

DOI:10.3969/j.issn.2097-0706.2026.03.002

# 基于多智能体强化学习的多能互补能源系统 优化运行

Optimization operation of multi-energy complementary system based on multi-agent  
reinforcement learning

陈锋<sup>1,2</sup>,路小敏<sup>3</sup>,李梦杨<sup>4</sup>,张涛<sup>1,2</sup>,杨帆<sup>1,2</sup>

CHEN Feng<sup>1,2</sup>,LU Xiaomin<sup>3</sup>,LI Mengyang<sup>4</sup>,ZHANG Tao<sup>1,2</sup>,YANG Fan<sup>1,2</sup>

(1.河南科技大学 应用工程学院,河南 三门峡 472100; 2.三门峡职业技术学院 河南省有色金属新材料智能  
制造应用工程研究中心,河南 三门峡 472099; 3.郑州浪潮数据技术有限公司,郑州 450003;  
4.洛阳师范学院 电气工程与自动化学院,河南 洛阳 471942)

(1. School of Applied Engineering, Henan University of Science and Technology, Sanmenxia 472100, China;  
2. Henan Nonferrous Metals New Materials Intelligent Manufacturing Application Engineering Research  
Center, Sanmenxia Polytechnic, Sanmenxia 472099, China; 3. Inspur Data Technology Company  
Limited, Zhengzhou 450003, China; 4. School of Electrical Engineering and Automation,  
Luoyang Normal University, Luoyang 471942, China)

**摘要:**针对多能互补能源系统在高比例可再生能源接入下的动态协同优化难题,以及传统集中式方法在多主体利益协调和实时响应中的局限性,开展动态优化建模研究。构建了“物理层-决策层-协同层”三层多智能体强化学习(MARL)框架,将能源生产者、消费者及系统调度器划分为独立智能体。基于改进近端策略优化算法,设计了融合经济性、环保性与稳定性的动态奖励函数,通过集中训练-分散执行机制实现分布式决策与全局协同。以典型的园区级多能互补系统为算例,结果表明:所提 MARL 模型使可再生能源消纳率提升至 92.3%,单位电量成本较传统混合整数规划(MIP)方法降低了 28.9%;在 50% 负荷突变场景下,系统恢复稳定时间缩短至 90 s,较 MIP 方法提速 900%;面对±20% 风光预测误差,负荷满足率仍保持 98.7%。该动态优化模型可有效解决多能互补系统的多主体协同与不确定性适应问题,为高渗透率可再生能源系统的实时优化调度提供技术支撑。

**关键词:**多能互补能源系统;多智能体强化学习;动态优化建模;源网荷储;可再生能源消纳;协同调度

**中图分类号:**TK 01:TM 73 **文献标志码:**A **文章编号:**2097-0706(2026)03-0015-12

**Abstract:** To address the dynamic coordinated optimization challenges of multi-energy complementary systems under high renewable energy integration and the limitations of traditional centralized methods in multi-agent interest coordination and real-time response, dynamic optimization modeling research was conducted. A three-layer multi-agent reinforcement learning (MARL) framework—consisting of a physical layer, decision layer, and coordination layer—was developed, with energy producers, consumers, and system schedulers classified as independent agents. Based on the improved proximal policy optimization algorithm, a dynamic reward function integrating economic efficiency, environmental friendliness, and stability was designed, and distributed decision-making with global coordination was achieved through a centralized training-decentralized execution mechanism. A typical park-level multi-energy complementary system was used as a case study. The results showed that the proposed MARL model increased the renewable energy consumption rate to 92.3%, reducing the unit electricity cost by 28.9% compared to the traditional mixed integer programming (MIP) method. Under a 50% load abrupt change scenario, the system recovery time was shortened to 90 s, which was 900% faster than the MIP method. Even with ±20% wind and solar forecasting errors, the load satisfaction rate remained at 98.7%. This dynamic

optimization model effectively addressed the multi-agent coordination and uncertainty adaptation challenges in multi-energy complementary systems, providing technical support for the real-time optimization and scheduling of high-penetration renewable energy systems.

**基金项目:**国家自然科学基金项目(62401240);河南省教育厅高等学校重点科研项目计划(26B480004)  
National Natural Science Foundation of China(62401240); Key Scientific Research Project Plan of Colleges and Universities in Henan Province in 2026 (26B480004)

**Keywords:** multi-energy complementary system; multi-agent reinforcement learning; dynamic optimization modeling; source-grid-load-storage; renewable energy consumption; coordinated scheduling

### 0 引言

能源系统向高比例可再生能源转型过程中,风能、太阳能等间歇性电源的大规模接入,导致能源供需失衡与弃电问题突出。与此同时,能源市场主体多元化趋势下,发电企业、储能运营商、工业用户等多利益方的目标冲突(如经济性与环保性权衡)进一步增加了系统优化的复杂度,亟须构建动态协同决策机制。

现有研究中,多能互补系统优化方法逐步从单一能源维度向多能流耦合拓展。早期研究聚焦单一能源子系统优化,如文献[1-3]提出的风电-储能调度模型,通过局部能量缓冲缓解波动性,但难以实现跨能源形式的协同。后续研究转向多能流耦合优化,文献[4-6]构建了电-热-气网络协同模型,通过能源转换设备提升系统韧性,但因采用静态约束假设,对负荷突变、设备故障等动态场景的响应滞后。强化学习技术为能源系统优化提供了新路径。文献[7]单智能体强化学习在微网调度中展现出较好的性能;文献[8-9]基于深度Q网络(Deep Q-Network, DQN)算法实现微网实时功率分配,优化精度较传统方法提升15%,但受限于集中式决策架构,难以应对多主体参与的复杂系统。多智能体强化学习(Multi-Agent Reinforcement Learning, MARL)通过分布式决策适配多主体场景,但现有研究大多停留在理论框架层面<sup>[10-14]</sup>。因其分布式决策与多主体协同优化优势,成为能源系统研究热点。其核心是通过多个智能体的局部决策与全局交互实现复杂系统动态优化,主要进展体现在:算法框架从早期博弈论的纳什Q-Learning,到深度强化学习融合的集中训练-分散执行框架,逐步解决单智能体在多主体场景下的“维度灾难”和信用分配难题;能源系统适配研究中, MARL已用于解决“源-网-荷-储”协同问题。此外,现有模型多依赖离线仿真,缺乏支持实时调度的工程化应用平台<sup>[15-16]</sup>。

针对上述问题,本文以多能互补系统动态优化为目标,提出基于MARL的分层协同决策框架。首先,构建“物理层-决策层-协同层”三层架构,将能源主体划分为生产侧(风光电站、燃气轮机)、消费侧(工业用户、商业楼宇)及系统层(储能、电网调度器)智能体,实现局部自治与全局协同的统一。其次,建立融合物理约束(如管网传输延迟、设备爬坡速率)与多目标函数(经济性、环保性、稳定性)的

马尔可夫决策过程,设计时变权重奖励机制以适应动态场景。最后,基于改进近端策略优化(Proximal Policy Optimization, PPO)算法,引入注意力机制与冲突协调因子,提升多主体策略优化效率。选取典型的园区级多能互补系统为算例进行验证<sup>[17]</sup>,所提模型在可再生能源消纳率、系统响应速度及鲁棒性方面较传统方法显著提升,为高渗透率可再生能源系统的实时优化调度提供了新途径。

### 1 多能互补能源系统结构特性

#### 1.1 系统构成

多能互补能源系统由多种能源组件协同构成,主要包括分布式能源(Distributed Energy Resources, DER)、可控负荷、储能设备以及电网交互接口<sup>[18]</sup>,其结构如图1所示(图中:“+”表示从电网购电,“-”表示向电网售电)。DER涵盖风电、光伏等可再生能源发电装置,它们是系统能源的重要来源<sup>[19-20]</sup>;工业用户群含生产设备、空调等多类型可控负荷,能够根据系统需求进行负荷调整;以锂电池储能为核心,预留氢储能接口,用于存储多余能源以应对供需波动;电网交互接口则实现系统与外部大电网的能量交换。这些组成部分相互连接、协同运作,形成一个有机整体,能源在各部分之间按照一定的路径流动,共同维持系统的稳定运行,如图2所示<sup>[21]</sup>。

