IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received August 29, 2019, accepted September 16, 2019, date of publication September 18, 2019,
date of current version September 30, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2942178

Data-Driven Distributionally Robust Chance-Constrained Unit Commitment With Uncertain Wind Power

ZHICHAO SHI, (Student Member, IEEE), **HAO LIANG**, (Member, IEEE), AND **VENKATA DINAVAHI**, (Senior Member, IEEE)
Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 1H9, Canada

Corresponding author: Zhichao Shi (zhichao1@ualberta.ca)

The work of Z. Shi was supported by the China Scholarship Council (CSC). This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

ABSTRACT The Unit commitment (UC) problem in power systems has been studied for a long time; however, many new challenges have emerged in the UC problem with the increasing penetration of renewable generation which is intermittent and uncertain. Compared with the common uncertainty modeling methods including stochastic programming and robust optimization, in this paper, we develop a data-driven distributionally robust chance-constrained (DDRC) UC model. The proposed two-stage UC model focuses on the commitment decision and dispatch plan in the first stage, and considers the worst-case expected cost for possible power imbalance or re-dispatch in the second stage. To capture the uncertainty of wind power distribution, a distance-based ambiguity set is designed which can be constructed in a data-driven manner. Based on the ambiguity set, the original complicated UC problem is reformulated to a tractable optimization problem which is then solved by the column-and-constraint generation (C&CG) algorithm. The performance of the the proposed approach is validated by case studies with different test systems including the IEEE 6-bus test system, modified IEEE 118-bus system and a practical-scale system, especially the value of data in controlling the conservativeness of the problem.

INDEX TERMS Ambiguity set, chance-constrained unit commitment, data-driven method, distributionally robust optimization (DRO).

NOMENCLATURE

The main notations used in this paper are listed below for quick reference.

A. INDICES AND SETS

$b/\mathcal{B}$	Index/set of buses.
$\mathcal{D}/\mathcal{D}_1$	Ambiguity set.
$i/\mathcal{I}$	Index/set of generation units.
$l/\mathcal{G}$	Index/set of transmission lines.
$t/\mathcal{T}$	Index/set of scheduling periods.

B. PARAMETERS

a_i, e_i, c_i	Coefficients of fuel cost function for generator i.
d_t^b	Load demand at bus b at time t.
$F_i(.)$	Fuel cost function of generator i.
K_l^b	Flow distribution factor for line l at bus b.
K_0	Number of pieces for the piecewise linear approximation.
L_l	Power flow capacity of transmission line l.
$p^n/\hat{p}^n$	True probability/nominal probability for a certain scenario.
$\overline{RD}_i$	Shut-down ramp-down rate limit for generator i.
$\overline{RU}_i$	Start-up ramp-up rate limit for generator i.
RU_i/RD_i	Ramp-up/ramp-down limit for generator i.
SU_i/SD_i	Start-up/shut-down cost of generator i.
T_i^{up}/T_i^{dn}	Minimum up-time/down-time for generator i.
$w_t^b(\xi)$	Wind power output at bus b at time t for scenario ξ.
$\underline{x}_i/\overline{x}_i$	Minimum/maximum output of generator i.
δ	Power imbalance tolerance level.
ϵ	Risk level in the chance constraint.
π_t^{gen}	Cost coefficient for generator up/down re-dispatch.

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

π_t^{ls}	Cost coefficient for load shedding.
θ	Tolerance level in the ambiguity set.
Γ	Budget of uncertainty in the uncertainty set.

C. VARIABLES

$d_t^{ls,b}$	Load shedding of load bus b at time t.
r_{it}^u/r_{it}^d	Up/down re-dispatch power at time t for unit i.
u_{it}/v_{it}	Binary start-up/shut-down variable for unit i at time t.
x_{it}	Output of generator i at time t.
y_{it}	Binary variable indicating the commitment status of unit i at time t.
α_t, β_t^n, γ_t^n, λ_t	Dual variables for the chance constraint reformulation at time t.
ρ_{bt}^+, ρ_{bt}^-	Binary variables used to define the uncertainty set of wind power.
σ_{bt}^{1+}, σ_{bt}^{1-}, σ_{btl}^{2+}, σ_{btl}^{2-}	Auxiliary variables to linearize bilinear terms in the second stage problem.
μ_t^1, μ_{it}^2, μ_{it}^3, μ_{it}^4, μ_{it}^5, μ_{tl}^6, μ_{tl}^7	Dual variables in the second stage problem.

I. INTRODUCTION

As a critical application in power system operation, unit commitment (UC) problem has been studied for a long time which aims to reduce system cost and improve reliability by optimal scheduling of generation units. In recent years, the deployment and utilization of renewable energy sources, especially wind energy, has increased significantly in existing power systems. However, due to the high-level penetration of intermittent and unpredictable renewable energy, new challenges about secure and reliable system operation also arise such as large reserve capacity demand, divergence of Area Control Error (ACE) and possible power imbalance [1]. Consequently, it is imperative to incorporate associated uncertainties into UC problems with renewable generation so that more reliable solutions can be attained.

Although UC problem is usually nonconvex and very difficult to solve, many efforts have been developed for UC problems with uncertainties in the past decades. In particular, the stochastic optimization methods [2] attract the most attention, which can be typically classified into two categories: stochastic programming and robust optimization methods. Stochastic programming is a traditional method to deal with data uncertainty and was first investigated to solve uncertain UC problems [3]. For example, in [4], a security-constrained UC (SCUC) problem with uncertain wind power is studied and the wind power is assumed to follow normal distribution. In [5], a stochastic UC model considering the uncertain load and outages is proposed, and the random variables are represented by scenarios trees. Generally, it is assumed that the probability distributions of random variables are known in stochastic programming methods, and the objective is to minimize the expected total system cost. However, the exact probability distribution is usually hard to be known in practice. In addition, stochastic programming methods suffer from heavy computational burden as substantial scenarios are required to comprehensively represent the probability distribution.

Robust optimization is another popular method to deal with uncertainties in UC problems. Compared with stochastic programming, the true probability distribution is not required in robust optimization, and the random variables are represented by some uncertainty sets. A vast number of literature about robust UC problems has been reported, such as the typical two-stage robust UC models [6]–[8] which consist of first-stage commitment decision and second-stage recourse action. In addition, multistage robust UC models have also been studied recently by taking into account the non-anticipativity of dispatch decisions [1], [9]. Compared with stochastic programming, robust optimization ignores the probabilistic information and tries to find the minimal cost under the worst-case scenario within the uncertainty set. Although the solution is robust against all uncertainty realizations, it may be over-conservative since the worst-case scenario rarely occurs in practice.

To address the shortcomings of stochastic programming and robust optimization methods, an alternative method, distributionally robust optimization (DRO), has attracted much attention recently [10]. The DRO method aims at optimizing an uncertain problem under the worst-case distribution from a so-called ambiguity set. The ambiguity set is a family of probability distributions which share some certain statistical information. Since partial distribution information is utilized in DRO method, the conservativeness of the solution from this method is between robust optimization and stochastic programming. In recent years, the DRO method has been widely applied to solve power system optimization problems, especially the moment-based DRO method [11]. For instance, in [12], a moment-based DRO model is proposed for UC problem, and linear decision rule is used to reformulate and solve the intractable problem. In [13], DRO approach is used to solve the contingency-constraint UC problem and the ambiguity set of contingency probability distributions is constructed based on available moment information. Similarly, the moment-based DRO method is used for co-optimization of energy and reserve dispatch in [14], and the problem is reformulated to a tractable semidefinite programming (SDP) problem.

Moment-based DRO method only considers moment information such as expectation and variance. However, the actual distribution contains more information than moments. In practice, a number of historical data of random variable are usually available from which we can obtain more valuable distribution information, e.g., an estimated distribution by data fitting. Therefore, distance-based or data-driven DRO methods have been investigated in some very recent studies [15]. In [16], a distance-based DRO model is studied for UC problem, and the ambiguity set is constructed based on

Kullback-Leibler (KL) divergence. Data-driven DRO method is also reported for UC problem [17] where the confidence band of cumulative distribution function (CDF) is used to construct the ambiguity set. Similarly, L_1 norm and L_∞ norm are used to construct confidence sets in a data-driven manner for stochastic UC problem [18], [19]. With the same confidence sets, a new duality-free decomposition method is proposed to solve the distributionally robust UC problem in [20]. Wasserstein metric based DRO method has also been studied for UC problem [21], [22] which constructs the ambiguity set based on Wasserstein ball. In addition, data-driven DRO method is also studied for reserve and energy scheduling problem [23] which considers the Wasserstein distance. According to abundant related works, DRO based optimization problems are usually very complicated and even intractable, and different ambiguity set construction methods lead to various problem reformulation methods.

