JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

Short-Term Wind Speed Forecast Based on Least Squares Support Vector Machine

Yanling Wang*, Xing Zhou*, Likai Liang*, Mingjun Zhang**, Qiang Zhang***, and Zhiqiang Niu****

Abstract

There are many factors that affect the wind speed. In addition, the randomness of wind speed also leads to low prediction accuracy for wind speed. According to this situation, this paper constructs the short-time forecasting model based on the least squares support vector machines (LSSVM) to forecast the wind speed. The basis of the model used in this paper is support vector regression (SVR), which is used to calculate the regression relationships between the historical data and forecasting data of wind speed. In order to improve the forecast precision, historical data is clustered by cluster analysis so that the historical data whose changing trend is similar with the forecasting data can be filtered out. The filtered historical data is used as the training samples for SVR and the parameters would be optimized by particle swarm optimization (PSO). The forecasting model is tested by actual data and the forecast precision is more accurate than the industry standards. The results prove the feasibility and reliability of the model.

Keywords

Cluster Analysis, Least Squares, Least Squares Support Vector Regression (LSSVR), Particle Swarm Optimization (PSO), Short-Time Wind Speed Forecasting

1. Introduction

The speed and direction of wind are affected by many factors with large randomness and the accuracy of wind prediction is necessary. But the current research results about wind speed forecasting is not very satisfactory. Wind speed forecasting methods can be divided into two categories. One is based on physical model and the other is based on historical data. The forecasting methods based on physical model commonly use numerical weather to forecast the wind speed [1]. Wind speed forecasting of numerical weather prediction is not for a certain wind turbines but for an area, so we need to solve how to map the wind speed of a unit area into the wind speed of a certain wind turbine generator [2]. The forecasting methods based on historical data usually use the Kalman filter method [3], artificial neural network [4], fuzzy logic method [5], time series method [6], support vector machine (SVM) [7], and so on. The above-mentioned method uses data correlation to predict wind speed, and the hugeness of historical data samples increases the computational complexity.

In this paper, we propose the method of least squares support vector machine (LSSVM) to solve the defects. This model uses Ward clustering analysis method to get training samples with high similarity. This model chooses RBF as kernel function, taking Gaussian radial basis function (RBF) with good simple structure into consideration, and optimizes the related parameters by using the particle swarm optimization (PSO). The quantity of samples is reduced by the classification and selection of historical data, and the prediction accuracy is improved. The actual data test results show the reliability and feasibility of the model, and also improve the forecast accuracy.

This paper is organized as follows: Section 2 introduces the model diagram of short-term wind speed forecast based on LSSVM. The prediction model and principle of SVM, clustering analysis, and PSO are described in detail. In Section 3, we give the improved prediction process based on LSSVM. Example results and performance analysis are presented in Section 4, and conclusions are provided in Section 5.

2. The Model of Short-Term Wind Speed Forecasting Based on LSSVM

Ward clustering analysis method is adopted to get training samples of the model based on LSSVM. It chooses the appropriate kernel function, and optimizes the related parameters by using the PSO to realize short-term speed forecasting. Its model diagram is shown in Fig. 1.
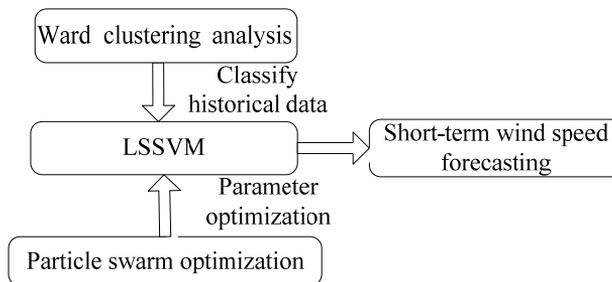

Fig. 1. Short-term wind speed forecasting model.

