Probabilistic modeling of output characteristics based on ECM algorithm for wind farms

Tingxiang Liu\textsuperscript{1}, Xiaoying Zhang\textsuperscript{1}, Kun Wang\textsuperscript{2}, Wei Chen\textsuperscript{1} and Xiaolan Wang\textsuperscript{1}

\textsuperscript{1}College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China
\textsuperscript{2}State Grid Gansu Electric Power Company Electric Power Research Institute, Lanzhou 730050, Gansu, China

Abstract: With the large-scale development of wind power, grasping the fluctuation characteristics of wind farm output has become a key link in wind power grid-connected operation. Because of the inherent defect in EM(expectation maximization) algorithm that is adopted to model wind farms, the fitting precision of WGMD(weighted Gaussian mixture distribution) is reduced. According to this problem, an improved probabilistic modeling method based on ECM(Expectation Constraint Maximization) algorithm is proposed in this paper. And several common distribution models are compared with simulate the output characteristics of wind farms. Simulation results show that the estimation of model parameters by the ECM algorithm can improve the simulation accuracy of the weighted Gaussian mixture model, that verifying the feasibility and effectiveness of the weighted Gaussian mixture probability model.

1. Introduction

Wind power has been greatly developed as a non-polluting renewable energy source in recent years. However, wind power generation has the characteristics of volatility, intermittence and randomness. Its large-scale interconnection will inevitably increase the uncertainty of the traditional power system operation, scheduling and risk assessment\textsuperscript{[1]}. Moreover, with the continuous increase in the scale of wind farms and the capacity of wind turbines, the proportion of wind power generation in the power grid is increasing, which inevitably brings many difficulties to the planning, design, security and stability analysis of power grids\textsuperscript{[2]}. Therefore, it is great practical significance to study the power output characteristics of the wind farm and establish an accurate probability model of wind power output, which can provide valuable basis for the planning of power system, risk analysis, reliability evaluation and economic dispatch\textsuperscript{[3]}.

There are two methods for probabilistic modeling of output characteristics of wind farms. One is based on a random sequence method, which uses the auto-regressive and moving average (ARMA) model in the time series method to simulate the change of the wind speed, and using matrix technology to generate multiple specified wind speed series to simulate the output-related characteristics of the wind farm\textsuperscript{[4-5]}. The other is based on the probability density function. Ding Ming uses Weibull distribution to model the wind turbine power model. According to the relationship between wind speed and wind power, wind power distribution characteristics can be obtained\textsuperscript{[6]}. References \textsuperscript{[7,8]} believes that using Beta distribution to describe wind power output prediction error is more reasonable and the fitting effect is better. However, the above single distribution function model cannot take into account the asymmetry distribution and multi-peak characteristics of the probability
distribution of prediction error.

Based on the above analysis, this paper starts with a large number of actual measured historical data from wind farms and establishes a weighted Gaussian mixture distribution (WGMD) for changes in wind power output. The method of estimating WGMD parameters is usually moment estimation method and maximum likelihood estimation method. Yan Hong adopts the traditional EM (Expectation Maximization) algorithm for solving unknown parameters of WGMD\cite{9}. However, the EM algorithm is mainly used to solve unknown parameters under incomplete data, and it is very easy to be affected by the initial conditions, resulting in the calculation result may not be able to converge to the global optimal value, which greatly affects the accuracy of the model\cite{10}. In order to overcome the deficiencies of EM algorithm, this paper uses ECM (Expectation Constraint Maximization) algorithm to estimate the parameters of WGMD.

2. Weighted gaussian mixture distribution

In recent years, Gaussian distribution has been used as a common tool in the statistical, computer and engineering fields for data analysis\cite{11}. WGMD is a model that is weighted by multiple complex data analyses in the case where a single Gaussian distribution cannot effectively represent complex data. It is not limited to a specific assumption of the probability density function form, and any probability density distribution can be approximated by the linear combination of several Gaussian density functions, and the probability density distribution of an arbitrary random variable can be accurately characterized\cite{12}.

