

DOI: 10.11949/0438-1157.

基于深度强化学习的绿氢生产过程调度优化

刘京鑫, 赵亮

(华东理工大学信息科学与工程学院, 上海 200237)

摘要: 针对风光发电波动导致电解槽负载频繁变化、弃电与购电并存等问题, 本文建立了光伏—风电—储能—电解槽耦合制氢过程模型, 并将在线调度表述为马尔可夫决策过程; 在多目标奖励中考虑产氢收益、购电成本、弃电抑制、负载平滑及储能荷电状态 (SOC) 安全约束, 采用软演员—评论家 (SAC) 算法获得最优调度策略。基于 NASA 气象数据与电价序列的算例表明, 所提方法在多场景下比专家规则与模型预测控制取得更高产氢量与可再生能源利用率, 并显著降低弃电, 且单步推理开销可满足 5 min 调度周期的实时应用需求。

关键词: 绿氢生产; 深度强化学习; 多目标优化; 储能系统; 可再生能源

中图分类号: TP 273.1

文献标志码: A

文章编号: 0438-1157 (XXXX) XX-0001-10

Deep reinforcement learning for scheduling of green hydrogen production process

LIU Jingxin, ZHAO Liang

(College of Information Science and Engineering, East China University of Science and Technology, Shanghai 200237, China)

Abstract: To address the issues of frequent load fluctuations of electrolyzers caused by volatile wind and solar power, as well as the coexistence of renewable energy curtailment and electricity purchase costs, this paper establishes a coupled photovoltaic-wind-energy storage-electrolyzer hydrogen production system model, and formulates the online scheduling problem as a Markov decision process. The multi-objective reward function comprehensively considers hydrogen production benefits, electricity purchase costs, renewable energy curtailment mitigation, load smoothing, and state-of-charge (SOC) security constraints, and the Soft Actor-Critic (SAC) algorithm is adopted to obtain the optimal scheduling strategy. Case studies based on NASA meteorological data and electricity price sequences demonstrate that the proposed method achieves higher hydrogen production and renewable energy utilization efficiency than expert rules and model predictive control under various scenarios, significantly reduces renewable energy curtailment, and the single-step inference cost can meet the real-time application requirements of a 5-minute scheduling cycle.

Keywords: green hydrogen production; deep reinforcement learning; multi-objective optimization; energy storage systems; renewable energy

收稿日期: 2026-03-28 修回日期: 2026-05-06

通信作者: 赵亮 (1979—), 男, 博士, 教授, lzhaol@ecust.edu.cn

第一作者: 刘京鑫 (2000—), 男, 硕士研究生, 1453597365@qq.com

基金项目: 国家自然科学基金面上项目 (62373154, 22178103)

引用本文: 刘京鑫, 赵亮. 基于深度强化学习的绿氢生产过程调度优化[J]. 化工学报, XXXX, XX(XX): 1–10

Citation: LIU Jingxin, ZHAO Liang. Deep reinforcement learning for scheduling of green hydrogen production process[J]. CIESC Journal, XXXX, XX(XX): 1–10

引 言

在“双碳”目标推动下,可再生能源(renewable energy, RE)制氢已成为实现氢能低碳化供给的重要路径^[1]。风电、光伏等绿色电力驱动电解水制氢,具有减少化石能源依赖和降低过程碳排放的潜力,但其大规模应用仍面临成本偏高和系统调度复杂等问题^[2]。尤其在风、光波动和电价时变条件下,如何实现绿氢生产过程的高效、经济与稳定运行,是制约其工程化应用的关键问题^[3]。

可再生能源制氢系统具有多约束耦合特征^[4]。一方面,风光发电存在波动性和间歇性^[5],易引起电解槽负载频繁变化^[6];另一方面,电解槽具有最小负载率约束和非线性效率特性^[7],储能系统又受到储能荷电状态(SOC)边界、充放电功率和效率等条件限制^[8]。因此,绿氢生产过程调度需要在产氢量、购电成本、弃电控制和运行安全之间进行动态权衡,本质上是连续决策、多目标协同的调度优化问题^[9]。

围绕绿氢生产过程的优化运行,已有研究主要采用混合整数规划、随机优化^[10]、鲁棒优化和模型预测控制等方法^[11],在容量配置、滚动优化和不确定性处理等方面取得一定进展^[12]。但是,这类方法通常依赖精确的数学模型,在处理电解槽效率非线性、多设备耦合和高维连续控制问题时仍存在一定局限^[13]。近年来,深度强化学习因具备交互学习和连续控制能力,在综合能源系统调度方面具有较好的应用潜力;但针对RE制氢场景,同时考虑电解槽运行特性、储能安全约束、负载平滑需求和RE消纳目标的统一建模与优化研究相对较少^[14]。

基于此,本文以光伏-风电-储能-电解槽耦合制氢过程为研究对象,建立考虑设备运行约束与风光波动特性的调度模型,并将其表述为马尔可夫决策过程;构造兼顾产氢收益、购电成本、弃电抑制、负载平滑和SOC安全的多目标奖励函数,采用软演员-评论家(SAC)算法求解调度策略;利用可再生能源和电价等数据开展多场景对比分析,验证所提方法在制氢性能、经济性和RE利用等方面的有效性。

