IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received January 27, 2019, accepted February 11, 2019, date of publication February 14, 2019, date of current version March 5, 2019.

*Digital Object Identifier 10.1109/ACCESS.2019.2899404*

# Probabilistic Estimation of Wind Power Ramp Events: A Data-Driven Optimization Approach

**YANG CAO[1], (Student Member, IEEE), WEI WEI[1], (Senior Member, IEEE), CHENG WANG[2], (Member, IEEE), SHENGWEI MEI[1], (Fellow, IEEE), SHAOWEI HUANG[1], (Member, IEEE), AND XUEMIN ZHANG[1], (Member, IEEE)**

[1]State Key Laboratory of Power Systems, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China
[2]State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources, North China Electric Power University, Beijing 102206, China

Corresponding authors: Shengwei Mei (meishengwei@tsinghua.edu.cn) and Shaowei Huang (huangsw@tsinghua.edu.cn)

This work was supported by the National Key R&D Program of China (technology and application of wind power/photovoltaic power prediction for promoting renewable energy consumption) under Grant 2018YFB0904200.

**ABSTRACT** Large-scale integration of renewable energy, such as wind and solar power generations, imposes an unprecedented challenge on power system operation, because wind/solar power output is volatile, while in the power system, the generation must balance load in real time. Failure of maintaining power balance may trigger contingency or even blackout, especially when renewable generation quickly increases or decreases during a certain period, which is called a ramp event. Since wind power cannot be predicted accurately, it is difficult to determine the incremental change in two consecutive time periods, not to mention the probability of a ramp event. This paper addresses this problem from the perspective of uncertainty quantification. The likelihood of a ramp event is cast as a data-driven robust probability inequality, which provides the probability of a random variable with unknown distribution belonging to a given polyhedron. To tackle the distributional uncertainty of wind output, we consider a collection of candidate distributions in an ambiguity set constructed from available data. The minimal requirements include the forecast value and the mean-absolute deviation, and the moment-based ambiguity set is comprised of all probability distributions that share the same values of mean and mean-absolute deviation. With more available historical data, a meaningful divergence-based ambiguity set can be set up which encapsulates all probability distributions that are close to an empirical distribution in the sense of Wasserstein metric. The proposed approach offers the probability upper bound of a ramp event in the worst-case wind power distribution, and the conservatism can be remarkably reduced when more historical data are at hand. The proposed methods are compared with the Gaussian mixture model, validating their effectiveness and advantage.

**INDEX TERMS** Data-driven optimization, forecast, ramp event, uncertainty, wind power.

## I. INTRODUCTION

Under the transition to a green and sustainable society, wind power has witnessed fast growth worldwide in the past decades. However, wind power is inherently stochastic because of the weather condition. Since generation must balance load instantly, a large change of power output in a short period, which is called wind power ramp event [1], [2], would bring great challenge to power system operation. Wind power ramp events could be divided into two categories based on the direction: upward ramp and downward ramp. Upward ramp events are usually caused by strong low pressure air systems, wind gusts, low-level jets and thunderstorms. On the contrary, downward ramp events occur due to the reversal of the aforementioned phenomenons [3]. In an upward ramp event, the system operator has to reduce the power output from conventional generators or curtail some wind generation; In a downward ramp event, the operator needs the support from the reserve capacity, or shed some load if the reserve is not adequate [4], [5]. All of these ramp events will increase operation costs and risks. As a result, better detection and forecasting of wind power ramp events could warn the operators and help them deploy proactive strategies to prevent risky failures.

The associate editor coordinating the review of this manuscript and approving it for publication was Canbing Li.

As an important technique, detecting wind power ramp event has been studied by many researchers. Reference [6] proposes a two-stage method to detect and categorize large wind ramps based on the wind farm data from the Wind Power Prediction Tool (WPPT) in Australia. A classifier model is developed in [7] based on support vector machine (SVM) to address the one-step and multi-step ahead classification of wind power ramp events. In [8], an optimal detection technique is proposed to identify wind ramps for large time series. A family of scoring functions with ramp definitions is given and a dynamic programming recursion is applied to detect ramp events. In [9], a swinging door algorithm (SDA) is applied to identify ramp events from historical data. This method requires only one parameter in its definition, and has significant advantages in computational efficiency and robustness against noisy data. Reference [10] adopts the SDA to extract ramp events from actual and forecasted wind power time series considering ramp magnitude, direction and duration. The relevant research is extended in [11], in which an optimized swinging door algorithm (OpSDA) is proposed for ramp detection. Compared to the SDA, the OpSDA could optimally determine the tunable SDA parameter and segregate the wind power output, so significantly improves the performance in detecting wind power ramp events.

Based on the ramp detection technique, the ramp forecasting technique could provide the warning information ahead of several minutes or hours and help the operators to schedule the system operation in advance. In general, the forecasting of wind power ramp events could be divided into two categories: deterministic forecasting and probabilistic forecasting. For the former one, [12] develops a multivariate time series model to forecast the wind power ramp rates based on data-mining algorithm. In [13], a temporal ramp forecasting model is presented using multiple numerical weather prediction (NWP) inputs, statistical processing and adaptive algorithms. A hybrid forecasting model called Orthogonal Test-Support Vector Machine (OT-SVM) is proposed in [14] to improve the accuracy in ramp forecasting. Based on the joint advantages of the orthogonal test and support vector machine (SVM), this technique considers the irregular characteristics of ramp events, and has high accuracy covering different input numbers and time resolutions. In [15], the physical meaning of wind power ramp events is clarified and a combined forecasting models based on the atomic sparse decomposition (ASD) and back propagation (BP) neural networks is proposed. Three forecasting models including atomic components self-prediction model, error component prediction model and combination prediction model are developed and compared through the real wind data in China. Based on the Wind Forecast Improvement Project (WFIP), [16] investigates the ramp forecasting accuracy from the aspects of ramp magnitude, direction and duration. The improved forecasting performance benefits the system operation in terms of economy and reliability.

