

Article

A Multi Time Scale Wind Power Forecasting Model of a Chaotic Echo State Network Based on a Hybrid Algorithm of Particle Swarm Optimization and Tabu Search

Xiaomin Xu *, Dongxiao Niu, Ming Fu, Huicong Xia and Han Wu

Received: 29 August 2015 ; Accepted: 20 October 2015 ; Published: 3 November 2015
Academic Editor: Guido Carpinelli

School of Economics and Management, North China Electric Power University, Beijing 102206, China; niudx@ncepu.edu.cn (D.N.); fm@ncepu.edu.cn (M.F.); 1111330121@ncepu.edu.cn (H.X.); wuhanncepu@gmail.com (H.W.)

* Correspondence: xuxm@ncepu.edu.cn; Tel.: +86-10-61773079; Fax: +86-10-61773308

Abstract: The uncertainty and regularity of wind power generation are caused by wind resources' intermittent and randomness. Such volatility brings severe challenges to the wind power grid. The requirements for ultrashort-term and short-term wind power forecasting with high prediction accuracy of the model used, have great significance for reducing the phenomenon of abandoned wind power, optimizing the conventional power generation plan, adjusting the maintenance schedule and developing real-time monitoring systems. Therefore, accurate forecasting of wind power generation is important in electric load forecasting. The echo state network (ESN) is a new recurrent neural network composed of input, hidden layer and output layers. It can approximate well the nonlinear system and achieves great results in nonlinear chaotic time series forecasting. Besides, the ESN is simpler and less computationally demanding than the traditional neural network training, which provides more accurate training results. Aiming at addressing the disadvantages of standard ESN, this paper has made some improvements. Combined with the complementary advantages of particle swarm optimization and tabu search, the generalization of ESN is improved. To verify the validity and applicability of this method, case studies of multitime scale forecasting of wind power output are carried out to reconstruct the chaotic time series of the actual wind power generation data in a certain region to predict wind power generation. Meanwhile, the influence of seasonal factors on wind power is taken into consideration. Compared with the classical ESN and the conventional Back Propagation (BP) neural network, the results verify the superiority of the proposed method.

Keywords: wind power output; multi time scale; seasonal factor; particle swarm optimization; tabu search; chaotic echo state network (ESN)

1. Introduction

The energy crisis is becoming more and more obvious with the continuous increase in energy consumption, overexploitation of traditional energy and exhaustion of fossil fuels. In order to ease the crisis, countries around the world are increasing the utilization of new energy sources and realizing the sustainable development of energy. As an important category of renewable energy, wind has the characteristics of huge reserves, renewability, wide distribution and non-polluting nature [1]. In recent years, as an important aspect of renewable energy utilization and the most mature new energy use technology, the penetration of wind energy in the power grid is increasing rapidly [2].

China has abundant wind energy resources. Besides, the government has also attached great importance to the development of wind power industry. China's wind power has entered a stage of rapid development since 2003. China's wind power installed capacity doubled for four consecutive years from 2006 to 2009. The new installed capacity of wind power in China exceeds 16 GW and the total installed capacity reached 41.83 GW in 2010, which surpassed the United States for the first time, ranking first in the World [3]. In 2012, China's total wind power installed capacity is still in the first place, and wind energy growth momentum in the United States is also relatively strong, with a new grid installed capacity is 13.1 GW; in Europe, Germany's wind power development is the best, followed by Spain and the United Kingdom. By the end of 2012, the national rankings of the top ten countries according to their total installed capacity of wind power was as shown in Figure 1.

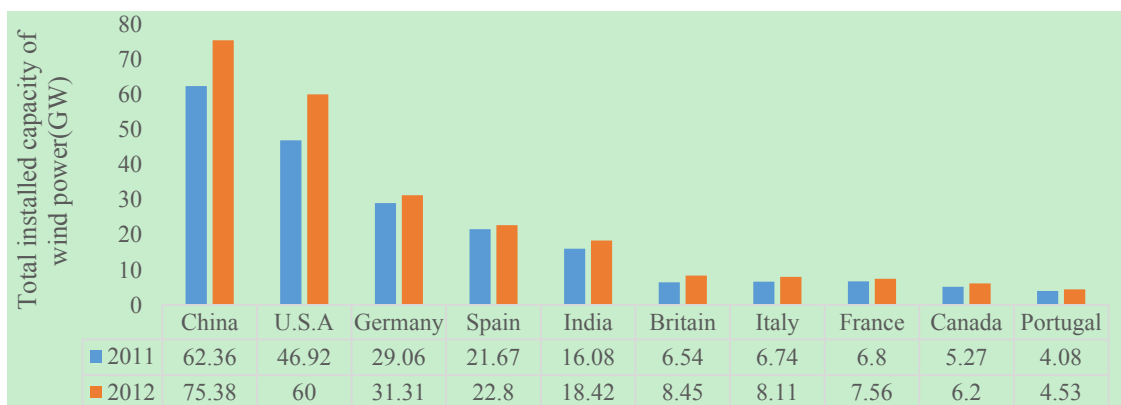


Figure 1. National ranking of the top ten countries by total installed wind power capacity.

Against the background of global economic development, the global wind power industry is growing rapidly. However, the intermittent nature and randomness of wind power determines the strong volatility of wind power. With the increasing number of wind farms and the installed capacity, the wind power volatility represents a huge challenge to economic security and increases the difficulty of the operation and management of power network dispatching companies once incorporated into the power grid. Accurate prediction of the wind power in advance can relieve the pressure of peaking power systems, which can effectively improve the ability to include wind power in the grid [4].

Many experts and scholars both at home and abroad have performed lots of theoretical studies and practical simulations. From the time scale point of view, wind power forecasting can be divided into long-term forecasting, short-term and ultrashort-term forecasting. Long-term forecasts take “week”, “month”, “year” as the units, and are mainly used for the wind farm design feasibility studies and the maintenance and commissioning of wind turbines. Different from the European wind power distributed access, most of China's wind power is connected through a centralized grid connection, so any power fluctuation has a greater impact on the power grid scheduling. The ultrashort-term prediction has been widely used in the optimization of spare capacity [5,6] and the economic load dispatch and control [7]. Short-term forecasting can reduce the abandoned wind and adjust the maintenance plans [3]. Short-term and ultrashort-term forecasting which take “hour”, “minute” as the units are particularly important.

Depending on the research method, the short-term wind power forecasting method can be roughly divided into physical, statistical and learning methods [3], seen in Figure 2.

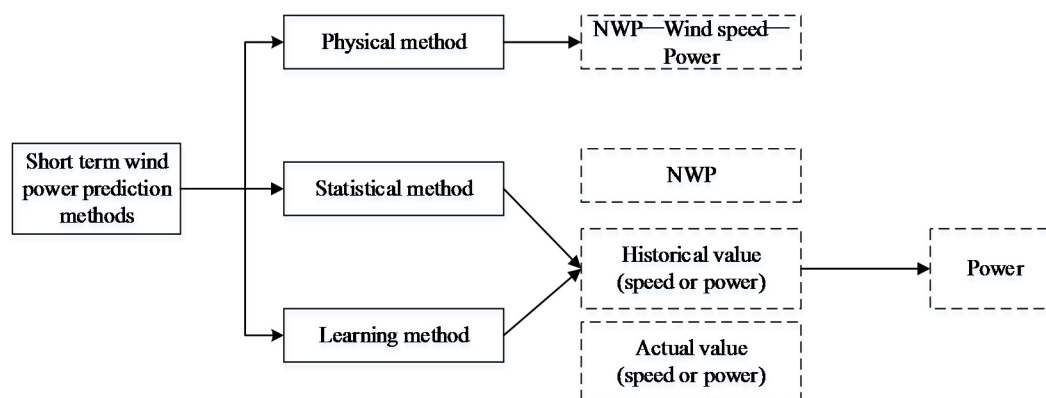


Figure 2. Short-term wind power prediction methods.

Physical methods are based on numerical weather forecast models (Numerical Weather Prediction—NWP). The wind turbine, wind speed, direction and other information about the hub height of the wind turbine are obtained, combined with physical information, and the power curve of the wind turbine is used to simulate the actual output power [8]. This method has no historical data support and is suitable for forecasting a new wind farm, but it has high accuracy requirements and data collection is difficult with high calculation demands and implementation costs.

Statistical methods establish a forecast model by finding the relationships between historical data and wind speed or power without considering the physical process of wind speed change. This relationship can be expressed in function form, such as Markov chains [4], regression analysis [9,10], exponential smoothing method [11], Kalman filtering method [12], ARMA model [13] and so on. Among them, the ARMA (p, q) model, used as the common statistical model, has high accuracy of analysis and prediction for stationary time series, however, due to the impact of the uncertain natural climate, wind energy has obvious trends, diversity and periodicity, which show non-stationary time series. Statistical models thus have great limitations. Compared with physical models, statistical models are relatively simple and have strong applicability, but they need a large amount of data support, which is not suitable for the prediction of new wind farms.

