A Copula-Based Conditional Probabilistic Forecast Model for Wind Power Ramps

Mingjian Cui, *Member, IEEE*, Venkat Krishnan, *Member, IEEE*, Bri-Mathias Hodge, *Senior Member, IEEE*, and Jie Zhang, *Senior Member, IEEE*

***Abstract*—Efficient management of wind ramping characteristics can significantly reduce wind integration costs for balancing authorities. By considering the stochastic dependence of wind power ramp (WPR) features, this paper develops a conditional probabilistic WPR forecast (cp-WPRF) model based on copula theory. The WPRs dataset is constructed by extracting ramps from a large dataset of historical wind power. Each WPR feature (e.g., rate, magnitude, duration, and start-time) is separately forecasted by considering the coupling effects among different ramp features. To accurately model the marginal distributions with a copula, a Gaussian mixture model is adopted to characterize the WPR uncertainty and features. The Canonical maximum likelihood method is used to estimate parameters of the multivariable copula. The optimal copula model is chosen based on the Bayesian information criterion from each copula family. Finally, the best conditions based cp-WPRF model is determined by predictive interval based evaluation metrics. Numerical simulations on publicly available wind power data show that the developed copula-based cp-WPRF model can predict WPRs with a high level of reliability and sharpness.**

***Index Terms*—Conditional probabilistic forecast, copula theory, Gaussian mixture model, wind power ramps.**

ACRONYMS

WPR	Wind power ramp.
WPRF	Wind power ramp forecasting.
SVM	Support vector machine.
p-WPRF	Probabilistic wind power ramp forecasting.
d-WPRF	Deterministic wind power ramp forecasting.
cp-WPRF	Conditional probabilistic wind power ramp forecasting.
OpSDA	Optimized swinging door algorithm.
$\mathcal{R}$	Ramp rate.
$\mathcal{D}$	Ramp duration.
$\mathcal{M}$	Ramp magnitude.
$\mathcal{S}$	Non-ramp duration or ramp start-time.
SWPO	Start-time wind power output.
GMM	Gaussian mixture model.
CDF	Cumulative distribution function.
PDF	Probability distribution function.
cPDF	Conditional probability density function.
CML	Canonical maximum likelihood.
PI	Predictive interval.
RRFE	Ramp rate forecast error.
BIC	Bayesian information criterion.
ACE	Average coverage error.
ASV	Average score value.
SEM	Synthetic evaluation metric.
PICP	Predictive interval coverage probability.
PINC	Predictive interval nominal confidence.

Manuscript received September 29, 2017; revised February 12, 2018 and May 2, 2018; accepted May 25, 2018. Date of publication May 29, 2018; date of current version June 19, 2019. This work was supported by the National Renewable Energy Laboratory under Subcontract XHQ-6-62546-01 (under the U.S. Department of Energy Prime under Contract DE-AC36-08GO28308). Paper no. TSG-01414-2017. *(Corresponding author: Jie Zhang.)*

M. Cui and J. Zhang are with the Department of Mechanical Engineering, University of Texas at Dallas, Richardson, TX 75080 USA (e-mail: mingjian.cui@utdallas.edu; jiezhang@utdallas.edu).

V. Krishnan and B.-M. Hodge are with the Power Systems Engineering Center, National Renewable Energy Laboratory, Golden, CO 80401 USA (e-mail: venkat.krishnan@nrel.gov; bri.mathias.hodge@nrel.gov).

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/TSG.2018.2841932

I. INTRODUCTION

WIND power ramps (WPRs) are caused by large fluctuations in wind speed in a short time period and can threaten the secure and stable operation of power systems [1]–[3]. However, wind power ramp forecasting (WPRF) is still challenging for system operators even though larger wind power penetrations are being seen in power systems worldwide [4], which makes WPRF significant for practical applications.

WPRF methods can be divided into deterministic forecasts (d-WPRF) and probabilistic forecasts (p-WPRF). The recent development of machine learning methods makes it possible to constitute deterministic WPRF models. For example, Liu *et al*. [5] developed a hybrid WPR forecasting model to combine an orthogonal test with a support vector machine (SVM). Zareipour *et al*. [6] presented a direct approach for predicting WPR events by using the SVM classifier. Cutler *et al*. [7] forecasted wind power ramps and evaluated the efficiency of the Wind Power Prediction Tool and the Mesoscale Limited Area Prediction System for ramp forecasts. Probabilistic WPRFs are expected to provide more information on forecasting uncertainties of wind power ramps. Accurate information of p-WPRF can help power system operations, such as (i) enabling credible wind-friendly flexible ramping products design in both unit commitment and economic dispatch models [8], and (ii) significantly helping

reduce wind integration cost by enabling wind power dispatchability and efficiently managing wind ramping characteristics. Most of existing p-WPRF methods can be further divided into two categories: two- and one-step methods. The two-step p-WPRF method is to first generate a massive number of wind power scenarios by using wind power forecasting techniques, and then detect all the possible WPRs. The probabilistic information of WPRF is obtained through the statistical analysis of those detected WPRs. For example, Li *et al*. [9] provided probabilistic WPR information using wind power scenarios generated from quantile forecasts. Cui *et al*. [10] utilized Neural Networks to generate wind power scenarios and derive the probabilistic distributions of ramp features. Due to the high dependence on wind power scenarios, the two-step p-WPRF method is generally computational expensive. The one-step p-WPRF method is to directly forecast ramp features based on historical measured WPR information and can eliminate impacts of generated wind power scenarios. For example, Taylor [11] used a multinomial logit structure and categorical distribution to directly estimate the ramp event probabilities for different thresholds.

Generally, WPRs can be characterized with four features: ramp rate, duration, magnitude, and start-time. However, most of current literature focuses on ramp rate forecasts while neglecting the stochastic dependence between different ramp features. To this end, a one-step copula-based conditional p-WPRF (cp-WPRF) model is developed in this paper, which is able to model conditional probabilistic forecasts of WPR features. Copula theory has been widely applied in dependence and uncertainty analyses. Zhang *et al*. [12] utilized the copula theory to formulate the conditional distributions of multiple wind farms' forecasting errors. Wang *et al*. [13] used the copula theory to describe the wind farms' dependence structure under multiple forecasting conditions. Copula theory has also been used for power system uncertainty analysis. For example, Wang *et al*. [14] utilized the high-dimensional copula theory for power system reserve requirement evaluation and wind power capacity credit assessment analysis.

In this paper, we seek to address two critical questions for balancing authorities with increasing wind power penetrations in power systems. Is it possible to quantitatively evaluate the probabilistic information of WPR features (e.g., rate, duration, magnitude, and start-time)? What is the impact of the stochastic dependence between WPRF uncertainties and different ramp features? This paper develops a cp-WPRF model to characterize key ramp features. The main contributions of this paper include: (i) using the Gaussian mixture model (GMM) to accurately fit the probability distributions of WPRF errors and ramp features; (ii) using copula to develop a cp-WPRF model to separately forecast each WPR feature considering the stochastic dependence of ramping features; and (iii) analyzing the probability information of conditional forecasts for ramp features.

The remainder of this paper is organized as follows. In Section II, a database of historical WPRs is constructed and ramp features are characterized. The stochastic dependence between WPRF uncertainty and variability is modeled in Section III. The development of the cp-WPRF model and evaluation metrics are described in Sections IV and V, respectively. Case studies and analysis performed on publicly available wind power data are present in Section VI. Concluding remarks are summarized in Section VIII.

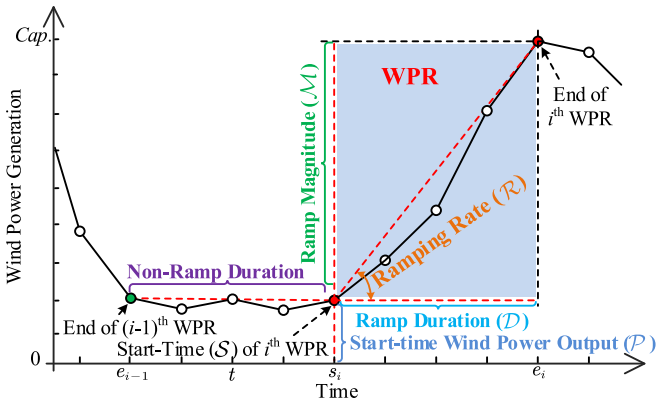


Fig. 1. A typical WPR represented by different ramp features.

II. Data Preparation and Characterization

Before performing the conditional forecasts of WPRs, a database of historical ramping features is constructed in Section II-A. The SVM method is used to generate d-WPRFs and construct a dataset of forecasting errors of different ramping features, as described in Section II-B. Uncertainty and variability of WPRFs are characterized in Section II-C for marginal distributions that can be used in copula-based conditional forecasts.

A. WPR Detection

The WIND Toolkit [15] is utilized to construct the historical WPRs database, and the optimized swinging door algorithm (OpSDA) [16] is used to automatically detect WPRs. In the OpSDA, the conventional swinging door algorithm (SDA) with a predefined value 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 ramp direction and relatively high ramp rates. A brief description of the OpSDA is introduced in this section, and more details can be found in [16]. A brief example of one WPR is illustrated in Fig. 1. As can be seen, one WPR consists of four ramp features and one auxiliary variable. The four ramp features are ramp rate, duration, magnitude, and non-ramp duration/start-time, which are represented by symbols $\mathcal{R}$, $\mathcal{D}$, $\mathcal{M}$, and $\mathcal{S}$, respectively. The auxiliary variable is the start-time wind power output (SWPO) represented by the symbol $\mathcal{P}$. Note that the ramp start-time can be calculated from the non-ramp duration. Since ramp features are more practical for power system operations than the auxiliary variable, the developed WPRF model in this paper is to forecast each ramp feature. Since four features and one auxiliary variable constitute one WPR, the stochastic dependence among them needs to be modeled to forecast any ramp feature. To characterize the mutual dependence of WPR features and the auxiliary variable, Copula

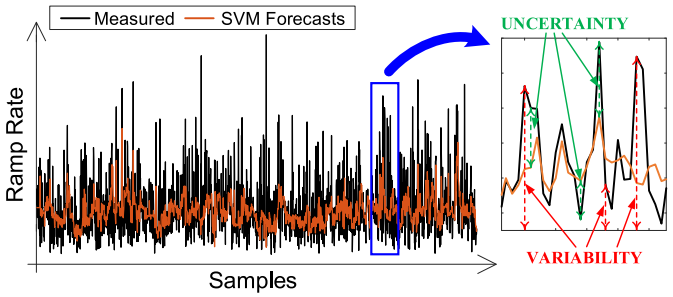


Fig. 2. Deterministic forecasting results of ramp rate using SVM.

theory is adopted for analytical analysis and introduced in Section III.