The aforementioned studies rely on a deterministic paradigm while the uncertainties of wind ramp events are neglected. Probabilistic approach provides statistical information of wind ramps and could help system operators make better dispatch decisions to cope with these risky events. In [17], a probabilistic wind ramp forecasting model is developed using large temporal scales information. An autoregressive logit model is proposed in [18] to simultaneously estimate the ramp event probabilities for different thresholds based on a multinomial logit structure and categorical distribution. Taking the ramp slope and phase forecasting errors into account, a short-term ramp forecasting model is proposed in [19], where a probability density function is estimated for the wind power and then is applied to provide probabilistic information of encountering wind ramp events. In [20], a neural network (NN) is proposed to model the stochastic process for wind power scenarios generation, and the distribution properties of ramp events are analyzed based on these possible scenarios. An analytical generalized Gaussian mixture model (GGMM) is developed in [21] to characterize the probability distributions of different ramp features, and then the cumulative distribution function of GGMM is applied to design a random number generator for ramp features. In [22], a logistic regression technique is applied to conduct a temporal ramp forecasting model. The temporal uncertainty of wind ramp events is addressed based on wind power scenarios generated from quantile forecasts of wind power, so as to provide additional probabilistic information for decision makers. In [23], a generalized Gaussian mixture model is developed to describe the wind power forecasting errors generated from a machine learning technique, and generate quantity of forecasting errors scenarios. The OpSDA is applied to conduct the probabilistic wind power ramp forecasting based on the generated scenarios under different weather and time conditions. By considering the stochastic correlation of different wind ramp features (magnitude, rate, start-time and duration), [24] investigates a conditional probabilistic wind power ramp forecasting model based on Copula theory. The Gaussian mixture model (GMM) and Bayesian information criterion are applied to fit and choose the optimal copula model for improved accuracy in prediction.

Above probabilistic approaches conduct statistical analysis based on an empirical distribution of wind power. However, on the one hand, calibrating an exact probability distribution requires enough historical data, which may not be available at hand; on the other hand, any parametric distribution may not completely fit real data, and the statistical analysis may be sensitive to the change in the distribution of uncertain data [25]. A possible way to overcome the above difficulties is to adopt distributionally robust optimization method [26], [27]. This method considers a family of distributions around the empirical distribution; the result provides a conservative estimation on the probability of target event, and is robust against perturbations in probability distribution.

In a distributionally robust optimization model, uncertainty is described by a probability density function (PDF); however, the PDF is inexact, and a family of candidate PDFs are taken into account, constituting a so-called ambiguity set. According to available information, the ambiguity set can be constructed by two methods. One uses moment information, such as those in [26] and [28], and the other incorporates an empirical distribution and divergence measure, such as those in [29] and [30]. This paper studies probabilistic forecasting of wind power ramp events from the uncertainty quantification perspective [31], [32]. Robust probability inequalities with two ambiguity sets are proposed to estimate the probability of ramp events which makes no reference to an exact wind power distribution. The first one only requires a point forecast and the mean-absolute deviation of forecast error and takes the unimodality of forecast error distribution into account; the second one needs an arbitrary number of historical realizations of wind power data. Both models are data-driven and give rise to computationally tractable convex optimization problems, which could be solved via off-the-shelf solvers.

The rest of this paper is organized as follows. The uncertainty quantification model for forecasting of wind power ramp events is explained in Section II, following which the two data-driven approaches are presented, including the moment-based one and its equivalent second-order cone program, as well as the divergence-based one and its equivalent linear or second-order cone program, leveraging the computational superiority of convex optimization. Case studies are conducted in Section III, in which the proposed methods are compared with Gaussian mixture model, demonstrating their effectiveness and advantages. Finally, conclusion is drawn in Section IV.

## II. PROPOSED MODELS

### A. UNCERTAINTY QUANTIFICATION MODEL

Wind power ramp takes place in a large-scale wind farm connected to a power transmission network. In the existing literature, a ramp event is defined in several ways. In [1] and [12], it refers to the large change of wind farm output over several periods (say, 20% capacity in 30 minutes or 1 hour) or the rapid change ratio of wind farm output in a short period. The difference between maximum and minimum wind power during a number of periods is also considered as a ramp event in [1].

In this paper, we study the ultra-short term forecast of single-period ramp defined in [1]: the change of wind power in two successive dispatch periods (usually one hour) exceeds a threshold. Assuming that the point wind power forecasts in future two successive periods are available, e.g., using the method in [33], it is easy to identify the incremental change of wind power output and thus detect a ramp event in a deterministic way. However, point forecast suffers from inaccuracy and could provide either conservative or optimistic results to the system operator. As a result, probabilistic ramp forecasting with confidence level information is highly desired, such as the conditional interval prediction method in [34]. Nevertheless, because a ramp event involves the output in two different periods, both of which are uncertain, it is difficult to extract the wind power movement from the conditional interval prediction method.

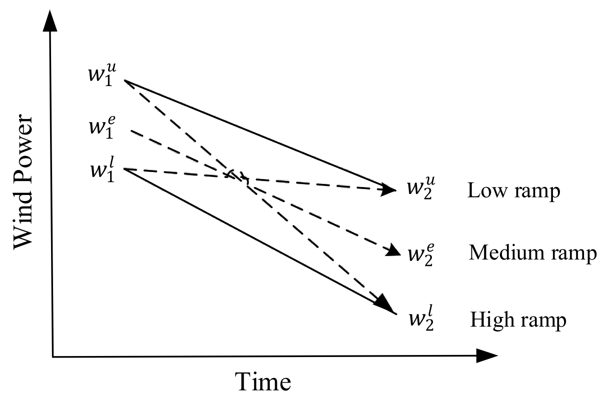



**FIGURE 1. Illustration of wind power ramp event.**

Fig. 1 illustrates a potential ramp event. The point forecasts in periods 1 and 2 are $w_1^e$ and $w_2^e$, respectively. Deterministic method compares the difference $w_1^e - w_2^e$ with a pre-specified threshold to judge a ramp event. However, as both values are uncertain and belong two certain intervals, the real incremental change could be either larger or smaller than $w_1^e - w_2^e$, jeopardizing the credibility of result. In the worst case, the movement $w_1^u - w_2^l$ constitutes a more severe ramp event; in the optimistic case, the movement $w_1^l - w_2^u$ may not be classified as a ramp event. In this regard, the uncertainty of $w_1$ and $w_2$ will be considered in this paper. We do not make assumptions on the exact probability distribution of forecast error; instead, we will make full use of distributional information recovered from limited data, and render a conservative estimation on the probability of a ramp event.

Let $\boldsymbol{w} = [w_1, w_2]^T$ be a vector of random variable, and $f(\boldsymbol{w})$ its probability density function (PDF). The probability of an upward ramp event can be expressed by

$$\Pr\{\boldsymbol{w} \in W\} \tag{1}$$

where $\Pr[\cdot]$ stands for the probability of an event, and

$$W = \left\{ \boldsymbol{w} \left| \begin{array}{c} 0 \le w_1 \le C \\ 0 \le w_2 \le C \\ w_2 - w_1 \ge R_U \end{array} \right. \right\} \tag{2}$$

where $C$ represents the capacity of the wind farm; $R_U$ denotes the upward ramp threshold. The probability of a downward ramp event can be set up in a similar way by replacing the last constraint with $w_1 - w_2 \ge R_D$, where $R_D$ denotes the downward ramp threshold.

