Global Energy Interconnection
Development and Cooperation Organization

ELSEVIER
Production And Hosting By Elsevier On Behalf Of KeAi
KeAi

Volume 3 Number 3 June 2020 (247-258)
DOI: 10.14171/j.2096-5117.gei.2020.03.006
Global Energy Interconnection
Contents lists available at ScienceDirect
https: //www.sciencedirect.com/journal/global-energy-interconnection

Full-length article

Probability distribution of wind power volatility based on the moving average method and improved nonparametric kernel density estimation

Peizhe Xin[1], Ying Liu[1], Nan Yang[2,3], Xuankun Song[1], Yu Huang[2]

1. State Grid Economic and Technological Research Institute Co. Ltd., State Grid Office District, North District of Future Science and Technology City, North Seven Changping District, Beijing 102209, P. R. China
2. Department of Hubei Provincial Collaborative Innovation Center for New Energy Microgrid, China Three Gorges University, Yichang, Hubei Province 443000, P. R. China
3. Stevens Institute of Technology, Hoboken, NJ 07030, USA

Scan for more details

Abstract: In the process of large-scale, grid-connected wind power operations, it is important to establish an accurate probability distribution model for wind farm fluctuations. In this study, a wind power fluctuation modeling method is proposed based on the method of moving average and adaptive nonparametric kernel density estimation (NPKDE) method. Firstly, the method of moving average is used to reduce the fluctuation of the sampling wind power component, and the probability characteristics of the modeling are then determined based on the NPKDE. Secondly, the model is improved adaptively, and is then solved by using constraint-order optimization. The simulation results show that this method has a better accuracy and applicability compared with the modeling method based on traditional parameter estimation, and solves the local adaptation problem of traditional NPKDE.

Keywords: Moving average method, Signal decomposition, Wind power fluctuation characteristics, Kernel density estimation, Constrained order optimization.

1 Introduction

In view of the continuous growth of the wind power access capacity, the fluctuation and randomness of the generated output power will introduce new challenges to the dispatching, planning, and safe and stable operation of power systems. Therefore, approaches needed to reduce the negative impact of grid-connected wind power on the power system have become a research hotspot. One of the effective ways to solve the aforementioned adverse effects is to analyze the wind power fluctuation characteristics (WPFC) and accurately establish its probability distribution model. This approach may also provide more safe and economical decision-making means for dispatch and planning agencies [1–3].

At present, scholars have conducted extensive research on WPFC modeling [4–9]. In [4], the author defined three

Received: 15 March 2020/ Accepted: 3 April 2020/ Published: 25 June 2020

Peizhe Xin
xinpeizhe@chinasperi.sgcc.com.cn

Ying Liu
liuying@chinasperi.sgcc.com.cn

Nan Yang
ynyyayy@ctgu.edu.cn

Xuankun Song
songxuankun@chinasperi.sgcc.com.cn

Yu Huang
357831347@qq.com

quantitative indicators based on the measured wind power output data, and analyzed WPFC from the aspects of the differential time scale and installed wind farm capacity. As the research progressed, some researchers began to use a mixed distribution of multiple functions as the prior distribution of wind power fluctuations. In [5], the author proposed a method of fitting WPFC with a Gaussian mixture model instead of a single distribution function model based on the measured data of a wind farm group. In [6], the author analyzed a large number of measured data of wind farm output, and used various probability distribution models to simulate the WPFC of wind farms. In [7], the authors proposed a hybrid t location–scale distribution model to describe the probability distribution of WPFC. In [8], the authors proposed a probability distribution model of a finite mixed Laplace distribution based on a large number of measured wind farm data. In [9], the authors proposed a hybrid Logistic distribution model to quantitatively describe WPFC, and validated it with measured data from a wind farm. Overall, prior publications used the theoretical distributions of one or more functions to fit the distribution curve of the wind power fluctuation index as the prior distribution, and then used historical data to estimate its parameters. Nevertheless, the universal applicability of the constructed model has been proven difficult to guarantee. Wind farms in China are extensively distributed throughout the country. Additionally, the WPFC in different regions may follow different forms of probability densities, and the modeling method based on parameter estimation is difficult to guarantee its universal applicability.

In view of this, some studies have proposed the modeling of WPFC with nonparametric estimation [10–14]. Compared with the parameter estimation method, this method directly builds a probabilistic model based on historical sample data, and does not need to make a prior judgment on the standard form of wind power fluctuation. Therefore, it is of great advantage to use nonparametric kernel density estimation (NPKDE) methods when it is not clear which standard parameter form the probability density of the wind power fluctuations will have. In [10], the authors first applied a statistical approach and the method of NPKDE to construct a distribution function of wind power prediction errors. Based on this, they established a benefit model based on wind farm wind energy loss and energy storage device planning considerations. In [11], the authors predicted the wind power based on the actual power curve, and then combined the distribution characteristics of the prediction errors to predict the error distribution characteristics using NPKDE. In [14], the authors first used the hybrid Copula function to characterize the correlations between multiple failure modes. Based on these, they constructed the corresponding probability density function and cumulative distribution function based on the NPKDE method. However, if NPKDE was used to model the WPFC, problems still persisted. Accordingly, the bandwidth is the only parameter that needs to be determined in the NPKDE model. Traditional NPKDE often uses the same bandwidth for all samples in a model. However, owing to the large wind power samples, different samples may be adapted to different bandwidths. Using only one bandwidth will lead to a low-degree of fitness at a certain point or in a certain interval of the sample, which is the problem of so-called low-local fitness.