2.1 Support Vector Regression Machine

SVM can not only solve the problem of classification but also deal with regression problems successfully, and get extensive application in the field of pattern recognition [8,9]. SVM is applied in classification domain by structural classification function and it can classify unknown samples according to the classification function [10-12]. The detailed algorithm is as below:

Training samples set $T = \{(x_i, z_i), i = 1, 2, \cdots, n\}$, $x_i \in R^n$ is the input vector, n is the number of samples and $z_i \in \{1, -1\}$ is the output category flag. The classification function in linear support vector classification machine is:

$$f_1(x) = \text{sgn}[\sum_{i=1}^{n} a_i^* z_i (x_i \cdot x) + b^*]$$ (1)

where x stands for unknown samples, a_i^* is the optimal regression coefficient, and b^* is the optimal deviation. Linear SVM can realize the classification of linearly separable data samples, but the linear SVM cannot classify the sample data with large overlap region. In order to solve the nonlinear problems, low dimensional input variables can be put into high dimensional space to seek the optimal classification plane. The inner product in high dimensional space makes the operation very complex. In order to solve this problem, kernel function is introduced [13]. Hidden computation of kernel function instead of inner product operation in high dimensional feature space can avoid the complicated calculation in high dimensional space. Based on this, the classification function of the nonlinear problem is:

$$f_1(x) = \text{sgn}[\sum_{i=1}^{n} a_i^* z_i K(x_i \cdot x) + b^*] \tag{2}$$

The use of kernel function makes the input function still operated in the original low dimensional space. Changing the parameters of the kernel function is equivalent to do corresponding adjustment for the mapping function, which is to change the distribution of data in high dimensional space. There are many kinds of kernel functions, and different kernel functions can be used to construct different types of nonlinear decision plane. The kernel functions which often be used are polynomial, linear kernel function, neural network kernel function, and Gaussian radial basis kernel function. In this paper, the kernel function model is Gaussian RBF:

$$K(x, x_i) = \exp[-\frac{\|x - x_i\|^2}{\sigma^2}] \tag{3}$$

Gaussian RBF has good simple structure, adaptability, strong generalization ability and wide domain of convergence. Arbitrary distribution of samples can be mapped to high dimensional space and the kernel parameter σ changes within the effective range, which does not make the space too complicated. It is easy to facilitate the optimization of the support vector regression (SVR).

SVM can be used to classify. In addition, it can also be applied in the field of function regression, namely SVR machine [14]. The use of SVM in the field of function regression is similar to the basic principle of support vector classification machine. Sample set Q contains an input parameter for n independent samples as the input vector, and training samples set is $\{(x_1, y_1), (x_2, y_2) \cdots, (x_n, y_n)\}$. $x_i \in R^n$, $y_i \in R$ are the corresponding target values, so the regression prediction function is:

$$f_2(x) = (\omega \cdot x) + b \tag{4}$$

where ω represents the complexity of the control model, and b is the deviation. Among them, $\omega \in R^n$, $b \in R$. The error control function is introduced to measure the difference between the predictive values and the actual values. There are many kinds of error control functions, such as linear function, Huber function, and ε-insensitive function. Different control functions can be used to build different SVR. ε-insensitive function not only reduces the complexity of the operation, but also has many good characteristics. Therefore, we choose the ε-insensitive function as the error control function:

$$|y_i - f_2(x_i)| = \begin{cases} 0, & if\ |y_i - f_2(x_i)| \le \varepsilon \\ |y_i - f_2(x_i)| - \varepsilon, & if\ |y_i - f_2(x_i)| > \varepsilon \end{cases} \tag{5}$$

where $f_2(x_i)$ represents the prediction value of the output of the regression prediction function, and the distance between the predicted value and the actual value of the $f_2(x)$, which is constructed in the end, should be less than ε. But it is impossible for all the samples to be fitted without error in the accuracy of ε and some sample points deviate from the predicted function frequently. The existence of errors is allowed in the regression function, besides in order to measure the degree of deviation from these points, the relaxation factor is further introduced. By introducing Lagrange function and Lagrange dual problem, the optimal a_i^* is obtained, and then the optimal ω^* and optimal b^* are obtained. The regression function is as following by solved:

$$f_2(x) = \sum_{i=1}^{n} (a_i^* - a_i)(x \cdot x_i) + b^* \tag{6}$$

For the nonlinear regression, first of all, the training data is mapped to high dimensional space from low dimensional, then the optimal regression function is got and the expression is shown in Eq. (7).

