Regulation Capacity Evaluation of Large-scale Residential Air Conditioners for Improving Flexibility of Urban Power Systems

Hongxun Hui[a,1], Peipei Yu[a,1], Hongcai Zhang[a,*], Ningyi Dai[a], Wei Jiang[b], Yonghua Song[a]

[a]*State Key Laboratory of Internet of Things for Smart City, University of Macau, Macao, 999078, China*
[b]*School of Electrical Engineering, Southeast University, Nanjing, 210096, China*

Abstract

The residential air conditioners (RACs) are increasing rapidly in urban power systems and have been widely considered as good regulation resources for improving the system flexibility and resiliency. However, in practical power systems, it is difficult to comprehensively acquire millions of RACs' operating data and buildings' thermal data, which makes the available regulation capacity of RACs tricky to evaluate. To address this issue, this paper proposes a Gaussian Mixture Model (GMM)-based evaluation method by utilizing partial easily observable data. First, a control framework of large-scale RACs is developed to provide regulation services for the power system. Based on the thermal-electrical models of RACs and buildings, a quantification method of the available regulation capacities is proposed under the premise of guaranteeing all the users' comfortable indoor temperatures. Considering the practical condition of insufficient data acquisition of large-scale heterogeneous RACs, a GMM-based evaluation method is designed to calculate the probability of semi-info RACs' expected regulation capacities by sampling a small portion of full-info RACs' characteristics. Moreover, the Expectation Maximization Algorithm and the Bayesian Information Criterion are employed to optimize the multi-dimensional parameters and the component number of the GMM, which significantly improve the evaluation accuracy with lower complexity. The proposed models and methods are verified in a demonstration project on demand response in China.

Keywords: demand response, regulation capacity, residential air conditioner, GMM

*Corresponding author
Email address: hczhang@um.edu.mo (Hongcai Zhang)
[1]Authors contribute equally.

1. Introduction

Distributed renewable energies (e.g., winds and solar photovoltaics) are increasing rapidly, which brings more fluctuating output power to urban power systems [1]. Hence, more flexibility is needed to consume intermittent renewable energies [2] and maintain the system balance [3]. With the gradual phasing-out of traditional generating units (e.g., thermal generators and gas turbines) [4], more attentions are shifted from supply-side to demand-side [5], i.e., regulating load resources to provide flexibility by utilizing Internet of Things technologies [6]. In particular, residential air conditioners (RACs) are one of the most important flexible loads in urban power systems [7], because RACs consume more than half of the electricity at home [8] and can be regulated within the users' comfortable indoor temperature constraints [9].

Many studies have been carried out to regulate RACs to provide flexibility [10]. For example, Lu [11] proposes a direct load control algorithm for RACs to provide balancing services, which can maintain all the users' desired indoor temperatures [12]. Chen et al. [13] model RACs as a virtual power plant and employ a reinforcement learning method to optimize the power consumption of RACs [14], which shows RACs have large regulation flexibility within the users' comfortable constraints [15]. Besides, Song et al. [16] model RACs as thermal batteries to work with lithium-ion batteries for improving the system flexibility, which verifies RACs can be compatible with traditional energy storage resources [17]. RACs are also equivalent to traditional thermal generating units to participate in the frequency regulation services [18], where RACs even show better regulation performance than traditional generators [19]. Apart from the benefits to power systems, users can also decrease the energy cost [20] or obtain profits [21] by authorizing their RACs to participate in the regulation services. Furthermore, many demonstration projects have also been carried out. For example, in New York of America, Avangrid company provides free self-install smart kits to customers, which can achieve the remote control of RACs to participate in regulation services. In Bornholm of Denmark, the EcoGrid EU project is implemented to control RACs and other small-scale loads to participate in the real-time electricity market [22]. Besides, a project by the name of friendly interactive system of supply and demand sides are carried out in Jiangsu

Province, China, where RACs are regulated to decrease the peak-valley differences of loads [23].

In most power system's operation framework, the system operator is responsible for globally optimizing different kinds of regulation resources (e.g., thermal power generators, hydraulic turbines, etc) [24]. For example, the optimization and dispatch interval is set as 15 min in many power systems around the world [25]. Before each round of dispatch (i.e., 15 min earlier), the thermal power generators, hydraulic turbines, etc should submit their available regulation capacities to the system operator [26], so that the power system can reserve sufficient regulation capacities for maintaining stability. However, the system operator generally does not directly optimize or control flexible loads due to the large individual number. Flexible loads are aggregated and regulated by a control center, which is usually held by a distribution system operator, a load serving entity, or an aggregator [27]. For convenience, the control center is uniformly called the aggregator of RACs in this paper [28]. The aggregator is responsible for evaluating the total available regulation capacity of RACs and submitting this data to the power system operator. However, the operating number and power of RACs are significantly affected by users' stochastic behaviors and the changeable weather [29]. Compared with the regulation capacities provided by traditional generating units, the regulation capacities of RACs are tricky to evaluate quantitatively and accurately. There are mainly three difficulties:

Large-scale number: The capacity of one thermal generating unit is around $60 \sim 100$ MW or even higher, while the rated power of one RAC is only $1 \sim 3$ kW [30]. Even for the central RAC systems that generating cooling capacity for the whole house, the rated power is only around $6.5 \sim 10$ kW, which is still far less than one generating unit. Therefore, in order to provide significant regulation capacity similar as one generating unit, up to 100,000 RACs may need to be aggregated to participate in the regulation. That is to say, the aggregator has to detect and process millions of parameters to evaluate the available regulation capacity in real time [31], which brings vast computational complexity.

Insufficient data: One RAC's regulation capacity is influenced by lots of factors [32], including the corresponding building's thermal capacity and resistance, the RAC's rated power and energy efficiency ratio (EER), the outdoor ambient temperature, the

real-time indoor temperature, and the user's comfortable temperature requirement [33]. Some data can be easily detected while some data are difficult to obtain. For example, the RAC's operating power and the building's indoor temperature can be easily detected by widely used smart meters and temperature sensors, respectively [34]. However, the building's thermal capacity and resistance are related to various parameters, such as the house area, volume, ceiling, wall materials, windows and cracks [35]. They are difficult to be comprehensively obtained by the aggregator, especially for millions of buildings [36]. Hence, in practical power systems, the aggregator has to evaluate the total regulation capacity of RACs based on partially observable data.

Heterogeneity: In urban power systems, the RACs, buildings and users are all highly heterogeneous [37]. For example, RACs have different brands, rated power and EERs. Buildings have different areas, volumes, wall materials and so on [38]. Users have different comfortable requirements on their indoor temperatures at different time of a day [39]. Combining these three kinds of varieties together makes the thermal and electrical models of RACs become extremely complicated [40]. More seriously, in the condition with millions of RACs, obtaining these heterogeneous data building-by-building is impractical. The heterogeneity significantly increases the evaluation difficulty of RACs' regulation capacity.

Confronted by the above difficulties, some evaluation methods are proposed and attempt to solve this problem. For example, Zhang et al. [41] develops a plug-and-play learning framework to automatically identify the thermal model of each thermal zone in a building without manual configuration. Nevertheless, this method is based on a large number of smart devices. The cost may be high and not appropriate for millions of RACs nowadays. Xie et al. [42] propose a moment estimation method and a probability density estimation method to calculate the RACs' regulation capacity. However, these two methods are based on the assumption of normally distributed users, RACs and buildings [43], which cannot deal with large-scale randomly distributed RACs. Cui et al. [44] develop a capacity-time evaluation framework for RACs to quantify the regulation capacity, while this method is based on the accurate thermal-electrical model parameters and cannot be used in the condition of insufficient data acquisition. Cai et al. [45] and Javed et al. [46] propose to utilize the most sophisticated artificial neural network for evaluating the regulation capacity, while the complexity and computation

efficiency increase the implementation difficulty in practical power systems.

To address the aforementioned issues, this paper proposes a Gaussian Mixture Model (GMM)-based evaluation method for RACs by utilizing partial easily observable data. First, a control framework of large-scale RACs is developed to provide regulation services for the power system. Based on the thermal-electrical models of RACs and buildings, a quantification method of the available regulation capacities is proposed under the premise of guaranteeing all the users' comfortable indoor temperatures. Considering the practical condition of insufficient data acquisition of large-scale heterogeneous RACs, a GMM-based evaluation method is designed to calculate the probability of semi-info RACs' expected regulation capacities by sampling a small portion of full-info RACs' characteristics. Moreover, the Expectation Maximization (EM) algorithm and the Bayesian Information Criterion (BIC) are employed to optimize the multi-dimensional parameters and the component number of the GMM, which significantly improve the evaluation accuracy with lower complexity. The proposed models and methods are verified using the realistic data of a demonstration project in China.

The remainder of this paper is organized as follows. Section 2 presents the regulation framework and thermal-electrical model of RACs. Section 3 formulates the GMM and the optimization methodologies for evaluating large-scale heterogeneous RACs' regulation capacity. Numerical studies and result analysis are presented in Section 4. Finally, Section 5 concludes this paper.

