Ultra-short-term Wind Power Output Prediction Based On LS-WMC Combined Model

Yanan Du1*, Zhengning Pang1 and Jingxian Qi1

1NARI GROUP CORPORATION/STATE GRID ELECTRIC POWER RESEARCH INSTITUTE, Nanjing, Jiangsu, 210003, China.
*Corresponding author’s e-mail: duyanan@sgepri.sgcc.com.cn

Abstract. In this paper, an ultra short term wind power output prediction method based on least square and weighted Markov chain model is proposed. Firstly, the wind farm output is fitted based on the least square method, and then the prediction accuracy is improved by using the weighted Markov chain model. At the same time, the partial exponential weighting method is used to improve the prediction accuracy of ultra short-term wind power output on the basis of reducing the generation frequency of transition probability matrix, so as to optimize and improve the prediction results. The results show that the method can improve the prediction accuracy.

1. Introduction

Wind power output has the characteristics of strong randomness and volatility [1-3]. Accurate wind power output prediction results can provide a good foundation for power grid security check, dynamic state estimation, security and stability analysis, reactive power optimization, plant layout control, etc. However, the wind power output is prone to mutation and poor stability, especially the seasonal influence leads to the existence of high-frequency fluctuation component and seasonal component in the wind power output data, and the poor signals have a great impact on the quality of the prediction sample data. At the same time, due to the wrong data and operation mode changes caused by data acquisition and transmission, it is easy to cause data mutation and data quality decline. Preprocessing of such volatile data includes empirical correction method, curve replacement method, interpolation method and wavelet analysis [4-7].

Wind power output prediction methods mainly include system output distribution prediction method and node output self variation prediction method. In reference [8], the prediction is based on the system output distribution, and the distribution coefficient is used to predict the output of each wind power. This method does not consider the inconsistency of output regions, resulting in low prediction accuracy. Literature [9] used state estimation to build wind power output model based on historical data, and dynamically adjusted the distribution coefficient, but lack of case verification. Literature [10] uses PSO algorithm to optimize BP neural network algorithm, so that it is not easy to fall into local minimum and enhance its generalization ability. In reference [11], the hybrid method of fuzzy system and artificial neural network is used to realize the wind power output prediction, and the output is divided into irrelevant output and related output based on the meteorological impact, so the training time is longer. Literature [12-13] uses a variety of single output prediction algorithms, such as support vector machine regression based on probability density, with low accuracy. In view of the complexity of the above model training and the low prediction accuracy, this paper uses the method of combining the least square method and the weighted Markov chain model to explore the change rule based on the historical wind
power output data, and carries on the data model fitting analysis and test for the specific wind power output data.

2. Least square fitting

Please follow these instructions as carefully as:

The least square fitting is based on the least square sum of errors, and the least square sum of distances between observation points and estimation points is minimized by selecting estimators. Suppose that the functional relationship between input X and output Y of the model is as follows:

$$Y = f(X; a_1, a_2, \ldots, a_k)$$  \hspace{2cm} (1)

Where $f$ is the mapping relation. For a given data set \{x_i, y_i\}, $i = 1, 2, \ldots, n$, the parameter $a_1, a_2, \ldots, a_k$ is estimated by using the set. Let the fitting polynomial be:

$$y = \hat{a}_0 + \hat{a}_1x + \ldots + \hat{a}_kx^k$$ \hspace{2cm} (2)

Where $\hat{a}_0, \hat{a}_1, \ldots, \hat{a}_k$ is the estimate of $a_1, a_2, \ldots, a_k$ respectively.