B. ε-SVM Based d-WPRF Model

Based on the historical WPR features data, a classic machine learning method, epsilon-insensitive support vector machine (ε-SVM), is used to generate d-WPRFs. Note that this paper does not aim to develop any new machine learning methods for WPRFs. SVM has been widely used in regression and classification problems [17], as well as for d-WPRFs. Given an observation WPR sample set $\{\mathbf{R}_t, \hat{R}_{t+1}\}_{t=1}^{\mathrm{N_{tr}}}$, each $\mathbf{R}_t \in \mathbb{R}^n$ represents the input variables of WPR samples at the current tth ramp, i.e., $\mathbf{R}_t = \left(R_{t-\mathrm{N_I}+1}, \ldots, R_{t-1}, R_t\right)$, and $\hat{R}_{t+1} \in \mathbb{R}$ represents the corresponding forecast at the next $(t+1)$th ramp with $\mathrm{N_{tr}}$ training points. $\mathrm{N_I}$ is the number of input variables. The nonlinear mapping is established to transfer the input data into a feature space with higher dimensions. The linear regression can be performed in this feature space, given by:

$$f(\mathbf{R}) = \left\langle w^{\mathrm{T}}, K(\mathbf{R}, \mathbf{R}_t)\right\rangle + b \tag{1}$$

where the coefficients w and b are estimated from the data. To map the high-dimensional feature space, the widely used radial basis function (RBF) is adopted as a kernel:

$$K(\mathbf{R}, \mathbf{R}_t) = \exp\left(-\gamma \|\mathbf{R} - \mathbf{R}_t\|^2\right). \tag{2}$$

To deal with infeasible constraints, slack variables ξ_t and ξ_t^* are introduced for each point. By minimizing the following risk function, all the variables are calculated:

$$\left\{w, b, \xi_t, \xi_t^*\right\} = \arg\ \min \frac{1}{2}\|w\|^2 + C\sum_{t=1}^{\mathrm{N_{tr}}}\left(\xi_t + \xi_t^*\right) \tag{3}$$

$$\xi_t,\ \ \xi_t^* \geq 0 \tag{4}$$

$$\hat{R}_{t+1} - \left(\left\langle w^{\mathrm{T}}, K(\mathbf{R}, \mathbf{R}_t)\right\rangle + b\right) \leq \varepsilon + \xi_t \tag{5}$$

$$\left(\left\langle w^{\mathrm{T}}, K(\mathbf{R}, \mathbf{R}_t)\right\rangle + b\right) - \hat{R}_{t+1} \leq \varepsilon + \xi_t^*. \tag{6}$$

As shown in Fig. 2, the SVM only depicts the variability of the ramp rate. However, it does not characterize the uncertainty in WPRF, which is discussed as follows.

C. Uncertainty and Variability Characterization of WPRF

Similar to the wind power, WPR features also present the uncertainty and variability characteristics [18]. WPRF errors are varying over time with different forecasting accuracies, which is the a proxy for the *uncertainty* (see the green brace in Fig. 2). The WPR features change frequently with time, which is the *variability* (see the red brace in Fig. 2). Taking the ramp rate forecast error (RRFE) as an example, Fig. 3 shows the scatter plots of joint distributions of RRFE and two representative WPR features: ramp rate and duration. Fig. 3(a) shows that RRFE is directly proportional to ramp rate, i.e., RRFE $\propto$ Rate. Fig. 3(b) shows that RRFE is inversely proportional to ramp duration, i.e., RRFE $\propto$ 1/Duration. This observation illustrates that RRFE correlates with WPR features, such as ramp rate and duration. However, it is still challenging to model the stochastic dependence analytically when considering multivariate marginal distributions. Though a correlation coefficient can characterize this relationship between RRFE and ramp features, it still cannot capture all of the dependence information. Copulas are efficient at describing the correlations of stochastically dependent variables, and

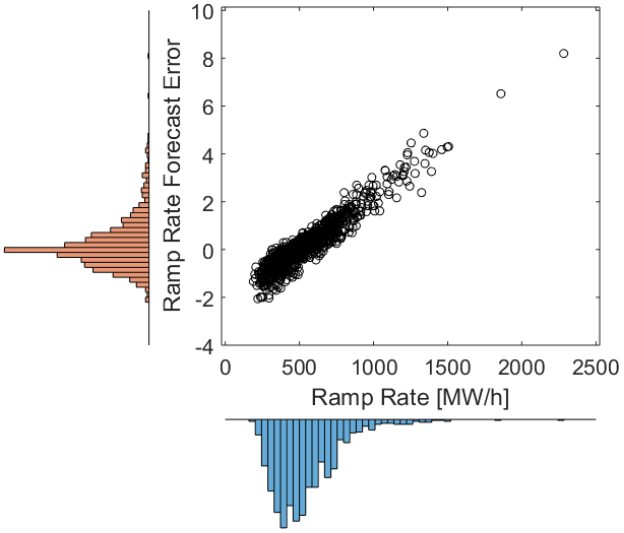


Fig. 3. Scatter plots of joint distributions for RRFE (y-axis) and two representative WPR features (x-axis). (a) joint distribution of RRFE and ramp rate; and (b) joint distribution of RRFE and ramp duration.

have thus been adopted to model the probabilistic relationship between WPRF errors and WPR features.

Currently, most papers in the literature use unimodal distributions (normal distribution [19]) or nonparametric distributions (kernel density estimation [13]) to model the marginal distributions of copula. However, unimodal distributions cannot accurately fit the irregular distributions of WPR features, and nonparametric distributions cannot be solved analytically. To characterize the uncertainty and variability of WPR features, the GMM distribution [20] is used in this paper to accurately model the multimodal probability distributions of WPRF errors and WPR features, respectively. The GMM distribution is formulated by:

$$\begin{aligned} f(x_r|\Gamma) &= \sum_{i=1}^{\mathrm{N_G}} \omega_i g_i(x_r|\mu_i, \sigma_i) \\ r \in \mathscr{R} &= \{\mathcal{R}, \mathcal{D}, \mathcal{M}, \mathcal{S}\} \end{aligned} \tag{7}$$

$$\int_{-\infty}^{+\infty} f(x_r|\Gamma) = \int_{-\infty}^{+\infty} \sum_{i=1}^{\mathrm{N_G}} \omega_i g_i(x_r|\mu_i, \sigma_i) = 1 \tag{8}$$

$$\Longrightarrow \sum_{i=1}^{\mathrm{N_G}} \omega_i \int_{-\infty}^{+\infty} g_i(x_r|\mu_i, \sigma_i) = 1 \tag{9}$$

where $\mathrm{N_G}$ is the total number of mixture components. $\mathscr{R}$ is the WPR features set, including the ramp rate ($\mathcal{R}$), ramp duration ($\mathcal{D}$), ramp magnitude ($\mathcal{M}$), and ramp start-time ($\mathcal{S}$). Γ defines the parameter set of all mixture components, i.e., $\Gamma = \{\omega_i, \mu_i, \sigma_i\}_{i=1}^{\mathrm{N_G}}$. σ is the standard deviation. μ is the mean value. ω is the weight. Each component $g(x_r|\mu, \sigma)$ conforms to a normal distribution, given by:

$$g(x_r|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_r - \mu)^2}{2\sigma^2}\right) \tag{10}$$

where the integral of a normal distribution equals unity. Thus, Eq. (9) becomes:

$$\sum_{i=1}^{\mathrm{N_G}} \omega_i = 1. \tag{11}$$

The GMM distribution has two unity attributes formulated in (8) and (11), which make it possible to use the expectation maximization algorithm to estimate all the parameters of mixture components. More detailed information about this algorithm can be found in [21]. The uncertainty and variability of WPR features are separately characterized by the GMM distribution with specific parameters.

III. Conditional Forecast of Uncertainty With Stochastic Dependence of Variability

Generally, the behavior of WPR uncertainties is affected by WPR variabilities, which is called *stochastic dependence* [12], [19]. In other words, the WPR uncertainty and variability are stochastically dependent on each other. To analytically characterize the stochastic dependence between WPR uncertainty and variability, Copula theory provides an effective way of capturing these correlations [22], [23]. Suppose that x_r is the rth uncertainty variable (WPRF errors) and $x_r \in \{x_{\mathcal{R}}, x_{\mathcal{D}}, x_{\mathcal{M}}, x_{\mathcal{S}}\}$ ($r \in \mathscr{R}$); y_c is the cth random variability (or condition) variable (WPR features and auxiliary variable) and $y_c \in \{y_{\mathcal{R}}, y_{\mathcal{D}}, y_{\mathcal{M}}, y_{\mathcal{S}}, y_{\mathcal{P}}\}$ ($c \in \mathscr{R} \cup \{\mathcal{P}\}$). The joint cumulative distribution function (CDF) $F_{X_rY_c}$ represents the stochastic dependence, given by:

$$F_{X_rY_c}(x_r, y_c) = F_{\mathrm{C}}\big(F_{X_r}(x_r), F_{Y_c}(y_c)\big) \tag{12}$$

where F_{X_r} and F_{Y_c} are the marginal CDFs of WPR uncertainty and variability that transform x_r and y_c into the uniform distributions, respectively. $F_{\mathrm{C}}(\cdot)$ is the copula CDF. In this way, Copula theory transforms the stochastic dependence problem into modeling F_{X_r}, F_{Y_c}, and $F_{\mathrm{C}}(\cdot)$.

