This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3076237, IEEE Transactions on Smart Grid

Fast Wasserstein-distance-based distributionally robust chance-constrained power dispatch for multi-zone HVAC systems

Ge Chen, *Student Member, IEEE,* Hongcai Zhang, *Member, IEEE,* Hongxun Hui, *Member, IEEE,* and Yonghua Song, *Fellow, IEEE*

***Abstract*—Heating, ventilation, and air-conditioning (HVAC) systems play an increasingly important role in the construction of smart cities because of their high energy consumption and available operational flexibility for power systems. To enhance energy efficiency and utilize their flexibility, strategic operation is indispensable. However, finding a desirable control policy for multi-zone HVAC systems is a challenging task because of unavoidable forecasting errors of ambient temperature and heat loads. This paper addresses this challenge by proposing a fast power dispatch model for multi-zone HVAC systems. A distributionally robust chance-constrained approach, which does not require the exact probability distributions of uncertainties, is employed to handle the uncertainties from forecasting errors. Both the uncertainty propagation among zones and accumulation over time are explicitly described based on the delicate indoor thermal model. Wasserstein distance is employed for the construction of ambiguity sets to improve the solution optimality. To overcome the computational intractability of Wasserstein-distance-based method, we first develop a time-efficient inner approximation for the objective function. A separation approach is then proposed to achieve the off-line calculation of uncertain parts in chance constraints. Numerical experiments prove that the proposed model can effectively achieve optimal power dispatch for HVAC systems with high computational efficiency.**

***Index Terms*—HVAC, multi-zone, distributionally robust optimization, chance-constrained optimization, Wasserstein distance.**

I. Introduction

WITH the ever-increasing requirements of energy saving and growing demands of indoor thermal comforts [1], heating, ventilation, and air-conditioning (HVAC) systems attract more and more attentions in recent years because of its high energy consumption in cities [2], [3]. The electricity consumed by space cooling of buildings reached 380 billion kWh in the USA, which was equivalent to around 10% of total electricity consumption in 2019 [4]. Moreover, due to the building thermal inertia (i.e. inherent ability of heating and cooling storage), part of HVAC loads can be shifted temporally with little impact on indoor thermal comforts [5]. This flexibility of HVAC loads can be leveraged to minimize the total energy cost of HVAC systems via planning the power schedule in advance [6], [7]. Thus, HVAC systems can be also regarded as an important demand response source for power systems [8]–[10]. Meanwhile, finding desirable operation strategies without violations of thermal comfort for HVAC systems becomes a crucial task for the construction of future smart energy systems.

A number of optimization models have been proposed for HVAC systems to improve their economical optimality and potentials for demand response. Reference [9] built an indoor thermal dynamic model to coordinate HVAC systems with energy storage and PV generation to reduce the total cost. Reference [10] established a thermal response model to simulate the operation of HVAC systems. The building thermal was served as energy storage to participate in demand response. Reference [11] proposed an aggregated HVAC model for demand response to reduce the energy cost. In [12], a comprehensive multi-zone HVAC model was developed to minimize the energy cost. Generally speaking, the works mentioned above can provide energy-efficient control policies for HVAC systems. However, these papers only considered deterministic parameters and ignored uncertainties from forecasting errors. In order to effectively utilize the flexibility of HVAC loads to provide demand response services or reduce energy costs, it is necessary to conduct ahead-of-time power scheduling that needs future environmental parameters, e.g., ambient temperature and heat loads (from indoor human activities). Although many efforts have been made to predict these parameters as accurately as possible [13], [14], forecasting errors still can not be fully eliminated. Ignoring these errors may deteriorate the operational performance of HVAC systems and harm the indoor thermal comforts. Thus, they shall be properly addressed during the operation of HVAC systems.

To consider the impacts of forecasting errors, many papers utilized uncertain programming techniques to find the optimal dispatch strategies for HVAC systems. In [15], a conditional value-at-risk (CVaR) based model was proposed to optimize operation of air conditioners. Stochastic programming was employed to handle the random outdoor temperatures. Reference [16] developed a chance-constrained demand response strategy for thermal appliances in HVAC systems under uncertain electricity prices and system loads. In [17], chance-constrained programming was used to balance the energy-saving and thermal discomforts. Reference [18] combined the scenario-based stochastic programming with Monte Carlo scenario generation to describe the uncertainties in ambient temperature and loads. References [19], [20] applied robust optimization to main-

This paper is funded in part by the Science and Technology Development Fund, Macau SAR (File no. SKL-IOTSC-2021-2023, and File no. 0137/2019/A3), and in part by the National Natural Science Foundation of China under Grant 52007200. (Corresponding author: *Hongcai Zhang*.)

G. Chen, H. Zhang, H. Hui and Y. Song are with the State Key Laboratory of Internet of Things for Smart City and Department of Electrical and Computer Engineering, University of Macau, Macao, 999078 China (email: hczhang@um.edu.mo).

tain the thermal comfort in a heat and electricity integrated energy system. Nevertheless, several challenges still remain. On the one hand, both stochastic and chance-constrained programmings require *a priori* knowledge of uncertainties. This information may be unavailable in many practical cases because we can only estimate it from historical data [21]. On the other hand, robust optimization, which requires that any realization in the uncertainty set should satisfy all constraints, often derives very conservative results [22].

Distributionally robust chance constrained (DRCC) optimization is proposed in [23] to overcome the challenges above. DRCC optimization supposes the underlying true distribution lies in an ambiguity set and optimizes objectives over all distributions in this ambiguity set. Unlike conventional stochastic and chance-constrained programmings, the DRCC method does not require exact probability distributions of uncertainties, and the solution is usually much less conservative compared with robust optimization [24]. Several papers have utilized DRCC optimization to handle uncertainties in HVAC systems. References [25], [26] minimized the cost of HVAC with a simplified thermal model. The moment-based DRCC optimization (**MDR**), which built the ambiguity set with moments, was employed and the chance constraints were reformulated into multiple linear constraints. In [27], HVAC systems were coordinated with intermittent PV generation and **MDR** was used to reformulate chance constraints. However, **MDR**'s ambiguity set of uncertainty based on moment information is often overly conservative [28].

In order to improve the solution optimality, the Wasserstein-distance-based DRCC method (**WDR**), which constructs the ambiguity set with Wasserstein distance, was proposed [28], [29]. **WDR** can provide excellent out-of-sample performance with controllable model conservativeness, and has been used to deal with the uncertain renewable generation in unit commitments [30]. Many reformulations have been proposed to convert the Wasserstein-distance-based chance constraints (WDR-CCs) into solvable forms. In [31], an exact mixed-integer reformulation of WDR-CCs was proposed based on a CVaR interpretation. To enhance the computational efficiency, an inner approximation was further developed to eliminate the integer variables. Reference [32] also proposed an exact mixed-integer form of WDR-CCs based on the corresponding CVaR interpretation. A linear outer approximation was further proposed to relax the quadratic terms based on McCormick envelope relaxation. Reference [33] focused on the WDR-CCs with uncertainties on the right-hand side (RHS). It leveraged a technique of quantile strengthening to significantly reduce the number of constraints. However, the above reformulations involve a significant number of additional auxiliary variables and constraints [29]. Since there are a large number of WDR-CCs in the focused optimal power dispatch problem, directly applying these reformulations may lead to computational intractability. To enhance the computational efficiency of **WDR**, reference [34] proposed a fully data-driven method to handle the uncertainties in an optimal power flow problem. Based on historical data, it constructed a hypercube to approximate the feasible set of WDR-CCs. Reference [35] adopted this hypercube-based method to describe the uncertainties in unit commitment and proposed a novel inner approximation for the objective to further improve the computational efficiency.

In general, the published papers have achieved remarkable progress on the power dispatch of HVAC systems. However, there are still two challenges. The first one is solution optimality. As aforementioned, **MDR** used in [25]–[27] can only derive overly conservative strategies. Although **WDR** can enhance the solution optimality, few papers adopt **WDR** to optimize multi-zone HVAC systems due to the computational intractability. Specifically, applying the CVaR-based reformulations proposed in [31]–[33] to the focused optimal power dispatch of multi-zone HVAC systems will introduce a significant number of auxiliary variables and additional constraints, so a huge computational resource is required, which may be unavailable in many practical cases. The time-efficient hypercube-based reformulation of **WDR** employed in [34], [35] can overcome the computational intractability of **WDR**, but is an inner approximation. Thus, the obtained strategy may not be so economical in practice. The second one is the uncertainty description. Due to building thermal inertia and heat exchange between adjacent zones, uncertainties will be propagated among zones and accumulated over time. However, most of the published papers including references [15]–[20], [25]–[27] mainly focused on the impacts of uncertainties on power consumption but paid little attention to the effects of uncertainties in indoor thermal comforts. Moreover, these papers usually assume that the indoor temperature was homogeneous so that neither the uncertainty accumulation over time nor propagation among zones were described.

To overcome the aforementioned challenges, we develop a fast DRCC power dispatch model for multi-zone HVAC systems under forecasting errors from ambient temperature and heat loads. The specific contributions are threefold:

1) We propose a detailed thermal dynamic model for multi-zone HVAC systems that explicitly describes the uncertainty accumulation and propagation. This model not only considered the heterogeneity of indoor temperatures among zones but also mathematically quantifies the impact of forecasting errors on indoor temperatures.
2) We first develop a **WDR**-based model to optimize the power schedule of multi-zone HVAC systems in the consideration of forecasting errors. Compared with **MDR**, the proposed method is less conservative with proper out-of-sample performance, especially in the case with a large amount of available historical data.
3) We develop a novel time-efficient reformulation of **WDR** for multi-zone HVAC systems to overcome its computation intractability. An inner approximation of objective function is first proposed by leveraging its specific structure to eliminate auxiliary variables and additional constraints. Then, based on Value-at-Risk (VaR), we propose a novel separation approach to separate the uncertain parts from decision variables in WDR-CCs. As a result, each individual WDR-CC can be exactly converted into a linear constraint. This method can not only perform an excellent computational efficiency but also achieve better optimality compared to the state-of-art hypercube-based method.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3076237, IEEE Transactions on Smart Grid

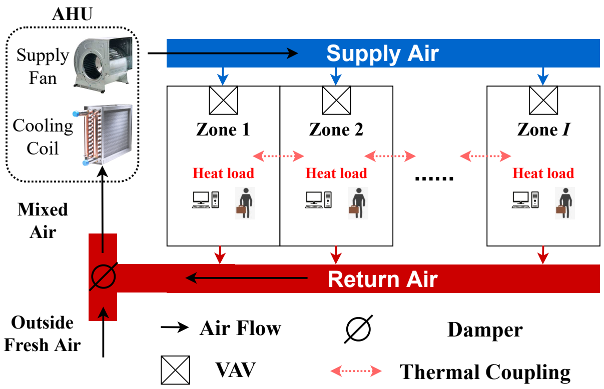


Fig. 1. Typical schematic of a multi-zone HVAC system.

Numerical experiments are conducted to validate the optimality, reliability, computational efficiency and scalability of the proposed method. Several state-of-art methods, including the CVaR-based **WDR** proposed in [31], hypercube-based **WDR** developed in [34] and **MDR** employed in [25]–[27] are introduced as our benchmarks to demonstrate the effectiveness of the proposed model.

The remaining parts are organized as follows. Section II describes the modeling methods of multi-zone HVAC systems. Section III presents the problem formulation. Section IV introduces reformulations of the DRCC model. Section V conducts simulations and Section VI concludes this paper.

II. Modeling of HVAC with Forecasting Errors

Fig. 1 illustrates a typical schematic of multi-zone HVAC systems. The indoor environments exchange heat with ambience and adjacent zones. To maintain thermal comfort, supply air is dispatched to each zone via variable air volume (VAV), absorbs heat from indoor environments, and becomes return air. The return air is mixed with the outside fresh air to adjust its CO_2 concentration. A damper is employed to control the fraction of fresh air. Driven by the supply fan, the mixed air enters the air handling unit (AHU), where it is cooled down by the cooling coil, and then supplied to different zones again.

