Short-Term Wind Speed Forecasting Based on Non-stationary Time Series analysis and ARCH model

Peng Lv
Department of Mathematics and Physics
North China Electric Power University
Changping District, Beijing
Email: lpeng1013@126.com

Lili Yue
Department of Mathematics and Physics
North China Electric Power University
Changping District, Beijing
Email: yll856@126.com

ABSTRACT — **Wind speed time series is nonlinear and non-stationary, and has time-varying variance. Therefore, Wind speed is often considered as one of the most difficult meteorological parameters to forecast. The proposed model is based on non-stationary time series theory and ARCH model. First, wind speed series is decomposed and reconstructed into approximate series and detailed series by wavelet analysis. Then use ARIMA model to analyze each part, simultaneously considering the heteroscedasticity effect of the residual series, the corresponding ARIMA-ARCH model is set up. The final forecasting wind speed values are the sum of the predicted approximate and detailed values. This proposed method is applied to forecast the actual wind speed data and verification results show it can improve the accuracy of wind speed forecasting.**

KEY WORDS — **short-term wind speed forecasting; non-stationary time series; wavelet analysis; ARCH model**

I. INTRODUCTION

Wind power energy is clean, inexhaustible, and free resource. Because of the lack of traditional energy and environment pollution, the development of wind power has been paid more attention by more and more countries [1]. Simulations of power generally are derived from simulations of speed, so wind speed forecasting is of more significance [2].

Many approaches for wind speed forecasting have been proposed at home and abroad, such as: Persistent forecasting, Kalman filters[3], Time series models[4], Regression analysis[5], ANN[6], Spatial correlation[7], SVM[8]. However, these methods have some shortcomings. Persistent forecasting method's prediction error is relatively large. Kalman filters method assumes the statistical properties of the noise are known, but estimating the properties in fact is quite hard. For the ARIMA model, low-order model's prediction accuracy is low and high-order model's parameter estimation is difficult. The training rate of ANN algorithm is slow, and it is easy to fall into local optimal solution. Spatial correlation method forecasts wind speed based on the similarity and delay of wind speed between spatially correlative sites, but it is difficult to collect data in applications. SVM algorithm converts a real problem into a quadratic convex programming problem with inequality constraints, then it's hard to calculate.

By the above discussion, we can see that choosing a appropriate model to predict wind speed is more important. This paper proposes a mean hourly wind speed forecasting model based on non-stationary time series, which takes the heteroscedasticity effect into account. The forecasting process can be performed in three distinct steps:

Step 1 For its non-stationary nature, the wind speed series $\{X_t\}$, is decomposed and reconstructed into approximate signal and detailed signals by wavelet transform.

Step 2 Establish models for each signal respectively using time series analysis. For the time-varying variance of wind speed series, considering the heteroscedasticity effect, the corresponding ARIMA-ARCH model is set up.

Step 3 Sum the prediction results of each signal to obtain

978-1-61284-774-0/11/$26.00 ©2011 IEEE

Authorized licensed use limited to: PUC MG. Downloaded on June 23,2021 at 01:21:55 UTC from IEEE Xplore. Restrictions apply.

the final forecasting wind speed.

II. BASIC THEORIES

A. Time Series Analysis

Random time series models include: AR, MA, ARMA, ARIMA [9, 10]. For a stationary time series, the typical linear model is ARMA. This model takes past values, prediction errors and a random term into account. An ARMA(p ,q) process of order p and q, is represented as:

$$\varphi(B)X_t = \theta(B)\varepsilon_t \tag{1}$$

$$\varphi(B) = 1 - \varphi_1 B - \cdots - \varphi_P B^P \tag{2}$$

$$\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q \tag{3}$$

Where, B is the backward shift operator; φ_i ($i=1,2,...,p$) are autoregressive parameters; θ_j ($j = 1,2,\cdots,q$) are the moving average parameters; and $\{\varepsilon_t\}$ is a white noise process. If q=0, then equation (1) becomes an AR model of order p; When p=0, the process becomes an MA model of order q.

If the time series is not stationary, it can be differenced d times until stationarity is achieved. Then the model is ARIMA(p, d, q), it can be written as:

$$\varphi(B)(1-B)^d X_t = \theta(B)\varepsilon_t \tag{4}$$

B. Wavelet Analysis

In 1988, Mallat illustrated the characteristics of Multi-resolution analysis from the perspective of space. Then the decomposition and reconstruction algorithm when signal passed different frequency filters was given[11].

1) Decomposition algorithm

Suppose V is the original signal. Based on Mallat algorithm V can be decomposed into detail signals at different scales $d_1, d_2, \cdots, d_J$ and a coarse approximation signal c_J. J is the level number of wavelet decomposition.

$$\begin{cases} c_j = Hc_{j-1} \\ d_j = Gd_{j-1} \end{cases} \tag{5}$$

Where, H and G are the coefficients of low-pass filter and high-pass filter.

