2009 International Conference on Energy and Environment Technology

One-Month Ahead Prediction of Wind Speed and Output Power Based on EMD and LSSVM

Wang Xiaolan
School of Electrical Engineering
Lanzhou University of Technology, Gansu, Lanzhou, 730050, PR China
wangzt@lut.cn

Li Hui
School of Electrical Engineering
Lanzhou University of Technology, Gansu, Lanzhou, 730050, PR China
toobage@163.com

Abstract —**Wind speed is a kind of non-stationary time series, it is difficult to construct the model for accurate forecast. The way improving accuracy of the model for predicting wind speed up to one-month ahead has been investigated using measured data recorded by wind farm. A forecasting method based on empirical mode decomposition (EMD) and least square support vector machine (LSSVM) is proposed in this paper. The non-stationary time series is decomposed into several intrinsic mode functions (IMF) and the trend term. The different LSSVM models to forecast each IMF are built up. These forecasting results of each IMF are combined to obtain the final forecasting result. Considering the power characteristics, unit efficiency and the operate condition of the generators, the one-month ahead forecasted output power of the wind power plant can be obtained.**

Keywords-wind speed forecasting; empirical mode decomposition(EMD); least square support vector machine (LSSVM); intrinsic mode function(IFM); wind power

I. INTRODUCTION

Wind energy is considered one of the most rapidly growing energy resources all over the world. It is expected that about 12% of the total world electricity demands to be supplied from wind energy resources by 2020[1]. Recently, the stand-alone capacity of the generating units and the total capacity of power generation about the large-scale wind farm have been improved highly. As the development of technology, the cost of the new power has already approached the conventional energy. The main method of exploitation and utilization the wind energy is the large-scale grid power generation at present. Wind speed and power prediction is an essential issues for: wind farms maintenance, optimal power flow between power network and wind farms, electricity marketing bidding, power system scheduling, and energy reserves and storages planning and scheduling [2-3].

At present, a number of different approaches have been applied to forecast wind speed and the power produced by wind farms. These methods can be divided into two categories. The first is the physical method, which use a lot of physical considerations to reach the best prediction precision and the second is the statistical method. Physical method has advantages in long-term prediction while statistical method does well in short-term prediction. Physical method[4], it uses physical considerations to predict the future speed and direction of wind, so the input variables will be the physical or meteorology information, such as description of orography, roughness, obstacles, meteo and so on; Statistical method [5], it has a high precision about short-term wind speed forecast. But, the precision of prediction for long term will be decreased obviously. Spatial correlation models[6], it takes the spatial relationship of different sites' wind speed into account. The wind speed time-series of the predicted points and its neighboring sites are employed to predict the wind speed. This method is very difficult in using because the measurement of many spatial correlated sites' wind speed values and timely transmission are all needed; Kalman filter [3], this method needs to know the statistical features of the noise, in fact, estimate the statistical features of the noise is a difficulty in the method; Time series method [7], one of the advantages of this method is that the less information are needed in constructing model, and it is very convenient to computing; Artificial neural network(ANN) [8] is one of the most widely used models in the last decade. But it is difficult to determine the reasonable network frame. It also has some intrinsic defects such as, slow study speed, local the minimum point in some areas and so on; Fuzzy logic method [9], the ability of forecast about fuzzy logic method is so weak and the fuzzy theory is still imperfect. The forecast ability of wind speed above mentioned methods are not very excellent and the mean absolute percentage error (MAPE) usually located from 25% to 40%.

Wind speed can be considered as a non-stationary time series, because of it has uncertain daily and seasonal cycle, it is difficult to construct the model for accurate forecast by traditional methods which based on the stationary signal. Considering the intrinsic characteristics of wind speed, a forecasting method based on empirical mode decomposition (EMD) and least square support vector machine (LSSVM) is proposed in this paper. At first, the wind speed signal was decomposed into several intrinsic mode functions (IMF) and the trend term. Secondly, use the LSSVM method to forecast each IMF and trend term, simultaneously, the different forecast models are built up. The last, forecasting results of each IMF and trend term are combined to obtain the final forecasting result. Considering the power characteristics, unit efficiency, and the operate condition of the device, the output power of the wind power plant can be obtained.

