# A Study of Wind Speed Prediction Based on Particle Swarm Algorithm to Optimize the Parameters of Sparse Least Squares Support Vector

Haijun Ren[1,2,*], Xin Lei[1,2], Ping Zhang[1,2]

[1] *School of Advanced Manufacture and Engineering*,
[2]Engineering Center of Automobile Electric and Embedded System,
Chongqing University of Posts and Communications
Chongqing,, China

***Abstract* — Wind speed forecasting can accurately improve prediction efficiency of wind power in wind farm, decrease failure probability of wind turbine, and extend life cycle. An innovative algorithm is proposed to optimize both the parameters of least squares support vector machine (LSSVM) and the procedure of finding sparse support vector. Firstly, the defects of support vector are analyzed. Then inequality constraints are replaced by equality constrains. Quadratic programming problem is transformed into linear equations through Lagrange method to solve goal function. Solving process for least squares support vector is deduced and the reasons why LSSVM does not possess the sparse property are analyzed. Parameters of the LSSVM model and the procedure of finding sparse support vector are optimized by particle swarm optimization (PSO) algorithm. Then based on actual wind speed, prediction performances of three kinds of forecasting methods, including sparse LSSVM optimized by particle swarm algorithm (SPSO-LSSVM), auto regressive moving average (ARMA) and artificial neural networks (ANN), are compared. The results show that the forecasting performance of SPSO-LSSVM is the best. The effectiveness of proposed algorithm is verified by simulation.**

***Keywords - wind speed prediction; sparse least squares support vector; particle swarm optimization; cross validation***

## I. INTRODUCTION

As the global environment is polluted, many sustainable energies and green powers are developing rapidly, including wind energy [1]. However, stochastic wind makes operational control of wind turbine very difficult and affects management and adjustment of power distribution system [2]. In order to solve the above problems, it is important that the wind speed should be forecasted accurately. Wind speed forecasting includes the short-term prediction at sampling period 10-second or 10-minute, the medium-term forecasting at time intervals hours, and the long-term prediction involving days [3]. The different timeframes have been selected based on the concrete fields.

The wind speed can be predicted by several kinds of means, mainly including statistic method, artificial intelligent method and hybrid prediction approach. Time series analysis is a main methods based on statistical theory, which is an approach to process dynamic data. Its mathematical model can be established by analyzing the historical relationships between input data and output data. Therefore, according to the constructed model, the output data of next step could be predicted based on the historical input data. There are several kinds of time series prediction methods, including auto regressive (AR), moving average (MA), ARMA, and their variants [4-7]. Kusiak [4] adopted time series approach to examine short-term wind behavior and compared exponential smoothing and data-driven models for wind power prediction. In this study, a double exponential smoothing model was proposed to forecast the wind speed. In addition, an evolutionary strategy algorithm was used to acquire the best smoothing constants. In the data-driven model, the parameters for forecasting wind speed are determined by boosting-tree algorithm. It is proved that the prediction performance of the former model is better than that of the later model. Cadenas [5] proposed an autoregressive integrated moving average (ARIMA) method for long-term wind speed prediction, which has strong robustness. However, there is a sensitivity feature in the seasonal ARIMA model when the curve is changed during the year. Furthermore, ARIMA is difficult to adapt to atypical values. Then it can not forecast the future values exactly. At the same time, the order of the stochastic model is very high.

The other important forecasting method is artificial intelligence algorithm like ANN, fuzzy inference algorithm, etc [8-10]. Pourmousavi Kani [8] introduced an ANN algorithm to predict short-term wind speed. To avoid over-training problem of ANN, a Markov chain (MC) was proposed, which could capture long-term trends in the wind speed data and store historical signals. So the prediction errors could be reduced. In this article, two ANNs were used. The primary prediction of short-term trend and wind speed were finished by the first ANN, which adopted a multi-layer perceptron. Then the transition probability matrixes with primary prediction values were calculated through MC. Consequently, the acquired primary prediction values were delivered to the second artificial networks. Finally, the accuracy of final wind speed prediction was improved based on these steps mentioned above. Li [9] compared the medium-term time intervals prediction performance of wind speed one hour ahead among three kinds of ANN which included adaptive linear neural network, back-propagation neural network, and radial basis function neural network. The relations between present time series and future data were analyzed through the autocorrelation function and the partial autocorrelation function. At different plots to collect wind speed, the historical data affected the future data to different extent. In this study, to reduce training time, the weight parameters were decided by LM optimization algorithm. The influences of different transmitting parameters on radial basis function neural network were studied. The effects of different learning rate on BP networks and ADLINE networks were analyzed.