2) Reconstruction algorithm

But after the wavelet decomposition, the points' number of detail signals and coarse approximation signal are halved compared with original signal V. So these decomposed signals have to be reconstructed by equation (6).

$$\begin{cases} C_j = \left(H^*\right)^j c_j \\ D_j = \left(H^*\right)^{j-1} G^* d_j \end{cases} \quad j = 1,2,\cdots,J \tag{6}$$

Where, H^* and G^* are the dual operators of H and G. D_j($j = 1,2,\cdots,J$) and C_J are the reconstructed signals of d_j($j = 1,2,\cdots,J$) and c_J. Then sum C_J, D_1, $\cdots$, D_J to obtain the original signal V.

$$V=D_1+D_2+\ldots D_J+C_J \tag{7}$$

C. ARCH Model

Using ARIMA model to simulate the non-stationary sequence, we suppose that the residual sequence $\{\varepsilon_t\}$ is white noise sequence, which requires its variance is a constant. But in fact this assumption is not always satisfied. Ignore the existence of heteroscedastic will lead to seriously underestimate the variance of residual series.

When the residual series with heteroscedastic, the forecasting errors at some points are relatively small, while at some points are relatively large. This variability shows there is a certain correlation in the variance of residual series. To estimate the correlation, Engle proposed a Autoregressive conditional heteroscedastic(ARCH) model in 1982[12]. Its main idea is: the conditional variance of the residual series depends on the previous values. In practice, the heteroscedasticity function of the residual series may have long-term autocorrelation. To amend this problem, Bollerslov proposed GARCH model in 1985. The GARCH(p, q) is given by

$$\begin{cases} \varepsilon_t = \sqrt{h_t e_t} \\ h_t = \eta_0 + \sum_{i=1}^{p} \eta_i h_{t-i} + \sum_{j=1}^{q} \lambda_j \varepsilon_{t-j} \end{cases} , e_t \xrightarrow{i,i,d} N(0,1) \tag{8}$$

Where, h_t is the conditional variance function of ε_t .If $p=0$, it is ARCH(q) model.

In the model, the residual series is no longer a random fluctuation, but has memory, thus the correlation of the residual variance is reflected.

D. Wavelet ARIMA-ARCH Model

This paper proposes a Wavelet ARIMA-ARCH model that combines wavelet, non-stationary time series and ARCH

Authorized licensed use limited to: PUC MG. Downloaded on June 23,2021 at 01:21:55 UTC from IEEE Xplore. Restrictions apply.

model, to forecast the wind speed.

The original wind speed time series is decomposed into detailed series at different scales $d_1, d_2, \cdots, d_J$ and a approximate series c_J, then obtain $D_j(j=1,2,\cdots,J)$ and C_J by reconstruction. Aimed at different time series(D_j and C_J), different ARIMA-ARCH models are established respectively to get their forecasting values, $\overline{C_J}$ and $\overline{D_j}$ ($j=1,2,\cdots,J$). Then, the last forecasting value of original wind speed can be expressed as

$$\overline{X_t} = \overline{C_J} + \sum_{j=1}^{J} \overline{D_j} \tag{9}$$

The forecasting flowchart of the proposed model is shown in Fig.1.

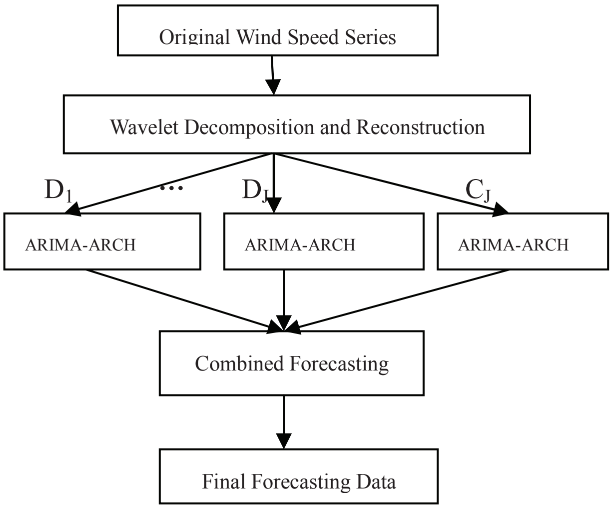


Fig.1 The modeling procedure

III. EXPERIMENTS AND COMPARISONS

Select the actual wind speed data of a wind farm as the research object. In experiment, 264 input/output data values are employed to simulate, we use the first 240 data values(the training data set) for the training, while the others are used as checking data for validating the proposed model.

