

崔在兴<sup>1</sup> 应雨龙<sup>1</sup> 李靖超<sup>2</sup> 王新友<sup>1</sup>

# 基于深度强化学习的综合能源系统低碳经济调度

## 摘要

综合能源系统 (IES) 能够实现多种能源形式的供应,但同时排放的大量 CO<sub>2</sub> 也影响着周边环境.针对综合能源系统的低碳经济调度问题,本文提出一种基于双延迟深度确定性策略梯度 (TD3) 算法的优化调度策略.首先,以调度运行成本最小为目标函数,建立考虑碳捕集技术和电转气技术的包含电、热、冷多能互补的综合能源系统模型;其次,引入碳交易机制,提高优化调度策略节能减排的积极性;然后,根据强化学习框架设计优化模型的状态空间、动作空间和奖励函数等,利用 TD3 算法中的智能体与环境互动,学习综合能源系统的运行策略;最后,利用历史数据对 TD3 算法的智能体进行训练,并对比线性规划和粒子群算法在不同场景下进行算例分析.结果表明,本文所提方法可以减少综合能源系统运行时的碳排放和运行成本,能够实现综合能源系统的低碳经济调度.

## 关键词

碳捕获系统;电转气;深度强化学习;双延迟深度确定性策略梯度;综合能源系统;碳交易机制

中图分类号 TM73;TK01

文献标志码 A

收稿日期 2023-06-08

资助项目 国家自然科学基金(62076160);上海市自然科学基金(21ZR1424700);上海市“科技创新行动计划”启明星项目(23QA1403800)

## 作者简介

崔在兴,男,硕士生,主要研究方向为综合能源系统能量管理. hanzo20160524@foxmail.com

应雨龙(通信作者),男,博士,副教授,主要研究方向为综合能源系统能量管理、能源互联网. yingyulong060313@163.com

1 上海电力大学 能源与机械工程学院,上海,200090

2 上海电机学院 电子信息学院,上海,201306

## 0 引言

由于经济快速发展,以及以煤炭为主、多能互补的能源生产与消费结构,中国已成为世界上最大的碳排放国<sup>[1]</sup>.为贯彻可持续发展战略,探索能源利用的低碳措施,实现综合能源系统的经济运行显得尤为重要<sup>[2]</sup>.

在综合能源系统(Integrated Energy System, IES)的电力供应方面,通过提升可再生能源供应占比可以降低碳排放水平,但是可再生能源如风能、光能因受天气影响较大具有不确定性.为提升可再生能源的消纳,文献[3]和文献[4]分别利用优化 IES 容量配置和增加先进储能装置的方法,提升了可再生能源的利用率,文献[5-6]验证了碳捕获系统(Carbon Capture System, CCS)技术路线的可行性和减排的有效性,并且进行了经济性分析.电转气(Power to Gas, P2G)技术则为解决可再生能源的消纳问题提供了新途径<sup>[7]</sup>.

文献[8]提出一种碳捕获系统和电转气设备的耦合模型,它考虑了燃气电厂的碳捕获,并通过算例验证了所提模型减排的有效性.文献[9]考虑到可再生能源的不确定性,提出一种结合碳捕获电厂和多能源需求响应的优化方案,可以提升可再生能源的消纳.文献[10]利用风电-光伏-氢能的一体化模型为 IES 提供稳定的氢能供应,减少了弃风、弃光,并且利用碳捕获技术减少 IES 的碳排放.文献[11]构建了加入碳捕获系统的电-气 IES 调度模型,并考虑了综合需求响应,由纵向需求确定同种能源可转移负荷量,由横向需求确定不同种类能源可替代的负荷量,但并未考虑碳交易机制.文献[12-13]利用生命周期评价方法对园区中能源链的碳足迹进行核算,但前者在模型中引入了奖惩阶梯型碳交易机制,后者则对比分析了不同碳交易价格对 IES 运行成本和碳排放的影响,结果表明,合理的碳交易市场能够促进园区的低碳发展.上述研究均使用传统优化算法,在涉及多种设备耦合的情况下,变量数目增多导致计算量呈指数增长,难以高效求解.同时,系统运行中会存在诸如需求波动、能源价格变动、设备性能变化等不确定性因素,传统优化算法难以有效应对.

以往解决不确定性问题的方法有随机优化<sup>[14]</sup>、区间优化<sup>[15-16]</sup>、鲁棒优化<sup>[17]</sup>、双层优化模型<sup>[18]</sup>等,其本质是将不确定性场景通过建模转化为确定性问题进行求解.基于数据驱动的人工智能技术具有更强大的自我学习、改进的能力,以及对研究的物理模型不敏感等特点<sup>[19]</sup>.

文献[20]使用深度Q网络对IES进行优化,但这种方法只能解决离散动作空间的问题,而实际运行动作具有连续性.文献[21]使用一种深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG)算法,能够作用于连续动作空间,但是Q值的过估计使得鲁棒性不够好.DDPG的改进版本——双延迟深度确定性策略梯度(Twin Delayed Deep Deterministic Policy Gradient, TD3)算法改善了过估计和高方差的问题,利用引入的两套神经网络的最小值减小过估计,利用延迟更新来减小目标函数高方差,以此提高算法稳定性.但是TD3算法对超参数非常敏感,学习率、经验池的大小有可能影响训练过程的稳定和收敛状态.另外,引入的双网络也有可能学习到的优化策略过于保守.

