Research on Modeling Method of Wind Power Fluctuation against Instability of Wind Power Generation

YaJun Rong, Ji Qi*

Institute of Electrical Engineering, Vanshan University
Qinhuangdao, Hebei Province, China

Abstract — Wind energy is a kind of clean and renewable energy, and wind power generation is the main form to utilize wind energy. Due to the instability of wind power generation, how to quantitatively describe wind power fluctuation has been a research hotspot at home and abroad. Based on this, this paper analyzes a large number of measured data aimed at the volatility series of wind power with 5 seconds' interval, and taking four groups of wind turbines as examples, it analyzes the situation of wind power fluctuation by several common probability distribution models and conducts optimal fitting using MATLAB software. The optimal probability distribution model is also used to respectively calculate the characteristic value of the four groups of wind turbines and the similarities and differences are analyzed. The optimal model location-scale is adopted to calculate the probability distribution parameters in 30 periods of the four wind power, and statistical models are used to make the probability parameter line chart of wind power fluctuation of different groups of wind turbines in different periods. The experimental results show that the modeling method aiming at wind power fluctuation can make short-term forecasting of wind speed of the wind farm and improve the accuracy of prediction.

Keywords - wind energy; wind power; fluctuation; modeling

I. INTRODUCTION

With energy and environmental problems become increasingly prominent, wind power as a clean renewable energy has been paid more and more attention. According to the statistic report of China Wind Energy Association, the installed wind capacity of China’s mainland has reached 75.3GW by the end of 2012. The newly issued China’s 12th five-year plan of renewable energy development and wind electricity development proposes that by 2017, China’s installed wind power integration will reach 100 million kilowatts and 200 million kilowatts in 2022.

Wind power generation is intermittent, variable and unpredictable, which makes its access threaten the reliability of the power grid [1]. Ultrashort fluctuations of wind power (a few minutes) have certain influence on the wind turbines; short-term fluctuations (a few hours or days) have certain influence on power grid dispatching; medium and long term fluctuations (weeks or months) have certain impact on wind farms or power grid maintenance plan [2]. Fluctuation of wind power output, if not controlled, will cause fluctuation and flickering of the network voltage, so many scholars are devoted to the study of fluctuation of wind power output [3]. At present, energy storage devices (pumped storage power plant, superconductor or flywheel) are used more in solving fluctuation of wind power output. The fluctuation characteristic of wind electricity is very important for the selection of capacity of energy storage devices and the formulation of control strategy [4]. The power of the wind turbine generators is mainly related to wind speed, so the uncertainty of wind makes it impossible for wind turbines to determine power generation according to the demands like conventional generators [5]. A wind farm usually has dozens or hundreds of wind turbines, and large wind power base is composed of dozens or even hundreds of wind farms. Therefore, the fluctuation of wind power has strong spatial and temporal differences, and large-scale wind power base usually needs access to the power grid to realize the transmission and consumption of wind power [6]. The random fluctuation of wind power is considered to be the main factor that brings adverse effects to the power grid [7].

Combined with the research status of wind power fluctuation characteristics, this paper summarizes the different quantitative indexes and classifies the methods used in research of wind power fluctuation. The wind fluctuation (i.e. from the trough to the crest and then to the trough) is adopted to study the distribution law of the power i of wind turbines. The dfittool of MATLAB software is used to conduct curve fitting of the data and determine the most consistent probability distribution. The problems needing to notice in the future research of wind power fluctuation are also put forward.

II. STATE OF THE ART

With increasing capacity of wind power grid integration, people are more concerned about the possible impact of the randomness of wind power. In fact, uncertainty and variability are inherent attributes of power system operation. Over the past century, the power system infrastructure construction and improvement operation level themselves are a continuous response to the uncertainty and variability [8]. Wind power access increases the uncertainty and
variability of the power system, but it does not change the two essential attributes of it.

The research of Alani et al in 2014 shows that the rapid changes of wind power, especially the minimum power changes, are mostly absorbed by mechanical inertia, thermal and hydraulic inertia and control deadband of the conventional hydro and thermal generating units, and that not all the wind power fluctuation will influence the power system [9]. Usually two methods are used to describe randomness of wind power: one is the method based on random sequence; Zafra et al respectively use the theory of Markov chain Monte Carlo and average model of band limited amplitude to conduct modeling for wind power sequence [10]. The other is Ge and others’ method based on probability density function [12]. In 2012, Wharton et al. studied the distribution characteristics of wind speed, wind power, wind turbines current harmonics and wind power prediction error [13]. It is generally believed that wind speed satisfies Weibull distribution. Provided that wind speed meets the Weibull distribution, we can get the distribution characteristics of wind power through the relationship between wind speed and wind turbine output power; but a large amount of measured data show that wind power does not meet the common probability distributions [14]. Vera et al, using Weibull distribution and normal distribution, fitted the probability density function of the different types of wind turbine current harmonics and described the prediction error of wind power prediction system based on neural network by normal distribution, but the probability density curve was greatly different from frequency distribution histograms [15]. The research of wind power fluctuation can lay the foundation for wind power prediction, and also provide possibility for the evaluation of system energy storage.

