Knowledge-Based Systems 56 (2014) 226–239

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting

Chao Ren [a,b], Ning An [b,c,*], Jianzhou Wang [a], Lian Li [a,b], Bin Hu [d], Duo Shang [e]

[a] School of Mathematics and Statistics, Lanzhou University, Lanzhou, Gansu, PR China
[b] Advanced Computing Lab, School of Information Science and Engineering, Lanzhou University, Lanzhou, Gansu, PR China
[c] Gerontechnology Lab, Hefei University of Technology, Hefei, Anhui, PR China
[d] School of Information Science and Engineering, Lanzhou University, Lanzhou, Gansu, PR China
[e] College of Engineering and Applied Science, Stony Brook University, Stony Brook, NY, USA

ARTICLE INFO

Article history:
Received 10 January 2013
Received in revised form 15 November 2013
Accepted 16 November 2013
Available online 23 November 2013

Keywords:
BP neural network
Input parameters selection
Particle swarm optimization algorithm
Wind speed
Wind forecasting

ABSTRACT

As a clean and renewable energy source, wind energy has been increasingly gaining global attention. Wind speed forecast is of great significance for wind energy domain: planning and design of wind farms, wind farm operation control, wind power prediction, power grid operation scheduling, and more. Many wind speed forecasting algorithms have been proposed to improve prediction accuracy. Few of them, however, have studied how to select input parameters carefully to achieve desired results. After introducing a Back Propagation neural network based on Particle Swam Optimization (PSO-BP), this paper details a method called IS-PSO-BP that combines PSO-BP with comprehensive parameter selection. The IS-PSO-BP is short for Input parameter Selection (IS)-PSO-BP, where IS stands for Input parameter Selection. To evaluate the forecast performance of proposed approach, this paper uses daily average wind speed data of Jiuquan and 6-hourly wind speed data of Yumen, Gansu of China from 2001 to 2006 as a case study. The experiment results clearly show that for these two particular datasets, the proposed method achieves much better forecast performance than the basic back propagation neural network and ARIMA model.

1. Introduction

Insufficient energy resources have become one of the major obstacles faced by the global sustainable development. As one of the most promising renewable energy sources, wind energy has increasingly drawn worldwide attention. According to the World Wind Energy Association [1], by the end of 2010, worldwide wind energy capacity reached 196,630 MW and all wind turbines installed worldwide can generate 430 TW hours per annum, which is 2.5% of the global electricity consumption. In some countries, wind energy has become one of the most important sources of electricity supply. The highest utilization rate of wind power can be found in three European countries: Denmark (21%), Portugal (18%) and Spain (16%). With the continuous growth of its capacity, wind power becomes increasingly important in the entire power system. To a larger extent, the output of the wind power is determined by the wind speed in wind farm. Because of its dependency on temperature, atmospheric pressure, elevation and terrain, wind is intermittent and volatility. This brings great uncertainty to wind speed, in turn to wind power output. Hence, accurate wind speed forecast is of great significance for the wind farm operation control, wind power prediction, power grid operation scheduling, planning and design of wind farms, and more. Especially, short-term and ultra-short-term wind speed forecast play an important role in the daily operation of wind farms and the power grid.

There have been a lot of researches on wind speed forecast [2–8]. Based on the time horizon of forecast, wind speed forecast can be grouped into three categories: ultra-short-term forecast, short-term forecast and mid-and-long term forecast. As to the short-term wind speed forecast models, there are four sub-categories [9]: the physical model, the statistical model, the spatial correlation model, and the artificial intelligence model.

The physical model forecasts the wind speed mainly through the terrain features, atmospheric pressure, ambient temperature and other meteorological information. Usually the physical model is only the first step in the forecast wind speed forecast, i.e. it is used as an auxiliary input for other statistical models. The statistical model is also known as stochastic time series model: it uses historical data, pattern recognition, parameter estimation, model validation and other tools to create a mathematical model to forecast wind speed. According to Jenkins's method [10], the statistical model can be further divided into the following categories: Autoregressive model (AR), Moving average model (MA), Autoregressive moving average model (ARMA) [11] and Autoregressive integrated moving average (ARIMA) [7]. Unlike other models, the spatial

* Corresponding author. Tel.: +86 180 1995 6087; fax: +86 551 2904642.
E-mail address: ning.an@ieee.org (N. An).

0950-7051/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2013.11.015

correlation model considers not only the wind speed of the given wind farm, but also wind speed at several adjacent locations. More specifically, it uses the spatial correlation between the wind speed data of several locations for the actual wind speed prediction. Recently, with the development of artificial intelligence techniques, many researchers proposed to apply them in wind speed forecast, including artificial neural network (ANN) [2,12–15], fuzzy logic method [16], Support Vector Machine (SVM) [17–19] and other mixed methods [20,21].

As a forecast technique, artificial neural network (ANN) has been widely used in many different domains [22–24], such as load forecasting [25,26] and wind speed forecast. Compared with other forecast methods, ANN methods are advantageous in terms of high data error tolerance, easy adaptability to online measurements, no need for excess information other than wind speed history. A Back Propagation Neural Network (BPNN) [27] is a typical artificial neural network. It is essentially a mapping function from input vector(s) to output vector(s) without knowing the correlation between the data. It can implement any complex nonlinear mapping function proved by mathematical theories, and approximate an arbitrary nonlinear function with satisfactory accuracy [28] . After learning the data trends from historical data, BPNN can be used effectively to forecast new data.

While BPNN algorithm tries to find the global solution of complex nonlinear functions, it depends on local searches which can lead to local minima and cause the failure of training. In addition, the trainability of BPNN can influence its predictability: normally the poor trainability means the poor predictability and the good trainability means the good predictability. When the improvement of trainability reaches a certain point, it can actually have a negative effect on the predictability of BPNN. This is so called overfitting phenomenon: we are learning the noise information contained inside the training dataset. Overfitting is a phenomenon in which a neural network gets very good at dealing with one data set at the expense of becoming very bad at dealing with other data sets. The overfitting problem refers to exceeding some optimal ANN size of ANN training that may finally result in worse predictive ability of a network [29].

Literature reviews showed us that while a large number of forecast methods have been studied, no particular one is found to be suitable for all data sets. As a point of illustration, Moghram and Rahman [30] reviewed five short-term load forecasting methods: (1) Multiple linear regression, (2) time series, (3) general exponential smoothing, (4) state space and Kallman filter, and (5) knowledge-based approach. After using each one of these methods to predict the hourly load of Southeastern utility, they compared predication results of summer loads and winter loads respectively in terms of precent error and found no one as the overall best method. For example, the transfer function model of time series approach mentioned earlier can achieve best forecast over the summer months, but get worst forecast over the winter months. Because of its strong dependency on historical data, the transfer function approach did not respond well to abrupt changes as did the knowledge-based approaches. Moghram and Rahman concluded that there is no one best approaches for all cases: we should analyse and understand model performance under specific conditions, and make incremental improvements based on knowledge gained.

Appropriately processing input samples is crucial for wind speed predication. In BPNN, different input dimensions and/or different size of training data can produce different forecast results with various accuracies. Few researchers have paid sufficient attention to this, however, until recently. Li and Shi [31] found that the selection of neural network type and parameters greatly affects the performance of wind speed forecasting by comparing three different artificial neural networks (ANNs). They also pointed out that data sources play an important role in selecting type of neural networks.

In this paper, we first introduce back propagation neural network based on particle swam optimization method [32] (PSO-BP for short), it has been widely used in the field of wind power [25,33,34]. Then we propose a method called IS-PSO-BP, which is short for Input parameter Selection (IS)-PSO-BP, to achieve better predict performance by combining the PSO-BP neural network with input parameters selection. Experiment results demonstrate that without carefully chosen parameters, BP neural network actually has better forecast performances than PSO-BP neural network. On the other hand, once appropriate parameters are selected, the forecast performance of the PSO-BP neural network is significantly better than the BP neural network. This clearly shows how important it is to select appropriate parameters to achieve a desired prediction performance.

The rest of the paper is organized as follows. After the real world data sets are introduced in Section 2, Section 3 provides a general description of BP neural network. In Section 4, the particle swarm optimization algorithm and PSO-BP model are discussed in detail. Particularly, we will introduce optimal parameters selection for PSO-BP model: this process includes two selection methods for dataset and one selection method for PSO parameters. The overall structure of IS-PSO-BP model will be given as well. Section 6 presents and discusses the experiment results. Finally, the conclusion remarks are provided.

2. Data set

As shown in Fig. 1, Jiuquan prefecture of Gansu province, which is located at the west end of Hexi Corridor, is part of the ancient Northern Silk Road. The entire Jiuquan prefecture stretches for 680 km from east to west and 550 km north to south, occupying 191,342 km^2; its population as of 2010 was 1,100,000 [35]. Within Jiuquan, Guazhou is known as "the World Storehouse of Wind Energy", and Yumen city is called "Wind Gap of the World". Taking advantage of its exceptional rich wind resource, Jiuquan started to build China's first 10-million-kilowatt wind power base that will grow into "Land Three Gorges" project. In accordance with the planning [36], the installed capacity of Jiuquan wind power base would reach 12.71 million kilowatts in 2020. Annually, assuming it generates 2300 h of full load power, the Jiuquan wind power base is expected to save about 9.72 million tons of standard coal, and reduce soot emissions, sulfur dioxide emissions and carbon dioxide emissions by about 13 million tons, 110,000 tons and 2930 tons respectively. These are obvious benefits in terms of energy saving and emission reduction. As discussed earlier, accurate forecasting daily and hourly average wind speed will help make this a reality in wind power base.

We collect the data of daily average wind speed in Jiuquan prefecture and 6-hourly wind speed data in Yumen city from 1 January 2001 to 31 December 2006 to study the performance of various forecasting methods. Figs. 2 and 3 show daily average wind speed of Jiuquan prefecture and 6-hourly wind speed of Yumen city for each year respectively. The forecasting methods under investigation will forecast the daily and 6-hourly average wind speed.

