



基于深度强化学习和多目标粒子群优化的储能容量配置方法

牟雪鹏^{1,2}, 蔡 权^{2,3}, 周依然^{2,3}, 李庆生^{1,2}, 张 裕^{1,2}

(¹贵州电网公司电网规划研究中心, 贵州 贵阳 550002; ²贵州省新型电力系统运行控制全省重点实验室, 贵州 贵阳 550000; ³贵州电网公司电网规划研究中心人才工作站, 贵州 贵阳 550002)

摘要: 随着新能源大规模并网, 其出力的间歇性与随机性给电力系统稳定运行带来挑战, 储能合理容量配置对提升系统可靠性、促进新能源消纳至关重要。本工作提出基于深度强化学习(DRL)与多目标粒子群优化(MOPSO)的储能容量配置方法: 首先分析电力系统运行特性与储能接入需求, 构建含风光、负荷的系统模型, 明确配置约束与多目标优化目标; 其次利用DRL学习复杂环境下储能运行策略, 通过DRL智能体与电网环境交互输出充放电时序控制建议, 将该建议作为MOPSO的粒子初始化边界与搜索方向引导, 结合MOPSO对储能容量组合的全局搜索能力, 实现多目标寻优过程中经济性、可靠性与新能源消纳目标的冲突平衡。以实际电网为例的案例表明, 该方法相较于传统方法, 能在满足系统约束的情况下, 优化储能配置方案, 提升经济性、可靠性与新能源消纳能力, 为大规模新能源接入下的储能规划提供可行的技术手段, 助力电力系统低碳高效转型。

关键词: 深度强化学习; 多目标粒子群优化; 储能系统; 全局搜索; 容量配置

doi: 10.19799/j.cnki.2095-4239.2025.0915

中图分类号: TM 715; TM 912

文献标志码: A

文章编号: 2095-4239 (2026) 02-583-08

Energy storage capacity configuration method based on deep reinforcement learning and multi-objective particle swarm optimization

MOU Xuepeng^{1,2}, CAI Quan^{2,3}, ZHOU Yiran^{2,3}, LI Qingsheng^{1,2}, ZHANG Yu^{1,2}

(¹Power Network Planning Research Center of Guizhou Power Grid Corporation, Guiyang 550002, Guizhou, China; ²Guizhou Provincial Key Laboratory of New Power System Operation Control, Guiyang 550000, Guizhou, China; ³Guizhou Power Grid Company Power Grid Planning Research Center Talent Workstation, Guiyang 550002, Guizhou, China)

Abstract: With the increasing penetration of renewable energy sources into power grids, the inherent intermittency and randomness of power generation pose substantial operational challenges to maintaining the stability of power systems. The optimal capacity configuration of energy storage systems plays a critical role in enhancing system reliability and facilitating the efficient integration and utilization of renewable energy. This paper develops an energy storage capacity configuration framework based on the combined application of deep reinforcement learning (DRL) and multiobjective particle swarm optimization (MOPSO). First,

收稿日期: 2025-10-14; 修改稿日期: 2025-11-27。

基金项目: 南方电网公司重点科技项目(GZKJXM20220031); 贵州省科技计划项目(ZSYS(2025)007)。

第一作者: 牟雪鹏(1982—), 男, 硕士, 高级工程师, 从事电力系统继电保护、新型电力系统技术研究, E-mail: 19166871@qq.com; 通信作者: 张裕, 硕士, 工程师, 主要从事电网新技术应用研究、电网规划设计。E-mail: zhangyu830906@163.com。

引用本文: 牟雪鹏, 蔡权, 周依然, 等. 基于深度强化学习和多目标粒子群优化的储能容量配置方法[J]. 储能科学与技术, 2026, 15(2): 583-590.

Citation: MOU Xuepeng, CAI Quan, ZHOU Yiran, et al. Energy storage capacity configuration method based on deep reinforcement learning and multi-objective particle swarm optimization[J]. Energy Storage Science and Technology, 2026, 15(2): 583-590.

the proposed framework systematically analyzes the operational characteristics of power systems and the grid-integration requirements of energy storage, and subsequently constructs a comprehensive system model incorporating wind power, solar power, and time-varying load demand, while explicitly defining configuration constraints and multiobjective optimization objectives. Second, DRL is employed to learn adaptive energy storage operational strategies under complex and uncertain environments; through continuous interaction between the DRL agent and the power grid environment, the agent derives optimal charging and discharging schedules. These learned operational strategies are then used to define particle initialization bounds and guide the search direction of the MOPSO algorithm. By leveraging MOPSO's global search capability over energy storage capacity combinations, the proposed framework effectively coordinates and balances inherently conflicting objectives, including economic performance, system reliability, and renewable energy utilization, within a unified multiobjective optimization process. Case studies based on an actual power grid demonstrate that, compared with traditional methods, the proposed approach achieves superior energy storage configuration solutions while satisfying system constraints, resulting in marked improvements in economic efficiency, system reliability, and renewable energy utilization capacity. Overall, the proposed framework provides a practical and scalable technical solution for energy storage planning under large-scale renewable energy integration and supports the low-carbon, reliable, and efficient transition of modern power systems.