2. Regulation Framework and Model of RACs

2.1. Regulation Framework of RACs

Fig. 1 shows the regulation framework of RACs. Each RAC is connected with a power line to get energy for generating cooling capacity. The aggregator is responsible for evaluating the available regulation capacity and regulating RACs to provide flexibility for the power system. When system stable operation is threatened by the fluctuating output power of renewable energies or the sudden disturbances of loads, the RACs will be controlled to adjust their operating power to assist the system balance. For example, it is assumed that a sudden reduction of renewable energies leads to the decrease of system frequency in Fig. 1. Then RACs will be controlled to reduce their operating power to provide up regulation services. The duration time is from t_{I} to t_{II}. The

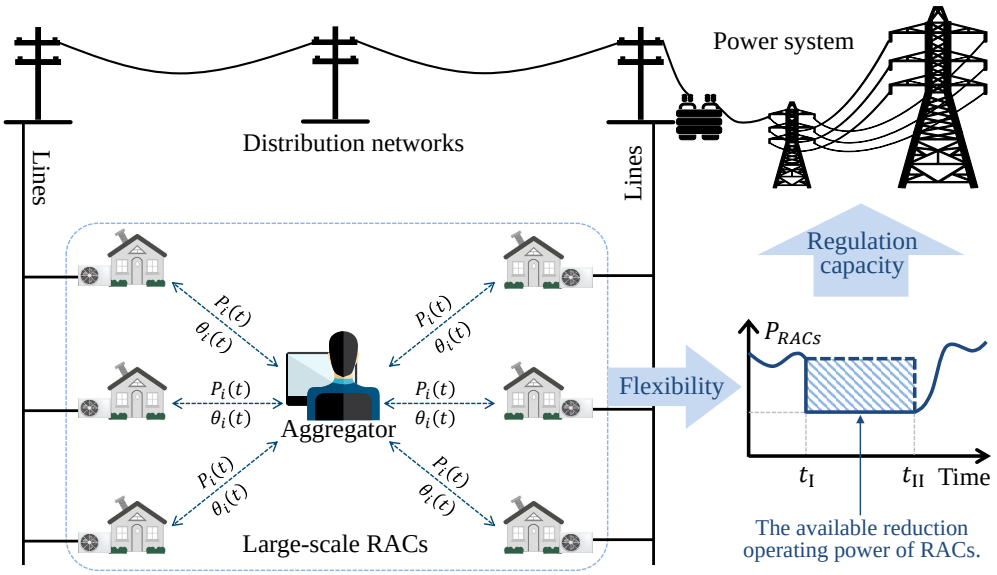

Figure 1: The regulation framework of RACs.

regulation requirement (including the regulation capacity and duration time) is based on the aggregator's submitted value before each round of dispatch (e.g., 15 min).

In order to increase the willingness of users for participating in regulation services, most aggregators tend to promise that the users' comfortable requirements on their indoor temperatures are guaranteed all the time. Hence, each RAC's regulation capacity and duration time should be constrained in some ranges to avoid uncomfortable indoor temperature. For example, if a building's indoor temperature wants to be maintained under 25℃, the corresponding RAC's regulation capacity and duration time may be constrained to be less than 1 kW and 15 min, respectively. Besides, the regulation capacity is also affected by the duration time. Under the same indoor temperature constraint and outdoor ambient temperature, this RAC's available regulation capacity may be smaller with the increase of the duration time. Therefore, the available regulation capacities of RACs should be evaluated considering the users' requirements, the buildings' characteristics, the real-time ambient temperature, and the regulation duration time, which will be illustrated in detail in the next two subsections 2.2 and 2.3.

Moreover, as described in Section 1, some data of the thermal-electrical models are difficult to be obtained by the aggregator. As shown in Fig. 1, here we choose two kinds of most common easily observable data to evaluate the regulation capacity: each RAC's operating power and the corresponding room's indoor temperature, which can be detected by widely used smart meters and temperature sensors. The specific GMM-

based evaluation method will be presented in detail in section 3.

2.2. Thermal-electrical Model of RACs

The thermal model of buildings installed with RACs can be expressed as [11]:

$$C_i \frac{\partial \theta_i(t)}{\partial t} = \frac{\theta_o(t) - \theta_i(t)}{R_i} - Q_i(t), \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \tag{1}$$

where $\theta_i(t)$ and $\theta_o(t)$ are the i-th building's indoor temperature (in °C) and the outdoor ambient temperature (in °C) at time t, respectively. Symbols C_i and R_i are the thermal capacity (in kJ/°C) and the thermal resistance (in °C/kW) of the i-th building, respectively. The prameter $\mathcal{I}$ is the set of RACs. Symbol Q_i is the cooling capacity (in kW) of the i-th RAC, which can be calculated as:

$$Q_i(t) = \eta_i P_i(t), \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \tag{2}$$

where η_i and $P_i(t)$ are the EER and operating power (in kW) of the i-th RAC, respectively. Generally, the values of EER distribute among $2.6 \sim 3.6$ [18]. With the incresae of EER, the same power energy can generate more cooling capacity.

In the stable operating state, the indoor temperature is generally maintained to be equal to the set value, i.e., $\theta_i(t) = \theta_i^{\text{set}}(t)$. Then the corresponding RAC's operating power can be calculated from Eqs. (1)-(2) as:

$$P_i(t) = \frac{\theta_o(t) - \theta_i^{\text{set}}(t)}{\eta_i R_i}, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}. \tag{3}$$

2.3. Regulation Capacity of RACs Within User Comfort Constraints

When RACs are controlled to provide regulation services for the power system, the most straight forward approach is to shut off all the RACs directly. Then the regulation capacity can be easily obtained by summarizing the RACs' initial operating power:

$$P_{\text{RACs}}(t) = \sum\nolimits_{i=1}^{N} P_i(t), \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}. \tag{4}$$

However, this evaluation method is impracticable, because this ON-OFF method probably has significant impact on users' indoor temperature comforts. Most research

and practical projects have shifted to regulating RACs within the users' comfort constraints. In this paper, it is assumed that all the users can set their comfortable indoor temperature ranges, i.e., $\theta_i \in [\theta_i^{\text{set}} - \theta_i^{\text{dev}},\ \theta_i^{\text{set}} + \theta_i^{\text{dev}}]$. Our problem is to evaluate the RAC's available regulation capacity within this comfortable indoor temperature range. Based on Eqs. (1)-(2), the indoor temperature deviation during the regulation process can be obtained by:

$$\int_{t_{\text{I}}}^{t_{\text{II}}} C_i \mathrm{d}\theta_i(t) = \int_{t_{\text{I}}}^{t_{\text{II}}} \frac{\theta_o(t) - \theta_i(t)}{R_i} \mathrm{d}t - \int_{t_{\text{I}}}^{t_{\text{II}}} \eta_i P_i^{\text{II}}(t) \mathrm{d}t, \quad \forall i \in \mathcal{I}, \forall t \in \mathcal{T}, \tag{5}$$

where $P_i^{\text{II}}(t)$ is the i-th RAC's operating power during the regulation period. It should be within the physical operational ranges, i.e., $P_i^{\text{II}}(t) \in [0, P_i^{\text{rated}}]$. It is assumed that the operating power $P_i^{\text{II}}(t)$ and the outdoor ambient temperature $\theta_o(t)$ keep constant during the regulation process (generally within 15 min). Then Eq. (5) can be calculated as:

$$P_i^{\text{II}} = \frac{\theta_o^{\text{reg}}}{\eta_i R_i} - \frac{\theta_i^{\text{II}} + \theta_i^{\text{I}}}{2\eta_i R_i} - \frac{C_i(\theta_i^{\text{II}} - \theta_i^{\text{I}})}{\eta_i T_D}, \quad \forall i \in \mathcal{I}, \tag{6}$$

where T_D is the regulation duration time (i.e., $T_D = t_{\text{II}} - t_{\text{I}}$); θ_o^{reg} is the outdoor ambient temperature during the regulation process; θ_i^{I} and θ_i^{II} are the indoor temperatures at the beginning and ending time of the regulation, respectively.

To evaluate the maximum regulation capacity, the indoor temperature after the regulation θ_i^{II} should be within the i-th user's comfort constraints, i.e., $\theta_i^{\text{II}} \in [\theta_i^{\text{set}} - \theta_i^{\text{dev}},\ \theta_i^{\text{set}} + \theta_i^{\text{dev}}]$. There are two typical regulation services, i.e., the up and down regulation services. The up regulation service is for increasing the power system frequency by decreasing the RACs' operating power. In this scenario, the i-th RAC's maximum regulation capacity can be obtained by allowing the indoor temperature to increase to the upper bound, i.e., $\theta_i^{\text{II}} = \theta_i^{\text{set}} + \theta_i^{\text{dev}}$. By contrast, the down regulation service is for decreasing the power system frequency by increasing the RACs' operating power. In this scenario, the i-th RAC's maximum regulation capacity can be obtained by allowing the indoor temperature to decrease to the lower bound, i.e., $\theta_i^{\text{II}} = \theta_i^{\text{set}} - \theta_i^{\text{dev}}$. Based on Eq. (6) and the i-th RAC's initial operating power, the i-th RAC's maximum up and down regulation

capacities can be calculated as follows:

$$\begin{cases} P_i^{\text{reg,up}} = P_i^{\text{I}} - P_i^{\text{II}} = P_i^{\text{I}} - \frac{\theta_o^{\text{reg}}}{\eta_i R_i} + \frac{\theta_i^{\text{set}} + \theta_i^{\text{dev}} + \theta_i^{\text{I}}}{2\eta_i R_i} + \frac{C_i(\theta_i^{\text{set}} + \theta_i^{\text{dev}} - \theta_i^{\text{I}})}{\eta_i T_D}, \quad \forall i \in \mathcal{I}, \\ P_i^{\text{reg,down}} = P_i^{\text{II}} - P_i^{\text{I}} = \frac{\theta_o^{\text{reg}}}{\eta_i R_i} - \frac{\theta_i^{\text{set}} - \theta_i^{\text{dev}} + \theta_i^{\text{I}}}{2\eta_i R_i} - \frac{C_i(\theta_i^{\text{set}} - \theta_i^{\text{dev}} - \theta_i^{\text{I}})}{\eta_i T_D} - P_i^{\text{I}}, \quad \forall i \in \mathcal{I}, \end{cases} \tag{7}$$

where P_i^{I} is the i-th RAC's operating power at time t_{I}, which can be easily detected by smart meters. Symbols $P_i^{\text{reg,up}}$ and $P_i^{\text{reg,down}}$ are the maximum available up and down regulation capacities provided by the i-th RAC, respectively. Hence, the total regulation capacities of RACs can be calculated by:

$$\begin{cases} P_{\text{RACs}}^{\text{reg,up}} = \sum_{i=1}^{N} P_i^{\text{reg,up}}, \quad \forall i \in \mathcal{I}, \\ P_{\text{RACs}}^{\text{reg,down}} = \sum_{i=1}^{N} P_i^{\text{reg,down}}, \quad \forall i \in \mathcal{I}. \end{cases} \tag{8}$$

3. Regulation Capacity Evaluation of Large-scale Heterogeneous RACs

3.1. Problem Formulation

As shown in Eqs. (6)-(8), the RAC's regulation capacity depends on the indoor temperature θ_i^{I} and θ_i^{II}, the building's thermal resistance R_i, the thermal capacity C_i, the RAC's EER η_i, the outdoor temperature θ_o^{reg}, and the regulation beginning time t_{I}, and the regulation duration time T_D. The RACs' operating power $P_i(t)$ and buildings' indoor temperature $\theta_i(t)$ can be monitored by widely used smart meters and temperature sensors, respectively. Hence, parameters P_i^{I}, θ_i^{I} and θ_o^{reg} can be detected at the beginning of the regulation. Moreover, the regulation signals (including the regulation beginning time t_{I} and the regulation duration time T_D) are from the aggregator. The comfortable indoor temperature ranges θ_i^{set} and θ_i^{dev} are preset values from users. Therefore, the above parameters can be detected by terminal devices easily and considered as known observable parameters.