Evaluating the probability in (1) requires a PDF of $\boldsymbol{w}$. For example, we may assume $\boldsymbol{w}$ obeys Gaussian distribution, and estimate the mean and variance via curve fitting methods [35]. However, such a parametric distribution may not be able to reflect the true distribution of wind power.

To circumvent the difficulty in PDF acquisition, we propose to consider a family of uncertain PDFs restricted in an ambiguity set $\mathbb{S}$, and aim to identify the most pessimistic outcome, giving rise to the following distributionally robust uncertainty quantification problem

$$\sup_{f(w)\in\mathbb{S}} \Pr\{\boldsymbol{w} \in W\} \tag{3}$$

In probability theory, such a problem is also known as a probability inequality [36]. Some famous examples in which first- and second-order moments are known include the univariate Chebyshev's inequality, Gaussian ineqaulity and their generalizations for multivariate cases [37], [38]. Problem (3) is actually a functional optimization problem. $\Pr[\cdot]$ defines a mapping from PDF $f(\boldsymbol{w})$ to a real number, and the decision variable is $f(\boldsymbol{w})$, which is a real-valued function. In other words, for every $\boldsymbol{w} \in W$, the value $f(\boldsymbol{w})$ is a decision variable. In this regard, the number of decision variables is infinite, and the feasible region is $\mathbb{S}$. If $\boldsymbol{w}$ follows any distribution contained in $\mathbb{S}$, the actual probability of ramp event should be no greater than the optimal value of (3). This prudent estimation could preserve sufficient security margin in the presence of PDF perturbations, and thus is acceptable in power system operation.

### B. MOMENT-BASED METHOD

#### UNIMODALITY

A probability distribution $f(\boldsymbol{w}) \in \mathbb{S}$ for $\boldsymbol{w} \in \mathbb{R}^k$ is unimodal with center $\boldsymbol{m}$ if $t^k f(B - \boldsymbol{m}/t)$ is non-decreasing in $t > 0$ for all Borel set $B \in \mathbb{B}(\mathbb{R}^k)$ [39].

The unimodality means that the PDF has a single maximum. Under above assumption, the moment-based ambiguity set is constructed as

$$\mathbb{S}_1 = \left\{ f(w) \,\middle|\, \begin{array}{l} \mathbb{E}_f(\boldsymbol{w}) = \boldsymbol{m}, \mathbb{E}_f(|\boldsymbol{w} - \boldsymbol{m}|) = \boldsymbol{\sigma} \\ f \text{ is unimodal with center } \boldsymbol{m} \end{array} \right\} \tag{4}$$

The moment-based ambiguity set $\mathbb{S}_1$ considers all PDFs that share the same mean value $\boldsymbol{m}$ and mean-absolute deviation $\boldsymbol{\sigma}$. In addition, we require the PDF $f(w)$ is unimodal.

Substituting $\mathbb{S}_1$ into problem (3), the resulting uncertainty quantification problem can be reformulated as the following form [32].

$$\inf\ 1 - \lambda + \boldsymbol{\sigma}^T \boldsymbol{\eta} \tag{5a}$$
$$\text{s.t. } \lambda \in \mathbb{R}, \quad \boldsymbol{\theta} \in \mathbb{R}^2, \ \boldsymbol{\eta} \in \mathbb{R}^2_+, \ \pi \in \mathbb{R}_+ \tag{5b}$$
$$-\boldsymbol{\eta} \le \boldsymbol{\theta} \le \boldsymbol{\eta}, \quad \lambda \le 1 \tag{5c}$$
$$-\frac{2}{3}\boldsymbol{\theta} \le \frac{2}{3}\boldsymbol{\theta} + \pi \boldsymbol{s} \le \frac{2}{3}\boldsymbol{\theta} \tag{5d}$$
$$\lambda \le \frac{27}{4}[\pi(R_U - \boldsymbol{s}^T \boldsymbol{m})]^{\frac{2}{3}} \tag{5e}$$

where $\boldsymbol{s} = [-1, 1]^T$; $\lambda$, $\boldsymbol{\theta}$, $\boldsymbol{\eta}$ and $\pi$ are all auxiliary variables.

In problem (5), the objective function (5a) and constraints (5b)-(5d) are all linear, while constraint (5e) brings computation difficulty due to the power item.

However, we find out that constraint (5e) is actually convex. Following some mathematical tricks in [40] (Section 2.3.1), (5e) can be reformulated as linear inequalities and second-order cones by introducing several auxiliary variables $y, t_1, t_2$.

$$\begin{aligned} y &= 4\lambda/27 \\ t_1 &\le \pi(R_U - \boldsymbol{s}^T \boldsymbol{m}) \\ \left\| t_2 \frac{y-1}{2} \right\|_2 &\le \frac{y+1}{2} \\ \left\| y \frac{t_1 - t_2}{2} \right\|_2 &\le \frac{t_1 + t_2}{2} \end{aligned} \tag{6}$$

To see their equivalence, expand two second-order cones, we get $t_2^2 \le y$, $y^2 \le t_1 t_2$; eliminating $t_2$ gives $y^3 \le t_1^2$ and thus

$$\lambda \le \frac{27}{4}\left[\pi(R_U - \boldsymbol{s}^T \boldsymbol{m})\right]^{\frac{2}{3}}$$

which is (5e).

Finally, probability inequality (3) with the moment-based ambiguity set $\mathbb{S}_1$ could be transformed to a second-order cone program with objective function (5a) and constraints (5b)-(5d) and (6), which could be solved efficiently.

### C. DIVERGENCE-BASED METHOD

The moment-based method has two input parameters: mean value and mean-absolute deviation, which can be recovered from very limited data or a rough guess. With the increasing of data availability, although the moment-based mode remains applicable, it may no longer be a good choice, because the distributional information disclosed by historical data is not fully utilized. To take full advantage of data, we consider an empirical distribution

$$f_0 = \frac{1}{N}\sum_{i=1}^{N} \delta_{\hat{\boldsymbol{w}}_i^0}$$

consisting of $N$ independent samples $\hat{\boldsymbol{w}}_i^0$, and each of them has a probability of $1/N$. $\delta_{\hat{\boldsymbol{w}}_i^0}$ denotes Dirac distribution concentrating unit mass at $\hat{\boldsymbol{w}}_i^0$.

