

Article

# Short-Term Wind Electric Power Forecasting Using a Novel Multi-Stage Intelligent Algorithm

Haoran Zhao <sup>1,2</sup>, Huiru Zhao <sup>1,2</sup> and Sen Guo <sup>1,2,\*</sup>

<sup>1</sup> School of Economics and Management, North China Electric Power University, Beijing 102206, China; haoranzhao0118@163.com (Ha.Z.); zhaohuiru@ncepu.edu.cn (Hu.Z.)

<sup>2</sup> Beijing Key Laboratory of New Energy and Low-Carbon Development, North China Electric Power University, Changping, Beijing 102206, China

\* Correspondence: guosen@ncepu.edu.cn; Tel.: +86-158-1142-4568

Received: 18 February 2018; Accepted: 15 March 2018; Published: 20 March 2018

**Abstract:** As the most efficient renewable energy source for generating electricity in a modern electricity network, wind power has the potential to realize sustainable energy supply. However, owing to its random and intermittent instincts, a high permeability of wind power into a power network demands accurate and effective wind energy prediction models. This study proposes a multi-stage intelligent algorithm for wind electric power prediction, which combines the Beveridge–Nelson (B-N) decomposition approach, the Least Square Support Vector Machine (LSSVM), and a newly proposed intelligent optimization approach called the Grasshopper Optimization Algorithm (GOA). For data preprocessing, the B-N decomposition approach was employed to disintegrate the hourly wind electric power data into a deterministic trend, a cyclic term, and a random component. Then, the LSSVM optimized by the GOA (denoted GOA-LSSVM) was applied to forecast the future 168 h of the deterministic trend, the cyclic term, and the stochastic component, respectively. Finally, the future hourly wind electric power values can be obtained by multiplying the forecasted values of these three trends. Through comparing the forecasting performance of this proposed method with the LSSVM, the LSSVM optimized by the Fruit-fly Optimization Algorithm (FOA-LSSVM), and the LSSVM optimized by Particle Swarm Optimization (PSO-LSSVM), it is verified that the established multi-stage approach is superior to other models and can increase the precision of wind electric power prediction effectively.

**Keywords:** wind electric power prediction; Beveridge–Nelson decomposition approach; LSSVM; grasshopper optimization algorithm

## 1. Introduction

With the continuous emergence of global warming, smog weather, and other environmental problems, the development of the conventional thermal power generation mode, which is unsustainable and has large pollutant emissions, has been limited by the renewable energy power generation mode due to its beneficial characteristics. Having been considered as the most efficient renewable energy source for generating electricity in modern power systems, wind power has experienced a rapid expansion throughout the world [1–3]. The global wind installed capacity reached approximately 539 GW at the end of 2017, of which 53 GW was newly installed in the year of 2017 [4]. Specifically, China has equipped 188 GW, which accounts for 34.9% of global wind installed capacity. According to the report ‘China Energy Outlook 2030’ issued by the China Energy Research Association, China will strive to achieve 250 GW of wind installed capacity by 2020, accounting for 12.5% of total installed capacity, and 450 GW by 2030 with 900 billion kWh of on-grid electricity [5].

Despite of the environmental benefits of wind energy [6], its stochastic and intermittent nature [7] will pose lots of challenges in power system operations and large scale penetration.

Wind-power-related uncertainties are able to put the security and reliability of the electricity network at risk with the increasing penetration of wind power [8,9]. Due to the non-storage of electric power, the power supply and power demand must reach a real-time balance [10]. One of the possible approaches to solve the above issues is to forecast wind power over an extensive time period accurately. Precise wind power forecasting can decrease operation costs (such as re-dispatch cost) and enhance the technological convergence of wind power into the existing power network [11]. For market operators, the precise forecasting of wind power production in a zone is essential to the balance and security of the power network [12].

Hence, extensive efforts have been made to improve wind power prediction precision in the literature, among which Numerical Weather Prediction (NWP) approaches [13,14] as well as statistical approaches [15–17] are utilized to predict wind power output. Although several optimization algorithms are utilized to improve wind power prediction accuracy, wind energy predictions are still subject to high forecasting errors, in the scope of [8%, 22%], due to various factors, such as the type of forecasting model, the geographic location, and the forecasting horizon [18].

The primary devotion of this research is to put forward a multi-stage intelligent algorithm for wind electric power forecasting that combines the Beveridge–Nelson (B-N) decomposition approach, the Least Square Support Vector Machine (LSSVM) model, and a newly proposed intelligent optimization algorithm called the Grasshopper Optimization Algorithm (GOA). In the data preprocessing stage, the B-N decomposition approach is utilized to disintegrate the wind electric energy data into a deterministic trend, a periodic term, and a stochastic component. In the forecasting stage, the future data of the three components are computed by the LSSVM, of which the parameters ' $c$ ' and the Radial Basis Function (RBF) kernel width ' $\sigma$ ' are optimally selected by the newly proposed optimization algorithm GOA. Forecasting results illustrate that the proposed multi-stage intelligent algorithm can effectively enhance the precision of wind electric power prediction. It can be safely said that this study represents the first time that the B-N decomposition method and the GOA algorithm have been employed in wind power prediction.

The other parts of this research are planned as below. Section 2 summarizes current studies about wind speed and power forecasting. The basic theory of the B-N decomposition approach, the LSSVM model, and the GOA model is introduced in Section 3. Section 4 proceeds with the framework illustration of the proposed multi-stage intelligent algorithm. Numerical results of GOA-LSSVM are analyzed in Section 5. Section 6 evaluates the prediction performance among LSSVM, FOA-LSSVM, PSO-LSSVM, and GOA-LSSVM. Section 7 draws conclusions.

## 2. Overview of Current Wind Speed and Power Forecasting

Research in the field of wind power prediction has made great contributions to the stable operation of a power network. Many models have been developed to accurately forecast future wind power, and tools are classified into two main-stream approaches, which are the physical approach and the statistical method. Additionally, an integration approach is also utilized to combine the merits of both methods.

### 2.1. Physical Forecasting Approach

The physical prediction approach usually employs concrete physical characteristics to simulate the on-site items for a wind farm location [19]. The physical approach conducts the enhancement of the NWP data to consider the on-site requirements through a downscaling approach on the basis of the physics of the lower atmospheric bound layer. In the downscaling method, the wind farms and their surroundings, such as the wind turbine power curve, are required to be described in detail. Then, the enhanced wind speed data can be fed into the wind power curve to compute the wind power output. If on-line data can be accessed, the output statistics of the model can decrease the error of the prediction [10]. Different from the statistical method, historical data are not necessary for the physical approach. However, obtaining the physical data is the primary disadvantage of the approach.

Risoe National Laboratory in Denmark developed the *Prediktor*, which considers local conditions through using the NWP prediction from a High-Resolution Limited Area Model [20,21]. The University of Oldenburg in Germany developed the *Previento*, which has a similar physical approach but employs a different NWP forecast from the *Lakelmodell* of the German Weather Service [22]. AWS True Wind Inc. in the USA developed the eWind, which has a similar physical approach to the *Prediktor* but employs a high-resolution boundary layer model as a numerical weather model to consider the local conditions [23].

The physical methods are developed on the basis of models employing essential physical regulations for the conservation of momentum, mass, and energy in air flows. These models deal with computational fluid dynamics for imitating the atmosphere. Although many computational fluid dynamics models can be accessed, they are all based on the same fundamental physical theories. The differences between them include how the grids are structured and scaled and how the numerical computations are performed [10].

## 2.2. Statistical Forecasting Approach

The alternative mainstream approach utilized to predict wind power is the statistical forecasting approach on the basis of the historical meteorological values and wind power data. This kind of method usually elaborates the relationship between wind power or speed prediction and impact factors, such as NWP and on-site evaluated data. Representative statistical methods contain traditional statistical methods and artificial intelligence techniques.

For the traditional statistical methods, time-series approaches are always utilized to forecast wind electric power or wind speed. The Auto-Regressive Moving Average (ARMA) approach [24] and the Auto-Regressive Integrated Moving Average (ARIMA) approach [25,26] are widely used to forecast wind power. The Autoregressive Conditional Heteroscedastic (ARCH) approach, integrated with the ARIMA approach and taking the heteroscedasticity effect into consideration, is utilized to predict sub-wind speed series calculated by wavelet decomposition. Forecasting results, which are the sum of the prediction values, illustrate that it can improve prediction precision [27]. Conventional statistical approaches are developed on the basis of classical linear statistical models. However, wind power is usually a nonlinear function considering its input characteristics [3]. Therefore, artificial intelligence technologies have been developed to enhance the prediction precision of wind power.

Artificial intelligence techniques have been widely utilized in many research studies to forecast wind power output [28]. The artificial neural networks (ANN) method with Gaussian procedure simulation has been put forward for short-run wind power prediction [29]. A Back Propagation (BP) neural network on account of Particle Swarm Optimization (PSO) has been conducted for input parameter selection to decrease errors in wind speed prediction [30]. A reduced support vector machine (RSVM) optimized by PSO has been utilized to conduct wind speed forecasting with feature selection [31]. Additionally, recurrent neural networks [32], Multi-Layer Perceptron (MLP) neural networks [17], and RBF neural networks [33] have been put forward for the prediction of wind power. Although neural networks and support vector machines can mimic nonlinear input functions, the learning capacity of a single method with conventional learning mechanisms has been constrained. To handle with this issue, combinations of neural networks, support vector machines, and other methods with optimization algorithms and data processing mechanisms have been put forward.