In summary, this study first uses the moving average method to extract the wind power fluctuations. It then uses an improved NPKDE method to model the extracted results. Subsequently, a constraint-order optimization algorithm is used to solve the model. The main research and innovation aspects of this study are as follows: 1. an improved NPKDE strategy is proposed, whereby the geometric distance is used as an index to identify the sample interval with a lower fitting accuracy, and its bandwidth is modified; 2. a method is proposed to model WPFC based on NPKDE. The advantages are: 1. the fixed, single bandwidth becomes adaptively adjusted by the sample interval, and the variable bandwidth improves the modeling accuracy of NPKDE; 2. the probability characteristics of the wind power fluctuation are modeled directly based on sample data, and there is no need to judge which standard distribution form it obeys. Therefore, it has higher accuracy and applicability. Finally, simulation results based on actual data obtained from a wind farm in Hubei have verified the correctness and effectiveness of the proposed modeling method.

2 Extraction of wind power fluctuation based on the method of moving average

Wind power contains continuous and minute components [15–17]. The continuous components undergo smaller fluctuations, exhibit longer fluctuation periods, and have minor influences on the stability of the power system and the accuracy of wind power forecasting. However, the minute component is characterized by a large fluctuation and a short-fluctuation period. When the installed capacity of wind power is large, the component will seriously affect the power quality of the power system. At present, the low-pass filters, wavelet decomposition, and empirical mode decomposition, are used to extract the momentum of the wind power wave. However, these methods cannot use the time scale of the data to decompose the signal. In addition, some problems are associated with these methods, such as the boundary effect [18–20].

Therefore, according to the basic concept of the algorithm for the separation of load components in [21], this study uses the method of moving average to eliminate the wind power fluctuation. First, a time window is selected with a specific length, and the arithmetic average of all the values within the window is estimated. The value of the center point of the window is then replaced with the estimated average value. Finally, the window is moved backward by the dot pitch. The aforementioned steps are repeated. Suppose the length of the moving average is K min. If K is an even number, then the minute component P_{Ht} and continuous component P_{Lt} of the wind power at time t are calculated using (1).

$$\begin{cases} P_{Lt} = (P_{t-(K/2-1)} + P_{t-(K/2-2)} + ... + P_t + ... + P_{t+K/2}) / K \\ P_{Ht} = P_t - P_{Lt} \\ t = \dfrac{K}{2}, \dfrac{K}{2}+1, ..., J - \dfrac{K}{2} \end{cases} \tag{1}$$

where P_{Lt} is the continuous component, P_{Ht} is the minute component, which is the amount of change superimposed on the continuous component, P_t is the measured average power at t min, t is the measurement point, and J is the total number of measurement points.

The length K of the moving time window is randomly selected, and its selection directly affects the extracted fluctuation components. If it is too long, the change trend of wind power will be reflected on the minute component P_{Ht}, and the minute component will no longer be a random variable. If it is too short, the fluctuation of wind power will be reflected on the continuous component P_{Lt}. In general, this can be selected according to experience, that is, 20 min is more suitable for general loads, and 30 min is suitable for large impact loads. Combined with the actual situation, this study selected the period of 20 min as the moving average sampling period [22].

3 Modeling of wind power fluctuation probability density based on improved NPKDE

3.1 Wind power fluctuation probability density model based on NPKDE

It is assumed that x_1, x_2, ..., x_n is an n-sample sequence denoting the wind power fluctuation. If $\hat{R}(x,b)$ is the probability density function of wind power fluctuation, then the NPKDE of this probability density function is estimated as,

$$\hat{R}(x,b) = \frac{1}{nb}\sum_{i=1}^{n} K(\frac{x-x_i}{b}) \tag{2}$$

where x_i represents the sample data of wind power fluctuation at the *ith* sampling point, b is the bandwidth, and $K(\bullet)$ is the kernel function.

To ensure the continuity of the estimated probability density function, the kernel function needs to be a symmetric smooth nonnegative function to meet the following characteristics,

$$\begin{cases} \int K(x,b)dx = 1 \\ \int xK(x,b)dx = 0 \\ \int x^2 K(x,b)dx = c \end{cases} \tag{3}$$

where c is a constant.

There are many kernel function choices. However, different kernel functions have little impact on accuracy [23–25]. Therefore, the kernel function selected in this study is a commonly used Gaussian function formulated as,

$$K(x) = \frac{1}{\sqrt{2}}\exp\left(-\frac{x^2}{2}\right) \tag{4}$$

Combination of (2) and (4) allows the formulation of the nonparametric kernel density estimation of the wind power probability density function as,

$$\hat{R}(x,b) = \frac{1}{\sqrt{2\pi}nb}\sum_{i=1}^{n}\exp[-\frac{1}{2}(\frac{x-x_i}{b})^2] \tag{5}$$

To evaluate the goodness-of-fit of the model, three evaluation indices are used as the valid evaluation indices of the model, namely the correlation coefficient r, root mean square error (RMSE), and mean absolute error (MAE). The smaller the RMSE and MAE are, the more accurate the model will be. The closer the correlation coefficient value is to unity, the better the model fit is [26, 27].

The model deviation (geometric distance) can be calculated as,

$$T_i = \hat{R}(x_i,b) - R(x_i,b) \tag{6}$$

where i = 1, 2, ..., m, where m is the number of groupings of the wind power fluctuation sample frequency histogram, $\hat{R}(x_i,b)$ is the ordinate of the *ith* histogram, and $\hat{R}(x_i,b)$ is the function value of the fitted probability density function.

