Short-term Wind Power Prediction with Signal Decomposition

Wang Lijie, Dong Lei, Gao Shuang, Liao Xiaozhong
School of Automation
Beijing Institute of Technology
Beijing, China

***Abstract*—Wind power is widely used to replace conventional power plant and reduce carbon emission. However, the variability and intermittency of wind makes the wind power output uncertain, which will bring great challenges to the electricity dispatch and the system reliability. So it is very important to predict the wind power generation. Two different signal decomposition methods are introduced into the prediction of wind power generation in this paper. One is wavelet transform (WT), and another is empirical mode decomposition (EMD). Both of them are good at decreasing the non-stationary behavior of the signal. ANN with the capacity of nonlinear mapping is used to model the decomposed time series. The prediction models WT-ANN and EMD-ANN are compared each other and a combined model based on them is tested. The wind power data from the Saihanba wind farm of China is used for this study.**

Keywords- wind power prediction, wavelet transform, empirical mode decomposition, combined model

I. INTRODUCTION

Wind power generation has rapidly developed over the last decade and global installed wind power continues to grow at around 28% per year. As the penetration of wind energy into grids has increased significantly, the intermittence and variability of the wind power output may bring several problems on the safety and reliability of power systems [1]. Reserve must be increased to balance generation and demand, and thus operational costs of the power plants will be increased correspondingly [2]. So it is very important to forecast the output of wind power to avoid the balancing problem.

According to different time scales, wind power forecasting can be divided into very short term, short term and medium term [3]. Very short-term forecast is in the range of one to a few hours and is used for regulating capacity at operator's real time control. Short-term forecast is in the range of ten hours to two days ahead and is interesting for the scheduling of conventional units at the distribution operators and trading in the day-ahead market at wind farm owners. Medium term is related to several days to one week and helps make maintenance planning of wind turbines.

Wind is persistent in nature, so methods that do not use numerical weather prediction (NWP) may get good forecast in very short-term wind power forecasting. The method consists of statistical models based on time series approach, such as ARMA, Kalman Filters [4], and intelligent models, such as neural networks [5] and support vector machines [6]. These models only take as inputs past values from the forecasted variables (wind speed or wind power) or/and explanatory variables (wind direction, temperature) [7]. The most common method is the persistence model which assumes that the prediction value of the wind power is the same to the latest measured one [8]. It is usually used as a benchmark to evaluate the performance of other advanced methods. Persistence model may be good enough for 1 to 3h ahead and statistical models will outperform the persistence model for forecast horizons between 3 and 6h. For time horizons greater than 3-6h, NWP should be included as inputs.

Short-term wind power forecasting requires meteorological predictions as inputs to forecast for horizons ranging from 6 to 72 hours. The methods are mainly divided into two groups. The first group is called physical approach, and it includes *downscaling* that transforms the forecasted wind information into wind at the turbine hub height by using physical considerations about terrain, *conversion* of wind power through manufacturer's power curve or estimated power curve, and *upscaling* from representative wind farms if a regional forecast is required [9]. The second group is called statistical approach, and it consists of finding the relationship between NWP, historical measurements, and generation output, usually employing recursive techniques. The statistical approaches approximate a "black-box", which includes most of the intelligent models, such as ANN, SVM, fuzzy logic, mixture of experts, nearest neighbour search [10].

Wavelet transform (WT), which can decompose the original signal into several time series that have simpler frequency components, has been successfully applied in the fields like data analysis and signal processing. It is also proposed for time series prediction combined with other models like neural networks [11]. Empirical mode decomposition (EMD) is another powerful multi -resolution signal decomposition technique for analyzing nonlinear and non-stationary signals. It decomposes a time series into components by empirically identifying the physical time scales intrinsic to the data. Each extracted mode, named intrinsic mode function (IMF), contains some basic properties [12].

In order to overcome the deficiency of traditional models in dealing with non-stationary signal, a neural network model with multi-decomposed-layer is presented in this paper. WT and EMD are respectively used to decompose the wind power signal and a series of stationary time series is got. Different corresponding neural networks are applied to model every decomposed time series. The forecast errors of these two models WT-ANN and EMD-ANN are smaller than the

978-1-4244-8039-5/11/$26.00 ©2011 IEEE

traditional persistence model. Simulation results also indicate that the EMD-ANN model is better than the WT-ANN model. In order to further improve the prediction precision, a combined model based on the WT-ANN and EMD-ANN is discussed.

