

Study on Modeling of Short-term Wind speed Forecasting based on Time Series Analysis

Yu ZHANG

Feixian School

Linyi University

Feixian, Shandong, 273400, P.R. China

Abstract—Wind speed forecasting is an important work in the design of wind farm planning. Because of the wind speed itself has a timing sequence and autocorrelation, the paper propose wind speed forecast model for wind farms based on time series analysis. In order to test the validity of time series analysis model, using ARMA (P, q) function. In the example, compared the distribution characteristics of wind speed prediction with the actual characteristics of wind speed distribution, verify the time series model proposed in this paper use for the feasibility of wind speed forecast.

Keywords: Time series; prediction; wind speed; wind farm;

I. INTRODUCTION

From the mature technology and economic feasibility, wind power has the best prospect in renewable energy. In recent years the power cost has been greatly decreased, wind power and conventional power generation has the potential to compete. With the adjustment of energy structure in our country, no pollution renewable wind power is paid more and more attention, and has drawn up the rapid development of wind power in support of relevant policies.

Assessment of wind farm wind energy resources is the development of wind power generation project based work. The wind speed forecasting is an important work in the design of wind farm planning^[1-2]. There are many methods of wind speed forecast, such as neural network method, time series method^[3-4]. But because of the time sequence also contains data sequence and data size, showing a dynamic process of the objective world, can reflect the objective world and change information, because the wind speed data with time order and discrete, so in this paper, using the time series analysis method in the prediction of wind speed. In the selected model, to determine the order of model parameter estimation and model; prediction of wind speed using the model to determine the appropriate order of the model and calculate the parameters of the model order after. On this basis, calculate the wind speed can directly reflect the distribution of wind energy resources of characteristic index, distribution, distribution and other parameters on the moon. Finally, by comparing the predicted data calculated by the monthly distribution, distribution and obtained by the historical wind speed data series measured value case, show

that the proposed autoregressive moving average model ARMA method is feasible.

II. STATISTICAL MODEL OF WIND SPEED TIME

A. Time series method

Models of random time series is divided into: autoregressive model (AR), moving average model (MA), the autoregressive moving average model (ARMA), autoregressive integrated moving average model (ARIMA)^[3]. Time series is described by the AR, MA, ARMA model is stationary time series, ARIMA model describes the non-stationary time series.

Time series describe by AR, MA, ARMA model is stationary time series, for the non-stationary time series, because the difference operation is information extraction ability determine the powerful, and many non-stationary sequence difference will show the properties of stationary series, so for nonstationary time series ARIMA model can be used for fitting. *ARIMA(p, d, q)* Model structure:

$$\varphi(B)(1-B)^d X_t = \theta(B)\varepsilon_t \quad (1)$$

Among them, B is the delay operator, D said the difference order, $\varphi(B) = 1 - \varphi_1 B - \dots - \varphi_p B^p$ is the autoregressive coefficients of the polynomial model, $\theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ is moving average coefficient polynomial model.

By formula (1) it is obvious that the essence of ARIMA model is the differential binding operations with ARMA model. This relationship is significant, indicating that any non stationary series as long as through the appropriate order differential operation can realize the difference after the smooth, can carry out the ARMA model fitting of difference sequence.

B. Stationarity definition and testing of time series

For time series {xt}, for any positive integer m, take $t_1, t_2, \dots, t_m \in T$, for any integer τ , its joint distribution function is:

$$F_{t_1, t_2, \dots, t_m}(x_1, x_2, \dots, x_m) = F_{t_1+\tau, t_2+\tau, \dots, t_m+\tau}(x_1, x_2, \dots, x_m) \quad (2)$$

Said time sequence {xt} is strictly stationary time series.

If the time series $\{x_t\}$ satisfies the following three conditions:

- 1) Taking $t \in T$, has $E(x_t^2) < \infty$;
- 2) Taking $t \in T$, has $E(x_t) = \mu$, μ is constant;
- 3) Taking $t, s, k \in T$, and $k + s - t \in T$, There exists a covariance function $\gamma(t, s) = \gamma(k, k + s - t)$;

Said $\{x_t\}$ is wide sense stationary sequence, also known as the second-order stationary.