在上述研究背景下,本文提出基于TD3算法对考虑了碳捕获系统和电转气的IES系统进行运行优化.首先,将碳捕获系统和电转气设备引入IES模型,以系统运行成本最低为目标函数.其次,利用深度强化学习算法进行仿真求解,验证所提模型与算法有效性与经济性.

## 1 IES建模

IES系统结构如图1所示.IES系统包含燃气轮机(Gas Turbine, GT)、燃气锅炉(Gas Boiler, GB)、热交换器(Heat Exchanger, HE)、吸收式制冷机(Absorption Refrigerator, AR)、电制冷机(Electricity-driven Refrigeration, ER)、风能(Power of Wind, PW)等.IES系统与上级电网和上级气网连接,天然气由上级气网购入,系统内配置有燃气轮机发电与风力发电供应电能,不足的电量由上级电网购入,多余的电量不考虑上网出售.用户侧的热负荷由燃气轮机余热和燃气锅炉提供,冷负荷由电制冷机和吸收式制冷机提供.CO<sub>2</sub>的捕获时间节点为燃烧后,捕获方式为化学吸收法,主要是利用胺溶液吸收CO<sub>2</sub>的特性,让烟气与胺溶液充分混合后,在分离器中加热分离出CO<sub>2</sub>.

### 1.1 目标函数

系统优化在考虑碳交易机制的同时,兼顾经济性与减少碳排放,以综合能源系统日运行总成本最小构建目标函数,包括系统向上级电网购电的总成本和系统向上级气网购气的总成本、碳交易成本、碳捕集成本.由于测试的系统运行时间跨度小,机组维护成本忽略不计.设备投资成本不是短期支出,且数

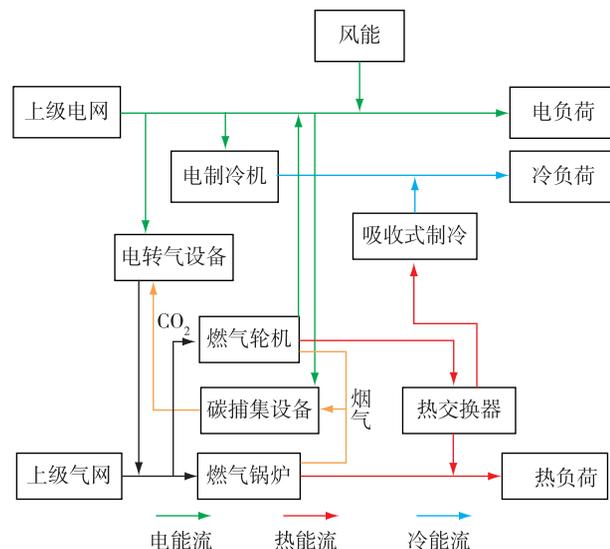


图1 综合能源系统结构

Fig. 1 Structure of integrated energy system

额偏大,不计入在内.目标函数表达式如下:

$$\min F = f_e + f_g + f_{cp} + f_{trade}, \quad (1)$$

$$f_e = \sum_t (E_{grid} C_e), \quad (2)$$

$$f_g = \sum_t ((V_{GT, gas} + V_{GB, gas}) C_{gas}), \quad (3)$$

$$f_{cp} = \sum_t \eta (V_{GT, gas} + V_{GB, gas}) \theta, \quad (4)$$

$$f_{trade} = \sum_t ((1 - \eta) (V_{GT, gas} + V_{GB, gas}) \theta - C_{free}) \alpha, \quad (5)$$

$$C_{free} = \sum PK. \quad (6)$$

式中: $F$ 为IES日运行总成本; $f_e$ 为IES向上级电网购电的总成本; $f_g$ 为IES向上级气网购气的总成本; $f_{trade}$ 为碳交易机制下IES碳排放权交易产生的成本; $f_{cp}$ 为碳捕集成本; $E_{grid}$ 为IES向上级电网购电的电量,单位kW·h; $C_e$ 为考虑市场需求响应的分时电价; $V_{GT, gas}$ 和 $V_{GB, gas}$ 分别为燃气轮机和燃气锅炉消耗的天然气的量,单位m<sup>3</sup>; $C_{gas}$ 为每立方米天然气的价格; $\eta$ 为碳捕集设备的综合捕集效率; $\theta$ 为标压下每立方米天然气燃烧所产生的CO<sub>2</sub>,单位kg/m<sup>3</sup>; $C_{free}$ 为有功单位免费分配的碳排放量,单位kg; $\alpha$ 为碳排放权交易价格,取300元/t; $P$ 为有功单位输出功率; $K$ 为相对应的分配系数.

### 1.2 约束条件

IES低碳经济调度的约束条件需要考虑电网、热网、冷网的功率平衡约束和系统内各个组件设备的运行约束.

### 1.2.1 GT 运行约束

$$P_{GT} = \eta_{GT} V_{GT, gas} q_g, \quad (7)$$

$$H_{GT} = \lambda_{GT} P_{GT}, \quad (8)$$

$$P_{GT, min} \leq P_{GT} \leq P_{GT, max}. \quad (9)$$

式中:  $P_{GT}$ ,  $H_{GT}$ ,  $\eta_{GT}$ ,  $\lambda_{GT}$  分别为燃气轮机的输出电功率、输出热功率、发电效率、热电比;  $q_g$  为天然气的低位热值, 单位  $\text{kJ}/\text{m}^3$ ;  $P_{GT, min}$  和  $P_{GT, max}$  分别为燃气轮机最小输出电功率和最大输出电功率, 单位  $\text{kW}$ .

