or hybrid models.

**Keywords:** ensemble empirical mode decomposition (EEMD); least squares support vector machine (LSSVM); principal component analysis (PCA); bat algorithm (BA); grey relational analysis

## 1. Introduction

Wind power has been identified as one of the most important and efficient renewable energy and has been extensively utilized throughout the world [1–3]. With the rapid development of wind power, the proportion of wind power in the whole power system is becoming larger. However, wind power is a rolling source of electrical energy due to the variability of wind speed, temperature and other factors. The uncertainty of wind power undoubtedly affects the power system stability and increases the operation cost of power systems [2]. Therefore, accurate forecasting approaches with respect to wind power generation have positive implications on power system planning for unit commitment and dispatch, and electricity trading in certain electricity markets.

There is abundant literature on wind power forecasting, most of which has been published in recent years. In contrast to the wealth of studies on wind speed prediction, there has been less research looking at wind power generation forecasting. The approaches of these studies can be classified into three categories: time series models [4–8], artificial intelligent algorithm models [9–17] and time-series artificial intelligent algorithm models [18]. Most of these approaches utilize time series analysis models,

including vector autoregressive (VAR) models [4,5], autoregressive moving average (AMRA) models, and autoregressive integrated moving average (ARIMA) models. Erdem [6] decomposed wind speed into lateral and longitudinal components with each component being represented by an ARMA model, then the predictive value results were obtained by accumulation. Liu [7] proposed an autoregressive moving average-generalized autoregressive conditional heteroscedasticity algorithm for modeling the mean and volatility of wind speed, with the model effectiveness being evaluated by multiple methods. The results suggested the proposed method effectively captured the characteristics of wind speed. Kavasseri [8] examined the use of ARIMA model to forecast wind speed. The simulation results indicated the forecasting accuracy of the proposed method outperformed the persistence models. Nevertheless, the wide implementation of time series models on wind power prediction can be problematic due to the poor nonlinear fitting capacity.

On the contrary, the adaptive and self-organized learning features of intelligent algorithms apparently facilitate the estimation of nonlinear time series. For instance, artificial neural network (ANN) [9,10] and least squares support vector machine (LSSVM) [11–14] are perceived to be highly effective methods in the field of wind power forecasting. Guo [11] successfully developed a hybrid seasonal auto-regression integrated moving average and LSSVM model to forecast the mean monthly wind speed. De Giorgi [12] developed a comparative study for the prediction of the power production of a wind farm using historical data and numerical weather predictions. The findings demonstrated that the hybrid approach based on wavelet decomposition with LSSVM significantly outperformed the hybrid artificial neural network (ANN)-based methods. Yuan [13] established a LSSVM model in the light of gravitational search algorithm (GSA) for short-term output power prediction of a wind farm. Compared with the back propagation (BP) neural network and support vector machine (SVM) model, the simulation results indicated that the GSA-LSSVM model had higher accuracy for short-term output power prediction. Wang [14] decomposed the non-stationary time series into several intrinsic mode functions (IMFs) and the corresponding, residue, then each sub-series was forecasted using diverse LSSVM models.

With the burgeoning use of artificial intelligence technology, many researchers have devoted increasing effort and time to delving into least squares support machine approaches. Since the performance of the prediction model depends on the regularization parameter and the kernel parameter of the LSSVM models, considerable research has established LSSVM models based on different intelligent algorithms for wind power prediction to attain satisfactory results [15–17]. Hu [15] introduced a modified quantum particle swarm optimization (QPSO) algorithm to select the optimal parameters of LSSVM, and the results suggested that the generalization capability and learning performance of LSSVM model were apparently enhanced. Sun [16] established a LSSVM model optimized by particle swarm optimization (PSO). The simulation results recognized that the proposed method can distinctly increase the predicting accuracy. Wang [17] constructed a LSSVM model where the parameters were tuned by a PSO method based on simulated annealing (PSOSA). A case study from four wind farms in Gansu Province, Northwest China was applied to corroborate the effectiveness of the hybrid model. Cai [18] utilized a time series model to select the input variables and multi-layer back propagation neural network and generalized regression neural network are applied it to conduct forecasting.

However, it can be concluded from the previous research that the PSO algorithm seems to suffer from the local optimum problem during the regularization parameter selection process. In order to overcome the weakness of existing algorithms, a novel global algorithm, namely, the bat algorithm (BA) originally was proposed by Yang in 2010, based on the echolocation behavior of bats [19]. With a good combination of the paramount advantages of PSO and genetic algorithm (GA), the superiority of the BA results from its simplification, powerful searching ability and fast convergence. Recently, a burgeoning number of studies focusing on the BA for parameter optimization have appeared [20–22]. Hafezi [20] explored a hybrid solution based on a BA to predict stock prices over a long term period. The model was examined through forecasting eight years of deutscher aktienindex (DAX) stock prices

and conceived as an appropriate tool for predicting stock prices. Senthilkumar [21] selected the best set of features from the initial sets using a BA, perceived as a the recent optimization algorithm for reducing the time consumption in detecting record duplication. Yang [22] exploited an efficient multi-objective optimization method in accordance with the BA to suppress critical harmonics and determine power factors for passive power filters (PPFs). Considering the excellent capacity of the BA during the process of parameter optimization, it is the purpose of the current study to select the two pertinent parameters of the LSSVM model and obtain the global optimal strategy using the BA method.

From the previous literature, it can be seen that the original series tend to be regarded as the independent variables pertaining to wind power forecasting. However, it might be difficult to attain satisfactory results due to the stochastic nature and complexity of wind power generation. In order to explore a successful forecasting model, the necessity of analyzing the features of the raw time series should be increasingly highlighted. Therefore, the decomposition of wind power generation series appears to be an indispensable part in improving the forecasting accuracy. Empirical mode decomposition (EMD), perceived as an efficient decomposition method, is employed to decompose the wind power series into diverse IMFs for prediction [23,24]. Bao [23] presented a short term wind power output prediction model and the prediction of short-term wind power was implemented by differential EMD and relevance vector machine (RVM). In [24] a hybrid prediction model of wind farm power using EMD, chaotic theory and grey theory was constructed. The ultimate results indicated that the proposed method had good prediction accuracy. From the presented literature, it is possible to see that sometimes EMD cannot correctly decompose the raw data sequences. The IMFs extracted by EMD have lost their physical meaning and weaken the regularity. To address the mode mixing issue of the EMD technology, an improved method called ensemble empirical mode decomposition (EEMD) was introduced by Wu and Huang in 2009 [25]. Wang [26] selected EEMD as a data-cleaning method aiming to remove the high frequency noise embedded in the wind speed series. In this study, the EEMD is applied to decompose the original wind power generation series into several empirical modes, and the simulation results are encompassed in comparison with EMD.

Furthermore, a wealth of variables have great influence on the forecasting accuracy and efficiency, and the literature on the input selection gives this scant regard. These studies tend to select inputs using personal experience alone. However, in this research reported here, principal component analysis (PCA) was conducted to select inputs. PCA, a multivariate data analysis technology, can transform a set of correlated variables into new uncorrelated variables, namely principal components, containing most of the comprehensive variability of the original dataset. Lam [27] conducted PCA to extract a 2-component model from five raw variables for modelling the electricity use in office buildings. The literature on the importance of input dimensionality reduction and the appropriate selection of modelling variables has been widely reported. Ndiaye [28] applied PCA to select nine variables from all available variables to predict the electricity consumption in residential dwellings. Hong [29] proposed a hybrid PCA neural network model to forecast the day-ahead electricity price.