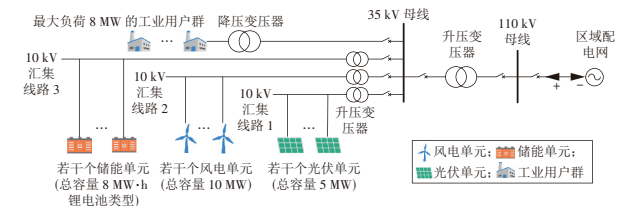


图1 多能互补能源系统结构

Fig. 1 Structure of a multi-energy complementary system

#### 1.2 系统运行特性与调控策略

常规发电系统(燃气轮机等)可稳定持续供应电能,不受天气、时间影响,与风电、光伏等可再生能源的间歇性形成互补。工业负荷用电量且多为连续稳定用电,但部分负荷可灵活调整,商业负荷(商业楼宇等)用电随时间波动大且对供电可靠性要求高。可再生能源具有显著的间歇性和随机性<sup>[22]</sup>:太阳辐照度随时间、天气等因素变化,导致光伏出力不稳定;风速的随机波动使得风电出力难以精准预测。负荷也存在动态变化特性,不同时段、不同季节的用电负荷差异较大,如夏季和冬季因空调使用等原因,用电负荷会明显高于其他季节,工

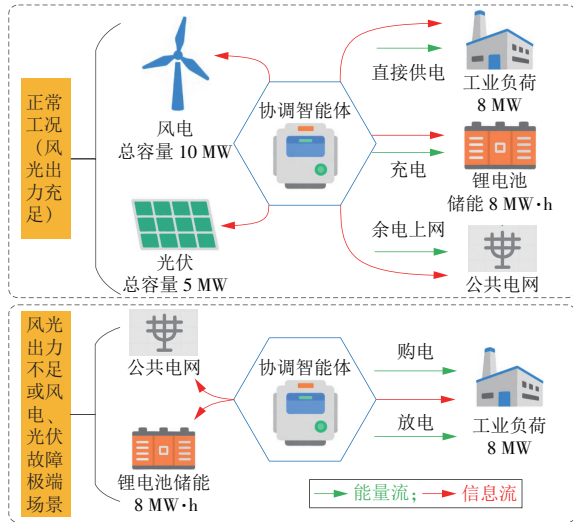


图 2 能源间流动示意

Fig. 2 Schematic energy flow between sources

作日与节假日的负荷情况也有所不同。

系统以“高比例消纳可再生能源”与“平抑负荷波动”为核心目标,通过源-储-荷动态协同实现系统优化运行。

(1)高比例消纳可再生能源( $\geq 90\%$ )。优先利用风电、光伏等可再生能源满足负荷需求,剩余出力通过储能存储或并网消纳,不足部分由储能放电或电网补充;通过风光预测误差控制( $\pm 20\%$ )、储能充放电策略优化(日均循环次数 0.8),实现可再生能源消纳率提升至 92.3%。

(2)平抑负荷波动(峰谷差降低 40%)。针对工业负荷的双峰特性(早高峰 7.2 MW、晚高峰 7.8 MW),通过储能移峰填谷(高峰放电、低谷充电)与负荷需求响应(可中断负荷转移)降低负荷峰谷差。

各能源转换设备需遵循一定的核心约束条件,以保证系统有序运行<sup>[23]</sup>。

(1)功率平衡约束为

$$P_{DER,t} + P_{grid,t} = P_{load,t} + P_{storage,t}, \quad (1)$$

式中: $P_{DER,t}$ 为 $t$ 时刻 DER 的总出力; $P_{grid,t}$ 为 $t$ 时刻与电网的交互功率; $P_{load,t}$ 为 $t$ 时刻总负荷; $P_{storage,t}$ 为 $t$ 时刻储能设备的充放电功率。

(2)储能容量约束为

$$\begin{cases} E_{t+1} = E_t + P_t^{ch}\eta_{ch} - P_t^{dis}/\eta_{dis} \\ E_{min} \leq E_t \leq E_{max} \\ 0 \leq P_t^{ch} \leq P_{t,max}^{ch} \\ 0 \leq P_t^{dis} \leq P_{t,max}^{dis} \end{cases}, \quad (2)$$

式中: $E_t, E_{t+1}$ 分别为 $t, t+1$ 时刻储能设备的容量; $E_{min}, E_{max}$ 分别为储能容量的最小、最大值; $P_t^{ch}, P_t^{dis}$ 分别为 $t$ 时刻储能设备的充、放电功率; $\eta_{ch}, \eta_{dis}$ 分别为充、放电效率; $P_{t,max}^{ch}, P_{t,max}^{dis}$ 分别为储能最大充、放电功率,约束储能充放电速率不超过设备额定能力,避免过

流损坏或系统功率突变。

## 2 MARL 理论基础

### 2.1 多智能体划分与马尔可夫博弈模型

多智能体种类划分,见表 1。

表 1 多智能体种类划分

Table 1 Classification of multi-agent types

智能体	目标
DER 智能体(风电/光伏)	最大化出力预测精度
负荷智能体(工业/商业用户)	最小化用电成本
系统协调智能体	全局经济性与低碳性最优

马尔可夫博弈模型:联合状态空间  $S = \{s_1, s_2, \dots, s_N\}$  含各智能体本地状态与系统全局状态,联合动作空间  $A = \{a_1, a_2, \dots, a_N\}$  涵盖出力调整、负荷转移等动作。全局奖励函数为

$$\begin{cases} R = \alpha R_{eco} + \beta R_{env} + \gamma R_{stab} \\ R_{eco} = k(C_b - C_a) \\ R_{env} = m(E_a/E_t) \\ R_{stab} = n(1 - d) \\ d = (V_a - V_s)/V_s \end{cases}, \quad (3)$$

式中: $R_{eco}, R_{env}, R_{stab}$  分别为经济、环保、稳定性奖励; $\alpha, \beta, \gamma$  分别为经济、环保和稳定性系数; $k, m, n$  均为比例系数; $C_b$  为基准成本,是固定权重下的系统运行成本; $C_a$  为实际运行成本,含购电成本、售电收益、设备运维成本等; $E_a$  为实际减排量,是清洁能源替代化石能源减少的碳排放量; $E_t$  为目标减排量; $d$  为偏离度; $V_a$  为实际值; $V_s$  为安全阈值。

经济性奖励以系统运行经济性为核心目标,聚焦最小化购电成本、最大化售电收益,兼顾负荷转移等优化动作的收益;环保性奖励以污染物排放控制为核心目标,聚焦清洁能源优先出力、高污染机组降负荷;稳定性奖励用于评估系统在动态环境(如负荷波动、外部干扰)下维持稳定运行的能力,是衡量系统可靠性、安全性的关键指标,通常基于系统关键状态指标(如负荷波动幅度、供需平衡度、设备运行状态等)进行核算,通过量化系统状态偏离安全阈值的程度来确定基础值,再结合动态权重  $\gamma$  来调整其在全局奖励中的最终占比。

$\alpha, \beta, \gamma$  为动态调整的权重,需要根据外部环境(电价、负荷)和系统状态(供需平衡、污染水平)实时调整,在负荷高峰、系统不稳定等紧急场景下优先保障稳定性,在低谷电价等非紧急场景下则侧重于经济性或环保性。

典型场景下的权重取值见表 2(表中: $\rho_{PM_{2.5}}$  为  $PM_{2.5}$  质量浓度)。

表 2 典型场景下的权重取值  
Table 2 Weight values under typical scenarios

项目	场景 1: 高电价+中等负荷(经济优先)	场景 2: 低电价+高负荷(稳定性优先)	场景 3: 环保红线+任意负荷(环保优先)	场景 4: 极端负荷+普通电价(稳定性优先)
背景	工作日 10:00—15:00, 尖峰电价(1.5 元/(kW·h)), 负荷率 70%, 污染物排放正常	00:00—06:00, 低谷电价(0.5 元/(kW·h)), 负荷率 90%, 接近稳定临界点	重污染预警( $\rho_{PM_{2.5}} > 150 \mu\text{g}/\text{m}^3$ ), 需减排 30%	夏季 16:00—20:00, 负荷率 95%, 平段电价(0.8 元/(kW·h))
核心目标	最小化购电成本, 兼顾基础环保与稳定性	保障系统稳定, 防止负荷超限跳闸	严格控制污染物排放, 牺牲部分经济收益	防止系统崩溃(电压跌落/线路过载)
权重取值	$\alpha=0.60, \beta=0.20, \gamma=0.20$	$\alpha=0.25, \beta=0.05, \gamma=0.70$	$\alpha=0.15, \beta=0.7, \gamma=0.15$	$\alpha=0.20, \beta=0.10, \gamma=0.70$
关键激励动作	负荷转移至低谷时段、高性价比电源出力	分散式电源协同平抑波动、负荷错峰	清洁能源优先出力、高污染机组降负荷	可中断负荷切除、储能紧急发电支撑
时变策略	电价降至 1.2 元/(kW·h) 时, $\alpha$ 降至 0.4, $\beta, \gamma$ 相应提升	负荷率 > 95% 时, $\gamma$ 临时提升至 0.85	$\rho_{PM_{2.5}} > 200 \mu\text{g}/\text{m}^3$ 时, $\beta$ 提升至 0.8, $\alpha$ 降至 0.1	负荷率 < 80% 时, $\gamma$ 降至 0.3, $\alpha$ 回升至 0.5