The three evaluation indicators are defined as follows,

$$T_{MAE} = \frac{1}{n}\sum_{i=1}^{n}|T_i| \tag{7}$$

$$T_{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}T_i^2} \tag{8}$$

$$r=\frac{\sum_{i=1}^{n}\left(\hat{R}(x_i,b)-\overline{R(x_i,b)}\right)^2}{\sum_{i=1}^{n}\left(R(x_i,b)-\overline{R(x_i,b)}\right)^2} \tag{9}$$

where n is the number of sample sequences, and $\overline{R(x_i,b)}$ is the average value of the ordinate values of the histogram.

3.2 Model improvement strategy

As indicated in (5), the existing NPKDE theory uses a fixed bandwidth b. Thus, only one b is obtained to estimate the minimum sum of fitness of all the data samples. This processing method may be associated with the situation in which the evaluation index is abnormally large for individual sample data. If the bandwidth b corresponding to the sample data is modified in a targeted manner, the adaptive bandwidth can be solved. Therefore, changing the originally fixed bandwidth yields a set of adaptive bandwidth vectors that can ensure the local adaptive characteristics of the constructed probability model, and further improves the model's goodness-of-fit. Therefore, on the basis of the aforementioned NPKDE method, the following improvement strategies are added. After the use of the bandwidth optimization model to obtain the optimal bandwidth b_Z, the applicability of sample spacing is discriminated. For any sample interval $l\in[l_1,l_2]$ (where $l_2>l_1$ and $l_1,l_2\in[1,n]$), if the following inequality is satisfied, there will be local adaptability problems in the sample interval.

$$T_l(b_z)\geqslant\lambda\overline{T(b_z)} \tag{10}$$

where $T_l(b_z)$ is the model deviation in any sample interval l, $\overline{T(b_z)}$ is the average model deviation of the entire sample space, and λ is the adjustment coefficient. The specific value of λ can be determined according to the actual test situation.

The mathematical expression of the mean geometrical distance $\overline{T(b_z)}$ is,

$$\overline{T(b_Z)}=\frac{1}{n}\sum_{i=1}^{n}T_i \tag{11}$$

For the interval of the local adaptability problem, the bandwidth adjustment model is constructed and the bandwidth matrix is modified to

$$b_l=\frac{n_lT(b_Z)_{mid}}{\sqrt{-2\ln\delta}}b_Z \tag{12}$$

where b_l is the bandwidth of the l sample interval, n_l is the number of samples in the sample interval, $T(b_z)_{mid}$ is the median of the geometric distance in the sample interval, and δ is the kernel function threshold.

Therefore, (5) can be modified in the following form, and an adaptive NPKDE model for the model is thus proposed,

$$\begin{aligned}\hat{R}(x,b)=&A_l\sum_{i=1}^{l_1}H(x_1,b)+A_2\sum_{i=l_1}^{l_2}H(x_2,b)+L+\\&A_k\sum_{i=l_k}^{l_{k+1}}H(x_k,b)+A_{k+1}\sum_{i=l_{k+1}}^{n}H(x_n,b)\end{aligned} \tag{13}$$

where

$$A_I=\omega_I\frac{1}{\sqrt{2\pi}(l_I-l_{I-1}+1)b_I}(l_{0=1}) \tag{14}$$

$$H(x_n,b)=exp[-\frac{1}{2}(\frac{x-x_i}{b_I})^2] \tag{15}$$

In (xx) the number of sampling intervals to be adjusted is k, and b_I is the modified bandwidth matrix of the interval I.

3.3 Bandwidth optimization model

In the NPKDE model, the selection of bandwidth b is important for kernel density estimation [28]. In this study, the test of fitness is used as a constraint condition, and is incorporated into the bandwidth optimization model. A constrained bandwidth optimization model is proposed as follows

$$\begin{aligned}&\min X_{\text{ise}}(b)=\int[\hat{R}(x,b)-R(x,b)]^2\mathrm{d}x\\&\text{s.t.}\quad \chi_b{}^2\leqslant\chi_{m-1}{}^2(\alpha)\end{aligned} \tag{16}$$

where $X_{\text{ise}}(\bullet)$ is the integrated mean-square-error, $R(x,b)$ is the true probability density function that takes the ordinate value of the frequency histogram, χ_b^2 is the test statistics of the NPKDE χ^2, and $\chi_{m-1}{}^2(\alpha)$ is the χ^2 distribution with m-1 degrees-of-freedom at the significant level α.

4 Solution of bandwidth optimization model based on constrained-order optimization

The sequential optimization method is an effective method used to solve complex optimization problems [29–30]. Traditional sequential optimization is generally used for unconstrained optimization problems. In this study, there are constraints in the bandwidth optimization model. Therefore, this study uses constrained sequential optimization based on the following steps:

Step 1: Determine the solution space Ω of the bandwidth value

Step 2: According to (3) and bandwidth value, calculate the wind power wave of the nonparametric kernel density estimation function value for momentum

Step 3: According to (17), calculate the number of solutions in the observation solution set

$$Prob(|\Theta \cap P| \geqslant s) = \sum_{j=k}^{\min\{g,p\}} \sum_{i=0}^{p-j} \frac{\begin{bmatrix} g \\ j \end{bmatrix}\begin{bmatrix} M-g \\ p-i-j \end{bmatrix}}{\begin{bmatrix} M \\ p-j \end{bmatrix}} \bullet \begin{bmatrix} p \\ i \end{bmatrix} q^{p-i}(1-q)^i \geqslant \eta \quad (17)$$

where $Prob(\bullet)$ is the alignment probability, Θ is a sufficiently good solution set, g is the number of real solutions, p is the number of solutions in the observed solution set P. Additionally, s shows that there are at least s truths, η shows the probability that the observation solution set P contains s sufficiently good solutions, and q is the probability that an adjective solution is actually observed in the solution space

Step 4: Select the χ^2 test as the approximate model, and find the p solutions that satisfy it in Ω to form the observation solution set P

Step 5: Select the objective function of (16) as the exact model, and use the exact model to compare the solutions in the solution set P in order; then select the first s solutions to be true solutions

Step 6: Use (7)–(9) as the criterion, selecting the optimal solution bz, and then use (10)–(12) to modify the bandwidth to obtain the final optimal bandwidth sequence.