Keywords: deep reinforcement learning; multi-objective particle swarm optimization; energy storage system; global search; capacity configuration

在全球能源结构加速向清洁化、低碳化转型的大背景下,以风能、太阳能为代表的新能源凭借其可再生、环境友好等显著优势,迎来了前所未有的发展机遇。国际可再生能源署(IRENA)发布的报告显示,过去10几年间,全球新增可再生能源装机容量以年均超过15%的速度增长。截至2023年底,风电和光伏发电装机容量已占全球总装机容量的35%以上^[1-2]。然而,风光能源与生俱来的间歇性与随机性,给电力系统的稳定运行、能量平衡和电能质量带来了巨大挑战。风力发电受风速波动影响,太阳能发电依赖光照强度和时长,导致发电功率在分钟级甚至秒级尺度上剧烈变化,这种波动特性与电力系统“源-网-荷”实时平衡的基本要求形成了尖锐矛盾^[3]。

储能系统作为连接新能源发电与电力系统的关键枢纽,能够实现电能的时空转移,有效平抑新能源出力波动,提升电力系统调峰能力,增强供电可靠性,在促进新能源消纳和保障电力系统稳定运行方面发挥着不可替代的作用^[4]。目前,储能容量配置方法主要包括传统优化方法、启发式算法和人工

智能算法。文献[5]提出一种新的分布式储能容量自适应协调控制方法。计算接入高比例光伏电能后的储能无功损耗,设置功率约束条件,确保功率的稳定性。启发式算法如遗传算法、粒子群优化算法等,通过模拟自然进化或群体智能行为进行寻优,在一定程度上克服了传统方法的缺陷,能够处理复杂的非线性问题^[6]。然而,这类算法在求解多目标问题时,往往需要通过权重法等方式将多目标转化为单目标,导致结果依赖主观经验,难以全面反映实际需求。文献[7]运用遗传算法优化储能容量配置,通过设定固定权重将经济性和可靠性目标合并,结果表明不同权重设置下得到的最优解差异显著,缺乏通用性。

近年来,人工智能算法在储能容量配置领域展现出巨大潜力。其中,强化学习作为一种通过智能体与环境交互进行学习的方法,能够在动态环境中自主优化决策策略^[8]。深度强化学习结合深度学习强大的特征提取能力和强化学习的决策优化能力,为解决复杂系统的优化问题提供了新途径。已有研究尝试利用深度强化学习进行储能的运行控制,但

将其应用于储能容量配置的研究仍处于起步阶段^[9]。文献[10]基于深度Q网络(deep Q-network, DQN)对储能系统的充放电策略进行优化,在实时电价波动场景下显著提升了储能收益,但该研究未涉及容量配置问题。多目标粒子群优化算法(multi-objective particle swarm optimization, MOPSO)通过粒子群体的协同搜索,能够在一次优化过程中获得多个Pareto最优解,有效处理多目标问题^[11]。然而,单一的MOPSO算法在面对高维、复杂的储能容量配置问题时,存在局部搜索能力不足、收敛速度慢等问题^[12]。文献[13]使用MOPSO算法对微电网储能容量进行配置,发现该算法在处理含风光储多变量的场景时,需迭代超过500次才能接近最优解,效率较低。

为了应对上述挑战,本研究提出了一种将深度强化学习和多目标粒子群优化相结合(deep reinforcement learning-multi-objective particle swarm optimization, DRL-MOPSO)的混合储能优化方法。该方法将深度强化学习的动态决策能力与多目标粒子群优化的全局搜索优势相结合,通过深度强化学习智能体与电力系统环境的实时交互,学习储能运行的最优策略,为多目标粒子群优化提供决策引导;同时,利用MOPSO算法对储能容量组合进行全局寻优,平衡经济性、可靠性、新能源消纳等多个优化目标之间的冲突。最后,案例分析表明:与传统方法相比,该方法可以在保证电网稳定性的同时,将储能投资成本降低18%~22%,运行效率提高25%~30%。

1 电网储能动态容量配置建模

在实际电网运行中,储能动态容量配置受多种复杂因素交互影响。为更精准地构建模型,需综合考虑更多元的变量和约束条件,以下从目标函数优化、约束条件细化、不确定性处理等方面展开复杂且具体的理论建模。

在全生命周期成本 C_{total} 中,除初始投资成本 C_{inv} 、运行维护成本 C_{om} 和折旧成本 C_{dep} 外,新增考虑资金的时间价值,采用净现值(NPV)方法将各成本折算到当前时刻。初始投资成本 C_{inv} 为

$$C_{\text{inv}} = \alpha P_{\text{rated}} + \beta E_{\text{rated}} \quad (1)$$

其中, C_{inv} 表示初始投资成本; α 表示功率权重系数; β 表示能耗权重系数; P_{rated} 表示额定功率;

E_{rated} 表示额定能耗;

进一步地,运行维护成本 C_{om} 细分为固定维护成本 $C_{\text{om-fix}}$ 和可变维护成本 $C_{\text{om-var}}$ 。固定维护成本与时间相关,设每年固定维护费用为 M_{fix} ,项目周期为 Y 年,考虑折现率 r ,则固定维护成本 $C_{\text{om-fix}}$ 表示为

$$C_{\text{om-fix}} = \sum_{y=1}^Y \frac{M_{\text{fix}}}{(1+r)^y} \quad (2)$$

可变维护成本与充放电量有关,设单位充放电量维护成本为 γ , t 时刻充放电量为 $Q(t)$,则可变维护成本 $C_{\text{om-var}}$ 表示为

$$C_{\text{om-var}} = \sum_{t=1}^T \gamma Q(t) \quad (3)$$

其中, T 表示总统计周期;