In this paper, the principal purpose of the experiment was to investigate a more accurate forecasting method for wind power generation. A hybrid model based on EEMD-principal component analysis (PCA)-least squares support vector machine (LSSVM)-bat algorithm (BA) was employed to forecast wind power generation. In addition, different models were developed using all available variables (least squares support machine-bat algorithm (LSSVM-BA), ensemble empirical mode decomposition-least squares support machine-bat algorithm (EEMD-LSSVM-BA)), and using only the variables deemed as significant by the PCA procedure (PCA-LSSVM-BA, ensemble empirical mode decomposition-principal component analysis-least squares support machine (EEMD-PCA-LSSVM), EEMD-PCA-LSSVM-BA). Therefore, a secondary aim of the present study was to determine whether an accuracy loss occurs when reducing the number of modelling variables using PCA. In comparison with the EMD method, the EEMD method can effectively mine the features of the original series through decomposing the series according to the difference of frequencies. First, the EEMD method was adopted to decompose the original wind power generation series to enhance the prediction

performance. Then, PCA was utilized to reduce the number of modelling inputs by identifying the significant variables maintaining most of the information present in the data set. Finally, LSSVM models were developed to predict the sub-series. Noticeably, in this work the two parameters of LSSVM were fine-tuned by the BA to ensure the generalization and the learning ability of LSSVM. The wind power generation forecasting values can be obtained according to the accumulation of the prediction values of all sub-series. To demonstrate the effectiveness of the proposed method, a case study from China was examined and the grey relational analysis was applied to evaluate the rationality of the forecasting series stemmed from the hybrid model from the perspective of geometric shape.

The advantages of the proposed hybrid model, which result in the better forecasting performance, can be summed up in the following several aspects: in the beginning, many single methods are applied to implement wind prediction using the original series directly, but the forecasting accuracy is not very satisfactory due to the influence of random noise in the raw series. In this study, EEMD is employed to preprocess the original wind power generation series to reduce the effect of random noise. Then, the determination of inputs in the proposed model is more novel. From previous papers, the selection of inputs is usually based on personal experience. However, wind power generation may be affected by many factors such as temperature, wind speed and installed capacity. Thus, the innovation of this paper is the application of PCA to select the proper inputs. Moreover, since artificial neural networks suffer from several disadvantages such as the occurrence of local minima, over fitting and slow convergence rate, LSSVM utilized in this study can improve the training speed for solving the problem. Unlike other LSSVM parameters optimization methods, which only utilize personal experience or traditional intelligent algorithms such as particle swarm optimization, the BA applied in this paper can avoid falling into local optimization and guarantee the generalization and the learning ability of LSSVM. Finally, grey relational analysis is utilized to demonstrate the superiority of the presented model considering the geometric shape of forecasting series and statistics. In brief, the novelty of the proposed model is described as follows: (a) a data preprocessing approach is explored to achieve the treatment of the original wind power generation series; (b) a PCA procedure is conducted to reduce the number of inputs without lowering the forecasting accuracy; (c) a LSSVM model with the relevant parameters optimized by BA is built to predict wind power generation; (d) grey relational analysis is adopted to cast light on the forecasting capacity of the proposed model.

The rest of this paper is organized as follows: Section 2 describes the modelling approaches of the proposed technique in detail. In Section 3 a hybrid model is constructed which is designed to predict wind power generation. Then, in Section 4 the proposed model is examined by a case study and an in depth comparison with other existing methods. Finally, Section 5 provides some conclusions of the whole research.

## 2. Methodology

### 2.1. Ensemble Empirical Mode Decomposition

EMD, originally proposed by Huang [30], is a powerful signal decomposition technology that aims to decompose complicated signals into several IMF components. However, sometimes EMD cannot correctly decompose the raw data sequences. These IMFs extracted by EMD have lost their physical meanings and weaken the regularity. Compared with EMD, EEMD has good performance in non-stationary signal decomposition. EEMD adds a white noise series to the raw signal  $x(t)$  to eliminate the mode mixing, obtaining the IMFs through the EMD procedures. The computational steps of the EEMD algorithm are described as follows:

*Step 1:* Calculate  $x^i(t) = x(t) + n^i(t)$ , where  $n^i(t)$  ( $i = 1, 2, 3, \dots, N$ ) represent the random white Gaussian noise series.

*Step 2:* Decompose the series  $x^i(t)$  using the EMD technology to obtain IMF modes  $imf_m^i(t)$  ( $m = 1, 2, 3, \dots, N$ ).

*Step 3:* Compute the mean of the corresponding series  $imf_m^i(t)$  as follows:

$$\overline{imf_m(t)} = \frac{1}{N} \sum_{i=1}^N imf_m^i(t) \quad (1)$$

Step 4: Repeat the above mean procedure to complete the process of EEMD. The decomposed results of the original signal series  $x(t)$  will be obtained as follows:

$$x(t) = \sum_{m=1}^k \overline{imf_m(t)} + r_k(t) \quad (2)$$

where  $\overline{imf_m(t)}$ , ( $m = 1, 2, 3, \dots, k$ ) are the IMFs decomposed by EEMD,  $r_k(t)$  denotes the corresponding residue.

### 2.2. Principal Component Analysis (PCA)

PCA based on population correlation coefficients is a statistical modelling technology which can identify the correlation among variables and generalize the data group in the light of particular linear combinations of variables, named principal components. In this study, PCA is employed to select the significant modelling inputs. Every principal component maintains interrelated variables resembling a data set. The first component indicates the paramount source of variance in the original data, and the other components account for the remaining variability. Details of the PCA method procedure are reported in [31].

### 2.3. Least Squares Support Vector Machine

The LSSVM, put forward by Suykens [32], is a variation of the standard support vector machine (SVM), adopting the loss function different from SVM and minimizing the square error. A quadratic programming problem can be transformed into linear equations through replacing inequality constraints with equality constraints, greatly reducing the computational complexity. In the LSSVM model, the training sample set  $S = \{(x_i, y_i) \mid i = 1, 2, 3, \dots, t\}$ ,  $x_i = R^n$ ,  $y_i = R$ . Then, the optimal decision function is framed by using the high dimensional feature space. The decision function can be expressed as follows:

$$f(x) = \omega^T \varphi(x) + b \quad (3)$$

where  $\varphi(x)$  represents the nonlinear mapping function from input space to high dimensional feature space,  $\omega$  is weight,  $b$  is bias.

The structural risk minimization can be described as follows:

$$R = \frac{1}{2} \|\omega\|^2 + c R_{\text{emp}} \quad (4)$$

where  $\|\omega\|^2$  suggests the complex degree of the model,  $c$  is the regularization parameter, controlling the degree of punishment beyond the error samples,  $R_{\text{emp}}$  is the empirical risk function, the objective function of LSSVM is obtained as follows:

$$\min Z(\omega, \xi) = \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^t \xi_i^2 \quad (5)$$

$$\text{s.t. } y_i = \omega \varphi(x_i) + \xi_i + b, \quad i = 1, 2, 3, \dots, t$$

where  $\xi_i$  is the error, the Lagrange function can be defined as follows:

$$L(\omega, b, \xi, \lambda) = \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^t \xi_i^2 - \sum_{i=1}^t \lambda_i (\omega \varphi(x_i) + \xi_i + b - y_i) \quad (6)$$

where  $\lambda_i (1, 2, 3, \dots, t)$  are the Lagrange multipliers

According to the Karush-Kuhn-Tucker (KKT) conditions, Equation (7) is shown as follows:



$$\begin{cases} \omega - \sum_{i=1}^t \lambda_i \xi_i^2 = 0, \\ \sum_{i=1}^t \lambda_i = 0, \\ \lambda_i - c \xi_i = 0, \\ \omega \varphi(x_i) + \xi_i + b - y_i = 0. \end{cases} \quad (7)$$