At present, although some researchers and institutions have carried out research on the characteristics of wind power, the conclusions of wind power characteristic analysis cannot be applied in power system planning, operation or control due to the inaccessibility to data of wind power and mature methods. The current research is aimed at the fluctuation of wind power with most of them being qualitative analysis, and there are many other features need to be explored except fluctuations of wind power. Therefore, it is urgently needed to quantitatively analyze the characteristics of wind power by mathematical methods.

III. METHODOLOGY

A. Experimental method

Wind power is a typical random power supply, so the sampling time directly affects the accuracy of the analysis results when analyzing the characteristics of wind power output. In addition, the length of sampling time also directly affects the speed of engineering applications. Random fluctuations of wind power are considered to be the main factors that bring adverse effects on the grid. To conduct modeling for wind power fluctuation against instability of wind power generation, we randomly choose four wind turbines to analyze what kinds of probability distribution the wind power fluctuation $P_{t}^{5s}(t_{j})$ of the turbine i conforms to within 30 days. The values are respectively calculated and verified, and similarities and differences of the four turbines are compared. Finally, the best probability distribution is determined to calculate the probability distribution parameters of the four wind power in 30 periods with a day as the width of time window and then it is examined. We try to compare wind power fluctuation probability distribution of different units and different periods and figure out its relationship with the overall distribution in the 30 days, thus researching on modeling method.

B. Model analysis

Within the 30 days, the probability distribution of wind power fluctuation $P_{t}^{5s}(t_{j})$ of unit i is analyzed, and fluctuation quantity y is chosen as the quantization index of wind power fluctuation, its formula is:

$$Y_{j} = \bar{X}_{i+1} - \bar{X}_{i}, \quad i=1,2, \ldots, q-1 \quad (1)$$

Where q is the ratio of the total observation time and observation intervals, and the observation interval in question one is 5 second. For the selection of wind turbines, we filter the five sets of data with the least data loss with EXCEL software, which are respectively the 7, 9, 11, 13 set of wind turbines. For the selection of data, MATLAB software is firstly used to conduct curve fitting for data in day 1 to 3 of the unit 7, and the results are shown in figure 2.

![Figure 1](image.png)  
**Figure 1** wind power generation

![Figure 2](image.png)  
**Figure 2** Seventh groups of wind power 1–3 curve fitting

As can be seen from Figure 2, the power of wind turbines varies with the change of wind speed and constantly fluctuates with time. So we choose a complete
one in the wind velocity fluctuations to be studied, which means to select a section of data from trough to crest and then to trough in the attachment. Then curve fitting is conducted on selected data, and MATLAB probability density dfit tool is used to obtain the power probability histogram and probability distribution map of normal distribution, t distribution, log-logistic distribution and Weibull distribution of the four turbines. And then a day is taken as the width of time window to calculate and examine the probability distribution parameters of the four wind power in 30 periods. We also compare wind power fluctuation probability distribution of different units and different periods and figure out its relationship with the overall distribution in the 30 days.

C. Probability distribution of the four sets of turbines

Figure 3 Probability distribution graph of seventh units

The four kinds of probability distribution of unit 7 are distribution with approximate data. The numeric feature of the four kinds of probability distribution can be calculated by MATLAB (see table 1).

<table>
<thead>
<tr>
<th></th>
<th>Log-Logistic</th>
<th>Normal distribution</th>
<th>Weibull</th>
<th>t location-scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean value</td>
<td>3280.12</td>
<td>3530.58</td>
<td>3541.15</td>
<td>3539.81</td>
</tr>
<tr>
<td>variance</td>
<td>2633210</td>
<td>2188350</td>
<td>2431810</td>
<td>3202530</td>
</tr>
<tr>
<td>log likelihood</td>
<td>-23365300</td>
<td>-23403700</td>
<td>-2326500</td>
<td>-23111600</td>
</tr>
</tbody>
</table>

From the data of table 1, it can be seen that the numeric feature of the four kinds of probability distribution is not quite different, the variance of normal distribution is the smallest, and that the value of log-likelihood function of the t location-scale distribution is the largest. We can also see from the graphic that the fitting effect of t location-scale is the best, so probability distribution of unit 7 is recommended as the distribution for t location-scale.