By contrast, other parameters of the buildings (i.e., R_i, C_i), and RACs (i.e., η_i) are hard to be obtained by the aggregator, especially when the number of RACs reaches million level. Here we propose using the GMM to evaluate the total regulation capcity of RACs (i.e., $P_{\text{RACs}}^{\text{reg}}$)[2] by sampling $\alpha\%$ parameters of RACs. The sampling share $\alpha\%$

[2]Note that a RAC in the operating state can decrease its power to provide up regulation services, which can also increase its power to provide down regulation services, as shown in the Eq. (7). Both

can be 1% or even less, so as to substantially reduce the difficulty of data acquisition.

For convenience, this paper labels the sampled $\alpha\%$ of RACs as full-info RACs with the set $\mathcal{I}_\alpha \subset \mathcal{I}$, and labels the other RACs as semi-info RACs with the set $\mathcal{I}_\beta \subset \mathcal{I}$. The total set of RACs is $\mathcal{I}$, which satisfies $\mathcal{I}_\alpha \bigcup \mathcal{I}_\beta = \mathcal{I}$ and $\mathcal{I}_\alpha \bigcap \mathcal{I}_\beta = \varnothing$. The regulation capacities of full-info RACs (i.e., $P_{i,\alpha}^{\text{reg}}$) in the set $\mathcal{I}_\alpha$ can be calculated based on Eqs. (6)-(8). The key problem is to evaluate the regulation capacities of semi-info RACs (i.e., $P_{i,\beta}^{\text{reg}}$) in the set $\mathcal{I}_\beta$ based on the following data:

1. Parameters P_i^{I}, θ_i^{I}, θ_i^{set} and θ_i^{dev} of all the RACs (i.e., $\forall i \in \mathcal{I}$).
2. Parameters R_i, C_i and η_i of the $\alpha\%$ of RACs (i.e., $\forall i \in \mathcal{I}_\alpha \subset \mathcal{I}$).
3. The total number of RACs N, the number of full-info RACs N_α, the number of semi-info RACs N_β, the outdoor ambient temperature θ_o^{reg}, the regulation beginning time t_{I}, and the regulation duration time T_D.

3.2. Developing GMM for RACs

First, the multi-dimensional vector of the full-info RACs' parameters in set $\mathcal{I}_\alpha$ can be formed as:

$$\boldsymbol{x}_i = [P_i^{\text{I}}, \theta_i^{\text{I}}, \theta_i^{\text{set}}, \theta_i^{\text{dev}}]^T, \quad \forall i \in \mathcal{I}_\alpha. \tag{9}$$

Then the sample set of full-info RACs can be expressed as:

$$\mathcal{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_{N_\alpha}]. \tag{10}$$

As for arbitrary vectors $\boldsymbol{x}$ in $\mathcal{X}$, the joint probability density function (PDF) can be expressed as:

$$\mathcal{N}(\boldsymbol{x}) = \frac{1}{(2\pi)^{d/2}|D|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{u})^T D^{-1}(\boldsymbol{x}-\boldsymbol{u})\right], \quad \forall \boldsymbol{x} \in \mathcal{X}, \tag{11}$$

where d is the dimension of the vector $\boldsymbol{x}$. Symbols $\boldsymbol{u}$ and D are the mean value and

these two regulation capacities can be evaluated by the proposed GMM method. To simplify the expression, we use $P_{\text{RACs}}^{\text{reg}}$ to indicate $P_i^{\text{reg,up}}$ and $P_i^{\text{reg,down}}$ in Section 3.

covariance matrix of the vector $\boldsymbol{x}$, which can be expressed as:

$$\boldsymbol{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_d \end{bmatrix}, \quad D = \begin{bmatrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1d} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{d1} & \delta_{d2} & \cdots & \delta_{dd} \end{bmatrix}. \tag{12}$$

The GMM is composed of multiple PDFs, which can be calculated by:

$$f(\boldsymbol{x}) = \sum\nolimits_{k=1}^{K} \pi_k \cdot \mathcal{N}(\boldsymbol{x}|\boldsymbol{u}_k, D_k), \tag{13}$$

$$\sum\nolimits_{k=1}^{K} \pi_k = 1, \quad \pi_k > 0, \quad \forall \boldsymbol{x} \in \mathcal{X}, \forall k \in \mathcal{K}, \tag{14}$$

where K is the total number of components in GMM. Symbol $\mathcal{N}(\boldsymbol{x}|\boldsymbol{u}_k, D_k)$ indicates the k-th PDF. Symbol π_k is the weight of the k-th PDF, which is also called the prior probability of choosing the k-th PDF. The weight π_k should be larger than 0, and the summation of all the weights should be equal to 1.

Based on the GMM in Eqs. (13)-(14), the regulation capacities of the i-th RAC in set $\mathcal{I}_\beta$ can be obtained by:

$$P_{i,\beta}^{\text{reg}} = \sum\nolimits_{k=1}^{K} p_k(\boldsymbol{x}_i) \cdot \overline{P_{k,\alpha}^{\text{reg}}}, \quad \forall i \in \mathcal{I}_\beta, \forall k \in \mathcal{K}, \tag{15}$$

where $p_k(\boldsymbol{x}_i)$ is the probability of the vector $\boldsymbol{x}_i$ belonging to the k-th PDF in GMM; $\overline{P_{k,\alpha}^{\text{reg}}}$ features the expected regulation capacity of RACs in set $\mathcal{I}_\alpha$ in the k-th component of the GMM. These two parameters can be calculated as follows:

$$p_k(\boldsymbol{x}_i) = \frac{\pi_k \cdot \mathcal{N}(\boldsymbol{x}_i|\boldsymbol{u}_k, D_k)}{f(\boldsymbol{x}_i)}, \quad \forall i \in \mathcal{I}_\beta, \forall k \in \mathcal{K}, \tag{16}$$

$$\overline{P_{k,\alpha}^{\text{reg}}} = \frac{\sum_{i=1}^{N_\alpha} P_{i,\alpha}^{\text{reg}} \cdot p_k(\boldsymbol{x}_i)}{\sum_{i=1}^{N_\alpha} p_k(\boldsymbol{x}_i)}, \quad \forall i \in \mathcal{I}_\alpha, \forall k \in \mathcal{K}. \tag{17}$$

Therefore, the regulation capacity of RACs in the set $\mathcal{I}_\beta$ can be evaluated by Eq. (15) according to the probability in Eq. (16) and the expected regulation capacity in Eq. (17).

3.3. EM Algorithm for Optimizing GMM Parameters

The regulation capacity evaluation method in Eqs. (15)-(17) is based on the GMM in Eqs. (11)-(14). However, the parameters (K, π_k, $\boldsymbol{u}_k$, D_k) in Eqs. (11)-(14) are unknown, because the sample set $\mathcal{X}$ of full-info RACs is an incomplete data set. Actually, we do not know the vector $\boldsymbol{x}_i$ belongs to which component in GMM and cannot use the maximum likelihood estimation algorithm [47]. To address this issue, the EM algorithm is proposed to obtain these multi-dimensional parameters in GMM. Specifically, there are two steps: the expectation-step (E-step) and the maximization-step (M-step).

In the **E-step**, a latent variable vector $\boldsymbol{z}$ with K dimensions is introduced, which can indicate the vector $\boldsymbol{x}_i$ comes from the k-th component. Hence, the data set $\mathcal{X}$ can be extended to a complete data set $\mathcal{Y}$:

$$\mathcal{Y} = [\boldsymbol{x}_i, \boldsymbol{z}_i] = [\boldsymbol{x}_i, z_{i,1}, z_{i,2}, ..., z_{i,K}], \quad \forall i \in \mathcal{I}_\alpha, \tag{18}$$

where $z_{i,k}$ can only be equal to 0 or 1. The summation is equal to 1, i.e., $\sum_{k=1}^{K} z_{i,k} = 1$. For example, when the data set $\mathcal{X}$ has three components (i.e., $K = 3$), the latent variable $\boldsymbol{z}_i = (1, 0, 0)$, $\boldsymbol{z}_i = (0, 1, 0)$, and $\boldsymbol{z}_i = (0, 0, 1)$ indicate the vector $\boldsymbol{x}_i$ is generated by the first, second, and third component in the GMM, respectively.