In the light of Equation (7), the optimization problem can be converted into the process of solving linear equations, which is presented as follows:

$$\begin{bmatrix} 0 & I^T \\ I & J + \frac{1}{c} \end{bmatrix} \begin{bmatrix} b \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \quad (8)$$

where  $I = [11\dots 1]^T$  is a  $t \times 1$  dimensional column vector,  $\lambda = [\lambda_1 \lambda_2 \dots \lambda_t]^T$ ,  $y = [y_1 y_2 \dots y_t]^T$ ,  $J_{ij} = \varphi(x_i)^T \varphi(x_j) = K(x_i, x_j)$ ,  $K$  is the kernel function which satisfies the condition of Mercer, the final form of LSSVM model emerges as follows:

$$f(x) = \sum_{i=1}^t \lambda_i K(x_i, x_j) + b \quad (9)$$

In this research, the radial basis function (RBF) is selected as the kernel function, as shown in Equation (10):

$$K(x_i, x_j) = \exp \left[ \frac{-\|x_i - x_j\|^2}{2\sigma^2} \right] \quad (10)$$

where  $\sigma^2$  is the parameter of the kernel function

Then, there are two parameters, the regularization parameter and the kernel parameter, determining the LSSVM model. In previous studies, experimental comparison, grid searching methods and cross validation methods were applied to optimize the two parameters, but they are time-consuming and inefficient. Therefore, this paper adopts a BA to optimize the two parameters, which can enhance and further the adaptability of the model and effectively improve the forecasting accuracy.

#### 2.4. Bat Algorithm (BA)

The BA is a novel meta-heuristic algorithm inspired by the echolocation behavior of bats. The BA offers an excellent way for optimization and classification in a powerful selection of complicated problems [19]. The basic flow of the BA can be generalized by the pseudo code listed in Algorithm 1.

---

##### Algorithm 1. Pseudo code of the Bat Algorithm.

---

- (1) Initialize the position of bat population  $x_i$  ( $i = 1, 2, \dots, n$ ) and  $v_i$
  - (2) Initialize pulse frequency  $f_i$  at  $x_i$ , pulse rates  $r_i$  and the loudness  $A_i$
  - (3) **While** ( $t <$  maximum number of iterations)
  - (4) Generate new solutions by adjusting frequency
  - (5) Update the velocities and solutions
  - (6) **If** ( $\text{rand} > r_i$ )
  - (7) Select a solution among the best solutions
  - (8) Generate a local solution around the selected best solution
  - (9) **End if**
  - (10) Generate a new solution by flying randomly
-

- 
- (11) **If** (rand <  $A_i$  &  $f(x_i) < f(x^*)$ )
  - (12) Accept the new solutions
  - (13) Increase  $r_i$  and reduce  $A_i$
  - (14) **End if**
  - (15) Rank the bats and find the current best  $x^*$
  - (16) **End while**
- 

### 2.5. Grey Relational Analysis

Based on the proximity measure similarity, the grey relational analysis theory was first proposed by Deng [33]. The purpose of the grey relational analysis is to examine whether the various series have a close relationship on the basis of the similarity degree of the geometric shape of the series. The higher the similarity degree is, the greater the correlation is. The basic steps of grey relational analysis are as follows:

*Step 1:* Define reference and comparison series

A reference time series can be defined as follows:

$$Y_0 = (Y_0(1), Y_0(2), \dots, Y_0(n)) \quad (11)$$

Then,  $t$  time series can be explained as follows:

$$Y_i = (Y_i(1), Y_i(2), \dots, Y_i(n)), i = 1, 2, \dots, t \quad (12)$$

*Step 2:* Dimensionless processing of time series:

$$\bar{Y}_i = \frac{1}{n} \sum_{m=1}^n Y_i(m) \quad (13)$$

$$S_i = \sqrt{\frac{1}{n-1} \sum_{m=1}^n (Y_i(m) - \bar{Y}_i)^2} \quad (14)$$

$$y_i(m) = \frac{Y_i(m) - \bar{Y}_i}{S_i} \quad (15)$$

*Step 3:* Compute the correlation coefficient:

$$r(y_0(k), y_i(k)) = \frac{\min_i \min_k |y_0(k) - y_i(k)| + \xi \max_i \max_k |y_0(k) - y_i(k)|}{|y_0(k) - y_i(k)| + \xi \max_i \max_k |y_0(k) - y_i(k)|} \quad (16)$$

where  $\xi$  is the distinguishing coefficient. In this work here, let  $\xi = 0.5$ .

*Step 4:* Calculate the grey relational degree:

$$r_i = \frac{1}{n} \sum_{k=1}^n r(y_0(k), y_i(k)) \quad (17)$$

where  $r_i$  is the grey relational degree of  $(y_0, y_i)$ , representing the similarity degree of the geometric shape of the series.

*Step 5:* Sort the grey relational degree

The grey relational degrees of series are ranked according to the size of the grey relational degrees. If  $r_i > r_k$ , then the similarity of the curve of  $i$  series to the curve of the reference series is higher than that of the  $k$  series.



### 3. Wind Power Generation Forecasting Model

In this section, the proposed model (EEMD-PCA-LSSVM-BA) is constructed in detail. The flowchart of the presented model is given in Figure 1. In addition, the diverse LSSVM models are developed by using all variables from the data set, and using only the variables previously deemed significant by PCA procedure. The forecasting accuracy of both methods is compared to determine whether the PCA procedure is successful in selecting significant inputs. The following four parts constitute the hybrid model.

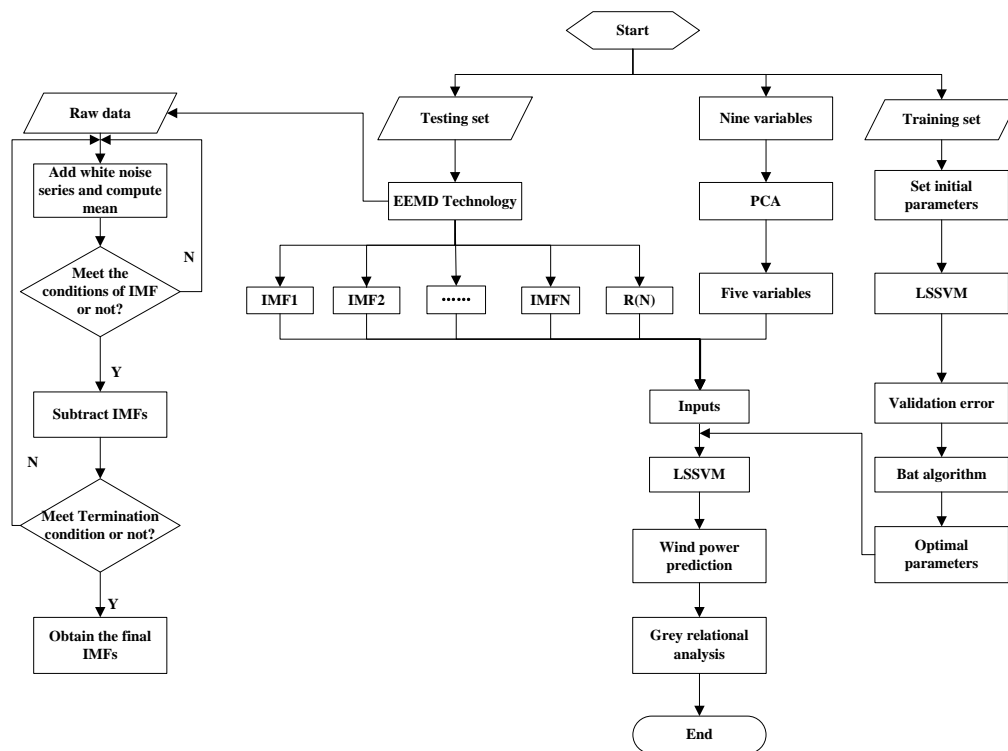


Figure 1. The flowchart of the proposed model.