### 1.2.2 GB 运行约束

$$H_{GB} = \eta_{GB} V_{GB, gas} q_g, \quad (10)$$

$$H_{GB, min} \leq H_{GB} \leq H_{GB, max}. \quad (11)$$

式中:  $H_{GB}$  和  $\eta_{GB}$  分别为燃气锅炉的输出热功率和产热效率;  $H_{GB, min}$  和  $H_{GB, max}$  分别为燃气锅炉输出热功率的下限和上限.

### 1.2.3 AR 运行约束

$$I_{AR} = \delta_{COP, AR} H_{AR}, \quad (12)$$

$$I_{AR, min} \leq I_{AR} \leq I_{AR, max}. \quad (13)$$

式中:  $I_{AR}$ ,  $\delta_{COP, AR}$ ,  $H_{AR}$  分别为吸收式制冷机的输出冷功率、制冷系数、输入热功率;  $I_{AR, min}$  和  $I_{AR, max}$  分别为吸收式制冷机的输出下限和上限.

### 1.2.4 ER 运行约束

$$I_{ER} = \delta_{COP, ER} E_{ER}, \quad (14)$$

$$I_{ER, min} \leq I_{ER} \leq I_{ER, max}. \quad (15)$$

式中:  $I_{ER}$ ,  $\delta_{COP, ER}$ ,  $E_{ER}$  分别为电制冷机的输出冷功率、制冷系数、输入电功率;  $I_{ER, min}$  和  $I_{ER, max}$  分别为输出冷功率的下限和上限.

### 1.2.5 功率平衡约束

$$I_{AR} + I_{ER} = I_{load}, \quad (16)$$

$$H_{GT} + H_{GB} = H_{load} + H_{AR}, \quad (17)$$

$$E_{grid} + P_{GT} + E_{pw} = E_{ER} + E_{load}. \quad (18)$$

式中:  $I_{load}$ ,  $H_{load}$ ,  $E_{load}$  分别为用户侧的冷负荷功率、热负荷功率、电负荷功率;  $E_{pw}$  为风能发电功率.

### 1.2.6 碳排放权交易机制

本文温室气体排放核算方法参考生态环境部公布的《企业温室气体排放核算方法与报告指南 发电设施》, 计算方法如下:

$$E_{F, i} = C_{C, i} O_{F, i} (44/12), \quad (19)$$

$$E_{carbon} = E_{grid} K_{grid}. \quad (20)$$

式中:  $E_{F, i}$  为第  $i$  种化石燃料的二氧化碳排放因子;  $C_{C, i}$  为第  $i$  种化石燃料的单位热值含碳量;  $O_{F, i}$  为第  $i$  种化石燃料的碳氧化率; 44/12 为二氧化碳与碳的相对分子质量之比;  $E_{carbon}$  和  $K_{grid}$  分别为电网碳排放量和排放因子, 其中电网排放因子与电网新能源供

电占比有关.

## 2 TD3 算法模型

强化学习 (Reinforcement Learning, RL) 问题可以用马尔可夫决策过程 (Markov Decision Process, MDP) 来解释. 如图 2 所示, 每次环境 (environment) 都可以用一个状态 (state) 来表示环境中产生的变化. 智能体 (agent) 根据状态值来选择行动, 环境会根据选取的行动来反馈奖励 (reward) 值, 并且基于选取的行动, 环境会相应改变, 生成新的状态. MDP 中的所有状态都具有马尔可夫性质, 即未来的状态及概率分布都仅依赖于当前状态, 与历史状态无关<sup>[22-23]</sup>.

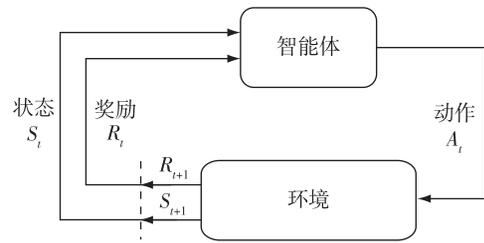


图 2 马尔可夫决策过程

Fig. 2 Markov decision process

MDP 由 5 个元素组成,  $M = \{S, A, P, R, \gamma\}$ , 其中:  $S$  为状态集合,  $A$  为动作集合,  $P$  为转移概率函数,  $R$  为奖励函数,  $\gamma$  为衰减函数. 转移概率函数代表的是智能体新状态和旧状态之间的转移概率:

$$P(s', r | s, a) = p \begin{bmatrix} S_{t+1} = s' \\ R_{t+1} = r | S_t = s \\ A_t = a \end{bmatrix}. \quad (21)$$

利用深度强化学习来解决 IES 的调度优化问题: 1) 把优化问题转化成可用 MDP 描述的过程; 2) 根据 IES 情况来选取可控制的对象作为动作空间, 合理地设计状态空间; 3) 根据目标函数设计奖励函数用来引导智能体的策略更新.

### 2.1 状态空间

在 IES 调度优化模型中, 状态空间要能够准确反映环境所处的状态, 包括需求侧的负荷、电价、气价及可再生能源出力. 状态空间的定义如式 (22) 所示:

$$s = [I'_{load}, H'_{load}, E'_{load}, C'_e, C'_{gas}, P'_{pw}]. \quad (22)$$

式中:  $C'_e$ ,  $C'_{gas}$ ,  $P'_{pw}$  分别为电价、气价和风能出力功率.