## 2.3. Combination Approach

The combination approach aims at integrating different approaches and maintaining the advantages of each approach. Combined wind forecasting methods usually include weighting-based combined methods [34–37], combined methods containing data pre-processing techniques [38–40], combined methods consisting of parameter optimization techniques [41–44], and combined methods comprising error processing techniques [45]. The study [46] proposed two univariate statistical prediction models on the basis of LSSVM, which are the univariate LSSVM approach and the integrated

model combining the LSSVM model with ARIMA. An integrated model integrating input chosen by Wavelet Transform (WT) and Support Vector Machine (SVM) has been proposed and its superiority has been verified in terms of wind speed forecasting [47].

Considering the high forecasting accuracy of the combination approach, this paper proposes a new multi-stage intelligent algorithm for wind electric power prediction. During the data preprocessing step, the B-N disintegration approach is used to disintegrate the nonlinear and unstable wind electric power data series into a deterministic component, a cyclic term, and a stochastic impact effect. In previous research, the B-N decomposition method has been used to discuss the influence of a financial crisis on GDP and electricity consumption [48,49]. For this research, we explore the possibility of applying the B-N decomposition method to a wind power forecasting issue. Since the three decomposed sequences obtained from the B-N decomposition method can reflect the changing tendency of history and the future, they are utilized to predict future wind electric power. In the prediction process, a new nature-inspired optimization algorithm, named the Grasshopper Optimization Algorithm (GOA) and established by Seyedali Mirjalili in 2017 [50], is utilized to optimally select two parameters of the LSSVM approach to enhance the forecasting precision of wind electric power. After comparing this proposed method with a single LSSVM model, FOA-LSSVM, and PSO-LSSVM, it is verified that GOA-LSSVM has superior performance over other approaches. Therefore, the proposed multi-stage intelligent algorithm for wind electric power forecasting has high prediction accuracy.

### 3. Basic Theory of the Proposed Methodology

#### 3.1. B-N Decomposition Approach

For the purpose of decomposing a first-order differential stable data sequence, the B-N decomposition approach was put forward, which verified that a data sequence with a first-order stable instinct could be resolved to a permanent component and a transitory term. For the permanent component, it is made up of the deterministic component and the random term. For the transitory term, it has a stable procedure with a zero mean value named the cyclic trend [51–54].

According to the B-N decomposition model and the calculation requirements, it is obligatory to test if the data sequence is first-order stable before decomposing the data series. If it is verified that the data sequence has a first-order difference, the details of the procedures of the B-N disintegration approach are as below [49,51,52].

Suppose the natural logarithm data of the hourly wind electric power to be  $\ln P$  under the first-order stable condition. To calculate the deterministic trend in accordance with the Wold theorem, we set:

$$\Delta \ln P_t = \mu + \varepsilon_t + \sum_{i=1}^{\infty} \lambda_i \varepsilon_{t-i} \quad (1)$$

where  $\Delta \ln P_t = \ln P_t - \ln P_{t-1}$ ,  $P_t$  demonstrates the wind electric power at time  $t$ ,  $\mu$  indicates the average data of  $\Delta \ln P_t$  in the long-term,  $\varepsilon_t \sim i.i.d.N(0, \sigma^2)$  (*i.i.d.* denotes independently and identically distributed),  $t$  implies the time point, and  $\lambda_i$  means the coefficient. The anticipated value of Equation (1) is:

$$E(\Delta \ln P_t) = E(\mu) + E(\varepsilon_t) + E\left(\sum_{i=1}^{\infty} \lambda_i \varepsilon_{t-i}\right) = E(\mu) \quad (2)$$

where  $E$  demonstrates the computational procedure of the anticipated value for every variable.

Since  $\Delta \ln P_t$  means the natural logarithm value of the wind electric power, its first-order difference indicates the proportion of increase of the wind electric power. In accordance with Equation (2), the mean value of the proportion of increase of wind electric power is the long-run proportion of increase of wind electric power. Based on the B-N disintegration logic, the deterministic component  $DT_t$  is disintegrated as [51]:

$$DT_t = \ln P_0 + \mu t \quad (3)$$

where  $DT_t$  illustrates the determinacy component at time  $t$ , and  $\ln P_0$  denotes the original data of the natural logarithm data of the wind electric power.

Moreover,  $\ln P_t$  is identified as the prediction value in terms of current information. Morley [55] noted that this data sequence was much more precisely forecasted through utilizing a stable univariate Auto Regressive (AR) Equation (1) method for its first differences, called:

$$(\Delta \ln P_t - \mu) = \varphi(\Delta \ln P_{t-1} - \mu) + \varepsilon_t \tag{4}$$

where  $\varphi$  is calculated by AR Equation (1) and  $|\varphi| < 1$ . Through considering the demonstrated Wold form in AR Equation (1), the minimum mean squared error (MMSE) of the  $j$ -period prediction of the first difference of  $\Delta \ln P_t$  is:

$$E_t[(\Delta \ln P_t - \mu)] = \varphi^j(\Delta \ln P_{t-1} - \mu). \tag{5}$$

Consistent with the disintegration identification of the B-N approach [51],  $T_t$ , the total tendency component of the data sequence, is determined as the MMSE forecasting of the long-run level for the sequence, which is computed as:

$$T_t = \lim_{j \rightarrow \infty} E_t[\ln P_{t+j} - j\mu] = \ln P_t + \lim_{j \rightarrow \infty} E_t[\Delta \ln P_{t+j} - \mu]. \tag{6}$$

Hence, combining Equations (5) and (6), and employing the infinite sum expression of geometric succession,  $T_t$  is computed as:

$$T_t = \ln P_t + \varphi / (1 - \varphi) * (\Delta \ln P_t - \mu). \tag{7}$$

Meanwhile, the cyclic tendency of  $\ln P_t$  is calculated as:

$$C_t = -\varphi / (1 - \varphi) * (\Delta \ln P_t - \mu) \tag{8}$$

where  $C_t$  indicates the cyclic tendency at time  $t$ .

The stochastic shock component is computed based on Equations (3), (7), and (8):

$$ST_t = \ln P_t - C_t - DT_t = \ln P_t + \varphi / (1 - \varphi) * (\Delta \ln P_t - \mu) - (\ln P_0 + \mu t) \tag{9}$$

where  $ST_t$  illustrates the stochastic component at time  $t$ .

### 3.2. The LSSVM Approach

LSSVM is an enhanced approach of SVM that alleviates the convex quadratic programming related to SVM [56–58]. LSSVM owns the merits of making slackness through an equality restraint as well as handling the regression issue as linear equations, which leads to faster training and higher precision and stability [59–61]. The fundamental methodology of the LSSVM approach is elaborated as follows [56,60].

Suppose a sequence of samples  $\{x_i, y_i\}_{i=1}^m$ , regarding  $x_i \in R_n$  as the input variable and  $y_i \in R_n$  as the outcome data for sample  $i$ . By utilizing a nonlinear function  $\varphi$ , the sample data are programmed to a high dimensional space, hence, to roughly depict it in a linear formula:

$$y_i = w^T \varphi(x) + b \tag{10}$$

where  $w$  indicates the weight vector and  $b$  implies the error.

In the initial space, LSSVM with an equality restraint is phrased as

$$\begin{cases} \min J(w, \xi) = \frac{1}{2} w^T w + \frac{1}{2} C \sum_{i=1}^m \xi_i^2 \\ y_i = w \varphi(x_i) + b + \xi_i \end{cases} \tag{11}$$

where  $C$  represents the regularization parameter and  $\xi_i$  means the slackness variable.

The Lagrangian function  $L$  can be adopted [62]:

$$L(w, b, \xi, a) = \frac{1}{2}w^T w + \frac{1}{2}C \sum_{i=1}^m \xi_i^2 - \sum_{i=1}^m a_i \{w^T \phi(x_i) + b + \xi_i - y_i\} \quad (12)$$

where  $a_i$  demonstrates Lagrange multiplier.

