

基于改进 TD3 的农村微电网源网荷储优化运行策略

陈国齐¹, 魏颖豪¹, 刘毅¹, 李梦锐¹, 熊马超¹, 王灿²

(1. 国网湖北省电力有限公司荆州供电公司, 湖北 荆州 431700; 2. 三峡大学电气与新能源学院, 湖北 宜昌 443002)

摘要: 农村微电网作为支撑“双碳”目标落地、赋能乡村振兴战略实施的重要载体, 通过整合分布式能源与传统发电资源, 能够有效提升系统整体的运行效率。为此, 提出了一种基于改进双延迟深度确定性策略梯度(twin delayed deep deterministic policy gradient, TD3)的农村微电网源网荷储优化运行策略。首先, 综合考虑运行成本和可再生能源利用率, 构建了含沼气的农村微电网源网荷储优化运行模型。其次, 构建了基于复合评分机制的注意力加权评论家网络(CASA-Critic), 动态提取并聚焦关键状态特征以提升价值网络的评估精度。然后, 采用动态感知策略, 基于训练阶段在时间进程与环境状态的不确定性特征动态调节动作探索强度, 改善算法在探索与利用之间的平衡。最后, 基于典型的农村微电网源网荷储运行场景进行仿真分析, 仿真结果表明, 相较于其他基于深度强化学习的优化运行策略, 所提方法具有更好的收敛性和寻优效果, 在提升系统运行经济性、促进可再生能源消纳方面具有明显的优越性。

关键词: 农村微电网; 优化运行; 注意力加权评论家网络; 动态感知策略; 深度强化学习

Optimal Operation Strategy for Source-Grid-Load-Storage in Rural Microgrids Based on Improved TD3

CHEN Guoqi¹, WEI Yinghao¹, LIU Yi¹, LI Mengrui¹, XIONG Machao¹, WANG Can²

(1. Jingzhou Power Supply Company, State Grid Hubei Electric Power Co., Ltd., Jingzhou, Hubei 431700, China; 2. College of Electrical Engineering and New Energy, China Three Gorges University, Yichang, Hubei 443002, China)

Abstract: As a crucial carrier for supporting the implementation of the "dual carbon" goals and empowering rural revitalization, rural microgrids can effectively enhance overall system operational efficiency by integrating distributed energy with traditional power generation resources. To this end, this paper proposes a optimal operation strategy for source-grid-load-storage in rural microgrids based on an improved twin delayed deep deterministic policy gradient (TD3) algorithm. Firstly, an optimal operation model for rural microgrids incorporating biogas is constructed by comprehensively considering operating costs and renewable energy utilization. Secondly, a compound assessment scored attention-weighted critic(CASA-Critic)is designed to dynamically extract and focus on key state features, thereby enhancing the evaluation accuracy of the value network. Then, a dynamic perception strategy is introduced to adaptively adjust the intensity of action exploration based on the temporal progression and environmental state uncertainties during training, improving the balance between exploration and exploitation. Finally, simulation analysis is conducted based on a typical rural microgrids source-grid-load-storage operation scenario. The results demonstrate that, compared to optimization strategies based on other deep reinforcement learning, the proposed method exhibits better convergence and optimization performance, while also showing significant superiority in improving system operational economy and promoting renewable energy accommodation.

Key words: rural microgrid; optimal operation; CASA-critic; dynamic perception strategy; deep reinforcement learning

0 引言

随着“双碳”目标的推进与农村能源结构的转型,农村微电网作为整合分布式能源、提升供电可靠性的重要载体正逐步成为推动乡村振兴的有效手段^[1-2]。然而,农村地区普遍存在能源分布不均、新能源的就地消纳能力差、供电不足等问题,给微电网的稳定运行带来严峻挑战。因此,如何适应农村微电网复杂环境并实现源网荷储的优化运行已成为当前研究的热点^[3-4]。

为进一步解决微电网的优化运行问题,文献[5]提出了一种考虑需求侧管理的风光燃储微电网两阶段优化调度策略,降低微电网运行成本与碳排放量,实现微电网的低碳经济运行。文献[6]提出了一种含有共享储能的微电网群分布鲁棒博弈优化调度方法,提高了微电网内部新能源的消纳能力。文献[7]基于改进白鲸优化算法提出了一种含风光燃储的微电网优化调度策略,在降低微电网系统运行的经济成本的同时提高了系统的可再生能源消纳比例。文献[8]提出了一种数据中心参与风光抽蓄互补微电网优化调度方法,能够有效平抑风光出力波动并提升系统经济效益。文献[9]提出了一种改进风光场景聚类联合虚拟储能的源网荷储低碳优化调度方法,有效降低了系统碳排放量与运行成本。然而,这类研究大多面向一般微电网场景,重点关注风、光、储和常规燃机之间的协调关系,对于农村微电网中具有可调度潜力的沼气资源关注不足,尚未充分体现农村地区生物质资源丰富及沼气兼顾供能与调节双重功能的应用特点。因此,如何结合农村场景并将沼气纳入源网荷储优化流程仍是值得进一步研究的问题。

近年来,深度强化学习(deep reinforcement learning, DRL)作为一种不依赖于精确系统模型的算法,凭借其自主学习与动态环境适应性的能力^[10-11],在微电网能量管理和优化调度方面得到广泛应用。文献[12]在考虑源荷不确定性的前提下提出了一种基于连续型深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法的微电网能量调度方法,进一步提升了微电网运行的经济性。文献[13]提出了一种动作双重深度Q网络(double deep Q-network, DDQN)的微电网能量在线优化调度策略,有效降低了微电网运行成本,延长了设备寿命。文献[14]提出

了一种基于联邦强化学习的微电网储能调度方法,提高了微电网自给率。文献[15]将多维动作近端策略优化算法(multi-dimensional action proximal policy optimization, MDA-PPO)应用到区块链中进行微电网经济调度,实现了可再生能源的最大化。上述研究表明,深度强化学习能够突破传统优化方法对精确模型和静态场景的依赖,在复杂动态环境中具备较强的自学习与滚动决策能力。但进一步分析可知,上述算法存在Q值高估或评估偏差累积现象,可能导致策略更新方向失真。因此,仅仅将通用DRL算法直接应用于农村微电网并不能保证评估精度能够达到理想效果。

为进一步抑制Q值网络中的过估计现象,提升策略收敛的准确性,文献[16]针对调度策略灵活性不足的问题提出了一种基于数据驱动的方式获得微电网调度策略,通过双延迟深度确定性策略梯度(twin delayed deep deterministic policy gradient, TD3)算法有效降低了系统运行成本和不平衡度。文献[17]提出了一种基于改进TD3算法的微电网低碳经济优化调度策略,降低了微电网峰值负荷和用户用能成本。文献[18]基于柔性行动器-评判器(soft actor-critic, SAC)框架的深度强化学习方法实现了自适应源荷不确定特征的微网经济调度。文献[19]运用优先回放机制改进SAC算法,进一步提高了算法决策的准确性。上述研究表明,围绕价值评估机制和样本利用方式进行改进是提升深度强化学习调度性能的重要方向。但上述研究仍存在两方面不足:其一,Critic网络大多采用全连接结构对状态特征进行统一处理,难以根据微电网运行状态动态识别并强化关键特征;其二,动作探索方式仍采用固定噪声进行探索,未能根据训练进程和环境状态变化动态调整探索强度,从而限制了训练效率和策略收敛质量。对于运行环境更复杂的农村微电网而言,上述问题会更加突出。

为了解决上述问题,本文提出了一种基于改进TD3的农村微电网源网荷储优化运行策略。首先,建立了含沼气的农村微电网源网荷储优化模型,并充分考虑可再生能源利用率和系统总运行成本;然后,针对传统TD3的Q值网络评估不准确的问题,引入注意力机制改进Critic网络动态提取并聚焦关键状态特征,以提升微电网优化运行策略的寻优能力和决策精度;同时,本文基于时间衰减项和空间感知项提出了一种动态感知策略,以提升算法初期

的训练效率,帮助算法在学习过程中平衡探索和利用的关系,进而生成更合理的农村微电网优化运行方案。最后,通过算例分析验证了本文所提改进TD3优化运行策略的有效性和优越性。

1 农村微电网系统源网荷储协同优化模型

本文所研究的农村微电网由风电机组、光伏机组、沼气发电机组、储能系统以及用户负荷组成。该微电网通过整合风光沼储等多能互补系统,有效解决了可再生能源发电与负荷需求不匹配的问题,进而实现了可再生能源的就地消纳。其具体结构如图1所示。

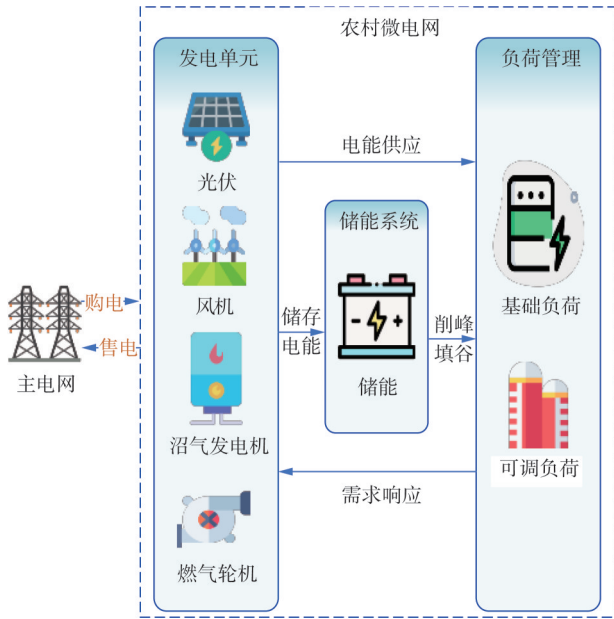


图1 农村微电网系统结构

Fig.1 Rural microgrids system architecture

1.1 目标函数

为实现农村微电网在运行过程中经济性和环保性需求,本文以系统运行总成本最低和可再生能源利用率最高作为目标优化函数。

1) 经济性目标函数

经济性目标函数以系统运行总成本最小化为目标,其表述如下。

$$\min F_1 = C_{\text{grid}} + C_{\text{fuel}} + C_{\text{om}} + C_{\text{dr}} + C_{\text{startup}} + C_{\text{bio, total}} \quad (1)$$