5 Data analyses

According to the measured data of a wind power field in Hubei Province, the simulation experiment was programmed in MATLAB.

5.1 Extraction of wind power volatility

From March 17 to April 19, 2009, 500 active output data were selected for analysis. The detailed wind power output data are shown in Appendix A. The sampling period of the data was 10 min and the wind farm consisted of 16 wind turbines rated at 850 kW. The moving average method was used to process the data of the first fan in this study, and the output comparison before and after processing is shown in Fig. 1.

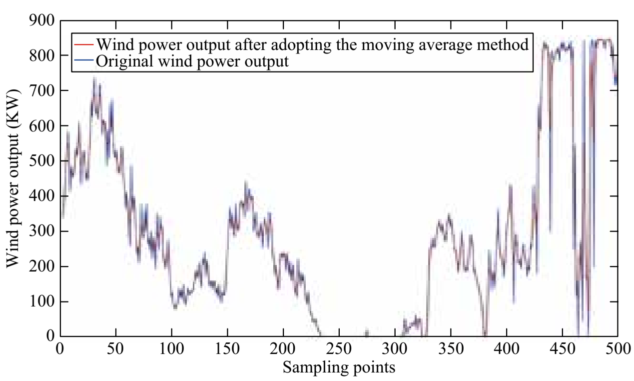


Fig. 1 Comparison of wind power sampled sequences

Fig. 1 shows that the moving average method can set a time window for the sampled data and estimate the local average. Furthermore, the fluctuating random error is filtered out, and the random fluctuation of wind power output is effectively reduced so as to obtain a smoother output result. The comparison of wind power fluctuation is shown in Fig. 2.

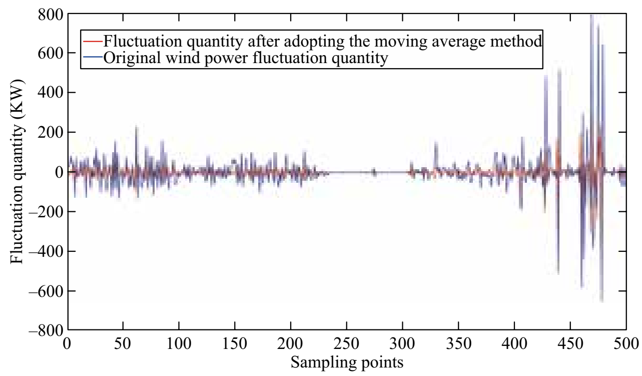


Fig. 2 Comparison of wind power fluctuations

Fig. 2 shows that the fluctuation of wind power output can be slowed down after the method of moving average is used. The wave momentum shows a smooth variational trend to effectively reduce the impact of wind power fluctuation on the accuracy of model construction.

5.2 Comparative analysis

5.2.1 Comparison of the results of NPKDE before and after improvement

The probabilistic density model of the wind power fluctuation component is established for the wind power volatility quantity extracted in this study. Compared with the simulation results of the traditional NPKDE method, the effectiveness of the improved modeling strategy proposed in this study is verified. By analyzing the MAE, RMSE, the three r indices, the fitting degree of the model was evaluated. The probability density curve is shown in Fig. 3. The results of the error operation are listed in Table 1.

Table 1 Comparison of evaluation indices before and after the improvement of the probability density model

Distribution model	Mean absolute error (*MAE*)	Root-mean-square error (*RMSE*)	*r*
Traditional nonparametric kernel density estimation method (NPKDE)	4.917×10^{-5}	3.104×10^{-4}	0.8665
Model proposed in this study	2.773×10^{-5}	1.704×10^{-4}	0.9774

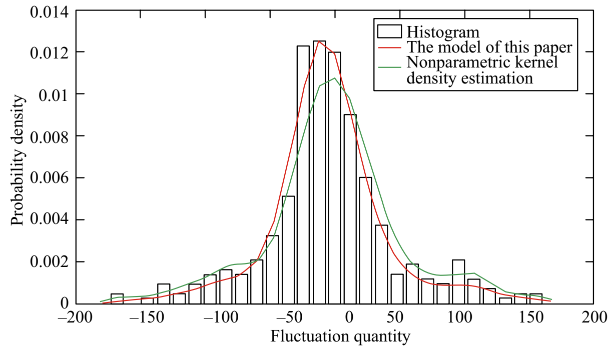


Fig. 3 Comparison of goodness-of-fit before and after the improvement of the NPKDE method

By analyzing the results of Fig. 3 and Table 1, it can be found that the method used in this study is better than the traditional NPKDE method. The MAE, RMSE, and r values of the model are reduced by 43.6%, 45.1%, and improved by 12.8%, compared with the traditional model, respectively. The reason is attributed to the fact that the traditional NPKDE method is designed to minimize the total error of the entire sample and obtain a fixed bandwidth value.

The method of this study not only focuses on the sum of errors, but also modifies the bandwidth of the local intervals where the error is large. Although it is possible to reduce the accuracy of some individual intervals, the accuracies of the sample intervals with the largest errors have been improved considerably. Thus, the overall modeling accuracy has been increased by this method (as shown in Table 1, the improved nonparametric correlation coefficient is much better than the traditional nonparametric correlation coefficient, and is close to unity). It can be observed that the overall modeling accuracy of the multivariable NPKDE method can be improved effectively by the adaptive adjustment of the bandwidth of the sample intervals with local adaptive problems.