Learning methods extract the relationship between the input and output using artificial intelligence methods. The models are usually nonlinear and cannot be expressed by an analytical expression. Commonly used models are neural network [14,15], wavelet analysis [16], support vector machine [17,18], etc.. Learning methods have strong adaptive ability, and the accuracy of the models can be improved by error correction. Wind power generation often shows irregular and uncertain behaviors. Thus, Bayesian theory is applied to wind power forecasting. Probability forecasting methods can give the probability of all the wind power generation in the next moment, providing more comprehensive forecast information, but they require a great quantity of data. The analysis and calculation process is more complex, and some data depends on the prior probability [19].

Many studies have judged the chaotic characteristics before prediction [20–22]. Literatures [20,21] verified that wind power output and wind speed time series have chaotic properties. According to the theory of nonlinear dynamics, chaotic time series in the short term are predictable, indicating that the wind power prediction is conducted with the chaos method. The traditional prediction methods for chaotic time series include the global method and local method. The global predictive method takes all state points into account and makes them the fitted objects to find out the rules to fit the prediction model. This method needs more historical data and a large amount of computation. It is difficult to calculate the phase space with high embedding dimensions, which limits its application. In contrast, the local method is more widely applied. Literature [21] proposes an improved weighted zero order local area prediction method, using the correlation degree of the phase point to find the near phase point, which improves the forecasting precision. A multivariable local prediction method of wind speed is proposed in [22]. Based on the principle

of phase relation, the multivariable time series data are screened to construct the multivariable phase space. Neural networks have been widely used in the prediction of chaotic sequences in the last few years [23–28]. Back Propagation (BP) neural network [23,24], radial basis function neural network [25], Elman neural network [26], GMDH type neural network [27] and wavelet neural network [28] have all been used to forecast chaotic sequences and have achieved different prediction effects in different areas. With its good nonlinear forecasting ability, the recurrent neural network has been the main tool for the prediction of chaotic time series, and has been widely recognized by the academic community [29]. However, the mathematical analysis is difficult in practical use, for it has the disadvantages of being time-consuming, with large computational demands and low training efficiency. As a new recurrent neural network, echo state network (ESN) has greatly improved stability, global optimality and training process complexity phase compared with the traditional neural network [30], attracting wide attention of scholars at home and abroad. It has been widely used in speech recognition, traffic control [26] and communication forecasting [31], but is rarely used in power system forecasting. In addition, the classical ESN has limited expression ability for nonlinear dynamic systems, and it is difficult to meet the requirement of precision of various forecasting problems in real situations, so its generalization ability is not strong.

In view of this, this paper presents a short-term wind power intelligent forecasting model for dealing with the characteristics of nonlinear power generation. Considering the uncertainties of wind power generation, the time series chaotic sequence is judged. The delay time and the embedding dimension are computed to reconstruct phase space; then, to overcome the shortcomings of the ESN network itself, this paper designs and trains an ESN to forecast short-term wind power by combining the two complementary optimization methods of particle swarm optimization and tabu search to improve the generalization ability of the model. The intelligent forecasting model proposed in this paper has the following innovations:

- (1) Firstly, the phase space reconstruction of chaotic time series is used to reduce the impact of the uncertainty of the original data;
- (2) Secondly, the ESN is introduced for prediction. As a new dynamic recurrent neural network, ESN has a unique simple training method and high precision training results. The introduction of reserve pool has overcome the problems of slow convergence and local minimum, which can bring about higher accuracy;
- (3) Considering the characteristics of the tabu search and particle swarm algorithm, the diversity principles of the tabu search algorithm are introduced into the mechanisms of “group share” and “random search” of particle swarm to make the search process have memory. The combination of the two algorithms provides complementary advantages to improve the prediction accuracy;
- (4) Considering the multitime scales of minutes, hours and days simultaneously, wind power predictions are carried out. Meanwhile, the seasonal factors are integrated into the model to analyze the characteristics of wind power, and the robustness and generalization ability of the proposed method is verified.

In the empirical analysis, compared with the classical ESN, the traditional BP network and the single method optimizing ESN, the proposed method achieves higher prediction accuracy. The main contents of this paper are as follows: Section 1 reviews and summarizes the wind power forecasting methods, and proposes the research methods and contents of this paper; Section 2 outlines the basic theory of the chaotic ESN and describes the structure and principle of the ESN; Section 3 introduces the optimization method of particle swarm optimization and tabu search and explains the principle and process of the combination optimization method in detail; Section 4 establishes the prediction process of the chaotic ESN based on particle swarm optimization and tabu search algorithms; Section 5 reports our experiments and relevant discussion; Concluding remarks are presented in Section 6.

Among them, $P(x_i)$ is the probability of event x_i , q is the count of events and states which might happen, logarithmic functions often take 2 to the base.

For time series x_i and delay time τ with the length of n , if the probability of x_i appearing in $\{x_i\}$ is set as $p(x_i)$, then the probability of $x_{i+\tau}$ appearing in $\{x_i\}$ is $P(x_{i+\tau})$, which is calculated by the frequency of occurrence in the corresponding time series. The joint probability of x_i and $x_{i+\tau}$ appearing in both two sequences is $P(x_i, x_{i+\tau})$, which can be obtained from the corresponding lattice on the plane $(x_i, x_{i+\tau})$. Thus, mutual information $I(X_i, X_{i+\tau})$ is a function $I(\tau)$ of delay time τ . The time delay of phase space reconstruction is the residence time when the function $I(\tau)$ reaches the minimum value for the first time.

2.1.2. False Nearest Neighbors Method (FNN) to Determine the Embedding Dimension

From the geometry point of view, a chaotic time series is the projection of the trajectory of the chaotic motion of the high-dimensional phase space in a one-dimensional space. In the projection process, the chaotic motion trajectory will be distorted. Points which are not adjacent in high-dimensional phase space may be adjacent when projected onto the one-dimensional axis, constituting false neighbor points. This is the reason why the chaotic time series is irregular. Reconstruction of phase space is actually the recovery of the chaotic motion trajectory from the chaotic time series. With the increase of embedding dimension m , the orbit of chaotic motion will gradually open, false neighbor points will be eliminated gradually and the trajectory of chaotic motion will be recovered, which is the starting point of false nearest neighbors method (FNN) [34].

In the d dimensional phase space, each phase point vector $x(i) = \{x(i), x(i + \tau), \dots, x(i + (d - 1)\tau)\}$, has a nearest neighbor point $x^{NN}(i)$ within certain distance and its distance can be described as:

$$R_d(i) = \| x(i) - x^{NN}(i) \| \tag{4}$$

When the dimension of phase space increasing from d to $d + 1$, the distance will change to $R_{d+1}(i)$, and:

$$R_{d+1}^2(i) = R_d^2(i) + \| x(i + \tau d) - x^{NN}(i + \tau d) \|^2 \tag{5}$$

If the value of $R_{d+1}(i)$ is much larger than that of $R_d(i)$, it can be considered the it is caused by the reason that the two adjacent points of the high dimensional chaotic attractor are projected to the lower dimensional orbit and become two adjacent points. Thus, such neighbor points are false.

Assuming that:

$$a_1(i, d) = \frac{\| x(i + \tau d) - x^{NN}(i + \tau d) \|^2}{R_d^2(i)} \tag{6}$$

If $a_1(i, d) > R_\tau$, then $x^{NN}(i)$ is the false nearest neighbor point of $x(i)$. Threshold value R_τ can be selected between [10, 50].

For the measured time series, we calculate the proportion of false nearest neighbors from the minimum embedding dimension, then increase d until the proportion of false nearest neighbors is less than 5 %, or the number of false nearest neighbors no longer decreases with the increase of d . At this time, the chaotic attractor can be thought to open completely and d is the embedding dimension. FNN is regarded as a quite effective method in terms of determining the embedding dimension of phase space dimension.

2.1.3. The Maximum Lyapunov Exponent for Chaotic Recognition

The value of the Lyapunov exponent reflects the shrinkage and tensile properties of the system in all directions. This paper calculates the Lyapunov exponent by means of a small data method to determine the chaos characteristics of the original sequence.

2.2. ESN

ESN is a new recurrent neural network composed of input, hidden and output layers [35]. The basic idea is to simplify the network training process, using large-scale recursive networks with random connections to replace the middle layer of a classic neural network. Among them, the hidden layer, called dynamic reservoir, is the core structure of the ESN network, and it is composed of relatively rich neurons with random and sparse connection relations, as shown in Figure 3.

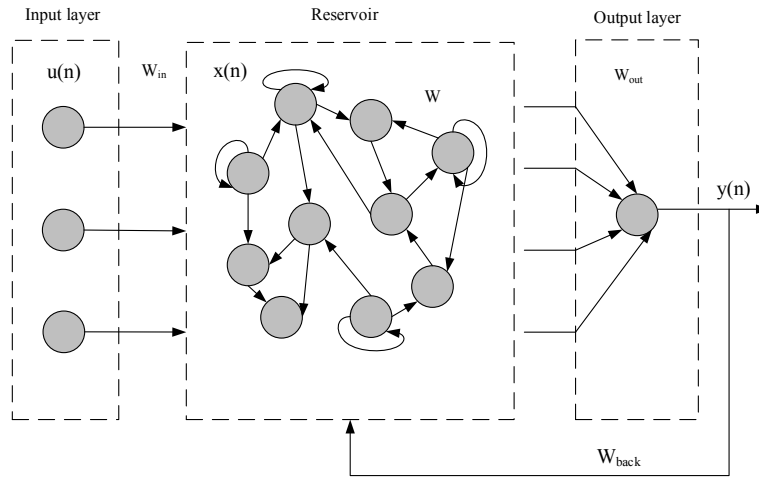


Figure 3. Echo state network (ESN) structure.