Figure 4 Probability distribution graph of ninth units

The probability distribution of unit 9 is consistent with the distribution of t location-scale, normal distribution, gamma distribution and Weibull distribution, and their numerical feature can be calculated through MATLAB software (See Table 2).
From Table 2, we can see that the variance of Gamma distribution is the smallest and that its log-likelihood function value is also the largest. It can also be seen from the probability distribution diagram that Gamma distribution curve is the closest to the fitting curve of the data. Therefore, unit 9 is recommended for Gamma distribution.

The probability distributions unit 11 is mostly consistent with Normal distribution, Weibull distribution, Logistic distribution and Birnbaum-Saunders distribution. MATLAB software is used to calculate its numerical feature (See Table 3).

From Table 3, it can be seen that the variance of Logistic distribution is the smallest and its log-likelihood function value is the largest. We can also see from the probability distribution diagram that the probability curve of Logistic distribution best match the fitting curve, so unit 11 is recommended for Logistic distribution.

The probability distributions unit 13 is more consistent with t distribution, normal distribution, Weibull
YaJun Rong et al: RESEARCH ON MODELING METHOD OF WIND…

distribution and logistic distribution, and the results calculated by MATLAB are shown in Table 4.

| Table 4 PROBABILITY DISTRIBUTION NUMERICAL CHARACTERISTIC TABLE FOR THIRTEENTH UNITS |
|-----------------|----------------|----------------|----------------|
|                  | t location-scale | Normal distribution | Weibull | Logistic |
| mean value       | 3233.77          | 3491.98          | 3497.22 | 3325.61 |
| variance         | 2418190          | 2007170          | 2096330 | 1917460 |
| Log likelihood   | -24581100        | -24662200        | -24560600 | -24596200 |

Table 4 shows that the variance of Logistic distribution and the log-likelihood function value of Weibull distribution are the smallest. Combined with the probability distribution in figure 5, we recommend unit 13 for Weibull distribution.

| Table 5 THE PROBABILITY DISTRIBUTION OF THE FOUR UNITS MOST ACCORD WITH |
|-----------------|----------------|----------------|----------------|
| Unit 7          | Gamma distribution | Logistic distribution | Weibull distribution |
| Unit 9          | Normal distribution | Weibull distribution | Logistic distribution |
| Unit 11         | Normal distribution | Logistic distribution | Weibull distribution |
| Unit 13         | Normal distribution | Logistic distribution | Weibull distribution |

Through the above analysis, we get the most suitable probability distribution of the four units (see Table 5).

**D. Modeling**

Normal distribution (N, $\sigma^2$): Let the probability density of continuous random variable $X$ be:

$$f(x) = \frac{1}{\sqrt{2\pi \sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (2)$$

Distribution function is:

$$F(x) = \frac{1}{\sqrt{2\pi \sigma}} \int_{-\infty}^{x} e^{-\frac{(y - \mu)^2}{2\sigma^2}} dy \quad (3)$$

Where $\mu$ is the mean value, and $\sigma^2$ is the variance.

T location-scale distribution: the random variable $T = \frac{x - \mu}{\sigma}$

$$T \sim \text{T location-scale} \quad (4)$$

It conforms to the degree of freedom $n$, and its distribution density function is:

$$f(x) = \frac{(n+1)^{1/2}}{\sigma \sqrt{n\pi}} \left[ n + \left( \frac{x - \mu}{\sigma} \right)^2 \right]^{-n/2} \quad (5)$$

Extreme value distribution: the probability density function of the extreme value distribution is described by the following formula:

$$f(x) = \frac{1}{\beta} \exp \left( -\frac{x - \alpha}{\beta} - \exp \left(-\frac{x - \alpha}{\phi} \right) \right) \quad (6)$$

Logistic distribution: the value of probability $p$ is between 0 and 1, and the relationship between $P$ and the independent variable $X$ is set up as:

$$p = \pi(x) = \frac{1}{\beta} e^{-\left(\frac{x - \alpha}{\beta}\right)} \exp(-\frac{x - \alpha}{\phi}) \quad (7)$$

Logistic regression function:

$$\pi(x) = \frac{\exp(\alpha + \beta x)}{1 + \exp(\alpha + \beta x)} \quad (8)$$
most consistency with the actual distribution law is found, that is the \( t \) location-scale curve.

After calculation, the characteristic values of each curve are shown in table 6:

<table>
<thead>
<tr>
<th></th>
<th>( t ) location-scale</th>
<th>Logistic</th>
<th>Normal distribution</th>
<th>Distribution of extreme value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu )</td>
<td>-0.0499</td>
<td>-0.0623</td>
<td>7.71\times10^{-5}</td>
<td>4.6801</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>4.0952</td>
<td>5.1093</td>
<td>10.1688</td>
<td>22.8389</td>
</tr>
<tr>
<td>( V )</td>
<td>1.3881</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Where \( t \) location-scale distribution has three characteristic values \( \mu, \sigma \) and \( V \); other distributions have only two characteristic values \( \mu \) and \( V \).