Part one: Data preprocessing. The EEMD approach is adopted to decompose the original wind power generation series into different IMFs. The aim of this technology is to diminish the non-stationary character of the series for the high-precision prediction.

Part two: Input selection. Using the PCA to reduce the number of modelling inputs without lowering the prediction accuracy, the procedure can efficiently mine the significant variables containing most of the overall variability present in the data sets.

Part three: Training and validation of model. In this study, wind power generation forecasting approach is in the light of LSSVM-BA model, the basic steps can be described as follows:

*Step 1: Parameter setting*

The main parameters of BA are initial population size  $n$ , maximum iteration number  $N$ , original loudness  $A$ , pulse rate  $r$ , location vector  $x$ , speed vector  $v$ , respectively.

*Step 2: Initialize population*

Initialize the bat populations position, each bat location strategy is a component of  $(\gamma, \sigma^2)$ , which can be defined as follows:

$$x = x_{\min} + \text{rand}(1, d) \times (x_{\max} - x_{\min}) \tag{18}$$

where the dimension of the bat population:  $d = 2$ .

### Step 3: Update parameters

Calculate the fitness value of population, find the current optimal solution and update the pulse frequency, velocity and position of bats as follows:

$$f_i = f_{\min} + (f_{\max} - f_{\min}) \times \beta \quad (19)$$

$$v_i^t = v_i^{t-1} + (x_i^t - x^*) \times f_i \quad (20)$$

$$x_i^t = x_i^{t-1} + v_i^t \quad (21)$$

where  $\beta$  denotes uniformly random numbers,  $\beta \in [0,1]$ ;  $f_i$  is the search pulse frequency of the bat  $i$ ,  $f_i \in [f_{\min}, f_{\max}]$ ;  $v_i^t$  and  $v_i^{t-1}$  are the velocities of the bat  $i$  at time  $t$  and  $t-1$ , respectively; further,  $x_i^t$  and  $x_i^{t-1}$  represent the location of the bat  $i$  at time  $t$  and  $t-1$ , respectively;  $x^*$  is the present optimal solution for all bats.

### Step 4: Update loudness and pulse frequency

Produce a uniformly random number  $\text{rand}$ , if  $\text{rand} > r_i$ , disturb the optimal strategy randomly and acquire a new strategy; if  $\text{rand} < A_i$  and  $f(x) > f(x^*)$ , then the new strategy can be accepted, the  $r_i$  and  $A_i$  of the bat are updated as follows:

$$A_i^{t+1} = \alpha A_i^t \quad (22)$$

$$r_i^{t+1} = r_i^0 [1 - \exp(-\gamma t)] \quad (23)$$

where  $\alpha$  and  $\gamma$  are constants.

### Step 5: Output the global optimal solution

The current optimal solution can be obtained depending on the rank of all fitness values of the bat population. Repeat the steps of Equation (19) to Equation (21) till the maximum iterations are completed and output the global optimal solution. Therefore, a wind power generation prediction model can be generated.

In addition, the LSSVM approach is employed to model the training set, and the mean square errors of the true values and forecasting values are adopted as the fitness functions of the BA. Then, the group of parameters of LSSVM is optimized by BA for the minimum fitness value. Finally, the LSSVM model with optimal parameters can be applied to predict the wind power generation.

Part four: Wind power generation forecasting. In this part, the LSSVM approach with the parameters optimized by the BA is employed to predict each series decomposed by EEMD. Then, the forecasting series of wind power generation can be obtained by accumulating the prediction values of each subsequence. After obtaining the prediction values through the presented hybrid model, grey relational analysis was developed to determine the forecasting performance of the hybrid model.

## 4. Case Study

### 4.1. Study Area and Data Set

In this paper, the selected study area is a wind farm: Zhangjiakou, which is located in northwest China-Hebei Province, featuring an abundant wind energy source. In this work the daily wind power generation data from 1 January 2015 to 28 October 2015 are chosen as the samples to illustrate the effective performance of the proposed model. The daily measurements of the nine variables of this period are average wind speed, daily mean temperature, highest temperature, equivalent utilization hours, lowest temperature, availability of fan, maximum wind speed, minimum wind speed and installed capacity, respectively. The total number of daily wind power generation data is 301. The series are divided into two parts: training set and testing set. Data from 1 January 2015 to 8 August 2015

accounting for approximately 73% of the data are selected as training set. The rest of the data are regarded as the testing set.

#### 4.2. Performance Criteria of Prediction Accuracy

In this paper, root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE) conceived as evaluation criteria are employed to assess the forecasting performance of the proposed model quantitatively:

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

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i| \quad (25)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - \hat{x}_i}{x_i} \right| \quad (26)$$

where  $x_i$  is the actual value at  $i$ , and  $\hat{x}_i$  is the corresponding predictive value.

#### 4.3. Selection of Modelling Inputs

The selection of the variables utilized as the modelling inputs plays a pivotal role in exploring a powerful forecasting model. The convergence problem and poor forecasting performance may appear when redundant variables or ones that offer little contribution to the model are utilized. Moreover, the variables can increase the effort to develop models [34]. In addition, too few variables may result in lower prediction accuracy resulting from the inability of the available inputs to explain the model output behavior [35]. A successful model should employ relatively few inputs containing enough relevant information to attain satisfactory forecasting precision.

In this study, PCA was applied to measure all the variables except the wind power generation—the variable to be forecasted. According to the theory of PCA, each component is expressed by a linear equation involving all variables, and in this equation every variable can obtain a coefficient. Variables that comprise most information present in the data set and have large coefficients in the first components can be perceived as significant variables due to the fact they have the most contribution to the overall data variability. In the method proposed here, variables with coefficients having an absolute value larger than 0.1 in the components which cumulatively explain at least 95% of the overall variability were conceived to be significant. The first five components extracted from the data set approximately explain 97% of the result, but the other four components explain less than 3%. Thus, variables having coefficients with absolute values greater than 0.1 from the first five components were considered significant. The data variability explained by the top five components and the absolute values of these variable coefficients are described in Tables 1 and 2 respectively.

**Table 1.** Data variability explained by the top five principal components (%).

Components	Per Component	Cumulative
Comp.1	41.4457	41.4457
Comp.2	25.9709	67.4166
Comp.3	17.8437	85.2602
Comp.4	7.2593	92.5195
Comp.5	4.6729	97.1924

From Table 2, it can be seen that average wind speed, daily mean temperature, equivalent utilization hours, availability of fan and maximum wind speed have the largest coefficients contributing

the most to the overall data variability. However, the phenomenon that some variables may have coefficients with absolute values slightly larger than 0.1 in the first principal components while they have much greater coefficients in the bottom components might occur. To address this issue, there needs to be a further analysis pertaining to the coefficients of the variables chosen previously. Specifically, a variable would not be perceived significant when the first five coefficients of the variable do not belong to the first five components. The ranking of the absolute values of variable correlation coefficients of the first five principal components can be noticeably indicated in Table 3. For instance, from Table 2 it can be seen that the variable including measurements of minimum wind speed has a coefficient superior to 0.1 in the component 4 which is one of the top five components. However, it can be seen from Table 3 that the all coefficients of this variable do not rank among the largest five in the first components. Therefore, this variable would not be regarded as significant.