A. Single Condition Based cp-WPRF

Because of the stochastic nature of WPR, a change of random WPRF errors would occur when the condition variables are altered, which is regarded as the stochastic dependence of WPR uncertainty with the variability. The joint distribution of WPRF errors and the conditional variable can be modeled by the copula function. The joint probability density function (PDF) $f_{X_rY_c}(x_r, y_c)$ is formulated in (28) in Appendix A-A with the marginal PDFs of x_r, y_c, and copula PDF $f_{\mathrm{C}}(\cdot)$. Given the point forecast of one single conditional variable is $y_c = \hat{R}_c$, the conditional PDF (cPDF) of WPRF errors can be expressed as:

$$\begin{aligned} f_{X_r|Y_c}\left(x_r|\hat{R}_c\right) &= f_{X_rY_c}\left(x_r, \hat{R}_c\right) \Big/ f_{Y_c}\left(\hat{R}_c\right) \\ &= f_{\mathrm{c}}\left(F_{X_r}(x_r), F_{Y_c}\left(\hat{R}_c\right)\right) f_{X_r}(x_r) \end{aligned} \tag{13}$$

where the WPR uncertainty variable (x_r) pool $x_r \in \{x_{\mathcal{R}}, x_{\mathcal{D}}, x_{\mathcal{M}}, x_{\mathcal{S}}\}$ includes the WPR uncertainties of four WPR features. The dependent condition (y_c) pool in $y_c \in \{y_{\mathcal{R}}, y_{\mathcal{D}}, y_{\mathcal{M}}, y_{\mathcal{S}}, y_{\mathcal{P}}\}$ includes all the possible variables that are correlated with the WPR uncertainty. In this paper, the dependent conditional variable pool consists of four WPR features and SWPO ($\mathcal{P}$). Note that the aforementioned pool definitions are not limited by this paper and can be extended by balancing authorities for further studies. Each uncertainty or condition variable is normalized by:

$$x = (x_{\mathrm{meas.}} - \mu)/\sigma \tag{14}$$

where μ and σ represent the mean value and standard deviation of uncertainty or condition variables, respectively.

B. Multiple Conditions Based cp-WPRF

Copula theory can also be used to establish the multiple conditions based cp-WPRF model by expanding (13). The joint PDF $f_{X_rY_1Y_2\cdots Y_c}(x_r, y_1, y_2, \ldots, y_c)$ is formulated in (29) in Appendix A-B. Given point forecasts of multiple conditional variables are $y_1 = \hat{R}_1, y_2 = \hat{R}_2, \ldots, y_c = \hat{R}_c$, the cPDF of WPRF errors, namely $f_{X_r|Y_1Y_2\cdots Y_c}(x_r|\hat{R}_1, \hat{R}_2, \cdots, \hat{R}_c)$, is expressed in (30) in Appendix A-B. Unlike the single condition based cp-WPRF, multiple conditional variables $(y_1, y_2, \ldots, y_c)$ are selected from the dependent conditions

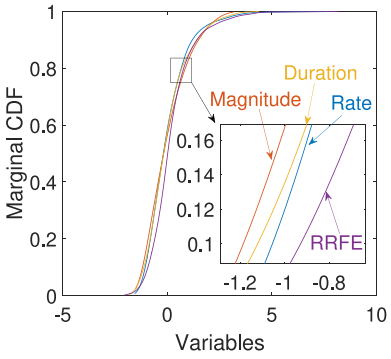


Fig. 4. CDF profiles of ramping features and RRFE.

Algorithm 1 Newton-Raphson Method for Generating the Lower Bounds of WPR Uncertainty

Initialization: obtain a random point ($\rho \in [0, 1]$) and evenly partition the WPRF errors into N_ρ regions ($[\underline{x_{r,\eta}}, \overline{x_{r,\eta}}]_{\eta=1}^{N_\rho}$). Decide the region where ρ is:
if $\overline{F_{X_r|Y_1\cdots Y_c}(x_{r,\eta})} < \rho < F_{X_r|Y_1\cdots Y_c}(\overline{x_{r,\eta}})$ **then**
  Return η, $\overline{x_r^0} = \underline{x_{r,\eta}}$; and the approximation x_r^1 is calculated, when iteration $l = 0$, by:
  if single condition based cp-WPRF **then**
    Choose the optimal copula by (17) and run the iterative formula in (31) in Appendix A-C;
  end
  else if multiple conditions based cp-WPRF **then**
    Choose the optimal copula by (17) and run the iterative formula in (32) in Appendix A-C.
  end
end
Calculate the lower bounds:
for Iteration l from 1 to 100 **do**
  if $|x_r^{l+1} - x_r^l| < \epsilon$ **then**
    The lower bound is returned as: $x_r^{\alpha_L} \approx (x_r^l + x_r^{l+1})/2$;
  else
    The iterative process is repeated by (31) and (32) in Appendix A-C.
  end
end

pool in (15) based on the constraint in (16). By varying the c value, all of the possible conditions can be considered.

$$y_1, y_2, \ldots, y_c \in \{y_{\mathcal{R}}, y_{\mathcal{D}}, y_{\mathcal{M}}, y_{\mathcal{S}}, y_{\mathcal{P}}\};\ c \in \mathscr{R} \cup \{\mathcal{P}\} \tag{15}$$

$$y_1 \neq y_2 \neq \cdots \neq y_c. \tag{16}$$

Overall, the conditional distribution of WPR uncertainty variables consists of the copula-based cPDF as the variant multiplier and the marginal GMM distribution as the base.

C. Optimal Copula Models Determination

To choose the optimal copula model, the Bayesian information criterion (BIC) is used to assess the performance of different copula models [24]. Thus, the optimal copula model is chosen by minimizing the BIC, formulated by:

$$\arg\ \min\ N_P \ln N_S - 2\ln\left[\sum_{t=1}^{N_S} \ln f_C\left(u_t, v_{1,t}, \ldots, v_{c,t}; \hat{\theta}\right)\right] \tag{17}$$

where N_P is the number of parameters in a copula model. N_S is the number of measured samples. For the Gaussian copula, $N_P = c(c+1)/2$. For the t copula, $N_P = 1 + c(c+1)/2$. For the Archimedean copula family, $N_P = 1$.

The optimal copula model is determined by BIC. In statistics, BIC is a criterion used for model selection among a finite set of models. The copula model with the minimum BIC is preferred. The minimum BIC is calculated using CDFs that transform the ramping features data into uniform distributions. Different CDF profiles of each ramping feature may generate a different optimal copula model. Fig. 4 illustrates the CDF profile differences of ramping features and RRFE, which shows that the optimal copula model could be different when used to forecast different ramping features.

IV. Development of CP-WPRF Model: An Example of Ramp Rate Forecasts

Based on the optimal copula model, the predictive intervals (PIs) of WPR uncertainties $[x_{r,t}^{\alpha_L}, x_{r,t}^{\alpha_U}]$ can be calculated by using the cPDF in (13) and (30). A PI ($I_{r,t}^{\beta}$) of the forecasted WPRs with a nominal coverage rate (1-β) can be expressed with the lower bound $\hat{R}_{r,t}^{\alpha_L}$ and the upper bound $\hat{R}_{r,t}^{\alpha_U}$ [25], given by:

$$I_{r,t}^{\beta} = \left[\hat{R}_{r,t}^{\alpha_L}, \hat{R}_{r,t}^{\alpha_U}\right] = \hat{R}_{r,t}^{\text{SVM}} + \underbrace{\left[x_{r,t}^{\alpha_L}, x_{r,t}^{\alpha_U}\right]}_{\text{UNCERTAINTY}} \tag{18}$$

where the lower and upper nominal proportions α_L and α_U equal to $\beta/2$ and $(1 - \beta/2)$, respectively. However, the inverse function of the cCDF cannot be analytically deduced. Alternatively, we use the Newton-Raphson method [20] to obtain the numerical solution. Taking the ramp rate forecasts as an example, the lower bound of ramp rate uncertainty is generated by using the pseudocode in Algorithm 1 based on the copula PDF and the lower nominal proportion α_L. The overall framework for generating the cp-WPRF of ramp rate is illustrated in Fig. 5, which mainly consists of four major steps: deterministic ramp rate forecast, marginal distribution fit, optimal copula model selection, and determining the best conditions based cp-WPRF model. The four major steps are described as follows:

- *Step 1:* Based on the measured WPR data of ramp features, a machine learning method (i.e., SVM) is used to separately generate deterministic forecasts of all WPR features and the corresponding WPRF errors (see Section II-B).
- *Step 2:* Each WPR feature and its forecasting errors are characterized by the GMM distribution (see Section II-C), which is used as the marginal distribution in copula models.

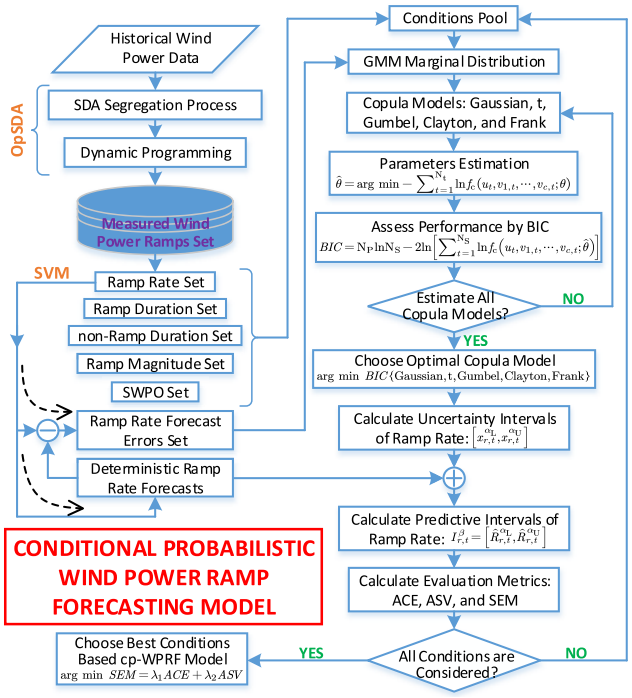


Fig. 5. The overall framework of the developed cp-WPRF model: an example of ramp rate forecasts.

- *Step 3:* Parameters of the copula models are estimated by the ML/CML method. The best copula model is chosen based on the minimum BIC (see Section III).
- *Step 4:* The PIs of ramp rate $I_{r,t}^{\beta}$ are calculated as the combination of the deterministic forecasts and the WPR uncertainties PIs (see (18)). The best conditions based cp-WPRF model is determined by the quality of PIs with evaluation metrics (see Section V).