The requirements of Karush–Kuhn–Tucker (KKT) for optimality are decided as

$$\begin{cases} \frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^m a_i \phi(x_i) \\ \frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^m a_i = 0 \\ \frac{\partial L}{\partial \xi_i} = 0 \rightarrow a_i = C \xi_i \\ \frac{\partial L}{\partial a_i} = 0 \rightarrow w^T \phi(x_i) + b + \xi_i - y_i \end{cases} \quad (13)$$

Eliminating the variables  $W$  and  $\xi_i$ , the optimized issue is converted into a linear equation as:

$$\begin{bmatrix} 0 & Q^T \\ Q & K + C^{-1}I \end{bmatrix} \begin{bmatrix} b \\ A \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (14)$$

where  $Q = [1, \dots, 1]^T$ ,  $A = [a_1, a_2, \dots, a_m]^T$ , and  $Y = [y_1, y_2, \dots, y_m]^T$ . Consistent with the Mercer's term, the Kernel function is written as:

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j). \quad (15)$$

Then, the LSSVM for regression is introduced as:

$$f(x) = \sum_{i=1}^m a_i K(x, x_i) + b. \quad (16)$$

As the RBF has fewer parameters to be assumed, it is chosen to be  $K(x, x_i)$  in this study and illustrated as:

$$K(x, x_i) = \exp\left\{-\|x - x_i\|^2 / 2\sigma^2\right\}. \quad (17)$$

Consequently, two parameters are required to be identified by a complicated search process, which are parameters ' $c$ ' and ' $\sigma$ ' [63,64]. The latest optimization algorithm GOA is utilized to identify the best values of these parameters.

### 3.3. Grasshopper Optimization Algorithm (GOA)

Inspired by different movement characteristics in larval and adult phases of grasshoppers, a novel optimization approach has been put forward by researcher Seyedali Mirjalili, namely the grasshopper optimization algorithm (GOA) [50]. In the nymph phase, the primary characteristic of a grasshopper swarm is slow movement and small steps, and they jump and move like rolling cylinders. Along their movement path, almost all vegetation can be eaten. Then, they step into the adult phase, in which long range and abrupt movement is the necessary characteristic of the grasshopper swarm [65,66]. The searching procedure of the GOA is divided into two types: exploration and exploitation. In exploitation, the search agents shift locally, while they tend to move abruptly in exploration [67]. Since a grasshopper performs these two functions naturally, it is of great significance to mathematically imitate this behavior and use the new nature-inspired optimization method to obtain optimal results.

According to the behaviors of grasshoppers, the GOA is composed of five phases listed as below.

**Step 1: Parameters setting.**

The primary parameters of the GOA comprise the number of grasshoppers  $SearchAgents\_no$ ; the number of variables  $dim$ ; the maximum number of iterations  $Max\_iteration$ ; the bottom limit  $lb = [lb_1, lb_2, \dots, lb_n]$ , and the upper limit  $ub = [ub_1, ub_2, \dots, ub_n]$  of variables.

**Step 2: Population initialization.**

The original population of grasshoppers in the GOA is set in a matrix as below:

$$G = \begin{bmatrix} g_{1,1} & g_{1,2} & \cdots & g_{1,d} \\ g_{2,1} & g_{2,2} & \cdots & g_{2,d} \\ \vdots & \vdots & \cdots & \vdots \\ g_{n,1} & g_{n,2} & \cdots & g_{n,d} \end{bmatrix} \tag{18}$$

where  $G$  indicates the position matrix of the grasshoppers,  $g_{ij}$  implies the  $j$ -th parameter's (dimension or variable) value of the  $i$ -th grasshopper, and  $i = 1, 2, \dots, n, j = 1, 2, \dots, d$ , and  $g_{ij}$  are computed through using the stochastic distribution listed in Equation (19).

$$G(i, j) = (ub(i) - lb(i)) \times rand() + lb(i) \tag{19}$$

where  $G(i, j)$  denotes the data of the  $i$ -th row and  $j$ -th column in the matrix,  $ub(i)$  and  $lb(i)$  respectively illustrate the upper limit and bottom limit of the  $i$ -th grasshopper, and  $rand()$  denotes the random value from a uniform distribution in the scope of  $[0, 1]$ .

**Step 3: Fitness function identification.**

For the purpose of assessing every grasshopper, the fitness function  $f[*]$  is required to be decided in an optimization procedure, and the matrix  $OG$  is applied to save the fitness data for grasshoppers described as below.

$$OG = \begin{bmatrix} OG_1 \\ OG_2 \\ \vdots \\ OG_n \end{bmatrix} = \left\{ \begin{array}{l} f[g_{11}, g_{12}, \dots, g_{1d}] \\ f[g_{21}, g_{22}, \dots, g_{2d}] \\ \vdots \\ f[g_{n1}, g_{n2}, \dots, g_{nd}] \end{array} \right\} \tag{20}$$

where  $OG$  represents the matrix for storing the fitness values of grasshoppers and  $n$  implies the amount of grasshoppers.

**Step 4: Iteration start.**

The moving behavior of grasshoppers is imitated by the following mathematical model [68]:

$$X_i = S_i + G_i + A_i \tag{21}$$

where  $X_i$  represents the  $i$ -th grasshopper's position,  $S_i$  means the social connection,  $G_i$  indicates the gravity force on the  $i$ -th grasshopper, and  $A_i$  implies the wind advection. On the purpose of providing random behavior, the equation is described as  $X_i = r_1 S_i + r_2 G_i + r_3 A_i$ , where  $r_1, r_2$ , and  $r_3$  are random digits in  $[0, 1]$ .

$$S_i = \sum_{\substack{j=1 \\ j \neq i}}^N s(d_{ij}) \hat{d}_{ij} \tag{22}$$

where  $d_{ij}$  represents the distance from the  $i$ -th to the  $j$ -th grasshopper, which can be computed as  $d_{ij} = |x_j - x_i|$ ,  $s$  implies a function employed to identify the intensity of social force, which is displayed in Equation (23), and  $\hat{d}_{ij} = \frac{x_j - x_i}{d_{ij}}$  denotes a unit vector from the  $i$ -th grasshopper to the  $j$ -th grasshopper.

The  $s$  function that identifies the social force is computed as below:

$$s(r) = f e^{-r} - e^{-r} \tag{23}$$

where  $f$  implies the attraction strength and  $l$  means the attractive length scale.

Figure 2 in Reference [50] illustrates how function  $s$  affects the social connection of grasshoppers and distances ranging from 0 to 15 are taken into account. Repulsion happens in the scope of  $[0, 2.079]$ . If a grasshopper keeps 2.079 units away from another one, there exists neither a repulsion nor an attraction force, which is named the comfort zone or comfortable distance. It can also be seen that the attraction strength raises from 2.079 units of distance to nearly 4 and reduces to 0. Alternating the parameters' values of  $l$  and  $f$  in Equation (23) leads to various social behaviors in artificial grasshoppers. The function  $s$  is re-drawn with various values of  $l$  and  $f$  independently. That figure illustrates that these two parameters change the attraction zone, comfort area, and repulsion area obviously.

Although the function  $s$  can classify the space between two grasshoppers into an attraction area, a comfort area, and a repulsion area, this function outputs values on the verge of zero with distances longer than 10. Therefore, this function cannot employ strong forces between grasshoppers with long distances. In order to deal with this problem, the distances between grasshoppers should be mapped in the interval of  $[1, 4]$ .

The  $G_i$  component in Equation (21) can be calculated as

$$G_i = -g\hat{e}_g \quad (24)$$

where  $g$  indicates the constant of earth gravitation, and  $\hat{e}_g$  represents a unity vector to the center of earth.

The  $A_i$  term in Equation (21) can be computed as

$$A_i = u\hat{e}_w \quad (25)$$

where  $u$  is constant drift and  $\hat{e}_w$  is a unit vector for the wind direction. Since larval grasshoppers do not have wings, their movements rely on wind direction to a great extent.

Substituting Equations (22), (24) and (25) into Equation (21),  $X_i$  is described as

$$X_i = \sum_{\substack{j=1 \\ j \neq i}}^N s(|x_j - x_i|) \frac{x_j - x_i}{d_{ij}} - g\hat{e}_g + u\hat{e}_w \quad (26)$$

where  $s(r) = fe^{-\frac{r}{l}} - e^{-r}$  and  $N$  represent the amount of grasshoppers.

However, this mathematical model brings the grasshoppers to achieve the comfort area quickly and the swarm does not verge on a particular point. Therefore, a revised model of this equation is put forward as below:

$$X_i^d = c \left( \sum_{\substack{j=1 \\ j \neq i}}^N c \frac{ub_d - lb_d}{2} s(|x_j^d - x_i^d|) \frac{x_j - x_i}{d_{ij}} \right) + \hat{T}_d \quad (27)$$

where  $ub_d$  and  $lb_d$  represent the superior and bottom limit in the  $D$ -th dimension,  $s(r) = fe^{-\frac{r}{l}} - e^{-r}$ ,  $\hat{T}_d$  indicates the data of the  $D$ -th dimension in the target, and  $c$  implies a reducing coefficient to lessen the repulsion region, comfort area, and attraction region. It should be pointed out that  $S$  is similar to the  $S_i$  section in Formula (21), and earth gravity has not been considered (not containing  $G_i$  term). It is assumed that the wind direction ( $A_i$  part) is towards a target ( $\hat{T}_d$ ).