However, $f_0$ is generally inaccurate, so we resort to an ambiguity set containing all PDFs that are close enough to the empirical distribution $f_0$. To quantify the distance between two PDFs, the Wasserstein metric is adopted in this paper, which is defined as

$$\begin{aligned} D_W(f, f_0) = \inf & \int_{\Xi^2} \left\| \boldsymbol{w} - \boldsymbol{w}^0 \right\|_p \Pi(d\boldsymbol{w}, d\boldsymbol{w}^0) \\ \text{s.t. } & \Pi \text{ is a joint distribution of } \boldsymbol{w} \\ & \text{and } \boldsymbol{w}^0 \text{ with marginals } \mathbb{Q} \text{ and } \mathbb{Q}_0 \end{aligned}$$

where $\|\cdot\|_p$ represents the $p$ norm on $\mathbb{R}^2$.

For two discrete distributions, the Wasserstein metric is given by

$$\begin{aligned} D_W(f, f_0) = \inf_{\pi \ge 0} & \sum_i \sum_j \pi_{ij} \left\| \boldsymbol{w}_j - \boldsymbol{w}_i^0 \right\|_p \\ \text{s.t. } & \sum\nolimits_j \pi_{ij} = p_i^0, \quad \forall i \\ & \sum\nolimits_i \pi_{ij} = p_j, \quad \forall j \end{aligned}$$

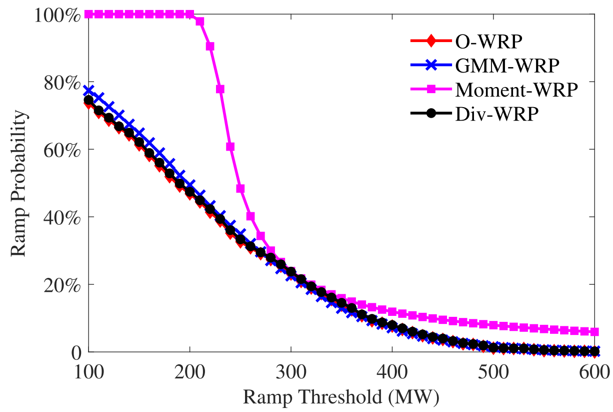



**FIGURE 2.** Ramp probabilities calculated by different models.

where $p_i^0$ and $p_j$ denote the probability of representative scenario $\boldsymbol{w}_i^0$ and $\boldsymbol{w}_j$.

Equipped with this Wasserstein distance measure, the ambiguity set can be constructed as

$$\mathbb{S}_2 = \{f(\boldsymbol{w}) | D_W(f, f_0) \leq d_w\} \quad (7)$$

where $d_w$ is a constant determining the size of the ambiguity set. When $d_w > 0$, there are infinitely many PDFs in ambiguity set (7); when $d_w$ approaches to 0, the ambiguity set (7) converges to the empirical distribution $f_0$. According to [41], if we choose

$$d_w = -\log(\alpha^*)/N \quad (8)$$

the confidence level of $f \in \mathbb{S}_2$ is at least $1 - \alpha^*$. Clearly, $d_w$ decreases when $N$ grows larger, which means that the more data we have, the more confident we are in the accuracy of empirical distribution $f_0$.

Kullback-Leibler (KL) divergence is another renowned distance metric for two probability distributions. For either metric, the distance threshold in the ambiguity set is a critical parameter in practical usage. For Wassertein metric, $d_w$ can be selected via equation (8), while for KL-divergence, the selection of such a parameter is a little bit tricky. Furthermore, the Wassertein metric based ambiguity set provides satisfactory out-of-sample performance guarantee [30]. So we choose Wasserstein metric in this paper.

Substituting $\mathbb{S}_2$ into problem (3), the resulting probability inequality can be reformulated as the following form [30].

$$\inf\ 1 - \frac{1}{N}\sum_{n=1}^{N} \beta_n + \gamma d_w \quad (9a)$$

$$\text{s.t. } \boldsymbol{\beta} \in \mathbb{R}^N, \quad \gamma \in \mathbb{R}_+, \ \boldsymbol{\tau} \in \mathbb{R}_+^N \quad (9b)$$

$$\beta_n \leq 1, \quad \forall n = 1, \ \ldots, N \quad (9c)$$

$$\beta_n + \tau_n \boldsymbol{s}\boldsymbol{w}_n \leq \tau_n R_U, \forall n = 1, \ldots, N \quad (9d)$$

$$\|\tau_n \boldsymbol{s}\|_q \leq \gamma, \forall n = 1, \ldots, N \quad (9e)$$

where $\boldsymbol{s} = [-1, 1]^T$; $\boldsymbol{\beta}, \gamma$ and $\boldsymbol{\tau}$ are all auxiliary variables; $\|\cdot\|_q$ is the dual norm of $\|\cdot\|_p$, where $p$ and $q$ satisfy $p^{-1} + q^{-1} = 1$.

**TABLE 1.** Ramp probabilities estimated by different models.

| $R_U$ (MW) | O-WRP | GMM-WRP | Moment-WRP | Div-WRP |
|---|---|---|---|---|
| 200 | 46.81% | 49.42% | 100% | 47.41% |
| 250 | 32.59% | 34.93% | 48.37% | 33.33% |
| 300 | 23.08% | 22.53% | 23.94% | 23.78% |
| 400 | 7.66% | 7.15% | 11.91% | 7.97% |
| 500 | 1.11% | 1.53% | 7.93% | 1.26% |
| 600 | 0 | 0.21% | 5.94% | 0.19% |
| Time (s) | / | 29.68 | 0.32 | 2.13 |

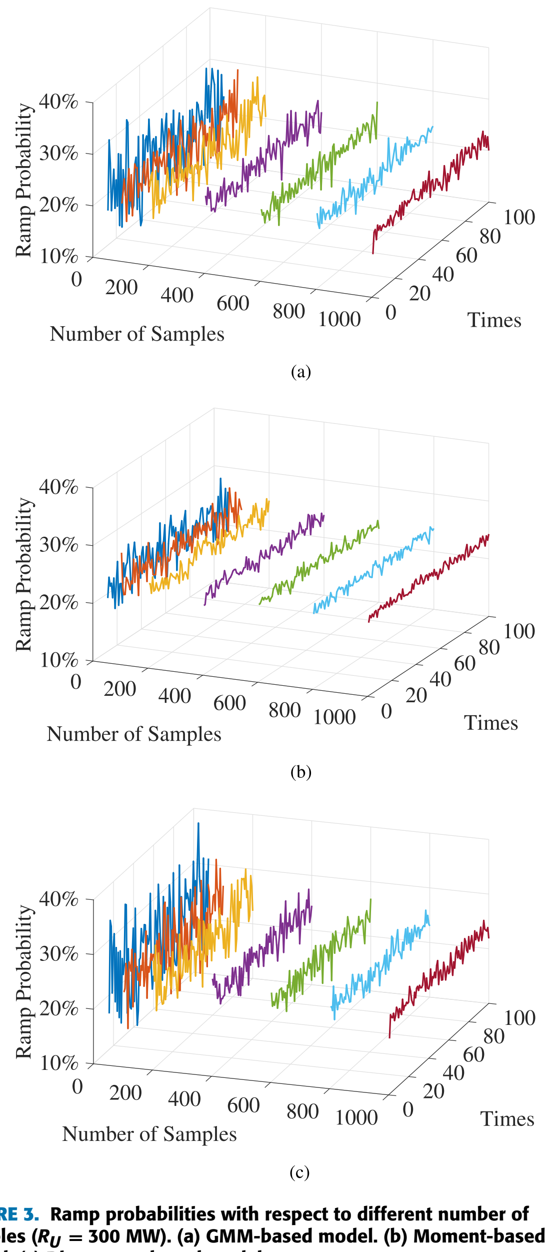



**FIGURE 3.** Ramp probabilities with respect to different number of samples ($R_U$ = 300 MW). (a) GMM-based model. (b) Moment-based model. (c) Divergence-based model.