V. CP-WPRF Evaluation Metrics

To evaluate the performance of cp-WPRF at different conditions, two PI-based metrics, namely reliability and sharpness [26]–[28], are adopted and briefly introduced in this section. Reliability indicates the correct degree of a cp-WPRF assessed by the hit percentage. Sharpness indicates the uncertainty conveyed by the cp-WPRF.

A. Reliability

The forecasted WPR features are expected to lie within the PI bounds with a prescribed probability termed as the nominal proportion. It is expected that the coverage probability of obtained PIs will asymptotically reach the nominal level of confidence (ideal case) over the full WPRs. PI coverage probability (PICP) is a critical measure for the reliability of the WPR PIs, formulated as $PICP = \frac{1}{\mathrm{N_t}} \sum_{t=1}^{\mathrm{N_t}} \phi_t^{\beta} \times 100\%$, where the indicator of PICP ($\phi_{r,t}^{\beta}$) is defined in (19). Theoretically, the PICP should be close to the corresponding PI nominal confidence (PINC). The average coverage error (ACE) [26] metric formulated in (20) should be as close to zero as possible. A smaller absolute ACE indicates more reliable PIs of WPRs.

$$\phi_{r,t}^{\beta} = \begin{cases} 1, & \hat{R}_{r,t} \in I_{r,t}^{\beta} \\ 0, & \hat{R}_{r,t} \notin I_{r,t}^{\beta} \end{cases} \tag{19}$$

$$ACE = \frac{1}{\mathrm{N_{SL}}} \sum_{j=1}^{\mathrm{N_{SL}}} \left| \frac{1}{\mathrm{N_t}} \sum_{t=1}^{\mathrm{N_t}} \phi_t^{\beta_j} - PINC^{\beta_j} \right| \times 100\% \tag{20}$$

where $\mathrm{N_{SL}}$ is the number of significance levels (SLs). $\mathrm{N_t}$ is the number of test samples.

B. Sharpness

Sharpness is related to the interval size of different SLs. The mean size of PIs (δ_r^{β}) at nominal coverage rate (1-β) is:

$$\delta_r^{\beta} = \frac{1}{\mathrm{N_t}} \sum_{t=1}^{\mathrm{N_t}} \left(x_{r,t}^{\alpha_\mathrm{U}} - x_{r,t}^{\alpha_\mathrm{L}} \right) \times 100\%. \tag{21}$$

The interval score $Sc_r^{\beta}(\hat{R}_{r,t})$ rewards narrow PIs and assesses a penalty if a target doesn't lie within estimated PIs.

$$Sc_r^{\beta}\left(\hat{R}_{r,t}\right) = \begin{cases} 2\beta\delta_r^{\beta}\left(\hat{R}_{r,t}\right) + 4\left(R_{r,t}^{\alpha_\mathrm{L}} - \hat{R}_{r,t}\right), & \text{if } \hat{R}_{r,t} < R_{r,t}^{\alpha_\mathrm{L}} \\ 2\beta\delta_r^{\beta}\left(\hat{R}_{r,t}\right), & \text{if } \hat{R}_{r,t} \in I_{r,t}^{\beta} \\ \beta\delta_r^{\beta}\left(\hat{R}_{r,t}\right) + 4\left(\hat{R}_{r,t} - R_{r,t}^{\alpha_\mathrm{U}}\right), & \text{if } \hat{R}_{r,t} > R_{r,t}^{\alpha_\mathrm{U}}. \end{cases} \tag{22}$$

The average score value (ASV) can be employed to comprehensively evaluate the overall skill of WPR PIs to assess the sharpness, given by:

$$ASV = \frac{1}{\mathrm{N_{SL}N_t}} \sum_{j=1}^{\mathrm{N_{SL}}} \sum_{t=1}^{\mathrm{N_t}} Sc_r^{\beta_j}\left(\hat{R}_{r,t}\right) \times 100\%. \tag{23}$$

Generally, smaller ACE and ASV values indicate better forecasting performance. To examine the trade-off between the reliability and sharpness metrics, the synthetic evaluation metric (SEM) is formulated by:

$$SEM = \lambda_1 \overline{ACE} + \lambda_2 \overline{ASV} \tag{24}$$

where $\overline{ACE}$ and $\overline{ASV}$ are normalized by (14). λ_1 and λ_2 are the weight coefficients (in this paper, $\lambda_1 = \lambda_2 = 0.5$). The weights could be adjusted based on the balancing authorities' (or other stakeholders') preferences between the two metrics.

C. Optimal Condition Determination

To select the optimal condition from the conditions pool, the objective function is constructed by minimizing the SEM metric for each correlated condition, given by:

$$\arg\min_{i \in \Omega} SEM_i = \lambda_{1,i} \overline{ACE}_i + \lambda_{2,i} \overline{ASV}_i \tag{25}$$

where Ω is the set of conditions pool.

Fig. 6 shows the procedure of selecting the optimal condition from the conditions pool, which is described as follows:

- *Step 1:* Prepare a conditions pool and choose the ith condition from the conditions pool.
- *Step 2:* Select the optimal copula model, calculate predictive intervals, and calculate evaluation metrics SEM_i in (25).

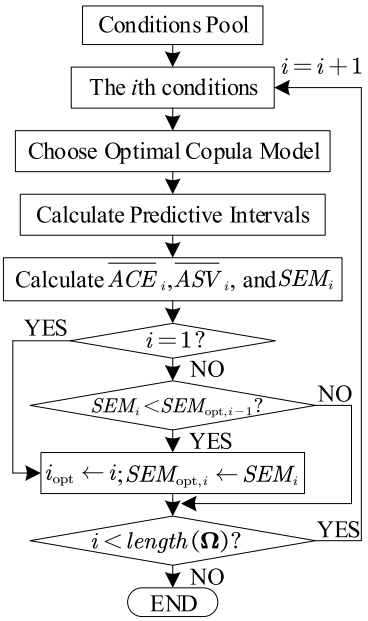


Fig. 6. Flowchart of selecting the optimal condition.

- *Step 3:* If i=1, the optimal SEM is set as $SEM_{\text{opt},1} = SEM_1$; otherwise, compare the ith SEM SEM_i with the optimal SEM $SEM_{\text{opt},i-1}$ in the last (i-1)th iteration.
- *Step 4:* If $SEM_i < SEM_{\text{opt},i-1}$, the optimal SEM is replaced by $SEM_{\text{opt},i} \leftarrow SEM_i$ and $i_{\text{opt}} \leftarrow i$; otherwise, this step is skipped and goes to *Step 5*.
- *Step 5:* Evaluate the termination condition. If the condition index i is smaller than the total number of conditions $i < length(\Omega)$, the index is updated by $i = i + 1$ and it returns to *Step 1:* otherwise, the iteration calculation is terminated. The optimal condition is selected as the i_{opt}th condition.

VI. Case Studies and Results

A. Test Case

The developed cp-WPRF model is evaluated using data from the Wind Integration National Dataset (WIND) Toolkit [15]. The data represents wind power generation from January 1st 2007 to December 31st 2012. The wind plants used in this analysis are located in the regions of Dallas, Miami, Chicago, Los Angeles, and New York with a 5-minute data resolution. The total rated wind power capacities are 10,028 MW, 9,555 MW, 9,974 MW, 10,119 MW, and 9,825 MW, respectively. There are approximately 1,585, 1,121, 1,819, 1,245, and 1,080 WPRs in each of the locations, respectively. The last 140 WPRs are used for testing. The remaining WPRs are used for training. All case studies are carried out using MATLAB 2016a on an Intel-i7-6600 2.6-GHz laptop with 16 GB of RAM memory. The door width of the OpSDA is set as 0.2% of the rated capacity.

B. Performance of the Developed cp-WPRF

The detected WPRs in the Dallas area are used to verify the effectiveness of the developed model in this case. There are a total of 1,585 WPRs detected by the OpSDA. Fig. 7 shows the cp-WPRF results for ramp rate at different nominal coverage rates (1-β). To evaluate the accuracy of the proposed model, we define two variables as *FP* and *FN*. *FP* is defined as the number of predicted ramps but not observed, which means the ramps are false alarms. *FN* is defined as the number of observed ramps but not predicted, which means the ramps are missed by the forecasting model. Two metrics are further defined based on the total number of ramps N_R: the rate of false alarm ramps (ROFA) and the rate of missed ramps (ROMR).

$$ROFA = \frac{FP}{N_R} \tag{26}$$

$$ROMR = \frac{FN}{N_R}. \tag{27}$$

TABLE I
Results of False Alarm Ramps and Missed Ramps With Different Ramp Rate Thresholds

Thresholds [MW/h]	False Alarm		Missed	
	FP	$ROFA$ [%]	FN	$ROMR$ [%]
190	10	7.14	6	4.29
200	8	5.71	5	3.57
210	7	5.00	5	3.57

As shown in Table I, ROFA with different ramp rate thresholds is in the range of 5%∼8%. ROMR with different ramp rate thresholds is in the range of 3%∼5%. The small ROFA and ROMR values verify the effectiveness of the developed cp-WPRF model.

Fig. 8 shows the performance of the developed copula-based cp-WPRF model for ramp rate. The joint conditional distribution of ramp duration and ramp rate forecast error (RRFE) is shown in Fig. 8(a), which illustrates the correlation between the RRFE and ramp duration. The cCDFs cluster shown in Fig. 8(b) is generated by the conditional distributions of the RRFE (X) for different ramp duration values ($\mathcal{D}$), abbreviated as '$X|\mathcal{D}$'. It shows that the cCDFs of RRFE present different shapes under conditions of different ramp duration values. Figs. 8(c) and 8(d) show the evaluation metrics of reliability and sharpness for the first fifteen correlated conditions that are chosen based on smaller SEM values, respectively. The $X|\mathcal{R}$ condition (where RRFE is only correlated with ramp rate) presents the closest coverage probability to the corresponding nominal proportion and relatively smaller interval scores for each nominal proportion. In Fig. 8(d), the higher interval scores are caused by conditions that are less correlated with the RRFE.