The sum of the distances from the given point to the fitting curve, that is, the sum of the squares of the deviations, is calculated as follows:

$$R = \sum_{i=1}^{n} [y_i - (\hat{a}_0 + \hat{a}_1x + \ldots + \hat{a}_kx^k)]^2$$ \hspace{2cm} (3)

The minimum value of $R$ can be calculated by solving the following conditions.

$$\frac{\partial R}{\partial a_1} = \frac{\partial}{\partial a_1} \sum_{i=1}^{n} [y_i - (\hat{a}_0 + \hat{a}_1x + \ldots + \hat{a}_kx^k)]^2 = 0$$

$$\frac{\partial R}{\partial a_2} = \frac{\partial}{\partial a_2} \sum_{i=1}^{n} [y_i - (\hat{a}_0 + \hat{a}_1x + \ldots + \hat{a}_kx^k)]^2 = 0$$

$$\ldots$$

$$\frac{\partial R}{\partial a_k} = \frac{\partial}{\partial a_k} \sum_{i=1}^{n} [y_i - (\hat{a}_0 + \hat{a}_1x + \ldots + \hat{a}_kx^k)]^2 = 0$$ \hspace{2cm} (4)

The results are as follows:

$$\begin{bmatrix}
1 & x_1 & \ldots & x_1^k \\
1 & x_2 & \ldots & x_2^k \\
\vdots & \vdots & \ddots & \vdots \\
1 & x_n & \ldots & x_n^k
\end{bmatrix}
\begin{bmatrix}
\hat{a}_0 \\
\hat{a}_1 \\
\vdots \\
\hat{a}_k
\end{bmatrix} =
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix}$$ \hspace{2cm} (5)

The estimated value of parameter $\hat{a}_j (j = 1, 2, \ldots, k)$ can be calculated by formula 5, and then the predicted value $\hat{y}_i$ at any time can be obtained by fitting function.

3. Markov chain prediction model

The Markov chain model is expressed as follows:

Suppose \{x_n, n \geq 0\} is a random process sequence, for any $i_0, i_1, i_2, \ldots, i_n, i_{n+1} \in S$, the state space s is a finite set of samples, and the probability of state transition at a certain time only depends on the state of the previous time.

$$P(X_{n+1} = i_{n+1}|X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1) = P(X_{n+1} = i_{n+1}|X_n = i_n)$$ \hspace{2cm} (6)

For any $i, j \in S$, $P\{X_{n+1} = j|X_n = i\} = p^{(1)}_{ij}(n)$ is called the one-step transition probability of $X_n$ at n, and $P^{(1)} = \{p^{(1)}_{ij}\}$ is the one-step state transition probability matrix of \{X_n\}. Similarly, for any $i, j \in S$, $P\{X_{n+k} = j|X_n = i\} = p^{(k)}_{ij}(n)$ is called the k-step transition probability of $X_n$ at n, and $P^{(k)} = \{p^{(k)}_{ij}\}$ is the k-step state transition matrix of \{X_n\}. Recorded as:
4. Least squares weighted Markov chain model

The least square method takes the least square sum of error as the criterion to predict wind power output. The model will learn the law of partial data noise, which has the problem of over fitting. The Markov chain model can reflect the random process characteristics of wind power output, which is more suitable for data prediction in the case of large volatility. Therefore, this paper combines the least square method with the weighted Markov model, so that the prediction method has the advantages of both. At the same time, a fixed frequency state transition matrix is used to improve the prediction speed, and the prediction results are optimized based on the partial exponential weighting method.

4.1. Partition state

The relative error $\varepsilon(k)$ between the wind power output value fitted by the least square method and the actual value is divided into m states. If $\varepsilon(k) \in (d_i, d_{i+1}), i = 1, 2, ..., m$, then the relative error of the $k$-th time point is in $E_i$ state, where, $d_i \leq d_{i+1}$ is the upper and lower bounds of state $E_i$.

4.2. Weight calculation

The Markov chain is weighted and improved based on autocorrelation coefficient, that is, the correlation coefficient is obtained by autocorrelation analysis of the original data. If the absolute value of autocorrelation coefficient is large, a large weight is given.

$$r_i = \frac{\sum_{k=i}^{n} (x_k - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=k}^{n} (x_k - \bar{x})^2}$$  \hspace{1cm} (8)

Where $r_i$ is the $k$-th autocorrelation coefficient; $x_i$ is the sample data at time $i$; $\bar{x}$ is the mean value of the sample data. By normalizing the autocorrelation coefficient, the weight of each order is obtained as follows:

$$w_i = \frac{|r_i|}{\sum_{i=1}^{m} |r_i|}$$  \hspace{1cm} (9)

Where $m$ is the maximum order calculated according to the prediction.