1 系统建模与调度问题描述

1.1 系统结构与关键单元建模

绿氢生产过程由光伏发电、风力发电、电池储能、碱性电解槽等单元构成,系统结构如图1所示。可再生能源发电优先用于驱动电解槽制氢;当风光

发电超过电解槽实时用电需求时,剩余电能优先用于储能充电,若储能达到充电上限,则形成弃电;当风光发电不足时,首先调用储能放电,若储能仍无法满足功率需求再从电网购电,以保证电解槽运行的连续性和氢产量的稳定性。



图1 可再生能源-储能-电解槽耦合制氢系统结构示意图

Fig.1 Schematic diagram of the renewable energy-storage-electrolyzer coupled hydrogen production system

在时段 t 内,系统功率平衡关系定义如下:

$$P_t^{pv} + P_t^w + P_t^g + P_t^{dis} = P_t^{el} + P_t^{ch} + P_t^{cur} \quad (1)$$

式中, P_t^{pv} 和 P_t^w 分别为光伏与风电功率, P_t^g 为电网购电功率, P_t^{dis} 和 P_t^{ch} 分别为储能放电功率与充电功率, P_t^{el} 为电解槽运行功率, P_t^{cur} 为弃电功率。

电解水制氢单元为绿氢生产的核心^[15],本文选用工程应用成熟的碱性电解槽^[16]。设电解槽额定功率为 $P^{elec,max}$ ^[17],时段 t 的运行负载率为 $L(t)$,则电解槽的实际功率为:

$$P^{elec}(t) = L(t)P^{elec,max} \quad (2)$$

式中, $L(t)$ 为电解槽负载率。由于碱性电解槽存在最小稳定运行负载率,为避免低负载下电极极化加剧和能效下降,设置最小负载率约束为^[18]:

$$L(t) \geq L^{min} \quad (3)$$

式中, L^{min} 为最小负载率^[19]。

满足最小稳定负载率约束的前提下,时段 t 的产氢速率 $H(t)$ 与电解槽功率关系如下^[20]:

$$H(t) = \frac{P^{elec}(t)}{E^{H_2}} \quad (4)$$

式中, E^{H_2} 为制氢能耗。

储能系统用来平衡短时功率,风光富余时充电以降低弃电,风光不足时放电以减少购电并维持电解槽稳定运行,直接影响RE利用率与系统经济性^[21]。设储能系统额定容量为 E^{bess} ,时段 t 的荷电状态为 $SOC(t)$,储能功率为 $P^{bess}(t)$ 。其中,设定 $P^{bess}(t) > 0$ 表示充电, $P^{bess}(t) < 0$ 表示放电,储能系统的荷电状态更新方程定义如下:

$$SOC(t+1) = SOC(t) + \frac{\Delta t}{E^{bess}} \begin{cases} \eta^{ch} P^{bess}(t) & P^{bess}(t) > 0 \\ \eta^{dis} P^{bess}(t) & P^{bess}(t) < 0 \end{cases} \quad (5)$$

式中, η^{ch} 和 η^{dis} 分别表示储能系统的充电效率和放电效率, Δt 为调度周期。

为保证储能系统安全运行并延缓电池衰减, 储能荷电状态需满足上下界约束:

$$SOC^{\min} \leq SOC(t) \leq SOC^{\max} \quad (6)$$

式中, $SOC^{\min}$ 和 $SOC^{\max}$ 分别为 SOC 下界和上界。

光伏和风力发电作为系统的外部输入, 其时序变化决定了 RE 供给水平。设时段 t 的太阳短波辐照度为 $G(t)$, 光伏综合转换效率为 η^p , 则光伏输出功率 P_t^p 可表示为:

$$P_t^p = \eta^p G(t) \quad (7)$$

设时段 t 的风速为 v^t , 风机切入风速、额定风速和切出风速分别为 v^{in} 、 v^r 和 v^{out} , 风电额定功率为 $P^{w, \max}$, 时段 t 的风电输出功率 P_t^w 定义为:

$$P_t^w = \begin{cases} 0 & v_t < v^{in} \\ P^{w, \max} \frac{v_t - v^{in}}{v^r - v^{in}} & v^{in} \leq v_t \leq v^r \\ P^{w, \max} & v^r < v_t \leq v^{out} \\ 0 & v^{out} < v_t \end{cases} \quad (8)$$

式(7)和式(8)描述了 RE 发电的变化规律; 电价影响系统在不同时间段的购电策略与运行经济性。因此, 绿氢生产过程调度并非静态的功率分配问题, 而是在风光波动、电价时变以及设备运行约束下, 实时协调电解槽负载、储能充放电及电网购电的动态优化问题。

1.2 调度目标与约束描述

在风光发电和电价随时间变化的条件下, 系统调度需要同时处理产氢量、购电成本、弃电水平和运行约束之间的关系。调度问题的输入、决策、输出关系如图2所示。

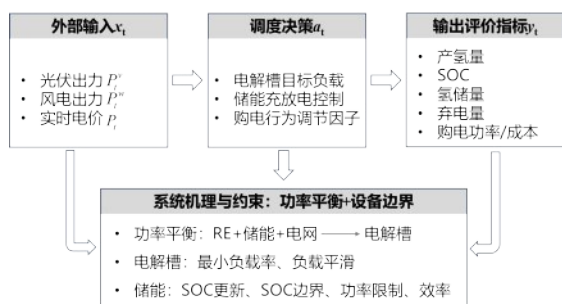


图2 调度优化问题的输入-决策-输出示意图

Fig.2 Input-decision-output diagram of the scheduling optimization problem

图2中, 提高电解槽负载有利于增加产氢, 但可能增加购电成本并带来更大的负载波动; 储能单元运行受荷电状态、充放电功率和效率限制。因此, 将该问题定义为在满足功率平衡和设备边界条件