3. BP neural network

As the BP neural network is used as a baseline method in our study, we briefly describe it in this section. The structure of a BP neural network is shown in Fig. 4. Each node in the network is a neuron whose function is to calculate the inner product of the input vector and weight vector by a nonlinear transfer function to get a scalar result. This particular network is a three-layer network: the input

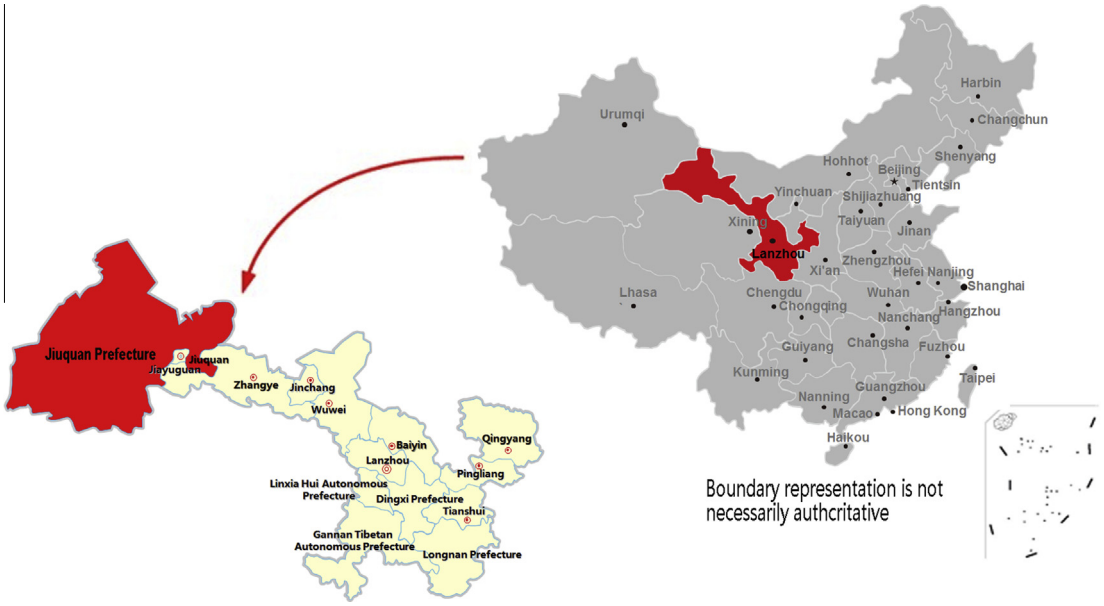

Fig. 1. Location of Jiuquan Prefecture within Gansu,China.

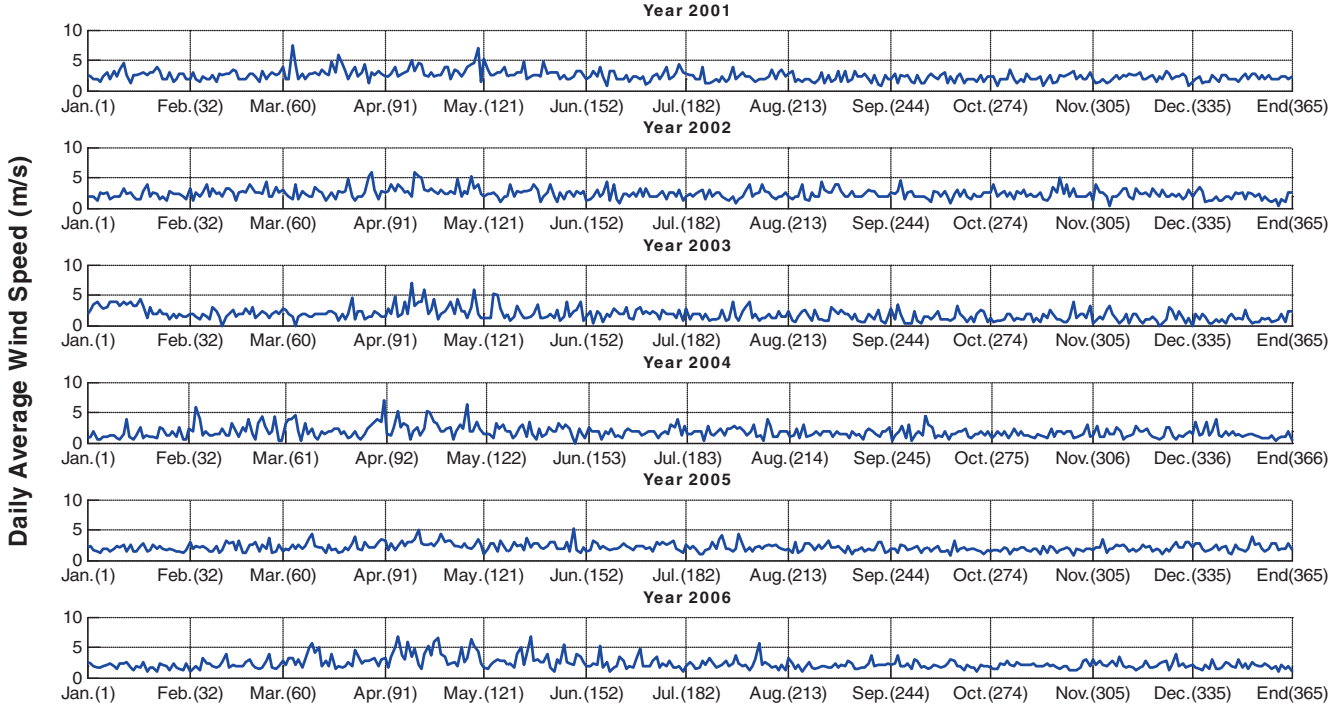

Fig. 2. Daily average wind speed of Jiuquan prefecture from 1 January 2001 to 31 December 2006.

layer, the hidden layer and the output layer. The five nodes in the input layer represent five history daily and hourly average wind speed value respectively. We choose five nodes as input layer is based on extensive experiments: when the input layer has five nodes, the forecast result is much better compared to other cases. The only one node in the output is the daily wind speed forecast value. As we know, the hidden layer affects the robustness of the neural network. To achieve better prediction results, we use the Hecht–Nelson method [37] to determine the node number of the hidden layer: when the node number of the input layer is n, the node number of the hidden layer is $2n + 1$. With n input neurons, $2n + 1$ hidden neurons, and one output neuron, the training process of BP network can be described as follows. To ensure the quality of forecast results, we adopt normalized method to treat the input and output data in advance of training the network, the formula is as follow:

$$X' = \{X'i\} = 2 \times \frac{Xi - Xi\min}{Xi\max - Xi\min} - 1, \quad i = 1,2,\ldots,n \; X' \subset [1,1] \tag{1}$$

where X_{imin} and X_{imax} are the minimum and maximum value of input array or output vectors, and X_i denotes the real value of each vector.

- Step 1. Calculate outputs of all hidden layer nodes.

$$y_j = f\left(\sum_i w_{ji}x_i + b_j\right) = f(net_j) \quad (i = 1,\ldots,n; j = 1,\ldots,2n+1) \tag{2}$$

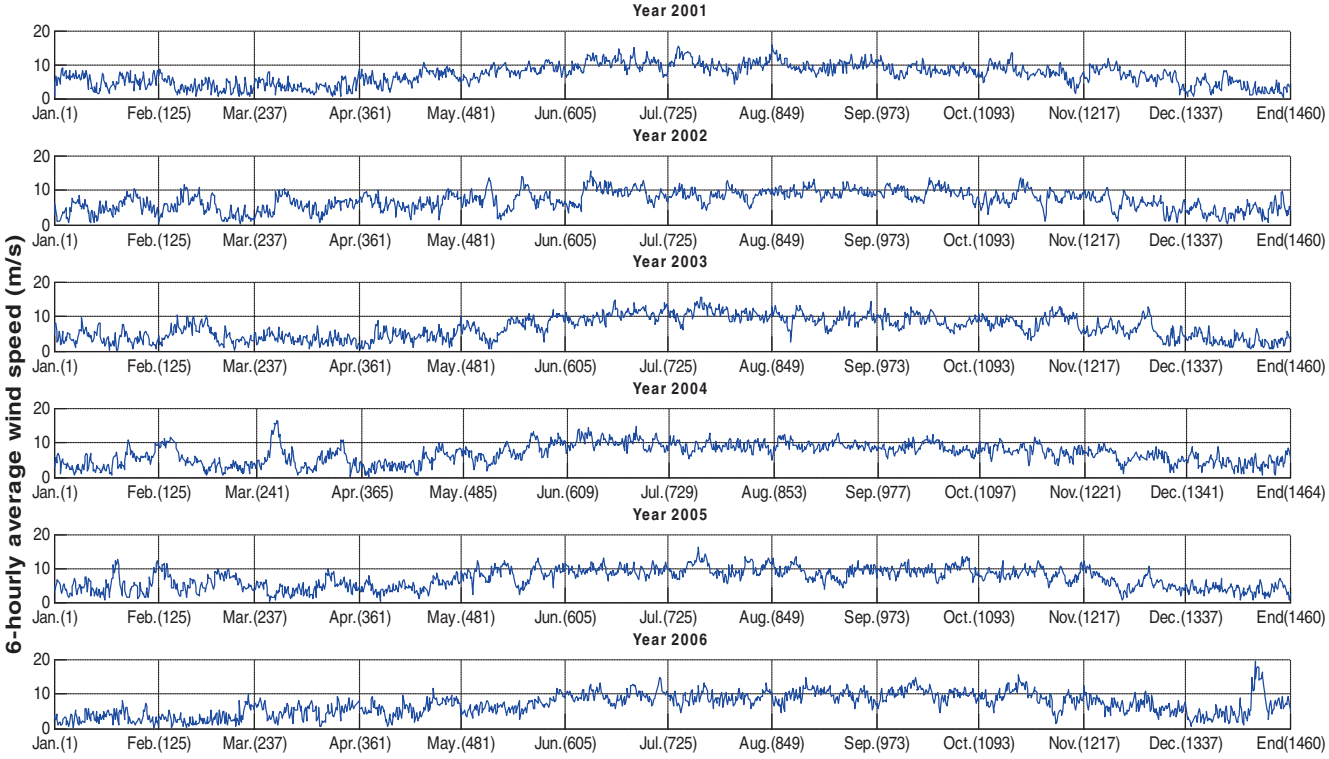

Fig. 3. 6-Hourly average wind speed of Yumen city from 1 January 2001 to 31 December 2006.

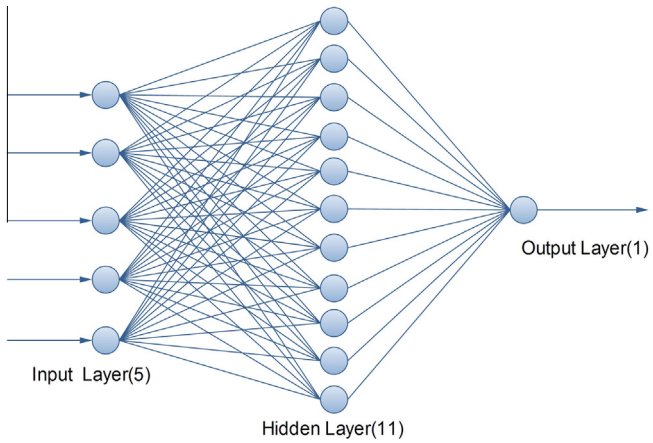

Fig. 4. The structure of the BP neural network.

$$net_j = \sum_i w_{ij} y_j + b_j \quad (j = 1, \ldots, 2n+1) \tag{3}$$

where the activation value of node j is net_j, w_{ji} represent the connection weight from input node i to hidden node j, b_j stands for the bias of the neuron j, y_j stands for the output of hidden layer node j, and f is the activation function of a node, which is usually a sigmoid function.