In addition, hybrid prediction method was employed to predict wind speed. Usually, the prediction performance of hybrid method is better than that of adopting only one forecasting method [11-12]. Monfared [11] proposed a novel approach combining fuzzy logic and neural networks to predict medium-term wind speed. The new strategy can increase forecasting accuracy when less rules is used in fuzzy logic and there are less neuron numbers and less learning time in artificial neural network. Liu [12] suggested a novel hybrid method that consisted of wavelet and time series analysis to forecast hourly wind speed. The prediction accuracy of adopted hybrid method is also better than that of classic time series analysis.

In prediction methods, support vector machine (SVM) possesses a special position. SVM is not only based on statistic theory but also has many features of machine learning. The problems of over-fitting and curse of dimensionality of neural networks could be declined by this machine algorithm. Particularly, if there are small amounts of sample data, the prediction ability of SVM is better than that of ANN. Therefore, this machine algorithm has been extensively studied and vigorously developed in the area of signal prediction. SVM can be transformed into LSSVM if factors of SVM are changed. Some parameter optimization algorithms related to them were studied [13-15]. Sun [13] proposed a novel short-term wind speed prediction algorithm. Firstly, original wind speed was decomposed into intrinsic mode functions with one residual series. Then LSSVM models were built based on each data set and the parameters of models were optimized by adopting BAT algorithm. Finally, each prediction data was added to achieve the ultimate predicted value. As prediction accuracy of LSSVM was affected by input data, the forecasting accuracy of model based on each subset was better than that of model based only on total data set. Mohandes [14] predicted long-term wind speed through SVM, which mapped the actual data into a high dimensional feature space. This mapping has a nonlinear feature and the data could be calculated with linear method. So the calculation of actual data was simplified. Actually, the mapping process is an optimization issue. The kernel function could be selected to describe the inner product in the D-dimensional feature space. Polynomial, Gaussian, and sigmoidal were usually adopted as kernel function, which transferred the data from the low dimensional feature space to the high dimensional feature space. The non-linear question was finally solved. There were wind speed data covering a period of twelve years, which were divided into training data, validation data and testing data. In terms of above datasets categorized, the author compared the forecasting performance of SVM with that of multilayer perceptron (MLP). The results indicated that the prediction errors of SVM were smaller than that of MLP.

As SVM has many merits, it is not only used in the field of wind speed prediction, but also utilized to predict stream flow, evaluate fault diagnosis of power transformers, build soft measurement model of chemical elements, etc [15-17]. The SVM is developed rapidly based on above applications. Nevertheless, the solution processing of SVM is very sophisticated. Then LSSVM is employed to solve some problems. The parameters of LSSVM are crucial to its performance. Therefore, many algorithms are used to optimize the parameters.

This paper herein proposes an innovative optimization method for predicting the wind speed at wind farm in North China. The normalizing factor and kernel function of LSSVM are optimized. Ten fold cross validation method is selected to determine fitness function during optimizing parameters. In case of finding sparse support vector, the sparse rate is considered as individual component in fitness function and its weight is increased. Parameters are optimized as the procedure of finding sparse support vectors is also optimized. Then a multidimensional optimization problem is constructed, which has been solved by PSO algorithm. After sampling of the realistic wind speed at wind farm in North China, this paper attempted to predict the wind speed by SPSO-LSSVM. Then the performance of forecasting method proposed was compared with those of ARMA and ANN. The suitability and effectiveness of the proposed models are thus confirmed.