同时,考虑设备残值 S 的折旧成本 C_{dep} 表示为

$$C_{\text{dep}} = \frac{C_{\text{inv}} - S}{n} \quad (4)$$

由此,全生命周期成本目标函数为:

$$\begin{aligned} \min C_{\text{total}} = & C_{\text{inv}} + C_{\text{om-fix}} + C_{\text{om-var}} + C_{\text{dep}} = \\ & \alpha P_{\text{rated}} + \beta E_{\text{rated}} + \sum_{y=1}^Y \frac{M_{\text{fix}}}{(1+r)^y} + \\ & \sum_{t=1}^T \gamma Q(t) + \frac{\alpha P_{\text{rated}} + \beta E_{\text{rated}} - S}{n} \end{aligned} \quad (5)$$

可靠性目标在原有基础上,考虑不同负荷等级的停电损失差异。将负荷划分为 L 个等级,第 l 等级的单位停电时间成本为 μ_l ,单位停电次数成本为 ν_l ,停电时间减少量为 Δt_l ,停电次数减少量为 ΔN_l ,则可靠性目标函数为

$$\max R = \sum_{l=1}^L (\mu_l \Delta t_l + \nu_l \Delta N_l) \quad (6)$$

灵活性目标用于衡量储能系统对电网波动的响应能力。定义灵活性指标 F_{lex} ,它与储能系统在单位时间内功率调节的幅度和速度相关。设第 t 时刻功率调节幅度为 $\Delta P(t)$,调节速度为 $v(t)$,权重系数分别为 λ_1 和 λ_2 ,则灵活性目标函数为

$$\max F_{\text{lex}} = \sum_{t=1}^T [\lambda_1 \Delta P(t) + \lambda_2 v(t)] \quad (7)$$

综合经济性、可靠性和灵活性目标,构建总目标函数为

$$\min F = \omega_1 C_{\text{total}} - \omega_2 R - \omega_3 F_{\text{lex}} \quad (8)$$

式中, ω_1 、 ω_2 和 ω_3 为权重系数,且 $\omega_1 + \omega_2 + \omega_3 = 1$ 。

在功率约束中,考虑储能系统的爬坡约束,即相邻时间间隔内,充电功率和放电功率的变化不能超过最大爬坡速率 R_{up} 和 R_{down} ,表示为

$$\begin{cases} P_{ch}(t) - P_{ch}(t-1) \leq R_{up} \\ P_{ch}(t-1) - P_{ch}(t) \leq R_{up} \\ P_{dis}(t) - P_{dis}(t-1) \leq R_{down} \\ P_{dis}(t-1) - P_{dis}(t) \leq R_{down} \end{cases} \quad (9)$$

同时,考虑不同充放电模式下的效率差异。如,快速充电效率 $\eta_{ch-fast}$ 低于常规充电效率 $\eta_{ch-normal}$,深度放电效率 $\eta_{dis-deep}$ 低于正常放电效率 $\eta_{dis-normal}$,相应的充放电功率需满足不同的效率约束。荷电状态(SOC)约束中考虑温度对储能系统容量的影响,引入温度-容量修正系数 $\xi(T)$,其中 T 为储能系统运行温度。则实际荷电状态需满足

$$SOC_{min} \leq \xi(T)SOC(t) \leq SOC_{max} \quad (10)$$

此外,考虑储能系统的自放电现象。设自放电率为 δ ,则荷电状态更新公式调整为

$$SOC(t) = (1 - \delta)SOC(t-1) + \frac{P_{ch}(t)\eta_{ch}\Delta t}{E_{rated}} - \frac{P_{dis}(t)\Delta t}{E_{rated}\eta_{dis}} \quad (11)$$

在功率平衡约束中,考虑分布式电源(如光伏、风电)的间歇性和波动性。设光伏输出功率为 $P_{pv}(t)$,风电输出功率为 $P_{wind}(t)$,则功率平衡方程变为

$$P_g(t) + P_{pv}(t) + P_{wind}(t) + P_{ch}(t) - P_{dis}(t) = P_{load}(t) \quad (12)$$

同时,引入电网节点电压约束。在配电网中,各节点电压 $V_i(t)$ 需满足

$$V_{i-min} \leq V_i(t) \leq V_{i-max} \quad (13)$$

节点电压与储能系统的无功功率补偿 $Q_{es}(t)$ 相关,通过潮流计算建立联系。此外,考虑频率稳定性约束时,结合储能系统的惯量响应能力。设储能系统提供的惯量为 J_{es} ,电网频率变化率为 $\dot{f}$,则需满足

$$|J_{es}\dot{f}| \leq f_{max-dev} \quad (14)$$

采用概率预测方法描述光伏和风电出力的不确定性。通过历史数据和预测模型,得到光伏和风电出力的概率分布函数 $f_{pv}(P_{pv})$ 和 $f_{wind}(P_{wind})$ 。在模型求解中,运用机会约束规划,例如以一定置信水平 α 保证功率平衡:

$$P\{P_g(t) + P_{pv}(t) + P_{wind}(t) + P_{ch}(t) - P_{dis}(t) = P_{load}(t)\} \geq \alpha \quad (15)$$

