

考虑频率-惯量协同约束的沙戈荒地区 光热-光伏联合电站智能调度

陈实, 晏红平, 臧天磊, 刘艺洪, 陈江平, 王舒灏, 李华强
(四川大学电气工程学院, 成都市 610065)

摘 要:【目的】针对我国“沙戈荒”大型新能源基地并网面临的高渗透率与低惯量特征,以及传统调度方法主要围绕稳态功率平衡和经济运行,对频率安全隐患关注不足的局限,本文提出一种计及广义惯量与频率变化率硬约束的光热-光伏联合电站优化调度方法。【方法】首先,构建包含光热同步机物理特性的频率安全边界模型,将优化调度建模为马尔可夫决策过程。其次,针对传统双延迟深度确定性策略梯度算法(twin delayed deep deterministic policy gradient, TD3)在长周期储能状态感知及稀疏频率越限样本学习上的不足,提出一种融合长短期记忆网络(long short-term memory, LSTM)与优先经验回放(prioritized experience replay, PER)的改进TD3(LSTM-PER-TD3)算法。【结果】基于改进IEEE-30及IEEE-57节点系统的多场景仿真表明:所提LSTM-PER-TD3算法在完整频率-惯量约束场景下实现了频率越限与备用短缺双零,且总运行成本与混合整数线性规划(mixed integer linear programming, MILP)方法得到的理论最优解仅相差2.94%。【结论】光热电站同时具备储能时移调节能力与同步机组物理惯量支撑双重优势,能够为高比例新能源基地筑牢频率安全保障;所构建的LSTM-PER-TD3优化算法,在兼顾运行安全性与经济性的前提下实现了机组高效调度,可为高比例新能源电力系统并网运行下的主动频率支撑与智能化经济调度提供理论依据与技术支撑。

关键词:沙戈荒地区;光热-光伏联合电站;频率稳定;频率-惯量约束;深度强化学习

中图分类号: TM734

文献标志码: A

文章编号: 1000-7229(2026)08-0101-18

DOI: 10. 12204/j. issn. 1000-7229. 2026. 08. 008

Optimized Dispatch for CSP-PV Hybrid Plants in Sandy-Gobi-Desert Regions Considering Frequency-Inertia Coordinated Constraints

CHEN Shi, YAN Hongping, ZANG Tianlei, LIU Yihong, CHEN Jiangping,
WANG Shuhao, LI Huaqiang

(College of Electrical Engineering, Sichuan University, Chengdu 610065, China)

ABSTRACT: [Objective] To address the high penetration and low-inertia characteristics arising from the grid integration of large-scale renewable energy bases in “sandy-gobi-desert” regions, as well as the limitations of traditional dispatch methods that primarily focus on steady-state power balance and economic operation with insufficient attention to frequency security risks, this paper proposes an optimal dispatch method for concentrated solar power-photovoltaic hybrid power plants considering generalized inertia and rate of change of frequency hard constraints. [Methods] First, the optimal dispatch of concentrated solar power-photovoltaic hybrid power plants is modeled as a markov decision process. This model incorporates dynamic security constraints, such as the rate of change of frequency and generalized inertia, to achieve strict control over grid frequency security boundaries. Second, to efficiently solve this highly nonlinear scheduling model and overcome the limitations of the conventional twin delayed deep deterministic policy gradient (TD3) algorithm—specifically addressing the low utilization efficiency of historical information and the insufficient learning of critical disturbance samples—an improved TD3 algorithm integrating long short-term memory (LSTM) networks and a prioritized experience replay (PER) mechanism is proposed. [Results] Multi-scenario simulation results based on modified IEEE-30 and IEEE-57 bus systems demonstrate that, under the complete frequency-inertia constrained scenario, the proposed LSTM-PER-TD3 algorithm achieves zero frequency-limit violations and zero reserve shortage, with a total operating cost only 2.94% above the mixed integer linear programming (MILP) theoretical optimum. [Conclusions] Concentrating solar power plants possess dual advantages of

thermal energy storage for time-shifting regulation and physical inertia support from synchronous units, which can firmly guarantee frequency security for high-penetration new energy bases. The developed LSTM-PER-TD3 optimization algorithm realizes efficient unit scheduling while balancing operational safety and economy. It provides theoretical foundations and technical support for active frequency support and intelligent economic scheduling of power systems with high-penetration new energy integration.

This work is supported by National Natural Science Foundation of China (No. 52577129, No. 52377115) and National Key R&D Program of China (No. 2025ZD0807300).

KEYWORDS: sandy-gobi-desert regions; concentrated solar power-photovoltaic hybrid power plant; frequency stability; frequency-inertia constraints; deep reinforcement learning

0 引言

全球能源转型与“双碳”目标的深入推进,正驱动新能源发电在电力系统中的占比持续攀升,成为保障能源安全与应对气候变化的关键路径^[1]。中国风电与光伏装机容量已分别突破 5 亿 kW 和 10 亿 kW,规模稳居全球首位^[2]。受地理资源禀赋影响,大型风电、光伏(photovoltaic, PV)基地主要集中于西北“沙戈荒”地区,该区域太阳能与风能资源富集,尽管已建成多条特高压直流外送通道,但送端电网“强直弱交”的结构性矛盾依然突出^[3]。随着本地常规火电逐步被替代,高比例电力电子设备并网导致送端系统呈现典型的“低短路比、低转动惯量”特征。在此类弱电网环境下,系统抗扰动能力显著削弱,传统输电模式面临严峻挑战,弃风弃光现象仍较突出^[4-5]。

为提升新能源消纳能力,现有实现路径主要聚焦两方面:一是优化电网运营管理策略^[6];二是部署储能技术,如电化学储能、抽水蓄能及熔盐储热等^[6-7]。此外,富余新能源电力用于电解水制氢、清洁供暖等多元化利用方式也被视为有效补充^[8-9]。然而,常规电化学储能多呈“跟网型”零惯量特征,仅能实现稳态能量时移,难以在暂态扰动期间提供系统急需的频率支撑。相比之下,光热发电(concentrated solar power, CSP)依托“集热-储热-汽轮机”的物理架构,在兼具大容量储能属性的同时,其同步发电机组可为电网提供真实的旋转惯量与电压支撑^[10]。面对暂态功率突变,CSP能够提供无控制时延的机电耦合刚性支撑,从物理机理上有效克服了风光逆变器模拟“虚拟惯量”时存在的响应滞后局限。此外,已有研究表明光热电站并网对改善系统暂态功角稳定性具有积极作用^[11]。

在 CSP 技术基础方面,已有研究围绕太阳直接辐射资源预测^[12]、熔融盐及氯化物储热介质与工程集成^[13-14],以及光热发电技术综述与发展展望^[15-16]等持续推进,为其规模化应用奠定基础。在 CSP 运行优化方面,文献[17]提出一种含光热电站的电-热综合能

源系统优化运行机制。文献[18]建立了光热电站深度调峰成本模型并提出补偿机制。文献[19]系统性分析了 CSP 在高比例可再生能源电力系统中的价值。在联合运行层面,文献[20]探讨了 CSP/PV/风电联合系统的最优容量配置与运行策略。文献[21]采用多目标优化方法研究 PV-CSP-风电储能系统的储热选型。文献[22]将调度问题拓展至双时间尺度,提出 CSP-PV 联合电站的动态协同优化调度策略。文献[23]则引入基于最大均值差异的分布鲁棒优化方法,考虑光热电站的灵活备用贡献。在多能互补与灵活运行层面,相关研究进一步拓展了光热电站的应用价值,涵盖弃电热储能的光伏-光热复合发电^[24]、计及需求响应的光热热电联供型微网分层调度^[25]、风-光-光热联合系统储热容量优化配置^[26]、含光热电站的需求响应虚拟电厂调度^[27],以及计及光热与氢储能的综合能源系统低碳优化运行^[28]。

然而,现有研究多集中于联合电站的经济性优化与常规备用调度,鲜有从频率安全与系统惯量支撑的视角构建调度模型。尤其在“沙戈荒”弱电网场景下,高比例新能源接入导致系统惯量水平显著降低,功率扰动事件后频率变化率(rate of change of frequency, RoCoF)与频率偏差极易超出安全限值^[29]。因此,面向大规模新能源基地安全运行需求,亟需发展融合频率-惯量约束的联合调度方法,并通过场景化仿真量化评估 CSP 在提供广义惯量与快速调频支撑方面的积极作用。

另一方面,高比例新能源优化调度问题因其高维度、强耦合及多重不确定性特征,对求解方法提出了更高要求。传统最优化算法在处理复杂动态约束时计算效率不足^[18-23],而启发式算法则易陷入局部最优且泛化能力有限^[30]。深度强化学习(deep reinforcement learning, DRL)凭借其对复杂时序决策问题的建模优势与自适应性,在电力系统调度领域展现出显著优势^[31-32]。其中,适用于连续动作空间的孪生延迟深度确定性策略梯度算法(twin delayed deep deterministic policy gradient, TD3)表现优异。文

献[33]把因果 Transformer-based 世界模型与 TD3 有机结合形成 CT-TD3 框架。文献[34]提出一种改进的 TD3 算法并引入动态平衡探索噪声机制。文献[35]在能量管理策略中改进 TD3 的经验回放机制,提出基于 Prioritized Sum-Tree 的经验回放策略。文献[36]在配电网电压调节问题中引入线性安全层与 TD3 结合。尽管这些研究在 TD3 算法改进上取得了进展,但改进多集中在经验回放策略、安全层投影或超参数调优上,对时序特征提取能力及其对策略决策的影响关注较少,且对稀有高风险扰动样本的主动采样策略或历史长期信息的结构化利用仍有待提升。

基于上述背景与挑战,本文针对“沙戈荒”地区新能源基地安全经济运行需求,提出一种考虑频率-惯量约束的光热-光伏联合电站优化调度方法。主要贡献如下:

1) 构建面向弱电网频率安全的广义惯量协同调度模型。针对新能源基地高比例电力电子设备带来的低惯量痛点,本文建立了计及光热电站同步机物理特性的广义惯量计算模型。将电加热器(electric heater, EH)的电热耦合作用与频率变化率、频率偏差等关键动态指标转化为调度硬约束。

2) 提出适应强动态约束的改进深度强化学习算法。针对频率安全约束下策略求解困难及关键故障样本稀缺的问题,提出融合长短期记忆网络(long short-term memory, LSTM)与优先经验回放(prioritized experience replay, PER)的改进 TD3 算法。

3) 量化揭示光热电站在新型电力系统中的惯量支撑价值。突破了传统研究多将光热电站视为单一储能调节资源的局限,重点挖掘其作为同步机物理支撑电源的特性。通过多场景仿真,验证了 CSP 储热系统与发电机组协同运行在提升系统广义惯量水平、遏制频率恶化方面的不可替代作用,为解决送端弱电网频率稳定难题提供了理论依据与技术参考。

1 考虑频率-惯量约束的光热-光伏联合电站调度模型

1.1 系统结构与研究对象建模

熔盐塔式光热发电系统因聚光比高、工作温度高及光热转换效率高特性,比传统新能源火电打捆模式,更适宜作为大规模可调度电源。本文研究对象为配置双熔盐储罐和朗肯循环发电单元的塔式 CSP 电厂,其由太阳能场、储热系统和动力循环系统3个核心子系统构成。系统运行模式可分为:启动、储热+发电、纯储热、放热发电、停机和防凝6种状态。

