A Wind Speed Prediction Model Based on ARIMA and Improved Kalman Filter Algorithm

Shengyuan Yuan, Yanxia Shen*

College of Internet of Things Engineering, Jiangnan University, Wuxi, Jiangsu, China

*Corresponding author: 3071789709@qq.com

Abstract. This article is dedicated to solving the problem of wind speed prediction. A time series analysis method based on the establishment of a differential autoregressive sliding model for simple training and a prediction model to obtain the state is proposed. The Kalman filter method with adaptive weighting coefficients is used to predict the equation, and the experimental results show that the composite algorithm can effectively reduce the prediction error.

1. Introduction

With the increase of global resource consumption and serious environmental pollution problems, the energy crisis has intensified, and countries around the world are paying more and more attention to the development and utilization of new energy [1]. Wind energy, a pollution-free green energy source, has become more and more important [2]. Wind energy is characterized by intermittency and uncertainty. The uncertainty of wind speed changes makes it difficult for wind turbines to continuously maintain their maximum power point [2]. How to accurately and effectively predict the wind speed to better serve the wind turbine and improve its stability [2] has become an important technical issue for wind power generation.

Currently commonly used short-term wind speed forecasting methods include continuous forecasting method [3], time series analysis method [3-4], support vector machine method [5], Kalman filter method [7-8] and wavelet. Analysis methods and various hybrid methods. Literature [3] established a wind speed prediction model using time series methods, but the average relative prediction error reached 15.7%, which is low in accuracy. Literature [6] uses the support vector machine method to predict wind speed, but it is necessary to collect data such as air pressure and temperature for training and optimization of parameters, which is a complex process. Literature [10] uses wavelet analysis and neural network method to predict wind speed, but the wavelet algorithm is building a suitable wavelet basis. There are greater difficulties in the above modeling, that is, the "black box" modeling makes the network learning process and model optimization more difficult [9]. Literature [9] established a time series sliding autoregressive model, but did not consider the divergence of the Kalman filter, and it is difficult to obtain satisfactory accuracy results. This paper proposes a Kalman filter prediction method based on time series sliding autoregressive model (ARIMA), and then adjusts the innovation weight to give a weighting coefficient to suppress divergence.
2. Principle Introduction

2.1. Prediction principle of time series analysis method
The commonly encountered series are non-stationary time series with obvious periodicity or trend [7]. There are three main steps in using time series analysis to build a random sequence model: selecting a model, determining the order of the model, and estimating parameters. This article mainly uses the sliding autoregressive model (ARMA).

After determining the type of model required, a specific model will be established and the sequence determined. This article chooses AIC criterion from the optimal criterion to determine the order. In the ARIMA regression model, the estimation standard is first determined to minimize the remaining estimated sum of squares, and then the coefficients of the model are determined by the least square method.

2.2. Basic principles of Kalman filtering
The core of Kalman filter method is to establish the state equation and observation equation, as follows:

\[ X(k+1) = \Phi(k+1,k)X(k) + \Gamma(k+1,k)w(k) \]  \hspace{1cm} (1)

\[ Z(k+1) = H(k+1)X(k+1) + v(k+1) \]  \hspace{1cm} (2)

Among them, the formula (1) is called the state equation, and the formula (2) is called the measurement equation. This article uses Kalman filter theory to solve the prediction problem. The final prediction recurrence equation can be found in Reference [9]:

3. Example modeling and simulation experiment

3.1. Time series analysis modeling
Sampling and training of wind speed measured data of wind farms in North China in 2012. The first 200 data among the 300 data are used as the sampling signal for modeling training and compared with the last 100 measurement data. The wind speed sampling signal and the differential autocorrelation function are shown in Figure 1 and Figure 2.

![Figure 1: Sampling point of wind speed](image1)

![Figure 2: Differential sequence signal autocorrelation function](image2)

It can be seen that the autocorrelation function of the sequence after the first-order difference can quickly decay to about 0 in a very short time, and constantly oscillate around 0. It can be seen that the wind speed series has been transformed into a stationary series after first-order difference. Use AIC criteria to determine the order and calculate on MATLAB. The results are shown in Table 1. The ARIMA model in this example is a (5, 1, 2) model, and the equation is:
\[ (1+0.1864B+0.4665B^2-0.5225B^3+0.2405B^4-0.1422B^5)X(t) = (1+1.0176B+0.9910B^2)a_t \]  

(3)

| \( \phi_1 \) | \( \phi_2 \) | \( \phi_3 \) | \( \phi_4 \) | \( \phi_5 \) | \( \theta_1 \) | \( \theta_2 \) | AIC |
|---|---|---|---|---|---|---|---|
| 0.1864 | 0.4665 | -0.5225 | 0.2405 | -0.1422 | 1.0176 | 0.9910 | -3.1488 |

Table 1. Parameters of ARIMA (5, 1, 2) model

At this time, the above formula can be used for prediction, and the actual wind speed curve and the predicted wind speed curve can be compared, as shown in Figure 3:

![Comparison of time series forecast results and actual results](image)

**Figure 3. Comparison of time series forecast results and actual results**

3.2. The establishment of Kalman equation of state and prediction equation

For the wind speed sliding autoregressive model of this example, for the convenience of modeling, use \( X_1(k) = X(k) \), \( X_2(k) = X(k-1) \), and so on, the model is expressed as (4):

\[ X_1(k+1) = 0.8136X_1(k) - 0.2801X_2(k) + 0.989X_3(k) - 0.763X_4(k) + 0.3827X_5(k) - 0.1422X_6(k) \]  

(4)