Sequence diagram test: the so-called timing diagram is a two-dimensional coordinate chart, usually the horizontal axis represents time and the vertical axis represents the sequence of values. According to the stationary time series of mean and variance is a constant of nature, timing diagram of stationary sequence should show that the sequence beginning at a constant value near random fluctuations, and fluctuations in the range of bounded characteristic. If the timing diagram of an observed sequence showed that the sequence has obvious tendency and periodicity, it is usually not a stationary sequence.

C. Model parameter estimation

Least squares estimation also make full use of the observation value sequence information, the estimation precision is very high, and does not require information about the distribution of sample, so in practice the most commonly used is the conditional least squares estimation method.

In ARMA(p,q), set $\bar{\beta} = (\varphi_1, \varphi_2, \dots, \varphi_p, \theta_1, \theta_2, \dots, \theta_q)$ (3)

Assume that the sequence of past unobserved values equal to zero, i.e. $x_t = 0, t \leq 0$

According to this assumption can obtain the finite expressions of the residual sequence:

$$\varepsilon_t = \frac{\varphi(B)}{\theta(B)} x_t = x_t - \sum_{i=1}^t \pi_i x_{t-i} \quad (4)$$

Then the residuals squared sum is:

$$Q(\bar{\beta}) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n \left(x_i - \sum_{i=1}^t \pi_i x_{t-i} \right)^2 \quad (5)$$

Through iterative method, the formula (5) reaches the minimum value that is the parameters of the conditional least squares estimation.

D. Adaptability test of the model

The so-called the adaptability of the model refers to the model to explain the dynamic nature of the system (i.e., the degree of correlation of data sequences). A good model should be able to extract the values observed in almost all the samples, sequence related information, i.e. fitting residuals will no longer contain any relevant information, as a white noise sequence, so the adaptability test of the model is the white noise test model residuals, using χ^2 test method to make statistics:

$$Q = N(N+2) \sum_{k=1}^{L(N)} \frac{\rho_k^2}{N-k} \quad (6)$$

In the model for the conditions, type (6) approximately obey $\chi^2(L(N) - p - q)$. Among them, N is the number of samples, ρ_k is the autocorrelation function of residual sequence, p,q is order of the model, $L(N) = [N/10]$ or \sqrt{N} .

In certain significant level α , if $Q \leq \chi_{1-\alpha}^2(L(N) - p - q)$, we accept the original hypothesis, that the residual series is white noise sequence, the fitting model significantly; if $Q > \chi_{1-\alpha}^2(L(N) - p - q)$, then we reject the null hypothesis, that the residual sequence of non white noise sequence, the fitting model is not significant.

III. THE CALCULATION STEPS

The main steps of predict the wind speed using autoregressive integrated moving average model ARIMA (P, D, q) are described below:

- 1) Input wind speed data sequence $\{x_t\}$, and carries on the differential transform, approximate stationary time series, denoted as $\{y_t\}$;
- 2) For the $\{y_t\}$, calculate its ACF and PACF, and draw the correlation diagram and the partial autocorrelation plot;
- 3) According to step 2) sequences truncated and tailing of ACF and PACF of $\{y_t\}$, combined with BIC criterion, choose the appropriate model to fit the samples of wind velocity;
- 4) To estimate the model parameters;
- 5) To test the effectiveness of the model. If the model does not meet the requirements, turn to step 3), fully considering the various possibilities, the establishment of a plurality of fitting model, choose one optimal model;
- 6) Using the fitting model to predict the samples of wind velocity;
- 7) the wind speed forecasting value were analyzed and measured wind speed value, in order to test the predictive ability of the model.

