文章编号:1674-7607(2026)02-0193-09 DOI:10.19805/j.cnki.jcspe.2026.240740

基于深度策略梯度的综合能源系统优化调度研究

王臣宇[1], 袁树福[1], 王志霄[1], 丁 吉[2], 孙 立[1]
(1. 东南大学 能源与环境学院,大型发电装备安全运行与智能测控国家工程研究中心,江苏南京 210096
2. 苏州同元软控信息技术有限公司,江苏苏州 215000)

摘 要:随着可再生能源的广泛应用,综合能源系统的运行复杂性和不确定性增大,传统的调度方法难以高效应对多能流、强耦合等问题。因此,构建了一个涵盖光伏、风力、氢气、蓄电池、热泵和燃料电池等设备的综合能源系统模型,结合深度强化学习深度策略梯度(DDPG)算法,以实现在不确定性工况下的能源调度优化。结果表明:DDPG 算法能够有效解决多能流协同调度问题,在不同负荷需求和能源供应条件下,实现了电、热、冷、氢的供需平衡和能源高效利用;通过训练,系统在面对不同的不确定性工况下展现出良好的调度能力和适应性,能够灵活应对能源的波动。

关键词:综合能源系统;深度确定性策略梯度算法;深度强化学习;能量调度;多能耦合
中图分类号:TM73 **文献标志码**:A **学科分类号**:480.80

Research on Optimal Scheduling of Integrated Energy Systems Based on Deep Deterministic Policy Gradient

WHANG Chenyu[1], *YUAN Shufu*[1], *WANG Zhixiao*[1], *DING Ji*[2], *SUN Li*[1]
(1. National Engineering Research Center of Power Generation Control and Safety, School of Energy and Environment, Southeast University, Nanjing 210096, China;
2. Suzhou Tongyuan Software & Control Technology Co., Suzhou 215000, Jiangsu Province, China)

Abstract: With the widespread application of renewable energy, the operational complexity and uncertainty of integrated energy systems have increased significantly. Traditional scheduling methods are insufficient to efficiently address issues such as multi-energy flows and strong coupling. Therefore, an integrated energy system model encompassing photovoltaic panels, wind turbines, hydrogen storage tanks, batteries, heat pumps, and fuel cells was developed. By combining the deep deterministic policy gradient (DDPG) algorithm, the energy scheduling optimization under uncertain conditions was achieved. Results show that the DDPG algorithm can effectively solve the problem of multi-energy flow coordinated scheduling, achieving supply-demand balance and efficient utilization of electricity, heat, cooling, and hydrogen under different load demands and energy supply conditions. Moreover, after training, the system exhibits excellent scheduling performance and adaptability when confronted with diverse uncertain scenarios, enabling flexible responses to energy fluctuations.

Key words: integrated energy system; deep deterministic policy gradient algorithm; deep reinforcement

收稿日期:2024-11-25 **修订日期**:2025-02-19
基金项目:江苏省科技厅资助科技项目(BK20240073,BA2023005)
作者简介:王臣宇(2001—),男,江苏南京人,硕士研究生,研究方向为综合能源调度。
孙 立(通信作者),男,教授,博士,E-mail:sunli12@seu.edu.cn。

learning; energy scheduling; multi-energy coupling

能源是人类社会发展的物质基础,随着光能、风能、氢能等可再生能源技术的不断进步,能源的多样性需求得到了有效满足。综合能源系统通过协调电、热、冷等多种能源形式,实现系统层面的互补与优化,从而提升能源利用效率[1]。

然而随着全球环保压力的加剧,中国长期依赖煤炭的能源结构已无法兼顾能源需求和环保要求。自 2020 年"双碳政策"提出以来,中国能源的绿色低碳转型已取得显著成效[2],新能源和清洁能源的快速发展为中国能源结构优化提供了重要支撑。2021 年,中国新能源发电量首次突破 1×10^4 亿 kW·h,展现出强劲的发展势头。但目前工业蒸汽主要依赖燃煤/燃气锅炉供热,导致污染排放高、余热资源浪费严重,尤其是 10 t 以下的小型燃煤锅炉已被环保政策禁止使用[3]。因此,发展基于蒸汽蓄热技术的电热耦合协同技术,特别是光伏绿电、谷电、余热等能源综合利用技术,具有重要的现实意义。

然而,综合能源系统中电、热、冷、氢等多能源协同调度的精准控制面临巨大的挑战。这种复杂性带来了较高的运行风险,且现有研究大多集中于基于数学模型的系统优化,忽略了系统的响应快速性、稳定性和抗震荡性等问题,难以解决系统运行的不稳定性[4-6]。

另一方面,基于深度学习的强化学习算法,如深度确定性策略梯度(DDPG)算法,已被证明在能源系统调度中具有很好的潜力,能有效提升系统的智能化和高效化。然而,当前算法在数据获取、模型可解释性和训练响应速度方面仍存在一定不足[7]。

本文基于 DDPG 算法,研究了电、热、冷、氢综合能源系统的优化调度问题,通过深度强化学习实现多能协同的高效调度,克服了现有方法在系统复杂性和不确定性方面的不足。与传统的数学建模方法不同,本文方法不仅关注系统的经济性优化,还强调了系统响应的快速性和稳定性,为解决综合能源系统的运行不稳定性问题提供了一种新思路。

1 综合能源系统的构建

电、热、冷、氢综合能源系统示意图如图 1 所示,主要包括光伏组件、风力发电机、电网、质子交换膜(PEM)电解槽、PEM 燃料电池、蓄电池、储氢罐、空气源热泵、电加热锅炉、蒸汽蓄热器、制冷机组、冰蓄冷装置等。