.

## II. LSSVM

### *A. SVM*

SVM is a small-sample method based on statistics studying theory, which consists of comprehensive theoretical system and the foundation of mathematics. If the input data and output data are obtained, the SVM can forecast the future output through regression model. As SVM can mapping nonlinear data from low-dimensional feature space to high-dimensional feature space, the data can be analyzed and calculated by linear method in the new space. Therefore, the SVM, which is used to solve the problems of nonlinear prediction based on regression model, can be presented as the follows [18]: If there is data set and it is $n$ -dimension vectors, some of the vectors of $l$ are selected as a sample, it can be shown as $(x_1, y_1), (x_2, y_2)...(x_l, y_l) \in R^n \times R$. Where $x_l$ is input variable, $y_l$ is output variable, $l = 1, 2, 3...l$ .The samples are mapped from original space $R^n$ to high-dimension space, which is $(\phi(x_1), \phi(x_2), ..., \phi(x_l))$. Then, the best optimization decision function is as follows:

$$y(x) = w \cdot \phi(x) + b \tag{1}$$

Where, $w$ is weight coefficient, $\phi(x)$ is mapping function, $b$ is deviation value.
According to the conversion formula, the data would be outputted through the linear method in the high-dimension space. Taking the minimum principle of structure risk into consideration, the solutions of $w$ and $b$ could be gained by minimizing the function:

$$R = \frac{1}{2}\|w\|^2 + c \cdot R_{emp} \tag{2}$$

Where $\|w\|^2$ is employed to control the complexity. $c$ is normalizing factor. $R_{emp}$ is error control function. Different SVM could be constructed if different loss functions are selected. The risk function can be formed by solving the optimization goal. The goal function is as follows:

$$\min J = \frac{1}{2}\|w\|^2 + c\sum_{i=1}^{l}(\xi_i + \xi_i^*) \tag{3}$$

Subject to:

$$\begin{cases} y_i - w\phi(x_i) - b \le \varepsilon + \xi_i \\ w\phi(x_i) + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \le 0, \varepsilon \le 0, \qquad i = 1, 2, ...l \end{cases} \tag{4}$$

Where $\varepsilon$ is estimation accuracy, $\xi_i$ and $\xi_i^*$ are slack variables.

The goal of above process is to solve a convex optimization problem. So the solving process of SVM is very complicated.

*B. LSSVM*

If the sum of squared errors is regarded as loss function and the inequality restrictions are changed into equality form, then the SVM can be transformed as LSSVM. The optimization problem is expressed as follows [19]:

$$\min J = \frac{1}{2}\|w\|^2 + \frac{1}{2}\gamma\sum_{i=1}^{l} e_i^2 \tag{5}$$

Subject to:

$$y_i = w\phi(x_i) + b + e_i, \quad i = 1, 2, ...l \tag{6}$$

The above optimization problem can be solved by Lagrange theory as follows:

$$L(w, b, e, \alpha) = J(w, e) - \sum_{i=1}^{l} \alpha_i \left\{ w\phi(x_i) + b + e_i - y_i \right\} \tag{7}$$

Where, $\alpha_i$ is Lagrange multiplier, $i = 1, 2, ...l$ , $\gamma$ is regularization parameter, $e_i$ is error .
The differentiation can be obtained according to optimization requirement:

$$\begin{cases} \frac{\partial L}{\partial w} = 0, \ w = \sum_{i=1}^{l} \alpha_l \phi(x_l) \\ \frac{\partial L}{\partial b} = 0, \ \sum_{i=1}^{l} \alpha_i = 0 \\ \frac{\partial L}{\partial e_i} = 0, \ \alpha_i = \gamma e_i \\ \frac{\partial L}{\partial \alpha_i} = 0, \ w\phi(x_i) + b + e_i - y_i = 0 \end{cases} \tag{8}$$