**Table 2.** Variable correlation coefficients of the first five principal components.

Variables	Coefficients in Components				
	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
Average wind speed	0.3315	0.3095	0.4210	0.4513	0.0668
Daily mean temperature	0.4918	0.3049	0.0932	0.1211	0.0076
Highest temperature	0.0862	0.0893	0.0671	0.0276	0.0600
Equivalent utilization hours	0.2703	0.5167	0.1130	0.1994	0.5545
Lowest temperature	0.0800	0.0092	0.0405	0.0141	0.0137
Availability of fan	0.1835	0.3512	0.5498	0.4711	0.3210
Maximum wind speed	0.2655	0.4906	0.1342	0.1503	0.7618
Minimum wind speed	0.0652	0.0313	0.0544	0.1145	0.0300
Installed capacity	0.0000	0.0000	0.0000	0.0000	0.0000

**Table 3.** The ranking of the absolute values of variable correlation coefficients of the first five principal components.

Variables	Ranking of the Absolute Values of Coefficients in Components				
	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
Average wind speed	2	4	2	2	4
Daily mean temperature	1	5	5	5	8
Highest temperature	6	6	6	7	5
Equivalent utilization hours	3	1	4	3	2
Lowest temperature	7	8	8	8	7
Availability of fan	5	3	1	1	3
Maximum wind speed	4	2	3	4	1
Minimum wind speed	8	7	7	6	6
Installed capacity	9	9	9	9	9

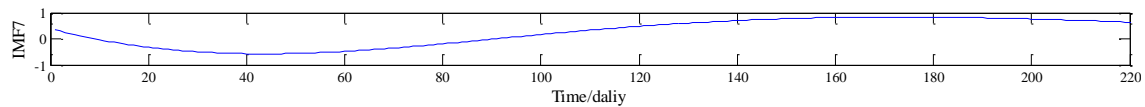
Based on the above procedure, in this study five significant variables are identified and utilized as the inputs of the models to forecast the wind power generation. The result of selected variables through PCA can be shown in Table 4.

**Table 4.** The selected variables through principal component analysis (PCA).

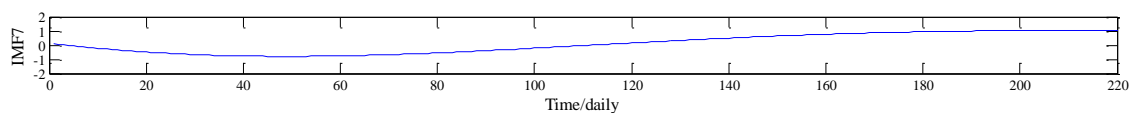
Variables	
Average wind speed	✓
Daily mean temperature	✓
Highest temperature	
Equivalent utilization hours	✓
Lowest temperature	
Availability of fan	✓
Maximum wind speed	✓
Minimum wind speed	
Installed capacity	

#### 4.4. Ensemble Empirical Mode Decomposition Results

To improve the forecasting performance of wind power generation series, in this study, EEMD is devised for decomposing the raw series. To corroborate the performance of EEMD, EMD is employed to decompose the original series. EMD is similar to EEMD, and the two technologies both decompose the raw wind power generation series into seven IMFs and one residue. However, for the same IMF7 from Figures 2 and 3 it can be seen that EEMD can retain the true information of the original data sequence to the utmost, and effectively suppress the occurrence of mode mixing and eliminate the noise, in line with the actual situation.



**Figure 2.** The ensemble empirical mode decomposition (EEMD) decomposed results of wind power generation of the training set.



**Figure 3.** The empirical mode decomposition (EMD) decomposed results of wind power generation of the training set.

#### 4.5. Selection of LSSVM Model

Previous studies on the LSSVM model for prediction demonstrate that the performance of the LSSVM approach relies on its parameters and the kernel function. The optimization of parameters is an indispensable part of any LSSVM model. The BA regarded as a population intelligent optimization algorithm offers a novel idea for searching the optimal parameters of LSSVM. In this paper, RBF is chosen as the kernel function of LSSVM algorithm, decreasing the complexity of the model and improving the training speed. Thus, the regularization parameter  $\gamma$  and kernel parameter  $\sigma^2$  can obtain the optimal values using the powerful automatic searching ability of BA. The main parameters of the BA are listed in Table 5. Table 6 shows the optimal parameters ( $\gamma$ ,  $\sigma^2$ ) of the LSSVM models obtained using the BA approach.

**Table 5.** Main parameters of bat algorithm (BA).

Parameters	Values	Parameters	Values
Initial population size	10	Minimum frequency	0
Initial loudness	0.25	Maximum frequency	5
Pulse rate	0.5	Max-iteration number	50

**Table 6.** The optimal parameters in the LSSVM model. Intrinsic mode function: IMF.

Components	$\gamma$	$\sigma^2$
IMF1	0.8200	3.5274
IMF2	71.8112	5.1095
IMF3	81.2527	6.7157
IMF4	18.2928	78.1796
IMF5	13.5707	34.3623
IMF6	4.8853	54.2762
IMF7	2.2090	6.4192
RES	1.2663	11.1730

Then, the forecasting performance of the LSSVM model with the parameters tuned by the BA is examined by using the testing set. The prediction errors of LSSVM are presented in Figure 4. From Figure 4, it can be seen that the error change of LSSVM is relatively steady, and only four errors exceed 5. Furthermore, the maximum error is  $-5.5741$  amongst all errors. Thus, it implies that the forecasting results can be considered acceptable.

In addition, in terms of EEMD-PCA-LSSVM-PSO, the max-iteration number of PSO is 200, the size of population is 20, the inertial factor ( $w$ ) is 0.6, the learning factors ( $c_1 = c_2$ ) are both 1.7. With respect to BPNN, the number of neuron in hidden layer is 13, the training number of BP is 200, and the learning rate is 0.04. Considering the EEMD-PCA-LSSVM approach, grid searching method and cross validation approach are applied to select the optimal regularization parameter ( $\gamma$ ) and the kernel parameter ( $\sigma^2$ ). For the grid searching method,  $\gamma = 2^{-10}$ – $2^{15}$ , the step of  $\gamma$  is 1;  $\sigma^2 = 2^{15}$ – $2^{-10}$ , the step of  $\sigma^2$  is 1.

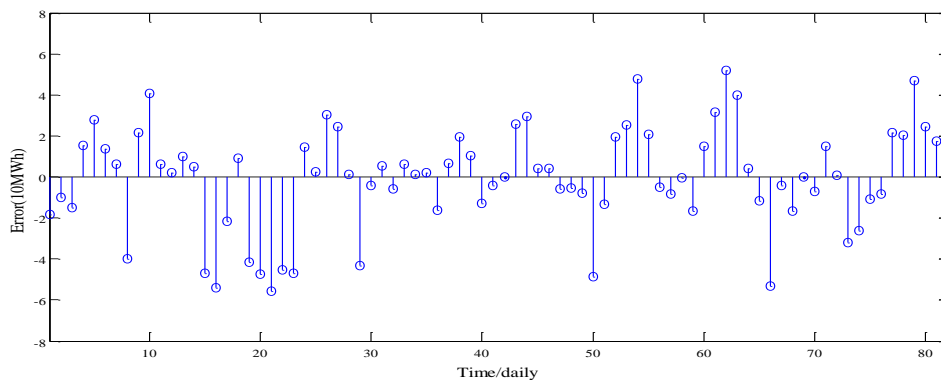


Figure 4. The prediction error of the proposed model.

#### 4.6. Comparative Analysis of Different Methods

To illustrate the excellent performance of the proposed model, this paper employs the LSSVM-BA, principal component analysis-least squares support machine-bat algorithm (PCA-LSSVM-BA), EEMD-PCA-LSSVM, EEMD-LSSVM-BA, empirical mode decomposition-principal component analysis-least squares support machine-bat algorithm (EMD-PCA-LSSVM-BA) and ensemble empirical mode decomposition-principal component analysis-least squares support machine-particle swarm optimization (EEMD-PCA-LSSVM-PSO) for comparison. Meanwhile, the single LSSVM and ARIMA models are developed to predict wind power generation. In addition, the back propagation neural network (BPNN) is utilized to forecast wind power generation. The comparison of prediction results with various models is shown in Table 7. Compared with other forecasting models, the proposed model displays better capacity on the prediction of wind power generation, capturing the characteristics of the wind power generation series, and achieving good forecasting performance.