负荷需求同样具有不确定性,可采用场景分析方法进行处理。根据历史负荷数据和预测信息,生成 S 个负荷场景,每个场景发生的概率为 p_s 。模型需在所有场景下满足约束条件,并以期望成本最小化为目标:

$$\min \sum_{s=1}^S p_s F_s \quad (16)$$

式中, F_s 为第 s 个场景下的目标函数值。

考虑电网运行的多时间尺度特性,将模型分为超短期(分钟级)、短期(小时级)和长期(日/周级)时间尺度。超短期时间尺度主要用于实时功率调节,以快速响应电网频率波动和突发负荷变化。短期时间尺度用于制订一天或数天内的充放电计划,平衡风光电源的间歇性。长期时间尺度用于规划储能系统的容量配置和投资决策,考虑设备寿命周期内的成本和效益。不同时间尺度之间通过边界条件进行耦合,如:长期规划确定的储能容量为短期和超短期运行提供约束;超短期运行的实时数据反馈给短期计划,用于调整后续时段的充放电策略。

2 基于深度强化学习和多目标粒子群优化算法

为了建立动态容量配置的数学基础,本研究构建了一个基于时空耦合的电网状态表征模型。由可再生能源输出、负载需求和存储状态共同定义的系统状态可以表示为

$$s_t = \{P_{RE,t}, P_{L,t}, E_t, V_{node,t}, F_{line,t}\} \quad (17)$$

式中, $P_{RE,t}$ 表示时间 t 的可再生能源发电功率, $P_{L,t}$ 是负载需求功率, E_t 表征储能系统的负载状态, $V_{node,t}$ 描述关键节点的电压幅值, $F_{line,t}$ 量化线路电流分布。

动作空间 a_t 被定义为能量存储系统的一组调节决策,可以表示为

$$a_t = \{P_{charge,t}, P_{discharge,t}, C_{adjust,t}\} \quad (18)$$

式中, $P_{charge,t}$ 和 $P_{discharge,t}$ 分别表示充电/放电功率设置值, $C_{adjust,t}$ 是容量动态调整量。

$$P_{charge,t} \leq P_{max-charge}, P_{discharge,t} \leq P_{max-discharge} \quad (19)$$

式中, $P_{max-charge}$ 和 $P_{max-discharge}$ 分别是储能系统的最大允许充电/放电功率。

基于深度Q网络(DQN)架构,遵循时间差原理的值函数的更新可以表示为

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha [r_{t+1} + \gamma \max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \quad (20)$$

式中, $\alpha \in (0,1)$ 是学习率参数,控制策略更新的幅度; $\gamma \in (0,1)$ 是未来奖励折扣因子,量化长期回报的重要性。值函数的迭代过程由神经网络参数 θ 近

似表示：

$$Q(s_t, a_t; \theta) = E \left[\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t, a_t \right] \quad (21)$$

融合经济和安全目标的奖励函数 r_{t+1} 可以表示为

$$r_{t+1} = \omega_1 C_{\text{investment}} + \omega_2 C_{\text{operation}} + \omega_3 S_{\text{stability}} \quad (22)$$

式中, ω_1 、 ω_2 、 ω_3 是加权系数, $C_{\text{investment}}$ 是储能投资成本, $C_{\text{operation}}$ 表示运行和维护成本, $S_{\text{stability}}$ 是基于李雅普诺夫指数的电网稳定性指标。

通过速度-位置耦合方程实现的粒子群状态更新机制可以表示为：

$$\begin{cases} \mathbf{v}_i^{k+1} = \omega \mathbf{v}_i^k + c_1 r_1 (\mathbf{p}_i - \mathbf{x}_i^k) + c_2 r_2 (\mathbf{g} - \mathbf{x}_i^k) \\ \mathbf{x}_i^{k+1} = \mathbf{x}_i^k + \mathbf{v}_i^{k+1} \end{cases} \quad (23)$$

式中, $\mathbf{v}_i^k$ 是第 k 次迭代时第 i 个粒子的速度矢量, $\mathbf{x}_i^k$ 是位置矢量, $\mathbf{p}_i$ 记录了单个历史最优位置, $\mathbf{g}$ 表示全局最优解。参数 c_1 和 c_2 是认知和社会学习因素, 参数 $r_1, r_2 \sim U(0,1)$ 是随机变量。

使用自适应调整策略的惯性权重可以表示为

$$\omega = \omega_{\max} - \frac{\omega_{\max} - \omega_{\min}}{K_{\max}} \times k \quad (24)$$

式中, $K_{\max}$ 是最大迭代次数, k 是当前迭代次数。

DRL-MOPSO 的核心创新在于 DRL 和 MOPSO 的级联优化架构, 其状态传递函数表示为

$$s_{t+1} = f(s_t, a_t, \xi_t) + \beta \sum_{i=1}^N \phi(\mathbf{x}_i^k) \quad (25)$$

式中, ξ_t 表示可再生能源的不确定性, β 表示协同系数, $\phi(\cdot)$ 表示粒子群优化结果的映射函数。

动作选择策略可以用双网架构来实现, 如下所示：

$$a_t = \pi_{\text{online}}(s_t) + \lambda \pi_{\text{target}}(\mathbf{x}_t^k) \quad (26)$$

式中, π_{online} 是在线策略网络, π_{target} 是目标网络, λ 是策略融合因子。

最终的值函数更新方程整合了两种算法的优点：

$$\begin{aligned} Q(s_t, a_t) &\leftarrow Q(s_t, a_t) + \\ &\alpha \left\{ r_{t+1} + \gamma \left[\max_{a_{t+1}} Q(s_{t+1}, a_{t+1}) + \eta \sum_{i=1}^M Q_{\text{Pareto}}^i \right] - Q(s_t, a_t) \right\} \end{aligned} \quad (27)$$