$$f_2(x) = \sum_{i=1}^{n} (a_i^* - a_i)K(x \cdot x_i) + b^* \tag{7}$$

The complexity of the SVR algorithm is affected by the number of samples, when the sample size increases, the computing speed will decline. Least squares support vector regression machine is on the basis of SVR algorithm by joining the squared error. Regression prediction function is:

$$f_2(x) = \omega^{\mathrm{T}} \varphi(x) + b \tag{8}$$

The $\varphi(x)$ in Eq. (8) is nonlinear function, which maps the low dimensional input samples to high dimensional space. The optimization problem for LSSVR is:

$$\min \quad J_p(\omega, e) = \frac{1}{2}\|\omega\| + \frac{1}{2}\gamma \sum_{i=1}^{n} e_i^2 \tag{9}$$

where J_p means the objective function that is optimized, γ represents the regularization parameter, and e_i represents the relaxation factor. The value of γ is correlated with the regression error of the model. The larger γ is, the smaller the regression error. By introducing the Lagrange function, the optimization problem is transformed into solving the linear system of equations. Finally, the nonlinear regression function of LSSVM model is obtained:

$$f_2(x) = \sum_{i=1}^{n} a_i K(x_i, x) + b \tag{10}$$

After determining the regression prediction function, the corresponding forecast data $f_2(x_i)$ will be obtained by the input data x_i. So the problem about SVR is converted into problem about solving equations of linear system.

2.2 Clustering Analysis

If the regularity of the training data used in the model is similar to the forecasting data, the prediction

error will be greatly reduced. So cluster analysis can be used to find similar historical data.

Clustering analysis method is used for quantitative analysis of things and the distinction among things is measured by numbers [15]. The things are divided into many groups and things within the same group are sorted in a class. The things in the same group have higher similarity while those of different groups don't, so the reliable classification basis is given from the angle of mathematical analysis. The system clustering method is applied to the historical data for SVR. In the beginning, each sample is one class in system clustering method and each step is to find the nearest two classes. There are many calculation methods for calculating distance. We choose the sum of squared residuals method (Ward method).

If

$$D_1 = \sum_{x_i \in G_1} (x_i - \overline{x_1})^{\mathrm{T}} (x_i - \overline{x_1}), D_2 = \sum_{x_j \in G_2} (x_j - \overline{x_2})^{\mathrm{T}} (x_j - \overline{x_2}) \tag{11}$$

$$D_{12} = \sum_{x_k \in G_1 \cup G_2} (x_k - \overline{x})^{\mathrm{T}} (x_k - \overline{x}) \tag{12}$$

and

$$\overline{x_1} = \frac{1}{n_1} \sum_{x_i \in G_1} x_i, \overline{x_2} = \frac{1}{n_2} \sum_{x_j \in G_2} x_j, \overline{x} = \frac{1}{n_1 + n_2} \sum_{x_k \in G_1 \cup G_2} x_k \tag{13}$$

so

$$D(G_1, G_2) = D_{12} - D_1 - D_2 \tag{14}$$

where G_1, G_2 is on behalf of the two kinds of classes, n_1 and n_2 are the number of sample points in G_1 and G_2. $D(G_1, G_2)$ is on behalf of the distance between the two classes. For the Ward method, if the distance between G_1, G_2 is enough short, they can be divided into the same class. If the distance between G_1, G_2 is very long, it makes $D(G_1, G_2)$ bigger. When the sample is classified, a sample will be seen as a class. The nearest two samples are merged into one class, namely the nature of the nearest two samples is the most close. At this point, n samples are divided into ($n - 1$) classes. Similarly, all samples will be consolidated into one class finally. The quantity of classes should be determined first in the application of the clustering method, and then samples should be divided by cluster analysis, until it obtains the classes determined.

2.3 Particle Swarm Optimization

LSSVR model can establish regression equation to forecast, but different parameters will have great influence on the model. In order to minimize the prediction error, the PSO is used to determine the optimal parameters [16,17].