ESN reservoirs have a large number of neurons, approximately 20–500, so they have a great short-term memory [31]. Assuming that an ESN network has M input units, N internal neurons, and L output units, then the input unit $u(n)$, internal state $x(n)$ and output unit $y(n)$ at the moment n are described as:

$$\begin{aligned}
 u(n) &= [u_1(n), u_2(n), \dots, u_M(n)]^T \\
 x(n) &= [x_1(n), x_2(n), \dots, x_N(n)]^T \\
 y(n) &= [y_1(n), y_2(n), \dots, y_L(n)]^T
 \end{aligned}
 \tag{7}$$

The reservoir state and output of ESN are updated according to the following formulas:

$$x(n + 1) = f(Wx(n) + W_{in}u(n) + W_{back}y(n))
 \tag{8}$$

$$\hat{y}(n + 1) = f_{out}(W_{out}[x(n + 1), u(n + 1), y(n)] + W_{bias}^{out})
 \tag{9}$$

W represents the connection weight matrix within the reservoir, and its sparse connection usually remains at 1%–5%, while the spectral radius ρ is usually less than 1 to ensure the echo state property and the dynamic memory capacity and stability of reservoir of ESN; W_{in} and W_{back} , respectively, represent the connection weight matrix of input and output to state variables; W_{out} represents the connection weight matrix of reservoirs, inputs and outputs to output, W_{bias}^{out} is a bias term of output, or may represent noise; $f = [f_1, f_2, \dots, f_N]$ indicates the internal neuron activation function, and $f_i (i = 1, 2 \dots N)$ normally takes hyperbolic tangent function. The output function is $f_{out} = [f_{out}^1, f_{out}^2, \dots, f_{out}^L]$. Typically, $f_{out}^i (i = 1, 2 \dots L)$ takes the identity function. In network training, the connection weight matrixes W , W_{in} and W_{back} connected to the reservoir are generated randomly, and will not change once generated, while the connection weight matrix W_{out} connected to the output is generated through testing. Therefore, the training purpose of ESN is to determine the output weight W_{out} [28].

When predicting a chaotic time series, we use the basic model with no output feedback, that is $W_{back} = 0$. Meanwhile, to simplify the calculation, we also assume that the connection weight of

input to output and output to output should be 0. The training procedure of output weight matrix W_{out} is as follows [26]:

- (1) Initialize the ESN. The weight matrixes W and W_{in} are generated randomly, elements of them are initialized to random numbers obeying $[0, 1]$ uniform distribution, and will not change once produced. The state vector $x(0)$ is set to zero;
- (2) Update the state vector of ESN reservoir. According to the formula Equation (4), calculate the new reservoir state vector $x(n + 1)$ with the given input vector $u(n)$ and state connection vector $x(n)$;
- (3) Calculate the output weight matrix W_{out} . Select all state connection vectors and desired output vectors after the time point $n = A_0$, then train the weight matrix W_{out} according to the following formulas:

$$\begin{aligned} n &= A_0, A_0 + 1, \dots, A \\ X &= [x(A_0), x(A_0 + 1), \dots, x(A)]^T \\ Y &= [y(A_0), y(A_0 + 1), \dots, y(A)]^T \\ (W_{out})^T &= X^{-1}Y \end{aligned} \tag{10}$$

- (4) After completing the training of output weight matrix W_{out} , set all the input vectors $u(n), (n = A + 1, \dots, N)$ after time point $n = A$ as the test input vectors and take them into the Equation (5) to get the predicted output vector $\hat{y}(n)$.

2.3. The Prediction Steps of Chaotic ESN

Specific prediction steps using chaos ESN are as follows:

- (1) Construct the network. Establish the network between an input matrix X and output vector Y , using the calculated embedding dimension m and delay time τ of chaotic time series as the input number [36,37]:

$$X = \begin{bmatrix} x_1 & x_{1+\tau} & \dots & x_{1+(m-1)\tau} \\ x_2 & x_{2+\tau} & \dots & x_{2+(m-1)\tau} \\ \vdots & \vdots & \ddots & \vdots \\ x_n & x_{n+\tau} & \dots & x_{n+(m-1)\tau} \end{bmatrix} \tag{11}$$

$$Y = \begin{bmatrix} x_{2+(m-1)\tau} \\ x_{3+(m-1)\tau} \\ \vdots \\ x_{n+1} \end{bmatrix} \tag{12}$$

- (2) Learning period. Select a certain length time series, and train the network using the matrix X as input and the vector Y as output. By adopting the output of the before process, compare the result with the target model. If there is any error, start the back-propagation process immediately, and correct the network weights in order to reduce error.
- (3) Forecasting period. The input of ESN is set to be $X' = [x_{n+1-(m-1)\tau}, x_{n+1-(m-2)\tau}, \dots, x_{n+1}]$, and then the output will be the actual prediction value $Y' = x_{n+2}$ of the sequence. The predicted value should be used in the iteration prediction, as the network input of a new round.

3. Theoretical Introduction of Optimization Algorithm

During the training of ESN network, over reliance on the training set is prone to appear as over fitting and reduce the generalization of the network. This paper optimizes the output weights through the hybrid complementary algorithm of particle swarm optimization and tabu search to improve generalization of the network.

3.1. Basic Theory of Particle Swarm Optimization

Particle swarm optimization (PSO) algorithm is an evolutionary algorithm. It originates from observing birds' behavior while searching for food. The conversion process of the motion of whole flock from disorderly to orderly comes from information shared by each individual bird in the flock, so as to find food [38].

The method optimized by PSO is to initialize a set of stochastic particles swam and find out the best method through multiple iterations. In the iteration process, each particle updates its direction and position constantly according to two extreme values. The first one is the optimal solution found by the particle itself, called individual extreme p_{best} , and the other is the current optimal solution found by the entire particle swarm, named global extreme g_{best} . At the beginning of the iterations, the position of each initialized particle is the individual extreme, while the best position of the particle swarm is the global extreme.

After all of the particles in the swarm complete the first iteration, After all particles finish the first iteration, the front and rear position of each particle should be compared, and the individual extreme with the optimal solution in this iteration should be updated if the new position is better than the previous one. Then, we need to get the optimal solution throughout the individual extremes of all particles in the swarm as the global extreme by comparison, and update the global extreme if the new one is better than the old one. The final global extreme obtained through these cycle iteration operations determines the optimal solution [39]. The diagram of the movement of particles is shown in Figure 4.

After obtaining individual extreme and global extreme in the process, each particle needs to update its velocity and position according to the following formulas:

$$v_{i,j+1} = wv_{i,j} + c_1 * \text{random}() * (p_{best_{i,j}} - P_{i,j}) + c_2 * \text{random}() * (g_{best_{i,j}} - P_{i,j}) \tag{13}$$

$$P_{i,j+1} = P_{i,j} + v_{i,j+1} \tag{14}$$

where, $v_{i,j}$ denotes the velocity of the i particle after j iterations, $P_{i,j}$ means the position of the i particle after j iterations, $p_{best_{i,j}}$ and $g_{best_{i,j}}$ are on behalf of the individual extreme and global extreme of the i particle after j iterations, w is the inertia weight of the updated speed to the speed of pre update, $\text{random}()$ is a random number within $(0, 1)$, c_1 and c_2 are learning factors within $(0, 2]$. The velocity of particles in the swarm is limited in $(0, v_{max})$, and the updated value should be replaced with v_{max} if it exceeds maximum v_{max} during the iteration process.

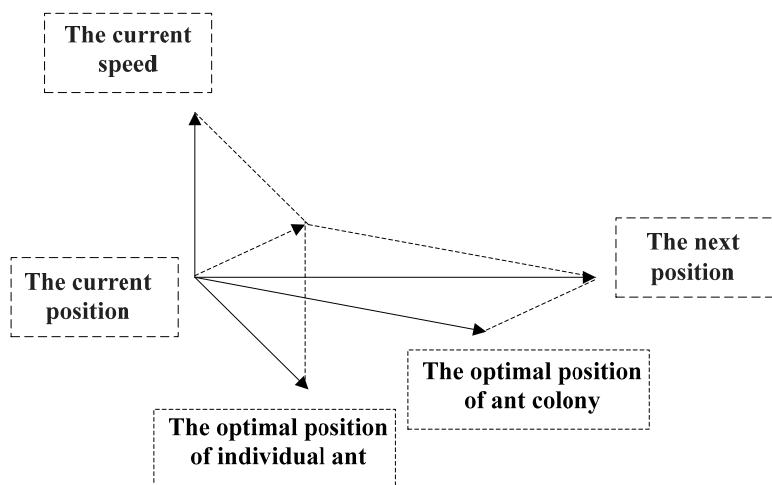


Figure 4. Movement of particles.