A. Modeling of indoor environments

We use i and t to index zones ($i \in \mathcal{I} = \{1, 2, 3, ..., I\}$) and time slot ($t \in \mathcal{T} = \{1, 2, 3, ..., T\}$). Based on [12], the indoor thermal dynamics of individual zone i is described by

$$\begin{aligned} C_i \frac{d\theta^{\text{in}}_{i,t}}{dt} =& \frac{\theta^{\text{out}}_t - \theta^{\text{in}}_{i,t}}{R_i} + \sum_{j \in \mathcal{I}/\{i\}} \frac{\theta^{\text{in}}_{j,t} - \theta^{\text{in}}_{i,t}}{R_{ij}} \\ & + q^{\text{h}}_{i,t} - q^{\text{c}}_{i,t}, \quad \forall i \in \mathcal{I}, \quad \forall t \in \mathcal{T}, \end{aligned} \tag{1}$$

where C_i is the heat capacity of zone i, in kWh/°C. The indoor and outdoor temperatures are denoted by $\theta^{\text{in}}_{i,t}$ and θ^{out}_t, in °C, respectively. Parameters R_i and R_{ij} represent the thermal resistances from the i-th zone to ambience and to the j-th zone, in °C/kW, respectively. Symbol $q^{\text{h}}_{i,t}$ is the heat load from indoor heat sources (i.e. humans and electric devices); $q^{\text{c}}_{i,t}$ is the cooling supply to the i-th zone, which is defined as

$$q^{\text{c}}_{i,t} = C_p \dot{m}_{i,t} (\theta^{\text{in}}_{i,t} - \theta_s), \quad \forall i \in \mathcal{I}, \quad \forall t \in \mathcal{T}, \tag{2}$$

where C_p is the specific heat capacity of air; $\dot{m}_{i,t}$ and θ_s are the mass flow rate and temperature of cooling air in the i-th zone, respectively. Note that θ_s is assumed as constant because it is usually determined by devices.

We use finite difference to convert Eq. (1) into

$$\begin{aligned} \theta^{\text{in}}_{i,t} =& a^{\text{in}}_i \theta^{\text{in}}_{i,t-1} + a^{\text{q}}_i (q^{\text{h}}_{i,t-1} - q^{\text{c}}_{i,t-1}) + a^{\text{out}}_i \theta^{\text{out}}_{t-1} \\ & + \sum_{j \in \mathcal{I}/\{i\}} b_{ij} \theta^{\text{in}}_{j,t-1}, \quad \forall i \in \mathcal{I}, \quad \forall t \in \mathcal{T}, \end{aligned} \tag{3}$$

where $a^{\text{in}}_i = 1 - \Delta t/(R_i C_i) - \sum_{j \in \mathcal{I}/\{i\}} \Delta t/(R_{ij} C_i)$, $a^{\text{q}}_i = \Delta t / C_i$, $a^{\text{out}}_i = \Delta t/(R_i C_i)$ and $b_{ij} = b_{ji} = \Delta t/(R_{ij} C_i)$. Symbol Δt denotes the length of time slot. Then, we can write the matrix form of Eq. (3) as:

$$\boldsymbol{\theta}^{\text{in}}_t = \boldsymbol{A}^{\text{in}} \boldsymbol{\theta}^{\text{in}}_{t-1} + \boldsymbol{A}^{\text{q}} (\boldsymbol{q}^{\text{h}}_{t-1} - \boldsymbol{q}^{\text{c}}_{t-1}) + \boldsymbol{a}^{\text{out}} \theta^{\text{out}}_{t-1}, \forall t \in \mathcal{T}, \tag{4}$$

where $\boldsymbol{\theta}^{\text{in}}_t$, $\mathbf{q}^{\text{h}}_t$, $\mathbf{q}^{\text{c}}_t$ and $\mathbf{a}_3$ are the vector form of $\theta^{\text{in}}_{i,t}$, $q^{\text{h}}_{i,t}$, $q^{\text{c}}_{i,t}$ and $a_{3,i}$, respectively. Matrices $\boldsymbol{A}^{\text{in}}$ and $\boldsymbol{A}^{\text{q}}$ are defined by

$$\boldsymbol{A}^{\text{in}} = \begin{bmatrix} a^{\text{in}}_1 & b_{12} & \cdots & b_{1I} \\ b_{21} & a^{\text{in}}_2 & \cdots & b_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ b_{I1} & b_{I2} & \cdots & a^{\text{in}}_I \end{bmatrix}, \tag{5}$$

$$\boldsymbol{A}^{\text{q}} = diag([a^{\text{q}}_1, a^{\text{q}}_2, \cdots, a^{\text{q}}_I]). \tag{6}$$

B. Modeling of chiller and pump

The power consumption of chiller can be obtained by [12]:

$$p^{\text{coil}}_t = \eta^{\text{coil}} \left(\beta \cdot \mathbf{1}^\intercal \boldsymbol{q}^{\text{c}}_t + C_p (1 - \beta)(\theta^{\text{out}}_t - \theta_s) \dot{m}^{\text{tot}}_t \right), \forall t \in \mathcal{T}, \tag{7}$$

where β is the fraction of return air delivered to AHU. Symbol η^{coil} is the reciprocal of coefficient of performance of chiller. Variable $\dot{m}^{\text{tot}}_t = \mathbf{1}^\intercal \dot{\boldsymbol{m}}_t, \forall t \in \mathcal{T}$, is the total mass flow rate at time t. In a large-scale HVAC system with many zones, the airflow is usually driven by multiple parallel supply fans, so the power consumption of fans can be expressed by

$$p^{\text{fan}}_t = N^{\text{fan}} \eta^{\text{fan}} \left(\frac{\dot{m}^{\text{tot}}_t}{N^{\text{fan}}}\right)^3 = \eta^{\text{fan}} \frac{(\dot{m}^{\text{tot}}_t)^3}{(N^{\text{fan}})^2}, \quad \forall t \in \mathcal{T}, \tag{8}$$

where N^{fan} denotes the number of supply fans and η^{fan} are the reciprocal of coefficient of performance of a single fan.

C. Uncertainty propagation and accumulation

Due to the heat exchange between adjacent zones and building thermal inertia, the forecasting errors of outdoor temperature and heat loads not only propagate among zones but also accumulate over time. To describe the uncertainty propagation and accumulation in multi-zone HVAC systems, we represent the outdoor temperature and heat loads by a deterministic forecasting value adding a random fluctuation:

$$\theta^{\text{out}}_t = \widehat{\theta}^{\text{out}}_t + \widetilde{\theta}^{\text{out}}_t, \quad \forall t \in \mathcal{T}, \tag{9}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3076237, IEEE Transactions on Smart Grid

$$\boldsymbol{q}_t^{\mathrm{h}} = \widehat{\boldsymbol{q}}_t^{\mathrm{h}} + \widetilde{\boldsymbol{q}}_t^{\mathrm{h}}, \quad \forall t \in \mathcal{T}. \tag{10}$$

The indoor temperature $\boldsymbol{\theta}_t^{\mathrm{in}}$ is thereafter also uncertain. We define two new variables $\widehat{\boldsymbol{\theta}}_t^{\mathrm{in}}$ and $\widetilde{\boldsymbol{\theta}}_t^{\mathrm{in}}$ to represent the deterministic and random parts of indoor temperature, as follows:

$$\widehat{\boldsymbol{\theta}}_t^{\mathrm{in}} = \boldsymbol{A}^{\mathrm{in}}\widehat{\boldsymbol{\theta}}_{t-1}^{\mathrm{in}} + \boldsymbol{A}^{\mathrm{q}}(\widehat{\boldsymbol{q}}_{t-1}^{\mathrm{h}} - \boldsymbol{q}_{t-1}^{\mathrm{c}}) + \boldsymbol{a}^{\mathrm{out}}\widehat{\theta}_{t-1}^{\mathrm{out}}, \forall t \in \mathcal{T}, \tag{11}$$

$$\widetilde{\boldsymbol{\theta}}_t^{\mathrm{in}} = \boldsymbol{\theta}_t^{\mathrm{in}} - \widehat{\boldsymbol{\theta}}_t^{\mathrm{in}}, \quad \forall t \in \mathcal{T}. \tag{12}$$

Based on Eqs. (4) and (11)-(12), the random part $\widetilde{\boldsymbol{\theta}}_t^{\mathrm{in}}$ can be further expressed by:

$$\begin{aligned}
\widetilde{\boldsymbol{\theta}}_t^{\mathrm{in}} &= \boldsymbol{A}^{\mathrm{q}}\widetilde{\boldsymbol{q}}_{t-1}^{\mathrm{h}} + \widetilde{\theta}_{t-1}^{\mathrm{out}}\boldsymbol{a}^{\mathrm{out}} + \boldsymbol{A}^{\mathrm{in}}(\boldsymbol{\theta}_{t-1}^{\mathrm{in}} - \widehat{\boldsymbol{\theta}}_{t-1}^{\mathrm{in}}) \\
&= \boldsymbol{A}^{\mathrm{q}}\widetilde{\boldsymbol{q}}_{t-1}^{\mathrm{h}} + \boldsymbol{A}^{\mathrm{in}}\boldsymbol{A}^{\mathrm{q}}\widetilde{\boldsymbol{q}}_{t-2}^{\mathrm{h}} + \widetilde{\theta}_{t-1}^{\mathrm{out}}\boldsymbol{a}^{\mathrm{out}} + \widetilde{\theta}_{t-2}^{\mathrm{out}}\boldsymbol{A}^{\mathrm{in}}\boldsymbol{a}^{\mathrm{out}} \\
&\quad + (\boldsymbol{A}^{\mathrm{in}})^2(\boldsymbol{\theta}_{t-2}^{\mathrm{in}} - \widehat{\boldsymbol{\theta}}_{t-2}^{\mathrm{in}}) \\
&= \cdots \\
&= \sum_{\tau=0}^{t-1}(\boldsymbol{A}^{\mathrm{in}})^{t-1-\tau}\boldsymbol{A}^{\mathrm{q}}\widetilde{\boldsymbol{q}}_{\tau}^{\mathrm{h}} + \sum_{\tau=0}^{t-1}\widetilde{\theta}_{\tau}^{\mathrm{out}}(\boldsymbol{A}^{\mathrm{in}})^{t-1-\tau}\boldsymbol{a}^{\mathrm{out}} \\
&\quad + (\boldsymbol{A}^{\mathrm{in}})^{t-1}\widetilde{\boldsymbol{\theta}}_0^{\mathrm{in}}, \quad \forall t \in \mathcal{T}.
\end{aligned} \tag{13}$$

Note that $\widetilde{\boldsymbol{\theta}}_0^{\mathrm{in}} = 0$ because the present indoor temperature can be measured, so we can write Eq. (13) as:

$$\widetilde{\boldsymbol{\theta}}_t^{\mathrm{in}} = \sum_{\tau=0}^{t-1}(\boldsymbol{A}^{\mathrm{in}})^{t-1-\tau}\left(\boldsymbol{A}^{\mathrm{q}}\widetilde{\boldsymbol{q}}_{\tau}^{\mathrm{h}} + \widetilde{\theta}_{\tau}^{\mathrm{out}}\boldsymbol{a}^{\mathrm{out}}\right), \quad \forall t \in \mathcal{T}, \tag{14}$$

Remark 1. *The summation operator in Eq. (14) indicates that the uncertain part of indoor temperature is not only determined by the latest forecasting errors but also affected by the previous ones (i.e. uncertainties accumulate over time). Because the off-diagonal elements of $\boldsymbol{A}^{\mathrm{in}}$ are nonzero, the uncertain part of indoor temperature in zone i is also influenced by the forecasting errors in other adjacent zones (i.e. uncertainties propagate among zones).*

The cooling supply can be expressed as the summation of deterministic and uncertain parts:

$$\widehat{q}_{i,t}^{\mathrm{c}} = C_p\dot{m}_{i,t}(\widehat{\theta}_{i,t}^{\mathrm{in}} - \theta_s), \quad \forall i \in \mathcal{I}, \quad \forall t \in \mathcal{T}, \tag{15}$$

$$\widetilde{q}_{i,t}^{\mathrm{c}} = C_p\dot{m}_{i,t}\widetilde{\theta}_{i,t}^{\mathrm{in}}, \quad \forall i \in \mathcal{I}, \quad \forall t \in \mathcal{T}. \tag{16}$$