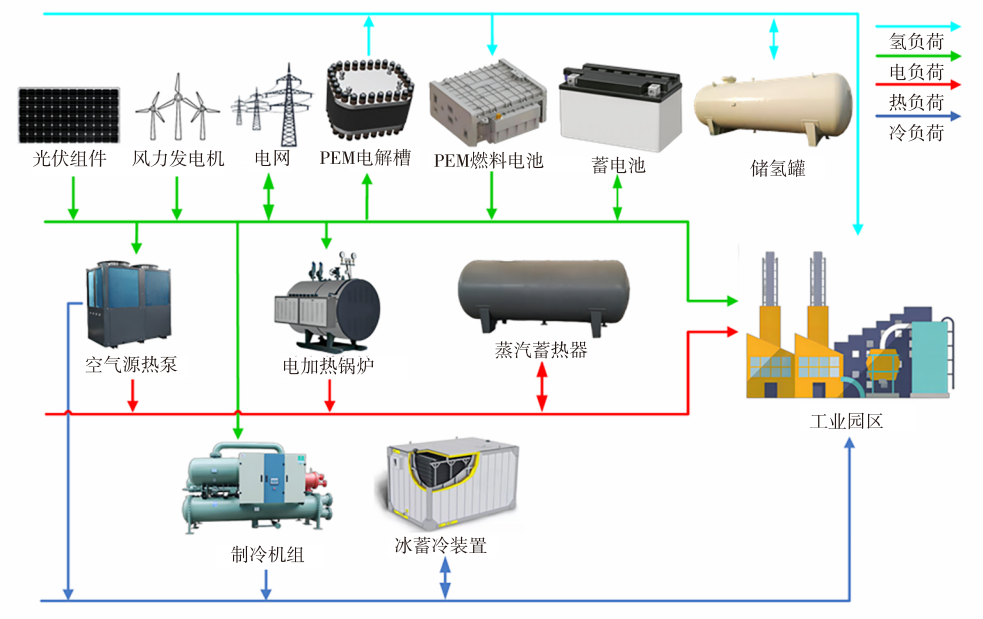


图 1 电、热、冷、氢综合能源系统示意图

Fig. 1 Diagram of an integrated energy system for electricity, heat, cooling and hydrogen

1.1 系统组成

1.1.1 光伏组件

选择光伏组件[8]作为部分电力来源,其输出电功率 $P_{E,PV}$ 可简化为

$$P_{E,PV}=\eta_{PV}\eta_{inv}A_{PV}I_0 \tag{1}$$

式中:η_{PV} 为光伏效率;η_{inv} 为逆变器效率;A_{PV} 为有效光

照面积，m^2；I_0 为太阳辐照度，W/m^2。

1.1.2　风力发电机

当风力机运行时，发电功率的方程[9]为

$$P_{w,WT}=\frac{1}{2}\rho\pi R^2V^3C_p(\beta,\lambda) \tag{2}$$

式中：$P_{w,WT}$为风功率，kW；ρ 为空气密度，kg/m^3；R 为风力机叶轮半径，m；C_p 为风能利用系数，根据贝兹理论得到最大风能利用率 $C_{pmax}\approx0.593$；λ 为叶尖速比；β 为桨距角；V 为风速，m/s。

风力发电机发电功率为

$$P_{E,WT,t}=\begin{cases}0, & 0\leqslant V<V_{in},V>V_{out}\\ P_{w,WT}, & V_{in}\leqslant V<V_{out}\end{cases} \tag{3}$$

式中：$P_{E,WT,t}$为 t 时刻风力发电机发电功率；V_{in}为启动风速；V_{out}为停机风速。

1.1.3　质子交换膜电解槽

本文选取质子交换膜电解槽[10]作为制氢设备，其具有高电解效率、快速响应和设计紧凑的优势。制氢速率 H_{ET}为

$$H_{ET}=\frac{\eta_F n_{ET} I_{ET}}{2F} \tag{4}$$

式中：η_F 为法拉第效率；n_{ET}为电解槽数目；I_{ET}为工作电流，A；F 为法拉第常数，C/mol。

1.1.4　质子交换膜燃料电池

质子交换膜燃料电池[11]的氢气消耗量通过法拉第定律计算：

$$H_{FC}=\frac{I_{FC}n_{FC}}{2F} \tag{5}$$

式中：H_{FC}为耗氢速率，mol/s；I_{FC}为工作电流，A；n_{FC}为电池数目。

燃料电池反应速率可以由电流密度 j 表示：

$$j=\frac{I_{FC}}{A_{cell}} \tag{6}$$

其中，A_{cell}为电池单元的反应面积，m^2。

燃料电池的热功率输出 $P_{H,FC}$表达式为

$$P_{H,FC}=P_{E,FC}\frac{\eta_T}{\eta_E} \tag{7}$$

式中：$P_{E,FC}$为燃料电池电功率，kW；η_T 为燃料电池热效率；η_E 为燃料电池电效率。

1.1.5　蓄电池

电池通常以电池组的模式运行[12]，其总功率可表示为

$$P_{E,B}=P_{E,b}\times n_1\times n_2 \tag{8}$$

式中：$P_{E,B}$为电池组总功率输出，kW；$P_{E,b}$为电池单元功率，kW；n_1 和 n_2 分别为电池行数和列数。

电池单元的电流 I_b 如下：

$$I_b=\frac{V_{oc}-\sqrt{V_{oc}^2-4R_bP_{E,b}}}{2R_b} \tag{9}$$

其中，开环电压 V_{oc}和电池单元内阻 R_b 均可视为定值。

电池单元的动态特性 S_{SOC}可表示为

$$S_{SOC_{t+1}}=S_{SOC_t}-\frac{I_{b,t}}{Q_B}\Delta T \tag{10}$$

式中：t 为调度时刻，$0\leqslant t\leqslant T$，T 为调度周期；ΔT 为调度步长，为 1 h；Q_B 为电池单元容量，Ah；$I_{b,t}$为 t 时刻电流，A，充电为负，放电为正。

1.1.6　储氢罐

本系统使用高压储氢罐（hydrogen storage tank，HST）模型[13]来储存电解槽制取的氢气，并且为燃料电池发电提供氢气来源。储氢罐的罐内压力状态 S_{TP}随着氢气的存入与释放而动态变化，可表示为

$$S_{TP_{t+1}}=S_{TP_t}-z_t\frac{RT_H}{V_H}H_{HST,t} \tag{11}$$

式中：$S_{TP_{H,t}}$为储氢罐在 t 时刻的罐内压力状态；$H_{HST,t}$为储氢罐在 t 时刻的氢气流速，mol/s；V_H 为储氢罐容积，L；T_H 为工作温度，K；R 为通用气体常数，J/(mol·K)；z_t 为氢气的压缩系数，可基于 NIST 数据并根据 Lemmon 方程计算。

$$z_t(S_{TP_t})=1+\sum_{i=1}^{9}a_i\left(\frac{T_0}{T_H}\right)^{b_i}(S_{TP_t})^{c_i} \tag{12}$$