In problem (9), the objective function (9a) and constraints (9b)-(9d) are all linear, except for the norm constraint (9e).

However, we find out that constraint (9e) gives rise to a linear one when $p = 1$ or $p = \infty$, and a second-order cone one when $p = 2$. In this paper, we choose $p = \infty$, $q = 1$ and

**TABLE 2. Statistical information of ramp probabilities with respect to different values of $N_S$ and $R_U$.**

| | | Mean value | | | Standard deviation | | | Maximum | | | Minimum | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $R_U$ (MW) | $N_S$ | GMM-WRP | Moment-WRP | Div-WRP | GMM-WRP | Moment-WRP | Div-WRP | GMM-WRP | Moment-WRP | Div-WRP | GMM-WRP | Moment-WRP | Div-WRP |
| 200 | 50 | 47.25% | 100% | 49.54% | 6.18% | 0% | 7.09% | 63.67% | 100% | 65.09% | 33.32% | 100% | 31.50% |
| | 100 | 47.36% | 100% | 48.37% | 4.60% | 0% | 4.80% | 58.95% | 100% | 59.75% | 37.34% | 100% | 37.08% |
| | 200 | 47.69% | 100% | 48.88% | 3.65% | 0% | 3.86% | 55.95% | 100% | 56.57% | 36.86% | 100% | 38.81% |
| | 400 | 48.08% | 100% | 48.64% | 2.41% | 0% | 2.40% | 53.69% | 100% | 55.17% | 42.54% | 100% | 40.75% |
| | 600 | 48.33% | 100% | 48.33% | 2.04% | 0% | 1.88% | 53.95% | 100% | 52.83% | 43.80% | 100% | 43.21% |
| | 800 | 48.33% | 100% | 48.22% | 1.89% | 0% | 1.75% | 53.84% | 100% | 52.96% | 43.85% | 100% | 43.94% |
| | 1000 | 48.23% | 100% | 48.00% | 1.55% | 0% | 1.52% | 51.68% | 100% | 51.50% | 45.06% | 100% | 44.60% |
| 250 | 50 | 32.98% | 47.77% | 36.21% | 5.48% | 5.15% | 6.03% | 46.06% | 61.56% | 52.50% | 21.70% | 37.46% | 19.33% |
| | 100 | 33.85% | 48.21% | 36.28% | 4.18% | 3.98% | 4.60% | 47.83% | 61.22% | 48.50% | 21.59% | 37.56% | 23.86% |
| | 200 | 34.64% | 47.55% | 36.22% | 3.07% | 2.67% | 3.61% | 45.21% | 53.85% | 48.67% | 27.49% | 39.96% | 27.87% |
| | 400 | 34.55% | 47.80% | 35.78% | 2.17% | 1.98% | 2.28% | 39.79% | 53.21% | 41.58% | 29.03% | 42.89% | 29.75% |
| | 600 | 34.78% | 47.78% | 35.87% | 1.81% | 1.66% | 1.90% | 39.34% | 52.45% | 41.33% | 30.34% | 44.01% | 32.17% |
| | 800 | 34.91% | 48.06% | 35.86% | 1.64% | 1.41% | 1.64% | 39.04% | 51.59% | 40.29% | 30.17% | 44.80% | 32.37% |
| | 1000 | 35.15% | 47.89% | 35.67% | 1.48% | 1.16% | 1.38% | 39.53% | 51.23% | 38.35% | 31.67% | 43.99% | 32.62% |
| 300 | 50 | 22.09% | 23.31% | 25.14% | 5.47% | 2.60% | 6.23% | 33.93% | 28.93% | 39.25% | 11.20% | 17.98% | 12.48% |
| | 100 | 22.04% | 23.69% | 24.19% | 3.36% | 1.88% | 4.23% | 28.81% | 29.25% | 33.08% | 13.36% | 18.84% | 14.26% |
| | 200 | 22.22% | 23.56% | 24.49% | 3.00% | 1.27% | 3.43% | 29.37% | 26.37% | 32.33% | 15.56% | 20.91% | 16.83% |
| | 400 | 22.65% | 23.66% | 24.38% | 2.00% | 0.92% | 2.06% | 27.90% | 25.77% | 29.67% | 15.95% | 21.51% | 19.50% |
| | 600 | 22.58% | 23.67% | 24.16% | 1.64% | 0.76% | 1.79% | 26.40% | 25.74% | 27.50% | 17.07% | 21.59% | 18.43% |
| | 800 | 22.62% | 23.72% | 24.10% | 1.47% | 0.68% | 1.52% | 27.91% | 25.05% | 28.09% | 18.33% | 21.89% | 20.47% |
| | 1000 | 22.23% | 23.62% | 23.81% | 1.34% | 0.58% | 1.25% | 25.41% | 24.94% | 26.43% | 18.18% | 22.27% | 20.17% |
| 400 | 50 | 6.43% | 11.89% | 8.47% | 3.12% | 1.41% | 3.77% | 13.98% | 16.85% | 18.21% | 0.28% | 9.32% | 0.48% |
| | 100 | 6.25% | 11.80% | 7.23% | 1.64% | 0.83% | 2.11% | 11.55% | 13.92% | 12.50% | 2.25% | 9.29% | 2.46% |
| | 200 | 6.58% | 11.84% | 7.65% | 1.46% | 0.64% | 1.91% | 10.22% | 13.27% | 13.21% | 3.97% | 10.41% | 3.93% |
| | 400 | 6.58% | 11.75% | 7.40% | 1.04% | 0.45% | 1.36% | 9.20% | 13.07% | 11.36% | 3.91% | 10.83% | 4.19% |
| | 600 | 6.75% | 11.85% | 7.67% | 0.86% | .370% | 1.00% | 8.65% | 13.06% | 10.19% | 5.13% | 10.95% | 5.89% |
| | 800 | 6.60% | 11.77% | 7.32% | 0.76% | 0.30% | 0.92% | 8.68% | 12.50% | 10.04% | 5.00% | 10.99% | 5.42% |
| | 1000 | 6.55% | 11.79% | 7.30% | 0.67% | 0.27% | 0.82% | 8.13% | 12.53% | 9.50% | 4.53% | 11.04% | 5.30% |
| 500 | 50 | 1.32% | 7.91% | 1.70% | 0.95% | 0.82% | 1.71% | 4.24% | 9.87% | 8.00% | 0.02% | 5.82% | 0.09% |
| | 100 | 1.39% | 7.84% | 1.77% | 0.66% | 0.65% | 1.19% | 3.10% | 9.26% | 6.12% | 0.15% | 6.49% | 0.12% |
| | 200 | 1.20% | 7.84% | 1.42% | 0.51% | 0.52% | 0.80% | 2.57% | 9.44% | 3.60% | 0.23% | 6.66% | 0.18% |
| | 400 | 1.28% | 7.80% | 1.49% | 0.36% | 0.32% | 0.54% | 2.65% | 8.46% | 3.03% | 0.62% | 7.01% | 0.31% |
| | 600 | 1.23% | 7.86% | 1.37% | 0.28% | 0.26% | 0.44% | 1.95% | 8.52% | 2.71% | 0.64% | 7.31% | 0.57% |
| | 800 | 1.21% | 7.87% | 1.38% | 0.29% | 0.20% | 0.43% | 2.38% | 8.57% | 2.55% | 0.67% | 7.44% | 0.53% |
| | 1000 | 1.15% | 7.84% | 1.35% | 0.20% | 0.18% | 0.28% | 1.71% | 8.22% | 2.00% | 0.51% | 7.41% | 0.64% |
| 600 | 50 | 0.18% | 5.79% | 0.25% | 0.21% | 0.70% | 0.37% | 0.78% | 7.94% | 2.46% | 0% | 4.26% | 0.03% |
| | 100 | 0.16% | 5.87% | 0.17% | 0.15% | 0.44% | 0.27% | 0.99% | 7.25% | 1.12% | 0% | 4.72% | 0.04% |
| | 200 | 0.15% | 5.87% | 0.19% | 0.10% | 0.34% | 0.27% | 0.49% | 6.78% | 1.05% | 0.01% | 5.19% | 0.03% |
| | 400 | 0.14% | 5.85% | 0.13% | % 0.07 | 0.23% | 0.15% | 0.34% | 6.32% | 0.61% | 0.03% | 5.31% | 0.02% |
| | 600 | 0.13% | 5.87% | 0.13% | 0.06% | 0.18% | 0.12% | 0.35% | 6.28% | 0.50% | 0.04% | 5.42% | 0.03% |
| | 800 | 0.12% | 5.88% | 0.10% | 0.04% | 0.14% | 0.10% | 0.27% | 6.20% | 0.44% | 0.05% | 5.51% | 0.02% |
| | 1000 | 0.12% | 5.86% | 0.14% | 0.04% | 0.13% | 0.10% | 0.29% | 6.10% | 0.50% | 0.06% | 5.58% | 0.03% |