Considering the advantage of distance-based DRO methods and the availability of historical wind power data, we study a data-driven distributionally robust chance-constrained (DDRC) UC problem under uncertainty in this work. The studied UC problem here is formulated as a two-stage model. In the first stage, with a chance constraint restricting the probability of power imbalance, the commitment decision and a base-case dispatch plan are determined, and in the second stage, we try to find the operational risk or expected re-dispatch cost caused by load curtailment or wind power spillage under the worst-case wind power distribution. Note that the proposed model is different from the common two-stage model which determines the first-stage commitment decision and second-stage recourse actions. In a traditional two-stage UC model, the valuable information is the commitment decisions, and the critical problem is how to obtain the recourse action with uncertainties [24]. Compared with the common UC model, the proposed model which considers the re-dispatch and load shedding decisions in the second stage can enhance the deployment of flexible resources such as fast-response generators and adjustable load demands.

Although a few works studied a similar two-stage model [1], [19], they adopted various uncertainty modeling methods. Compared with the popular stochastic programming or robust methods in previous literature, DRO method is studied in this work, and the uncertainty of wind power distribution is captured by a distance-based ambiguity set, more specifically, a set with the form of L_1 norm which can be constructed from historical data. Based on the ambiguity set, the proposed complex DDRC UC problem can be reformulated into a tractable optimization problem, thus solved by some existing decomposition algorithms such as the column-and-constraint generation (C&CG) algorithm. Compared with [19], an improved second-stage model is studied to include both system and nodal uncertainty in this work, and we adopt a new ambiguity set to model the uncertainty which also leads to different derivation. Although the data-driven ambiguity set is also studied in [18], several differences can be identified in this work by comparison. In the two-stage model of this paper, we consider both the commitment decision and base-case dispatch plan in the first stage and a chance constraint is introduced to restrict the possible power imbalance, while in [18], only commitment decision is considered in the first stage and the power balance constraint is used. For solution method, we propose a new problem reformulation method to deal with the worst-case expectation. In addition, our reformulation method of the second-stage objective can better show the value of data in controlling the conservatism of the problem. Correspondingly, the contributions of this work are summarized as follows:

1) We propose a data-driven distributionally robust chance-constrained two-stage UC model in this work which determines the commitment decision and base-case dispatch plan in the first stage and minimizes the re-dispatch cost due to possible power imbalance in the second stage. Specifically, a chance constraint is used to restrain the power imbalance in the first stage, and the re-dispatch cost resulted from load curtailment or wind power spillage is considered in the second stage. This is a new model by combining the new DRO technique and two-stage chance-constrained model compared with those in previous literature.
2) A new problem reformulation method is proposed with the studied distance-based ambiguity set. Particularly, based on the proposed ambiguity set, the original complicated UC problem is reformulated into a tractable two-stage optimization problem which can be solved in a decomposition framework, i.e., the second-stage objective function is transformed into a convex combination of conditional value-at-risk (CVaR) and worst-case cost.
3) According to the available historical wind power data size, the constructed ambiguity set can be adjusted, thus the conservativeness of the solution can also be altered accordingly. In addition, the new reformulation method helps explicitly reveal the value of additional data in reducing the conservatism of the problem, and we can flexibly acquire the corresponding stochastic problem and robust problem.

The remainder of this paper is organized as follows. The proposed DDRC UC problem as well as the ambiguity set is mathematically described in Section II. Section III proposes the solution methodology for the UC problem which includes problem reformulation method and the introduction of C&CG solution algorithm. In Section IV, we conduct case studies based on the IEEE 6-bus test system and modified IEEE 118-bus system and a practical-scale 319-bus system to validate the effectiveness of the proposed approach. Finally, the conclusions are drawn in Section V.

II. PROBLEM FORMULATION

In this section, we first formulate the DDRC two-stage UC problem which includes various constraints and objective function. The ambiguity set and its construction are then introduced to capture the uncertain distribution of renewable generation (i.e., wind power).

A. UC MATHEMATICAL MODEL

Two-stage UC models are widely studied in previous literature, and most of them focus on commitment decision in the first stage and optimal recourse in the second stage. For example, two-stage robust UC problems are studied in [25], [26], and two-stage UC problems considering distributional uncertainty are investigated in [13], [27]. In these works, the here-and-now decision variables are usually commitment decision and wait-and-see decision variables are recourse action. By contrast, we develop a new DDRC two-stage UC model in this work which considers the traditional UC model with a chance constraint in the first stage and attains the best corrective actions by minimizing the expected re-dispatch cost in the second stage. In addition, two-stage energy and reserve dispatch problems are also widely studied [28] [29] which considers the base-case dispatch plan in the first stage and re-dispatch in the second stage, while UC decisions are not covered in these works. From this aspect, the proposed UC model can be regarded as a combination of traditional two-stage UC problem with the energy and reserve dispatch problem, and it is solved with the new DRO method instead of the previous stochastic or robust optimization methods. The detailed formulation of the proposed model is as follows:

$$\min \sum_t \sum_i [SU_i u_{it} + SD_i v_{it} + F_i(x_{it})] + \max_{P \in \mathcal{D}} E_P[Q(y,u,v,x,\xi)] \tag{1}$$

$$s.t. \; -y_{i(t-1)} + y_{it} - y_{ih} \le 0, \quad \forall t \in \mathcal{T},\ \forall i \in \mathcal{I},\ 1 \le h-(t-1) \le T_i^{up} \tag{2}$$

$$y_{i(t-1)} - y_{it} + y_{ih} \le 1, \quad \forall t \in \mathcal{T},\ \forall i \in \mathcal{I},\ 1 \le h-(t-1) \le T_i^{dn} \tag{3}$$

$$-y_{i(t-1)} + y_{it} - u_{it} \le 0, \quad \forall t \in \mathcal{T},\ \forall i \in \mathcal{I} \tag{4}$$

$$y_{i(t-1)} - y_{it} - v_{it} \le 0, \quad \forall t \in \mathcal{T},\ \forall i \in \mathcal{I} \tag{5}$$

$$y_{it},\ u_{it},\ v_{it} \in \{0,1\}, \quad \forall t \in \mathcal{T},\ \forall i \in \mathcal{I} \tag{6}$$

$$\underline{x}_i y_{it} \le x_{it} \le \overline{x}_i y_{it}, \quad \forall t \in \mathcal{T},\ \forall i \in \mathcal{I} \tag{7}$$

$$x_{it} - x_{i(t-1)} \le \overline{RU}_i u_{it} + RU_i y_{i(t-1)}, \quad \forall t \in \mathcal{T},\ \forall i \in \mathcal{I} \tag{8}$$

$$x_{i(t-1)} - x_{it} \le \overline{RD}_i v_{it} + RD_i y_{it}, \quad \forall t \in \mathcal{T},\ \forall i \in \mathcal{I} \tag{9}$$

$$-L_l \le \sum_{b \in \mathcal{B}} K_l^b \Big(\sum_i x_{it}^b + w_t^b(\xi) - d_t^b\Big) \le L_l, \quad \forall t \in \mathcal{T},\ l \in \mathcal{G} \tag{10}$$

$$Pr\Big(-\delta \le \sum_i x_{it} + \sum_b w_t^b(\xi) - \sum_b d_t^b \le \delta\Big) \ge 1-\epsilon, \quad \forall t \in \mathcal{T} \tag{11}$$

where the objective function in (1) contains the start-up cost, shut-down cost, fuel cost and worst-case expected penalty cost. Constraints (2) and (3) represent the minimum up-time and minimum down-time constraints, respectively. Constraints (4) and (5) restrict the start-up and shut-down operation, respectively. Constraint (6) lists the binary variables representing the generators' statuses of on/off, start-up and shut-down. Constraint (7) represents the generation capacity limits. Constraints (8) and (9) enforce the ramp-up and ramp-down rates, respectively. Constraint (10) denotes the power transmission line capacity limits which is from DC power flow model. Note that the variable x_{it}^b represents the output of the ith generator located at bus b at time t. Constraint (11) defines the chance constraint for possible power imbalance, and the small violation probability should be less than a predefined risk level.