A. Wavelet Decomposition and Reconstruction

1) Choose wavelet function. *db(N)* wavelet functions have been applied to several fields in dealing with non-stationary time series Because of the characteristics of compactly supported orthogonal. So we select *db(3)* as wavelet basis function.

2) Confirm the scales of wavelet decomposition. The more scales the original signal is decomposed, the better stationary the decomposed signals are, but great errors will be brought about at the same time. In this paper, the number of scales is 3($J=3$).

The original wind speed time series and the 3-layer wavelet decomposition and reconstruction results (C_3, D_3, D_2, D_1), are shown in Fig. 2

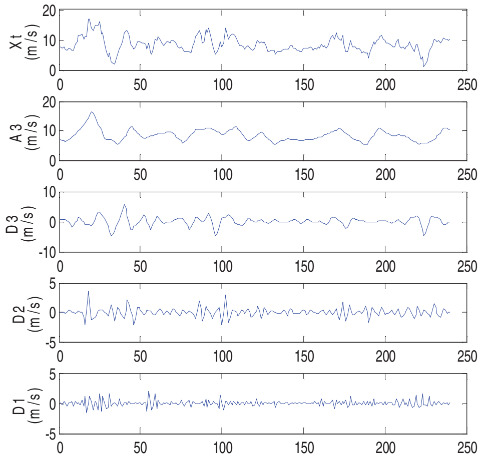


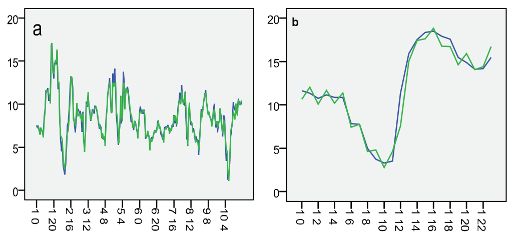
Fig.2 Original wind speed series and decomposed series by wavelet analysis

B. Establish ARIMA-ARCH Model in Every Wavelet-scale Domain

Considering the heteroscedasticity effect, establish different ARIMA-ARCH model for C_3, D_3, D_2, D_1. Suppose the forecasting values of D_j and C_3 are $\overline{D_j}$ and $\overline{C_3}$, then the forecasting values of original wind speed can be expressed as

$$X = \overline{D_1} + \overline{D_2} + \overline{D_3} + \overline{C_3} \tag{10}$$

The simulation and forecasting results of the original wind speed series are shown in Fig.3.



(a) fitting result; (b)forecasting result

(X-axis: wind speed data(m/s); Y-axis: time(h))

Authorized licensed use limited to: PUC MG. Downloaded on June 23,2021 at 01:21:55 UTC from IEEE Xplore. Restrictions apply.

Fig.3 Simulation and forecasting results of the original series b wavelet ARIMA-ARCH model

C. *Comparison of Different Models*

With the same wind speed training data set, use the traditional ARIMA model, wavelet ARIMA model, ARIMA-ARCH model, and wavelet ARIMA-ARCH model to forecast it respectively. Then the prediction results and forecasting errors of these models are shown in Fig.4

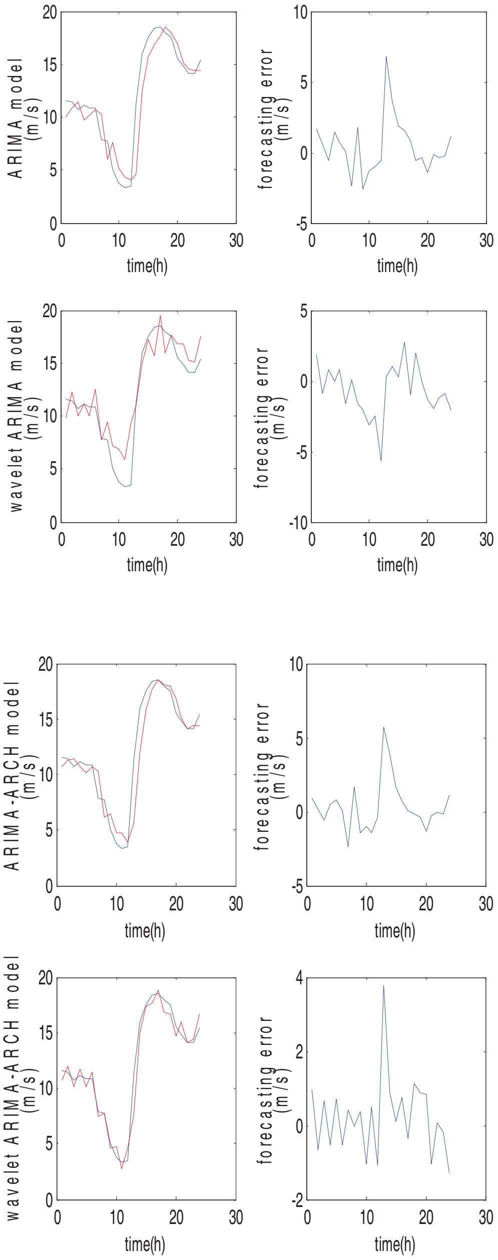


Fig.4 Simulation and forecasting results of the wind speed by different models

D. *Forecasting Results Evaluation*

In this study, an evaluation index, namely mean relative error (MRE), is used to measure the forecasting precision. The index formula is represented as follows