After $e_i$ and $w$ are eliminated, the realization equation can be expressed as follows:

$$\begin{bmatrix} 0 & I^T \\ I & \delta + \gamma^{-1} I \end{bmatrix} \cdot \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix} \tag{9}$$

Where, $\delta_{ij} = K(x_i, x_j) = \phi(x_i)\phi(x_j)$, $i, j = 1, 2, 3...N$, $K(x_i, x_j)$ is kernel function for LSSVM, $I$ is unit matrix, $I^T$ is transport matrix of $I$. So it can be gained as follows:

$$f(x) = \sum_{i=1}^{l} \alpha_i K(x_i, x) + b \tag{10}$$

From the equations (8), it is gained that Lagrange multiplier $\alpha_i$ is proportionate to errors. Then almost every input sample will be included in the regression model. Therefore, LSSVM would lose its sparse-feature. LSSVM can be acquired as follows:

Step1: Normalize sample data and build input-output matrix;
Step2: Initialize least squares support vector;
Step3: Optimize parameters;
Step4: Employ the optimized parameters to train model;
Step5: Predict wind speed.

## III. SPARSE SUPPORT VECTOR AND PARAMETERS OPTIMIZATION FOR LSSVM BASED ON PSO

### *A. Sparse Processing*

Sparse-feature is just that a part of data in sample are considered as support vectors, which include large of feature of sample and they can replace samples in main character[20]. To take advantage of LSSVM and to make itself to possess sparse-feature, a random constant is given to each training sample from zero to one, which indicates the probability of considering the training sample as support vector [20]. The data over defined random constant indicates the corresponding sample as support vector and it will be put into training sample set. Otherwise, it will be put into testing sample set as testing data. Another fitness function, including sparse ratio, training error and testing error, should be defined. To improve calculation effectiveness, mean square value of errors and sparse ratio are considered as optimization goals. Furthermore, sparse ratio is considered as individual influence factor. In processing of training model, fitness function can be calculated and the sample sets can be gained from corresponding to the minimum fitness value, in which the number of sample sets must be less than the total sample sets. Therefore, the constructed LSSVM model is given sparse-feature.

Fitness function is as follows:

$$fitness = \min\left(\frac{1}{m}\sum_{i=1}^{m}\left(\frac{\hat{y}_i - y_i}{y_i}\right)^2 + \frac{1}{n-m}\sum_{j=1}^{n-m}\left(\frac{\hat{y}_j - y_j}{y_i}\right)^2 + \frac{m}{n}\right) \quad (11)$$

Where, $n$ is the number of training samples, $m$ is the sparse sample number reserved, $y_i$ is training sample output, $\hat{y}_i$ is training sample value estimated, $y_j$ is testing sample output, $\hat{y}_j$ is testing sample value estimated.

Sparse process is described briefly as follows:

Step1: Produce random number of [0, 1] as index, which are given to every original sample data.

Step2: Select preserved samples. Compare the index of every original sample data with constant (it is 0.5 here). If the given index is bigger than 0.5, the corresponding sample data will be put into preserved data set as training data, which is a support vector. If the given index is less than 0.5, the corresponding sample data will be put into testing data set.

Step3: Use the new training data to build LSSVM model and calculate fitness function;

Step4: Compare the new value of fitness function with the old value, then, preserve the less one. If the termination criterion is not satisfied, go back to Step1, otherwise, go to Step5.

Step5: The termination criterion is satisfied. Terminate the processing of optimization and output the results.

### *B. PSO*

PSO is a swarm intelligent method mimicking birds' pursuing, which adjusts individual's performance through information sharing between swarms and individual's own experience. By doing so, the best optimization goal can be obtained. The original values of PSO are many random particles and their best optimization solution can be gained by iteration procedure. Particles can trace their individual maximum values and the global maximum value to update their places and their speed. The formula of places' updating and speeds' updating of particles are as follows [21]:

$$v_{id}^{k+1} = \omega v_{id}^{k} + d_1 r_1 (p_{id}^{k} - x_{id}^{k}) + d_2 r_2 (p_{gd}^{k} - x_{id}^{k}) \quad (12)$$

$$x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k} \quad (13)$$

Where, $d_1$ and $d_2$ are both constants; $r_1$ and $r_2$ are both random number; $x_{id}$ is position of particle; $v_{id}$ is speed of particle; $p_{id}$ is the best present position of particle; $p_{gd}$ is the best global position of particle; $k$ is the number of iterations.