the constraint (9e) becomes the following linear form.

$$2\tau_n \leq \gamma, \quad \forall n = 1, \ldots, N \tag{10}$$

Finally, probability inequality (3) with the divergence-based ambiguity set $\mathbb{S}_2$ could be transformed to a linear program with objective function (9a) and constraints (9b)-(9d) and (10), which could be solved efficiently.

## III. CASE STUDY

The proposed models are compared with a Gaussian mixture model (GMM) and tested using real data from wind farms in north China. The dataset includes hourly point forecasts and observed outputs from January 1st, 2006 to December 31st, 2015. To examine the probability of a 200 MW ramp event, we choose two consecutive hours in which wind power forecasts rest in the interval [495 MW, 505 MW] and [695 MW, 705 MW], respectively. We recover 1083 data pairs from the dataset. In case studies, only the probability of upward ramp events is estimated, and downward ramp can be treated in the same way. All experiments are conducted on a laptop with Intel i5-7300HQ CPU and 8G memory. Linear programs and second-order cone programs

**TABLE 3. Performances of different models.**

| | GMM-based model | Moment-based model | Divergence-based model |
|---|---|---|---|
| Data requirement | Abundant | Little | Moderate |
| Accuracy | Relatively accurate | Relatively conservative | Slightly conservative |
| Robustness | No guarantee | Guaranteed | Guaranteed |
| Computational burden | High | Very low | Low |

are established by YALMIP in MATLAB R2018a, and solved by CPLEX 12.8.

For the sake of comparison, the GMM is designed to estimate ramp probability. GMM is a mixture of several Gaussian distributions and could characterize the uncertainties obeying arbitrary distributions. The detailed GMM application in ramp forecasting is described in Appendix. In this paper, three Gaussian distributions are included in the GMM distribution. The wind ramp probability calculated from the GMM-based model, moment-based model and divergence-based model are termed as GMM-WRP, Moment-WRP and Div-WRP, respectively. The observed wind power ramp probability (O-WRP) from historical data is termed as O-WRP. The results with different ramp thresholds $R_U$ are compared in Fig. 2 and Table 1. It is observed that GMM-WRP and Div-WRP are very close to the O-WRP while Moment-WRP is always larger than them. This is because that GMM-based model and divergence-based model fully utilize the dispersion of historical data, while the moment-based one only considers two distribution parameters and neglect more useful information. Nevertheless, due to the consideration for the unimodality of forecast errors, the Moment-WRP decreases quickly with the increasing of ramp threshold $R_U$, and becomes comparable to GMM-WRP and Div-WRP when $R_U > 280$ MW. In Table 1, GMM-WRP is smaller than O-WRP when $R_U = 300$ MW and 400 MW, while Div-WRP is always larger than O-WRP. This indicates that compared to GMM-based model, the proposed divergence-based model could provide a more conservative but reliable estimation to ensure the safe operation of power system. In addition, the average computation time for estimating ramp probability with one given threshold is listed in Table 1. GMM-based model consumes much longer time than the proposed data-driven models, because its calculation consists of two steps: parameter estimation and Monte Carlo based probability calculation.

In the data-driven approaches, the number of historical samples is a critical factor that significantly influences the estimation results. To investigate the impact, we firstly generate 10000 wind power scenarios based on the aforementioned 1083 data pairs using the scenario generation method in [42]. Next, we randomly select a set of samples with a given number of elements from the 10000 scenarios, and test the performances of three models. This procedure is repeated 100 times. Results are shown in Fig. 3 for $R_U = 300$ MW. With the increasing of sampled data, the volatilities of three models all decrease, indicating that the number of samples has a notable effect on the stability of probability estimation. In addition, the volatility of moment-based model is apparently smaller than those of GMM-based model and divergence-based model, because it uses the least historical data information and thus is barely affected by the variation of sample number.