下, 对电解槽负载、储能充放电和电网购电进行协同调节的动态优化问题, 决策变量包括电解槽负载率、储能充放电量和购电调节因子。

需要说明的是, 本文目的是验证深度强化学习求解绿氢生产过程调度问题的有效性, 采用面向调度决策的等效建模方法, 模型可以描述风光波动、电价时变、电解槽负载约束等对调度结果的影响, 反映制氢过程在调度层面的运行特性。模型未考虑电解槽启停损耗、设备寿命衰减、氢储运过程、预测误差等因素, 研究结果更适用于方法验证与调度方案分析, 而不适用于工业装置的操作优化。

1.3 马尔可夫决策过程建模

为描述风光-储能-电解制氢系统的连续调度过程, 将其表述为马尔可夫决策过程 $M=(S, A, P, R)^{[22]}$ 。其中, 状态 S 表示当前时段的系统运行信息, 动作 A 表示调度指令, 状态转移 P 反映系统在环境输入和控制作用下向下一时刻的演化过程, 奖励 R 用于评价当前决策的运行效果。

状态空间由光伏功率、风电功率、电价、电解槽当前负载率、储能荷电状态等变量构成, 用于表征当前时刻的供能条件和运行状态; 动作空间包括电解槽负载率、储能充放电量和购电调节因子。策略输出动作后, 环境空间结合设备约束和功率平衡关系计算实际执行结果, 并更新下一时段状态。

奖励函数协调长期运行中产氢、成本和约束之间的关系, 定义为:

$$r_t = r^{H_2} - r^{\text{cost}} - r^{\text{cur}} - r^{\text{var}} - r^{\text{SOC}} \quad (9)$$

式中, r^{H_2} 为产氢收益项, r^{cost} 为购电成本项, r^{cur} 为弃电惩罚项, r^{var} 为负载变化惩罚项, r^{SOC} 为储能越界惩罚项, 分别定义为:

$$r^{H_2} = k_h m_t^{H_2} \quad (10)$$

$$r^{\text{cost}} = k_c P_t^e \Delta t \quad (11)$$

$$r^{\text{cur}} = k_r P_t^{\text{cur}} \Delta t \quad (12)$$

$$r^{\text{var}} = k_l |L_t - L_{t-1}| \quad (13)$$

$$r^{\text{SOC}} = k_s [\max(0, SOC_{\min} - SOC_t) + \max(0, SOC_t - SOC_{\max})] \quad (14)$$

式中, $m_t^{H_2}$ 为时段 t 内的产氢量, P_t^e 为时段 t 的电价, P_t^e 为电网购电功率, P_t^{cur} 为弃电功率, L_t 为电解槽负载率, SOC_t 为储能荷电状态; k_h 、 k_c 、 k_r 、 k_l 、 k_s 分别为产氢收益、购电成本、弃电抑制、负载平滑和 SOC 安全约束对应的权重系数。

在满足最小稳定负载率约束的前提下, 时段 t 内的产氢量定义为:

$$m_t^{H_2} = H(t) \Delta t = \frac{P^{\text{elec}}(t)}{E^{H_2}} \Delta t \quad (15)$$

式中, $P^{dec}(t)$ 为时段 t 的电解槽输入功率, E^{H_2} 为单位质量氢气的综合制取能耗。

综合各项奖励, 系统调度问题可转化为连续状态、连续动作条件下的序贯决策问题, 并采用 SAC 算法求解。统计周期 T 内的主要评价指标定义如下:

$$M^{H_2} = \sum_{i=1}^T m_i^{H_2} \quad (16)$$

$$C^{grid} = \sum_{i=1}^T p_i P_i^g \Delta t \quad (17)$$

$$c^{H_2} = \frac{C^{grid}}{M^{H_2}} \quad (18)$$

$$E^{cur} = \sum_{i=1}^T P_i^{cur} \Delta t \quad (19)$$

$$\eta^{RE} = 1 - \frac{E^{cur}}{\sum_{i=1}^T (P_i^{pw} + P_i^w) \Delta t} \quad (20)$$

$$R^{avg} = \frac{1}{T-1} \sum_{i=2}^T |L_i - L_{i-1}| \quad (21)$$

式中, M^{H_2} 为总产氢量, C^{grid} 为总购电成本, c^{H_2} 为单位产氢成本, E^{cur} 为总弃电量, η^{RE} 为可再生能源利用率, R^{avg} 为平均负载爬坡。 R^{avg} 数值越小, 表示电解槽负载变化越平稳。

2 求解方法与实验设置

2.1 SAC 求解框架

针对建立的连续状态、连续动作调度问题, 采用 SAC 算法求解^[23]。SAC 是属于离线策略的强化学习方法^[24], 在优化累计回报的同时引入熵正则项, 能够兼顾策略学习与动作探索, 适用于电解槽负载、储能充放电和购电水平等连续控制变量的联合优化^[25]。

在训练过程中, 策略网络根据当前状态输出动作, 价值评估网络用于估计状态-动作对的价值, 二者通过与环境交互不断迭代更新^[26]。智能体在每个调度周期接收由风光发电、电价、电解槽负载率和储能荷电状态等信息构成的系统状态, 输出电解槽目标负载率、储能充放电控制量和购电调节因子, 环境结合设备约束和功率平衡关系计算实际执行结果, 并返回奖励值和下一时刻状态^[27]。为提高训练稳定性与样本利用效率, 采用经验回放机制、双 Q 网络结构和目标网络软更新策略。训练完成后, 利用策略网络实现系统在不同运行场景下的连续调度决策。训练框架如图 3 所示。