CSP-PV 联合发电系统的拓扑架构如图1所示。

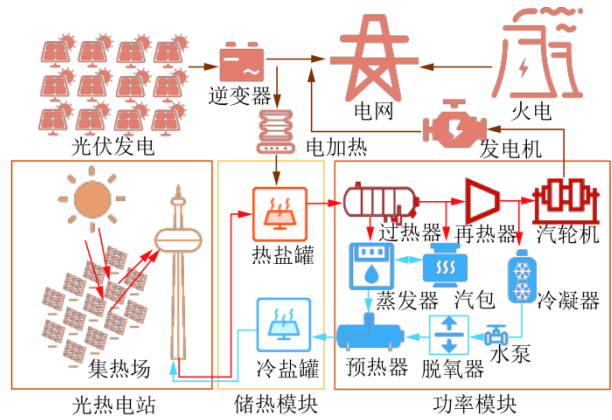


图1 光热-光伏联合电站系统架构

Fig. 1 CSP-PV hybrid power plant system architecture

1) 光热发电系统。

定日镜场作为 CSP 电站的核心集热单元,负责将直接法向辐照度汇聚反射至吸热器表面。反射过程存在4类光学损失:余弦损失、阴影和遮挡损失、大气衰减损失和截断损失。太阳能场热功率输出建模为:

$$Q_{SF}^t = C_{SF} \cdot \eta_{SF} \cdot I_{DN}^t \quad (1)$$

式中: Q_{SF}^t 为太阳能场输出热功率; C_{SF} 为有效集热面积; η_{SF} 为集热系统综合效率; I_{DN}^t 为直接法向辐照度。

储热系统通过熔盐介质实现时空平移与解耦:低温熔盐经泵输送至吸热器加热后存入高温罐;高温熔盐经蒸汽发生器(steam generation system, SGS)换热降温后返回低温罐,完成热-电转换循环。动力循环输入热能表示为:

$$Q_{PB, in}^t = Q_{TES, dis}^t \cdot \eta_{dis} + Q_{PB}^t \quad (2)$$

$$P_{CSP}^t = Q_{PB, in}^t \cdot \eta_{PB} \quad (3)$$

式中: $Q_{PB, in}^t$ 为时刻 t 输入动力循环系统的总热能; $Q_{TES, dis}^t$ 为时刻 t 储热系统释放的热能; η_{dis} 为储热罐(thermal energy storage, TES)放热效率; Q_{PB}^t 为时刻 t 太阳能场输入的热能; P_{CSP}^t 为时刻 t 光热电站输出的电功率; η_{PB} 为动力循环系统的热电转换效率。

2) 光伏发电系统及电加热器。

光伏阵列的输出功率主要取决于光电转换效率、组件运行温度及光伏板表面的入射辐照度。光伏输出功率模型为:

$$P_{PV}^t = C_{PV} \cdot \eta_{PV} \cdot I_{GN}^t \quad (4)$$

式中: P_{PV}^t 为时刻 t 光伏电站输出的电功率; C_{PV} 为光伏电站的光伏板面积; η_{PV} 为光电转换效率; I_{GN}^t 为时刻 t 地表法向辐照度。

EH 模块有助于将多余的 PV 电力转化为热能,随后存储在 TES 子系统中。这一过程将 EH 定位为连

接 CSP 和 PV 系统的能量耦合枢纽。

$$Q_{EH}^t = P_{EH}^t \cdot \eta_{EH} \quad (5)$$

式中: Q_{EH}^t 为时刻 t 电加热器输出到 TES 的热能; P_{EH}^t 为时刻 t 电加热器消耗的弃用光伏电力; η_{EH} 为电加热器电热转换效率。

1.2 系统惯量计算

电力系统惯量表征了系统在遭受功率扰动瞬间,通过释放旋转动能以抑制频率突变的能力,是维持频率稳定的物理基础。在“沙戈荒”高比例新能源电网中,传统同步机组占比下降导致系统惯量水平显著降低。而在高比例新能源接入的“沙戈荒”电网中,为了更全面地评估系统抗扰动能力,本文定义系统广义惯量为包含影响系统频率变化率的所有惯量形式的集合。具体而言,它不仅包含剩余火电机组的固有惯量,更纳入了 CSP 通过其同步发电机组提供的物理旋转惯量。

$$H_{\text{sys}}^t = \sum_{g=1}^{N_G} H_G^g \cdot P_G^{g,t} + \sum_{c=1}^{N_C} H_{\text{CSP}}^c \cdot P_{\text{CSP}}^{c,t} \quad (6)$$

$$M_H^t = \frac{H_{\text{sys}}^t}{f_0} \quad (7)$$

式中: H_{sys}^t 为时刻 t 系统的惯量; H_G^g 、 H_{CSP}^c 分别为火电机组和 CSP 机组的惯量常数; $P_G^{g,t}$ 、 $P_{\text{CSP}}^{c,t}$ 分别为时刻 t 投入运行的火电机组和 CSP 机组的额定容量; N_G 、 N_C 分别为火电机组和 CSP 机组数量; M_H^t 为系统惯性常数; f_0 为系统的标称频率。

需要指出的是,本文定义的广义惯量特指由系统内火电与光热电站提供的真实物理旋转惯量,以区别于风光等设备模拟的虚拟惯量。不同于逆变器受限于锁相环与控制算法的响应滞后,光热电站依托机电耦合机制,可在暂态扰动瞬间提供零时延的刚性惯量支撑。因此,本文选用无迟滞的物理惯量作为评估基准,以准确刻画弱电网暂态初期的真实抗扰底线。

1.3 频率安全约束

为防止功率扰动引发频率崩溃或触发低频减载 (under-frequency load shedding, UFLS), 本文依据国家标准划定频率安全阈值, 构建了如下约束模型。

1) 频率变化率约束。

高比例新能源环境下, 过大的频率变化率极易引发分布式电源防孤岛保护误动脱网或锁相环失步。参考 GB 38755—2019^[37] 及 GB/T 40595—2021^[38], 本文将 $|\text{RoCoF}| \leq 0.5 \text{ Hz/s}$ 作为保障并网稳定性的硬性约束, 通过优化系统物理惯量支撑来实现:

$$\left| \frac{df}{dt} \right| = \left| \frac{\Delta P_{\text{loss}}^t}{2H_{\text{sys}}^t} \cdot f_0 \right| \leq R_{\text{RoCoF max}} \quad (8)$$

式中: ΔP_{loss}^t 为突发事件功率扰动量; $R_{\text{RoCoF max}}$ 为允许的最大频率变化率限值。

2) 频率偏差约束。

为严格保证系统在 $N-1$ 故障情况下不触发第二道防线。依据 GB/T 40596—2021^[39], 本文采用预防性控制思路, 在动作阈值基础上预留安全裕度, 将暂态频率最低点安全约束设定为 49.5 Hz , 引入调速器爬坡受限模型^[40], 确保考虑滞后效应后的频率最低点仍满足安全阈值:

$$R_{G, \text{PFR}}^t \leq 2r_G \frac{M_H^t (f_0 - f_{\text{db}} - f_{\text{UFLS}})}{\Delta P_{\text{loss}}^t} \quad (9)$$

$$R_{\text{CSP}, \text{PFR}}^t \leq 2r_{\text{CSP}} \frac{M_H^t (f_0 - f_{\text{db}} - f_{\text{UFLS}})}{\Delta P_{\text{loss}}^t} \quad (10)$$

$$\sum_{g=1}^{N_G} R_{G, \text{PFR}}^{g,t} + \sum_{c=1}^{N_C} R_{\text{CSP}, \text{PFR}}^{c,t} \geq \Delta P_{\text{loss}}^t \quad (11)$$

式中: $R_{G, \text{PFR}}^t$ 、 $R_{\text{CSP}, \text{PFR}}^t$ 分别为常规火电厂和 CSP 电站提供的一次调频备用容量; f_{db} 为系统调速器死区; f_{UFLS} 为系统低频减载装置的动作阈值; r_G 、 r_{CSP} 分别为常规火电厂和 CSP 电站提供一次调频响应的最大爬坡速率; $R_{G, \text{PFR}}^{g,t}$ 、 $R_{\text{CSP}, \text{PFR}}^{c,t}$ 分别为常规火电机组和 CSP 机组提供的一次调频备用容量。

需要说明的是, 本文提出的预防性调度策略主要针对 $N-1$ 单一故障。若发生 $N-2$ 或直流双极闭锁等极端大扰动, 导致系统物理惯量需求超出调控极限时, 系统将依规启动第二、三道防线以维持稳定, 不在本文预防性调度的求解范围内。

1.4 光热-光伏联合电站的优化目标和约束

1) 目标函数。

本文综合考虑光伏、光热发电运行维护成本、火电机组发电成本、一次调频备用成本、旋转备用成本、弃光成本以及切负荷成本, 并以系统成本最低为目标, 建立系统调度优化成本模型。

$$\begin{aligned} \min F = & \sum_{t=1}^{N_T} \sum_{g=1}^{N_G} C_G^g (P_G^{g,t} + R_{G, \text{PFR}}^{g,t}) + \\ & \sum_{t=1}^{N_T} \sum_{c=1}^{N_C} C_{\text{CSP}}^c (P_{\text{CSP}}^{c,t} + R_{\text{CSP}, \text{PFR}}^{c,t}) + \\ & \sum_{t=1}^{N_T} \sum_{g=1}^{N_G} C_{G, \text{R}}^g (R_{G, \text{up}}^{g,t} + R_{G, \text{down}}^{g,t}) + \\ & \sum_{t=1}^{N_T} \sum_{c=1}^{N_C} C_{\text{CSP}, \text{R}}^c (R_{\text{CSP}, \text{up}}^{c,t} + R_{\text{CSP}, \text{down}}^{c,t}) + \\ & C_{\text{RE}, \text{cur}} \left(\sum_{t=1}^{N_T} \sum_{c=1}^{N_C} P_{\text{CSP}, \text{cur}}^{c,t} + \sum_{t=1}^{N_T} \sum_{p=1}^{N_P} P_{\text{PV}, \text{cur}}^{p,t} \right) + \\ & C_{\text{L}, \text{cur}} \sum_{t=1}^{N_T} \sum_{l=1}^{N_L} P_{\text{L}, \text{cur}}^{l,t} + \sum_{t=1}^{N_T} C_{\text{EH}} \cdot P_{\text{EH}}^t \end{aligned} \quad (12)$$

式中: C_G^g 、 C_{CSP}^c 分别为火电机组和 CSP 机组的单位运

行成本; $C_{G,R}^g$ 、 $C_{CSP,R}^c$ 分别为火电机组和 CSP 机组的单位旋转备用成本; $R_{G,up}^{g,t}$ 、 $R_{G,down}^{g,t}$ 分别为 t 时刻火电机组提供的正负旋转备用; $R_{CSP,up}^{c,t}$ 、 $R_{CSP,down}^{c,t}$ 分别为光热机组提供的正负旋转备用; $C_{RE,cur}$ 为弃光惩罚成本系数; $C_{L,cur}$ 为弃负荷单位惩罚成本; C_{EH} 为电加热器单位运行成本; $P_{CSP,cur}^{c,t}$ 、 $P_{PV,cur}^{p,t}$ 分别为光热、光伏未利用的功率; $P_{L,cur}^{l,t}$ 为切负荷功率。

2) 约束条件。

约束条件包括功率平衡、机组运行上下限与爬坡、储热系统容量与充放热状态、系统正负旋转备用容量等,具体公式见附录 A。

2 基于改进 TD3 的调度优化算法

2.1 马尔可夫决策过程建模

将每个决策周期内的电源、负载及其他不受调度策略控制的量近似为常数,该调度模型可看作一个序贯决策模型,进一步转化成马尔可夫决策过程 (Markov decision process, MDP), 包括四个关键要素: 状态空间 s 、动作空间 a 、状态转移概率 p 和奖励函数 r 。考虑到电力系统包含高度非线性潮流方程及机组动态特性, 构建显式、精确的状态转移概率矩阵极其困难。因此, 本文采用无模型 (model-free) 强化学习方法, 将电力系统仿真环境视为“黑盒”, 智能体可以通过与环境交互产生的时序样本数据来隐式地学习状态转移概率。智能体会根据观察到的环境当前状态采取适当的动作, 并获取相应的奖励。决策的最终目标为最大化累计奖励函数。

