

DOI: 10.19783/j.cnki.pspc.250196

基于深度强化学习的区域综合能源系统主动调节 灵活性规则提取及可解释优化调度

窦嘉铭¹, 王小君¹, 司方远¹, 刘兆燕², 席嫣娜², 刘翌¹, 盛康玲¹, 段玉戈¹

(1. 北京交通大学电气工程学院, 北京 100044; 2. 国网北京市电力公司, 北京 100031)

摘要: 区域综合能源系统调度需充分挖掘主动调节能力以应对新能源波动及多变负荷。针对传统方法依赖精确建模、难以适应高不确定性、缺乏主动调节动态解析与策略可解释性的问题, 提出主动调节灵活性规则提取及可解释调度强化学习方法。首先基于设备调节边界、响应速率及耦合关系, 解析量化电、热等子系统设备功率调节量等灵活性规则指标。其次, 设计融合主动调节灵活性物理规则奖励函数, 将其嵌入改进深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法框架, 在策略更新过程中引入设备运行边界约束惩罚与灵活性激励, 通过动态约束构建、学习率自适应及策略可视化等增强策略物理一致性与可解释性。仿真表明, 所提方法相对二次规划、粒子群等方法调节能力指标分别提升 11.08%、15.86%, 且通过灵活性规则对日调节能力进行提取, 为调度策略提供可追溯的物理依据和人机协同支持。

关键词: 主动调节能力; 综合能源系统; 灵活性; 强化学习; 可解释人工智能; 规则提取

Extraction of active regulation flexibility rules and interpretable optimal scheduling for regional integrated energy systems based on deep reinforcement learning

DOU Jiaming¹, WANG Xiaojun¹, SI Fangyuan¹, LIU Zhaoyan², XI Yanna², LIU Zhao¹, SHENG Kangling¹, DUAN Yuge¹

(1. School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China;

2. State Grid Beijing Electric Power Company, Beijing 100031, China)

Abstract: The scheduling of regional integrated energy systems must fully exploit active regulation capability to cope with new energy fluctuation and diverse load conditions. Traditional methods rely heavily on precise modeling, struggle with high uncertainty, and lack dynamic analysis of active regulation as well as interpretability of scheduling strategies. To address these challenges, this paper proposes an active flexibility regulation rule extraction and explainable reinforcement learning method. First, based on the equipment regulation boundaries, response rates, and coupling relationships, flexibility metrics such as power regulation capacities of electrical and thermal subsystem components, are quantitatively analyzed. Second, a reward function integrating the physical rules of active regulation flexibility is designed and embedded into an improved deep deterministic policy gradient (DDPG) framework. During policy updates, device operation constraints and flexibility incentives are incorporated. Dynamic constraint construction, adaptive learning rate adjustment, and policy visualization are adopted to enhance physical consistency and interpretability of the learning process. Simulation results show that the proposed method improves the regulation capability by 11.08% and 15.86% compared with quadratic programming and particle swarm optimization, respectively. Moreover, the extracted flexibility rules enable interpretable day-ahead regulation capability analysis, providing traceable physical insights and supporting human-AI collaborative decision-making in scheduling strategies.

This work is supported by the Science and Technology Project of the Headquarters of State Grid Corporation of China (No. 520201250010-075-ZN).

Key words: active regulation capability; integrated energy system; flexibility; reinforcement learning; explainable AI; rule extraction

基金项目: 国家电网有限公司总部科技项目资助(520201250010-075-ZN)

0 引言

区域综合能源系统(integrated energy systems, IES)有望通过能量枢纽与多能网络,整合冷热电联产、分布式发电、储能、能源转换设备等可控机组,带来灵活多能源转换存储以及快速爬坡调节能力^[1-4]。其可在一定程度上缓解新能源发电的不稳定性和波动性,并通过高效整合与利用更多的可再生能源,适应广泛且多变的能源需求,从而进一步提升系统多能源供需调控能力^[5-8]。

区域 IES 的多能源供需调控能力体现在对复杂能源网络的全面感知、智能分析和精准控制上。其中,主动调节能力是实现多能供需调控与多主体能量共享的关键支撑与前提^[9]。与电力系统主要通过需求响应及负荷管理提升调节能力和容量的方式不同^[10],IES 异质能源在生产、输送、转换、储存以及最终使用过程中紧密交织。多能互补的特点使得影响主动调节性能和容量的关键因素变得尤为复杂,难以准确刻画。为此,须考虑在多元灵活资源聚合下对 IES 主动调节能力的构成要素和关键灵活性规则进行动态解析和量化评估,并以此为依据,制定后续自治协调优化和能量共享策略^[9]。

目前区域 IES 调度研究已取得显著进展。文献[11]提出考虑功率可调节裕度的 IES 多时间尺度调度模型,分析了净负荷预测误差和可调度机组对可调节能力的影响。文献[12]引入了电转气、碳捕集利用与封存、掺氢燃气等灵活性技术,建立了考虑奖惩碳交易的调度模型。文献[13]提出不确定量波动范围灵活性调度指标,构建了运行成本最小和可容纳室外温度和光照强度不确定波动范围最大的 IES 多目标区间优化调度模型。文献[14]则基于心理学原理构建可转移电负荷和可替代负荷区间响应模型,并采用线性规划求解。文献[15]建立考虑供需双侧响应和碳交易机制的氢能 IES 鲁棒调度模型,通过鲁棒优化处理可再生能源不确定性。综合来看,尽管现有求解方法取得了一定成果,但其多数依赖于精确数学建模,通过场景缩减、典型曲线,简化或凸化求解模型非线性项和不确定性^[16],系统发生变化时需要重新建模求解。而 IES 高效运行需要高度灵活性和快速响应的运行规则,此类方法在应对高不确定性和强随机性时,往往难以为继^[17]。同时上述方法在高维非凸空间计算效率受限,存在时效性较差的问题^[18-19]。

因此,为解决含主动调节能力评估的 IES 运行

优化问题,学者们正在积极探寻新方法。新质生产力下出现以强化学习(reinforcement learning, RL)为代表的 IES 调度求解方法^[19-20]。文献[21]指出 IES 规划问题模型复杂、变量众多,属于高维混合整数非线性规划,关键需求是多目标的高效优化,往往采用运筹学优化与 RL 相结合的研究思路。文献[17]针对 IES 间歇性能源和负荷不确定性逐步增强的问题,验证了 RL 在不同时间尺度以及不确定性环境下的可行性与有效性。文献[22]提出计及可再生能源和负荷需求不确定性的 IES 能量管理 RL,其能够自适应对源荷随机波动做出响应。文献[23]提出利用图 RL 实现 IES 非欧拓扑信息实现优化调度。总的来看,不同于传统方法输出精确解,RL 通过输出策略映射,可根据 IES 的不同状态生成动态调整的调度策略,具有自适应性强、学习能力高、无需精确建模、端到端学习等优势。

然而,作为有潜力的解决方案,RL 面向 IES 主动调节灵活性规则构建研究仍处于起步阶段,尚未形成统一理论体系与方法框架,亟待解决的问题包括:

1) 新型电力系统背景下,区域 IES 需要更强的动态调节和柔性响应^[24]。RL 具备实时探索与应对高不确定性的特点^[25-27],但现有研究多侧重策略优化与效果提升本身,普遍缺乏与调节能力、灵活性规则等能源物理知识规则的融合范式;

2) 能源调度过程对策略透明性、可追溯性与人机协同能力的要求日益增强。基于 RL 的调度优化结果未来有望辅助调度员进行决策,但目前受限于缺乏有效工具分析和展示决策背后的主动调节机制与调控逻辑^[28]。

基于上述考虑,本文提出基于 RL 的区域 IES 主动调节灵活性规则提取及可解释优化调度方法,主要贡献包括:

1) 面向跨区域 IES 能量互济需求,动态解析与量化表达主动调节需求,形成多种主动调节灵活性规则,为后续 RL 调度提供知识嵌入基础;

2) 设计主动调节能力灵活性物理规则嵌入 RL 调度算法,在确定性策略梯度智能体更新基础上,进行灵活性规则信息嵌入、约束建模、调节能力可视化等改进,实现调度策略生成,比现有方法调节能力提升 3%~15%;

3) 构建 IES 灵活性物理规则驱动的 RL 可解释性框架,通过动态可视化与规则化表达主动调节能力,为黑盒调度提供了一定程度可追溯的调控路径解析与人机协同决策支持。

1 问题描述

1.1 主动调节能力与灵活性规则

1.1.1 主动调节能力定义

主动调节能力是指 IES 通过整合和协调本地的多元灵活资源, 对电力、热能、天然气等多能供需平衡进行本地自治调节的能力^[9]。如图 1 所示, 该能力能够使系统根据电力供需平衡功率曲线(即图 1 绿色虚线)进行必要的调整, 同时调节前后功率曲线形成的面积差(即图 1 红色部分)即为 IES 在能量共享框架下的主动调节量 $P_{\text{regulation}}$, 如式(1)所示。

$$P_{\text{regulation}} = \int_{t_1}^{t_2} |P_{\text{real}(t)} - P_{\text{theory}(t)}| dt \quad (1)$$

