Wind power forecasting based on time series ARMA model

Jingmin Wang¹, Qingwei Zhou² and Xueting Zhang²

¹ Beijing Key Laboratory of New Energy and Low-Carbon Development, North China Electric Power University, Beijing, 102206, China; wangjmtc@163.com
² North China Electric Power University, Baoding 071003, China

Abstract. Destruction of the ecological environment such as global warming and poor atmospheric pollution has forced carbon emission reduction and renewable energy utilization to accelerate. Under this circumstance, wind power generation as a clean renewable energy source has been valued by many countries and has increased globally, and domestic wind power generation is also being developed and promoted on a large scale. In order to further study the intermittent and unstable wind power, this paper uses the autoregressive moving average (ARMA) model based on time series algorithm to predict short-term wind power. The ARMA model is established based on the time series of wind power historical data, and the data is subjected to ADF test to prove the rationality of the selected data. The prediction model is analyzed and verified by using the measured wind power data of a wind farm from 0:00 am to 23:45 on the 1st of the month. In this paper, the prediction error is analyzed by MAE, RMSE and MdAPE. The analysis results show that the proposed model has less prediction error. The wind power prediction method based on ARMA model can effectively predict the wind power.

1. Introduction

Due to the instability and intermittent nature of wind power generation, the current difficulty in integrating wind energy has led to a large number of wind abandonment phenomena and wasted a large amount of renewable energy. When large-scale wind power is combined with the power system, it is necessary to consider whether the wind power fluctuation exceeds the balance ability of the power system[1]. Therefore, managing the intermittent nature of the operation and control of existing wind power systems becomes critical.

An effective solution to predict the future value of wind power generation is a wind power forecasting system with high precision and full functionality to help solve the problem of large-scale wind power grid-connected operation. Among them, wind power output prediction technology, especially short-term power prediction, needs an accurate prediction technology as a basis to help predict wind power generation changes and reduce the impact of wind power on power system security[2]. As wind power penetration is increasing, more sophisticated and innovative methods are needed.

Domestic wind power forecasting research started late, but in recent years, under the correct guidance of the national power sector, about 300 wind farms have used China's independently developed wind power forecasting system, and its development situation is relatively optimistic. However, the inadequacies are inevitable, especially the rapid development of its speed, which has resulted in the weight of the wind power industry. Although the cumulative installed capacity of wind power is still developing continuously, its growth rate is obviously slowing down. In the future power
generation strategy, China's development from speed-scale-oriented development to quality-efficiency-oriented development has become the trend of the times[3]. Domestic scholars have also made remarkable achievements in the research work on wind power prediction, and a large number of research methods have emerged.

2. Summary of Wind Power Forecasting Model

So far, there are a number of methods that are used to predict wind power, which are classified according to different classification criteria. According to different prediction methods, it can be divided into three types: statistical prediction method, intelligent prediction method and physical prediction method.

Among them, the statistical prediction method uses mathematical statistical methods to establish the mapping relationship between wind speed, wind direction and ambient temperature and wind power. The commonly used methods are continuous method, time series method, Kalman filter method, regression analysis method and exponential smoothing[4]. The intelligent prediction method uses artificial intelligence method to establish the mapping relationship between input and output, and predicts the wind power. The commonly used methods include neural network method, support vector machine method and wavelet analysis method[5]. A multi-objective optimization model for wind power interval prediction based on wavelet neural network was proposed. The probabilistic selection effect and constraint reduction strategy of basic multi-objective artificial bee colony algorithm were improved, and the reliability of wind power interval prediction was improved[6]. The physical prediction method establishes a mathematical relationship based on meteorological data such as wind speed, wind direction, and ambient temperature, and information, altitude, and terrain around the wind farm to predict wind power. Discuss the overall framework of wind power cluster power forecasting system from four aspects: clustering and single wind farm, from four aspects: forecasting process, data source, data flow direction and physical level[7]. A combined prediction model based on error analysis and correction, studied the distribution law of wind power prediction error, and constructed a combined model to improve prediction accuracy through error correction.

The autoregressive moving average model (ARMA) has high prediction accuracy in predicting the accuracy of short-term wind prediction. It describes the random data features well and analyzes historical wind speed data[8]. The research in this paper focuses on the prediction of existing wind farm data using the ARMA model of time series analysis to reduce the prediction error. After passing the stationarity and data processing, the wind power time series model is identified according to the statistical characteristics of the wind power sequence and the statistical characteristics of different time series models. In the fixed process of the model, the AIC criterion function is adopted. The parameters were estimated by least squares method, and the model was used to predict the ultra-short-term wind power.

