

DOI: 10.12158/j.2096-3203.2025.03.003

深度强化学习驱动的风储系统参与能量-调频市场竞价策略

李钟平, 向月

(四川大学电气工程学院, 四川 成都 610065)

摘要:在电力市场环境下,风储系统通过参与能量市场和调频市场实现经济性提升和电网调频调峰辅助功能,但竞价策略需要解决风储竞价能量-调频双市场协同优化等关键问题。为此,文中提出一种基于深度强化学习驱动的风储系统参与能量-调频市场竞价策略,以应对不完全信息市场环境下的风储系统竞价策略。首先,构建风储系统参与能量-调频市场交易框架,阐明各市场主体的竞价与运营策略;然后,针对不同调频资源的响应能力差异,引入实时调频性能得分模型,并建立风储系统竞价模型;最后,为求解不完全信息市场环境下的多主体随机博弈问题,采用具备无模型学习能力的多智能体深度强化学习方法,处理多主体竞价博弈关系。仿真结果表明,文中所提方法能够有效为风储系统参与能量-调频市场制定竞价策略,在保证高收敛稳定性的同时显著提升经济性收益,并有效支持电网的调频调峰需求。

关键词:风储系统;能量-调频市场;深度强化学习;实时调频性能得分;演员-评论家;多主体竞价博弈

中图分类号:TM732

文献标志码:A

文章编号:2096-3203(2025)03-0030-13

0 引言

随着全球一次能源的持续消耗和“双碳”目标的推进,可再生能源发展备受瞩目^[1-4]。然而,新能源渗透率的不断提高导致电力系统惯量降低,使电网频率稳定性面临巨大挑战^[5-9]。我国风电场接入电力系统技术规定^[10]中明确要求包括风电在内的各类发电企业应积极提供电力辅助服务,鼓励其通过市场化的方式提供调频等电力辅助服务,以保障电力系统的稳定和安全。以山东电力市场为例,2023年新能源发电参与现货市场比例已突破10%,随着电力市场规则的日益成熟,风储系统在电力市场中的参与度显著提升,可通过协同参与能量交易和频率调节服务实现更高的收益^[11-12]。

现有研究表明,风机可通过虚拟惯量控制^[13]、转子动能控制^[14]和综合惯性控制^[15]等方法提供频率响应,虽然风电参与调频辅助服务市场具备一定技术可行性,但由于风电出力具有随机性、波动性等不确定性问题^[16-17],其调节精度不高、调频性能指标欠佳,在调频市场缺乏竞争力,因此风电单独参与调频市场实际可行性极低^[18-19]。相较而言,储能系统凭借良好的四象限工作特性^[20]和快速双向功率调节能力^[21],可有效平抑风电功率波动,能够与风电集群组成风储系统参与电力市场。现阶段针对风储系统参与电力市场已开展广泛研究,文献^[22]建立一种多时间尺度市场投标优化模型,通过

预留储能系统功率降低风储投标偏差风险。文献^[23]提出基于市场剩余需求曲线联合建模,制定风储系统日前市场的随机优化投标。文献^[24]以电力现货市场日前日内双结算模式为基础,建立风储系统日前投标和实时平衡的滚动自调节模型。然而,上述研究中风储系统往往被视作市场价格接受者,无法参与市场竞价环节,因此其参与市场的积极性降低。

此外,上述研究仅适用于完全信息市场环境^[25],即各市场主体对其他主体报价决策信息已知,整个市场信息处于透明状态,但显然该假设并不适用于电力市场的实际交易模式。因此文中研究风储系统参与能量-调频市场应尽可能贴合电力市场实际运行状态,即各市场主体处于不完全信息市场环境,对其余竞争成员报价决策信息未知。在此背景下,具备强大无模型学习能力的深度强化学习方法^[26-27]对此类不完全信息市场环境下的博弈问题适用性极高。针对多市场主体竞价博弈均衡问题,多智能体深度强化学习不仅继承了深度强化学习方法的高维数据处理和复杂策略学习能力^[28-29],其多智能体系统还可以将各市场主体作为智能体,在算法模型之间进行信息交流和内部交互,其训练效果比单智能体系统更好。文献^[30]采用演员-评论家(actor-critic, AC)框架来实现多智能体间内部交互,提出了一种基于AC框架的多智能体协同决策模型。文献^[31]提出一种高维协同多智能体软演员-评论家(soft actor-critic, SAC)算法,利用高斯分布生成连续动作空间,以解决传统强化学习维数灾