5.2.2 Precision and fitness comparison between the parameter estimation method of the mixed distribution function and the adaptive NPKDE

The precision and fitness of wind power volatility of the proposed method is verified by using the parameter estimation method and the modeling method based on mixed logistic, mixed t location–scale, and mixed Gaussian distributions, respectively. The sampled data represent the active power data of the entire wind farm, and the studied time period was from March 17, 2015, to April 19, 2015. The sampling period was 10 min. The comparison results of the precision and fitness of different model evaluation indicators are listed in Table 2. The comparison curve of precision for probability density is shown in Fig. 4, and the fitness comparison curve is shown in Fig. 5.

Table 2 Comparison of the precision and fitness of different model evaluation indicators

Distribution model	*MAE*		*RMSE*		*r*	
	Precision	Fitness	Precision	Fitness	Precision	Fitness
Mixed Gaussian	7.015×10^{-5}	8.486×10^{-5}	3.772×10^{-4}	4.271×10^{-4}	0.9631	0.9534
Mixed t location–scale	6.881×10^{-5}	8.443×10^{-5}	3.696×10^{-4}	4.202×10^{-4}	0.9769	0.9687
Mixed logistic	5.473×10^{-5}	8.167×10^{-5}	3.146×10^{-4}	4.248×10^{-4}	0.9332	0.9295
Model proposed in this paper	2.773×10^{-5}	6.138×10^{-5}	1.704×10^{-4}	3.393×10^{-4}	0.9774	0.9954

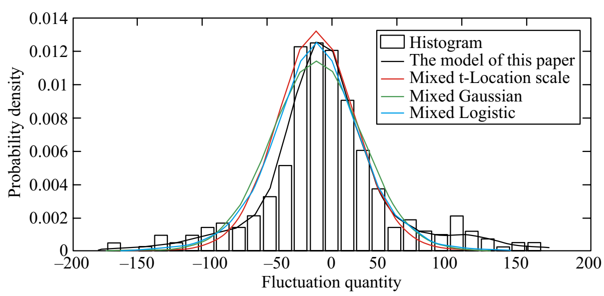


Fig. 4 Comparison of the precisions of different model evaluation indicators

From Fig. 4 and Table 2, it can be observed that among the four probability distributions of wind farm power fluctuation, the model in this study has the best fit, and the three indices are all the best. The r values of the mixed Gaussian and mixed logistic distributions were only 0.9631 and 0.9332. It can be observed that if the prior distribution selection was erroneous, the parameter estimation method could not easily obtain a better modeling accuracy. Although the correlation coefficient of the mixed t-location distribution was close to the model in this study, its MAE increased by 59.7%, and the RMSE increased by 53.9% compared with the model proposed in this study. Based on the analysis of the results, it can be concluded that the improved NPKDE method has a higher modeling accuracy than the mixed parameter estimation method. The main reason is that the method is driven by sample data to directly model the probability distribution. It is not necessary to select the distribution function form of the sample distribution in advance, while the modeling precision does not depend on the selection result of the distribution function that is only related to the bandwidth selection.

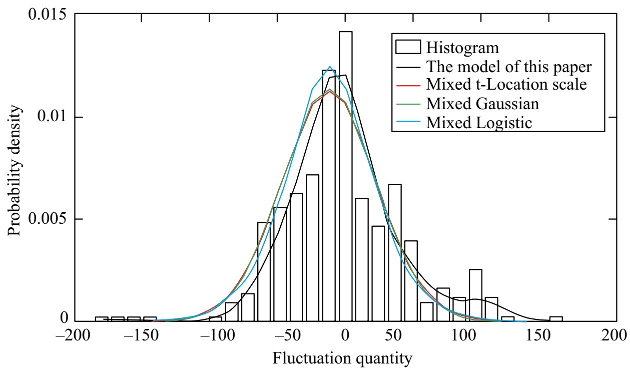


Fig. 5 Comparison of the fitness values of different model evaluation indicators

Fig. 5 and Table 2 show that the model in this study yielded the best fitting effect among the four probability distributions of wind power volatility, and its r value was 0.9954. However, the modeling accuracy of the mixed t location–scale, the mixed Gaussian, and the mixed logistic distributions, were much lower, and their r values were only 0.9534, 0.9687, and 0.9295, respectively. Therefore, the improved NPKDE method proposed in this study is more applicable than the mixed parameter estimation method. The main reason is that the parameter estimation method needs to determine the form of probability distribution function in advance, but the probability distribution of different wind farms may follow different forms. Thus, the parameter estimation of different wind farms with the same distribution function may reduce the goodness-of-fit of the model. Therefore, the method proposed in this study can ensure the precision and fitness of modeling.

5.2.3 Comparison of Operation Time in Constructing Different Models

To test the calculation efficiency of the model constructed in this study, the calculation times of different models were compared, and the specific results are listed in Table 3.

Table 3 Comparison of operation times in the construction of different models

Distribution model	Operation time/s
Mixed Gaussian	87.14
Mixed t location–scale	89.53
Mixed logistic	85.77
Traditional NPKDE	83.94
Model proposed in this study	86.98

Table 3 shows that the operation time of the model constructed by the method in this study has not been increased considerably when the modeling accuracy and the overall goodness-of-fit was improved. Compared with the traditional NPKDE method with the shortest operation time, the operation time of this method was only increased by 3.04 s. It can be concluded that the method in this study improves the modeling accuracy and goodness-of-fit, but with the improvement of modeling accuracy and goodness-of-fit, the operation time of the model did not increase significantly. These findings conform to the actual requirements of the modeling process.