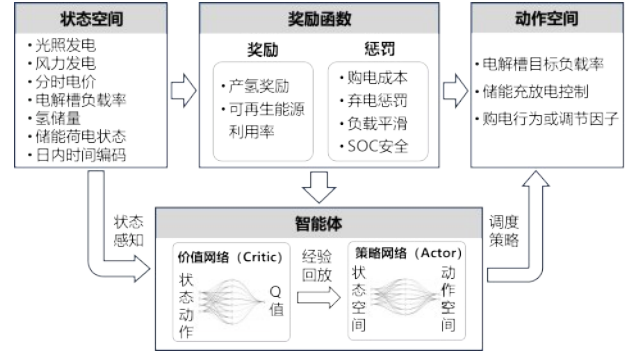


图3 基于SAC的调度策略训练框架

Fig.3 SAC-based Scheduling Policy Training Framework

2.2 实验设置与评价指标

实验中, 电解槽额定功率为 2000 kW, 最小负载率为 0.3; 储能容量为 2000 kWh, 最大充放电功率为 500 kW, SOC 运行范围为 0.3 ~ 0.9, 充放电效率均为 0.9; 调度时间步长设为 5 min。风光发电与电价均以时序形式输入环境, 用于构造滚动调度过程^[28]。其中, 电价序列采用具有日内周期波动特征的时序信号构建: 设基准电价为 0.5 元/kWh, 并引入幅值为 0.3 的周期波动项, 同时叠加随机扰动, 以表征并网运行条件下购电成本的短时变化; 为避免极端值影响训练稳定性, 电价取值进一步限制在 0.3 - 2.0 元/kWh 区间, 以构造具有峰谷特征和随机波动的短时电价环境, 检验调度策略对时变购电成本的适应能力^[29]。

奖励函数参数设定如表 1 所示。各权重系数是在保证奖励分项数量级相对平衡的基础上, 经多组预实验筛选得到。

为对比策略在不同资源条件下的适应性, 构建高光伏、高风电和低 RE 三类典型场景, 并设置多日综合测试集用于整体对比评估, 各场景划分见表 2。各场景采用相同的设备参数配置, 场景间差异仅体现在风光发电和电价输入序列不同, 保证不同场景调度结果的可比性。

选择专家规则(Greedy)和模型预测控制(MPC)与所提 SAC 方法在相同风光发电、电价输入和设备约束条件下进行对比^[30]。Greedy 方法依据当前时刻风光发电、储能状态和设备约束执行局部决策: 优先利用 RE 发电驱动电解槽运行; 当 RE 发电高于电解槽需求时, 剩余电能先储能充电, 超过储能部分记为弃电; 当 RE 发电不足时, 优先调用储能放电, 仍不足时再由电网购电; 电解槽负载依据可用功率在设备约束范围内即时调整。MPC 方法采用滚动

表1 奖励函数各组成项及参数设置依据

Table 1 Basis for the components and parameter settings of the reward function

奖励项	作用	参数/阈值	设置依据
产氢收益项	提升产氢量	$k_h=9.00$	主目标项,保证策略不过度保守
购电成本惩罚项	抑制高价购电	$k_e=1.50$	与电价和购电功率耦合,体现经济性约束
弃电惩罚项	提高RE消纳水平	$k_r=0.05$	防止放弃可用绿电
负载平滑项	限制电解槽负载剧烈波动	$k_f=0.80$	反映设备平稳运行需求
SOC安全惩罚	防止储能过充过放	SOC上下界	对应储能单元物理运行边界
SOC区间激励	保持储能调节裕度	目标区间0.4-0.6	提高后续时段调节灵活性
电价引导项	强化低价用电、高价抑制购电倾向	电价阈值0.3、0.5、0.8元·kWh ⁻¹	体现分时电价运行经济性
运行边界软惩罚	避免过高/过低负载	负载阈值0.85/0.65	抑制极端运行偏好

表2 实验场景与数据集划分

Table 2 Experimental Scenarios and Dataset Partitioning

场景编号	场景名称	特征描述	用途
S1	高光伏场景	光伏占比高、日内峰值明显	典型行为分析
S2	高风电场景	风功率高、波动性强	典型行为分析
S3	低RE场景	风光均偏低、购电压力大	鲁棒性验证

优化思想,在每个调度时刻基于未来预测窗口内的风光发电和电价预测值构建短时优化问题,以购电成本、弃电惩罚和负载波动惩罚的加权和最小为目标,并满足功率平衡、电解槽负载上下限、储能SOC边界及充放电功率等约束;优化求解得到未来窗口内的控制序列后,仅执行首个控制动作,在下一时刻基于更新状态重新求解。本文MPC的滚动步长与调度周期一致,均为5 min;预测时域设为24步,对应2 h短时滚动优化窗口。MPC与SAC采用相同的参数设置。强化学习策略采用“训练—测试”分离方式,各方法在同一测试集上运行,并输出相同评价指标^[31]。

3 结果与分析

3.1 综合性能对比

为评价所提SAC方法在多目标约束下的综合性能,对比其与随机策略(Random)、Greedy与MPC

等方法的性能。评价指标包括总产氢量(式(16))、单位产氢成本(式(18))、RE利用率(式(20))与弃电量(式(19)),多指标对比如图4所示。

由图4可知,SAC方法在多数场景下性能较好。与Greedy方法相比,SAC能够根据风光发电、电价和储能状态的变化动态调整电解槽负载与购电水平,同时提高产氢量和RE利用率;与MPC相比,SAC不依赖精确模型,仍能获得稳定的运行结果,在多场景测试中适应性较好。