For Equation (27), the first term (the sum) takes other grasshoppers' position into account. The second term ( $\hat{T}_d$ ) mimics their trends to move towards the food. The inner  $c$  is devoted to the decrease of attraction or repulsion strength between grasshoppers rational to the generation number, while the outer  $c$  shrinks the search coverage around the *target* with the increasing of the generation

count. For the aim of balancing exploration and exploitation, the parameter  $c$  should be reduced to be rational to the generation number. This procedure encourages exploitation with the increasing of generation. The coefficient  $c$  decreases the comfort area rational to the iteration number and can be computed as

$$c = c_{\max} - l \frac{c_{\max} - c_{\min}}{L} \quad (28)$$

where  $c_{\max}$  indicates the maximum data,  $c_{\min}$  implies the minimum data,  $l$  represents the present generation, and  $L$  denotes the total amount of generations. In this method, 1 and 0.00001 are employed for  $c_{\max}$  and  $c_{\min}$ .

#### Step 5: Optimal selection.

In this mathematical model, grasshoppers are required to move towards a target gradually over a number of iterations. However, in a real search space, the global optimum is exactly unknown. Hence, it is necessary to search for a target for grasshoppers in each step of a generation. In the GOA, the fittest grasshopper (the one with the best objective value) is assumed to be the target, which can help the GOA to store the most promising target in each generation and require grasshoppers to move towards the target.

During the iteration, the positions of search agents will be updated based on Equation (27). Additionally, the best target position obtained to date is updated in every generation. Positions are required to be updated iteratively until they satisfy the terminal criterion. The location and fitness value of the target are ultimately feedback as the best solution for the global optimum.

### 4. The Proposed Multi-Stage Intelligent Algorithm for Predicting Wind Electric Power

A novel multi-stage intelligent algorithm is put forward to improve wind electric power prediction accuracy. Since hourly wind electric power data is non-stationary with periodic changing tendency and under the interference of random factors, the B-N disintegration method is utilized to disintegrate the virginal hourly wind electric power data into a deterministic trend, a cyclic component, and a stochastic impact effect. Then, the GOA-LSSVM approach is employed to predict three disintegrated data sequences, and final prediction wind electric power data are calculated. For LSSVM, parameters ' $\sigma$ ' and ' $c$ ' are required to be decided before predicting three decomposed data sequences. To enhance the prediction precision, the GOA is applied to identify the optimum values of them.

The details of the procedures of the proposed multi-stage intelligent algorithm for wind electric power prediction are summarized as follows.

#### Step 1: Unit root test.

In accordance with the basic rules of the B-N disintegration approach, we need to test if the logarithmic sequence of the original data sequence is first-order stable. The Augmented Dickey–Fuller (ADF) approach is utilized to carry out the stability examination.

#### Step 2: Initial data disintegration.

After confirming the first-order stability of the original data's logarithmic sequence, the initial data are disintegrated into a deterministic trend, a cyclic component, and a random impact effect according to Equations (3), (8) and (9).

#### Step 3: Set Parameters.

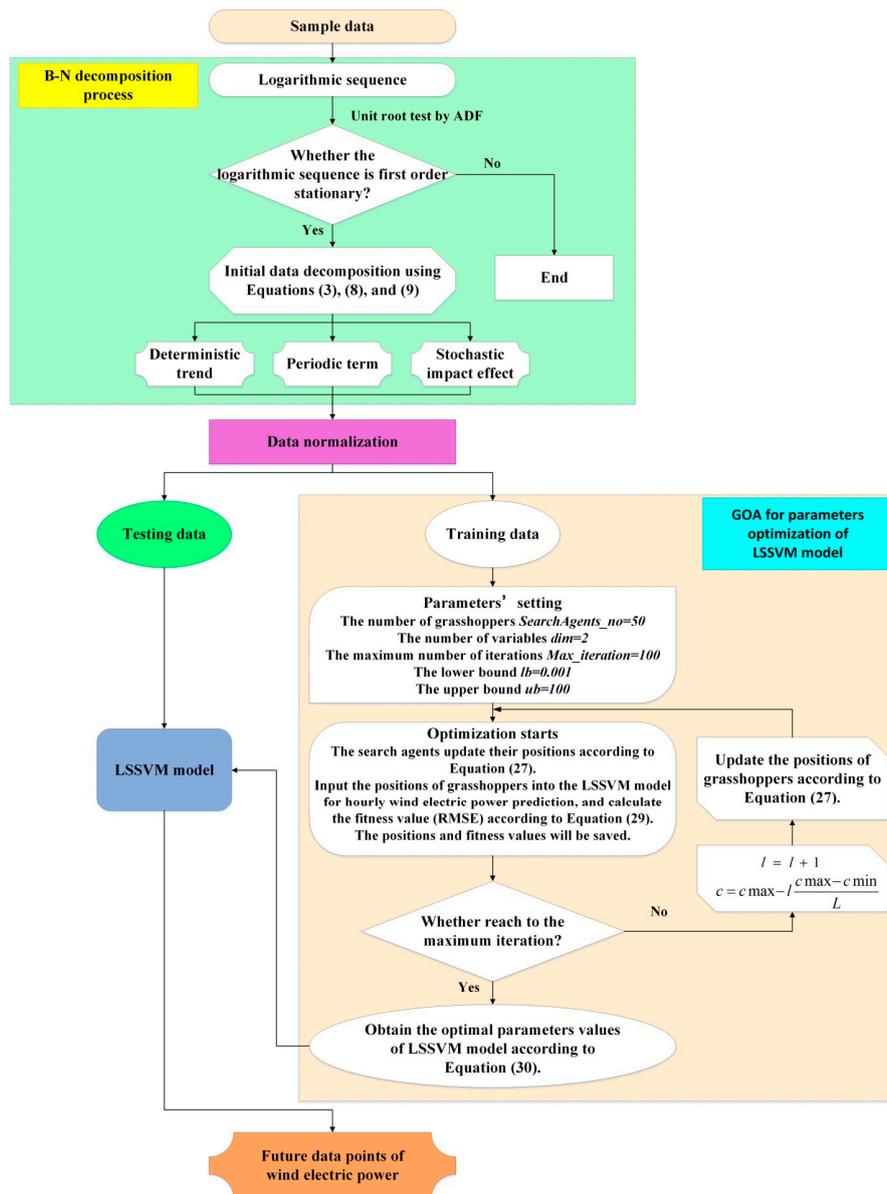
In the GOA, five parameters are required to be set, embodying the amount of grasshoppers  $SearchAgents\_no$ ; the amount of variables  $dim$ ; the maximum amount of iterations  $Max\_iteration$ ; and the bottom limit  $lb = [lb_1, lb_2, \dots, lb_n]$  and the upper limit  $ub = [ub_1, ub_2, \dots, ub_n]$  of variables. We set  $SearchAgents\_no = 50$ ,  $dim = 2$ ,  $Max\_iteration = 100$ ,  $lb = 0.001$ , and  $ub = 100$ .

#### Step 4: Optimization starts.

Through employing the GOA to identify the optimum values of the two parameters in LSSVM, the fitness function  $f[*]$  is required to be decided. The Root Mean Square Error (RMSE) expressed as Equation (29) is applied to compose the fitness function.

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (x(k) - \hat{x}(k))^2} \tag{29}$$

where  $x(k)$  is the objective data of wind electric power at time  $k$  and  $\hat{x}(k)$  is the prediction data of wind electric power at time  $k$ .



**Figure 1.** The procedure of the established multi-stage intelligent algorithm for wind electric power forecasting. B-N: Beveridge–Nelson; ADF: augmented Dickey–Fuller; GOA: grasshopper optimization algorithm; LSSVM: least squares support vector machine; RMSE: root mean square error.

The GOA optimizes the two parameters in LSSVM through generating a set of stochastic solutions. Search agents renovate their locations in terms of Equation (27). The position of the best target found

to date is updated in every generation. The distances between grasshoppers need to be normalized in the interval of [1, 4] in every iteration. The locations of grasshoppers are employed to imply two parameters in LSSVM, namely  $g_{i,1} = \sigma$  and  $g_{i,2} = c$ . Supposing that the objective data sequence  $\{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$  is employed in the first generation, the fitting sequence  $\{\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(n)\}$  is computed in terms of the LSSVM. Then, the fitness function is identified minimizing the RMSE of the prediction value, written by

$$RMSE = \min \sqrt{\frac{1}{n} \sum_{k=1}^n (x(k) - \hat{x}(k))^2}. \quad (30)$$

Positions and fitness values will be saved. Then, the remaining 99 iterations can be performed in accordance with the optimization procedure of the GOA.

#### Step 5: Optimization termination.

Through the whole optimization procedure, various values of RMSEs are computed with varying parameter values, and the minimum RMSEs can be discovered when the optimization procedure is finished. The optimum values of ' $\sigma$ ' and ' $c$ ' can be gained through using the GOA in terms of Equation (30). Then, the optimized LSSVM can be established, and the future data points of wind electric power can be predicted.

The procedure of the proposed multi-stage intelligent algorithm for wind electric power prediction is illustrated in Figure 1.

## 5. Numerical Results and Analysis for GOA-LSSVM

### 5.1. B-N Decomposition Results

In this research, hourly wind electric power data of one province in Northwest China will be fed into the proposed multi-stage intelligent algorithm for future wind electric power forecasting. According to the available data, 216 h (9 days) data will be treated as the training sample, and 168 h (7 days) data will be managed to be the testing sample to measure the performance of the established model.