Each particle represents a possible solution of optimization problem in the particle swarm algorithm group and the location of the food is on behalf of the optimal solution. The group finds the optimal solution in the N-dimensional space. The PSO algorithm first initializes a set of random particles and searches the optimal solution by multiple iterations. In each iteration, the particles update themselves by tracking two extremes [18]: the individual extremes p_{best} and the global extremes g_{best}. The individual extremum is the optimal position where a particle has been passed. The global extremum is the best one in the optimal positions of all the particles in the individual particle group. The particle updates its

speed and position according to the above two extremes. Assuming an existing particle swarm, at time t, the coordinates of each particle are $\overline{x_i}(t) = (x_i^1, x_i^2, \cdots, x_i^n, \cdots, x_i^N)$ and the velocities of the particles are $\overline{v_i}(t) = (v_i^1, v_i^2, \cdots, v_i^n, \cdots, v_i^N)$. The coordinate position $\overline{x_i}(t)$ and the velocity $\overline{v_i}(t)$ are adjusted at time $t+1$ according to Eq. (15).

$$
\begin{aligned}
&\overline{v_i}(t+1) = \overline{v_i}(t) + c_1 r_1(\overline{p_{\text{best}}}(t) - \overline{x_i}(t)) + c_2 r_2(\overline{g_{\text{best}}}(t) - \overline{x_i}(t)) \\
&\overline{x_i}(t+1) = \overline{x_i}(t) + \overline{v_i}(t+1) \\
&\begin{cases} v_i^n = v_{\max}, & v_i^n > v_{\max} \\ v_i^n = -v_{\max}, & v_i^n < -v_{\max} \end{cases}
\end{aligned}
\tag{15}
$$

where c_1 and c_2 represent the acceleration constants of the particles, and the range of values is usually [0,2]. r_1 and r_2 are two random numbers that are evenly distributed in [0,1]. $\overline{p_{\text{best}}}(t)$ is the mean of the individual extremes and $\overline{g_{\text{best}}}$ is the mean of the global extremes.

3. Forecasting Process

This part designs a process for forecasting model of short-term wind speed that is based on theories about least square support vector regression machine [12], Ward clustering algorithm, and PSO introduced above. We get specific wind speed for one actual wind farm according to the process of the model and give detailed analysis.

This model is based on LSSVM. It uses Ward clustering analysis to get training samples for the model through the historical wind data classification and selection. We choose the appropriate kernel function, and use PSO algorithm to optimize the relevant parameters. Then the model realizes the relatively accurate forecasting of short-term wind speed, and the model prediction process is as follows:

Step 1. The model uses the Ward clustering analysis to classify historical data, and then we will choose historical data whose characteristics is similar to the data used for forecasting as training samples of LSSVR.

Step 2. The training sample data should be normalized and rationally preprocessed, then it is trained by LSSVR. Finally, the regression prediction model is established. The normalized equation is such as Eq. (16).

$$
\tilde{x}_i = \frac{x_i - x_{i\min}}{x_{i\max} - x_i} (i = 1, 2, \cdots, n)
\tag{16}
$$

where $\tilde{x}_i$ is the normalized value, x_i is the measured wind speed value, $x_{i\min} = \min(x_i)$, $x_{i\max} = \max(x_i)$.

Step 3. The reasonable parameters of PSO will be determined to optimize the parameters of the LSSVR.

Step 4. LSSVR regression model whose parameters are optimized should be tested, and it is evaluated by using the test data to calculate the precision.

Step 5. If the test does not meet the precision requirement, we should determine parameters by retraining LSSVR again. Else the regression forecasting model will be determined, then we can forecast wind speed by input data we have.

4. Example Analysis

4.1 Practical Application of Clustering Analysis

Firstly, we determine and classify historical data through Ward cluster analysis, then give the total number n of classes to be separated. The model uses the Ward clustering analysis to classify historical data that is 30 days before the day we want to forecast. The sample data by SVR do not have high similarity with the predicted wind speed data, so we decided to divide the data into three categories. Then we will choose historical data whose characteristic is similar to the data used for forecasting as training samples of LSSVR. Now 10 groups of historical data and predicted input data are taken as examples to demonstrate the Ward cluster analysis method. Ten groups of historical wind speed data and predicted input data are shown in Fig. 2.

