Frequency Regulation Capacity Offering of District Cooling System Based on Reinforcement Learning

Peipei Yu, Hongxun Hui, Hongcai Zhang, Chao Huang, Yonghua Song
State Key Laboratory of Internet of Things for Smart City, University of Macau, Macao, China
hongxunhui@um.edu.mo, hczhang@um.edu.mo

2022 IEEE Power & Energy Society General Meeting (PESGM) | 978-1-6654-0823-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/PESGM48719.2022.9916851

***Abstract*—With the development of commercial buildings in modern cities, the district cooling system (DCS) is rapidly increasing due to its high efficiency for providing cooling services to multiple buildings. By utilizing buildings' inherent thermal inertia, DCS has huge potential to participate in the electricity market and provide regulation capacity. However, offering the DCS's regulation capacity ahead of the operating hour is quite challenging. Its available capacity is changing with time significantly due to multiple commercial buildings' stochastic cooling demand and the electricity market's uncertain signals. To address this issue, this paper proposes a strategy framework to offer the DCS's available regulation capacity for achieving the maximum revenue while respecting the users' comfortable indoor temperature requirements. First, the DCS's revenue model is developed based on its regulation capacity and performance score in the electricity market. Then, the regulation capacity offering strategy in each time slot is formulated as a Markov Decision Process (MDP). On this basis, the deep determined policy gradient algorithm is implemented to iterate the policy in the MDP to obtain the optimal results, which requires no knowledge of the uncertainties or physical model. Finally, we use the realistic RegA frequency regulation signals from PJM market to validate that the proposed strategy is effective in evaluating the system's available capacity with high-quality performance.**

***Index Terms*—Regulation capacity offering, demand response, district cooling system, deep reinforcement learning.**

I. Introduction

With the increasing penetration of intermittent and uncertain renewable generation in power systems, more regulation resources are required [1]. Traditionally, the regulation capacity is provided by supply-side resources (e.g., thermal and hydro generators), while these resources are being phased out and may cause insufficient regulation capacities in the near future. With the development of information and communication technologies [2], demand-side resources (DSRs) are paid more attention to provide regulation services for power systems. This paper focuses on one type of emerging DSR, i.e., *district cooling system* (DCS). First, DCS is increasing rapidly in modern cities due to its high efficiency to provide cooling services for multiple commercial buildings [3]. Thus DCS will take a large share of modern cities' power consumption and have huge regulation potential. Second, one common DCS possesses a quite large cooling capacity (up to 100MW) and can be regarded as a natural aggregator to directly participate in the electricity market [4]. Third, DCS can modulate its power continuously over a wide range of time duration with negligible impacts on the indoor temperature utilizing buildings' inherent thermal inertia [5]. Therefore, DCS is a promising DSR while has been rarely studied previously.

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. 0137/2019/A3).

The DCS's main motivation for providing regulation capacity is to obtain extra revenues with negligible impact on the indoor temperature [6]. In most electricity markets, revenues are proportional to the regulation capacity offer and performance score [7], i.e., a larger capacity offer or a higher performance score can bring more revenues to the regulation service provider. However, these two values go against each other at most time. A too large capacity offer may lead to a bad regulation performance and get a quite low performance score, because the building's thermal inertia is limited and subject to its comfort requirements. If the score is lower than the market's access requirement, the DCS may be banned from participating in the market anymore [7]. By contrast, a smaller regulation capacity undoubtedly can decrease the revenue even though the DCS gets a higher performance score. However, it is quite challenging for DCS to give a adequate capacity offering strategy ahead of the operating hour considering the score constraints and buildings' comfort requirements because of the following reasons:

1) *Uncertainty*: The stochastic regulation signals from the electricity market are hard to predict, because it is based on the real-time balance of the whole power system. Thus, it is difficult to ensure the actual regulation performance. Furthermore, the random human behaviors and ambient temperature can directly influence the buildings' cooling demand and indoor temperature, which increase the difficulty to identify the available capacities for maintaining the comfort during the regulation process.
2) *Complexity*: A DCS is a networked system that serves for over ten commercial buildings within a radius of 2 km. It has complex thermal dynamics that include cooling power generating, transmitting and consuming processes whose parameters cannot be obtained accurately in practice. Thus the DCS is hard to be modelled.
3) *Continuity of decision-making*: The DCS's regulation capacity in each hour can directly impact the next hour's system operating states (e.g., the DCS's power consumption, the building's indoor temperature), which will further affect the next hour's available regulation capacity [8]. Thus, the optimal capacity offering strategy

should consider the impacts among different hours, so as to achieve maximum revenues in a long term.