The numerical results of all correlated conditions for ramp rate are summarized in Table II. It indicates that the $X|\mathcal{R}$ condition shows the smallest SEM value, which presents the optimal trade-off between reliability and sharpness compared to other correlated conditions. It also means in this case the ramp rate probabilistic forecasts can be significantly improved by considering the stochastic dependence between the ramp rate uncertainty and measured ramp rates. Note that the aforementioned results are based on the specific data used in this

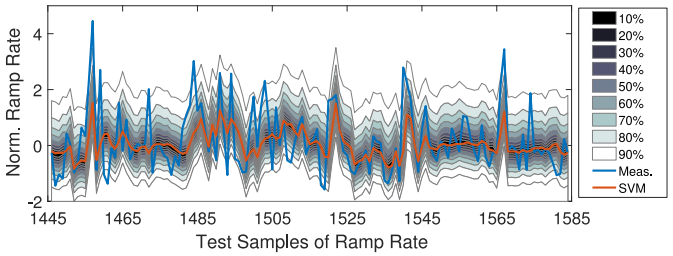


Fig. 7. cp-WPRF results for ramp rate at different nominal coverage rates.

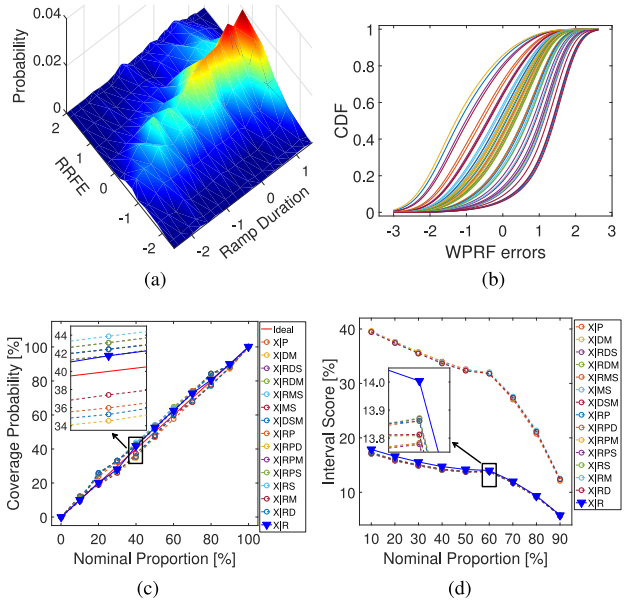


Fig. 8. Performance of the copula-based cp-WPRF model for ramp rate. (a) joint PDF of duration & rate error; (b) cCDFs cluster; (c) reliability; and (d) sharpness.

section. However, by using the developed cp-WPRF model in this paper, system operators and users are able to determine the best condition (i.e., features combination) when forecasting a ramping feature.

The numerical results in Table II also verify that there is no strong correlation between the forecasting accuracy and the number of ramping features considered. This is because some features may be helpful for improving the reliability metric (a smaller ACE value), whereas considering others may be helpful for improving the sharpness metric (a smaller ASV value). However, the ranks of features with the smallest ACE and ASV are completely different. Thus, it is still challenging to specify features which are important for improving both the reliability and sharpness metrics. As shown in Table II, the $X|\mathcal{RPS}$ condition performs the best at improving the sharpness metric with the smallest ASV value (12.89), and the $X|\mathcal{R}$ condition performs best at improving the reliability metric with the smallest ACE value (1.12).

C. Comparisons of Different Probabilistic WPRF Models

To verify the effectiveness of the developed cp-WPRF model, four probabilistic WPRF models are used for comparisons. The detailed information on the four models is described as below and also summarized in Table III.

TABLE II
CP-WPRF RESULTS OF ALL CORRELATED CONDITIONS FOR RAMP RATE

Conditions Pool	Optimal Copula Model	BIC [$\times 10^4$]	ACE [%]	ASV [%]	SEM
$X\|\mathcal{RPDMS}$	t	-2.37	2.73	12.94	0.02
$X\|\mathcal{PDMS}$	Gaussian	-2.12	2.32	30.08	0.48
$X\|\mathcal{RPDM}$	t	-1.17	2.66	12.94	-0.06
$X\|\mathcal{RPDS}$	t	-1.51	2.66	12.94	-0.06
$X\|\mathcal{RDMS}$	t	-2.31	2.73	12.96	0.02
$X\|\mathcal{RPMS}$	t	-1.12	2.73	12.93	0.02
$X\|\mathcal{RPD}$	Frank	-3.16	2.33	12.90	-0.49
$X\|\mathcal{RPM}$	t	-3.16	2.33	12.90	-0.49
$\boldsymbol{X\|\mathcal{RPS}}$	**t**	**-3.17**	**2.33**	**12.89**	**-0.49**
$X\|\mathcal{RDM}$	t	-1.08	2.60	12.97	-0.15
$X\|\mathcal{RDS}$	t	-1.44	2.60	12.97	-0.15
$X\|\mathcal{RMS}$	t	-1.05	2.60	12.97	-0.15
$X\|\mathcal{PDM}$	Gaussian	-9.21	2.58	30.23	0.82
$X\|\mathcal{PDS}$	t	-1.26	2.45	30.21	0.66
$X\|\mathcal{DSM}$	Gaussian	-2.06	1.69	30.11	-0.33
$X\|\mathcal{PMS}$	Gaussian	-8.77	2.33	30.21	0.50
$X\|\mathcal{RP}$	t	-2.45	2.33	13.00	-0.48
$X\|\mathcal{RD}$	t	-2.48	2.14	12.95	-0.74
$X\|\mathcal{RM}$	Gumbel	-2.48	2.19	12.95	-0.67
$X\|\mathcal{RS}$	t	-2.48	2.26	12.95	-0.59
$X\|\mathcal{PD}$	t	-6.99	2.84	30.31	1.17
$X\|\mathcal{PM}$	t	-6.95	2.84	30.31	1.17
$X\|\mathcal{PS}$	t	-6.97	2.78	30.30	1.08
$X\|\mathcal{DM}$	Gaussian	-8.52	1.85	30.22	-0.12
$X\|\mathcal{DS}$	t	-1.18	1.96	30.23	0.02
$X\|\mathcal{MS}$	Gaussian	-8.08	1.81	30.25	-0.17
$\boldsymbol{X\|\mathcal{R}}$	**Gumbel**	**-2.55**	**1.12**	**13.34**	**-2.03**
$X\|\mathcal{P}$	t	-5.77	1.88	30.15	-0.09
$X\|\mathcal{D}$	Clayton	-3.25	2.20	29.95	0.32
$X\|\mathcal{M}$	Clayton	-3.26	2.27	29.94	0.40
$X\|\mathcal{S}$	Clayton	-3.27	2.40	29.94	0.57

Note that symbols of '$\mathcal{R}$', '$\mathcal{D}$', '$\mathcal{M}$', '$\mathcal{S}$', and '$\mathcal{P}$' represent the conditions of ramp rate, ramp duration, ramp magnitude, ramp start-time, and SWPO, respectively. For example, '$X|\mathcal{RD}$' means the ramp rate forecast error (X) is correlated with the ramp rate ($\mathcal{R}$) and ramp duration ($\mathcal{D}$).

TABLE III
DESCRIPTIONS OF PROBABILISTIC WPRF MODELS

WPRF Models	Descriptions		
	Marginal Distribution	WPRF Error Uncertainty	Copula Dependence
Model 1	GMM	–	✘
Model 2	GMM	✘	–
Model 3	normal	✘	✘
Model 4 (*proposed*)	GMM	✘	✘

- *Model 1:* only consider the ramp features' dependence without modeling WPRF error uncertainties.
- *Model 2:* only considering WPRF error uncertainties without modeling the ramp features' dependence.
- *Model 3:* both the ramp features' dependence and WPRF error uncertainties are considered by using the normal marginal distribution.
- *Model 4 (proposed):* both the ramping features' dependence and WPRF error uncertainties are considered by using the GMM marginal distribution.

Two cases are studied to compare the performance of four probabilistic WPRF models: (i) *Case 1:* WPRF results of *different ramp features* in *the same region*; and (ii) *Case 2:* WPRF results of *the same ramp feature* in *different regions*.

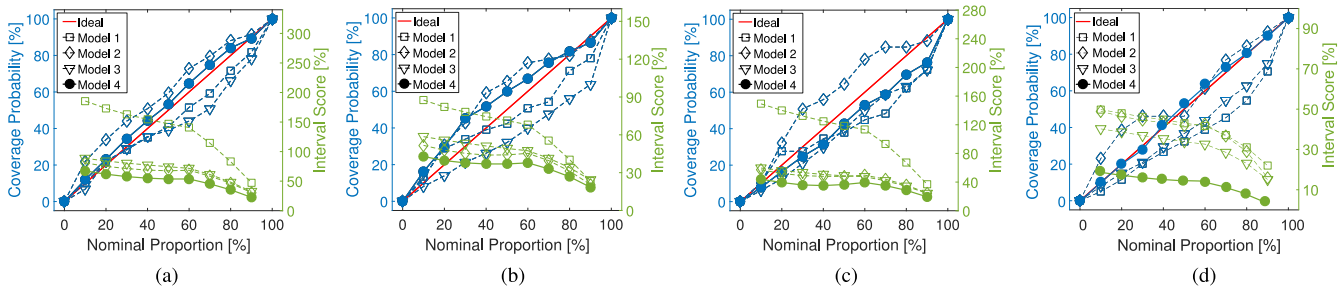


Fig. 9. Comparison of different models of conditional probabilistic forecasts for ramp features in the Dallas area. (a) magnitude; (b) duration; (c) start-time; and (d) rate.