式中: F_1 为系统运行总成本; C_{grid} 为微电网与主网的交互成本; C_{fuel} 为燃气轮机的燃料成本; C_{om} 为所有设备的运维成本; C_{dr} 为激励型需求响应的补偿

成本; C_{startup} 为燃料机组启停成本; $C_{\text{bio, total}}$ 为沼气成本。各类成本具体计算方式如式(2)所示。

$$\begin{cases} C_{\text{grid}} = \sum_{t=1}^T [C_{\text{buy}}(t)P_{\text{buy}}(t) - C_{\text{sell}}(t)P_{\text{sell}}(t)] \\ C_{\text{fuel}} = \sum_{t=1}^T c_{\text{fuel, MT}} \cdot F_{\text{fuel, MT}}(t) \\ C_{\text{om}} = \sum_i k_{\text{om, } i} P_i(t) \\ C_{\text{dr}} = \sum_{t=1}^T c_{\text{curt}} P_{\text{curt}}(t) \Delta t \\ C_{\text{startup}} = \sum_{t=1}^T c_{\text{su, MT}} \cdot I_{\text{su, MT}}(t) \\ C_{\text{bio, total}} = \sum_{t=1}^T [C_{\text{bio, base}}(t) + C_{\text{bio, dispatch}}(t)] \end{cases} \quad (2)$$

式中: $C_{\text{buy}}(t)$ 、 $C_{\text{sell}}(t)$ 分别为 t 时刻向主网的购电价格和售电价格; $P_{\text{buy}}(t)$ 、 $P_{\text{sell}}(t)$ 分别为 t 时刻向主网的购电价格和售电价格; T 为完整的一个调度周期,本文取值为 24 h; $c_{\text{fuel, MT}}$ 为燃气轮机单位燃料成本; $k_{\text{om, } i}$ 为第 i 台设备的单位功率运维成本系数; $P_i(t)$ 为第 i 台设备的单位功率; $F_{\text{fuel, MT}}(t)$ 为 t 时刻燃气的消耗量; Δt 为单个调度周期的持续时间,本文取值为 1 h; $P_{\text{curt}}(t)$ 为负荷削减量; $I_{\text{su, MT}}(t)$ 为启停标识变量,启动时为 1,停机时为 0; c_{curt} 为单位功率的加负荷成本; $c_{\text{su, MT}}$ 为单次启停成本; $C_{\text{bio, base}}(t)$ 为 t 时刻沼气的基础成本; $C_{\text{bio, dispatch}}(t)$ 为 t 时刻沼气的调度成本。

为更精确地描述农村微电网中沼气能源的时变特性与调度,本文从基础成本与动态调度成本两个维度构建其成本模型,具体表述如式(3)所示。

$$\begin{cases} C_{\text{bio, base}} = \sum_{t=1}^T \left(\alpha \cdot p_{\text{raw}} \cdot m_{\text{raw}}(t) + \beta \cdot p_{\text{elec}}(t) \cdot p_{\text{process}}(t) \right) \cdot \Delta t \\ C_{\text{bio, dispatch}} = \sum_{t=1}^T \left[V_{\text{store}}(t) \cdot p_{\text{store}} \cdot \delta_{\text{loss}}(t) + p_{\text{penalty}} \cdot \max(0, V_{\text{plan}}(t) - V_{\text{actual}}(t)) \right] \end{cases} \quad (3)$$

式中: α 为原料-沼气转化系数; p_{raw} 为单位沼气原料的收集成本; $m_{\text{raw}}(t)$ 为 t 时刻制备沼气的原料总质量; β 为制备沼气能耗折算系数; $p_{\text{elec}}(t)$ 为 t 时刻微电网内的用电单价; $p_{\text{process}}(t)$ 为 t 时刻沼气制气设备的运行功率; $V_{\text{store}}(t)$ 为 t 时刻沼气罐内储气体积; p_{store} 为单位沼气的存储成本; $\delta_{\text{loss}}(t)$ 为 t 时刻沼气在储存过程中的损耗率; p_{penalty} 为单位体积沼气供需缺口惩罚单价; $V_{\text{plan}}(t)$ 为 t 时刻沼气调度的计划用量;

$V_{\text{actual}}(t)$ 为 t 时刻沼气调度的计划实际用量。

2) 环保性目标函数

该函数以可再生能源消纳水平最高为目标, 表示清洁能源在用电需求的占比, 具体如式(4)所示。

$$\max F_2 = \frac{\sum_{t=1}^T [P_{\text{PV}}(t) + P_{\text{WT}}(t) + P_{\text{G}}(t)] \Delta t}{\sum_{t=1}^T P_{\text{load}}^{\text{total}}(t) \Delta t} \times 100\% \quad (4)$$

式中: F_2 为可再生能源利用率; $P_{\text{PV}}(t)$ 为 t 时刻光伏产生的功率; $P_{\text{WT}}(t)$ 为 t 时刻风机产生的功率; $P_{\text{G}}(t)$ 为 t 时刻沼气发电机产生的功率; $P_{\text{load}}^{\text{total}}(t)$ 为微电网中负荷所需要的功率之和。

3) 加权目标函数

为了便于求解目标函数, 将多目标问题通过加权转换为单目标问题以便求解^[20], 具体如式(5)所示。

$$\min F = \omega_1 \cdot \frac{F_1 - F_{1,\min}}{F_{1,\max} - F_{1,\min}} - \omega_2 \cdot \frac{F_2 - F_{2,\min}}{F_{2,\max} - F_{2,\min}} \quad (5)$$

式中: F 为加权后的总成本; ω_1 、 ω_2 分别为总运行成本和可再生能源利用率的权重系数; $F_{1,\max}$ 、 $F_{2,\max}$ 分别为各自场景下的总运行成本和可再生能源利用率的最大值; $F_{1,\min}$ 、 $F_{2,\min}$ 分别为各自场景下的总运行成本和可再生能源利用率的最小值。

1.2 约束条件

为确保农村微电网优化运行模型的可行性和高效性, 需遵循以下关键约束条件。

1) 功率平衡约束

$$P_{\text{PV}}(t) + P_{\text{WT}}(t) + P_{\text{G}}(t) + P_{\text{bat}}(t) + P_{\text{buy}}(t) + P_{\text{MT}}(t) = P_{\text{load}}^{\text{total}}(t) + P_{\text{sell}}(t) \quad (6)$$

式中: $P_{\text{bat}}(t)$ 为 t 时刻储能系统产生的功率; $P_{\text{buy}}(t)$ 为 t 时刻微电网中购电功率; $P_{\text{MT}}(t)$ 为 t 时刻微型燃气轮机产生的功率; $P_{\text{sell}}(t)$ 为 t 时刻微电网中售电功率。

2) 设备运行约束

(1) 沼气发电机

$$\begin{cases} P_{\text{G}}^{\min} \leq P_{\text{G}}(t) \leq P_{\text{G}}^{\max} \\ |P_{\text{G}}(t) - P_{\text{G}}(t-1)| \leq \Delta P_{\text{G}}^{\max} \end{cases} \quad (7)$$

式中: $P_{\text{G}}^{\min}$ 、 $P_{\text{G}}^{\max}$ 分别为沼气发电机允许的最小稳定出力 and 最大稳定出力; $\Delta P_{\text{G}}^{\max}$ 为相邻两段期间沼气发电机出力允许的最大变化量。

(2) 微型燃气轮机

$$\begin{cases} P_{\text{MT}}^{\min} \cdot I_{\text{MT}}(t) \leq P_{\text{MT}}(t) \leq P_{\text{MT}}^{\max} \cdot I_{\text{MT}}(t) \\ |P_{\text{MT}}(t) - P_{\text{MT}}(t-1)| \leq \Delta P_{\text{MT}}^{\max} \\ \sum_{\tau=t-T_{\text{on}}^{\min}+1}^t I_{\text{MT}}(\tau) \geq T_{\text{on}}^{\min} \cdot [I_{\text{MT}}(t) - I_{\text{MT}}(t-1)] \\ \sum_{\tau=t-T_{\text{off}}^{\min}+1}^t [1 - I_{\text{MT}}(\tau)] \geq T_{\text{off}}^{\min} \cdot [I_{\text{MT}}(t-1) - I_{\text{MT}}(t)] \end{cases} \quad (8)$$

式中: $P_{\text{MT}}^{\min}$ 、 $P_{\text{MT}}^{\max}$ 分别为燃气轮机允许的最小稳定出力 and 最大稳定出力; $I_{\text{MT}}(t)$ 为微型燃气轮机在时间 t 的启停状态; $\Delta P_{\text{MT}}^{\max}$ 为相邻两段期间微型燃气轮机出力允许的最大变化量; τ 为时间常数, 表示最小持续运行及最小停机时间; $I_{\text{MT}}(\tau)$ 为在时间 τ 下微型燃气轮机的运行状态; $T_{\text{on}}^{\min}$ 为微型燃气轮机一次启动后必须连续运行的最短时间; $T_{\text{off}}^{\min}$ 为微型燃气轮机一次停运后必须连续停机的最短时间。