## 2.2 动作空间

在强化学习中,动作空间一般设计为模型中的变量,包括燃气轮机发电功率、燃气锅炉热功率、电制冷机和吸收式制冷机的冷功率、上级电网购入的电功率.如式(23)所示:

$$a = [P'_{CT}, H'_{GB}, I'_{ER}, I'_{AR}, E'_{grid}]. \quad (23)$$

本文所提出的模型中,上述变量之间存在时序耦合特性,选取式(23)作为动作空间难以处理模型中的能量平衡约束.同时,为简化模型的复杂程度,设计电制冷机出力与燃气轮机出力作为动作空间,如式(24)所示:

$$a = [P'_{CT}, I'_{ER}]. \quad (24)$$

## 2.3 奖励函数与衰减系数

为了达到 IES 调度优化经济成本最低的目标策略,设计目标函数的负值作为奖励函数,  $w$  为出力约束的惩罚值,以指导智能体的学习. 即时奖励如式(25)所示:

$$r = -(f_e + f_g + f_{trade} + f_{cp}) + w. \quad (25)$$

## 2.4 TD3 算法流程

TD3 是 Actor-Critic 框架下的一种确定性深度强化学习算法<sup>[24]</sup>, Actor 网络决定智能体采取的动作, Critic 网络指导 Actor 网络输出的动作从而使得奖励最大化. 为了更加准确地估计  $Q(s, a)$  的值, TD3 算法有 3 方面的优化:

- 1) 采用两套 Critic 网络, 估计目标值时选取较小值. 双重网络主要用来抑制网络估计过高问题.
- 2) 目标策略平滑正则化. 估计时在下一个状态上加入噪声扰动, 以获得更准确的价值估计.
- 3) 延迟更新. Critic 网络更新频率高于 Actor 网络, 保证 Actor 网络的训练更稳定.

基于 TD3 算法的综合能源系统调度结构如图 3 所示. 其中: Actor 网络的输入输出分别为状态变量和动作  $a$ ; Critic 网络中的输入输出分别为状态变量、动作变量和值函数  $Q(s, a)$ . 智能体与 IES 互动得出样本数据  $\{s, a, r, s'\}$  存储在经验池中, 在学习策略评估时, 随机从经验池中采样  $N$  组样本数据进行梯度更新.

## 3 案例仿真

### 3.1 基础参数设置

为验证本文所提 IES 低碳经济配置方案的有效性和经济性, 基于北方某地区 IES 历史数据进行案例分析. 向上级电网购电价格如表 1 所示, 向上级电网购气价格为 2.9 元/ $m^3$ . IES 内各设备的参数如表

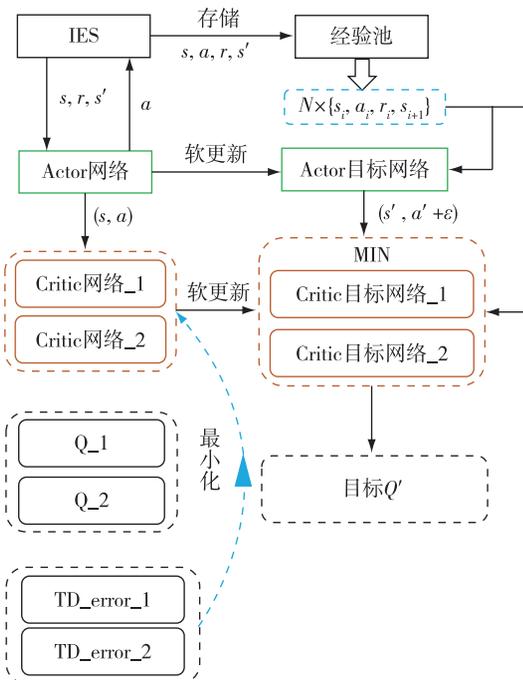


图 3 TD3 算法结构

Fig. 3 TD3-based reinforcement learning framework

2 所示. 图 4 是对 IES 在 24 h 内需求侧的电、热、冷负

表 1 分时电价

时段	价格/(元/(kW·h))
00:00—07:00	0.42
23:00—24:00	
07:00—10:00	0.62
15:00—18:00	
21:00—23:00	
10:00—15:00	0.82
18:00—21:00	

表 2 综合能源系统各设备特性参数

设备类型	特性参数	取值
燃气轮机	发电效率( $\eta_{CT}$ )	0.357
	热电比( $\lambda_{CT}$ )	1.23
	最大输出功率/kW	200
燃气锅炉	热效率( $\eta_{GB}$ )	0.95
	最大输出功率/kW	300
吸收式制冷	制冷系数( $\delta_{COP,AR}$ )	0.7
	最大输出功率/kW	100
电制冷机	制冷系数( $\delta_{COP,ER}$ )	3.83
	最大输出功率/kW	120

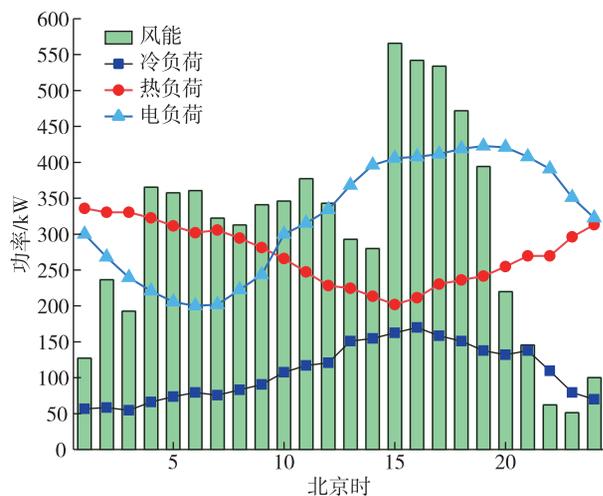


图4 新能源出力与负荷需求的预测曲线

Fig. 4 Forecast of new energy output and load demand curves

荷和供应侧的新能源出力的预测曲线.碳捕集成本和碳存储成本参考文献[5]中的结论.碳交易中对有功单位免费分配的碳排放额度参考文献[13]中的数据.TD3算法中的神经网络的激活函数为 ReLU. Actor 和 Critic 网络学习率分别为 0.001 和 0.002, 衰减系数为 0.95.