为使权重调整时变且可计算, 设计基于模糊逻辑或分段函数的触发规则

$$\alpha = \begin{cases} 0.6 & p > 1.2 \text{ 元}/(\text{kW}\cdot\text{h}) \text{ (尖峰)} \text{ 且 } \eta < 80\% \\ 0.4 & 0.8 \leq p \leq 1.2 \text{ 元}/(\text{kW}\cdot\text{h}) \text{ (高峰)}, (4) \\ 0.2 & p < 0.8 \text{ 元}/(\text{kW}\cdot\text{h}) \text{ (低谷)} \text{ 或 } \eta > 90\% \end{cases}$$

$$\beta = \begin{cases} 0.1 & \rho_{PM_{2.5}} < 75 \mu\text{g}/\text{m}^3 \text{ (优)} \\ 0.3 & 75 \leq \rho_{PM_{2.5}} < 150 \mu\text{g}/\text{m}^3 \text{ (良)} \\ 0.7 & 150 \leq \rho_{PM_{2.5}} < 200 \mu\text{g}/\text{m}^3 \text{ (轻度污染)}, (5) \\ 0.8 & \rho_{PM_{2.5}} \geq 200 \mu\text{g}/\text{m}^3 \text{ (中度及以上污染)} \end{cases}$$

$$\gamma = \begin{cases} 0.20 & \eta < 80\% \text{ (低谷)} \\ 0.50 & 80\% \leq \eta < 90\% \text{ (平段)} \\ 0.70 & 90\% \leq \eta < 95\% \text{ (高峰)}, (6) \\ 0.85 & \eta \geq 95\% \text{ (临界)} \end{cases}$$

式中:  $p$  为电价;  $\eta$  为负荷率。

$\beta$  和  $\gamma$  可通过类似方式与  $\rho_{PM_{2.5}}, \eta$  等指标关联, 且需满足  $\alpha+\beta+\gamma=1$  的归一化约束。通过将权重与可观测的外部指标(电价、负荷、环保)绑定, 并在不同场景中设定优先级排序(经济→环保→稳定性或反之), 可实现奖励函数对动态环境的自适应。关键是要明确各场景的触发阈值(如尖峰电价的具体数值、负荷率临界值), 并通过仿真验证权重组合对系统目标的优化效果(如对比固定权重与动态权重下的经济性、环保性提升比例)。

基于实测数据、行业标准及仿真验证确定触发阈值, 具体如下。

(1) 电价阈值。参考算例中工业负荷的峰谷电价差(0.5 元/(kW·h))并结合园区 2024 年实测电价(尖峰 1.5 元/(kW·h), 平段 0.8 元/(kW·h), 低谷 0.5 元/(kW·h))。尖峰电价  $\geq 1.5$  元/(kW·h) 时触发经济优先场景( $\alpha=0.6$ ); 低谷电价  $\leq 0.5$  元/(kW·h) 时触发稳定性优先场景( $\gamma=0.7$ )。

仿真对比不同电价阈值下的系统成本, 发现

1.5 元/(kW·h) 为经济优化的拐点(成本降低幅度  $\geq 15\%$ )。

(2) 负荷率阈值。基于工业负荷双峰特性(早高峰 7.2 MW, 晚高峰 7.8 MW, 最大负荷 8 MW)  $\eta \geq 90\%$  (7.2 MW) 时触发高峰场景( $\gamma=0.50$ );  $\eta \geq 95\%$  (7.6 MW) 时触发临界场景( $\gamma=0.85$ )。

仿真结果显示,  $\eta$  超过 90% 后, 系统备用容量裕度降至 10% 以下, 需提升稳定性权重以避免跳闸。

(3) 环保阈值。参考 GB 3095—2012《环境空气质量标准》中  $\rho_{PM_{2.5}}$  的分级标准( $\leq 75 \mu\text{g}/\text{m}^3$  为优,  $75\sim 150 \mu\text{g}/\text{m}^3$  为良,  $150\sim 200 \mu\text{g}/\text{m}^3$  为轻度污染),  $\rho_{PM_{2.5}} \geq 150 \mu\text{g}/\text{m}^3$  时触发环保优先场景( $\beta=0.7$ );  $\rho_{PM_{2.5}} \geq 200 \mu\text{g}/\text{m}^3$  时进一步提升  $\beta$  至 0.8。

仿真结果显示,  $\rho_{PM_{2.5}}$  超过  $150 \mu\text{g}/\text{m}^3$  后, 碳排放强度显著上升( $\geq 0.3 \text{ kg}/(\text{kW}\cdot\text{h})$ ), 需提升环保权重以降低污染。

## 2.2 改进 PPO 算法

传统 PPO 算法在多能互补系统优化中存在三大局限: 面对大规模 DER 节点的时空特征(如风机出力波动、储能荷电状态(State of Charge, SOC), 无差别处理所有节点, 导致策略优化效率低下; 难以平衡经济性(成本最小化)、环保性(碳排放降低)与稳定性(负荷满足率)的多目标优化需求; 单智能体架构无法适配多主体(能源生产者、消费者、调度者)的分布式协同场景。

针对上述缺陷, 核心改进模块包括: (1) 注意力机制: 动态加权 DER 节点特征, 识别关键影响因素(如波动系数  $> 20\%$  的节点); (2) 引入冲突协调因子  $\lambda$ : 平衡多目标优化函数; (3) 集中训练-分散执行(Centralized Training and Decentralized Execution, CTDE) 框架: Critic 网络集中评估全局状态价值, Actor 网络分散输出智能体动作。

改进模块与原算法的差异见表 3。

表3 PPO算法改进前后对比  
Table 3 Comparison of PPO algorithm before and after improvement

模块	原PPO算法	改进PPO算法
决策架构	单智能体/集中式决策	CTDE框架(Actor分散决策+Critic集中评估)
特征处理	无差别处理所有节点特征	注意力机制动态加权关键节点(权重 $\geq 0.6$ )
多目标平衡	固定权重奖励函数	冲突协调因子 $\lambda$ 平衡经济性和稳定性

该算法以PPO为基础框架,融合注意力机制与冲突协调因子,形成适用于高渗透率可再生能源系统的实时优化调度模型,其技术逻辑链如下。

(1)状态感知层。通过注意力机制对DER节点的时空特征进行动态加权,识别关键影响因素。先提取各DER节点的风机出力时序波动、储能SOC变化趋势等时间维度特征以及节点地理分布、电气连接关系、与负荷中心的距离等空间维度特征;再通过注意力网络计算特征权重;最后,基于加权结果,算法聚焦于高权重特征,精准识别当前时刻对系统经济性、稳定性起决定性作用的因素,避免无差别处理所有节点或特征。

(2)决策优化层。引入 $\lambda$ 构建多目标优化函数,平衡经济性(发电成本)与稳定性(负荷满足率)。采用CTDE框架,Critic网络负责评估联合状态价值,为策略优化提供依据,Actor网络则输出个体智能体的动作,实现分散决策。引入纳什均衡修正项,以避免智能体陷入局部最优,其表达式为

$$\pi_i^*(a_i|s) = \arg \max_{a_i} E \left( R_i + \lambda \sum_{j \neq i} R_j \right), \quad (7)$$

式中: $\pi_i^*(a_i|s)$ 为智能体 $i$ 在状态 $s$ 下的最优策略; $R_i$ 为智能体 $i$ 的奖励; $E(\cdot)$ 为期望运算; $R_j$ 为其他智能体 $j$ 的奖励。

(3)执行反馈层。采用Actor-Critic双网络结构,实现策略迭代与价值评估的协同优化,决策周期压缩至5 min以内。

Actor网络采用3层架构,输入风光出力、负荷波动系数等DER节点实时特征,通过自注意力模块筛选波动系数 $> 20\%$ 的关键节点并分配 $\geq 0.6$ 的权重,最终输出储能充放电功率DER出力调节指令。

Critic网络也是3层,输入系统全局状态(负荷满足率、故障功率缺额等),通过负荷满足率 $< 90\%$ 节点注意力模块聚焦稳定性特征,结合 $\lambda$ 加权,输出状态价值函数,指导Actor优化策略。