1) 状态空间。

t 时刻的状态空间 s^t 是有关当前观测环境的信息, 本文包含了气象参数、系统参数和上一个时段的动作 a^{t-1} 。

$$s^t = [I_{DN}^t, I_{GN}^t, P_L^t, S_{OC}^t, a^{t-1}] \quad (13)$$

式中: P_L^t 为系统总负荷; S_{OC}^t 为 TES 内部容量状态 (state of charge, SOC)。

2) 动作空间。

动作空间是智能体根据环境状态空间作出的设备输出决策, 本文包含 PV-CSP 联合电站出力, 火电机组出力和电加热功率等。

$$a^t = [P_{CSP}^{c,t}, P_{PV}^{p,t}, P_G^{g,t}, R_{G,PFR}^{g,t}, R_{CSP,PFR}^{c,t}, R_{G,down}^{g,t}, R_{CSP,up}^{c,t}, R_{CSP,down}^{c,t}, P_{EH}^t, P_{L,cur}^{l,t}] \quad (14)$$

3) 奖励函数。

奖励是智能体根据运营目标从当前动作中获得的收益, 对于本文含光热-光伏联合电站系统的优化调度, 奖励为综合运行成本 F 与越限惩罚项 F_{punish} 的

负加权。

$$r^t = -(F + F_{punish}) \quad (15)$$

2.2 孪生延迟深度确定性策略算法

相较于基于模型的优化方法, DRL 在不确定环境中展现出更强的适应性与鲁棒性。深度确定性策略梯度算法作为面向连续动作空间的 DRL 方法, 采用 Actor-Critic (A-C) 架构, 并利用经验回放和目标网络提升稳定性。然而, A-C 架构下的典型算法 DDPG 因固有的动作值高估和目标值波动问题限制了其性能。为解决上述问题, 本文采用 TD3 算法, 该算法通过以下三项关键技术提升基于策略梯度的深度强化学习性能。

1) 引入孪生 Critic 网络结构。

构建两个独立的 Critic 网络, 通过取二者输出的最小值更新策略。

$$L(\theta^{\omega^i}) = \frac{1}{N} \sum_{i=1}^N (Q^i(s^t, a^t | \theta^{\omega^i}) - y^t)^2 \quad (16)$$

$$y^t = r^t + \gamma \min_{i=1,2} Q^i(s^{t+1}, a^t | \theta^{\omega^i}) \quad (17)$$

式中: $L(\theta^{\omega^i})$ 为第 i 个 Critic 网络的损失函数; θ^{ω^i} 为第 i 个 Critic 网络的权重参数; N 为历史经验回放池的样本对数; $Q^i(s^t, a^t | \theta^{\omega^i})$ 为当前网络输出; y^t 为 t 时刻的目标 Q 值; r^t 为 t 时刻环境即时奖励; γ 为奖励折扣因子。

2) 实施延迟策略更新机制。

策略网络在 Q 网络更新后按设定的时间间隔更新, 从而缓解因过于频繁更新而导致的参与者网络中的策略波动。

$$\begin{cases} \theta^{\omega'} \leftarrow \tau \theta^{\omega} + (1 - \tau) \theta^{\omega'} \\ \theta^{\omega'} \leftarrow \tau \theta^{\omega} + (1 - \tau) \theta^{\omega'} \end{cases} \quad (18)$$

式中: θ^{ω} 、 $\theta^{\omega'}$ 为 Critic 网络和 Actor 网络的参数; τ 为软更新系数; θ^{ω} 、 $\theta^{\omega'}$ 为目标 Critic 网络和目标 Actor 网络的参数。

3) 目标策略平滑正则化。

为抑制目标值估计方差, 在目标 Actor 网络输出中注入截断高斯噪声。

$$a^t = \mu^t(s^{t+1} | \theta^{\omega'}) + \varepsilon, \varepsilon \sim \text{clip}(\mathcal{N}(0, \sigma), -c, c) \quad (19)$$

式中: ε 为遵循高斯分布的噪声值; $\text{clip}(\cdot)$ 为截断操作, 把噪声限制在 $[-c, c]$ 区间内; $\mathcal{N}(0, \sigma)$ 为均值为 0, 标准差为 σ 的正态高斯噪声。

2.3 改进 TD3 算法架构设计

在光热-光伏联合电站调度场景中, 原始 TD3 算法面临两大关键挑战, 显著增加了智能体的训练难度。其一, 强时序依赖建模不足。光照、负荷、储热状态、频率/惯量约束满足情况等系统状态和调度决

策具有显著的时序关联性,传统TD3算法仅基于单步状态进行决策,难以有效提取状态序列中隐含的长短期时序特征,导致策略在处理频率-惯量约束等强历史依赖性问题时效果不佳。其二,关键事件学习效率低。均匀采样机制忽略了不同样本对网络参数更新

的价值差异,而调度过程中涉及系统安全或高成本/高效益的关键决策事件相对稀少却至关重要,均匀采样导致这些关键事件的学习机会不足,拖慢收敛速度并影响最终策略性能。为有效应对上述挑战,本文对TD3算法进行了针对性改进,如图2所示。

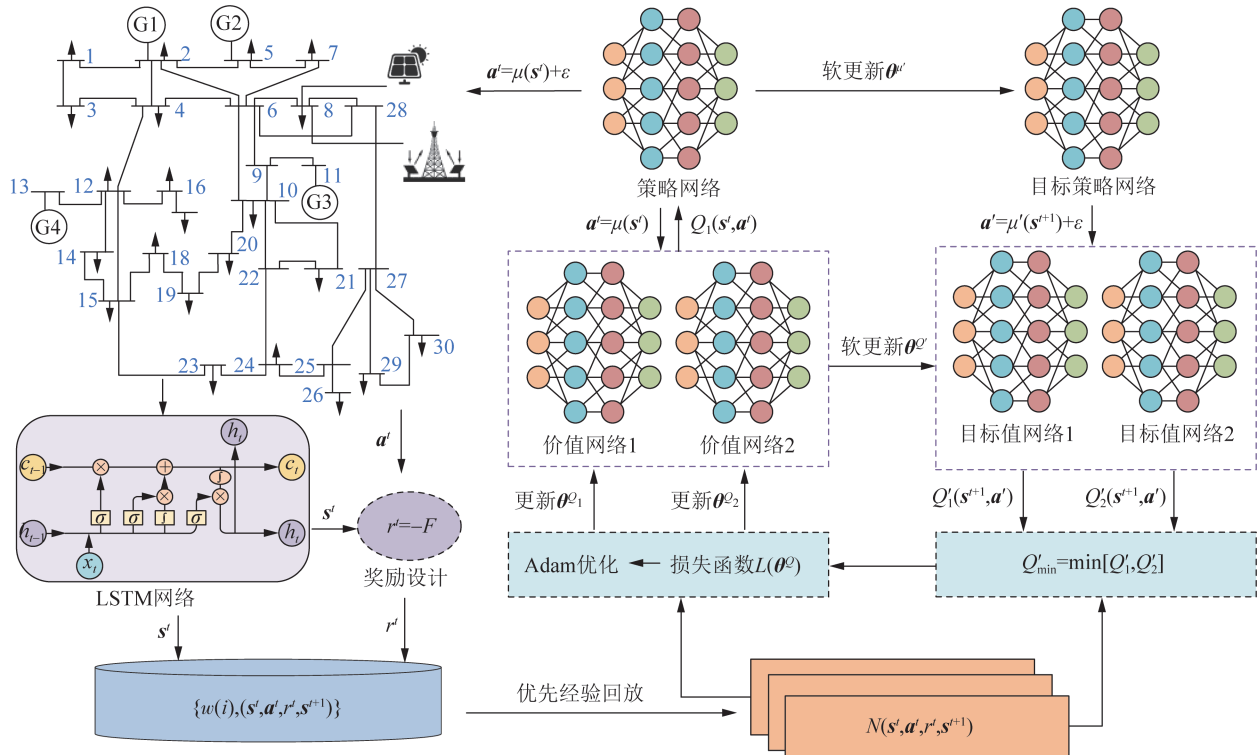


图2 基于改进TD3算法的光热-光伏联合电站调度优化框架图

Fig. 2 Optimization framework for scheduling of CSP-PV hybrid plant based on improved TD3 algorithm

1) 嵌入LSTM时序特征提取器。

构建基于LSTM的时序特征提取模块,替代原始TD3网络中全连接层对单步状态的直接映射。LSTM层以滑动时间窗口内的连续状态序列作为输入向量。利用其特有的门控机制,LSTM能够有效提取隐含在历史状态轨迹中的长短期时序特征,并生成包含系统演化规律的隐含状态向量。隐含状态向量被同时输入Actor网络和Critic网络,提升策略对系统时序演化特性的建模能力,使其能够更有效地处理频率-惯量约束等依赖于历史状态的问题。

2) 集成优先经验回放。

鉴于均匀采样机制在处理稀疏高价值样本时的低效性,本文引入优先经验回放策略。该策略的核心思想是依据经验样本对网络参数更新的价值差异进行非均匀采样。具体而言,采用样本的时间差分误差绝对值作为衡量其价值的指标。时间差分误差较大的样本表明当前网络对该状态的评估准确度较低,往往蕴含系统状态跃变或边界

约束触发等高价值信息,因此被赋予更高的采样优先级。

$$p(i) = \frac{|\delta(i)|^\alpha + \xi}{\sum_{i=1}^M (|\delta(i)|^\alpha + \xi)} \quad (20)$$

式中: $p(i)$ 表示样本采样概率; $|\delta(i)|$ 为样本的时间差分误差绝对值;超参数 α 用于在纯贪婪采样和均匀采样之间取得平衡,本文设定为0.6; ξ 是一个小的正常数,用于确保每个样本至少有一个小的优先级; M 为经验池中的样本总数。

此外,为修正因非均匀采样导致的动作价值分布偏差,在计算梯度更新时引入重要性采样权重 $w(i)$ 进行加权校正。

$$w(i) = \left(\frac{1}{M} \cdot \frac{1}{p(i)} \right)^\beta \quad (21)$$

式中:超参数 β 控制校正强度,其初始值设为0.4,并随训练进程线性增长至1.0。网络参数更新时,样本的梯度更新量需乘以 $w(i)$ 。经验池底层采用Sum-Tree数据结构,以支持高效的优先级采样与权重

更新。

调度算法的完整执行流程伪代码如表1所示。

表1 考虑频率-惯量协同约束的 LSTM-PER-TD3 调度算法执行流程

Table 1 Execution flow of the LSTM-PER-TD3 scheduling algorithm considering frequency-inertia collaborative constraints

初始化 Actor 网络 θ^a 与两个 Critic 网络的权重 θ^c , 以及对应的目标网络权重; 初始化优先经验回放池, LSTM 历史步长, PER 超参数, 延迟更新频率
for episode = 1 to M do
重置电网调度环境, 获取初始状态 s' ;
构建初始历史状态序列;
for step $t = 1$ to T do
通过 Actor 网络和 LSTM 提取特征选择动作, 并加入探索噪声;
$a' = \mu'(s^{t+1} \theta^a) + \varepsilon, \varepsilon \sim \text{clip}(\mathcal{N}(0, \sigma), -c, c)$;
下发调度指令, 获取奖励 r' 和下一时刻状态;
将下一时刻状态压入序列末尾, 移除最旧状态, 得到新序列;
将状态转移元组存入经验池, 并赋予当前最大优先级;
if 经验池中样本数 > 批次大小 then
按概率中抽取 K 个样本;
计算每个样本的重要性采样权重;
计算目标动作;
计算目标 Q 值;
计算时间差分误差 s ;
更新 Critic 网络, 最小化加权损失函数;
时间差分误差更新经验池中对应样本的优先级;
if $t \bmod \tau == 0$ then
通过最大化 Q 值更新 Actor 网络参数;
软更新目标网络;
$\theta^c \leftarrow \tau \theta^c + (1 - \tau) \theta^c$
$\theta^a \leftarrow \tau \theta^a + (1 - \tau) \theta^a$
end if
end if
end for
end for

