Statistical Representation of Wind Power Ramps Using a Generalized Gaussian Mixture Model

Mingjian Cui, *Member, IEEE*, Cong Feng, *Student Member, IEEE*, Zhenke Wang, and Jie Zhang, *Senior Member, IEEE*

***Abstract*—Wind power ramps are significantly impacting the power balance of the system operations. Understanding the statistical characteristics of ramping features would help power system operators better manage these extreme events. Toward this end, this paper develops an analytical generalized Gaussian mixture model (GGMM) to fit the probability distributions of different ramping features. The nonlinear least-squares method with the trust-region algorithm is adopted to optimize the tunable parameters of mixture components. The optimal number of mixture components is adaptively solved by minimizing the Euclidean distance between the GGMM and the actual histogram distribution. The probability distribution of ramping features is generally truncated due to the ramp definition with a specific threshold. Thus, a sign function is utilized to truncate the GGMM distribution. Then, the cumulative distribution function of the GGMM is analytically derived and utilized to design a random number generator for ramping features. Numerical simulations on publicly available wind power data show that the parametric GGMM can accurately characterize the irregular and multimodal distributions of ramping features.**

***Index Terms*—Generalized Gaussian mixture model, probability distribution, statistical analysis, wind power ramps.**

I. Introduction

With the increase of wind power penetration in the power grid, the uncertainty and variability of wind power have drawn more and more attention, especially under extreme weather conditions [1], [2]. As an important type of extreme events, wind power ramps (WPRs) have been investigated in recent studies [3]–[5]. WPRs have a serious impact on the power balance of the system, and may lead to an instability of the power system frequency, load shedding, and other reliable operations. Statistical analysis would help power system operators better understand the characteristics of ramping features, thereby assisting them to manage these extreme ramping events. Wind power ramping features generally consist of ramping duration, rate, and magnitude. These ramping features have been used to adjust the system scheduling in the California Independent System Operator (CAISO) Balancing Authority [6], [7], which makes the statistical analysis more desired and practical [8].

Currently, there are few studies in the literature focusing on accurately characterizing the parametric distributions of wind power ramping features, which are apt to be practically integrated into the power system scheduling models like the chance-constrained economic dispatch (ED) and unit commitment (UC) [9]. Sevlian *et al.* [10] characterized and analyzed ramping magnitude, duration, and rate by empirical distributions. But the empirical distribution is a step function with discrete (rather than continuous) probability values, which cannot be analytically expressed. Cui *et al.* [11] depicted the ramping feature statistics by using the kernel smoothing probability density (ksdensity) estimate. However, the ksdensity distribution was a nonparametric model, which limits the practical use of the distribution. Ganger *et al.* [12] utilized the Fréchet distribution (a generalized extreme value distribution) to fit the empirical wind power ramping magnitude. However, the Fréchet distribution is a unimodal distribution that cannot accurately capture the multimodal distribution.

The Gaussian mixture model (GMM) has been widely used to fit the multimodal distribution in the statistics community, and recently been applied in the renewable energy areas [13]–[15]. The GMM specializes in characterizing the multimodal and irregular probability distribution. Ke *et al.* [13] customized the GMM by three Gaussian functions and utilized the GMM to approximate the probability density function (PDF) of wind power generation with triple probability peaks. Singh *et al.* [14] represented all irregular PDFs of load using GMM in various distribution system applications. Valverde *et al.* [15] proposed the use of GMM to represent non-Gaussian correlated wind power output and aggregated load demands for modeling the probabilistic load flow. However, the widely used GMM is constrained by two unity constraints. The first constraint is that the integral of each mixture component (standard normal distribution) over the entire sample space must strictly equal unity. The other constraint is that the sum of all weights must strictly equal unity as well. These constraints limit the practical use of GMM in distribution fitting. WPRs are highly nonlinear and uncertain, and likely present multi-mode in the distribution of ramping features. To accurately characterize the distribution of ramping features, this paper develops a generalized Gaussian mixture model (GGMM) to fit the PDFs and

Manuscript received November 18, 2016; revised January 23, 2017, March 13, 2017, April 19, 2017, and June 5, 2017; accepted July 11, 2017. Date of publication July 14, 2017; date of current version December 14, 2017. This work was supported by the National Renewable Energy Laboratory under Subcontract XGJ-6-62183-01 (under the U.S. Department of Energy Prime Contract DE-AC36-08GO28308). Paper no. TSTE-00897-2016. *(Corresponding author: Jie Zhang.)*

The authors are with the Department of Mechanical Engineering, The University of Texas at Dallas, Richardson, TX 75080 USA (e-mail: mingjian.cui@utdallas.edu; cong.feng1@utdallas.edu; zhenke.wang@utdallas.edu; jiezhang@utdallas.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSTE.2017.2727321

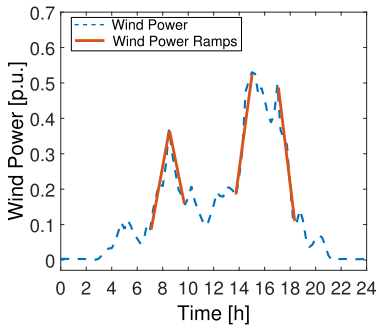


Fig. 1. An example of wind power ramps in one day.

cumulative distribution functions (CDFs) of different ramping features.

In addition to the parametric GMM distribution, one nonparametric (i.e., ksdensity) and two more parametric distributions (i.e., normal and hyperbolic), widely used in statistical analysis, are selected to verify the effectiveness of the GGMM in this paper. Zhang *et al.* [16] used the nonparametric ksdensity distribution in wind speed distribution characterization and renewable energy forecasting. Makarov *et al.* [6] utilized the normal distribution to design the random number generator and generate load forecasting errors. Hodge *et al.* [17] deployed the hyperbolic distribution to analyze and characterize wind and load forecasting errors.

The developed GGMM model is expected to accurately model the distributions of different ramping features, therefore being used in a variety of power system operations. The advantage of the GGMM approach is that different types of ramping feature distributions are fairly represented as a combination of multiple Gaussian components. The main contributions of this paper include: (i) developing a GGMM methodology to fit the irregular and multimodal distributions of wind power ramping features; (ii) deducing the analytical CDF expression of the GGMM; and (iii) designing a random number generator based on the analytical CDF for wind power ramping features.

The organization of this paper is as follows. In Section II, a wind power ramp extraction method using an optimized swinging door algorithm is briefly introduced. Section III presents the analytical expressions of the probability and cumulative distributions of the developed GGMM. Case studies and results analysis performed on publicly available wind power data are discussed in Section IV. Concluding remarks and future work are given in Section VI.

II. Wind Power Ramping Features Extraction

An optimized swinging door algorithm (OpSDA) [11] is used to detect all the WPRs in historical wind power data. A brief example of wind power ramp detection in one day is illustrated in Fig. 1. In the OpSDA, the swinging door algorithm (SDA) with a predefined parameter, ε, is first applied to segregate the wind power data into multiple discrete segments. Then dynamic programming is used to merge adjacent segments with the same ramping direction and relatively high ramping rate. Subintervals that satisfy the ramping rules are rewarded by a score function; otherwise, their score is set to zero. The current subinterval is retested as above after being combined with the next subinterval. This process is performed recursively to the end of dataset. Finally, significant wind power ramps with the maximum score are successfully extracted.

A positive score function, S_c, is designed based on the length of the interval segregated by the SDA. Given a time interval (u, v) of all discrete time points and an objective function, S_1, of the dynamic programming, one WPR is extracted by maximizing the objective function:

$$S_1(u,v) = \max_{u<k<v} [S_c(u,k) + S_1(k,v)], \quad u < v \tag{1}$$

s.t.

$$S_c(u,v) > S_c(u,k) + S_c(k+1,v), \quad \forall u < k < v \tag{2a}$$

$$S_c(u,v) = (u-v)^2 \times RL(u,v), \quad \forall u < k < v \tag{2b}$$

where the positive score function S_c conforms to a superadditivity property in (2). Ramping rule, RL, is defined as the change in wind power magnitude without ramp duration limits [18], [19]. Based on (1)-(2), the optimization process can proceed recursively as follows. Assuming that the number of all WPRs is M ($\forall m : 1 < m < M$); the WPR interval set $\xi = \{E_1, \ldots, E_m, \ldots, E_M\}$ is the set of intervals, $E_m = (s_m, e_m)$; and the non-WPR interval set $\bar{\xi} = \{\overline{E_1}, \ldots, \overline{E_m}, \ldots, \overline{E_M}\}$ is the set of intervals $\overline{E_m} = (\overline{s_m}, \overline{e_m})$. For $\forall(\overline{s_m}, \overline{e_m}) \in \bar{\xi}$ and $\forall u, v : \overline{s_m} < u < v < \overline{e_m}$, then:

$$RL(u,v) = 0 \tag{3a}$$

$$S_c(u,v) = 0 \tag{3b}$$

$$S_1^*(\overline{s_m}, \overline{e_m}) = 0 \tag{3c}$$

Considering the super-additivity property in (2a), $S_1^*(s_m, e_m)$ equals $S_c(s_m, e_m)$. An optimal sequence of WPR and non-WPR can be presented as $\Theta = \{\overline{E_1}, E_1, \ldots, \overline{E_m}, E_m, \overline{E_{m+1}}, E_{m+1}, \ldots, E_{M-1}, \overline{E_M}\}$, for a given wind power series with L points. The solution to (1) being $S_1^*(\overline{s_1}, \overline{e_M})$ is shown as:

$$\begin{aligned} S_1^*(\overline{s_1}, \overline{e_M}) = & \max_{\overline{s_1}<k_1<k_2<\cdots<k_{u-1}<k_u<\overline{e_M}} \Big\{ S_c(\overline{s_1}, k_1) \\ & + S_c(k_1+1, k_2) + \cdots + S_c(k_{u-1}+1, k_u) \Big\} \\ & + \max_{k_u+1<k_u<\overline{e_M}} S_1(k_u+1, \overline{e_M}) \end{aligned} \tag{4}$$