The data of the 10-day historical wind speed is classified by Ward clustering method. The historical data of 10 days and a set of input forecast data are divided into three classes, named I, II, III type respectively. The classification results are shown in Table 1.
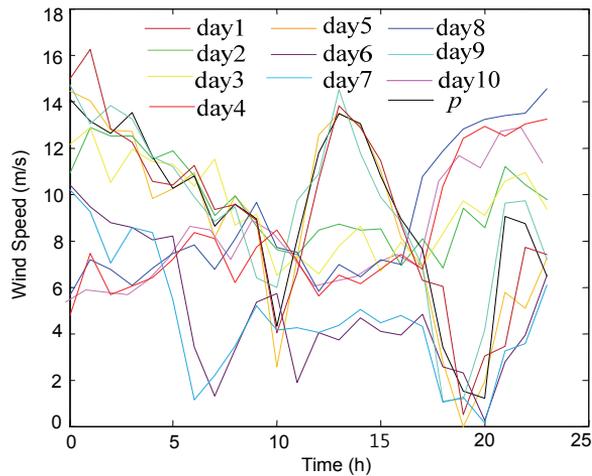

Fig. 2. Data of the 10-day historical wind speed.

Table 1. History data classification results

Category	Date of wind speed
I	day6, day7
II	day1, day5, day9, p
III	day2, day3, day4, day8, day10

We use number $\mathrm{day}1, \mathrm{day}2, \cdots, \mathrm{day}10$ to represent 10-day historical wind speed data respectively, and p stands for input data used to forecast. As seen in the Table 1, the input data is assigned to class, that is to say, when we use the input data to realize the forecast of short-term wind speed, the historical data whose label is day1, day5, day9 should be used as the training sample of LSSVR. Practical prediction used 30 groups historical data, but only 10 groups of historical data was used to display

Ward clustering analysis method just for convenience. The trends in wind speed data for categories I, II, and III are shown in Fig. 3, Fig. 4, and Fig. 5 respectively.
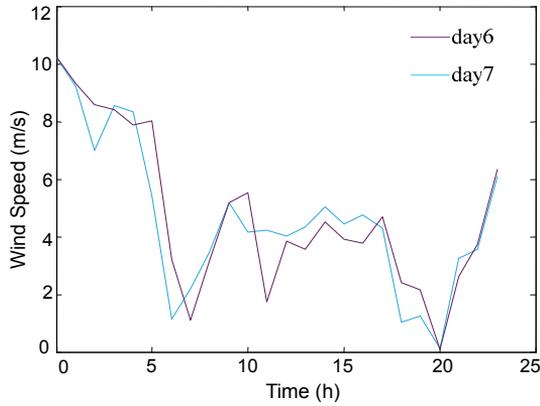

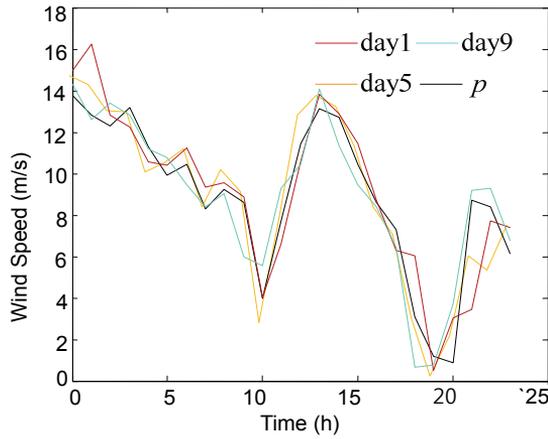**Fig. 3.** Daily variation of wind speed curve of category I.

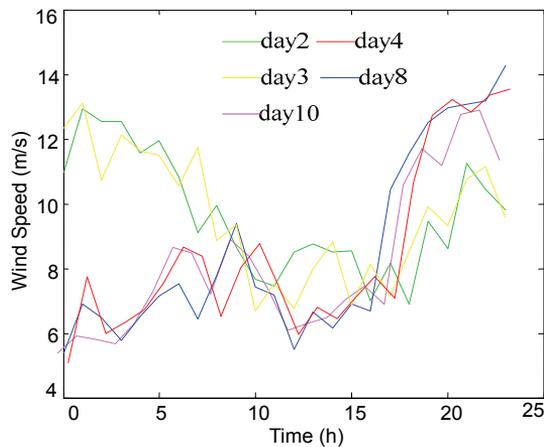**Fig. 4.** Daily variation of wind speed curve of category II.