3.2 方法稳健性分析

为进一步验证所提SAC方法的稳定性与参数设定合理性,从训练随机性和奖励参数两个方面进行分析,包括奖励参数敏感性分析和多随机种子稳定性验证。

3.2.1 奖励参数敏感性分析 奖励函数参数设定会影响强化学习策略的行为。为分析关键参数对调度结果的影响,选取产氢收益系数、购电成本惩罚系数和负载平滑惩罚系数作为敏感性分析对象,并以基准参数为中心,分别设置0.5倍、1.0倍和1.5倍三个水平,比较不同参数下总产氢量、单位产氢成本、RE利用率、弃电量和平均负载爬坡(式(21))的变化情况,结果如表3所示。

由表3可知,奖励权重变化会使策略在不同优化目标之间产生差异。增大产氢收益系数,策略倾

表3 关键奖励参数敏感性分析结果

Table 3 Sensitivity analysis results of key reward parameters

参数设置	产氢收益系数		购电成本惩罚系数		负载平滑惩罚系数		基准参数
	0.5倍	1.5倍	0.5倍	1.5倍	0.5倍	1.5倍	
总产氢量/kg	3525.6	3621.8	3619.1	3614.9	3606.2	3632.3	3594.2
单位产氢成本/(元·kg ⁻¹)	10.10	10.58	10.46	10.35	10.46	10.51	10.49
RE利用率/%	92.9	93.5	93.8	93.8	93.5	93.7	93.2
弃电量/kWh	14466.0	12866.3	12231.1	12203.5	12745.2	12509.8	14214.8
平均负载爬坡	1.319	0.939	1.210	0.995	1.115	0.942	1.105



(a) 产氢量对比



(b) 单位成本对比



(c) 可再生能源利用率对比



(d) 弃电量对比

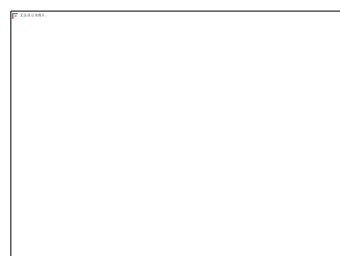
图4 冬季与夏季测试集多指标对比

Fig.4 Multi-Indicator Comparison of Winter and Summer Test Sets

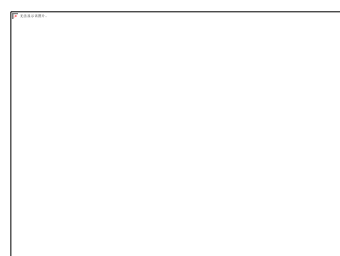
向于维持较高电解槽负载,总产氢量和RE利用率提高,弃电量降低,但单位产氢成本上升;增大购电成本惩罚系数,策略会抑制高价购电行为,单位产氢成本下降;增大负载平滑惩罚系数,平均负载爬坡降低,说明负载调节更加平稳,但可能削弱策略对RE波动的快速响应能力。综合来看,各参数扰动下主要指标未出现剧烈变化,说明所提奖励函数具有一定鲁棒性。

3.2.2 多随机种子稳定性验证 在相同训练集、测试集和超参数设置下,选取3个随机种子独立训练

SAC策略,分别在冬季和夏季测试集上分析产氢量、单位产氢成本、RE利用率和弃电量,结果如图5所示。



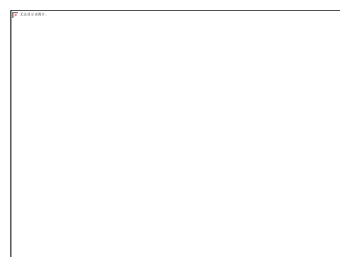
(a) 产氢量对比



(b) 单位成本对比



(c) RE利用率对比



(d) 弃电量对比

图5 SAC多随机种子稳定性验证结果

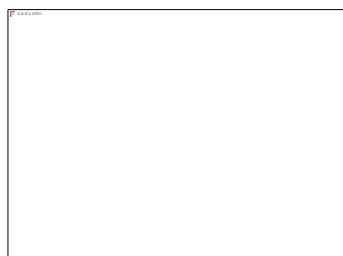
Fig.5 Stability verification results of SAC under multiple random seeds

由图5可知,不同随机种子下各指标变化幅度较小。冬季测试集中,产氢量为 3232.5 ± 9.0 kg、单位产氢成本为 11.894 ± 0.016 元/kg、RE利用率为 $96.04 \pm 0.18\%$ 、弃电量为 5952.2 ± 272.5 kWh;夏季测试集中,对应指标分别为 4024.3 ± 6.2 kg、 9.520 ± 0.031 元/kg、 $91.55 \pm 0.16\%$ 和 $18\ 448.2 \pm 354.6$ kWh。

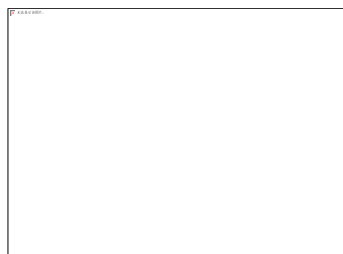
冬季与夏季测试集中,产氢量、单位产氢成本和RE利用率保持较高一致性,变化系数均不超过0.33%;弃电量虽有一定波动,但整体差异较小,冬、夏季变化系数分别为4.58%和1.92%。这表明所提SAC方法对随机初始化不敏感,调度结果具有较好的稳定性和可重复性。