(3) 储能系统

$$\begin{cases} P_{\text{bat}}^{\text{dis},\min} \leq P_{\text{bat}}(t) \leq P_{\text{bat}}^{\text{ch},\max} \\ S_{\text{OC}}^{\min} \leq S_{\text{OC}}(t) \leq S_{\text{OC}}^{\max} \\ S_{\text{OC}}(t) = S_{\text{OC}}(0) \end{cases} \quad (9)$$

式中: $P_{\text{bat}}^{\text{ch},\max}$ 、 $P_{\text{bat}}^{\text{dis},\min}$ 分别为储能系统允许的最大充电功率和最小放电功率; $S_{\text{OC}}^{\min}$ 、 $S_{\text{OC}}^{\max}$ 分别为储能系统允许的最小和最大荷电状态, 防止储能系统过放和过充; $S_{\text{OC}}(0)$ 、 $S_{\text{OC}}(t)$ 分别为调度期初和调度期末的储能系统的荷电状态。此约束保证储能系统在调度周期内能量守恒, 不影响下一个周期的调度。

3) 电网交互约束

$$\begin{cases} 0 \leq P_{\text{buy}}(t) \leq P_{\text{grid}}^{\max} \\ 0 \leq P_{\text{sell}}(t) \leq P_{\text{grid}}^{\max} \end{cases} \quad (10)$$

式中 $P_{\text{grid}}^{\max}$ 为微电网与主网连接点的最大交换功率。

4) 需求响应约束

$$\begin{cases} 0 \leq P_{\text{curt}}(t) \leq P_{\text{curt}}^{\max} \\ \sum_{t=1}^T P_{\text{curt}}(t) \leq E_{\text{curt}}^{\text{total}} \end{cases} \quad (11)$$

式中: $P_{\text{curt}}^{\max}$ 为时间 t 内允许的最大负荷削减量; $E_{\text{curt}}^{\text{total}}$ 为整个调度周期内允许负荷削减总量的上限值, 防止对用户造成过度影响。

2 改进 TD3 的农村微电网源网荷储优化运行算法

2.1 马尔可夫模型设计

农村微电网优化运行问题在数学上可表述为一

种涵盖多能源出力调节与电能交互的复杂过程, 其本质是一个包含高维状态空间、连续动作空间以及多重约束的优化问题。因此本文将上述优化问题转化为部分可观测马尔可夫决策过程。该过程由四元组 $\{s, a, p, r\}$ 构成^[21]。其中 s 表示环境的全局状态集合, 包含描述所有设备及环境的信息; $a = \{a_1, a_2, \dots, a_n\}$ 为所有设备的联合动作空间, 其中, n 为设备的数量, a_i 表示第 i 个设备的动作空间; P 为状态转移函数, 描述从当前状态 s 到下一个状态 s' 的概率分布; $r = \{r_1, r_2, \dots, r_n\}$ 为奖励函数, r_i 为设备 i 在当前状态执行联合动作后所获得的即时奖励。为了更精确描述马尔可夫决策过程, 其关键元组状态空间、动作空间及奖励函数表示如下。

1) 状态空间

在 t 时刻智能体所能观测的环境信息包括光伏出力 $P_{PV}(t)$ 、风机出力 $P_{WT}(t)$ 、负荷功率 $P_{load}^{total}(t)$ 、储能系统的荷电状态 $S_{OC}(t)$ 、与主网交换的功率 $P_{buy}(t)$ 、设备启停量 $I_{MT}(t)$ 和电网电价 $C_{price}(t)$ 。因此, t 时刻决策的状态空间 s_t 定义为:

$$s_t = [P_{PV}(t), P_{WT}(t), P_{load}^{total}(t), S_{OC}(t), P_{buy}(t), I_{MT}(t), C_{price}(t)] \quad (12)$$

2) 动作空间

在农村微电网中, 动作空间包括沼气发电机出力 $P_G(t)$ 、燃气轮机出力 $P_{MT}(t)$ 、储能系统充放电功率 $P_{bat}(t)$ 以及负荷削减量 $P_{curt}(t)$ 。因此, t 时刻的动作空间 Ω_t 具体表述如下。

$$\Omega_t = [P_{G,t}, P_{MT,t}, P_{bat,t}, P_{curt,t}] \quad (13)$$

3) 奖励函数

环境依据决策动作触发状态转移, 并通过状态观测反馈即时奖励信号。智能体通过约束动作空间优化策略, 以实现长期累积奖励最大化。因此, 本文 t 时刻智能体执行动作后获得的即时奖励 r_t 的表达式为:

$$r_t = - \left(F + \rho \cdot \frac{\sum (\Delta P_i(t) + \Delta R_i(t))}{P_{total} + R_{total}} \right) \quad (14)$$

式中: F 为总加权成本, 鼓励智能体降低运行成本和提高可再生能源消纳比例; P_{total} 、 R_{total} 分别为总功率和总爬坡功率上限; ρ 为惩罚系数, 确保任何约束违反导致的惩罚远高于其他惩罚; $\Delta P_i(t)$ 为设备 i 的出力越限功率; $\Delta R_i(t)$ 为设备 i 的爬坡越限功率。

2.2 改进 TD3 算法设计

双延迟深度确定性策略梯度 (twin delayed deep deterministic policy gradient, TD3) 算法作为深度强化学习领域的重要延伸, 能够解决微电网中各种调度及优化运行问题^[22]。该算法通过独特的 Actor-Critic 架构能够有效处理高维连续状态空间和动作空间, 实现多尺度的协调优化^[23]。尽管 TD3 算法在理论上极具优势, 但 TD3 在面向农村微电网这类具有强随机性特征的复杂系统时仍存在两大局限: 其一, TD3 的 Critic 网络依赖全连接层处理状态特征, 无法适应系统的动态变化需求, 导致系统评估存在偏差; 其二, TD3 采用固定方差的高斯噪声进行探索, 未能结合训练进程与环境状态调整探索强度, 限制了探索效率与策略优化的平衡。针对 TD3 存在的不足, 本文构建了基于复合评分机制的注意力加权^[24] Critic 网络 (compound assessment scored attention-weighted critic, CASA-Critic) 和动态感知策略, 以提升算法在复杂场景下的评估准确性和决策效率。

1) CASA-Critic 网络

传统 TD3 算法的 Critic 网络在处理高维连续状态-动作空间时, 通常采用全连接神经网络结构对输入特征进行统一映射, 其 Q 值函数可表述为:

$$Q(s_t, a_t) = f_{FC}(\phi(s_t, a_t)) \quad (15)$$

式中: $Q(s_t, a_t)$ 为动作价值函数, 表示智能体处于状态 s_t 、执行动作 a_t 后, 按照当前策略继续进行所能获得的期望回报; $\phi(s_t, a_t)$ 表示处于状态 s_t 、执行动作 a_t 后的拼接特征; $f_{FC}(\phi(s_t, a_t))$ 为多层全连接神经网络构成的非线性映射。

上述结构虽然具有较强的表示能力, 但其全连接层对所有输入特征进行均匀处理, 无法区分不同特征维度在系统运行过程中的相对重要性, 难以有效聚焦关键运行状态, 导致 Q 值估计存在系统性偏差, 进而影响策略更新的准确性与收敛效率。为进一步提升 Critic 网络对关键状态特征的动态识别与聚焦能力, 本文采用复合评分机制及注意力权重计算相结合的方式, 评估各维特征的重要性, 使该网络能够动态聚焦于系统关键状态特征。具体改进步骤如下。

首先, 通过特征映射层将状态-动作对转换为特征矩阵, 以便提取状态-动作对中的潜在特征,

为后续注意力机制提供丰富的特征表示。

$$\mathbf{H} = \mathbf{W}_h [s_t \| a_t] + \mathbf{b}_h \quad (16)$$

式中: $[s_t \| a_t]$ 为 t 时刻状态空间 s_t 和执行动作 a_t 的特征拼接; $\mathbf{W}_h$ 、 $\mathbf{b}_h$ 分别为特征映射矩阵及偏置向量, 其中 h 为特征映射层; $\mathbf{H}$ 为原始状态-动作对转换为高维特征表示且 $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_i, \dots, \mathbf{h}_n]$; $\mathbf{h}_i$ 为矩阵 $\mathbf{H}$ 中的第 i 个特征向量。

其次, 为量化各个特征维度重要性, 构建复合评分机制, 计算每个特征向量的得分, 并通过 Softmax 函数对得分进行归一化, 得到每个特征的注意力权重, 以动态评估各特征维度的重要程度。每个特征向量的注意力权重为:

$$\begin{cases} e_i = \text{LeakyReLU}(\mathbf{W}_m \mathbf{h}_i + \mathbf{b}_m) \\ \alpha_i = \frac{\exp(e_i)}{\sum_{j=1}^n \exp(e_j)} \end{cases} \quad (17)$$

式中: $\mathbf{W}_m$ 为将 $\mathbf{h}_i$ 向量映射到注意力评分层的注意力权重参数, 其中, i 为特征向量的个数; $\mathbf{b}_m$ 为注意力偏置项; m 为注意力评分层; LeakyReLU 为激活函数在负值区域的斜率; α_i 为第 i 个特征向量 $\mathbf{h}_i$ 的注意力权重; e_i 为第 i 个特征向量 $\mathbf{h}_i$ 注意力得分; $\exp(e_i)$ 为对注意力得分 e_i 的取指运算。