- Step 2. Calculate output data of neural network.

$$O_1 = f_o\left(\sum_j w_{oj} y_j + b_o\right) \quad (j = 1, \ldots, 2n+1) \tag{4}$$

where w_{oj} represent the connection weight from hidden node j to output node o, b_o stands for the bias of the neuron, O_1 stands for the output data of network, and f_o is the activation function of output layer node.

- Step 3. Minimize the global error E via the training algorithm.

$$E = \frac{1}{2}\sum (O_1 - R_t)^2 \tag{5}$$

R_t is the real output of training data.

4. BP neural network optimized by Particle Swarm Optimization algorithm (PSO)

4.1. Particle Swarm Optimization algorithm

Inspired by the flocking behavior, the PSO algorithm is an evolutionary computation technique that is often used for optimization. In the basic version of the PSO algorithm, a swarm of particles keeps moving around in a search-space according to a few simple formulae, and a satisfactory solution will eventually emerge at the end. Suppose the search space is a D-dimensional space and the number of particle swarm is n, then the particle i of swarm can be represented by a D-dimensional vector: $x_i = (x_{i1}, x_{i2}, \ldots, x_{iD})$ $(i = 1, 2, \ldots, n)$ and the speed of particle i is $V_i = (V_{i1}, V_{i2}, \ldots, V_{iD})$ $(i = 1, 2, \ldots, n)$. The algorithm can be described as follow:

- Step 1. Randomly initialize a group of m particles with positions and velocities.
- Step 2. Calculate the fitness value of each particle.
- Step 3. Calculate the position of best fitness value of each particle from its historical movement, which called *pbest*.
- Step 4. Calculate the position of best fitness value of all particles from global historical movement, which called *gbest*.
- Step 5. Update particles' speed and position using the next two formulas.

$$V_i^{t+1} = \omega^t V_i^t + c_1\tau_1(pbest^t - x_i^t) + c_2\tau_2(gbest^t - x_i^t) \tag{6}$$

$$X_i^{t+1} = X_i^t + V_i^{t+1} \tag{7}$$

Here, V_i^t is the velocity of particle i at iteration t. As V_i^t is uncontrollable, a particle will cycle beating in the problem space [27], in order to inhibit the erratic beating, the speed is often limited to a value within $[-v_{max}, v_{max}]$ [38]; x_i^t represents the position of i particle i at iteration t; τ_1 and τ_2 are two uniform random number from [0,1]; c_1 and c_2 are learning factor which also called acceleration constants as they control how far a particle can move in a single iteration. Generally, $c_1 = c_2 = 2$ were used [4,38]. As previously introduced, $pbest^t$ and $gbest^t$ are the individual best position and the global best position of all particles at iteration t respectively. The variable ω^t is inertia weight at iteration t which is defined by Eq. (8) below. A larger inertia weight led to the global exploration and a smaller inertia weight tends to facilitate the local exploration to fine-tune the current search area.

$$\omega^t = \omega_{max} - (\omega_{max} - \omega_{min}) \times \frac{iter}{itmax} \tag{8}$$

where ω_{max} and ω_{min} are maximum and minimum of inertia weight respectively, which are suggested to be 0.9 and 0.4 respectively [39]; *iter* is the iteration t; *itmax* is maximum of iteration number.

- Step 6. If the end condition is not satisfied, loop back to step 2 again. The end condition usually is a predetermined *itmax* value or adapts fitness value.

4.2. Chaotic search process

We use chaos theory with the PSO method to avoid it falling into local optima early in the search process. This chaotic search procedure can be described as follow:

- Step 1. Calculate particles' average distance D(t) and fitness varianceσ^2 using Eqs. (9) and (10). If $D(t) < \alpha$ *or* $\sigma^2 < \beta$, go to step 2.

$$D(t) = \frac{1}{N \bullet L} \bullet \sum_{i=1}^{N} \sqrt{\sum_{d=1}^{N} (p_{id} - \overline{p_d})^2} \tag{9}$$

$$\sigma^2 = \sum_{i=1}^{N} \left(\frac{f_i - f_{avg}}{f} \right)^2 \tag{10}$$

where N is the population size of particles. L represents the diagonal maximum length in searching space. D is the dimension of the solution space. P_{id} is the dth dimension coordinate value of particle i and $\overline{P_d}$ is the average coordinate value of the dth dimension. The variable f_i represents the fitness value of particle i, f_{avg} is the average fitness value of all particles and f is normalization scale factor.

- Step 2. Perform the chaotic search. Randomly initialize a chaos variabley_0, using y_0 to generate chaotic sequences until the chaos iteration $n > M$, using the best point to randomly replace a particle. Then go to verify the end condition of PSO procedure.

4.3. BP neural network optimized by Particle Swarm Optimization algorithm

The BP neural network optimized by Particle Swarm Optimization algorithm is also called PSO-BP algorithm, it takes the weights and biases of neurons trained as one particle for PSO algorithm. The fundamental idea of PSO-BP algorithm can be described as follows:

- Step 1. Normalize the training dataset and the testing dataset into [−1 1].
- Step 2. Randomly initialize a group of m particles with the number is m, including positions and speed velocities.
- Step3. Compute every particle's fitness value: Referring to complexity of training dataset, LM algorithm or the Conjugate gradient algorithm is used to train the BP neural network. When the dataset is simple, the LM algorithm is used, otherwise Conjugate algorithm is chosen. Update the weights and biases of each neuron using current *gbest* value. The performance function selects the MSE function: where MSE denotes the mean sum of squares of the network errors.
- Step 4. Update particles' speed and position using the Eqs. (6) and (7).
- Step 5. If the end condition is not satisfied, go to step 2 again, or else, go to step 6. The end station usually is set to a previously determined *itmax* or adapts fitness value.
- Step 6. Update the weights and biases of BP neural network by PSO algorithm and the network can be used for forecasting.
- Step 7. Renormalize the forecasting results from [−1 1].

5. Optimal parameters selection for BP neural network based on particle swarm optimization

In order to achieve better performance, we denote the hybrid approach by integrating Input parameters Selection method and PSO-BP neural network (IS-PSO-BP model for short). The first step of IS-PSO-BP model is dealing with PSO parameters and input dataset selection. The flow chart of IS-PSO-BP model is shown in Fig. 9.

5.1. PSO parameters selection

Experimenting with Jiuquan wind speed dataset, we discover the best fitness values which are plotted in Fig. 5. The max iteration number was set to 1000 in the experiments and the plots are similar, the PSO method begins to converge at the 300th iteration. To ensure the better forecast result, we set the iteration number from 10 to 300, and the step length is 10. As the results shown, not the greater of iteration number is set, the forecast result is better. For different datasets, the max iteration number is different. So we combine PSO parameters selection method with PSO-BP neural network.

5.2. Input datasets selection

For the same dataset, different input dataset selection methods lead to different forecast results. As we pay more attention to the input dataset selection, two methods are used in this study: lateral data selection and longitudinal data selection.

5.2.1. Lateral data selection

The first data selection method we used is called lateral data selection. Firstly, the original dataset be numbered from 1 to N according to time sequence. Secondly, use IS-PSO-BP model to determine the dimension (d) of input dataset, which is useful to select training samples and testing samples. Let us suppose i be the starting point of the sample selection. For example, the first group of training sample input is a vector from $i + 1$ to $i + d$, and the output is $i + d + 1$; the second group of training sample input is a vector from $i + d + 1$ to $2d + i$, and the output is $i + 2d + 1$; so the kth group of training sample input is a vector from $i + d(k - 1) + 1$ to $i + kd$, and the output is $i + kd + 1$. Testing samples are selected according to the following rules: the first group of testing sample input is a vector from $i + 2$ to $i + d + 1$, and the output is $i + d + 2$; the second group of testing sample input is a vector from $i + d + 2$ to $i + 2d + 1$, and the output is $i + 2d + 2$; so the kth group of testing sample input is a vector from $i + d(k - 1) + 2$ to $i + kd + 1$, and the output is $i + kd + 2$. The training dataset and testing dataset selection procedure for Jiuquan daily wind speed dataset are shown in Fig. 6, and the input dimension is five.

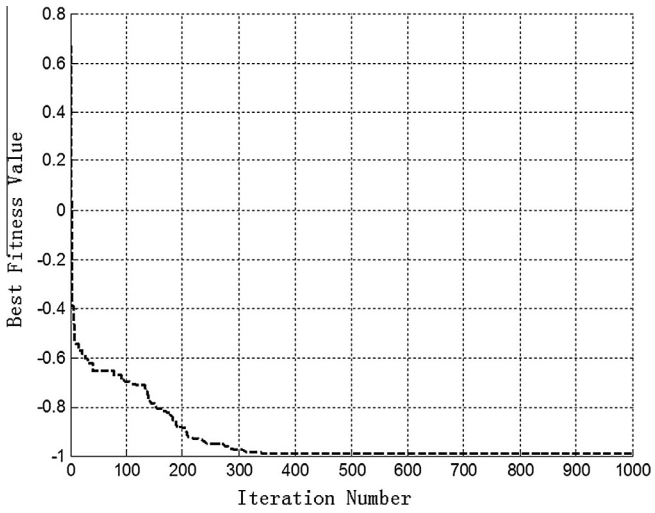

Fig. 5. The plot of best fitness value.

5.2.2. Longitudinal data selection

The other data selection method we used is longitudinal data selection. Firstly, it divides the original dataset into subsets according to a particular date of the month. For example, January 1st 2001, February 1st 2001, until December 1st 2006 was all in the first subset; similar for other 30 subsets for daily wind speed data. For the 6-hourly wind speed data, hour 0 January 1st 2001, hour 0 February 1st 2001, until hour 0 December 1st 2006 were all in the first subset; similar for other 123 subsets. Secondly, it uses the dimension (d) of input dataset determined by IS-PSO-BP model to select training samples and testing samples. Let us suppose i be the starting point of the sample selection. For example, the first group of training sample input is a vector from $i+1$ to $i+d$, and the output is $d+i+1$ from the first subset; the second group of training sample input is a vector from $i+1$ to $i+d$, and the output is $i+d+1$ from the second subset; so the kth group of training sample input is a vector from i to $i+d$, and the output is $i+d+1$ from the kth subset. Testing samples are selected according to the following rules: the first group of testing sample input is a vector from $i+2$ to $i+d+1$, and the output is $i+d+2$ from the first subset; the second group of testing sample input is a vector from $i+2$ to $i+d+1$, and the output is $i+d+2$ from the second subset; so the kth group of testing sample input is a vector from $i+2$ to $i+d+1$, and the output is $i+d+2$ from the kth subset. The longitudinal selection procedure of training dataset and testing dataset for Jiuquan daily wind speed dataset are shown in Fig. 7, and for Yumen hourly wind speed dataset are shown in Fig. 8, and the input dimension here is five. This method includes both horizontal characteristics and longitudinal information.