III. Generalized Gaussian Mixture Model

A. GGMM Description

The PDF of the GGMM is a weighted sum of a finite number of Gaussian functions. It is characterized by the number of mixture components, weights, mean values, and standard deviations of each component, and formulated as:

$$f_{\mathrm{G}}(x|\mathrm{N_G};\mathbf{\Gamma}) = \sum_{i=1}^{\mathrm{N_G}} \omega_i g_i(x|\mu_i, \sigma_i), \quad \forall x \in \mathscr{X}, \forall i \in \mathscr{I} \tag{5}$$

where $\mathscr{X}$ is the data set of a random variable, x, with a total number of N_x. $\mathscr{I}$ is the set of mixture components with a total number of $\mathrm{N_G}$. $\mathbf{\Gamma}$ is the overall parameter matrix. Each

TABLE I
MATHEMATICAL DIFFERENCES BETWEEN THE GMM AND THE DEVELOPED GGMM

Constraints	GMM	GGMM
(i)	$\omega_i \geq 0, \forall i \in \mathscr{I}$	$-\infty \leq \omega_i \leq +\infty, \forall i \in \mathscr{I}$
(ii)	$\sum_{i=1}^{\mathrm{N_G}} \omega_i \equiv 1$	$-\infty \leq \sum_{i=1}^{\mathrm{N_G}} \omega_i \leq +\infty$
(iii)	$\int_{-\infty}^{+\infty} g_i(x\|\mu,\sigma)dx \equiv 1, \forall i \in \mathscr{I}$	$\int_{-\infty}^{+\infty} g_i(x\|\mu,\sigma)dx > 0, \forall i \in \mathscr{I}$

vector of $\mathbf{\Gamma}$ (γ_i) defines a mixture component of the GGMM, i.e., $\mathbf{\Gamma} = \{\gamma_i : \gamma_i = \{\omega_i, \mu_i, \sigma_i\}\}_{i=1}^{\mathrm{N_G}}$. ω is the weight. μ is the mean value. σ is the standard deviation. The function of each component $g(x|\mu,\sigma)$ conforms to a Gaussian function:

$$g(x|\mu,\sigma) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{6}$$

The developed GGMM in this paper has improved the GMM that is currently used in the literature. For the GMM methodology, the expectation maximization (EM) algorithm is utilized to obtain all the solution parameters [13]–[15]. However, it must be strictly subject to three constraints: (i) all the weights of mixture components must be nonnegative; (ii) the sum of all weights equals one; and (iii) the integral of each mixture component (standard normal distribution) equals one. For the developed GGMM, all the three constraints are not required, which means more mixture Gaussian components, even with negative weights, can be integrated in a more generalized way. Mathematical differences between the GMM and the developed GGMM are listed in Table I.

B. Parameter Estimation by the Nonlinear Least-Squares Method

The nonlinear least square (NLS) method is adopted to estimate all the parameters (ω, μ, and σ) of mixture components of the GGMM. The number of these mixture components is fixed during each step of parameter estimation. Thus in this section, the PDF is simplified as $f_{\mathrm{G}}(x|\mathrm{N_G};\mathbf{\Gamma}) \xrightarrow{\mathrm{N_G}} f_{\mathrm{G}}(x|\mathbf{\Gamma})$. Suppose a set of data points for ramping features (x) and their actual probabilities (p), i.e., (x_1, p_1), ..., (x_j, p_j), ..., $(x_{\mathrm{N}_x}, p_{\mathrm{N}_x})$, the objective function, S_2, of the NLS is to minimize the sum of the squares of fitting residuals, given by:

$$S_2 = \sum_{j=1}^{\mathrm{N}x} r_j^2 = \sum_{j=1}^{\mathrm{N}x} \left[p_j - f_{\mathrm{G}}(x_j|\mathbf{\Gamma})\right]^2 \tag{7}$$

where r_j is the residual of the jth data point.

Taking the mean value, μ, as an example, the minimum value of S_2 occurs when the gradient is zero, given by:

$$\frac{\partial S_2}{\partial \mu_i} = 2\sum_{j=1}^{\mathrm{N}_x} r_j \frac{\partial r_j}{\partial \mu_i} = -2\sum_{j=1}^{\mathrm{N}_x} r_j \frac{\partial f_{\mathrm{G}}(x_j|\mathbf{\Gamma})}{\partial \mu_i} = 0 \tag{8}$$

Since the model contains $3 \times \mathrm{N_G}$ parameters, there are $3 \times \mathrm{N_G}$ gradient equations. Then, each mean value of the GGMM, μ_i, is refined iteratively by the successive approximation:

$$\mu_i \approx \mu_i^{k+1} = \mu_i^k + \Delta\mu_i \tag{9}$$

where k is an iteration number and the increment, $\Delta\mu_i$, is known as the step size. At each iteration, the GGMM is linearized by approximating to a first-order Taylor series expansion:

$$f_{\mathrm{G}}(x|\mathbf{\Gamma}) \approx f_{\mathrm{G}}(x|\mathbf{\Gamma}^k) + \sum_{i=1}^{\mathrm{N_G}} \frac{\partial f_{\mathrm{G}}(x|\mathbf{\Gamma}^k)}{\partial \omega_i}\Delta\omega_i + \sum_{i=1}^{\mathrm{N_G}} \frac{\partial f_{\mathrm{G}}(x|\mathbf{\Gamma}^k)}{\partial \mu_i}\Delta\mu_i + \sum_{i=1}^{\mathrm{N_G}} \frac{\partial f_{\mathrm{G}}(x|\mathbf{\Gamma}^k)}{\partial \sigma_i}\Delta\sigma_i \tag{10}$$

where the Jacobian matrix, $\mathbf{J}$, is defined as a set of the derivatives $\partial f_{\mathrm{G}}(x|\mathbf{\Gamma}^k)/\partial\omega_i$, $\partial f_{\mathrm{G}}(x|\mathbf{\Gamma}^k)/\partial\mu_i$, and $\partial f_{\mathrm{G}}(x|\mathbf{\Gamma}^k)/\partial\sigma_i$. Each derivative is analytically deduced by:

$$J_{ij}^{\omega} = \frac{\partial f_{\mathrm{G}}(x_j|\mathbf{\Gamma}^k)}{\partial \omega_i} = e^{-\frac{(x_j-\mu_i^k)^2}{2(\sigma_i^k)^2}} \tag{11a}$$

$$J_{ij}^{\mu} = \frac{\partial f_{\mathrm{G}}(x_j|\mathbf{\Gamma}^k)}{\partial \mu_i} = \frac{\omega_i^k(x_j-\mu_i^k)}{(\sigma_i^k)^2} e^{-\frac{(x_j-\mu_i^k)^2}{2(\sigma_i^k)^2}} \tag{11b}$$

$$J_{ij}^{\sigma} = \frac{\partial f_{\mathrm{G}}(x_j|\mathbf{\Gamma}^k)}{\partial \sigma_i} = \frac{\omega_i^k(x_j-\mu_i^k)^2}{(\sigma_i^k)^3} e^{-\frac{(x_j-\mu_i^k)^2}{2(\sigma_i^k)^2}} \tag{11c}$$

$$\mathbf{J} = [\mathbf{J}^{\omega}\ \mathbf{J}^{\mu}\ \mathbf{J}^{\sigma}]^{\mathbf{T}},\ \forall J_{ij}^{\omega} \in \mathbf{J}^{\omega}, \forall J_{ij}^{\mu} \in \mathbf{J}^{\mu}, \forall J_{ij}^{\sigma} \in \mathbf{J}^{\sigma} \tag{11d}$$

Since the iterative residuals are given by: $\Delta p_j = p_j - f_{\mathrm{G}}(x_j|\mathbf{\Gamma}^k)$, the original residuals are rearranged by:

$$\begin{aligned} r_j &= (p_j - f_{\mathrm{G}}(x_j|\mathbf{\Gamma}^k)) + (f_{\mathrm{G}}(x_j|\mathbf{\Gamma}^k) - f_{\mathrm{G}}(x_j|\mathbf{\Gamma})) \\ &= \Delta p_j - \sum_{s=1}^{\mathrm{N_G}} J_{sj}^{\omega}\Delta\omega_s - \sum_{s=1}^{\mathrm{N_G}} J_{sj}^{\mu}\Delta\mu_s - \sum_{s=1}^{\mathrm{N_G}} J_{sj}^{\sigma}\Delta\sigma_s \end{aligned} \tag{12}$$

Then substituting these expressions in (11) and (12) into the gradient equations in (8), we can rearrange and get the normal equations:

$$\sum_{j=1}^{\mathrm{N}_x}\sum_{s=1}^{\mathrm{N_G}} J_{ij}^{\mu}(J_{sj}^{\omega}\Delta\omega_s + J_{sj}^{\mu}\Delta\mu_s + J_{sj}^{\sigma}\Delta\sigma_s) = \sum_{j=1}^{\mathrm{N}_x} J_{ij}^{\mu}\Delta p_j \tag{13a}$$

$$\sum_{j=1}^{\mathrm{N}_x}\sum_{s=1}^{\mathrm{N_G}} J_{ij}^{\omega}(J_{sj}^{\omega}\Delta\omega_s + J_{sj}^{\mu}\Delta\mu_s + J_{sj}^{\sigma}\Delta\sigma_s) = \sum_{j=1}^{\mathrm{N}_x} J_{ij}^{\omega}\Delta p_j \tag{13b}$$

$$\sum_{j=1}^{\mathrm{N}_x}\sum_{s=1}^{\mathrm{N_G}} J_{ij}^{\sigma}(J_{sj}^{\omega}\Delta\omega_s + J_{sj}^{\mu}\Delta\mu_s + J_{sj}^{\sigma}\Delta\sigma_s) = \sum_{j=1}^{\mathrm{N}_x} J_{ij}^{\sigma}\Delta p_j \tag{13c}$$

The normal equations are written in the matrix notation as:

$$\mathbf{J^T J}\begin{bmatrix}\boldsymbol{\Delta\omega}\\ \boldsymbol{\Delta\mu}\\ \boldsymbol{\Delta\sigma}\end{bmatrix} = \mathbf{J^T}\boldsymbol{\Delta p} \tag{14}$$

Since the estimated initial parameters may be far from the optimum, Equation (14) is improved by using the trust-region

algorithm [20], given by:

$$[\mathbf{J^T J} + \lambda \mathrm{diag}(\mathbf{J^T J})] \begin{bmatrix} \boldsymbol{\Delta\omega} \\ \boldsymbol{\Delta\mu} \\ \boldsymbol{\Delta\sigma} \end{bmatrix} = \mathbf{J^T \Delta p} \tag{15}$$

where $\mathrm{diag}(\mathbf{J^T J})$ is the diagonal matrix with the same diagonal as $\mathbf{J^T J}$ and λ is used to control the trust-region size.