其中，a_i、b_i、c_i 均为常数，具体取值见表 1。标准参考温度 $T_0=100$ K，$R=8.314\ 5$ J/(mol·K)。

表 1　a_i、b_i、c_i 取值表

Tab. 1　Values of a_i, b_i and c_i

i	a_i	b_i	c_i
1	0.058 846 000 0	1.325	1.00
2	−0.061 361 110 0	1.870	1.00
3	−0.002 650 473 0	2.500	2.00
4	0.002 731 125 0	2.800	2.00
5	0.001 802 374 0	2.938	2.42
6	−0.001 215 070 7	3.140	2.63
7	$0.958\ 842\ 000\ 0\times10^{-4}$	3.370	3.00
8	$-0.110\ 904\ 000\ 0\times10^{-6}$	3.750	4.00
9	$0.126\ 440\ 300\ 0\times10^{-9}$	4.000	5.00

1.1.7　热泵模型

空气源热泵（heat pump，HP）可以在高负荷下提供峰值热功率[14]。热泵的性能系数 η_{COP}与温差相关[15]，其经验关系可拟合为

$$\eta_{COP}=h_1-h_2\Delta T_{HP}+h_3\Delta T_{HP}^2,\quad 15\leqslant\Delta T_{HP}\leqslant60 \tag{13}$$

式中：ΔT_{HP} 为热泵制冷/制热目标温度 T_D/T_H 与当前环境温度 T_a 的差值；h_1、h_2 和 h_3 为拟合系数。

热泵的供热功率可以表示为

$$P_{H,HP}=P_{E,HP}\eta_{COP_H} \tag{14}$$

式中：$P_{H,HP}$ 为热泵热功率，kW；$P_{E,HP}$ 为对应电功率；η_{COP_H} 为供热性能系数。

热泵的供冷功率可以表示为

$$P_{C,HP}=P_{E,HP}\eta_{COP_C} \tag{15}$$

式中：$P_{C,HP}$ 为热泵制冷功率，kW；η_{COP_C} 为供冷性能系数。

1.1.8 电加热锅炉

本系统配置了电加热锅炉(electric-heated boiler，EB)作为将电能转换为热能的热力设备。电加热锅炉不会产生污染和噪声，且占地面积较小，其热效率高达 98%以上，具有良好的能量转换性能。

电加热锅炉的供热功率可表示为

$$P_{H,EB,t}=\eta_{EB}P_{E,EB,t} \tag{16}$$

式中：$P_{H,EB,t}$ 为电加热锅炉热功率，kW；$P_{E,EB,t}$ 为电加热锅炉耗电功率，kW；η_{EB} 为电加热锅炉的转换效率。

1.1.9 蒸汽蓄热器

为描述蒸汽蓄热器[16]的动态特性，引入与电池类似的储热状态，定义为

$$S_{HSD_{t+1}}=S_{HSD_t}-\frac{P_{H,TES,t}}{H_c}\Delta T \tag{17}$$

式中：$P_{H,TES,t}$ 为 t 时刻热泵充热(负)或放热(正)功率，kW；H_c 为热泵最大储热量，kW·h。S_{HSD} 被限制在 0～1 之间。

1.1.10 制冷机组

本系统选取的制冷机组(refrigeration air condition chiller，RAC)可以将电能转换为制冷量。制冷机组具有较高的能效比，可以实现高效的能量转化。

制冷机组的制冷功率可表示为

$$P_{C,RAC,t}=\eta_{COP_{RAC}}P_{E,RAC,t} \tag{18}$$

式中：$P_{C,RAC,t}$ 为制冷功率，kW；$P_{E,RAC,t}$ 为制冷机组耗电功率，kW；$\eta_{COP_{RAC}}$ 为制冷机组的能效比。

1.1.11 冰蓄冷装置

为描述冰蓄冷装置[17](cool energy storage，CES)的动态特性，引入与蒸汽蓄热器类似的储冷状态，定义为

$$S_{CSD_{t+1}}=S_{CSD_t}-\frac{P_{C,CES,t}}{C_c}\Delta T \tag{19}$$

式中：$P_{C,CES,t}$ 为 t 时刻冰蓄冷装置的充冷(负)或放冷(正)功率，kW；C_c 为冰蓄冷装置最大蓄冷量，kW·h。S_{CSD} 被限制在 0～1。

1.2 运行流程及策略

该综合能源系统主要包括冷、热、电、氢 4 种能源。其中，提供电能的设备主要包括光伏组件、风力发电机、PEM 燃料电池和电网，蓄电池进行电能的存储。提供氢能的设备主要包括 PEM 电解槽，储氢罐进行氢能的存储。提供热能的设备主要包括空气源热泵、电加热锅炉，蒸汽蓄热器进行热能的存储。提供冷能的设备主要包括空气源热泵、制冷机组，冰蓄冷设备进行冷能的存储。该综合能源系统运行策略如下：(1)系统优先消纳光伏组件、风力发电机产出的电能；(2)当冷、热、电、氢 4 种能源多于工业园区的相应负荷需求时，将会以相应形式的能量储存在储能设备中；(3)空气源热泵将同时产出热能和冷能 2 种形式的能量；(4)储能设备的储能量将控制在相应的容量范围内，可通过换热和相应的设备进行转化；(5)该工业园区用热形式主要是热水和暖气(热空气)，统一为热负荷的形式，多余的热能作为蓄热水的预热，通过电加热锅炉加热成高温蒸汽储存在蒸汽蓄热器中。

1.3 优化调度构建

本系统具有良好的调度特性，蓄电池、蒸汽蓄热器、冰蓄冷装置和储氢罐均有工作状态和非工作状态，工作状态既有充能过程也有释能过程，可以实现电、热、冷、氢的供需平衡。接入电网可以在用电高峰时购电，在电产能过剩时向电网售电。由于本文主要研究源荷不确定下的能量调度，所以本系统不考虑配置成本。因此，调度优化的目标函数以及相应约束构建如下。

1.3.1 目标函数

目标函数主要包括系统的运行成本 $C_{OM,t}$、从电网购电成本 $C_{G,t}$ 以及储能设备状态 $C_{S,T}$：

$$\begin{aligned}J=&\min\sum_{t=1}^{T}(C_{OM,t}+C_{G,t})+\sum_{i=1}k_iC_{S,T}=\\&\min\sum_{t=1}^{T}(\sum_{i=1}^{N}c_iP_{OM,i,t}+p_tP_{E,G,t})+\\&(k_1\mid S_{SOC_T}-S_{SOC_0}\mid+k_2\mid S_{TP_T}-S_{TP_0}\mid+\\&k_3\mid S_{HSD_T}-S_{HSD_0}\mid+k_4\mid S_{CSD_T}-S_{CSD_0}\mid)\end{aligned} \tag{20}$$