5.3. BP neural network based on PSO algorithm with Parameters Selection

To obtain better forecast accuracy, we propose the novel hybrid PSO-BP algorithm with a parameters selection method. For the basic PSO-BP neural network, forecast accuracy is relevant to the parameters, such as the iteration max number, input dimension and training number setting. If the parameters are not suitable, the result of PSO-BP neural network will be no better than the traditional BP neural network. But if the parameters are trained, the result will be enhanced. We provide the overall workflow of IS-PSO-BP model in Fig. 9, and its procedure can be broken down into the following steps:

- Step 1. Determine the topology of the network. In this study, we use three-layer network.
- Step 2. Load the input dataset, finish parameter initialization including data selection method chosen (DataType), dimension (Dim) and training number setting (TrainNum). If DataType is chosen to 1, the TrainNum is 31/124, using the

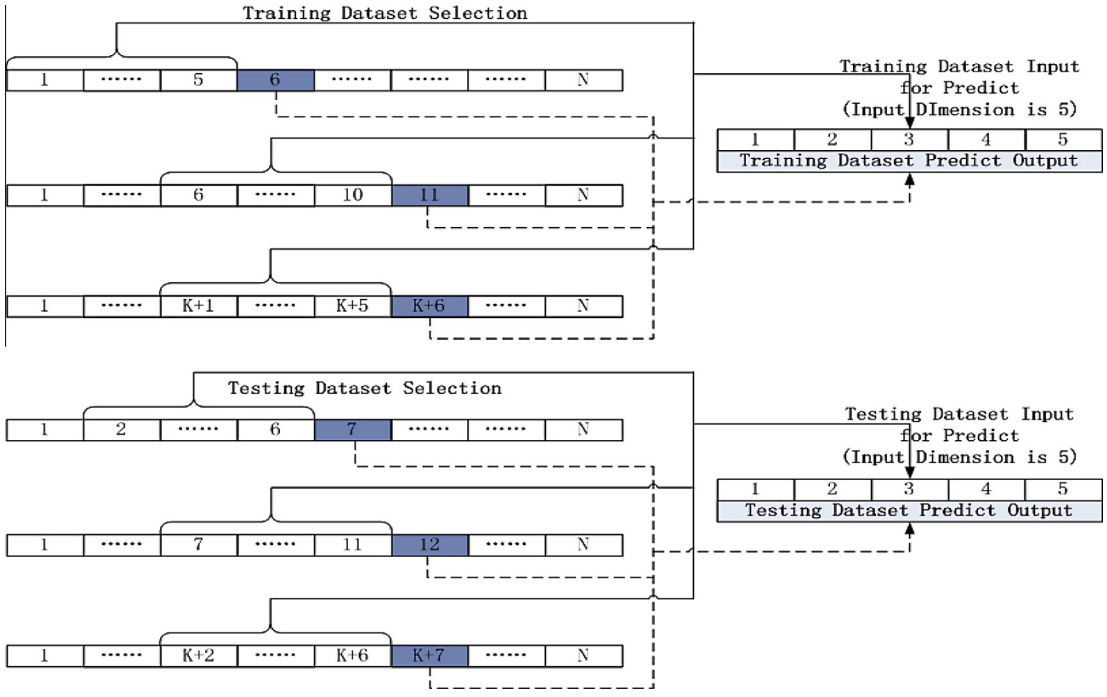

Fig. 6. Illustrate the lateral data selection procedure.

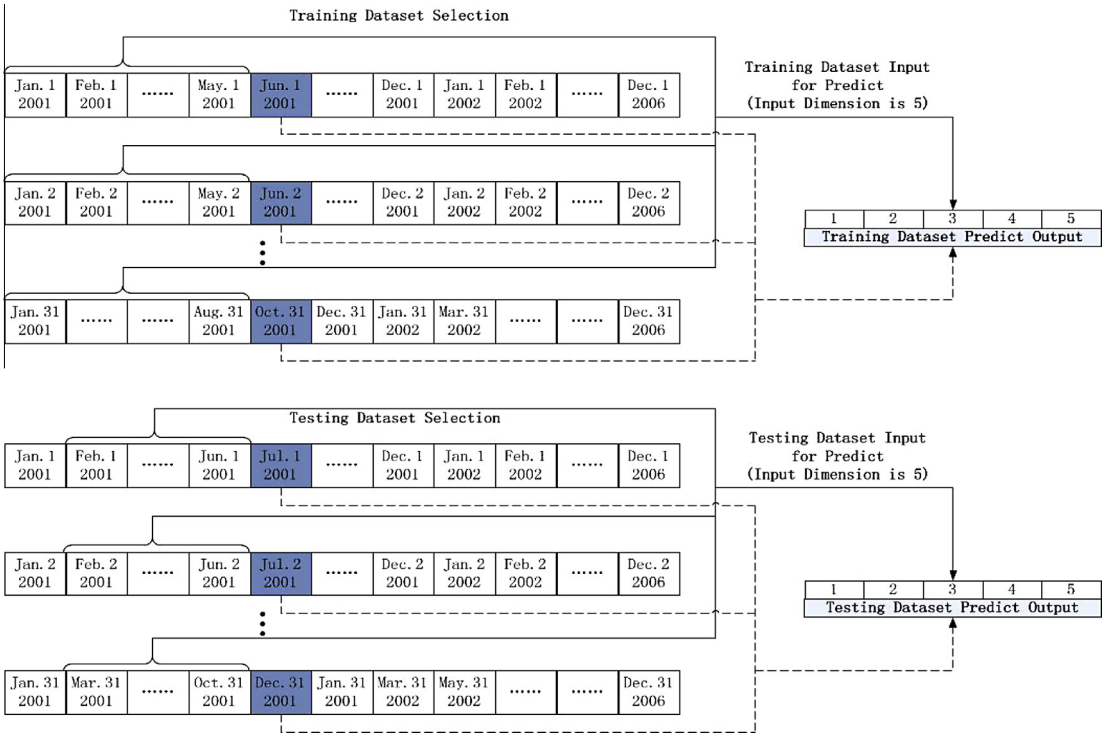

Fig. 7. Illustrate the longitudinal data selection procedure for daily wind speed data.

longitudinal data selection method; otherwise, the DataType is setting to 2, the TrainNum will be from 50 to 70, using the lateral data selection method. In all cases, the Dim will be from 5 to 15.

- Step 3. Determine a group of parameters according to step two. Initialize the parameters of BP neural network: the weights (W_{ij}) and bias (b_i) on each node.
- Step 4. Use the function that calculates the forecasting error of basic BP neural network as the fitness function of CPSO. CPSO algorithm is used to train the W_{ij} and b_i.
- Step 5. Use the parameters calculated by PSO for the BP neural network. Then calculate the testing sample data, record the results.
- Step 6. Update the parameters which PSO-BP used according to the Step 2 until all the conditions calculated. If the loop termination condition does not meet, initialize the parameters and go to Step 3, the new result will be calculated. When the loop condition reached, the flow will go to Step 7.
- Step 7. Evaluate all results with IS-PSO-BP model by comparing performance metrics with basic BP neural network considering all the conditions mentioned in Step 2. Finally, the model will output the best dimension, iteration max number for each training number of the data.
- Step 8. Initialize the network with the parameters obtained by Step 7 and used for the new data forecasting.

6. Simulation results of the case study

6.1. Performance metrics of forecast accuracy

To evaluate the forecast accuracy, many performance measure methods have been used, but no one can be recognized as the universal standard method. Therefore, we need to cover some performance metrics to comprehensively understand the algorithm characteristics. Four metrics are used in this study: mean error (AE), mean absolute error (MAE), mean square error (MSE), and mean absolute percentage error (MAPE).

$$AE = \frac{1}{N}\sum_{n=1}^{N}(y_n - \hat{y}_n) \tag{11}$$

$$MAE = \frac{1}{N}\sum_{n=1}^{N}|y_n - \hat{y}_n| \tag{12}$$

$$MSE = \frac{1}{N}\sum_{n=1}^{N}(y_n - \hat{y}_n)^2 \tag{13}$$

$$MAPE = \frac{1}{N}\sum_{n=1}^{N}\left|\frac{y_n - \hat{y}_n}{y_n}\right| \times 100\% \tag{14}$$

where y_n and $\hat{y}_n$ represent observed value and predictive value of nth data for performance evaluation, N is the total number of data used for performance evaluation and comparison.

AE is the average forecast error of n times forecast results. MAE is average absolute forecast error of n times forecast results. Because the prediction error may be positive and negative, the AE cannot reflect the level of error, this problem can be avoided by MAE. MSE is the average of the prediction error squares, it can evaluate the change of the predict model; the smaller of MSE value, the better of prediction model. MAPE is a measure of the accuracy of predict method for performance evaluation and comparison in statistics.

6.2. Comparisons of BPNN, PSO-BP using lateral data selection model

To select more effective forecasting method, we firstly use BP and PSO-BP neural network algorithm to forecast the wind speed of Jiuquan area discussed before. For BP neural network, basic gradient descent algorithm is used for network training, learning velocity is set to 0.1, the maximum number of training is 10,000,