Fig. 5. Daily variation of wind speed curve of category III.

We can see that the change trend of data of day1, day5, day9, p is consistent enough through observing the change trend of wind speed data of classes I, II, and III. So the Ward clustering analysis method can realize the accurate classification of historical data, and the classification results are satisfactory.

4.2 Parameter Optimization Based on Particle Swarm Optimization

The parameters of the least squares support vector regression machine will change the performance of SVM, and then affect the prediction accuracy of the model. However, the SVM itself cannot determine the optimal parameters. In order to optimize the selection of parameters, PSO algorithm is adopted to improve the prediction accuracy of the model.

During the learning of least squares support vector regression machine, part of data is selected as training data, another part of the data is used to test. Then the model is evaluated according to the test results. Finally, the model chooses the best parameters. The measurement criterion is the fitness function. We select the average relative error as the fitness function, such as Eq. (17).

$$f_{\text{MSE}} = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} \tag{17}$$

In Eq. (17), f_{MSE} is on behalf of the fitness function value, y_i stands for actual wind speed and $\hat{y}_i$ represents the wind speed that is forecasted. Therefore, there are total n sets of data.

The parameters of this model which need to be optimized are: the regularization parameter γ and kernel parameter σ. First, a group of random particles should be initialized. Each particle represents a set of values of the regularization parameter and kernel parameter. The initial position is an arbitrary particle, and the arbitrary particle is trained by LSSVR. We calculate the value of fitness function. A set of optimal regularization parameter and kernel parameter can be found by iterative trained, thus the optimization of LSSVR parameters can be realized.

4.3 Selection of Input Data

When predicting the wind speed at a certain time, it is necessary to select a certain amount of historical data as input variables. The number of input independent variables will also have a certain impact on the accuracy of the forecast. Inadequate input variables cause the relationship between the predicted output data and the input data cannot be fully reflected. However, adequate input variables lead to data redundancy, which makes the inherent relationship between the data cannot be adequately represented. In order to select the input independent variables, the average wind speed data for 3 hours, 5 hours, and 8 hours in succession are selected as the input. The selection of the input independent variables is determined by observing the prediction error and the analysis results are shown in Table 2.

Table 2. The comparison of input variables

Input variable	Average relative error
3 hours before the wind speed	0.3749
5 hours before the wind speed	0.1622
8 hours before the wind speed	0.2874

It can be seen that the prediction accuracy is the highest when the wind speed data is selected as the input independent variable for 5 hours before the prediction time, that is to say, when the wind speed data for t moment was predicted, $t-1$, $t-2$, $t-3$, $t-4$, $t-5$ time wind speed data as input, can realize the prediction of short-term wind speed.

4.4 Model Test

Now we forecast short-term wind speed of 24 hours for a certain farm (short-term prediction means predicting every 15 minutes). We select eight days whose wind speed is similar to data for prediction by Ward clustering analysis as the training sample of LSSVR, so it is total 8×24×4 sets of training data and 24×4 sets of predicted data. Related parameters of PSO algorithm are shown in Table 3.

The value of forecasted wind speed and actual wind speed is shown in Fig. 6.