III. Problem Formulation

A. Objective function

In **WDR**, the objective is defined as the expectation of energy cost with the worst-case distribution among ambiguity set to ensure that the thermal comfort requirements can be satisfied under the uncertainties with underlying true distribution:

$$\min_{\boldsymbol{x}} \sup_{\mathbb{P}\in\mathcal{P}} \; \mathbb{E}_{\mathbb{P}}\left[\sum_{t\in\mathcal{T}} c_t^{\mathrm{ep}} \cdot \Delta t \cdot \left(p_t^{\mathrm{coil}}(\boldsymbol{x},\boldsymbol{\xi}) + p_t^{\mathrm{fan}}(\boldsymbol{x},\boldsymbol{\xi})\right)\right], \tag{17}$$

where $\boldsymbol{x} = [(\dot{\mathbf{m}}_t)^\intercal, (\theta_t^{\mathrm{in}})^\intercal, (\mathbf{q}_t^{\mathrm{c}})^\intercal, \forall t \in \mathcal{T}]$ denotes the variable vector and $\boldsymbol{\xi}$ denotes the vector of uncertain parameters. We assume that $\boldsymbol{\xi}$ follows distribution $\mathbb{P}$, which is unknown. However, based on historical data, we can safely estimate that $\mathbb{P}$ belongs to a set of distributions $\mathcal{P}$, i.e. $\mathbb{P} \in \mathcal{P}$. $\mathcal{P}$ is called the ambiguity set of uncertain distribution $\mathbb{P}$. Parameter c_t^{ep} is the electricity price. The vector $\boldsymbol{\xi}$ is constructed as:

$$\boldsymbol{\xi}^\intercal = \left[(\boldsymbol{\xi}^{\mathrm{in}})^\intercal, (\boldsymbol{\xi}^{\mathrm{out}})^\intercal\right], \boldsymbol{\xi} \in \mathbb{R}^{IT+T}, \tag{18}$$

$$(\boldsymbol{\xi}^{\mathrm{in}})^\intercal = \left[(\widetilde{\boldsymbol{\theta}}_1^{\mathrm{in}})^\intercal, (\widetilde{\boldsymbol{\theta}}_2^{\mathrm{in}})^\intercal, \cdots, (\widetilde{\boldsymbol{\theta}}_T^{\mathrm{in}})^\intercal\right], \boldsymbol{\xi}^{\mathrm{in}} \in \mathbb{R}^{IT}, \tag{19}$$

$$(\boldsymbol{\xi}^{\mathrm{out}})^\intercal = \left[\theta_1^{\mathrm{out}}, \theta_2^{\mathrm{out}}, \cdots, \theta_T^{\mathrm{out}}\right], \boldsymbol{\xi}^{\mathrm{out}} \in \mathbb{R}^{T}, \tag{20}$$

where $\widetilde{\boldsymbol{\theta}}_t^{\mathrm{in}}$ is obtained via Eq. (14) based on forecasting errors.

B. Constraints

1) Indoor thermal dynamics: Given in Eqs. (11) and (15).

2) Thermal comfort: Indoor temperatures should maintain in a comfortable region to guarantee the thermal comfort:

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left(\widehat{\theta}_{i,t}^{\mathrm{in}} + \widetilde{\theta}_{i,t}^{\mathrm{in}} \geq \underline{\theta}^{\mathrm{in}}\right) \geq 1-\epsilon, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \tag{21}$$

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left(\widehat{\theta}_{i,t}^{\mathrm{in}} + \widetilde{\theta}_{i,t}^{\mathrm{in}} \leq \overline{\theta}^{\mathrm{in}}\right) \geq 1-\epsilon, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}. \tag{22}$$

Eqs. (21)-(22) guarantee that with probability at least $1-\epsilon$, the indoor temperature stays in the comfortable region.

3) Device limitations: The mass flow rate to each zone is limited by the local VAV:

$$\underline{m}_i^{\mathrm{vav}} \leq \dot{m}_{i,t} \leq \overline{m}_i^{\mathrm{vav}}, \quad \forall i \in \mathcal{I}, \quad \forall t \in \mathcal{T}. \tag{23}$$

The total mass flow rate is restricted by the supply fan

$$\dot{m}_t^{\mathrm{tot}} \leq \overline{m}^{\mathrm{fan}}, \quad \forall t \in \mathcal{T}. \tag{24}$$

C. Ambiguity Set

Let the random vector $\boldsymbol{\xi}$ supported by $\Xi \in \mathbb{R}^{IT+T}$, the Wasserstein ambiguity set $\mathcal{P}$ is defined as

$$\mathcal{P} = \left\{\mathbb{P}\{\widehat{\boldsymbol{\xi}} \in \Xi\} = 1 : W(\mathbb{P}, \mathbb{P}_{\widehat{\boldsymbol{\xi}}}) \leq \delta\right\}, \tag{25}$$

where $\mathbb{P}_{\widehat{\boldsymbol{\xi}}}$ represents the empirical probability distribution based on historical data and $\delta > 0$ denotes the Wasserstein radius. The Wasserstein distance $W(\mathbb{P}_1, \mathbb{P}_2)$ is defined as

$$W(\mathbb{P}_1, \mathbb{P}_2) = \inf_{\mathbb{Q}}\left\{\int_{\Xi\times\Xi} \|\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2\|\mathbb{Q}(d\boldsymbol{\xi}_1, d\boldsymbol{\xi}_2)\right\}, \tag{26}$$

where $\mathbb{Q}$ is a joint distribution of $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ with marginal distributions $\mathbb{P}_1$ and $\mathbb{P}_2$. Operator $\|\cdot\|$ represents the norm and we use $L1$-norm in this paper. According to Eq. (25), the Wasserstein radius determines the Wasserstein radius determines the conservativeness of solutions. Thus, a desirable control strategy requires a proper Wasserstein radius. The detailed steps for calculating the Wasserstein radius will be given in Section V.

IV. Solution Methodology

Because of the bilinear terms in constraints (2), expectation term in objective function (17), and probabilities in WDR-CCs (21)-(22), the problem is highly non-convex and intractable. In this section, we first reformulate the objective and WDR-CCs into tractable forms and then approximate the bilinear terms with linear constraints so that the whole problem can be solved by off-the-shelf solvers efficiently.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3076237, IEEE Transactions on Smart Grid

A. Reformulation of Objective

The objective function (17) can be rewritten as

$$\min_{\boldsymbol{x}} \left\{ \widehat{f}(\boldsymbol{x}) + \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}(\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi})) \right\}, \tag{27}$$

where $\widehat{f}$ and $\widetilde{f}$ are the deterministic and random parts of objective, and they are defined as

$$\begin{aligned}\widehat{f}(\boldsymbol{x}) = \sum_{t\in\mathcal{T}} \Big(& p_t^{\text{fan}} + \eta^{\text{coil}}\beta\cdot\mathbf{1}^\intercal\widehat{\mathbf{q}}_t^{\text{c}} \\ &+\eta^{\text{coil}}C_p(1-\beta)(\widehat{\theta}_t^{\text{out}}-\theta_s)\dot{m}_t^{\text{tot}}\Big)\, c_t^{\text{ep}}\Delta t,\end{aligned} \tag{28}$$

$$\begin{aligned}\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi}) &= \sum_{t\in\mathcal{T}} c_t^{\text{ep}}\Delta t\eta^{\text{coil}}C_p\left(\beta(\widetilde{\boldsymbol{\theta}}_t^{\text{in}})^\intercal\dot{\boldsymbol{m}}_t + (1-\beta)\theta_t^{\text{out}}\dot{m}_t^{\text{tot}}\right) \\ &= \eta^{\text{coil}}\Delta t C_p(\boldsymbol{c}^{\text{ep}})^\intercal\left(\beta\boldsymbol{M}^{\text{z}}\boldsymbol{\xi}^{\text{in}} + (1-\beta)\boldsymbol{M}^{\text{tot}}\boldsymbol{\xi}^{\text{out}}\right),\end{aligned} \tag{29}$$

where $\boldsymbol{c}^{\text{ep}}$ is the vector form of electricity price. Matrix $\boldsymbol{M}^{\text{z}}$ and $\boldsymbol{M}^{\text{tot}}$ are defined as

$$\boldsymbol{M}^{\text{z}} = \begin{bmatrix} (\dot{\boldsymbol{m}}_1)^\intercal & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & (\dot{\boldsymbol{m}}_2)^\intercal & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & (\dot{\boldsymbol{m}}_T)^\intercal \end{bmatrix} \in \mathbb{R}^{T\times IT}, \tag{30}$$

$$\boldsymbol{M}^{\text{tot}} = diag\left[\dot{m}_1^{\text{tot}}, \dot{m}_2^{\text{tot}}, \cdots, \dot{m}_T^{\text{tot}}\right] \in \mathbb{R}^{T\times T}. \tag{31}$$

We use f_{WC} to denote the expectation of $\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi})$ in the worst case, and f_{WC} is equivalent to the following optimization problem by a strong dual reformulation (for detailed proof, please refer to reference [28]):

$$\begin{aligned}&f_{\text{WC}}(\boldsymbol{x}) = \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}_{\mathbb{P}}(\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi})) = \\ &\inf_{\lambda\geq 0}\left\{\lambda\delta + \frac{1}{N}\sum_{n=1}^{N}\sup_{\boldsymbol{\xi}\in\Xi}(\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi}) - \lambda\|\boldsymbol{\xi}-\boldsymbol{\xi}_n\|_1)\right\}\end{aligned} \tag{32}$$

where $\{\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_N\}$ denotes the sample set of random vector $\boldsymbol{\xi}$. We introduce epigraphical auxiliary variables $s_n, \forall n\in\mathcal{N}$ ($\mathcal{N}=\{1,2,...,N\}$) to reformulate (32) as:

$$f_{\text{WC}}(\boldsymbol{x}) = \inf_{\lambda\geq 0}\left\{\lambda\delta + \frac{1}{N}\sum_{n=1}^{N}s_n\right\}, \tag{33}$$

$$\begin{aligned}\text{s.t.:}\quad & s_n \geq \widetilde{f}(\boldsymbol{x},\underline{\boldsymbol{\xi}}) - \lambda\|\underline{\boldsymbol{\xi}}-\boldsymbol{\xi}_n\|_1, && \forall n\in\mathcal{N}, \\ & s_n \geq \widetilde{f}(\boldsymbol{x},\overline{\boldsymbol{\xi}}) - \lambda\|\overline{\boldsymbol{\xi}}-\boldsymbol{\xi}_n\|_1, && \forall n\in\mathcal{N}, \\ & s_n \geq \widetilde{f}(\boldsymbol{x},\boldsymbol{\xi}_n), && \forall n\in\mathcal{N},\end{aligned} \tag{34}$$

where the bounds $\underline{\boldsymbol{\xi}}$ and $\overline{\boldsymbol{\xi}}$ of random vector can be approximated by samples. When we regard $\boldsymbol{\xi}$ as variables, the function $\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi}) - \lambda\|\boldsymbol{\xi}-\boldsymbol{\xi}_n\|$ is affine in the two intervals $[\underline{\boldsymbol{\xi}},\boldsymbol{\xi}_n]$ and $[\boldsymbol{\xi}_n,\overline{\boldsymbol{\xi}}]$. In other words, the supremum in Eq. (32) is always obtained at the extreme points (i.e. at $\boldsymbol{\xi}_n$, $\underline{\boldsymbol{\xi}}$ or $\overline{\boldsymbol{\xi}}$). Thus, this supremum can be equivalent to constraints (34) after introducing epigraphical variables.

Eqs. (33)-(34) convert the worst-case expectation in (32) into a linear form, which can be solved by common off-the-shelf solvers. However, this conventional reformulation involves $N+1$ auxiliary variables (i.e. λ and $s_n, \forall n\in\mathcal{N}$) and $3N+1$ additional constraints (i.e. $\lambda>0$ and Eq. (34)), leading to a huge computational burden.