Firstly, the B-N decomposition model should be applied to decompose the logarithmic sequence of the initial 384 ( $384 = 216 + 168$ ) h wind electric power data. Before decomposition, it is necessary to carry out a unit root test in accordance with the ADF test to examine if the logarithmic series is first-order stationary. As listed in Table 1, the data sequence of the wind electric power is stable after a first-order difference, which satisfies the requirement of the B-N disintegration approach that the data sequence is first-order co-integrated. Therefore, Equations (3), (8) and (9) can be employed to decompose the logarithmic sequence of original 384 h wind electric power data, and then the deterministic component, cyclic term, and stochastic component in logarithmic form are obtained. Since the deterministic term, cyclic component, and random term are in logarithmic form, we need to transform them into natural numbers to apply the GOA-LSSVM model. The decomposition results of these three terms in natural numbers are displayed in Figures 2 and 3. The deterministic trend means that the time sequence increases steadily with the change of time. The periodic term indicates the periodic variation of the time series. The stochastic component implies the impact of unexpected and unobservable random factors on the time series.

**Table 1.** The ADF test results of the logarithmic series of the original 384 h wind electric power data.

Series	Test Form (C,T,K)	ADF Test Value	p Value	Conclusion
lnP	(N,N,1)	-1.336253	0.1681	Unstable
$\Delta \ln P$	(N,N,0)	-12.94301	0.0000	stable

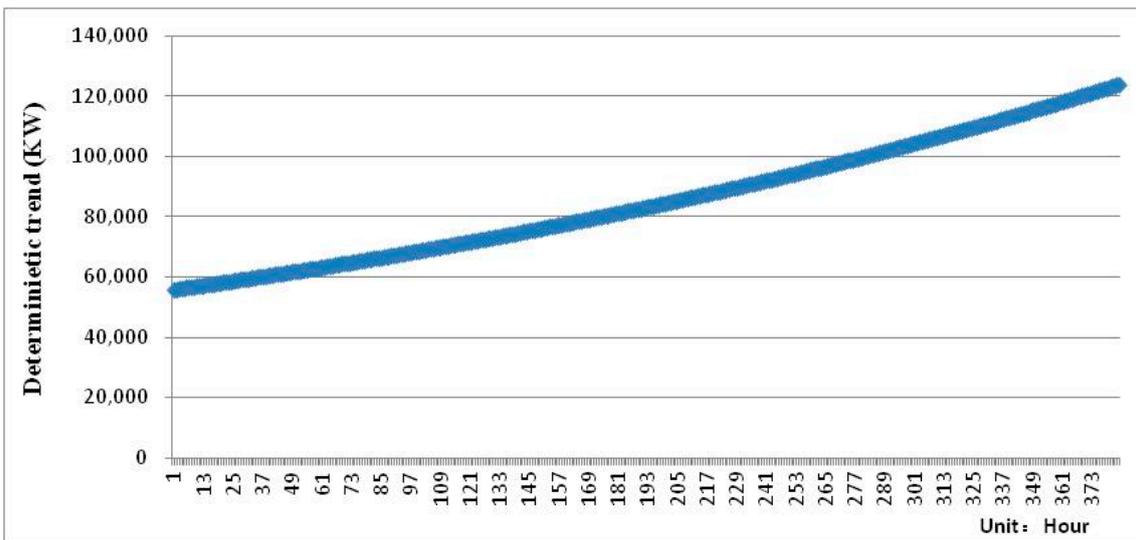


Figure 2. Disintegration results of deterministic term.

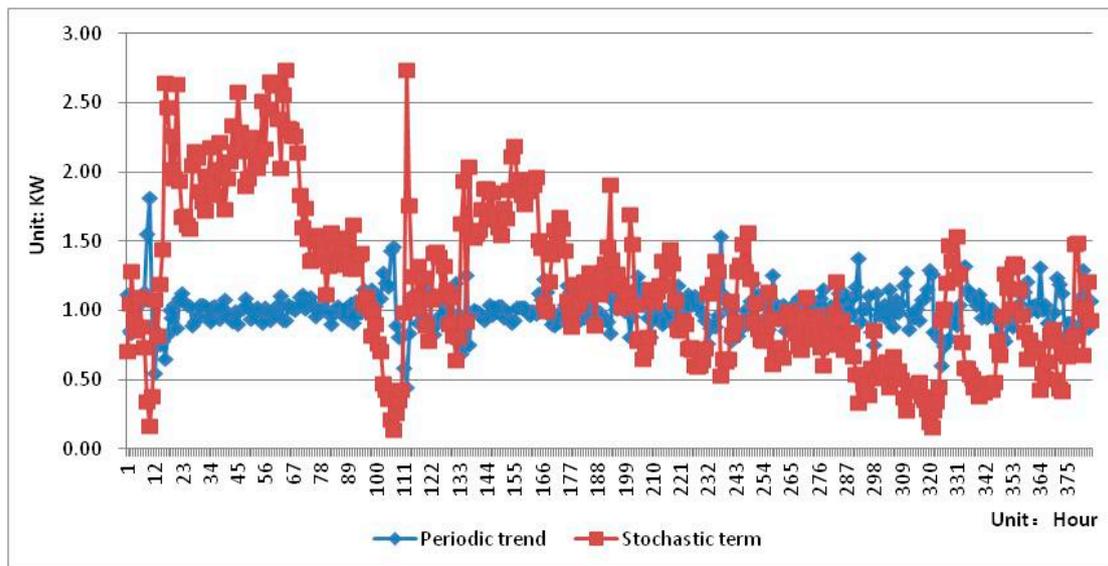


Figure 3. Disintegration results of cyclic component and stochastic component.

5.2. GOA-LSSVM Prediction Results

The GOA-LSSVM approach is employed to predict the three components, respectively, and the final wind electric power value is obtained by multiplying the predicted deterministic trend, cyclic component, and random trend. Based on the data available, through using the GOA-LSSVM model to forecast the deterministic trend, the deterministic trend in the last moment and the deterministic trend at the same time yesterday will be fed into GOA-LSSVM. By utilizing the GOA-LSSVM model to forecast the periodic component, the periodic component in the last moment and the periodic component at the same time yesterday are handled as the input items for GOA-LSSVM. By utilizing the GOA-LSSVM approach to forecast the stochastic trend, the stochastic trend in the last moment and the stochastic trend at the same time yesterday are managed as the input terms for GOA-LSSVM.

The forecasting procedure of GOA-LSSVM should start with normalizing the sample data in the scope of  $[0, 1]$  utilizing the following formula:

$$z = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (31)$$

where  $x_{\min}$  and  $x_{\max}$  denote the minimum and maximum data of every input sequence.

In the GOA-LSSVM approach, the parameters ' $c$ ' and ' $\sigma$ ' of the LSSVM approach will be identified by the GOA method. Through the optimization process, the optimal values of  $\sigma^2$  and  $c$  of LSSVM for determinacy component prediction, cyclic term prediction, and random component prediction are determined by Equation (30), and the optimum values of them are illustrated in Table 2.

In the testing phrase, the optimum values of  $\sigma^2$  and  $c$  will be applied to calculate the forecasting values of the deterministic trend, the periodic component, and the stochastic term, respectively. Through multiplying these three terms, the final wind electric power values are obtained. The values of three error criteria are applied to discuss the prediction performance, which are the Root Mean Square Error ( $RMSE$ , computed in accordance with Equation (29)), the Normalized- $RMSE$  calculated based on [69,70], and the Mean Absolute Percentage Error ( $MAPE$ ) calculated as below:

$$MAPE = \frac{1}{n} \sum_{k=1}^n \left| \frac{x(k) - \hat{x}(k)}{x(k)} \right| \times 100\% \quad (32)$$

where  $x(k)$  indicates the objective data at time  $k$  and  $\hat{x}(k)$  means the prediction data at time  $k$ .

The values of  $RMSE$ , normalized- $RMSE$ , and  $MAPE$  for GOA-LSSVM of each day through calculating the average hourly  $RMSE$ , normalized- $RMSE$ , and  $MAPE$  values are shown in Table 3. The value of  $MAPE$  on the first day is the highest, which is 9.54%, the value of  $MAPE$  on the third day is the lowest, which is 1.52%, and the values of  $MAPE$  on other days range from 3% to 4.6%. The average value of  $MAPE$  is 4.28%, which is much lower than the prediction error in the current literature, which ranges from 8% to 22% [18]. For the normalized- $RMSE$  values, all of them are less than 10%, which implies that the established model has excellent performance in wind electric power prediction [70]. This proves that the GOA-LSSVM can effectively improve the forecasting accuracy.

**Table 2.** Best values of  $\sigma^2$  and  $c$  of LSSVM for determinacy component prediction, cyclic term prediction, and random component prediction.