Moreover, from Table 7, it can be seen that the hybrid model (EEMD-PCA-LSSVM-BA) has the highest accuracy compared with the ARIMA, BPNN, LSSVM, LSSVM-BA, PCA-LSSVM-BA, EEMD-PCA-LSSVM, EEMD-LSSVM-BA, EMD-PCA-LSSVM-BA and EEMD-PCA-LSSVM-PSO models. For instance, the proposed model achieves reductions of 44.44%, 42.99%, 41.64%, 38.14%, 28.42%, 16.58%, 20.98%, 10.93% and 9.58% in total MAPE compared with the ARIMA, BPNN, LSSVM, LSSVM-BA, PCA-LSSVM-BA, EEMD-PCA-LSSVM, EEMD-LSSVM-BA, EMD-PCA-LSSVM-BA and EEMD-PCA-LSSVM-PSO models. The abatements of MAE, RMSE and MAPE are evidently listed in Table 8. The computational formulas of the abatements of MAE, RMSE and MAPE can be defined as follows:

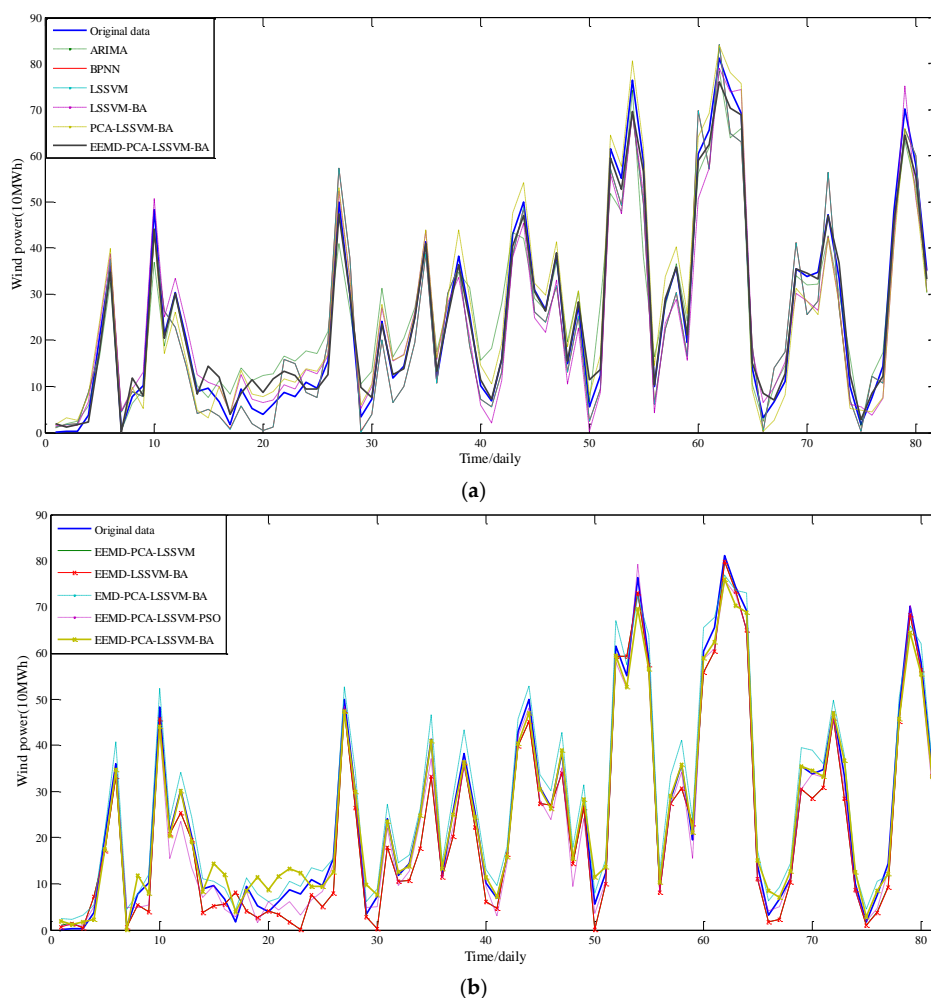
$$\frac{\text{MAE}_{\text{comparative model}} - \text{MAE}_{\text{EEMD-PCA-LSSVM-BA}}}{\text{MAE}_{\text{comparative model}}} \times 100\% \quad (27)$$

$$\frac{\text{RMSE}_{\text{comparative model}} - \text{RMSE}_{\text{EEMD-PCA-LSSVM-BA}}}{\text{RMSE}_{\text{comparative model}}} \times 100\% \quad (28)$$

$$\frac{\text{MAPE}_{\text{comparative model}} - \text{MAPE}_{\text{EEMD-PCA-LSSVM-BA}}}{\text{MAPE}_{\text{comparative model}}} \times 100\% \quad (29)$$

where the comparative models represent ARIMA, BPNN, LSSVM, LSSVM-BA, PCA-LSSVM-BA, EEMD-PCA-LSSVM, EEMD-LSSVM-BA, EMD-PCA-LSSVM-BA and EEMD-PCA-LSSVM-BA models; the initial values of MAE, RMSE and MAPE can be obtained from Table 7.

In addition, the absolute errors between the real values and the estimated values may be captured on the basis of Figure 5. From subfigure (a) of Figure 5, it can be seen that the single ARIMA, BPNN, and LSSVM models have poor performance in forecasting wind power generation, revealing the inability of the single models to address the comprehensive features of the original wind power generation series. In contrast, the prediction values obtained by the hybrid LSSVM methods can be acceptable. Furthermore, subfigure (b) of Figure 5 shows that the hybrid models of LSSVM with the parameters optimized by intelligent algorithms have great advantages in wind power generation forecasting. Most importantly, the LSSVM-BA method outperforms the LSSVM-PSO model.



**Figure 5.** The forecasting results with different models. (a) autoregressive integrated moving average (ARIMA), BPNN, LSSVM, least squares support machine-bat algorithm (LSSVM-BA), PCA-LSSVM-BA and EEMD-PCA-LSSVM-BA models; (b) EEMD-PCA-LSSVM, ensemble empirical mode decomposition-least squares support machine-bat algorithm (EEMD-LSSVM-BA), EMD-PCA-LSSVM-BA, EEMD-PCA-LSSVM-PSO and EEMD-PCA-LSSVM-BA models.