$$MRE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{X_t - \overline{X_t}}{X_t}\right| \tag{11}$$

Where, N is the number of forecasting periods, X_t and $\overline{X_t}$ separately are the actual and the forecasting time series values at period t.

By calculating, the mean relative errors of traditional ARIMA model, wavelet ARIMA model, ARIMA-ARCH model separately are 17.85%, 12.38%, 13.52%. However, the MRE of wavelet ARIMA-ARCH model is only 8.72%, which is significantly lower than the compared models. Besides, from Fig.4, we can see each model has larger prediction errors in 10 to 14 times. But, we can also see from Tab.1, the prediction errors of the improved wavelet ARIMA-ARCH model in these points are significantly lower. The results indicate this algorithm develops the precision greatly and possesses certain actual value in real-time dispatch.

Tab.1 Errors of wind speed forecasting by different models from 10 to 14

time (h)	ARIMA	wavelet ARIMA	ARIMA -ARCH	wavelet ARIMA -ARCH
10	-1.06	-2.54	-1.41	0.49
11	-6.56	-5.73	-3.41	-1.1
12	6.71	4.3	5.72	3.76
13	3.54	0.98	2.9	0.88
14	1.85	0.23	1.67	0.09

IV. CONCLUSION

1) The hybrid model combined wavelet transform and time series analysis is a very effective method for the nonlinear, non-stationary time series. By wavelet transform, wind speed series which is very random, is divided into profile and detail parts. Thus, the variation of wind speed series is clearer, more conducive to establish model and to improve the forecast accuracy.

2) As the variance of wind speed series is time-varying,

Authorized licensed use limited to: PUC MG. Downloaded on June 23,2021 at 01:21:55 UTC from IEEE Xplore. Restrictions apply.

the proposed model, which considers the heteroscedasticity effect, is just based on "changing variance". So it overcomes the shortcomings of traditional ARIMA model and makes the prediction accuracy improved.

3) In the future studies, we can take the impact of atmospheric temperature, air density, terrain shape and other physical factors on wind speed into account to improve accuracy.

REFERENCES

[1] Li Jing, Song Jia-hui, Wang Wei-sheng. Modeling and dynamic simulation of variable speed wind turbine with large capacity[J].Proceeding of the CSEE, 2004,24(6):100-105.

[2] Fan Shu，Liao J R，Yokoyama R，et a1．Forecasting the wind generation using a two-stage network based on meteorological information[J]．IEEE Transactions on Energy Conversion，2009，24(2):474-482．

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

[4] Kamal L，Jafri Z. Time series models to simulate and forecast hourly averaged wind speed in Quetta Pakistan[J].Solar Energy，1997,61(1):23-32.

[5] Wang Xiao-lan, Li Hui. Effective wind speed forecasting in annual prediction of output power for wind farm[J].Proceeding of the CSEE, 2010,30(8):117-122.

[6] Chao Wang，Wenjun Yan. Short-term wind speed prediction of wind farms based on improved particle swarn optimization algorithm and neural network[C]. 2010 International Conference on Mechanic Automation and Control Engineering，2010.

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

[8] Ji Guo-rui, Han Pu, Zhai Yong-jie. Wind speed forecasting based on support vector machine with forecasting error estimation[C].The 6th International Conference on Machine and Cybernetics, Dalian, China, 2007:2745-2739.

[9] Wei Wei-xian.Conditions of forecast combinations under non-stationary time series of data and their applications[J].Systems engineering and electronics, 1998:30-33.

[10]Wang Yang. Application of time series analysis[M].Beijing：Renmin University of China Press, 2005:41-95.

[11]Cheng Zheng-xing. Wavelet Algorithms and Applications[M].Xi'an：Xi'an Jiaotong University Press, 1998:1-10.

[12] He Yu, Gao Shan, Chen Hao. Wind speed forecast for wind farms based on ARMA-ARCH model[J].Jiangsu Electrical Engineering, 2009,28(3):1-4

Authorized licensed use limited to: PUC MG. Downloaded on June 23,2021 at 01:21:55 UTC from IEEE Xplore. Restrictions apply.