In the particle swarm optimization algorithm process, particles share the global extreme value with other particles within the group. This one-way flow of shared information and data makes the whole search process follow the group within the current optimal solution. Therefore, the initial particle swarm optimization algorithm has fast global convergence capability.

3.2. Tabu Search

The tabu search algorithm adopts a neighborhood search algorithm through imitating human's memory to avoid local optimal solutions. The algorithm must be able to accept an inferior solution, that is to say, the solution obtained from iteration is not necessarily better than the previous one, but once the inferior solution is accepted, it is possible for an iteration to fall into a circle [40]. The basic theory is to operate on the current solution through the application of a moving operator when giving an initial solution. A set of neighborhoods of the current solution will be generated after that. Then an optimal solution is searched as the current solution through the neighborhoods. The above operation is repeated until the convergence condition is satisfied. To avoid falling into a search loop, a flexible "memory" technology is utilized in the search: setting a taboo table of designated length so that it is propitious to record and select optimization procedures which have been carried out before, as well as to guide the next search direction. Once the algorithm puts some move into the tabu table, it will be banned. In this way, it is possible to prevent algorithm from revisiting a solution which has been already visited in the last several iterations, which averts the circulation and helps the algorithm avoid local optimal solutions. The tabu table will be updated in the iterative process. After iteration for a certain number of times, a move which is beyond the length of the taboo list will be shifted out of the table and the solution is released from it. Such a rule is called contempt criteria [41]. In general, tabu search has several key elements as follows:

- (1) Encoder mode: the initial solution of the TS can be randomly generated. It can also be readily obtained by means of heuristics based on the properties of the problem at hand;
- (2) The choice of candidate solutions: the number of candidate solutions should be firstly determined followed by the choice of the optimal candidate solution. In principle, all neighborhood solutions of the current solution should be executed. However, when the scale of a problem is large, there will be a great number of neighborhood solutions. Considering the efficiency of neighborhood search, it is better to merely select subsets of the neighborhood. As for the selection of best candidate solution, the general approach is to choose the best one that meets the principle of contempt or non-taboo in the set of candidate solutions;
- (3) Tabu table and its length: tabu table usually adopts a FIFO queue and its length can be fixed or change adaptively;
- (4) Contempt principles: it often uses simple contempt principles that if a solution is better than the optimum, its tabu properties are ignored. The solution are selected as the current solution;
- (5) Termination condition: the common condition is the maximum number of functions or the maximum continuous iteration number when the best solution can sustain.

It is advantageous for tabu search to accept inferior solutions through the search process and it has strong mountain climbing ability, which can avoid slipping into local optimal spot. The drawback is that the initial has strong dependence. Iterative search process is moving a solution to another one, which reducing the efficiency of obtaining global optimal solution.

3.3. Hybrid Algorithm Based on Particle Swarm Algorithm and Tabu Search

The particle swarm algorithm belongs to iterative random search algorithm in nature, which can optimize a variety of functions effectively. It also has the advantage of simple operation and strong global searching ability. Great results are demonstrated in continuous optimization and discrete optimization problems. However, as the execution progresses, the algorithm can easily fall into a

local extreme point. Meanwhile, the search precision value is not high and the convergence velocity is slow. It is prone to the precocious phenomenon.

The tabu search algorithm has a fast convergence rate and strong local search ability, but its search performance relies heavily on the initial value which is already given. A better initial solution often makes the tabu search algorithm converge to the global optimal solution quickly while a poor one may greatly diminish its convergence velocity. Thus, another heuristic algorithms is often applied to provide a better initial solution, in order to improve the performance of the search algorithm.

In consideration of the characteristics of the two algorithms, this paper introduces the diversity thought of tabu search algorithm into the mechanisms of “group share” and “random search” of the particle swarm to make the search process have memory. Owing to the better globe search capability of the particle swarm algorithm, it is available in the first stage. The tabu search algorithm can work in the later stage in consideration of its stronger local search capability. The TS algorithm with its strong local search ability is introduced in the later period of the particle swarm algorithm. Consequently, we can leverage their comparative advantages successfully and readily solve the difficulty of hard convergence.

To improve search efficiency, this paper assumes that the key in the combination of these two algorithms is to escape this local optimal state through the tabu algorithm when the particle swarm algorithm tends to precocious phenomenon. However, as it is difficult to determine whether the precocious phenomenon takes place, we adopt the sample variance of fitness to judge precociousness referring to [42]:

$$m < \sigma/\sigma_1 < n \quad (15)$$

where, σ is the variance of sample’s fitness after particle swarm optimization; σ_1 is the sample variance after the last particle swarm optimization; m and n are two numbers approaching to 1. It is possible to deem that the particle swarm algorithm is going to be premature when the changes of variance generated from two adjacent optimizations are small. There is no theoretical method to determine the values of m and n . If the neighborhood of $[m, n]$ is too large, the program is likely to jump outside global search. To achieve the best effects, this paper sets $m = 0.9, n = 1.1$ through multiple operations.

4. Chaotic ESN Optimized by Particle Swarm and Tabu Search

Neural networks have powerful parallel processing and nonlinear mapping capability. In other words, they can combine various approaches to process non-linear signals and tools. This can be used to study time series, predict and control for unknown power system. As the chaotic time series has regularity in the interior which arises from the nonlinearity, it indicates some correlation among time series in the time delay state space, which makes the system seems have a memory function, while it is also difficult to express this rule by conventional analytical methods. This information processing manner is exactly what the neural networks have; therefore, neural networks are adaptive to predict chaotic time series.

The ESN is a new type of large-scale recurrent neural network proposed by Jaeger and Haas [35] in 2004. Currently, this model has been widely used to predict and study chaotic time series [26–29]. However, ESN also has its own shortcomings. In many instances, ESN’s state matrix would be ill-conditioned, causing excessive amplitude of output weights and influencing the model’s accuracy and generalization ability. To solve the problem, this paper proposes a combined optimization model to optimize the ESN’s output weights based on particle swarm and tabu searching to improve the robustness of the model. Figure 5 and the specific forecasting steps and are shown as below.

- (1) Determine the delay time τ and embedding dimension m according to the mutual information method and FNN method to reconstruct the phase sequence of wind power generation;

- (2) Use the reconstructed data as sample data, separate the training set and test set and set the parameters of the model, including: particle swarm size M , iteration number, tabu search length, etc.;
- (3) Initialize the parameters of the particle swarm optimization and determine the initial position and speed to find the initial self-optimal position of each particle, that is, the position of each particle itself;
- (4) Calculate the fitness of particle swarm, and find out p_{best} and g_{best} ;
- (5) Update the position and speed of all particles according to Equations (13) and (14);
- (6) According to Equation (15), judge that whether the particle tends to premature. If yes, then go to the next step of tabu search; otherwise, determine whether the particle swarm optimization meet the convergence condition, if yes, then output results, otherwise, go to Step (3);
- (7) Set the current solution optimized by particle swarm as the initial solution of the tabu search, use the neighborhood function of the current global extremum to produce a number of neighborhood values, and select some candidate solutions with the highest fitness;
- (8) Judge each candidate using the contempt criterion, if satisfactory, replace the current solution p_{best} with the candidate solution and update the tabu list and optimal solution, then go to the Step (6), otherwise, proceed to the next step;
- (9) Judge whether the candidate solution is in the tabu list, select the best state instead of the current solution, find the corresponding tabu object to replace the taboo object which firstly entered the tabu table;
- (10) Judge whether meet the maximum iteration number, if yes, then output optimal solution, otherwise, go to Step (3).