### 3.2 仿真结果及分析

为验证通过 TD3 算法训练得到的智能体是否能够在满足需求侧的负荷需求的情况下,对 IES 进行优化调度,并与线性规划 (Linear Programming, LP) 和经典的启发式算法粒子群算法 (Particle Swarm Optimization, PSO) 进行对比.3 种算法都统一使用本文提出的综合能源系统低碳调度模型,并设置以下两种场景.

- 1) 场景 1: 不考虑碳捕集和电转气设备;
- 2) 场景 2: 碳捕集和电转气设备协同运行.

场景 1 的运行结果如表 3 所示,图 5 为场景 1 下 TD3 算法电能流调度情况.

图 5 展示了一个确定性场景,其中包含了一天内的冷、热、电负荷和新能源出力数据.在问题能够使用线性规划去描述的情况下,理论上系统运行成本应当满足  $LP < TD3 < PSO$ ,而表 3 的仿真结果也验

表 3 场景 1 不同方法仿真结果比较

Table 3 Comparison of simulation results using different methods under scenario 1

算法	总成本/元	碳交易成本/元	碳排放/kg	计算时间/s
LP	3 364.9	188.5	2 409.5	1
PSO	4 564.6	255.6	3 442.3	102
TD3	3 364.9	188.5	2 409.5	0.7

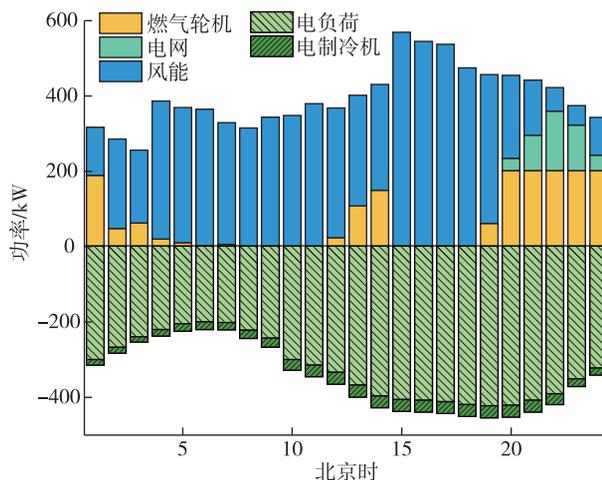


图 5 场景 1 下 TD3 算法电能流调度情况

Fig. 5 Electric energy flow scheduling of TD3 algorithm under scenario 1

证了这一理论.以 PSO 算法得出的结果为基准,TD3 算法的总成本减少了 26.3%,在碳排放量、碳交易成本上,TD3 对比 PSO 分别减少了 30% 和 26.2%,在面对多目标时,PSO 算法处于劣势.由图 5 可以看出,在满足需求侧负荷的情况下,且没有引入碳捕集和电转气设备时存在弃风现象,不能完全消纳可再生能源.

表 4 为场景 2 运行结果.从表 4 中可以看出,碳捕集和电转气设备协同运行的情况下,IES 的碳交易成本从支出变为了利润.对比场景 1 的 3 种算法的碳排放量均有不同程度减少,PSO 算法减少最多,为 2 130.5 kg.

图 6 为场景 2 不同方法耗气量对比,可以看出 TD3 耗气量小于 PSO,略大于 LP.图 7 为不同方法电

表 4 场景 2 不同方法仿真结果比较

Table 4 Comparison of simulation results using different methods under scenario 2

算法	总成本/元	碳捕集成本/元	碳存储成本/元	碳交易成本/元	碳排放/kg	计算时间/s
LP	3 362.6	562.6	75.00	-290.7	958.0	1
PSO	4 352.4	726.4	90.00	-367.1	1 311.8	102
TD3	3 367.6	582.1	77.81	-310.4	746.4	0.7