II. THEORY AND METHOD

A. Empirical Mode Decomposition

Empirical Mode Decomposition (EMD), proposed by Huang *et al*. [10], is a new method for decomposing multicomponent signals. It has been found many immediate applications in a variety of problems covering geophysical and biomedical engineering. And, the method has been improved by Huang *et al* at 1999. EMD utilizes empirical knowledge of oscillations intrinsic to a

978-0-7695-3819-8/09 $26.00 © 2009 IEEE
DOI 10.1109/ICEET.2009.571

time series in order to represent them as a superposition of components with well defined instantaneous frequencies. These components are called intrinsic mode functions (IMF). Huang *et al* have defined IMFs as a class of functions that satisfy two conditions: (1) In the whole data set, the number of extrema and the number of zero-crossings must be either equal or differ at most by one; (2) At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

The detail decomposition process is given by reference [11].

B. Least Squares Support Vector Machines

Support vector machines (SVM) for classification and nonlinear function estimation, as introduced by Vapnik [12] and further investigated by many others, is an important methodology in the area of neural networks and nonlinear modeling. A modified version of SVM classifiers, Least Squares SVM (LSSVM) classifiers [13], was proposed in Suykens and Vandewalle (1999). A two-norm was taken with equality instead of inequality constraints so as to obtain a linear set of equations instead of a quadratic programming problem in the dual space. The formulation of LSSVM is introduced by reference [13].

III. MODELING PROCESS BASED ON EMD AND LSSVM

Because of the wind speed is a kind of non- stationary time series, in order to improve the forecast precision, the forecast method based on EMD and LSSVM is proposed in this paper. To begin with, the time series were decomposed into several IMFs and the trend term. What is more, build up different models to forecasting each IMF and trend term based on LSSVM. The last, final forecasting result can be obtained by means of combine the forecast results of each IMF and trend term. Fig.(1) shows the detail process of the modeling.

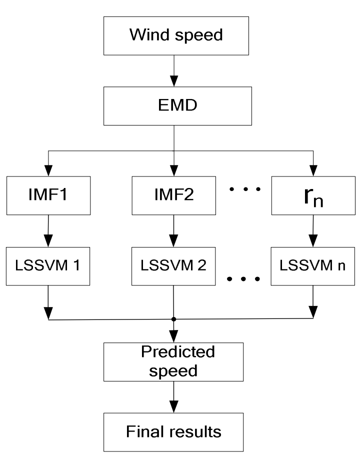

Fig.1 Prediction modeling process

IV. APPLICATION INSTANCE AND ANALYSIS

A. Date Structures

Hexi Corridor in northwest of China, stores a lot of wind energy. The Development and Reform Commission in China has been started to construct a typical base which would be produce 10 million kilowatts wind energy per year in Jiuquan. Until 2010, the installed capacity of the wind power base will be reach to 5 million kilowatts, and until 2015 the installed capacity will reach or beyond to 12 million kilowatts. So, it is very significant to accurate forecast the wind speed and power of Jiuquan wind farm. In this paper, the data used in the research were collected at a large wind farm in Jiuquan. The date contains 120 groups and each group stands for the daily average wind speed (from January 1th, 2006 to April 30th, 2006) which are used as the training samples; and the other 31 groups daily average wind speed (from May 1th to 31th, 2006) consider as the forecast samples. The forecast model will be established under the training samples to predicting the test samples. The forecasting results can demonstrate the performance of model.

B. Decomposed the Time Series Based on EMD

According to the algorithm of the EMD, MATLAB7.0 software is used as a simulate tool, so the signal can be decomposed into several IMFs and trend term. Fig. (2) shows the decomposed results. In Fig.(2), IMF1 is the highest frequency term and the IMF6 is the lowest frequency term, $r6$ is the trend term. If $X(t)$ represents the original signal, c_i represents the IMFi, r_n represents the trend term, the results decomposed in Fig. (2) can be express as follows:

$$X(t) = \sum_{i=1}^{n} c_i + r_n \qquad (1)$$

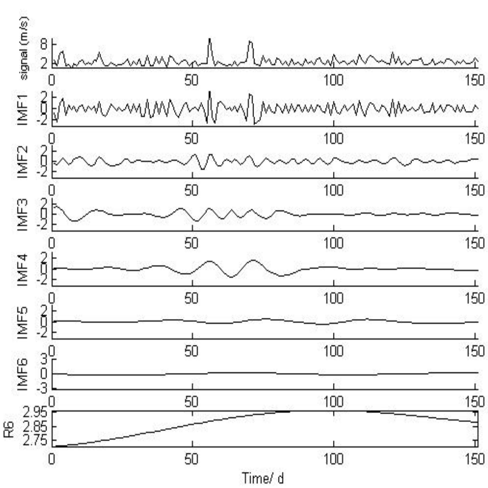

Fig.2 Original signal and its EMD decomposition

C. LSSVM Model Prediction

In accordance with the theory of LSSVM, model should be built under the training samples. Utilize the 120 groups daily average wind speed (from January 1th, 2006 to April 30th, 2006) as the training set. And, the other 31 groups daily average wind speed (from May 1th to 31th, 2006) consider as the test set. In this paper, after grid search and cross validate, RBF kernel function was selected $k(x_i, x_j) = \exp(-\|x_i - x_j\|^2/\sigma^2)$, $\sigma^2 = 0.002$, $\gamma = 103$. MATLAB7.0 software is used as a simulate tool to complete the simulation experiment. Fig. (3) shows the forecasted result of the LSSVM model.

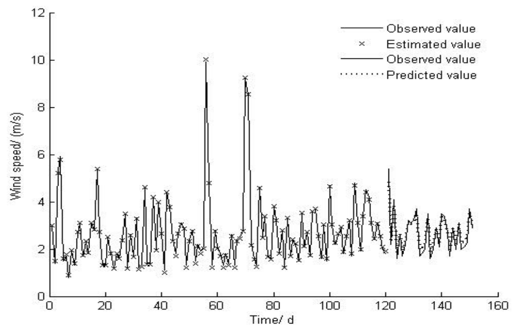

Fig.3 Forecasting of wind speed with LS-SVM

D. EMD-LSSVM Hybrid Model Prediction

According to the different characteristics of IMFs and trend term, different LSSVM kernel functions should be selected. It is so important to choose the suitable kernel function because it can strongly influenced the performance of LSSVM. RBF kernel function is the best choice for forecasting the IMFi. In order to make the computation process convenient, the same kernel parameters ($\sigma^2 = 0.002, \gamma = 103$) would be selected. When forecast the trend term ($r6$), polynomial kernel function would be selected and the parameter d =3.

These forecasted results of each IMF and trend term are combined to obtain the final forecasted results. In order to demonstrate the effectiveness of the model, another hybrid model using EMD and RLS has been built to forecast the short-term wind speed. Using the measured date from large wind farm, the date contains 61 groups and each group stands for the daily average wind speed (from March 1th, 2006 to April 30th, 2006), as the training set. And, the other 31 groups daily average wind speed (from May 1th to 31th, 2006) considering as the test set will be predicted. Fig.(4) shows the forecasted results of the three models (LSSVM model、EMD-LSSVM model and EMD-RLS Model).

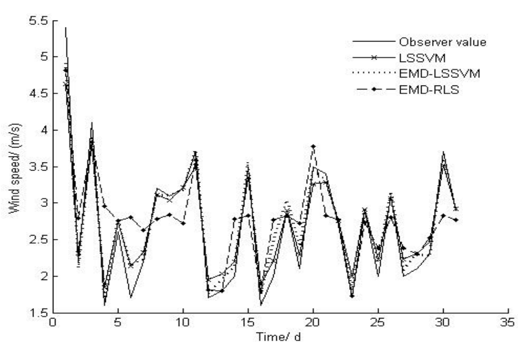

Fig.4 The final forecasting based on three models

E. Error Analysis

It is very significant to select the reasonable error analyze methods to judge the results. In this paper, the mean absolute percentage error (MAPE) and root-mean-squares of errors (RMSE) had been selected as the evaluation indicators.

$$MAPE = \frac{1}{M}\left[\sum_{i=1}^{M}\left|\frac{r_i - f_i}{r_i}\right| \times 100\%\right] \quad (2)$$

where r_i represents the original wind speed time series, f_i represents the forecasted wind speed time series, M represents the number of the time series points.

$$RMSE = \sqrt{\frac{1}{M}\sum_{k=1}^{M}[y(k) - \hat{y}(k)]^2} \quad (3)$$

where $y(k)$ represents the original wind speed time series, $\hat{y}(k)$ represents the forecasted wind speed time series, M represents the number of the time series points. Fig. (5) shows the relative errors of the three models.