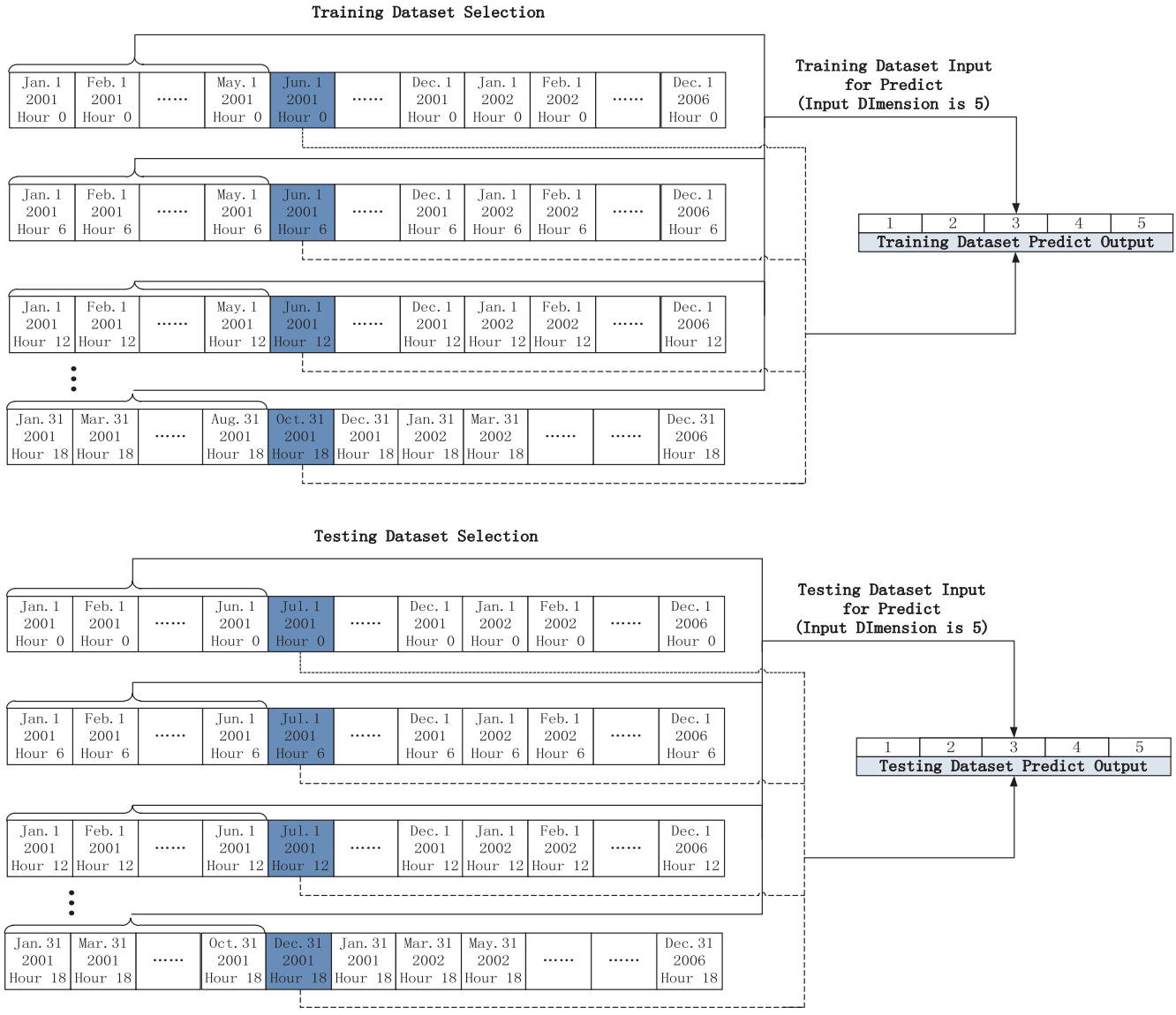

Fig. 8. Illustrate the longitudinal data selection procedure for hourly wind speed data.

and the training requirements precision is 0.01. The experimental results of the BP neural network algorithm and PSO-BP neural network algorithm using lateral data selection model will be presented in this section. It should be pointed out that each experiment was independently carried out 50 times and the typical results (AE, MAE, MSE and MAPE) were recorded in the Table 1. It can be seen that the two models using random numbers of previous observations and training samples, the BP model performs the better in terms of all the two statistics on average. More importantly, Table 1 indicates that different parameters of two models indeed lead to different performances. When the parameters are suitable, the PSO-BP model can perform better than the BP model. So we use IS-PSO-BP model as improved method to forecast the wind speed.

6.3. Comparisons of BPNN, IS-PSO-BP using lateral data selection model

To validate the improvement, we forecast the wind speed based on two different methods, and compare performance metrics in different situations. Using the IS-PSO-BP method, we find that setting the input dimension to 5 gives a better result for Jiuquan data. Consequently, we set the input dimension to 5 for all the experiments on Jiuquan wind speed data. We first select the training and testing dataset using lateral data selection method, with the training data size from 50 to 70, step-size is 5, start position i is set to 0.

Forecast accuracy results of BPNN model and IS-PSO-BP model with different training numbers are shown in Table 2. The best max iteration number for each training data was calculated by IS-PSO-BP model. For example, when the training data size is 50, the max iteration is set to 90; max iteration is 80 for training data size 55; and so on. From the result, we can see the IS-PSO-BP model can achieve better performance than the BPNN Model. It can also be observed that the bigger training sample size does not guarantee better performance. Meanwhile, the lateral data selection model may be suitable for some situations, but the longitudinal data selection model can be also used as data selection model. The following chapter will give the compared results between BPNN model and PSO-BP model.

6.4. Comparisons of BPNN, PSO-BP using longitudinal data selection model

After testing the lateral data selection method, start position i is set to 0, we next use longitudinal data selection method. The

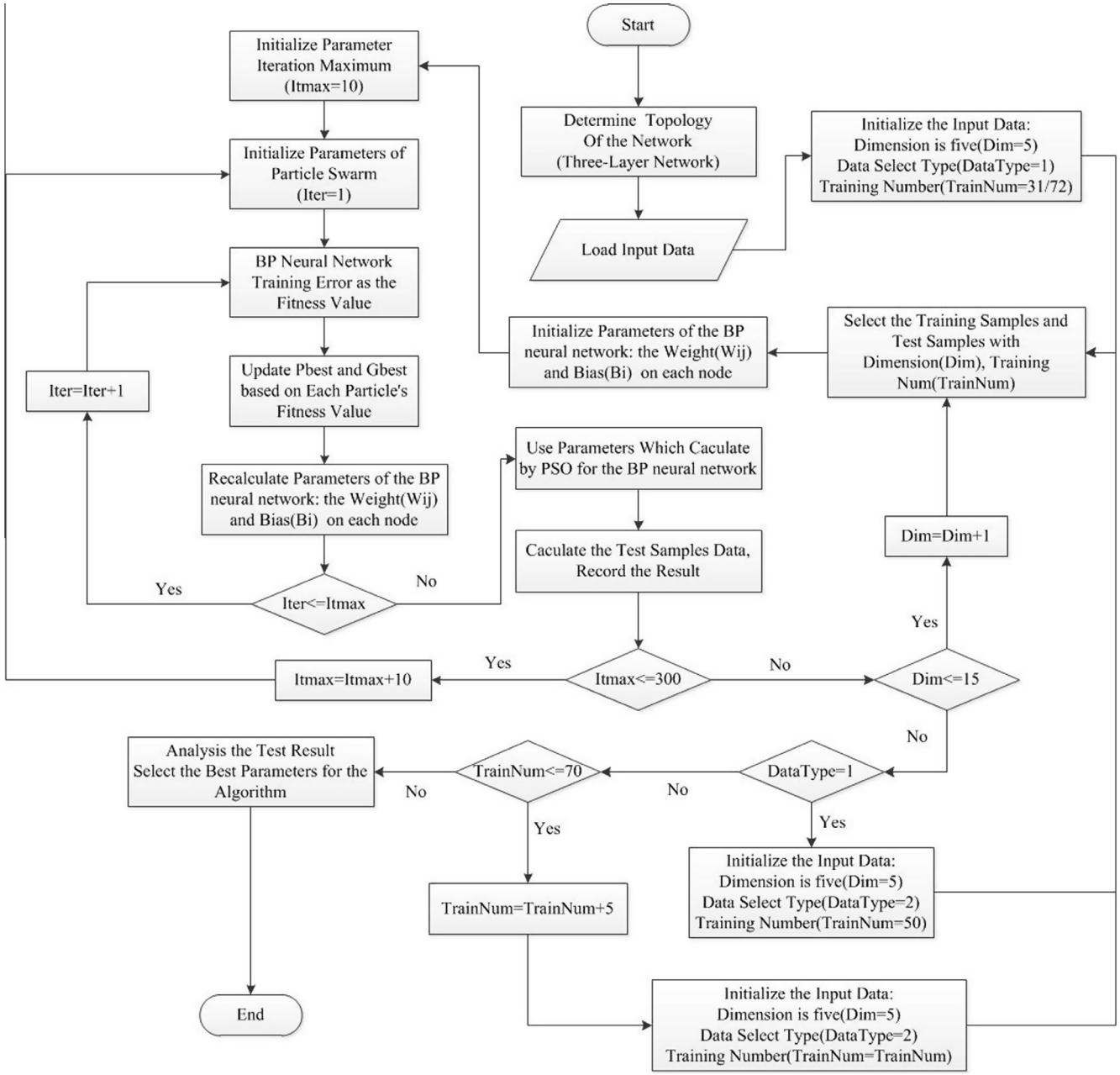

Fig. 9. Flow chart of IS-PSO-BP model.

Table 1
Comparisons of the PSO-BP neural network and BP neural network models about forecast accuracy using lateral data selection model.

Input dimension	Train number	Max. iteration	PSO-BP neural network				BP neural network			
			AE	MAE	MSE	MAPE (%)	AE	MAE	MSE	MAPE (%)
5	50	40	0.72	0.92	1.1	27.51	1.12	1.3	2.1	34.63
5	50	290	0.81	1.01	1.52	28.67	1.16	1.25	2	33.37
7	50	50	0.18	0.95	1.64	45.92	0.84	1	1.34	31.62
9	50	190	1.71	1.82	4.14	42.36	1.63	1.71	3.77	41.58
9	50	200	2.21	2.31	8.2	44.77	1.55	1.78	3.96	43.39
14	60	120	−0.52	1.08	1.97	87.55	0.61	0.95	1.26	32.32
14	60	130	0.38	1.16	1.96	48.36	0.75	1.07	1.64	33.88
14	65	110	0.05	0.92	1.37	41.92	0.58	0.92	1.17	31.79
14	65	120	0.30	0.94	1.41	37.84	0.58	0.92	1.17	31.81
6	70	180	−0.44	0.70	0.80	41.28	0.27	0.75	0.89	29.04
6	70	290	−0.27	0.88	1.15	50.35	0.20	0.67	0.67	27.69

Table 2
Comparisons of the two models about forecast accuracy with different training number using Jiuquan daily wind speed data.

Train number	Max iteration	IS-PSO-BP neural network				BP neural network			
		AE	MAE	MSE	MAPE (%)	AE	MAE	MSE	MAPE (%)
50	90	0.72	0.96	1.3	27.83	1.17	1.26	2.03	33.36
55	80	0.88	1.05	1.58	28.84	1.1	1.28	2.07	34.19
60	80	0.86	1.05	1.71	28.42	1.25	1.32	2.23	34.06
65	120	0.7	0.9	1.15	26.36	1.23	1.3	2.17	33.78
70	160	0.61	0.87	1.01	26.88	1.21	1.3	2.15	34.01

Table 3
Comparisons of the PSO-BP neural network and BP neural network models about forecast accuracy using longitudinal data selection model.