To evaluate the regulation potential of HVAC, some researchers have proposed relevant methodologies. D. Xie et al. [9] develops a probability density estimation method to evaluate the regulation capacity of large-scale aggregated HVAC systems, which is based on an individual HVAC model. J. Cai et al. [10] identifies the maximum available capacity based on a pseudo-optimization method, while the proposed strategy relies heavily on the accurate and steady model parameters. M. B. Anwar et al. [11] presents an extensive multi-perspective method to assess the capacity of the aggregated residential HVAC, where the capacity market needs to be explicitly modeled. X. Li et al. [12] formulates the flexibility assessment problem as a quadratic programming to minimize the cost, based on the detailed thermal model of the aggregated buildings. All the aforementioned evaluation methods requires the accurate system model, which can not work for DCS. Besides, their objectives mainly focus on the cost while the performance score is not taken into account.

To address the above issues, we propose a model-free method to evaluate the available DCS regulation capacity based on deep reinforcement learning (DRL). DRL has achieved great success in challenging decision-making problems, which can address the uncertainty and continuity problem efficiently [13]. Compared with the aforementioned literature, this paper's main contributions include:

1) We develop a model for DCS to achieve maximum revenues in frequency regulation markets, which considers the requirements from both the market performance score and buildings' temperature comfort.
2) We formulate the capacity offering problem as a Markov Decision Process (MDP). The designed state space and reward function in the MDP can effectively balance the trade-off between the DCS's revenue and the penalty risk due to bad performance.
3) We develop a model-free strategy based on the deep determined policy gradient algorithm. The proposed strategy requires little knowledge of the accurate system model and uncertainty distributions.

II. Hour-ahead Regulation Capacity Offering

This section describes an hour-ahead regulation capacity offering problem of DCS for participating in the regulation market, in which the DCS is assumed as a market price-taker. The market environment is based on the PJM market, which can be also adapted to other markets. The regulation market closes 60 minutes before the operating hour, and all participators' regulation capacity offers should be determined before the market closes [14].

A. Modelling of DCS

As shown in Fig. 1, a DCS is composed of one energy station, multiple pipelines and buildings. Chillers in the energy station, as the system's main power consumer, produce chilled water for buildings through pipelines. There are two isolated water loops in the DCS to transport thermal energies. The first water loop is the water cycle in pipelines, and the second water loop is the water cycle in each building. The heat exchanging process between two water loops is executed in each building's heat exchanger. Based on the energy balance, chillers' power consumption can be formulated as:

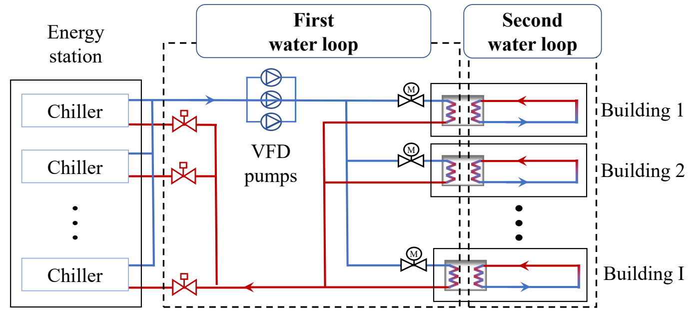

Fig. 1. The operating framework of DCS.