After determining the commitment decision and base-case dispatch plan, the system operators conduct re-dispatch strategy by considering all uncertainty realizations in the second stage. Discrete scenarios are often used to replace the continuous distribution to solve the difficult numerical computation. In this work, we assume that the uncertain parameter ξ has a finite support, i.e., there are a finite number of realizations (e.g., scenarios $\xi_1, \xi_2, \ldots \xi_N$) for the uncertain wind power output [18]. However, the true probability distribution is unknown here and is described by the ambiguity set.

In formulation (1), the operational risk for the second stage problem is considered which also represents the expected penalty cost or re-dispatch cost [23]. This cost is caused by load curtailment or over-generation of the system with the reveal of uncertain wind power. Note that wind power curtailment is not considered since finite scenarios are assumed in this work as mentioned above. In other words, we need to re-adjust the generation or consider load shedding with possible worst-case wind power in the second stage. Additionally, we should note that both system-level and nodal-level uncertainty modeling should be investigated to identify the real worst case. Specifically, we have the following formulation for the second stage problem:

$$Q(y,u,v,x,\xi) = \min \sum_t [\pi_t^{gen} \sum_i (r_{it}^u + r_{it}^d) + \pi_t^{ls} \sum_b d_t^{ls,b}] \tag{12}$$

$$s.t. \; \sum_i (r_{it}^u - r_{it}^d) + \sum_b d_t^{ls,b} = \sum_b d_t^b - \sum_i x_{it} - \sum_b w_t^b(\xi), \quad \forall t \in \mathcal{T},\ \forall i \in \mathcal{I} \tag{13}$$

$$x_{it} + r_{it}^u \le \overline{x}_i y_{it}, \quad \forall t \in \mathcal{T},\ i \in \mathcal{I} \tag{14}$$

$$x_{it} - r_{it}^d \ge \underline{x}_i y_{it}, \quad \forall t \in \mathcal{T},\ i \in \mathcal{I} \tag{15}$$

$$0 \le r_{it}^u \le RU_i, \quad \forall t \in \mathcal{T},\ i \in \mathcal{I} \tag{16}$$

$$0 \le r_{it}^d \le RD_i, \quad \forall t \in \mathcal{T},\ i \in \mathcal{I} \tag{17}$$

$$\sum_{b \in \mathcal{B}} K_l^b \Big(\sum_i x_{it}^b + w_t^b(\xi) - d_t^b + \Delta p_t^b\Big) \le L_l, \quad \forall t \in \mathcal{T},\ l \in \mathcal{G} \tag{18}$$

$$\sum_{b \in \mathcal{B}} K_l^b \Big(\sum_i x_{it}^b + w_t^b(\xi) - d_t^b + \Delta p_t^b\Big) \ge -L_l, \quad \forall t \in \mathcal{T},\ l \in \mathcal{G} \tag{19}$$

$$\Delta p_t^b = \sum_i (r_{it}^u - r_{it}^d) + d_t^{ls,b}, \quad d_t^{ls,b} \ge 0,\ \forall\, t \in \mathcal{T} \tag{20}$$

where the objective function in (12) includes the possible re-dispatch cost of generators and load shedding cost. Constraint (13) denotes the system power balance with re-dispatch action. The re-dispatch amount of each unit should be limited by the available generator capacity as given in (14)-(15) and the corresponding ramp rate given in (16)-(17). The set of constraints (18) and (19) show the transmission line flow considering the adjusted dispatch action as given in (20). Note that the re-dispatch variables in (20) correspond to the generator units in (19) and the symbol b is omitted here for consistency. In addition, we assume that all generators can flexibly adjust their output in the re-dispatch process with corresponding ramping rate in this work [1].

In the objective function, the fuel cost function of generation units $F_i(x_{it})$ is typically a non-decreasing quadratic function. For practical and computational purposes, the quadratic fuel cost function is usually approximated by a piece-wise linear function. In this work, we use the following piece-wise linear model for fuel cost calculation [30], [31]:

$$0 \le \hat{x}_{it}^k \le \Delta x_i^k y_{it}, \quad \forall k = 1, \ldots, K_0 \tag{21a}$$
$$\Delta x_i^k = \frac{\overline{x}_i - \underline{x}_i}{K_0} \tag{21b}$$
$$x_{i,ini}^k = (k-1)\Delta x_i^k + \underline{x}_i \tag{21c}$$
$$x_{i,fin}^k = \Delta x_i^k + x_{i,ini}^k \tag{21d}$$
$$C_{i,ini}^k = a_i(x_{i,ini}^k)^2 + e_i x_{i,ini}^k + c_i \tag{21e}$$
$$C_{i,fin}^k = a_i(x_{i,fin}^k)^2 + e_i x_{i,fin}^k + c_i \tag{21f}$$
$$s_i^k = \frac{C_{i,fin}^k - C_{i,ini}^k}{\Delta x_i^k} \tag{21g}$$
$$x_{it} = \underline{x}_i y_{it} + \sum_k \hat{x}_{it}^k \tag{21h}$$
$$F_i(x_{it}) = a_i(\underline{x}_i)^2 + e_i \underline{x}_i + c_i y_{it} + \sum_k s_i^k \hat{x}_{it}^k \tag{21i}$$

where new variable $\hat{x}_{it}^k$ is introduced, and its relationship with the decision variable x_{it} is described in (21h). In addition, K_0 is the number of pieces, and s_i^k is the slope of each linear piece. The coefficients a_i, e_i and c_i are dependent on the physical characteristic of the generators.

B. AMBIGUITY SET CONSTRUCTION

As discussed above, the true probability distribution of wind power is unknown and ambiguous in practice. However, we can get partial information about the true distribution from available historical data and construct the ambiguity sets to capture the uncertainty of distribution. In this work, we focus on a distance-based ambiguity set which has the following form:

$$\mathcal{D} = \{P \in \mathcal{P} : \text{dist}(P, P_0) \le \theta\} \tag{22}$$

where $\mathcal{P}$ is the set of all possible distributions, P_0 and θ are the nominal distribution and divergence tolerance level, respectively. The dist() function denotes a distance measure between two distributions, such as the KL divergence [16]. Since discrete distribution for wind power is considered in this study, we adopt the L_1-norm distance to construct the ambiguity set. In addition, the nominal distribution derived from historical data tends to converge to the true distribution under L_1-norm as the data size increases [18]. Consequently, the ambiguity set used in this work, denoted by $\mathcal{D}_1$, can be expressed as follows:

$$\mathcal{D}_1 = \{p \in [0,1]^N : \sum_{n=1}^{N} |p^n - \hat{p}^n| \le \theta, \sum_{n=1}^{N} p^n = 1\} \tag{23}$$

where p^n and $\hat{p}^n$ are the true probability and nominal probability respectively corresponding to index n, and N is the number of scenarios. Note that the ambiguity set $\mathcal{D}_1$ is a specific set compared with the general form $\mathcal{D}$, and the true probability in this set is unknown which can be described with the nominal probability estimated from historical data.

To construct the set $\mathcal{D}_1$, a critical step is the determination of nominal distribution and tolerance level. For the nominal or reference distribution, we can derive it with nonparametric estimation method in a data-driven manner. Specifically, assuming that there are A historical data samples available in total, we can estimate the reference distribution with a histogram. For example, according to the number of scenarios, we can construct a histogram with N bins. Count the number of data samples in each bin, say, $A_1, A_2, \ldots, A_N$ and $A = \sum_{n=1}^{N} A_n$, then we can use the frequency A_n/A in each bin as the nominal probability $\hat{p}^n$. For simplicity, we denote the nominal distribution as $P_0 = (\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_N)$. Note that the nominal distribution from the histogram is only an estimation of the true distribution and it may be obtained from other methods.

Based on the size of available historical data, we can also define proper tolerance level (i.e., θ) to construct effective ambiguity set. Following the above histogram approach, we can determine the tolerance level according to the Proposition 8 in [32] as follows:

$$\theta = \sqrt{\chi^2_{N-1,1-\tilde{\alpha}}/A} \tag{24}$$

where $1 - \tilde{\alpha}$ is the confidence level that the data-driven ambiguity set $\mathcal{D}_1$ with nominal distribution P_0 and θ contains the unknown real distribution. From (24), we can find that the value of θ is mainly determined by the confidence level and historical data size. Since the value of θ decreases as the size of historical data increases, the true distribution becomes much closer to the reference distribution. When the data size goes to infinity, θ will be zero. In addition, since $\sum_{n=1}^{N} |p^n - \hat{p}^n| \le \sum_{n=1}^{N}(p^n + \hat{p}^n) = 2$, we can obtain the θ value limit which should fall into [0,2] interval. The effect of θ value will be further analyzed in the numerical experiments.