Parameters	LSSVM for Deterministic Trend	LSSVM for Periodic Term	LSSVM for Stochastic Trend
$\sigma^2$	1.0581	1.2414	8.8182
$c$	95.053	94.877	95.452

**Table 3.** The values of  $RMSE$ , normalized- $RMSE$ , and mean absolute percentage error ( $MAPE$ ) of each day.

Error Criteria	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Average Value
$RMSE(KW)$	7812.16	4600.74	1591.12	2657.38	3100.32	3828.40	3546.86	3876.71
Normalized- $RMSE$	9.68%	5.05%	1.87%	4.63%	4.21%	4.32%	3.84%	4.80%
$MAPE$	9.54%	4.07%	1.52%	3.78%	4.58%	3.05%	3.37%	4.28%

## 6. Forecasting Performance Assessment

To assess the prediction performance, three compared prediction models have been chosen, including LSSVM, FOA-LSSVM, and PSO-LSSVM. For these three models, the input variables and the output variable are the same as GOA-LSSVM, while the difference among them is the parameters' selection mechanism.

For the LSSVM model, the deterministic trend, periodic component, and stochastic term will be forecasted taking the term's value in the last moment and the component's value at the same

time yesterday for each trend from 1 to 216 h as input variables. Values of parameters  $\sigma^2$  and  $c$  are identified [71], which are listed in Table 4. Additionally, the hourly wind electric power for the future 168 h can be forecasted.

**Table 4.** Values of  $\sigma^2$  and  $c$  for different compared models.

Compared Models	LSSVM		FOA-LSSVM		PSO-LSSVM	
	$\sigma^2$	$c$	$\sigma^2$	$c$	$\sigma^2$	$c$
Deterministic component	0.1882	188.21	0.1638	128.93	0.1627	165.13
Periodic term	0.1694	130.57	0.1778	142.03	0.1831	125.72
Stochastic trend	0.1275	130.79	0.1989	120.50	0.1427	151.29

For the Fruit-fly Optimization Algorithm (FOA)-LSSVM approach, the parameters  $\sigma^2$  and  $c$  of LSSVM for determinacy trend prediction, cyclic component prediction, and random component prediction are optimally selected by FOA. Before iteration optimization, the initial parameters of FOA are assumed to be:  $maxgen = 100$ ,  $sizepop = 20$ ,  $(X\_axis, Y\_axis) \subset [-50, 50]$ ,  $FR \subset [-10, 10]$ ,  $X\_axis = rands(1, 2)$ , and  $Y\_axis = rands(1, 2)$ . The optimum values of  $\sigma^2$  and  $c$  selected by FOA for these three terms' forecasting are shown in Table 4.

For the PSO-LSSVM approach, the parameters  $\sigma^2$  and  $c$  of LSSVM for determinacy trend prediction, cyclic component prediction, and random term prediction are iteratively identified by PSO. Before optimization, the original parameters of PSO are supposed to be:  $maxgen = 300$ ,  $swarm\ size = 30$ ,  $particle\ size = 2$ , the minimum of the particle = [0.01, 1000], the maximum of the particle = [0.1, 1000], the minimum of velocity = -500, the maximum of velocity = 500, and learning factors  $c1 = 1.5$  and  $c2 = 1.7$ . The optimum  $\sigma^2$  and  $c$  selected by PSO for these three terms' forecasting are demonstrated in Table 4.

According to the optimum  $\sigma^2$  and  $c$  for determinacy component prediction, cyclic term forecasting, and random component prediction in different models, the future 168 h wind electric power values are calculated through multiplying the prediction data of the three components. The MAPE values, normalized RMSE values, and RMSE values of each day through calculating the average hourly MAPE, normalized RMSE, and RMSE values for various approaches are computed to assess the prediction performance of the various prediction approaches and are illustrated in Table 5. The LSSVM model has the poorest performance compared to other models with optimization algorithms, and the average value of MAPE is 15.61%, which is the only one larger than 10%. The GOA-LSSVM model is superior to the other models with a 4.28% average MAPE value; except for the MAPE value on the first day, the MAPE values on other days are all smaller than 4.6%, specifically, the MAPE value on the third day is only 1.52%. The forecasting performance of the PSO-LSSVM model ranks second with an 8.00% average MAPE value; however, the MAPE values on the last two days are all more than 10%, which is much higher than the values of GOA-LSSVM. The forecasting performance of the FOA-LSSVM approach ranks third with a 9.89% average MAPE value. For the RMSE values, the average values of the LSSVM model and the FOA-LSSVM model are larger than 10,000 KW, which indicate that there exists a large gap between the forecasted data and the objective data of wind electric power. The smallest average RMSE value is 3876.71 KW, which belongs to the GOA-LSSVM model, followed by 8046.46 KW of the PSO-LSSVM model. For the normalized-RMSE values, the average values of GOA-LSSVM and PSO-LSSVM are both less than 10%, which demonstrate that both of these two models have excellent performance in wind electric power prediction, but the values on Day 1, Day 6, and Day 7 of PSO-LSSVM are all larger than 10%, which are much greater than GOA-LSSVM. Most of the normalized-RMSE values of LSSVM and FOA-LSSVM are larger than 10%, which verify that the established approach effectively enhances the prediction accuracy.

**Table 5.** Values of *MAPE*, *RMSE*, and normalized-*RMSE* (*N-RMSE*) of each day for different models.

Compared Models		Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Average Value
GOA-LSSVM	<i>MAPE</i>	9.54%	4.07%	1.52%	3.78%	4.58%	3.05%	3.37%	4.28%
	<i>RMSE</i>	7812.16	4600.74	1591.12	2657.38	3100.32	3828.4	3546.86	3876.71
	<i>N-RMSE</i>	9.68%	5.05%	1.87%	4.63%	4.21%	4.32%	3.84%	4.80%
LSSVM	<i>MAPE</i>	23.11%	9.24%	6.21%	15.24%	29.60%	14.04%	11.84%	15.61%
	<i>RMSE</i>	19,684.52	10,271.47	6863.51	9909.27	16,500.29	12,079.6	13,223.24	12,647.41
	<i>N-RMSE</i>	24.39%	11.27%	8.05%	17.26%	22.43%	13.62%	14.31%	15.90%
FOA-LSSVM	<i>MAPE</i>	13.80%	8.12%	7.96%	9.30%	5.40%	8.87%	15.81%	9.89%
	<i>RMSE</i>	13,873.89	8971.39	7819.88	9659.2	4418.03	9693.03	17,567.19	10,286.09
	<i>N-RMSE</i>	17.19%	9.84%	9.18%	16.82%	6.01%	10.93%	19.01%	12.71%
PSO-LSSVM	<i>MAPE</i>	7.79%	3.55%	3.27%	4.88%	7.42%	16.23%	12.88%	8.00%
	<i>RMSE</i>	10,484.87	4087.68	3680.11	3988.71	6726.3	13,669.15	13,688.39	8046.46
	<i>N-RMSE</i>	12.99%	4.48%	4.32%	6.95%	9.14%	15.41%	14.81%	9.73%

Compared with PSO-LSSVM and FOA-LSSVM, GOA-LSSVM has a better forecasting performance. This is because in the GOA the next location of a grasshopper is identified on the basis of its current location, the global best, and the location of all other grasshoppers according to Equation (27), while PSO, as the most well-regarded swarm intelligent optimization algorithm in the current literature, updates the positions of particles with regard to present location, personal best, and the global best. For FOA-LSSVM, the fruit flies renew their positions according to the location of the best fruit fly with the best smell. This demonstrates that in PSO and FOA the other search agents make no contribution to the update of a search agent's position, while all grasshoppers' status are involved in the position updating of each search agent [50]. Thus, the GOA performs better than PSO and the FOA on the optimization issue. Therefore, it can be concluded that the proposed multi-stage intelligent algorithm with the B-N decomposition method and the LSSVM model optimized by the newly proposed evolutionary algorithm GOA for hourly wind electric power forecasting is effective and practical.

## 7. Conclusions

Wind power is of great significance among low-carbon and low-emissions renewable energy technologies, which are promising to realize sustainable energy supply. The accurate prediction of wind power makes a great contribution to the extensive scale penetration of wind power into a power network considering the stochastic and intermittent characteristics of wind energy. Therefore, a multi-stage intelligent algorithm for wind electric power prediction is put forward in this research to increase the prediction precision of wind power. This established multi-stage intelligent algorithm combines the B-N decomposition approach, the LSSVM model, and a newly proposed intelligent optimization algorithm GOA together. In the data preprocessing stage, considering the stochastic tendency of wind electric power, the B-N disintegration approach is utilized to disintegrate the initial wind electric power data into a determinacy trend, a cyclic component, and a random component. Then, LSSVM is utilized to predict the future values of these three terms, of which the parameters ' $c$ ' and ' $\sigma$ ' are optimally selected by the GOA. The final future wind electric power data are calculated through multiplying the predicted values of the deterministic trend, cyclic component, and random component. For the purpose of evaluating the forecasting performance, the LSSVM approach, the FOA-LSSVM approach, and the PSO-LSSVM approach are selected as the compared methods. Through conducting comparative analysis, it can be concluded that the proposed multi-stage intelligent algorithm for wind electric power forecasting performs the best with a 4.28% forecasting error, which is much lower than the prediction error in the current literature, which ranges from 8% to 22% [18]. Although the proposed multi-stage intelligent optimization algorithm consists of complicated computation steps, including data decomposition, parameter optimization, and future data-point prediction, and this process may have a high computational cost, the proposed multi-stage intelligent algorithm can improve forecasting accuracy significantly and can be of great help for electricity power system management

and power dispatching. Therefore, the proposed multi-stage intelligent algorithm is an effective and attractive optimization algorithm for wind electric power prediction. In the future research, this proposed model can also be combined with the Numerical Weather Prediction method in using physical data to further increase the forecasting precision of wind power, and it can also be utilized for other prediction studies, such as power load prediction in the short term and photovoltaic power generation prediction. Additionally, considering the intense demand for quantitative information of wind power uncertainty, the research of wind power prediction will gradually turn from point prediction to probabilistic prediction [72–74]. Therefore, in our future research, we will focus on exploring appropriate probabilistic forecasting approaches to enhance the prediction accuracy of wind electric power prediction.