Table 2 shows the statistical information for estimated ramp probabilities with respect to different values of $R_U$ and numbers of sampled data $N_S$. When more sampled data are used, the standard deviation and difference between maximum and minimum for ramp probabilities decrease significantly. When $N_S \leq 200$, GMM-WRP and Div-WRP are more sensitive compared with Moment-WRP, because the dispersion information used in these two models must be based on a certain number of samples, while the moments are less sensitive with moderate data. When $N_S \geq 400$, GMM-WRP and Div-WRP perform much better than Moment-WRP in terms of estimation accuracy. Nevertheless, GMM-WRP could be either optimistic or pessimistic, while Div-WRP is always conservative, which could be observed from the mean value when $R_U = 300$ MW and 400 MW.

In summary, for the system operator, moment-based model is recommended when we only have limited historical data. Even in the case that there is no data at all, the model could also be applied if we provide proper guess of mean value and mean-absolute deviation. On the contrary, in the case that there are plenty of data, divergence-based model is recommended because it provides a more accurate result but still with distributional robustness guarantee to limit the probability upper bound of a ramp event.

## IV. CONCLUSION

This paper proposes data-driven probability inequality models to estimate the probability upper bound of wind power ramp events with a given ramp threshold. Moment information and unimodality of wind power are utilized to construct the moment-based ambiguity set; Wasserstein metric is applied to describe the distance between two probability distributions for constructing the divergence-based ambiguity set. The uncertainty quantification models are developed based on the above sets, and show some appealing features: first, they make no reference to the exact distributions of uncertain factors; second, they rely on computationally tractable convex optimization. Finally, performances of the discussed models in this paper are summarized in Table 3.

## APPENDIX

Suppose the probability distribution function (PDF) $f_0(\boldsymbol{w})$ represents the reference multivariate distribution of actual wind power outputs $\boldsymbol{w} = [w_1, w_2]$, and Gaussian mixture model (GMM) is used to estimate $f_0(\boldsymbol{w})$. It is characterized by the weights, mean vectors and covariance matrixes, and formulated as follows:

$$f_0(\boldsymbol{w}) = \sum_{k=1}^{K} \pi_k N(\boldsymbol{w}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \tag{11}$$

where $K$ denotes the number of Gaussian distributions; $\pi_k$ denotes the nonnegative weight of $k$ th Gaussian distribution, requiring $\sum_{k=1}^{K} \pi_k = 1$; $N(\boldsymbol{w}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$ denotes the $k$th Gaussian distribution, defined as follows:

$$N(\boldsymbol{w}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{1}{(2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(\boldsymbol{w}-\boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1}(\boldsymbol{w}-\boldsymbol{\mu}_k)} \tag{12}$$

where $\boldsymbol{\mu}_k$ denotes the mean vector; $\boldsymbol{\Sigma}_k$ denotes the covariance matrix; $n$ denotes the dimension of variable $\boldsymbol{w}$.

Based on model (11), assuming that we have a series of wind power outputs samples $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_M$, the parameters $\pi_k$, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ of $f_0(\boldsymbol{w})$ could be estimated by expectation maximization (EM) algorithm [43]. After the reference distribution $f_0(\boldsymbol{w})$ is acquired, the Monte Carlo simulation is applied to calculate the probability $\text{Pr}_0\{w \in W\}$ with respect to $f_0(\boldsymbol{w})$ [44].