6 Conclusions

A wind power fluctuation modeling method was proposed in this study based on the method of moving average and adaptive NPKDE. To improve the precision of the constructed model of this study, an algorithm was used to solve it based on constraint ordinal optimization. The specific conclusions are as follows:

(1) The method of moving average can reduce the random wind power output fluctuations in this study so as to effectively reduce the impact of these fluctuations on the accuracy of the constructed model

(2) The improved NPKDE method proposed in this study can solve the problem of large interval errors of local samples. Based on the bandwidth correction strategy, multiple adaptive bandwidths can replace the single, fixed bandwidth of traditional NPKDE, and can improve the goodness-of-fit of the NPKDE method

(3) Compared with the traditional modeling method based on parameter estimation, the improved NPKDE based on modeling method is more accurate and applicable.

Appendix A

Table Wind power fluctuation at various sampling points

Sampling points	Wind power output	Sampling points	Wind power output	Sampling points	Wind power output	Sampling points	Wind power output	Sampling points	Wind power output
1	339.6875	101	99.5	201	234.5625	301	−1.1875	401	321.5
2	365.1875	102	79.875	202	189.5	302	−1.125	402	342.5
3	408.8125	103	80.3125	203	240.9375	303	−2	403	355
4	488.4375	104	90.1875	204	220.8125	304	−1.5625	404	436
5	533.875	105	99	205	176.3125	305	0.125	405	414.9375
6	588.3125	106	139.375	206	214.25	306	5.1875	406	287.5
7	540.5	107	112	207	160.4375	307	−1.875	407	97.6875
8	416.1875	108	135.3125	208	187.4375	308	−0.8125	408	272.9375
9	486.5	109	107.625	209	207.25	309	53.6875	409	296.4375
10	456.5625	110	150.0625	210	132.8125	310	37.9375	410	260.25
11	462.6875	111	100.5625	211	125.25	311	6.25	411	208.4375
12	480.0625	112	92.1875	212	114.4375	312	23.8125	412	211.9375
13	537.5625	113	114.375	213	223.1875	313	29	413	217.9375
14	517	114	140.5	214	180.8125	314	26.625	414	195.625
15	547.1875	115	126.5625	215	145.625	315	40.5625	415	255.375
16	612.9375	116	119.875	216	184.3125	316	39.9375	416	265.75
17	564.4375	117	132.4375	217	160.1875	317	35.0625	417	228.5
18	436.1875	118	131	218	177.625	318	36.5625	418	222.9375
19	512	119	136.0625	219	140.625	319	62.875	419	174.125
20	531.75	120	160.875	220	107.4375	320	21.4375	420	200.625
21	529.125	121	188.4375	221	94.625	321	45.875	421	188.9375
22	446.625	122	167.625	222	128.3125	322	53	422	267.5
23	466.3125	123	184.1875	223	95.5	323	52.625	423	251.75
24	447.9375	124	211.4375	224	68.75	324	13.25	424	312
25	522.6875	125	163.5	225	69.5	325	−2.125	425	433.625
26	554	126	142.25	226	49.1875	326	−2.8125	426	363.6875
27	672.125	127	222.6875	227	42.3125	327	−1.9375	427	347.8125
28	629.5	128	211.9375	228	47.3125	328	−2.375	428	181.75
29	695.375	129	242.125	229	51.6875	329	16	429	662.5625
30	738.5	130	173.6875	230	34.6875	330	76.8125	430	600
31	638.0625	131	220.875	231	18.1875	331	224.875	431	591.0625
32	644.3125	132	202.125	232	16.125	332	254.4375	432	725.1875
33	621.75	133	146.125	233	9.5	333	258.8125	433	810.125
34	714.4375	134	148.6875	234	−0.1875	334	282.75	434	837.6875
35	721.75	135	160.75	235	−2	335	253.6875	435	842.0625
36	626.5	136	135.875	236	−2.25	336	256.125	436	825.625