II. Signal Decomposition Theory

A. Wavelet Transform

Wavelet Transform is a new tool for time-frequency analysis. Multi-resolution approximation by wavelet basis functions is a technique for representing a function on many different scales, which are formed by scaled and translated mother wavelet [13]. The Continuous Wavelet Transform (CWT) of a signal $x(t)$ is defined as

$$WT_x(a,b)=\frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty}x(t)\psi^*(\frac{t-b}{a})dt=\langle x(t),\psi_{a,b}(t)\rangle \tag{1}$$

Where $\psi(t)$ is the mother wavelet, and other wavelets

$$\psi_{a,b}(t)=\frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}) \tag{2}$$

are its dilated and translated versions, where a and b are the dilation parameter and translation parameter respectively. The discrete WT (DWT), instead of CWT, is used in practice. Calculations are made for chosen subset of scales and positions. This scheme is conducted by using filters and computing so called approximations and details. The approximations are the high-scale, low frequency components of the signal. The details are the low-scale, high-frequency components. The DWT coefficients are computed using the equation

$$WT_x(j,k)=\int_{-\infty}^{+\infty}x(t)\psi_{j,k}(t)dt \tag{3}$$

Where $a=a_0^j$, $b=ka_0^jb_0$, $j,k\in Z$.

The decomposition (filtering) process can be iterated, so that one signal is broken down into many lower resolution components. This is called the Mallat wavelet decomposition tree shown in Fig. 1.

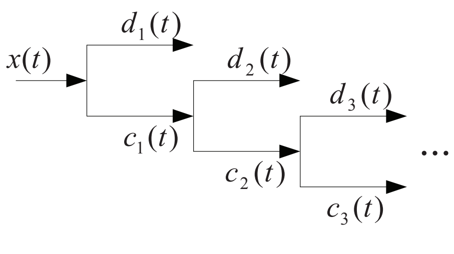

Figure 1. Mallat wavelet decomposition tree

B. Empirical Mode Decomposition

Empirical mode decomposition (EMD) is a method of decomposing a signal in the time domain. The decomposition is based on the direct extraction of the signal energy associated with various intrinsic modes in different time scales. The principle of the method is to adaptively decompose a given signal into a collection of intrinsic mode functions (IMFs) by a shifting procedure. The decomposition is useful for analyzing non-linear and non-stationary signals.

An IMF is a function which satisfies the following two conditions [14]:

(1) The number of zero point and the number of extreme in the whole sampled data must be equal or differ by at most 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, which means the signal is locally symmetric about the time axis.

For any given signal $x(t)$, the process of EMD is as follows:

Step 1: After the extreme of the signal are identified, all the local maxima are connected by a cubic spline line as the upper envelope $u_1(t)$ and all the local minima are connected as the lower envelope $v_1(t)$.

Step 2: The mean of the envelopes is calculated by $m_1(t)=(u_1(t)+v_1(t))/2$. The difference between the original data $x(t)$ and $m_1(t)$ is the first component $h_1(t)$, i.e. $h_1(t)=x(t)-m_1(t)$.

Step 3: If $h_1(t)$ satisfies all the requirement of IMF, the first IMF component of the signal can be designated as $IMF_1(t)=h_1(t)$; Otherwise, $h_1(t)$ is treated as the original data and Step 1 ~ 3 are repeated many times until an IMF is obtained.

Step 4: The first residue can be separated from the first IMF by $r_1(t)=x(t)-IMF_1(t)$

Step 5: The decomposing process can stop when the residue $r_1(t)$ becomes so small or a monotonic function. Otherwise, the residue $r_1(t)$ is treated as the new data and Step 1 ~ 5 are repeated many times.

In the end, n IMF components and one residue $r_n(t)$ are obtained and the original signal $x(t)$ can be represented as the following equation

$$x(t)=\sum_{i=1}^{n}IMF_i(t)+r_n(t) \tag{4}$$

III. Wind Power Prediction Model Based on Signal Decomposition and ANN

A. Model Structure and Algorithm

Wavelet transform decomposes the signal into different frequency channels. The frequency of the decomposed series is more unitary than the original signal, so the decomposed time series is much simpler than the original time series and treated as approximately stationary time series. EMD decomposes the signal into a collection of IMFs, which can act as the intrinsic oscillating mode to express the local properties of the data.