III. SOLUTION METHODOLOGY

In this section, we introduce the solution methodology which first reformulates the original UC problem including transformation of the chance constraints and objective function.

With the proposed reformulation technique, the original complex UC problem becomes tractable, and the conservativeness can be controlled flexibly in a data-driven manner. Then, based on the problem structure, a decomposition algorithm (i.e., C&CG algorithm) is introduced to solve the problem.

A. PROBLEM REFORMULATION

To solve the UC problem, we need to focus on two critical points in the problem formulation. One is the data-driven chance constraint for power imbalance, and the other is the two-level objective function. Since the distribution of wind power is ambiguous and defined within a distance-based ambiguity set, worst-case distribution should be considered to ensure the chance constraint. In other words, the chance constraint in (11) can be recast as follows:

$$\min_{P\in\mathcal{D}_1} Pr(-\delta \le \sum_i x_{it} + \sum_b w_t^b(\xi) - \sum_b d_t^b \le \delta) \ge 1-\epsilon, \quad \forall t \in \mathcal{T}. \tag{25}$$

Based on the studied discrete distribution, the above inequality can be further reformulated as [19]

$$\min_{P\in\mathcal{D}_1} \sum_{n=1}^{N} p_t^n \cdot \mathbb{I}_{[-\delta,\delta]}(\sum_i x_{it} + \sum_b w_t^b(\xi^n) - \sum_b d_t^b) \ge 1-\epsilon, \ \forall t \in \mathcal{T} \tag{26}$$

where $\mathbb{I}_{[-\delta,\delta]}(\sum_i x_{it} + \sum_b w_t^b(\xi^n) - \sum_b d_t^b)$ represents an indicator function. It equals 1 if $-\delta \le \sum_i x_{it} + \sum_b w_t^b(\xi^n) - \sum_b d_t^b \le \delta$, otherwise it is 0. Note that similar idea can also be used to deal with robust chance constraint with other distance based ambiguity set such as the KL divergence based set and CVaR method may be investigated to derive a conservative constraint [33].

To simplify the notation, we can introduce binary variables z_t^n to replace the indicator function, i.e., $z_t^n = \mathbb{I}_{[-\delta,\delta]}(.)$. Then, the above constraint (26) can be reformulated using big-M method as follows:

$$-\delta - (1-z_t^n)M \le \sum_i x_{it} + \sum_b w_t^b(\xi) - \sum_b d_t^b \le \delta + (1-z_t^n)M, \quad \forall t \in \mathcal{T}, \ \forall n \tag{27}$$

$$\min_{p_t^n} \sum_{n=1}^{N} p_t^n z_t^n \ge 1-\epsilon \tag{28}$$

$$\sum_{n=1}^{N} |p_t^n - \hat{p}_t^n| \le \theta \tag{29}$$

$$\sum_{n=1}^{N} p_t^n = 1, \quad \forall t \in \mathcal{T}, \ \forall n \tag{30}$$

$$p_t^n \ge 0, \quad \forall t \in \mathcal{T}, \ \forall n \tag{31}$$

where constraints (29)-(31) are derived from the predefined ambiguity set $\mathcal{D}_1$. For constraint (29), we can introduce an auxiliary variable q_t^n to eliminate the absolute operation, i.e., let $q_t^n = |p_t^n - \hat{p}_t^n|$, which leads to the following equivalent formulas [18]:

$$\sum_{n=1}^{N} q_t^n \le \theta \tag{32}$$

$$q_t^n \ge p_t^n - \hat{p}_t^n \tag{33}$$

$$q_t^n \ge \hat{p}_t^n - p_t^n. \tag{34}$$

Considering the minimization operation in constraint (28), we can try to transform the minimization into maximization by duality theory which helps remove the operation in the inequality. Based on the constraints (32)-(34), (30)-(31) and the minimization objective, the dual results can be deduced as follows:

$$\max_{\alpha_t,\beta_t^n,\gamma_t^n,\lambda_t} \ -\alpha_t\theta + \sum_{n=1}^{N}(-\beta_t^n\hat{p}_t^n + \gamma_t^n\hat{p}_t^n) + \lambda_t \tag{35}$$

$$-\beta_t^n + \gamma_t^n + \lambda_t \le z_t^n, \quad \forall t \in \mathcal{T}, \ \forall n \tag{36}$$

$$-\alpha_t + \beta_t^n + \gamma_t^n \le 0, \quad \forall t \in \mathcal{T}, \ \forall n \tag{37}$$

$$\alpha_t, \ \beta_t^n, \ \gamma_t^n \ge 0, \ \lambda_t \text{ unrestricted}, \quad \forall t \in \mathcal{T}, \ \forall n \tag{38}$$

where α_t, β_t^n, γ_t^n and λ_t are corresponding dual variables.

Then the constraints (28)-(31) can be replaced with the following constraints:

$$-\alpha_t\theta + \sum_{n=1}^{N}(-\beta_t^n\hat{p}_t^n + \gamma_t^n\hat{p}_t^n) + \lambda_t \ge 1-\epsilon, \quad \forall t \in \mathcal{T} \tag{39}$$

$$\text{Constraints } (36)-(38). \tag{40}$$

In addition to the data-driven chance constraint, the objective function in the proposed UC model also involves the uncertain distribution which hinders the optimization of the problem. To reformulate the objective function, we need to focus on the second-level objective, i.e., the worst-case expected penalty cost $\max_{P\in\mathcal{D}_1} E_P[Q(\boldsymbol{x},\xi)]$. Note that $\boldsymbol{x}$ is used to represent the decision vector for notation brevity. According to Theorem 1 in [32], we can get an equivalent reformulation with the ambiguity set $\mathcal{D}_1$ as follows:

$$\max_{P\in\mathcal{D}_1} E_P[Q(\boldsymbol{x},\xi)] = (1-\frac{\theta}{2})\text{CVaR}_{\theta/2}^{P_0}[Q(\boldsymbol{x},\xi)] + \frac{\theta}{2}\max_{\xi} Q(\boldsymbol{x},\xi) \tag{41}$$

where $\text{CVaR}_{\theta/2}^{P_0}[Q(\boldsymbol{x},\xi)]$ denotes the conditional value-at-risk of $Q(\boldsymbol{x},\xi)$ with respect to the nominal distribution P_0 with confidence level $\theta/2$. In addition, the CVaR is defined as below [34]:

$$\text{CVaR}_{\theta/2}^{P_0} = \min_{\phi} \phi + \frac{1}{1-\theta/2} E_{P_0}[Q(\boldsymbol{x},\xi)-\phi]^+ \tag{42}$$

where ϕ is a new free variable and $[Q(\boldsymbol{x},\xi)-\phi]^+ = \max\{Q(\boldsymbol{x},\xi)-\phi, 0\}$.

By substituting the CVaR, the worst-case expectation objective can be further written as follows:

$$\max_{P\in\mathcal{D}_1} E_P[Q(\boldsymbol{x},\xi)] = \min_{\phi}\{(1-\theta/2)\phi + E_{P_0}[Q(\boldsymbol{x},\xi)-\phi]^+\} + \frac{\theta}{2}\max_{\xi} Q(\boldsymbol{x},\xi). \quad (43)$$

In the above formulation, $E_{P_0}[.]^+$ can be obtained based on the estimated nominal distribution P_0. For $\max_{\xi} Q(\boldsymbol{x},\xi)$, it is actually a max-min function. Thus, the final reformulated objective function including the commitment cost has a min-max-min form. Then the reformulated problem can be regarded as a common two-stage robust problem which can be solved by a decomposition algorithm.

B. C&CG DECOMPOSITION ALGORITHM

As discussed above, the reformulated problem is a two-stage optimization problem that can be solved in a decomposition framework. In this study, we investigate the C&CG algorithm [35] to solve the problem which creates a master problem and subproblem. Given a unit commitment decision and base-case dispatch plan, the worst-case uncertainty is captured in the subproblem. Meanwhile, new variables and constraints are generated in the subproblem and fed back to the master problem. The algorithm iterates until all uncertainties can be guarded against.