continue

Sampling points	Wind power output	Sampling points	Wind power output	Sampling points	Wind power output	Sampling points	Wind power output	Sampling points	Wind power output
37	577.625	137	161.8125	237	−1.3125	337	282.25	437	833.5625
38	619.75	138	157.75	238	−2.3125	338	307.375	438	802.5
39	613.6875	139	124.375	239	−2	339	320	439	808
40	608.1875	140	139.625	240	−2.125	340	334.0625	440	301.5
41	488.8125	141	112.9375	241	−2	341	319.25	441	818.3125
42	590.875	142	150.125	242	−2.0625	342	300.75	442	800.1875
43	508.3125	143	119	243	−2.0625	343	322.5625	443	789.9375
44	665.1875	144	98.25	244	−2.25	344	295.8125	444	798.875
45	582.75	145	130	245	−2.125	345	275.5	445	812.125
46	677.875	146	120.6875	246	−2	346	273.25	446	822.375
47	544.0625	147	119.5	247	−2.0625	347	323.0625	447	825
48	513.625	148	158.6875	248	−1.125	348	341.625	448	788.25
49	511.5625	149	221.625	249	−1.1875	349	351.6875	449	830.3125
50	533.3125	150	267.5	250	−1.1875	350	314.5	450	841.25
51	471.625	151	332.3125	251	−1.3125	351	335.75	451	819.0625
52	469.4375	152	370.8125	252	−1.0625	352	301.25	452	842
53	466.375	153	310.6875	253	−1	353	287.375	453	821.9375
54	542.1875	154	353.0625	254	−1.125	354	246.875	454	812.75
55	530.375	155	300	255	−1.125	355	245.3125	455	821.3125
56	452.3125	156	286.9375	256	−1.1875	356	219.1875	456	830.1875
57	434.4375	157	381.3125	257	−1	357	183.6875	457	822
58	367.375	158	310.4375	258	−1.3125	358	212.1875	458	844.1875
59	412	159	303.75	259	−1.1875	359	243.625	459	790.125
60	382.0625	160	311.6875	260	−1.125	360	301.875	460	827.8125
61	308.3125	161	383.875	261	−2.1875	361	278.8125	461	249.1875
62	259	162	389.375	262	−2.6875	362	226	462	547.6875
63	488.9375	163	380.875	263	−1.25	363	187.75	463	115.8125
64	357.1875	164	378.875	264	−1.3125	364	194.1875	464	162.1875
65	399.8125	165	341.375	265	−0.6875	365	188.3125	465	−3.5
66	385.8125	166	443	266	0.5	366	208.9375	466	219.5625
67	314.6875	167	386.9375	267	0.75	367	248.5	467	274.25
68	238	168	357.9375	268	−1.125	368	292.3125	468	222.4375
69	257.5625	169	423.5625	269	−1.125	369	286	469	46.375
70	211.3125	170	403.6875	270	−1	370	264.125	470	845.9375
71	248.9375	171	391.4375	271	−1.1875	371	183.75	471	461.8125
72	377.1875	172	402.6875	272	−1.125	372	191.75	472	166.6875
73	292.4375	173	301.5625	273	−1.1875	373	173.25	473	170.3125
74	282.1875	174	322.4375	274	1.125	374	163.125	474	−2.875

continue

Sampling points	Wind power output	Sampling points	Wind power output	Sampling points	Wind power output	Sampling points	Wind power output	Sampling points	Wind power output
75	363.75	175	355.875	275	21.125	375	142.625	475	92.0625
76	300.5	176	298.0625	276	−0.6875	376	137.25	476	829.4375
77	381.75	177	283.5	277	−1	377	111.875	477	840.4375
78	266.1875	178	286.1875	278	−1.9375	378	100.1875	478	848.5625
79	279.3125	179	326.875	279	−1.375	379	74.6875	479	200.625
80	245.875	180	305.375	280	−1.0625	380	3.5	480	844.5625
81	303.625	181	253.6875	281	−1.4375	381	−3.0625	481	826.375
82	223.8125	182	318.9375	282	−1.375	382	−4.125	482	849.9375
83	284.875	183	341.0625	283	−1.3125	383	47.9375	483	847.5
84	262.8125	184	326.5625	284	−1.125	384	142.3125	484	850.25
85	194.5	185	310.875	285	−1.8125	385	246.3125	485	848.5
86	353.8125	186	253.25	286	−1.125	386	197.3125	486	850.25
87	293.6875	187	355.6875	287	−1	387	133.375	487	850.5625
88	245.875	188	351.8125	288	−1.0625	388	186.25	488	841.25
89	344.8125	189	308.875	289	−1.0625	389	177	489	823.1875
90	311.375	190	234.875	290	−1.0625	390	141.1875	490	848
91	315.5625	191	228.375	291	−1.0625	391	239.625	491	847.8125
92	230.1875	192	178.4375	292	−1	392	282.5	492	850
93	205.9375	193	205	293	−1.4375	393	366.625	493	849.875
94	197.1875	194	174	294	−1.3125	394	266	494	848.75
95	214.8125	195	135.9375	295	−1.125	395	230.5	495	800.375
96	277.5	196	143.125	296	−1.125	396	225.4375	496	836.875
97	221	197	234.4375	297	−1.1875	397	215.625	497	788.6875
98	162.8125	198	239.125	298	−1.1875	398	170.8125	498	719.25
99	115.75	199	229.75	299	−1.125	399	210.6875	499	743.5
100	125.25	200	241.125	300	−1.1875	400	292	500	735

Acknowledgements

This work was supported by Science and Technology project of the State Grid Corporation of China "Research on Active Development Planning Technology and Comprehensive Benefit Analysis Method for Regional Smart Grid Comprehensive Demonstration Zone," and National Natural Science Foundation of China (51607104).

References

[1] Yao J.G, Yang S.C, Wang K, et al. (2014) Framework and strategy design of demand response scheduling for balancing wind power fluctuation. Automation of Electric Power Systems, 38(9): 85-92

[2] Huang K.Y, Huang Y.C (2004) Integrating direct load control with interruptible load management to provide instantaneous reserves for ancillary services. IEEE Transactions on Power Systems, 19(3):1626-1634

[3] P. Mukherjee, V.V. Rao (2020) Effective location of SMES for power fluctuation mitigation of grid connected doubly fed induction generator, Journal of Energy Storage, 29

[4] Wan S.T, Wan J (2015) Research on wind power fluctuation characteristics based on quantitative index and probability density distribution. Acta Energiae Solaris Sinica, 36(02): 362-368

[5] Cui Y, Yang H.W, Li H.B (2016) Probability density distribution function of wind power fluctuation of a wind farm group based

on the gaussian mixture model. Power System Technology, 40(04): 1107-1112

[6] Lin H, Sun L.C, Chang X.Q (2014) A probabilistic model to simulate wind power output fluctuation of a certain wind farm cluster in Xinjiang region. Power System Technology, 38(06): 1616-1620

[7] Yang M, Dong J.C (2016) Study on characteristics of wind power fluctuation based on mixed distribution model. Proceedings of the CSEE, 36(S1): 69-78