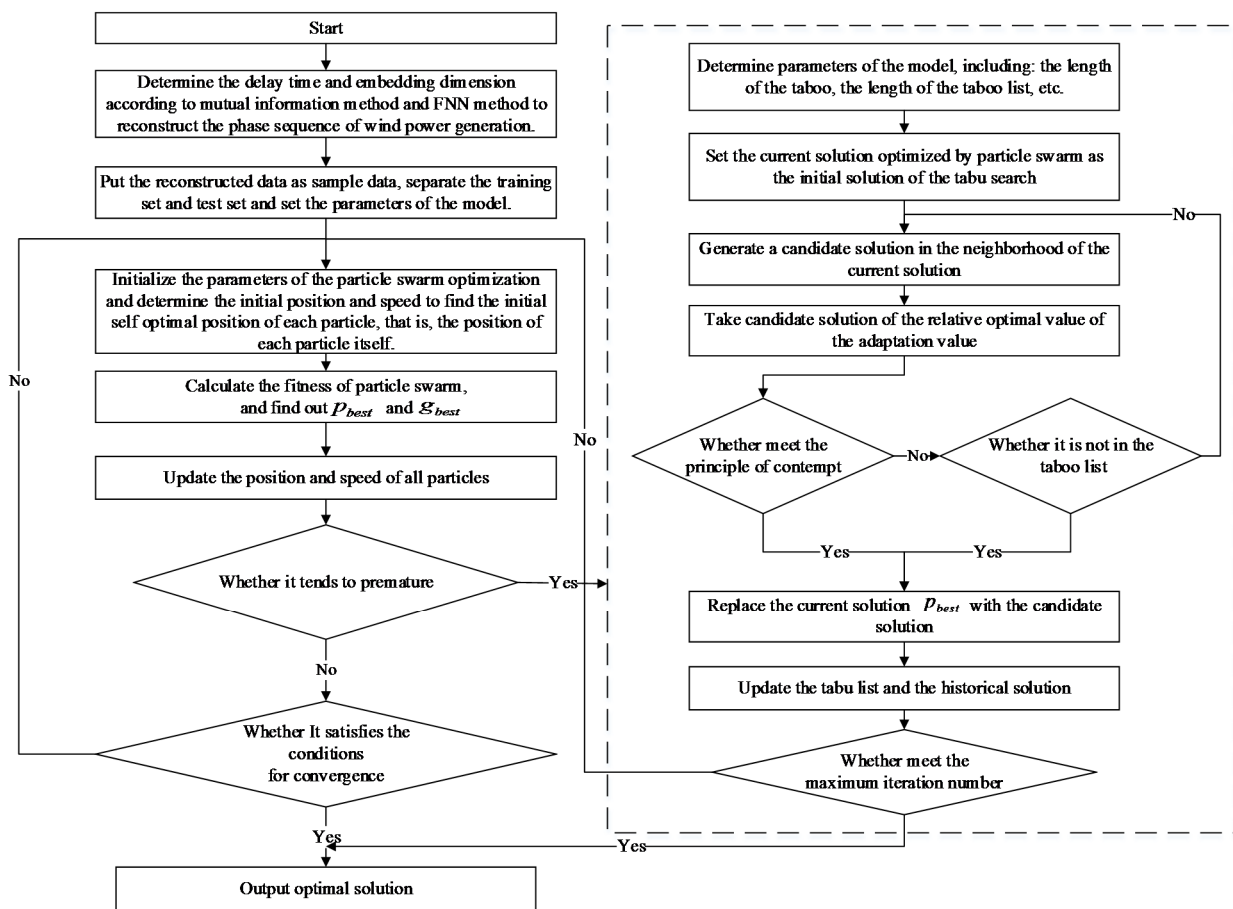


Figure 5. Flow chart of chaos ESN optimized by particle swarm and tabu search algorithm.

5. Empirical Analysis

5.1. Multi Time Scale Forecasting of Short-term Wind Power Generation without Considering Seasonal Factors

Ultrashort-term and short-term wind power forecasting with high prediction accuracy of the model requirements, have great significance for reducing the phenomenon of abandoned wind power, optimizing the conventional power generation plan, adjusting maintenance schedules and developing real-time monitoring systems. In this paper, we consider different time scales and predict the wind power output. Through reading the literature, we can find that for the time scale of wind power generation minutes, hours and days are usually chosen [5–7,16,19–21]. Thus, without considering the seasonal factors, we select a set data from the Jilin wind farm in the second quarter in 2013 for 5 min' ultrashort-term and hour-ahead short-term wind power forecasting. The prediction of day-ahead wind power generation will be discussed in the next section in the seasonal forecasting of wind power.

5.1.1. Ultrashort-Term Wind Power Generation Based on the Time Scale of 5 Min

Wind power output in a certain wind farm from 1 to 5 June 2013 was selected as a sample for analysis, sampling every 5 min and collecting 288 points a day to form a dataset with a total of 1430 sample points for empirical analysis. Among them, the first 1152 data are set as training samples, and the last 278 data in the last four days are used as test samples.

The original data should be normalized before the network study:

$$\bar{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (16)$$

After the normalized treatment, the values of each variable are between $[0, 1]$, which eliminates the dimensional effect. We select the mutual information method for delay time to reconstruct the chaotic time series. As we know, the time when the mutual information function reaches the minimum time for the first time is the optimal value, so as shown in Figure 6. $\tau = 4$ is the optimal delay time.

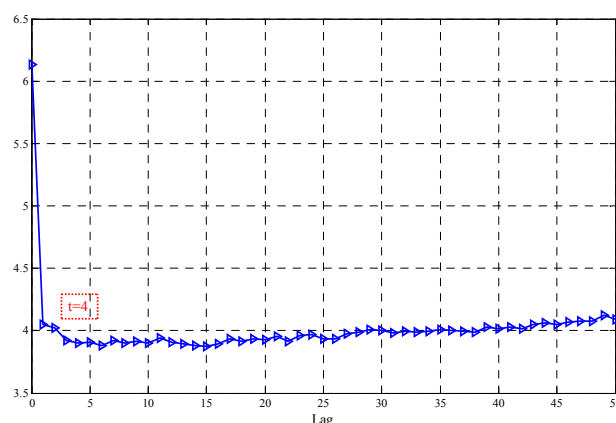


Figure 6. Optimal delay time calculated by mutual information method.

The embedding dimension is determined by the FNN false nearest neighbor point method. When the false nearest neighbor points no longer decreases with the increase of m , the minimum embedding dimension is m . It can be seen from Figure 7 that all discriminant curves of discriminant 1, 2 and the joint curve display when m is greater than or equal to 5, no longer decrease with the increase of m , and tend to be stable. Thus the minimum embedding dimension is $m = 5$.

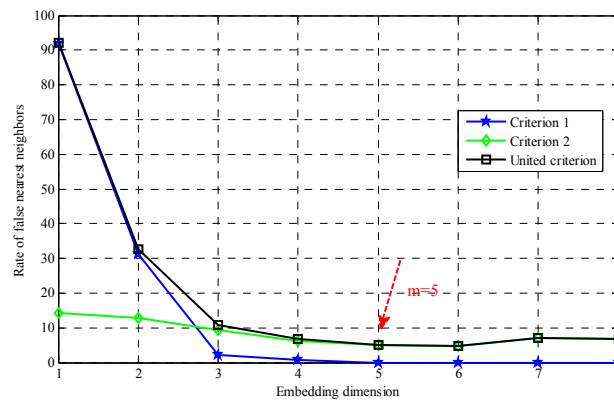


Figure 7. FNN method for minimum embedding dimension.

The phase are reconstructed with the time delay $\tau = 4$ and the embedding dimension $m = 5$. Using the small data volume method, we know that the maximum Lyapunov exponent is $\lambda_1 = 3.1$. The result of $\lambda_1 > 0$ verifies the time series of wind power generation has chaotic properties, so it can be predicted by the chaos method.

The reconstructed training samples are put into the ESN network, choosing the pool dimension of 500 and maintaining a 3% sparse connection. As the spectral radius of the connection weight matrix has no uniform standards and algorithms, the optimal value is found through many experiments. This paper determines a spectral radius of 7.54, pool dimension of 500 and maintain sparse connection of 3% through the calculations.

The activation function of the pool is tansig and the output unit adopts a linear activation function. The output weights of the training method refer to the particle swarm and tabu search training process. At the same time, the standard ESN network and BP neural network are used for comparison. The network parameters of the standard ESN network are the same as ESN network optimized by PSO and TS algorithm (PSOTS-ESN), and the output weights use the least squares method.

Various error metrics between the actual and forecasting data have been defined to assess the forecasting performance. In our experiments, mean absolute percentage error (MAPE), root mean square error (RMSE) and mean absolute error (MAE) are introduced to appraise and compare the different simulation results:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{Y}_i - Y_i}{Y_i} \right| \quad (17)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (18)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (19)$$

Based on the parameter settings, the results of the prediction of the wind power output by the particle swarm tabu search optimized ESN and the comparison methods are illustrated in Figures 8–10 respectively. Overall, the forecasting trends predicted by the four models are consistent with the actual value. The PSOTS-ESN algorithm and the standard PSO-ESN algorithm are closer to the actual curve, while the deviation of BP forecasting curve is relatively large, especially at the peak point of the curve. Error evaluations of the different algorithms are presented in Table 1.

Table 1. Error and computational time comparison of the different algorithms. Echo state network optimized by particle swarm optimization and tabu search algorithm: PSOTS-ESN; particle swarm optimization echo state network: PSO-ESN; echo state network: ESN; back propagation: BP; mean absolute percentage error: MAPE; mean absolute error: MAE; root mean square error: RMSE.

| Algorithm | MAPE (%) | MAE (MW) | RMSE (MW) | Computational time (s) |
|-----------|----------|----------|-----------|------------------------|
| PSOTS-ESN | 3.97 | 8.82 | 2.97 | 144 |
| PSO-ESN | 5.51 | 12.7 | 3.49 | 152 |
| ESN | 7.79 | 14.43 | 3.75 | 155 |
| BP | 8.6 | 15.8 | 4.8 | 201 |

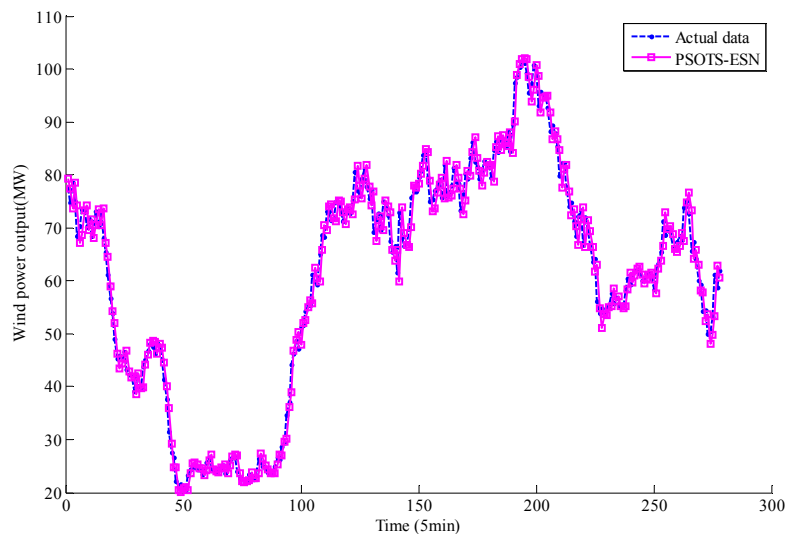


Figure 8. Curves of proposed model forecasting results and actual wind power output.