3 算例分析

3.1 参数设置

本文选用改进的 IEEE-30 节点系统作为仿真测试平台。为模拟“沙戈荒”大型新能源基地并网场景, 对标准系统拓扑进行了如下修正: 在节点 7 处接入一套集成了光热-光伏-储热-电加热的联合发电系统。该系统包含 100 MW 光热发电机组、200 MW 光伏阵列, 并配置最大容量 3500 MWh 的储热系统及 120 MW 电加热器以提供灵活性支撑, 详细参数见附录表 B1 和 B2。为模拟新能源对传统火电的替代场景, 停运了原系统节点 2 处的同步发电机组, 仅保留其余 4 台火电机组作为基础电源与备用支撑。仿真环境采用典型“沙戈荒”地区的历史光照辐射及负荷预测数据作为系统输入, 相关曲线见附录图 B1。在频率安全与备用参数设置方面, 设定系统额定频率 f_0

为 50 Hz, 一次调频死区 f_{db} 为 0.033 Hz^[41]。设定单一最大功率扰动为负荷的 10%, 参考《电力系统安全稳定导则》及相关研究^[41], 设定正、负旋转备用需求分别为总负荷的 5% 和 2%。需要说明的是, 针对光伏突增风险, 本文模型允许通过快速弃光及光热转储热运行来消纳过剩功率, 因此 2% 的负荷备用主要用于应对负荷侧的小幅波动, 足以满足平衡需求。

算法基于 TensorFlow2.0 深度学习框架搭建。算法的核心网络结构与关键超参数分别如附录 B 表 B3 和 B4 所示。考虑到光热电站储热系统具有显著的热惯性, 且云层遮挡导致的光照波动具有短时持续性, 调度策略需要感知过去一段时间内的系统状态演变趋势。本文进行了多组序列长度的量化对比实验, 详见附录 C 表 C1, 最终将 LSTM 的时间接收窗口设定为 4 h。此外, 为保障算法采样的无偏性与收敛稳定性, 本文针对优先经验回放机制中的关键超参数进行了敏感性分析与整定, 详见附录 C 图 C1。

3.2 LSTM-PER-TD3 算法性能分析

为全面验证所提 LSTM-PER-TD3 算法的有效性, 本文在改进的 IEEE-30 节点系统上对 4 种 TD3 变体算法在 1500 个训练轮次中的收敛性能进行了比较, 如图 3 所示。其中, TD3 为基准算法; P-TD3 为仅引入优先经验回放机制的变体; L-TD3 为仅嵌入 LSTM 时序特征提取模块的变体; PL-TD3 (即本文所提 LSTM-PER-TD3) 为同时融合上述两种机制的完整改进算法。

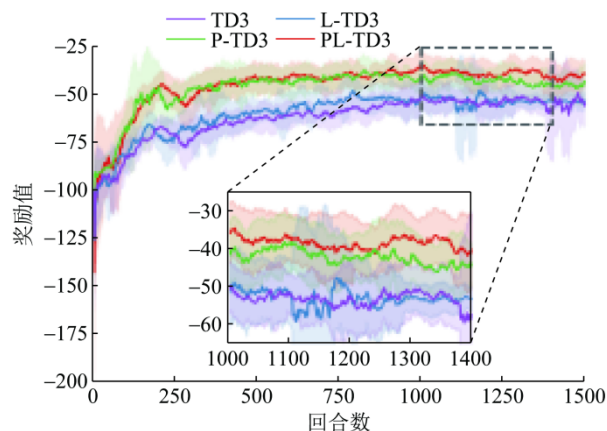


图3 改进 IEEE-30 节点算例的平均奖励对比结果

Fig. 3 Comparison of average rewards for the improved IEEE-30 bus system

由图 3 可见, 4 种算法的平均奖励收敛曲线呈现出清晰的性能层级。在前 300 回合, 引入优先经验回放机制的算法的奖励曲线以最陡峭的斜率迅速上升, 明显快于其他两种算法。这表明 PER 机制在早期学习阶段展现出最高的样本利用效率, 能够帮助

智能体快速找到一条可行的路径。相较之下,L-TD3和TD3算法的学习速度相对平缓。

进入 300 ~ 1500 回合后,各算法的性能差异进一步显现。本文提出的 LSTM-PER-TD3 算法曲线始终保持在最高水平,其平均奖励曲线稳定在其他曲线之上,并最终收敛到一个最佳的性能区间。与之相比,P-TD3 曲线虽然在早期表现优异,但最终收敛值低于 LSTM-PER-TD3。而 L-TD3 和 TD3 曲线则收敛到更低的奖励值,且 TD3 算法的曲线波动最为剧烈。

为定量对比各算法的收敛性能,本文进一步统计了最终平均奖励、收敛回合数与稳定性等关键指标,如表 2 所示。

表 2 各算法变体在训练过程中的性能指标对比
Table 2 A comparison of performance metrics of each algorithm variant during the training process

算法变体	最终平均奖励	90% 性能回合数	50% 性能回合数	标准差
TD3	-54.663	788	89	17.987
P-TD3	-44.557	390	93	7.834
L-TD3	-53.001	610	70	3.902
PL-TD3	-39.924	202	28	8.436

由表 2 可知,LSTM-PER-TD3 算法的最终平均奖励为-39.924,相较于基准 TD3 算法的-54.663 提升了 36.9%。即便是与仅引入优先经验回放的 P-TD3 算法相比,LSTM-PER-TD3 算法仍表现出 11.6% 的性能增益。这验证了融合时序特征提取与优先回放机制后,算法能够有效突破局部最优,收敛至质量更高的调度策略。

在早期学习效率方面,LSTM-PER-TD3 算法仅需 28 个训练回合即可达到最终性能的 50%,显著优于基准 TD3 算法的 89 回合、P-TD3 算法的 93 回合及 L-TD3 算法的 70 回合。该结果有力证实了所提算法在训练初期具有极高的样本利用效率。在收敛速度方面,LSTM-PER-TD3 算法同样表现出显著优势,达到 90% 性能指标仅需 202 回合,远少于基准 TD3 算法的 788 回合。上述数据共同表明,改进算法不仅加速了训练收敛过程,更从根本上提升了最终策略的优化质量。

渐近收敛的稳定性指标揭示了算法在收敛速度与稳健性之间的权衡关系。尽管 LSTM-PER-TD3 算法的标准差优于基准 TD3 算法,但略高于仅引入 LSTM 的 L-TD3 算法。这一现象表明,虽然 PER 机制通过聚焦高差分误差样本加速了学习进程,但也可能引入额外的估计方差;而 LSTM 对时序信息的有效提取则为算法提供了基础的稳健性保障,使得 L-TD3

算法在多次实验中表现出最小的波动性。

3.3 与传统 MILP 优化方法的对比分析

为进一步评估本文所提 DRL 算法的性能水平,本节引入经典的混合整数线性规划(mixed integer linear programming, MILP)方法作为基准进行对比。MILP 模型基于全局完美预测信息,旨在寻找系统运行的理论最优解。考虑到非线性 AC 潮流约束在 MILP 中求解困难,本文建立了基于系统功率平衡与线性化频率安全约束的 MILP 模型,并使用开源求解器 PuLP 进行求解。为了保证对比的公平性与物理有效性,本文采用离线求解-在线校验的策略,即将 MILP 优化的调度指令回代至 Pandapower 非线性仿真环境中,计算其实际运行成本并检测是否存在越限情况。对比结果如表 3 所示。从经济性角度分析,MILP 方法凭借其对未来信息的完全掌握及全局寻优能力,实现了最低的运行成本。本文提出的 LSTM-PER-TD3 算法在仅依赖局部观测信息的情况下,能够收敛至 98.033 万元,与理论最优解仅相差 2.94%。

表 3 LSTM-PER-TD3 与 MILP 基准方法的性能对比
Table 3 Performance comparison between LSTM-PER-TD3 and MILP benchmark

方法	总运行成本/万元	与最优解差距/%	频率/备用越限次数	在线决策响应时间/s
MILP	95.230	0	0	离线优化
本文	98.033	+2.94	0	1

进一步从算法决策原理剖析,这 2.94% 的成本差异本质上反映了两种优化范式在应对不确定性时的策略行为分化。MILP 模型建立在全局完美预测信息的确定性假设之上,倾向于采取紧贴物理约束边界的极致运行策略以追求经济效益最大化。然而,本文的 DRL 智能体身处充满源荷波动的不确定性交互环境中,无法预知未来的精确状态。为了严格规避频率越限带来的高额惩罚,智能体在策略演化过程中自发内化了“风险厌恶”机制,主动在动作决策中留出了更宽裕的安全裕度。因此,这种略显保守的策略所带来的成本增加,并非算法求解精度的缺陷,而是为对冲环境不确定性、确保系统频率绝对安全所支付的合理“安全溢价”。

从应用维度分析,尽管 MILP 效果略优,但其严重依赖于简化的线性物理模型,忽略了网损与电压约束,且随着网络节点增加,计算复杂度呈指数级上升,难以满足实时调度需求。相比之下,本文的 DRL 方法直接与非线性 AC 环境交互,能够隐式地处理复杂的电网物理特征,且训练完成后具备秒级的在线

响应速度,更加适用于高比例新能源电网的实时智能调度。

3.4 系统优化调度结果分析

为验证所提调度策略的有效性,本节基于高比例可再生能源渗透的独立测试系统开展了24 h日前优化调度仿真,结果如图4—8所示。

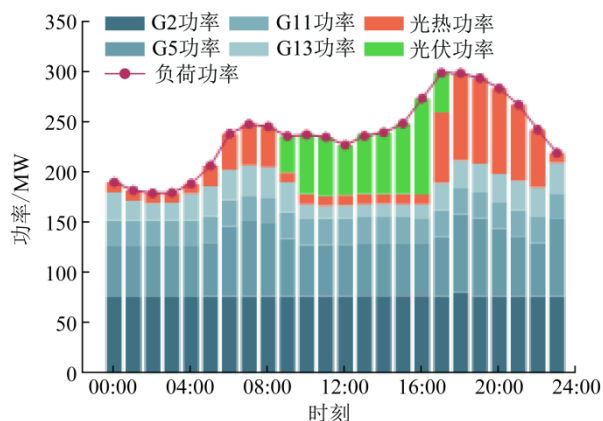


图4 典型日日前调度优化结果

Fig. 4 Optimal day-ahead scheduling results for a typical day

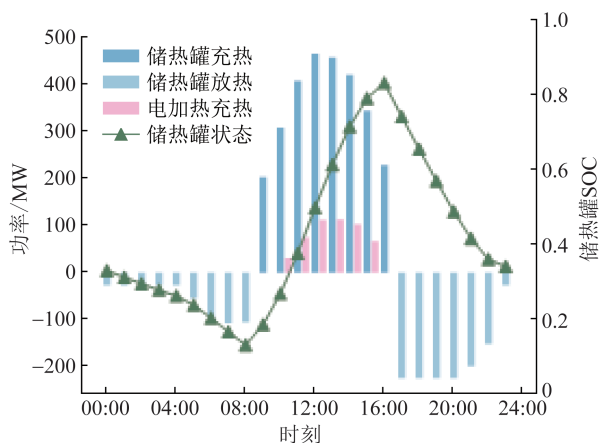


图5 储热罐充放热功率以及对应的SOC变化曲线

Fig. 5 Charging/discharging power and SOC variation curve of thermal storage tank

如图4所示,所提调度策略有效提升了系统的灵活调节能力,实现了对各类资源的优化利用。在时序分布上,常规火电机组主要承担夜间及凌晨时段的基荷与部分腰荷,为系统提供基础功率支撑。日间09:00—16:00,光伏机组出力占据主导地位,大幅降低了系统对化石能源的依赖。值得注意的是,CSP充分发挥了其灵活调节特性:在午间光伏出力高峰期,CSP机组维持最小技术出力或停机备用,将集热场捕获的热能优先注入储热系统;待17:00—21:00光伏出力骤降而负荷攀升至峰值时,CSP迅速增加出力以弥补功率缺额。该策略有效平抑了净负荷“鸭子曲线”带来的爬坡压力,显著增强了晚高峰时段的