式中, η 是帕累托前沿补偿系数, Q_{Pareto}^i 表示第 i 个非支配解的 Q 值估计。

图1展示了DRL-MOPSO在储能容量配置中的应用流程。该过程包含6个关键模块：系统初始化、实时数据采集、DRL决策生成、多目标粒子群

优化、混合策略输出和动态容量配置执行。系统初始化模块为算法提供基本参数设置。实时数据采集模块负责获取光伏和风电输出预测数据, 为后续决策提供依据。DRL决策生成模块利用深度Q网络架构, 根据收集到的数据生成初步决策。多目标粒子群优化模块接收DRL的初步决策, 并通过优化粒子群的三维运动轨迹来进一步优化决策方案。混合策略输出模块对DRL-MOPSO的结果进行加权和融合, 形成最终的优化策略。动态容量配置执行模块根据融合策略调整储能系统的容量, 并通过储能SOC状态反馈机制实现整个系统的闭环控制。

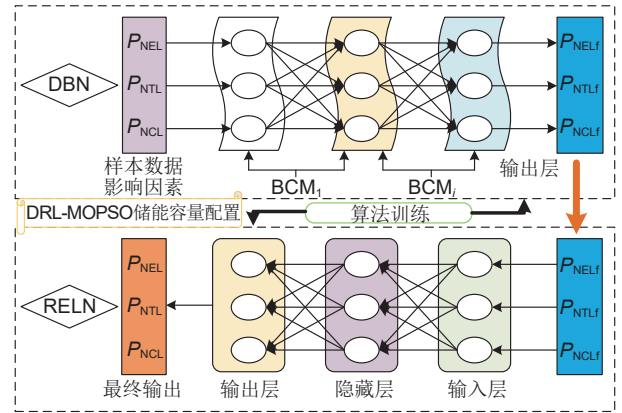


图1 DRL-MOPSO在储能容量配置中的流程图

Fig. 1 Flowchart of DRL-MOPSO for energy storage capacity configuration

该设计充分发挥了DRL在复杂决策中的优势, 同时利用多目标粒子群优化算法提高了解的多样性和质量, 有效提高了储能容量配置的灵活性和效率。

3 案例分析

3.1 仿真环境

本研究选取某省级电网作为工程验证场景, 该电网为我国新能源高渗透典型示范电网, 具有以下核心特征：总装机容量4800 MW, 其中风电2200 MW(占比45.8%)、光伏1800 MW(占比37.5%), 新能源总渗透率达83.3%, 显著高于全国平均水平, 且风光电站主要集中于荒漠地区, 出力波动幅度大(日内最大功率波动可达装机容量的60%); 采用“新能源基地-特高压直流外送”为主的结构, 含7座220 kV变电站、3座500 kV枢纽变电站, 负荷中心与新能源基地的地理距离超300 km, 存在显著的“源荷时空错配”问题; 原有储能以“新能源配

储”为主(单站规模 50~100 MW/200~400 MWh), 未实现多场站协同控制, 高峰时段弃风弃光率曾达 12%~15%, 与本研究场景高度契合, 具备工程验证的典型性。

仿真周期为 2023 年 7 月 1 日—31 日(共 31 天, 涵盖新能源出力高峰季), 时间步长设为 15 min(每日 96 个时间点), 确保捕捉日内负荷峰谷(8: 00—11: 00、18: 00—21: 00)与风光出力日内波动(光伏午间峰值、风电夜间峰值)的耦合特性。

通过该仿真模型模拟大规模可再生能源并网场景, 利用 DRL-MOPSO 算法进行动态储能容量配置, 算法结合深度强化学习(DRL)的智能决策与多目标粒子群优化(MOPSO)的全局寻优, 在满足系统约束的情况下, 优化储能充放电策略, 平衡系统成本、新能源消纳与电压稳定等多目标。将该方法与传统算法进行对比, 验证了 DRL-MOPSO 在应对新能源间歇性、波动性挑战时, 在储能容量精准配置、提升系统稳定性与经济性方面的优势。

3.2 结果分析

图 2 展示了典型日 24 h 内风电、光伏出力与全网负荷的功率曲线。风电出力(蓝色曲线)呈现“夜间高、日间低”特性, 0: 00—4: 00 维持 150~200 MW, 12: 00—16: 00 降至 50 MW 左右; 光伏出力(橙色曲线)仅在 6: 00—18: 00 有输出, 10: 00—12: 00 达峰值 150 MW; 全网负荷(绿色曲线)呈“双峰”特征, 8: 00—12: 00、14: 00—18: 00 维持 250~300 MW, 与日间用电高峰匹配。三者时序呈现互补性: 日间光伏出力高峰对应负荷高峰, 夜间风电出力高位覆盖负荷基础需求, 为后