The weighted Gauss mixture model is used to fit the probability density function of wind farm output as:

$$ f(x, \Theta) = \sum_{m=1}^{M} \alpha_m G_m(x / \theta_m) = \sum_{m=1}^{M} \alpha_m \frac{1}{\sqrt{2\pi \sigma^2_m}} e^{-\frac{(x-\mu_m)^2}{2\sigma^2_m}} \quad (1) $$

Where, $\Theta = \{\theta_m = (\alpha_m, \mu_m, \sigma_m), i = 1,2,...,M\}$ is model parameters, $M$ is the number of model parameters, $\alpha_m$, $\mu_m$, $\sigma^2_m$ are the weights, mean values and variances of the $m$ components of the weighted Gaussian mixture probability model respectively, and $X$ is measured data of wind farm output, furthermore, $\sum_{m=1}^{M} \alpha_m = 1$, $\alpha_m \geq 0$.

For the weighted Gaussian mixture model of wind farm output, the measured data of wind farm output is the observed data, which is called the incomplete data; Whereas each output power of a wind farm is unobservable, it is called lost data. Let $\omega \in \{1, 2,...,M\}$ denote the category of wind farm output, and the complete data is $X = \{x_i = (y_i, \omega_i), i = 1,2,...,N\}$.

3. Calculate unknown parameters by ECM algorithm

3.1 EM algorithm

EM algorithm is a method of estimating parameters that has developed rapidly and widely in recent years. It is an iterative method proposed by Dempster\cite{10}, which simplifies the calculation of maximum likelihood estimation. When maximization is not explicitly expressed, the generalized EM algorithm is given, which is GEM (General Expectation Maximization) algorithm. Meng et al. proposed a special GEM algorithm called ECM algorithm\cite{13}, which preserves the simplicity and stability of the EM algorithm.

The EM algorithm is mainly used to solve the problem of maximum likelihood estimation of unknown parameters under incomplete data. Assume that $Y_{obs}$ is the observation data, $Y_{mis}$ is the missing data, $\Theta$ is the unknown parameter, and the complete data is $Y = (Y_{obs}, Y_{mis})$. Let $f(\Theta | Y_{obs})$ denotes the posterior distribution density function of $\Theta$ based on observation data $Y$, $f(\Theta | Y_{obs}, Y_{mis})$ denotes the posterior distribution density function for $\Theta$ after adding data $Y_{mis}$, $f(Y_{mis} | \Theta, Y_{obs})$ denotes the conditional...
distribution density function of potential data $Y_{mis}$ under given $\theta$ and observed data $Y_{obs}$. The purpose of parameter estimation is to observe the mode of the posterior distribution $|\theta, Y_{obs}|$. The EM algorithm remembers $\theta^{(m)}$ as the estimated value of the posterior mode at the beginning of the $m$th iteration. Then the two steps of the $(m+1)$th iteration are the following:

**E step:** Impute the expectation of the $\log f(\theta | Y_{obs}, Y_{mis})$ conditional distribution of $Y_{mis}$, then remove $Y_{mis}$ through calculus:

$$Q(\theta | \theta^{(m)}, Y_{obs}) = \int \log(f(\theta | Y_{obs}, Y_{mis}))f(Y_{mis} | \theta^{(m)}, Y_{obs})dY_{mis} \quad (2)$$

**M step:** Determine $\theta^{(m+1)}$ by maximizing the imputed log-likelihood $Q(\theta | \theta^{(m)}, Y_{obs})$:

$$Q(\theta^{(m+1)} | \theta^{(m)}, Y_{obs}) = \max_{\theta} Q(\theta | \theta^{(m)}, Y_{obs}), \text{for all } \theta \in \Theta \quad (3)$$

After getting $\theta^{(m+1)}$, an iteration of $\theta^{(m)}$ to $\theta^{(m+1)}$ is formed. The above steps are iterated to

$$\|\theta^{(m+1)} - \theta^{(m)}\| \text{ or } \|Q(\theta^{(m+1)} | \theta^{(m)}, Y_{obs}) - Q(\theta^{(m)} | \theta^{(m)}, Y_{obs})\|$$

is small enough to stop.