After estimating the parameters of each mixture component of the GGMM, the next step is to determine the optimal number of mixture components, $\mathrm{N_{G,opt}}$, by minimizing the Euclidean distance between the actual PDF, $\mathrm{PDF_A}$, and the PDF of GGMM, f_G, i.e., $f_\mathrm{G}(x|\mathrm{N_G}) \xrightarrow{\mathrm{N_{G,opt}}} f_\mathrm{G}(x)$. The objective function is formulated as:

$$\min \sqrt{\sum_{x \in \mathscr{X}} [f_\mathrm{G}(x|i) - \mathrm{PDF_A}]^2}, \quad i = 1, 2, ..., \mathrm{N_G} \tag{16}$$

C. Truncated GGMM

Normally, the probability distribution of ramping features is truncated due to the ramp definition with a specific threshold. Hence, a sign function, $sign(x)$, is utilized in this section to truncate the aforementioned PDF of the GGMM, given by:

$$sign(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{17}$$

Then the final truncated PDF of the GGMM, $f(x)$, is analytically formulated as:

$$f(x) = f_\mathrm{G}(x) \times \mathrm{sign}(x - \mathrm{Tr}) \tag{18}$$

where Tr is the threshold for defining wind power ramps. In this paper, a wind power ramp is defined as the change in wind power output larger than 20% of the rated capacity without constraining the ramping duration and rate. The threshold of ramping magnitude, $\mathrm{Tr_M}$, equals 0.2. The threshold of ramping duration, $\mathrm{Tr_D}$, equals 0. The threshold of ramping rate, $\mathrm{Tr_R}$, is calculated by $\mathrm{Tr_M}/(\max(\mathrm{Dr}))$, where $\max(\mathrm{Dr})$ represents the maximum value of ramping duration. This wind power ramp definition is predefined empirically and has been widely used in the literature, such as [10], [11], [18], [21], [22].

D. Analytical Expression of CDF of the GGMM

The cumulative distribution, F_G, is another essential statistic metric to analyze ramping features due to its monotonicity, which is analytically expressed as:

$$\begin{aligned} F_\mathrm{G}(x|\mathrm{N_G};\boldsymbol{\Gamma}) &= \int_{-\infty}^{x} \sum_{i=1}^{\mathrm{N_G}} \omega_i e^{-\frac{1}{2}\left[\frac{t-\mu_i}{\sigma_i}\right]^2} dt \\ &= \sum_{i=1}^{\mathrm{N_G}} \left[\frac{\sqrt{\pi}}{2}\omega_i\sigma_i erf\left(\frac{\mu_i - x}{\sigma_i}\right)\right] + \mathrm{C} \end{aligned} \tag{19}$$

where erf is the Gaussian error function and defined as:

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{20}$$

Equation (19) is an indefinite integral with a constant C, which can be solved by (21). Since the ramping features are normalized into the range [0, 1], it can be derived that $F_\mathrm{G}(x < 0) = 0$ and $F_\mathrm{G}(x > 1) = 1$. Hence, we use $x = -0.1$ (i.e., $F_\mathrm{G}(-0.1) = 0$) to obtain the constant, C, given by:

$$\mathrm{C} = F_\mathrm{G}(-0.1) - \sum_{i=1}^{\mathrm{N_G}} \left[\frac{\sqrt{\pi}}{2}\omega_i\sigma_i erf\left(\frac{-0.1 - \mu_i}{\sigma_i}\right)\right] \tag{21}$$

Considering the sign function in (17), the final CDF of the GGMM, $F(x)$, is analytically formulated as:

$$F(x) = F_\mathrm{G}(x) \times \mathrm{sign}(x - \mathrm{Tr}) \tag{22}$$

Algorithm 1: Newton-Raphson method for generating the random number of ramping features.

1 Initialization: uniformly sample one point and evenly partition a ramping feature, x, into N regions $([\underline{x_n}, \overline{x_n}]_{n=1}^{\mathrm{N}})$ $r \in [0, 1]$. Decide the region where r is:
2 **if** $F_\mathrm{G}(\underline{x_n}) < r < F_\mathrm{G}(\overline{x_n})$ **then**
3 Return n, $x_0 = \overline{x_n}$; and the approximation x_1 is calculated when $k = 0$, by:

$$\begin{aligned} x_{k+1} &= x_k - \frac{F_\mathrm{G}(x_k) - r}{F'_\mathrm{G}(x_k)} = x_k - \frac{F_\mathrm{G}(x_k) - r}{f_\mathrm{G}(x_k)} \\ &= x_k - \frac{\sum_{i=1}^{\mathrm{N_G}} \left[\frac{\sqrt{\pi}}{2}\omega_i\sigma_i erf(\frac{\mu_i - x_k}{\sigma_i})\right] + \mathrm{C} - r}{\sum_{i=1}^{\mathrm{N_G}} \omega_i e^{-\frac{1}{2}\left[\frac{x_k - \mu_i}{\sigma_i}\right]^2}} \end{aligned} \tag{23}$$

4 **end**
5 Output the random number of one ramping feature:
6 **for** Iteration k *from* 1 to 100 **do**
7 **if** $|x_k - x_{k-1}| < \epsilon$ **then**
8 The sampled ramping feature is returned: $\hat{x} \approx x_k$;
9 **else**
10 The iterative process is repeated using (23).
11 **end**
12 **end**

E. Application of GGMM: Random Number Generator (RNG)

One important application of the ramping features statistical characterization is to generate random and realistic ramping features for the power system operations study. In this section, we design an RNG based on the GGMM (RNG-GGMM). The inverse transform method using Monte Carlo sampling has been widely used to generate the random number [23]. In this paper, the RNG-GGMM method utilizes the inverse function of the CDF of GGMM, formulated as:

$$\hat{x} = F_\mathrm{G}^{-1}\left(\sum_{i=1}^{\mathrm{N_G}} \left[\frac{\sqrt{\pi}}{2}\omega_i\sigma_i erf\left(\frac{\mu_i - x}{\sigma_i}\right)\right] + \mathrm{C}\right) \tag{24}$$

However, the inverse function in (24) cannot be analytically deduced. Instead of obtaining the analytical solution, we use the Newton-Raphson method to obtain a numerial solution. The

pseudocode of the ramping features generation process is illustrated in Algorithm 1. The process repeats until the range is less than the stopping threshold ϵ, given by:

$$|x_k - x_{k-1}| < \epsilon \tag{25}$$

where ϵ is set as 1×10^{-8}. Finally, the random number of a ramping feature is generated.

IV. Case Studies and Results

A. Test Cases and Benchmarks

The proposed GGMM distribution is evaluated and analyzed based on the Wind Integration National Dataset (WIND) Toolkit [24]. The WIND Toolkit data represents wind power generation and forecasts spanning from January 1st 2007 to December 31st 2012, sampled every 5 minutes. Five regions located around Dallas, Chicago, New York, Los Angeles, and Miami are selected for the case studies. Two rated capacities with approximately 10GW and 2 GW are chosen, including almost 3,500 and 700 wind sites, respectively. Five time resolutions (5-, 15-, 30-, 45-minute, and 1-hour) are selected for the case studies. The total number of samples is 10,696,960, which is sufficiently large for statistical analysis of ramping features. The door width of OpSDA is set as 5% of the rated capacity.

Based on the analysis of a large number of experiments, we find that the ramping duration and rate usually present unimodal PDF distributions, while the ramping magnitude presents a multimodal distribution. For the simplicity of comparison, we study the statistical characterizations of the ramping duration and rate in Section IV-C, and those of the ramping magnitude with more simulations in Section IV-D.

B. Ensemble Metrics for Assessing the Fitting Performance

To verify the performance of GGMM with different distributions, a suit of widely used metrics are adopted to assess the distribution accuracy. These metrics include the Chi-square goodness-of-fit (GOF), correlation coefficient, normalized root mean square error (NRMSE), maximum absolute error (MaxAE), mean absolute error (MAE), Kolmogorov-Smirnov test integral (KSIPer), and fourth root mean quartic error (4RMQE). The correlation coefficient is a measure of the correlation between the actual PDF and the PDF of fitting distributions. NRMSE is suitable for evaluating the overall accuracy of the fit while penalizing large fitting errors in a square order. MaxAE is suitable for evaluating the largest fitting error. MAE is suitable for evaluating uniform fitting errors. KSIPer evaluates the statistical similarity between the actual PDF and the PDF of fitting distributions. 4RMQE is suitable for evaluating the overall accuracy of the fit while penalizing large fitting errors in a quartic order. A smaller value indicates a better forecast for most of the metrics, except for the correlation coefficient. Detailed information about these metrics can be found in [16].