3. Wind power prediction time series model

3.1 Basic form of time series analysis model

There are three basic types of models for time series analysis: AR (Auto-Regressive) model, MA (Moving Average) model, and ARMA (Auto Regressive Moving Average) model.

\[ Y_t = \sum_{i=1}^{p} \phi_i Y_{t-i} + u_t \]  \hspace{1cm} (1)

\[ Y_t = u_t - \sum_{j=1}^{q} \theta_j u_{t-j} \]  \hspace{1cm} (2)

\[ Y_t = \sum_{i=1}^{p} \phi_i Y_{t-i} + u_t - \sum_{j=1}^{q} \theta_j u_{t-j} \]  \hspace{1cm} (3)
Where $Y_t$ is the value of time series $t$ of wind power; it is an autoregressive parameter, $u_t$ is a random interference term at time $t$, forming a white noise sequence; $p$ is an autoregressive order; $q$ is the moving sliding average order.

The AR model and the MA model can be considered as special cases of the ARMA model. When $q=0$, the ARMA($p,0$) model is the $p$-order autoregressive model, when $p=0$, the ARMA($0,q$) model is the $q$-order moving average model. The essence is the ARMA model. The AR($p$) model is a reflection of the past state of the system itself, and the MA($q$) model is used to collect and analyze the effects of the noise sequence. The autoregressive moving average model (ARMA) treats the data formed over time as a random time series. The value at the $t$ moment in the sequence is not only related to the value at time $(t-1)$, but also $(t-1)$. The amount of random interference at the moment is related, thereby establishing a model to predict future values.

### 3.2 ARMA model building

#### 3.2.1 Data stability analysis

The ARMA model requires that the time series of the main part must be stationary. Before the modeling, the ADF unit root's stationarity test should be performed on the acquired time series. For non-stationary time series, it needs to be smoothed by differential processing. The ADF unit root test is implemented by Eviews software, and the stability of the time series is judged according to the calculation result. In general, compare the ADF statistic with the critical value at 5% of the significance level. If the ADF statistic is smaller than the critical value of the test, the unit root does not exist, indicating that the original sequence is a stationary sequence. If the ADF statistic is larger than the critical value of the test, the unit root exists, indicating that the original sequence is a non-stationary sequence.

#### 3.2.2 Model recognition

For the stationary time series $Y_t$, according to the numerical trend of its autocorrelation function (ACF) and partial correlation function (PACF), it can be judged which model is most suitable in AR model, MA model and ARMA model, the discriminant formula is as follows.

$$\text{ACF}(k) = \frac{\rho(t, t+k)}{\sqrt{\text{Da}(t) \text{Da}(t+k)}}$$

$$\text{PACF}(k) = \frac{E[a(t) - E(a(t))[a(t-k)-E(a(t-k))]]}{E[(a(t-k) - E(a(t-k)))^2]}$$

#### 3.2.3 Model ordering and parameter estimation

The order of the model determines the autoregressive order $p$ and the sliding average order $q$. By autocorrelation and partial autocorrelation analysis of time series samples, the autocorrelation and partial autocorrelation curves are used to determine the order of the model, and then several models of different orders are selected for analysis. The model is based on the Akaike information criterion (AIC criterion). The order is as follows:

$$Z = N \ln \sigma^2 + 2(p + q)$$

Where $Z$ is the AIC function value; $\sigma^2$ is the variance estimate of the model residual sequence; $N$ is the sample length. The smaller the $Z$ value, the better the fitting effect. The least squares estimation method is used to estimate the parameters. Under the condition of known sample sequence values, the nonlinear least squares method is used to find the autoregressive parameters and the moving average parameters that minimize the sum of the residuals.

#### 3.2.4 Model suitability test

In the ARMA model, the basis of parameter estimation is to assume that the random interference term is a white noise sequence. There are many methods for testing white noise. The most commonly used
is the autocorrelation function test. If the residual sequence is not autocorrelated, the residual is a white noise sequence and the model is valid.

4. Case study

4.1 Raw data
The measured wind power data of a wind farm from 0:00 am to 14:45 pm on the 1st of the month was selected, and the sampling interval was 15 min. As the original data, the daily load curve of the wind power was predicted on the 15th. In the specific forecasting process, the influence of factors such as machine failure and human operation on wind speed is not considered, and the prediction work is based purely on historical wind power data.

4.2 Data processing and model building
The Q statistic tests the autocorrelation function of the wind power history sequence exponentially decaying, but the decay rate is very slow, so the original sequence is considered to be non-stationary. To eliminate sequence trends and reduce data fluctuations, a first-order natural logarithm difference is generated to generate a new sequence r. The t-statistic of the ADF test = -2.7964, which is less than the t-statistic threshold of the test level of 1%, 5%, and 10%, and the probability value P corresponding to the t-statistic is very small, so the null hypothesis that the sequence has a unit root is rejected. That is, the original data is stable.