$$P_\tau^{\text{ch}} = \frac{Q_\tau^{\text{ch}}}{\text{COP}}, \quad \forall \tau, \tag{1}$$

$$Q_\tau^{\text{ch}} = c^{\text{w}}(T_\tau^{\text{ch,r}} - T_\tau^{\text{ch,s}}) \sum\nolimits_{i \in \mathcal{I}} m_{i,\tau}^{\text{I}}, \quad \forall \tau, \tag{2}$$

where P_τ^{ch} and Q_τ^{ch} are chillers' electricity power and cooling power at time τ, in kW, respectively. The parameter COP is the chiller's coefficient of performance. Symbol c^{w} is the heat capacity of water, in kJ/(kg·°C). Symbols $T_\tau^{\text{ch,r}}$ and $T_\tau^{\text{ch,s}}$ are the supply water temperature and return water temperature of the chiller at time t, in °C. The set I denotes the group of all buildings. Symbol $m_{i,\tau}^{\text{I}}$ is the ith building's mass flow rate in the first water loop at time t, in m^3/s.

The chiller's return water temperature $T_\tau^{\text{ch,r}}$ in Eq. (2) is determined by all buildings' return water temperatures $T_{i,\tau}^{\text{I,r}}$ and corresponding mass flow rates $m_{i,\tau}^{\text{I}}$, which is given as:

$$T_\tau^{\text{ch,r}} = \frac{\sum_{i \in \mathcal{I}} m_{i,\tau}^{\text{I}} T_{i,\tau}^{\text{I,r}}}{\sum_{i \in \mathcal{I}} m_{i,\tau}^{\text{I}}}, \quad \forall \tau. \tag{3}$$

Here, the building's return water temperature $T_{i,\tau}^{\text{I,r}}$ is determined by the exchanging heat $Q_{i,\tau}$ between two water loops in Fig. 1. The $Q_{i,\tau}$ depends on the heat exchanger's performance, which can be calculated by:

$$Q_{i,\tau} = k_i F_i \frac{(T_{i,\tau}^{\text{II,r}} - T_{i,\tau}^{\text{I,s}}) - (T_{i,\tau}^{\text{II,s}} - T_{i,\tau}^{\text{I,r}})}{\ln[(T_{i,\tau}^{\text{II,r}} - T_{i,\tau}^{\text{I,s}})/(T_{i,\tau}^{\text{II,s}} - T_{i,\tau}^{\text{I,r}})]}, \forall i \in \mathcal{I}, \forall \tau, \tag{4}$$

where k_i and F_i are heat exchangers' heat transfer coefficient and surface area, respectively. The temperatures $T_i^{\text{II,r}}$ and $T_i^{\text{II,s}}$ are the ith building's return and supply water temperatures in the second water loop, respectively.

Based on the Eq. (6), the energy balance can be formulated to express the temperature changes of the supply and return water, as follows:

$$Q_{i,\tau} = m_{i,\tau}^{\text{I}} c^{\text{w}} (T_{i,\tau}^{\text{I,r}} - T_{i,\tau}^{\text{I,s}}) \cdot \eta_i^{\text{I}}, \quad \forall i \in \mathcal{I}, \forall \tau, \tag{5}$$

$$Q_{i,\tau} = m_{i,\tau}^{\text{II}} c^{\text{w}} (T_{i,\tau}^{\text{II,r}} - T_{i,\tau}^{\text{II,s}}), \quad \forall i \in \mathcal{I}, \forall \tau, \tag{6}$$

978-1-6654-0823-3/22/$31.00 ©2022 IEEE

where η_i^{I} is the heat transfer efficiency from the first water loop to the second water loop; $m_{i,\tau}^{\mathrm{II}}$ is the mass flow rate in the second water loop at time τ.

Finally, the building's water cycle supplies cooling capacity to maintain comfortable indoor temperature. The thermal dynamic process can be given as:

$$D_i \frac{\partial T_{i,\tau}^{\mathrm{A}}}{\partial \tau} = \frac{T_\tau^{\mathrm{O}} - T_{i,\tau}^{\mathrm{A}}}{R_i} + \zeta_{i,\tau} - \eta_i^{\mathrm{II}} Q_{i,\tau}, \forall i \in \mathcal{I}, \forall \tau, \quad (7)$$

where $T_{i,\tau}^{\mathrm{A}}$ is the ith building's indoor temperature. Parameters D_i and R_i are the thermal capacity (in kJ/°C) and thermal resistance (in °C/W) of each building, respectively. Symbol $\zeta_{i,\tau}$ is the indoor heat loads (e.g., human behaviors and equipment), in kW; η_i^{II} is the heat transfer efficiency from the second water loop to the building.