然后, 基于注意力权重对原始特征进行加权聚合, 生成聚焦关键信息的特征表示如式(18)所示。

$$\mathbf{h}_m = \sum_{i=1}^n \alpha_i \mathbf{h}_i \quad (18)$$

式中 $\mathbf{h}_m$ 为经过注意力权重加权后得到的特征向量。

最后, 将聚合后的特征输入至后续全连接网络, 输出动作价值估计如式(19)所示。

$$Q(s, a) = f_{\text{FC}}(\mathbf{h}_{\text{att}}) \quad (19)$$

式中: $Q(s, a)$ 为在状态 s 、动作 a 下的 Q 值; f_{FC} ($\cdot$) 为全神经网络计算。

上述 CASA-Critic 网络通过复合评分机制动态计算的注意力权重实现了特征选择与信息聚焦。该机制使 Critic 网络能够自适应地强化与当前系统状态关联度高的特征, 显著提升 Q 值估计的准确性与对动态环境的适应性, 为策略梯度的高效更新提供了更为可靠的价值评估基础。

2) 动态感知策略

传统 TD3 算法在动作探索阶段通常采用固定方差的高斯噪声扰动策略输出, 其动作生成方式如式

(20) 所示。

$$a_t = \pi_{\theta}(s_t) + \epsilon, \epsilon \sim \mathcal{N}(0, \sigma_{\text{fixed}}) \quad (20)$$

式中: a_t 为 t 时刻的探索动作; $\pi_{\theta}(s_t)$ 为 Actor 网络输出的确定性动作; θ 为当前 Actor 网络参数; σ_{fixed} 为预设固定噪声的标准差; ϵ 为加入动作中的高斯探索噪声。

上述方法虽然实现简单, 但在面向农村微电网这类动态系统时存在明显不足。首先, 固定方差的探索机制难以反映训练进程, 导致训练后期仍存在不必要的随机扰动, 降低了策略的收敛效率; 其次, 未考虑系统状态的不确定性, 在状态空间不同区域采用相同的探索强度, 难以实现探索与利用的有效平衡, 易导致策略陷入局部最优或收敛迟缓。针对上述不足, 本文通过融合时间衰减与状态感知项^[24-27], 提出了一种动态感知策略以动态调整探索强度。其改进步骤如下。

首先, 为降低历史陈旧样本对当前策略更新的影响, 引入时间衰减机制, 对经验样本赋予不同权重。设经验回放池中第 i 条样本的采样时间为 t_i , 当前训练时刻为 t , 则其时间衰减权重定义为:

$$\omega_i = \exp(-\varphi \cdot (t - t_i)) \quad (21)$$

式中: ω_i 为第 i 个样本的时间权重; φ 为时间衰减系数, 用于调节历史信息的衰减速度。

基于上述时间衰减权重给定历史状态样本集合 $S = \{(s^{(i)}, \omega_i) | i = 1, 2, \dots, N\}$ 。其中, $s^{(i)}$ 为历史状态样本集合中的第 i 个状态向量。在训练开始前, 利用历史运行数据或经验回放池的预填充样本, 对状态空间进行离线聚类建模, 结合时间衰减权重, 对状态空间进行集中式加权聚类分析, 提取系统的典型运行模式。聚类完成后得到一组典型状态中心, 其构建方式如式(22)所示。

$$\mathbf{c}_k = \frac{\sum_{s^{(i)} \in c_k} \omega_i s^{(i)}}{\sum_{s^{(i)} \in c_k} \omega_i} \quad (22)$$

式中: K 为聚类簇数量; c_k 为第 k 类状态集合; $\mathbf{c}_k$ 为第 k 类状态的加权聚类中心向量。

其次, 在当前时刻 t 智能体观测到系统状态 s_t 时, 通过计算该状态与所有典型状态中心之间的最小欧氏距离, 构建状态感知距离指标定义其与历史典型状态之间的欧式距离如式(23)所示。

$$d_t = \min_{k=1, \dots, K} \|s_t - \mathbf{c}_k\|_2 \quad (23)$$

式中 d_t 为当前状态相对于历史运行模式的陌生程度,距离越大表示当前状态越偏离已有经验。

为将状态感知结果引入动作探索过程,对距离指标进行归一化处理,并结合时间衰减项构建动态感知权重。其构建方式如式(24)所示。

$$\begin{cases} \eta_t = \frac{d_t - d_{\min}}{d_{\max} - d_{\min}} \\ \lambda_t = \lambda_{\min} + (\lambda_{\max} - \lambda_{\min}) \cdot \eta_t \cdot \exp(-\beta \Delta t) \end{cases} \quad (24)$$

式中: $d_{\max}$ 和 $d_{\min}$ 分别为历史距离的最大值和最小值; η_t 为状态陌生度因子; λ_t 为动态感知权重,用于调节状态陌生度在奖励或价值估计中的影响程度; $\lambda_{\min}$ 、 $\lambda_{\max}$ 分别为动态感知权重下限和上限; Δt 为时间间隔,表示当前状态与最近一次相似状态出现之间的时间间隔。

然后,通过归一化距离和动作探索噪声对总体探索强度的贡献程度构建综合状态项,形成既考虑训练进程又响应环境变化的动态探索策略。其构建方式如式(25)所示。

$$\begin{cases} \sigma_t = \sigma_{\min} + (\sigma_{\max} - \sigma_{\min}) \cdot \eta_t \\ a_t = \pi_{\theta}(s_t) + \epsilon_t, \epsilon_t \sim \mathcal{N}(0, \sigma_t) \end{cases} \quad (25)$$

式中: σ_t 为综合状态项; $\sigma_{\max}$ 、 $\sigma_{\min}$ 分别为动态探索噪声的上下限; ϵ_t 为 t 时刻动态探索噪声的标准差服从正态分布的动态感知因子; a_t 为 t 时刻服从动态感知策略探索动作。

最后,在完成探索动作生成后,基于TD3的双Critic结构与目标网络机制,对下一状态-动作对的价值进行评估,构建目标Q值并更新当前价值函数。为此,目标Q值的计算可进一步改写为:

$$y_t = r_t - \lambda_t d_t + \gamma \min_{i=1,2} Q_{\theta'_i}(s_{t+1}, \pi_{\phi'}(s_{t+1})) \quad (26)$$

式中: y_t 为 t 时刻的目标Q值, r_t 为 t 时刻的即时奖励; γ 为折扣因子; $Q_{\theta'_i}(s_{t+1}, \pi_{\phi'}(s_{t+1}))$ 为第 i 个目标Critic网络在下一状态 s_{t+1} 时对应的动作价值估计, s_{t+1} 为智能体执行当前动作后得到的 $t+1$ 时刻环境状态; $\pi_{\phi'}(s_{t+1})$ 为目标Actor网络根据下一状态 s_{t+1} 输出的目标动作; θ'_i 为第 i 个目标Critic网络参数; ϕ' 为目标Actor网络参数。

上述机制与传统固定噪声探索相比,具有双重自适应能力:时间衰减机制保障了从充分探索到高效利用的平稳过渡;状态感知机制则在系统运行不确定性高时增强探索,在状态稳定时减少不必要的扰动。基于该机制生成的探索动作能够更有效地覆

盖高价值状态区域,避免在熟悉环境中过度扰动。同时,结合双目标CASA-Critic网络计算的目标Q值能够在探索过程中提供更为准确的价值评估,减少因随机噪声引入的Q值估计偏差,有助于算法在农村微电网复杂动态环境中更快收敛至更优策略。

2.3 算法框架及流程

结合上述改进方法得到本文设计的基于改进TD3的微电网优化运行策略流程如图2所示,其具体步骤如下。

1)初始化环境状态和所有状态集合;初始化Actor当前策略网络、两个CASA-Critic价值网络和目标策略网络Actor的参数;初始化动态感知策略参数;设置经验回放缓冲区大小、最大训练回合数、目标网络软更新系数等超参数。

2)根据当前智能体的观测值,通过CASA-Actor网络结合动态感知策略式(25)输出动作,执行动作后观测环境反馈的状态转移及奖励,并将完整经验元组(状态、动作、奖励、下一状态)存入经验回放缓冲区中。

3)当经验回放缓冲区中的数据量足够时,从缓冲区中随机采样一个包含 N 个经验元组的小批量数据,用于网络更新。

4)基于经验元组,根据式(26)计算目标Q值 y ,并通过最小化时序差分误差损失函数更新两个CASA-Critic当前价值网络的参数。

5)每隔设定的延迟更新步数,基于当前采样的数据,通过确定性策略梯度上升方向更新Actor当前策略网络参数,提升期望累积回报。

6)采用软更新目标价值网络和目标策略网络参数。

7)判断算法是否达到最大训练回合数,若满足条件则结束训练并输出微电网最优运行策略,否则重复步骤2)~6)继续训练,直至输出最优运行策略。

3 算例分析

3.1 算例设置

为验证基于改进TD3的农村微电网源网荷储优化运行策略的优越性,基于改进IEEE-37节点微电网系统进行仿真实验,该系统由光伏(photovoltaic, PV)、风机(wind power, WP)、燃气轮机(micro-gas turbine, MT)、沼气发电机(biogas generator,