Based on the complete data set $\mathcal{Y}$, the posterior probability of the vector $\boldsymbol{x}_i$ generated by the k-th component in GMM can be expressed as $p_k(\boldsymbol{z}_k = 1|\boldsymbol{x}_i)$. For simplifying the following expression, we use $\gamma_{i,k}$ to denote the $p_k(\boldsymbol{z}_k = 1|\boldsymbol{x}_i)$. Based on the Bayesian formula and the total probability formula, the $\gamma_{i,k}$ can be calculated by:

$$\begin{aligned}
\gamma_{i,k} &\triangleq p_k(z_k = 1|\boldsymbol{x}_i) \\
&= \frac{P(z_k = 1) \cdot p_k(\boldsymbol{x}_i|z_k = 1)}{p_k(\boldsymbol{x}_i)} \\
&= \frac{P(z_k = 1) \cdot p_k(\boldsymbol{x}_i|z_k = 1)}{\sum_{k=1}^{K} P(z_k = 1) \cdot p_k(\boldsymbol{x}_i|z_k = 1)} \\
&= \frac{\pi_k^j \cdot \mathcal{N}(\boldsymbol{x}_i|\boldsymbol{u}_k^j, D_k^j)}{\sum_{k=1}^{K} \pi_k^j \cdot \mathcal{N}(\boldsymbol{x}_i|\boldsymbol{u}_k^j, D_k^j)}, \\
&\quad \forall i \in \mathcal{I}_\alpha, \forall j \in \mathcal{J}, \forall k \in \mathcal{K},
\end{aligned} \tag{19}$$

where $p_k(\boldsymbol{x}_i|z_k = 1)$ is the prior probability; π_k^j is the probability of the vector $\boldsymbol{x}_i$ coming from the k-th component; j indicates the iteration times of the optimization for

calculating π_k^j, $\boldsymbol{u}_k^j$ and D_k^j. Note that the samples in each component are assumed to follow a normal distribution, i.e., $p_k(\boldsymbol{x}_i|z_k=1)=\mathcal{N}(\boldsymbol{x}_i|\boldsymbol{u}_k^j, D_k^j)$. Parameters π_k^j, $\boldsymbol{u}_k^j$ and D_k^j are variables to be solved.

Based on the $\gamma_{i,k}$ in Eq. (19), we can obtain the Q-function of the GMM:

$$\begin{aligned}\mathcal{Q} &= E_{\boldsymbol{z}}[\ln p(\mathcal{X}, \boldsymbol{z}|\pi^j, \boldsymbol{u}^j, D^j)|\mathcal{X}, \pi^j, \boldsymbol{u}^j, D^j] \\ &= \sum_{k=1}^{K}\left(\sum_{i=1}^{N_\alpha}\gamma_{i,k}\ln\pi_k^j + \sum_{i=1}^{N_\alpha}\gamma_{i,k}\ln\mathcal{N}(\boldsymbol{x}_i|\boldsymbol{u}_k^j, D_k^j)\right), \\ &\quad \forall i\in\mathcal{I}_\alpha, \forall j\in\mathcal{J}, \forall k\in\mathcal{K},\end{aligned} \tag{20}$$

where $E_{\boldsymbol{z}}$ is the expectation of the conditional probability distribution $p(\boldsymbol{z}|\mathcal{X}, \pi^j, \boldsymbol{u}^j, D^j)$ for the latent variable vector $\boldsymbol{z}$ under samples $\mathcal{X}$ and parameters $(\pi^j, \boldsymbol{u}^j, D^j)$; $\mathcal{Q}$ is the Q-function, which will be utilized in the next M-step.

In the **M-step**, the iteration model in the next time is calculated by maximizing the Q-function, which is expressed as:

$$\left[\pi_k^{j+1}, \boldsymbol{u}_k^{j+1}, D_k^{j+1}\right] = \arg\max(\mathcal{Q}), \quad \forall j\in\mathcal{J}, \forall k\in\mathcal{K}. \tag{21}$$

Get derivative with respect to the variables $(\pi_k^j,\ \boldsymbol{u}_k^j,\ D_k^j)$ and let the equations be equal to 0:

$$\frac{\partial\mathcal{Q}}{\partial\pi_k^j}=0, \quad \frac{\partial\mathcal{Q}}{\partial\boldsymbol{u}_k^j}=0, \quad \frac{\partial\mathcal{Q}}{\partial D_k^j}=0, \tag{22}$$

we can get the optimized parameters $(\pi_k^{j+1},\ \boldsymbol{u}_k^{j+1},\ D_k^{j+1})$ for the $(j+1)$-th iteration:

$$\pi_k^{j+1} = \frac{1}{N_\alpha}\sum\nolimits_{i=1}^{N_\alpha}\gamma_{i,k}, \forall i\in\mathcal{I}_\alpha, \forall j\in\mathcal{J}, \forall k\in\mathcal{K}, \tag{23}$$

$$\boldsymbol{u}_k^{j+1} = \frac{\sum_{i=1}^{N_\alpha}\gamma_{i,k}\boldsymbol{x}_i}{\sum_{i=1}^{N_\alpha}\gamma_{i,k}}, \quad \forall i\in\mathcal{I}_\alpha, \forall j\in\mathcal{J}, \forall k\in\mathcal{K}, \tag{24}$$

$$D_k^{j+1} = \frac{\sum_{i=1}^{N_\alpha}\gamma_{i,k}\|\boldsymbol{x}_i-\boldsymbol{u}_k^j\|^2}{\sum_{i=1}^{N_\alpha}\gamma_{i,k}}, \forall i\in\mathcal{I}_\alpha, \forall j\in\mathcal{J}, \forall k\in\mathcal{K}. \tag{25}$$

The EM algorithm can converge after several repetitions of the E-step and M-step. The ending criterion of the iteration is that the Q-function does not have significant increase, i.e., $\mathcal{Q}^{j+1}-\mathcal{Q}^j\le\varepsilon$.

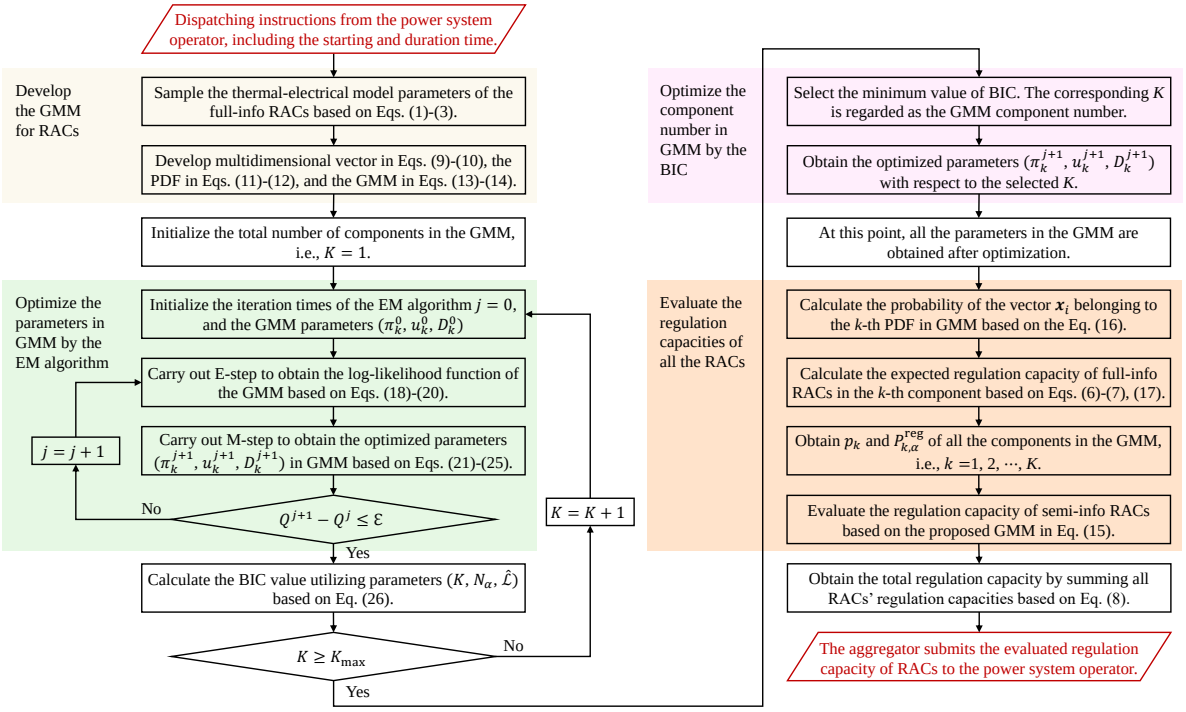

Figure 2: The implementation procedure of the GMM-based evaluation method.

3.4. BIC for Optimizing GMM Component Number

Based on the Eq. (18)-(25), the parameters of the GMM can be optimized. However, the total number of components in the GMM (i.e., K) is not optimized. Here, the BIC is employed to select the best component number K. The BIC can generally be regarded as a fitting degree for selecting the best model of all the samples, which is expressed as:

$$\text{BIC} = K \ln N_\alpha - 2 \ln \mathcal{L}. \tag{26}$$

Eq. (26) includes two items. The first item is composed of the total number of components K and the number of full-info RACs N_α, which features the complexity. The second item in the Eq. (26) is related to the log-likelihood function value $\ln \mathcal{L}$, which features the accuracy. In other words, the BIC model considers the complexity and accuracy of the GMM, which will be smaller with the increase of $\ln \mathcal{L}$ and the decrease of K and N_α. Therefore, the GMM is considered to be better with a smaller BIC value to decrease the complexity and increase the accuracy. The whole implementation procedure of the proposed GMM-based regulation capacity evaluation method for RACs is shown in Fig. 2.

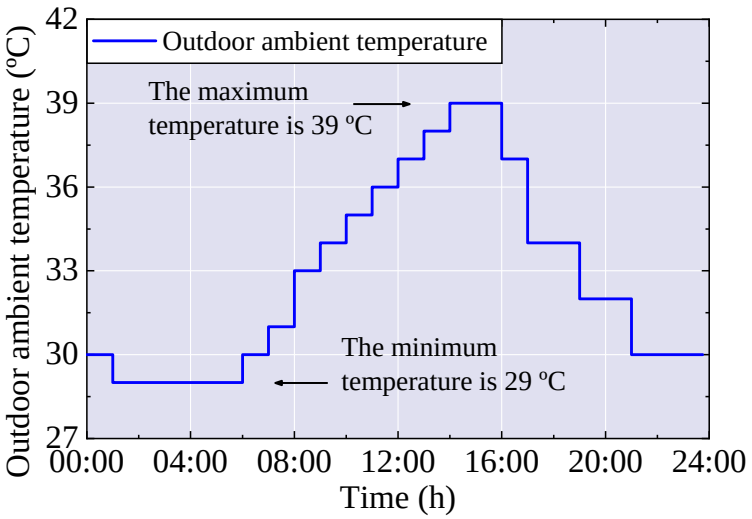

Figure 3: The outdoor ambient temperature.