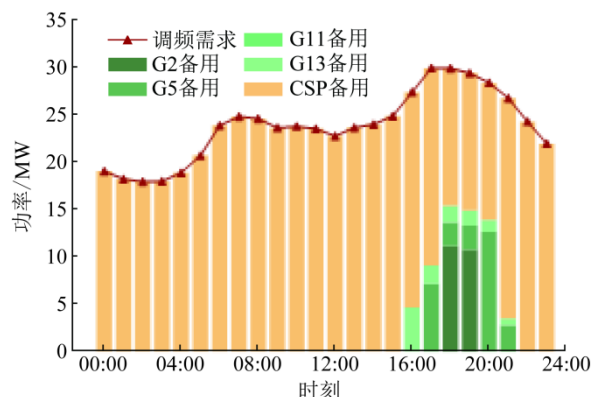


图6 一次调频备用提供计划

Fig. 6 Primary frequency regulation reserve provision schedule

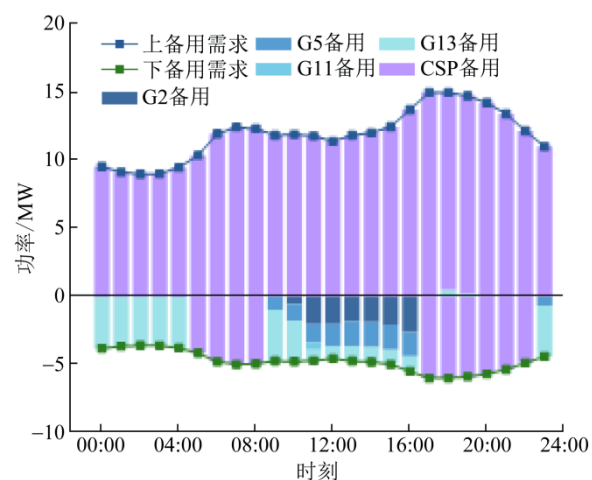


图7 上下旋转备用提供计划

Fig. 7 Up/down spinning reserve provision schedule

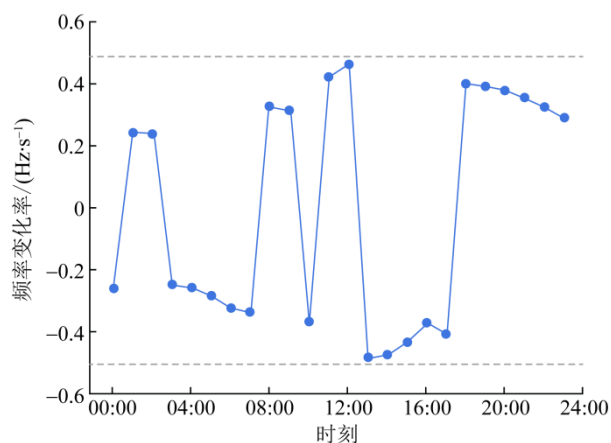


图8 系统频率变化率的日内动态曲线

Fig. 8 Intra-day dynamic profile of the system rate of change of frequency

供电可靠性。

图5揭示了TES作为系统灵活性核心的运行机制。TES呈现出鲜明的“日充夜放”模式,实现了能量在时间维度上的高效转移。在充电阶段09:00—15:00, TES的SOC从0.1左右的低位稳步攀升至接近

0.95的上限。此阶段的能量来源体现了多能互补优势:不仅吸纳了来自CSP镜场的直接热能,还通过EH回收了原本受阻的弃光电量,实现了可再生能源的高效消纳。在放电阶段05:00—08:00及17:00—22:00, TES释放热能驱动CSP发电,特别是在晚高峰时段提供了关键的功率支撑。此外,清晨05:00—08:00实施的主动预放电策略有效降低了初始SOC,为日间吸纳光热资源预留了充足裕度,这是避免因储能饱和和引发弃能风险的关键预见性调度措施。

所提策略在满足能量平衡的同时,精确地满足了系统的动态安全备用需求。图6展示了在一次调频备用方面,CSP电站凭借其同步发电机的惯量特性和快速响应能力,在绝大部分时段,尤其是可再生能源大发时段承担了主要的调频任务,使得火电机组能更稳定地运行在经济区间。仅在夜间等少数时段,火电机组作为补充参与调频,确保了调频备用时刻满足需求。图7展示了在旋转备用方面,调度策略具备动态经济分配的特性。CSP在夜间凭借其丰富的热能储备提供了主要的上备用,而在白天则由火电机组承担更多责任。上下备用的供给总量始终大于需求曲线,保证了系统应对负荷波动的能力。图8表明调度策略在所考虑工况下能够维持RoCoF在安全范围以内,体现出一定的动态支撑能力。

3.5 对比方案设置

为验证本文所建考虑频率-惯量约束的调度模型及其求解算法的有效性,本节设置了4种不同场景进行对比分析:

场景一(无约束):仅考虑功率平衡约束,以总运行成本最低为优化目标。该方案结果理论上经济性最优,但未计及系统安全运行要求,用作基准参考。

场景二(仅备用):在场景一的基础上,增加系统正/负旋转备用容量约束。

场景三(仅频率):在场景一的基础上,增加系统广义惯量、频率变化率及频率偏差安全约束。

场景四(完整模型):本文所提模型,同时考虑功率平衡、旋转备用及频率-惯量安全约束。

不同约束场景下光热-光伏联合电站调度结果如表4所示。

在系统可靠性方面,场景一虽维持了功率平衡,但引发了5次频率越限并伴随42.3 MWh的负备用短缺;场景二消除了备用短缺,但频率越限仍达3次;场景三实现了零频率越限,却仍存在36.2 MWh的备用短缺;而场景四则同时达成零频率越限与零备用短缺的双重清零,验证了备用容量与频率-惯量约束的互补性与协同价值。

表4 不同约束场景下光热-光伏联合电站调度方案对比
Table 4 Comparison of CSP-PV hybrid plant dispatch under different constraint scenarios

评估指标	指标值			
	场景一	场景二	场景三	场景四
总运行成本/万元	64.965	81.683	81.094	98.033
发电成本/万元	64.965	65.808	64.988	65.824
旋转备用成本/万元	0	15.875	0	15.995
一次调频备用成本/万元	0	0	16.106	16.214
频率越限事件次数/次	5	3	0	0
正备用短缺量/MWh	0	0	0	0
负备用短缺量/MWh	42.3	0	36.2	0

在经济性维度,4种场景发电成本相近,总成本差异主要源于辅助服务采购:场景二/四因配置旋转备用产生显著成本,场景三/四因调用一次调频备用产生成本。场景一至四总成本分别为64.965万元、81.683万元、81.094万元、98.033万元,量化揭示了系统安全性提升与运行成本增加之间的耦合权衡机理。边际分析显示:场景一到二的成本增量15.875万元基本对应旋转备用投入,场景一到三的增量16.106万元主要源于一次调频备用,场景三到四的增量15.995万元则用于补齐旋转备用缺额,以满足全维度的系统安全约束。综上,单一依赖备用或频率-惯量约束均无法保障全时域安全;完整模型通过储能协同提供容量与动态支撑,在可量化成本下实现零越限与零短缺,为调度决策在经济性与多维度安全性间的最优权衡提供了实证依据与方法论支撑。

3.6 扩展算例分析

为进一步验证所提LSTM-PER-TD3算法在大规模电力系统中的泛化能力与求解稳定性,本文构建了拓扑结构更为复杂的改进IEEE-57节点测试系统进行扩展实验。该系统在节点7处接入了一套包含400 MW光伏、200 MW光热及7000 MWh储热系统的联合电站,以模拟“沙戈荒”大型新能源基地场景。详细的系统参数设置及机组配置见附录D表D1—D3。

图9展示了算法在该系统下的训练收敛性能。相较于30节点系统,网络拓扑约束导致的可行域收窄及状态空间维度增加,算法进入平稳收敛期所需的训练回合数延长至约350回合。尽管训练初期因探索高维边界约束而出现震荡,但得益于PER机制对高价值样本的有效挖掘,策略最终实现了渐近收敛且奖励值波动极小,证明了算法在复杂高维环境下的鲁棒性。

关于该大规模系统在典型日下的详细优化调度

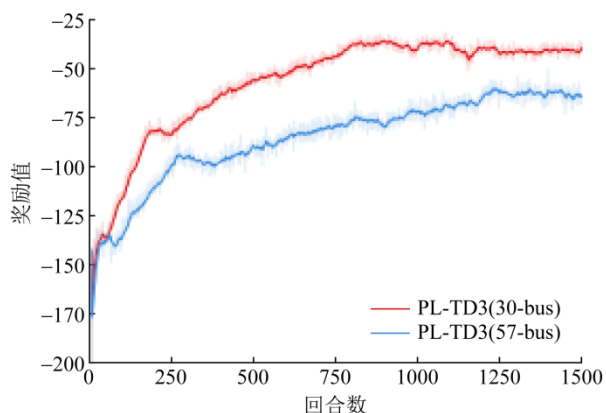


图 9 不同算例奖励对比结果

Fig. 9 Comparison results of rewards in different test systems

结果,包括各机组出力堆叠图、储热系统SOC变化、备用分配及频率安全校验曲线,见附录D图D1至图D5。仿真结果表明,在大规模网络约束下,所提策略依然能够有效协调源-储-网资源,在应对8%负荷阶跃扰动时,成功将RoCoF限制在安全范围内,实现了经济性与安全性的双重优化。

3.7 调度策略鲁棒性分析

尽管频率安全阈值依据国标设定为定值,但为了验证所提策略在不同强度扰动下的适应能力,本文设置了不同幅度的负荷阶跃扰动场景,对系统的频率响应指标进行了压力测试,结果如表5所示。

表 5 不同扰动幅度下的系统频率安全指标与鲁棒性分析
Table 5 Frequency security indicators and robustness analysis under different disturbance magnitudes

扰动幅度/%	功率缺额极值/MW	RoCoF 极值/(Hz·s ⁻¹)	安全指标越限情况
5	15	0.18	无越限
10	30	0.40	无越限
15	45	0.61	越限,未切负荷
20	60	0.81	越限,触发切负荷

表5显示,在扰动幅度为5%至10%的中小负荷阶跃场景下,调度策略通过协同光热与火电机组的一次调频及物理惯量支撑,成功将RoCoF极值控制在0.40 Hz/s以内,实现了零越限的安全运行目标。当扰动幅度增大至15%、功率缺额达45 MW时,RoCoF升至0.61 Hz/s。此时,改进TD3算法能够深度挖掘光热储能的快速爬坡潜力进行紧急功率支援,虽各项指标轻微突破预防控制阈值,但成功避免了低频减载装置动作,体现了极限工况下的主动调节韧性。仅在面对20%的极端扰动及60 MW的巨大功率缺额时,因超出源侧调频物理极限导致RoCoF剧增至0.81 Hz/s,系统正确触发切负荷等第二道防线以阻

断频率崩溃。综上,所提策略不仅在10%以内的常态扰动下具备强鲁棒性,在极端工况下也能通过多源协同响应提供有效缓冲。

4 结 论

本文针对“沙戈荒”高渗透率新能源基地面临的低惯量危机,提出一种计及频率-惯量协同约束的光热-光伏联合电站深度强化学习调度策略,实现了系统动态安全与运行经济性的帕累托均衡。所提策略能够在确保频率安全的前提下有效权衡系统经济性与安全性,且改进算法在不同规模系统下展现出优于传统TD3算法的收敛速度与鲁棒性。主要结论如下:

1)光热电站具备无可替代的广义惯量支撑价值。CSP不仅具备大容量时移储能特性,其内置的同步发电机组在突发功率扰动时,可为弱电网提供无迟滞的刚性物理支撑。将CSP由单纯储能单元转变为广义惯量支撑电源纳入协同调度,从物理机理上有效阻断了暂态频率越限与切负荷风险。

2)在极低惯量场景下,坚守频率安全底线存在显性的经济成本。仿真表明,为满足频率与备用的双重安全约束,系统调度策略会呈现出主动规避风险的保守性特征,需支付约20%的安全溢价。这种以适度的经济让步换取系统绝对可靠性与动态安全裕度的机制,符合实际工程中安全保底的调度逻辑。

3)改进的数据驱动算法有效攻克了高维强时序耦合约束下的求解瓶颈。融合长短期时序感知与优先经验回放的改进TD3算法,克服了传统数学规划面临的维度灾难及原生算法关键样本利用率低的缺陷,为复杂电网的在线自适应安全防御与智能化决策提供了高效工具。

本文在建模时主要基于固定的电网物理拓扑假设,尚未充分计及实际工程运行中因故障隔离或检修引发的网络拓扑动态演化。后续研究将进一步考虑电网拓扑动态重构对系统惯量空间分布的影响,并拓展至多源微网集群交互条件下的分布式协同调度。

利益冲突声明 (Conflict of Interests):

所有作者声明不存在利益冲突。

作者贡献声明 (Authors' Contributions):

陈实、臧天磊进行了研究设计,晏红平、刘艺洪进行了模型代码编写和实验数据分析,李华强、陈江平、王舒灏参与了论文写作指导和修订。所有作者均阅读并同意了论文终稿内容。

5 参考文献

- [1] 张智刚, 康重庆. 碳中和目标下构建新型电力系统的挑战与展望[J]. 中国电机工程学报, 2022, 42(8): 2806-2818.
- Zhang Zhigang, Kang Chongqing. Challenges and prospects for constructing the new-type power system towards a carbon neutrality future[J]. Proceedings of the CSEE, 2022, 42(8): 2806-2818.
- [2] 国家能源局. 2025年全国电力工业统计数据简报[EB/OL]. (2025-07-31) [2025-10-20]. <https://www.nea.gov.cn/20250731/36b5b3b74d344325bb67f38b8e943c18/c.html>.
- National Energy Administration. Briefing on national power industry statistics in 2025[EB/OL]. (2025-07-31) [2025-10-20]. <https://www.nea.gov.cn/20250731/36b5b3b74d344325bb67f38b8e943c18/c.html>.
- [3] 王杨, 罗抒予, 姚凌翔, 等. 面向大型新能源基地的太阳能光热发电规划研究综述: 场景、模型与发展方向[J]. 电网技术, 2025, 49(7): 2712-2724.
- Wang Yang, Luo Shuyu, Yao Lingxiang, et al. A review of concentrating solar power planning for large-scale renewable energy bases: scenarios, models and development perspectives[J]. Power System Technology, 2025, 49(7): 2712-2724.
- [4] 李湃, 王伟胜, 黄越辉, 等. 大规模新能源基地经特高压直流送出系统中长期运行方式优化方法[J]. 电网技术, 2023, 47(1): 31-40.
- Li Pai, Wang Weisheng, Huang Yuehui, et al. Method on optimization of medium and long term operation modes of large-scale renewable energy power base through UHVDC system[J]. Power System Technology, 2023, 47(1): 31-40.
- [5] 沙韵, 周明, 杨宏基, 等. 考虑光热电站和直流联络线灵活性的高比例新能源互联系统优化运行[J]. 电网技术, 2020, 44(9): 3306-3313.
- Sha Yun, Zhou Ming, Yang Hongji, et al. Interconnected power system optimal operation with renewable generation considering flexibility of concentrating solar power plants & HVDC Tie-line[J]. Power System Technology, 2020, 44(9): 3306-3313.
- [6] 陈实, 郭正伟, 周步祥, 等. 考虑5G基站储能可调度容量的有源配电网协同优化调度方法[J]. 电网技术, 2023, 47(12): 5225-5237.
- Chen Shi, Guo Zhengwei, Zhou Buxiang, et al. Coordinated optimization scheduling for active distribution networks considering schedulable capacity of energy storage for 5G base stations[J]. Power System Technology, 2023, 47(12): 5225-5237.
- [7] 陈实, 唐国登, 刘艺洪, 等. 基于条件风险强化学习的梯级水光蓄联合优化调度[J]. 电力建设, 2025, 46(9): 84-97.
- Chen Shi, Tang Guodeng, Liu Yihong, et al. Condition-based risk reinforcement learning for joint optimal scheduling of cascade hydropower and solar-powered reservoirs[J]. Electric Power Construction, 2025, 46(9): 84-97.
- [8] 孙惠娟, 段伟男, 陈俐, 等. 基于乐观行动-评判深度强化学习的含氢综合能源系统低碳经济调度[J]. 电网技术, 2024, 48(5): 1873-1883.
- Sun Huijuan, Duan Weinan, Chen Li, et al. Low-carbon economic scheduling of hydrogen integrated energy system based on optimistic actor-critic deep reinforcement learning[J]. Power System Technology, 2024, 48(5): 1873-1883.
- [9] Liu J J, Meng X Y, Wu J T. Data-driven optimal scheduling for integrated electricity-heat-gas-hydrogen energy system considering demand-side management: a deep reinforcement learning approach[J]. International Journal of Hydrogen Energy, 2025, 103: 147-165.
- [10] Gou H X, Ma C, Liu L. Optimal wind and solar sizing in a novel hybrid power system incorporating concentrating solar power and considering ultra-high voltage transmission[J]. Journal of Cleaner Production, 2024, 470: 143361.
- [11] 苏志军, 夏世威, 李庚银, 等. 考虑光热电站接入的电力系统暂态功角稳定性分析[J]. 高压电器, 2025, 61(5): 228-236.
- Su Zhijun, Xia Shiwei, Li Gengyin, et al. Transient power angle stability analysis of power system considering the access of concentrated solar power plant[J]. High Voltage Apparatus, 2025, 61(5): 228-236.
- [12] 周灿, 周育才, 谭艳祥, 等. 基于聚类和RIME-LightGBM的光热电站太阳直接法向辐射预测[J]. 电力科学与技术学报, 2025, 40(6): 241-249.
- Zhou Can, Zhou Yucai, Tan Yanxiang, et al. Prediction of direct normal irradiance from photovoltaic power plants based on clustering and RIME-LightGBM[J]. Journal of Electric Power Science and Technology, 2025, 40(6): 241-249.
- [13] 慈俊昌. 熔融盐储热技术在光热发电领域的工程应用进展[J]. 南方能源建设, 2025, 12(5): 85-99.
- Ci Junchang. Progress in the engineering application of molten salt thermal storage technology in the field of solar thermal power generation[J]. Southern Energy Construction, 2025, 12(5): 85-99.
- [14] 邹博, 任建地, 许道明, 等. 氯化物熔盐储热技术应用于新能源发电的研究进展[J]. 发电技术, 2025, 46(5): 872-884.
- Zou Bo, Ren Jiandi, Xu Daoming, et al. Recent developments on the application of chloride molten salt heat storage technology to new energy power generation[J]. Power Generation Technology, 2025, 46(5): 872-884.
- [15] 龙潇, 张晋宾, 陈令特. 未来能源技术展望[J]. 发电技术, 2025, 46(4): 651-693.
- Long Xiao, Zhang Jinbin, Chen Lingte. Prospects for future energy technologies[J]. Power Generation Technology, 2025, 46(4): 651-693.
- [16] 孙丽平, 宋子恒, 马骏, 等. 中国太阳能光热发电技术研究综述[J]. 南方能源建设, 2025, 12(5): 70-84.
- Sun Liping, Song Ziheng, Ma Jun, et al. Research overview of solar thermal power technology in China[J]. Southern Energy Construction, 2025, 12(5): 70-84.
- [17] 杨宏基, 周明, 武昭原, 等. 含光热电站的电-热能源系统优化运行机制[J]. 电网技术, 2022, 46(1): 175-184.
- Yang Hongji, Zhou Ming, Wu Zhaoyuan, et al. Optimal operation of electro-thermal energy systems with concentrated solar power plant[J]. Power System Technology, 2022, 46(1): 175-184.
- [18] 崔杨, 于世鹏, 张节潭, 等. 考虑光热电站调峰补偿的高比例新能源电力系统经济调度[J]. 中国电机工程学报, 2023, 43(13): 4922-4934.