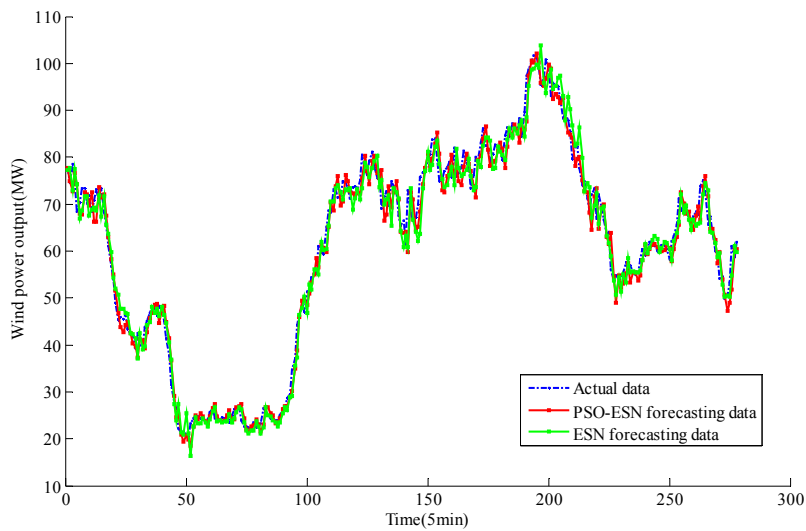


Figure 9. Curves of comparison model forecasting results.

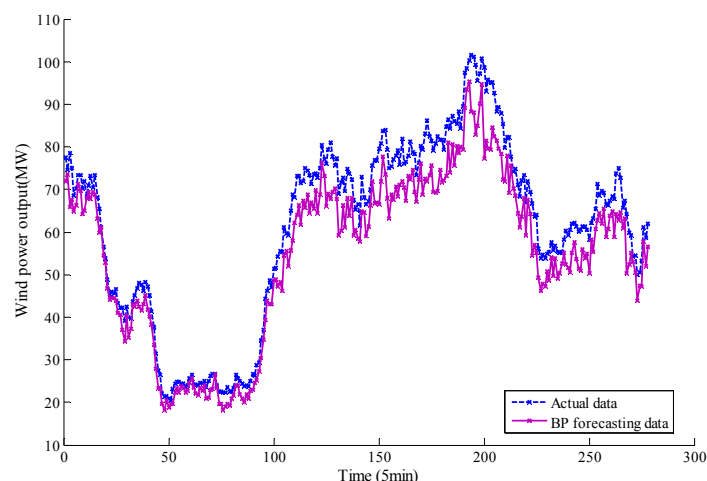


Figure 10. Curves of Back Propagation (BP) model forecasting results.

The error of MAPE, MAE and RMSE of PSOTS-ESN is the smallest, which improves the effect of short-term wind power forecasting. The RMSE and MAE of PSOTS-ESN method are 2.97 MW and 8.82 MW, respectively. The maximum is BP with the value of 4.8 MW and 15.8 MW. Compared with other methods, the proposed method has less computation time and the convergence speed of the algorithm is improved.

From the above tables and figures, it can be concluded that:

- (a) The result of Figures 8 and 9 and Table 1 denote that the proposed PSOTS-ESN model in this paper has more obvious advantages than the PSO-ESN model. The error metrics and computation time of the PSOTS-ESN are all less than those of the PSO-ESN model. The main reason is the fact that the hybrid optimization algorithm makes up for the defects of any single algorithm and chaotic time reconstruction can eliminate the chaos of the original data without regularity which improves the accuracy and the generalization ability of the model and makes for a better prediction effect.
- (b) Comparing Figures 8–10 we can find that the forecasting precision of the combined model is higher than that of the single model. As the combination forecasting model can realize the complementary advantages of different algorithms to improve the accuracy of the algorithm, the prediction of a single model has a large limitation.
- (c) From Figures 9 and 10 and Table 1, we can find that the forecasting precision of the standard ESN is higher than that of the BP neural network. The conclusion can be drawn as follows: for the forecasting of non-stationary time series, the processing power of ESN is better than that of BP, and the prediction accuracy is higher.

5.1.2. Short-Term Wind Power Generation Based on the Time Scale of 1 H

In wind power prediction, hourly prediction is the most common way used. As short-term wind power forecasting has high precision requirements and has important practical significance for the power grid enterprise scheduling, this paper selects 352 consecutive hours of wind power output from a wind farm data for testing to validate the generalization ability of the proposed method. The first 282 data points are used as training set and the latter 70 data are as testing set. The same analysis can be made as proposed in Section 5.1.1.

Figure 11 displays the forecast curves of the proposed PSOTS-ESN method and the comparison with the two BP and ESN model methods. Figures 12 and 13 present the histograms of the relative errors of the three methods.

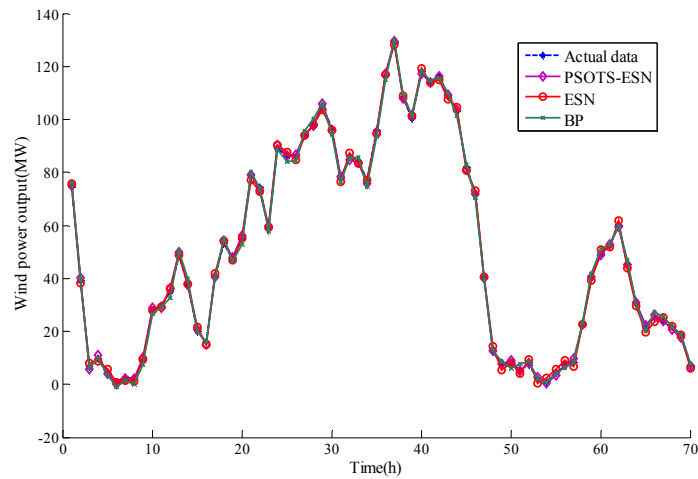


Figure 11. Curves of forecasting model results of echo state network optimized by particle swarm optimization and tabu search algorithm (PSOTS-ESN), ESN and BP.

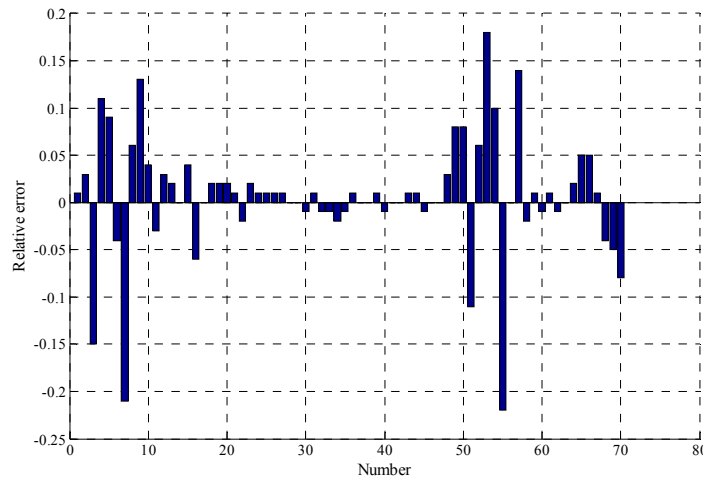


Figure 12. Error distribution histogram of PSOTS-ESN.

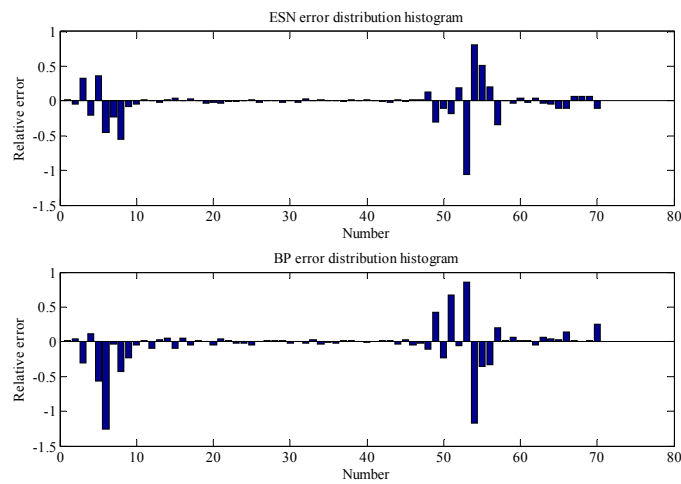


Figure 13. Error distribution histogram of ESN and BP.