Combining (1) and (43), we can acquire the following objective function:

$$\min \sum_t\sum_i[SU_iu_{it}+SD_iv_{it}+F_i(x_{it})]+(1-\theta/2)\phi + E_{P_0}[Q(\boldsymbol{x},\xi)-\phi]^+ + \frac{\theta}{2}\max_{\xi}\min_{r_{it}^u,r_{it}^d,d_t^{ls,b}}\sum_t\Big[\pi_t^{gen}\sum_i(r_{it}^u+r_{it}^d)+\pi_t^{ls}\sum_b d_t^{ls,b}\Big]. \quad (44)$$

Based on this objective, we can decompose the reformulated problem with the related constraints, and the master problem (MP) is as follows:

$$\begin{aligned}\min \ & \sum_t\sum_i[SU_iu_{it}+SD_iv_{it}+F_i(x_{it})]+(1-\theta/2)\phi + E_{P_0}[Q(\boldsymbol{x},\xi)-\phi]^+ + \eta\\ s.t.\ & \text{Constraints } (2)-(9),\ (21) \text{ and } (27),\\ & \text{Constraints } (10), (13)-(20),\ \forall n,\\ & \text{Constraints } (36)-(39),\ \text{Optimality cuts,}\end{aligned}$$

where η represents the optimal value of subproblem and the optimality cuts are derived from subproblem. Note that the CVaR or $E_{P_0}[\cdot]^+$ term in the objective function can be solved as a scenario-based stochastic optimization problem [28]. Moreover, all the constraints involved in MP are linear with continuous or integer variables. Therefore, the MP is a mixed integer linear programming (MILP) problem that can be solved by off-the-shelf solvers.

To formulate the subproblem (SP), we first dualize the second-stage problem with constraints (13)-(20) to eliminate the inner minimization in (44). For the second stage problem, we need to find the worst-case scenario in the finite support set of uncertain parameter ξ. Actually, the finite support set in this work can be extended to be the common interval uncertainty set, and the worst-case scenario is usually the extreme point of this convex set. Thus, we can define the following adjustable uncertainty set for the second stage problem:

$$\mathcal{U} = \Big\{w_t(\xi) = \bar{w}_t^b + \hat{w}_t^{b+}\rho_{bt}^+ - \hat{w}_t^{b-}\rho_{bt}^-,\ \rho_{bt}^+ + \rho_{bt}^- \le 1,\ \Big[\sum_b\sum_t(\rho_{bt}^+ + \rho_{bt}^-)\Big]/\Gamma_{\max} \le \Gamma, (\rho_{bt}^+,\rho_{bt}^-)\in\{0,1\}\Big\} \quad (45)$$

where $\bar{w}_t^b$ is the forecasted mean value of wind power, $\hat{w}_t^{b+}$ and $\hat{w}_t^{b-}$ are the corresponding deviation from the upper bound and lower bound in the finite support set, ρ_{bt}^+ and ρ_{bt}^- are auxiliary binary variables. In addition, the normalized budget of uncertainty Γ is used here [1]. Note that Γ is set to 1 in the proposed DDRC UC model to find the worst case scenario of the second stage problem. Then the subproblem can be written as below:

$$\begin{aligned}f_{sp} = (\theta/2)\max \sum_t\Big\{&\mu_t^1(\sum_b d_t^b - \sum_i x_{it} - \sum_b \bar{w}_t^b)\\ &+\sum_b \hat{w}_t^{b+}\sigma_{bt}^{1+} + \sum_b \hat{w}_t^{b-}\sigma_{bt}^{1-} + \sum_i \mu_{it}^2(-\overline{x}_i y_{it} + x_{it})\\ &+\sum_i \mu_{it}^3(\underline{x}_i y_{it} - x_{it}) + \sum_i(-\mu_{it}^4 RU_i - \mu_{it}^5 RD_i)\\ &+\sum_l \mu_{tl}^6(-L_l + L_{tl}^0) + \sum_l \mu_{tl}^7(-L_l - L_{tl}^0)\\ &+\sum_l\sum_b \hat{w}_t^{b+}\sigma_{btl}^{2+} + \sum_l\sum_b \hat{w}_t^{b-}\sigma_{btl}^{2-}\Big\}\end{aligned} \quad (46)$$

$$\mu_t^1 - \mu_{it}^2 - \mu_{it}^4 + \sum_l K_l^b(-\mu_{tl}^6 + \mu_{tl}^7) \le \pi_t^{gen}, \quad \forall t,\ \forall i \quad (47)$$

$$-\mu_t^1 - \mu_{it}^3 - \mu_{it}^5 + \sum_l K_l^b(\mu_{tl}^6 - \mu_{tl}^7) \le \pi_t^{gen}, \quad \forall t,\ \forall i \quad (48)$$

$$\mu_t^1 + \sum_l K_l^b(-\mu_{tl}^6 + \mu_{tl}^7) \le \pi_t^{ls}, \quad \forall t,\ \forall i \quad (49)$$

$$\sigma_{bt}^{1+} \le M\rho_{bt}^+,\ \sigma_{bt}^{1+} \le -\mu_t^1 + M(1-\rho_{bt}^+) \quad (50)$$

$$\sigma_{bt}^{1-} \le M\rho_{bt}^-,\ \sigma_{bt}^{1-} \le \mu_t^1 + M(1-\rho_{bt}^-) \quad (51)$$

$$\sigma_{btl}^{2+} \le M\rho_{bt}^+,\ \sigma_{btl}^{2+} \le (\mu_{tl}^6 - \mu_{tl}^7)K_l^b + M(1-\rho_{bt}^+) \quad (52)$$

$$\sigma_{btl}^{2-} \le M\rho_{bt}^-,\ \sigma_{btl}^{2-} \le (-\mu_{tl}^6 + \mu_{tl}^7)K_l^b + M(1-\rho_{bt}^-) \quad (53)$$

$$\rho_{bt}^+ + \rho_{bt}^- \le 1,\ (\rho_{bt}^+,\rho_{bt}^-)\in\{0,1\}, \quad \forall t,\ \forall b \quad (54)$$

$$\sum_b\sum_t(\rho_{bt}^+ + \rho_{bt}^-)/\Gamma_{\max} \le \Gamma \quad (55)$$

$$(\mu_{it}^2, \mu_{it}^3, \mu_{it}^4, \mu_{it}^5, \mu_{tl}^6, \mu_{tl}^7) \ge 0 \quad (56)$$

$$L_{tl}^0 = \sum_b K_l^b(\sum_i x_{it}^b + \bar{w}_t^b - d_t^b) \quad (57)$$

where (46) is the dual objective function and (47)-(57) are corresponding constraints.

In the objective function, there are some bilinear terms due to the introduction of dual variables which are given below:

$$-\sum_{b} \hat{w}_t^{b+} \mu_t^1 \rho_{bt}^+ + \sum_{b} \hat{w}_t^{b-} \mu_t^1 \rho_{bt}^- \tag{58}$$

$$\sum_{l}\sum_{b} \hat{w}_t^{b+} \rho_{bt}^+ (\mu_{tl}^6 - \mu_{tl}^7) K_l^b - \sum_{l}\sum_{b} \hat{w}_t^{b-} \rho_{bt}^- (\mu_{tl}^6 - \mu_{tl}^7) K_l^b. \tag{59}$$

These bilinear terms have been linearized using big-M technique which introduces the auxiliary variables σ_{bt}^{1+}, σ_{bt}^{1-} and σ_{btl}^{2+}, σ_{btl}^{2-} along with corresponding constraints (50)-(53). Then we can solve the MP and SP iteratively until the algorithm terminates. Specifically, if $f_{sp} \leq \eta$, the algorithm terminates and the optimal decisions are obtained, otherwise, the following cut is fed back to the MP for next iteration:

$$\eta \geq (\theta/2) \sum_{t} \left[\pi_t^{gen} \sum_{i} (r_{it}^u + r_{it}^d) + \pi_t^{ls} \sum_{b} d_t^{ls,b} \right]. \tag{60}$$

More details about the algorithm can be found in related references [19], [23], [35].

IV. NUMERICAL RESULTS

To validate the performance of the proposed approach, case studies based on IEEE test systems and a practical system are conducted in this section. First, the IEEE 6-bus test system with a wind farm is studied to illustrate the effectiveness of the proposed model and approach. Then a modified IEEE 118-bus test system is investigated for the scalability and potential practical application of the proposed approach. In addition, we also study a practical 319-bus system with a large number of uncertainty sources to further verify the scalability of the proposed approach. Related simulation experiments are implemented in MATLAB environment with YALMIP toolbox [36] and GUROBI solver on a desktop which has an an Intel Core i7-6700 CPU 3.40 GHz and 8 GB of RAM.