IV. THE EXAMPLE ANALYSIS

A. The establishment of ARIMA model

1) Stationarity test of samples wind velocity series
The ARMA model is only suitable for fitting the stationary series, so first of all to the wind series are stationary test. In this paper, using the stationary series typically have short term correlation, through the stationary test of the autocorrelogram samples of wind velocity series. The original wind speed sequence diagram and correlation coefficient map and partial autocorrelation coefficient is shown in figure 1. Wind speed sequence diagram, the self-closed graph (ACF) and partial autocorrelation graph (PACF) after one order difference is as shown in figure 2.

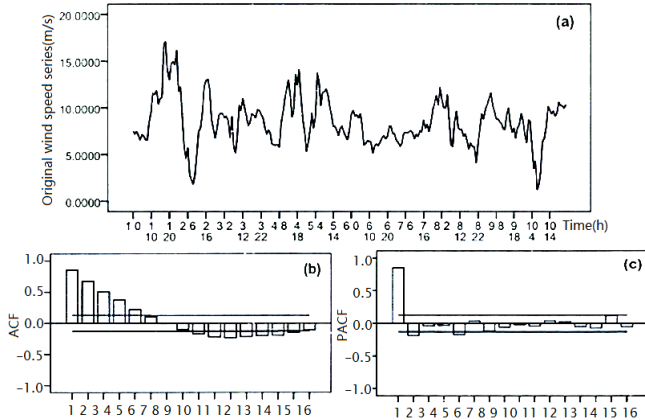


Fig.1 The original wind speed sequence diagram(a) and correlation coefficient(b) and partial correlation coefficient map graph(c)

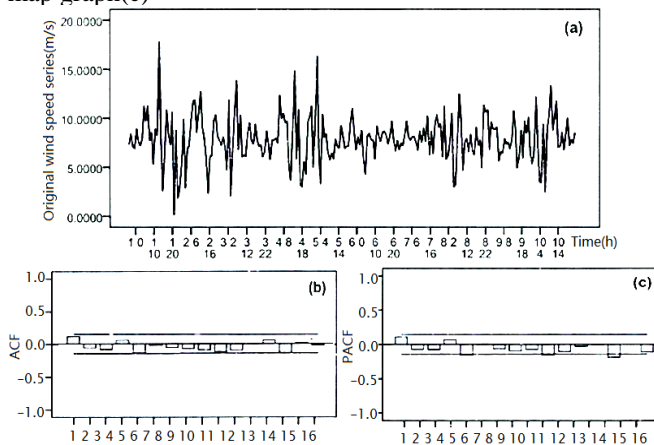


Fig.2 Wind speed sequence diagram(a) and ACF(b) and PACF(c) after one order difference

B. Preliminary estimation model

According to the autocorrelation coefficient of the wind speed series after difference and partial autocorrelation coefficient of the tailing and censored, can wind speed forecast model for sequence of preliminary rough estimate. Table 1 gives the samples of wind velocity correlation coefficient and partial correlation coefficient after difference.

Table.1 The correlation coefficient and partial correlation coefficient of wind series after difference

lag	1	2	3	4	5	6	7	8
ACF	0.097	-0.057	-0.079	0.049	-0.122	-0.023	-0.054	-0.071
PACF	0.097	-0.067	-0.068	0.061	-0.145	0.006	-0.061	-0.089
lag	9	10	11	12	13	14	15	16
ACF	-0.087	-0.106	-0.088	-0.002	0.045	-0.119	0.011	-0.022
PACF	-0.066	-0.135	-0.093	-0.027	-0.006	-0.172	-0.003	-0.101

C. To determine the precise estimation of model order and parameters of the model

The model has carried on the preliminary estimates, and need to precise order of the model, and the model parameters are determined. For each model preliminarily determined, by using BIC criterion and model matching value to select the

relatively optimal ARIMA model. The different ARIMA models corresponding BIC value and model matching value (fit value) is as shown in table 2. The model matching values are expressed as:

$$\text{Fit} = \frac{\sum [1 - (\text{wind speed forecasting value} - \text{wind speed measured value}) / (\text{wind speed measured values} - \text{the average wind speed})]}{\text{wind speed measured values} - \text{the average wind speed}}$$