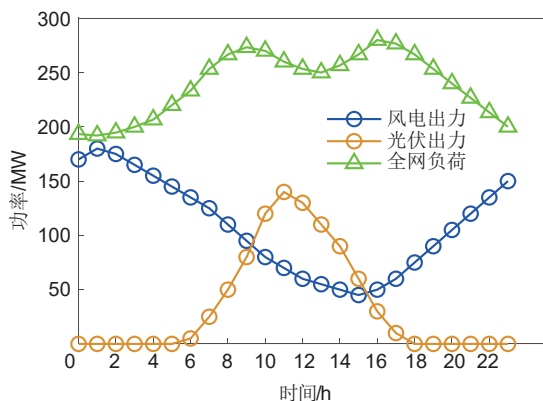


图 2 风光出力 and 负荷曲线

Fig. 2 Output power curves of wind power, photovoltaic power and load

续储能配置与功率平衡分析提供了数据支撑。

图 3 对比了 DRL-MOPSO、DRL 与传统 PSO 在多目标优化性能上的差异。实验结果显示, DRL-MOPSO 在系统总成本、可再生能源利用率、电压稳定性三项核心指标中表现最优: 如图 3(a), DRL-MOPSO 方法的系统总成本为 100.50×10^6 元, 较 DRL、传统 PSO 分别减少 9.70×10^6 元、 19.65×10^6 元, 成本优化率分别为 8.8%、16.4%; 如图 3(b), DRL-MOPSO 方法的可再生能源利用率达 90.75%, 较 DRL、传统 PSO 分别提升 5.25%、10.45%, 体现了其新能源消纳的资源整合能力; 如图 3(c), DRL-MOPSO 方法的电压稳定性指标(标准化稳定系数)为 95.80, 较 DRL、传统 PSO 分别提高 2.40、5.60, 印证了其电网动态稳定保障能力。这一优势源于 DRL 实时运行策略的智能学习与 MOPSO 全局多目标寻优的协同, 破解了传统算法在多目标权衡中的局限性。

图 4 以 2023 年 7 月 1 日—31 日(共 31 天)的数据为例, 对比 DRL-MOPSO、DRL、传统 PSO 3 种方法在动态容量配置中的日度性能差异。图 4(a)为 3 种方法的边际投资成本日度分布, DRL-MOPSO 全月边际成本维持在 30~45 元/MWh, 其中 7 月 7 日—15 日(高温负荷高峰段)边际成本较 DRL 降低 12%~16%, 较传统 PSO 降低 18%~22%。图 4(b)为储能充放电效率日度分布, DRL-MOPSO 全月效率稳定在 $88.6\% \pm 0.8\%$, 仅在 7 月 10 日—13 日(阴雨导致风光出力波动增大)小幅降至 87.5%; 而 DRL 和传统 PSO 受出力波动影响更显著, 全月平均效率较 DRL-MOPSO 分别低 6.3% 和 10.8%。图 4(c)为电压波动抑制率日度分布, DRL-MOPSO 全月电压波动抑制率达 $92.3\% \pm 1.5\%$, 在 7 月 10 日(负荷骤增 20%)、7 月 25 日(风电出力骤降 35%)等极端场景中, 抑制率仍维持在 89% 以上; 而 DRL 和传统 PSO 在极端场景下抑制率分别降至 83.2% 和 77.1%, 全月平均抑制率较 DRL-MOPSO 低 6.6% 和 13.1%, 验证了 DRL-MOPSO 在月度长周期内的稳定性优势。

图 5 对算法的收敛特性进行了量化分析。如图 5(a)所示, DRL-MOPSO 在迭代初期(100 次以内)即达到稳定的多目标适应度值, 而 MOPSO 和 DRL 需要约 250 次迭代才能达到相近优化水平, DRL-MOPSO 相对 MOPSO 和 DRL 的收敛效率提

升幅度达40%~62.5%，如图5(b)所示。从帕累托前沿探索效果来看，DRL-MOPSO通过深度强化学习提供的状态-动作先验知识引导粒子群搜索，使解集覆盖度达到0.89(归一化值)，较单一算法提

升23%~37%，如图5(c)所示。这一特性表明DRL-MOPSO方法在处理高维复杂优化问题时，具备更优的全局搜索能力和解集多样性维持能力。

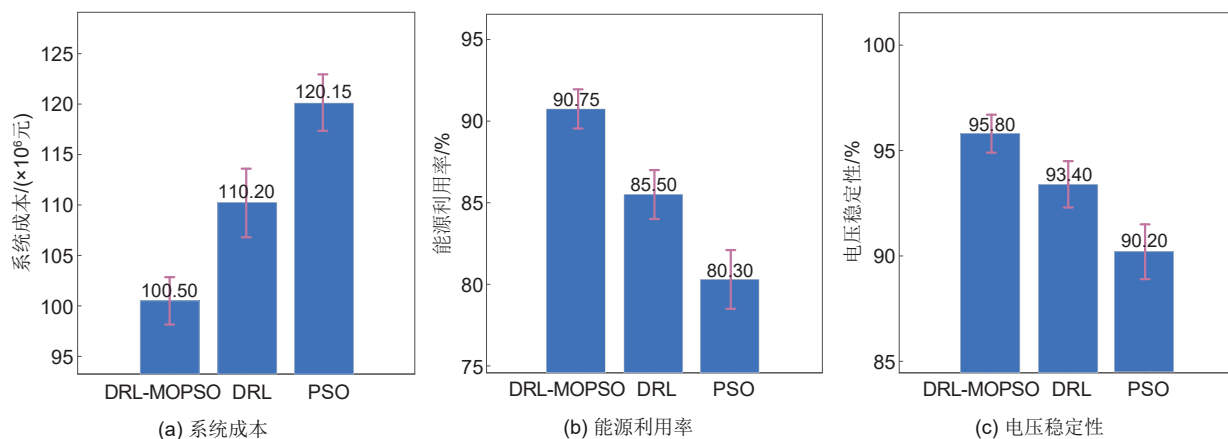


图3 DRL-MOPSO、DRL和传统PSO 3种方法在多目标优化性能方面的比较

Fig. 3 Comparison of multi-objective optimization performance among DRL-MOPSO, DRL and traditional PSO methods