B. Regulation Capacity Offering by DCS in Market

During the t-th operating hour, the DCS should regulate its power consumption to follow the market signal, and then get the hourly revenue r_t^{M} according to its actual regulation performance, as follows:

$$r_t^{\mathrm{M}} = C_t^{\mathrm{ch}} s_t^{\mathrm{M}} p_t, \quad \forall t, \quad (8)$$

where C_t^{ch} and s_t^{M} are the capacity offer and regulation performance score of the DCS, respectively; p_t represents the clearing price in the regulation market. The specific score calculation rule is based on PJM manual [14]. Considering that one DCS's capacity is not large enough to influence the price, we regard the DCS as a market price-taker. Therefore, in this paper, the DCS aims to maximize the value of $C_t^{\mathrm{ch}} s_t^{\mathrm{M}}$. Furthermore, considering that last hour's capacity offer will significantly influence the current hour's offering strategy, the DCS's objective is to maximize the total cumulative revenue during the continuous 24 hours as follows:

$$\max_{C_t^{\mathrm{ch}}} \sum_{j=1}^{24} C_t^{\mathrm{ch}} s_t^{\mathrm{M}} p_t, \quad (9)$$

$$\text{s.t.: } s_t^{\mathrm{M}} \geq \underline{s}^{\mathrm{M}}, \Delta T_{i,t} \leq \Delta \overline{T_i}, \quad \forall i \in \mathcal{I}, \forall t \quad (10)$$

where $\underline{s}$ and $\Delta\overline{T_i}$ are the requirements of markets' performance score and buildings' comfortable temperature range, respectively. It is noted that a regulation market has a minimum performance score to access (e.g., $s^{\mathrm{M}} \geq 0.75$ in PJM), which further presents a challenge to the capacity offering – a too large capacity offer may cause a quite low score that may let the DCS to be kicked out of market; while a too small capacity offer may cause a low revenue. This problem will be addressed in the following section.

III. Online Capacity Offering Based on Deep Determined Policy Gradient

This section formulates the capacity offering problem of Eq. (9) as a Markov Decision Process (MDP) and adopts a model-free strategy to achieve the optimal capacity offer.

As shown in Fig. 2, according to the state, the agent determines the capacity offering before the market closes. Then the agent receives the next state and hourly revenue after the operating hour, which is stored as historical data to update the agent's offering strategy.

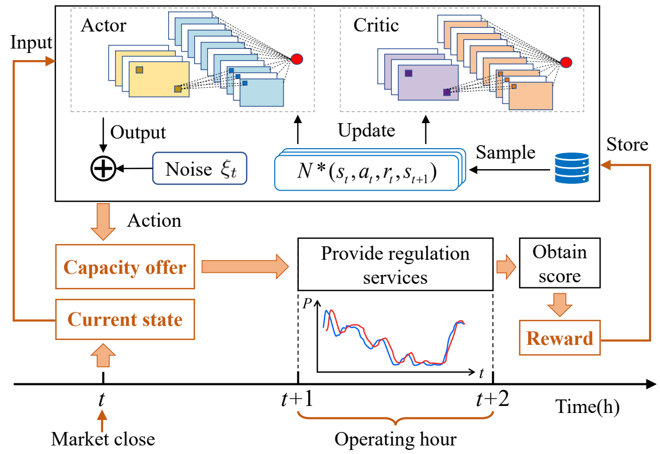

Fig. 2. The DDPG Scheme of Capacity Offering Problem.