式中: $P_{\text{real}(t)}$ 为调节后的实际功率; $P_{\text{theory}(t)}$ 为调节前的基准功率。一般而言, 主动调节量是在时间段 $[t_1, t_2]$ 内, IES 通过其多元资源调节所能实现的功率曲线调整总量。然而由于涉及多种能源形式在生产、传输、转换、存储与消费等环节的复杂耦合, IES 主动调节能力难以直接观测或统一定量。因此, 需通过对系统和设备的动态行为过程物理建模, 进行动态抽取与量化评估。

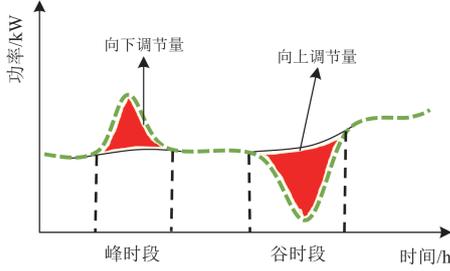


图 1 主动调节能力示意图

Fig. 1 Schematic diagram of active regulation capabilities

1.1.2 主动调节灵活性规则提取

为实现主动调节能力结构化建模, 对灵活性规则进行了提取, 刻画 IES 各类设备的调节边界、调节速度、响应特性与耦合关系等关键因素。“灵活性规则”即用于结构化表达各类可调节设备在不同能源系统中的调节能力边界、调节速度限制、运行优先级和互联强度等物理条件与行为特征。从能源类型视角, 灵活性规则可分为电、热、气等多个子系统; 而从供需方向来看, 又可分为供给型设备与消耗型设备^[29]。本文研究的 IES 包含热电联产(combined heat and power, CHP)、燃气锅炉(gas boiler, GB)、电锅炉(electric boiler, EB)等可控能源转换设备, 下文将以日前调度为例开展灵活性规则分析。

1) 能源消耗型设备功率调节量

$$T_{u,j,t}^d = \begin{cases} \min(P_{u,j,t} - P_{u,j,\min}, r_{u,j}^- \Delta t), & d = \text{down} \\ \min(r_{u,j}^+ \Delta t, P_{u,j,\max} - P_{u,j,t}), & d = \text{up} \end{cases} \quad (2)$$

式中: d 为上下调方向; $T_{u,j,t}^d$ 为 t 时段消耗型设备对子系统 j (在本文中为电和热)的可调节量; $P_{u,j,t}$ 为 t 时段消耗型设备在子系统 j 的实际消耗功率; $P_{u,j,\max}$ 、 $P_{u,j,\min}$ 分别为消耗型设备在子系统 j 的最大、最小消耗功率; $r_{u,j}^+$ 、 $r_{u,j}^-$ 分别为消耗型设备的最大向上、向下爬坡速率; Δt 为时间间隔。

2) 能源供给型设备功率调节量建模

$$T_{j,u,t}^d = \begin{cases} \min(P_{j,u,t} - P_{j,u,\min}, r_{j,u}^- \Delta t), & d = \text{up} \\ \min(r_{j,u}^+ \Delta t, P_{j,u,\max} - P_{j,u,t}), & d = \text{down} \end{cases} \quad (3)$$

式中: $T_{j,u,t}^d$ 为 t 时段供给型设备对子系统 j 的可调节量; $P_{j,u,t}$ 为 t 时段供给型设备在子系统 j 的实际生产功率; $P_{j,u,\max}$ 、 $P_{j,u,\min}$ 分别为供给型设备在子系统 j 的最大、最小消耗功率; $r_{j,u}^+$ 、 $r_{j,u}^-$ 分别为供给型设备的最大向上、向下爬坡速率。

上述两类功率调节量规则属基础规则, 为进一步全面评估系统层面调节潜力, 设计下面 5 种灵活性规则指标作为补充。

3) 能源互联及区域调度能力 (energy interconnection and regional dispatching capability, EIRDC)

EIRDC 表示 IES 通过能源网络能够调度的最大能源容量与系统最大用能量的比值, 可量化区域间的协同和调度灵活性, 记为 r_{EIRDC} , 如式(4)所示。

$$r_{\text{EIRDC}} = \frac{P_{\text{grid},t}}{P_{\text{grid}}^{\max}} \times 100\% \quad (4)$$

式中: $P_{\text{grid},t}$ 为 t 时段 IES 与能源供应商交互功率;

P_{grid}^{\max} 为允许交互功率上限。

4) 储能占比 (energy storage capacity ratio, ESCR)

ESCR 用于衡量 IES 储能资源的配置程度, 记为 r_{ESCR} , 如式(5)所示。

$$r_{\text{ESCR}} = \frac{H_{\text{BES}}}{H_{\text{BES}}^{\max}} \times 100\% \quad (5)$$

式中: H_{BES} 为当前储能设备的实际储能容量, 可通过当前储能状态和储能设备的总容量计算; H_{BES}^{\max} 为系统中所有储能设备总容量。

5) 惯性渗透率 (system inertial penetration rate, SIPR)

SIPR 用于衡量 IES 中惯性特性设备在整体供能中的占比, 记为 r_{SIPR} , 如式(6)所示。

$$r_{\text{SIPR}} = \frac{P_{\text{Inertia}}}{P_{\text{total}}} \times 100\% \quad (6)$$

式中: P_{Inertia} 为供能惯性设备供能总量; P_{total} 为系统供能总量。

6) 热电耦合动态响应时间(CHP dynamic response time, CDRT)

考虑到热力学对电热类 IES 灵活性具有一定影响, 提出 CDRT, 记为 r_{CDRT} , 用于量化 CHP 等设备响应速度动态惯性特性, 通过设备在电/热功率输出上的实际变化速率与其爬坡能力的比值, 衡量其从接收到调节指令到稳定输出目标所需最短时间, 如式(7)所示。

$$r_{\text{CDRT}} = \max \left(\frac{|P_{\text{elec}}^{\text{new}} - P_{\text{elec}}^{\text{old}}|}{E_{\text{ramp}}^{\text{elec}}}, \frac{|H_{\text{heat}}^{\text{new}} - H_{\text{heat}}^{\text{old}}|}{E_{\text{ramp}}^{\text{heat}}} \right) \quad (7)$$

式中: $E_{\text{ramp}}^{\text{elec}}$ 和 $E_{\text{ramp}}^{\text{heat}}$ 分别为电热功率爬坡速率(kW/min); $P_{\text{elec}}^{\text{old}}$ 、 $P_{\text{elec}}^{\text{new}}$ 分别为设备在调节指令前与达到目标时的电功率输出; $H_{\text{heat}}^{\text{old}}$ 、 $H_{\text{heat}}^{\text{new}}$ 分别为设备在调节指令前与达到目标时的热功率输出。

7) 热解耦调节裕度(electro-thermal decoupling margin, ETDM)

IES 中广泛存在“以热定电”运行方式, 在 CHP 中表现为热电耦合刚性约束。基于热力学原理, ETDM 刻画当前热负荷下 CHP 电功率最小出力限制, 记为 r_{ETDM} , 其计算公式为

$$r_{\text{ETDM}} = \frac{P_{\text{elec}}^{\text{max}} - P_{\text{elec}}^{\text{min}}(H_{\text{heat}})}{P_{\text{elec}}^{\text{max}}} \quad (8)$$

式中: $P_{\text{elec}}^{\text{min}}(H_{\text{heat}})$ 表示在当前热功率输出 H_{heat} 下, CHP 所需维持的最小电功率出力。ETDM 越高, 表明当前调度状态下电热耦合对电功率调节的限制越小, 系统具备更高的电能调节自由度。

1.2 IES 优化调度数学模型

IES 调度运行模式与区域内分布式发电、联合供热购电、多能源转换及多能源负荷等息息相关。为此, 本节设计 IES 调度数学模型, 包括目标函数与约束条件两部分。

1.2.1 目标函数

目标函数为区域 IES 运行成本 F 最优。

$$F = \min(F_{\text{e}}^{\text{out}} + F_{\text{g}}^{\text{out}}) \quad (9)$$

$$F_{\text{e}}^{\text{out}} = \sum_{t=1}^T (p_{\text{e},t}^{\text{buy}} \cdot \rho_{\text{e},t}^{\text{buy}} - p_{\text{e},t}^{\text{sell}} \cdot \rho_{\text{e},t}^{\text{sell}}) \quad (10)$$

$$F_{\text{g}}^{\text{out}} = \sum_{t=1}^T (p_{\text{g},t}^{\text{buy}} \cdot \rho_{\text{g},t}^{\text{buy}}) \quad (11)$$

式中: $F_{\text{e}}^{\text{out}}$ 为 IES 从主电网购电成本; $F_{\text{g}}^{\text{out}}$ 为 IES