C. Statistical Comparisons of Ramping Duration and Rate

Ramping duration and rate are two important ramping features. In this section, the probability distributions of ramping duration and rate are characterized and analyzed by using the GGMM. Fig. 2(a) compares the probability distribution of ramping duration by using five distributions. There are five Gaussian components for the GGMM that optimally fit the distribution of ramping duration. Due to the irregular and asymmetric characteristics of the ramping duration distribution, the parametric models of the normal and hyperbolic distributions fail to track the actual probability values, especially for the peak values. However, the GMM, GGMM, and the nonparametric distribution (ksdensity) can fit the actual histogram distribution well. This is specifically illustrated in Table II with numerical metrics. A larger value of the correlation coefficient and smaller values of other metrics indicate a better fit. The GMM, GGMM, and ksdensity distributions show better performance than other parametric distributions, and the GGMM provides an equal-to-better performance comparing to the GMM and ksdensity distributions, especially in terms of the KSIPer indicator. It means the GGMM can characterize more statistical similarity to the actual histogram distribution for the ramping duration.

The probability distribution of ramping rate is a truncated distribution as shown in Fig. 2(b), where the truncation threshold is 3.54 MW/min. Four Gaussian components are found to optimally fit the actual histogram distribution of ramping rate. It is shown that the GGMM characterizes the peak of probability better than other distributions. Table III lists the metrics for the fitting performance of different distributions. Among all parametric models, the GGMM performs better than the normal and hyperbolic distributions for all metrics, and shows an equal-to-better performance comparing to the GMM. Comparing with the nonparametric ksdensity model, the GGMM can also provide an equal-to-better performance however in an analytical way.

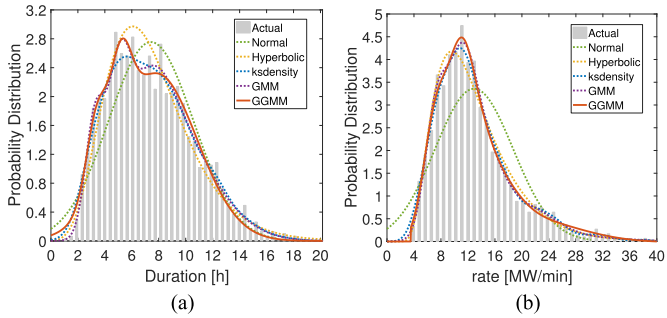


Fig. 2. PDFs of ramping duration and rate using five distributions. (a) PDFs of ramping duration. (b) PDFs of ramping rate.

TABLE II
Metrics of Five Distributions Estimated for Ramp Duration

Metrics	Distributions				
	Normal	Hyper.	ksdensity	GMM	GGMM
Correlation coeff.	0.95	0.98	0.99	0.99	0.99
Chi - square	4.26	0.68	0.46	0.47	0.42
NRMSE	0.10	0.07	0.05	0.05	0.05
MaxAE	0.29	0.18	0.14	0.13	0.11
MAE	0.07	0.05	0.03	0.03	0.03
KSIPer	11.37	17.70	12.06	11.88	10.24
4RMQE	0.14	0.10	0.07	0.07	0.06

TABLE III
METRICS OF FIVE DISTRIBUTIONS ESTIMATED FOR RAMP RATE

Metrics	Distributions				
	Normal	Hyper.	ksdensity	GMM	GGMM
Correlation coeff.	0.92	0.98	0.99	0.99	0.99
Chi - square	8.82	0.43	0.25	0.40	0.14
NRMSE	0.11	0.04	0.03	0.03	0.03
MaxAE	0.31	0.16	0.11	0.09	0.07
MAE	0.07	0.02	0.02	0.02	0.02
KSIPer	21.26	6.54	6.29	5.88	5.29
4RMQE	0.16	0.07	0.05	0.05	0.04

TABLE IV
METRICS OF FIVE DISTRIBUTIONS ESTIMATED FOR RAMP MAGNITUDE

Metrics	Distributions				
	Normal	Hyper.	ksdensity	GMM	GGMM
Correlation coeff.	0.88	0.89	0.96	0.97	0.98
Chi - square	4.09	3.49	1.41	1.07	0.56
NRMSE	0.16	0.16	0.09	0.08	0.06
MaxAE	0.41	0.45	0.22	0.27	0.18
MAE	0.12	0.11	0.07	0.06	0.04
KSIPer	33.16	30.14	23.73	17.16	10.28
4RMQE	0.22	0.21	0.12	0.12	0.09

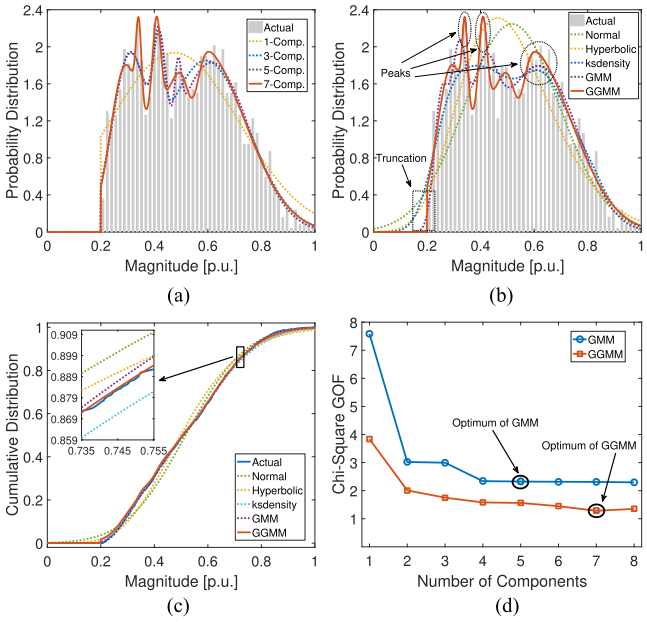


Fig. 3. Comparisons of PDFs, CDFs, and Chi-square goodness-of-fit. (a) PDFs of different components. (b) PDFs of different distributions. (c) CDFs of different distributions. (d) Chi-square GOF.

D. Statistical Comparisons of Ramping Magnitudes

In addition to ramping duration and rate, the magnitude is another important ramping feature. Three cases have been studied by using the developed GGMM:

Case 1: Comparison of five distributions in a *single region* at a *single time resolution*.

Case 2: Comparison of the PDF fitting performance for the GMM and GGMM in a *single region* at *multiple time resolutions*.

Case 3: Comparison of the PDF fitting performance for the GMM and GGMM in *multiple regions* at a *single time resolution*.

Case 1: The region around Dallas, Texas with 711 wind sites is chosen as a representative to compare the fitting performance. Fig. 3 compares the PDF, CDF, and Chi-square goodness-of-fit for five distributions. As shown in Fig. 3(a), seven Gaussian components are found to optimally fit the distribution of ramping magnitude. There are totally 21 parameters (3×7) in the GGMM. For the actual histogram distribution, there are three peaks located around 0.32 p.u., 0.41 p.u., and 0.65 p.u., as shown in Fig. 3(b). It means that the wind power output changes occur at these three values with a high probability, which are more critical and informative for power system operations. For example, ramping reserve requirements [25] could be designed by considering these three peak values instead of only one peak. Since the normal and hyperbolic distributions conform to the unimodal distribution, only one single peak (0.5 and 0.45 p.u., respectively) is depicted with the highest probability. This ill-information may mislead power system operators to mainly cope with the WPRs with magnitudes spanning from 0.45 p.u. to 0.5 p.u., and neglect other significant wind power ramps around 0.32 p.u., 0.41 p.u., and 0.65 p.u.. Though the nonparametric model (ksdensity) is well-known in fitting the irregular distributions, it presents a worse performance than both the GMM and GGMM from visual inspection. Besides, the nonparametric nature of ksdensity restricts its application in practice. This is because the analytical expressions of both PDF and CDF of the distribution are generally required in stochastic power system operations, such as chance-constrained constraints in ED or UC problems. The comparison of CDF fitting performance is illustrated in Fig. 3(c). It is also shown that both the GMM and GGMM follow the actual CDF curve significantly better than other distributions.

To compare the fitting performance of the GMM and GGMM, the Chi-square (χ^2) statistics [14] is used to measure the goodness-of-fit. A smaller value indicates a better fit. As seen in Fig. 3(d), the GGMM has a smaller Chi-square goodness-of-fit than the GMM when using the same number of mixture components. Note that the optimal number of mixture components of the GMM is five in this case study.

Another interesting finding in fitting the actual histogram distribution is the truncation part as shown in the left tail area in Fig. 3(b). Due to the definition of WPRs (20% of the rated capacity), all the ramping magnitudes are greater than or equal to 0.2 p.u.. For the ramping magnitude less than 0.2 p.u., the occurrence probability should be zero. Under this circumstance, the GGMM distribution performs much better than any other distributions due to its intrinsic truncation process, which also verifies the effectiveness of the GGMM.

For quantitative comparison, Table IV lists the fitting metrics for different distributions. Regarding the correlation coefficient metric in green, the GGMM shows the largest value. Regarding other metrics in blue, the GGMM shows the smallest values. This numerically indicates that the GGMM outperforms other

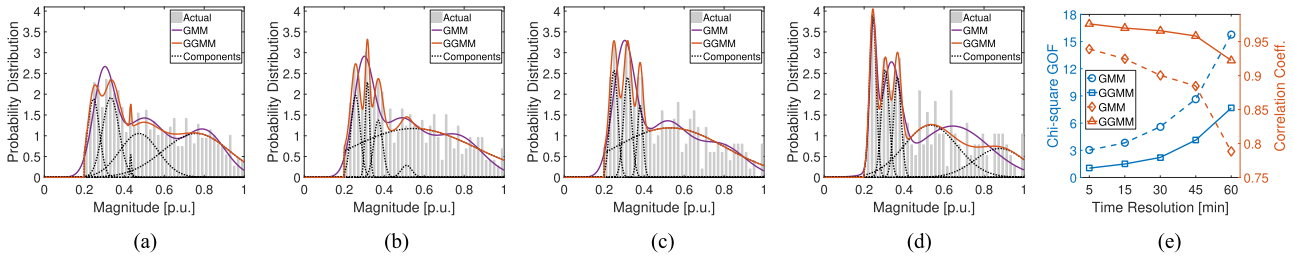


Fig. 4. Comparisons of the PDF fitting performance and metrics of the GMM and GGMM in Dallas region at multiple time resolutions. (a) 15-minute. (b) 30-minute. (c) 45-minute. (d) 1-hour. (e) Metrics.

parametric distributions, and even performs better than the non-parametric distribution (ksdensity) as a parametric model.