A. Preliminaries of an MDP

As a formulation of sequential decision making, MDP is defined as a tuple $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma \rangle$, where $\mathcal{S}$ is a state space, $\mathcal{A}$ is an action space, $\mathcal{P}$ is a transition probability function, $\mathcal{R}$ is a reward function and $\gamma \in [0,1]$ is a discount factor. In an MDP, the decision maker is defined as an agent. At each time stot $t \in [0,T]$, the agent observes the DCS's operating state $\boldsymbol{s}_t \in \mathcal{S}$ and determines an action $\boldsymbol{a}_t \in \mathcal{A}$. Then it receives a corresponding reward $R_t \in \mathbb{R}$ after executing the action. The system state turns to $\boldsymbol{s}_{t+1}$ from $\boldsymbol{s}_t$, based on the transition probability $\mathcal{P}(\boldsymbol{s}_{t+1}|\boldsymbol{s}_t, \boldsymbol{a}_t)$. The sequence of states and actions $\{\boldsymbol{s}_0, \boldsymbol{a}_0, \boldsymbol{s}_1, ..., \boldsymbol{a}_{T-1}, \boldsymbol{s}_T\}$ is expressed by ι.

The determined mapping from a state space $\mathcal{S}$ to an action space $\mathcal{A}$ is defined as a policy π: $\mathcal{S} \to \mathcal{A}$. The objective of the agent is to find an optimal policy that maximizes the expected total reward $J(\pi)$ as:

$$\max_\pi J(\pi) = \mathop{\mathbb{E}}_{\iota \sim \pi} \left[\sum_{t=0}^{T} \gamma^t \mathcal{R}_t \right]. \quad (11)$$

To calculate the expected reward of a single action, an important action-value function $Q^\pi(\boldsymbol{s}, \boldsymbol{a})$ is defined as:

$$Q^\pi(\boldsymbol{s}, \boldsymbol{a}) = \mathop{\mathbb{E}}_{\iota \sim \pi} \left[\sum_{t=0}^{T} \gamma^t \mathcal{R}_t | \boldsymbol{s}_0 = \boldsymbol{s}, \boldsymbol{a}_0 = \boldsymbol{a} \right], \quad (12)$$

where $Q^\pi(\boldsymbol{s}, \boldsymbol{a})$ denotes the expected reward taking action $\boldsymbol{a}$ at state $\boldsymbol{s}$ following policy π.

B. Formulating Capacity Offering Problem as an MDP

In the DCS capacity offering problem, the system bidder is treated as the agent who interacts with the regulation market. The action determined by the agent at each time step is offering the regulation capacity of the next operating hour:

$$\boldsymbol{a}_t = \left[C_t^{\mathrm{ch}} \right], \quad \forall t \in \mathcal{T}, \quad (13)$$

where the scale of the action space $|\mathcal{A}|$ is 1. The state of the DCS environment considers the information of both the market signal and buildings' indoor temperature, which is useful for prediction. The state is defined as:

$$\boldsymbol{s}_t = \left[t, p_{t-1}, P_{t-1}^{\text{ch}}, \mathbb{E}[\Delta T_t], \mathbb{E}[\sigma_{t-1}^{\text{s}}], f(\sigma_{t-1}^{\text{s}})\right]^{\intercal}, \forall t \in \mathcal{T}, \tag{14}$$

$$\Delta T_t = \sum\nolimits_{i \in \mathcal{I}} \Delta T_{i,t}, \quad \forall t \in \mathcal{T}, \tag{15}$$

$$f(\sigma_{t-1}^{\text{s}}) = \max_x \left| \int_0^x \sigma_x^{\text{s}} \mathrm{d}x \right|, \quad \forall t \in \mathcal{T}, \tag{16}$$

where t^{H} is the hour of the time t; ΔT represents the average temperature deviation of all buildings' indoor temperature; σ_t^{s} is the market signal. $\mathbb{E}[\Delta T_{i,t-1}]$ reflects buildings' available regulation range for the next time step. $\mathbb{E}[\sigma_{t-1}^{\text{s}}]$ is the average market signal in the last operating hour. The function $f(\sigma_{t-1}^{\text{s}})$ represents the maximum accumulative adjustment in the power consumption during last operating hour.