Seen from Figures 12 and 13 and Table 2, the method presented in this paper has higher accuracy. The vast majority of the predictive error of the proposed method is within 20%, and 70% of the

relative error is within 5%. At the same time, the numerical values of the three error functions are less than the contrast methods. The prediction results are satisfactory. In the comparison with ESN and BP, the relative error of the prediction results are both within 50%, the overall error of ESN is less than BP. The result shows that the effect of ESN prediction is better than BP model for short-term non-stationary sequence. From the final results, the proposed method demonstrate the excellent performance in the wind power forecasting based on the time scale of 1 h with strong generalization ability and robustness.

Table 2. Error comparison of different algorithms for 1 h forecasting.

| Algorithm | MAPE (%) | MAE (MW) | RMSE (MW) |
|-----------|----------|----------|-----------|
| PSOTS-ESN | 5.9012 | 0.68 | 0.63081 |
| ESN | 9.7283 | 1.03 | 1.3733 |
| BP | 9.0579 | 1.15 | 1.8571 |

5.2. Multi Time Scale Forecasting of Short-term Wind Power Generation considering Seasonal Factors

The wind power output depends mainly on the local wind resources, and the characteristics of wind resources mainly refers to the change of wind speed. There are obvious differences in wind speed and wind direction during the different seasons in China. Consequently, the seasonal factor has an important influence on the wind power forecasting. With the changes of the seasonal cycle, the wind speed in the same season and the same month has a certain periodicity. Based on this feature, wind power output also has a certain periodicity. The seasonal analysis of wind power can be predicted according to the rules of different seasons to improve the prediction accuracy.

The empirical analysis is conducted with the daily data of wind power output for the four quarters from 2014 to 2015 for a wind farm in Jilin. Each quarter selects the previous 70 data points for the training set and the last 20 data points for the testing set, getting the following forecasting results in Figure 14, where, the blue dotted line in Figure 14 represents the original value and the red solid line shows the predicted value.

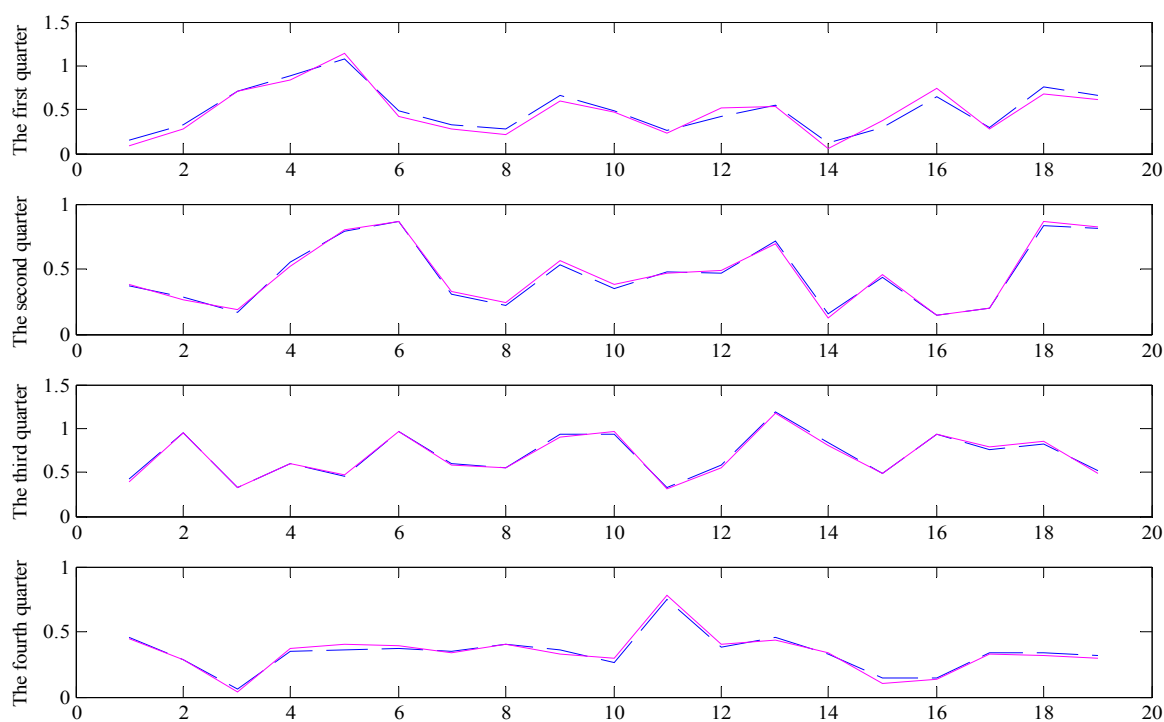


Figure 14. Four quarters' forecasting results by PSOTS-ESN model.

Table 3 shows the error function values predicted by different methods in different seasons, while the error values of forecasting results without considering the seasonal factors are also listed. From the quarterly forecasting results, it can be seen that the forecasting error the quarter and the fourth quarter are relatively small, while the forecasting error in the summer and fall are larger. This occurs because the wind speed and wind direction in summer and autumn are not stable, causing a weak regularity of the wind power output. In the seasonal wind power generation based on the time scale of one day, the prediction error of the proposed method is the smallest and the prediction accuracy is the highest, which verifies the generalization ability and robustness of the proposed method.

In this paper, we take the sample without considering the seasonal factors as a comparison. As can be seen from the results, the prediction error of the PSOTS-ESN method in this paper is relatively smaller, but the overall prediction error is far higher than for the quarterly forecasting results. The wind power generation is influenced by the seasons. Through the above analysis, it can be known that wind power generation has a certain seasonality and periodicity, which is advantageous to improve the overall forecast accuracy by taking into account the seasonal wind power generation.

Table 3. Error comparison of different algorithms for seasonal forecasting.

| Error | First quarter | | | Second quarter | | | Third quarter | | |
|-------|--------------------|--------------------|--------------------|---------------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| | PSOTS-ESN | ESN | BP | PSOTS-ESN | ESN | BP | PSOTS-ESN | ESN | BP |
| MAPE | 3.231 | 4.5 | 5.98 | 8.2 | 10.3 | 12.6 | 8.8 | 9.2 | 10.7 |
| RMSE | 6×10^{-4} | 1×10^{-3} | 2×10^{-3} | 5.5×10^{-4} | 1×10^{-3} | 2×10^{-3} | 8×10^{-4} | 3×10^{-3} | 2×10^{-2} |
| MAE | 2×10^{-2} | 3×10^{-2} | 3×10^{-2} | 2×10^{-2} | 2×10^{-2} | 3×10^{-2} | 5×10^{-2} | 4×10^{-2} | 5×10^{-2} |
| Error | Fourth quarter | | | Non-seasonal factor forecasting | | | - | - | - |
| | PSOTS-ESN | ESN | BP | PSOTS-ESN | ESN | BP | - | - | - |
| MAPE | 5.52 | 9.6 | 13.5 | 11.79 | 13.5 | 18.4 | - | - | - |
| RMSE | 5×10^{-4} | 1×10^{-2} | 2×10^{-3} | 3×10^{-2} | 0.4 | 0.8 | - | - | - |
| MAE | 1×10^{-2} | 2×10^{-2} | 4×10^{-2} | 0.19 | 0.34 | 0.61 | - | - | - |

The unit of MAPE is %, unit of RMSE and MAE is MW.

6. Conclusions

In this paper, a novel composite PSOTS-ESN technique for ultrashort-term and short-term wind power forecasting systems is proposed. The wind power generation results are affected by wind speed, wind direction, temperature and other natural climate factors, showing a certain disorder and regularity. The chaos time series method takes advantage in dealing with uncertain and regularity time series. It can order the chaotic time series into a new sequence through the delay time and the embedding dimension to meet the short-term forecast demands. The combination of particle swarm optimization and tabu search show the advantages of the two algorithms, and also make up for the deficiency of the single algorithm in the implementation process, making the optimization results more accurate. This paper selects the wind power output data from a certain wind farm for multi time scale forecasting. Meanwhile, the influence of seasonal factors on wind power is taken into consideration. The results show that wind power has a seasonal pattern and the original wind power time series demonstrate chaotic characteristics. The proposed PSOTS-ESN is used to predict the multi time scale wind power output. The experiments demonstrate that the method proposed in this paper outperforms other models for the same task and has strong generalization ability and robustness. It is shown that the proposed model is effective to predict wind power time series with its short-time memory ability and achieves satisfactory results.

Acknowledgments: This work was supported by Natural Science Foundation of China (Project no.71471059). It was also funded by the Fundamental Research Funds for the Central Universities (2015XS41) and the Beijing municipal construction project. The authors would like to thank the anonymous reviewers for their valuable comments, which greatly helped us to clarify and improve the contents of the paper.

Author Contributions: In this research activity, all the authors were involved in the data analysis and preprocessing phase, simulation, results analysis and discussion, and manuscript preparation. All authors have approved the submitted manuscript.