To improve the computational performance, we propose an inner approximation of $f_{\text{WC}}(\boldsymbol{x})$ to eliminate the auxiliary variables and additional constraints. The following proposition demonstrates the expression of this inner approximation.

Proposition 1. *Function $f_{\text{IA}}(\boldsymbol{x})$ is an inner approximation of $f_{\text{WC}}(\boldsymbol{x})$ in Eq. (33), where*

$$f_{\text{IA}}(\boldsymbol{x}) = \eta^{\text{coil}}\Delta t C_p\cdot(c_t^{\text{ep}})^\intercal\boldsymbol{m}^{\text{tot}}\cdot\delta + \widetilde{f}(\boldsymbol{x},\boldsymbol{\xi}^{\text{avg}}), \tag{35}$$

and $\boldsymbol{\xi}^{\text{avg}} = \frac{1}{N}\sum_{n=1}^{N}\boldsymbol{\xi}_n$ is the mean of the selected samples.

Proof: See Appendix A.

Remark 2. *The proposed f_{IA} in Eq. (35) has much higher computational efficiency than the original reformulation f_{WC} in Eq. (33), because no auxiliary variable or additional constraint will be involved.*

B. Reformulation of WDR-CCs

As mentioned in Section I, references [31]–[33] proposed CVaR-based reformulations to convert intractable WDR-CCs into solvable forms. These CVaR-based reformulations are suitable for tasks with a small number of WDR-CCs. However, applying these reformulations to the focused optimal power dispatch of HVAC systems may lead to computational intractability because a significant number of auxiliary variables and additional constraints will be introduced. To overcome the aforementioned computational intractability, we propose a novel separation method based on value-at-risk (VaR) to separate the uncertain parts from decision variables in WDR-CCs. The (1-ϵ)-VaR of $f(\boldsymbol{\xi})$ is defined as

$$\begin{aligned}\mathbb{P}\text{-VaR}_{1-\epsilon}(f(\boldsymbol{\xi})) &= \inf_g\left\{g\in\mathbb{R}|\mathbb{P}\left(f(\boldsymbol{\xi})\leq g\right)\geq 1-\epsilon\right\}, \\ &= \inf_g\left\{g\in\mathbb{R}|\mathbb{P}\left(g\leq f(\boldsymbol{\xi})\right)\leq\epsilon\right\}.\end{aligned} \tag{36}$$

By applying VaR on Eq. (21), we can separate the uncertain parts from decision variables

$$\begin{aligned}&\inf_{P\in\mathcal{P}}\mathbb{P}\left(\widetilde{\theta}_{i,t}^{\text{in}} + \widehat{\theta}_{i,t}^{\text{in}} \geq \underline{\theta}^{\text{in}}\right) \geq 1-\epsilon, \\ \Leftrightarrow &\sup_{P\in\mathcal{P}}\mathbb{P}\left(0\leq -\widetilde{\theta}_{i,t}^{\text{in}} - \widehat{\theta}_{i,t}^{\text{in}} + \underline{\theta}^{\text{in}}\right) \leq \epsilon, \\ \Leftrightarrow &\sup_{P\in\mathcal{P}}\mathbb{P}\text{-VaR}_{1-\epsilon}\left(-\widetilde{\theta}_{i,t}^{\text{in}} - \widehat{\theta}_{i,t}^{\text{in}} + \underline{\theta}^{\text{in}}\right) \leq 0, \\ \Leftrightarrow &\sup_{P\in\mathcal{P}}\mathbb{P}\text{-VaR}_{1-\epsilon}\left(-\widetilde{\theta}_{i,t}^{\text{in}}\right) \leq \widehat{\theta}_{i,t}^{\text{in}} - \underline{\theta}^{\text{in}}.\end{aligned} \tag{37}$$

The last equivalent transformation in Eq. (37) is due to the translation invariant of VaR. The term in the left-hand side of the last inequality in Eq. (37) represents the uncertain part and its value is not relevant to any decision variable. We use lower margin $r_{i,t}^L$ to denote the uncertain part $\sup_{P\in\mathcal{P}}\mathbb{P}\text{-VaR}_{1-\epsilon}\left(-\widetilde{\theta}_{i,t}^{\text{in}}\right)$, and the value of $r_{i,t}^L$ can be calculated by the following proposition.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3076237, IEEE Transactions on Smart Grid

Proposition 2. *The value of $r^L_{i,t}$ is equivalent to the optimal value of the following problem*

$$\begin{cases} \min\limits_{\boldsymbol{v}_1, v_2, \boldsymbol{v}_3, r^L_{i,t}} \quad r^L_{i,t} \\ \textit{s.t.:} \quad \delta v_2 + \dfrac{1}{N} \sum\limits_{n \in \mathcal{N}} v_{1,n} \le \epsilon \\ \qquad v_{1,n} \ge 1 - v_{3,n}(r^L_{i,t} - \widetilde{\theta}^{\text{in}}_{i,t,n}), \quad \forall n \in \mathcal{N}, \\ \qquad v_2 \ge v_{3,n}, \quad \forall n \in \mathcal{N}, \\ \qquad \boldsymbol{v}_1 \succeq 0, \quad \boldsymbol{v}_3 \succeq 0, \end{cases} \tag{38}$$

where $\boldsymbol{v}_1, v_2, \boldsymbol{v}_3$ are auxiliary variables. The subscript n in $\widetilde{\theta}_{i,t,n}$ represents that this uncertain part of indoor temperature is calculated by substituting the n-th sample of forecasting errors into Eq. (14).

Proof: See Appendix B.

Problem (38) is non-convex due to the bilinear term $v_{3,j} r^L_{i,t}$ in the second constraint. Nevertheless, Eq. (38) is a small-scale problem and its optimal value can be found efficiently by employing line search among $r^L_{i,t}$.

Similarly, Eq. (22) can be reformulated as

$$\sup_{P \in \mathcal{P}} \mathbb{P}\text{-VaR}_{1-\epsilon}\left(\widetilde{\theta}^{\text{in}}_{i,t}\right) \le \overline{\theta}^{\text{in}} - \widehat{\theta}^{\text{in}}_{i,t}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}. \tag{39}$$

We define $r^U_{i,t} = \sup_{P \in \mathcal{P}} \mathbb{P}\text{-VaR}_{1-\epsilon}\left(\widetilde{\theta}^{\text{in}}_{i,t}\right)$. By replacing $-\widetilde{\theta}^{\text{in}}_{i,t}$ with $\widetilde{\theta}^{\text{in}}_{i,t}$ in Eq. (38), we can easily obtain the value of $r^U_{i,t}$.

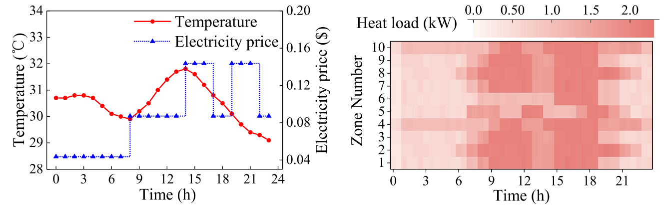
Remark 3. *The lower and upper margins $r^L_{i,t}$ and $r^U_{i,t}$ are independent of decision variables, so they can be calculated offline. As a result, the computational efficiency of problem (38) has no effect on the solving time of the power schedule optimization for HVAC systems. Once we obtain the values of $r^L_{i,t}$ and $r^U_{i,t}$ and substitute them into WDR-CCs, the uncertain parts in WDR-CCs can be treated as fixed parameters. Then, each individual WDR-CC degenerates into a linear constraint. Compared with the conventional* ***WDR*** *used in [30], [34] which reformulated each WDR-CC into a group of additional linear constraints, the proposed reformulation can achieve much better computational performance with no need for any auxiliary variable and additional constraint.*

C. Relaxation of bilinear constraints

By adopting the McCormick envelope (MCE) [36], the bilinear constraints (15) is relaxed into four linear constraints

$$\begin{cases} w_{i,t} - \underline{m}^{\text{vav}}_i(\theta^{\text{in}}_{i,t} - \theta_s) - \dot{m}_{i,t}(\underline{\theta}^{\text{in}} - \theta_s) \ge -\underline{m}^{\text{vav}}_i(\underline{\theta}^{\text{in}} - \theta_s), \\ w_{i,t} - \overline{m}^{\text{vav}}_i(\theta^{\text{in}}_{i,t} - \theta_s) - \dot{m}_{i,t}(\overline{\theta}^{\text{in}} - \theta_s) \ge -\overline{m}^{\text{vav}}_i(\overline{\theta}^{\text{in}} - \theta_s), \\ w_{i,t} - \overline{m}^{\text{vav}}_i(\theta^{\text{in}}_{i,t} - \theta_s) - \dot{m}_{i,t}(\underline{\theta}^{\text{in}} - \theta_s) \le -\overline{m}^{\text{vav}}_i(\underline{\theta}^{\text{in}} - \theta_s), \\ w_{i,t} - \underline{m}^{\text{vav}}_i(\theta^{\text{in}}_{i,t} - \theta_s) - \dot{m}_{i,t}(\overline{\theta}^{\text{in}} - \theta_s) \le -\underline{m}^{\text{vav}}_i(\overline{\theta}^{\text{in}} - \theta_s), \end{cases}$$
$$\forall i \in \mathcal{I}, \quad \forall t \in \mathcal{T}, \tag{40}$$

where $w_{i,t}$ is the auxiliary variable. Note MCE is a relaxation technique and may introduce relaxation errors. In the target system, numerical experiments show that the relaxation errors are negligible, which will be discussed in Section V.



(a) (b)

Fig. 2. One-day (a) nominal outdoor temperature, electricity price, and (b) heat loads (10 zones)

TABLE I
PARAMETERS IN CASE STUDY

Parameters	Value	Parameters	Value
$\underline{\theta}^{\text{in}}$	24°C	C_i	1.188kWh/°C
$\overline{\theta}^{\text{in}}$	28°C	R_i	7.5°C/kW
$\underline{m}^{\text{vav}}_i$	0	R_{ij}(adjacent)	22.5°C/kW
$\overline{m}^{\text{vav}}_i$	0.5 kg/s	R_{ij}(not adjacent)	0
β	0.8	Δt	0.5h
η^{coil}	0.28	η^{fan}	0.72kW·(kg/s)$^{-3}$

After all the reformulations above, the final formulation of our optimization problem is converted into a convex form:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \widehat{f}(\boldsymbol{x}) + f_{\text{IA}}(\boldsymbol{x}), \\ \text{s.t.:} \quad & \text{Eqs. (7)-(8), (11), (23)-(26), (28), (35), (37), (39)-(40).} \end{aligned}$$

Remark 4. *In summary, based on Proposition 1, we can reformulate the complex objective function into a simple linear form. By using Proposition 2, we can calculate the uncertain parts of WDR-CCs in offline and replace all sophisticated WDR-CCs with simple linear constraints. By adopting the MCE, the bilinear constraints are approximated by linear ones. As a result, the whole optimization problem becomes a simple linear programming that can be efficiently solved by off-the-shelf solvers.*

V. CASE STUDY

A. System Configuration

The case study is based on a 10-zone HVAC system. The number of fans N^{fan} and maximum total mass flow rates $\overline{m}^{\text{fan}}$ are set as 1 and 2.25kg/s, respectively. The nominal outdoor temperature, electricity price, and heat loads are shown in Fig. 2. Both of the two risk parameters ϵ and ϵ_w are set as 0.1, respectively. Other parameters are listed in Table I. A receding horizon optimization schema is employed to make the simulation results more practical. The optimization horizon is set as 12h, while the update interval (UI) is set as 1h. In other words, we update the operation strategy every hour, and the latest information of uncertainties can be easily involved in each update process.

The construction of ambiguity set requires samples of forecasting errors. Since the load/temperature predictions are not the focus of this paper, we randomly generate 10,000 normally distributed samples with zero expectation for simplicity. This manner is also widely used in many other published papers [25], [30], [37]. The standard deviations of ambient temperature and heat load forecasting errors are set as 1°C and 0.1kW during the data generation, respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3076237, IEEE Transactions on Smart Grid

All simulations are performed on an Intel(R) Core(TM) 8700 3.20GHz CPU with 16 GB memory. The corresponding optimization problem is built by CVXPY [38] and solved by GUROBI. The solving time limitation is set as 3600s.