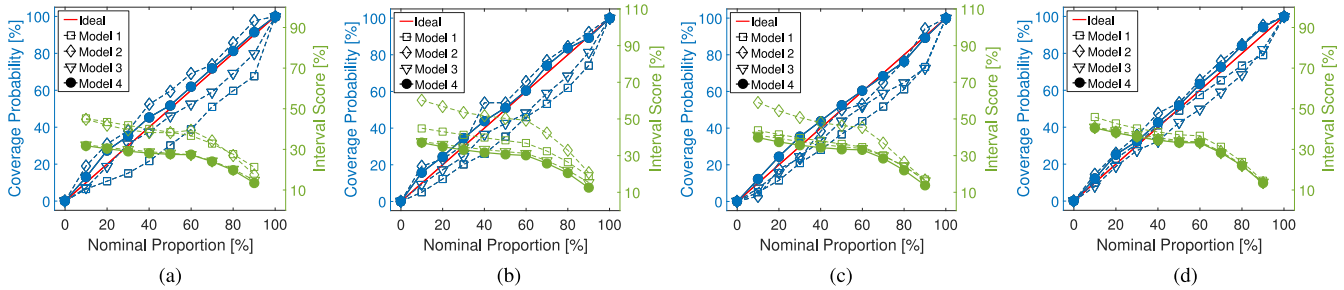


Fig. 10. Comparison of different models of cp-WPRF for ramp rate in multiple regions. (a) MIA; (b) CHI; (c) NYC; and (d) LA.

Case 1: The WPR data of Section VI-B in the Dallas area is used as an extensive study in this case, i.e., different ramp features with multiple WPRF models. For ramp rate, Model 4 uses the selected $X|\mathcal{R}$ condition as the best cp-WPRF model with the optimal Gumbel copula. For ramp magnitude, Model 4 uses the selected $X|\mathcal{RMS}$ condition as the best cp-WPRF model with the optimal t copula. For ramp duration, Model 4 uses the selected $X|\mathcal{DR}$ condition as the best cp-WPRF model with the optimal Gaussian copula. For ramp start-time, Model 4 uses the selected $X|\mathcal{DRSP}$ condition as the best cp-WPRF model with the optimal t copula.

For the simplicity of comparison, Fig. 9 shows the coverage probabilities (reliability) and interval scores (sharpness) for different ramp features. As can be seen, the coverage probability curve of Model 4 (the blue solid line) is the closest to the ideal nominal proportion line (the red solid line) in all cases. It is also shown that Model 4 has the smallest interval scores for each nominal proportion. Particularly, Model 4 performs much better in terms of the interval score for ramp rate as shown in Fig. 9(d). This is probably because the stochastic dependence between the ramp rate feature and its uncertainty is simultaneously considered in Model 4. For a better illustration, Table IV lists the numerical results of evaluation metrics for different ramp features. Specifically, the reliability metric ranges of Model 1, 2, 3, and 4 are 4%–13%, 6%–14%, 7%–13%, and 1%–7%, respectively. The sharpness metric ranges of Model 1, 2, 3, and 4 are 40%–134%, 35%–63%, 31%–69%, and 30%–51%, respectively. The final SEM ranges are 0.3–1.04, 0.09–0.28, -0.15–0.42, and -1.14–-0.85, respectively. For all four ramp features, Model 4 presents the smallest ACE, ASV, and SEM values. This is because Model 4 considers both the WPRF error uncertainty and the stochastic dependence of different ramp features, compared to both Model 1 and Model 2. In addition, the accurate characterization of marginal distributions in Model 4 can significantly improve probabilistic WPRF metrics, compared to Model 3.

TABLE IV
COMPARATIVE RESULTS FOR DIFFERENT RAMP FEATURES

Features	Metrics	Model 1	Model 2	Model 3	Model 4
Rate	ACE [%]	12.46	6.64	9.45	**1.64**
	ASV [%]	40.18	38.97	31.58	**30.05**
	SEM	1.02	*0.27*	-0.15	**-1.14**
Mag.	ACE [%]	4.87	8.40	7.53	**2.83**
	ASV [%]	134.30	63.03	68.74	**50.26**
	SEM	0.53	*0.28*	0.18	**-0.99**
Dur.	ACE [%]	6.86	13.25	12.86	**6.34**
	ASV [%]	65.76	35.98	46.13	**34.48**
	SEM	0.30	*0.12*	0.42	**-0.85**
Start-time	ACE [%]	9.37	9.35	8.85	**5.89**
	ASV [%]	110.80	45.35	46.24	**34.82**
	SEM	1.04	*0.09*	-0.04	**-1.09**

The impact of the copula-based stochastic dependence is further analyzed by comparing Model 2 and Model 4. For all WPR features, evaluation metrics are better with improved SEM values of 1.41 [=0.27-(−1.14)], 1.27 [=0.28-(-0.99)], 0.97 [=0.12-(−0.85)], and 1.18 [=0.09-(−1.09)]. These findings would help balancing authorities efficiently manage WPRs. For example, a better forecast accuracy of ramp magnitude can be used to design more reliable ramping products in the electricity market [29].

Case 2: The ramp rate data of the four regions in Miami, Chicago, New York, and Los Angeles is used to verify the robustness of the developed cp-WPRF model. Fig. 10 shows the coverage probabilities and interval scores for ramp rate forecasts in these regions. It is found that Model 4 presents the closest coverage probability curve (the blue solid line)

TABLE V
COMPARATIVE RESULTS FOR RAMP RATE IN MULTIPLE REGIONS

Regions	Metrics	Model 1	Model 2	Model 3	Model 4
MIA	ACE [%]	13.51	6.93	4.65	**2.65**
	ASV [%]	36.38	35.67	26.31	**25.71**
	SEM	1.16	0.40	-0.65	**-0.91**
CHI	ACE [%]	10.56	4.55	5.72	**2.61**
	ASV [%]	35.78	46.42	29.97	**28.45**
	SEM	0.73	0.50	-0.34	**-0.89**
NYC	ACE [%]	10.69	3.42	7.03	**2.27**
	ASV [%]	33.91	42.67	32.65	**30.93**
	SEM	0.53	0.41	-0.07	**-0.86**
LA	ACE [%]	2.78	4.33	5.66	**2.58**
	ASV [%]	34.55	31.45	31.71	**31.07**
	SEM	0.37	-0.06	0.48	**-0.79**

to the ideal nominal proportion line (the red solid line) in all cases. This model also shows the smallest interval scores for each nominal proportion. Numerical results of evaluation metrics for ramp rate in different regions are illustrated in Table V. Specifically, the reliability metric ranges of Model 1, 2, 3, and 4 are 2%−14%, 3%−7%, 4%−8%, and 2%−3%, respectively. The sharpness metric ranges of Model 1, 2, 3, and 4 are 33%−37%, 31%−47%, 26%−33%, and 25%−32%, respectively. The SEM ranges are 0.37−1.16, -0.06−0.5, -0.65−0.48, and -0.91−-0.79, respectively. For all regions, Model 4 presents the smallest ACE, ASV, and SEM values. This is because Model 4 considers both the WPRF error uncertainty and the stochastic dependence of other ramp features. By comparing Model 2 and Model 4, all four regions show significant improvements in the evaluation metrics. The improved SEM values in Miami, Chicago, New York, and Los Angeles are 1.31 [=0.40-(−0.91)], 1.39 [=0.50-(−0.89)], 1.27 [=0.41-(−0.86)], and 0.73 [=-0.06-(−0.79)], respectively.

Another interesting finding is that the interval score difference between Model 3 and Model 4 in Fig. 9 is more significant than that in all four regions in Fig. 10. This is because the GMM marginal distribution can fit the WPRF error uncertainties *significantly better* than the normal marginal distribution. However, for other four regions in Fig. 10, the GMM marginal distribution can fit the WPRF error uncertainties only *slightly better* than the normal marginal distribution. To compare the fitting performance of the normal and GMM distributions for WPRF errors, Table VI illustrates the Chi-square (χ^2) statistics [20] to measure the goodness-of-fit. As can be seen, for the Dallas data, the fitting performance of GMM is about 72% better than that of the normal distribution. However, for other four regions (MIA, CHI, NYC, and LA), the slight improvements of using GMM are approximately 21%∼27%, compared to the normal distribution.

D. Robustness Analysis of cp-WPRF Models

To verify the robustness of the developed model with different forecasting methods, three autoregressive moving average model (ARMA) methods, i.e., ARMA(1,1), ARMA(2,1), and ARMA(3,1), are used as a comparison with the SVM-based

TABLE VI
χ^2 STATISTICS COMPARISON OF GMM AND NORMAL DISTRIBUTIONS FOR RAMP RATE AT FIVE REGIONS

Distributions	MIA	CHI	NYC	LA	DAL
normal	3.11	2.98	3.28	3.31	3.82
GMM	2.37	2.15	2.62	2.54	1.07
Improvement by GMM	23%	27%	21%	23%	72%

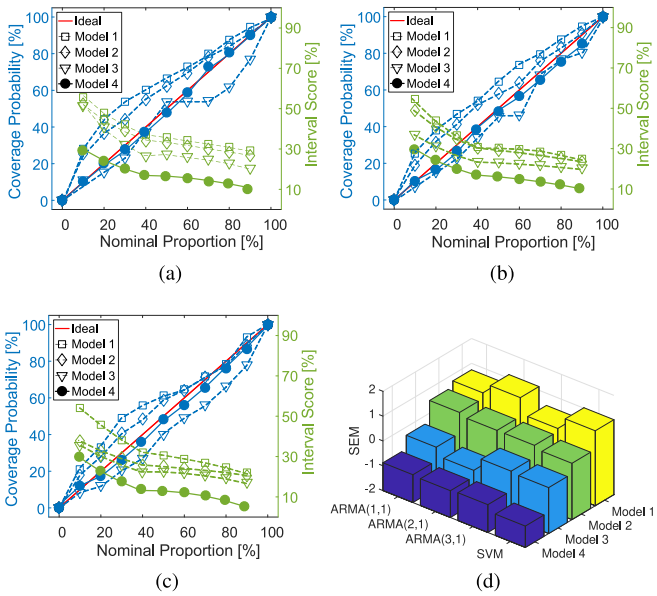


Fig. 11. Performance of different forecasting methods based cp-WPRF models. (a) ARMA(1,1)-based cp-WPRF; (b) ARMA(2,1)-based cp-WPRF; (c) ARMA(3,1)-based cp-WPRF; and (d) SEM values of multiple models.