The process of optimizing particles is described as follows:

Step 1: Produce random initial particles and define their speeds.

Step2: Produce initial swarms and take them as initial parameters to build training model. Furthermore, ten fold cross validation is carried out and initial fitness function is calculated.

Step3: Initialize maximum individual value and calculate global maximum value.

Step4: Find the best speed and position of particle through iterative computation. Then update the individual maximum value and the swarm maximum value. If the termination criteria are not satisfied, go back to Step4, otherwise, go to Step5.

Step5: If the termination criteria are satisfied, terminate the process of solving and output the results.

As it could not only optimize regularization factor and kernel parameter, but also rarefy support vectors needed, so, the problem can be solved as multi-dimensional parameter optimization. The former $n$ -dimensional parameters are taken as vectors rarefied and the rear $w$ -dimensional parameters are taken as regularization factor and kernel parameter optimization. To avoid over-fitting and underfitting, cross validation is employed to determine fitness function when the parameters are optimized. Then the forecasting precision of model can be improved. The formula is as follows:

$$fitness = \frac{1}{k}\sum_{i=1}^{k}\sum_{j=1}^{n}\frac{1}{n}(y_i - \hat{y}_i)^2 \tag{14}$$

Where, $k$ is the number of cross validation, $k \geq 2$.

To evaluate prediction wind speed, forecasting values should be compared with realistic wind speed. There are several kinds of methods that are usually used to evaluate forecasting performance, which include the mean absolute error (MAE), the root mean squared error (RMSE), and the mean absolute percentage error (MAPE) [13]. They are as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right| \tag{15}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{16}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \times 100\% \tag{17}$$

Where, $y_i$ is observed value, $\hat{y}_i$ is predicted value, $n$ is sample size.

## IV. CASE STUDY AND RESULTS ANALYSIS

To verify the effectiveness of proposed algorithm, the realistic data of wind speed in North China were collected every 10 minutes from the January 1, 2012 to December 31, 2012. In the study, the total actual data were classified into four groups, including spring, summer, autumn and winter that correspond to March 5 to March 8, May 31 to June 3, September 26 to September 29, December 3 to December 6 in 2012 respectively. Through training model and predicting the wind speed ahead based on these data, the obtained results would be reliable. Because there were too much data, 500 consecutive data were selected in each group as experimental data, in which the former 416 data were training data while the later 84 data were testing data. Under training models, the training data set were divided into ten groups, in which each group data should be considered as validation data, while the rest nine groups data should be training data. Then ten models could be obtained. As the mean forecasting effectiveness of ten models was employed as fitness function, finally the training efficiency could be improved.

The four groups of wind speed data in spring, summer, autumn and winter were trained through SPSO-LSSVM. The constructed models were used to predict the testing data after they were trained. To know the performance of SPSO-LSSVM, prediction results are compared, including ANN, ARMA and SPSO-LSSVM.

Now, wind speed deviation between sample times $t + \Delta t$ and $t$ is defined as transient variation value of wind speed. Then the variation amplitudes of wind speed in spring, summer, autumn and winter are not the same, of which the mean absolute fluctuation deviations can be calculated as follows:

$$y_{fde} = \frac{\sum_{i=1}^{p}\left|x_i - \bar{x}_i\right|}{p} \tag{18}$$

Where, $y_{fde}$ is the mean absolute fluctuation deviation, $x_i$ is testing sample data, $\bar{x}_i$ is the mean of testing sample data, $p$ is the number of testing sample data.