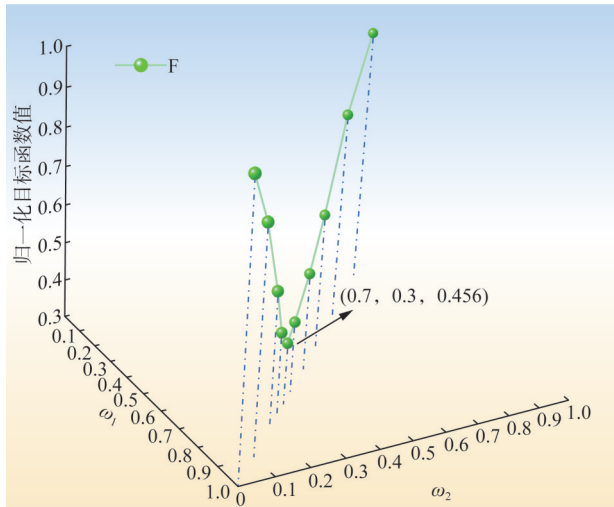


图4 参数灵敏度分析

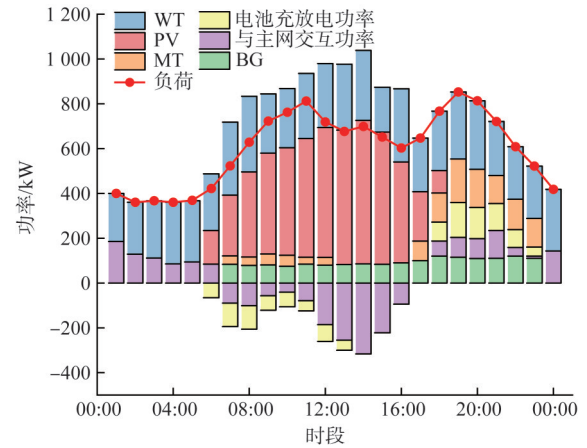
Fig. 4 Parameter sensitivities analysis

再生能源利用率权重系数,主要调节环保性目标;两者共同决定归一化目标函数值 F 的大小,决定系统的综合优化性能。分析图4曲线走势可以看出,当 $\omega_1 < 0.6$ 或 $\omega_1 < 0.8$ 时, F 值逐渐增大表明系统优化性能逐渐下降。而 ω_1 在 $[0.65, 0.75]$ 区间内, F 值变化较为平缓,表明在此范围内选取权重对最终优化目标的扰动较小。当 $\omega_1 < 0.7$, $\omega_2 < 0.3$ 时,归一化目标函数值 F 达到最小值为0.456,表明该权重下系统的综合性能最优,能够在经济性与环保性之间取得最佳平衡。因此本文选取 $\omega_1 < 0.7$, $\omega_2 < 0.3$ 作为目标函数值的权重系数。

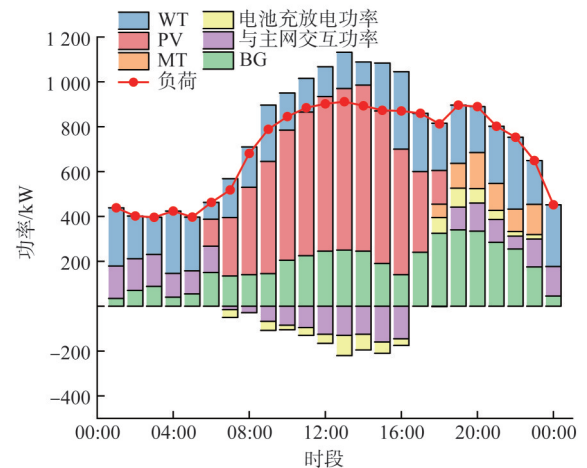
3.3 运行结果分析

为验证本文所提策略的有效性,本节保留训练好的Actor和Critic网络参数,并将其应用于图3所示的农村微电网实际调度决策中。将一天划分为24个时段进行仿真,利用CREST生成的数据作为测试集对所提算法在农村微电网环境下进行调度测试,运行结果如图5所示。

图5(a)展示了沼气淡季时(冬季)农村微电网的出力情况。在00:00—05:00时段,负荷需求量较低,电网电价处于低谷时段,风机出力不足,微电网优先购电满足剩余负荷需求;该时段沼气发电机不出力,一方面是因为电网分时电价较低,从主网购电的经济性优于沼气发电机,另一方面则是因为暂停沼气发电机运行有利于储存沼气资源,为后续电价较高或负荷高峰时段预留调节能力。在08:00—16:00时段负荷需求量较高,风电和光伏机组也



(a) 沼气淡季农村微电网出力情况



(b) 沼气旺季农村微电网出力情况

图5 农村微电网调度结果

Fig. 5 Dispatching results of the rural microgrid

处于高出力水平,可通过优先调度风机、光伏机组出力及沼气出力满足大部分负荷需求,多余电力将售卖给主网或对储能电池进行充电,从而减少燃气轮机的发电量。由于该阶段风光出力相对较高,沼气机组在淡季并不承担持续性主供电任务,而是主要发挥补充供能和灵活调节作用,以提高有限沼气资源的利用效率。在18:00—23:00时段出现第二个负荷高峰且电网电价处于峰时阶段,微网减少与主网的交易,通过储能放电、沼气发电机和燃气轮机机组出力满足微网内部电功率平衡。该时段内储能系统放电和沼气出力减少了对燃气轮机的需求;相较于白天时段,沼气机组在该阶段出力明显增加,说明在淡季资源受限条件下,沼气更多被配置到风光不足、负荷较高的关键时段,用于及时补偿功率缺口。储能系统则释放白天储存的电能,与沼

气机组共同出力减少了对燃气轮机和主网购电的需求。这种调度策略有助于优化农村微电网内的资源配置,提高系统的灵活性和经济性,同时可以增加可再生能源的利用率。

图 5(b)展示了沼气旺季时(夏季)农村微电网的出力情况。不同于沼气淡季,农忙时节负荷需求增加,而且太阳辐射更强、光伏出力水平更高,沼气资源供给也相对更充足。在 00:00—05:00 时段,微电网负荷较低,风机虽有一定出力但整体水平有限,因此系统主要通过主网购电满足剩余负荷需求。由于该时段电价较低,优先购电具有更好的经济性,同时保留小部分沼气出力,减少沼气发电机的频繁启停,并为后续高负荷时段平滑提升出力提供条件。在 06:00—17:00 时段,随着光伏出力快速上升并在中午前后达到峰值,此时,光伏成为该阶段的主要供电来源,而风电和沼气机组则承担补充供电与功率调节功能。该时段系统优先消纳风电和光伏,在满足本地负荷需求的基础上,多余电能用于储能充电或向主网售电,体现出旺季条件下农村微电网较强的本地平衡能力。在 18:00—23:00 时段,光伏出力为 0,系统进入晚高峰阶段。此时通过储能放电、沼气发电机持续出力以及微燃机协同运行实现功率平衡,从而有效降低了对主网购电和微燃机出力的依赖。

总体来看,沼气旺季下农村微电网更加突出沼气电源的稳定支撑作用。与沼气淡季相比,旺季场景中沼气机组在白天和傍晚阶段承担了更多出力任务,并与光伏、风电及储能形成了良好的互补关系,有助于提高系统运行的经济性、灵活性。

3.4 算法性能分析

3.4.1 不同算法性能对比

为了验证本文所提策略的有效性和优越性,本文分别设置了 3 种调度策略。策略 A: 基于改进 TD3 的农村微电网优化运行策略。策略 B: 文献[12]所提基于 DDPG 的农村微电网优化运行策略。策略 C: 文献[17]所提基于 TD3 的农村微电网优化运行策略。

将所构建的网络用于农村微电网优化运行之前,需要基于历史运行数据对算法网络进行集中式训练。图 6—8 分别展示了基于 3 种不同深度强化学习的方法在训练过程中获得的奖励、算法稳定性及综合性能对比。

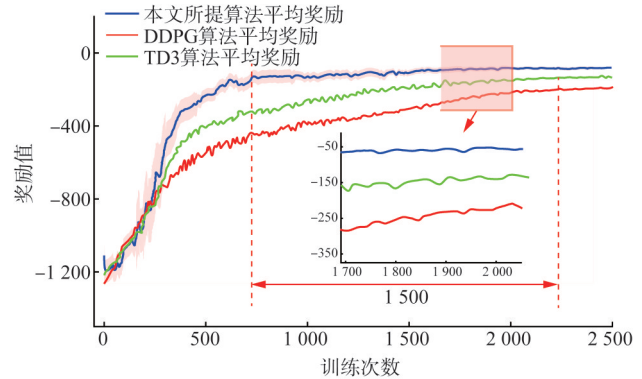


图 6 基于 3 种不同深度强化学习的方法在训练过程中获得的奖励

Fig. 6 Rewards obtained during training for three different deep reinforcement learning approaches

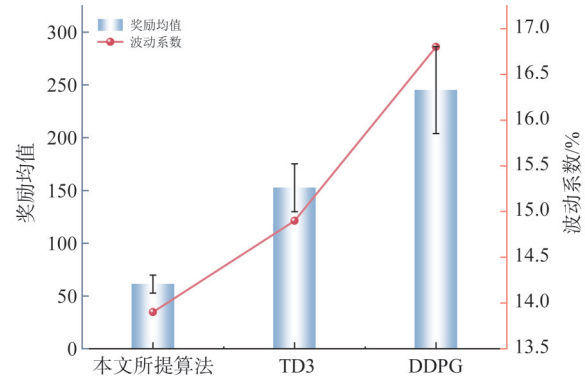


图 7 算法稳定性对比

Fig. 7 Comparison of algorithm stabilities

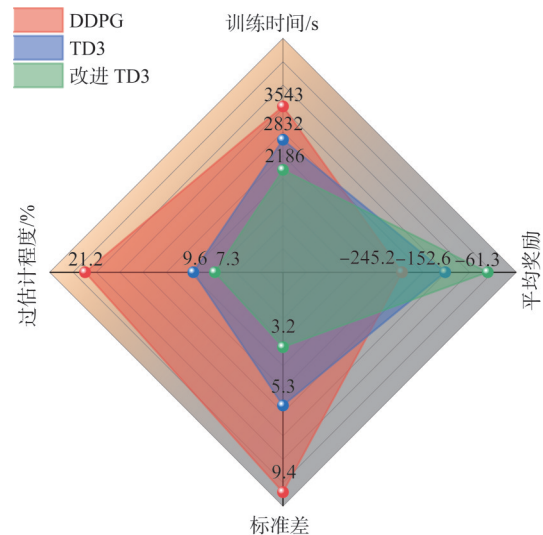


图 8 算法综合性能对比

Fig. 8 Comparison of algorithm comprehensive performance

由图 6 可知,从训练初期开始,策略 A 所提算法由于数据的不确定性使得奖励值出现较大波动,

这是因为智能体在初始阶段尚未熟悉环境。而随着训练的不断进行,智能体通过经验池逐步学习,奖励值不断增大,在训练轮数约1 700次后,逐渐趋于稳定,最终奖励值接近-60,说明改进算法在训练中实现了良好的学习效果。策略B所提算法在整体上累积奖励值增长最慢,在训练轮数达到约2 100次后逐步趋于平稳,其奖励值最低,说明该学习策略并不理想。策略C所提算法在表现出次优的效果,在训练轮数达到约3 000次后逐步趋于平稳,尽管其奖励值稳定在较高水平,但仍略低于策略A。