The DCS bidder aims at increasing the cumulative revenue from regulation markets while keeping the buildings comfortable. In addition, to be eligible to participate in the market, the regulation performance score s^{M} should be above the minimum requirement $\underline{s}^{\text{M}}$. The reward function is defined as:

$$\mathcal{R}_t = \kappa_t C_t^{\text{ch}} s_t^{\text{M}} p_t - \beta(1-\kappa_t)\frac{C_t^{\text{ch}}}{s_t^{\text{M}}}, \quad \forall t \in \mathcal{T}, \tag{17}$$

$$\kappa_t = \left(1 + \frac{s_t^{\text{M}} - \underline{s}^{\text{M}}}{|s_t^{\text{M}} - \underline{s}^{\text{M}}|}\right)/2, \quad \forall t \in \mathcal{T}, \tag{18}$$

where κ_t is designed to judge whether s_t^{M} satisfies the required value $\underline{s}^{\text{M}}$ (i.e., when $s_t^{\text{M}} \geq \underline{s}^{\text{M}}$, $\kappa_t = 1$; Otherwise, $\kappa_t = 0$). Symbol β is the penalty coefficient of the score violation.

It is noted that when κ_t equals to 1, the reward function $\mathcal{R}_t$ represents the cumulative revenue in the market. In contrast, when κ_t equals to 0, the reward function $\mathcal{R}_t$ is the negative of the $\frac{C_t^{\text{ch}}}{s_t^{\text{M}}}$, which guides the agent to reduce the capacity offer C_t^{ch} or improve the regulation performance score s_t^{M}.

IV. Case Study

A. Test System

The test system is modelled based on a realistic DCS in Hengqin, China. The energy station supplies chilled water for 12 buildings, with a cooling capacity 144 MW. Based on the DCS's technical guidelines, the coefficient of performance (COP) and heat transfer coefficient (k_i) are 5 and 4.5 kW/(m^2·°C), respectively. The efficiency in two water loops ($\eta_i^{\text{I}}, \eta_i^{\text{II}}, \forall i \in \mathcal{I}$) are assumed to be 90%. The supply chilled water temperature $T_t^{\text{ch,s}}$ is a constant with 3°C, $\forall t$. The heat capacity (c^{w}) is 4.2 kJ/(kg·°C), according to the national standard in China. Buildings' thermal capacity (D_i) and thermal resistance (R_i) are both proportional to their floor areas ranging from 100,000 m^2 to 300,000 m^2. The indoor set temperature ($T_{i,\tau}^{\text{set}}$) is distributed in 20∼23°C. The comfortable temperature range is assumed $T_{i,\tau}^{\text{set}} \pm 1$°C when DCS provides regulation services (i.e., $|\Delta T_{i,\tau}| \leq 1$).

The outdoor temperature T_t^{O} is collected from the realistic data in Zhuhai, China from January 1, 2020 to December 30, 2020. The market regulation instruction (σ_t^{s}) adopts the realistic RegA signal from PJM in year 2020, as shown in Fig. 3. In PJM, the capacity offer of the demand resource (C_t^{ch}) should be larger than 1MW, and the performance score (s_t^{M}) should be higher than 0.75. The score is calculated based on the market rule. In addition, the hourly base power (P_t^{base}) of each participator is defined by PJM, using the participator's average power consumption of the same hour in last five days.

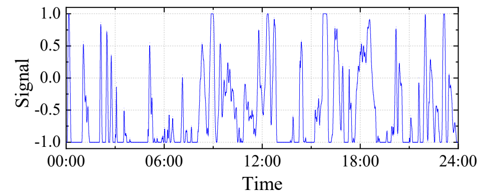

Fig. 3. The regulation signal from PJM in one day.

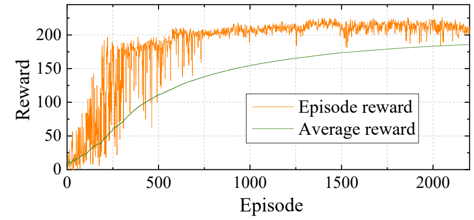

Fig. 4. The reward during the training process.

B. Training and Online Control

The objective of our agent is to give an hour-ahead capacity offer for continuous 24 hours to obtain more cumulative revenue in Eq. (9). The key parameters of the DRL algorithm are set as Table I. Both the actor and critic networks are composed of one input layer, two hidden layers (256x128) and one output layer. The simulation is implemented using Python with an Intel core i7 CPU @3.0 GHz and 16GB memory.