Table 3. Parameters of particle swarm optimization algorithm

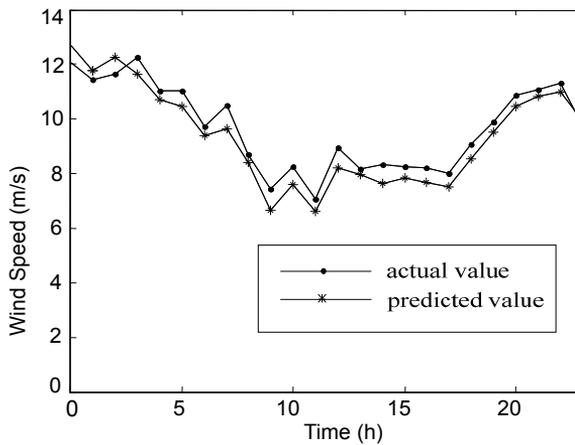| Population size n	Maximum iterations	Acceleration constant c_1	Acceleration constant c_2
20	200	1.5	1.7

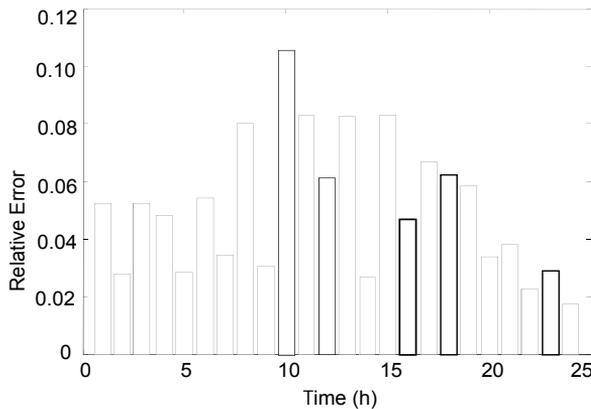**Fig. 6.** Predicted value of wind speed and actual value of wind speed.

Fig. 7. Relative error of prediction point.

In Fig. 6, the predicted result of wind speed is close to the actual value, and the model can accurately predict the wind speed. The relative error of each prediction point is calculated, and the relative error of each prediction point is shown in Fig. 7.

As seen from Fig. 7, the prediction error of the model is stable. The average relative error of the estimates is 5.11% and the maximum relative error is not more than 14%. For forecasting of short-term wind speed, the industry standard of relative error issued by the State Grid should be less than 20% [20], and the prediction data we get meets the industry standard. It proves that the prediction of the model performances is good and the model is reliable.

5. Conclusion

In this paper, we have proposed short-term wind speed forecast based on LSSVM. Compared with the traditional short-term wind speed forecasting methods, the proposed model is applied in regression field to find the regression relationship between historical data and forecast data. When choosing the training sample, the model uses cluster analysis innovatively to select appropriate historical data rather than chooses a large number of historical data casually as sample.

The model observably improves the quality of the training samples, hence the accuracy and reliability of the model can be improved. Now, the view of SVM cannot select the optimal parameters by itself, so it is combined with PSO algorithm to optimize the parameters. Hence, the model's prediction error can be limited at about 5%, and the result of the model can get great improvement compared with the traditional methods.

Acknowledgement

References

[1] Y. Q. Liu, Y. M. Wang, and D. Infield, "Numerical weather prediction wind correction methods and its impact on computational fluid dynamics based wind power forecasting," *Journal of Renewable and Sustainable Energy*, vol. 8, no. 3, pp. 1-14, 2016.

[2] S. Tohidi, A. Rabiee, and M. Parniani, "Influence of model simplifications and parameters on dynamic performance of grid connected fixed speed wind turbines," in *Proceedings of the 19th International Conference on Electrical Machines*, Rome, Italy, 2010, pp. 1-7.

[3] D. P. Bertsekas, "Incremental least squares methods and the extended Kalman filter," *SIAM Journal on Optimization*, vol. 6, no. 3, pp. 807-822, 1996.

[4] R. J. Kuligowski and A. P. Barros, "Localized precipitation forecasts from a numerical weather prediction model using artificial neural networks," *Weather & Forecasting*, vol. 13, no. 4, pp. 1194-1204, 1998.

[5] L. Landberg, "Short-term prediction of local wind conditions," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 89, no. 3-4, pp. 235-245, 2001.

[6] S. V. Sobrad and S. H. Bajantri, "Wind speed forecasting based on time series analysis, wavelet analysis and GARCH analysis," *International Journal of Advance Research in Engineering, Science & Technology*, vol. 3, no. 6, pp. 356-361, 2016.

[7] H. Q. Wang, "Short term wind speed prediction based on support vector machine," *Natural Science*, vol. 29, no. 1, pp. 16-18, Mar. 2009.