Form Table 7, it can be concluded that: (a) among all the forecasting models, the proposed EEMD-PCA-LSSVM-BA model achieves the best performance. Especially, compared with



PCA-LSSVM-BA and EMD-PCA-LSSVM-BA models, it can be found that EEMD technology can apparently enhance the forecasting ability of wind power generation series with regard to the evaluation indexes of MAE, RMSE and MAPE. For instance, the MAPE of EEMD-PCA-LSSVM-BA is 21.43%, but PCA-LSSVM-BA and EMD-PCA-LSSVM-BA models are 29.94% and 24.06%, respectively; (b) in comparison with EEMD-LSSVM-BA, the EEMD-PCA-LSSVM-BA model has better forecasting performance, demonstrating that the PCA procedure is successful in selecting significant inputs. Also, the PCA-LSSVM-BA model using the PCA-selected inputs slightly outperforms the LSSVM-BA model using all the inputs, corroborating the excellent ability of PCA procedure; (c) this study establishes two improved LSSVM models, and the performance of LSSVM based on BA is superior to the EEMD-PCA-LSSVM-PSO model concerning the three criteria of MAE, RMSE and MAPE. For instance, the MAE of EEMD-PCA-LSSVM-BA is 2.0298, while the MAE of EEMD-PCA-LSSVM-PSO is 2.4356. There seems to be a paramount reason for this phenomenon, namely that the BA adopts the major advantages of the existing intelligent algorithms in some way, combining the amazing echolocation behavior of bats, while particle swarm optimization is a special case of the BA in simplified form; (d) the improved LSSVM models have better performance than single LSSVM approach. The primary reason may be that the process of automatic searching is added to the improved LSSVM model, which equips the LSSVM model with better learning and generalization ability to acquire the global optimal solution easily; (e) in comparison with the ARIMA model merely using the raw wind power generation series, the improved LSSVM methods and the Neural network model(BPNN) are more powerful than the ARIMA model, which proves that the intelligent approaches have more research value and development space than the statistical models in the realm of wind power generation forecasting.

Furthermore, the time durations of the computing about different approaches are described in Table 9. In this study, a computer equipped with an Intel<sup>®</sup> Core<sup>™</sup> i3-3110M processor CPU @ 2.40 GHz, 4 GB RAM and the 64 bit Windows 7 operating system (OS) was used. Also, MATLAB R2014a was applied to write all programs of this paper.

From Table 9, it can be seen that the single models such as ARIMA and LSSVM take less time compared with the hybrid models. However, the prediction accuracy of single models is lower than the hybrid models. Thus, it is reasonable to adopt more accurate wind power generation approaches taking a little more time for the security of the power system. In addition, the forecasting time of EEMD-PCA-LSSVM-BA is smaller than EEMD-PCA-LSSVM-PSO. It suggests that BA can reduce the time of parameter optimization of LSSVM effectively.

#### 4.7. The Results of Grey Relational Analysis

In the current work, grey relational analysis is applied to verify whether the curve of the prediction result from the presented model has the highest similarity to the curve of the actual wind power generation series. Table 10 describes the grey relational degrees among the forecasting results from different models. From Table 10, it can be seen apparently that the proposed EEMD-PCA-LSSVM-BA model has the greatest similarity to the true wind power generation curve from the perspective of the geometric shape of the series.

**Table 7.** The comparison of prediction results with different models.

Indexes	Forecasting Methods									
	ARIMA	BPNN	LSSVM	LSSVM-BA	PCA-LSSVM-BA	EEMD-PCA-LSSVM	EEMD-LSSVM-BA	EMD-PCA-LSSVM-BA	EEMD-PCA-LSSVM-PSO	EEMD-PCA-LSSVM-BA
MAE (10 MWh)	4.5965	4.5038	4.4376	3.8998	3.8402	3.1324	3.3559	2.9335	2.4356	2.0298
RMSE (10 MWh)	5.8381	5.1245	4.9528	4.3537	4.0914	3.3859	3.9463	3.1906	2.8756	2.7117
MAPE (%)	38.57%	37.59%	36.72%	34.64%	29.94%	25.69%	27.12%	24.06%	23.70%	21.43%

**Table 8.** The abatements of MAE, RMSE and MAPE reductions in comparison with model except EEMD-PCA-LSSVM-BA.

Indexes	The Proportion of Reductions									
	ARIMA	BPNN	LSSVM	LSSVM-BA	PCA-LSSVM-BA	EEMD-PCA-LSSVM	EEMD-LSSVM-BA	EMD-PCA-LSSVM-BA	EEMD-PCA-LSSVM-PSO	EEMD-PCA-LSSVM-BA
MAE (10 MWh)	55.84%	54.93%	54.26%	47.95%	47.14%	35.20%	39.52%	30.81%	16.66%	16.66%
RMSE (10 MWh)	53.55%	47.08%	45.25%	37.72%	33.72%	19.91%	31.29%	15.01%	5.70%	5.70%
MAPE (%)	44.44%	42.99%	41.64%	38.14%	28.42%	16.58%	20.98%	10.93%	9.58%	9.58%

**Table 9.** The time durations of the computing about different approaches.

CPU time (s)	Forecasting Methods									
	ARIMA	BPNN	LSSVM	LSSVM-BA	PCA-LSSVM-BA	EEMD-PCA-LSSVM	EEMD-LSSVM-BA	EMD-PCA-LSSVM-BA	EEMD-PCA-LSSVM-PSO	EEMD-PCA-LSSVM-BA
Time (s)	1.045 s	61.284 s	2.660 s	55.977 s	50.500 s	44.063 s	53.165 s	48.172 s	72.074 s	48.277 s

**Table 10.** The grey relational degree of various models.

Grey Relational Degree									
ARIMA	BPNN	LSSVM	LSSVM-BA	PCA-LSSVM-BA	EEMD-PCA-LSSVM	EEMD-LSSVM-BA	EMD-PCA-LSSVM-BA	EEMD-PCA-LSSVM-PSO	EEMD-PCA-LSSVM-BA
0.7954	0.7902	0.7898	0.7924	0.8064	0.8007	0.8045	0.8157	0.8209	0.8331

## 5. Conclusions

In order to enhance the forecasting accuracy wind power generation efficiently, a hybrid model is framed in this study. First, an EEMD technology was employed to decompose the original wind power generation series. Then, principal component (PCA) was applied to select the significant modelling inputs: five significant variables were selected from nine available inputs. Next, the relevant parameters of the proposed model were optimized by a BA. Finally, the presented method with favorable learning ability and generalization was developed to predict wind power generation. The simulation results and grey relational analysis indicate that the proposed hybrid model performs better than ARIMA, BPNN, LSSVM, LSSVM-BA, PCA-LSSVM-BA, EEMD-PCA-LSSVM, EEMD-LSSVM-BA, EMD-PCA-LSSVM-BA and EEMD-PCA-LSSVM-PSO models.

The superiority of the proposed hybrid model over other models may be accounted for by following aspects: (a) the forecasting performance of wind power generation series can be greatly augmented by using an EEMD method; (b) the simplified model using a reduced number of inputs selected by PCA procedure is more accurate than the models using all the inputs. This could suggest that variables not considered significant not only do not bring valuable information to the input set, but also add noise and unnecessary variability affecting the forecasting accuracy of models; (c) the parameters of the LSSVM models play an essential role in wind power generation prediction. Therefore, in this paper the BA is employed to optimize the parameters of the LSSVM model; (d) the hybrid model can comprehensively capture the characteristics of the raw wind power generation series, whilst the single models can only tap into the limited features of the original series. In this sense, it might be rational to see that the proposed hybrid model performs better than the other single or hybrid models regarding the criteria of MAE, RMSE and MAPE. In addition, the larger grey relational values also confirm that the proposed model outperforms the other models from the perspective of the geometric shape of forecasting series and statistics. Thus, the current method is a credible and promising algorithm for wind power generation prediction.

Regarding some limitations of this study, further research is necessary. Due to the unavailability of a reliable numerical weather prediction system, the meteorological data such as pressure, relative humidity and air density cannot be obtained. Therefore, this study only selected nine variables as alternative variables of the modelling inputs. The other relevant variables that affect wind power generation need to be investigated in further research.