为了量化评估不同强化学习算法在训练过程中的稳定性,本文引入奖励均值波动系数作为评价指标。该系数定义为训练后期最后500回合奖励值的标准差与均值之比,以反映算法收敛后策略输出的波动程度,其结果如图7所示。由图7可知,本文所提改进TD3算法的波动系数为13.8%,分别低于传统TD3算法的14.5%与DDPG算法的16.7%。该结果表明,改进TD3在训练后期具有更稳定的收敛特性与更低的输出波动,这主要得益于CASA-Critic网络对关键状态特征的聚焦评估降低了价值估计方差,以及动态感知策略减少了后期不必要的随机扰动。因此,改进TD3算法在实际农村微电网随机运行环境中表现出更强的稳定性。

图8展示了算法性能对比的各项指标,包括训练时间、过估计程度、标准差和平均奖励。其中过估计程度是深度强化学习中的一个关键评估指标,用以衡量预测Q值相对于实际Q值的偏差程度,其值越低,表明评估越准确。标准差反映了算法训练过程中的稳定性。较小的标准差表示算法训练过程稳定。

从图8可以看出,改进后的TD3算法在多个关键指标优于其他算法;首先,在训练时间方面,改进后的算法为2 186 s,体现了其较高的计算效率,原因在于引入复合评分和注意力加权机制使Critic网络能够更精准地关注状态-动作对中的关键信息,减少了在无关特征上的计算。其次,在过估计程度方面,DDPG算法目标Q值的计算本质上是一个最大化操作。当存在误差的情况下,这种操作会系统性地放大被高估的Q值,导致偏差从目标值传回当前值,并不断累积。TD3算法存在两个独立的Critic网络。在计算目标Q值时,选取这两个网络

输出的最小值,减小了过估计程度;而本文所提算法通过注意力权重,有区分地聚焦于与Q值估计最相关的关键信息,进一步减少了过估计程度。关于标准差方面,改进TD3算法的值最小,这主要归因于动态感知策略。该策略结合了时间衰减和状态感知项,在训练初期进行充分探索,在后期及状态不确定性低时减少了不必要的随机扰动。最后,改进TD3在平均奖励方面达到了-61.3,远超TD3的-152.6和DDPG的-245.2,这得益于其准确的价值估计和更高效的策略搜索。综上所述,改进TD3在计算效率、评价准确性、训练稳定性和策略优化方面表现更出色,具备更强的实用潜力。

3.4.2 算法灵敏度分析

为验证算法参数设置的合理性,本文针对动态探索噪声上下限开展灵敏度分析。考虑到农村微电网优化运行具有明显的时序耦合特征,不同时段储能荷电状态、沼气可用量及购售电行为会对后续调度结果产生持续影响,因此探索噪声范围将直接影响智能体的长期决策能力与训练收敛质量。基于此,分别对动态探索噪声范围下的训练表现进行对比,结果如图9所示。

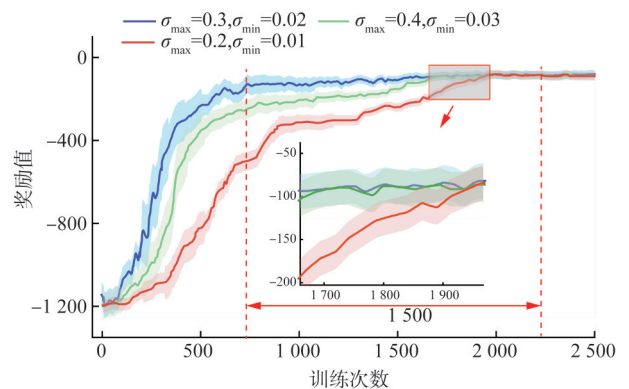


图9 不同动态探索噪声范围下智能体平均奖励值变化

Fig. 9 Changes in the average reward values of the agent under different dynamic exploration noise ranges

由图9可知,不同动态探索噪声上下限组合对算法训练过程具有较明显影响。虽然3组参数下的奖励值在训练后期均逐步趋于稳定,最终收敛水平整体较为接近,但不同参数组合在训练前期的奖励提升速度以及进入稳定区间的快慢方面仍存在较明显差异。当 $(\sigma_{\max}, \sigma_{\min}) = (0.3, 0.02)$ 时,算法在前期能够较快摆脱低奖励状态,奖励值提升速度相对更快,且中后期波动较小,整体收敛过程更为平

稳；当 $(\sigma_{\max}, \sigma_{\min}) = (0.4, 0.03)$ 时，算法同样能够保持较平稳的收敛过程，但其奖励提升速度略慢于前者，进入稳定区间的时间稍晚；当 $(\sigma_{\max}, \sigma_{\min}) = (0.2, 0.01)$ 时，由于噪声幅值偏小，算法前期探索能力相对不足，奖励提升速度较慢，进入稳定收敛区间所需训练轮次相对更多。综合来看， $(\sigma_{\max}, \sigma_{\min}) = (0.3, 0.02)$ 在保证训练稳定性的同时具有更快的收敛速度，因此本文选取该组参数作为动态探索噪声的最终设置。

3.5 考虑有无沼气运行的效益分析

为了准确评估沼气发电在农村微电网中的综合价值，本节采用改进TD3的农村微电网源网荷储优化运行策略对有无沼气两种情景进行对比分析。图10—11分别展示了有、无沼气方案各类能源占比，表4综合对比了无沼气方案各类成本。

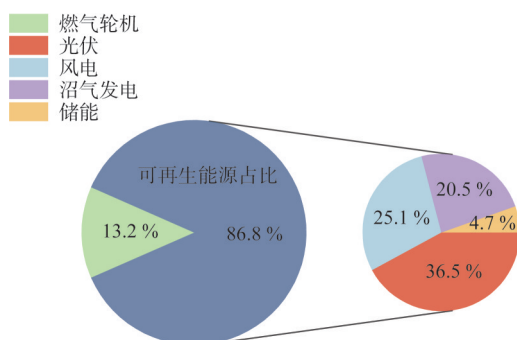


图10 有沼气方案各类能源占比

Fig. 10 Proportion of various energy sources of the scheme with biogas

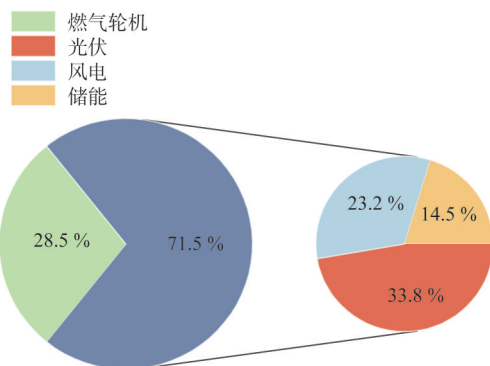


图11 无沼气方案各类能源占比

Fig. 11 Proportion of various energy sources of the scheme without biogas

由表4可知，燃料成本变化从6 140.67元降低至4 850.50元，这主要是因为沼气作为廉价的可再生能源替代了部分价格较高的天然气。其次，设备

启停成本从822.35元降低至246.89元，这表明沼气发电作为稳定可控的分布式电源，有效平抑了系统的功率波动，减少了微型燃气轮机的频繁启停，降低了操作成本。同时，当沼气发电机组投入运行时，减少了向主网购电的需求，这导致了运维成本增加和与主网交互的成本减少。尽管有沼气方案的运维成本和与主网交互成本较高，但有沼气方案的总运行成本为14 526.82元，相较于无沼气方案降低了13.84%。由图10—11可知，有沼气方案的可再生能源占比高达86.8%，而无沼气方案降则占71.5%，两者相差15.3%。这一差距直接体现了沼气发电作为可再生能源的直接贡献。在有沼气方案中，风光资源得以充分消纳，因为沼气与储能共同为系统提供了必要的灵活调节能力。在无沼气方案中，尽管风电与光伏的占比之和与有沼气方案相近，但储能放电占比大幅提升至14.5%。这表明系统为了维持稳定，不得不更加依赖储能系统进行深度调节，这不仅加大了储能的运行压力，也可能导致在极端情况下产生“弃风弃光”现象。综上所述，有沼气方案在农村微电网的实际运行中更具潜力。

表4 有无沼气方案成本对比

Tab. 4 Cost comparison of schemes with and without biogas

项目	元	
	有沼气方案	无沼气方案
燃料成本	4 839.50	6 140.67
与主网交互成本	5 620.18	6 883.42
运维成本	3 811.25	3 013.19
启停成本	246.89	822.35
总运行成本	14 526.82	16 859.67