B. Benchmarks

We implement seven models (i.e. M1-M7 shown in Table II) to validate the benefits of the proposed method. The first model M1 is the proposed one. The second model M2 is used for comparison to demonstrate the superiority of the proposed inner approximation f_{IA}.

We further implement the CVaR-based **WDR** employed in [30], [31], i.e., M3, to highlight the benefits of the proposed separation methods. M3 converts each individual WDR-CC into multiple linear constraints via CVaR approximation. For example, the generic form of WDR-CCs in our problem can be expressed as

$$\inf_{\mathbb{P}\in\mathcal{P}} \mathbb{P}\left(\boldsymbol{\alpha}^{\intercal}\boldsymbol{x} \leq \boldsymbol{\gamma}^{\intercal}\boldsymbol{\xi} + d\right) \geq 1-\epsilon. \tag{41}$$

In M3, the WDR-CC (41) is reformulated as

$$\begin{cases} g \geq 0, \boldsymbol{\psi} \succeq 0, \\ \epsilon N g - \mathbf{1}^{\intercal}\boldsymbol{\psi} \geq \delta N, \\ \dfrac{\boldsymbol{\gamma}^{\intercal}\boldsymbol{\xi}_n + d - \boldsymbol{\alpha}^{\intercal}\boldsymbol{x}}{\|\boldsymbol{\gamma}\|_*} \geq g - \psi_n, \forall n \in \mathcal{N}, \end{cases} \tag{42}$$

where $g \in \mathbb{R}$ and $\boldsymbol{\psi} \in \mathbb{R}^N$ are auxiliary variables. Employing M3 will involve $(N+1)$ variables (i.e. g and $\psi_n, \forall n \in \mathcal{N}$) and $2(N+1)$ constraints (i.e. Eq. (42)) for every WDR-CC.

The forth model M4 is the hypercube-based **WDR** proposed in [34], [35]. M4 constructs a hypercube with the smallest volume to inner approximate the feasible set of each WDR-CC. In M4, the uncertainty set can be expressed as a hypercube

$$\mathcal{Z} = \{-l \cdot \mathbf{1} \leq \boldsymbol{\zeta} \leq l \cdot \mathbf{1}\}, \tag{43}$$

where l is the side length of the constructed hypercube and $\boldsymbol{\zeta}$ is the standardized version of $\boldsymbol{\xi}$, as follows

$$\boldsymbol{\zeta} = \boldsymbol{\Sigma}^{-1/2}(\boldsymbol{\xi} - \boldsymbol{\mu}), \tag{44}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the sample mean and covariance of $\boldsymbol{\xi}$, respectively. To reduce the conservativeness of solutions, M4 minimizes l to make the uncertainty set as small as possible

$$\min_{l} \quad l, \qquad \text{s.t.:} \sup_{\mathbb{P}^s\in\mathcal{P}^s} \mathbb{P}^s(\boldsymbol{\zeta} \notin \mathcal{Z}) \leq \epsilon, \tag{45}$$

in which, symbols $\mathbb{P}^s$ and $\mathcal{P}^s$ are the true distribution of $\boldsymbol{\zeta}$ and ambiguity set, respectively. Reference [34] further proved that the term $\sup_{\mathbb{P}^s_m\in\mathcal{P}^s_m} \mathbb{P}^s_m(\boldsymbol{\zeta} \notin \mathcal{Z}_m)$ is equal to $\kappa\delta + \frac{1}{N}\sum_{n=1}^{N}\left(1-\kappa\left(l_m - \|\boldsymbol{\zeta}_n\|_\infty\right)^+\right)^+$, where $(\cdot)^+ = \max\{0,\cdot\}$. By solving Eq. (45), we can get the optimal side length l and the desirable hypercube. Then, M4 re-expresses the obtained hypercube as a convex hull of its vertices and restricts that all vertices must satisfy the constraints inside the WDR-CC. Finally, Eq. (41) is reformulated as

$$\boldsymbol{\alpha}^{\intercal}\boldsymbol{x} \leq \boldsymbol{\gamma}^{\intercal}\boldsymbol{\xi}^{(k)} + d, \quad \forall k = \{1, 2, \cdots, 2^m\}, \tag{46}$$

where $\boldsymbol{\xi}^{(k)}$ is calculated by $\boldsymbol{\xi}^{(k)} = \boldsymbol{\Sigma}^{1/2}(\boldsymbol{\zeta}^{(k)} + \boldsymbol{\mu})$, and $\boldsymbol{\zeta}^{(k)}$ is the vertex of the constructed hypercube. Parameter m is the dimension of uncertainties (i.e. $\boldsymbol{\xi} \in \mathbb{R}^m$). Since the uncertainties in Eqs. (21) and (22) are one-dimensional, two new convex constraints are generated to replace the original two WDR-CCs. Thus, M4 requires no additional constraint.

The fifth benchmark is **MDR**, in which, the uncertain parameters in its objective is replaced by their mean and the WDR-CC (41) is reformulated as

$$\boldsymbol{\alpha}^{\intercal}\boldsymbol{x} + \sqrt{\frac{1-\epsilon}{\epsilon}}\sqrt{(\boldsymbol{\gamma})^{\intercal}\boldsymbol{\Sigma}\boldsymbol{\gamma}} \leq d + (\boldsymbol{\gamma})^{\intercal}\boldsymbol{\mu}, \tag{47}$$

where $\boldsymbol{\Sigma}$ is the covariance of the random vector.

The sixth method M6 is robust optimization. It requires that the thermal comforts should be always satisfied for any realization in the support of random parameters, so its objective is expressed as

$$\min \widehat{f}(\boldsymbol{x}) + \max_{\boldsymbol{\xi}\in\Xi} \widetilde{f}(\boldsymbol{x}, \boldsymbol{\xi}). \tag{48}$$

The last benchmark, M7, is a risk neutral model that directly ignores the uncertainties.

M1-M4 are **WDR** methods and require proper Wassertein radius to control the conservativeness of results. In M1-M3, a two-fold cross validation is employed to find the best radius. The steps are summarized as follows [29]: (i) randomly partition available samples into training and test sets; (ii) solve the problem in Proposition 2 with $\delta \in [0.001, 0.002, \cdots, 0.1]$; (iii) estimate the violation probability based on the constraints in VaR (i.e., $\mathbb{P}\left(g \leq f(\xi)\right)$) in test sets; (iv) repeat (ii)-(iii) 10 times and find the 90-percentile of the violation probability; (v) output the smallest δ such that its 90-percentile violation is less than the given confidence level ϵ_w, where ϵ_w is the risk probability that the given Wassertein radius can not cover the true distribution of uncertainties. In M4, the Wasserstein radius is directly estimated based on historical data instead of using cross validation to align with references [34], [35].

C. Optimality and computational performance

Table III summarizes the results of the aforementioned seven models. The UI and size of the sample set are set as 1h and 100, respectively. Among the four **WDR** models, the proposed model M1 not only shows the highest computational efficiency but also achieves great optimality with proper reliability. Compared with M2, more than 66.6% of solving time can be saved with only 0.09% cost increasing. M3 induces more conservative results than M1 due to the inner approximations in both objective and constraints. Moreover, because M3 requires many auxiliary variables and additional constraints, the solving time of M3 reaches 1.28s, which is almost two orders of magnitude greater than that of M1. M4 performs similar computational efficiency as the proposed M1. However, the hypercube-based reformulation is an inner approximation. Moreover, according to our numerical experiments, the Wasserstein radius of M4 (directly estimated from data without cross-validation) is also considerably larger than those in M1-M3 (obtained by cross-validation). Thus, the solution of M4 is much more conservative than that of M1. The **MDR** model,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3076237, IEEE Transactions on Smart Grid

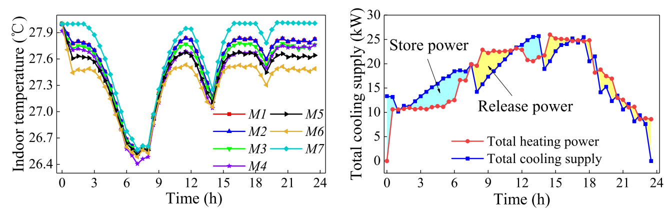
TABLE II
DESCRIPTIONS OF IMPLEMENTED MODELS

Models	Types	Reformulations of Objective Function			Reformulations of WDR-CCs			
		Formulations	Auxiliary Variables[1]	Additional Constraints[1]	Formulations	Methods	Auxiliary Variables[2]	Additional Constraints[2]
M1	WDR	$\widehat{f}(\boldsymbol{x}) + f_{\mathrm{IA}}(\boldsymbol{x})$	0	0	Eqs. (37)-(39)	VaR-based	0	0
M2	WDR	$\widehat{f}(\boldsymbol{x}) + f_{\mathrm{WC}}(\boldsymbol{x})$	N+1	3N+1	Eqs. (37)-(39)	VaR-based	0	0
M3	WDR	$\widehat{f}(\boldsymbol{x}) + f_{\mathrm{IA}}(\boldsymbol{x})$	0	0	Eq. (42)	CVaR-based	2IT(N+1)	4IT(N+1)
M4	WDR	$\widehat{f}(\boldsymbol{x}) + f_{\mathrm{IA}}(\boldsymbol{x})$	0	0	Eq. (46)	Hypercube-based	0	0
M5	MDR	$\widehat{f}(\boldsymbol{x}) + \widetilde{f}(\boldsymbol{x}, \boldsymbol{\mu})$	0	0	Eq. (47)	Moment-based	0	0
M6	RO	Eq. (48)	0	0	-	-	0	0
M7	RN	$\widehat{f}(\boldsymbol{x})$	0	0	-	-	0	0

[1] Values in columns “Auxiliary Variables” and “Additional Constraints” in “Reformulation of objective function” refer to the numbers of auxiliary variables and additional constraints introduced by reformulating the objective with the corresponding model.
[2] Values in columns “Auxiliary Variables” and “Additional Constraints” in “Reformulation of WDR-CCs” refer to the numbers of auxiliary variables and additional constraints introduced by reformulating WDR-CCs with the corresponding model.

TABLE III
RESULTS OF DIFFERENT MODELS (N=100, UI=1H)

Models	Energy Cost ($)	Solving Time (s)	Reliability (%)	Max Relaxation Error (°C)
M1	31.63	0.04	96.17	0.11
M2	31.60	0.12	96.05	0.11
M3	32.06	1.28	98.76	0.11
M4	32.92	0.04	99.53	0.11
M5	32.89	0.04	99.89	0.11
M6	35.86	0.04	100	0.11
M7	30.54	0.04	58.52	0.11



(a) (b)

Fig. 3. (a) Indoor temperatures of M1-M7 and (b) total heating power and cooling supply of M1 with N=100 and UI = 1h.

M5, only restricts the moments of uncertain parameters, so the ambiguity set includes any distribution satisfying the given moments. Thus, M5 shows worse optimality than all the four **WDR** models (i.e. M1-M4). The reliability of M6 reaches 100%, but the energy cost is much higher than those in other models. M7 can not meet the requirement of confidence level (i.e reliability $> 90\%$), although it achieves the lowest energy cost. Thus, M6 and M7 are not suitable for dispatching the cooling power in HVAC systems. We also present the results of “Max Relaxation Error” in Table III to investigate the impact of errors introduced by MCE approximation. “Max Relaxation Error” refers to the maximum absolute difference between the calculated indoor temperature based on MCE and its true value calculated by the actual thermal dynamic model in Section II. The maximum relaxation errors are always kept at around 0.11°C. Considering the temperature difference for cooling supply (i.e. $\widehat{\theta}^{\mathrm{in}}_{i,t} - \theta_s$ in Eq. (2)) is maintained at around 10°C in our simulation, such a low error will not significantly affect the energy cost and can be ignored in practice. These results confirm the superiority of the proposed model on optimality and computational efficiency.

