Forecasting of Wind Speed Based on Wavelet Analysis and Support Vector Machine

Zhou Songlin*, Mao Meiqin*, Liuchen chang**
* Research Center for Photovoltaic System Engineering Ministry of Education, Hefei University of Technology,Hefei 230009, China;
** University of New Brunswick，Canada

Abstract--- **A new forecasting method of wind speed based on wavelet analysis and support vector machine is proposed in this paper. By the method, firstly, the original wind speed sequences are decomposed into coarse components and detail components, and secondly, every wavelet components are separately forecasted with corresponding support vector machine models. Finally, the forecasting results of original wind speed series are achieved by using wavelet reconstruction .The experimental results by simulation prove that this method is capable of improving generalization performance and forecasting precision. The mean absolute percentage error (MAPE) reaches 8.21%.**

*Index Terms--***Wind speed forecasting; Support vector machine; Wavelet analysis; Generalization performance**

I. Introduction

The output power of wind power plants has a close relation with wind speed. Wind speed presents strong randomness. When it exceeds a certain threshold, it will bring about a remarkable impact on power quality and reliability [1] [2]. High forecasting accuracy of wind speed has a practical significance for utility company to regulate dispatching plan. At present, the wind speed forecasting error in wind plant is about 20%-40% [3]. This error is not only related to forecasting method, but also related to forecasting cycle, position and wind characteristic.

There have been various methods developed for forecasting of wind speed. Specifically, as wavelet multi-resolution analysis method is used to decompose the wind speed sequence into coarse components and various detail components, a higher forecasting accuracy can be obtained.

Reference [4] associated wavelet with auto -regressions (AR) model to forecast wind speed and obtained higher accuracy than only using AR model. Due to the limitation of predicting ability of AR model itself, the accuracy of this method was not so satisfied. Support vector machine (SVM) is a new forecasting method based on structural risk minimization principle, which has better generalization performance than ANN and AR. In recent years, SVM is also used to forecast wind speed. Reference [5] applied SVM to predict short-time wind speed, the mean absolute percent error (MAPE) of which was 24%. Though the accuracy was higher than ANN, it was not satisfactory. Reference [6] proposed an approach of wind speed forecasting with least squares support vector machine (LS-SVM) and achieved a good effect. Its MAPE reached 12.53%. However, wind speed time series have periodicity and strong randomness, which comprise various spectrums, so the forecasting system adopting SVM alone was not able to provide the acceptable accuracy of prediction. To improve prediction accuracy further, a new method based on both wavelet analysis and SVM is proposed in this paper. Multi-resolution analysis method is applied to decompose the wind speed sequence at a certain extent into coarse components and detail components firstly, and then different forecasting models are established based on the corresponding characteristics of various wavelet components. Finally, the forecasting results of all components are reconstructed into the final results. The method is used to forecast the wind speed in a wind power plant and a higher forecasting accuracy is obtained.

This work is supported by Special Fund of the National Basic Research Program of China (No. 2009CB219708).

II. Wavelet Decomposition And Reconstruction

By using wavelet multi-resolution analysis[7], discrete sequence can be decomposed into low-frequency coarse coefficient $c_1(t)$ and high-frequency detail coefficient $d_1(t)$, and then $c_1(t)$ can further be decomposed. By this transformation, we get a set of coefficients which represent detail signals and coarse signals at different levels. Decomposition coefficients can be expressed as:

$$c_{j,k} = \sum_m h(m-2k)c_{j-1} \tag{1}$$

$$d_{j,k} = \sum_m g(m-2k)d_{j-1,k} \tag{2}$$

where k,m are the shift parameters, $c_{j,k}$ is the low-frequency coefficient, $d_{j,k}$ is the high-frequency coefficient, $h(m-2k)$, $g(m-2k)$ are high-pass filter and low-pass filter respectively..