收稿日期: 2024-12-08; 修回日期: 2025-03-04

基金项目: 国家自然科学基金资助项目(U2166211)

和局部最优等问题。文献[32]提出基于改进SAC的多主体联合竞价模型,通过建立点对点风电交易市场有效规避风电偏差惩罚,结合样本降重方法有效避免局部最优问题,为高风电渗透率电力系统市场竞价交易问题提供解决方案。

综上所述,目前阶段关于风储系统参与电力市场的风储系统往往以价格接受者的身份参与调频市场,随着电力系统风电渗透率的不断提高,该市场机制难以吸引大型风储系统参与调频市场,不利于系统总调频响应能力的提升。对此,文中提出风储系统参与能量-调频市场交易框架,市场内各主体均参与竞价,形成多主体自由竞价的机制。此外,现有部分研究仅局限于完全信息市场环境下风储系统竞价策略优化,显然这与市场运行过程中成员报价决策信息不对外公开的实际情况不符,因此针对不完全信息市场环境下多主体竞价博弈问题,文中提出一种多智能体软演员-评论家(multi-agent soft actor-critic, MASAC)驱动的风储系统竞价模型,同时引入实时调频性能得分模型,用以体现调频资源的优异程度。最后通过算例仿真分析验证文中所提方法的有效性与优越性,为风储系统参与电力市场竞价策略制定提供参考,有效助力电网调频调峰。

1 风储系统参与能量-调频市场交易框架

为研究多市场主体参与能量-调频市场的竞价策略制定,文中详细阐述所提出的电力市场机制,制定各主体竞价运营策略和能量-调频市场联合出清等市场交易流程,风储系统与多个火电厂以平等地位参与电力市场,能量市场按各主体报价直接出清;调频市场则根据各调频资源的性能指标进行竞标价格排序,最终按照排序后综合价格联合出清。

1.1 市场机制

文中结合美国宾夕法尼亚-新泽西-马里兰州(Pennsylvania-New Jersey-Maryland, PJM)电力市场机制,制定风储系统参与电力市场的规则,其参与市场范围涉及能量市场和调频市场,文中侧重于各市场主体在电能量交易和调频辅助服务交易过程中的竞价策略制定。

在市场申报环节,各市场主体(包括风储系统和多个独立火电厂)向调度中心申报能量-调频市场的投标信息,各市场主体报价方式虽然有所差异,但整体上都以自身效益最大化为目标,向调度中心分别申报能量市场和调频市场的竞标价格及容量,然后调度中心以总用能成本最小化为目标,在考虑

物理约束、满足负荷需求及调频需求的前提下,形成日前能量-调频市场的出清结果。电力市场实时交易期间各市场主体需上报系统实时出力,以供计算各主体双市场中标电量收益和未完成中标计划偏差惩罚。图1为风储系统参与能量-调频市场的交易框架。

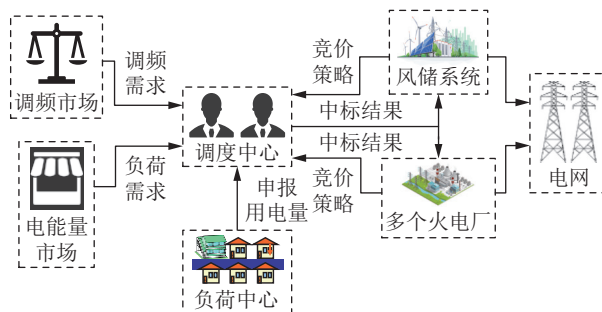


图1 风储系统参与能量-调频市场的交易框架

Fig.1 The trading framework for joint participation of wind-storage system in energy and frequency regulation markets