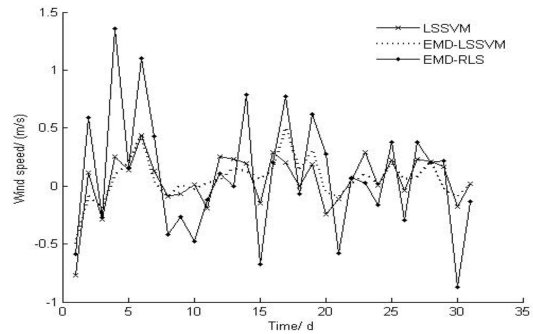

Fig.5 Forecasting errors based on thee models

The comparison of the MAPE and RMSE for the proposed three models (LSSVM model、EMD-LSSVM model and EMD-RLS model) are recorded in Table I.

TABLE I.
THE COMPARISON OF FORECASTING ERRORS

Algorithm	*MAPE*(%)	*RMSE*
LSSVM	7.923	0.236
EMD-LSSVM	5.731	0.188
EMD-RLS	14.12	0.418

The scope of the MAPE for three models changes from 5.731 to 14.12 (%) and the RMSE changes from 0.188 to 0.418. According to the forecasted errors, it is clearly see that EMD-LSSVM model is the optimal model because of the lowest errors, not only the MAPE (5.731%) but also the RMSE (0.188); secondly, the LSSVM model; the EMD-RLS model is not the ideal model in this paper, because of the poor forecasted errors. The simulation experiment demonstrates EMD-LSSVM model has the ability in improving the predict precision towards the non-stationary time series.

F. Output Power Prediction

Many factors can influence the output power of wind farm, the main factors, such as wind speed, wind direction, air pressure, temperature and humidity and so on. These factors can be described by data series, then, the data series and the forecasting wind speed can compose the basic input data space of the artificial neural network (ANN). Thus, the output of the ANN will be the output power of the every unit. Considering the unit efficiency, operate condition of the device and the wind farm losses, the forecast value of the output power of the wind farm can be obtained. Fig. (6) shows the detail structure of the output power forecasting system.

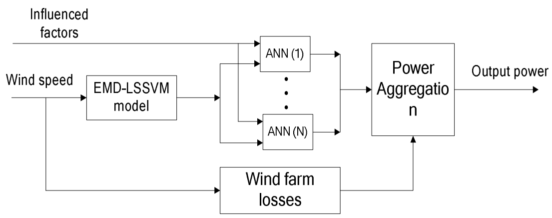

Fig.6 Structure of the power forecasting system.

In order to enhance the learning ability and the forecast precision of the ANN, the data should be normalized. Such as wind speed normalization, wind direction normalization, air pressure, temperature and humidity normalization, and so on. Fig. (7) shows the basic structure of ANN in Fig. (6).

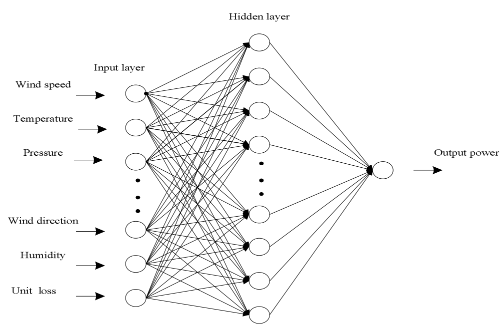

Fig.7 Structure of the ANN

In this paper, the simulation system based on the wind turbine with rated power 600*KW* has been structured. The simulation parameters: wind turbine radius 22 *m*; normal rated wind speed 17*m/s;* cut-in wind speed 3*m/s*; cut-out wind speed 25*m/s*; rated speed 30*r/min*; rated voltage 690*V*; frequency 50/60*Hz*; swept area 1452 m^2 ; air density 1.293 kg/m^3 ; wind energy utilization coefficient 0.44.

According to the predicted wind speed (May 1th to 31th, 2006), using the output power forecasting system, the daily output power of the wind farm can be obtained. TABLE II shows the final forecasting results.