A. IEEE 6-BUS TEST SYSTEM

The studied IEEE 6-bus test system is illustrated in Fig. 1. Detailed parameters about this system can be found in [37]. In addition, a wind farm is connected to the system at bus 4 which is the source of uncertainty. For wind power data, we can generate the historical data from a certain assumed distribution such as the multivariate normal distribution [18]. In this work, the forecasted mean wind power is acquired from the IESO website [38] which is properly scaled and shown in Fig. 2, and the variance is 0.3 of the mean. Then the historical data can be generated by Monte Carlo simulation. Note that the historical data can be directly collected in practice. In addition, the number of bins is set to be 5, $K_0 = 5$ for the piecewise linear fuel cost function, the penalty re-dispatch and load shedding cost coefficient π_t^{gen} and π_t^{ls} are 50 \$/MW and 100 \$/MW, respectively, and the time horizon is 24 hours.

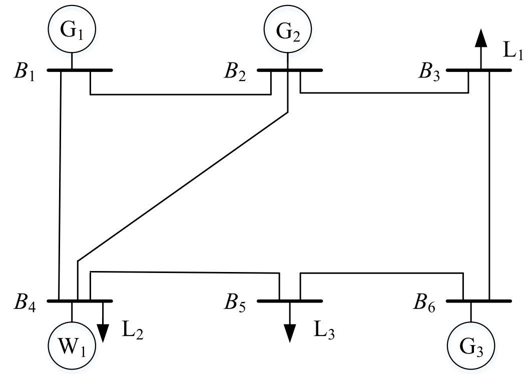

FIGURE 1. Structure of IEEE 6-bus test system.

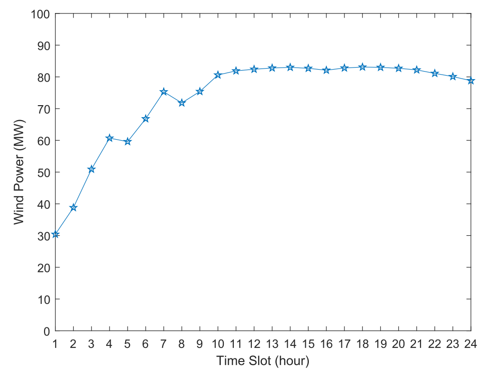

FIGURE 2. Mean wind power output.

1) INFLUENCE OF HISTORICAL DATA SIZE

For the proposed data-driven UC problem, we first investigate the influence of historical data size. With different data size, different nominal distributions and thus the ambiguity sets can be generated. In this case, we set the confidence level in the ambiguity set to be $1 - \tilde{\alpha} = 95\%$ and let the data size A vary from 50 to 5000. The corresponding results including total system cost and tolerance level θ are summarized in Table 1. From this table, we can see that θ value decreases with the increment of historical data which means that the ambiguity set $\mathcal{D}_1$ shrinks and the problem becomes less conservative. Thus, the total system cost decreases with the rise of data size as shown in the table. This result verifies the importance of data for reducing the uncertainty, i.e., more information can be attained with more data available. Actually, θ will become 0 and the set $\mathcal{D}_1$ degrades to a singleton (i.e., the nominal distribution) as the data size goes to infinity. In addition, the average time for the simulation experiment is about 13.27s.

2) INFLUENCE OF CONFIDENCE LEVEL IN AMBIGUITY SET

To construct the ambiguity set in a data-driven manner, we also need to pay attention to the effects of confidence level $(1 - \tilde{\alpha})$ in addition to the data size. In this section, we fix the data size to be 1000 and study the variation of total system cost with different confidence level ranging from 0.6 to 0.95.

TABLE 1. Influence of data size on system cost.

Data size (A)	Total cost ($\times 10^4$\$)	θ
50	4.9131	0.4356
100	4.8527	0.3080
500	4.7475	0.1378
1000	4.7231	0.0974
2000	4.7031	0.0689
5000	4.6850	0.0436

TABLE 2. Influence of confidence level on system cost.

Confidence level	Total cost ($\times 10^4$\$)	θ
0.6	4.6943	0.0636
0.7	4.7014	0.0698
0.8	4.7080	0.0774
0.9	4.7127	0.0882
0.95	4.7231	0.0974

Similarly, the corresponding system cost and the θ values are reported in Table 2. As can be seen from this table, θ becomes larger as the confidence level increases which also results in a larger ambiguity set. A larger ambiguity set can cover the true distribution with a higher chance, and this reliability is achieved by increasing the total cost, i.e., the problem becomes more conservative.

3) COMPARISON WITH OTHER METHODS

One of the significant advantages of the proposed data-driven solution methodology is that the conservativeness of the problem can be adjusted depending on the amount of available data. As discussed above, the data size has an effect on the θ value which determines the ambiguity set and conservativeness. With the proposed problem reformulation method, we can easily compare the proposed distributionally robust UC problem with the common stochatic UC (SUC) problem and Robust UC (RUC) problem by adjusting the θ value. From (41), we can derive that the worst-case expected cost becomes the worst-case cost when θ is set to be 2 or CVaR when θ is 0 which converges to the expectation [32]. In other words, the corresponding two-stage SUC or RUC problem can be obtained by setting θ to be 0 or 2, and they are used for comparison purpose here. In addition, adjustable robust optimization is also a popular method in solving UC problems with uncertainties [1], [7]. Therefore, we also consider the adjustable robust UC (ARUC) problem as a benchmark model here. Note that the ARUC model in this work is derived from adjusting the normalized budget of uncertainty Γ in (45), i.e., Γ is set to 0.2, 0.5 and 0.8, respectively. Thus, the ARUC here is actually a data-driven model considering the distributional uncertainty. Taking $A = 1000$ as an example, we can compare the proposed distributionally robust UC (DRUC) problem with SUC, RUC and ARUC problem, and the result is shown in Fig. 3. Note that the results of SUC can be regarded as optimal with known probability

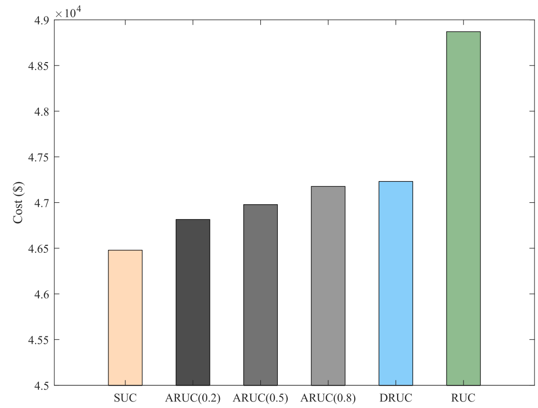

FIGURE 3. Comparison with other methods for 6-bus system.

distribution of wind power. From this figure, it can be seen that the conservativeness of the proposed UC problem is between those of SUC and RUC, which also validates the flexibility of the data-driven method in controlling the conservativeness. In addition, the ARUC problem, with a lower cost, is less conservative than the DRUC problem since it is derived based on DRUC model, and the conservatism of ARUC model decreases as the parameter Γ becomes smaller.

To better show the benefits of the proposed approach, we implemented an out-of-sample assessment of the commitment and base-case dispatch decisions obtained from different models mentioned above. Out-of-sample assessment is widely used to compare the performance of the solution in related references [13], [17]. More specifically, we first solved the first-stage UC problem and obtained the corresponding decisions. Then, we fixed the first-stage decisions, and the second-stage re-dispatch problem was solved with 300 randomly generated scenarios from the ambiguity set [28]. By calculating the average second-stage cost, we can compare the out-of-sample performance of different models. The boxplot result of the second-stage simulated cost is given in Fig. 4. As shown in this figure, the proposed DRUC model has the lowest average (median) second-stage cost which represents better performance. For the RUC model, the higher cost is caused by the great down re-dispatch and there is almost no load shedding for this case. Note that the second-stage cost is related to the cost coefficients, and the difference may become more significant by setting a larger penalty cost coefficients.