4. Case Study

4.1. Test System

The test system is based on a realistic demonstration project in China, in which the total number of RACs and corresponding buildings N are 100,000. The buildings' thermal parameters are based on the design standard for residential buildings (JGJ134-2010) [48]. The EER of RACs η_i distributes among $2.6 \sim 3.6$. The set temperature of each RAC θ_i^{set} distributes among $18 \sim 27$℃ according to heterogeneous users' requirements [49]. The maximum deviation of the indoor temperature θ_i^{dev} is 2℃ from the set values. The sample parameter $\alpha\%$ is set as 1%, i.e., 1000 samples of full-info RACs. At the beginning of the regulation, the indoor temperature is generally near the set value in most buildings, i.e., $\theta_i^{\text{I}} \approx \theta_i^{\text{set}}$. Hence, without loss of generality, we can simplify the multi-dimensional vector of RACs' parameters $\boldsymbol{x}_i$ in Eq. (9) into a two-dimensional vector $\boldsymbol{x}_i = [P_i^{\text{I}}, \theta_i^{\text{I}}]$. Besides, considering the outdoor temperature has significant influence on the regulation capacity of RACs as shown in Eq. (1), here we use the differences between the outdoor and indoor temperature (i.e., $\theta_o^{\text{reg}} - \theta_i^{\text{I}}$) to replace θ_i^{I} to characterize the RAC's regulation capacity, i.e., $\boldsymbol{x}_i = [P_i^{\text{I}}, \theta_o^{\text{reg}} - \theta_i^{\text{I}}]$. The outdoor ambient temperature adopts the realistic test data in the demonstration project with 15 min resolution, as shown in Fig. 3. Three typical case studies are carried out in this paper:

Case 1: It is assumed that the system has a peak power at noon, and RACs are required to provide up regulation services. The beginning time is at 12:30 pm. The regulation duration period T_D is 15 min. The outdoor ambient temperature θ_o^{reg} is 37℃.

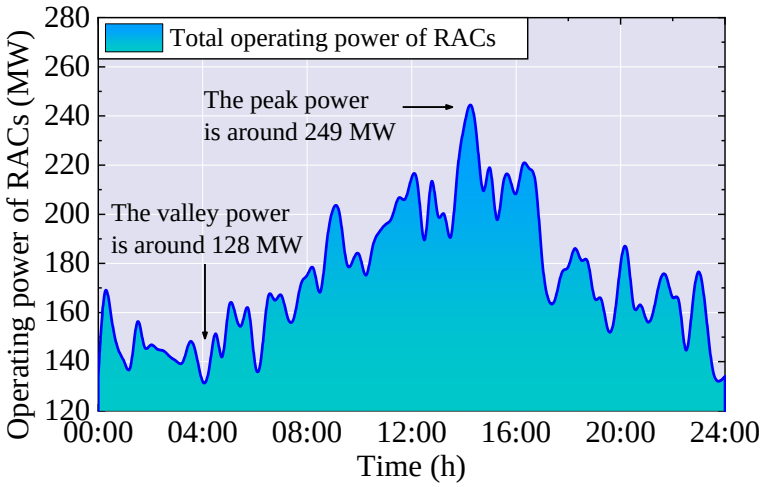

Figure 4: The initial total operating power of RACs.

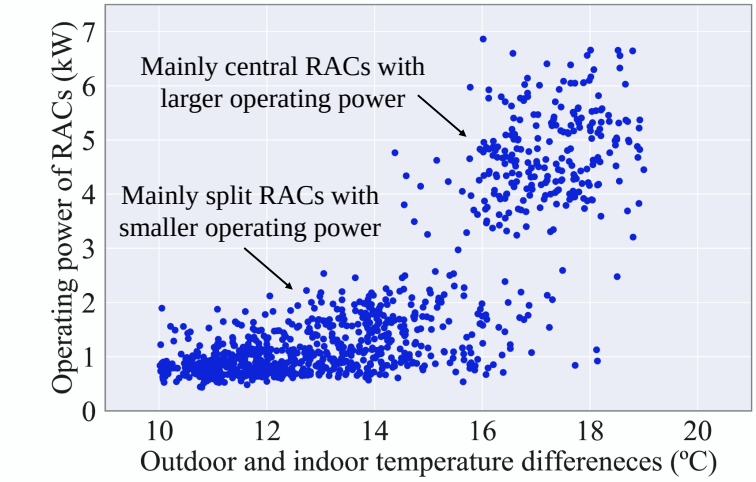

Figure 5: The distribution of RACs' operating power and outdoor-indoor temperature differences in Case 1.

Case 2: It is assumed that the system has a valley power at night, and RACs are required to provide down regulation services. The beginning time is at 04:00 am. The regulation duration period T_D is 15 min. The outdoor ambient temperature θ_o^{reg} is 29℃.

Case 3: Based on the fluctuating outdoor ambient temperature in Fig. 3 and the RACs' real-time operating power in Fig. 4, continuously evaluate the available up and down regulation capacities.

The test system is implemented using Python with an Intel core i7-9700 CPU @3.00 GHz with 16.0GB RAM.

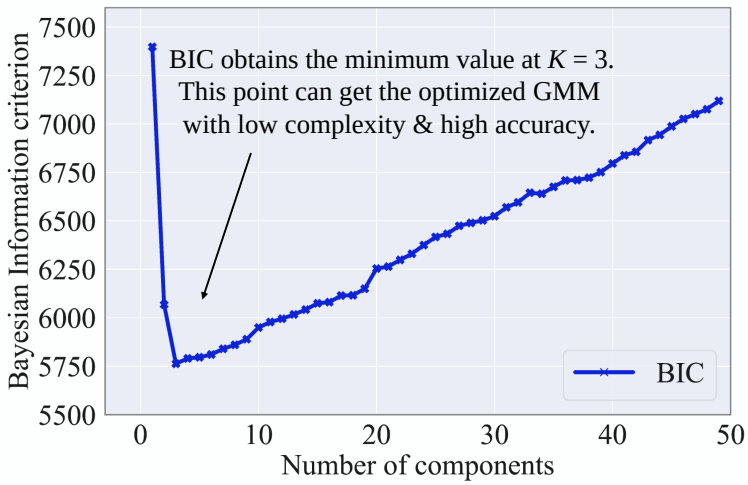

Figure 6: Bayesian information criterion for determining the component number of GMM in Case 1.

4.2. Evaluation Results of Up Regulation Capacity

The up regulation service begins at 12:30 pm and lasts for 15 min. Fig. 5 shows the initial distribution of full-info RACs. It can be seen that the operating power distributes among $0 \sim 7$kW. The temperature differences between the outdoor and indoor distribute among $10 \sim 19$℃ (i.e., the users' set temperatures are among $18 \sim 27$℃ considering the outdoor temperature is 37℃). On the whole, there are mainly two kinds of RACs: i) small split RACs with the operating power of $1 \sim 2.5$kW for cooling single rooms; ii) central RACs with a larger operating power of $3.5 \sim 7$kW for cooling all the rooms of a house.

In order to select the best component number in the GMM, the BIC is calculated based on Eq. (26). As shown in Fig. 6, the BIC can get the minimum value when the component number K is set as 3. That is to say, $K = 3$ can get the best balance between the accuracy and complexity of the GMM. When the component number K is larger than 3, the accuracy of the GMM can get improved, while the complexity gets increased more quickly. It can further impact the generalization ability and the computational efficiency, especially when there are millions of RACs. To illustrate this problem, the results of RACs based on GMMs with different component numbers are shown in Fig. 7, Fig. 8 and Table 1.

As shown in Fig. 7, there are three components. The first kind of RACs in component $k = 0$ represent some small split RACs with the operating power of $1 \sim 1.5$ kW. These RACs may be mainly used in small bedrooms with a relatively higher set temperature

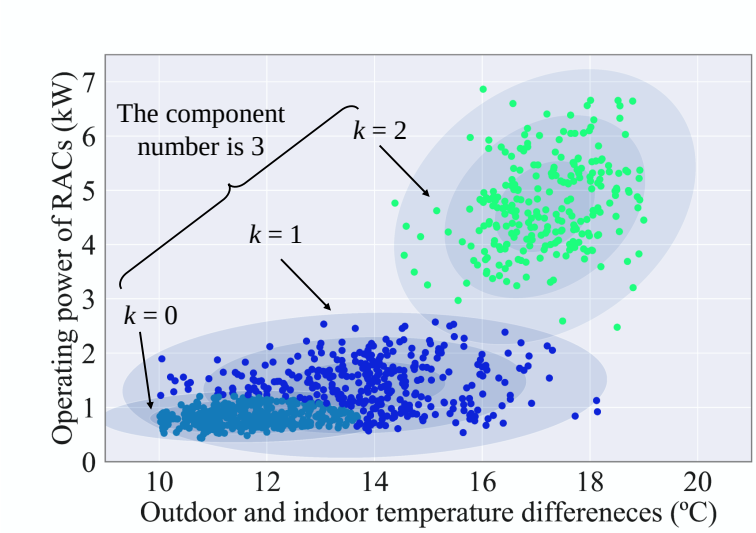

Figure 7: The GMM with 3 components for evaluating RACs' regulation capacity in Case 1.

(23 $\sim$ 27℃). The second kind of RACs in component $k = 1$ represent some small split RACs with the operating power of 1 $\sim$ 3 kW. They are similar with the first kind of RACs while with wider set temperature ranges (20 $\sim$ 26℃), which may be mainly used in living rooms or some bedrooms with lower set temperature. The RACs in component $k = 2$ in Fig. 7 have larger operating power (3.5 $\sim$ 7kW), which are mainly central RACs. Compared with split RACs in the previous two components, the central RACs generally have lower set temperature and consume more energy. Because the air tightness and thermal insulation of large spaces are generally lower than that of small rooms.