TABLE II
THE COMPARISON OF FORECASTING RESULTS

Time (*d*)	Wind speed (*m/s*)	Power (*kw*)	Time (*d*)	Wind speed (*m/s*)	Power (*kw*)
1	4.907	237.641	17	2.498	31.367
2	2.121	18.917	18	3.033	56.127
3	3.909	120.153	19	2.410	28.165
4	1.694	9.777	20	3.445	82.242
5	2.798	44.053	21	3.308	72.784
6	2.139	19.688	22	2.718	40.372
7	2.230	22.303	23	1.809	11.928
8	3.102	60.062	24	2.912	49.695
9	3.107	60.353	25	2.196	21.298
10	3.174	64.317	26	3.154	63.121
11	3.725	104.002	27	2.081	18.138
12	1.755	10.874	28	2.312	24.855
13	1.954	15.007	29	2.265	23.373
14	2.126	19.329	30	3.607	94.358
15	3.560	90.725	31	2.909	49.503
16	1.742	10.639			

The forecasting results demonstrate the forecasting system proposed in this paper has the advanced capacity to predicting the output power of the wind farm one-month ahead. The RMSE, both observed value and predicted value of the output power, is about 10%.

V. CONCLUSION

In this paper, The LSSVM model, EMD-LSSVM model and EMD-RLS models for forecasting wind power were built. The one-month ahead wind speed of the wind farm have been predicted by each of the models. The error analysis shows that EMD-LSSVM model is the optimal model for wind speed forecasting, due to the lowest MAPE (5.731%) and the RMSE (0.188). The simulation experiment shows that the hybrid model (EMD-LSSVM) is an effective method towards the non-stationary time series forecasting. Utilizing the predicted value of the wind speed and considering the influencing factors, such as the local humidity, temperature, air pressure and the losses of the wind turbines, the output power forecasting can be obtained. Simulation results show that the RMSE of the one-month ahead forecasting of output power is about 10%. Thus, it is an effective scheme to the output power forecasting of the wind farm.

REFERENCES

[1] European Wind Energy Association. (2002).Wind force 12, [Online]. Available: http://www.ewea.org/doc/ WindForce12.pdf.

[2] Alexiadis MC, Dokopoulos PS, Sahsamanoglou HS, Manousaridis IM. Short-term forecasting of wind speed and related electrical power. Solar Energy 1998;63(1):61-8.

[3] Bossanyi E A. Short-term Wind Prediction Using Kalman Filters[J]. Wind Engineering, 1985, 9(1): 1-8.

[4] Landberg L. Short-term prediction of the power production from wind farms. Journal of Wind Engineering and Industrial Aerodynamics 1999;80:207-20.

[5] Notis C, Trettel D, Aquino J, Piazza T, Taylor L, Trask D. Learning to forecast wind at remote sites for wind energy applications. PNL-4318, Pacific Northwest Laboratory, 1983.

[6] Alexiadis M C,Dokopoulos P S,Sahsamanoglou H S. Wind speed and power forecasting based on spatial correlation models[J].IEEE Trans on Energy Conversion,1999,14(3):836-842.

[7] Guoyang W, Yang X, Shasha W. Discussion about short-term forecast of wind speed on wind farm. Jilin Electric Power 2005;181(5):21-4.

[8] Wang X, Sideratos G, Hatziargyriou N, Tsoukalas LH. Wind speed forecasting for power system operational planning. In: Proceedings of the 8th International Conference on Probablitistic Methods Applied to Power System; 2004.

[9] Damousis IG, Alexiadis MC, Theocharis JB, Dokopoulos PS. A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation.Energy Conversion 2004 ;19(2):352-61.

[10] Huang N E,Shen Z,Long S R.The Empirical Mode Decomposition and The Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis[J].Proc R Soc London A, 1998, 454: 903-99.

[11] G. Rilling, P. Flandrin, and P. Goncalves. On empirical mode decomposition and its algorithms, in Proc. of IEEE-EUSIPCO workshop, NSIP-03, June 2003.

[12] Vapnik V N. The Nature of Statiscal Learning Theory[M]. New York:Springer Verlag,l999.

[13] Suykens J A K,Vandewalle J. Least Squares Support Vector Machine Classifiers[J].Neural Processing Letters, 1999,9(3):293-300.