Even though some IMFs still have non-stationary behavior, the effect between them is been isolated, which will try to decrease the influence of non-stationary behavior on prediction. A series of stationary data are obtained after the signal decomposition process. These series are relatively easier than the original signal. Considering the complexity and high nonlinearity of the wind power system, BP-NN (Back Propagation) method is applied to model every time series. Therefore, a wind power prediction model combining signal decomposition with ANN is presented in Figure 2.

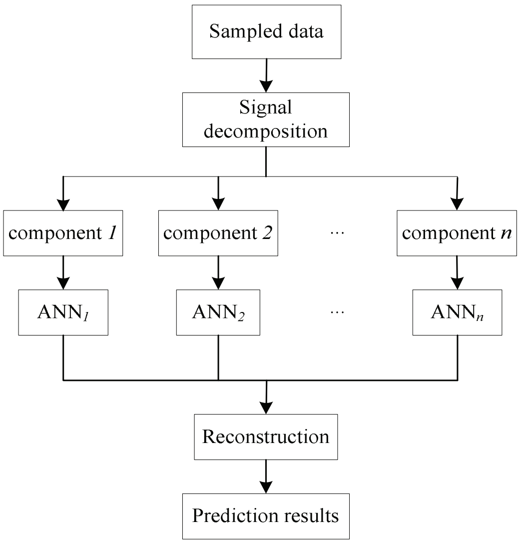

Figure 2. Wind power prediction model based on signal decomposition and ANN

The sampled data of wind power are decomposed into different components. These components are detail parts $d_i(t)(i=1,\cdots,j)$ and an approximate part $c_j(t)$ by wavelet transform, or intrinsic mode function $IMF_i(t)(i=1,\cdots,k)$ and a residue $r_k(t)$ by EMD. There is no correlation between *j* and *k*.

j is the decomposition scale and established before wavelet transformation. Once the decomposition scale j and the wavelet function are established, they are used for all sampled data. Every part is analyzed to identify the chaotic characteristic respectively. Then, different ANN models based on phase space reconstruction [15] can be constructed to match with every decomposed time series.

k is the number of IMF and determined by the process of adaptive EMD, so *k* is unfixed with the change of the sampled data. For example, nine IMFs are obtained when the sampled data is from 1 to 720, but ten IMFs are got when the sampled data is from 2 to 721. ANN is used to model every IMF. Then as the new data appears, the number of IMF after decomposition may vary, so rolling learning is applied to rebuild ANNs according to the new decomposition results. This method can continuously adjust the number of ANN and update the weights to make the model track the inherent change of the time series instantly.

The final prediction value is a combination of the prediction results from all ANN models. They are combined by choosing appropriate coefficients for minimizing prediction error. For example, the detail part $d_1(t)$ with high frequency represents stochastic component and can be neglected in reconstruction.

B. Experiment Results

The sampled wind power generation time series of the Saihanba wind farm is from 2005. 12. 1 to 2006. 1. 31. The time series of the first month are used for training the hybrid model, and then the model is used to predict the second month wind power generation for one hour ahead.

Wavelet transform and EMD are respectively applied to decompose the time series. In wavelet transform, db4 function is chosen as mother wavelet and the decomposition scale is four. The largest Lyapunov exponents of every decomposed time series are positive and all series can be predicted using ANN based on phase space reconstruction. In EMD, the time window is the data of one month. Every time, the nearest 720 data is used for adaptive decomposition and corresponding ANNs are built for every decomposed component. As the true data of next hour is got, the decomposition and modeling above are repeated to ensure the networks to learn the latest changes.

Wind power prediction obtained by the model using signal decomposition and ANN is presented in Fig. 3. On the whole, every prediction method has its advantage and disadvantage at different points and the prediction results are similar. For example, the EMD-ANN model is better than the WT-ANN model at the 8th point, but the WT-ANN is better than the EMD-ANN at the 10th point.

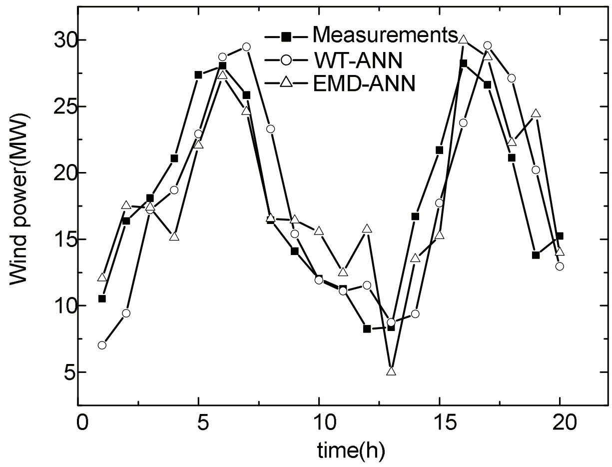

Figure 3. Prediction results of wind power generation using signal decomposition and ANN