3.6 不同方案的运行结果分析

为进一步验证本文所提优化运行策略的优越性，将本文所提基于改进TD3的农村微电网源网荷储优化运行策略(方案A)、文献[17]所提基于深度强化学习的微电网运行优化方法(方案B)、文献[12]所提考虑源荷不确定性下微电网能量调度的深度强化学习策略(方案C)和文献[7]所提基于改进白鲸优化算法的微电网优化调度策略(方案D)进行对比。对比结果如图12所示。

根据图12的结果分析，本文提出的方案A在运行经济性与可再生能源利用效率方面均表现出显

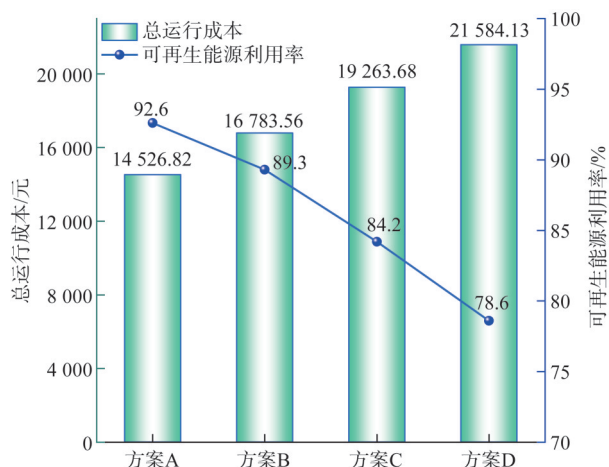


图 12 各方案的对比结果

Fig. 12 Comparison results of different schemes

著优势。具体而言, 方案 A 的系统总运行成本为 14 526.82 元, 相较于方案 B、C 和 D, 分别降低了 13.45%、24.59% 和 32.70%。在可再生能源利用率方面, 4 种方案的结果分别为 92.6%、89.3%、84.2% 和 78.6%, 本文所提方案 A 分别较其他方案提升了 3.3%、8.4% 和 14.0%。上述性能差异主要源于各算法在机制方面区别。方案 D 的改进白鲸优化算法属于传统启发式优化方法, 在处理高维、动态变化的微电网运行状态时, 存在计算复杂度高、收敛速度慢的局限, 容易陷入局部最优解。方案 C 采用的 DDPG 算法, 虽具备处理非线性与不确定性问题的能力, 但其 Critic 网络在 Q 值估计中存在明显的过估计偏差, 导致策略评估失准, 运行成本未能达到最优。方案 B 采用的 TD3 算法通过双 Critic 网络结构与目标策略平滑机制, 有效抑制了 Q 值过估计, 但仍存在低估偏差, 影响策略更新效率, 限制模型收敛至最优性能。方案 A 在方案 B 的基础上进一步引入基于复合评分机制的注意力加权 Critic 网络与动态感知策略。复合评分和注意力加权机制增强了网络对系统关键状态特征的感知与聚焦能力, 从而提升了 Q 值估计的准确性; 动态感知策略则通过动态调整探索强度, 促进智能体在训练过程中更高效地发掘高价值策略。综上所述, 本文所提出的基于改进 TD3 的源网荷储优化策略在运行经济性、可再生能源消纳能力方面均具明显的优越性。

4 结论

针对农村微电网的优化运行问题, 本文构建了包含沼气的农村微电网源网荷储优化运行模型, 提出了一种基于改进 TD3 的农村微电网优化运行策略。通过算例分析验证了所提策略的有效性和优越性, 并得出以下结论。

1) 沼气发电作为一种本地化、可调度的可再生能源对系统优化运行效果显著。旺季时, 沼气主要发挥稳定支撑作用; 淡季时, 则主要承担关键时段补偿与灵活调节功能, 从而减少了系统对燃气轮机的依赖。

2) 相较于 DDPG 与 TD3 算法, 本文所提出的改进 TD3 算法在训练效率、评估准确性、收敛稳定性等方面均表现出显著优势。

3) 基于改进 TD3 的运行策略可以实现更优的农村微电网源网荷储优化调度。相较于其他 3 种方案, 本文所提策略在系统总运行成本与可再生能源利用率两方面均表现最优。

4) 相较于无沼气方案, 有沼气方案将系统的可再生能源占比提升至 86.8%, 使总运行成本降低 13.77%。

本文方法在实际应用中仍存在一定局限性。由于农村微电网初期建设成本较高, 且该方法对系统感知与通信能力有一定要求, 同时其收益主要体现在运行优化和综合效益提升的长期过程中, 因此更适用于具备一定智能化基础和新能源开发条件的农村场景。未来随着农村电力基础设施持续完善及智能化运行体系不断发展, 本文所提方法在典型农村场景中具备良好的经济应用前景和推广潜力。

参考文献

- [1] 凌凯, 王灿, 张高瑞, 等. 考虑路径损耗的热电联供型微网三层能量优化策略[J]. 中国电力, 2023, 56(2): 102–113.
LING Kai, WANG Can, ZHANG Gaorui, et al. Three-layer energy optimization strategy for CHP microgrid[J]. Electric Power, 2023, 56(2): 102–113.
- [2] 张琳, 谢洪途, 赵路路, 等. 混合能源直流微电网能源优化管控策略研究[J]. 电力系统保护与控制, 2024, 52(3): 141–151.
ZHANG Lin, XIE Hongtu, ZHAO Lulu, et al. Energy optimization and control strategy for a hybrid energy DC microgrid[J]. Power System Protection and Control, 2024, 52(3): 141–151.
- [3] 王灿, 张雪菲, 凌凯, 等. 基于区间概率不确定集的微电网两阶段自适应鲁棒优化调度[J]. 中国电机工程学报, 2024, 44

- (5): 1750 – 1764.
WANG Can, ZHANG Xuefei, LING Kai, et al. Two-stage adaptive robust optimal scheduling based on the interval probability uncertainty set for microgrids [J]. Proceedings of the CSEE, 2024, 44(5): 1750 – 1764.
- [4] 张金良, 程佳. 含沼气能发电的农村多微网系统容量配置双层优化模型[J]. 电力建设, 2025, 46(10): 73 – 87.
ZHANG Jinliang, CHENG Jia. Dual layer optimization model for capacity configuration of rural multi-microgrid system with biogas energy generation [J]. Electric Power Construction, 2025, 46(10): 73 – 87.
- [5] 张程, 匡宇, 刘佳静, 等. 考虑需求侧管理的风光燃储微网两阶段优化调度[J]. 电力系统保护与控制, 2022, 50(24): 13 – 22.
ZHANG Cheng, KUANG Yu, LIU Jiajing, et al. Two-stage optimal scheduling of a wind, photovoltaic, gas turbine, fuel cell and storage energy microgrid considering demand-side management [J]. Power System Protection and Control, 2022, 50(24): 13 – 22.
- [6] 臧云帆, 夏晟, 李嘉文, 等. 含共享储能的微电网群分布鲁棒博弈优化调度方法[J]. 电力系统保护与控制, 2023, 51(24): 90 – 101.
ZANG Yunfan, XIA Sheng, LI Jiawen, et al. A robust game optimization scheduling method for shared energy storage micro electric network group distribution [J]. Power System Protection and Control, 2023, 51(24): 90 – 101.
- [7] 杨程峰. 基于改进白鲸优化算法的微电网优化调度研究[D]. 吉林: 东北电力大学, 2025.
- [8] 岳佳明, 钟建伟, 朱占春, 等. 数据中心参与风光抽蓄互补微电网优化调度[J]. 电工电气, 2025(5): 13 – 20, 71.
YUE Jiaming, ZHONG Jianwei, ZHU Zhanchun, et al. Data center participates in optimal scheduling of a wind-solar-pumped-storage hybrid microgrid [J]. Electrotechnics Electric, 2025(5): 13 – 20, 71.
- [9] 姚明明, 张新, 杨培宏, 等. 基于改进风光场景聚类联合虚拟储能的源网荷储低碳优化调度[J]. 电力系统保护与控制, 2024, 52(15): 115 – 130.
YAO Mingming, ZHANG Xin, YANG Peihong, et al. Low-carbon optimal scheduling of source-grid-load-storage based on improved wind-solar scene clustering combined with virtual energy storage [J]. Power System Protection and Control, 2024, 52(15): 115 – 130.
- [10] GRONAUER S, DIEPOLD K. Multi-agent deep reinforcement learning: a survey [J]. Artificial Intelligence Review, 2022, 55(2): 895 – 943.
- [11] 张帆, 武东昊, 陈玉萍, 等. 多智能体深度强化学习的分布式园区综合能源系统经济调度策略[J]. 电力系统及其自动化学报, 2022, 34(12): 18 – 26.
ZHANG Fan, WU Donghao, CHEN Yuping, et al. Economic scheduling strategy for integrated energy system in distributed parks based on multi-agent deep reinforcement learning [J]. Proceedings of the CSU-EPSC, 2022, 34(12): 18 – 26.
- [12] 马冲冲, 王一铮, 王坤, 等. 考虑源荷不确定性下微电网能量调度的深度强化学习策略[J]. 高技术通讯, 2023, 33(1): 79 – 87.
MA Chongchong, WANG Yizheng, WANG Kun, et al. Deep reinforcement learning based dispatch strategy for microgrid energy management considering uncertainty of source and load [J]. Chinese High Technology Letters, 2023, 33(1): 79 – 87.
- [13] 林文硕, 李欣, 周荣生, 等. 基于深度强化学习的微电网低碳优化运行方法[J]. 现代电力, 2025, 42(6): 1131 – 1142.
LIN Wenshuo, LI Xin, ZHOU Rongsheng, et al. A low-carbon optimal operation method for microgrids based on deep reinforcement learning [J]. Modern Electric Power, 2025, 42(6): 1131 – 1142.
- [14] 王齐. 基于联邦强化学习的微电网储能调度方法的研究与应用[D]. 哈尔滨: 黑龙江大学, 2025.
- [15] 王琳道, 刘大明. 基于改进PPO算法的微电网交易及调度研究[J]. 计算机仿真, 2025, 42(5): 84 – 91.
WANG Lindao, LIU Daming. Research on microgrid transaction and scheduling based on improved PPO algorithm [J]. Computer Simulation, 2025, 42(5): 84 – 91.
- [16] 曾磊, 丁泉, 陈孝煜, 等. 基于深度强化学习的微电网运行优化方法[J]. 浙江电力, 2025, 44(6): 31 – 40.
ZENG Lei, DING Quan, CHEN Xiaoyu, et al. A DRL-based optimization method for microgrid operation [J]. Zhejiang Electric Power, 2025, 44(6): 31 – 40.
- [17] WANG C, ZHANG J, WANG A, et al. Prioritized sum-tree experience replay TD3 DRL-based online energy management of a residential microgrid [J]. Applied Energy, 2024 (368): 123471.1 – 123471.10.
- [18] 罗永建, 刘承锡, 董旭柱. 基于深度强化学习架构的多能互补微网日前经济调度研究[J]. 武汉大学学报(工学版), 2023, 56(11): 1393 – 1404.
LUO Yongjian, LIU Chengxi, DONG Xuzhu. Day ahead economic dispatching of multi energy complementary micro grid based on deep reinforcement learning architecture [J]. Engineering Journal of Wuhan University, 2023, 56(11): 1393 – 1404.
- [19] 雷强, 武鹏荣, 李振文. 基于改进SAC算法的微电网优化调度[J]. 电气传动, 2026, 56(3): 81 – 87.
LEI Qiang, WU Pengrong, LI Zhenwen. Optimal scheduling of microgrid based on improved SAC algorithm [J]. Electric Drive, 2026, 56(3): 81 – 87.
- [20] 张彤彤. 改进的多目标粒子群算法在分布式光伏电源优化配置中的应用[J]. 电子技术, 2025, 54(7): 290 – 291.
ZHANG Tongtong. Application of Improved multi-objective particle swarm optimization algorithm in distributed photovoltaic power optimization configuration [J]. Electronic Technology, 2025, 54(7): 290 – 291.
- [21] 杨迪, 王辉, 杨睿彤, 等. 基于自适应奖励函数-多智能体深度强化学习的移动储能时空调度策略[JOL]. 中国电机工程学报, 1 – 13 [2025 – 8–22]. <https://doi.org/10.13334/j.0258-8013.pcsee.251329>.
YANG Di, WANG Hui, YANG Ruitong, et al. Spatio-temporal dispatch strategy of mobile energy storage based on multi-agent deep reinforcement learning with adaptive reward function [JOL]. Proceedings of the CSEE, 1 – 13 [2025 – 8–23]. <https://doi.org/10.13334/j.0258-8013.pcsee.251329>.
- [22] 李嘉文, 余涛, 张孝顺, 等. 基于改进深度确定性梯度算法的AGC发电功率指令分配方法[J]. 中国电机工程学报, 2021, 41(21): 7198 – 7212.
LI Jiawen, YU Tao, ZHANG Xiaoshun, et al. AGC power generation command allocation method based on improved deep deterministic policy gradient algorithm [J]. Proceedings of the CSEE, 2021, 41(21): 7198 – 7212.
- [23] 梁涛, 柴露露, 谭建鑫, 等. 基于深度强化学习算法的氢耦合电-热综合能源系统优化调度[J]. 电力自动化设备, 2025, 45