### 3.2 ECM algorithm

The basic idea of the ECM algorithm is to use several simple conditions to maximize the steps instead of the complex M steps in the EM algorithm. Each step of the Conditional Maximization (CM) step is under the constraint of the corresponding parameter, so that the Q function defined by the EM algorithm reaches a maximum [14]. Let $G = \{g_s(\theta); s = 1, 2, ..., S\}$ be set of $S(\geq 1)$ preselected constraint functions for the parameter $\theta$. The parameter $\theta$ is divided into sub-vectors $\theta_1, \theta_2, ..., \theta_S$. With ECM, the M-step is replaced by $S$ CM steps in the $m$th iteration, marked as $(1) (2) \cdot \cdot \cdot (S)$ CM step.

**CM step:** Find $\theta_{s(S)}$ such that

$$Q(\theta_{s(S)} | \theta^{(m)}, Y_{obs}) = \max_{\theta} Q(\theta | \theta^{(m)}, Y_{obs}) \quad (4)$$

Where $\theta \in \Theta^{(m)} = \{\theta \in \Theta : g_s(\theta) = g(\theta^{(m)}); s = 1, 2, ..., S\}$, the next iterate $\theta^{(s+1)} = \theta_{s(S)}$ is obtained and an iteration is completed.

**E step:** Impute the expectation of the $\log f(\theta | Y_{obs}, Y_{mis})$ conditional distribution of $Y_{mis}$, then remove $Y_{mis}$ through calculus, as shown in formula (2);

**sth CM step:** Find $\theta^{(s(S)+S)}$ such that

$$Q(\theta^{(s(S)+S)} | \theta^{(m)}, Y_{obs}) = \max_{\theta} Q(\theta | \theta^{(m)}, Y_{obs}) \quad (4)$$

Where, $\theta \in \Theta^{(m)} = \{\theta \in \Theta : g_s(\theta) = g(\theta^{(m)}); s = 1, 2, ..., S\}$, the next iterate $\theta^{(s+1)} = \theta^{(s(S)+S)}$. The rational behind the CM steps is that in problems where maximizing $Q(\theta | \theta^{(m)})$ over $\theta \in \Theta$ is difficult, it may be impossible to choose $G$ so that it is simple to maximize over $\theta \in \Theta^{(m)}$ for $s = 1, 2, ..., S$.

### 3.3 WGMD parameter estimation

When the observed value $y = (y_1, y_2, ..., y_N)$ of the weighted Gaussian mixture model describing the wind farm output is given, the likelihood function of the probability distribution is:

$$L(\theta) = \sum_{n=1}^{N} \log \sum_{m=1}^{M} \alpha_m G_m(y | \theta_m) \quad (5)$$

**E step:** Impute the conditional expectation of a complete data likelihood function, $\theta^{(s)}$ is the parameter estimate of the $s$th iteration, $F(\omega | y, \theta^{(s)})$ is the posterior probability of $\omega$:

$$Q(\theta | \theta^{(s)}) = E[\log f(y, \omega | \theta) | y, \theta^{(s)}]$$

$$= \sum_{n=1}^{N} F(\omega | y, \theta^{(s)}) \log f(y, \omega | \theta) \quad (6)$$

**CM step:** Impute $\theta^{(s+1)} = (\theta_1^{(s+1)}, \theta_2^{(s+1)}, ..., \theta_S^{(s+1)})$ such that...
\[ Q(\theta^{(s+1)} | \theta^{(s)}) = \max_{\theta} Q(\theta | \theta^{(s)}) \quad (7) \]