As to investigate the overall performance of the models, the prediction errors of every model are shown in Table I. NMAE represents the normalized mean absolute error and NRMSE represents the normalized root mean square error. It is obvious that the errors of WT-ANN and EMD-ANN are smaller than the persistence model, which is owed to the signal decomposition process. The error of EMD-ANN model is smaller than the WT-ANN model. It is because that EMD is based on the local properties of the signal, but WT needs to choose the decomposition scale and wavelet function first, so the decomposition results of EMD is more precise than the WT.

TABLE I. NMAE AND NRMSE OF DIFFERENT PREDICTION MODELS

Models	NMAE	NRMSE
Persistence	12.55%	15.44%
WT-ANN	11.33%	13.76%
EMD-ANN	10.20%	13.58%

IV. Wind Power Prediction with Combined Model

Wind power prediction of large wind farm using only a single model brings the risk of high and costly errors. Particularly in the case of extreme events, individual models can go wrong. Considering the prediction results of WT-ANN and EMD-ANN, one is better at some points, but another is better at the other points. Therefore, a combined model for wind power prediction based on WT-ANN and EMD-ANN is proposed.

A. Combination Method

$x(t)$ $(t=1,2,...,n)$ is the measured wind power generation at time *t*. Assume there are p different prediction models. And the predicted values of respective model is $\hat{x}_1(t), \hat{x}_2(t)\cdots, \hat{x}_p(t)$. The linear combination model can be written as

$$\hat{x}(t) = l_1\hat{x}_1(t) + l_2\hat{x}_2(t) + \cdots + l_p\hat{x}_p(t) \tag{5}$$

Where the $L=(l_1, l_2, \cdots, l_p)^T$ are the weights of each model and

$$l_1 + l_2 + \cdots + l_p = 1 \tag{6}$$

The prediction error of the model i at time *t* is

$$e_{it} = x(t) - \hat{x}_i(t)\,,\ i=1,2,\cdots,p\,,\ t=1,2,\cdots,n \tag{7}$$

Then the error matrix shows as

$$E = [(e_{it})_{p\times n}][(e_{it})_{p\times n}]^T \tag{8}$$

The quadratic sum of error of the linear combination model is

$$J = \sum_{t=1}^{n}[x(t) - \hat{x}(t)]^2 = \sum_{t=1}^{n}(\sum_{i=1}^{p} l_i e_{it})^2 = L^T EL \tag{9}$$

The minimum of the quadratic sum of error is selected as objective function, then combination weights can be obtained by calculating the optimal function (10)

$$\min J = L^T EL \qquad s.t. R^T L = 1, R = (1,1,\cdots,1)^T \tag{10}$$

B. Combined Prediction of Wind Power

WT-ANN and EMD-ANN are chosen to predict wind power respectively. The optimal combination weights of the two prediction models are obtained by calculating the optimal function (10), $l_1 = 0.48624$, $l_2 = 0.51376$. The combination prediction results are shown in Fig. 4. It can be seen that this simple combination approach improves the prediction accuracy very significantly compared to the results of the single models. The NMAE and NRMSE of the combination model is 9.57% and 11.89%, which are both smaller than the single models shown in Table I. Therefore, the combined model can synthesize information and fuse prediction deviation in different single models, resulting in prediction accuracy improved.

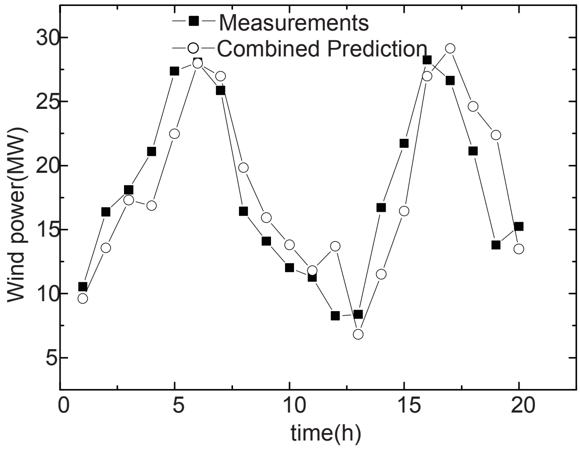

Figure 4. Prediction results of wind power generation using combined model

V. Conclusions

Two models for wind power prediction are presented based on signal decomposition and ANN. One model is WT-ANN, and another one is EMD-ANN. Both of the models are validated for the case of the Saihanba wind farm. The obtained results show that the two new models with signal decomposition have better prediction properties than the simple persistence model, and the EMD-ANN is more precise than the WT-ANN.