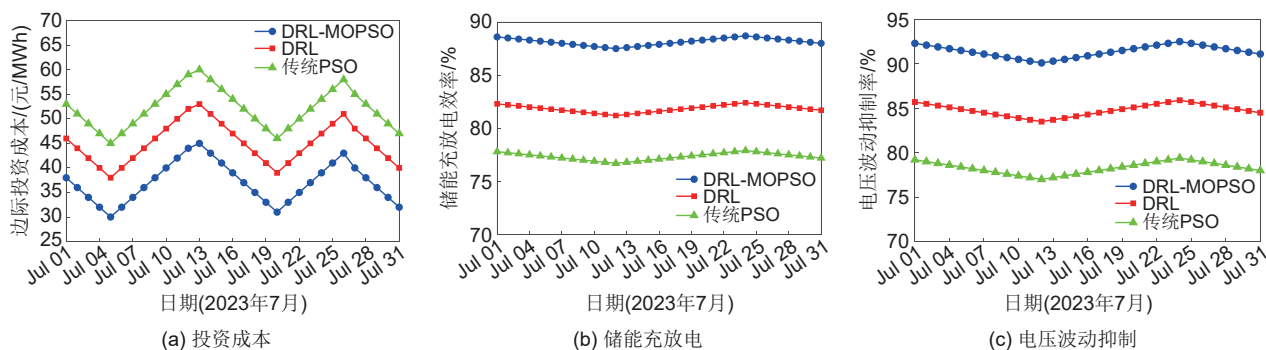


图4 动态容量配置的时间效应比较分析

Fig. 4 Comparative analysis of the temporal effects of dynamic capacity allocation

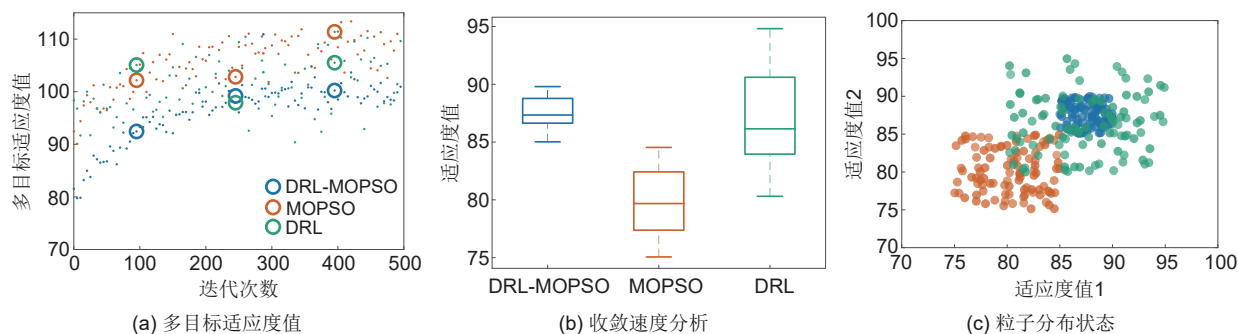


图5 算法收敛特性的量化分析

Fig. 5 Quantitative analysis of the convergence characteristics of the algorithms

4 结 论

本研究提出了基于深度强化学习和多目标粒子群优化的储能容量配置方法,旨在解决大规模可再生能源并网场景中的动态储能容量配置问题。通过深度整合DRL的动态决策能力和MOPSO的全局搜索优势,该方法有效地应对了可再生能源的不确定性和可变性,实现了投资成本、运行效率和电网稳定性等多目标的综合优化。分析结果表明,DRL-MOPSO在系统总成本、可再生能源利用率和电压稳定性等关键指标上具有显著优势。特别是在24 h时间窗口的动态容量配置中,DRL-MOPSO在投资成本分布、储能充放电效率和电压波动抑制率方面表现出优异的性能,这表明了其在实际应用中的潜力。

尽管DRL-MOPSO取得了显著成果,但仍有进一步优化的空间。未来的工作可能侧重于提高该算法在大规模系统中的计算效率,以及增强其在极端天气条件下对可再生能源输出波动的适应性。