B. IEEE 118-BUS TEST SYSTEM

In this section, a case study with the modified 118-bus test system [39] is conducted to verify the scalability and potential application of the proposed approach for large systems. In this case, three wind farms are connected to the system at buses 10, 30 and 50. With the same parameter settings and analysis method, we can obtain the corresponding simulation results as given in Table 3, 4 and Fig. 5, respectively. From these results, similar conclusions can be attained which are omitted here. Additionally, the out-of-sample assessment for

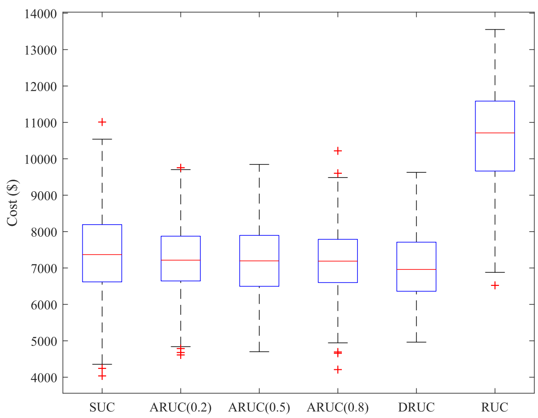

FIGURE 4. Out-of-sample assessment result for 6-bus system.

TABLE 3. Influence of data size for 118-bus system.

Data size (A)	Total cost ($)	θ
50	1150931.7	0.4356
100	1148830.2	0.3080
500	1146150.4	0.1378
1000	1145678.2	0.0974
2000	1145239.2	0.0689
5000	1144773.4	0.0436

TABLE 4. Influence of confidence level for 118-bus system.

Confidence level	Total cost ($)	θ
0.6	1145084.8	0.0636
0.7	1145179.4	0.0698
0.8	1145233.9	0.0774
0.9	1145449.2	0.0882
0.95	1145678.2	0.0974

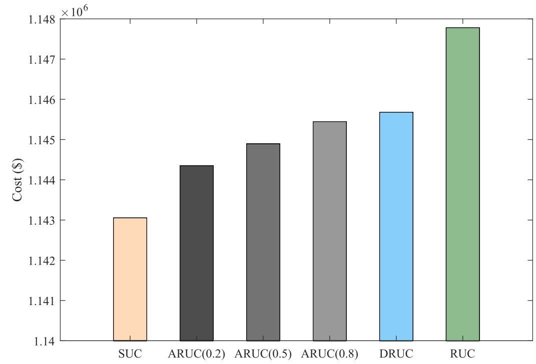

FIGURE 5. Comparison with other methods for 118-bus system.

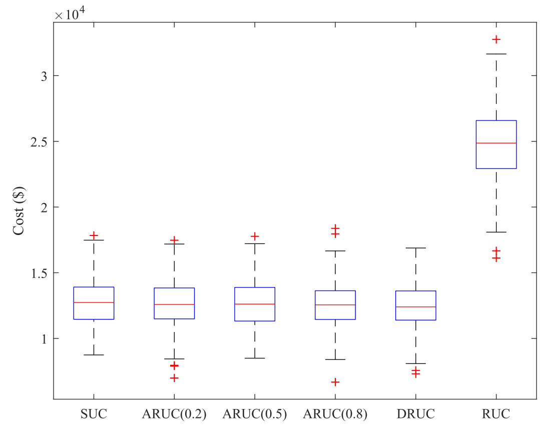

FIGURE 6. Out-of-sample assessment result for 118-bus system.

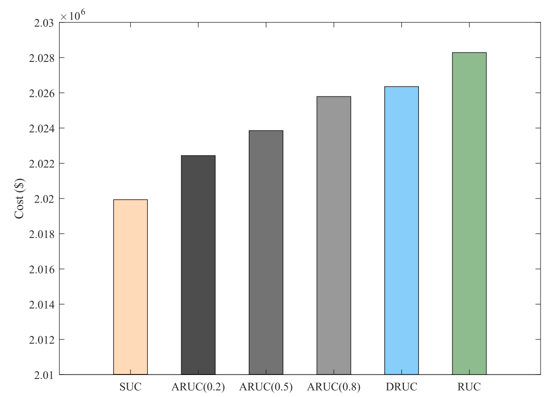

FIGURE 7. Comparison with other methods for 319-bus system.

this system is also carried out, and the boxplot result is shown in Fig. 6. From this figure, we can also see the benefit of the proposed approach by comparing the average cost. This proves the effectiveness of the proposed approach for large-scale systems. In addition, the average time for the experiment with this system is about 1147.63s which is also acceptable in practice.

C. PRACTICAL-SCALE POWER SYSTEM

To further evaluate the performance of the proposed approach in a practical-scale power system, a real-world provincial 319-bus system located in Northeast China is studied in this section [40]. There are 65 units and 431 branches in this system, and the detailed data can be found in [16], [40]. In order to model a large number of uncertainty sources, 10 wind farms are considered here which are assumed to share the same support set for wind power data. For simplicity, we only conducted the performance comparison of different models with this practical system, and the influence of the data size and confidence level in the ambiguity set is omitted. The corresponding experiment results are shown in Fig. 7 and Fig. 8. As can be seen from these results, the cost comparison has a similar trend with those in previous case studies. Therefore, we can acquire similar conclusions for this practical-scale power system which shows the scalability of the proposed approach. In addition, due to the scale increase of this practical system, the resulting model is also a large-scale complex model with many variables and constraints, and the average time for this experiment is about 5754.52s. Actually, this simulation time is used to obtain the day-ahead commitment and base-case dispatch plan, and for the

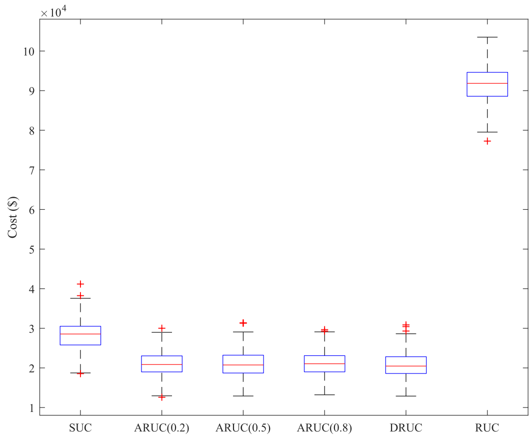

FIGURE 8. Out-of-sample assessment result for 319-bus system.

second-stage problem or real-time re-dispatch, the average simulation time is only about 3.77s, which is quite fast in practice. Consequently, the proposed approach is applicable for large-scale practical system. In addition, it is also worth studying the proposed method with larger test systems for future research where faster algorithms may be investigated to deal with the long computation time.

V. CONCLUSION

In this work, a two-stage data-driven distributionally robust chance-constrained unit commitment problem is studied which determines the commitment decision and basic dispatch plan in the first stage, and considers the worst-case expected power imbalance or re-dispatch cost in the second stage. The chance constraint is used to restrain the possible energy imbalance. Different from the moment-based ambiguity set, we constructed a distance-based ambiguity set to capture the uncertainty of wind power distribution, and this set can be derived in a data-driven environment. Numerical results show that the system cost decreases with more available historical data, for example, the cost decreases from \$4.9131 $\times$ 10^4 with data size 50 to \$4.6850 $\times$ 10^4 with data size 5000 for IEEE 6-bus test system, and that the conservativeness of the problem can be controlled by tuning the data size and confidence level in the ambiguity set. In addition, the effectiveness and flexibility of the proposed data-driven approach is also verified by the comparison with SUC, ARUC and RUC problems, for example, the total cost of the proposed DRUC problem with data size 1000 is \$4.7231 $\times$ 10^4 for IEEE 6-bus test system which is between the cost of SUC problem (\$4.6478 $\times$ 10^4) and RUC problem (\$4.8870 $\times$ 10^4).

Although two-stage UC models are the most common structure in existing literature, the extended models, i.e., multi-stage UC models have also begun to attract public attention recently. For example, the multi-stage robust UC problem has been investigated in some works [1], [9] which takes the non-anticipativity of uncertainty into account and applies the affine policies to solve the problem. Consequently, a potential research topic based on this work is the multi-stage UC problem considering distributional uncertainty. The corresponding solution methodology for multi-stage distributionally robust UC model is also worth studying in the future.

REFERENCES

[1] M. I. Alizadeh, M. P. Moghaddam, and N. Amjady, "Multistage multiresolution robust unit commitment with nondeterministic flexible ramp considering load and wind variabilities," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 872–883, Apr. 2018.

[2] Q. P. Zheng, J. Wang, and A. L. Liu, "Stochastic optimization for unit commitment—A review," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1913–1924, Jul. 2015.