As mentioned in Section I, indoor environments can store heat/cooling power locally with little impact on thermal comfort due to the inherent thermal inertia. Fig. 3(a) shows the indoor temperature variations of M1-M7. Two evident temperature drops occur in 03:00-07:00 and 12:00-14:00. Observing that the electricity price will go up at 08:00 and 15:00, the system wants to store some power in the hours with low electricity price for later use, so the cooling supply increases and becomes larger than the heating power (i.e. heat load plus heat transfer from ambience), as shown in Fig. 3(b). Then, during the following high electricity price period, the stored power is released to compensate for cooling demands, so there are two obvious temperature rises at 09:00-11:30 and 14:00-16:00. Note that if we control the HVAC system without considering future expectations, the indoor temperature will keep at its maximum allowable value (i.e. $\overline{\theta}^{\mathrm{in}}$) to minimize the heat transfer from ambience. As a result, the system will not arbitrage electricity price differences utilizing its thermal inertia. These results confirm the effectiveness of utilizing building thermal inertia for demand response.

As shown in Fig. 3(a), the indoor temperature of M6 is always the lowest, which demonstrates that the strategy obtained by M6 is the most conservative. On the contrary, the policy of M7 is the riskiest because the corresponding indoor temperature of M7 keeps at the highest level. Among the five DRCC models, the indoor temperature of the proposed model is almost the same as that of M2 and is slightly higher compared to M3-M5. This result further illustrates that the proposed model requires a smaller temperature margin and is less conservative than M3-M5. The optimal temperature differences at each time slot among different models are large enough to be perceived by advanced temperature sensors with a sufficient high resolution [39], [40] (e.g. the optimal temperature difference between M6 and M7 reaches 0.6°C, while the resolution of the sensor proposed in [39] is less than 0.1°C).

D. Influence factors

1) Sample size: Fig. 4 illustrates the results of five DRCC models (i.e. M1-M5) with different sizes of the sample set. Note there is no result for M3 when N>500 because “out of memory” occurs, which demonstrates the computational

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3076237, IEEE Transactions on Smart Grid

intractability of the CVaR-based **WDR** method. With the increase of sample size, both the energy costs and reliability obtained by M1-M4 decrease, but the reliability is always greater than the required confidence level. When the sample size is small, the distribution information of uncertainty is incomplete. Thus, a large Wasserstein radius is required to cover the underlying true distribution of uncertainties, leading to a very conservative solution. When more historical data collected, the estimated distribution of uncertainty will be more accurate. As a result, a small Wasserstein radius is enough to cover the true distribution and a more economical strategy can be obtained. Similar to Table III, both M3 and M4 get much more conservative results compared with the proposed model M1 no matter how large the size of sample size is. The cost of M1 is slightly higher than that of M2, but the difference decreases gradually with the increase of sample size. Increasing sample number can decrease the Wasserstein radius. Thus, the difference between the first term of two objectives, i.e., f_{WC} in Eq. (33) and f_{IA} in Eq. (35), can be reduced. The cost and reliability of M5 almost stay constant and keep at a relatively unnecessary conservative level under different sizes of the sample set. The performance of **MDR** can be usually improved with the increase of the sample number, but the results of M5 almost keeps constant in our simulation as the sample number is growing. As aforementioned, MDR utilizes the first- and second-order moments to construct ambiguity sets. Once the moment information is determined, the ambiguity set is fixed, and the conservativeness of the solution is also decided. Because the samples used to construct ambiguity set is selected randomly, the expectation $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ of the random variable $\boldsymbol{\xi}$ barely changes with the sample set size grows from 50 to 1,000, so the corresponding results almost keep constant. Compared with the conventional **WDR** models M2 and M3, the superiority of the proposed model M1 in computational efficiency becomes more significant with the increase of sample size. This is because the numbers of auxiliary variables and additional constraints required by M2 and M3 grow rapidly with the increase of the sample number. Moreover, the computational efficiency of M1 can always maintain at a very high level and is not affected by the size of the sample set. Similarly, all relaxation errors are relatively small, which indicates the effectiveness of MCE used in Section IV-C.

2) Update interval (UI): Table IV shows the results of M1 under different UIs. With the increase of UI, the energy cost grows while the reliability decreases. According to Eq. (14), the effects of forecasting errors on indoor temperatures accumulate over time. Once the control policy is updated, the uncertainty and relaxation error will be reset because the indoor temperature at this moment can be measured directly. The longer UI is, the more uncertainty will accumulate, and the higher margin $r^L_{i,t}$ is required, as shown in Fig. 5(a). To meet the larger margin $r^L_{i,t}$, a higher energy cost is required. Meanwhile, due to the building thermal inertia, the relaxation error can be accumulated over time and reach a high level eventually in the case with a large UI, as shown in Fig. 5(b). This high relaxation error harms the reliability of the control policy and makes the solution riskier.

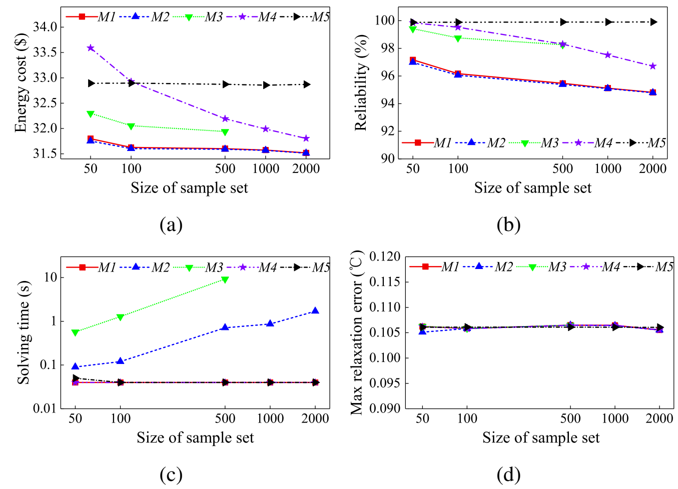


Fig. 4. Results of (a) energy costs, (b) reliability, (c) solving times, and (d) maximum relaxation errors obtained by different models. Note there is no result of M3 when N >500 due to the out-of-memory issue.

TABLE IV
RESULTS OF M1 WITH DIFFERENT UPDATE INTERVALS (N=100)

Update interval (h)	Energy Cost ($)	Solving time (s)	Reliability (%)	Max Relaxation Error (°C)
1	31.63	0.04	96.17	0.11
2	32.28	0.04	95.62	0.28
3	32.84	0.04	95.02	0.42
4	33.01	0.04	94.34	0.51
6	34.29	0.04	92.77	0.58

E. Importance of considering uncertainty propagation and accumulation

We further conduct a comparison between the following two models to demonstrate the importance of considering the uncertainty propagation and accumulation

1) M1: the proposed model;
2) M1*: M1 without considering the uncertainty propagation among zones and accumulation over time are ignored (i.e. the uncertain part of the indoor temperature is calculated by $\widetilde{\boldsymbol{\theta}}^{\mathrm{in}}_t = \boldsymbol{A}^{\mathrm{q}}\widehat{\boldsymbol{q}}^{\mathrm{h}}_{t-1} + \widetilde{\theta}^{\mathrm{out}}_{t-1}\boldsymbol{a}^{\mathrm{out}}$).

The size of the sample set and UI are set as 100 and 1h. The corresponding results are shown in Fig. 6. M1* ignores the uncertainty propagation and accumulation, so it underestimates the impact of uncertainties on indoor temperature according to Eq. (14). Although the energy cost of M1* is lower than that in M1, its reliability can not always satisfy the required confidence level (i.e. reliability≥90%). For example, the reliability of M1* is only 73.84% and much less than the given value in the case of UI=6h. With the increase of the UI, more uncertainty accumulates over time, so the aforementioned underestimation on the impacts of forecasting

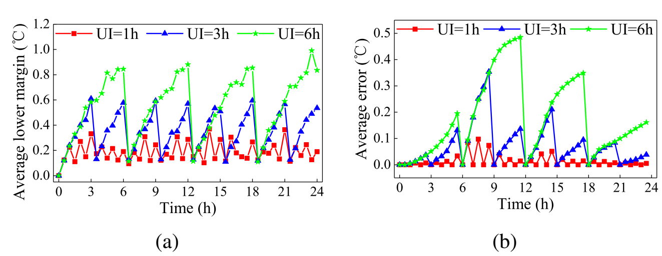


Fig. 5. Average (a) lower margin and (b) relaxation error under different UIs

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3076237, IEEE Transactions on Smart Grid

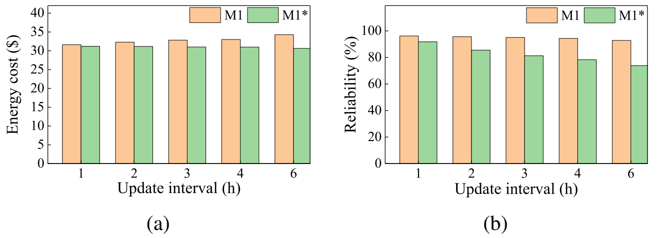


Fig. 6. (a) Energy costs and (b) reliability of M1 and M1* with different UIs (N=100).

errors becomes more significant. Thus, a larger UI indicates higher differences in the energy cost and reliability between the two models. These results demonstrate that considering the uncertainty propagation and accumulation is necessary for ensuring the reliability of indoor thermal comfort.

F. Out-of-Sample performance

In practice, the true distributions of forecasting errors are unknown. In this section, we implement three cases with different sample generators to demonstrate that the proposed model can provide desirable out-of-sample performance no matter what the distributions of uncertainties are. These cases are summarized as follows:

1) Case 1: samples are generated by uniform distributions, i.e., $\widetilde{\theta}_{i,t}^{\text{out}} \sim \text{U}(0,3)$ and $\widetilde{q}_{i,t}^{\text{h}} \sim \text{U}(0,0.3)$;
2) Case 2: samples are generated by Laplace distributions, i.e., $\widetilde{\theta}_{i,t}^{\text{out}} \sim \text{Laplace}(0,1)$ and $\widetilde{q}_{i,t}^{\text{h}} \sim \text{Laplace}(0,0.1)$;
3) Case 3: samples are generated by logistic distributions, i.e., $\widetilde{\theta}_{i,t}^{\text{out}} \sim \text{Logistic}(0,1)$ and $\widetilde{q}_{i,t}^{\text{h}} \sim \text{Logistic}(0,0.1)$.

A total of 10,000 samples are generated in each case. Then, 100 samples are randomly selected from the generated sample set to construct the ambiguity set.

Table V lists the results of the three cases. The proposed model can meet the requirement of confidence level with much lower energy cost than those in M3, M4, M5 in all cases. Moreover, M1 always performs excellent computational efficiency no matter which distribution is used to generate samples. The solving time of the proposed model is almost two orders of magnitude lower than that in M3, which indicates the great computational performance of the proposed VaR-based separation method in Eqs. (37)-(39). Compared with M2, employing the proposed model can reduce about 66.7% of solving time with a negligible cost increase, i.e. around \$0.02, which confirms the excellent time efficiency of the proposed inner approximation in Eq. (35). Although the solving times of M4 and M5 are similar to that in M1, their energy costs are much larger. These results indicate the great optimality and time-efficiency of the proposed model.

G. Scalability

We implement seven case studies with different zone numbers to demonstrate the scalability of the proposed model. The size of the sample set and UI are fixed at 2000 and 1h, respectively. The simulation results are listed in Table VI. The corresponding limitations of total mass flow rates $\overline{m}^{\text{fan}}$ are 2.25kg/s, 4.5kg/s, 11.25kg/s, 22.5kg/s, 45kg/s, 112.5kg/s and 225kg/s, respectively. Compared with M2, the proposed model M1 can save more than 98% of solving time with the negligible cost increasing (around 0.02%) no matter how large the zone number is. When M3 is employed, out of memory issue always occurs in each case due to the newly introduced auxiliary variables and additional constraints. Conversely, even in a large-scale HVAC system case (I=1000), the solution of M1 can be derived in a very short time (34.28s). Moreover, although the computational efficiency of M4 reaches the same high level as M1, the corresponding energy cost and reliability are more conservative. These results confirm the superiority of the proposed model in optimality and computational efficiency.