- Cui Yang, Yu Shipeng, Zhang Jietan, et al. Economic dispatch of high-proportion renewable energy power system considering peak-shaving compensation of concentrating solar power plant [J]. Proceedings of the CSEE, 2023, 43(13): 4922-4934.
- [19] Du E S, Zhang N, Hodge B M, et al. The role of concentrating solar power toward high renewable energy penetrated power systems [J]. IEEE Transactions on Power Systems, 2018, 33(6): 6630-6641.
- [20] Xiao Y L, Zou C Z, Dong M Q, et al. Feasibility study: economic and technical analysis of optimal configuration and operation of a hybrid CSP/PV/wind power cogeneration system with energy storage [J]. Renewable Energy, 2024, 225: 120273.
- [21] Liu H T, Zhai R R, Patchigolla K, et al. Multi-objective optimisation of a thermal-storage PV-CSP-wind hybrid power system in three operation modes[J]. Energy, 2023, 284: 129255.
- [22] Hu B J, Cai F L, Tai N L, et al. Dual-time scale optimal dispatch of the CSP-PV hybrid power plant considering dynamic operation [J]. Energy, 2024, 306: 132488.
- [23] Duan J D, Gao Q, Xia Y R, et al. MMD-DRO based economic dispatching considering flexible reserve provision from concentrated solar power plant[J]. Energy, 2024, 308: 132875.
- [24] 刘佳佳, 巨星. 弃电储能光伏-光热复合发电系统技术经济性分析[J]. 发电技术, 2025, 46(4): 807-817.
- Liu Jiajia, Ju Xing. Techno-economic analysis of photovoltaic-concentrating solar power hybrid system for thermal energy storage of electricity curtailment[J]. Power Generation Technology, 2025, 46(4): 807-817.
- [25] 左超文, 王樊云, 陈洁. 考虑需求响应的光热电站热电联供型微网分层优化调度[J]. 分布式能源, 2024, 9(2): 63-73.
- Zuo Chaowen, Wang Fanyun, Chen Jie. Hierarchical optimization scheduling of combined heat and power microgrid for photothermal power station considering demand response [J]. Distributed Energy, 2024, 9(2): 63-73.
- [26] 解佗, 孙丹阳, 张刚, 等. 计及风光不确定性的风-光-光热联合发电系统中光热电站储热容量优化配置[J]. 智慧电力, 2024, 52(2): 32-39.
- Xie Tuo, Sun Danyang, Zhang Gang, et al. Optimal capacity configuration of thermal storage within CSP plant in wind-PV-CSP hybrid power generation system considering uncertainty of wind and photovoltaic power[J]. Smart Power, 2024, 52(2): 32-39.
- [27] 王辉, 李乃慧, 梁凌, 等. 耦合系统下含光热电站的需求响应虚拟电厂优化调度[J]. 电测与仪表, 2025, 62(12): 30-40.
- Wang Hui, Li Naihui, Liang Ling, et al. Optimal scheduling of virtual power plant with concentrating solar power based on demand response in coupled system [J]. Electrical Measurement & Instrumentation, 2025, 62(12): 30-40.
- [28] 郑连华, 文中, 邱智武, 等. 计及光热电站和氢储能的综合能源系统低碳优化运行[J]. 电测与仪表, 2025, 62(2): 9-15, 34.
- Zheng Lianhua, Wen Zhong, Qiu Zhiwu, et al. Low-carbon optimized operation of integrated energy system considering solar thermal power plants and hydrogen storage [J]. Electrical Measurement & Instrumentation, 2025, 62(2): 9-15, 34.
- [29] 刘翔宇, 张礼浩, 顾雪平, 等. 计及调峰调频安全约束的低惯量电网日前调度模型[J]. 中国电机工程学报, 2025, 45(21): 8390-8403.
- Liu Xiangyu, Zhang Lihao, Gu Xueping, et al. Day ahead dispatching model of low-inertia power grid considering security constraints of peak shaving and frequency regulation [J]. Proceedings of the CSEE, 2025, 45(21): 8390-8403.
- [30] Chen S S, Gao C W, Chen M, et al. An optimal demand response strategy for communication base stations with installing nanoenergy storage system [J]. Nanoscience and Nanotechnology Letters, 2020, 12(6): 841-848.
- [31] 张有兵, 林一航, 黄冠弘, 等. 深度强化学习在微电网系统调控中的应用综述[J]. 电网技术, 2023(7): 2774-2787.
- Zhang Youbing, Lin Yihang, Huang Guanhong, et al. Review on applications of deep reinforcement learning in regulation of microgrid systems[J]. Power System Technology, 2023(7): 2774-2787.
- [32] Duan J J, Shi D, Diao R S, et al. Deep-reinforcement-learning-based autonomous voltage control for power grid operations [J]. IEEE Transactions on Power Systems, 2020, 35(1): 814-817.
- [33] 陈实, 朱亚斌, 刘艺洪, 等. 基于世界模型深度强化学习的含风电电力系统低碳经济调度[J]. 电网技术, 2024, 48(8): 3143-3154.
- Chen Shi, Zhu Yabin, Liu Yihong, et al. Low-carbon economic dispatch of wind-containing power systems based on world model deep reinforcement learning[J]. Power System Technology, 2024, 48(8): 3143-3154.
- [34] Qiu Y, Zhou S Y, Xia D, et al. Local integrated energy system operational optimization considering multi-type uncertainties: a reinforcement learning approach based on improved TD3 algorithm [J]. IET Renewable Power Generation, 2023, 17(9): 2236-2256.
- [35] Wang C, Zhang J H, Wang A Q, et al. Prioritized sum-tree experience replay TD3 DRL-based online energy management of a residential microgrid[J]. Applied Energy, 2024, 368: 123471.
- [36] Hou S R, Fu A H, Duque E M S, et al. DistFlow safe reinforcement learning algorithm for voltage magnitude regulation in distribution networks [J]. Journal of Modern Power Systems and Clean Energy, 2025, 13(1): 300-311.
- [37] GB 38755—2019 电力系统安全稳定导则[S].
- GB 38755—2019 Code on security and stability for power system[S].
- [38] GB/T 40595—2021 并网电源一次调频技术规定及试验导则[S].
- GB/T 40595—2021 Guide for technology and test on primary frequency control of grid-connected power resource[S].
- [39] GB/T 40596—2021 电力系统自动低频减负荷技术规定[S].
- GB/T 40596—2021 Technical rules for automatic under-frequency load shedding in electric power systems[S].
- [40] Chávez H, Baldick R, Sharma S. Governor rate-constrained OPF for primary frequency control adequacy [J]. IEEE Transactions on Power Systems, 2014, 29(3): 1473-1480.
- [41] 胡梦凯, 张高航, 李凤婷, 等. 考虑频率安全及调峰需求的风电高渗透系统两阶段优化调度方法[J]. 电网技术, 2025, 49(10): 4365-4375.
- Hu Mengkai, Zhang Gaoang, Li Fengting, et al. A two-stage

optimal scheduling method for high penetration level of wind power system considering frequency security and peaking demand [J].
Power System Technology, 2025, 49(10): 4365-4375.

收稿日期: 2026-02-02 修回日期: 2026-03-02

作者简介:



陈实

陈实(1977),男,博士,副教授,博士生导师,研究方向为电力系统信息化及智能化、优化控制运行;

晏红平(2000),男,硕士研究生,研究方向为电力系统信息化及智能化;

臧天磊(1986),男,博士,副教授,博士生导师

师,通信作者,研究方向为多能流系统能量管理与优化控制、能源信息物理系统,E-mail:zangtianlei@126.com;

刘艺洪(1997),男,博士研究生,研究方向为电力系统信息化及智能化、优化控制运行;

陈江平(2000),男,硕士研究生,研究方向为电力系统信息化及智能化;

王舒灏(2000),男,硕士研究生,研究方向为电力系统信息化及智能化;

李华强(1965),男,博士,教授,博士生导师,主要研究方向为电压稳定和控制、电力市场与技术经济等。

(编辑 张小飞)

附录 A

1) 功率平衡约束。

系统需要保证运行时都满足负荷需求:

$$\begin{cases} P_{PV,i}^t + P_{GEN,i}^t + P_{CSP,i}^t = P_{L,i}^t + \\ U_i^t \sum_{j \in i} U_j^t (G_{ij} \cos \theta_{ij}^t + B_{ij} \sin \theta_{ij}^t) \\ Q_{GEN,i}^t + Q_{PV}^t = Q_{L,i}^t + \\ U_i^t \sum_{j \in i} U_j^t (G_{ij} \sin \theta_{ij}^t - B_{ij} \cos \theta_{ij}^t) \end{cases} \quad (A1)$$

式中: $P_{PV,i}^t$ 、 $P_{GEN,i}^t$ 、 $P_{CSP,i}^t$ 分别为光伏、火电、光热发电机组在时刻 t 注入节点 i 的有功功率; $Q_{GEN,i}^t$ 、 Q_{PV}^t 分别为火电、光伏发电机组在时刻 t 注入节点 i 的无功功率; $P_{L,i}^t$ 、 $Q_{L,i}^t$ 分别为节点 i 在时刻 t 的负荷有功、无功功率需求; θ_{ij}^t 为时刻 t 两节点间的相角差; G_{ij} 、 B_{ij} 为两节点间的电导和电纳; U_i^t 、 U_j^t 分别为节点 i 、 j 在时刻 t 的电压。

2) 系统机组运行约束: 包括出力上下限约束和爬坡约束。

$$\begin{cases} P_G^{g,\min} \leq P_G^{g,t} \leq P_G^{g,\max} \\ P_{CSP}^{c,\min} \leq P_{CSP}^{c,t} \leq P_{CSP}^{c,\max} \\ 0 \leq P_{PV}^t \leq P_{PV}^{\max} \\ 0 \leq P_{EH}^t \leq P_{EH}^{\max} \end{cases} \quad (A2)$$

式中: $P_G^{g,\max}$ 、 $P_G^{g,\min}$ 分别为火电发电机组有功功率上、下限; $P_{CSP}^{c,\max}$ 、 $P_{CSP}^{c,\min}$ 分别为光热发电机组有功功率上、下限; $P_{PV}^{\max}$ 、 $P_{EH}^{\max}$ 为光伏机组、电加热器有功功率上限。

3) 储热模块充放热状态约束。

TES 系统是 CSP 区别于其他可再生能源的关键组成部分:

$$Q_{TES}^t = (1 - \eta_{TES})Q_{TES}^{t-1} + Q_{TES, ch}^t \eta_{ch} - Q_{TES, dis}^t + Q_{EH}^t \quad (A3)$$

$$Q_{TES}^t = Q_{TES}^{t=0} \quad (A4)$$

$$Q_{TES, ch}^t \cdot Q_{TES, dis}^t = 0 \quad (A5)$$

式中: Q_{TES}^t 为 t 时刻储热罐储存的热能; η_{TES} 表示储热损失系数; η_{ch} 表示储热罐储热效率; $Q_{TES, ch}^t$ 为 t 时刻储热罐吸收的热能; $Q_{TES}^{t=0}$ 、 $Q_{TES}^{t=T}$ 分别为调度日内初始时

刻和最后时刻的储热系统储热量。

$$S_{oc}^{\min} \leq S_{oc}^t \leq S_{oc}^{\max} \quad (A6)$$

$$S_{oc}^t = \frac{Q_{TES}^t}{Q_{TES}^{\max}} \quad (A7)$$

式中: $S_{oc}^{\min}$ 、 $S_{oc}^{\max}$ 分别为储热罐容量状态上、下限; $Q_{TES}^{\max}$ 为储热罐的额定容量。

4) 系统旋转备用功率约束。

当新能源在峰值负荷向下波动时,光热发电可以利用更多的热能来增加功率,实现提供正旋转储备。当新能源在谷负荷处向上波动时,光热发电通过主动控制降低汽轮机的输出,可用热能储存在 TES 中,以备后期使用,实现提供负旋转储备。本文考虑 CSP 和火电机组联合提供旋转储备。

$$\begin{cases} R_{G, up}^{g,t} + R_{CSP, up}^{c,t} \geq R_{up}^{t, \min} \\ R_{G, down}^{g,t} + R_{CSP, down}^{c,t} \geq R_{down}^{t, \min} \end{cases} \quad (A8)$$

式中: $R_{down}^{t, \min}$ 、 $R_{up}^{t, \min}$ 分别为系统的正负旋转储备的最低要求。

$$\begin{cases} 0 \leq R_{CSP, up}^{c,t} \leq \min(P_{CSP}^{c, \max} - P_{CSP}^{c,t}, r_{CSP}^{c, up}) \\ 0 \leq R_{CSP, down}^{c,t} \leq \min(P_{CSP}^{c,t} - P_{CSP}^{c, \min}, r_{CSP}^{c, down}) \\ Q_{TES}^{c,t} - \sum_{t=1}^{N_t} \frac{R_{CSP, up}^{c,t}}{\eta_{PB}} \geq Q_{TES}^{c, \min} \\ Q_{TES}^{c,t} + \sum_{t=1}^{N_t} \frac{R_{CSP, down}^{c,t}}{\eta_{PB}} + \sum_{t=1}^{N_t} P_{EH}^t \cdot \eta_{EH} \leq Q_{TES}^{c, \max} \end{cases} \quad (A9)$$

式中: $r_{CSP}^{c, up}$ 、 $r_{CSP}^{c, down}$ 分别为第 c 台 CSP 机组上、下坡速率上限。

$$\begin{cases} P_G^{g,t} + R_{G, up}^{g,t} - (P_G^{g,t-1} - R_{G, down}^{g,t-1}) \leq r_G^{g, up} \\ P_G^{g,t} - R_{G, down}^{g,t} - (P_G^{g,t-1} + R_{G, up}^{g,t-1}) \geq -r_G^{g, down} \\ 0 \leq R_{G, up}^{g,t} \leq \min(P_G^{g, \max} - P_G^{g,t}, r_G^{g, up}) \\ 0 \leq R_{G, down}^{g,t} \leq \min(P_G^{g,t} - P_G^{g, \min}, r_G^{g, down}) \end{cases} \quad (A10)$$

式中: $r_G^{g, up}$ 、 $r_G^{g, down}$ 分别为第 g 台常规火电机组上、下坡速率上限。

附录 B

表 B1 常规发电机组关键参数
Table B1 Parameters of conventional generator units

发电机(位置)	最小出力/MW	最大出力/MW	爬坡限制/ (MW·h ⁻¹)	一次调频爬坡率/ (MW·s ⁻¹)	等效惯性常数/s	发电成本/ (元/MWh)
G2(Bus1)	75	120	30	2.2	6	144
G5(Bus4)	50	80	20	1.8	4	126
G11(Bus10)	25	50	15	1.2	5.1	162
G13(Bus12)	12	30	10	1.2	5.5	108

表 B2 新能源联合电站设备参数
Table B2 Parameters of hybrid renewable energy power plant equipment