On the contrary, the original signal can be reconstructed as Eq. (3)

$$c_{j-1,k} = \sum_k c_{j,k}h(m-2k) + \sum_k d_{j,k}g(m-2k) \tag{3}$$

III. Support Vector Machine

Given a set of data $G = \{(x_i, d_i)\}_{i=1}^N$ (where, x_i is the input vector; d_i is the actual value, and N is the

Authorized licensed use limited to: Aarhus University. Downloaded on March 28,2022 at 20:04:48 UTC from IEEE Xplore. Restrictions apply.

total number of data patterns), the regression function is formulated as follows.

$$f(x) = < w \cdot \Phi(x) > +b \tag{4}$$

Where, x is the feature of the inputs, ω and b are coefficients. Its constraint equation is:

$$\begin{cases} y_i - < w \cdot \Phi(x_i) > -b \le \varepsilon \\ < w \cdot \Phi(x_i) > +b - y_i \le \varepsilon \end{cases} \tag{5}$$

where $i = 1,2,\cdots,k$.

According to statistical learning theory, the optimization objectives should be the term $\min(\frac{1}{2}\|w\|^2)$. Taking into account the tolerance error, we introduce the relaxation factor $\xi_i^* \ge 0$ and $\xi_i \ge 0$, then Eq. (5) is turned to Eq. (6).

$$\begin{cases} y_i - < w \cdot \Phi(x_i) > -b \le \varepsilon + \xi_i \\ < w \cdot \Phi(x_i) > +b - y_i \le \varepsilon + \xi_i^* \end{cases} \tag{6}$$

The optimization objective becomes

$$J = \min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{k}(\xi_i + \xi_i^*),$$

where, C is balance factor, which indicates a compromise between the smoothness of f and the tolerance error that is larger than ε. C is used to play a role of regulation between improving generalization ability and reducing error. Lagrange function is introduced as Eq. (7)

$$\begin{aligned} L(w,b,\xi_i,\xi_i^*) &= \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*) - \sum_{i}^{l}\alpha_i[\varepsilon + \xi_i \\ &- y_i + (< w \cdot \Phi(x_i) > +b)] - \sum_{i=1}^{1}\alpha_i^*[\varepsilon + \xi_i^* + y_i \\ &- (< w \cdot \Phi(x_i) > +b)] - \sum_{i=1}^{l}(n_i\xi_i + \eta_i^*\xi_i^*) \end{aligned} \tag{7}$$

where w, b, ξ_i, ξ_i^* are original variables, α_i, α_i^*, η_i, η_i^* are dual variables, and $\alpha_i, \alpha_i^*, \eta_i, \eta_i^* \ge 0$.

The dual problem is given by

$$\begin{aligned} \max Q(\alpha - \alpha^*) &= \frac{1}{2}\sum_{i,j=1}^{l}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)(\Phi(x_i), \Phi(x_j)) \\ &- \varepsilon\sum_{i=1}^{l}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{l}y_i(\alpha_i - \alpha_i^*) \end{aligned} \tag{8}$$

$$subject\ to \begin{cases} \sum_{i=1}^{l}(\alpha_i - \alpha_i^*) = 0 \\ \alpha_i, \alpha_i^* \in [0, C] \end{cases} \tag{9}$$

After introducing the kernel function $K(x,x')$, the parameter w and the function to be estimated are obtained as Eq. (10).

$$\begin{cases} w = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)\Phi(x_i) \\ f(x) = \sum_{i=1}^{l}(\alpha_i - \alpha_i^*)K(x_i, x) + b \end{cases} \tag{10}$$

In Eq. (8), Lagrange multipliers α_i and α_i^* satisfy the equality

$(\alpha_i - \alpha_i^*) \ne 0$, $\alpha_i \times \alpha_i^* = 0, \alpha_i \ge 0, \alpha_i^* \ge 0$, $i = 1,2,\cdots,l$.

Those vectors with $\alpha_i \ne 0$ are called support vectors, which contribute to the final solution.

The choice of kernel function is essential and it directly affects the realization results of the algorithm. By comparison, the radial basis kernel function as Eq. (11) is the best choice in this article.

$$K(x,x') = \exp\left[-\frac{\|x - x'\|^2}{\sigma^2}\right] \tag{11}$$

IV. Wavelet-SVM Prediction

In the paper, a set of daily average wind speed data of a wind farm [9], comprising 285 days in 2004, are adopted to validate the method above. Fig.1 is the curve of original wind speed sequence. The inserted dimension of samples is set to be 5 by experiments, which means that previous 5 data will be used to predict the next 6th data. Hence, the 200 wind speed samples constructed from 1 to 205 are used to train, the 80 wind speed samples constructed from 201 to 285 are used to predict. The simulation experiments are performed by MATLAB7.4 in terms of the following steps:

1).To decompose original wind speed sequence with wavelet Here, two-scale wavelet decomposition of wind speed sequence is carried out by applying bi-orthogonal wavelet db5, and three coefficients a2, d1, d2 are obtained.