1thCM: Calculate \( \theta^{(1)} = \arg \max_{\theta} Q(\theta | \theta^{(0)}) \) through a given \( \theta = \theta^{(0)}, ..., \theta_s = \theta^{(s)}; \)

2thCM: Calculate \( \theta^{(2)} = \arg \max_{\theta} Q(\theta | \theta^{(1)}) \) through a given \( \theta = \theta^{(1)}, \theta_j = \theta^{(j)}, j = 3, ..., s; \)

After (S-1)th calculations;

SthCM: Calculate \( \theta^{(s)} = \arg \max_{\theta} Q(\theta | \theta^{(s-1)}) \) through a given \( \theta = \theta^{(s-1)}, \theta_j = \theta^{(j)}, \theta_s = \theta^{(s)}; \)

According to Bayes rule\(^{[11]}\), such that:

\[ \omega_m = \frac{\alpha_m G_n(y_n | \theta_m)}{f(y_n | \theta_m)} \quad (8) \]

According to the method in [15], \( \alpha_m \) is solved as:

\[ \alpha_m = \frac{1}{N} \sum_{n=1}^{N} \omega_m \quad (9) \]

4. Evaluation index

Comparing the error parameter of a model with the actual fitting effect, there are root mean squared error (RMSE), the sum of squares due to error (SSE) and the coefficient of determination (R-square). The closer RMSE and SSE are to 0, the better the model fits the original data. And the R-square value range is \([0,1]\), the closer to 1, the better the interpretation of the original data, the better the model fitting effect.

RMSE: Also called the standard error, which reflects the error between the fitted data and the original data. Calculated as follows:

\[ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \quad (10) \]

SSE: The sum of the squared errors of the corresponding points of the fitting data and the original data. Calculated as follows:

\[ SSE = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 \quad (11) \]

R-square: Characterizes the fit of the data by changes in the data. Calculated as follows:

\[ R-square = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (12) \]

Where, \( \hat{y}_i \) is the sample data; \( y_i \) is the original data; \( \bar{y}_i \) is the average value of the original data.

5. Example illustrating ECM

In this paper, the probability density distribution characteristic of wind power fluctuation is obtained by measuring the output power data of wind farm in a certain area abroad. A probability model of a weighted Gaussian mixture model is established to compare the fitting effect of other single distribution function models and verification of its effectiveness in the study of wind power characteristics. The EM algorithm and ECM algorithm were used to estimate the parameters of the Gaussian mixture model, and the ECM algorithm was verified the accuracy of the model.

According to the wind farm power distribution and probability density distribution curve, Weibull distribution, \( t \) location-scale distribution and normal distribution are used to fit wind power fluctuation characteristics. The fitting effect diagram is shown in Figure 1.
As can be seen from Figure 1, the single distribution function cannot accurately fit the "tailing" property of the probability density distribution characteristics of wind power fluctuations, and the fitting effect is poor. Therefore, when a single wind farm has fluctuating random characteristics, the output fluctuation characteristics of the wind farm group will no longer satisfy a single specific distribution function curve.

For WGMD modeling, target distributions are usually modeled using no more than 5th-order models. Estimate the probability distribution of wind power output based on measured data, and optimize the fitting effect by increasing the number of Gaussian distributions. This paper uses the 2nd and 5th order WGMD to get the wind power probability density distribution curve, as shown in Figure 2.