3.3 典型场景调度结果分析

为进一步分析SAC方法在不同资源条件下的调度结果,选取高光伏场景和低RE场景,分别对应RE供给较充足和供给不足两类工况。在高光伏场景下,系统日间光伏发电功率较高,RE供给在中午前后达到峰值,如图6所示。



(a) 可再生能源功率



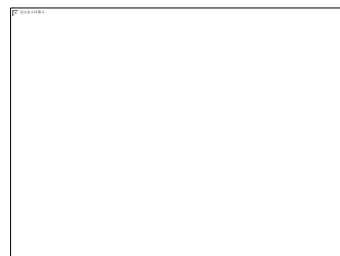
(b) 电网电价



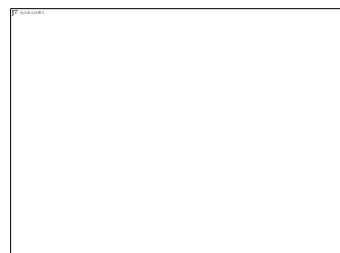
(c) 电解槽负载调节



(d) 储能系统状态



(e) 购电与弃电



(f) 氢气累计产量

图6 高光伏场景下SAC策略调度结果

Fig.6 Scheduling result of SAC under the high-PV scenario

图6中,SAC方法在RE充足时提高电解槽负载率,储能充电以吸收富余电能;随着光伏发电功率降低,储能放电和适量购电,维持电解槽连续运行。该场景下,电解槽负载率较高,储能SOC处于安全范围,购电功率变化平滑,说明该策略能够根据RE发电变化及时调整负载分配与储能状态,提高RE利用率并保持系统运行稳定。

在低RE供给场景下,风光发电偏低,可用绿电明显减少,调度重点由富余能量回收转向有限资源下的供电协调,如图7所示。



(a) 可再生能源功率



(b) 电网电价



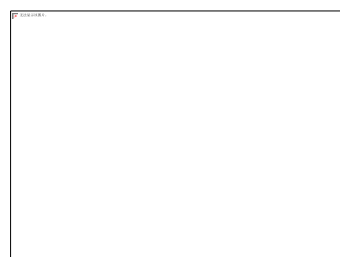
(c) 电解槽负载调节



(d) 储能系统状态



(e) 购电与弃电



(f) 氢气累计产量

图7 低RE场景下SAC策略调度结果

Fig.7 Scheduling result of SAC under the low-RE scenario

由图7可知,SAC方法根据RE供给水平、电价变化和储能状态对负载进行调节,资源不足时降低电解槽负载,结合储能放电与分时购电保障制氢过程连续进行。该场景下风光发电较低,优先消纳RE,弃电量接近于0;储能SOC在约束范围内,未出现快速逼近边界的情况;累计产氢量增长速度低于高光伏场景,但仍能连续产氢,表明SAC在RE不足时能够在产氢需求、购电成本和运行安全之间达到较平衡的调度结果。

3.4 消融与工程效益分析

为进一步说明所提方法中关键设计对调度结果的影响,在相同参数设置下进行消融实验,分析储能和电解槽负载调节对调度结果的作用。消融实验结果如表4所示。

表4 消融实验结果对比

Table 4 Comparison of Ablation Study Results

变体	产氢量/kg	单位成本/(元·kg ⁻¹)	RE利用率/%	弃电量/kWh	平均奖励
无储能系统	1242.9	16.58	96.9	1601.3	9.80
固定负荷	1253.9	16.64	97.5	1294.4	15.34
SAC完整版	1322.1	15.49	97.9	1074.2	15.19

由表4可知,完整版SAC策略在产氢量、RE利用率、弃电控制和运行平稳性之间达到均衡。削弱储能能力,系统对风光波动的缓冲作用明显下降,富余RE难以转移至后续时段,弃电量增加;在资源不足时段更依赖电网补电。当进一步削弱电解槽负载调节能力后,系统对外部扰动的响应灵活性降低,电解槽运行容易偏离有利工况,影响制氢效率和调度效果。

在工程效益方面,比较了不同策略的经济性与环境效益,结果如表5所示。

由表5可知,三种方法在RE参与下均能显著降低电网侧间接排放。SAC方法年产氢量和年碳减排量最高,验证了其在扩大制氢规模方面的优势;SAC

表5 不同策略的碳排放与减排效益对比

Table 5 Comparison of Carbon Emissions and Mitigation Benefits Across Different Strategies

方案	年产氢量/kg	年购电量/ 万kWh	碳强度/(kg CO ₂ ·kg ⁻¹ H ₂)	碳减排量/t CO ₂	减排 率/%
Greedy	100358.9	210.01	12.16	2849.73	70.02
MPC	102904.7	210.32	11.87	2927.62	70.55
SAC	116477.8	258.73	12.91	3128.76	67.55