Table 7 shows the fit index of \( t \) location-scale distribution is the smallest, which means \( t \) location-scale distribution is best suited to describe the distribution characteristics of wind power and reflects the correctness of the calculation of the above characteristic values. The fitting index reflects 98% of significance level. \( t \) location-scale distribution is used to fit the probability density function of the four units. Its results are shown in table 8:

<table>
<thead>
<tr>
<th></th>
<th>( t ) location-scale</th>
<th>Logistic</th>
<th>Normal</th>
<th>Extremum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitting index ( I )</td>
<td>0.0220</td>
<td>0.0384</td>
<td>0.0447</td>
<td>0.1047</td>
</tr>
</tbody>
</table>

Table 2 shows that the fit index of \( t \) location-scale distribution is the smallest, which means \( t \) location-scale distribution is best suited to describe the distribution characteristics of wind power and reflects the correctness of the calculation of the above characteristic values. The fitting index reflects 98% of significance level. \( t \) location-scale distribution is used to fit the probability density function of the four units. Its results are shown in table 8:

<table>
<thead>
<tr>
<th></th>
<th>Seven fan</th>
<th>Nine fan</th>
<th>Eleven fan</th>
<th>Thirteen fan</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mu )</td>
<td>5.60e-6</td>
<td>-0.4993</td>
<td>-0.0267</td>
<td>0.0036</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>6.99e-6</td>
<td>4.0952</td>
<td>7.7011</td>
<td>14.4201</td>
</tr>
<tr>
<td>( V )</td>
<td>0.1337</td>
<td>1.3881</td>
<td>1.5341</td>
<td>1.5671</td>
</tr>
</tbody>
</table>

The results show that different units have different characteristic values, but the three characteristic values are in line with the law of \( t \) location-scale probability density function. Therefore, though different characteristic values reflect the spatial difference of wind power characteristics, the power fluctuations caused by temporal difference are consistent with the \( t \) location-scale distribution, which further proves that it is right to describe the distribution characteristics of wind power with \( t \) location-scale distribution.

Though the above \( t \) location-scale probability distribution, we calculate the probability distribution parameters in 30 periods of the four wind power with a day as the width of time window. 120 groups of data are acquired, which are placed in the appendix. We make the fold line of probability distribution parameters of wind power fluctuation of different units (space) and in different periods (time) using these data.

The characteristic value of the same unit in the 30 periods is constantly changing, which reflects the uncertainty and randomness of wind power fluctuation. The characteristic value of different units in the same period is also different, which is because of the influence of spatial difference: the difference in wind characteristics in different space leads to the difference in probability characteristic values. However, different curves have the same trend, and the distribution and degree of monotone of extreme points are roughly the same, so it is relevant to some extent.

Taking \( V \) as example, the characteristic values of 30 days’ overall distribution of four turbines are 0.1337, 1.3881, 1.5341 and 1.5671. The characteristic values of each period are reflected on the overall value through time effects, making their changes lead to the changes of overall values. Therefore, the power fluctuations of each period are linked with the overall changes to influence each other.

This shows that the observation time can be shortened to save manpower and material resources, making its characteristic value fluctuate within the range of accuracy. This also shows that the record period of power change may be extended appropriately: besides seconds, minutes can also be tried as unit steps. This will solve the problem that the second-level data of all wind power are difficult to be centrally recorded due to limitations of data storage and management while the accuracy meets the requirements.

V. CONCLUSIONS

As wind is formed in air flow caused by air pressure difference, the wind direction and strength are changing. Therefore, wind power generation is fluctuant, intermittent and random. To solve this problem, based on the analysis of a large amount of measured data, this paper adopts several common probability distribution models to analyze the wind power fluctuation, finding that \( T \) location-scale model can be used to describe the probability distribution of wind power fluctuations. The analysis results show that \( t \) location-scale probability distribution model, with good fitting effect, can describe the wind power fluctuations pretty well and that it can take into consideration the average change process and the corresponding factors with strong predictive ability and accurate results. Meanwhile, the requirement for data processing capability is demanding; it can find out the result with external help, being strongly theoretical. The actual measurement proves that modeling on wind power fluctuations against wind electricity instability can conduct short-term forecasting of the wind strength and speed in the wind farms by physical simulation and scientific statistical methods, thus satisfying the requirements of wind power dispatching department for wind electricity dispatch and promoting the development of wind energy utilization.
REFERENCES