从天然气供应商购买天然气成本; $p_{\text{e},t}^{\text{buy}}$ 、 $p_{\text{e},t}^{\text{sell}}$ 分别为 t 时段 IES 从主电网买入、卖出的电量; $\rho_{\text{e},t}^{\text{buy}}$ 、 $\rho_{\text{e},t}^{\text{sell}}$ 分别为 t 时段 IES 从主电网处买入、卖出电量单价; $p_{\text{g},t}^{\text{buy}}$ 为 t 时段 IES 从天然气供应商购买天然气的量; $\rho_{\text{g},t}^{\text{buy}}$ 为 t 时段 IES 从天然气供应商处购买天然气单价。

1.2.2 约束条件

区域 IES 调度过程需满足一定安全约束, 作为开展经济性灵活调整的前提。

1) 设备安全约束

设备安全约束包括充放电约束、上下限约束等, 如式(12)和式(13)所示。

$$\begin{cases} c_t^{\text{soc},\text{min}} \leq c_t^{\text{soc}} \leq c_t^{\text{soc},\text{max}} \\ \eta_{\text{BES}} = \eta_{\text{ch}}, \quad p_{\text{e},t}^{\text{BES}} < 0 \\ \eta_{\text{BES}} = \eta_{\text{dis}}^{-1}, \quad p_{\text{e},t}^{\text{BES}} \geq 0 \\ c_0^{\text{soc}} = k \\ c_t^{\text{soc}} = c_{t-1}^{\text{soc}} - \eta_{\text{BES}} p_{\text{e},t}^{\text{BES}} b_{\text{BES}}^{-1} \Delta t \\ p_{\text{h}}^{\text{EB},\text{min}} \leq p_{\text{h},t}^{\text{EB}} \leq p_{\text{h}}^{\text{EB},\text{max}} \\ p_{\text{h}}^{\text{GB},\text{min}} \leq p_{\text{h},t}^{\text{GB}} \leq p_{\text{h}}^{\text{GB},\text{max}} \\ p_{\text{e}}^{\text{CHP},\text{min}} \leq p_{\text{e},t}^{\text{CHP}} \leq p_{\text{e}}^{\text{CHP},\text{max}} \\ p_{\text{e}}^{\text{BES},\text{min}} \leq p_{\text{e},t}^{\text{BES}} \leq p_{\text{e}}^{\text{BES},\text{max}} \end{cases} \quad (12)$$

式中: c_t^{soc} 、 $c_t^{\text{soc},\text{max}}$ 、 $c_t^{\text{soc},\text{min}}$ 分别为储能荷电状态及其上、下限; c_0^{soc} 为储能荷电状态初始值; k 为 SOC 初始状态; η_{BES} 为充放电系数; η_{ch} 、 η_{dis} 分别为储能(battery energy storage, BES)充、放电效率; $p_{\text{e},t}^{\text{BES}}$ 为 BES 充放电功率; b_{BES} 为 BES 容量; $p_{\text{h},t}^{\text{EB}}$ 为 EB 提供的热功率; $p_{\text{h}}^{\text{EB},\text{max}}$ 、 $p_{\text{h}}^{\text{EB},\text{min}}$ 分别为 EB 热功率上、下限; $p_{\text{h},t}^{\text{GB}}$ 为 GB 提供的热功率; $p_{\text{h}}^{\text{GB},\text{max}}$ 、 $p_{\text{h}}^{\text{GB},\text{min}}$ 分别为 GB 热功率上、下限; $p_{\text{e},t}^{\text{CHP}}$ 为 CHP 提供的电功率; $p_{\text{e}}^{\text{CHP},\text{max}}$ 、 $p_{\text{e}}^{\text{CHP},\text{min}}$ 分别为 CHP 电功率上、下限; $p_{\text{e}}^{\text{BES},\text{max}}$ 、 $p_{\text{e}}^{\text{BES},\text{min}}$ 分别为 BES 充放电功率上、下限。

2) 电、热功率平衡安全约束

$$p_{\text{e},t}^{\text{load}} = p_{\text{e},t}^{\text{buy}} + p_{\text{e},t}^{\text{PV}} + p_{\text{e},t}^{\text{CHP}} + p_{\text{e},t}^{\text{BES}} - p_{\text{e},t}^{\text{EB}} \quad (14)$$

$$p_{\text{h},t}^{\text{load}} = p_{\text{h},t}^{\text{CHP}} + p_{\text{h},t}^{\text{GB}} + p_{\text{h},t}^{\text{EB}} \quad (15)$$

式中: $p_{\text{e},t}^{\text{load}}$ 为电负荷需求; $p_{\text{e},t}^{\text{PV}}$ 为光伏提供的电功率; $p_{\text{e},t}^{\text{EB}}$ 为 EB 消耗的电功率; $p_{\text{h},t}^{\text{load}}$ 为 IES 热负荷需求; $p_{\text{h},t}^{\text{CHP}}$ 为 CHP 提供的热功率。

3) 能源交易传输安全约束

对主动调节能力的评估可支撑后续 IES 跨区域互补互济。因此, 灵活性规则嵌入的调度求解结果

应为后续园区与外部供应商或外部园区交互功率, 该交互功率应符合传输约束, 即

$$\left| P_{\text{grid},t} \right| \leq P_{\text{grid}}^{\max} \quad (16)$$

1.3 IES 调度马尔可夫决策模型设计及其灵活性规则嵌入奖励函数

本节构建嵌入主动调节灵活性规则马尔可夫决策过程模型(Markov decision process, MDP), 在考虑经济性的同时, 融合物理调节边界、灵活性规则等信息, 在一定程度上增强算法和策略自解释性。

1) 状态空间设计

调度智能体所处环境是指其实际工作场所, 即 1.2 节系统及约束(式(7)一式(14))。智能体须在每一时刻感知环境状态 s_t , 做出调度决策, 获得环境相应奖励。状态空间即是对智能体当前时刻所处环境的表示, 因此其包含了影响调度行为中较为核心的变量, 如式(17)所示。

$$s_t = \{p_{e,t}^{\text{PV}}, p_{e,t}^{\text{load}}, p_{h,t}^{\text{load}}, c_t^{\text{soc}}\} \quad (17)$$

2) 动作空间设计

动作空间 a_t 定义了智能体可做出的决策变量, 在 IES 调度中即智能体在每个时间步可选择的所有可能动作集合。考虑到调度目标是通过调节各机组实现供需平衡与灵活经济运行, 因此动作空间设计为各机组设备出力, 如式(18)所示。

$$a_t = \{p_{e,t}^{\text{CHP}}, p_{h,t}^{\text{CHP}}, p_{h,t}^{\text{GB}}, p_{e,t}^{\text{EB}}, p_{h,t}^{\text{EB}}, p_{e,t}^{\text{BES}}\} \quad (18)$$

3) 主动调节灵活性物理规则嵌入的奖励函数设计

奖励函数定义了智能体在环境中采取某一动作后获得的“反馈”, 是引导智能体学习“好行为”与“坏行为”的关键。考虑到调度策略在经济性、系统灵活性与运行边界之间的协同, 提出主动调节能力灵活性物理规则嵌入的奖励函数 R_t , 如式(19)所示。

$$R_t = \beta_1 F_t + \beta_2 \ln \left(1 + \left(\frac{T_{u,j,t}^d + T_{j,u,t}^d + r_{\text{ESCR}} + r_{\text{SIPR}}}{r_{\text{EIRDC}} + r_{\text{CDRT}} + r_{\text{ETDM}}} \right) \right) - \beta_3 (u_{\text{CHP}} + u_{\text{GB}} + u_{\text{EB}} + u_{\text{SOC}} + u_{\text{BES}} + u_{\text{Grid}}) \quad (19)$$

式中: β_1 、 β_2 、 β_3 分别为经济奖励、灵活性激励与边界惩罚系数; F_t 表示 t 时段的购电费用, 衡量经济运行效果; u_{CHP} 、 u_{GB} 、 u_{EB} 、 u_{SOC} 、 u_{BES} 、 u_{Grid} 分别表示各设备的越限惩罚。第二项为灵活性激励, 通过多个反映系统物理调节能力的指标构建加权, 体现当前时刻灵活性水平。与传统 RL 单一最小化运行成本的目标函数不同, 本文在奖励机制中显式引入主动调节能力驱动物理灵活性指标簇, 以实现对设备运行边界、响应潜力、调节能力的量化。第三项为 CHP 越限、GB 越限等运行边界惩罚。因购

电越少则奖励越大, 经济性奖励取负号。灵活性奖励采用对数函数构造非线性激励, 可在提升系统灵活性时提供正向反馈, 又能在趋近饱和时实现边际递减, 避免过度激励。惩罚项取负号即让智能体规避不合规行为, 引导策略向安全、可行的方向优化。