![Figure 1. Raw data ADF test](image)

The autocorrelation coefficient and the partial autocorrelation coefficient are calculated according to the formula for the wind power sequence of the first two weeks. The result is shown in Fig. 2. The autocorrelation coefficient of the sequence is gradually decreasing, and it is decreasing until the fourth step of the lag. The partial autocorrelation function shows the statistical spikes at the 1st and 2nd lags. Therefore, the following two types are estimated. Model form: ARMA (4, 2) and ARMA (4, 1).

![Figure 2. Calculation results of autocorrelation coefficient and partial autocorrelation coefficient of data](image)

The probability values corresponding to the goodness of fit of the two ARMA models are very small, indicating that the model is significant overall and the fitting effect of the model is particularly good. The estimation results give that the modulo polynomial reciprocal roots of the AR process and the MA process are all less than one, so the two models of the above estimation can be considered to be stationary and reversible.
Comparing the two models, the adjusted coefficient of ARMA (4, 1) is greater than ARMA (4, 2). The AIC and SC criteria of ARMA(4,1) are both smaller than the corresponding values of ARMA(4,2), so the model ARMA(4,1) is considered to be better than the model ARMA(4,2), and the final model is ARMA (4, 1) model.

(3) Applicability test
The autocorrelation coefficient and the partial correlation coefficient of the residual sequence approach 0, and the residual sequence appears as non-autocorrelation, and the residual is considered to be a self-noise sequence. The probability of the Q statistic of the residual sequence starts to approach 0 after the lag period is greater than 11, and the final result can be approximated as the self-noise sequence, indicating that the model is valid.

4.3 Analysis of prediction results
According to the ARMA (4,1) model, the wind power of the 15th day is predicted by the actual wind power data of the selected wind farm for two weeks at the beginning of the month. The prediction result is shown in Fig. 3.

![Figure 3. Comparison of wind power predicted and actual values on the 15th](image)

The solid line is the actual value and the dotted line is the predicted value. It can be seen from Figure 3 that the predicted power curve on 15th, is basically fitted to the actual power curve. Comprehensive evaluation of the results of short-term wind power forecasting is one of the important components of forecasting and an important basis for adopting forecasting methods. It is unscientific to use only some error indicators for judgment. Therefore, a comprehensive and comprehensive evaluation of the prediction results must be evaluated in a variety of error forms to judge the feasibility, validity and accuracy of each prediction method.

\[
MAE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100% 
\]

(7)

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left( y_i - \hat{y}_i \right)^2} \times 100% 
\]

(8)

\[
M_dAPE = \text{median} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100% 
\]

(9)

| Method | Error |
|--------|-------|
| MAE    | 9.14% |
| RMSE   | 30.24%|
| MdAPE  | 9.86% |
5. Conclusion

Wind power prediction can not only optimize, secure, and control power grid scheduling, but also has important significance for reducing the operating cost of the power grid and increasing the value of power in the power market. The time series method focuses on the time series characteristics of wind power historical data, and makes short-term predictions and comprehensively evaluates the prediction effects. The results of the example show that the accuracy of the ARMA model prediction method is high when dealing with ultra-short-term prediction, but the prediction accuracy decreases with the extension of the prediction time. In practical applications, adjustments should be made according to usage to improve accuracy, and attention should be paid to the preprocessing of raw data. Current research is limited to ultra-short-term predictions, and other time-scale prediction methods will be the focus of follow-up research.

References
[1] Gu Xingkai, Fan Gao, Wang Xiaoyong. Overview of wind power forecasting technology [J]. Grid Technology, 2007, (S2): 335-338.
[2] Fan Gao, Pei Jie, Xin Yaozhong. Current situation and prospect of wind power forecasting [J]. China Power, 2011, 44 (6): 38-41.
[3] Xue Yusheng, Zhao Junhua. Comments on short-and ultra-short-term wind power forecasting [J]. Power system automation, 2015, 39 (6): 141-151.
[4] Qian Zheng, Pei Yan, Cao Lixiao, et al. Overview of wind power forecasting methods [J]. High voltage technology, 2016, (04): 1047-1060.
[5] Hong Cui, Lin Weiming, Wen Buying. A review of wind speed and wind power forecasting methods in the electric field [J]. Grid and clean energy, 2011, 27 (1): 60-66.
[6] Chen Jie, Shen Yanxia, Lu Xin, et al. A multi-Objective intelligent optimization prediction method for wind power probability interval [J]. Power grid technology, 2016, (08): 2281-2287.
[7] Peng Xiaosheng, Xiong Lei, Wen Jinyu, et al. Summary of the improved method for short and ultra short term power forecasting of wind power cluster [J]. Chinese Journal of Electrical Engineering, 2016, (23): 6315-6326.
[8] Wang Guoquan, Wang Sen, Liu Huayong. Study on short-term wind speed forecasting method of electric field [J]. Renewable energy, 2014, 32 (8): 1134-1139.