Case 2: Fig. 4 compares the PDF fitting performance of the GMM and GGMM in Dallas region at multiple time resolutions: 15-, 30-, 45-minute, and 1-hour. Different mixture components are also illustrated in this figure. It is seen that the GMM presents a worse performance of fit than the GGMM from visual inspection. Numerical comparisons of two evaluation metrics are shown in Fig. 4(e) (Chi-square GOF at the left axis and correlation coefficient at the right axis). For all time resolutions, the GGMM shows smaller Chi-square GOFs and larger correlation coefficients than the GMM, which verifies the effectiveness of the GGMM distribution. We also find that the Chi-square GOF values increase and the correlation coefficient values decrease along with the increase of time resolution. It indicates that the increasing time resolution makes the fitting performance worse. This is because the total number of WPRs is reduced when the time resolution increases. The reduced number of WPRs makes the actual histogram distribution more irregular and fluctuate. Thereby it is more challenging to fit.

Another interesting finding is the change of distribution shapes with the increase of time resolution. For the 1-hour time resolution, more ramping magnitudes are concentrated in the range [0.2, 0.4] with high probabilities. For the 15-minute time resolution, the distribution is not as concentrated as the 1-hour resolution case around lower peaks of probability. It means WPRs are relatively more uncertain and unstable at the small time resolution. In power system operations, the real-time economic dispatch (RTED) is run at a 5-minute time resolution; and the day-ahead unit commitment (DAUC) is run at a 1-hour time resolution [26]. Under this circumstance, WPRs could impact the scheduling in RTED more significantly than that in DAUC. This finding would help balancing authorities to make better plans to manage WPRs with high renewable penetrations, such as utilizing the ramping feature information (e.g., ramping product market design [22]) in system operations, or mitigate ramping events (e.g., ramping control using energy storage systems [27]).

Case 3: Table V shows the parameters of GGMM components for five regions. Negative Components are highlighted for cases: Dallas-10GW and Miami-10GW. The results of PDF fitting performance for the GMM and GGMM in the five regions are illustrated in Fig. 5. Individual components of each GGMM distribution are also shown in this figure. From visual inspection, the GGMM presents a better fitting performance than the GMM. The GGMM is also capable of fitting one negative component

TABLE V
PARAMETERS OF GGMM COMPONENTS FOR FIVE REGIONS: NEGATIVE COMPONENTS ARE HIGHLIGHTED

Cases	GGMM Components		
	Weight Coefficients (ω)	Mean Values (μ)	Standard Deviations (σ)
D-2GW	(1.93, 0.60, 1.45, 0.97, 1.35)	(0.28, 0.37, 0.46, 0.62, 0.80)	(0.09, 0.01, 0.10, 0.08, 0.14)
C-2GW	(1.52, 0.94, 1.22, 1.43)	(0.26, 0.39, 0.33, 0.42)	(0.03, 0.01, 0.04, 0.44)
N-2GW	(2.53, 2.84, 1.76, 1.16, 0.72)	(0.31, 0.26, 0.36, 0.55, 0.44)	(0.02, 0.04, 0.04, 0.26, 0.06)
L-2GW	(1.11, 1.93, 0.68, 1.45, 1.98)	(0.46, 0.31, 0.41, 0.25, 0.48)	(0.01, 0.07, 0.03, 0.01, 0.20)
M-2GW	(1.48, 2.46, 1.17, 2.72)	(0.30, 0.34, 0.44, 0.27)	(0.01, 0.08, 0.21, 0.04)
D-10GW	(1.04, 0.844, 1.36, **-0.52**, 2.01)	(0.41, 0.34, 0.29, 0.54, 0.57)	(0.02, 0.01, 0.07, 0.03, 0.23)
C-10GW	(2.24, 1.02, 1.60, 0.56, 0.64, 1.29)	(0.26, 0.45, 0.36, 0.53, 0.82, 0.59)	(0.04, 0.03, 0.06, 0.01, 0.07, 0.22)
N-10GW	(2.43, 1.61, 1.17, 1.20, 0.84)	(0.27, 0.33, 0.38, 0.49, 0.46)	(0.04, 0.02, 0.03, 0.31, 0.09)
L-10GW	(1.28, 1.13, 2.24, 1.93, 1.71, 0.57, 1.59)	(0.55, 0.46, 0.31, 0.67, 0.42, 0.59, 0.51)	(0.01, 0.004, 0.07, 0.11, 0.04, 0.01, 0.06)
M-10GW	(3.82, 2.50, 1.64, 1.07, **-0.79**)	(0.28, 0.36, 0.41, 0.52, 0.27)	(0.04, 0.02, 0.09, 0.21, 0.005)

Note that the abbreviations of "D", "C", "N", "L", and "M" represent the regions of "Dallas", "Chicago", "New York", "Los Angeles", and "Miami", respectively.

for the cases of Dallas-10GW (Fig. 5(f)) and Miami-10GW (Fig. 5(j)). The extended negative component could help fit the actual histogram distribution more accurately.

Numerical comparisons of two evaluation metrics are illustrated in Fig. 6 (Chi-square GOF with the bar graph at the left axis and correlation coefficient with the stem graph at the right axis). The range of correlation coefficient is [0.976, 0.994] when using the GGMM, and it decreases to [0.939, 0.987] when using the GMM. The range of Chi-square GOF is [0.61, 9.93] when using the GGMM, and it increases to [2.21, 16.06] when using the GMM. For all different regions, the GGMM shows smaller Chi-square GOFs and larger correlation coefficients than the GMM, which verifies the effectiveness of the GGMM distribution.

Moreover, comparing to the distributions of ramping duration and rate in Section IV-C, it is shown that the developed GGMM performs significantly well at fitting the probability distributions of ramping magnitude. For the ramping duration and rate, the GGMM can provide better fitting performance than unimodal distributions (the normal and hyperbolic), and equal-to-better

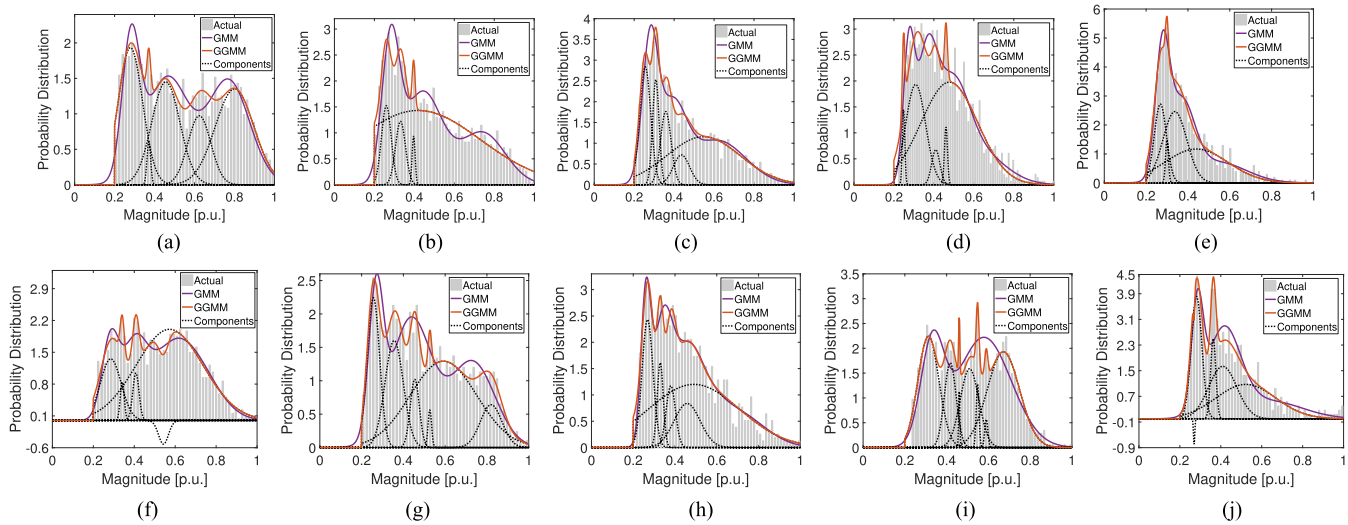


Fig. 5. PDF fitting performance for the GMM and GGMM in multiple regions with two rated capacities. (a) DAL-2GW. (b) CHI-2GW. (c) NYC-2GW. (d) LA-2GW. (e) MIA-2GW. (f) DAL-10GW. (g) CHI-10GW. (h) NYC-10GW. (i) LA-10GW. (j) MIA-10GW.

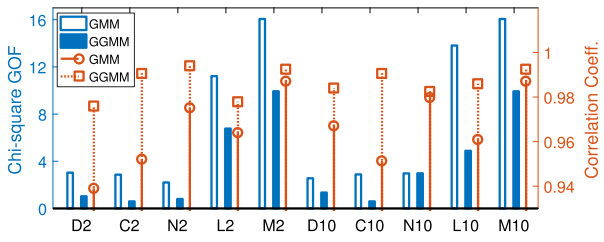


Fig. 6. PDFs of ramping magnitude using GMM and GGMM in five regions.

performance comparing to the GMM and the nonparametric ksdensity distribution.

E. GGMM Application I: Random Number Generator

The developed RNG-GGMM is used to generate a large number of random ramping features. To evaluate the effectiveness of the RNG-GGMM, a quantile-quantile (Q-Q) plot is used to compare the quantiles of the generated random number versus the quantiles of the actual ramping feature data. An accurate distribution ensures that the Q-Q plot appears linear and close to the diagonal line. Taking the ramping magnitude as an example, Fig. 7 compares the Q-Q plots of five distributions using the RNG. It is shown that the RNG-GGMM is distributed significantly close to the diagonal line, especially for the intervals at the bottom left and top right corners. This is because the analytical CDF of GGMM accurately fits the actual CDF as shown in Fig. 3(c).