Although the solving times in Table VI are much lower than the UI (i.e. 1h), these results still demonstrate the necessity of developing time-efficient reformulations for **WDR**. Firstly, the memory space is limited in practice. However, even in the case with 10 zones, the conventional CVaR-based model M3 can not be conducted because of the out-of-memory issue. In the large-scale case with 1000 zones, out-of-memory issue also occurs in the calculation of M2. Conversely, the proposed model M1 can work well in the large-scale case with 1000 zones. Secondly, one individual building may contain hundreds of rooms. Every room should be controlled properly so that the total energy cost can be minimized with little indoor thermal discomfort. Unfortunately, we observe that the solving time of M2 grows exponentially with the increase of zone number according to Table VI. In the case with 500 zones, the solving time of M2 has already reached to 592s. For a large case with more zones, we can imply that the solving time of M2 may be greater than the UI even if we have sufficient memory space. Thus, time-efficient reformulation is indispensable in practice.

TABLE V
RESULTS WITH DIFFERENT SAMPLE GENERATORS (N=100, UI=1H)

Case	Model	Energy Cost (\$)	Solving Time (s)	Reliability (%)
1	M1	32.66	0.04	96.15
	M2	32.64	0.12	96.11
	M3	33.02	1.42	98.08
	M4	34.06	0.05	99.39
	M5	34.79	0.04	99.99
2	M1	32.15	0.04	96.13
	M2	32.12	0.12	96.05
	M3	32.74	1.24	98.17
	M4	33.47	0.04	98.84
	M5	33.84	0.04	99.69
3	M1	23.56	0.04	96.21
	M2	23.52	0.11	96.15
	M3	23.52	1.49	98.06
	M4	30.71	0.04	98.79
	M5	22.48	0.04	99.76

VI. CONCLUSIONS

This paper propose a fast power dispatch model for multi-zone HVAC systems considering forecasting errors of ambient temperature and heat loads. The uncertainty propagation among zones and accumulation over time are fully described to quantify the corresponding effects on indoor thermal comforts. DRCC method, which does not require *a priori* knowledge of distribution information about uncertainties, is employed to handle the forecasting errors. Wasserstein distance is utilized

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3076237, IEEE Transactions on Smart Grid

TABLE VI
PERFORMANCE OF M1-M4 WITH DIFFERENT NUMBER OF ZONES (N=2000, UI=1H)

Number of Zones (I)		10	20	50	100	200	500	1000
Number of fans (N^{fan})		1	2	5	10	20	50	100
Energy Cost ($)	M1	31.52	61.34	150.45	305.33	609.38	1518.82	3028.07
	M2	31.51	61.33	150.42	305.26	609.25	1518.42	out of memory
	M3	out of memory						
	M4	31.80	61.91	151.83	308.17	615.09	1532.70	3054.42
Solution time (s)	M1	0.04	0.11	0.21	0.65	1.51	13.37	34.28
	M2	1.70	5.29	22.59	69.82	211.93	592.24	out of memory
	M4	0.04	0.12	0.24	0.59	1.47	13.58	31.79
Reliability (%)	M1	94.81	94.82	94.87	94.84	94.76	95.09	95.38
	M2	94.79	94.79	94.84	94.81	94.74	94.99	out of memory
	M4	96.71	96.74	96.74	96.75	96.72	96.81	97.01

to construct the ambiguity set to improve the optimality of solutions. Since the conventional CVaR-based **WDR** involves a significant number of auxiliary variables and additional constraints, this paper first proposes an inner approximation for objective to eliminate all auxiliary variables or constraints. Then, each individual WDR-CC is reformulated into one linear constraint by separating the uncertain parts from decision variables. Simulation results demonstrate that more than 98% of solving time can be reduced by the proposed inner approximation with negligible cost increase in the case with a large size of sample size. Numerical experiments confirm that applying the proposed separation method can achieve much better computational efficiency, e.g., the solving time of the proposed model is at least two orders of magnitude lower compared with the conventional CVaR-based **WDR**. Simulation results also validate that the proposed model can derive a less conservative strategy with proper reliability compared to the state-of-art hypercube-based model.

APPENDIX A

Proof of Proposition 1: Note the uncertain part $\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi})$ defined in Eq. (29) can be expressed as:

$$\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi}) = \boldsymbol{h}^\intercal\boldsymbol{\xi} = (\boldsymbol{h}^{\text{in}})^\intercal\boldsymbol{\xi}^{\text{in}} + (\boldsymbol{h}^{\text{out}})^\intercal\boldsymbol{\xi}^{\text{out}}, \tag{49}$$

$$\boldsymbol{h}^{\text{in}} = \eta^{\text{coil}}\Delta t C_p \cdot \beta \boldsymbol{c}_e^\intercal \boldsymbol{M}^{\text{z}}, \tag{50}$$

$$\boldsymbol{h}^{\text{out}} = \eta^{\text{coil}}\Delta t C_p \cdot (1-\beta) \boldsymbol{c}_e^\intercal \boldsymbol{M}^{\text{tot}}. \tag{51}$$

According Eq. (51), we observe that $\boldsymbol{h} \succeq 0$ because both the coefficients and variables are nonnegative. Moreover, it is obvious that $(\underline{\boldsymbol{\xi}} - \boldsymbol{\xi}) \preceq 0$, so we have

$$\frac{\widetilde{f}(\boldsymbol{x},\underline{\boldsymbol{\xi}}) - \widetilde{f}(\boldsymbol{x},\boldsymbol{\xi})}{\|\underline{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_1} = \frac{\boldsymbol{h}^\intercal(\underline{\boldsymbol{\xi}} - \boldsymbol{\xi})}{\|\underline{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_1} \le 0, \quad \forall \boldsymbol{\xi} \in \Xi, \tag{52}$$

Because $(\overline{\boldsymbol{\xi}} - \boldsymbol{\xi}) \succeq 0$, we can also prove that

$$\frac{\widetilde{f}(\boldsymbol{x},\overline{\boldsymbol{\xi}}) - \widetilde{f}(\boldsymbol{x},\boldsymbol{\xi})}{\|\overline{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_1} = \frac{\boldsymbol{h}^\intercal(\overline{\boldsymbol{\xi}} - \boldsymbol{\xi})}{\|\overline{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_1} = \frac{\sum_{l\in\mathcal{L}} h_l(\overline{\boldsymbol{\xi}} - \boldsymbol{\xi})_l}{\sum_{l\in\mathcal{L}} (\overline{\boldsymbol{\xi}} - \boldsymbol{\xi})_l}$$
$$\le \sum_{l\in\mathcal{L}} \frac{h_l(\overline{\boldsymbol{\xi}} - \boldsymbol{\xi})_l}{(\overline{\boldsymbol{\xi}} - \boldsymbol{\xi})_l} = \boldsymbol{1}^\intercal\boldsymbol{h} = \eta^{\text{coil}}\Delta t C_p \cdot \boldsymbol{c}_e^\intercal \boldsymbol{m}^{\text{tot}}, \quad \forall \boldsymbol{\xi} \in \Xi, \tag{53}$$

where the subscript l represents the component of vectors and $\mathcal{L} = \{1, 2, \cdots, IT + T\}$. We add an additional constraint $\lambda \ge \boldsymbol{1}^\intercal\boldsymbol{h} \ge 0$, then the original constraint $\lambda \ge 0$ can be eliminated. By substituting this additional constraint into Eqs. (52)-(53), we can get

$$\lambda \ge \max\left\{\frac{\boldsymbol{h}^\intercal(\underline{\boldsymbol{\xi}} - \boldsymbol{\xi})}{\|\underline{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_1}, \frac{\boldsymbol{h}^\intercal(\overline{\boldsymbol{\xi}} - \boldsymbol{\xi})}{\|\overline{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_1}, \right\}, \quad \forall \boldsymbol{\xi} \in \Xi. \tag{54}$$

The inequality above can be further converted into

$$\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi}) \ge \widetilde{f}(\boldsymbol{x},\overline{\boldsymbol{\xi}}) - \lambda\|\overline{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_1, \quad \forall \boldsymbol{\xi} \in \Xi, \tag{55}$$

$$\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi}) \ge \widetilde{f}(\boldsymbol{x},\underline{\boldsymbol{\xi}}) - \lambda\|\underline{\boldsymbol{\xi}} - \boldsymbol{\xi}\|_1, \quad \forall \boldsymbol{\xi} \in \Xi. \tag{56}$$

Thus, the constraints Eq. (34) can be eliminated and the worst-case expectation $f_{\text{WC}}(\boldsymbol{x})$ can be written as

$$f_{\text{IA}}(\boldsymbol{x}) = \inf_{\lambda} \left\{\lambda\delta + \frac{1}{N}\sum_{n=1}^{N}\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi}_n)\right\}, \tag{57}$$

$$\text{s.t.:} \quad \lambda \ge \boldsymbol{1}^\intercal\boldsymbol{h}. \tag{58}$$

Because of the additional constraint (58), the objective f_{IA} is an inner approximation of $f_{\text{WC}}(\boldsymbol{x})$ in Eq. (32). Since the function $\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi}_j)$ is linear, we have $\frac{1}{N}\sum_{n=1}^{N}\widetilde{f}(\boldsymbol{x},\boldsymbol{\xi}_n) = \widetilde{f}(\boldsymbol{x},\boldsymbol{\xi}^{\text{avg}})$, where $\boldsymbol{\xi}^{\text{avg}}$ is the expectation of the N samples. Then, by substituting the optimal solution of the problem above (i.e. $\lambda = \boldsymbol{1}^\intercal\boldsymbol{h}$) into Eq. (57), we can express f_{IA} as the form in Proposition 1.

APPENDIX B

Proof of Proposition 2: According to Eq. (36), the value of $r_{i,t}^L$ is equivalent to the optimal value of the following problem

$$\begin{cases} \min\limits_{r_{i,t}^L} \quad r_{i,t}^L, \\ \text{s.t.:} \quad \sup\limits_{\mathbb{P}\in\mathcal{P}} \mathbb{P}(r_{i,t}^L \le -\widetilde{\theta}_{i,t}^{\text{in}}) \le \epsilon, \end{cases} \quad \forall i \in \mathcal{I}, \quad \forall t \in \mathcal{T}, \tag{59}$$

The violation probability $\mathbb{P}(r_{i,t}^L \le -\widetilde{\theta}_{i,t}^{\text{in}})$ can be converted into an expectation form

$$\mathbb{P}(r_{i,t}^L \le -\widetilde{\theta}_{i,t}^{\text{in}}) = \mathbb{E}(\mathbb{I}(\theta_{i,t}^{\text{in}} \in \Omega_{i,t})), \ i \in \mathcal{I}, \ \forall t \in \mathcal{T}, \tag{60}$$

$$\Omega_{i,t} = \{\widetilde{\theta}_{i,t}^{\text{in}} \mid \widetilde{\theta}_{i,t}^{\text{in}} \le -r_{i,t}^L\}, \ \forall i \in \mathcal{I}, \ \forall t \in \mathcal{T}, \tag{61}$$

where $\mathbb{I}(\cdot)$ denotes the indicator function (i.e. if $\widetilde{\theta}_{i,t}^{\text{in}} \in \Omega_{i,t}$, the value of $\mathbb{I}(\cdot)$ is 1; otherwise, the value is zero). We employ Corollary 5.3 in [29] to reformulate the worst-case expectation

$\sup_{\mathbb{P}\in\mathcal{P}} \mathbb{E}(\mathbb{I}(\theta_{i,t}^{\text{in}} \in \Omega_{i,t}))$ as:

$$\begin{cases} \min\limits_{\boldsymbol{v}_1,v_2,\boldsymbol{v}_3} \quad \delta \cdot v_2 + \dfrac{1}{N} \sum\limits_{n\in\mathcal{N}} v_{1,n} \\ \text{s.t.:} \quad v_{1,n} \geq 1 - v_{3,n}(r_{i,t}^L - \widetilde{\theta}_{i,t,n}^{\text{in}}), \quad \forall n \in \mathcal{N}, \\ \quad v_2 \geq v_{3,n}, \quad \forall n \in \mathcal{N}, \\ \quad \boldsymbol{v}_1 \succeq 0, \quad \boldsymbol{v}_3 \succeq 0, \end{cases} \tag{62}$$

where $\widetilde{\theta}_{i,t,n}^{\text{in}}$ can be calculated by substituting the n-th samples of the heat load and ambient temperature forecasting errors into Eq. (14). Finally, we can get the reformulation of $r_{i,t}^L$ in Proposition 2 by substituting Eq. (62) into Eq. (59).