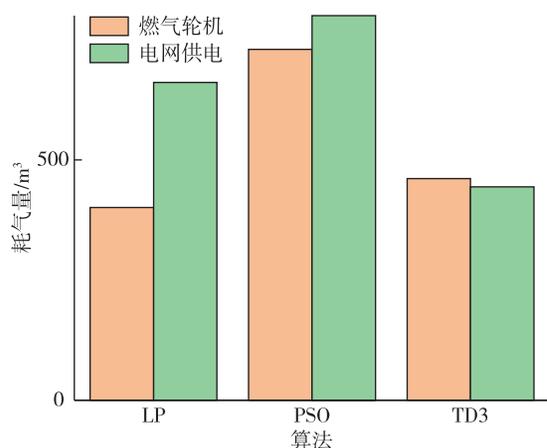


图6 场景2不同方法耗气量对比

Fig. 6 Comparison of gas consumption between different methods under scenario 2

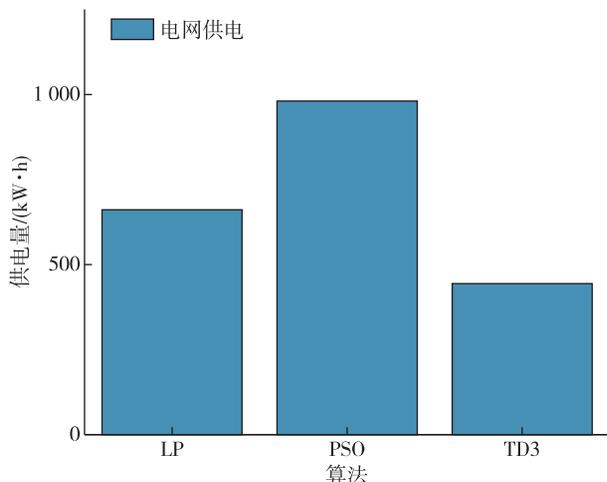


图7 场景2不同方法电网供电量对比

Fig. 7 Comparison of grid power supply between different methods under scenario 2

网供电量对比,TD3明显小于另外两种方法.碳排放包含两部分,一部分是IES中燃气轮机和燃气锅炉中燃烧后经过碳捕集向大气排放的 $\text{CO}_2$ ,另一部分则是来自电网供电的间接排放,这部分不考虑碳捕集.综合图6、图7的数据对比两部分的差异,TD3对比LP总成本略高,但碳排放减少约22%,说明TD3可以有效应对多目标问题.

图8—10为TD3算法优化得出的系统电能、热能和冷能的能量流动关系.可以看出碳捕集与电转气设备的耦合运行解决了场景1中存在的弃风现象,能量的生产和消费做到了实时平衡.图11显示了风能全部被消纳,碳捕集系统与电转气设备协同运行的情况下,显著提高了风电的利用率.

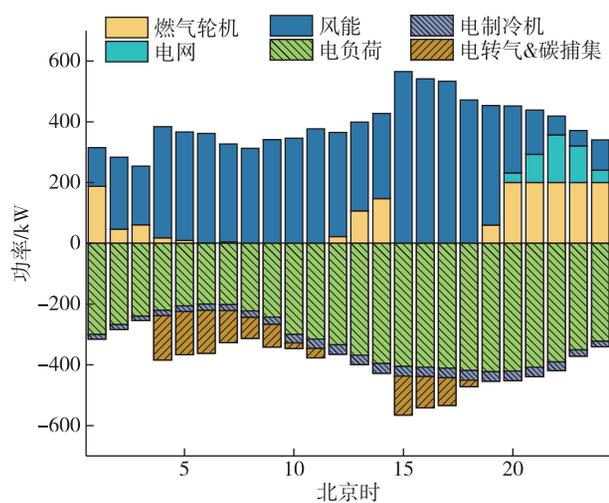


图8 系统电能流动关系

Fig. 8 System electrical energy flow

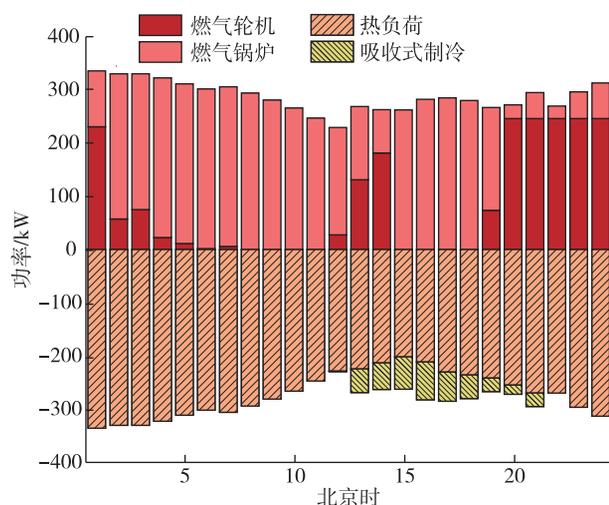


图9 系统热能流动关系

Fig. 9 System thermal energy flow

## 4 结论

本文探讨了降低IES碳排放水平的低碳措施,并提出一种基于深度强化学习方法考虑碳捕集和电转气的优化调度方案.它既可以减少二氧化碳排放,又能够提高可再生能源的利用率.通过案例仿真分析对比,得到如下结论:碳捕集系统与电转气设备协同运行的情况下,可以显著提高风电的利用率,TD3算法优化策略在经济成本上略高于线性规划(LP)算法,但碳排放减少约22%,可以实现IES经济性与低碳环保的双目标,满足低碳经济发展的需求.