2 基于主动调节灵活性规则信息嵌入的深度强化学习 IES 调度方法

2.1 整体框架

考虑到 IES 可调对象差异大、数量多、动作输出难以统一、智能体参数调整复杂, 提出主动调节灵活性规则信息嵌入的智能调度方法。如图 2 所示, 该方法基于知识-数据融合思想, 对深度确定性策略梯度(deep deterministic policy gradient, DDPG)^[30-33]进行改进, 引入灵活性规则与设备调节能力可视化模块, 构建融合物理边界、经济目标与灵活性激励 RL。

1) 在决策层, 整体系统建模为调度智能体。智能体获取 IES 当前状态 $s_{t-1}(+n_{t-1})$ (其中 n_{t-1} 为从零均值的高斯分布 $N(0, \sigma^2)$ 中抽取的随机数, 用于模拟不确定性的高斯扰动), 输出 IES 中每台机组下一时段实时计划策略 a_t , 如式(20)所示。

$$a_t = f_{\text{actor}}(s_{t-1}; \theta_A) + N(0, \sigma^2) \quad (20)$$

式中: f_{actor} 为 Actor 网络; θ_A 为 Actor 网络参数。

2) 在优化层, Critic 网络对策略 a_{t-1} 进行评价, 估计当前策略下的动作价值 $Q(s_t, a_t)$, 如式(21)所示。

$$Q(s_t, a_t) = f_{\text{critic}}(s_t, a_t; \theta_C) \quad (21)$$

式中: f_{critic} 为 Critic 网络; θ_C 为 Critic 网络参数, 采用高斯扰动调整调度动作策略。更新 Critic 时, 利用目标 Critic 网络计算下一状态目标 Q 值。上述过程中, 由于传统 RL 无法直接对机组调节能力进行判断, 提出利用主动调节能力评估规则式(2)一式(8)计算调节能力价值加权, 构成总奖励得分 $Q(s_t, a_t)$ 。同时, 考虑日前调度问题特点, 设置 24 h 优化循环与调节规则嵌入模块: 在循环过程中, 计算获取每日调度策略及日调度情况下每小时各机组调节能力。

3) 设计经验池回放更新网络参数, 将每次智能体与环境交互数据存储在经验池, 并从经验池随机抽取数据更新网络参数, Critic 损失函数 $L(\theta_C)$ 更新为

$$L(\theta_C) = \mathbb{E}[(R_{t+1} + \gamma_1 \cdot Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t))^2] \quad (22)$$

式中: $\mathbb{E}[\cdot]$ 为数学期望; R_{t+1} 为即时奖励; γ_1 为折扣因子。

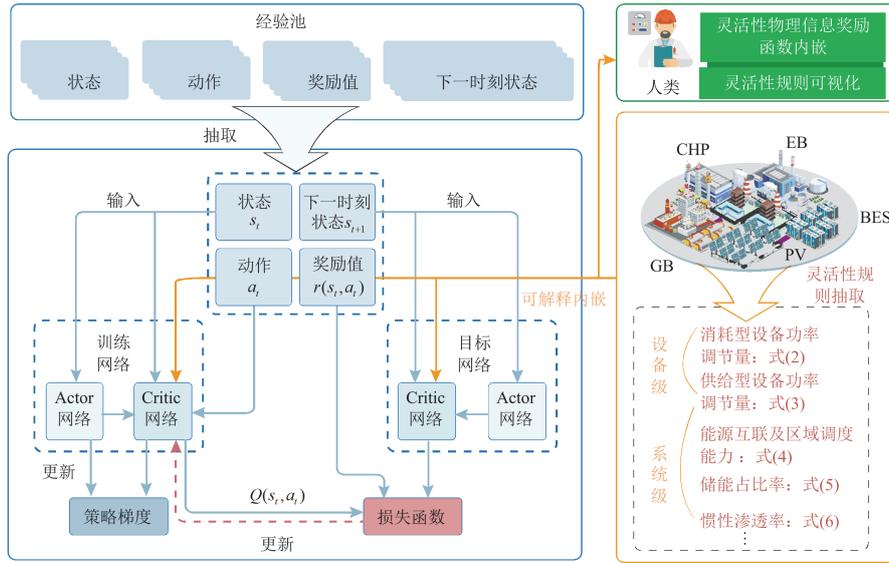


图2 主动调节能力灵活性规则嵌入的RL IES调度框架

Fig. 2 IES dispatching framework based on RL with flexibility rules of active regulation capabilities

4) 开展策略网络更新，其目标是提升当前策略在 Critic 下的动作价值，通过最大化动作价值函数来构成 Actor 网络损失函数 $L(\theta_A)$ ，进而优化策略网络参数 θ_A 。

$$L(\theta_A) = -\mathbb{E}[Q_c(s_t, \mu(s_t; \theta_A))] \quad (23)$$

式中： $\mu(s_t; \theta_A)$ 为当前状态下 Actor 网络输出的动作；负号表示该优化目标为最大化动作价值函数。

5) 为了提高训练稳定性，采用软更新策略对目标网络参数逐步迭代。目标 Actor 网络参数与 Critic 网络参数更新公式分别为

$$\theta_{A'} \leftarrow \tau \cdot \theta_A + (1 - \tau) \cdot \theta_{A'} \quad (24)$$

$$\theta_{C'} \leftarrow \tau \cdot \theta_C + (1 - \tau) \cdot \theta_{C'} \quad (25)$$

式中： $\theta_{A'}$ 和 $\theta_{C'}$ 分别为更新后的 Actor 和 Critic 网络参数； τ 为软更新系数，用于控制更新速率， $\tau \in (0, 1)$ 。

2.2 灵活性规则嵌入策略改进

为使 RL 适应含调节能力评估的区域 IES 调度复杂性，进一步提升物理一致性，在 DDPG 算法基础上，围绕设备约束建模、学习率调节策略、策略结果可解释等环节进行改进。

1) 动态约束构建及惩罚设计

DDPG 本质上为无约束优化，其目标是最大化奖励，一般难以实现动态约束。本节构建基于设备运行特性与系统调节需求的动作惩罚，显式将其嵌入奖励函数中。以 c_t^{soc} 为例，为防止出现 $c_{t+1}^{\text{soc}} > c^{\text{soc,max}}$ 或 $c_{t+1}^{\text{soc}} < c^{\text{soc,min}}$ 的越限，设计惩罚 u_{soc} ，如式(26)所示。

$$u_{\text{soc}} = \lambda_{\text{soc}} \cdot |c_{t+1}^{\text{soc}} - c^{\text{soc,min}}| \quad (26)$$

式中： λ_{soc} 为惩罚系数，用于调整惩罚项。对于设备出力越限、非允许时间段动作等异常情况，均引入类似惩罚项，并叠加至奖励函数中。

此外，为进一步保证策略输出在物理范围内，动作输出阶段引入边界裁剪，对智能体生成的连续动作 a_t 进行约束。具体而言，对于每个动作维度 a_t^i ，均施加如下额外约束。

$$a_t^i = \min(a_{\text{max}}^i, \max(a_{\text{min}}^i, a_t^i)) \quad (27)$$

式中： a_{min}^i 、 a_{max}^i 分别为第 i 个动作对应的物理最小、最大边界。

2) 学习率自适应调节

传统 DDPG 采用固定学习率开展参数优化，此设置早期可导致策略探索不足、收敛速度较慢，而后期阶段又可能陷入局部最优、陷入局部最优。为此，设计基于训练轮次的学习率衰减，按照预设周期动态调整策略网络与价值网络学习率 O_{new} ，即

$$O_{\text{new}} = O_{\text{init}} \cdot \gamma_2^{\lfloor \frac{m}{n} \rfloor} \quad (28)$$

式中： O_{init} 为初始学习率； γ_2 为衰减系数； m 为当前训练轮次； n 为学习率更新间隔。

3) 活性规则嵌入流程

为提升所提方法在 IES 调度中的工程适应性与可解释性，基于主动调节能力建模与灵活性指标嵌入机制，构建“行为可视、逻辑可回溯”的训练测试流程。首先， t 时段 Actor 网络根据当前观测 IES 状态 s_t 并生成调度动作 a_t 。随后，IES 环境根据此时调度动作进行状态转移，得到下一时刻状态 s_{t+1} 。由于 IES 中 PV 出力与电热负荷本身具有随机性，

在状态转移时引入扰动采样, 构造面向扰动场景的随机运行轨迹。随后, 环境给出状态转移后的奖励 $r(s_t, a_t)$, 并将其与当前状态、动作、下一状态构成四元组 $(s_t, a_t, r(s_t, a_t), s_{t+1})$ 存入经验池。每日调度周期最后一小时对整日 24 h 总调节能力进行汇总评估, 生成调节行为轨迹图与指标变化曲线, 实现调控过程调节能力可视化, 辅助调度员理解。最后, 将 Actor 网络生成的策略与系统状态一同输入 Critic 网络, 后者输出状态-动作值函数 $Q(s_t, a_t)$ 用于指导策略优化。调度员可基于历史经验与相关指标, 对策略行为进行决策评估与人机协同干预。上述流程在梯度下降机制下循环迭代, 逐步实现调度智能体自适应学习与行为策略的可解释优化。