Based on the Eq. (17), the above three components in Fig. 7 have different expected regulation capacities, which are $\overline{P_{0,\alpha}^{\mathrm{reg,up}}} = 0.264$ kW, $\overline{P_{1,\alpha}^{\mathrm{reg,up}}} = 0.179$ kW, and $\overline{P_{2,\alpha}^{\mathrm{reg,up}}} = 0.590$ kW. The $\overline{P_{1,\alpha}^{\mathrm{reg,up}}}$ is less than $\overline{P_{0,\alpha}^{\mathrm{reg,up}}}$, because the set temperatures of most RACs in component $k = 1$ are lower than the set values in component $k = 0$. To maintain a lower indoor temperature, the RAC have to keep operating in relatively higher power. Moreover, the central RACs have the largest expected regulation capacity $\overline{P_{2,\alpha}^{\mathrm{reg,up}}}$, because they have larger operating power. Actually, compared with the 3 $\sim$ 6 times operating power than the split RACs in component $k = 0$, the central RACs only have about twice available regulation capacities of the split RACs. This is because the low air tightness and thermal insulation of large spaces equate to decreasing the thermal resistance, which is unfavorable for RACs to provide regulation capacities.

To illustrate the effect of component number in GMM, we also shows the GMM results with the component number of 15 in Fig. 8. It achieves a more precise classifica-

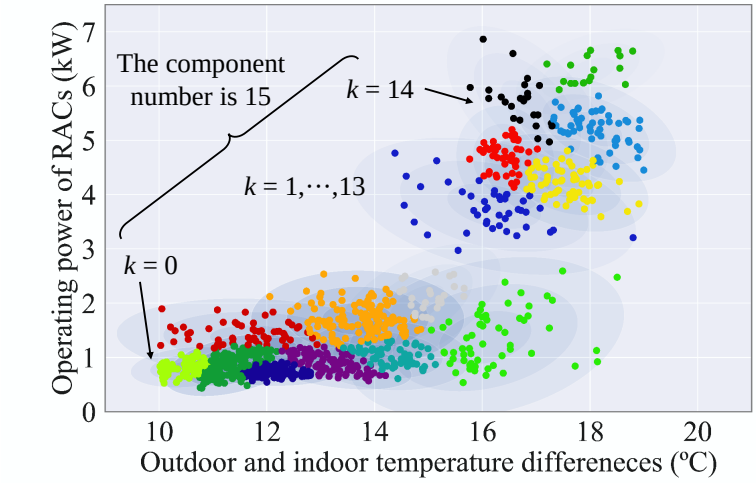

Figure 8: The GMM with 15 components for evaluating RACs' regulation capacity in Case 1.

Table 1: Evaluation Results of Up Regulation Capacities in Case 1

K	Up regulation capacity/MW	Evaluation accuracy/%	Computation time/s	Increase of time/%
3	31.20	98.48%	24.50	N/A
5	31.19	98.45%	27.63	12.78%
10	31.29	98.76%	32.04	30.77%
15	31.10	98.17%	37.44	52.82%
20	31.02	97.91%	43.03	75.63%

tion of different kinds of RACs, while it can decrease the generalization ability and the computational efficiency. As shown in Table 1, based on the optimized GMM with the component number $K = 3$, the evaluation accuracy of the regulation capacity can reach 98.48% by utilizing 24.50s computation time. This illustrates the effectiveness of the proposed GMM method for evaluating the RACs' up regulation capacity. Furthermore, with the increase of component number, the evaluation accuracy is almost unchanged or even lower because of insufficient generalization ability in $K = 20$ scenario. More seriously, the computation time is increased significantly to be more than 75.63%. This illustrates the effectiveness of the EM algorithm and BIC method for determining appropriate parameters in GMM.

4.3. Evaluation Results of Down Regulation Capacity

The down regulation service begins at 04:00 am and lasts for 15 min. Fig. 9 shows the initial distribution of full-info RACs. Compared with Fig. 5, the operating power becomes smaller and distributes among $0 \sim 5$kW. Especially some central RACs shift

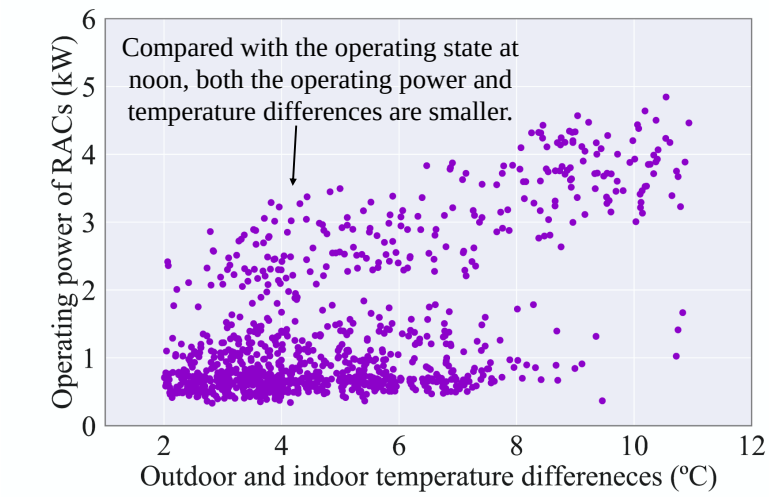

Figure 9: The distribution of RACs' operating power and outdoor-indoor temperature differences in Case 2.

their operating power from 3.5 $\sim$ 7kW at noon to 2 $\sim$ 5kW at night. The main reason is the temperature differences between the outdoor and indoor become smaller from 10 $\sim$ 19℃ in Fig. 5 to 2 $\sim$ 12℃ in Fig. 9. It contributes to the decrease of the operating power.

Similar with the up regulation in Case 1, the BIC is employed to select the best component number in GMM. As shown in Fig. 10, the BIC can get the minimum value when the component number K is set as 5. On this basis, we can obtain the GMM, as shown in Fig. 11. The split RACs are clustered into 3 components, i.e., $k = 0 \sim 2$. The central RACs are clustered into 2 components, i.e., $k = 3 \sim 4$. Based on the Eq. (17), the expected regulation capacities of RACs in the above five components are as follows: $\overline{P_{0,\alpha}^{\text{reg,down}}} = 0.323$ kW, $\overline{P_{1,\alpha}^{\text{reg,down}}} = 0.805$ kW, $\overline{P_{2,\alpha}^{\text{reg,down}}} = 0.728$ kW, $\overline{P_{3,\alpha}^{\text{reg,down}}} = 0.279$ kW, $\overline{P_{4,\alpha}^{\text{reg,down}}} = 0.185$ kW. Then we can obtain the evaluated regulation capacity of RACs, as shown in Table 2.

To illustrate the effect of component number in GMM, we also shows the GMM results with the component number of 20 in Fig. 12. It achieves a more precise classification of different kinds of RACs, while it can decrease the generalization ability and the computational efficiency. As shown in Table 2, based on the optimized GMM with the component number $K = 5$, the evaluation accuracy of the regulation capacity can reach 98.69% by utilizing 27.94s computation time. This illustrates the effectiveness of the proposed GMM method for evaluating the RACs' down regulation capacity.

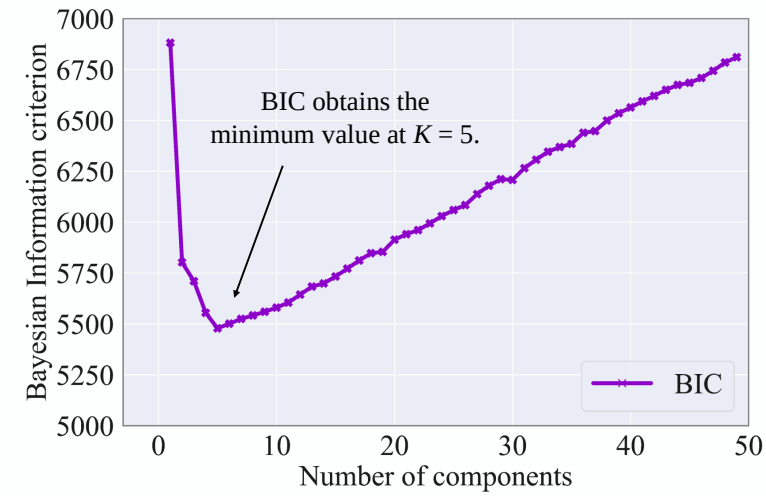

Figure 10: Bayesian information criterion for determining the component number of GMM in Case 2.

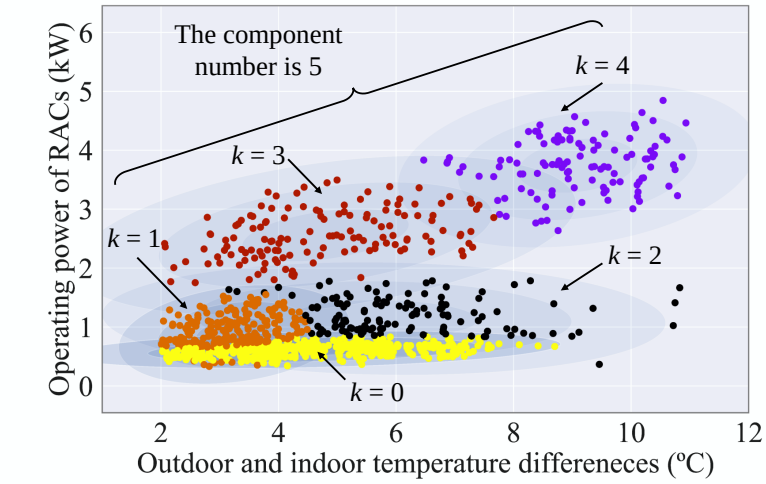

Figure 11: The GMM with 5 components for evaluating RACs' regulation capacity in Case 2.

Furthermore, with the increase of component number, the evaluation accuracy is almost unchanged, while the computation time is increased significantly to be more than 56.87%. This illustrates the effectiveness of the EM algorithm and BIC method for determining appropriate parameters in GMM.