F. GGMM Application II: Probabilistic Reserve Sizing

For power system operations, statistical analysis has been used for dynamic reserve sizing based on fitted CDFs [28], [29]. In this case, the developed GGMM can also be used to characterize the CDF of wind power forecasting errors (WPFEs). Both the dynamic upward (R^{up}) and downward (R^{dn}) reserves are sized to cover a certain cumulative probability of WPFEs termed as the design reliability (DR), as shown in Fig. 8. Required reserves are dynamically formulated and determined

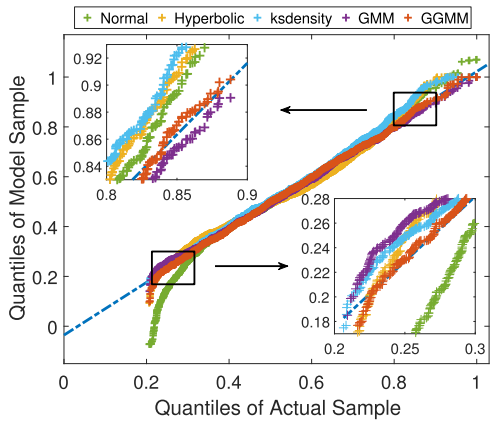


Fig. 7. Q–Q plots of different distributions using the RNGs.

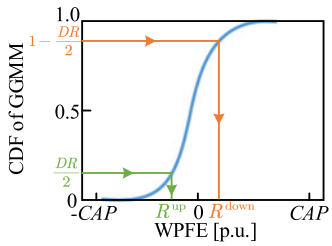


Fig. 8. Diagram of determining the dynamic reserves using the CDF of GGMM. *CAP* is the installed wind power capacity.

by:

$$F\left(R^{\mathrm{up}}\right) \leq \frac{DR}{2} \Rightarrow R^{\mathrm{up}} \leq F^{-1}\left(\frac{DR}{2}\right) \tag{26}$$

$$R^{\mathrm{up}} = \min\left(0,\ F^{-1}\left(\frac{DR}{2}\right)\right) \tag{27}$$

$$F\left(R^{\mathrm{dn}}\right) \geqslant 1 - \frac{DR}{2} \Rightarrow R^{\mathrm{dn}} \geqslant F^{-1}\left(1 - \frac{DR}{2}\right) \tag{28}$$

$$R^{\mathrm{dn}} = \max\left(0,\ F^{-1}\left(1 - \frac{DR}{2}\right)\right) \tag{29}$$

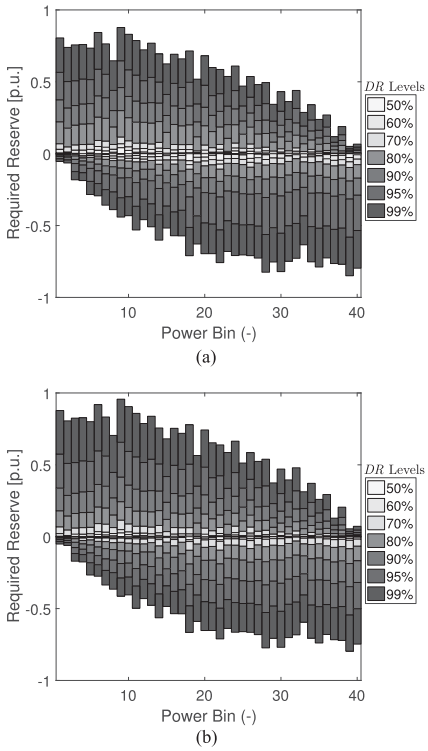


Fig. 9. Downward (positive values) and upward (negative values) reserves per power bin (40 bins in total) for different DR levels (50%, 60%, 70%, 80%, 90%, 95%, and 99%). (a) Original data. (b) GGMM distribution.

where the CDF F is analytically characterized by the developed GGMM distribution model based on WPFEs. The required reserves are numerical solutions of the inverse CDF F^{-1}, which are iteratively calculated by the Newton-Raphson method based on the analytical CDF F.

Based on the aforementioned probabilistic reserve sizing method in power system operations, the requirement amount of upward and downward reserves are calculated to ensure a certain DR level. The wind power generation and forecasts data are collected from WIND Toolkit [24] with 631,296 samples. The actual reliability is defined as the ratio of (a) the number of time steps that sufficient capacity is scheduled to cover the observed wind power forecasts, to (b) the total number of samples in the time series [28], [29]. Fig. 9 shows the results of required downward (positive values) and upward (negative values) reserves per power bin (40 bins in total) using the original data and the developed GGMM distribution. The DR levels are chosen in a range from 50% to 99%. The required reserves directly calculated by the original data are slightly different from those calculated by the GGMM distribution, due to the difference in fitting the probability distributions of both the PDF and CDF. In addition, both results show that higher reliability levels require more reserves.

Table VI compares the actual reliability with the DR based on four distributions (normal, hyperbolic, GMM, and GGMM). The normal distribution reaches only one low DR level (60%). The hyperbolic distribution reaches two low DR levels (50% and 60%). The GMM reaches three DR levels (50%, 60%, and 70%). When using the original data, the highest DR level that can be reached is 90%. For higher DR levels (95% and 99%),

TABLE VI
DESIGN AND ACTUAL RELIABILITY FOR RESERVES BASED ON THE DEVELOPED GGMM DISTRIBUTION

DR [%]	Actual Reliability [%]				
	Normal	Hyperbolic	GMM	Original Data	GGMM
50	49.9431	**50.6248**	**50.2324**	**50.2323**	**50.1244**
60	**60.0138**	**60.2182**	**60.1542**	**60.3974**	**60.7414**
70	68.2173	69.1524	**70.5149**	**70.2232**	**70.3143**
80	78.3257	79.2654	79.2437	**80.3394**	**80.1242**
90	88.4754	89.3449	89.1274	**90.2246**	**90.7824**
95	94.3871	94.5874	94.2784	94.9876	**95.1354**
99	98.6418	98.2551	98.6752	98.8669	**99.3548**

TABLE VII
DESIGN AND ACTUAL RELIABILITY FOR RESERVES WITH DIFFERENT NUMBERS OF SAMPLES (6 K, 60 K, AND 600 K)

DR [%]	Actual Reliability [%]					
	Original Data			GGMM		
	6 k	60 k	600 k	6 k	60 k	600 k
50	**50.0763**	**50.1239**	**50.2323**	**50.0653**	**50.1059**	**50.1244**
60	58.3489	**60.2874**	**60.3974**	**60.2123**	**60.4294**	**60.7414**
70	68.8533	**70.1987**	**70.2232**	**70.0853**	**70.1686**	**70.3143**
80	78.8564	79.3443	**80.3394**	**80.0654**	**80.1019**	**80.1242**
90	88.4634	89.6781	**90.2246**	**90.1743**	**90.2864**	**90.7824**
95	93.8341	94.5345	94.9876	94.8879	**95.0757**	**95.1354**
99	97.6217	98.7452	98.8669	98.1864	98.8649	**99.3548**

only the GGMM distribution yields sufficient reserves to ensure the theoretical DR level. Comparing to other distributions (normal, hyperbolic, and GMM), the GGMM can fit tails and peaks of the WPFE probability distribution, which are critically important in the determination of reserves, especially for higher DR levels. Comparing to the original data, the GGMM can calculate the required reserves based on the continuity of its analytical PDF and CDF with highly accurate fitting performance. However, the original data consist of finite samples with a discrete probability distribution (rather than continuous), which affects the accuracy of the inverse CDF in (27) and (29). Table VII compares the results with different numbers of samples. More DR levels are reached with the increase of the number of samples. When the number of samples is $\sim$ 6 k, the original data reaches only one low DR level (50%). When the number of samples is $\sim$ 600 k, higher DR levels are reached (50%–90%). The reached DR levels using the original data are very sensitive to the number of samples, since smaller samples of original data can only generate fewer discrete probability values for both the PDF and CDF. Comparing to the original data, the reached DR levels using the GGMM distribution are less sensitive to the number of samples. This is because the analytical PDF and CDF of the GGMM distribution can accurately fit the probability distribution of the original data even with fewer samples, such as $\sim$ 6 k samples.

V. DISCUSSION AND EXTENSION

It is still challenging to directly apply the GMM distribution into the power system operation models, even though the op-

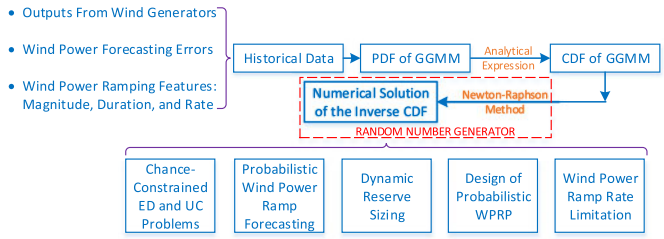


Fig. 10. Schematic diagram of the applications of the GGMM distribution.

erators are informed well about the essential meaning of the positive weights of the GMM distribution. From the perspective of practical applications, the fitting accuracy that directly affects the economic and reliability benefits of power system operations will be more important. In this paper, we have performed sufficient case studies to verify that the GGMM distribution shows a significantly better fitting performance than the GMM distribution. This is because the GGMM is no longer subject to the GMMs constraints: (i) all the weights must be nonnegative; (ii) the sum of all weights must equal to one. More mixture Gaussian components are adopted in the developed GGMM to help better fit the actual histogram distribution. From the perspective of practical applications, the essential meaning of *negative weights* of the developed GGMM is to help ensure a better fitting performance compared to the GMM distribution. The research in this paper can be further extended. First, the developed GGMM distribution model can be used in applications of power system operations. Two representative examples are discussed here: chance-constrained ED/UC problems, and probabilistic wind power ramp forecasting. Second, we discuss how to use the GGMM-based ramping feature models in the development of a new ramping reserve product.