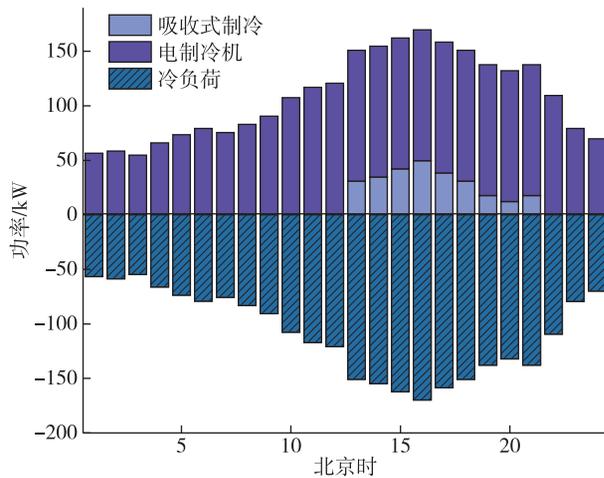


图 10 系统冷能流动关系

Fig. 10 System cold energy flow

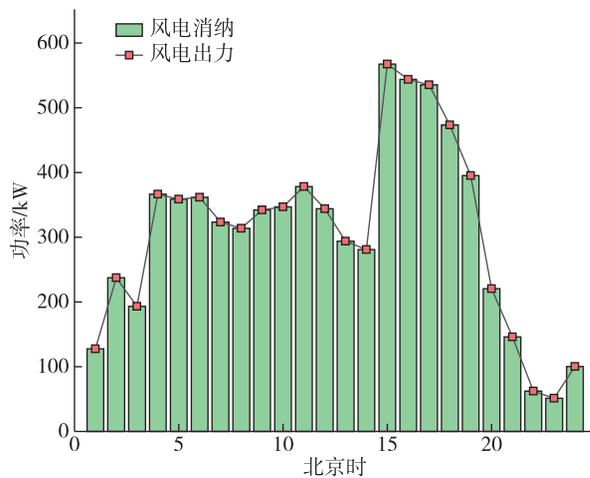


图 11 可再生能源消纳情况

Fig. 11 Accommodation of renewable energy

参考文献

References

[ 1 ] Xiong W M, Wang Y, Mathiesen B V, et al. Heat roadmap China; new heat strategy to reduce energy consumption towards 2030[J]. Energy, 2015, 81: 274-285

[ 2 ] 韩肖清, 李廷钧, 张东霞, 等. 双碳目标下的新型电力系统规划新问题及关键技术[J]. 高电压技术, 2021, 47(9): 3036-3046  
HAN Xiaoqing, LI Tingjun, ZHANG Dongxia, et al. New issues and key technologies of new power system planning under double carbon goals[J]. High Voltage Engineering, 2021, 47(9): 3036-3046

[ 3 ] 侯健敏, 丁苏云, 余威杰, 等. 包含多投资主体的综合能源系统容量配置[J]. 南京信息工程大学学报(自然科学版), 2023, 15(6): 703-711  
HOU Jianmin, DING Suyun, YU Weijie, et al. Capacity

allocation of integrated energy system involving multiple investors[J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2023, 15(6): 703-711

[ 4 ] 门家凯, 邱建龙, 陈向勇. 基于先进绝热压缩空气储能区域综合能源系统优化策略[J]. 南京信息工程大学学报(自然科学版), 2022, 14(3): 361-367  
MEN Jiakai, QIU Jianlong, CHEN Xiangyong. Integrated energy optimization strategy based on advanced adiabatic compressed air energy storage [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2022, 14(3): 361-367

[ 5 ] 吴其荣, 陶建国, 范宝成, 等. 燃煤电厂开展大规模碳捕集的技术路线选择及经济敏感性分析[J]. 热力发电, 2022, 51(10): 28-34  
WU Qirong, TAO Jianguo, FAN Baocheng, et al. Technical route selection and economic sensitivity analysis of large-scale carbon capture in coal-fired power plant [J]. Thermal Power Generation, 2022, 51(10): 28-34

[ 6 ] Franki V, Višković A, Šapić A. Carbon capture and storage retrofit: case study for Croatia [J]. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 2021, 43(24): 3238-3250

[ 7 ] 窦迅, 赵文浩, 郎伊紫禾, 等. 计及电转气技术的天然气-电力耦合系统运行研究综述[J]. 电网技术, 2019, 43(1): 165-173  
DOU Xun, ZHAO Wenhao, LANG Yizihe, et al. A review of operation of natural gas-electricity coupling system considering power-to-gas technology [J]. Power System Technology, 2019, 43(1): 165-173

[ 8 ] 李欣, 刘立, 黄婧琪, 等. 含耦合 P2G 和 CCS 的园区级综合能源系统优化调度[J]. 电力系统及其自动化学报, 2023, 35(4): 18-25  
LI Xin, LIU Li, HUANG Jingqi, et al. Optimal scheduling of park-level integrated energy system with coupling of P2G and CCS [J]. Proceedings of the CSU-EPSA, 2023, 35(4): 18-25

[ 9 ] Yang D F, Xu Y, Liu X J, et al. Economic-emission dispatch problem in integrated electricity and heat system considering multi-energy demand response and carbon capture technologies [J]. Energy, 2022, 253: 124153

[ 10 ] Wang S X, Wang S M, Zhao Q Y, et al. Optimal dispatch of integrated energy station considering carbon capture and hydrogen demand [J]. Energy, 2023, 269: 126981

[ 11 ] 田丰, 贾燕冰, 任海泉, 等. 考虑碳捕集系统的综合能源系统“源-荷”低碳经济调度[J]. 电网技术, 2020, 44(9): 3346-3355  
TIAN Feng, JIA Yanbing, REN Haiquan, et al. “source-load” low-carbon economic dispatch of integrated energy system considering carbon capture system [J]. Power System Technology, 2020, 44(9): 3346-3355