Let (4) among them \( X_2(k + 1) = X_1(k) \), \( X_3(k + 1) = X_2(k) \). And so on, the state equation of the system can be expressed as (5):

\[
\begin{bmatrix}
X_1(k+1) \\
X_2(k+1) \\
X_3(k+1) \\
X_4(k+1) \\
X_5(k+1) \\
X_6(k+1)
\end{bmatrix} = \begin{bmatrix}
0.8136 & -0.2801 & 0.9890 & -0.7630 & 0.3827 & -0.1422 \\
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix} \begin{bmatrix}
X_1(k) \\
X_2(k) \\
X_3(k) \\
X_4(k) \\
X_5(k) \\
X_6(k)
\end{bmatrix} + \begin{bmatrix}
1 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix} \begin{bmatrix}
\mathbf{w}(k+1)
\end{bmatrix}
\]

(5)

Take \( \mathbf{X}(0|0) = [0] \), \( \mathbf{P}(0|0) = 10\mathbf{I} \), \( \mathbf{R}(k) = [1] \), \( \mathbf{Q}(k) = [1] \) based on experience. The prediction result is shown in Figure 4. Due to certain errors in the establishment of the system model, the common Kalman filtering method will be different, resulting in inaccurate prediction results. In order
to suppress this phenomenon. In order to pursue higher precision, consider using dynamic form instead of single form, select fading factor through the covariance estimation of innovation sequence and theoretical value constructor to adjust the weight of innovation, and assign variable coefficients $\lambda_k$ (The fading factor) Introduced into the Kalman filter equation to dynamically increase or decrease the influence of the current measured value on the state variable in the recursive process, and improve the influence of incorrect system modeling and unknown interference characteristics on the filtering accuracy. The required accuracy can be achieved at any stage.

The value of $\lambda_k$ is in accordance with:

$$\lambda_k = \max \left\{ 1, \frac{tr(C_k)}{tr(C_{\hat{k}})} \right\} \quad (6)$$

Among them, $\hat{C}_k$ is the predicted value of $C_k$, can be expressed as $\hat{C}_k = \lambda_k H P_{k/k-1} H^T + R$. The experimental results are shown in Figure 5. The comparison of experimental data is shown in Table 2.

| Types                              | actual value | Time series forecast | Normal Kalman filter value | Fixed weighting coefficient Kalman filter value | Dynamic weighting coefficient Kalman filter value |
|------------------------------------|--------------|----------------------|----------------------------|-----------------------------------------------|-----------------------------------------------|
| Average wind speed                 | 15.5906      | 15.5807              | 15.5813                    | 15.6141                                       | 15.5856                                       |
| Relative prediction error          | ——           | 0.0099               | 0.0093                     | 0.0235                                        | 0.005                                         |
| Average percentage forecast error  | ——           | 5.10%                | 4.67%                      | 6.37%                                         | 2.10%                                         |
4. Conclusions
From the experimental results in Table 2 and Figure 8, it can be seen that the data predicted by the hybrid Kalman filter method is more accurate than the data predicted by the pure time series analysis method. The ordinary Kalman filter method will appear to be a certain amount in the wind speed jump phase. There will be a certain "divergence phenomenon" in the later stage. If a fixed weighting coefficient is given, there will be a "jumping phenomenon" with large errors in the early stage. The hybrid forecasting method improved by the dynamic weighting coefficient method reduces the average percentage forecast error to 2.10%. It has a stronger degree of fit no matter in the early wind speed change phase or in the later wind speed steady phase. It shows that this hybrid method is feasible.

References
[1] Li Pengyu, Zheng Tianlong, Li Lin, Ma Yingqun, Sun Xu, Liu Junxin. An appropriate technique for treating rural wastewater by a flow step feed system driven by wind-solar hybrid power. [J]. Environmental Research, 2020, 187.
[2] Zengye Fu. Discussion on new energy power generation wind power technology [J]. Science and Technology Innovation, 2019(36): 145-146.
[3] Shaoji Zhang, Zeng Jie, Zhang Hua, Application of time series model in wind speed prediction of wind farms [J]. Water Resources and Hydropower Technology, 2016.1000-0860 (2016) 12-0132-04.
[4] Hanmo Sun. Research on wind speed prediction of wind turbine based on ARMA model [D]. North China Electric Power University (Beijing) North China Electric Power University, 2011.
[5] Manping Fan, Zhou Dong. Short-term wind speed prediction based on improved particle swarm optimization LS-SVM [J]. Journal of Electric Power, 2020, 35(02): 123-128+142.
[6] Wind speed prediction and analysis based on Kalman filtering [D]. Degen Xin, Lanzhou University, 2017.
[7] Yufang Chang, Zhang Li, Xie Hao, Liu Guangyu. Combined forecast of short-term wind speed using wavelet analysis and neural network [J]. Journal of Huaqiao University (Natural Science Edition), 2019, 40(04): 556-560
[8] Chunbo Xiu, Ren Xiao, Yanqing Li, Mingfeng Liu. Short-term prediction method of wind speed series based on Kalman filter [J]. Transactions of the Chinese Society of Electrical Engineering, 2014, 29(02): 253-259.
[9] Difu Pan, Liu Hui, Yanfei Li. Wind farm wind speed prediction optimization model based on time series analysis and Kalman filter algorithm [J]. Power System Technology, 2008, 32 (7): 82-86.
[10] Wangyan Qiu, Rongbing Li, Jianye Liu. General aviation GNSS/micro-inertial integrated navigation algorithm based on improved adaptive fading Kalman filter [J]. Electronic Measurement Technology, 2020, 43(10): 95-100