**Acknowledgments:** This study is funded by the Major State Research and Development Program of China under Grant No. 2016YFB0900501.

**Author Contributions:** The research concept was proposed by Huiru Zhao. The empirical analysis and the written of the manuscript were conducted by Haoran Zhao. The manuscript was revised by Sen Guo.

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

## References

1. Montoya, F.G.; Manzano-Agugliaro, F.; López-Márquez, S.; Hernández-Escobedo, Q.; Gil, C. Wind turbine selection for wind farm layout using multi-objective evolutionary algorithms. *Expert Syst. Appl.* **2014**, *41*, 6585–6595. [CrossRef]
2. Manzano-Agugliaro, F.; Alcayde, A.; Montoya, F.G.; Zapata-Sierra, A.; Gil, C. Scientific production of renewable energies worldwide: an overview. *Renew. Sustain. Energy Rev.* **2013**, *18*, 134–143. [CrossRef]
3. Chitsaz, H.; Amjady, N.; Zareipour, H. Wind power forecast using wavelet neural network trained by improved Clonal selection algorithm. *Energy Convers. Manag.* **2015**, *89*, 588–598. [CrossRef]
4. World Wind Energy Association. Available online: [www.wwindea.org](http://www.wwindea.org) (accessed on 18 December 2017).
5. China Energy Research Association. *China Energy Outlook 2030*; Economic Management Press: Beijing, China, 2016.
6. Hernández-Escobedo, Q.; Manzano-Agugliaro, F.; Zapata-Sierra, A. The wind power of Mexico. *Renew. Sustain. Energy Rev.* **2010**, *14*, 2830–2840. [CrossRef]
7. Hernandez-Escobedo, Q.; Manzano-Agugliaro, F.; Gazquez-Parra, J.A.; Zapata-Sierra, A. Is the wind a periodical phenomenon? The case of Mexico. *Renew. Sustain. Energy Rev.* **2011**, *15*, 721–728. [CrossRef]
8. Xu, Y.; Dong, Z.Y.; Xu, Z.; Meng, K.; Wong, K.P. An intelligent dynamic security assessment framework for power systems with wind power. *IEEE Trans. Ind. Inf.* **2012**, *8*, 995–1003. [CrossRef]
9. Hu, P.; Karki, R.; Billinton, R. Reliability evaluation of generating systems containing wind power and energy storage. *IET Gener. Trans. Distrib.* **2009**, *3*, 783–791. [CrossRef]
10. Jung, J.; Broadwater, R.P. Current status and future advances for wind speed and power forecasting. *Renew. Sustain. Energy Rev.* **2014**, *31*, 762–777. [CrossRef]
11. Hernández-Escobedo, Q.; Saldaña-Flores, R.; Rodríguez-García, E.R.; Manzano-Agugliaro, F. Wind energy resource in Northern Mexico. *Renew. Sustain. Energy Rev.* **2014**, *32*, 890–914. [CrossRef]
12. Rohrig, K.; Lange, B. Improvement of the power system reliability by prediction of wind power generation. In Proceedings of the Power Engineering Society General Meeting, Tampa, FL, USA, 24–28 June 2007.
13. Chen, N.; Qian, Z.; Nabney, I.T.; Meng, X. Wind power forecasts using Gaussian processes and numerical weather prediction. *IEEE Trans. Power Syst.* **2014**, *29*, 656–665. [CrossRef]
14. Khalid, M.; Savkin, A.V. A method for short-term wind power prediction with multiple observation points. *IEEE Trans. Power Syst.* **2012**, *27*, 579–586. [CrossRef]
15. Kavasseri, R.G.; Seetharaman, K. Day-ahead wind speed forecasting using f-ARIMA models. *Renew. Energy* **2009**, *34*, 1388–1393. [CrossRef]
16. González-Minguez, C.; Muñoz-Gutiérrez, F. RETRACTED: Wind prediction using Weather Research Forecasting model (WRF): A case study in Peru. *Energy Convers. Manag.* **2014**, *81*, 363–373. [CrossRef]

17. Methaprayoon, K.; Yingvivatanapong, C.; Lee, W.J.; Liao, J.R. An integration of ANN wind power estimation into unit commitment considering the forecasting uncertainty. *IEEE Trans. Ind. Appl.* **2007**, *43*, 1441–1448. [[CrossRef](#)]
18. Monteiro, C.; Bessa, R.; Miranda, V.; Botterud, A.; Wang, J.; Conzelmann, G. *Wind Power Forecasting: State-Of-The-Art 2009*; Argonne National Laboratory (ANL): Lemont, IL, USA, 2009.
19. Lange, M.; Focken, U. *Physical Approach to Short-Term Wind Power Prediction*; Springer: Berlin, Germany, 2006.
20. Landberg, L. Short-term prediction of local wind conditions. *J. Wind Eng. Ind. Aerodyn.* **2001**, *89*, 235–245. [[CrossRef](#)]
21. Landberg, L. Short-term prediction of the power production from wind farms. *J. Wind Eng. Ind. Aerodyn.* **1999**, *80*, 207–220. [[CrossRef](#)]
22. Focken, U.; Lange, M.; Waldl, H.P. Previento-a wind power prediction system with an innovative upscaling algorithm. In Proceedings of the European Wind Energy Conference, Copenhagen, Denmark, 2–6 July 2001; p. 276.
23. McGowin, C. California wind energy forecasting system development and testing. Phase 1: Initial testing. In *EPRI Final Report*; Electric Power Research Institute: Palo Alto, CA, USA, 2003.
24. Erdem, E.; Shi, J. ARMA based approaches for forecasting the tuple of wind speed and direction. *Appl. Energy* **2011**, *88*, 1405–1414. [[CrossRef](#)]
25. Sfetsos, A. A novel approach for the forecasting of mean hourly wind speed time series. *Renew. Energy* **2002**, *27*, 163–174. [[CrossRef](#)]
26. Liu, H.; Tian, H.Q.; Chen, C.; Li, Y.F. A hybrid statistical method to predict wind speed and wind power. *Renew. Energy* **2010**, *35*, 1857–1861. [[CrossRef](#)]
27. Lv, P.; Yue, L. Short-term wind speed forecasting based on non-stationary time series analysis and ARCH model. In Proceedings of the 2011 International Conference on Multimedia Technology (ICMT), Hangzhou, China, 26–28 July 2011; pp. 2549–2553.
28. Banos, R.; Manzano-Agugliaro, F.; Montoya, F.G.; Gil, C.; Alcayde, A.; Gómez, J. Optimization methods applied to renewable and sustainable energy: A review. *Renew. Sustain. Energy Rev.* **2011**, *15*, 1753–1766. [[CrossRef](#)]
29. Blonbou, R. Very short-term wind power forecasting with neural networks and adaptive Bayesian learning. *Renew. Energy* **2011**, *36*, 1118–1124. [[CrossRef](#)]
30. Ren, C.; An, N.; Wang, J.; Li, L.; Hu, B.; Shang, D. Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting. *Knowl.-Based Syst.* **2014**, *56*, 226–239. [[CrossRef](#)]
31. Kong, X.; Liu, X.; Shi, R.; Lee, K.Y. Wind speed prediction using reduced support vector machines with feature selection. *Neurocomputing* **2015**, *169*, 449–456. [[CrossRef](#)]
32. Barbounis, T.G.; Theocharis, J.B. Locally recurrent neural networks for long-term wind speed and power prediction. *Neurocomputing* **2006**, *69*, 466–496. [[CrossRef](#)]
33. Sideratos, G.; Hatzigiorgiou, N. Using radial basis neural networks to estimate wind power production. In Proceedings of the 2007 IEEE Power Engineering Society General Meeting, Tampa, FL, USA, 24–28 June 2007.
34. Ramirez-Rosado, I.J.; Fernandez-Jimenez, L.A.; Monteiro, C.; Sousa, J.; Bessa, R. Comparison of two new short-term wind-power forecasting systems. *Renew. Energy* **2009**, *34*, 1848–1854. [[CrossRef](#)]
35. Han, S.; Liu, Y.; Li, J. Wind power combination prediction based on the maximum information entropy principle. In Proceedings of the 2012 World Automation Congress (WAC), Puerto Vallarta, Mexico, 24–28 June 2012.
36. Li, G.; Shi, J.; Zhou, J. Bayesian adaptive combination of short-term wind speed forecasts from neural network models. *Renew. Energy* **2011**, *36*, 352–359. [[CrossRef](#)]
37. Vaccaro, A.; Mercogliano, P.; Schiano, P.; Villacci, D. An adaptive framework based on multi-model data fusion for one-day-ahead wind power forecasting. *Electr. Power Syst. Res.* **2011**, *81*, 775–782. [[CrossRef](#)]
38. Bhaskar, K.; Singh, S.N. AWNN-assisted wind power forecasting using feed-forward neural network. *IEEE Trans. Sustain. Energy* **2012**, *3*, 306–315. [[CrossRef](#)]
39. Tascikaraoglu, A.; Uzunoglu, M.; Vural, B. The assessment of the contribution of short-term wind power predictions to the efficiency of stand-alone hybrid systems. *Appl. Energy* **2012**, *94*, 156–165. [[CrossRef](#)]