在上述过程中, 传统 DDPG 虽能实现策略生成, 但其策略学习缺乏明确的决策依据与调控路径, 导致策略的可追溯性和可解释性不足, 难以满足人机协同的分析需求。为此, 本文通过嵌入灵活性规则挖掘 IES 的调节能力资源, 旨在提升模型的自解释能力。AI 可解释性即 AI 能够以人类易于理解的方式, 传达其在学习过程中所挖掘的数据特征及领域知识, 可分为建模前解释、训练后解释、模型自解释三类^[34], 常见技术包括敏感性分析、可视化、因果推理、局部近似等。在本文中, 主要涉及训练后解释和模型自解释两类: 训练后解释利用模型输出和中间结果进行分析与归因, 开展灵活性规则的记录与分项奖励的分析, 实现策略调节能力时间轨迹的可视化; 模型自解释则体现为模型结构本身具备一定的可解释性, 结构化奖励函数和训练过程灵活性规则使模型具备一定“内嵌解释性”, 其每一项与调控行为、系统状态之间具有明确关系, 人类可通过奖励组成回溯每一步动作的逻辑依据, 加强理解。

3 算例分析

本节通过算例仿真验证所提区域 IES 主动调节灵活性规则提取及可解释优化调度方法的有效性。电、热负荷和光伏出力曲线如图 3 所示。所提方法的参数见表 1—表 4。

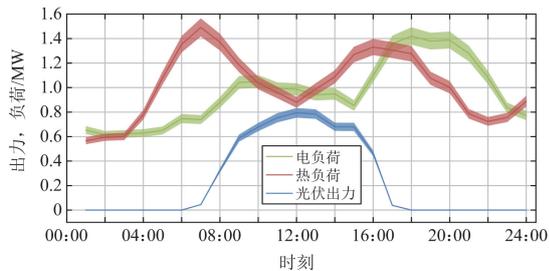


图 3 电、热负荷和光伏出力

Fig. 3 Load, thermal and PV output

表 1 设备电热功率

设备类型	最小电(热)功率/MW	最大电(热)功率/MW
CHP	0	0.4
BES	-2.5	2.5
GB	0	0.3
EB	0	0.8

表 2 设备的其他运行参数

设备类型	参数	值
CHP	热电比	1.2
	电效率	0.35
GB	效率	0.8
EB	输出效率	0.95
	充放电效率	0.95
BES	荷电状态最小值	0.2
	荷电状态最大值	0.8
	折旧系数	0.05
	容量/MWh	1

表 3 能源价格

时段	购电电价/(元/kWh)	售电电价/(元/kWh)
峰时段	1.3	0.8
平时段	0.8	0.2
谷时段	0.2	0

表 4 训练参数设置

参数	取值	参数	取值
训练轮次	1000	奖励折扣因数	0.99
每轮周期	24	探索率	0.03
价值网络层数	3	策略网络层数	5
价值网络学习率	0.005	策略网络学习率	0.001
软更新参数	0.005	批处理大小	64
缓冲区大小	10 000	—	—

3.1 离线训练情况分析

3.1.1 调度智能体训练情况分析

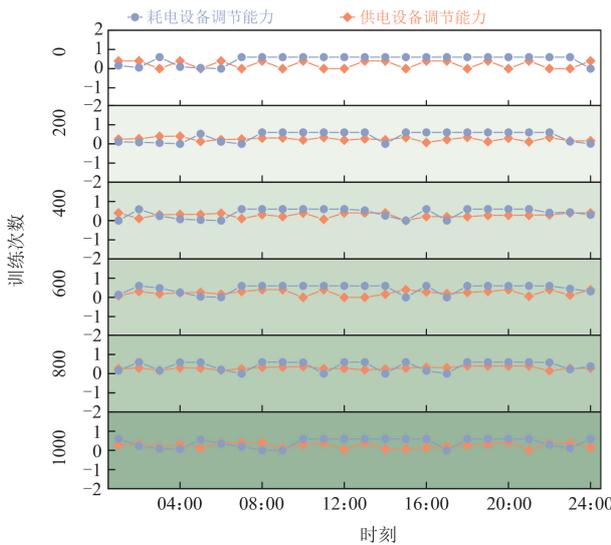
分析训练情况发现, 奖励值随训练轮次呈现波动但总体上升, 智能体逐步学习如何更有效地进行调度决策。同时, 决策奖励函数的不断提升得益于智能体对 IES 状态的适应和策略的优化。训练初期, 奖励值在 0 附近振荡, 调度智能体在探索 IES 和学习最佳行动策略中尚未达到平衡。随着足够的探索和学习, 调度智能体开始展现出更稳定的策略, 奖励值持续增长, 最后达到了+50 左右的奖励。

3.1.2 训练过程调节能力变化分析

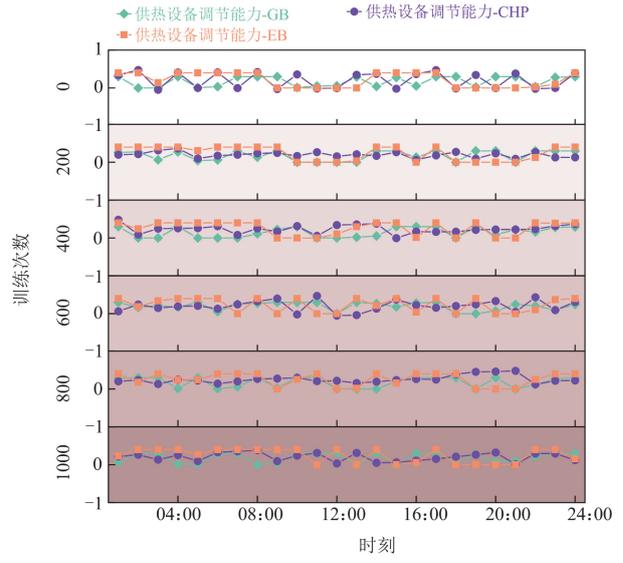
调度智能体训练过程调节能力变化分析如图 4

所示。训练初期，调节能力普遍较低。供电设备调节能力在某些情况下甚至为 0，智能体尚未学习到有效的调节策略。EB 等供热设备训练前期调节能力较低，且随着训练的进行，设备调节能力有一定提升。供电设备调节能力在多次训练后有了微小的增长。对于 CHP 和 EB 等供热设备，训练中期的热能调节能力表明 IES 开始理解如何更有效地利用这些设备来满足热能需求。在训练后期，调度智能体做出较好决策，调节能力在一定程度上得到了提

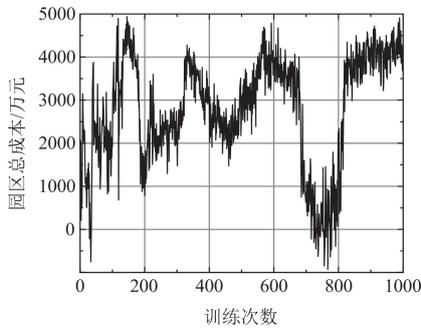
升。耗电设备的调节能力趋于稳定，CHP 和 EB 等供热设备在训练后期的调节能力相对前期有一定提高。同时，为展示所提方法的可观测性，图 4(d)、图 4(e)分别展示训练 500 轮时 CDRT 和 ETDM 的评估结果。可以看出，CDRT 在 24 h 调度训练时呈现一定波动，保持在 0.1~0.6 min；ETDM 整体灵活性较高，在 08:00 时几乎完全解耦，CHP 在低热负荷条件下具备较大电功率调节空间。



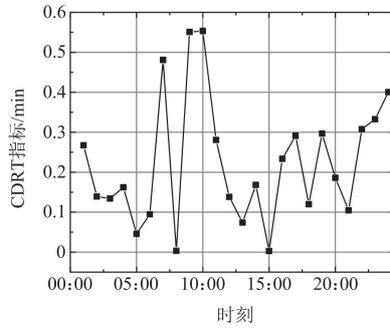
(a) 电力系统



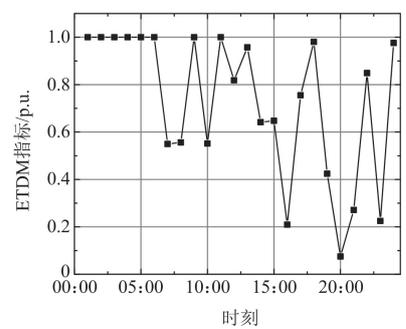
(b) 热力系统



(c) 成本变化曲线



(d) 训练500轮CDRT灵活性评估



(e) 训练500轮ETDM灵活性评估

图 4 调度决策生成过程中的 IES 调节能力变化情况可视化

Fig. 4 Changes in IES active regulation capability during dispatch decision generation process