In order to fuse prediction deviation in different single models and reduce the prediction errors, a combined model for wind power prediction based on the WT-ANN and the EMD-ANN is proposed. The combination model shows much better performance compared with the single models.

Acknowledgment

Authors greatly acknowledge the financial support provided by the National Science Foundation of China under the AWARD Number 50777003 and The China Power Investment Corporation science & technology project (2009065ZDGKJX).

References

[1] Bart C.Ummels, Madeleine Gibescu, Engbert Pelgrum, Wil L.Kling, Arno J.Brand, "Impacts of wind power on thermal generation unit commitment and dispatch," IEEE transactions on energy conversion, vol.22, no.1, pp.44-51, 2007.

[2] Hannele Holttinen, "Handling of wind power forecast errors in the Nordic power market," The 9th International conference on probabilistic methods applied to power systems, Stockholm, Sweden, 2006.

[3] Michael Negnevitsky, Cameron W.Potter, “Innovative short-term wind generation prediction techniques,” Power engineering society general meeting, Montreal, 2006.

[4] Pan Di-Fu, Liu Hui, Li Yan-Fei, “A wind speed forecasting optimization model for wind farms based on time series analysis and kalman filter algorithm,” Power System Technology, vol.32, no.7, pp.82-86, 2008.

[5] L. Shuhui, D. C. Wunsch, E. A. O'Hair, and M. G. Giesselmann, “Using neural networks to estimate wind turbine power generation,” IEEE Trans. Energy Conversion, vol. 16, no.3, pp. 276-282, Sept. 2001.

[6] Du Ying, Lu Ji-Ping, Li Qing, Deng Ying-Ling, “Short-term wind speed forecasting of wind farm based on least square-support vector machine,”Power System Technology, vol.32, no 15, pp. 62-66, 2008.

[7] C. Monteiro, R. Bessa, V. Miranda, A. Botterud, J. Wang, and G. Conzelmann, “Wind power forecasting: state-of-the-art 2009,” Argonne national laboratory, US.

[8] Negnevitsky M, Johnson P, Santoso S, “Short Term Wind Power Forecasting Using Hybrid Intelligent Systems,” IEEE Power Engineering Society General Meeting, 2007.

[9] G. Giebel, L. Landberg, G. Kariniotakis, R. Brownsword, “State-of-the-art on methods and softwore tools for short-term prediction of wind energy production,” European wind energy conference & exhibition EWEC, Madrid, Spain, 2003.

[10] Ernst B, Oakleaf B, Ahlstrom M L, et al, “Predicting the Wind,” IEEE Power & Energy Magazine, vol.11, pp.79-89, 2007.

[11] Xieping Gao, Fen Xiao, Jun Zhang, and Chunhong Cao, “Short-term prediction of chaotic time series by wavelet networks,” The 5th World Congress on Intelligent Control and Automation, 2004, vol. 3, pp. 1931 - 1935.

[12] Z.Lu, J.S.Smith, Q.H.Wu, J.Fitch, “Empirical mode decomposition for power quality monitoring,” IEEE/PES Transmission and Distribution conference & exhibition: Asia and Pacific, Dalian, 2005.

[13] P. Cristea, R. Tuduce, and A. Cristea, “Time series prediction with wavelet neural networks,” The 5th Seminar on neural network applications in electrical engineering, Belgrade, Sept. 2000.

[14] Abdelkhalek Bouchikhi, Abdel-Ouahab Boudraa, Salem Benramdane, El-Hadji Diop, “Empirical mode decomposition and some operators to estimate instantaneous frequency: a Comparative study,” The 3rd International Symposium on Communication, Control and Signal Processing (ISCCSP), Malta, Mar.2008.

[15] Dong Lei, Wang Lijie, Hu Shi, Gao Shuang, and Liao Xiaozhong, “Prediction of Wind Power Generation based on Chaotic Phase Space Reconstruction Models,” The 7th international conference on power electronics and drive system (PEDS), Bangkok, Thailand, Nov. 2007.