The schematic diagram of the applications of the GGMM distribution is shown in Fig. 10. Historical data can be collected from outputs from wind generators, wind power forecasting errors, and wind power ramping features. The PDF of GGMM is formulated based on these historical data. The CDF of GGMM is analytically expressed as the integral of the GGMM PDF. Since the inverse CDF of GGMM cannot be analytically deduced, the Newton-Raphson method is used to obtain numerical solutions based on the analytical GGMM CDF. This is also termed as the random number generator (see Section IV-E). Finally, the numerical solutions can be used in applications of power system operations as follows.

A. Extension I: Application of the Developed GGMM

1) Chance-Constrained Optimization Problem: Chance-constrained ED and UC problems are widely studied in the literature [9], [30]–[32]. Taking Reference [30] as an example, the probabilistic constraints of the regulation reserve are expressed by the chance-constrained formulas, given by:

$$\Pr\left\{\sum_{i=1}^{NI} r_i \geq \sum_{j=1}^{NW}\left(w_j - w_{av,j}\right)\right\} \geq c \tag{30}$$

where r_i represents the available amount of regulation reserve provided by conventional generator i with the total number NI. w_j is the power output of wind generator (or wind farm) j with the total number NW. $w_{av,j}$ is the actual available power output of wind generator (or wind farm) j. $\Pr\{\cdot\}$ represents the probability that should be greater than the confidence level c. To retain a sequential linear programming formulation, the chance-constraint in (30) is converted to a linear and deterministic expression, deduced as:

$$\sum_{j=1}^{NW} w_j - \sum_{i=1}^{NI} r_i \leq G^{-1}(1-c) \tag{31}$$

where G is the analytical CDF of outputs of wind generators. G^{-1} is the inverse CDF of the GGMM distribution. $G^{-1}(1-c)$ is a constant (numerical solution of the inverse function G^{-1}) which is preprocessed before running ED and UC optimization models. It can be iteratively calculated by the Newton-Raphson method based on the analytical CDF G, as illustrated in Algorithm 1.

2) Probabilistic Wind Power Ramp Forecasting: The developed GGMM can also be used in probabilistic WPR forecasting in both the real-time market (5-minute time resolution) and the day-ahead market (1-hour time resolution). For the real-time market, it takes approximately 3$\sim$ 4 minutes to generate the ramping forecasts, which will be updated every 15 minutes. For the day-ahead market, it takes approximately 10$\sim$ 15 minutes to generate the ramping forecasts, which will be updated every 24 hours. More detailed information on the GGMM-based probabilistic wind power ramping forecast can be found in [33]. First, an ensemble machine learning technique is developed to forecast the basic wind power forecasting scenario and calculate the historical forecasting errors. The GGMM distribution is used to fit the PDF of forecasting errors. The inverse transform method based on Monte Carlo sampling and the analytically deduced CDF is used to generate a massive number of forecasting error scenarios. A wind power forecasting scenario is generated by adding the basic ensemble forecasting data with each individual forecasting error scenario. Each individual wind forecast is put into the OpSDA algorithm to extract all the significant wind power ramps. Therefore, probabilistic wind power ramping forecasts are generated and analyzed, such as the ramping magnitude and start-time.

B. Extension II: Application of GGMM-Based Probabilistic Ramping Features

1) Ramping Magnitude: The GGMM-based probabilistic ramping magnitude could be applied to design new types of reserve products, such as the design of probabilistic wind power ramping product (WPRP). Taking the upward WPRP as an example based on the optimization model in [22], the ramping magnitude could be used to formulate the probabilistic constraints of upward flexible ramping reserve requirement (32) and then deduced to be linear and deterministic by using the

GGMM distribution in (33).

$$\begin{cases} \Pr\left\{\sum_{i=1}^{NI} fu_i + \sum_{j=1}^{NW} UP_j^{\mathrm{M}} \geq UR\right\} \geq c_{\mathrm{M}} \\ \Pr\left\{\sum_{i=1}^{NI} fu_i + \sum_{j=1}^{NW} UP_j^{\mathrm{M}} \geq UR\right\} \\ = 1 - \Pr\left\{UR - \sum_{i=1}^{NI} fu_i \geq \sum_{j=1}^{NW} UP_j^{\mathrm{M}}\right\} \\ = 1 - F_{\mathrm{M}}\left(UR - \sum_{i=1}^{NI} fu_i\right) \end{cases} \tag{32}$$

$$\Longrightarrow UR - \sum_{i=1}^{NI} fu_i \leq F_{\mathrm{M}}^{-1}(1 - c_{\mathrm{M}}) \tag{33}$$

where fu_i is the scheduled flexible up-ramping reserve of conventional generator i. UP_j^{M} is the up-WPRP provided by the ramping magnitude of wind generator (or wind plant) j. UR is the total flexible up-ramping reserve requirements of the system. F_{M}^{-1} is the inverse function of the GGMM CDF, F_{M}, of ramping magnitude. c_{M} is the confidence level which the probability of the ramping magnitude constraint should be greater than.

2) Ramping Duration and Rate: The power system operators need to decide at what rates ($\overline{R}$) the WPRs should be limited during operations. The GGMM distributions of ramping duration and rate could be used in power system optimization models (such as the model proposed in [10]) in an analytical manner. Given a ramp with random variables of ramping duration D and rate R, the compensation price ρ_1 (MW/hour) of the curtailed wind power when ramping rate is greater than the rate limit $\overline{R}$, and the cost price ρ_2 (MW/hour) of fast generators to compensate a WPR with the rate limit $\overline{R}$, the total cost compensated to the wind plant is $\rho_1(R-\overline{R})_+ D^2$ and the total cost of fast generators is $\rho_2(\overline{R}-R)_+ D^2$. The optimization model is to determine an optimal $\overline{R}$ that minimizes the total expected cost, given by:

$$\min\ E\left[\rho_1(R-\overline{R})_+ D^2 + \rho_2(\overline{R}-R)_+ D^2\right] \tag{34}$$

This problem can be solved by using the GGMM distributions of D and R. Assuming that the random variables R and D^2 are independent, this problem is deduced as:

$$\min\ (\rho_1 - \rho_2) E[R]\, E\left[D^2\right] + (\rho_2 - \rho_1)\overline{R} E\left[D^2\right] \tag{35}$$

where the expect value of R is $E[R] = \int_{-\infty}^{+\infty} x f_{\mathrm{R}}(x)\mathrm{d}x$ and f_{R} is the GGMM PDF of ramping rate. Taking a two component GGMM as an example, $E[R]$ can be analytically expressed as:

$$\begin{aligned} E[R] &= \int_{-\infty}^{+\infty} x\left[\omega_1 e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} + \omega_2 e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}}\right]\mathrm{d}x \\ &= \frac{\omega_1\mu_1\sigma_1\sqrt{\pi}}{2i} erfi\left[\frac{(x-\mu_1)i}{\sigma_1}\right] - \frac{\omega_1\sigma_1^2}{2} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \\ &\quad + \frac{\omega_2\mu_2\sigma_2\sqrt{\pi}}{2i} erfi\left[\frac{(x-\mu_2)i}{\sigma_2}\right] - \frac{\omega_2\sigma_2^2}{2} e^{-\frac{(x-\mu_2)^2}{2\sigma_2^2}} \end{aligned} \tag{36}$$

where $efri$ is the imaginary error function and expressed as $efri(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{t^2}dt$. The expect value of D^2 is $E\left[D^2\right] = \int_{-\infty}^{+\infty} y^2 f_{\mathrm{D}}(y)\mathrm{d}y$ and f_{D} is the GGMM PDF of ramping duration. Similarly taking a two component GGMM as an example, $E\left[D^2\right]$ can be analytically expressed as:

$$\begin{aligned} &E\left[D^2\right] \\ &= \int_{-\infty}^{+\infty} y^2\left[\omega_1 e^{-\frac{(y-\mu_1)^2}{2\sigma_1^2}} + \omega_2 e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}\right]\mathrm{d}y \\ &= \frac{\omega_1\sigma_1\sqrt{\pi}(2\mu_1^2+\sigma_1^2)}{4i} erfi\left[\frac{(y-\mu_1)i}{\sigma_1}\right] - \frac{\omega_1\sigma_1^2 y}{2} e^{-\frac{(y-\mu_1)^2}{2\sigma_1^2}} \\ &\quad + \frac{\omega_2\sigma_2\sqrt{\pi}(2\mu_2^2+\sigma_2^2)}{4i} erfi\left[\frac{(y-\mu_2)i}{\sigma_2}\right] - \frac{\omega_2\sigma_2^2 y}{2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}} \\ &\quad - \frac{\omega_1\mu_1\sigma_1^2}{2} e^{-\frac{(y-\mu_1)^2}{2\sigma_1^2}} - \frac{\omega_2\mu_2\sigma_2^2}{2} e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}} \end{aligned} \tag{37}$$

VI. Conclusion

This paper developed a generalized Gaussian mixture model (GGMM) to characterize the probability density functions (PDFs) and cumulative distribution functions (CDFs) of wind power ramping features. The GGMM was analytically expressed as a parametric form. First, the non-linear least square method with the trust-region algorithm was adopted to estimate all the parameters of mixture components. Second, the optimal number of mixture components was adaptively solved by minimizing the Euclidean distance between the GGMM and the actual histogram distribution. The CDF of GGMM was analytically deduced and used to design a random number generator (RNG). Numerical simulations on publicly available wind power data showed that:

1) Among all parametric models, the GGMM performed better than the normal and hyperbolic distributions for all metrics, and presented an equal-to-better performance comparing to the GMM.
2) Comparing with the nonparametric ksdensity model, the GGMM could provide an equal-to-better performance but in a parametric way.
3) The GGMM performed better in fitting the probability distribution of ramping magnitude, and equal-to-better performance in fitting the probability distributions of ramping duration and rate, comparing to the GMM.
4) The GGMM could accurately fit the actual CDF. The RNG based on the GGMM also performed effectively.