40. Liu, Y.; Shi, J.; Yang, Y.; Lee, W.J. Short-term wind-power prediction based on wavelet transform–support vector machine and statistic-characteristics analysis. *IEEE Trans. Ind. Appl.* **2012**, *48*, 1136–1141. [[CrossRef](#)]
41. Flores, J.J.; Graff, M.; Rodriguez, H. Evolutive design of ARMA and ANN models for time series forecasting. *Renew. Energy* **2012**, *44*, 225–230. [[CrossRef](#)]
42. Salcedo-Sanz, S.; Ortiz-García, E.G.; Pérez-Bellido, Á.M.; Portilla-Figueras, A.; Prieto, L. Short term wind speed prediction based on evolutionary support vector regression algorithms. *Expert Syst. Appl.* **2011**, *38*, 4052–4057. [[CrossRef](#)]
43. Amjady, N.; Keynia, F.; Zareipour, H. Short-term wind power forecasting using ridgelet neural network. *Electr. Power Syst. Res.* **2011**, *81*, 2099–2107. [[CrossRef](#)]
44. Han, S.; Li, J.; Liu, Y. Tabu search algorithm optimized ANN model for wind power prediction with NWP. *Energy Procedia* **2011**, *12*, 733–740. [[CrossRef](#)]
45. Louka, P.; Galanis, G.; Siebert, N.; Kariniotakis, G.; Katsafados, P.; Pytharoulis, I.; Kallos, G. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. *J. Wind Eng. Ind. Aerodyn.* **2008**, *96*, 2348–2362. [[CrossRef](#)]
46. Zhang, Q.; Lai, K.K.; Niu, D.; Wang, Q. Wind park power forecasting models and comparison. In Proceedings of the 2012 Fifth International Joint Conference on Computational Sciences and Optimization (CSO), Harbin, China, 23–26 June 2012; pp. 27–31.
47. Liu, D.; Niu, D.; Wang, H.; Fan, L. Short-term wind speed forecasting using wavelet transform and support vector machines optimized by genetic algorithm. *Renew. Energy* **2014**, *62*, 592–597. [[CrossRef](#)]
48. Shaoping, W.; Jin, H. Trend-cycle Decomposition and Stochastic Impact Effect of Chinese GDP. *Econ. Res. J.* **2009**, *4*, 65–76.
49. Zhao, H.; Zhao, H.; Guo, S.; Li, F.; Hu, Y. The Impact of Financial Crisis on Electricity Demand: A Case Study of North China. *Energies* **2016**, *9*, 250. [[CrossRef](#)]
50. Saremi, S.; Mirjalili, S.; Lewis, A. Grasshopper Optimisation Algorithm: Theory and application. *Adv. Eng. Softw.* **2017**, *105*, 30–47. [[CrossRef](#)]
51. Beveridge, S.; Nelson, C.R. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the ‘business cycle’. *J. Monet. Econ.* **1981**, *7*, 151–174. [[CrossRef](#)]
52. Nelson, C.R.; Plosser, C.R. Trends and random walks in macroeconomic time series: Some evidence and implications. *J. Monet. Econ.* **1982**, *10*, 139–162. [[CrossRef](#)]
53. Campbell, J.Y.; Mankiw, N.G. *Permanent and Transitory Components in Macroeconomic Fluctuations*; National Bureau of Economic Research: Cambridge, MA, USA, 1987.
54. Campbell, J.Y.; Mankiw, N.G. *Are Output Fluctuations Transitory*; National Bureau Economic Research: Cambridge, MA, USA, 1986.
55. Morley, J.C. A state–space approach to calculating the Beveridge–Nelson decomposition. *Econ. Lett.* **2002**, *75*, 123–127. [[CrossRef](#)]
56. Vapnik, V.N.; Vapnik, V. *Statistical Learning Theory*; Wiley: New York, NY, USA, 1998.
57. Suykens, J.A.K.; Lukas, L.; Van Dooren, P.; De Moor, B.; Vandewalle, J. Least squares support vector machine classifiers: A large scale algorithm. In Proceedings of the European Conference on Circuit Theory and Design (ECCTD), Stresa, Italy, 29 August–2 September 1999; Volume 99, pp. 839–842.
58. Suykens, J.A.K.; Vandewalle, J. Least squares support vector machine classifiers. *Neural Process. Lett.* **1999**, *9*, 293–300. [[CrossRef](#)]
59. Deo, R.C.; Kisi, O.; Singh, V.P. Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. *Atmos. Res.* **2017**, *184*, 149–175. [[CrossRef](#)]
60. Zhao, H.; Guo, S.; Zhao, H. Energy-Related CO<sub>2</sub> Emissions Forecasting Using an Improved LSSVM Model Optimized by Whale Optimization Algorithm. *Energies* **2017**, *10*, 874. [[CrossRef](#)]
61. Cherkassky, V.; Mulier, F.M. *Learning from Data: Concepts, Theory, and Methods*; John Wiley & Sons: Hoboken, NJ, USA, 2007.
62. Kisi, O. Pan evaporation modeling using least square support vector machine, multivariate adaptive regression splines and M5 model tree. *J. Hydrol.* **2015**, *528*, 312–320. [[CrossRef](#)]
63. Deo, R.C.; Tiwari, M.K.; Adamowski, J.F.; Quilty, J.M. Forecasting effective drought index using a wavelet extreme learning machine (W-ELM) model. *Stoch. Environ. Res. Risk Assess.* **2016**, 1–30. [[CrossRef](#)]

64. Deo, R.C.; Wen, X.; Qi, F. A wavelet-coupled support vector machine model for forecasting global incident solar radiation using limited meteorological dataset. *Appl. Energy* **2016**, *168*, 568–593. [[CrossRef](#)]
65. Simpson, S.J.; McCaffery, A.; HAeGELE, B.F. A behavioural analysis of phase change in the desert locust. *Biol. Rev.* **1999**, *74*, 461–480. [[CrossRef](#)]
66. Rogers, S.M.; Matheson, T.; Despland, E.; Dodgson, T.; Burrows, M.; Simpson, S.J. Mechanosensory-Induced behavioural gregarization in the desert locust *Schistocerca gregaria*. *J. Exp. Biol.* **2003**, *206*, 3991–4002. [[CrossRef](#)] [[PubMed](#)]
67. Eiben, A.E.; Schippers, C.A. On evolutionary exploration and exploitation. *Fundam. Inform.* **1998**, *35*, 35–50.
68. Topaz, C.M.; Bernoff, A.J.; Logan, S.; Toolson, W. A model for rolling swarms of locusts. *Eur. Phys. J. Spec. Top.* **2008**, *157*, 93–109. [[CrossRef](#)]
69. Loague, K.; Green, R.E. Statistical and graphical methods for evaluating solute transport models: Overview and application. *J. Contam. Hydrol.* **1991**, *7*, 51–73. [[CrossRef](#)]
70. Santos, R.D.; Boote, K.J.; Sollenberger, L.E.; Neves, A.L.; Pereira, L.G.; Scherer, C.B.; Gonçalves, L.C. Simulated optimum sowing date for forage pearl millet cultivars in multilocation trials in Brazilian semi-arid region. *Front. Plant Sci.* **2017**, *8*, 2074–2085. [[CrossRef](#)] [[PubMed](#)]
71. De Brabanter, K.; Karsmakers, P.; Ojeda, F.; Alzate, C.; De Brabanter, J.; Pelckmans, K.; De Moor, B.; Vandewalle, J.; Suykens, J.A.K. *LS-SVMLab Toolbox User's Guide*; ESAT-SISTA Technical Report; Katholieke Universiteit Leuven: Leuven, Belgium, 2011.
72. Hong, T.; Pinson, P.; Fan, S.; Zareipour, H.; Troccoli, A.; Hyndman, R.J. Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. *Int. J. Forecast.* **2016**, *32*, 896–913. [[CrossRef](#)]
73. Zhang, Y.; Wang, J. K-nearest neighbors and a kernel density estimator for GEFCom2014 probabilistic wind power forecasting. *Int. J. Forecast.* **2016**, *32*, 1074–1080. [[CrossRef](#)]
74. Zhang, Y.; Wang, J.; Wang, X. Review on probabilistic forecasting of wind power generation. *Renew. Sustain. Energy Rev.* **2014**, *32*, 255–270. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).