Input dimension	Train number	Max iteration	PSO-BP neural network				BP neural network			
			AE	MAE	MSE	MAPE (%)	AE	MAE	MSE	MAPE (%)
5	31	110	−0.04	0.38	0.21	18.76	0.11	0.42	0.25	19.20
5	31	180	−0.01	0.44	0.26	21.93	0.03	0.37	0.19	17.60
6	31	200	−0.32	0.76	1.05	37.59	0.14	0.65	0.58	23.04
6	31	210	−0.20	0.78	1.03	36.17	0.16	0.62	0.54	22.18
9	31	90	0.30	0.87	1.05	33.47	0.68	1.01	1.53	30.70
9	31	110	0.50	1.00	1.47	37.00	0.53	0.86	1.02	28.55
10	31	220	−0.17	0.83	1.10	39.21	0.46	0.85	1.10	30.36
10	31	240	−0.24	0.86	1.25	44.35	0.57	0.97	1.68	31.53
11	31	250	0.22	0.72	0.72	57.78	0.67	0.77	0.89	30.32
11	31	280	0.27	0.79	0.89	44.56	0.51	0.77	0.84	32.24
13	31	250	0.61	1.61	3.78	47.69	1.46	2.08	6.34	45.16
13	31	260	1.83	2.65	10.14	60.15	1.38	2.00	5.91	44.60

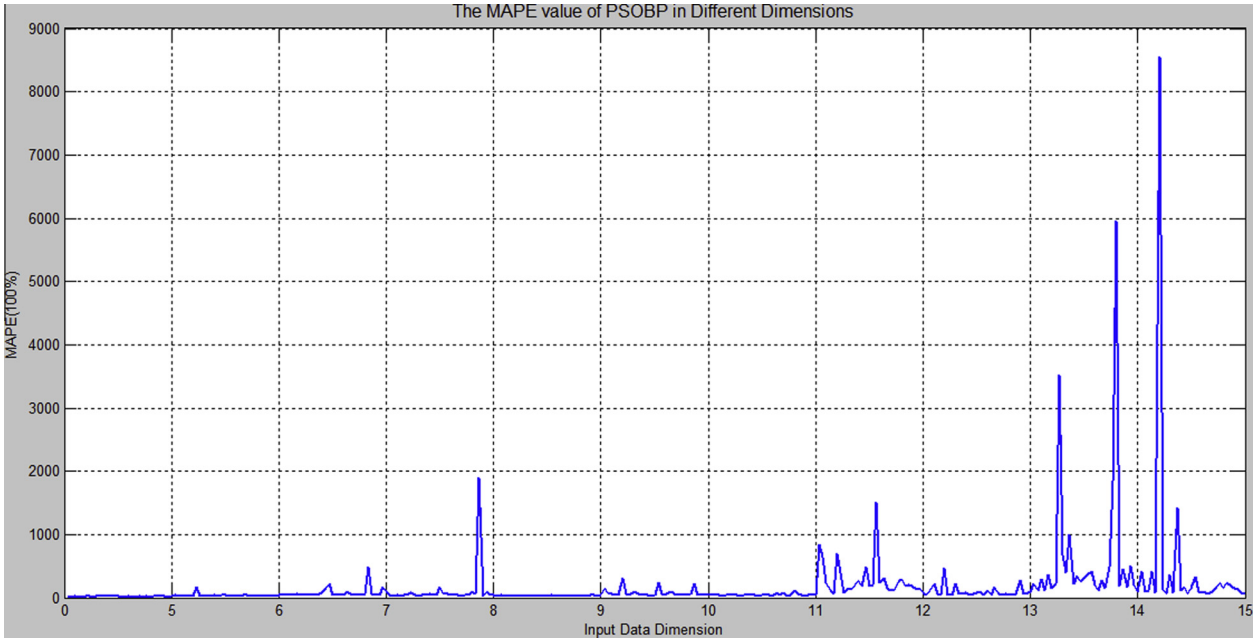

Fig. 10. The MAPE value of PSO-BP model in different dimensions.

experimental results of the BP model and PSO-BP model using longitudinal data selection model will be shown in Table 3. It should be pointed out that each experiment was randomly carried out 31 times. It can be seen that the two models using random input dimension of previous observations and random max iteration for PSO-BP model, the BP model performs better in terms of all the two statistics on average. The results are similar with the lateral one. In the next section, we will show the forecasting results calculated by IS-PSO-BP model and BP neural network model.

6.5. Comparisons of BPNN, IS-PSO-BP,ARIMA using longitudinal data selection model

Figs. 10–12 illustrate results of experiments using longitudinal data selection experiment, and these figures have the same display format: *X*-axis represents the dimension from 5 to 15 with step-size 1 and *Y*-axis represents the MAPE value in terms of percentage. On *X*-axis, the range [0,5] corresponds to dimension 5, the range [5,6] corresponds to dimension 6, and so on. Within each range, i.e. each dimension, the number on *X*-axis corresponds to max

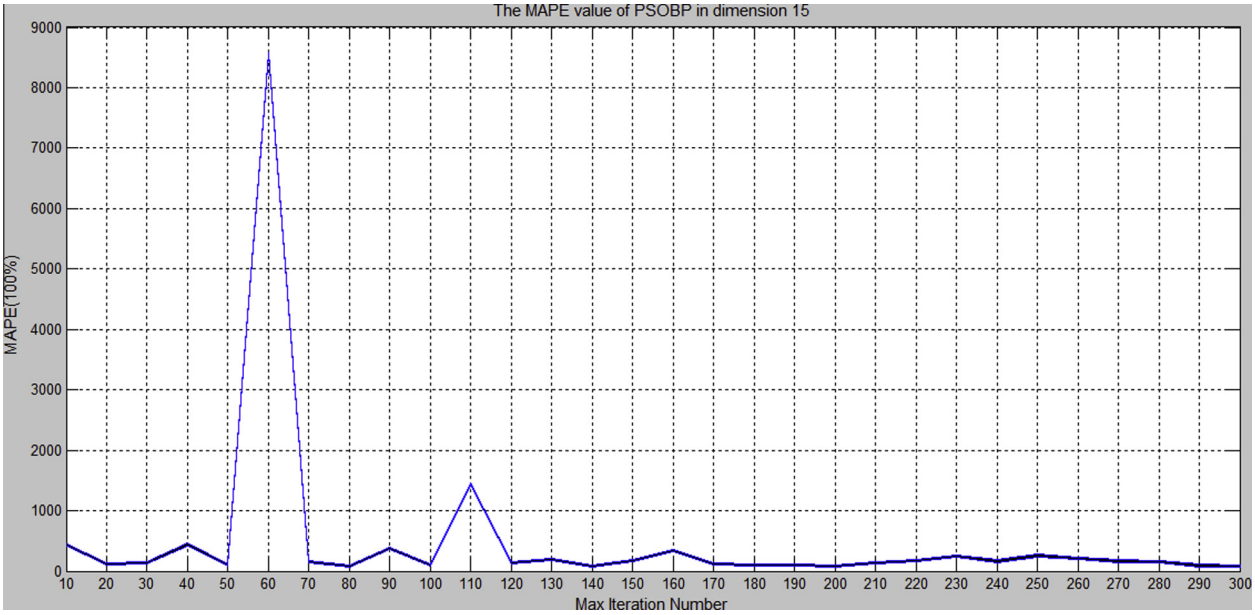

Fig. 11. The MAPE value of PSO-BP model in dimension 15.

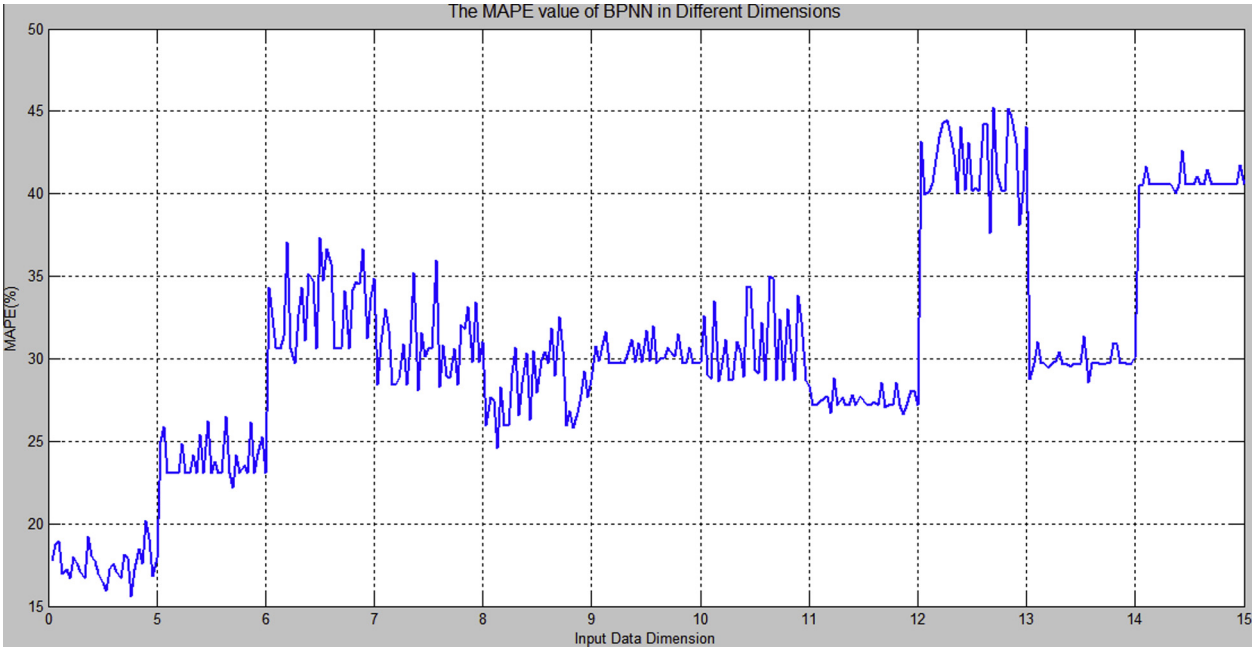

Fig. 12. The MAPE value of BPNN model in different dimensions.

iteration which ranges from 10 to 300 with step-size 10. On *Y*-axis, each MAPE value is an absolute average of 31 daily forecasting data using Jiuquan daily wind speed dataset. While Figs. 10 and 12 display the MAPE plot at different dimensions with different max iteration numbers using the IS-PSO-BP model and BPNN model respectively, Fig. 11 details the MAPE values at different iteration when thee dimension is 15 for the IS-PSO-BP model. As shown in Figs. 10 and 12, the input dimension has a significant impact on the prediction accuracy, which has the best value when the input dimension is 5. The ARIMA model in our paper is used as a comparison model. ARIMA is generally referred to as an ARIMA(p,d,q) model where parameters p, d, and q are non-negative integers that refer to the order of the autoregressive, integrated, and moving average parts of the model respectively. On the other hand, for identifying the orders of ARIMA(p,d,q) model, autocorrelation and partial auto correlation graphs which are drawn based on $1 \sim M$ lag numbers provide information about the AR and MA orders (p and q) [40]. Concretely, p is determined as the number of the front a few significant coefficients in the partial autocorrelation graph and similarly q is determined as the number of the front a few significant coefficients in the autocorrelation graph. According to the figure of ACF and PACF, we set the parameters of ARIMA model for two dataset respectively, ARIMA$(2,1,3)$ for Jiuquan daily wind speed data and ARIMA$(3,1,4)$ for Yumen hourly wind speed data. For Jiuquan daily wind speed dataset, Table 4 further illustrates the comparative results using BP neural model, ARIMA$(2,1,3)$ model and IS-PSO-BP model with input dimension is 5. The best max iteration is 10 calculated by IS-PSO-BP model.