4.3. Calculate the predicted value

Selecting the $k$ time points closest to the prediction time, taking the state where the relative error of the $K$ time points is located as the initial state, and combining with the corresponding transition probability matrix, the state probability $P_i(k), i \in E$ at that time can be predicted; The weighted sum of the prediction probabilities of the same state is taken as the final transition probability.

$$P_i = \sum_{k=1}^{m} w_i \times P_i(k)$$  \hspace{1cm} (10)

Then the optimized prediction value is obtained as follows:

$$\hat{y} = \hat{y} \left[1 - \frac{1}{2}(d_i + d_{i+1})\right]$$  \hspace{1cm} (11)

4.4. Exponential weighted average

The generation frequency of state transition matrix is set as $p$, and the random sequence $\{x_n\}, n \geq 0$ is described as $y$, if $t = kp, k \geq 0$, formula (11) is used to predict the result; Otherwise, the predicted value
of wind power output at time $t$ is regarded as the weighted average of the weighted predicted value at time $t-1$ and the unweighted predicted value at time $t$.

$$
\hat{Y}_t = \begin{cases} 
\hat{y}_t, & t = kp \\
(1-\theta)\hat{Y}_{t-1} + \theta \hat{y}_t, & \text{other}
\end{cases}
$$

5. Data experiment and analysis

In this paper, the least square weighted Markov chain model is used to predict the wind power output. Firstly, the prediction time period is determined. Then, based on the date value, the wind power output data in the first $k$ cycles and $s$ measurement points which are closer to the prediction date are selected to construct the historical data series, and then the least square method is used to fit it. The fitting relative error is divided into several states based on the value of error results, and then the state transition matrix of Markov chain model is constructed. The prediction results are optimized by probability weighting method. At the same time, according to the frequency setting of state transition matrix generation, the partial exponential weighting method is used to realize the wind power output prediction of non set cycle time.

In order to reduce the instability of wind power output data, based on the output data of a wind farm in Hebei Province on a specific date in recent one year, this paper uses the least squares weighted Markov model to predict the output data of the wind farm. Firstly, by analyzing the wind power output change rate and average output change rate, the abnormal points are identified, and the interpolation method is used to preprocess the original bus output data.

5.1. Least square prediction

The data after preprocessing are fitted by least square method. Taking partial sampling points as an example, the fitting results are as follows:

| Num | Actual wind power output (MW) | Forecast of wind power output (MW) | Relative error (%) | State |
|-----|-------------------------------|-----------------------------------|-------------------|-------|
| 1   | 1594                          | 1462.5                            | -8.2496           | E3    |
| 2   | 1533.9                        | 1462.5                            | -4.6548           | E4    |
| 3   | 1470.1                        | 1320.6                            | -10.1693          | E5    |
| 4   | 1340                          | 1305.1                            | -2.6044           | E3    |
| 5   | 1431.5                        | 1320.6                            | -7.7568           | E2    |
| 6   | 1254.3                        | 1314.8                            | 4.82340           | E1    |
| 7   | 897                           | 887.6                             | -1.0479           | E3    |
| 8   | ...                           | ...                               | ...               | ...   |

According to the actual situation of relative error and Markov chain state division method, the relative error is divided into five states, and the division standard is shown in Table 2. The results of relative error division are shown in the status column of Table 1.

| State | State space (%) |
|-------|-----------------|
| E1    | (-10,-5)        |
| E2    | (-5,0)          |
| E3    | (0,4)           |
| E4    | (4,8)           |
| E5    | (8,12)          |
5.2. Weighted Markov chain prediction

According to the numerical value of the state column in Table 1, the state transition matrix of 1 to 5 steps is calculated. The autocorrelation coefficients and weights of each order are calculated according to equations (8) and (9), as shown in Table 3.