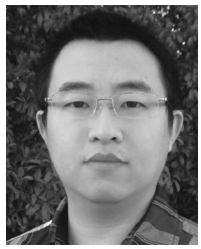
Acknowledgment

The authors would also like to thank the anonymous reviewers for their constructive suggestions to this research.

References

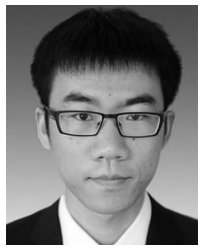
[1] Q. Xu, N. Zhang, C. Kang, Q. Xia, D. He, C. Liu, Y. Huang, L. Cheng, and J. Bai, "A game theoretical pricing mechanism for multi-area spinning reserve trading considering wind power uncertainty," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1084–1095, Mar. 2016.

[2] H. Jiang, Y. Zhang, J. J. Zhang, D. W. Gao, and E. Muljadi, "Synchrophasor-based auxiliary controller to enhance the voltage stability of a distribution system with high renewable energy penetration," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 2107–2115, Jul. 2015.
[3] Y. Qi and Y. Liu, "Wind power ramping control using competitive game," *IEEE Trans. Sustain. Energy*, vol. 7, no. 4, pp. 1516–1524, Oct. 2016.
[4] M. Cui, D. Ke, Y. Sun, D. Gan, J. Zhang, and B.-M. Hodge, "Wind power ramp event forecasting using a stochastic scenario generation method," *IEEE Trans. Sustain. Energy*, vol. 6, no. 2, pp. 422–433, Apr. 2015.
[5] M. Cui, D. Ke, Y. Sun, and D. Gan, "Statistical scenarios forecasting of wind power ramp events using modified neural networks," *J. Mod. Power Syst. Clean Energy*, vol. 3, no. 3, pp. 371–380, Jul. 2015.
[6] Y. V. Makarov, C. Loutan, J. Ma, and P. De Mello, "Operational impacts of wind generation on California Power Systems," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 1039–1050, May 2009.
[7] Q. Wang and B.-M. Hodge, "Enhancing power system operational flexibility with flexible ramping products: A review," *IEEE Trans. Ind. Informat.*, 2016, to be published, doi: 10.1109/TII.2016.2637879.
[8] M. Cui, J. Zhang, C. Feng, A. R. Florita, Y. Sun, and B.-M. Hodge, "Characterizing and analyzing ramping events in wind power, solar power, load, and netload," *Renew. Energy*, vol. 111, pp. 227–244, Oct. 2017.
[9] H. Wu, M. Shahidehpour, Z. Li, and W. Tian, "Chance-constrained day-ahead scheduling in stochastic power system operation," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1583–1591, Jul. 2014.
[10] R. Sevlian and R. Rajagopal, "Detection and statistics of wind power ramps," *IEEE Trans. Power Syst.*, vol. 28, no. 4, pp. 3610–3620, Nov. 2013.
[11] M. Cui, J. Zhang, A. R. Florita, B.-M. Hodge, D. Ke, and Y. Sun, "An optimized swinging door algorithm for identifying wind ramping events," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 150–162, Jan. 2016.
[12] D. Ganger, J. Zhang, and V. Vittal, "Statistical characterization of wind power ramps via extreme value analysis," *IEEE Trans. Power Syst.*, vol. 29, no. 6, pp. 3118–3119, Nov. 2014.
[13] D. Ke, C. Chung, and Y. Sun, "A novel probabilistic optimal power flow model with uncertain wind power generation described by customized Gaussian mixture model," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 200–212, Jan. 2016.
[14] R. Singh, B. C. Pal, and R. A. Jabr, "Statistical representation of distribution system loads using Gaussian mixture model," *IEEE Trans. Power Syst.*, vol. 25, no. 1, pp. 29–37, Feb. 2010.
[15] G. Valverde, A. Saric, and V. Terzija, "Probabilistic load flow with non-Gaussian correlated random variables using Gaussian mixture models," *IET Gener. Transmiss. Distrib.*, vol. 6, no. 7, pp. 701–709, 2012.
[16] J. Zhang, A. Florita, B.-M. Hodge, S. Lu, H. F. Hamann, V. Banunarayanan, and A. M. Brockway, "A suite of metrics for assessing the performance of solar power forecasting," *Sol. Energy*, vol. 111, pp. 157–175, Jan. 2015.
[17] B.-M. Hodge, D. Lew, and M. Milligan, "Short-term load forecasting error distributions and implications for renewable integration studies," in *Proc. IEEE 5th Green Technol. Conf.*, Denver, CO, USA, 2013, pp. 435–442.
[18] J. Zhang, M. Cui, B.-M. Hodge, A. Florita, and J. Freedman, "Ramp forecasting performance from improved short-term wind power forecasting over multiple spatial and temporal scales," *Energy*, vol. 122, pp. 528–541, Mar. 2017.
[19] C. Kamath, "Associating weather conditions with ramp events in wind power generation," in *Proc. Power Syst. Conf. Expo.*, Phoenix, AZ, USA, 2011, pp. 1–8.
[20] J. J. Moré and D. C. Sorensen, "Computing a trust region step," *SIAM J. Sci. Statist. Comput.*, vol. 4, no. 3, pp. 553–572, 1983.
[21] C. Ferreira, J. Gama, L. Matias, A. Botterud, and J. Wang, "A survey on wind power ramp forecasting," Argonne Nat. Lab., DuPage County, IL, USA, Tech. Rep. ANL/DIS-10-13, Dec. 2010.
[22] M. Cui, J. Zhang, H. Wu, and B.-M. Hodge, "Wind-friendly flexible ramping product design in multi-timescale power system operations," *IEEE Trans. Sustain. Energy*, vol. 8, no. 3, pp. 1064–1075, Jul. 2017.
[23] P. Glasserman, *Monte Carlo Methods in Financial Engineering*. New York, NY, USA: Springer, 2003.
[24] C. Draxl, A. Clifton, B.-M. Hodge, and J. McCaa, "The wind integration national dataset (WIND) Toolkit," *Appl. Energy*, vol. 151, pp. 355–366, Aug. 2015.
[25] C. Wang, P. B. Luh, and N. Navid, "Ramp requirement design for reliable and efficient integration of renewable energy," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 562–571, Jan. 2017.
[26] E. Ela and M. O'Malley, "Studying the variability and uncertainty impacts of variable generation at multiple timescales," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1324–1333, Aug. 2012.
[27] B. Zhou and T. Littler, "Local storage meets local demand: A technical solution to future power distribution system," *IET Gener. Transmiss. Distrib.*, vol. 10, no. 3, pp. 704–711, 2016.
[28] K. Bruninx and E. Delarue, "A statistical description of the error on wind power forecasts for probabilistic reserve sizing," *IEEE Trans. Sustain. Energy*, vol. 5, no. 3, pp. 995–1002, Jul. 2014.
[29] K. Bruninx, K. Van den Bergh, E. Delarue, and W. D'haeseleer, "Optimization and allocation of spinning reserves in a low-carbon framework," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 872–882, Mar. 2016.
[30] Z. Zhang, Y. Sun, D. W. Gao, J. Lin, and L. Cheng, "A versatile probability distribution model for wind power forecast errors and its application in economic dispatch," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3114–3125, Aug. 2013.
[31] L. Roald, S. Misra, T. Krause, and G. Andersson, "Corrective control to handle forecast uncertainty: A chance constrained optimal power flow," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1626–1637, Mar. 2017.
[32] D. Pozo and J. Contreras, "A chance-constrained unit commitment with an security criterion and significant wind generation," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2842–2851, Aug. 2013.
[33] M. Cui, C. Feng, Z.Wang, J. Zhang, Q.Wang, A. R. Florita, V. Krishnan, and B.-M. Hodge, "Probabilistic wind power ramp forecasting based on a scenario generation method," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Chicago, IL, USA, 2017, pp. 1–5.



Mingjian Cui (S'12–M'16) received the B.S. and Ph.D. degrees in electrical engineering and automation from Wuhan University, Wuhan, China, in 2010 and 2015, respectively.

He is currently a Research Associate, as a Postdoctoral, with the University of Texas at Dallas, Richardson, TX, USA. He is also a Visiting Scholar with the Transmission and Grid Integration Group, National Renewable Energy Laboratory, Golden, CO, USA, from 2014 to 2015. His research interests include renewable energy forecasting, power system operation and control, unit commitment, economic dispatch, optimization modeling, electricity market, data analytics, and statistical analysis.



Cong Feng (S'17) received the B.S. degree in power engineering from Wuhan University, Wuhan, China, in 2014, and the M.S. degree from the University of Texas at Dallas, Richardson, TX, USA, in 2017. He is currently a Ph.D student of the Department of Mechanical Engineering at the University of Texas at Dallas.

His research interests include wind power forecasting, machine learning, and data analytics.

Zhenke Wang received the B.S. degree in aerospace engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2013. He is currently a master student of the Department of Mechanical Engineering at the University of Texas at Dallas, Richardson, TX, USA.

His research interests include wind power forecasting and energy storage.

Jie Zhang (M'13–SM'15) received the B.S. and M.S. degrees in mechanical engineering from Huazhong University of Science and Technology, Wuhan, China, in 2006 and 2008, respectively, and the Ph.D. degree in mechanical engineering from Rensselaer Polytechnic Institute, Troy, NY, USA, in 2012.

He is currently an Assistant Professor of the Department of Mechanical Engineering at The University of Texas at Dallas. His research interests include multidisciplinary design optimization, complex engineered systems, big data analytics, wind and solar forecasting, renewable integration, energy systems modeling, and simulation.