4.4. Continuously Dynamic Evaluation of Up and Down Regulation Capacities

Based on the outdoor temperature in Fig. 3 and the RAC's operating power in Fig. 4, the available up and down regulation capacities can be evaluated continuously, as shown in Fig. 13[3]. It can be seen that the up regulation capacity is almost equal to the down

[3]Note that the continuously evaluated regulation capacities assume the RACs do not provide regulation service at the previous period. That is to say, the evaluation in Fig. 14 is based on the initial

Table 2: Evaluation Results of Down Regulation Capacities in Case 2

K	Down regulation capacity/MW	Evaluation accuracy/%	Computation time/s	Increase of time/%
5	37.80	98.69%	27.94	N/A
10	37.73	98.51%	33.52	19.97%
15	37.81	98.72%	39.27	40.55%
20	37.79	98.67%	43.83	56.87%

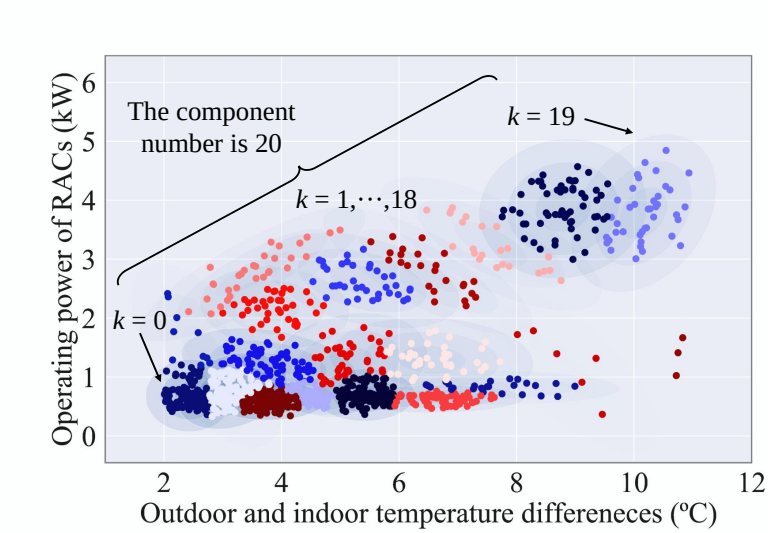

Figure 12: The GMM with 20 components for evaluating RACs' regulation capacity in Case 2.

regulation capacity at night (i.e., 00:00∼08:00 and 20:00∼24:00), because the maximum increasing and decreasing deviations of the indoor temperature are both set as 2℃ (i.e., $\theta_i^{\text{dev}} = 2$℃). Besides, the outdoor temperature is relatively stable at night, which causes similar up and down regulation capacities at night.

By contrast, the up regulation capacity is obviously less than the down regulation capacity in the day (i.e., the shadow area during 08:00∼20:00). In other words, the RACs' operating power can be increased larger to provide down regulation service in the day, while it can be decreased smaller to provide up regulation service in the day. The reason is that the difference between the indoor and outdoor temperatures will be higher with the decrease of the indoor temperature, which leads to a larger heat conductivity. In other words, the high outdoor ambient temperature in the hot day can lead to the increase of the indoor temperature more easily, which constrains the

operating power before the regulation. For example, if the RACs are controlled to provide regulation service at time t, the available up and down regulation capacities at time $t+1$ should be re-evaluated based on the updated RACs' operating states after the regulation.

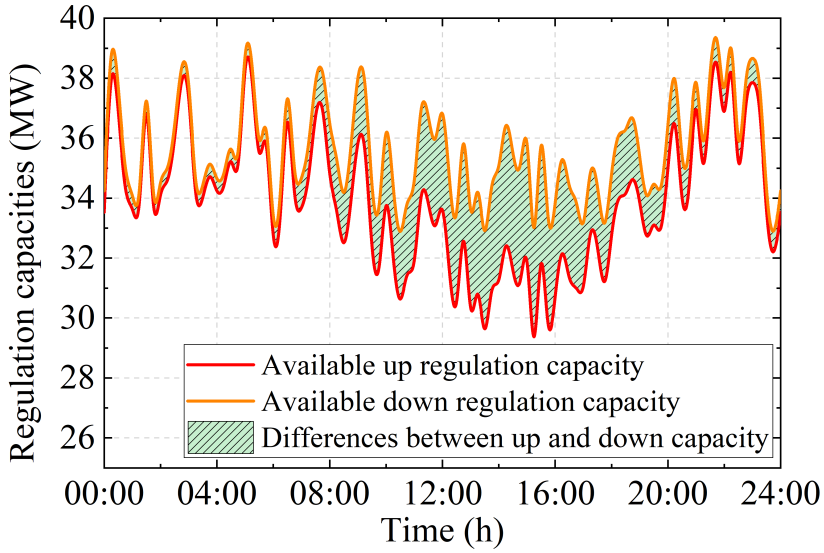

Figure 13: The available up and down regulation capacities of RACs.

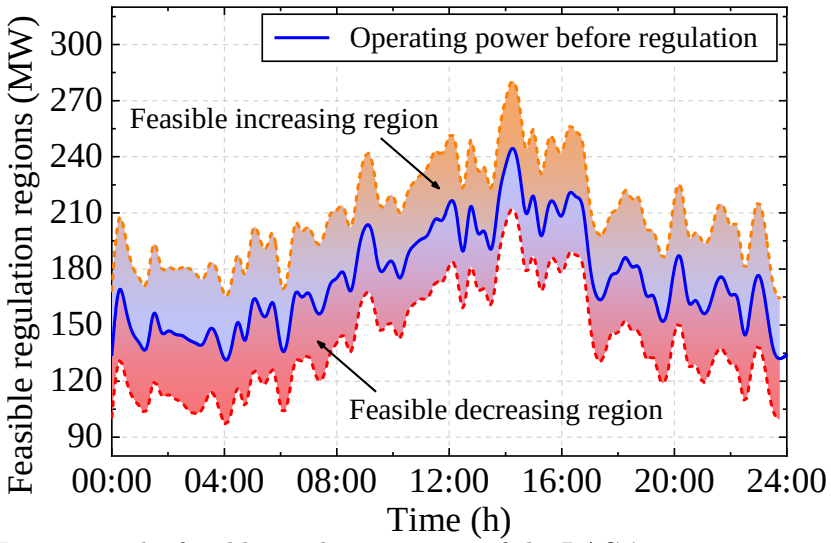

Figure 14: The feasible regulation regions of the RACs' operating power.

decrease of the RAC's operating power in the cooling state. This conclusion warns that in practical power systems, the up regulation capacity from RACs is more likely to be insufficient in the day. The power system operator probably needs to reserve larger up regulation capacity from other resources to guarantee the system stability.

Based on the available up and down regulation capacities in Fig. 13, we can obtain the feasible regulation regions of RACs' operating power, as shown in Fig. 14. The increasing and decreasing regions are the flexibility of RACs, which can be used directly by the power system operator to guide the regulation services.

Table 3: Evaluation Results of Up Regulation Capacities in Different Sampling Shares

$\alpha\%$	K	Up regulation capacity/MW	Evaluation accuracy/%	Computation time/s	Increase of time/%
0.1%	3	30.65	96.74%	22.15	−9.59%
1%	3	31.20	98.48%	24.50	N/A
10%	4	31.43	99.21%	32.27	31.71%

4.5. Effect Analysis of Sampling Shares of Full-info RACs

To illustrate the impact of $\alpha\%$, three scenarios with different $\alpha\%$ are compared in this subsection, including $\alpha\% = 0.1\%$, 1%, and 10%. The corresponding samples of full-info RACs are set as 100, 1000 and 10,000. Similar with Case 1, it is assumed that the system has a peak power at noon, and RACs are required to provide up regulation services. The beginning time is at 12:30 pm. The regulation duration period T_D is 15 min. The outdoor ambient temperature θ_o^{reg} is 37℃.

The results are shown in Table 3. The up regulation capacities are evaluated as 30.65 MW, 31.20 MW and 31.43 MW, respectively. The evaluation accuracy increases from 96.74% to 99.21% when the sample parameter $\alpha\%$ is increased from 0.1% to 10%. It verifies the evaluation error of the RACs' regulation capacity generally becomes smaller with the increase of $\alpha\%$, because more realistic data of full-info RACs can be obtained by the aggregator and less semi-info RACs need to be evaluated. For example, if $\alpha\%$ is set as 100%, all the parameters will be detected and the evaluation results can be regarded as the accurate value. However, a larger $\alpha\%$ also increases the computation time, which is increased to be more than 31.71% in the $\alpha\% = 10\%$ scenario. More seriously, the construction cost for obtaining 10,000 samples of full-info RACs is 10 times than obtaining 1000 samples in the $\alpha\% = 1\%$ scenario. Therefore, the proposed GMM method can deal with different sample parameters while the choice principle depends on a balance between the evaluation accuracy and the cost. In practical projects, if the requirement on the evaluation accuracy is higher than 95%, the sample parameter $\alpha\%$ can be set as $0.1\% \sim 10\%$.

5. Conclusion

This paper proposes a GMM-based evaluation method of the RACs' regulation capacity by utilizing partial easily observable data. The control framework and thermal-

electrical models of large-scale RACs are developed to provide regulation services for the power system. The proposed GMM-based evaluation method considers the practical condition of insufficient data acquisition and employs the EM algorithm to optimize the multi-dimensional parameters. The BIC is utilized to optimize the component number in GMM. The results show that the regulation capacity can be evaluated with more than 98% accuracy within 30s, which can satisfy the application requirement of day-ahead and intra-day dispatch. The proposed GMM-based evaluation method can promote the utilization of demand-side regulation resources in the near future smart grid paradigm.

6. Acknowledgement

This paper is funded in part by the Science and Technology Development Fund, Macau SAR (File no. SKL-IOTSC(UM)-2021-2023, and File no. 0003/2020/AKP).