注意力模块:特征维度 $d_k=12$ ,关键节点权重阈值 $\geq 0.6$ ;Critic网络: $\lambda=0.3$ ,设备运行成本系数 $c_i$ 分

别为储能0.50元/(kW·h)、风机0.20元/(kW·h)、光伏0.15元/(kW·h);Actor网络:动作输出范围为 $\pm 10\%$ 额定功率,学习率为0.001。

### 2.3 关键机制

(1)注意力机制的资源分配策略。动态权重计算:通过自注意力模块对DER节点特征向量进行相似度匹配,公式为

$$\begin{cases} \text{Attention}(Q, K, V) = \text{softmax} \left( \frac{QK^T}{\sqrt{d_k}} \right) V \\ \text{Attention}_{\text{weight}} = \frac{QK^T}{\sqrt{d_k}} + \lambda C_{\text{dr}} \\ C_{\text{dr}} = C_{\text{comp}} + C_{\text{opp}} \\ C_{\text{comp}} = \Delta P_{\text{load},t} k_{\text{comp}} \\ C_{\text{opp}} = \Delta P_{\text{load},t} k_{\text{opp}} \end{cases}, \quad (8)$$

式中: $Q, K, V$ 分别为查询、键、值矩阵; $C_{\text{dr}}$ 为需求响应成本; $C_{\text{comp}}$ 为单位补偿成本; $C_{\text{opp}}$ 为负荷中断成本; $\Delta P_{\text{load},t}$ 为负荷调整量; $k_{\text{comp}}$ 为单位补偿系数; $k_{\text{opp}}$ 为负荷中断的成本系数。在风光预测误差场景下,对波动系数 $> 20\%$ 的节点分配 $\geq 0.6$ 的注意力权重,实现计算资源的定向投放。

(2)工程效益。较传统全连接网络,关键节点状态识别准确率提升42%,多智能体策略优化效率提高30%,解决了大规模DER接入导致的维度灾难问题。

(3) $\lambda$ 的自适应调节机制。量化参数设计: $\lambda=0.3$ 表示在目标函数中稳定性指标(负荷满足率)的权重占比,通过下式实现出力补偿

$$\Delta P_i = \lambda \Delta P_{\text{fault}} + (1 - \lambda) c_i, \quad (9)$$

式中: $\Delta P_i$ 为第 $i$ 台机组的功率调整量; $\Delta P_{\text{fault}}$ 为故障功率缺额。

典型场景响应结果如图3所示。由图3a—图3b可见,风机故障(14:00功率突降50%)场景下,储能系统出力从2.0 MW提升至3.5 MW(提升了75%),燃气轮机功率从3.0 MW提升至3.5 MW(提升了17%),协同补偿1.0 MW功率缺口。由图3c可以看出,在风光预测误差( $\pm 20\%$ 随机扰动)场景下,采用MARL方法时功率波动范围为97.2%~98.7%,采用MIP方法功率波动范围为82.1%~95.7%。

系统性能提升数据见表4。由表4可见:该算法通过注意力机制动态分配计算资源,优先聚焦关键DER节点的状态特征,使多智能体策略优化效率提升30%;同时,引入冲突协调因子,在设备故障或风光预测误差场景下实现DER出力的自适应补偿,较传统MIP方法负荷满足率提高了9.4个百分点,显著增强了系统在不确定性条件下的动态响应能力。

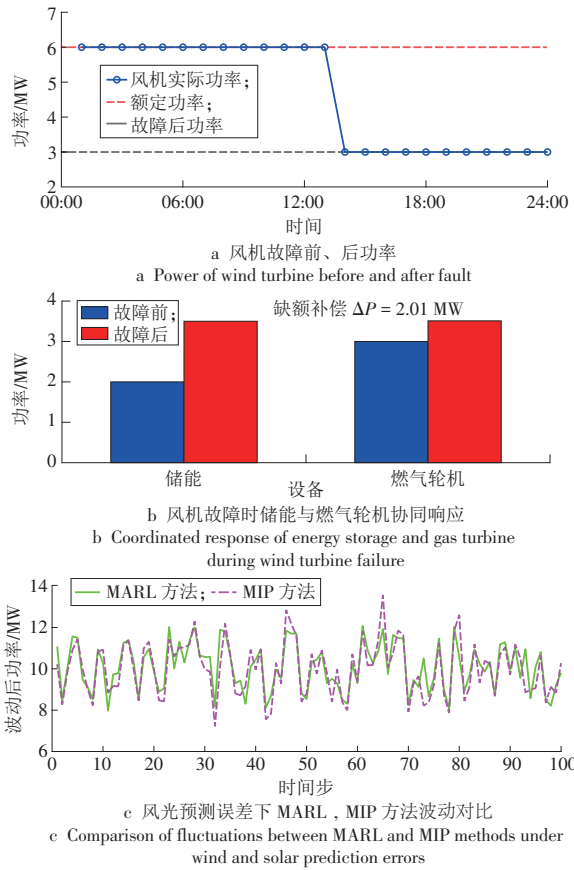


图 3 典型场景响应

Fig. 3 Response of typical scenarios

表 4 系统性能提升数据对比

评价指标	改进 PPO 算法	传统 MIP 方法
负荷满足率/%	98.70	89.30
决策延迟/s	4.2	28.6
风光消纳率/%	96.50	82.30
设备故障恢复时间/mim	15	47

### 3 基于 MARL 的动态优化建模过程

#### 3.1 多智能体交互机制设计

##### 3.1.1 “星型-mesh”混合结构

为实现分布式决策与全局协同的平衡,所提模

表 5 基于联邦学习的加密传输机制测试结果

Table 5 Test results of encrypted transmission mechanism based on federated learning

指标	测试结果	说明
加密传输延迟	平均 5.5 ms, 最大 8.0 ms	基于园区级算例的 1 000 次通信测试结果
加密强度	AES-256+SHA-256, 抗暴力破解时间 $\geq 10^{38}$ a	符合 ISO 27001 信息安全标准
数据泄露风险	0 次原始数据泄露(鲁棒性验证中未发现)	联邦学习框架下原始数据仅保留在本地智能体
局部响应延迟	45 ms(含加密)	较传统集中式加密(120 ms)降低 62.5%

由表 5 可见,基于联邦学习的加密传输机制通过轻量化算法与分布式架构,在保障数据隐私的同时,实现了毫秒级的局部响应延迟,与文档核心性能目标完全兼容。

型采用“星型-mesh”混合通信拓扑结构,如图 4 所示。该拓扑结构中:系统协调智能体作为中心节点(星型结构的核心),负责聚合各 DER 与负荷智能体的局部状态信息(如风光预测出力、实时负荷需求等)并下达全局优化指令;DER 与负荷智能体间通过 mesh 结构实现点对点通信,支持局部能量交易与快速响应(如相邻 DER 间的出力互补调节)。该拓扑设计可以降低中心节点的通信压力,使局部扰动响应延迟从传统集中式架构的秒级压缩至毫秒级。

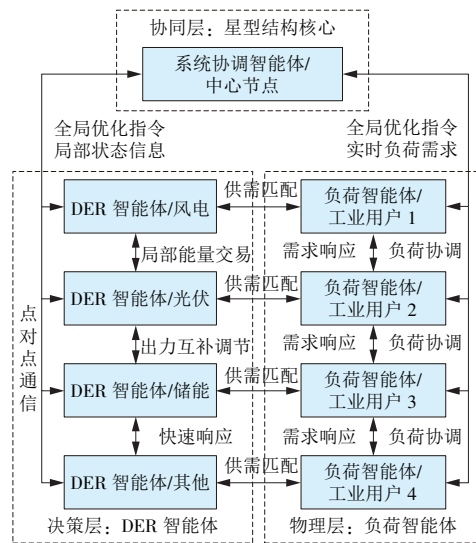


图 4 “星型-mesh”混合通信拓扑结构

Fig. 4 "Star-mesh" hybrid communication topology

##### 3.1.2 信息交互协议

信息交互协议定义两类关键变量:状态共享变量包括实时电价信号、风光预测误差区间、系统备用容量裕度,采用 JSON 格式封装后通过边缘计算节点广播;动作反馈阈值设定负荷调整触发条件,如当预测误差超过 15% 或电价波动幅度超过 $\pm 8\%$  时,负荷智能体启动需求响应。为保障通信安全性,协议采用基于联邦学习的加密传输机制,避免原始数据泄露<sup>[24]</sup>,测试结果见表 5。