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

References

1. Wang, J.Z.; Hu, J.M.; Ma, K.L. A self-adaptive hybrid approach for wind speed forecasting. *Renew. Energy* **2015**, *78*, 374–385. [[CrossRef](#)]
2. Zhang, K.F.; Yang, G.Q.; Chen, H.Y.; Wang, Y.; Ding, X. The wind power prediction error estimation based on data feature extraction method. *Autom. Electr. Power Syst.* **2014**, *38*, 22–27. (In Chinese)
3. Fang, J.X. Study on Prediction Model of Short-Term Wind Speed and Wind Power. Master's Thesis, Beijing Jiaotong University, Beijing, China, 2011.
4. Carpinone, A.; Giorgio, M.; Langella, R.; Testa, A. Markov chain modeling for very-short-term wind power forecasting. *Electr. Power Syst. Res.* **2015**, *122*, 152–158. [[CrossRef](#)]
5. Xue, Y.S.; Yu, C.; Zhao, J.H.; Li, L.; Liu, X.; Wu, Q.; Yang, G. A review on short-term and ultra short term wind power prediction. *Autom. Electr. Power Syst.* **2015**, *39*, 141–151.
6. Wang, Y. Research on Ultra Short-term Wind Power Forecasting and its Ensuing Integration Transient Tripping Control Issues. Doctoral thesis, Zhejiang University, Hangzhou, China, 2014. (In Chinese)
7. Yan, H.M.; Li, X.J.; Ma, X.M.; Hui, D. Wind power output schedule tracking control method of energy storage system based on ultra-short term wind power prediction. *Power Syst. Technol.* **2015**, *39*, 432–439.
8. Gong, W.Y.; Meyer, F.J.; Liu, S.Z.; Hanssen, R.F. Temporal filtering of inSAR data using statistical parameters from NWP models. *IEEE Geosci. Remote. Sens. Soc.* **2015**, *5*, 4033–4044. [[CrossRef](#)]
9. Suomalainen, K.; Pritchard, G.; Sharp, B.; Yuan, Z.; Zakeri, G. Correlation analysis on wind and hydro resources with electricity demand and prices in New Zealand. *Appl. Energy* **2015**, *137*, 445–462. [[CrossRef](#)]
10. Karadas, M.; Celik, H.M.; Serpen, U.; Toksoy, M. Multiple regression analysis of performance parameters of a binary cycle geothermal power plant. *Geothermics* **2015**, *54*, 68–75. [[CrossRef](#)]
11. Cadenas, E.; Jaramillo, O.A.; Rivera, W. Analysis and forecasting of wind velocity in chetumal, quintana roo, using the single exponential smoothing method. *Renew. Energy* **2015**, *35*, 925–930. [[CrossRef](#)]
12. Gao, P. Application of Kalman filter in wind power prediction. *Guide Bus.* **2012**, *13*, 269–270. (In Chinese)
13. Zhang, J.Y.; Wang, C. Application of ARMA Model in Ultra-short Term Prediction of Wind Power. In Proceedings of the 2013 International Conference on Computer Sciences and Applications, Wuhan, China, 14–15 December 2013.
14. Tagliaferri, F.; Viola, I.M.; Flay, R.G.J. Wind direction forecasting with artificial neural networks and support vector machines. *Ocean Eng.* **2015**, *97*, 65–73. [[CrossRef](#)]
15. Zeng, M.; Li, S.L.; Wang, L.; Xue, S.; Wang, R.C. Wind power prediction model based on combinatorial optimization algorithm of ARMA model and BP neural network. *East China Electr. Power* **2013**, *41*, 347–352.
16. Shi, H.T.; Yang, J.L.; Ding, M.S.; Wang, J.M. A short term wind power prediction method based on wavelet decomposition and BP neural network. *Autom. Electr. Power Syst.* **2011**, *35*, 44–48.
17. Zhang, Q.; Lai, K.K.; Niu, D.X.; Wang, Q.; Zhang, X. A fuzzy group forecasting model based on least squares support vector machine (LS-SVM) for short-term wind power. *Energies* **2012**, *5*, 3329–3346. [[CrossRef](#)]
18. Zeng, J.W.; Qiao, W. Short-term solar power prediction using a support vector machine. *Renew. Energy* **2013**, *52*, 118–127. [[CrossRef](#)]
19. Tang, X. Short-Term Wind Speed Prediction Based on Bayesian Inference. Master's Thesis, North China Electric Power University, Beijing, China, 2013. (In Chinese)
20. Wang, L.J.; Liao, X.Z.; Gao, S.; Dong, L. Chaos Characteristics Analysis of Wind Power Generation Time Series for a Grid Connecting Wind Farm. *Trans. Beijing Inst. Technol.* **2007**, *27*, 1077–1080.
21. Luo, H.Y.; Liu, T.Q.; Li, X.Y. Chaotic forecasting method of short-term wind speed in wind farm. *Power Syst. Technol.* **2009**, *33*, 67–71.
22. Guo, C.X.; Wang, Y.; Shen, Y.; Wang, M.; Cao, Y. Multivariate local prediction method for short-term wind speed of wind farm. *Proc. CSEE* **2012**, *32*, 24–31.
23. Chen, M. *Chaotic Time Series Based on BP Neural Network*; Central South University: Changsha, China, 2007.
24. 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](#)]
25. Zhang, Z.Y.; Wang, T.; Liu, X.G. Melt index prediction by aggregated RBF neural networks trained with chaotic theory. *Neurocomputing* **2014**, *131*, 368–376. [[CrossRef](#)]

26. Luo, Z. The comparative study of traffic flow prediction based on ESN and Elman neural network. *J. Hunan Univ. Technol.* **2013**, *27*, 67–72. (In Chinese)
27. Zhao, X.M.; Song, Z.H.; Li, P. The improved GMDH-type neural network and its application to forecasting chaotic time series. *J. Circuits Syst.* **2012**, *1*, 13–17.
28. Song, T.; Li, H. Chaotic time series prediction based on wavelet echo state network. *Acta Phys. Sin.* **2012**, *61*, 1–7.
29. Wang, J.M. The Method of Nonlinear Time Series Prediction Based on Echo State Network. Ph.D. Thesis, Harbin Institute of Technology, Harbin, China, 2011. (In Chinese)
30. Jaeger, H.; Haas, H. Harnessing non-linearity: Predicting chaotic systems and saving energy in wireless communication. *Science* **2004**, *304*, 78–80. [[CrossRef](#)] [[PubMed](#)]
31. Niu, D.X.; Ji, L.; Wu, H.M. Prediction of daily maximum load based on the Bayesian framework and echo state network. *Power Syst. Technol.* **2012**, *11*, 1–6. (In Chinese)
32. Zheng, Y.K. Short-term Load Forecasting Based on Phase Space Reconstruction and Support Vector Machine. Ph.D. Thesis, Southwest Jiaotong University, Chengdu, China, 2008. (In Chinese)
33. Guo, Z.H.; Chi, D.Z.; Wu, J. A new wind speed forecasting strategy based on the chaotic time series modelling technique and the Apriori algorithm. *Energy Convers. Manag.* **2014**, *84*, 140–151. [[CrossRef](#)]
34. Marin, C.I.; Arias, A.E.; Artigao Castillo, M.M. A distributed memory architecture implementation of the False Nearest Neighbors method based on distribution of dimensions. *J. Supercomput.* **2012**, *59*, 1596–1618. [[CrossRef](#)]
35. Liu, D.; Wang, J.L.; Wang, H. Short-term wind speed forecasting based on spectral clustering and optimized echo state networks. *Renew. Energy* **2015**, *78*, 599–608. [[CrossRef](#)]
36. Deng, H.L.; Li, X.Q. Stock price inflection point prediction method Based on chaotic time series analysis. *Stat. Decis.* **2007**, *5*, 19–20.
37. Niu, D.X.; Lu, Y.; Xu, X.M.; Li, B. Short-Term Power Load Point Prediction Based on the Sharp Degree and Chaotic RBF Neural Network. *Math. Prob. Eng.* **2015**, *2015*. [[CrossRef](#)]
38. Niu, D.X.; Zhao, L.; Zhang, B.; Wang, H.F. Application of particle swarm grey model in power load forecasting. *Chin. J. Manag. Sci.* **2007**, *15*, 69–73. (In Chinese)
39. Niu, D.X.; Li, J.C.; Li, J.Y. Middle-long power load forecasting based on particle swarm optimization. *Comput. Math. Appl.* **2009**, *57*, 1883–1889. [[CrossRef](#)]
40. Wang, S.L.; Li, Z.T.; Xing, M. Genetic algorithm and tabu search algorithm for the optimization of neural network structure. *Comput. Appl.* **2007**, *27*, 1426–1429.
41. Qiu, H.; Ming, W.; Zhang, Z.Z. A tabu search algorithm for the multi-period inspector scheduling problem. *Comput. Oper. Res.* **2015**, *59*, 78–93.
42. Yao, J.; Fang, Y.J.; Chen, G. Application of hybrid genetic and tabu search algorithm in load distribution unit. *Proc. CSEE* **2010**, *30*, 95–100.



© 2015 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/>).