3.2 在线测试结果分析

所提方法在线测试结果如图 5 和图 6 所示，下面分别按照谷时段、峰时段进行分析。

3.2.1 谷时段分析

谷时段处于 01:00—07:00，CHP 热、电功率普遍较低，夜间系统的能源需求相对较小。CHP 这一时段运行在较低负荷下，但仍能提供一定量热能、电能。例如 01:00 时热功率为 0.157 MW，电功率

为 0.188 MW，低负荷下 CHP 仍能有效进行能源转换。BES 在谷时段电功率为负值，主要进行放电操作，其释放能量以满足 IES 需求或作为备用能源。例如 01:00 时 BES 电功率为-0.418 MW，被用来平衡 IES 电负荷或提供必要的支持。GB 和 EB 热功率在谷时段也相对较低，在需求较低的时段，系统减少了对这些设备的依赖，以降低运行成本。然而，GB 机组在 07:00 时的热功率有所增加，与谷时段

结束时能源需求的逐渐增加有关。

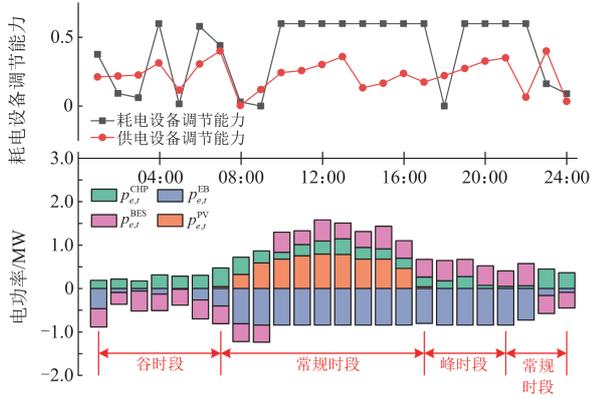


图 5 电力系统机组出力计划与调节能力评估

Fig. 5 Power system unit output planning and regulation capability

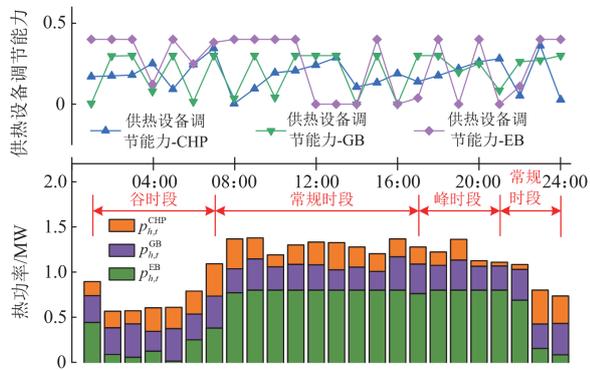


图 6 热力系统机组出力计划与调节能力评估

Fig. 6 Heat system unit output planning and regulation capability

对谷时段灵活性分析如下：耗电设备灵活性在谷时段表现出较高水平，特别是在 01:00 和 06:00 时，耗电设备能够灵活地调整其运行，以减少电能使用。供电设备灵活性在谷时段相对较低(接近 0)，特别是在 02:00—03:00，由于夜间需求降低，导致供电给设备运行在较低负荷或部分关闭状态。在热能设备灵活性方面，CHP 灵活性在谷时段展示出一定波动，04:00 时达到 0.092 MW，07:00 时降低至 0.003 MW。GB 灵活性在谷时段大多时间接近 0.3 MW，其在夜间能以相对固定水平提供热能，但调节变化有限。EB 灵活性在谷时段显示出较高数值，尤其在 01:00 和 06:00，分别为 0.400 MW 和 0.381 MW。EB 在夜间具有较高调节能力，能够根据热需求的变化灵活地调整其运行状态。

综合来看，谷时段耗电设备和 EB 在夜间具有较高的灵活性，能够根据需求的变化进行有效调整。而 CHP 和 GB 灵活性相对较低。此外，CHP 的灵活性在夜间表现出一定波动，与热需求的不稳定性

有关。为提高谷时段的系统灵活性，未来可以考虑更精细的需求响应，或对 CHP 和 EB 技术进行升级，以提高其在低需求时段的运行效率和灵活性等。

3.2.2 峰时段分析

峰时段为 18:00—22:00，此时 CHP 热功率和电功率都有所增加，IES 正积极响应能源需求增加，这与实际相符。

峰时段灵活性分析如下。首先，供电设备展示出一定灵活性波动。刚进入峰时段时，供电设备灵活性为 0.274 MW，IES 有较高的能力调整电能的供应以应对需求的增加。然而在 22:00 时，供电设备的灵活性降低至 0.162 MW，系统电能供应能力有所下降。热能供给方面，在峰时段开始时 CHP 具有较高的热能灵活性，随着时间的推移其灵活性逐渐降低。GB 灵活性在峰时段相对稳定。另外，不同类型的供热设备在峰时段的灵活性也有所不同。耗电设备的灵活性在峰时段也表现出一定的变化。18:00 时耗电设备灵活性为 0.600 MW，而 22:00 时降低至 0.162 MW，峰时段用户对电能的需求更加刚性，难以通过需求侧管理来提高灵活性。此外，EB 灵活性在峰时段也值得关注。例如 18:00 时 EB 的灵活性为 0.013 MW，而在 22:00 时升高至 0.400 MW，EB 在峰时段开始时放电能力有限，但随着时间推移，其放电能力有所提高。

综合来看，虽然 IES 在峰时段具有一定的调节能力，但不同类型设备在不同时段的灵活性表现不一，且整体灵活性随着时间推移有所下降。为提高峰时段的系统灵活性，需要采取优化 CHP 和 EB 运行策略、加强需求侧管理提高用户端的电能消耗灵活性等措施。

3.3 与现有研究的对比及可解释性讨论

3.3.1 与现有研究对比

为验证所提方法的有效性，将所提方法与遗传算法(genetic algorithm, GA)、粒子群算法(particle swarm optimization, PSO)、蚁群算法(ant colony optimization, ACO)、差分进化算法(differential evolution, DE)、灰狼优化算法(grey wolf optimizer, GWO)等启发式算法以及二次规划等数学方法进行对比，其结果如表 5 所示。

综合来看，所提方法经济成本略高于部分启发式算法，但在约束违反度与调节能力两个指标上展现出一定优势。约束违反度较 GA 与 PSO 降低 42.7%、64.4%，表明所提方法在长期策略学习过程中内嵌系统运行边界及物理可行域有一定效果。综上，所提方法的调节能力指标为 33.16，略高于其他

表 5 不同方法结果对比

Table 5 Results comparison of different methods

方法	经济成本/万元	约束违反度/p.u.	调节能力
所提方法	2674.194	1112.66	33.16
GA	2369.311	1941.87	28.62
PSO	2385.333	3126.57	32.11
ACO	2425.612	2716.57	31.58
DE	2357.541	2042.51	30.17
GWO	2394.832	2182.71	28.05
二次规划	2975.308	1245.45	29.85

算法，其相较二次规划、粒子群算法等分别提升 11.08%、15.86%，体现出所提方法在调度中对多能系统灵活性挖掘能力与主动调配能力稍强。进一步分析 RL 智能体训练过程发现，奖励函数引入越限惩罚与灵活性的正向激励在策略学习中起到关键作用。通过每时刻动作执行后的 CDRT、ETDM、

ESCR、SIPR 等灵活性指标进行动态评估，并将其作为 RL 优化的部分，调度智能体能够在策略更新过程中逐步学习“安全可行”与“高灵活性”之间的权衡。此外，环境中包括多维调节行为与约束反应，使得调度智能体在长期学习中形成对约束边界的内在敏感性，一定程度上避免传统 RL 因缺乏物理规则指导而出现策略跳变与短视问题。

3.3.2 可解释性讨论

所提方法与现有方法(如文献[29]等)在提供动态和直观监测 IES 调节能力的对比见表 6。可以看出，传统方法求解在处理动态环境和实时调节需求时受限于预设模型和假设，难以迅速适应变化。而经典 RL 虽在整体优化上性能较好，但其过于关注算法效率和准确性，而忽视了实时监测和设备行为的物理行为，对设备在不同运行条件下的调节能力缺乏具体表征，使得调度决策缺乏针对性和前瞻性。