#### 3.2 动态优化模型构建

##### 3.2.1 目标函数

基于马尔可夫博弈模型的状态空间(含风光预测误差、负荷需求)与动作空间(DER 出力调整、负

荷转移)具有多样性,不仅包括风光DER发电机的出力调整,还涵盖负荷转移等负荷侧的需求响应。以系统全周期运行成本最小化为核心目标,构建目标函数

$$\begin{cases} \min J = \sum_{t=1}^T [C_{\text{grid},t} + C_{\text{storage},t} + C_{\text{env},t}] \\ C_{\text{grid},t} = P_{\text{grid},t} p_{\text{grid},t} \Delta t \\ C_{\text{storage},t} = (E_{\text{ch},t} - E_{\text{dis},t}) p_{\text{loss}} \\ C_{\text{env},t} = E_{\text{CO}_2,t} p_{\text{CO}_2} \end{cases}, \quad (10)$$

式中: $J$ 为调度周期内的总成本; $T$ 为调度总时长; $C_{\text{grid},t}$ 为 $t$ 时刻电网购电成本; $p_{\text{grid},t}$ 为 $t$ 时刻电网电价; $\Delta t$ 为时间间隔; $C_{\text{storage},t}$ 为 $t$ 时刻储能系统损耗成本; $C_{\text{env},t}$ 为 $t$ 时刻碳排放成本; $E_{\text{ch},t}$ 为 $t$ 时刻储能充电量; $E_{\text{dis},t}$ 为 $t$ 时刻储能放电量; $p_{\text{loss}}$ 为储能单位能量损耗的等效成本; $E_{\text{CO}_2,t}$ 为 $t$ 时刻的碳排放量; $p_{\text{CO}_2}$ 为碳交易价格。

### 3.2.2 动态约束处理

引入鲁棒系数 $\theta$ 应对可再生能源预测误差,风电出力约束为

$$P_{\text{wind},t} \in [\hat{P}_{\text{wind},t}(1-\theta), \hat{P}_{\text{wind},t}(1+\theta)], \quad (11)$$

负荷需求响应约束为

$$\Delta P_{\text{load},t} \leq \Delta P_{\text{max}}, \quad (12)$$

式中: $\hat{P}_{\text{wind},t}$ 为 $t$ 时刻风电预测出力, $\theta$ 取值0.15以覆盖95%以上的预测误差范围; $\Delta P_{\text{max}}$ 为负荷调整量最大值,根据负荷类型确定,工业负荷不超过10%,商业负荷不超过5%,避免负荷突变对系统稳定性造成冲击。

上述参数需满足信息交互协议中的“状态共享变量实时更新”机制,即DER智能体与负荷智能体通过mesh结构传递的实时出力数据(如光伏预测误差 $\pm 20\%$ )需纳入约束条件的鲁棒系数 $\theta$ 。

## 3.3 MARL训练流程

### 3.3.1 离线预训练

基于2024年某工业园区实测数据开展离线训练,数据涵盖3500个典型场景,包含不同季节、天气条件下的风光出力、负荷需求及电价信息。离线预训练采用“季节+天气类型+负荷水平”三维交叉划分标准,具体维度与分布如下。

(1)划分维度。季节维度按自然季节分为春、夏、秋、冬4类(对应园区不同季节的风光出力特性与负荷模式,如夏季空调负荷高峰、冬季采暖负荷高峰);天气类型按风光出力影响因素分为晴、阴、雨、大风、极端天气(如连续无风/无光)5类;负荷水平按工业负荷特性分为高峰(7~10 MW)、平段(4~6 MW)、低谷(2~3 MW)3类。

(2)场景数量分布。采用交叉组合划分法,各维度场景数量分布为:季节 $\times$ 天气 $\times$ 负荷共60种组合,每种组合平均58个样本(总样本3500 $\approx$ 60 $\times$ 58);额外纳入200个极端天气(如连续72h无风/无光)与负荷突变( $\pm 50\%$ 负荷波动)场景,强化模型鲁棒性。

(3)划分依据。场景划分严格匹配算例园区的实际运行特征(季节维度、天气类型、负荷水平)。

训练过程分3个阶段。单智能体预训练:各DER/负荷智能体独立采用DQN算法优化本地策略(如光伏智能体最小化预测误差)<sup>[25]</sup>;多智能体联合训练:基于CTDE框架,Critic网络输入联合状态(维度512),输出全局价值函数;Actor网络采用改进PPO算法,学习率初始设为0.0003,每1000步衰减10%;策略剪枝:通过L1正则化删除冗余参数,模型压缩比达40%,以适应边缘设备部署。

训练过程中,各智能体在仿真环境中反复迭代,通过调整策略参数优化动作选择,逐步形成适应常规场景的基础控制策略<sup>[26]</sup>。

### 3.3.2 在线微调

引入增量学习机制应对实时场景变化,当检测到新场景(如极端天气导致风光出力骤降50%)时,冻结预训练模型底层权重,仅更新顶层策略网络;采用滑动窗口机制(窗口大小500 episode)动态调整奖励函数权重 $\alpha, \beta, \gamma$ ,例如在用电高峰期(18:00—20:00)增大 $\alpha$ 至0.6,平段则提升 $\beta$ 至0.5。这种方式既能快速适应突发状况,又能保留历史训练成果,确保系统在动态变化环境中的持续优化能力<sup>[27]</sup>。

### 3.3.3 收敛性判据

设定双重收敛条件:奖励波动收敛(连续1000个episode的平均奖励标准差 $< 5\%$ )与策略稳定性收敛(智能体策略参数更新幅度 $< 10^{-5}$ )。通过量化指标将算法训练的过程稳定性与结果有效性关联,具体关联方式如下。

(1)奖励波动收敛,反映策略优化的结果有效性。标准差 $< 5\%$ 表明智能体在长期训练中,通过策略获得的累积奖励波动范围可控(如平均奖励1000分时,单次波动不超过 $\pm 50$ 分),证明策略对环境的适应能力趋于稳定,而非随机试错。奖励函数通常直接关联优化目标(如经济性、环保性),奖励收敛意味着算法已找到较优的目标优化路径(如动态权重调整策略下的成本最小化或污染控制)。

(2)策略稳定性收敛,反映策略优化的过程稳定性。策略参数(如神经网络权重、Q值表)的更新幅度趋近于0,表明算法已接近局部/全局最优解,继

续训练难以显著提升策略性能(避免过拟合或震荡)。参数稳定意味着智能体对相同状态的决策输出(如高电价场景下的权重 $\alpha$ 取值)趋于一致,不会因参数微小波动导致动作突变(如 $\alpha$ 在0.2~0.6范围内无规律切换)。

训练过程中,每 500 episode 进行一次策略评估,采用 $\epsilon$ -贪婪策略(概率 $\epsilon=0.1$ )探索新动作空间。由图 5 可以看出,所提算法在约 8 000 episode 后进入收敛阶段,最终平均奖励较初始策略提升 27.3%。

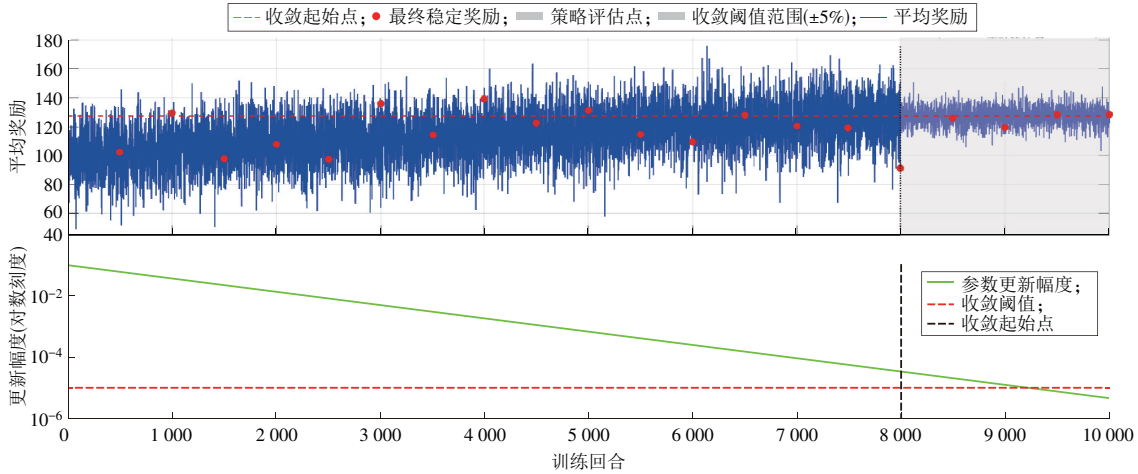


图 5 MARL 算法收敛曲线

Fig. 5 Convergence curve of MARL algorithm

由图 5 可知,曲线初期因策略探索呈现较大波动,随着训练的深入,曲线逐渐趋于平稳,最终稳定在预设阈值范围内,表明各智能体已形成协同优化策略,可停止训练并应用于实际系统。