TABLE VII
OPTIMAL COPULA MODELS AND BEST CONDITIONS OF DIFFERENT FORECASTING METHODS BASED CP-WPRF MODELS

Forecasting Method	Best Condition	Optimal Copula Model
ARMA(1,1)	$X\|\mathcal{RDM}$	t
ARMA(2,1)	$X\|\mathcal{RMS}$	Frank
ARMA(3,1)	$X\|\mathcal{MR}$	t
SVM	$X\|\mathcal{R}$	Gumbel

cp-WPRF model in Fig. 9(d). Fig. 11 shows the performance of cp-WPRF models using various ARMA forecasting methods. As can be seen in Figs. 11(a)–11(c), the developed cp-WPRF model (Model 4) consistently presents the closest coverage probability curve (the blue solid line) and the smallest interval score (the green solid line), when using different ARMA methods. Specifically, Fig. 11(d) compares the SEM values of cp-WPRF models using different forecasting methods. The developed cp-WPRF model (Model 4) consistently provides the smallest SEM values (see the dark blue bars). These observations verify that the developed cp-WPRF is robust with different forecasting methods.

Table VII illustrates optimal copula models and best conditions of different forecasting methods based on cp-WPRF models. It shows that the optimal copula model may change with the forecasting method. This is because the probability distributions of ramp rates forecasted by the distinct forecasting method are different. These probability distributions could

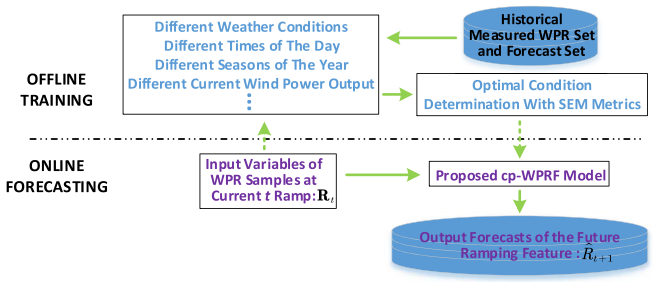


Fig. 12. Overall framework of the online ramping forecasting and the offline training for determining the optimal condition.

generate different CDF values of $u = F(x_r)$, which can impact the estimated parameters in (33) and the performance of the optimal copula model with different BIC values in (17).

VII. Discussion

For practical applications, the developed cp-WPRF model can be executed in two steps: offline training and online forecasting. Fig. 12 shows the overall framework of the online ramping forecasting and the offline training for determining the optimal condition. In the first step of offline training, operators could adopt the developed procedure with SEM metrics to determine the optimal condition based on historical measured and forecasted WPR sets. The optimal condition determination procedure is further classified based on different scenarios. These scenarios are divided by weather conditions (such as wind speed), time of day, season of year, and current wind power output. More details about scenarios can be found in our previous study in [8] and [30]. For each scenario, the optimal conditions are pre-determined and prepared for online forecasting. In the second step of online forecasting, operators put WPR samples of the current tth ramp ($\mathbf{R}_t$) into the proposed cp-WPRF model, and the optimal condition is chosen from the offline training stage based on the particular scenario to which the input variables of the WPR samples belong. Finally, the future ramp $\hat{R}_{t+1}$ can be forecasted online by using the current ramp information and the offline pre-determined optimal condition.

VIII. Conclusion

This paper developed a conditional probabilistic wind power ramp forecast (cp-WPRF) model based on Copula theory. Based on the WIND Toolkit wind power data, the optimized swinging door algorithm (OpSDA) was first used to construct the wind power ramps (WPRs) dataset. The epsilon-insensitive support vector machine (ε-SVM) was then adopted to separately generate deterministic forecasts for each WPR feature and calculate corresponding forecasting errors. To characterize the conditional distributions of WPR forecasting errors, Copula theory was used to model the stochastic dependence with different ramp features based on the marginal distribution using the Gaussian mixture model (GMM). Numerical simulations and comparisons on four probabilistic WPRF models showed that:

(i) By considering the stochastic dependence between the ramp rate uncertainty and historical ramp rates, probabilistic forecasts of ramp rate could be significantly improved with best evaluation metrics.
(ii) The developed cp-WPRF model could enhance the probabilistic forecasting accuracy of different ramping features: ramp rate, duration, magnitude, and start-time.
(iii) The developed cp-WPRF model also provided high probabilistic forecasting accuracy and robustness in different regions.

In the future, this research can be further improved by: (i) developing probabilistic wind power ramp products in the electricity market design; and (ii) considering the uncertainty of space-time dependencies of various nearby locations and look-ahead times.

Appendix A

A. Single Condition

The joint PDF with the marginal PDFs of the WPR uncertainty variable (x_r), the dependent condition variable (y_c), and the copula PDF ($f_C(\cdot)$) is given by:

$$\begin{aligned} f_{X_rY_c}(x_r, y_c) &= \frac{\partial^2 F_{X_rY_c}(x_r, y_c)}{\partial x_r \partial y_c} \\ &= \frac{\partial^2 F_C\big(F_{X_r}(x_r), F_{Y_c}(y_c)\big)}{\partial x_r \partial y_c} \frac{\partial F_{X_r}(x_r)}{\partial x_r} \frac{\partial F_{Y_c}(y_c)}{\partial y_c} \\ &= f_C\big(F_{X_r}(x_r), F_{Y_c}(y_c)\big) f_{X_r}(x_r) f_{Y_c}(y_c). \end{aligned} \tag{28}$$

B. Multiple Conditions

The joint PDF $f_{X_rY_1Y_2\cdots Y_c}(x_r, y_1, y_2, \ldots, y_c)$ is given by:

$$\begin{aligned} &f_{X_rY_1Y_2\cdots Y_c}(x_r, y_1, y_2, \ldots, y_c) \\ &\quad = \frac{\partial^2 F_C\big(F_{X_r}(x_r), F_{Y_1}(y_1), F_{Y_2}(y_2), \ldots, F_{Y_c}(y_c)\big)}{\partial x_r \partial y_1 \partial y_2 \cdots \partial y_c} \\ &\qquad \times \prod_{k=1}^{c} \frac{\partial F_{Y_k}(y_k)}{\partial y_k} \frac{\partial F_{X_r}(x_r)}{\partial x_r} \\ &\quad = f_C\big(F_{X_r}(x_r), F_{Y_1}(y_1), F_{Y_2}(y_2), \ldots, F_{Y_c}(y_c)\big) \\ &\qquad \times \prod_{k=1}^{c} f_{Y_k}(y_k) f_{X_r}(x_r). \end{aligned} \tag{29}$$

The cPDF of WPRF errors $f_{X_r|Y_1Y_2\cdots Y_c}(x_r|\hat{R}_1, \hat{R}_2, \cdots, \hat{R}_c)$ is expressed by:

$$\begin{aligned} f_{X_r|Y_1\cdots Y_c}\left(x_r|\hat{R}_1, \cdots, \hat{R}_c\right) &= \frac{f_{X_rY_1\cdots Y_c}\left(x_r, \hat{R}_1, \cdots, \hat{R}_c\right)}{f_{Y_1\cdots Y_c}\left(\hat{R}_1, \cdots, \hat{R}_c\right)} \\ &= \frac{f_C\left(F_{X_r}(x_r), F_{Y_1}\left(\hat{R}_1\right), \ldots, F_{Y_c}\left(\hat{R}_c\right)\right)}{f_C\left(F_{Y_1}\left(\hat{R}_1\right), \ldots, F_{Y_c}\left(\hat{R}_c\right)\right)} f_{X_r}(x_r). \end{aligned} \tag{30}$$

C. Iterative Process of Algorithm 1

The iterative formula of single condition based cp-WPRF is run by:

$$\begin{aligned} x_r^{l+1} &= x_r^l - \frac{F_{X_r|Y_c}\left(x_r^l|y_c=\hat{R}_c\right)-\alpha_{\mathrm{L}}}{F'_{X_r|Y_c}\left(x_r^l|y_c=\hat{R}_c\right)} \\ &= x_r^l - \frac{F_{X_r|Y_c}\left(x_r^l|y_c=\hat{R}_c\right)-\alpha_{\mathrm{L}}}{f_{X_r|Y_c}\left(x_r^l|y_c=\hat{R}_c\right)} \\ &= x_r^l - \frac{\int_{-\infty}^{x_r^l} f_{\mathrm{C}}\left(F_{X_r}(\tau), F_{Y_c}\left(\hat{R}_c\right)\right) f_{X_r}(\tau)\mathrm{d}\tau - \alpha_{\mathrm{L}}}{f_{\mathrm{C}}\left(F_{X_r}\left(x_r^l\right), F_{Y_c}\left(\hat{R}_c\right)\right) f_{X_r}\left(x_r^l\right)}. \end{aligned} \tag{31}$$

The iterative formula of multiple conditions based cp-WPRF is run by:

$$\begin{aligned} x_r^{l+1} &= x_r^l - \frac{F_{X_r|Y_1\cdots Y_c}\left(x_r^l|y_1=\hat{R}_1,\ldots,y_c=\hat{R}_c\right)-\alpha_{\mathrm{L}}}{f_{X_r|Y_1\cdots Y_c}\left(x_r^l|y_1=\hat{R}_1,\ldots,y_c=\hat{R}_c\right)} \\ &= x_r^l - \frac{\int_{-\infty}^{x_r^l} f_{\mathrm{C}}\left(F_{X_r}(\tau), F_{Y_1}\left(\hat{R}_1\right),\ldots,F_{Y_c}\left(\hat{R}_c\right)\right) f_{X_r}(\tau)\mathrm{d}\tau - \alpha_{\mathrm{L}}}{f_{\mathrm{C}}\left(F_{X_r}\left(x_r^l\right), F_{Y_1}\left(\hat{R}_1\right),\ldots,F_{Y_c}\left(\hat{R}_c\right)\right) f_{X_r}\left(x_r^l\right)}. \end{aligned} \tag{32}$$

D. Parameters Estimation of Optimal Copula Models

Copula models are divided into two families: Elliptical and Archimedean. The Elliptical copula family consists of the Gaussian and t copulas. The Archimedean copula family consists of the Gumbel, Clayton, and Frank copulas. For all five copula models in the single condition (i.e., bivariate type) and the Elliptical copula family in multiple conditions (i.e., multivariable type), the Maximum Likelihood method can be used to estimate the copula models' parameters, which has been implemented by the fitting function copulafit built into the Statistics and Machine Learning Toolbox in MATLAB [31]. However, for the multivariable Archimedean copula family, it is still challenging to solve the sole parameter θ by using the current MATLAB toolbox. Hence, we adopt the Canonical Maximum Likelihood (CML) method that is implemented based on the empirical CDF of samples. Given that $u = F(x_r)$, $v_1 = F(y_1)$, $\cdots$, and $v_c = F(y_c)$, the objective of CML is expressed by:

$$\begin{aligned} \hat{\theta} &= \arg\ \min -\sum_{t=1}^{\mathrm{N_S}} \ln f_{\mathrm{C}}\left(u_t, v_{1,t},\ldots,v_{c,t};\theta\right) \\ &= \arg\ \min -\sum_{t=1}^{\mathrm{N_S}} \ln \frac{\mathrm{d}F_{\mathrm{C}}\left(u=u_t, v_1=v_{1,t},\ldots,v_c=v_{c,t};\theta\right)}{\mathrm{d}u\mathrm{d}v_1\cdots\mathrm{d}v_c} \end{aligned} \tag{33}$$

where $f_{\mathrm{C}}(u_t, v_{1,t},\ldots,v_{c,t};\theta)$ is the multivariable copula PDF that can be calculated by the multivariable copula CDF: Clayton copula in (34), Gumbel copula in (35), and Frank copula in (36). A single-variable bounded nonlinear minimization function fminbnd embedded in MATLAB [32] is used as the optimization solver.