| k  | $r_k$ | $w_k$ |
|----|-------|-------|
| 1  | 0.7054| 0.3718|
| 2  | 0.5719| 0.2812|
| 3  | 0.4981| 0.2904|
| 4  | 0.1384| 0.0149|
| 5  | -0.0345| 0.0417|

The state transition matrix is updated step by step, and the weighted Markov chain model is used to modify the training result of the least square method. The curve of the prediction result is shown in Figure 1. In order to further illustrate the effectiveness of the proposed method, this paper compares the proposed method (LS-WMC) with the least square fitting method (LS), and selects 20 samples for experiments. The error of 20 experiments is shown in Figure 2.

6. Conclusion

Based on the time series and instability of wind power output, this paper uses the least square method to fit the wind power output curve. Based on the fitting error, the weighted Markov chain model is used to divide the state and predict the maximum transition probability, and the initial fitting value is optimized. In order to reduce the generation frequency of the state transition matrix, the partial exponential weighting method is used to optimize, so as to make the prediction error smaller and the accuracy higher, so as to make the wind power output prediction result better, and provide reference for the auxiliary decision-making of power grid security check and operation mode change.

References

[1] Xu Pengchao, Li Yan, Zhao Yanlei. Short-term wind power prediction based on BP neural network with adaptive mutation bat optimization algorithm. Electrical Measurement & Instrumentation, 2021,58(04):121-127.

[2] Yang Mao, Dai Bozhi, Liu Lei. A Review of Wind Power Probabilistic Prediction. Journal of Northeast Dianli University(Natural Science Edition), 2020,40(02):1-6.

[3] Peng Xiaosheng, Xiong Lei, Wern Jinyu, etc. A Summary of the State of the Art for Short-
term and Ultra-short-term Wind Power Prediction of Regions. Proceedings of the CSEE, 2016, 36(23): 6315-6326+6596.

[4] Zhang Hanchao, Kuang Honghai, Wang Jianhui, etc. Application of Error Correction Algorithm in Short-term Power Forecasting of Wind Power. The Journal of New Industrialization, 2019, 9(02): 1-9.

[5] Chen Hao, Zhang Jianzhong, Xu Chao, Tan Fenglei. Short-term wind power forecast based on MOSTAR model. Power System Protection and Control, 2019, 47(01): 73-79.

[6] Hu Yawei, Wang Xiao, Chao Qin, etc. A Forecasting Accuracy Improvement Method for Wind Power Based on Phase and Level Errors Translating and Interpolating Correction. Power System Technology, 2015, 39(10): 2758-2765.

[7] Li Jun, Zhuo Zeying. Wind Power Short-term Forecasting Based on Empirical Wavelet Transform and Multiple Kernel Learning. Information and Control, 2018, 47(04): 437-447.

[8] Xu Shanshan, Power system operation mode prediction and its application[D]. North China Electric Power University, 2006.

[9] Zhang Wei, Deng Yuanchang. Short-Term Wind Speed and Wind Power Prediction Based on the Grey-Markov Chain. Electric Power, 2013, 46(02): 98-102.

[10] Mei Yu, Wang Honglei, Wang Jin. Short-term Wind Power Prediction Simulation Based on Pso-bp Combined Improved Model. Computer engineering & Software, 2020, 41(12): 7-10.

[11] Yang Mao, Qi Yue, Mu Gang, etc. Wind power prediction based on adaptive neuro-fuzzy inference system. Electrical Measurement & Instrumentation, 2015, 52(14): 6-10.

[12] Meng Shengwei, Feng Xiaoxiao, Pang Jingyue, etc. Probabilistic Wind Power Prediction Based on LS-SVM and Kernel Density Estimation, Computer Measurement & Control, 2019, 27(12): 34-38.

[13] Yang Ping, Ye Chao. A Combination Forecasting Model for Short-term Wind Power Combination Based on Time Varying Probability Weight. Guangdong Electric Power, 2016, 29(10): 1-6.