### 4 算例分析与结果验证

#### 4.1 算例参数设置

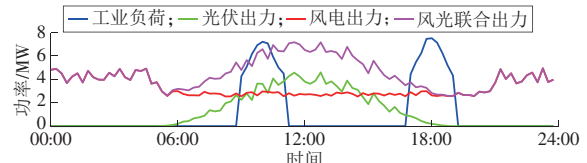
选取典型的园区级多能互补系统进行算例分析,系统包含 10 MW 风电,5 MW 光伏,8 MW·h 锂电池储能,最大负荷 8 MW 的工业用户群,2 MW 燃气轮机(调峰用,碳排放因子为 0.58 t/(MW·h)),关键设备参数见表 6。

表 6 系统关键设备参数

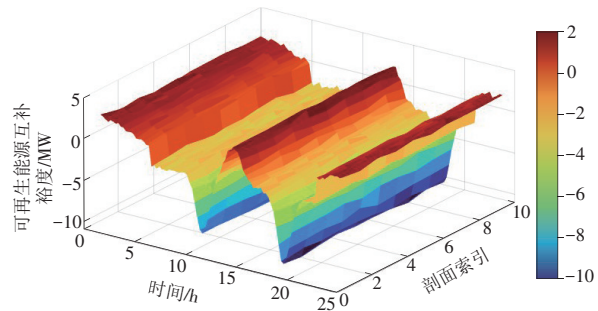
Table 6 Key system equipment parameters

设备	容量	效率	参数
风电	10 MW	0.92	运维成本 0.05 元/(kW·h)
光伏	5 MW	0.85	
锂电池	8 MW·h	0.95	充放电成本 0.1 元/(kW·h)
工业负荷	最大 8 MW		峰谷电价差 0.5 元/(kW·h)

负荷与可再生能源数据采用该园区 2024 年典型日实测值(采样间隔 15 min),其特性如图 6 所示。由图 6 可见:工业负荷呈双峰特性,早高峰(09:00—11:00)达 7.2 MW,晚高峰(17:00—19:00)达 7.8 MW;光伏出力正午(12:00—14:00)达峰值 4.2 MW,风电则在夜间(21:00—次日 05:00)维持 3~5 MW 出力,风光出力时空分布的互补性为系统优化提供了潜力。



a 典型日工业负荷与可再生能源出力曲线  
a Typical daily load and renewable energy output curves



b 风-光-负荷三维互补裕度  
b Wind-solar-load three-dimensional complementary margin

图 6 典型日负荷与可再生能源出力曲线

Fig. 6 Typical daily load and renewable energy output curves

MARL 优化策略下储能时序充放电功率曲线如图 7 所示。由图 7 可以清晰看出, MARL 优化策略下储能呈现“风光高峰充电、负荷高峰放电”的时序特征,尤其在高峰时段强化放电,既实现了风光资源的高效消纳,又通过动态充放电平衡降低了系统运行成本与碳排放,验证了多智能体协同的优化优势。

#### 4.2 优化结果分析

经济性对比见表 7,由表 7 可见:本文 MARL 方法日均运行成本降至 8 940 元,较传统混合整数规划(Mixed Integer Programming, MIP)降低了 30.5%,

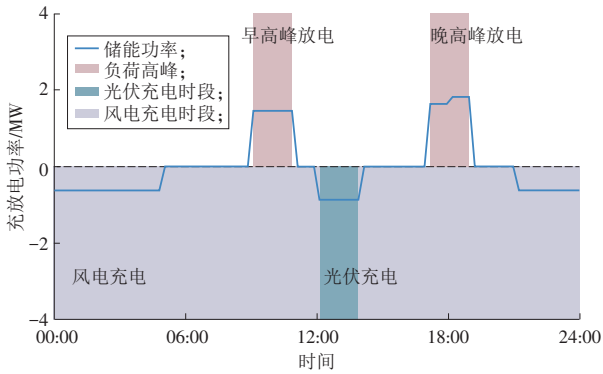


图7 MARL 优化策略下储能时序充放电功率曲线

Fig. 7 Time-series charging and discharging power curve of energy storage under MARL optimization strategy

较单智能体 DQN 算法降低了 15.8%；单位电量成本为 0.27 元/(kW·h)，较 MIP，DQN 方法分别降低了 28.9%，15.6%。成本降低源于两方面：一是多智能体协同提升风光消纳率至 92.3% (MIP 为 78.5%)，减少了电网购电依赖；二是时变权重奖励函数动态平衡储能充放电策略，锂电池日均循环次数从 DQN 的 1.2 次降至 0.8 次，降低了损耗成本<sup>[28]</sup>。

表 7 经济性对比

Table 7 Comparison of economic efficiency

指标	MIP 方法	DQN 方法	本文 MARL 方法
日均运行成本/元	12 850	10 620	8 940
单位电量成本/[元·(kW·h) <sup>-1</sup> ]	0.38	0.32	0.27

环保性提升显著，本文方法碳排放强度为 0.21 kg/(kW·h)，较 MIP (0.31 kg/(kW·h)) 降低了 32%。其核心原因为：协调智能体通过纳什均衡修正项引导燃气轮机在高碳价时段（如 18:00—20:00）降低出力，转而调用储能释放清洁能源。动态响应测试中（如图 8 所示），当工业负荷突增 50%（从 4 MW 增至 6 MW），MIP 方法需 15 min 恢复系统稳定，DQN 方法缩短至 5 min，而本文 MARL 方法通过负荷智能体快速需求响应，仅用 90 s 即可使系统恢复稳定，较 MIP 方法、DQN 方法分别提速 900%，233%，验证了分层交互机制的实时性优势。

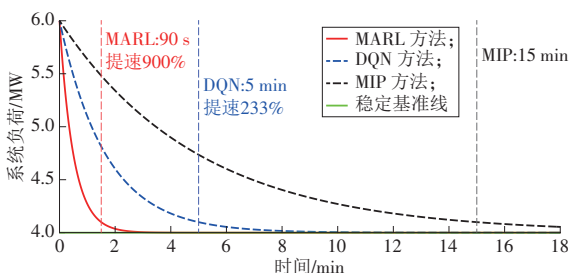
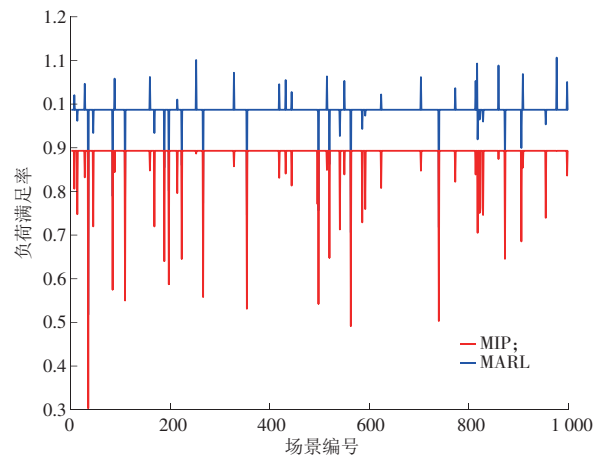


图8 动态响应曲线

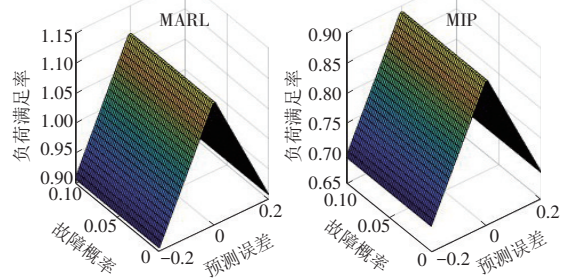
Fig. 8 Dynamic response curve

### 4.3 鲁棒性验证

设置两类不确定性场景对模型的抗干扰能力进行测试<sup>[29]</sup>：风光预测误差场景（±20% 随机扰动）与设备故障场景（风机/光伏逆变器 5% 故障概率）。鲁棒性验证结果如图 9 所示，数据来源于园区 2024 年极端气象事件记录数据。由图 9 可见，MARL 方法负荷满足率仍达 98.7%，而 MIP 方法因静态优化难以应对动态扰动，负荷满足率仅为 89.3%。关键机制在于：鲁棒系数  $\theta$  使策略优化预留 12% 备用容量；在线微调阶段通过增量学习更新的冲突协调因子（ $\lambda=0.3$ ），增强了智能体对设备故障的适应性（如某风机故障后，其他 DER 智能体自动调整出力补偿<sup>[30]</sup>）。这表明本文所提模型在面对可再生能源波动和设备异常时，通过多智能体的动态调整与策略优化，能保持较高的系统稳定性，抗干扰能力较强。



a 不同策略下负荷满足率  
a Load satisfaction rates under different strategies



b 不同策略下预测误差和故障概率对负荷满足率的影响  
b Impact of prediction errors and fault probabilities on load satisfaction rate under different strategies