2).To construct and process samples

The 200 training samples and 80 test samples are extracted separately from the coarse component and various detail components. After that, all samples should be normalized.

3).To establish the objective functions list as Eq.7.

4).To use grid search method to select the best parameters ε, C and σ^2.

The selection of ε, C and σ^2 has a great influence on SVM regression estimation accuracy . The parameters ε, C and σ^2 are determined by grid search in which cross calculation method is adopted. The optimal selection of those parameters will effectively avoid over-fitting phenomenon. The optimization results are showed in Table 1.

5).To train network and predict wind speed.

The SVM models are established for three wavelet coefficients. A rolling-based forecasting procedure and a one-step-ahead forecasting policy are applied in the example. So the network will reflect the latest change of

Authorized licensed use limited to: Aarhus University. Downloaded on March 28,2022 at 20:04:48 UTC from IEEE Xplore. Restrictions apply.

wind velocity. Three error functions such as mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE) are adopted to evaluate the forecast accuracy.

$$E_{MAE} = \frac{1}{N}\sum\left|W_R - W_F\right|; \tag{12}$$

$$E_{MAPE} = \frac{1}{N}\sum\frac{\left|W_R - W_F\right|}{W_R}; \tag{13}$$

$$E_{RMSE} = \sqrt{\frac{1}{N}\sum(W_R - W_F)^2}\ . \tag{14}$$

Where N is the number of samples, W_R is the measured value, W_F is predict value.

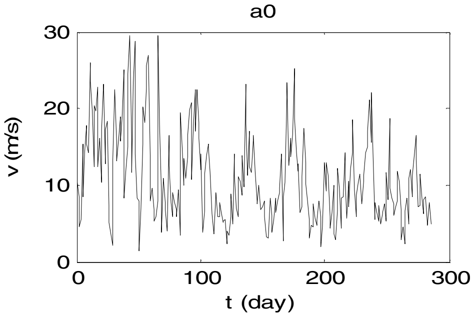


Fig.1.Original wind speed sequence

Table.1 Parameter optimization results

data	C	σ^2	ε	MSE
a2	97	0.0625	0.0032	0.0054
d2	41	0.0078	0.0057	0.0179
d1	12	0.125	0.0145	0.0152

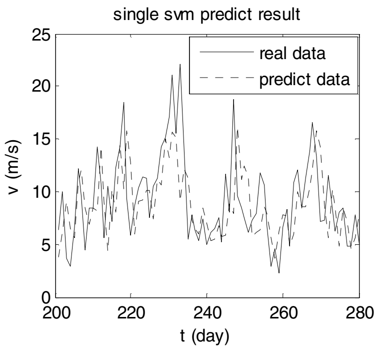
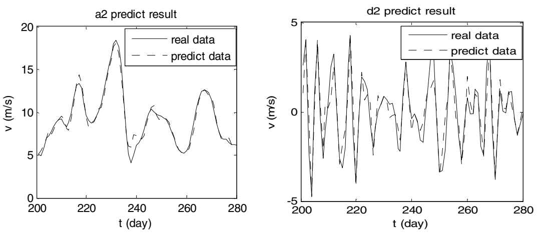
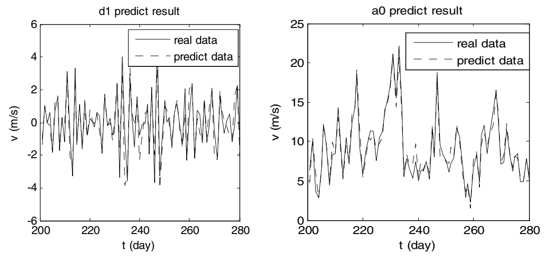