[ 12 ] 李嘉祺, 陈艳波, 陈来军, 等. 工业园区综合能源系统低碳经济优化运行模型[J]. 高电压技术, 2022, 48(8): 3190-3200

- LI Jiaqi, CHEN Yanbo, CHEN Laijun, et al. Low-carbon economy optimization model of integrated energy system in industrial parks [J]. High Voltage Engineering, 2022, 48(8):3190-3200
- [13] 王泽森,石岩,唐艳梅,等.考虑LCA能源链与碳交易机制的综合能源系统低碳经济运行及能效分析[J].中国电机工程学报,2019,39(6):1614-1626,1858  
WANG Zesen, SHI Yan, TANG Yanmei, et al. Low carbon economy operation and energy efficiency analysis of integrated energy systems considering LCA energy chain and carbon trading mechanism [J]. Proceedings of the CSEE, 2019, 39(6):1614-1626, 1858
- [14] Qiao X B, Zou Y, Li Y, et al. Impact of uncertainty and correlation on operation of micro-integrated energy system [J]. International Journal of Electrical Power & Energy Systems, 2019, 112:262-271
- [15] 李鹏,王加浩,黎灿兵,等.计及源荷不确定性与设备变工况特性的园区综合能源系统协同优化运行方法[J].中国电机工程学报,2023,43(20):7802-7812  
LI Peng, WANG Jiahao, LI Canbing, et al. Collaborative optimal scheduling of the community integrated energy system considering source-load uncertainty and equipment off-design performance [J]. Proceedings of the CSEE, 2023, 43(20):7802-7812
- [16] 曾博,徐富强,刘裕,等.考虑可再生能源与需求响应协同增效的能量枢纽多目标区间优化规划方法[J].中国电机工程学报,2021,41(21):7212-7225  
ZENG Bo, XU Fuqiang, LIU Yu, et al. Multi-objective interval optimization approach for energy hub planning with consideration of renewable energy and demand response synergies [J]. Proceedings of the CSEE, 2021, 41(21):7212-7225
- [17] Ceseña E A M, Mancarella P. Energy systems integration in smart districts: robust optimisation of multi-energy flows in integrated electricity, heat and gas networks [J]. IEEE Transactions on Smart Grid, 2018, 10(1):1122-1131
- [18] Wang S X, Dong Y C, Zhao Q Y, et al. Bi-level multi-objective joint planning of distribution networks considering uncertainties [J]. Journal of Modern Power Systems and Clean Energy, 2021, 10(6):1599-1613
- [19] 杨挺,赵黎媛,王成山.人工智能在电力系统及综合能源系统中的应用综述[J].电力系统自动化,2019,43(1):2-14  
YANG Ting, ZHAO Liyuan, WANG Chengshan. Review on application of artificial intelligence in power system and integrated energy system [J]. Automation of Electric Power Systems, 2019, 43(1):2-14
- [20] 刘俊峰,陈剑龙,王晓生,等.基于深度强化学习的微能源网能量管理与优化策略研究[J].电网技术,2020,44(10):3794-3803  
LIU Junfeng, CHEN Jianlong, WANG Xiaosheng, et al. Energy management and optimization of multi-energy grid based on deep reinforcement learning [J]. Power System Technology, 2020, 44(10):3794-3803
- [21] Yang T, Zhao L Y, Li W, et al. Dynamic energy dispatch strategy for integrated energy system based on improved deep reinforcement learning [J]. Energy, 2021, 235:121377
- [22] Botteghi N, Poel M, Brune C. Unsupervised representation learning in deep reinforcement learning: a review [J]. arXiv e-Print, 2022, arXiv:2208.14226
- [23] Ladosz P, Weng L L, Kim M, et al. Exploration in deep reinforcement learning: a survey [J]. Information Fusion, 2022, 85:1-22
- [24] 马丽新,刘磊,刘晨.安全屏障机制下基于SAC算法的机器人导航系统[J].南京信息工程大学学报(自然科学版),2023,15(2):201-209  
MA Lixin, LIU Lei, LIU Chen. Robot navigation system based on SAC with security barrier mechanism [J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2023, 15(2):201-209

## Low-carbon economic dispatch of integrated energy system based on deep reinforcement learning

CUI Zaixing<sup>1</sup> YING Yulong<sup>1</sup> LI Jingchao<sup>2</sup> WANG Xinyou<sup>1</sup>

<sup>1</sup> College of Energy and Mechanical Engineering, Shanghai University of Electric Power, Shanghai 200090, China

<sup>2</sup> School of Electronic Information Engineering, Shanghai Dianji University, Shanghai 201306, China

**Abstract** Integrated energy system (IES) enables the supply of multiple forms of energy, but the large amount of carbon dioxide it emitted affects the surrounding environment. Here, an optimal scheduling approach based on Twin Delayed Deep Deterministic Policy Gradient (TD3) is proposed for low-carbon economic scheduling of IES. First, taking the minimum operation cost as the objective function, an IES model with multiple complementary energies of electricity, heat and cold is established considering carbon capture technology and power-to-gas technology. Second, a carbon trading mechanism is introduced to stimulate the enthusiasm of energy conservation and emission reduction

under optimal scheduling. Then, according to the reinforcement learning framework, the state space, action space and reward function of the optimization model are designed, and the agents in the TD3 algorithm are used to interact with the environment to explore strategies and learn the IES operation strategies. Finally, the historical data are used to train the agents of TD3 algorithm, and the linear programming and particle swarm optimization are compared under different scenarios. The results show that the proposed approach can reduce the IES carbon emission and operating cost, thus realizing the low-carbon economic dispatch of the integrated energy system.

**Key words** carbon capture system (CCS); power-to-gas (P2G); deep reinforcement learning; twin delayed deep deterministic policy gradient (TD3); integrated energy system (IES); carbon trading mechanism