设备类型	参数	数值
储热系统(TES)	最大储热容量/MWh	3500
	最大充热功率/MW	900
	最大放热功率/MW	250
	SOC 上限	0.95
	SOC 下限	0.10
	初始 SOC	0.35
	自放热损失率	0.031
光热发电(CSP)	额定发电功率/MW	100
	光-热-电转换效率	0.38
	爬坡限制/(MW·h ⁻¹)	80
	一次调频爬坡率/(MW·s ⁻¹)	5.0
	运维成本/(元/MWh)	36
	等效惯性常数/s	4.0
光伏发电(PV)	额定装机容量/MW	200
	运维成本/(元/MWh)	28.8
电加热器(EH)	额定功率/MW	120
	电-热转换效率	0.95
频率安全参数	一次调频评估窗口/s	7.5

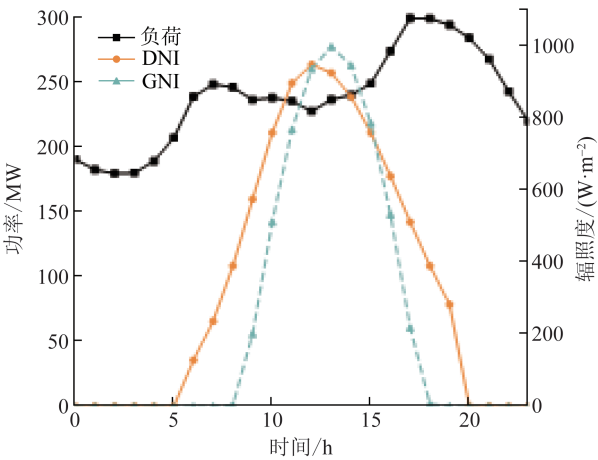


图 B1 负荷曲线与太阳辐照度预测数据

Fig. B1 Load curve and predicted data of solar irradiance

表 B3 LSTM-PER-TD3算法网络结构
Table B3 Network structure of the LSTM-PER-TD3 algorithm

网络名称	网络结构	神经元维度
共享记忆网络	输入层	15(状态维度)
	LSTM层+ReLu	128
Actor网络	全连接层+ReLu	256
	全连接层+ReLu	256
	全连接层+Tanh(输出)	11(动作维度)
Critic网络	(记忆网络输出,动作)拼接输入	128+11
	全连接层+ReLu	256
	全连接层+ReLu	256
	全连接层(输出)	1(Q值)

表 B4 改进 TD3算法超参数设置
Table B4 Key Hyperparameter settings of improved TD3 algorithm

网络	学习率	折扣因子	软更新率	经验回放池容量	训练总回合数
Actor	0.000 1	0.99	0.005	300 000	1500
Critic	0.001 0	0.99	0.005	300 000	1500

附录 C

1)LSTM输入序列长度选取依据。

算法性能进行了对比测试,测试结果如附表 C1

选取 2 h、4 h、6 h 和 8 h 四种不同序列长度对

所示。

表 C1 不同 LSTM 输入序列长度下的算法性能对比
Table C1 Algorithm performance comparison under different LSTM input sequence lengths

LSTM输入序列长度/h	收敛所需平均回合数	收敛后系统总运行成本/万元	单步决策平均耗时/ms
2	1450	102.451	14.2
4	1120	98.033	17.8
6	1380	99.127	25.5
8	1800	104.680	34.6

由表 C1 可知,当输入序列较短时,智能体难以完整提取储热系统长周期的能量时移特征,决策存在“短视”现象,导致系统运行成本偏高。反之,当输入序列过长时,庞大的状态空间不仅增加了单步计算耗时,其引入的远期冗余噪声更严重干扰了网络梯度的稳定更新,导致收敛回合数剧增甚至陷入次优策略。实验结果表明,选取 4 h 能够在时序感知完整性与计算效率之间取得最优平衡。

2) 优先经验回放关键超参数整定分析。

为确定最佳参数组合,本文在保持其他参数不变的条件下,对 α 与 β 组成的 16 组参数网格进行了交叉验证,将各组参数收敛后的系统总运行成本绘制成参数敏感性热力图,如附图 C1 所示。

测试结果表明:当 α 过小,趋近于均匀采样;当 α 过大,导致过度拟合少数极端样本时,算法均无法获得理想的调度策略;同时, β 值需与 α 良好匹配以消除梯度偏差。实验结果表明,当设定 $\alpha=0.6,\beta=0.4$ 时,智能体表现出最优的综合寻优性能。

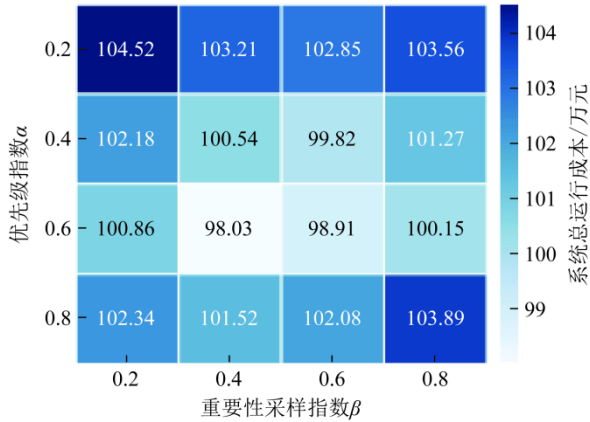


图 C1 不同 α 与 β 组合下的总运行成本参数敏感性热力图

Fig. C1 Parameter sensitivity heatmap of total operation cost under different combinations of α and β

附录 D

表 D1 常规发电机组关键参数 (IEEE-57 节点测试系统)

Table D1 Parameters of conventional generator units (IEEE 57-bus test system)

发电机编号	最小出力/MW	最大出力/MW	爬坡限制/ (MW·h ⁻¹)	一次调频爬坡率/ (MW·s ⁻¹)	等效惯性常数/s	发电成本/ (元/MWh)
G0	50	100	25	1.2	5.0	144
G1	70	140	40	2.2	6.0	158
G2	50	100	25	1.2	5.0	144
G3	50	100	30	1.8	5.5	180
G4	210	410	100	2.5	6.5	166

表 D2 新能源联合电站设备参数 (IEEE-57 节点测试系统)

Table D2 Parameters of hybrid renewable energy power plant equipment (IEEE 57-bus test system)

设备类型	参数	数值
储热系统 (TES)	最大储热容量/MWh	7000
	最大充热功率/MW	1500
	最大放热功率/MW	500
	SOC 上限	0.95
	SOC 下限	0.10
	初始 SOC	0.50
	小时自放热损失率	0.031
光热发电 (CSP)	额定发电功率/MW	200
	光-热-电转换效率	0.38
	小时爬坡限制/(MW·h ⁻¹)	80
	一次调频爬坡率/(MW·s ⁻¹)	5.0
	运维成本/(元/MWh)	36
	等效惯性常数/s	4.0
光伏发电 (PV)	额定装机容量/MW	400
	运维成本/(元/MWh)	28.8
电加热器 (EH)	额定功率/MW	250
	电-热转换效率	0.95
频率安全参数	一次调频评估窗口/s	7.5

表 D3 LSTM-PER-TD3算法网络结构(IEEE-57节点测试系统)
Table D3 Network structure of the LSTM-PER-TD3 algorithm (IEEE 57-bus test system)

网络名称	网络结构	神经元维度
共享记忆网络	输入层	15
	LSTM层+ReLU	256
Actor网络	全连接层+ReLU	512
	全连接层+ReLU	512
	全连接层+Tanh(输出)	11
Critic网络	(记忆网络输出,动作)拼接输入	256+11
	全连接层+ReLU	512
	全连接层+ReLU	512
	全连接层(输出)	1

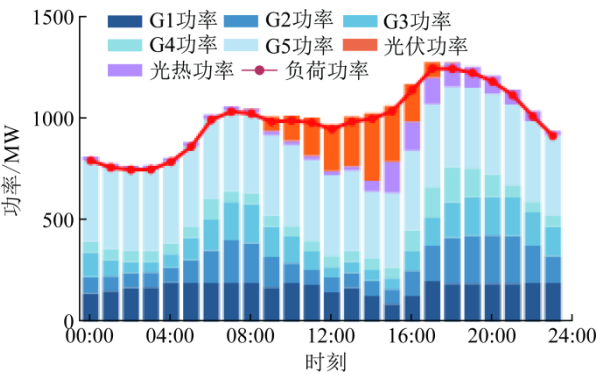


图 D1 典型日日前调度优化结果
Fig. D1 Optimal day-ahead scheduling results for a typical day

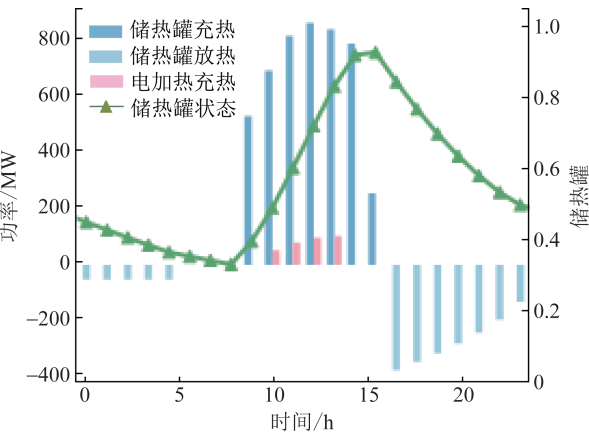


图 D2 储热罐充放热功率以及对应的SOC变化曲线
Fig. D2 Charging/discharging power and SOC variation curve of thermal storage tank

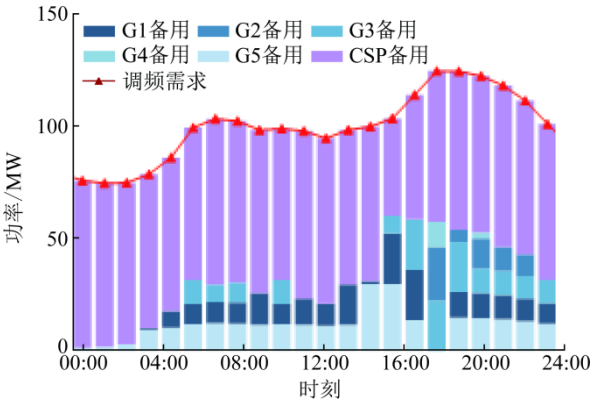


图 D3 一次调频备用提供计划
Fig. D3 Primary frequency regulation reserve provision schedule

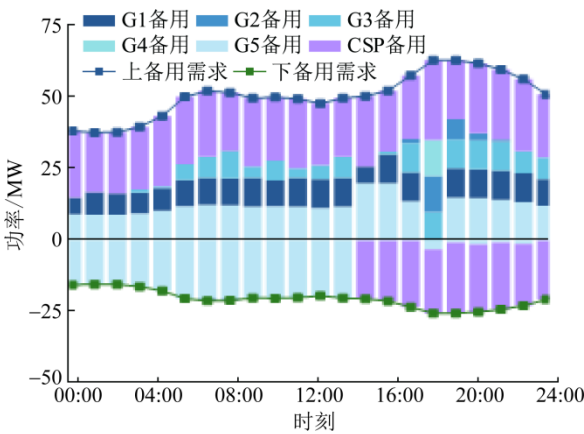


图 D4 上下旋转备用提供计划
Fig. D4 Up/down spinning reserve provision schedule

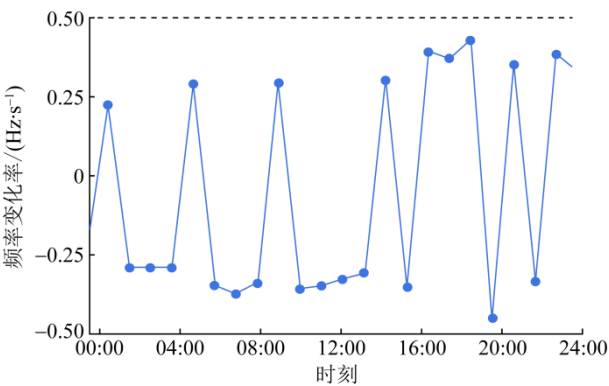


图 D5 系统频率变化率的日内动态曲线
Fig. D5 Intra-day dynamic profile of the system rate of change of frequency