**Acknowledgments:** This study is supported by the Fundamental Research Funds for the Central Universities (2015ZD33).

**Author Contributions:** Qunli Wu designed this paper and made overall guidance; Chenyang Peng wrote the whole manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Wan, C.; Xu, Z.; Pinson, P.; Dong, Z.Y.; Wong, K.P. Optimal prediction intervals of wind power generation. *IEEE Trans. Power Syst.* **2014**, *29*, 1166–1174. [[CrossRef](#)]
2. Saleh, A.E.; Moustafa, M.S.; Abo-Al-Ez, K.M.; Abdullah, A.A. A hybrid neuro-fuzzy power prediction system for wind energy generation. *Int. J. Electr. Power Energy Syst.* **2016**, *74*, 384–395. [[CrossRef](#)]
3. Foley, A.M.; Leahy, P.G.; Marvuglia, A.; McKeogh, E.J. Current methods and advances in forecasting of wind power generation. *Renew. Energy* **2012**, *37*, 1–8. [[CrossRef](#)]
4. Ewing, B.T.; Kruse, J.B.; Schroeder, J.L.; Smith, D.A. Time series analysis of wind speed using var and the generalized impulse response technique. *J. Wind Eng. Ind. Aerodyn.* **2007**, *95*, 209–219. [[CrossRef](#)]
5. Hill, D.; Bell, K.R.W.; McMillan, D.; Infield, D. A vector auto-regressive model for onshore and offshore wind synthesis incorporating meteorological model information. *Adv. Sci. Res.* **2014**, *11*, 35–39. [[CrossRef](#)]
6. Erdem, E.; Shi, J. ARMA based approaches for forecasting the tuple of wind speed and direction. *Appl. Energy* **2011**, *88*, 1405–1414. [[CrossRef](#)]

7. Liu, H.; Erdem, E.; Shi, J. Comprehensive evaluation of ARMA–GARCH(-M) approaches for modeling the mean and volatility of wind speed. *Appl. Energy* **2011**, *88*, 724–732. [[CrossRef](#)]
8. Kavasseri, R.G.; Seetharaman, K. Day-ahead wind speed forecasting using *f*-ARIMA models. *Renew. Energy* **2008**, *34*, 1388–1393. [[CrossRef](#)]
9. Huang, D.Z.; Gong, R.X.; Gong, S. Prediction of wind power by chaos and BP artificial neural networks approach based on genetic algorithm. *J. Electr. Eng. Technol.* **2015**, *10*, 41–46. [[CrossRef](#)]
10. Li, G.; Shi, J. On comparing three artificial neural networks for wind speed forecasting. *Appl. Energy* **2010**, *87*, 2313–2320. [[CrossRef](#)]
11. Guo, Z.H.; Zhao, J.; Zhang, W.Y.; Wang, J.Z. A corrected hybrid approach for wind speed prediction in Hexi Corridor of China. *Energy* **2011**, *36*, 1668–1679. [[CrossRef](#)]
12. De Giorgi, M.; Campilongo, S.; Ficarella, A.; Congedo, P. Comparison between wind power prediction models based on wavelet decomposition with least-squares support vector machine (LS-SVM) and artificial neural network (ANN). *Energies* **2014**, *7*, 5251–5272. [[CrossRef](#)]
13. Yuan, X.; Chen, C.; Yuan, Y.; Huang, Y.; Tan, Q. Short-term wind power prediction based on LSSVM–GSA model. *Energy Convers. Manag.* **2015**, *101*, 393–401. [[CrossRef](#)]
14. Wang, X.; Li, H. One-month ahead prediction of wind speed and output power based on EMD and LSSVM. In Proceedings of the International Conference on Energy and Environment Technology, Guilin, China, 16–18 October 2009; pp. 439–442.
15. Hu, Z.Y.; Liu, Q.Y.; Tian, Y.X.; Liao, Y.F. A short-term wind speed forecasting model based on improved QPSO optimizing LSSVM. In Proceedings of the International Conference on Power System Technology (POWERCON), Chengdu, China, 20–22 October 2014.
16. Sun, B.; Yao, H.T. The short-term wind speed forecast analysis based on the PSO-LSSVM predict model. *Power Syst. Prot. Control* **2012**, *40*, 85–89.
17. Wang, J.Z.; Wang, Y.; Jiang, P. The study and application of a novel hybrid forecasting model—A case study of wind speed forecasting in china. *Appl. Energy* **2015**, *143*, 472–488. [[CrossRef](#)]
18. Cai, K.; Tan, L.N.; Li, C.L.; Tao, X.F. Short term wind speed forecasting combing time series and neural network method. *Power Syst. Technol.* **2008**, *32*, 82–85.
19. Yang, X.-S. A new metaheuristic bat-inspired algorithm. *Nat. Inspir. Coop. Strateg. Optim.* **2010**, *284*, 65–74.
20. Hafezi, R.; Shahrabi, J.; Hadavandi, E. A bat-neural network multi-agent system (BNNMAS) for stock price prediction: Case study of dax stock price. *Appl. Soft Comput.* **2015**, *29*, 196–210. [[CrossRef](#)]
21. Senthilkumar, P.; Vanitha, N.S. A unified approach to detect the record duplication using bat algorithm and fuzzy classifier for health informatics. *J. Med. Imaging Health Inform.* **2015**, *5*, 1121–1132. [[CrossRef](#)]
22. Yang, N.C.; Le, M.D. Optimal design of passive power filters based on multi-objective bat algorithm and pareto front. *Appl. Soft Comput.* **2015**, *35*, 257–266. [[CrossRef](#)]
23. Bao, Y.; Wang, H.; Wang, B.N. Short-term wind power prediction using differential EMD and relevance vector machine. *Neural Comput. Appl.* **2014**, *25*, 283–289. [[CrossRef](#)]
24. An, X.; Jiang, D.; Zhao, M.; Liu, C. Short-term prediction of wind power using EMD and chaotic theory. *Commun. Nonlinear Sci.* **2012**, *17*, 1036–1042. [[CrossRef](#)]
25. Wu, Z.H.; Huang, N.E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Adv. Adapt. Data Anal.* **2009**, *1*, 1–41. [[CrossRef](#)]
26. Wang, J.; Jiang, H.; Han, B.; Zhou, Q. An experimental investigation of FNN model for wind speed forecasting using EEMD and CS. *Math. Probl. Eng.* **2015**, *2015*. [[CrossRef](#)]
27. Lam, J.C.; Wan, K.K.W.; Cheung, K.L.; Yang, L. Principal component analysis of electricity use in office buildings. *Energy Build.* **2008**, *40*, 828–836. [[CrossRef](#)]
28. Ndiaye, D.; Gabriel, K. Principal component analysis of the electricity consumption in residential dwellings. *Energy Build.* **2011**, *43*, 446–453. [[CrossRef](#)]
29. Hong, Y.-Y.; Wu, C.-P. Day-ahead electricity price forecasting using a hybrid principal component analysis network. *Energies* **2012**, *5*, 4711–4725. [[CrossRef](#)]
30. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [[CrossRef](#)]
31. Jolliffe, I. *Principal Component Analysis*, 2nd ed.; Springer: New York, NY, USA, 2002; pp. 299–335.

32. Suykens, J.A.K.; Vandewalle, J. Recurrent least squares support vector machines. *IEEE Trans. Circuits Syst. I Fundam. Theory Appl.* **2000**, *47*, 1109–1114. [[CrossRef](#)]
33. Deng, J.L. *Introduction to Grey System Theory*; International Academic Publishers: Beijing, China, 1989; pp. 1–24.
34. Back, A.D.; Trappenberg, T.P. Selecting inputs for modeling using normalized higher order statistics and independent component analysis. *IEEE Trans. Neural Netw.* **2001**, *12*, 612–617. [[CrossRef](#)] [[PubMed](#)]
35. May, R.; Dandy, G.; Maier, H. Review of input variable selection methods for artificial neural networks. In *Artificial Neural Networks-Methodological Advances and Biomedical Applications*; Suzuki, K., Ed.; INTECH Open Access Publisher: Rijeka, Croatia, 2011; pp. 19–44.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).