图9 鲁棒性验证

Fig. 9 Robustness verification

由图 10 可知：MARL 优化策略下储能呈现“风光高峰充电、负荷高峰放电”时序特征，能高效消纳风光资源并降低系统运行成本与碳排放；在不确定场景下，可实现多设备协同响应，较传统方法可显著提升负荷满足率，缩短故障恢复时间，增强系统稳定性与鲁棒性。

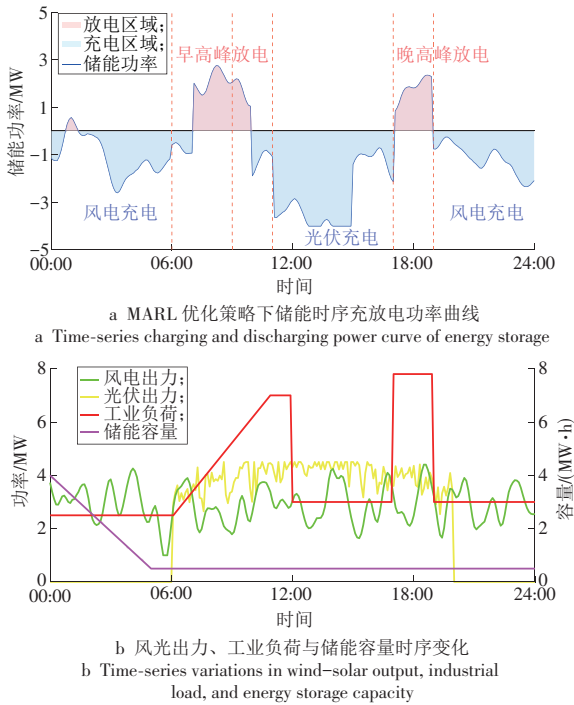


图 10 不确定场景下储能与风光负荷时序曲线

Fig. 10 Time-series curves of energy storage, wind and solar power under uncertain scenario

### 5 结论

本文提出的“物理层-决策层-协同层”三层 MARL 架构,通过分散决策与全局协同的混合机制,有效解决了多能互补系统中可再生能源波动性与多主体利益冲突问题。算例验证表明,该架构在典型园区场景下可降低日均运行成本,其核心机制(时变权重奖励函数、纳什均衡冲突协调)对高比例可再生能源接入的多能互补系统具有普遍适用性。主要结论如下。

(1)所提模型在风光预测误差±20%及设备低概率故障场景下表现出较强的鲁棒性(负荷满足率达 98.7%),但在极端天气(如连续 72 h 无风无光)下或多设备并发故障时,储能与电网交互的备用容量仍显不足。

(2)通过“星型-mesh”通信拓扑实现毫秒级决策,缓解了“维度灾难”,使多主体协同效率提升了 15.8%,增强了对电价波动、负荷突变场景的适应性。

(3)通过优化储能与电网交互的备用容量配置、纳入氢能储能等新兴柔性资源并建模协同优化、强化极端场景鲁棒性设计、多能互补冗余和风光预测动态调整,解决了极端天气下储能支撑不足问题,显著增强了电力系统的抗风险能力与整体韧性,为复杂环境下的能源安全稳定运行提供了可靠保障。

### 参考文献:

[1]肖玲娟,李润祥.以风电为主的新能源电站源-储协调优化调度方法[J].自动化与仪器仪表,2025(7):97-102.  
XIAO Lingjuan, LI Runxiang. Optimization and scheduling method for source storage coordination of new energy power stations mainly based on wind power [J]. Automation & Instrumentation, 2025(7): 97-102.

[2]郭少臣.风电储能系统中的能量管理效率技术优化路径[J].电气技术与经济,2024(12):69-72.  
GUO Shaochen. Technical optimization path of energy management efficiency in wind energy storage system [J]. Electrical Equipment and Economy, 2024(12): 69-72.

[3]赵书强,吴博,李志伟,等.风电-储能参与调频的高比例风电电力系统运行经济性分析[J].南方电网技术,2023,17(4):69-76,89.  
ZHAO Shuqiang, WU Bo, LI Zhiwei, et al. Operational economic analysis of high-proportion wind power system with wind power and energy storage participating in frequency regulation [J]. Southern Power System Technology, 2023, 17(4): 69-76, 89.

[4]徐楠,陈斌,黄伟,等.基于滚动优化的电-热-气-冷系统多时间尺度低碳运行[J].电网与清洁能源,2024,40(7):95-106,113.  
XU Nan, CHEN Bin, HUANG Wei, et al. Multi-timescale low-carbon operation of the electric-heat-gas-cooling combined supply system based on rolling optimization [J]. Advances of Power System & Hydroelectric Engineering, 2024, 40(7): 95-106, 113.

[5]周益民,杨博,胡袁炜骥,等.考虑绿证-碳交易的多能互补综合能源系统电-热-气协同低碳优化调度[J].电网技术,2025,49(6):2428-2435.  
ZHOU Yimin, YANG Bo, HU Yuanweiji, et al. Electricity-heat-gas collaborative low-carbon optimal dispatch of multi-energy complementary integrated energy system considering green certificate and carbon emission trading [J]. Power System Technology, 2025, 49(6): 2428-2435.

[6]刘毓伶,赵兴勇,刘立.含信息间隙决策理论的电热气联合系统优化调度[J].电力系统及其自动化学报,2024,36(12):45-53.  
LIU Yuling, ZHAO Xingyong, LIU Li. Optimization and scheduling of power, heat and gas combination system based on information gap decision theory [J]. Proceedings of the CSU-EPSA, 2024, 36(12): 45-53.

[7]徐新宇,秦伟,李楠,等.基于强化学习的马尔科夫决策过程电网侧储能系统的智能调度优化[J].电工技术,2025(13):94-96.  
XU Xinyu, QIN Wei, LI Nan, et al. Intelligent scheduling optimization of power grid side energy storage system based