式中：J 为目标函数；c_i 和 $P_{OM,i,t}$ 分别为系统中设备的度电运行成本和 t 时刻功率；N 为设备数目；p_t 为电网购电成本；k_1、k_2、k_3 和 k_4 均为权重系数；下标 0 表示初始时刻。

1.3.2 运行约束

$$\begin{aligned}&P_{E,PV,t}+P_{E,WT,t}+P_{E,G,t}+P_{E,FC,t}+P_{E,B,t}=\\&P_{E,L,t}+P_{E,RAC,t}+P_{E,HP,t}+P_{E,EB,t}+P_{E,ET,t}\end{aligned} \tag{21}$$

$$P_{\mathrm{H,HP},t}+P_{\mathrm{H,EB},t}+P_{\mathrm{H,TES},t}=P_{\mathrm{H,L},t} \tag{22}$$

$$P_{\mathrm{C,HP},t}+P_{\mathrm{C,RAC},t}+P_{\mathrm{C,CES},t}=P_{\mathrm{C,L},t} \tag{23}$$

$$H_{\mathrm{ET},t}+H_{\mathrm{HST},t}=H_{\mathrm{FC},t}+H_{\mathrm{L},t} \tag{24}$$

$$P_{\mathrm{E,RAC,min}}\leqslant P_{\mathrm{E,RAC},t}\leqslant P_{\mathrm{E,RAC,max}} \tag{25}$$

$$P_{\mathrm{E,FC,min}}\leqslant P_{\mathrm{E,FC},t}\leqslant P_{\mathrm{E,FC,max}} \tag{26}$$

$$P_{\mathrm{E,HP,min}}\leqslant P_{\mathrm{E,HP},t}\leqslant P_{\mathrm{E,HP,max}} \tag{27}$$

$$P_{\mathrm{E,EB,min}}\leqslant P_{\mathrm{E,EB},t}\leqslant P_{\mathrm{E,EB,max}} \tag{28}$$

$$P_{\mathrm{E,ET,min}}\leqslant P_{\mathrm{E,ET},t}\leqslant P_{\mathrm{E,ET,max}} \tag{29}$$

$$H_{\mathrm{HST,min}}\leqslant H_{\mathrm{HST},t}\leqslant H_{\mathrm{HST,max}} \tag{30}$$

$$\begin{cases}0\leqslant C_{\mathrm{in}}\leqslant C_{\mathrm{in,max}}\\0\leqslant C_{\mathrm{out}}\leqslant C_{\mathrm{out,max}}\end{cases} \tag{31}$$

$$S_{\mathrm{SOC_{min}}}\leqslant S_{\mathrm{SOC}_t}\leqslant S_{\mathrm{SOC_{max}}} \tag{32}$$

$$S_{\mathrm{TP_{min}}}\leqslant S_{\mathrm{TP}_t}\leqslant S_{\mathrm{TP_{max}}} \tag{33}$$

$$S_{\mathrm{HSD_{min}}}\leqslant S_{\mathrm{HSD}_t}\leqslant S_{\mathrm{HSD_{max}}} \tag{34}$$

$$S_{\mathrm{CSD_{min}}}\leqslant S_{\mathrm{CSD}_t}\leqslant S_{\mathrm{CSD_{max}}} \tag{35}$$

$$\begin{cases}S_{\mathrm{SOC}_0}=0.5(S_{\mathrm{SOC_{min}}}+S_{\mathrm{SOC_{max}}})\\S_{\mathrm{HSD}_0}=0.5(S_{\mathrm{HSD_{min}}}+S_{\mathrm{HSD_{max}}})\\S_{\mathrm{CSD}_0}=0.5(S_{\mathrm{CSD_{min}}}+S_{\mathrm{CSD_{max}}})\\S_{\mathrm{TP}_0}=0.5(S_{\mathrm{TP_{min}}}+S_{\mathrm{TP_{max}}})\end{cases} \tag{36}$$

其中，式(21)～式(24)分别表示为电能、热能、冷能和氢能平衡约束，kW；$P_{\mathrm{E,L},t}$、$P_{\mathrm{H,L},t}$、$P_{\mathrm{C,L},t}$和 $H_{\mathrm{L},t}$分别表示电、热、冷和氢负荷需求，kW；$P_{\mathrm{E,G},t}$为与电网交互电功率，kW，$P_{\mathrm{E,G},t}>0$ 表示从电网购电，$P_{\mathrm{E,G},t}<0$ 表示向电网售电；$P_{\mathrm{E,EF},t}$为电解槽耗电功率。式(25)～式(35)表示设备运行约束，其中式(31)表示充放电速率约束。式(36)为对应储能设备的初始储能状态。下标 t 表示 t 时刻，min 表示最小值，max 表示最大值。

2　DDPG 算法训练过程

2.1　探索噪声与马尔可夫决策过程

DDPG 是经典的深度强化学习算法，可用来解决综合能源系统连续控制问题。为了更好地探索决策过程中的未知环境，本文选择了适用于惯性系统的奥恩斯坦-乌伦贝克(Ornstein-Uhlenbeck，OU)噪声[18]，并设置噪声衰减以达到探索和利用之间的平衡，具体的数学表达式如下：

$$\mathrm{d}N_t=\theta(\mu-N_t)\mathrm{d}t+\sigma\mathrm{d}W_t \tag{37}$$

$$a_t=P(s_t\mid\theta_{\mathrm{p}})+N_t \tag{38}$$

式中：N_t 为噪声值；θ 为均值回归速率；μ 为均值；σ 为波动率；W_t 为布朗运动。式(38)表示噪声添加过程，a_t 为动作，$P(s_t\mid\theta_{\mathrm{p}})$为 DDPG 中的 Actor 网络，$\theta_{\mathrm{p}}$为网络参数，$s_t$ 为系统状态。OU 过程的时序相关性使系统能够在同方向上进行多步探索以积累训练经验，从而有效提高了惯性系统的探索效率。

强化学习任务通常表示为马尔科夫决策过程(Markov decision process，MDP)，以五元组(S，A，$\boldsymbol{P}$，R，γ)定义，其中 S 为状态空间，A 为动作空间，$\boldsymbol{P}$为状态转移矩阵，R 为奖励函数，γ 为折扣因子。由于本文采用的是无模型的 DDPG 算法，因此不涉及状态转移矩阵。基于电热协同系统特性，本文定义了具体的状态空间、动作空间和奖励函数。