Fig.2. SVM predict result



(a) (b)



(c) (d)

Fig.3. Wavelet-SVM predict results

Table.2 Comparison of prediction methods

Predict method	Predict object	MAE (m/s)	RMSE (m/s)	MAPE (%)
Wavelet-SVM	a2	0.34	0.42	4.28
	d1	0.44	0.65	74.85
	d2	0.48	0.64	62.62
	a0	0.57	0.72	8.21
SVM	a0	1.23	1.43	15.72

Fig.3 (a), (b) and (c) are the predicted curves of a2, d2 and d1 respectively, and we obtain the predicted curve of original wind speed sequence a0 as Fig.3 (d) by wavelet reconstruction. It can be seen that the curve a2 represents wind trend, it seems relatively smooth and periodic. Therefore it is easy to be predicted accurately and its MAPE reaches 4.28%. The values of detail coefficients d1, d2 are not only relatively small but also comprise massive random elements and interference elements. So their forecasting accurate is lower, and their MAPE reach 74.85% and 62.62% respectively. However, the value of d1 and d2 are so small that they make little effect on the overall prediction accuracy. Clearly, the prediction accuracy of a0 mostly depends on that of coarse component a2. Therefore, when SVM prediction models are constructed, the structure and parameters of SVM model for coarse component a2 should be paid more attention to.

Fig.2 is the SVM prediction result without wavelet decomposition. Table 2 provides comparison results between SVM and Wavelet-SVM. It is indicated that, by applying wavelet transform, the spectrums of linear components and random components of high frequency in wind speed information will show a clearly separate feature which will be helpful to forecast. Simulation experiments prove that Wavelet-SVM has higher prediction precision and better generalization ability than SVM.

V. CONCLUSIONS

In this paper, a Wavelet-SVM model is proposed as a viable alternative to forecast wind speed. A comparative study of two models is evaluated. Results from two studied cases indicate that Wavelet-SVM model provides more accurate forecasting results than the SVM model. The superior forecasting ability of the proposed model is due to the following two reasons. First, the use of wavelet multi-resolution in data preprocessing increases the forecasting performance of the proposed mode.

Second, minimizing the structural risk improves the

Authorized licensed use limited to: Aarhus University. Downloaded on March 28,2022 at 20:04:48 UTC from IEEE Xplore. Restrictions apply.

generalization ability of the Wavelet-SVM models. In practical applications, however, the range of predictive values and their confidence interval are more meaningful than the predictive values themselves.

REFERENCES

[1] Alexiadis M. Dokopoulos P, Sahsamanoglou H *et al*. Short term forecasting of wind speed and related electrical power [J]. *Solar Energy*, 1998, 63(1):61-68.

[2] Bossanyi E A. Short-term wind prediction using Kalman filters [J]. *Wind Engineering*, 1985, 9(1): 1-8.

[3] Liu Bin，Su Hongye, Chu Jian. Predictive control algorithm based on least squares support vector machines[J]. *Control and Decision2004*, 19(12): 1399-1402(in Chinese).

[4] Kariniotakis G, Stavrakakis G, Nogaret E. Wind power forecasting using advanced neural network models [J]. *IEEE Trans Energy Conversion*, 1996, 11(4): 762-767.

[5] Yang Xiuyuan, Xiao Yang, Chen Shuyong. Wind speed and generated power forecasting in wind farm [J]. *Proceedings of the CSEE*, 2005, 25 (11): 1-5(in Chinese).

[6] Peng Yuhua. Wavelet Transform and Engineering Application [M].Beijing: *Science Press*, 1999.

[7] Deng Naiyang, Tian Yingjie. New method for data mining-support vector machine [M]. Beijing: *Science Press*, 2004.

[8] Zhao Dengfu, Wang Meng, Zhang Jiangshe, Wang Xifan. A Support vector machine approach for short term load forecasting [J]. *Proceedings of the CSEE,* 2002, 4 (4): 26-30(in Chinese).

[9] http://www.knmi.nl/samenw/hydra.

Authorized licensed use limited to: Aarhus University. Downloaded on March 28,2022 at 20:04:48 UTC from IEEE Xplore. Restrictions apply.