参 考 文 献

- [1] 陈昀, 刘继雷, 刘蓓, 等. 基于不确定性天气数据的光伏发电系统电量分时段预测[J]. 中国测试, 2023, 49(S2): 76-83.
CHEN Yun, LIU Jilei, LIU Bei, et al. Timephased prediction of photovoltaic power generation system based on uncertain weather data[J]. China Measurement & Test, 2023, 49(S2): 76-83.
- [2] 周校聿, 刘尧, 鲍福增, 等. 百兆瓦级储能参与电网双重辅助服务调度的联合优化模型[J]. 电力系统自动化, 2021, 45 (19): 60-69.
ZHOU Xiaoyu, LIU Rao, BAO Fuzeng, et al. Joint optimization model for hundred-megawatt-level energy storage participating in dual ancillary services dispatch of power grid[J]. Automation of Electric Power Systems, 2021, 45(19): 60-69.
- [3] 闫震, 刘强, 李会彬, 等. 基于混合储能荷电状态的光伏微网功率优化管理方法[J]. 储能科学与技术, 2025, 14(5): 2067-2077.
YAN Z, LIU Q, LI H B, et al. Power optimization management method for photovoltaic microgrids based on the state of charge of hybrid energy storage systems[J]. Energy Storage Science and Technology, 2025, 14(5): 2067-2077.
- [4] 朱兰, 董凯旋, 唐陇军, 等. 计及同步机惯性与储能虚拟惯性价值的电能、惯性及一次调频联合优化出清模型[J]. 中国电机工程学报, 2024, 44 (19): 7543-7554. DOI:10.13334/j.0258-8013.pcsee.230306.
ZHU Lan, DONG Kaixuan, TANG Longjun, et al. Joint optimal clearing model for electric energy, inertia and primary frequency response considering synchronous inertia and energy storage virtual inertia values[J]. Proceedings of the CSEE, 2024, 44(19): 7543-7554. DOI:10.13334/j.0258-8013.pcsee.230306.
- [5] 柳长发, 付立衡, 张增丽, 等. 高比例光伏接入的分布式储能容量自适应协调控制方法[J]. 储能科学与技术, 2024, 13(8): 2696-2703.
LIU C F, FU L H, ZHANG Z L, et al. Adaptive coordinated control method for distributed energy storage capacity with high proportion of photovoltaic access[J]. Energy Storage Science and Technology, 2024, 13(8): 2696-2703.
- [6] 商立群, 闵鹏波, 张建涛. 基于改进PSO算法的光伏阵列MPPT研究[J]. 传感器与微系统, 2024, 43(8): 35-39.
SHANG Liqun, MIN Pengbo, ZHANG Jiantao. Research on MPPT of photovoltaic array based on improved PSO algorithm[J]. Transducer and Microsystem Technologies, 2024, 43(8): 35-39.
- [7] 刘凡, 李凤婷, 张高航, 等. 计及循环寿命和运营策略的风电汇集区域储能电站优化配置[J]. 电力系统保护与控制, 2023, 51(8): 127-139. DOI: 10.19783/j.cnki.pspc.221188.
LIU F, LI F T, ZHANG G H, et al. Optimal configuration of storage power stations in a wind power gathering area considering cycle life and operation strategy[J]. Power System Protection and Control, 2023, 51(8): 127-139. DOI: 10.19783/j.cnki.pspc.221188.
- [8] 阙凌燕, 蒋正威, 杨力强, 等. 基于改进遗传算法的源网荷储协同控制方法[J]. 沈阳工业大学学报, 2023, 45(6): 612-618.
QUE Lingyan, JIANG Zhengwei, YANG Liqiang, et al. Cooperative control method based on improved genetic algorithm for source network load storage[J]. Journal of Shenyang University of Technology, 2023, 45(6): 612-618.
- [9] 马骞, 肖亮, 程冰, 等. 基于强化学习-模型预测控制(RL-MPC)的分布式储能协同一次调频控制方法[J]. 储能科学与技术, 2025, 14(08): 3138-3148.
MA Qian, XIAO Liang, CHENG Bing, et al. Cooperative primary frequency modulation control method for distributed energy storage based on reinforcement learning model predictive control[J]. Energy Storage Science and Technology, 2025, 14(8): 3138-3148.
- [10] 周楠, 樊玮, 刘念, 等. 基于需求响应的光伏微网储能系统多目标容量优化配置[J]. 电网技术, 2016, 40(6): 1709-1716. DOI: 10.13335/j.1000-3673.pst.2016.06.015.
ZHOU N, FAN W, LIU N, et al. Battery storage multi-objective optimization for capacity configuration of PV-based microgrid considering demand response[J]. Power System Technology, 2016, 40(6): 1709-1716. DOI: 10.13335/j.1000-3673.pst.2016.06.015.
- [11] 卢秀丽, 安娟华, 冀松, 等. 基于粒子群优化的WSN多目标3D无人机运动规划[J]. 传感器与微系统, 2025, 44(11): 153-158.
LU X L, AN J H, JI S, et al. Multi-objective 3D UAV motion planning in WSN based on article swarm optimization[J]. Transducer and Microsystem Technologies, 2025, 44(11): 153-158.
- [12] 苏明健, 肖宝弟, 岳丽丽. 基于改进PSO-SA算法的城轨列车ATO节能优化研究[J]. 传感器与微系统, 2023, 42(10): 64-67, 76. DOI: 10.13873/J.1000-9787(2023)10-0064-04.
SU M J, XIAO B D, YUE L L. Research on ATO energy saving optimization of urban rail train based on improved PSO-SA algorithm[J]. Transducer and Microsystem Technologies, 2023, 42 (10): 64-67, 76. DOI: 10.13873/J.1000-9787(2023)10-0064-04.
- [13] 叶林, 王凯丰, 赖业宁, 等. 低惯量下电力系统频率特性分析及电池储能调频控制策略综述[J]. 电网技术, 2023, 47(2): 446-462. DOI: 10.13335/j.1000-3673.pst.2022.1269.
YE L, WANG K F, LAI Y N, et al. Review of frequency characteristics analysis and battery energy storage frequency regulation control strategies in power system under low inertia level[J]. Power System Technology, 2023, 47(2): 446-462. DOI: 10.13335/j.1000-3673.pst.2022.1269.