2.1.1　状态空间

$$S=\{\sin t,\cos t,P_{\mathrm{E,PV},t},P_{\mathrm{E,WT},t},P_{\mathrm{E,L},t},P_{\mathrm{H,L},t},P_{\mathrm{C,L},t},H_{\mathrm{L},t},S_{\mathrm{SOC}_t},S_{\mathrm{TP}_t},S_{\mathrm{HSD}_t},S_{\mathrm{CSD}_t}\} \tag{39}$$

2.1.2　动作空间

$$A=\{H_{\mathrm{FC},t},P_{\mathrm{E,RAC},t},P_{\mathrm{E,HP},t},P_{\mathrm{E,EB},t},P_{\mathrm{E,ET},t}\} \tag{40}$$

2.1.3　奖励函数

$$R=-(C_{\mathrm{OM},t}+C_{\mathrm{G},t}+C_{\mathrm{S,T}}+l_1\mid P_{\mathrm{E,out},t}\mid+l_2\mid P_{\mathrm{H,out},t}\mid+l_3\mid P_{\mathrm{C,out},t}\mid+l_4\mid H_{\mathrm{out},t}\mid-l_5 f(p_t,S_{\mathrm{TP}_t},P_{\mathrm{E,ET},t})-l_6 f(S_{\mathrm{CSD}_t},P_{\mathrm{E,RAC},t})) \tag{41}$$

式中：$P_{\mathrm{E,out},t}$、$P_{\mathrm{H,out},t}$、$P_{\mathrm{C,out},t}$和 $H_{\mathrm{out},t}$分别为蓄电池、蒸汽蓄热器、冰蓄冷装置、储氢罐调度后对应功率即电、热、冷、氢的过剩(大于 0)或不足值(小于 0)；l_1～l_6 为根据训练经验和效果整定的权重；$f(p_t,S_{\mathrm{TP}_t},P_{\mathrm{E,ET},t})$和 $f(S_{\mathrm{CSD}_t},P_{\mathrm{E,RAC},t})$分别表示在电价低价时鼓励从电网购电进行电解水制氢弥补储氢罐氢气消耗以及进行制冷机组制冷弥补冰蓄冷设备冷量消耗。奖励函数中的变量均是决策变量。

2.2　DDPG 算法训练流程

DDPG 算法的训练流程如下：

(1) 神经网络初始化

根据状态空间和动作空间，分别建立策略网络(P 网络)和评价网络(Q 网络)，如图 2 所示。除输入层外，P 网络和 Q 网络的层数分别为 11 和 10。P 网络为 12 输入(状态)、5 输出(动作)，输出层的激活函数选择 Sigmoid 映射到区间 0～1 内，便于动作的放缩还原；Q 网络为 17 输入(状态＋动作)、1 输出

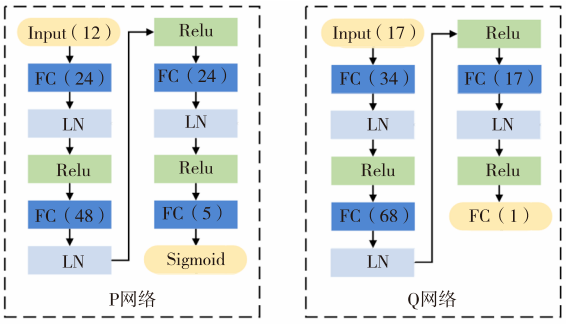


图 2　DDPG 策略网络和评价网络结构示意图

Fig. 2　Diagram of the structure of DDPG policy network and critic network

(奖励)。

(2) 算法参数设置

训练参数设置:设置网络的学习率 α_p 和 α_Q,目标网络更新系数 τ,调度周期 T,训练回合数 N_e,折扣因子 γ,经验回收区 N_{Memory},训练批量大小 $N_{Batchsize}$。

OU 噪声参数设置:值回归速率 θ,均值 μ,波动率 σ,最大波动 σ_{max},最小波动 σ_{min},噪声衰减周期 T_{Nd}。

奖励函数参数设置:系数 $k_1 \sim k_4$ 和 $l_1 \sim l_6$。

(3) 训练场景设置

每个训练回合开始前,在典型日场景和设备安全运行可行域内,对荷源数据的均值添加不超过 30%的正态随机波动,不断生成新的典型日训练场景至训练结束。

(4) MDP 训练循环

在每个训练开始前,获得系统的初始状态 S_0;对于每个时间步的状态 s_t,策略网络基于 $P(s_t|\theta_p|)$输出动作 a_t,系统执行动作后获得奖励 r_t,并得到下一步状态 s_{t+1};将训练经验(s_t, a_t, r_t, s_{t+1})保存至经验回放池,更新当前状态,重复 MDP 至末时刻,完成一回合的训练。

(5) 网络参数更新

当经验回放池里的经验参数满足批量大小时,每个时间步都会进行网络参数更新。更新策略网络和评价网络,软更新目标网络。

3 仿真结果与分析

3.1 训练场景和系统设置

本研究的训练回合数为 200,需要随机生成 200 组训练场景,电负荷、热负荷、冷负荷、氢负荷、光伏发电和风力发电的均值与波动范围如图 3 所示。其中,实线为负荷均值,阴影为数据可波动范围。光伏发电和风力发电的均值由太阳辐照以及风速计算得出,对应数据见图 4~图 5 所示。

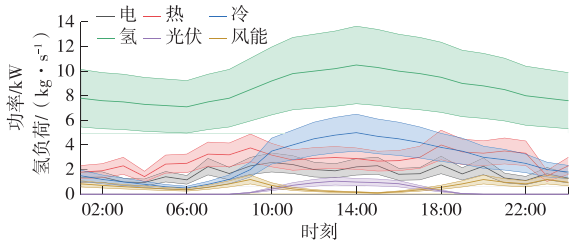


图 3 随机工况均值和波动范围

Fig. 3 Mean and range of random operating conditions

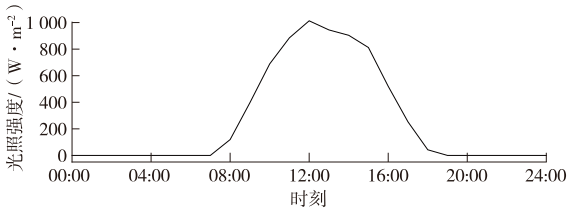


图 4 太阳辐照度逐时图

Fig. 4 Hourly solar irradiance chart

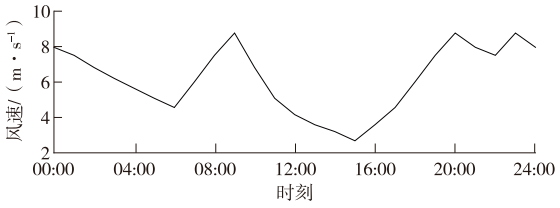


图 5 风速逐时图

Fig. 5 Hourly wind speed chart

根据典型日工况负荷数据[19],综合能源系统中各设备的规格和安全性约束见表 2。

表 2 设备主要参数[4,15]