Based on the formula (18), the mean absolute fluctuation deviations of the wind speed in spring, summer, autumn and winter are 1.37m/s 、 2.41m/s 、 0.73m/s 、 2.36m/s respectively. The results from Table 1, Table 2 and Table 3 show that the prediction maximal MAE, maximal RMAE and maximal MAPE values appear in summer and take the first place. The values in winter are in the second place. The error values in spring take the third place, while the minimal values occur in autumn. Then it can be considered that the prediction accuracy will decline as the transient fluctuation deviations of wind changes more. From Figure 1 to Figure 4, these figures show that the forecasting ability of SPSO-LSSVM is usually better than that of ANN and ARMA. When normalization factor and kernel parameter are decided different values, the prediction values would change. Sometimes, the variation amplitude is very big.

In the process of constructing LSSVM model, finding correct parameters is so difficult only based on experiences. If like that, more time will be taken and the final results will be unexpected values. As through the PSO algorithm, appropriate parameters can be automatically found out, thus the efficiency will be enhanced. Usually, the best parameters can be also easier gained through PSO algorithm. On the other hand, each input data is included in the model of PSO-LSSVM. Therefore, if the sample data is more, the model is more complicate and will be limited by more factors. So the generalization ability of model will decline. Then a sparse model is built based on JIAO method. The numbers of four preserved samples are 175, 168, 185 and 195 respectively. The corresponding sparse rates are about 58%, 60%, 56% and 53%. As main feature information are included in preserved samples, based on which simpler model can be built. Furthermore, the procedure of modeling is faster. Finally, the generalization ability and prediction accuracy of model are both improved.