REFERENCES

[1] E. Gonzalez-Romera, M. A. Jaramillo-Moran, and D. Carmona-Fernandez, "Monthly electric energy demand forecasting based on trend extraction," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1946–1953, 2006.

[2] Y.-J. Kim, D. H. Blum, N. Xu, L. Su, and L. K. Norford, "Technologies and magnitude of ancillary services provided by commercial buildings," *Proc. IEEE*, vol. 104, no. 4, pp. 758–779, 2016.

[3] Z. Afroz, G. Shafiullah, T. Urmee, and G. Higgins, "Modeling techniques used in building hvac control systems: A review," *Renew. Sust. Energ. Rev.*, vol. 83, pp. 64–84, 2018.

[4] B. P. Center, "Annual energy outlook 2020," 2020.

[5] A. Kathirgamanathan, M. De Rosa, E. Mangina, and D. P. Finn, "Data-driven predictive control for unlocking building energy flexibility: A review," *Renew. Sust. Energ. Rev.*, vol. 135, p. 110120, 2021.

[6] L. Yu, T. Jiang, and Y. Zou, "Online energy management for a sustainable smart home with an hvac load and random occupancy," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 1646–1659, 2019.

[7] L. Yu, D. Xie, C. Huang, T. Jiang, and Y. Zou, "Energy optimization of hvac systems in commercial buildings considering indoor air quality management," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5103–5113, 2019.

[8] L. Zhang, N. Good, and P. Mancarella, "Building-to-grid flexibility: Modelling and assessment metrics for residential demand response from heat pump aggregations," *Appl. Energy*, vol. 233, pp. 709–723, 2019.

[9] J. A. Pinzon, P. P. Vergara, L. C. Da Silva, and M. J. Rider, "Optimal management of energy consumption and comfort for smart buildings operating in a microgrid," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 3236–3247, 2018.

[10] A.-Y. Yoon, Y.-J. Kim, and S.-I. Moon, "Optimal retail pricing for demand response of hvac systems in commercial buildings considering distribution network voltages," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5492–5505, 2018.

[11] R. Adhikari, M. Pipattanasomporn, and S. Rahman, "An algorithm for optimal management of aggregated hvac power demand using smart thermostats," *Appl. Energy*, vol. 217, pp. 166–177, 2018.

[12] Y. Yang, G. Hu, and C. J. Spanos, "Hvac energy cost optimization for a multizone building via a decentralized approach," *IEEE Trans. Autom. Sci. Eng.*, 2020.

[13] R. Ramakrishna, A. Scaglione, V. Vittal, E. Dall'Anese, and A. Bernstein, "A model for joint probabilistic forecast of solar photovoltaic power and outdoor temperature," *IEEE Trans. Signal Process.*, vol. 67, no. 24, pp. 6368–6383, 2019.

[14] D. Geysen, O. De Somer, C. Johansson, J. Brage, and D. Vanhoudt, "Operational thermal load forecasting in district heating networks using machine learning and expert advice," *Energy Build.*, vol. 162, pp. 144–153, 2018.

[15] Z. Wu, S. Zhou, J. Li, and X. Zhang, "Real-time scheduling of residential appliances via conditional risk-at-value," *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1282–1291, 2014.

[16] Y. Huang, L. Wang, W. Guo, Q. Kang, and Q. Wu, "Chance constrained optimization in a home energy management system," *IEEE Trans. Smart Grid*, vol. 9, no. 1, pp. 252–260, 2016.

[17] K. Ma, G. Hu, and C. J. Spanos, "Energy management considering load operations and forecast errors with application to hvac systems," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 605–614, 2016.

[18] E. Grover-Silva, M. Heleno, S. Mashayekh, G. Cardoso, R. Girard, and G. Kariniotakis, "A stochastic optimal power flow for scheduling flexible resources in microgrids operation," *Appl. Energy*, vol. 229, pp. 201–208, 2018.

[19] S. Lu, W. Gu, S. Zhou, S. Yao, and G. Pan, "Adaptive robust dispatch of integrated energy system considering uncertainties of electricity and outdoor temperature," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4691–4702, 2019.

[20] S. Lu, W. Gu, K. Meng, and Z. Y. Dong, "Economic dispatch of integrated energy systems with robust thermal comfort management," *IEEE Trans. Sustain. Energy*, 2020.

[21] J. Goh and M. Sim, "Distributionally robust optimization and its tractable approximations," *Oper. Res.*, vol. 58, no. 4-part-1, pp. 902–917, 2010.

[22] B. Fanzeres, A. Street, and L. A. Barroso, "Contracting strategies for renewable generators: A hybrid stochastic and robust optimization approach," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1825–1837, 2014.

[23] 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.

[24] Z. Shi, H. Liang, S. Huang, and V. Dinavahi, "Distributionally robust chance-constrained energy management for islanded microgrids," *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2234–2244, 2018.

[25] Y. Du, L. Jiang, C. Duan, Y. Li, and J. Smith, "Energy consumption scheduling of hvac considering weather forecast error through the distributionally robust approach," *IEEE Trans. Ind. Informat.*, vol. 14, no. 3, pp. 846–857, 2017.

[26] Y. Wang, Y. Du, C. Duan, H. Xu, and L. Jiang, "Data-driven distributionally robust energy consumption scheduling of hvac based on disjoint layered ambiguity set," in *2019 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, IEEE, 2019.

[27] Y. Zhang, J. Dong, T. Kuruganti, S. Shen, and Y. Xue, "Distributionally robust building load control to compensate fluctuations in solar power generation," in *2019 American Control Conference (ACC)*, pp. 5857–5863, IEEE, 2019.

[28] R. Gao and A. J. Kleywegt, "Distributionally robust stochastic optimization with wasserstein distance," *arXiv preprint arXiv:1604.02199*, 2016.

[29] P. M. Esfahani and D. Kuhn, "Data-driven distributionally robust optimization using the wasserstein metric: Performance guarantees and tractable reformulations," *Math. Programming*, vol. 171, no. 1-2, pp. 115–166, 2018.

[30] A. Zhou, M. Yang, M. Wang, and Y. Zhang, "A linear programming approximation of distributionally robust chance-constrained dispatch with wasserstein distance," *IEEE Trans. Power Syst.*, 2020.

[31] W. Xie, "On distributionally robust chance constrained programs with wasserstein distance," *Math. Programming*, pp. 1–41, 2019.

[32] R. Ji and M. A. Lejeune, "Data-driven distributionally robust chance-constrained optimization with wasserstein metric," *J Glob Optim*, pp. 1–33, 2020.

[33] N. Ho-Nguyen, F. Kılınç-Karzan, S. Küçükyavuz, and D. Lee, "Distributionally robust chance-constrained programs with right-hand side uncertainty under wasserstein ambiguity," *arXiv preprint arXiv:2003.12685*, 2020.

[34] C. Duan, W. Fang, L. Jiang, L. Yao, and J. Liu, "Distributionally robust chance-constrained approximate ac-opf with wasserstein metric," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4924–4936, 2018.

[35] 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, 2019.

[36] G. P. McCormick, "Computability of global solutions to factorable nonconvex programs: Part i—convex underestimating problems," *Math. Programming*, vol. 10, no. 1, pp. 147–175, 1976.

[37] Y. Zhou, M. Shahidehpour, Z. Wei, Z. Li, G. Sun, and S. Chen, "Distributionally robust co-optimization of energy and reserve for combined distribution networks of power and district heating," *IEEE Trans. Power Syst.*, vol. 35, no. 3, pp. 2388–2398, 2019.

[38] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, "A rewriting system for convex optimization problems," *J. Control Decis.*, vol. 5, no. 1, pp. 42–60, 2018.

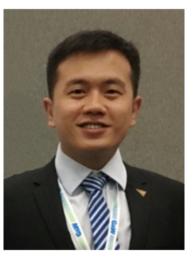
[39] M. A. Pertijs, K. A. Makinwa, and J. H. Huijsing, "A cmos smart temperature sensor with a 3σ inaccuracy of$\pm$0.1 c from-55 c to 125 c," *IEEE J Solid-State Circuits, 40 (12)*, 2005.

[40] C. Barth, M. Bonura, and C. Senatore, "High current probe for ic(b,t) measurements with ±0.01 k precision: Hts current leads and active temperature stabilization system," *IEEE Trans. Appl. Supercond.*, vol. 28, no. 4, pp. 1–6, 2018.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TSG.2021.3076237, IEEE Transactions on Smart Grid

Ge Chen (S'20) received the B.S. degree from Huazhong University of Science and Technology, Wuhan, China, in 2015 and the M.S. degree from Xi'an Jiaotong University, both in thermodynamic engineering. He is currently working toward the Ph.D. degree at University of Macau, Macau, China. His research interests include Internet of Things for smart energy, optimal operation and data-driven optimization under uncertainty.

Hongcai Zhang (S'14–M'18) received the B.S. and Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2013 and 2018, respectively. He is currently an Assistant Professor with the State Key Laboratory of Internet of Things for Smart City and Department of Electrical and Computer Engineering, University of Macau, Macao, China. In 2018-2019, he was a postdoctoral scholar with the Energy, Controls, and Applications Lab at University of California, Berkeley, where he also worked as a visiting student researcher in 2016. His current research interests include Internet of Things for smart energy, optimal operation and optimization of power and transportation systems, and grid integration of distributed energy resources.



Hongxun Hui (S'17–M'20) received both the Ph.D. and B. Eng degrees in electrical engineering from Zhejiang University in 2020 and 2015, respectively. He is currently a Post-doctoral Fellow with the State Key Laboratory of Internet of Things for Smart City, University of Macau. His research interests include modelling and optimal control of demand side resources in smart grid, the electricity market considering demand response, and the uncertainty analysis brought by flexible loads and renewable energies.

Yonghua Song (F'08) received the B.E. and Ph.D. degrees from the Chengdu University of Science and Technology, Chengdu, China, and the China Electric Power Research Institute, Beijing, China, in 1984 and 1989, respectively, all in electrical engineering. He was awarded DSc by Brunel University in 2002, Honorary DEng by University of Bath in 2014 and Honorary DSc by University of Edinburgh in 2019. From 1989 to 1991, he was a Post-Doctoral Fellow at Tsinghua University, Beijing. He then held various positions at Bristol University, Bristol, U.K.; Bath University, Bath, U.K.; and John Moores University, Liverpool, U.K., from 1991 to 1996. In 1997, he was a Professor of Power Systems at Brunel University, where he was a Pro-Vice Chancellor for Graduate Studies since 2004. In 2007, he took up a Pro-Vice Chancellorship and Professorship of Electrical Engineering at the University of Liverpool, Liverpool. In 2009, he joined Tsinghua University as a Professor of Electrical Engineering and an Assistant President and the Deputy Director of the Laboratory of Low-Carbon Energy. During 2012 to 2017, he worked as the Executive Vice President of Zhejiang University, as well as Founding Dean of the International Campus and Professor of Electrical Engineering and Higher Education of the University. Since 2018, he became Rector of the University of Macau and the director of the State Key Laboratory of Internet of Things for Smart City. His current research interests include smart grid, electricity economics, and operation and control of power systems. Prof. Song was elected as the Vice-President of Chinese Society for Electrical Engineering (CSEE) and appointed as the Chairman of the International Affairs Committee of the CSEE in 2009. In 2004, he was elected as a Fellow of the Royal Academy of Engineering, U.K. In 2019, he was elected as a Foreign Member of the Academia Europaea.