表 6 与现有含灵活性评估的调度方法对比

Table 6 Comparison with existing scheduling methods with flexibility assessment

对比方法	文献[29]方法	文献[35]方法	文献[2]方法	本文方法
调度策略求解方法	CPLEX 求解器	CPLEX 求解器	交替方向乘法	深度强化学习
调度策略生成方式	在线学习	在线学习	在线学习	离线训练+在线预测
策略调节能力可观测性	不可观察	不可观察	不可观察	可观察，逐时输出设备调节能力
动态场景适应性	弱，须重新求解	弱，须重新求解	弱，须重新求解	中，在一定程度上自适应应对扰动变化
动态场景调节能力评估	须根据不同场景重新评估调节能力	须根据不同场景重新评估调节能力	须根据不同场景重新评估调节能力	根据不同情景自动评估调节能力
与调度员的交互性	弱，不支持交互	弱，不支持交互	弱，不支持交互	强，支持训练过程策略实施评估与干预反馈
调节能力评估全面性	一般	一般	一般	一般
调节方法可扩展性	须重新设计，不可直接扩展	须重新设计，不可直接扩展	须重新设计，不可直接扩展	可直接修改参数或场景扩展
调节能力信息可读性	各能源子系统可调节边界	各能源子系统功率调节裕度	总灵活性裕度	各能源子系统中机组功率调节裕度

相比之下，所提方法除具备 RL 端到端决策能力，还通过引入灵活性规则实现策略生成过程中的物理一致性建模与调节逻辑透明化。其具体优势包括：1) 在策略可解释性方面，传统方法通常仅输出最终的调度结果，缺乏对调度背后逻辑的刻画；而所提方法通过在奖励函数中嵌入上/下调边界规则等灵活性评估指标，并在调度过程中输出各设备的调节能力曲线，使得调度行为与物理规则之间的映射关系清晰可见，提升了可解释性与可追溯性。2) 在调节能力可观测性方面，显式记录每一时刻主动调节能力边界，并动态更新，使得调度员可实时掌握系统的“调节余地”，这是传统优化方法和黑盒式 RL 方法所缺乏的。3) 从安全性角度，RL 方法策

略生成可能出现“策略漂移”、“设备越限”等不可控现象，造成调度失效甚至设备损坏；所提方法设计设备运行边界等物理规则作为边界裁剪软惩罚嵌入奖励函数，通过灵活性规则实现对策略行为的物理校验，在一定程度上提升了调度策略的安全性。

未来工作可进一步研究基于安全约束的 RL 训练机制(如 Safe RL、Constraint RL 等)，并结合专家系统进行策略筛选，以在保证智能调度能力的同时，实现调度过程的可控、安全、可信与可审计。

4 结论

本文提出一种主动调节能力规则提取嵌入的 IES 强化学习经济调度方法，主要结论如下。

1) 本文从框架方面证明了 RL 方法可辅助机理模型对 IES 主动调节能力进行实时动态解析, 挖掘自主灵活性规则。

2) 基于区域 IES 主动调节灵活性规则嵌入的调度计划可部署于辅助决策系统, 以进一步提升综合能源主动调节能力规则及其可解释性, 并为后续 IES 能量资源高效共享与利用奠定基础。

未来可从调节能力物理规则全面性、物理信息强化学习等方面进一步研究 IES 智能调度灵活性。

参考文献

- [1] 邱玥, 陆帅, 陆海, 等. 综合能源系统灵活性: 基本内涵、数学模型与研究框架[J]. 电力系统自动化, 2022, 46(17): 16-43.
 QIU Yue, LU Shuai, LU Hai, et al. Flexibility of integrated energy systems: basic connotation, mathematical models, and research framework[J]. Automation of Electric Power Systems, 2022, 46(17): 16-43.
- [2] 胡福年, 张彭成, 周小博, 等. 计及灵活性资源的综合能源系统源荷协调优化调度[J]. 中国电力, 2024, 57(5): 2-13.
 HU Funian, ZHANG Pengcheng, ZHOU Xiaobo, et al. Coordinated optimal scheduling of source and load in integrated energy system considering flexible resources[J]. Electric Power, 2024, 57(5): 2-13.
- [3] 杜洁, 张海锋, 王佳蕊, 等. 计及用户热耐受度的区域综合能源系统韧性评估及提升研究[J]. 全球能源互联网, 2024, 7(1): 3-13.
 DU Jie, ZHANG Haifeng, WANG Jiarui, et al. Resilience assessment and enhancement oriented human heat tolerance of regional integrated energy systems[J]. Journal of Global Energy Interconnection, 2024, 7(1): 3-13.
- [4] 李崎勇, 赵新哲, 郑一飞, 等. 基于纳什谈判考虑能源共享的区域综合能源系统优化配置[J]. 电力系统保护与控制, 2023, 51(5): 22-32.
 LI Jiyong, ZHAO Xinzhe, ZHENG Yifei, et al. Optimal configuration of a regional integrated energy system considering energysaving based on Nash negotiation[J]. Power System Protection and Control, 2023, 51(5): 22-32.
- [5] WANG Tao, HUANG Zhu, YING Ruixuan, et al. A low-carbon operation optimization method of ETG-RIES based on adaptive optimization spiking neural P systems[J]. Protection and Control of Modern Power Systems, 2024, 9(5): 162-177.
- [6] 高云波, 上官小丹, 董景, 等. 计及气网约束的区域综合能源系统惯性资源配置研究[J]. 供用电, 2024, 41(4): 35-44, 52.
 GAO Yunbo, SHANGGUAN Xiaodan, DONG Jing, et al. Research on inertial resource configuration of the regional integrated energy system considering gas network constraints[J]. Distribution & Utilization, 2024, 41(4): 35-44, 52.
- [7] LIU Chao, WANG Weiru, LI Jing, et al. Optimized scheduling of integrated energy systems for low carbon economy considering carbon transaction costs[J]. Global Energy Interconnection, 2024, 7(4): 377-390.
- [8] 范宏, 鲁家阳, 陆骁霄. 考虑激励型需求响应的多区域综合能源系统协同规划[J]. 电测与仪表, 2023, 60(9): 117-124, 187.
 FAN Hong, LU Jiayang, LU Xiaoxiao. Coordinated planning of multi-region integrated energy system considering incentive demand response[J]. Electrical Measurement & Instrumentation, 2023, 60(9): 117-124, 187.
- [9] 司方远, 张宁, 韩英华, 等. 面向多元灵活资源聚合的区域综合能源系统主动调节能力评估与优化: 关键问题与研究架构[J]. 中国电机工程学报, 2024, 44(6): 2097-2119.
 SI Fangyuan, ZHANG Ning, HAN Yinghua, et al. Fundamental problems and research framework for assessment and optimization of the functional regulation capacity of the regional integrated energy system under the aggregation of diversified and flexible resources[J]. Proceedings of the CSEE, 2024, 44(6): 2097-2119.
- [10] 刘又榕, 林顺富, 沈运帷, 等. 计及电动汽车参与多元需求响应的微电网多时间尺度优化调度模型[J]. 电力建设, 2023, 44(10): 51-62.
 LIU Yourong, LIN Shunfu, SHEN Yunwei, et al. Multi-time-scale optimization scheduling model of microgrid with electric vehicles participating in multiple demand response[J]. Electric Power Construction, 2023, 44(10): 51-62.
- [11] 李志颖, 王致杰, 王鸿. 考虑功率可调节裕度的区域综合能源系统多时间尺度优化调度[J]. 太阳能学报, 2025, 46(2): 255-261.
 LI Zhiying, WANG Zhijie, WANG Hong. Multi-time scale optimal scheduling of regional integrated energy system considering power adjustable margin[J]. Acta Energetica Sinica, 2025, 46(2): 255-261.
- [12] 周永旺, 许灿城, 蔡政彤, 等. 考虑奖惩碳交易的含灵活性设备联合运行的综合能源系统优化[J/OL]. 上海交通大学学报: 1-25[2025-07-09]. <https://doi.org/10.16183/j.cnki.jsjtu.2024.291>
 ZHOU Yongwang, XU Cancheng, CAI Zhengtong, et al. Optimization of integrated energy system with joint operation of flexible equipment considering reward and penalty carbon trading[J/OL]. Journal of Shanghai Jiao