Tab. 2 Main parameters of devices[4,15]

设备	规格	约束
光伏	$A_{PV}=10\ m^2$ $\eta_{PV}=11\%$ $\eta_{inv}=95\%$	
风力发电机	$R=15\ m$ $C_p=0.40$ $\rho=1.225\ kg/m^{-3}$	
电解槽	$n_{ET}=1$ $\eta_F=0.95$	$P_{E,ET,max}=10\ kW$, $P_{E,ET,min}=1\ kW$
燃料电池	$P_{E,FC,r}=5\ kW$	$P_{E,FC,max}=5\ kW$, $P_{E,FC,min}=0.5\ kW$
电池	$Q_b=100\ Ah$ $n_1=n_2=13$	$S_{SOC_{max}}=0.9$, $S_{SOC_{min}}=0.2$
储氢罐	$V_H=5\ m^3$ $T_H=223.15\ K$	$S_{TP_{max}}=7$, $S_{TP_{min}}=3.5$
热泵	$P_{E,HP,r}=10\ kW$ $\eta_{COP_H}=4.1$ $\eta_{COP_C}=3.2$	$P_{E,HP,max}=10\ kW$, $P_{E,HP,min}=0\ kW$
电加热锅炉	$\eta_{EB}=0.98$	$P_{E,EB,max}=5\ kW$, $P_{E,EB,min}=0\ kW$
蓄热罐	$H_c=5\ kW\cdot h$	$S_{HSD_{max}}=1$, $S_{HSD_{min}}=0$
制冷机组	$P_{E,RAC,r}=5\ kW$ $\eta_{COP_{RAC}}=3.5$	$P_{E,RAC,max}=5\ kW$, $P_{E,RAC,min}=0\ kW$
冰蓄冷装置	$C_c=4\ kW\cdot h$	$S_{CSD_{max}}=0.9$, $S_{CSD_{min}}=0.2$

所选取的电价为工商业用电(35 kV 两部制)。其中,低谷时段(00:00—08:00)为 0.268 7 元/(kW·h),平时段(11:00—17:00、22:00—24:00)为 0.642 0 元/(kW·h),高峰时段(08:00—11:00、17:00—22:

00)为 1.1040 元/(kW·h)。

3.2　算法实现与参数调整

DDPG 算法的主要参数见表 3。

表 3　DDPG 算法主要参数设置

Tab. 3　Main parameters of DDPG

参数	取值	参数	取值	参数	取值
α_P	0.001	σ_{max}	0.35	k_3	4.0
α_Q	0.001	T_{Nd}	4 800	k_4	4.0
τ	0.001	l_1	1.8	$P_{E,out,min}$	0
N_e	200	l_2	1.5	$P_{E,out,max}$	0.3
γ	0.99	l_3	1.5	$P_{H,out,min}$	0
N_{Memory}	4 000	l_4	1.3	$P_{H,out,max}$	0.6
$N_{Batchsize}$	96	l_5	1.0	$P_{C,out,min}$	0
μ	0	l_6	1.0	$P_{C,out,max}$	0.6
θ	0.15	k_1	5.0	$H_{out,min}$	0
σ_{min}	0.001	k_2	4.0	$H_{out,max}$	0.3

在调度研究中，燃料电池和热泵的功率通过 Sigmoid 函数的 0～1 输出值进行缩放，而电池和蓄热罐的储能状态更新通过当前时刻的电热功率平衡来计算。若无法保证功率平衡，则计算当前电热功率的过剩或不足。在参数调整方面，储能奖励项的参数可根据调度结果适当调整，但不宜过大，以免影响上一时刻功率平衡；能量过剩与不足的奖励值均要大于燃料成本奖励值以确保能量效率；由于热泵的效率较高，热过剩奖励项系数应稍大于热不足奖励项系数，而电能则相反，从而避免算法无法收敛。

3.3　典型日工况下实验结果与分析

DDPG 累计训练曲线如图 6 所示。由图 6 可知，前 25 回合累积奖励收敛速度较快，随着 OU 噪声的不断减小，在 173 回合收敛。

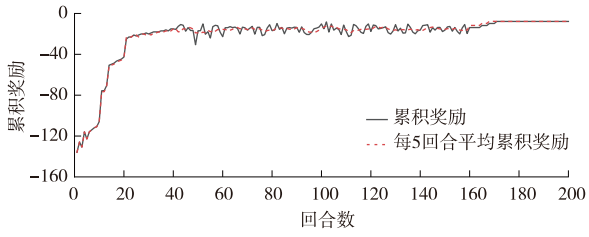


图 6　DDPG 累积奖励训练曲线

Fig. 6　Cumulative reward training curves of DDPG

由图 7 和图 8 可知，4 种储能设备调度结果均在合理工况内，表明该调度策略可以确保能量的有效储存和释放，消纳系统多余能量。其中，储热设备波动较为均匀，由于该典型日热负荷需求较均匀，储热设备对于热能的释放与消纳较少；蓄电池存储状态在 5 点前呈现明显下降，随后上升，这是因为 5 点前无太阳辐射，无法进行光伏发电，蓄电池释放电量较多；高压储氢罐在上午和下午时刻有明显的氢能释放现象，这是因为该时刻氢能负荷需求较大；蓄冷设备在中午时刻释放冷能较多，这是因为中午环境温度较高，对于冷能的需求较大。