方法年购电量为258.73万kWh、单位产氢碳强度为12.91 kg CO₂·kg⁻¹ H₂,高于Greedy和MPC,反映了产氢规模与购电间接排放之间的关系。

综合消融结果以及工程效益分析可以看出,本

文构建的状态空间、动作空间和奖励函数能够较好支撑风光—储能—电解水制氢系统的协同调度,所提方法有效平衡了产氢能力、运行经济性和低碳效益。

4 结论

本文针对并网型风光—储能—电解水制氢系统在风光波动和电价时变条件下面临的动态调度问题,提出了基于SAC方法的连续调度方法,建立了包含光伏、风电、储能和电解槽的调度模型,并通过状态空间、动作空间和多目标奖励函数设计,实现了对电解槽负载、储能充放电及电网购电行为的协同调度优化。算例结果表明,所提方法在高光伏、高风电和低RE等不同场景下均能够获得较好的综合调度效果,在产氢能力、购电成本、RE利用率和运行平稳性之间实现较合理的平衡;与Greedy和MPC方法相比,SAC策略表现出更好的场景适应性和连续控制能力,消融实验进一步说明储能协同和电解槽连续负载调节是提升系统综合性能的重要因素。

本文建立的是面向调度优化的等效模型,未考虑电解槽启停损耗、寿命衰减、氢储运过程等因素,未来将进一步考虑更复杂的设备机理、预测误差以及多时间尺度协调等问题,以提高方法在实际工程场景中的适用性。

符号说明

a_t —— 动作向量
 $C(t)$ —— 时段 t 购电成本,元
 C^{grid} —— 总购电成本,元
 c^{H_2} —— 单位产氢成本,元 $\cdot\text{kg}^{-1}$
 E^{bess} —— 储能额定容量, kWh
 E^{sur} —— 总弃电量, kWh
 E^{H_2} —— 单位质量氢气综合制取能耗, $\text{kWh}\cdot\text{kg}^{-1}$
 $G(t)$ —— 时刻 t 的太阳短波辐照度, $\text{kW}\cdot\text{m}^{-2}$
 $H(t)$ —— 时刻 t 的产氢速率, $\text{kg}\cdot\text{h}^{-1}$
 k_b, k_c, k_r, k_p, k_s —— 奖励函数权重系数
 $L(t)$ —— 电解槽负载率
 L^{min} —— 电解槽最小稳定负载率
 $m_t^{H_2}$ —— 时段 t 内的产氢量, kg
 M^{H_2} —— 总产氢量, kg
 $P^{\text{bess}}(t)$ —— 储能功率(正为充电、负为放电), kW
 $P^{\text{ch}}(t)$ —— 储能充电功率, kW
 $P^{\text{cur}}(t)$ —— 弃电功率, kW
 $P^{\text{dis}}(t)$ —— 储能放电功率, kW

$P^{\text{lec}}(t)$ —— 电解槽耗电功率, kW
 $P^{\text{lec,max}}$ —— 电解槽额定功率, kW
 $P^{\text{e}}(t)$ —— 电网购电功率, kW
 $P^{\text{pv}}(t)$ —— 光伏出力, kW
 $P^{\text{w}}(t)$ —— 风电出力, kW
 $P^{\text{r}}(t)$ —— 可再生能源总出力, kW
 p_t —— 时段 t 的电价,元 $\cdot\text{kWh}^{-1}$
 R^{avg} —— 平均负载爬坡
 r_t —— 时段 t 即时奖励
 $\text{SOC}(t)$ —— 储能荷电状态
 $\text{SOC}^{\text{min}}, \text{SOC}^{\text{max}}$ —— SOC下限与上限
 s_t —— 时段 t 状态向量
 v_t —— 时刻 t 的风速, $\text{m}\cdot\text{s}^{-1}$
 $v^{\text{in}}, v^{\text{r}}, v^{\text{out}}$ —— 风机切入、额定、切出风速, $\text{m}\cdot\text{s}^{-1}$
 ΔL^{max} —— 相邻时段最大负载变化幅度
 $\eta^{\text{ch}}, \eta^{\text{dis}}$ —— 储能充/放电效率
 η^{pv} —— 光伏综合转换效率
 η^{RE} —— 可再生能源利用率
 γ —— 折扣因子

参考文献

- [1] 国家发展和改革委员会,国家能源局. 氢能产业发展中长期规划(2021–2035年)[EB/OL]. 2022. https://zfxxgk.nea.gov.cn/2022-03/23/c_1310525630.htm.
- [2] IEA. Global Hydrogen Review 2024[R]. Paris: International Energy Agency, 2024. <https://www.iea.org/reports/global-hydrogen-review-2024>.
- [3] Hydrogen Council. Hydrogen Scaling Up: A Sustainable Pathway for the Global Energy Transition[R]. Brussels: Hydrogen Council, 2017. <https://hydrogencouncil.com/en/newsroom/>.
- [4] Aminaho E N, Aminaho N S, Aminaho F. Techno-economic assessments of electrolyzers for hydrogen production[J]. Applied Energy, 2025, **399**: 126515.
- [5] Sayed-Ahmed H, Toldy Á I, Santasalo-Aarnio A. Dynamic operation of proton exchange membrane electrolyzers: Critical review[J]. Renewable and Sustainable Energy Reviews, 2024, **189**: 113883.
- [6] Nguyen E, Olivier P, Péra M C, et al. Impact of short-term intermittent operation on experimental industrial PEM and alkaline electrolyzers[J]. International Journal of Hydrogen Energy, 2025, **126**: 516–530.
- [7] Minnah P, Pope K. Degradation modelling and the impact of intermittent operation on proton exchange membrane electrolyzers[J]. International Journal of Hydrogen Energy, 2026, **202**: 152962.
- [8] Johnsen A G, Mitridati L, Zarrilli D, et al. The value of ancillary services for electrolyzers[J]. Computers & Chemical Engineering, 2026, **204**: 109360.
- [9] Zeng A D, Wang Z R, Wang J W, et al. Optimal scheduling of electricity – hydrogen – heat integrated energy system considering electrolyzer dynamic temperature and pipeline network heat storage[J]. International Transactions on Electrical Energy Systems, 2025, **2025**: 9478361.