## REFERENCES

[1] C. Kamath, "Understanding wind ramp events through analysis of historical data," in *Proc. IEEE PES Transmiss. Distrib. Conf. Expo.*, Apr. 2010, pp. 1–6.
[2] C. Ferreira, J. Gama, L. Matias, A. Botterud, and J. Wang, "A survey on wind power ramp forecasting," Argonne Nat. Lab. (ANL), Argonne, IL, USA, Tech. Rep. ANL/DIS-10-13, 2011.
[3] C. Kamath, "Associating weather conditions with ramp events in wind power generation," in *Proc. IEEE/PES Power Syst. Conf. Expo. (PSCE)*, Mar. 2011, pp. 1–8.
[4] L. Yang and C. Zhang, "Study on control strategy of wind power ramping," *Electr. Meas. Instrum.*, vol. 52, no. 7, pp. 30–34, Apr. 2015.
[5] N. Navid and G. Rosenwald, "Market solutions for managing ramp flexibility with high penetration of renewable resource," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 784–790, Oct. 2012.
[6] 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, Int. J. Prog. Appl. Wind Power Convers. Technol.*, vol. 10, no. 5, pp. 453–470, Sep. 2007.
[7] 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*, Jul. 2011, pp. 1–3.
[8] 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.
[9] A. Florita, B.-M. Hodge, and K. Orwig, "Identifying wind and solar ramping events," in *Proc. IEEE Green Technol. Conf.*, Apr. 2013, pp. 147–152.
[10] J. Zhang, A. Florita, B.-M. Hodge, and J. Freedman, "Ramp forecasting performance from improved short-term wind power forecasting," in *Proc. ASME Int. Design Eng. Tech. Conf. Comput. Inf. Eng. Conf.*, Aug. 2014, p. V02AT03A022.
[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] H. Zheng and A. Kusiak, "Prediction of wind farm power ramp rates: A data-mining approach," *J. Solar Energy Eng.*, vol. 131, no. 3, p. 031011, Aug. 2009.
[13] B. Greaves, J. Collins, J. Parkes, and A. Tindal, "Temporal forecast uncertainty for ramp events," *Wind Eng.*, vol. 33, no. 4, pp. 309–319, Jun. 2009.
[14] Y. Liu, Y. Sun, D. Infield, Y. Zhao, S. Han, and J. Yan, "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.
[15] M. Cui, Y. Sun, and D. Ke, "Wind power ramp events forecasting based on atomic sparse decomposition and BP neural networks," *Autom. Electr. Power Syst.*, vol. 38, no. 12, pp. 6–11, Jun. 2014.
[16] 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.
[17] A. Bossavy, R. Girard, and G. Kariniotakis, "Forecasting uncertainty related to ramps of wind power production," in *Proc. Eur. Wind Energy Conf. Exhib. (EWEC)*, vol. 2, Apr. 2010, pp. 1–9.
[18] 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, 2016.
[19] C. Ferreira, J. Gama, V. Miranda, and A. Botterud, "Probabilistic ramp detection and forecasting for wind power prediction," in *Reliability and Risk Evaluation of Wind Integrated Power Systems*. Berlin, Germany: Springer, 2013, pp. 29–44.
[20] 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.
[21] M. Cui, F. Cong, Z. Wang, and Z. Jie, "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.
[22] Y. Li, P. Musilek, E. Lozowski, C. Dai, T. Wang, and Z. Lu, "Temporal uncertainty of wind ramp predictions using probabilistic forecasting technique," in *Proc. IEEE 2nd Int. Conf. Big Data Comput. Service Appl. (BigDataService)*, Mar. 2016, pp. 166–173.
[23] 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.
[24] M. Cui, V. Krishnan, B.-M. Hodge, and J. Zhang, "A copula-based conditional probabilistic forecast model for wind power ramps," *IEEE Trans. Smart Grid*, to be published.
[25] D. Bertsimas and A. Thiele, "A robust optimization approach to inventory theory," *Oper. Res.*, vol. 54, no. 1, pp. 150–168, Feb. 2006.
[26] E. Delage and Y. Ye, "Distributionally robust optimization under moment uncertainty with application to data-driven problems," *Oper. Res.*, vol. 58, no. 3, pp. 595–612, 2010.
[27] W. Wiesemann, D. Kuhn, and M. Sim, "Distributionally robust convex optimization," *Oper. Res.*, vol. 62, no. 6, pp. 1358–1376, 2014.
[28] J. Goh and M. Sim, "Distributionally robust optimization and its tractable approximations," *Oper. Res.*, vol. 58, no. 4, pp. 902–917, Aug. 2010.
[29] R. Jiang and Y. Guan, "Data-driven chance constrained stochastic program," *Math. Program.*, vol. 158, no. 1, pp. 291–327, 2016.
[30] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations," *Math. Program.*, vol. 171, nos. 1–2, pp. 115–166, Sep. 2018.
[31] H. Owhadi, C. Scovel, T. J. Sullivan, M. McKerns, and M. Ortiz, "Optimal uncertainty quantification," *SIAM Rev.*, vol. 55, no. 2, pp. 271–345, 2010.
[32] G. A. Hanasusanto, V. Roitch, D. Kuhn, and W. Wiesemann, "A distributionally robust perspective on uncertainty quantification and chance constrained programming," *Math. Program.*, vol. 151, no. 1, pp. 35–62, 2015.
[33] Z. Li, Y. Lin, Y. Zhao, X. Song, J. Teng, and J. Jin, "Short-term wind power prediction based on extreme learning machine with error correction," *Protect. Control Mod. Power Syst.*, vol. 1, no. 1, p. 1, Jun. 2016.
[34] P. Pinson and G. Kariniotakis, "Conditional prediction intervals of wind power generation," *IEEE Trans. Power Syst.*, vol. 25, no. 4, pp. 1845–1856, Nov. 2010.
[35] J. A. Rice, *Mathematical Statistics and Data Analysis*. Scarborough, ON, Canada: Nelson Education, 2006.
[36] D. Bertsimas and I. Popescu, "Optimal inequalities in probability theory: A convex optimization approach," *SIAM J. Optim.*, vol. 15, no. 3, pp. 780–804, 2005.
[37] L. Vandenberghe, S. Boyd, and K. Comanor, "Generalized Chebyshev bounds via semidefinite programming," *SIAM Rev.*, vol. 49, no. 1, pp. 52–64, 2007.
[38] B. P. G. Van Parys, P. J. Goulart, and D. Kuhn, "Generalized Gauss inequalities via semidefinite programming," *Math. Programm.*, vol. 156, no. 1, pp. 271–302, 2015.
[39] S. Dharmadhikari and K. Joag-Dev, *Unimodality, Convexity, and Applications*. Amsterdam, The Netherlands: Elsevier, 1988.

[40] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, vol. 2. Philadelphia, PA, USA: SIAM, 2001.
[41] N. Fournier and A. Guillin, "On the rate of convergence in Wasserstein distance of the empirical measure," *Probab. Theory Related Fields*, vol. 162, nos. 3–4, pp. 707–738, Aug. 2015.
[42] P. Pinson and R. Girard, "Evaluating the quality of scenarios of short-term wind power generation," *Appl. Energy*, vol. 96, pp. 12–20, Aug. 2012.
[43] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Process. Mag.*, vol. 13, no. 6, pp. 47–60, Nov. 1996.
[44] N. Metropolis and S. Ulam, "The Monte Carlo method," *J. Amer. Statist. Assoc.*, vol. 44, no. 247, pp. 335–341, 1949.

**YANG CAO** (S'18) received the B.Sc. degree in electrical engineering from Tsinghua University, Beijing, China, in 2017, where he is currently pursuing the Ph.D. degree.

His research interests include renewable power generation, and planning and operation of multi-carrier energy systems.

**WEI WEI** (M'15–SM'18) received the bachelor's and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2008 and 2013, respectively, where he was a Postdoctoral Researcher, from 2013 to 2015.

He was a Visiting Scholar with Cornell University, Ithaca, NY, USA, in 2014, and with Harvard University, Cambridge, MA, USA, in 2015. He is currently an Associate Professor with Tsinghua University. His research interests include applied optimization and energy system economics.

**CHENG WANG** (M'18) received the B.Sc. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2012 and 2017, respectively.

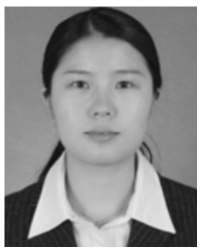
From 2015 to 2016, he was a Visiting Ph.D. Student with the Argonne National Laboratory, Argonne, IL, USA. He is currently a Lecturer with North China Electric Power University, Beijing. His research interest includes optimization and control of integrated energy systems. He was an Outstanding Reviewer of the IEEE Transactions on Power Systems, in 2016.

**SHENGWEI MEI** (F'15) received the B.Sc. degree in mathematics from Xinjiang University, Urumqi, China, the M.Sc. degree in operations research from Tsinghua University, Beijing, China, and the Ph.D. degree in automatic control from the Chinese Academy of Sciences, Beijing, in 1984, 1989, and 1996, respectively.

He is currently a Professor with Tsinghua University. His research interests include power system complexity and control, game theory, and its application in power systems.

**SHAOWEI HUANG** (M'11) received the B.S. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2006 and 2011, respectively, where he is currently an Associate Professor with the Department of Electrical Engineering and Applied Electronic Technology.

His research interest includes parallel and distributed computing and their applications to power systems.



**XUEMIN ZHANG** (M'06) received the B.S. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2001 and 2006, respectively, where she is currently an Associate Professor with the Department of Electrical Engineering.

Her research interest includes power system analysis and control, especially stabilization control, cascading failure modeling, and mitigation.