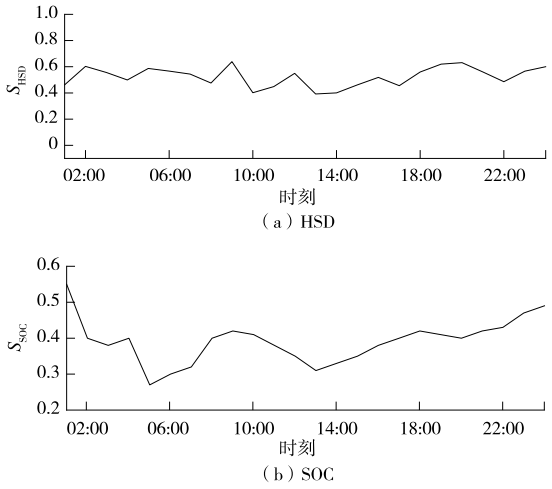


图 7　HSD 和 SOC 的调度结果

Fig. 7　Scheduling results of HSD and SOC

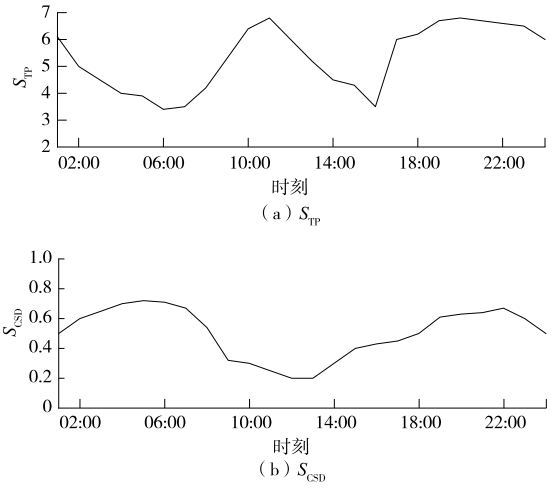


图 8　TP 和 CSD 的调度结果

Fig. 8　Scheduling results of TP and CSD

由图 9 可以看出，热泵电功率受冷、热负荷共同作用的影响较大，尤其是在 12:00 之后，随着建筑物冷、热负荷需求的同步升高，系统所需输出的功率随之显著增加，整体变化趋势更加明显。从图 10 可知，当蓄热器蓄热度较低时，由于需要补充更多热量以提高其储能水平，电加热锅炉的功率需求相对较高；而当蓄热度达到较高水平后，蓄热器内部已积累了足够热量，因此对外部热源的依赖减少，功率需求随之降低。结合图 11 与图 4 可知，在光照强度较高时，系统会优先利用光伏发电以提高可再生能源利用率；而当光照较弱或完全无光照时，则需通过燃料

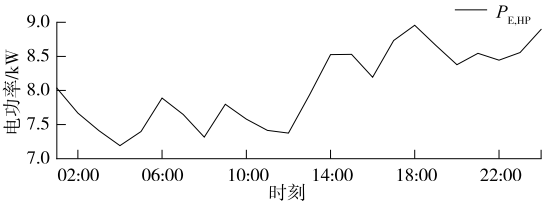


图 9 热泵调度结果

Fig. 9 Scheduling results of heat pump

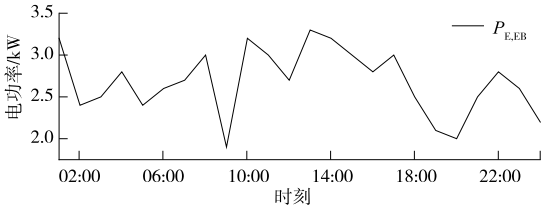


图 10 电加热锅炉调度结果

Fig. 10 Scheduling results of electric boiler

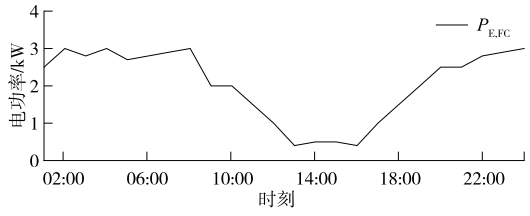


图 11 燃料电池调度结果

Fig. 11 Scheduling results of fuel cell

电池进行补充以保障系统的稳定运行。

3.4 不确定性工况下实验结果与分析

对系统在不同不确定性下的实验结果进行分析，如图 12、图 13 所示。可以看出，在 10%、20%、30%不确定性工况下，系统均有良好的实施调度能

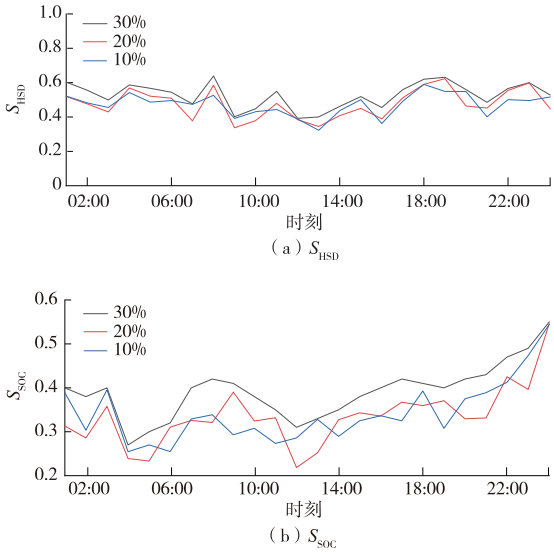


图 12 不确定性工况下 S_{HSD} 和 S_{SOC} 的调度结果

Fig. 12 Scheduling results of S_{HSD} and S_{SOC} under uncertain conditions

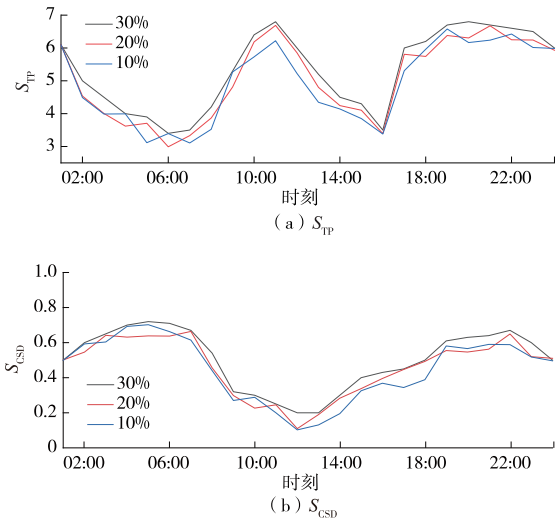


图 13 不确定性工况下 S_{TP} 和 S_{CSD} 的调度结果

Fig. 13 Scheduling results of S_{TP} and S_{CSD} under uncertain conditions

力和系统适应性。基于深度强化学习的调度策略不仅具备较强的鲁棒性，而且能够在复杂和不确定的工况下灵活应对各种负荷需求变化，从而提高综合能源系统的整体运行效率和稳定性。对比不同不确定性发现，20%不确定性下，调度结果波动最大，此时系统的预测模型和调度算法可能会出现一定的误差，尤其是在高不确定性的时间段。过低的不确定性(10%)通常意味着系统可以较精准地预测需求，调度策略较为平稳；而在 30%不确定性情况下，系统会采取更加保守的应对策略，准备处理更大波动，因此整体波动可能相对较小。20%的不确定性通常意味着预测误差的波动和系统响应之间的交互作用最为显著。