The agent's reward during the training process is illustrated in Fig. 4, where the higher reward means the better decision

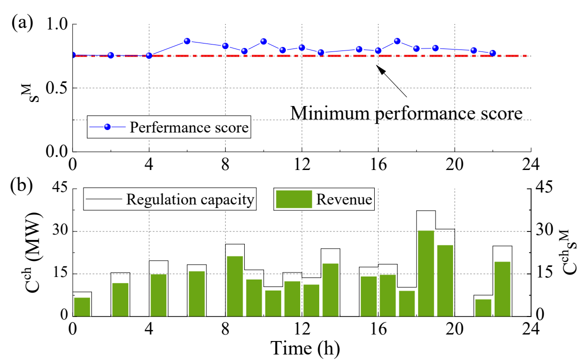

Fig. 5. Capacity offer and performance score with the proposed DRL strategy.

TABLE I
PARAMETERS OF THE DRL CONTROLLER.

Symbols	Definitions	Values
β	Violation penalty coefficient	0.1
γ	Discount factor	0.9
K	Mini batch size	100
u	Mean of exploration noise	0
b	Standard deviation of exploration noise	0.3

978-1-6654-0823-3/22/$31.00 ©2022 IEEE

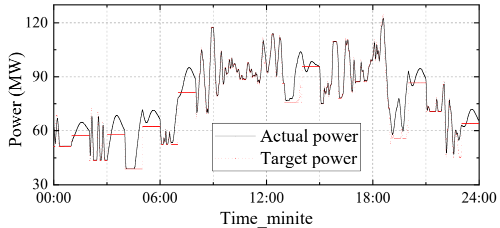

Fig. 6. Power consumption with regulation services.

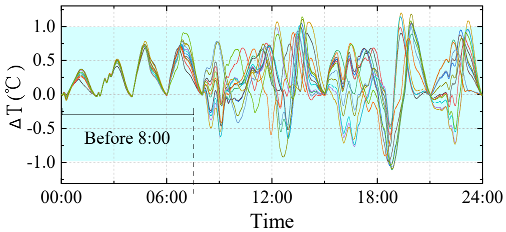

Fig. 7. Building temperature deviation with regulation services.

made by the agent. To improve the training efficiency, the agent ends one episode when its score is lower than the required value ($\underline{s}^{\mathrm{M}}$=0.75). According to the curve of the episode reward, the reward converges to a stable and optimal value after 500 episodes.

After training, the well-trained agent is applied in the online controlling of DCS to offer hour-ahead capacities intelligently for continuous 24 hours. The agent's hourly capacity offer and performance score are shown in Fig. 5. As shown in Fig. 5(a), all performance scores are larger than the required minimum score 0.75, and satisfy the market requirement during the regulation process. It guarantees the DCS's qualification to be eligible for the market in the long term. It can be seen from Fig. 5(b) that the DCS provides regulation services (i.e., C_t^{ch}>0) for 17 hours, which means the DCS does not participate in the market during the other 7 hours with $C_t^{\mathrm{ch}} = 0$. Because sometimes the continuous regulation may lead to a bad regulation performance and obtain a low score, it is better to quit the market for an hour temporally and be well-prepared for the next hour by recovering buildings' indoor temperature.

C. Control Results Analysis

Based on the capacity offer strategy, Fig. 6 shows the DCS's power consumption in a day. The black curve represents the system's real power consumption and the red curve shows the target power consumption. When DCS participates in the market, the hourly target power is related to this hour's base power, regulation signal and corresponding capacity offer. When DCS quits the market ($C_t^{\mathrm{ch}} = 0$), the target power equals to the base power. It can be seen that two power curves fits well when the DCS provides regulation services.