Table.2 the BIC value and fit value of each model

Model	ARIMA(1,1,1)	ARIMA(1,1,2)	ARIMA(1,1,3)	ARIMA(1,1,4)	ARIMA(1,1,5)
BIC	0.709	0.687	0.736	0.763	0.707
fit	2025.34	2040.08	2025.22	2025.33	2039.69
Model	ARIMA(2,1,1)	ARIMA(2,1,2)	ARIMA(2,1,3)	ARIMA(2,1,4)	ARIMA(2,1,5)
BIC	0.735	0.634	0.694	0.745	0.734
fit	2025.52	2040.52	2039.85	2030	2039.74

Therefore, the establishment of ARIMA (2,1,1) model to predict the samples of wind velocity. Using the least squares estimate parameters of the model, as shown in table 3.

Table.3 ARIMA (2,1,1) prediction model parameters of wind speed

AR Parameter		MA Parameter
φ_1	φ_2	θ_1
1.016	-0.176	0.991

D. Adaptability test

After getting the fitting model, need to examine the model adaptability, namely residual test sequence model is white noise sequence, using the QLB statistic to test the adaptability of the model. Calculations were performed using SPSS software, the results is as shown in table 4.

Table.4 Adaptability test of ARMA(2,1,1) model

Ljung-Box Q(18)		
Q_{LB}	DF	P
18.436	15	0.240

As seen from table 4, statistics of P value is 0.240, significantly greater than 0.05, so that the residual series fitting model belongs to pure random sequence, namely the fitting model is significantly effective.

E. Wind speed forecasting and error analysis

The establishment of ARIMA (2,1,1) model, using it to forecast the future of wind speed, and data and measured wind speed are compared and analyzed, in order to test the predictive ability of the model. The fitting results, the prediction model of the effect and the prediction error as shown in figure 3. The (a) is a comparison chart of measured wind velocity and wind speed forecast model; (b) is the model prediction error.

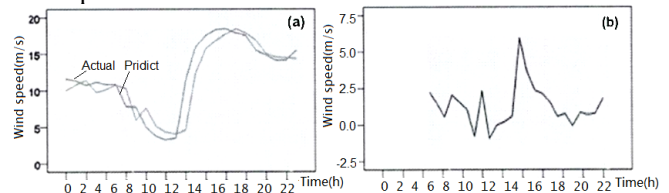


Fig.3 3 The forecast effect and forecast error of ARIMA(2,1,1) model

V. CONCLUSIONS

The paper use the time series modeling to simulate the actual wind speed data, and acquire the model of wind speed forecast, and have achieved good results, that the time sequence model is appropriate for wind speed prediction, can reflect the wind farm wind speed distribution characteristics in the future. But the prediction effect is not very ideal, especially in wind speed fluctuation is large, low forecasting precision.

REFERENCE

- [1] Wang Xiaolan, Li Hui. Research on effective wind in power output of wind farm annual forecast [J]. Chinese of the CSEE, 2010,30 (8): 117-122.
- [2] Wang Xiaolan, Wang Mingwei. Forecasting of short term wind speed based on wavelet decomposition and least squares support vector machine [J]. Power system technology, 2010, 34 (1): 179-184.
- [3] Gong Songjian, Yuan Yuhao, Wang Li, Zhang Ming. Short term wind speed forecasting PSO optimization based on LS-SVM [J]. Renewable energy, 2011,29 (2): 22-27.
- [4] Du Ying, Lu Jiping, Li Qing, Deng Yingling. Short term wind speed forecasting based on least squares support vector machine [J]. power system technology, 2008,32 (15): 62-66.
- [5] Du Libo. Research on the method of wind speed estimation based on RS-SVM [D]. Xi'an University of Science And Technology, 2009