- [10] Zhang L Q, Dai W Z, Zhao B, et al. Multi-time-scale economic scheduling method for electro-hydrogen integrated energy system based on day-ahead long-time-scale and intra-day MPC hierarchical rolling optimization[J]. *Frontiers in Energy Research*, 2023, **11**: 1132005.
- [11] Amores E, Rodríguez J, Carreras C. Influence of operation parameters in the modeling of alkaline water electrolyzers for hydrogen production[J]. *International Journal of Hydrogen Energy*, 2014, **39**(25): 13063–13078.
- [12] Masampally V S, Mitra K, Runkana V. Modeling and simulation of hydrogen production in a PEM water electrolyzer at high current densities[J]. *Energy Conversion and Management*, 2026, **352**: 121110.
- [13] Zhang Y T, Zhang T C, Liu S W, et al. Multi-energy complementary optimal scheduling based on hydrogen gas turbine considering the flexibility of electrolyser[J]. *IET Renewable Power Generation*, 2024, **18**(12): 1972–1985.
- [14] Cardona P, Costa-Castelló R, Roda V, et al. Model predictive control of an on-site green hydrogen production and refuelling station[J]. *International Journal of Hydrogen Energy*, 2023, **48**(47): 17995–18010.
- [15] Ulleberg Ø. Modeling of advanced alkaline electrolyzers: a system simulation approach[J]. *International Journal of Hydrogen Energy*, 2003, **28**(1): 21–33.
- [16] Abdin Z, Webb C J, Gray E M. Modelling and simulation of an alkaline electrolyser cell[J]. *Energy*, 2017, **138**: 316–331.
- [17] Ashraf H, Kawai E, Mae M, et al. Advancing electrochemical modelling of PEM electrolyzers through robust parameter estimation with the weighted mean of vectors algorithm[J]. *Multiscale and Multidisciplinary Modeling, Experiments and Design*, 2025, **8**(9): 397.
- [18] Mößle P, Herrmannsdörfer T, Welzl M, et al. Economic optimization for the dynamic operation of a grid connected and battery-supported electrolyzer[J]. *International Journal of Hydrogen Energy*, 2025, **100**: 749–759.
- [19] Zheng Y, You S, Bindner H W, et al. Optimal day-ahead dispatch of an alkaline electrolyser system concerning thermal – electric properties and state-transitional dynamics[J]. *Applied Energy*, 2022, **307**: 118091.
- [20] Voulkopoulos X, Dimitriadis C N, Georgiadis M C. Optimal scheduling of a RES – Electrolyzer aggregator in electricity, hydrogen and green certificates markets[J]. *International Journal of Hydrogen Energy*, 2024, **78**: 1099–1107.
- [21] Wang J B, Wen J F, Wang J R, et al. Coordinated scheduling of wind-solar-hydrogen-battery storage system for techno-economic-environmental optimization of hydrogen production[J]. *Energy Conversion and Management*, 2024, **314**: 118695.
- [22] Haarnoja T, Zhou A, Abbeel P, et al. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor[EB/OL]. 2018: arXiv: 1801.01290. <https://arxiv.org/abs/1801.01290>.
- [23] Haarnoja T, Zhou A, Hartikainen K, et al. Soft actor-critic algorithms and applications[EB/OL]. 2018: arXiv: 1812.05905. <https://arxiv.org/abs/1812.05905>.
- [24] Bäuml A, Benterki A, Meng J W, et al. Energy management strategies based on soft actor critic reinforcement learning with a proper reward function design based on battery state of charge constraints[J]. *Journal of Energy Storage*, 2024, **90**: 111797.
- [25] Zhao Y J, Lin F, Yang Z P. Soft actor-critic-based energy management strategy for hybrid energy storage system in urban rail transit DC traction power supply system[J]. *Electric Power Systems Research*, 2026, **252**: 112372.
- [26] Lu W C, Gao Y, Sun Z, et al. An Improved soft actor – critic framework for cooperative energy management in the building cluster[J]. *Applied Sciences*, 2025, **15**(16): 8966.
- [27] Van Quang T, Doan D T. Online transfer learning (OTL) for accelerating deep reinforcement learning (DRL) for building energy management[J]. *Journal of Building Performance Simulation*, 2025: 1–20.
- [28] RMI. 2024 China Power Market Outlook: 10 Key Trends for Market Players[R]. RMI, 2024. https://rmi.org/wp-content/uploads/dlm_uploads/2025/01/final-2024-China-Power-Market-Outlook-10-Key-Trends-for-Market-Players-2501117.pdf.
- [29] Li D F, Li Z H, Li X J, et al. Optimization method of time-of-use electricity price for the cost savings of power grid investment[J]. *Frontiers in Energy Research*, 2024, **12**: 1361116.
- [30] Dong L, Huang Y, Xu X, et al. Research on priority scheduling strategy for smoothing power fluctuations of microgrid Tie-lines based on PER-DDPG algorithm[J]. *IET Generation, Transmission & Distribution*, 2024, **18**(20): 3221–3233.
- [31] Tang D G, Ge P X, Yuan C Q, et al. Optimal management of coupled hydrogen-electricity energy systems at ports by multi-time scale scheduling[J]. *Applied Energy*, 2025, **391**: 125885.