E. Archimedean Copula Family

The multivariable Clayton copula PDF:

$$f_{\mathrm{C}}(u, v_1,\ldots,v_c;\theta) = \left[\max\left(u^{-\theta}+v_1^{-\theta}+\cdots+v_c^{-\theta}-c;0\right)\right]^{-1/\theta}. \tag{34}$$

The multivariable Gumbel copula PDF:

$$f_{\mathrm{C}}(u, v_1,\ldots,v_c;\theta) = \exp\left\{-\left[(-\ln u)^{\theta}+(-\ln v_1)^{\theta}+\cdots+(-\ln v_c)^{\theta}\right]^{1/\theta}\right\}. \tag{35}$$

The multivariable Frank copula PDF:

$$f_{\mathrm{C}}(u, v_1,\ldots,v_c;\theta) = -\frac{1}{\theta}\ln\left\{1+\frac{\left[\exp(-\theta u)-1\right]\left[\exp(-\theta v_1)-1\right]\cdots\left[\exp(-\theta v_c)-1\right]}{\left[\exp(-\theta)-1\right]^{c}}\right\}. \tag{36}$$

Acknowledgment

The authors would also like to thank the anonymous reviewers for their constructive suggestions to this research. This work was authored in part by Alliance for Sustainable Energy, LLC, the manager and operator of the National Renewable Energy Laboratory for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Wind Energy Technology Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

References

[1] 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.

[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] M. Cui and J. Zhang, "Estimating ramping requirements with solar-friendly flexible ramping product in multi-timescale power system operations," *Appl. Energy*, vol. 225, pp. 27–41, Sep. 2018.

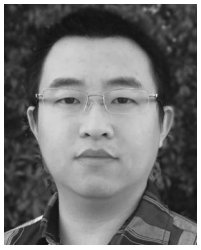
[4] Q. Xu *et al.*, "A game theoretical pricing mechanism for multi-area spinning reserve trading considering wind power uncertainty," *IEEE Trans. Power Syst.*, vol. 31, no. 2, pp. 1084–1095, Mar. 2016.

[5] Y. Liu *et al.*, "A hybrid forecasting method for wind power ramp based on orthogonal test and support vector machine (OT-SVM)," *IEEE Trans. Sustain. Energy*, vol. 8, no. 2, pp. 451–457, Apr. 2017.

[6] H. Zareipour, D. Huang, and W. Rosehart, "Wind power ramp events classification and forecasting: A data mining approach," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, San Diego, CA, USA, 2011, pp. 1–3.

[7] N. Cutler, M. Kay, K. Jacka, and T. S. Nielsen, "Detecting, categorizing and forecasting large ramps in wind farm power output using meteorological observations and WPPT," *Wind Energy*, vol. 10, no. 5, pp. 453–470, Sep./Oct. 2007.

[8] M. Cui, J. Zhang, Q. Wang, V. Krishnan, and B.-M. Hodge, "A data-driven methodology for probabilistic wind power ramp forecasting," *IEEE Trans. Smart Grid*, to be published. [Online]. Available: https://ieeexplore.ieee.org/document/8068999/, doi: 10.1109/TSG.2017.2763827.
[9] Y. Li *et al.*, "Temporal uncertainty of wind ramp predictions using probabilistic forecasting technique," in *Proc. IEEE Second Int. Conf. Big Data Comput. Service Appl. (BigDataService)*, Oxford, U.K., 2016, pp. 166–173.
[10] M. Cui *et al.*, "Wind power ramp event forecasting using a stochastic scenario generation method," *IEEE Trans. Sustain. Energy*, vol. 6, no. 2, pp. 422–433, Apr. 2015.
[11] J. W. Taylor, "Probabilistic forecasting of wind power ramp events using autoregressive logit models," *Eur. J. Oper. Res.*, vol. 259, no. 2, pp. 703–712, Jun. 2016.
[12] N. Zhang, C. Kang, Q. Xia, and J. Liang, "Modeling conditional forecast error for wind power in generation scheduling," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1316–1324, May 2014.
[13] Z. Wang, W. Wang, C. Liu, Z. Wang, and Y. Hou, "Probabilistic forecast for multiple wind farms based on regular vine copulas," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 578–589, Jan. 2018.
[14] Y. Wang *et al.*, "An efficient approach to power system uncertainty analysis with high-dimensional dependencies," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2984–2994, May 2018.
[15] 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.
[16] M. Cui *et al.*, "An optimized swinging door algorithm for identifying wind ramping events," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 150–162, Jan. 2016.
[17] A. Ul Haque, M. H. Nehrir, and P. Mandal, "A hybrid intelligent model for deterministic and quantile regression approach for probabilistic wind power forecasting," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1663–1672, Jul. 2014.
[18] 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.
[19] G. Papaefthymiou and D. Kurowicka, "Using copulas for modeling stochastic dependence in power system uncertainty analysis," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 40–49, Feb. 2009.
[20] M. Cui, C. Feng, Z. Wang, and J. Zhang, "Statistical representation of wind power ramps using a generalized Gaussian mixture model," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 261–272, Jan. 2018.
[21] 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.
[22] Y. Wang *et al.*, "Dependent discrete convolution based probabilistic load flow for the active distribution system," *IEEE Trans. Sustain. Energy*, vol. 8, no. 3, pp. 1000–1009, Jul. 2017.
[23] Y. Wang *et al.*, "Data-driven probabilistic net load forecasting with high penetration of behind-the-meter PV," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 3255–3264, May 2018.
[24] Y. He and A. Kusiak, "Performance assessment of wind turbines: Data-derived quantitative metrics," *IEEE Trans. Sustain. Energy*, vol. 9, no. 1, pp. 65–73, Jan. 2018.
[25] P. Pinson, H. A. Nielsen, J. K. Møller, H. Madsen, and G. N. Kariniotakis, "Non-parametric probabilistic forecasts of wind power: Required properties and evaluation," *Wind Energy*, vol. 10, no. 6, pp. 497–516, 2007.
[26] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1033–1044, May 2014.
[27] C. Gallego-Castillo, R. Bessa, L. Cavalcante, and O. Lopez-Garcia, "On-line quantile regression in the RKHS (reproducing Kernel Hilbert space) for operational probabilistic forecasting of wind power," *Energy*, vol. 113, pp. 355–365, Oct. 2016.
[28] C. Feng, M. Cui, B.-M. Hodge, and J. Zhang, "A data-driven multi-model methodology with deep feature selection for short-term wind forecasting," *Appl. Energy*, vol. 190, pp. 1245–1257, Mar. 2017.
[29] 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.
[30] M. Cui *et al.*, "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.
[31] Mathworks, Inc. *Copula Distributions and Correlated Samples*. Accessed: Feb. 27, 2018. [Online]. Available: https://www.mathworks.com/help/stats/copulafit.html
[32] Mathworks, Inc. *Functions/Optimization/fminbnd*. Accessed: Feb. 27, 2018. [Online]. Available: https://www.mathworks.com/help/matlab/ref/fminbnd.html



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

He is currently a Post-Doctoral Research Associate with Southern Methodist University, Dallas, TX, USA. He was a Post-Doctoral Research Associate with the University of Texas at Dallas, Richardson, TX, USA, in 2016 and 2017. He was also a Visiting Scholar from 2014 to 2015 with the Transmission and Grid Integration Group, National Renewable Energy Laboratory, Golden, CO, USA. He has published over 50 journal and conference papers. His research interests include power system operation, wind and solar forecasts, machine learning, data analytics, and statistics.

Venkat Krishnan (M'15) received the M.S. and Ph.D. degrees from Iowa State University, Ames, IA, USA, in 2007 and 2010, respectively. His research interests include electricity system capacity expansion planning, power transmission and distribution systems stability and security assessments, markets, and energy storage integration.

He is currently a Senior Engineer with the Power System Design and Studies Group, Power System Engineering Center, National Renewable Energy Laboratory, Golden, CO, USA.

Bri-Mathias Hodge (M'10–SM'17) received the B.S. degree in chemical engineering from Carnegie Mellon University, Pittsburgh, PA, USA, in 2004, the M.S. degree from the Process Design and Systems Engineering Laboratory of Åbo Akademi, Turku, Finland, in 2005, and the Ph.D. degree in chemical engineering from Purdue University, West Lafayette, IN, USA, in 2010.

He is currently the Manager of the Power System Design and Studies Group, National Renewable Energy Laboratory, Golden, CO, USA. His current research interests include energy systems modeling, simulation, optimization, and wind power forecasting.

Jie Zhang (M'13–SM'15) received the B.S. and M.S. degrees in mechanical engineering from the 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 with the Department of Mechanical Engineering, University of Texas at Dallas, Richardson, TX, USA. His research interests include multidisciplinary design optimization, complex engineered systems, big data analytics, wind and solar forecasting, renewable integration, and energy systems modeling and simulation.