- on reinforcement learning Markov decision process [J]. *Electric Engineering*, 2025(13): 94-96.
- [8]张硕,袁春辉,李英姿,等.新型电力系统演化趋优下多元负荷绿色协同市场行为DQN仿真模型[J].*电力建设*, 2025, 46(7): 191-204.  
ZHANG Shuo, YUAN Chunhui, LI Yingzi, et al. DQN simulation model of green cooperative market behavior of multi-load users for optimal evolution of new-type power system [J]. *Electric Power Construction*, 2025, 46 (7) : 191-204.
- [9]周雪松,韩静,马幼捷,等.基于DQN算法的直流微电网负载接口变换器自抗扰控制策略[J].*电力系统保护与控制*, 2025, 53(1): 95-103.  
ZHOU Xuesong, HAN Jing, MA Youjie, et al. Active disturbance rejection control strategy of a DC microgrid load interface converter based on a DQN algorithm [J]. *Power System Protection and Control*, 2025, 53(1): 95-103.
- [10]王仁浚,高红均,罗龙波,等.基于深度强化学习的新配电网系统优化运行研究综述[J].*电力自动化设备*, 2025, 45(9): 152-164.  
WANG Renjun, GAO Hongjun, LUO Longbo, et al. Review of research on new distribution system optimization operation based on deep reinforcement learning [J]. *Electric Power Automation Equipment*, 2025, 45 (9) : 152-164.
- [11]青辰,魏震波,刘洋,等.基于双时间尺度模型预测控制的灵活性资源动态调度[J].*高压电器*, 2025, 61(5): 31-40, 52.  
QING Chen, WEI Zhenbo, LIU Yang, et al. Dynamic scheduling of flexible resources based on dual-time scale model predictive control [J]. *High Voltage Apparatus*, 2025, 61(5): 31-40, 52.
- [12]江美慧,许镇江,张其朴,等.面向综合能源系统的综合需求响应策略及其建模技术[J].*电力建设*, 2024, 45(12): 65-82.  
JIANG Meihui, XU Zhenjiang, ZHANG Qipu, et al. Integrated demand response strategy and modeling technology for integrated energy systems [J]. *Electric Power Construction*, 2024, 45(12): 65-82.
- [13]蒋明轩,卞艺衡,李更丰,等.面向能源互联网的电-碳-氢耦合交易市场研究综述[J].*电力建设*, 2025, 46(8): 150-165.  
JIANG Mingxuan, BIAN Yiheng, LI Gengfeng, et al. Review of the research on the electricity-carbon-hydrogen coupling trading market under the energy Internet [J]. *Electric Power Construction*, 2025, 46(8): 150-165.
- [14]李民,周博宇,蒋雷雷,等.基于多智能体强化学习的电池一致性充放电控制策略[J].*武汉大学学报(理学版)*, 2025, 71(3): 395-403.  
LI Min, ZHOU Boyu, JIANG Leilei, et al. Battery consistency charge-discharge control strategy based on multi-agent reinforcement learning [J]. *Journal of Wuhan University (Natural Science Edition)*, 2025, 71 (3) : 395-403.
- [15]徐慧慧,田云飞,赵宇洋,等.基于条件生成对抗网络与多智能体强化学习的配电网可靠性评估方法[J].*中国电力*, 2025, 58(4): 230-236.  
XU Huihui, TIAN Yunfei, ZHAO Yuyang, et al. A reliability assessment method for distribution networks based on conditional generative adversarial network and multi-agent reinforcement learning [J]. *Electric Power*, 2025, 58(4): 230-236.
- [16]李峥嵘,刘雨欣,朱晗,等.面向“双碳”目标的光储一体化建筑集群电力协同调度方法研究综述[J].*建筑节能*, 2023, 51(8): 1-10.  
LI Zhengrong, LIU Yuxin, ZHU Han, et al. Review on power cooperative scheduling method for building cluster integrated PV and storage dual to the carbon peaking and carbon neutrality goals [J]. *Building Energy Efficiency*, 2023, 51(8): 1-10.
- [17]于小唐,覃智君,梁志坚.需求响应的园区综合能源系统优化运行方式研究[J].*广西大学学报(自然科学版)*, 2024, 49(4): 788-798.  
YU Xiaotang, QIN Zhijun, LIANG Zhijian. Research on optimization operation mode of park integrated energy system for demand response [J]. *Journal of Guangxi University (Natural Science Edition)*, 2024, 49 (4) : 788-798.
- [18]葛文超,任洪波,朱跃钊,等.镇级可再生能源多能互补系统优化配置与调度探索[J].*太阳能学报*, 2025, 46(3): 168-176.  
GE Wenchao, REN Hongbo, ZHU Yuezhao, et al. Optimized allocation and scheduling of renewable energy in multi-energy complementary system at township level [J]. *Acta Energetica Sinica*, 2025, 46 (3) : 168-176.
- [19]洪春雪,肖海平,谭甲群,等.考虑容量电价的风光火储能源基地多目标优化调度[J].*综合智慧能源*, 2025, 47(7): 1-11.  
HONG Chunxue, XIAO Haiping, TAN Jiaqun, et al. Multi-objective optimal schedule of a wind-photovoltaic-thermal-storage energy base considering capacity tariffs [J]. *Integrated Intelligent Energy*, 2025, 47(7): 1-11.
- [20]钟永洁,王紫东,左建勋,等.计及多时段尺度与地域分层的多能互补系统经济调度[J].*综合智慧能源*, 2024, 46(4): 52-59.  
ZHONG Yongjie, WANG Zidong, ZUO Jianxun, et al. Economic dispatch of multi-energy complementary systems considering multi-period scales and regional stratification [J]. *Integrated Intelligent Energy*, 2024, 46(4): 52-59.

- [21]王心玉, 李金航, 陈衡, 等. 同时参与绿证交易-碳交易的区域多能互补电力系统优化调度[J]. 动力工程学报, 2025, 45(2): 315-324.  
WANG Xinyu, LI Jinhang, CHEN Heng, et al. Optimization scheduling of regional multi-energy complementary power system participating simultaneously in green certificate trading-carbon trading [J]. Journal of Chinese Society of Power Engineering, 2025, 45 (2) : 315-324.
- [22]曹猛, 解超, 尹纯亚, 等. 计及源荷不确定性的多综合能源微网协同优化运行策略[J]. 电网与清洁能源, 2025, 41(3): 112-124.  
CAO Meng, XIE Chao, YIN Chunya, et al. Co-optimized operation strategy for multiple integrated energy microgrids considering source-load uncertainty [J]. Power System and Clean Energy, 2025, 41(3): 112-124.
- [23]何良策, 王宇, 卢志刚, 等. 面向多区域综合能源系统低碳运行的共享电-氢储能优化配置[J]. 电力系统保护与控制, 2025, 53(18): 52-63.  
HE Liangce, WANG Yu, LU Zhigang, et al. Optimal allocation of shared electricity-hydrogen storage for low-carbon operation of multiple regional integrated energy systems [J]. Power System Protection and Control, 2025, 53(18): 52-63.
- [24]蒋灵慧, 冯霞, 崔凯平, 等. 面向联邦学习的油电混合电动汽车站点推荐算法研究[J]. 重庆理工大学学报(自然科学版), 2025, 39(3): 113-119.  
JIANG Linghui, FENG Xia, CUI Kaiping, et al. Research on hybrid electric vehicle station recommendation algorithm based on federated learning [J]. Journal of Chongqing Institute of Technology, 2025, 39 (3) : 113-119.
- [25]吴润泽, 霍金鑫, 郭昊博. 基于DQN的电力协同计算与缓存的任务卸载策略[J]. 电力建设, 2024, 45(8): 149-158.  
WU Runze, HUO Jinxin, GUO Haobo. DQN-based task offloading strategy for power co-computing and caching [J]. Electric Power Construction, 2024, 45(8): 149-158.
- [26]乔艺林, 王楚通, 熊厚博, 等. 基于发电侧共建共享的季节性-短期混合储能系统优化配置[J]. 可再生能源, 2025, 43(5): 654-662.  
QIAO Yilin, WANG Chutong, XIONG Houbo, et al. Optimization of seasonal and short-term hybrid energy storage system based on co-construction and sharing of power generation side [J]. Renewable Energy Resources, 2025, 43(5): 654-662.
- [27]马爽, 张弘毅, 马炜翔, 等. 基于沙戈荒地区源荷随机特性的多能源优化运行研究综述[J]. 高电压技术, 2025, 51(2): 507-519.  
MA Shuang, ZHANG Hongyi, MA Weixiang, et al. Research review on the optimal operation of multi-energy system in gobi and desert areas based on stochastic characteristics of source-load [J]. High Voltage Engineering, 2025, 51(2): 507-519.
- [28]黄晟, 潘丽君, 屈尹鹏, 等. 基于改进深度Q网络的无预测风电场日前拓扑优化[J]. 电力系统自动化, 2025, 49(2): 122-132.  
HUANG Sheng, PAN Lijun, QU Yinpeng, et al. Day-ahead topology optimization for wind farm without forecasting based on improved deep Q network [J]. Automation of Electric Power Systems, 2025, 49 (2) : 122-132.
- [29]孙浩, 邢作霞, 吴维宁, 等. 计及电力市场交易机制的风-光-储-氢混合电厂配置策略研究[J]. 储能科学与技术, 2025, 14(7): 2801-2812.  
SUN Hao, XING Zuoxia, WU Weining, et al. Research on configuration strategies for wind-solar-battery-hydrogen hybrid power plants considering electricity market trading mechanisms [J]. Energy Storage Science and Technology, 2025, 14(7): 2801-2812.
- [30]吴晨曦, 倪索引, 郑静, 等. 考虑能碳溯源的风储协同调度研究[J]. 电网与清洁能源, 2025, 41(2): 84-92.  
WU Chenxi, NI Suoyin, ZHENG Jing, et al. Research on wind storage-energy co-scheduling considering energy-carbon traceability [J]. Power System and Clean Energy, 2025, 41(2): 84-92.

(本文责编:刘芳)

收稿日期:2025-07-18;修回日期:2025-12-08  
上网日期:2026-03-12;附录网址:www.iieneg.cn

## 作者简介:

陈锋(1987),男,副教授,高级工程师,硕士,从事高比例新能源新型电力系统优化运行与智能控制等方面的研究, chenfeng@smxpt.edu.cn;

路小敏(1989),女,工程师,硕士,从事电力大数据技术等各方面的研究,703399594@qq.com;

李梦杨(1988),女,副教授,博士,从事非线性系统鲁棒控制、多智能体、工业智能、无人驾驶等方面的研究, limengyang8801@163.com;

张涛(1995),男,工程师,硕士,从事电力电子与电力传动等方面的研究,474872683@qq.com;

杨帆(1989),女,讲师,硕士,从事跨区域电力联网背景下人力资源统筹配置等方面的研究,18239861350@163.com。