[8] Zhou T.H, Chen L, Li J (2017) Wind power fluctuation characteristic analysis based on finite Laplace mixture model. Power System Technology, 41(02): 543-550

[9] Yang M, Ma J, Li C.F (2017) A mixed logistic distribution model for wind power fluctuation characteristics. Power System Technology, 41(5): 1376-1382

[10] Ye R.L, Guo Z.Z, Liu R.Y, et al (2014) A method for designing optimal energy storage system based on analysis of wind power forecast error. Automation of Electric Power Systems, 38(16): 28-34

[11] Zhou S.L, Mao M.Q, Su Jian.H (2011) Short-term forecasting of wind power and N on-parametric confidence interval Estimation. Proceedings of the CSEE, 31(25): 10-16

[12] Sun J.B, Wu X.S, Zhang B.H (2013) Wind Power interval prediction based on non-parametric kernel density estimation. Water Resources and Power, 31(9): 233-235

[13] Yang N, Cui J.Z, Zhou Z (2016) Research on nonparametric kernel density estimation for modeling of wind power probability characteristics based on fuzzy ordinal optimization. Power System Technology, 40(02): 335-340

[14] Sun Y.T, Luo L.F, Zhang Q, Qin X.R (2019) Reliability analysis of stochastic structure with multi-failure modes based on mixed Copula. Engineering Failure Analysis, 105: 930-944

[15] Feng J.X, Liang J, Feng Y.K (2015) Optimized calculation of wind farm energy storage capacity based on the fluctuation characteristics of min power component of wind power. Power System Protection and Control, 43(03): 1-6

[16] Lu P.M, Wen B.Y (2014) Study on static voltage stability of hydraulic wind power integrated with power grid. Power System and Clean Energy, 30(05): 85-91

[17] Li L, Liu T.Q, Chen Z.H, et al (2015) Ultra-short-term wind power forecasting based on EEMD and ARCH. Electrical Measurement & Instrumentation, 52(18): 16-21

[18] Zhan Z.Q, Fu J.X. (2019) Research on coordinated control of energy management of wind-storage combined system with smooth wind power. Sichuan Electric Power Technology, 42 (05): 29-36

[19] Wang X.L, Li Hui (2012) Multiscale prediction of wind speed and output power for the wind farm. Journal of Control Theory and Applications, 10(2): 251-258

[20] Yan G.G, Xiong H, Yang M (2013) Study on real-time cooperative prediction based on wavelet theory for wind power. Journal Of Northeast Dianli University, 33(Z1): 120-125

[21] Wang D.X (2004) Analysis of AGC regulation demand in power system operation. Automation of Electric Power Systems, (08): 6-9

[22] Ma S.L, Jiang X.P, Ma H.M, et al. (2014) Capacity allocation of hybrid energy storage system to suppress wind power fluctuations. Power System Protection and Control, 42(08): 108-114

[23] Epanechnikov V A (1969) Nonparametric estimation of a multidimensional probability density. Theory of Probability and Its Application, (14): 156-161

[24] An Z.W, Xin Y (2015) Sensitivity analysis of reliability of gear in wind turbine gear-box based on important sampling method. Journal of Lanzhou University of Technology, 41(3): 36-40

[25] James W. Taylor, Jooyoung Jeon (2015) Forecasting wind power quantiles using conditional kernel estimation. Renewable Energy, 80

[26] Yadav Dhyan Chandra, Pal Saurabh (2019) To generate an ensemble model for women thyroid prediction using data mining techniques. Asian Pacific journal of cancer prevention : APJCP, 20(4): 1275-1281

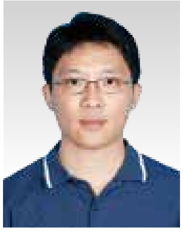
[27] Fanelli Andrea, Vonberg Frederick W, LaRovere Kerri L, et al. (2019) Fully automated, real-time, calibration-free, continuous noninvasive estimation of intracranial pressure in children. Journal of neurosurgery. Pediatrics, 1-11

[28] Boscolo Riccardo, Pan Hong, Roychowdhury Vwani P (2004) Independent component analysis based on nonparametric density estimation. IEEE Transactions on Neural Networks, 15(1): 55-65

[29] Ho Y.C, Sreenivas R S, Vakili P (1992) Ordinal optimization of discrete event dynamic systems. Theory and Applications, 2(2): 61-88

[30] Lina Reichenberg, Filip Johnsson, Mikael Odenberger (2014) Dampening variations in wind power generation—the effect of optimizing geographic location of generating sites. Wind Energy, 17(11): 1631-1643

Biographies



Peizhe Xin received his Master's degree at the University of Electronic Science and Technology, China. He is working at the State Grid Beijing Institute of Economics and Technology. His research interests include intelligent power grid and power communication network planning.

Ying Liu received her Master's degree at the North China Electric Power University. She is working at the State Grid Beijing Institute of Economics and Technology Her research interests include Power system analysis and control, smart grid evaluation and design.

Nan Yang received his Ph.D. degree at the Wuhan University. He is working in the Department of Hubei Provincial Collaborative Innovation Center for New Energy Microgrid, China Three Gorges University. His research interests include power system operation and control, power system unit combination, artificial intelligence and big data, and power system planning.

Xuankun Song received her Bachelor's degree at the North China Electric Power University. She is working at the State Grid Beijing Institute of Economics and Technology. Her research interests include power system analysis and control, and smart grid evaluation and design.

Yu Huang received his Master's degree at the China Three Gorges University. His research interests include intelligent power grid and power communication network planning.

(Editor Zhou Zhou)