[8] R. AhilaPriyadharshini and S. Arivazhagan, "Object recognition based on local steering kernel and SVM," *International Journal of Engineering, Transactions B: Applications*, vol. 26, no. 11, pp. 1281-1288, 2013.

[9] A. Goel and M. Pal, "Stage-discharge modeling using support vector machines," *IJE Transactions A: Basics*, vol. 25, no. 1, pp. 1-9, 2012.

[10] R. Li, Q. Chen, and H. R. Xu, "Wind speed forecasting method based on LS-SVM considering the related factors," *Power System Protection And Control*, vol. 38, no. 21, pp. 146-151, 2010.

[11] G. M. Zhang, Y. H. Yuan, and S. J. Gao, "A predictive model of short-term wind speed based on improved least squares support vector machine algorithm," *Journal of Shanghai Jiao Tong University*, vol. 45, no. 8, pp. 1125-1129, 2011.

[12] B. Sun and H. T. Yao, "The short-term wind speed forecast analysis based on the PSO-LSSVM predict model," *Power System Protection and Control*, vol. 40, no. 5, pp. 85-89, 2012.

[13] J. Zhang, "The research on short-term forecast of wind farm wind speed method based on support vector machine (SVM)," M.S. theses, Huaqiao University, Fujian, China, 2012.

[14] J. Zeng and H. Zhang, "wind speed forecasting model based on least squares Support Vector Machine," *Power System Technology*, vol. 33, no. 18, pp. 144-147, 2009.

[15] H. W. Peng, "The research on enterprise accident prevention and control level based on Ward clustering method," *Business Herald*, vol. 9, pp. 263-264, 2012.

[16] B. B. Xu, Z. G. Zhang, X. L. Zheng, X. T. He, and Q. Jing, "Application overview of least squares support vector machine in short term wind speed forecasting," *Electrical Technology*, vol. 6, pp. 22-25, 2013.

[17] H. Bagheri Tolabi, M. H. Moradi and F. Bagheri Tolabi, "New technique for global solar radiation forecast using bees algorithm," *IJE Transactions B: Applications*, vol. 26, no. 11, pp. 1385-1392, 2013

[18] N. Li, "Particle swarm optimization (PSO) algorithm and its application research," *Agricultural Network Information*, vol. 9, pp. 146-148, 2010.

[19] *Specification of Wind Power Prediction Function (Q/GDW 588-2011)*. Beijing: China Electric Power Press, 2011.

Yanling Wang https://orcid.org/0000-0003-0467-8356

She received her Ph.D. degree from Shandong University, China and now teaches in School of Mechanical Electrical and Information Engineering at Shandong University, China. Her current research areas are smart grid, power grid transmission capacity, power system operation and control.

Xing Zhou https://orcid.org/0000-0001-7094-1250

She is a postgraduate in School of Mechanical Electrical and Information Engineering at Shandong University, China. She is major in circuit and system. Her main research interests include current-carrying transmission and automation of electric power systems.

Likai Liang https://orcid.org/0000-0003-1271-6906

She received her Ph.D. degree from School of Electrical Engineering from Shandong University, China in 2013. She teaches as an associate professor in School of Mechanical Electrical and Information Engineering at Shandong University, China. Her current research interests include power system operation and control.

Mingjun Zhang https://orcid.org/0000-0002-8090-5963

He graduated from Mudanjiang Electric Power Technology School in 1989. He now works as an assistant engineer in State Grid Jiamusi Power Supply Co. Ltd. His current research areas are smart grid and new energy power generation.

Qiang Zhang https://orcid.org/0000-0003-1457-5166

He received her Ph.D. degree from School of Electrical Engineering from Shandong University in 2007. He is currently a Senior Engineer with Shandong Electric Power Dispatching Control Center, Jinan, China. His current research interests include electric power system theory, operation and control.

Zhiqiang Niu https://orcid.org/0000-0001-6494-7795

He received his M.S. degree from Shandong University and now works in State Grid Weihai Power Supply Company. His current research interests include power network dispatching automation, and power system optimal operation