As can be seen from Figure 2, the 2nd-order and the 5th-order WGMD can well fit the probability density distribution curve of wind power fluctuations, and with the increase of the weighted order, the fitting accuracy is also improving. Table 1 gives the index values of the evaluation of the fitting effect to quantify the fitting effect of the two models and the commonly used distribution functions.

| Fitting model     | \(I_{\text{RMS}}\) | \(I_{\text{MSE}}\) | \(R^2\)    |
|-------------------|------------------|-----------------|-----------|
| t distribution    | 0.4              | 25.05           | 0.57      |
| Normal distribution | 0.8              | 21.67           | 0.63      |
| Weibull distribution | 0.25             | 12.95           | 0.88      |
| 2nd-order WGMD    | \(1.8 \times 10^{-3}\) | \(5.1 \times 10^{-4}\) | 0.92      |
| 5th-order WGMD    | \(5.2 \times 10^{-4}\) | \(3.6 \times 10^{-5}\) | 0.98      |

As shown in Table 1, the \(I_{\text{RMS}}\) values of the three commonly used distribution functions are all large, \(I_{\text{MSE}}\) is much larger than 0, and the fitting error is relatively large. In comparison, the \(R^2\) of
Weibull distribution reaches 0.88, and the fitting effect is better, and $I_{SSR}$ and $I_{BSE}$ are the minimum among the three distributions. Comparing the $I_{SSR}$ and $I_{BSE}$ index values of WGMD are close to 0, $R^2$ index values are all above 0.9, in which the $R^2$ value of fifth-order WGMD reaches 0.98. The above analysis shows that when characterizing wind power based on probability density function, WGMD is more suitable for fitting wind power fluctuation characteristics, and the fifth-order WGMD fitting effect is better.

In the parameter estimation, because the EM algorithm calculates the fifth-order model parameters, the parameter estimation error is large, the local optimization results are obvious, and it is easy to appear in the iteration loop. Therefore, the EM algorithm and ECM algorithm are used to compare the parameters of the third-order WGMD. The probability density function of wind power based on the third-order Gaussian mixture model is shown in formula (13). The parameters are estimated as shown in Table 2.

$$f(x) = \alpha_1 \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} + \alpha_2 \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} + \alpha_3 \frac{1}{\sqrt{2\pi}\sigma_3} e^{-\frac{(x-\mu_3)^2}{2\sigma_3^2}} \quad (13)$$

Figure 3 uses the probability density distribution of the model determined by the EM algorithm and the ECM algorithm respectively.

![Figure 3. EM algorithm and ECM algorithm fitting distribution curve](image)

Table 2. Wind power probability model parameter estimation

| Model parameters | $\alpha_1$ | $\alpha_2$ | $\alpha_3$ | $\mu_1$ | $\mu_2$ | $\mu_3$ | $\sigma_1$ | $\sigma_2$ | $\sigma_3$ |
|------------------|------------|------------|------------|--------|--------|--------|----------|----------|----------|
| EM algorithm     | 0.27       | 0.51       | 0.19       | 33.67  | 25.06  | 18.79  | 5.06     | 10.93    | 3.22     |
| ECM algorithm    | 0.26       | 0.55       | 0.22       | 29.77  | 28.07  | 15.66  | 5.37     | 12.61    | 3.18     |

As shown in Figure 3, the probability distribution of WGMD based on ECM algorithm and actual wind power is more consistent and the residual is smaller. The residuals at 0-20 MW based on the EM algorithm are larger, and the wind farm power is mostly concentrated here. Therefore, the EM algorithm has a much poorer overall ECM fitting effect. The evaluation index values of the two algorithms are shown in Table 3.

| Fitting model | $I_{SSR}$ | $I_{BSE}$ | $R^2$ |
|---------------|----------|----------|------|
| EM algorithm  | $4.2\times10^{-4}$ | $5.5\times10^{-5}$ | 0.95 |
| ECM algorithm | $3.6\times10^{-4}$ | $1.3\times10^{-5}$ | 0.99 |