References

[1] Zhang R, Jiang T, Bai L, Li G, Chen H, Li X, et al. Adjustable robust power dispatch with combined wind-storage system and carbon capture power plants under low-carbon economy. Int J Electr Power Energy Syst 2019;113:772–81.

[2] Xu X, Yan Z, Shahidehpour M, Li Z, Yan M, Kong X. Data-driven risk-averse two-stage optimal stochastic scheduling of energy and reserve with correlated wind power. IEEE Trans Sustain Energy 2019;11(1):436–47.

[3] Jiang T, Li Z, Jin X, Chen H, Li X, Mu Y. Flexible operation of active distribution network using integrated smart buildings with heating, ventilation and air-conditioning systems. Appl Energy 2018;226:181–96.

[4] Zhou Q, Shahidehpour M, Li Z, Xu X. Two-layer control scheme for maintaining the frequency and the optimal economic operation of hybrid AC/DC microgrids. IEEE Trans Power Syst 2018;34(1):64–75.

[5] Siano P. Demand response and smart grids—a survey. Renew Sust Energy Rev 2014;30:461–78.

[6] Hui H, Ding Y, Shi Q, Li F, Song Y, Yan J. 5G network-based Internet of Things for demand response in smart grid: A survey on application potential. Appl Energy 2020;257:113972.

[7] Xu S, Chen X, Xie J, Rahman S, Wang J, Hui H, et al. Agent-based modeling and simulation for the electricity market with residential demand response. CSEE J Power and Energy Syst 2020;7(2):368–80.

[8] Chen G, Zhang H, Hui H, Song Y. Fast Wasserstein-distance-based distributionally robust chance-constrained power dispatch for multi-zone HVAC systems. IEEE Trans Smart Grid 2021;12(5):4016–28.

[9] Shao C, Ding Y, Wang J, Song Y. Modeling and integration of flexible demand in heat and electricity integrated energy system. IEEE Trans Sustain Energy 2017;9(1):361–70.

[10] Jiang T, Wu C, Zhang R, Li X, Chen H, Li G. Flexibility clearing in joint energy and flexibility markets considering TSO-DSO coordination. IEEE Trans Smart Grid 2022;Early Access.

[11] Lu N. An evaluation of the HVAC load potential for providing load balancing service. IEEE Trans Smart Grid 2012;3(3):1263–70.

[12] Wang J, Huang S, Wu D, Lu N. Operating a commercial building hvac load as a virtual battery through airflow control. IEEE Trans Sustain Energy 2020;12(1):158–68.

[13] Chen T, Cui Q, Gao C, Hu Q, Lai K, Yang J, et al. Optimal demand response strategy of commercial building-based virtual power plant using reinforcement learning. IET Gener Transm Distrib 2021;15(16):2309–18.

[14] Chen G, Zhang H, Hui H, Dai N, Song Y. Scheduling thermostatically controlled loads to provide regulation capacity based on a learning-based optimal power flow model. IEEE Trans Sustain Energy 2021;12(4):2459–70.

[15] Hong J, Hui H, Zhang H, Dai N, Song Y. Distributed control of large-scale inverter air conditioners for providing operating reserve based on consensus with nonlin-

ear protocol. IEEE Int Things J 2022;doi: 10.1109/JIOT.2022.3151817; Early Access.

[16] Song M, Gao C, Yan H, Yang J. Thermal battery modeling of inverter air conditioning for demand response. IEEE Trans Smart Grid 2017;9(6):5522–34.

[17] Wang H, Yan Z, Shahidehpour M, Zhou Q, Xu X. Optimal energy storage allocation for mitigating the unbalance in active distribution network via uncertainty quantification. IEEE Trans Sustain Energy 2020;12(1):303–13.

[18] Hui H, Ding Y, Zheng M. Equivalent modeling of inverter air conditioners for providing frequency regulation service. IEEE Trans Ind Electron 2018;66(2):1413–23.

[19] Hui H, Ding Y, Chen T, Rahman S, Song Y. Dynamic and stability analysis of the power system with the control loop of inverter air conditioners. IEEE Trans Ind Electron 2020;68(3):2725–36.

[20] Siano P, Sarno D. Assessing the benefits of residential demand response in a real time distribution energy market. Appl Energy 2016;161:533–51.

[21] Chen T, Pourbabak H, Liang Z, Su W. An integrated evoucher mechanism for flexible loads in real-time retail electricity market. IEEE Access 2017;5:2101–10.

[22] Ding Y, Pineda S, Nyeng P, Østergaard J, Larsen EM, Wu Q. Real-time market concept architecture for EcoGrid EU—A prototype for European smart grids. IEEE Trans Smart Grid 2013;4(4):2006–16.

[23] Hui H, Ding Y, Lin Z, Siano P, Song Y. Capacity allocation and optimal control of inverter air conditioners considering area control error in multi-area power systems. IEEE Trans Power Syst 2020;35(1):332–45.

[24] Ding Y, Xie D, Hui H, Xu Y, Siano P. Game-theoretic demand side management of thermostatically controlled loads for smoothing tie-line power of microgrids. IEEE Trans Power Syst 2021;36(5):4089–101.

[25] Hui H, Siano P, Ding Y, Yu P, Song Y, Zhang H, et al. A transactive energy framework for inverter-based hvac loads in a real-time local electricity market considering distributed energy resources. IEEE Trans Ind Informat 2022;doi: 10.1109/TII.2022.3149941; Early Access.

[26] Jiang T, Yuan C, Bai L, Chowdhury BH, Zhang R, Li X. Bi-level strategic bidding model of gas-fired units in the interdependent electricity and natural gas markets. IEEE Trans Sustain Energy 2021;doi: 10.1109/TSTE.2021.3110864; Early Access.

[27] Shi Q, Cui H, Li F, Liu Y, Ju W, Sun Y. A hybrid dynamic demand control strategy for power system frequency regulation. CSEE J Power and Energy Syst 2017;3(2):176–85.

[28] Zhou Q, Li Z, Wu Q, Shahidehpour M. Two-stage load shedding for secondary control in hierarchical operation of islanded microgrids. IEEE Trans Smart Grid 2019;10(3):3103–11.

[29] Shi Q, Li F, Liu G, Shi D, Yi Z, Wang Z. Thermostatic load control for system frequency regulation considering daily demand profile and progressive recovery. IEEE Trans Smart Grid 2019;10(6):6259–70.

[30] Hu Q, Li F, Fang X, Bai L. A framework of residential demand aggregation with financial incentives. IEEE Trans Smart Grid 2016;9(1):497–505.

[31] Ahmad T, Zhang D. Using the internet of things in smart energy systems and networks. Sustain Cities and Society 2021;68:102783.

[32] Kim YJ. Optimal price based demand response of HVAC systems in multizone office buildings considering thermal preferences of individual occupants buildings. IEEE Trans Ind Informat 2018;14(11):5060–73.

[33] Shi Q, Chen CF, Mammoli A, Li F. Estimating the profile of incentive-based demand response (IBDR) by integrating technical models and social-behavioral factors. IEEE Trans Smart Grid 2020;11(1):171–83.

[34] Wang Y, Chen Q, Hong T, Kang C. Review of smart meter data analytics: Applications, methodologies, and challenges. IEEE Trans Smart Grid 2019;10(3):3125–48.

[35] Li W, Du P, Lu N. Design of a new primary frequency control market for hosting frequency response reserve offers from both generators and loads. IEEE Trans Smart Grid 2017;9(5):4883–92.

[36] Hu Q, Li F. Hardware design of smart home energy management system with dynamic price response. IEEE Trans Smart grid 2013;4(4):1878–87.

[37] Huang C, Zhang H, Song Y, Wang L, Ahmad T, Luo X. Demand response for industrial micro-grid considering photovoltaic power uncertainty and battery operational cost. IEEE Trans Smart Grid 2021;12(4):3043–55.

[38] Yan B, Long E, Meng X. Dynamic thermal reaction analysis of wall structures in various cooling operation conditions. Energy Convers Manag 2015;105:872–9.

[39] Yan B, Long E, Meng X, Zhang Y, Hou D, Du X. Influence of user behavior on unsatisfactory indoor thermal environment. Energy Convers Manag 2014;86:1–7.

[40] Zhang D, Zhu H, Zhang H, Goh HH, Liu H, Wu T. An optimized design of residential integrated energy system considering the power-to-gas technology with multifunctional characteristics. Energy 2022;238:121774.

[41] Zhang X, Pipattanasomporn M, Chen T, Rahman S. An IoT-based thermal model learning framework for smart buildings. IEEE Int Things J 2019;7(1):518–27.

[42] Xie D, Hui H, Ding Y, Lin Z. Operating reserve capacity evaluation of aggregated heterogeneous TCLs with price signals. Appl Energy 2018;216:338–47.

[43] Morales JM, Perez-Ruiz J. Point estimate schemes to solve the probabilistic power flow. IEEE Trans Power Syst 2007;22(4):1594–601.

[44] Cui W, Ding Y, Hui H, Lin Z, Du P, Song Y, et al. Evaluation and sequential dispatch of operating reserve provided by air conditioners considering lead–lag rebound effect. IEEE Trans Power Syst 2018;33(6):6935–50.

[45] Cai M, Pipattanasomporn M, Rahman S. Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques. Appl Energy 2019;236:1078–88.

[46] Javed F, Arshad N, Wallin F, Vassileva I, Dahlquist E. Forecasting for demand response in smart grids: An analysis on use of anthropologic and structural data and short term multiple loads forecasting. Appl Energy 2012;96:150–60.

[47] Cui M, Feng C, Wang Z, Zhang J. Statistical representation of wind power ramps using a generalized gaussian mixture model. IEEE Trans Sustain Energy 2017;9(1):261–72.

[48] Hui H, Ding Y, Luan K, Chen T, Song Y, Rahman S. Coupon-based demand response for consumers facing flat-rate retail pricing. CSEE J Power and Energy Syst 2022;Early Access.

[49] Xie K, Hui H, Ding Y, Song Y, Ye C, Zheng W, et al. Modeling and control of central air conditionings for providing regulation services for power systems. Appl Energy 2022;315:119035.