- (1): 59 – 66.
- LIANG Tao, CHAI Lulu, TAN Jianxin, et al. Optimal scheduling of hydrogen coupled electrothermal integrated energy system based on deep reinforcement learning algorithm [J]. Electric Power Automation Equipment, 2025, 45(1): 59 – 66.
- [24] 林成, 吕俊兵, 张裕, 等. 基于 LSTM-GBRT 模型的省级区域碳排放预测方法[J]. 南方电网技术, 2026, 20(6): 168 – 178.
- LIN Cheng, LÜ Junbing, ZHANG Yu, ZHANG Zhong. Provincial-level area carbon emission forecasting method based on LSTM-GBRT model [J]. Southern Power System Technology, 2026, 20(6): 168 – 178.
- [25] 王宏, 郭志涛, 张森, 等. 融合 Kalman 滤波的双注意力 LSTM 剩余使用寿命预测网络[J]. 现代电子技术, 2025, 48(22): 160 – 166.
- WANG Hong, GUO Zhitao, ZHANG Sen, et al. Dual-attention LSTM RUL prediction network with Kalman filter [J]. Modern Electronic Technique, 2025, 48(22): 160 – 166.
- [26] 罗林霖, 王霄, 何志琴, 等. 基于滚动模态分解和 GCN-DaBiLSTM 的综合能源系统多元负荷预测模型[J]. 广东电力, 2025, 38(9): 130 – 138.
- LUO Linlin, WANG Xiao, HE Zhiqin, et al. Multi-load forecasting model for integrated energy system based on rolling mode decomposition and GCN-DaBiLSTM [J]. Guangdong Electric Power, 2025, 38(9): 130 – 138.
- [27] 崔星, 李晋国, 张照贝, 等. 基于改进粒子群算法优化 LSTM 的短期电力负荷预测[J]. 电测与仪表, 2024, 61(1): 131 – 136.
- CUI Xing, LI Jinguo, ZHANG Zhaobei, et al. The short-term power load forecasting based on NIWPSO-LSTM neural network [J]. Electrical Measurement & Instrumentation, 2024, 61(1): 131 – 136.
- [28] 崔志坤, 王清亮, 刘景青, 等. 基于模糊集理论和序关系分析法的电力市场指标体系研究[J]. 电测与仪表, 2024, 61(5): 31 – 37.
- CUI Zhikun, WANG Qingliang, LIU Jingqing, et al. Evaluation of the electricity markets index system based on fuzzy set theory and order relationship analysis [J]. Electrical Measurement & Instrumentation, 2024, 61(5): 31 – 37.
- [29] 王斯琪, 曹昉, 姚力. 计及多级市场代理购电成本传导风险的分时电价定价模型[J]. 中国电力, 2025, 58(11): 25 – 37.
- WANG Siqi, CAO Fang, YAO Li. Time-of-use pricing model considering the risk of cost transmission in multi-level market agent electricity purchase [J]. Electric Power, 2025, 58(11): 25 – 37.

收稿日期: 2026-01-26; 网络首发日期: 2026-06-24

作者简介:

陈国齐(1992), 男, 工程师, 硕士, 研究方向为微电网优化调度, 516503824@qq.com;

魏颖豪(1984), 男, 高级工程师, 硕士, 研究方向为微电网能量管理, yinghao2586@163.com;

王灿(1987), 通信作者, 男, 副教授, 博士, 研究方向为微电网优化运行与控制, xfcancan@163.com。

(上接第 66 页 Continued from Page 66)

- ZHANG Qunli, ZHANG Jinliang, LIU Yimo, et al. Analysis of flexible load regulation capability in electric heating system with phase-change for energy storage using photovoltaic and grid power [J]. Renewable Energy Resources, 2024, 42(9): 1179 – 1188.
- [29] BAI C, LI Q, ZHENG X, et al. Dynamic weighted-gradient descent method with smoothing momentum for distributed energy management of multi-microgrids systems [J]. IEEE Transactions on Smart Grid, 2023, 14(6): 4152 – 4168.
- [30] HAN J, FANG Y, LI Y, et al. Optimal planning of multi-microgrid system with shared energy storage based on capacity leasing and energy sharing [J]. IEEE Transactions on Smart Grid, 2025, 16(1): 16 – 31.
- [31] VERA E G, CAÑIZARES C A, PIRNIA M, et al. Two-stage stochastic optimization model for multi-microgrid planning [J]. IEEE Transactions on Smart Grid, 2023, 14(3): 1723 – 1735.
- [32] 林顺富, 邱蕾, 周波, 等. 面向配电系统灵活性提升的多微电网分布鲁棒协同优化调度[J]. 电网技术, 2026, 50(6): 2591 – 2604.
- LIN Shunfu, QIU Lei, ZHOU Bo, et al. Distributed robust coordinated optimal dispatching of multi-microgrids for distribution system flexibility enhancement [J]. Power System Technology, 2026, 50(6): 2591 – 2604.
- [33] 钱帅帅, 苏珈, 杜欣慧, 等. 计及源-荷不确定性的综合能源系统分布鲁棒优化[J/OL]. 南方电网技术, 1 – 13 [2025 – 12 – 03]. <https://link.cnki.net/urlid/44.1643.TK.20251202.1752.024>.
- QIAN Shuaishuai, SU Jia, DU Xinhui, et al. Robust optimization of integrated energy system distribution taking into account source-load uncertainty [J]. Southern Power System Technology, 1 – 13 [2025 – 12 – 03]. <https://link.cnki.net/urlid/44.1643.TK.20251202.1752.024>.

收稿日期: 2025-06-26; 网络首发日期: 2026-01-05

作者简介:

于鹤洋(1987), 男, 工程师, 硕士, 研究方向为用电管理, yuheyang@csg.cn;

金鑫(1985), 男, 高级工程师, 硕士, 研究方向为用电通信, jinxin1@csg.cn;

曹望璋(1993), 男, 工程师, 硕士, 研究方向为电力需求侧管理, caowz@csg.cn。