系统在尝试应对这种中等不确定性的情况时，可能会频繁调整调度策略，导致较大波动。因此，系统既不能完全适应低不确定性带来的稳定性，又无法充分预见 30%更高不确定性下的极端变化，造成调度的过度响应或滞后响应。

4 结论

(1) 本文建立的综合能源系统模型充分考虑了系统运行状态与连续的设备调度，能够全面反映该综合能源系统的设备动态运行过程。

(2) 通过处理复杂的连续控制问题，DDPG 算法能够有效实现综合能源系统中电、热、冷、氢等多种能量的协同调度，提高了系统的整体性能和适应性。

(3) 在不同不确定性条件下，DDPG 算法展示了较高的实时调度能力和系统适应性，表明其在应

对多能流和强耦合的能源系统中具有优势。

本文主要研究了基于 DDPG 的综合能源系统多能协同调度，能量涉及电、热、冷、氢。今后的研究将考虑加强更多强化学习算法提升算法收敛速度与准确性。

参考文献：

[1] 李时，王斌，齐红涛，等. 分布式园区综合能源系统经济调度监测与分析[J]. 电气自动化，2024，46(1)：97-100，103.
LI Shi，WANG Bin，QI Hongtao，et al. Economic dispatch monitoring and analysis of integrated energy system in distributed park[J]. Electrical Automation，2024，46(1)：97-100，103.

[2] 王建. 煤电节能减排升级与改造行动启动[J]. 中国设备工程，2014(10)：10.
WANG Jian. Launch of energy conservation and emission reduction upgrades and transformation actions for coal power[J]. China Plant Engineering，2014(10)：10.

[3] 苏艳. 双碳政策下产业园区的低碳模式探索[J]. 住宅与房地产，2023(30)：22-25.
SU Yan. Exploration of low-carbon models for industrial parks under the dual-carbon policy[J]. Housing and Real Estate，2023(30)：22-25.

[4] 陈良. 考虑电热综合需求响应的园区综合能源系统协同优化配置研究[D]. 吉林：东北电力大学，2023.

[5] 赵振宇，邓涵予. 阶梯碳下考虑耦合电转气-碳捕集和热电联产灵活输出的综合能源系统优化调度[J]. 动力工程学报，2024，44(8)：1287-1297.
ZHAO Zhenyu，DENG Hanyu. Optimal scheduling of integrated energy systems considering coupled power to gas，carbon capture system，and combined heating and power flexible outputs under stepped carbon[J]. Journal of Chinese Society of Power Engineering，2024，44(8)：1287-1297.

[6] 韩中合，马立，段宇轩，等. 基于多元储能的分布式能源系统优化调度方法研究[J]. 动力工程学报，2024，44(2)：317-327.
HAN Zhonghe，MA Li，DUAN Yuxuan，et al. Research on optimal scheduling method of distributed energy system based on multi-energy storage[J]. Journal of Chinese Society of Power Engineering，2024，44(2)：317-327.

[7] 邱勇，王瑾，张树斌，等. 基于 Modelica 与 Julia 的气—热—电综合能源系统动态特性分析[J]. 软件导刊，2023，22(10)：33-41.
QIU Yong，WANG Jin，ZHANG Shubin，et al. Dynamic characteristics analysis of gas-heat-electricity integrated energy system based on Modelica and Julia[J]. Software Guide，2023，22(10)：33-41.

[8] SALAZAR A，BERZOY A，SONG Wenzhan，et al. Energy management of islanded nanogrids through nonlinear optimization using stochastic dynamic programming[J]. IEEE Transactions on Industry Applications，2020，56(3)：2129-2137.

[9] 李俊. 风力发电机组发电性能影响因素分析[J]. 机械制造，2024，62(5)：32-34，43.
LI Jun. Influencing factor analysis of power generation performance of wind turbine unit[J]. Machinery，2024，62(5)：32-34，43.

[10] LISO V，SAVOIA G，ARAYA S S，et al. Modelling and experimental analysis of a polymer electrolyte membrane water electrolysis cell at different operating temperatures[J]. Energies，2018，11(12)：3273.

[11] STAFFELL I. Zero carbon infinite cop heat from fuel cell CHP[J]. Applied Energy，2015，147：373-385.

[12] GONG Xianzhi，XIONG Rui，MI C C. Study of the characteristics of battery packs in electric vehicles with parallel-connected lithium-ion battery cells[J]. IEEE Transactions on Industry Applications，2015，51(2)：1872-1879.

[13] ZHENG Jinyang，ZHANG Xin，XU Ping，et al. Standardized equation for hydrogen gas compressibility factor for fuel consumption applications[J]. International Journal of Hydrogen Energy，2016，41(15)：6610-6617.

[14] TAHIR M F，CHEN Haoyong，MEHMOOD K，et al. Integrated energy system modeling of China for 2020 by incorporating demand response，heat pump and thermal storage[J]. IEEE Access，2019，7：40095-40108.

[15] AKMAL M，FOX B. Modelling and simulation of underfloor heating system supplied from heat pump[C]//2016 UKSim-AMSS 18th International Conference on Computer Modelling and Simulation. Cambridge，UK：IEEE，2016.

[16] YU Chen，HUANG Yongping，ZHANG Chengbin. Role of metal foam in solidification performance for a latent heat storage unit[J]. International Journal of Energy Research，2020，44(1)：2110-2125.

[17] 唐明，李玉杰，王林杰，等. 冰蓄冷装置参与能量互补的微网自律协同控制研究[J]. 电力工程技术，2022，41(3)：118-124，132.
TANG Ming，LI Yujie，WANG Linjie，et al. Autonomous coordinated control for microgrids considering multi-energy complementary with ice storage devices

（下转第 212 页）