Fig. 7 shows all buildings' temperature deviations from their set values ($\Delta T_{i,\tau}$) in a day. The blue area in the figure represents the comfortable deviation range, where $\Delta T_{i,\tau} \in [-1, 1]$. It can be seen that the temperature deviations are always smaller than 1°C (i.e., within the blue area), which satisfies the buildings' comfortable requirements. Moreover, when buildings' temperature deviation is uncomfortable (i.e., out of blue area), the building will quit AI controller and ignore the regulation signal to recover its own indoor temperature. Furthermore, it can be seen from Fig. 7 that the indoor temperature deviations fluctuate less before 8:00 am. The first reason is that the ambient temperature is relatively stable compared with the following hours. The second reason is that each building's indoor heat loads (e.g., human behaviors and ventilation rates) are almost constant at night, which causes the similar indoor temperature deviations.

V. Conclusion

This paper proposes an hour-ahead capacity offering strategy for the DCS to provide regulation services. To cope with the uncertain market signal and complex physical model, the proposed strategy adopts a model-free DRL algorithm to the continuous decision-making problem. The case study results show that the well-trained agent can determine a proper capacity offer to maximize the total cumulative revenue while satisfying the performance requirement from the market and the comfort requirement from buildings.

References

[1] A. Q. Al-Shetwi, M. Hannan, K. P. Jern, M. Mansur, and T. Mahlia, "Grid-connected renewable energy sources: Review of the recent integration requirements and control methods," *J. Clean. Prod.*, vol. 253, p. 119831, 2020.

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

[3] Y. Li, Y. Rezgui, and H. Zhu, "District heating and cooling optimization and enhancement – towards integration of renewables, storage and smart grid," *Renew. Sust. Energ. Rev.*, vol. 72, pp. 281–294, 2017.

[4] B. Mohandes, M. S. E. Moursi, N. Hatziargyriou, and S. E. Khatib, "A review of power system flexibility with high penetration of renewables," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 3140–3155, 2019.

[5] M. Cai, M. Pipattanasomporn, and S. Rahman, "Day-ahead building-level load forecasts using deep learning vs. traditional time-series techniques," *Appl. Energy*, vol. 236, pp. 1078–1088, 2019.

[6] H. Zhang, Z. Hu, E. Munsing, S. J. Moura, and Y. Song, "Data-driven chance-constrained regulation capacity offering for distributed energy resources," *IEEE Trans. Smart Grid*, vol. 10, no. 3, pp. 2713–2725, 2019.

[7] Y. Xiao, Q. Su, F. S. S. Bresler, R. Carroll, J. R. Schmitt, and M. Olaleye, "Performance-based regulation model in pjm wholesale markets," in *2014 IEEE PES General Meeting*, pp. 1–5, 2014.

[8] L. Fabietti, T. T. Gorecki, F. A. Qureshi, A. Bitlislioğlu, I. Lymperopoulos, and C. N. Jones, "Experimental implementation of frequency regulation services using commercial buildings," *IEEE Trans. Smart Grid*, vol. 9, no. 3, pp. 1657–1666, 2018.

[9] D. Xie, H. Hui, Y. Ding, and Z. Lin, "Operating reserve capacity evaluation of aggregated heterogeneous tcls with price signals," *Appl. Energy*, vol. 216, pp. 338–347, 2018.

[10] J. Cai and J. E. Braun, "A regulation capacity reset strategy for hvac frequency regulation control," *Energy Build.*, vol. 185, pp. 272–286, 2019.

[11] M. B. Anwar and M. O'Malley, "Strategic participation of residential thermal demand response in energy and capacity markets," *IEEE Trans. Smart Grid*, pp. 1–1, 2021.

[12] X. Li, W. Li, R. Zhang, T. Jiang, H. Chen, and G. Li, "Collaborative scheduling and flexibility assessment of integrated electricity and district heating systems utilizing thermal inertia of district heating network and aggregated buildings," *Appl. Energy*, vol. 258, p. 114021, 2020.

[13] Z. Zhang, D. Zhang, and R. C. Qiu, "Deep reinforcement learning for power system applications: An overview," *CSEE J. Power Energy Syst.*, vol. 6, no. 1, pp. 213–225, 2020.

[14] PJM Manual 11: Energy Ancillary Services Market Operations (Revision: 177). [Online]. Available: https://www.pjm.com/-/media/documents/manuals/m11.ashx. [Accessed: Nov. 01, 2021].

978-1-6654-0823-3/22/$31.00 ©2022 IEEE