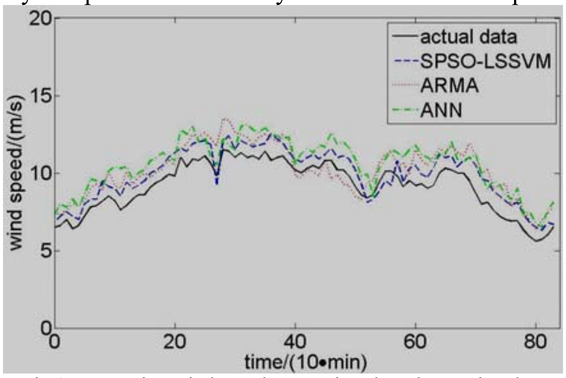


Fig.1 Forecasting wind speed comparison based on spring data

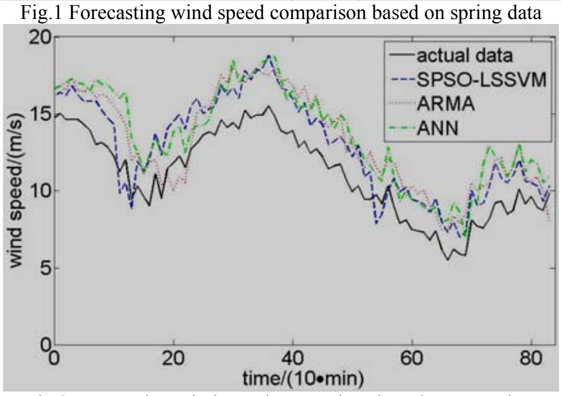


Fig.2 Forecasting wind speed comparison based summer data

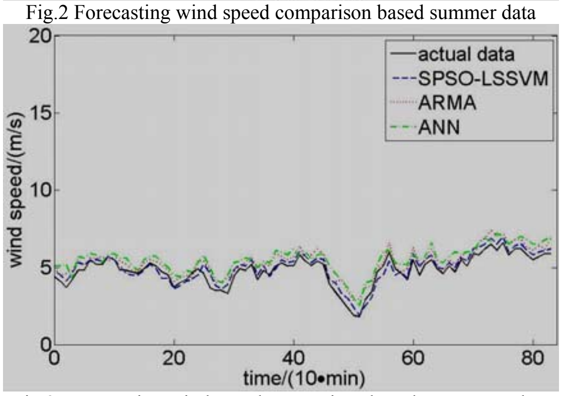


Fig.3 Forecasting wind speed comparison based on autumn data

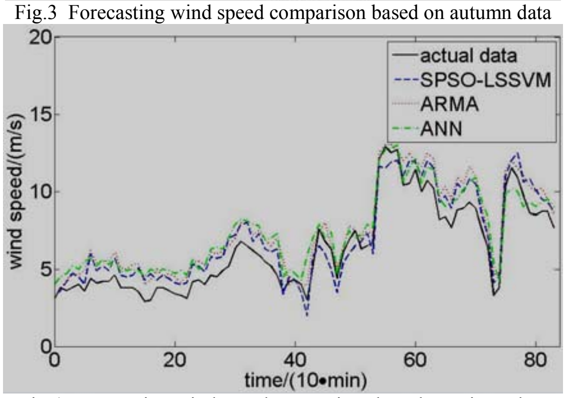


Fig.4 Forecasting wind speed comparison based on winter data

TABLE I MAE (M/S) COMPARING WITH THREE KINDS OF DIFFERENT ALGORITHM USING FOUR GROUPS OF DATA

| **Seasons** | **SPSO-LSSVM** | **ARMA** | **ANN** |
|---|---|---|---|
| Spring | 0.85 | 1.20 | 1.35 |
| Summer | 1.96 | 2.34 | 2.51 |
| Autumn | 0.30 | 0.53 | 0.70 |
| Winter | 0.95 | 1.28 | 1.39 |

TABLE II RMSE (M/S) COMPARING WITH THREE KINDS OF DIFFERENT ALGORITHM USING FOUR GROUPS OF DATA

| **Seasons** | **SPSO-LSSVM** | **ARMA** | **ANN** |
|---|---|---|---|
| Spring | 0.92 | 1.32 | 1.45 |
| Summer | 2.05 | 2.47 | 2.68 |
| Autumn | 0.34 | 0.60 | 0.78 |
| Winter | 1.04 | 1.38 | 1.50 |

TABLE III MAPE (%) COMPARING WITH THREE KINDS OF DIFFERENT ALGORITHM USING FOUR GROUPS OF DATA

| **Seasons** | **SPSO-LSSVM** | **ARMA** | **ANN** |
|---|---|---|---|
| Spring | 9.67 | 13.78 | 16.52 |
| Summer | 18.47 | 22.38 | 23.96 |
| Autumn | 6.90 | 12.46 | 16.25 |
| Winter | 17.20 | 21.17 | 23.02 |

## V. CONCLUSIONS

This paper adopts PSO algorithm to optimize the parameters of LSSVM and the procedure of finding sparse support vector. The fitness function of parameter optimization is confirmed by cross validation. The realistic wind speed in North China in 2012 are selected as samples which are divided into four groups to train and test model. The results of prediction show that: (1) The model of LSSVM, which parameters and the procedure of finding sparse support vector are optimized by PSO, can improve the prediction accuracy, though more training time will be consumed. (2) Each input variable is included in LSSVM model, which makes the structure of the model more complicated. Furthermore, uncertainty will be raised if more input data are included in model. By dint of the sparse support vector to build LSSVM model, the number of input data is reduced and the structure of model is simplified. Then, the generalization and prediction accuracy of model are enhanced. (3) Fluctuation of wind speed affects the accuracy of forecasting wind speed. The greater the wind speed fluctuation is the lower the prediction precision is while the smaller the wind speed fluctuation is the higher the prediction accuracy is. (4) If different sample data are used, the prediction accuracies with the same forecasting algorithm are different.

## ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China (Grant No.51405052), the China Scholarship Council(CSC) and Youth Foundation of

Chongqing University of Posts and Telecommunications (Grant No. A2012-88).