- Tong University: 1-25[2025-07-09]. <https://doi.org/10.16183/j.cnki.jsjtu.2024.291>
- [13] 盛焯, 林舜江, 梁炜焜, 等. 考虑灵活性的城市综合能源系统区间优化调度[J]. 中国电机工程学报, 2025, 45(10): 3723-3737.
SHENG Xuan, LIN Shunjiang, LIANG Weikun, et al. Interval optimal dispatch of urban integrated energy system considering the flexibility[J]. Proceedings of the CSEE, 2025, 45(10): 3723-3737.
- [14] 闫梦阳, 李华强, 王俊翔, 等. 计及综合需求响应不确定性的园区综合能源系统优化运行模型[J]. 电力系统保护与控制, 2022, 50(2): 163-175.
YAN Mengyang, LI Huaqiang, WANG Junxiang, et al. Optimal operation model of a park integrated energy system considering uncertainty of integrated demand response[J]. Power System Protection and Control, 2022, 50(2): 163-175.
- [15] 毕锐, 王孝淦, 袁华凯, 等. 考虑供需双侧响应和碳交易的氢能综合能源系统鲁棒调度[J]. 电力系统保护与控制, 2023, 51(12): 122-132.
BI Rui, WANG Xiaogan, YUAN Huakai, et al. Robust dispatch of a hydrogen integrated energy system considering double side response and carbon trading mechanism[J]. Power System Protection and Control, 2023, 51(12): 122-132.
- [16] 黎静华, 朱梦姝, 陆悦江, 等. 综合能源系统优化调度综述[J]. 电网技术, 2021, 45(6): 2256-2272.
LI Jinghua, ZHU Mengshu, LU Yuejiang, et al. Review on optimal scheduling of integrated energy systems[J]. Power System Technology, 2021, 45(6): 2256-2272.
- [17] 蔺伟山, 王小君, 孙庆凯, 等. 不确定性环境下基于深度强化学习的综合能源系统动态调度[J]. 电力系统保护与控制, 2022, 50(18): 50-60.
LIN Weishan, WANG Xiaojun, SUN Qingkai, et al. Dynamic dispatch of an integrated energy system based on deep reinforcement learning in an uncertain environment[J]. Power System Protection and Control, 2022, 50(18): 50-60.
- [18] 胡维昊, 曹迪, 黄琦, 等. 深度强化学习在配电网优化运行中的应用[J]. 电力系统自动化, 2023, 47(14): 174-191.
HU Weihao, CAO Di, HUANG Qi, et al. Application of deep reinforcement learning in optimal operation of distribution network[J]. Automation of Electric Power Systems, 2023, 47(14): 174-191.
- [19] 包义辛, 徐楞赞, 杨强. 基于 MADDPG 算法的建筑群柔性负荷优化调控方法[J]. 综合智慧能源, 2023, 45(7): 61-69.
BAO Yixin, XU Luoyun, YANG Qiang. Optimized control method for flexible load of a building complex based on MADDPG reinforcement learning[J]. Integrated Intelligent Energy, 2023, 45(7): 61-69.
- [20] 戴彦, 王刘旺, 李媛, 等. 新一代人工智能在智能电网中的应用研究综述[J]. 电力建设, 2018, 39(10): 1-11.
DAI Yan, WANG Liuwang, LI Yuan, et al. A brief survey on applications of new generation artificial intelligence in smart grids[J]. Electric Power Construction, 2018, 39(10): 1-11.
- [21] 韩笑, 郭剑波, 蒲天骄, 等. 电力人工智能技术理论基础与发展展望(一): 假设分析与应用范式[J]. 中国电机工程学报, 2023, 43(8): 2877-2891.
HAN Xiao, GUO Jianbo, PU Tianjiao, et al. Theoretical foundation and directions of electric power artificial intelligence (I): hypothesis analysis and application paradigm[J]. Proceedings of the CSEE, 2023, 43(8): 2877-2891.
- [22] 王新迎, 赵琦, 赵黎媛, 等. 基于深度 Q 学习的电热综合能源系统能量管理[J]. 电力建设, 2021, 42(3): 10-18.
WANG Xinying, ZHAO Qi, ZHAO Liyuan, et al. Energy management approach for integrated electricity-heat energy system based on deep Q-learning network[J]. Electric Power Construction, 2021, 42(3): 10-18.
- [23] 吕金玲, 王小君, 窦嘉铭, 等. 考虑运行状态信息的综合能源系统图强化学习优化调度[J]. 电力系统保护与控制, 2024, 52(2): 1-14.
LÜ Jinling, WANG Xiaojun, DOU Jiaming, et al. Optimal dispatch of an integrated energy system based on graph reinforcement learning considering operation state information[J]. Power System Protection and Control, 2024, 52(2): 1-14.
- [24] 孙大雁, 耿建, 杨胜春, 等. 新型电力系统环境下煤电与新能源等多元资源协调优化运行及其挑战[J]. 电网技术, 2024, 48(8): 3105-3113.
SUN Dayan, GENG Jian, YANG Shengchun, et al. Challenges of coordinated and optimized operation of coal power and new energy and other diversified resources in new power system[J]. Power System Technology, 2024, 48(8): 3105-3113.
- [25] 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 the application of artificial intelligence in power systems and integrated energy systems[J]. Automation of Electric Power Systems, 2019, 43(1): 2-14.
- [26] 高冠中, 姚建国, 严嘉豪, 等. 基于多智能体深度强化学习的配-微网协同优化调度研究[J]. 智慧电力, 2024,

- 52(9): 80-87.
GAO Guanzhong, YAO Jianguo, YAN Jiahao, et al. Collaborative optimization scheduling of distribution network and microgrids based on multi agent deep reinforcement learning[J]. Smart Power, 2024, 52(9): 80-87.
- [27] 周翔, 王继业, 陈盛, 等. 基于深度强化学习的微网优化运行综述[J]. 全球能源互联网, 2023, 6(3): 240-257.
ZHOU Xiang, WANG Jiye, CHEN Sheng, et al. Review of microgrid optimization operation based on deep reinforcement learning[J]. Journal of Global Energy Interconnection, 2023, 6(3): 240-257.
- [28] 王小君, 窦嘉铭, 刘墨, 等. 可解释人工智能在电力系统中的应用综述与展望[J]. 电力系统自动化, 2024, 48(4): 169-191.
WANG Xiaojun, DOU Jiaming, LIU Zhao, et al. Review and prospect of explainable artificial intelligence and its application in power systems[J]. Automation of Electric Power Systems, 2024, 48(4): 169-191.
- [29] 李志颖, 王鸿, 王致杰. 基于多能源耦合的区域综合能源系统主动调节能力评估[J]. 上海电机学院学报, 2023, 26(6): 330-336.
LI Zhiying, WANG Hong, WANG Zhijie. Evaluation of active regulation capability of regional integrated energy systems based on multi-energy coupling[J]. Journal of Shanghai Dianji University, 2023, 26(6): 330-336.
- [30] 李付强, 张文朝, 潘艳, 等. 基于改进深度确定性策略梯度算法的电压无功优化策略[J]. 智慧电力, 2024, 52(5): 1-7, 30.
LI Fuqiang, ZHANG Wenchao, PAN Yan, et al. Reactive voltage optimization strategy based on improved depth deterministic strategy gradient algorithm[J]. Smart Power, 2024, 52(5): 1-7, 30.
- [31] 杨挺, 赵黎媛, 刘亚闯, 等. 基于深度强化学习的综合能源系统动态经济调度[J]. 电力系统自动化, 2021, 45(5): 39-47.
YANG Ting, ZHAO Liyuan, LIU Yachuang, et al. Dynamic economic dispatch of integrated energy systems based on deep reinforcement learning[J]. Automation of Electric Power Systems, 2021, 45(5): 39-47.
- [32] 朱涛, 海迪, 李文云, 等. 基于深度确定性策略梯度算法的配电网最优电压实时控制[J]. 电网与清洁能源, 2024, 40(6): 121-129.
ZHU Tao, HAI Di, LI Wenyun, et al. Real-time control of optimal voltage in distribution networks based on deep deterministic policy gradient algorithm[J]. Power System and Clean Energy, 2024, 40(6): 121-129.
- [33] 周雪松, 张心茹, 赵淦宇, 等. 基于DDPG算法的微网负载端接口变换器自抗扰控制[J]. 电力系统保护与控制, 2023, 51(21): 66-75.
ZHOU Xuesong, ZHANG Xinru, ZHAO Hanyu, et al. Active disturbance rejection control of microgrid load-side interface converter based on DDPG algorithm[J]. Power System Protection and Control, 2023, 51(21): 66-75.
- [34] 蒲天骄, 乔骥, 赵紫璇, 等. 面向电力系统智能分析的机器学习可解释性方法研究(一): 基本概念与框架[J]. 中国电机工程学报, 2023, 43(18): 7010-7030.
PU Tianjiao, QIAO Ji, ZHAO Zixuan, et al. Research on interpretable methods of machine learning applied in intelligent analysis of power system (part I): basic concept and framework[J]. Proceedings of the CSEE, 2023, 43(18): 7010-7030.
- [35] 李志颖, 王鸿, 王致杰. 考虑功率可调节裕度的综合能源系统云边协同优化[J]. 上海电机学院学报, 2024, 27(3): 125-130.
LI Zhiying, WANG Hong, WANG Zhijie. Cloud-edge collaborative optimization of integrated energy systems considering power adjustment margin[J]. Journal of Shanghai Dianji University, 2024, 27(3): 125-130.

收稿日期: 2025-02-27; 修回日期: 2025-07-14

作者简介:

窦嘉铭(1996—), 男, 博士研究生, 主要研究方向为电力系统智能调度及其可解释性; E-mail: djmmjddjmmjd@163.com

王小君(1978—), 男, 博士, 教授, 博士生导师, 主要研究方向为电力系统分析与控制、综合能源系统优化运行; E-mail: xjwang1@bjtu.edu.cn

司方远(1992—), 男, 通信作者, 博士, 副教授, 主要研究方向为低碳综合能源系统的优化规划、运行与控制。E-mail: fysi@bjtu.edu.cn

(编辑 周金梅)