Similarly, for Yumen hourly wind speed dataset, Table 5 illustrates the comparative results using BP neural model, ARIMA(3,1,4) model and IS-PSO-BP model with input dimension is 5. The best max iteration here is 20 calculated by IS-PSO-BP model.

6.6. Discussion for the experiment results

We have done 1980 times of experiments using IS-PSO-BP model to determine the input parameters of prediction methods under two data selection models for Jiuquan daily wind speed dataset. It can also be observed in Fig. 10 that not the training is sufficient, the predictive result is better. For the PSO-BP model, the max iteration number is bigger, the fitness value is smaller. But the MAPE value of forecast result does not decrease with fitness value, an appropriate max iteration number can avoid over-fitting and lead to a better result. Through the IS-PSO-BP can determine the best max iteration number for the PSO-BP model.

Meanwhile, Figs. 10 and 12 demonstrate that when predicting daily wind speed data from the history data, more input dimensions do not necessarily translate into better results. Here, the input dimension N denotes that wind speed data of immediate preceding N days are used to predicate the wind speed of the current day. As shown in Figs. 10 and 12, the MAPE values are generally larger when the input dimension is 15 than otherwise. In particular, the worst MAPE value in Fig. 10 is close to 9000% when the input dimension is 15 and its max iteration number is 60 as shown in Fig. 11. There are at least three reasons for this: (1) while a larger input dimension leads to a more complex neural network, across various dimensions our experiments keep the training data size at 31 and this inflexibility might not be sufficient for complex neural networks; (2) when the max iteration number of the PSO method is not chosen appropriately, the iteration could end prematurely causing the neural network ill-suited for making the prediction; and (3) different neural networks might need different kinds of sample data for accurate prediction. Clearly, selecting appropriate parameters is crucial for PSO-BP model. In our particular case of Jingquan wind speed dataset, the mean value of MAPE values at different max iteration numbers is smallest when the input dimension is 5 compared to other dimension numbers as shown in Figs. 10 and 12. Hence, for Jiuquan wind speed data set, the most desired input dimension number is 5.

Tables 2 and 4 have experimental results of two types of forecasting methods under two data selection models. When using lateral data selection method, IS-PSO-BP model has smaller minimum and maximum MAPE value (26.36,28.84) than BPNN model (33.36,34.19). Similarly, when using longitudinal data selection method, the MAPE value (15.51) of IS-PSO-BP model is smaller than MAPE value (21.10) of BPNN and ARIMA(2,1,3) model(19.55). It is easy to see that the longitudinal data selection method archive better forecast accuracy than the lateral one. Although both models are designed for predicting daily wind speed value, not for a curve prediction, the longitudinal method not the lateral one can be changed to forecast a month curve as shown in Fig. 13 using Jiuquan daily wind speed dataset, and the actual data were shown in Table 6. The 18-Day wind curve forecasting of Yumen 6-hourly wind speed data using longitudinal data selection method is shown in Fig. 14. Above all, the longitudinal data selection method is more suitable than the lateral method in wind speed prediction.

Table 4
Comparisons of the three models about forecast accuracy with training number is 31 using Jiuquan daily wind speed data.

Model	Training MSE	AE	MAE	MSE	MAPE (%)
IS-PSO-BP	0.46	−0.08	0.16	0.17	15.51
BP	–	0.13	0.41	0.22	21.10
ARIMA(2,1,3)	–	−0.10	0.37	0.18	19.55

Table 5
Comparisons of the three models about forecast accuracy with training number is 124 using Yumen hourly wind speed data.

Model	Training MSE	AE	MAE	MSE	MAPE (%)
IS-PSO-BP	1.83	1.27	1.55	3.67	21.02
BP	–	2.56	2.63	9.42	35.19
ARIMA(3,1,4)	–	−2.69	2.76	9.95	37.08

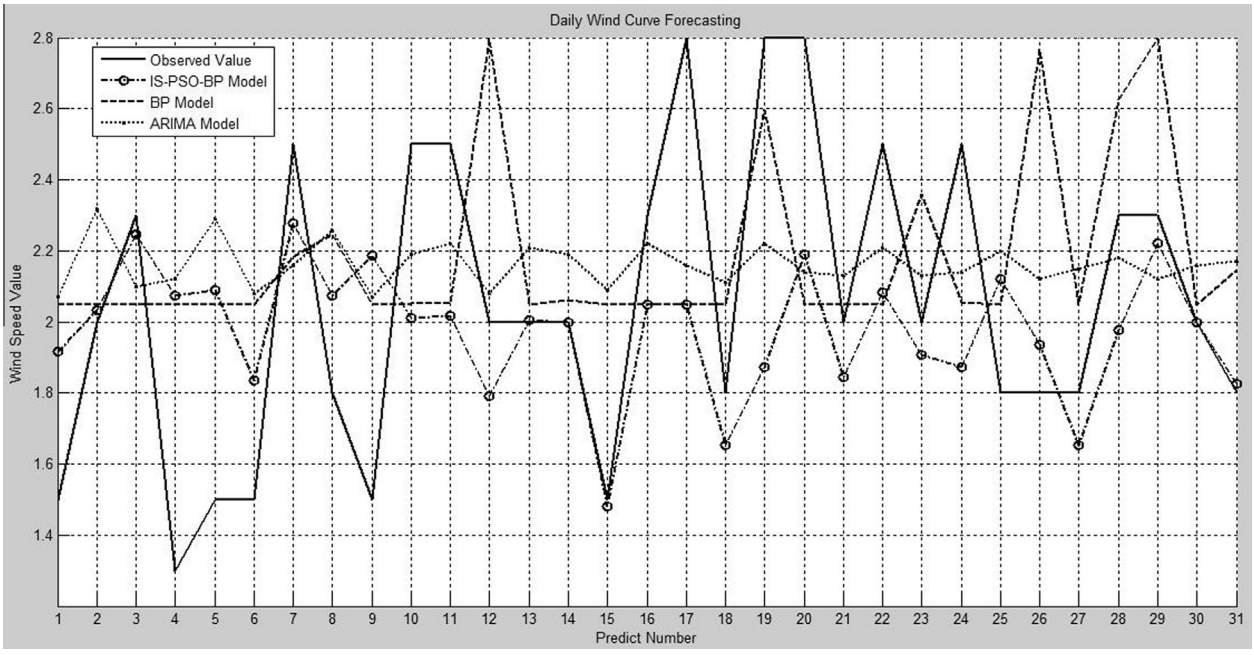

Fig. 13. Monthly wind curve forecasting of Jiuquan daily wind speed data using longitudinal data selection method.

Table 6
Monthly wind speed forecasting of three models using Jiuquan daily wind speed data.

No.	Observed value	IS-PSO-BP value	ARIMA value	BPNN value	No.	Observed value	IS-PSO-BP value	ARIMA value	BPNN value
1	1.5	1.9168	2.07	2.05	17	2.8	2.0474	2.16	2.0502
2	2	2.0339	2.32	2.05	18	1.8	1.6536	2.11	2.05
3	2.3	2.2473	2.1	2.05	19	2.8	1.8738	2.22	2.5962
4	1.3	2.0740	2.12	2.05	20	2.8	2.1905	2.14	2.05
5	1.5	2.0881	2.29	2.05	21	2	1.8433	2.13	2.05
6	1.5	1.8353	2.08	2.05	22	2.5	2.0845	2.21	2.05
7	2.5	2.2784	2.16	2.1838	23	2	1.9068	2.13	2.3597
8	1.8	2.0749	2.26	2.2468	24	2.5	1.8726	2.14	2.0551
9	1.5	2.1863	2.07	2.0501	25	1.8	2.122	2.2	2.0501
10	2.5	2.0125	2.19	2.0533	26	1.8	1.9364	2.12	2.7659
11	2.5	2.0158	2.22	2.0506	27	1.8	1.6538	2.15	2.05
12	2	1.7916	2.08	2.7963	28	2.3	1.9765	2.18	2.6247
13	2	2.0040	2.21	2.05	29	2.3	2.2223	2.12	2.8
14	2	1.9992	2.19	2.0622	30	2	1.9998	2.16	2.05
15	1.5	1.4818	2.09	2.05	31	1.8	1.8261	2.17	2.1457
16	2.3	2.0496	2.22	2.05					

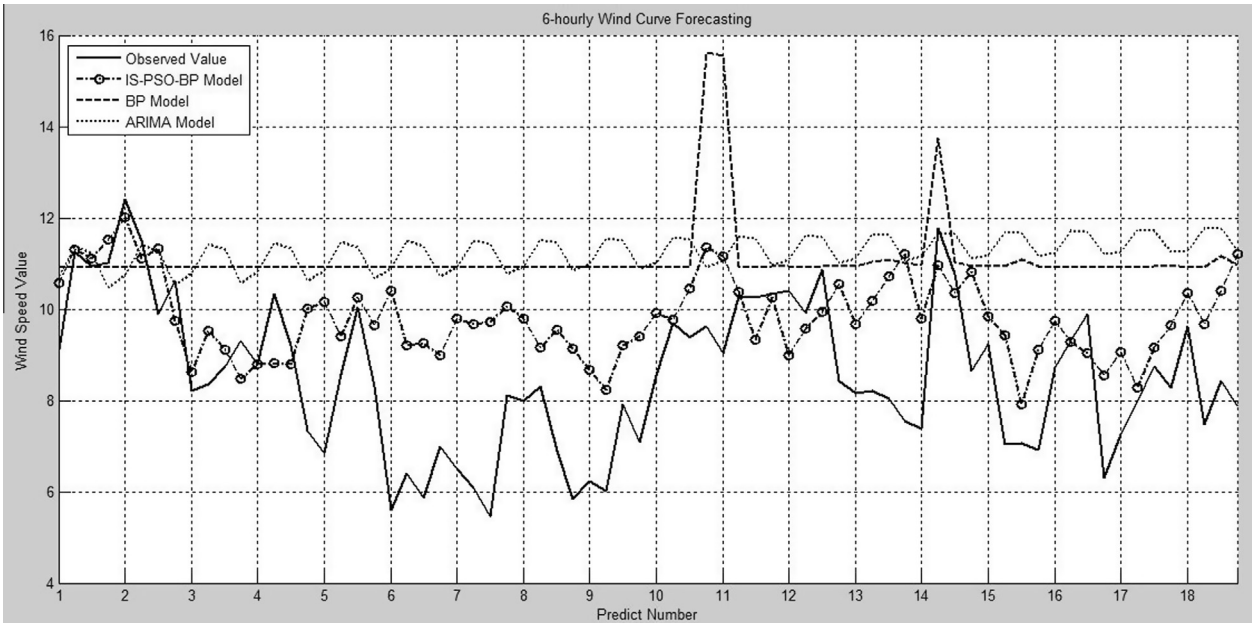

Fig. 14. 18-Day wind curve forecasting of Yumen 6-hourly wind speed data using longitudinal data selection method.