The EM algorithm is used to estimate the unknown parameters, which can easily be affected by the initial conditions, so that the calculation results can not converge to the global optimal value. Instead, the ECM algorithm uses conditional maximization steps to replace the maximum step of EM, so that every step of the maximization is under the corresponding parameter constraints, so that the Q function defined by the EM algorithm reaches the maximum, the optimization parameters of the model can be accurately evaluated.
6. Conclusion
In this paper, the probability density function of the WGMD fitted wind power output fluctuation characteristics is established. The parameter estimation of WGMD is carried out by ECM algorithm, and a wind farm is used as an example to verify. The concrete results are as follows:

(1) With the large-scale development of wind power, the volatility of wind farm output has become a key link in wind power grid operation. This paper analyzes the measured data of a wind farm in a foreign country, establishes WGMD and compares traditional modeling methods. It is verified that WGMD is very suitable for describing wind farms with randomness and volatility, and significantly improves the simulation accuracy of wind farm output characteristics.

(2) For the parameter estimation of the weighted Gaussian mixture distribution model, the commonly used EM algorithm is vulnerable to the initial value and falls into the local optimum. This paper used the ECM algorithm to replace the one maximization of the EM algorithm by multiple condition maximization. The disadvantages of the EM algorithm are avoided, so that the final parameter estimation converges to the global optimum and the model accuracy is guaranteed.

Acknowledgements
The authors also want to give their thanks to the Natural Science Foundation of China.
(No1. 51867015, No2.51767017)

References:
[1] Y S Xue, X Lei, F Xue, et al. A review on impacts of wind power uncertainties on power systems[J]. PCSEE, 34(29), 5029-5040(2014).
[2] Y Qiao, Z X Lu. Wind farms active power control considering constraints of power grids[J]. AEPS, 33(22), 88-93(2009).
[3] H F Liang, D W Cao, B Liu, et al. Modeling and error analysis of wind farm probability distribution model[J]. Journal of North China Electric Power University, 44(3), 8-14(2017).
[4] S Y Xie, X L Wang, Y Qu, et al. Influence of different wind speed simulation methods on power system reliability assessment under the framework of segmentation multi-objective risk analysis[J]. PCSEE, 33(31), 81-90(2013).
[5] R. Billinton. Time-series models for reliability evaluation of power systems including wind energy[J]. MICROELECTRON RELIAB, 36(9), 1253-1261(1996).
[6] M Ding, Y C Wu, L J Zhang. Study on calculation method of wind speed probability distribution parameters in wind farms[J]. PCSEE, 25(10), 107-110(2005).
[7] H Bludszuweit, J A Dominguez-Navarro, A Llombart. Statistical analysis of wind power forecast error[J]. IEEE T POWER SYST, 23(3), 983-991(2008).
[8] A Fabbri, T Román, J Abbad, et al. Assessment of the cost associated with wind generation prediction errors in a liberalized electricity market[J]. IEEE T POWER SYST, 20(3), 1440-1446(2005).
[9] H Yan, L C Sun, X Q Chang. The probabilistic modeling of the fluctuation characteristics of wind power output in Xinjiang [J]. POWER SYSTEM TECHNOLOGY, 38(6), 1616-1620(2014).
[10] A P Dempster, N M Laird, D B Rubin. Maximum likelihood from incomplete data via the EM algorithm[J]. J R STAT SOC, 39(1), 1-38(1977).
[11] H X Gao. Applied Multivariate Statistical Analysis[M]. Beijing: Peking University Press(2005).
[12] Y Zhao, J Wang Y J Xiong, et al. Gaussian mixture model of random variables in power system reliability assessment[J]. AEPS, 40(1), 66-71,80(2016).
[13] X L Meng, D B Rudin. Maximum likelihood estimation via the ECM algorithm: A general framework[J]. BIOMETRIKA, 80(2), 267-278(1993).
[14] X L Meng, D B Rudin. Maximum likelihood estimation via the ECM algorithm: Computing the asymptotic variance[J]. STAT SINICA, 55-75(1995).
[15] D M Titterington, A F M Smith. *Statistical analysis of finite mixture distribution*[M]. New York: John Wiley & Sons, 179-191(1985).