[3] S. Takriti, J. R. Birge, and E. Long, "A stochastic model for the unit commitment problem," *IEEE Trans. Power Syst.*, vol. 11, no. 3, pp. 1497–1508, Aug. 1996.

[4] J. Wang, M. Shahidehpour, and Z. Li, "Security-constrained unit commitment with volatile wind power generation," *IEEE Trans. Power Syst.*, vol. 23, no. 3, pp. 1319–1327, Aug. 2008.

[5] L. Wu, M. Shahidehpour, and T. Li, "Stochastic security-constrained unit commitment," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 800–811, May 2007.

[6] R. Jiang, J. Wang, and Y. Guan, "Robust unit commitment with wind power and pumped storage hydro," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 800–810, May 2012.

[7] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 52–63, Feb. 2013.

[8] C. Shao, X. Wang, M. Shahidehpour, X. Wang, and B. Wang, "Security-constrained unit commitment with flexible uncertainty set for variable wind power," *IEEE Trans. Sustain. Energy*, vol. 8, no. 3, pp. 1237–1246, Jul. 2017.

[9] A. Lorca and X. A. Sun, "Multistage robust unit commitment with dynamic uncertainty sets and energy storage," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 1678–1688, May 2017.

[10] J. Goh and M. Sim, "Distributionally robust optimization and its tractable approximations," *Oper. Res.*, vol. 58, no. 4, pp. 902–917, Aug. 2010.

[11] 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.

[12] P. Xiong, P. Jirutitijaroen, and C. Singh, "A distributionally robust optimization model for unit commitment considering uncertain wind power generation," *IEEE Trans. Power Syst.*, vol. 32, no. 1, pp. 39–49, Jan. 2017.

[13] C. Zhao and R. Jiang, "Distributionally robust contingency-constrained unit commitment," *IEEE Trans. Power Syst.*, vol. 33, no. 1, pp. 94–102, Jan. 2018.

[14] W. Wei, F. Liu, and S. Mei, "Distributionally robust co-optimization of energy and reserve dispatch," *IEEE Trans. Sustain. Energy*, vol. 7, no. 1, pp. 289–300, Jan. 2016.

[15] 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.

[16] Y. Chen, Q. Guo, H. Sun, Z. Li, W. Wu, and Z. Li, "A distributionally robust optimization model for unit commitment based on Kullback–Leibler divergence," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 5147–5160, Sep. 2018.

[17] C. Duan, L. Jiang, W. Fang, and J. Liu, "Data-driven affinely adjustable distributionally robust unit commitment," *IEEE Trans. Power Syst.*, vol. 33, no. 2, pp. 1385–1398, Mar. 2018.

[18] C. Zhao and Y. Guan, "Data-driven stochastic unit commitment for integrating wind generation," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2587–2596, Jul. 2016.

[19] A. Bagheri, C. Zhao, and Y. Guo, "Data-driven chance-constrained stochastic unit commitment under wind power uncertainty," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Chicago, IL, USA, Jul. 2017, pp. 1–5.

[20] T. Ding, Q. Yang, X. Liu, C. Huang, Y. Yang, X. Liu, C. Huang, Y. Yang, M. Wang, and F. Blaabjerg, "Duality-free decomposition based data-driven stochastic security-constrained unit commitment," *IEEE Trans. Sustain. Energy*, vol. 10, no. 1, pp. 82–93, Jan. 2019.

[21] R. Zhu, H. Wei, and X. Bai, "Wasserstein metric based distributionally robust approximate framework for unit commitment," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2991–3001, Jul. 2019.
[22] W. Hou, R. Zhu, H. Wei, and H. TranHoang, "Data-driven affinely adjustable distributionally robust framework for unit commitment based on Wasserstein metric," *IET Gener. Transmiss. Distrib.*, vol. 13, no. 6, pp. 890–895, 2018.
[23] L. Yao, X. Wang, C. Duan, J. Guo, X. Wu, and Y. Zhang, "Data-driven distributionally robust reserve and energy scheduling over Wasserstein balls," *IET Gener., Transmiss. Distrib.*, vol. 12, no. 1, pp. 178–189, 2017.
[24] H. Ye and Z. Li, "Robust security-constrained unit commitment and dispatch with recourse cost requirement," *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3527–3536, Sep. 2016.
[25] Y. An and B. Zeng, "Exploring the modeling capacity of two-stage robust optimization: Variants of robust unit commitment model," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 109–122, Jan. 2015.
[26] C. Lee, C. Liu, S. Mehrotra, and M. Shahidehpour, "Modeling transmission line constraints in two-stage robust unit commitment problem," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1221–1231, May 2014.
[27] Y. Zhang, X. Han, M. Yang, B. Xu, Y. Zhao, and H. Zhai, "Adaptive robust unit commitment considering distributional uncertainty," *Int. J. Elect. Power Energy Syst.*, vol. 104, pp. 635–644, Jan. 2019.
[28] X. Xu, Z. Yan, M. Shahidehpour, Z. Li, M. Yan, and X. Kong, "Data-driven risk-averse two-stage optimal stochastic scheduling of energy and reserve with correlated wind power," *IEEE Trans. Sustain. Energy*, to be published.
[29] Y. Chen, W. Wei, F. Liu, and S. Mei, "Distributionally robust hydro-thermal-wind economic dispatch," *Appl. Energy*, vol. 173, pp. 511–519, Jul. 2016.
[30] M. Carrión and J. M. Arroyo, "A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1371–1378, Aug. 2006.
[31] A. Soroudi, *Power System Optimization Modeling in GAMS*. Cham, Switzerland: Springer, 2017.
[32] R. Jiang and Y. Guan, "Risk-averse two-stage stochastic program with distributional ambiguity," *Oper. Res.*, vol. 66, no. 5, pp. 1390–1405, 2018.
[33] Y. Cao, W. Wei, J. Wang, S. Mei, M. Shafie-Khah, and J. P. S. Catalao, "Capacity planning of energy hub in multi-carrier energy networks: A data-driven robust stochastic programming approach," *IEEE Trans. Sustain. Energy*, to be published.
[34] R. Rockafellar and S. Uryasev, "Conditional value-at-risk for general loss distributions," *J. Banking Finance*, vol. 26, no. 7, pp. 1443–1471, Jul. 2002.
[35] B. Zeng and L. Zhao, "Solving two-stage robust optimization problems using a column-and-constraint generation method," *Oper. Res. Lett.*, vol. 41, no. 5, pp. 457–461, 2013.
[36] J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Proc. IEEE Int. Symp. CACSD*, Sep. 2004, pp. 284–289.
[37] *IEEE 6-Bus Test System Data*. Accessed: Nov. 10, 2018. [Online]. Available: http://motor.ece.iit.edu/data/JEAS 6bus.doc
[38] *Independent Electricity System Operator (IESO)*. Accessed: Nov. 5, 2018. [Online]. Available: http://www.ieso.ca/en/Power-Data/Data-Directory
[39] *IEEE 118-Bus System Data*. Accessed: Nov. 10, 2018. [Online]. Available: http://motor.ece.iit.edu/data/JEAS 118bus.doc
[40] Z. Li, W. Wu, J. Wang, B. Zhang, and T. Zheng, "Transmission-constrained unit commitment considering combined electricity and district heating networks," *IEEE Trans. Sustain. Energy*, vol. 7, no. 2, pp. 480–492, Apr. 2016.

ZHICHAO SHI (S'17) received the B.Sc. degree from Nankai University, Tianjin, China, in 2013, and the M.Sc. degree from the National University of Defense Technology, Changsha, China, in 2015. He is currently pursuing the Ph.D. degree in electrical and computer engineering with the University of Alberta, Edmonton, AB, Canada. His current research interests include microgrid, computational intelligence and optimization, and power system optimization problems.

HAO LIANG (S'09–M'14) received the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, in 2013. He has been an Assistant Professor with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada, since 2014. His current research interests include smart grid, wireless communications, and wireless networking.

VENKATA DINAVAHI (S'94–M'00–SM'08) received the B.Eng. degree in electrical engineering from the National Institute of Technology (NIT), Nagpur, India, in 1993, the M.Tech. degree in electrical engineering from the Indian Institute of Technology (IIT) Kanpur, India, in 1996, and the Ph.D. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2000. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada. His research interests include real-time simulation of power systems and power electronic systems, electromagnetic transients, device-level modeling, large-scale systems, and parallel and distributed computing.