By comparing three types of forecasting models, we can see the IS-PSO-BP model is robust and can obtain better performance, as shown by the smaller MAPE value and MSE value. For example, in Table 4 the MAPE of IS-PSO-BP is 15.51%, BP model is 21.1% and ARIMA(2,1,3) model is 19.55%; the MSE of IS-PSO-BP model is 0.17, BP model is 0.22 and ARIMA(2,1,3) model is 0.18. In Table 5 the MAPE of IS-PSO-BP is 21.02%, BP model is 35.19% and ARIMA(3,1,4) model is 37.08%; the MSE of IS-PSO-BP model is 3.67, BP model is 9.42 and ARIMA(3,1,4) model is 9.95. As shown in Figs. 13 and 14, using ARIMA model to forecast longer horizon, the forecasting results can be seen as a straight line or a slight fluctuation of frequency curve. The wind speed data is fluctuant, and the ARIMA forecasting results are equivalent to the average wind speed data. This could explain why the forecasting performance of ARIMA model was not so bad, but the forecasting results cannot show the wave characteristics of wind speed.

It is noteworthy that, to verify the performance of our proposed model, we compare it with BPNN model and ARIMA model on two types of wind speed data (hourly and daily). But for predicting the wind speed at another place, the input parameters of IS-PSO-BP need to be calculated again as there are not uniform parameters for different datasets. As to the execution clock time of our proposed method, the forecast took about 5.5 h for both Jingquan and Yumen datasets.

7. Conclusions

Globally, wind energy has become one of the fastest growing clean and renewable energy sources. To integrate wind energy into the power system, it is important to predicate the wind power generation. Due to the continuous fluctuation of wind resources, it remains a large challenge to accurate and reliable forecast short-term wind speed. In this paper, we first use PSO-BP model and BP model to forecast the wind speed data with randomly generated parameters of PSO-BP model. We propose two input dataset selection methods to use historical data for training and testing: the lateral data selection and longitudinal data selection methods. Comparing the forecast accuracy of these models shows that different parameters directly influence the forecast accuracy. This leads

us to combine PSO-BP model with parameter selection method called IS-PSO-BP model to achieve better performance. This parameter selection method includes input dataset selection method, input data dimension determined method and PSO parameters confirmed method, and it outputs optimized parameters for the next forecast. Finally, we use the Jiuquan daily wind speed dataset and Yumen hourly wind speed dataset as two real world cases to compare forecast performance of BP model, ARIMA model and IS-PSO-BP model, and the latter method shows significant improvement. From the experiment, we found the longitudinal data selection method is more suitable than the lateral method in wind speed prediction, and the longitudinal one is more useful in fact.

Acknowledgment

This work was supported in part by National Science & Technology Pillar Program of China under Grant 2013BAH19F01, by National Natural Science Foundation of China under Grant Nos. 61073193, 70673030, 90924025, and by Hefei University of Technology under Grant Nos. 407-037036 and 2012HGZY0031.

References

[1] World Wind Energy Association, World Wind Energy Report, 2012 <http://www.wwindea.org/home/images/stories/worldwindenergyreport2010_s.pdf> (accessed 09.08.2012).
[2] Z.-H. Guo, J. Wu, H.-Y. Lu, J.-Z. Wang, A case study on a hybrid wind speed forecasting method using BP neural network, Knowledge-Based Systems 24 (2011) 1048–1056.
[3] M. Alexiadis, P. Dokopoulos, H. Sahsamanoglou, Wind speed and power forecasting based on spatial correlation models, IEEE Transactions on Energy Conversion 14 (1999) 836–842.
[4] J.F. Kennedy, J. Kennedy, R.C. Eberhart, Swarm Intelligence, Morgan Kaufmann Pub., 2001.
[5] K. Cai, L.-N. Tan, C.-L. Li, X.-F. Tao, Short-term wind speed forecasting combing time series and neural network method, Power System Technology 8 (2008) 024.
[6] I.G. Damousis, M.C. Alexiadis, J.B. Theocharis, P.S. Dokopoulos, A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation, IEEE Transactions on Energy Conversion 19 (2004) 352–361.
[7] R.G. Kavasseri, K. Seetharaman, Day-ahead wind speed forecasting using f-ARIMA models, Renewable Energy 34 (2009) 1388–1393.
[8] E. Cadenas, W. Rivera, Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA–ANN model, Renewable Energy 35 (2010) 2732–2738.
[9] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, Z. Yan, A review on the forecasting of wind speed and generated power, Renewable and Sustainable Energy Reviews 13 (2009) 915–920.
[10] G. Box, Box and Jenkins: time series analysis, forecasting and control. A Very British Affair autofilled 161 (1994).
[11] J.L. Torres, A. Garcia, M. De Blas, A. De Francisco, Forecast of hourly average wind speed with ARMA models in Navarre (Spain), Solar Energy 79 (2005) 65–77.
[12] G. Li, J. Shi, J. Zhou, Bayesian adaptive combination of short-term wind speed forecasts from neural network models, Renewable Energy 36 (2011) 352–359.
[13] T. Barbounis, J. Theocharis, Locally recurrent neural networks for long-term wind speed and power prediction, Neurocomputing 69 (2006) 466–496.
[14] T. Barbounis, J. Theocharis, Locally recurrent neural networks for wind speed prediction using spatial correlation, Information Sciences 177 (2007) 5775–5797.
[15] Z. Guo, W. Zhao, H. Lu, J. Wang, Multi-step forecasting for wind speed using a modified EMD-based artificial neural network model, Renewable Energy 37 (2012) 241–249.
[16] Y.-Y. Hong, H.-L. Chang, C.-S. Chiu, Hour-ahead wind power and speed forecasting using simultaneous perturbation stochastic approximation (SPSA) algorithm and neural network with fuzzy inputs, Energy 35 (2010) 3870–3876.
[17] J. Zhou, J. Shi, G. Li, Fine tuning support vector machines for short-term wind speed forecasting, Energy Conversion and Management 52 (2011) 1990–1998.
[18] S. Salcedo-Sanz, Á.M. Pérez-Bellido, A. Portilla-Figueras, L. Prieto, Short term wind speed prediction based on evolutionary support vector regression algorithms, Expert Systems with Applications 38 (2011) 4052–4057.
[19] M. Mohandes, T. Halawani, S. Rehman, A.A. Hussain, Support vector machines for wind speed prediction, Renewable Energy 29 (2004) 939–947.
[20] G. Li, J. Shi, Application of Bayesian model averaging in modeling long-term wind speed distributions, Renewable Energy 35 (2010) 1192–1202.
[21] G. Li, J. Shi, Applications of Bayesian methods in wind energy conversion systems, Renewable Energy 43 (2012) 1–8.
[22] E. Hadavandi, H. Shavandi, A. Ghanbari, Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting, Knowledge-Based Systems 23 (2010) 800–808.
[23] A. Khosravi, S. Nahavandi, D. Creighton, Quantifying uncertainties of neural network-based electricity price forecasts, Applied Energy 112 (2013) 120–129.
[24] A. Khotanzad, H. Elragal, T.-L. Lu, Combination of artificial neural-network forecasters for prediction of natural gas consumption, IEEE Transactions on Neural Networks 11 (2000) 464–473.
[25] Z. Bashir, M. El-Hawary, Applying wavelets to short-term load forecasting using PSO-based neural networks, IEEE Transactions on Power Systems 24 (2009) 20–27.
[26] Z. Xiao, S.-J. Ye, B. Zhong, C.-X. Sun, BP neural network with rough set for short term load forecasting, Expert Systems with Applications 36 (2009) 273–279.
[27] D.E. Rumelhart, G.E. Hintont, R.J. Williams, Learning representations by back-propagating errors, Nature 323 (1986) 533–536.
[28] L. Zhang, K. Wu, Y. Zhong, P. Li, A new sub-pixel mapping algorithm based on a BP neural network with an observation model, Neurocomputing 71 (2008) 2046–2054.
[29] I.V. Tetko, D.J. Livingstone, A.I. Luik, Neural network studies. 1. Comparison of overfitting and overtraining, Journal of Chemical Information and Computer Sciences 35 (1995) 826–833.
[30] I. Moghram, S. Rahman, Analysis and evaluation of five short-term load forecasting techniques, IEEE Transactions on Power Systems 4 (1989) 1484–1491.
[31] G. Li, J. Shi, On comparing three artificial neural networks for wind speed forecasting, Applied Energy 87 (2010) 2313–2320.
[32] J.-R. Zhang, J. Zhang, T.-M. Lok, M.R. Lyu, A hybrid particle swarm optimization–back-propagation algorithm for feedforward neural network training, Applied Mathematics and Computation 185 (2007) 1026–1037.
[33] E. Assareh, M. Behrang, A. Ghanbarzadeh, The integration of artificial neural networks and particle swarm optimization to forecast world green energy consumption, Energy Sources, Part B: Economics, Planning, and Policy 7 (2012) 398–410.
[34] F. Sha, F. Zhu, S.N. Guo, J.T. Gao, Based on the EMD and PSO-BP neural network of short-term load forecasting, Advanced Materials Research 614 (2013) 1872–1875.
[35] Introduction of Jiuquan City on Wikipedia, 2013 <http://zh.wikipedia.org/wiki/%E9%85%92%E6%B3%89%E5%B8%82> (accessed 12.05.2013).
[36] The Central People's Government of the People's Republic of China, 2013 <http://www.gov.cn/jrzg/2009-08/08/content_1386664.htm> (accessed 12.05.2013).
[37] R. Hecht-Nielsen, Kolmogorov's mapping neural network existence theorem, in: Proceedings of the International Conference on Neural Networks, Publishing, 1987.
[38] E. Ozcan, C.K. Mohan, Particle swarm optimization: surfing the waves, in: Proceedings of the 1999 Congress on Evolutionary Computation, 1999, CEC 99, Publishing, 1999.
[39] Y. Shi, R.C. Eberhart, Empirical study of particle swarm optimization, in: Proceedings of the 1999 Congress on Evolutionary Computation, 1999, CEC 99, Publishing, 1999.
[40] V.Ş. Ediger, S. Akar, ARIMA forecasting of primary energy demand by fuel in Turkey, Energy Policy 35 (2007) 1701–1708.