## REFERENCES

[1] Shamshirband Shahaboddin, Petkovic Dalibor, Saboohi Hadi, et al. "Wind turbine power coefficient estimation by soft computing methodologies:Comparative study", Energy Conversion and Management, No.81, pp. 520-526,2014.
[2] Giorgi Maria Grazia De, Campilongo Stefano, Ficarella Antonio, et al. "Comparison between wind power prediction models based on wavelet decomposition with least-support vector machine(LS-SVM) and artificial neural network(ANN)", Energies, No.7, pp. 251-5272, 2014.
[3] Kusiak Andrew, Zhang Zijun, Verma Anoop. "Prediction, operations, and condition monitoring in wind energy", Energy, No.60, pp. 1-12, 2013.
[4] Kusiak Andrew, Zhang Zijun. "Short-horizon prediction of wind power: a data-driven approach", IEEE Transactions on Energy Conversion, No.25, pp. 1112-1122, 2010.
[5] Cadenas Erasmo, Rivera Wilfrido. "Wind speed forecasting in the south coast of Oaxaca, Mexico", Renewable Energy, NO.31, pp. 2116-2128, 2007.
[6] Gomes Pedro, Castro Rui. "Wind speed and wind power forecasting using statistical models: autoregressive moving average(ARMA) and artificial neural networks(ANN)", International Journal of Sustainable Energy Development, No.1, pp. 41-50, 2012.
[7] Erdem E, Shi J. "ARMA based approaches for forecasting the tuple of wind speed and direction", Applied Energy, No.88, pp. 1405-1414, 2011.
[8] Pourmousavi Kani S.A, Ardehali M.M. "Very short-term wind speed prediction: a new artificial neural network-Markov chain model", Energy Conversion and Management, No.52, pp. 738-745,2011.
[9] Li Gong, Shi Jing. "On comparing three artificial neural networks for wind speed forecasting", Applied Energy, No.87, pp. 2313-2320, 2010.
[10] Barbounis Thanasis G, Theocharis John B, Alexiadis Minas C et al. "Long-Term wind speed and power forecasting using local recurrent neural network models", IEEE Transactions on Energy Conversion, No.21, pp. 273-284, 2006.
[11] Monfared M, Rastegar H, Kojabadi Hm. "A new strategy for wind speed forecasting using artificial intelligent methods", Renewable Energy, No.34, pp. 845-848, 2009.
[12] Liu H, Tian H, Chen C. "A hybrid statistical method to predict wind speed and wind power", Renewable Energy, No.35, pp. 1857-1861, 2010.
[13] Sun Wen, Liu Mohan, Liang Yi. "Wind speed forecasting based on FEEMD and LSSVM optimized by the bat algorithm", Energies, No.8, pp. 6585-6607, 2015.
[14] Mohandes M.A., Halawani T.O., Rehman S. et al. "Support vector machines for wind speed prediction", Renewable Energy, No.29, pp. 939-947, 2004.
[15] Ch Sudheer, Anand Nitin, Panigrahi B.K. et al. "Streamflow forecasting by SVM with quantum behaved particle swarm optimization", Neurocomputing, No.101, pp. 18-23, 2013.
[16] Zheng H.B., Liao R.J., Grzybowski S.. "Fault diagnosis of power transformers using multi-class least square support vector machines classifiers with particle swarm optimization", IET Electr. Power Appl, No.5, pp. 691-696, 2015.
[17] Liu Ruilan, Xu Yan, Rong Zhou. "Modeling soft sensor based on sparse least square support vector machine", CIESC Journal, No.66, pp. 1402-1406, 2015.
[18] Thukaram D., Khincha H.P., Vijaynarasimha H.P.. "Artificial neural network and support vector machine approach for locating faults in radial distribution systems", IEEE Transactions on Power Delivery, No.20, pp. 710-721, 2005.
[19] Zhou Junyi, Shi Jing, Li Gong. "Fine tuning support vector machines for short-term wind speed forecasting", Energy Conversion and Management, No.52, pp. 1990-1998, 2011.
[20] Zhao Yongping, Sun Jianguo. "Fast method for sparse least squares support vector regression machine", Control and Decision, No.23, pp. 1347-1352, 2008.
[21] Mei Fei, Mei Jun, Zheng Jianyong, et al. "Application of particle swarm fused KFCM and classification model of SVM for fault diagnosis of circuit breaker", Proceedings of the CSEE, No.33, pp. 134-141, 2013.