1.2 各市场主体的竞价运营策略

在能量-调频市场中,各市场主体采用自主竞价的方式满足市场负荷及调频需求,并以市场用能成本最小化为目标进行双市场联合出清。

多个火电厂作为能量-调频市场的竞争者,各自作为独立的市场主体参与竞价,考虑机组自身运行情况、发电成本和调节性能等因素,作为传统调频资源独立参与双市场联合出清,从而有效平衡双市场出清价格并增强市场竞争性。

风储系统作为一个整体,须合理分配其在能量市场和调频市场的功率份额。由于风电功率具有不确定性,而储能系统具有快速响应特性,风储联合运行可有效弥补风电单独参与电力市场的不足,从而提升其在能量-调频市场竞价中的竞争力。区别于调频性能较弱的传统火电厂,风储系统在制定竞价策略时更倾向于将功率份额放至调频市场以获得更为稳定的调频收益。风储系统报价目标为自身效益最大化,综合考虑风电出力水平、储能荷电状态(state of charge, SOC)以及功率等因素,向市场提交报价报量。

2 风储系统参与能量-调频市场竞价模型

风储系统以自身效益最大化为目标参与电力市场。在能量市场中,须申报竞标价格和竞标容量策略,在调频市场中,须申报竞标容量价格、竞标调频价格及竞标容量,同时调度中心在调频市场中须考虑调频资源响应能力等综合调频性能指标,依据各主体竞价策略实现联合出清。

2.1 收益模型

各市场主体的总收益由两部分组成, 能量市场的电能量交易收益和调频市场的辅助服务收益。文中收益模型制定过程中引入偏差惩罚机制, 以约束市场主体随意制定竞价策略的行为, 确保其中标容量与各自机组日内实际出力相匹配。

2.1.1 能量市场收益

储能作为风储系统的关键组成部分, 不再单独参与能量市场的购售电交易, 而是与风电集群联合交易, 为风储系统在能量-调频市场中获取更大收益。各市场主体参与能量市场的收益指市场主体通过日前投标报价, 经能量市场出清后, 计算各市场主体所获得收益和实际出力与中标电量间偏差惩罚费用, 形成各市场主体能量市场收益函数。能量市场各市场主体总体收益如下:

$$F_{en} = \sum_{t=1}^N \lambda_{MC,t}^{en} (p_{WE,t}^{en} + p_{T,t}^{en}) \Delta t - F_{en}^{pen} \quad (1)$$

$$F_{en}^{pen} = \sum_{t=1}^N \rho (|p_{WE,t}^{en} - P_{WE,MC,t}^{en}| + |p_{T,t}^{en} - P_{T,MC,t}^{en}|) \Delta t \quad (2)$$

式中: F_{en} 为能量市场总收益; N 为调度时段总数; $\lambda_{MC,t}^{en}$ 为第 t 个调度时段的能量市场价格; $p_{WE,t}^{en}$ 、 $p_{T,t}^{en}$ 分别为第 t 个调度时段风储系统和火电厂在电能量市场的实际出力; Δt 为时段间隔; F_{en}^{pen} 为偏差惩罚费用; $P_{WE,MC,t}^{en}$ 、 $P_{T,MC,t}^{en}$ 分别为第 t 个调度时段风储系统和火电厂在电能量市场的中标容量; ρ 为偏差惩罚系数。

2.1.2 调频市场收益

参照美国 PJM 市场中调频市场规则, 市场成员获得收益分为调频容量收益和调频性能收益, 调频市场引入实时调频性能得分模型, 并将各主体竞标价格按排序后综合价格进行出清, 因此调频市场各主体中标价格不再为统一值, 调频市场各市场主体总体收益如下:

$$F_{reg} = F^{cap} + F^{perf} - F_{reg}^{pen} \quad (3)$$

$$F^{cap} = \sum_{t=1}^N (K_{WE,t}^{perf} \lambda_{WE,MC,t}^{cap} P_{WE,t}^{reg} + K_{T,t}^{perf} \lambda_{T,MC,t}^{cap} P_{T,t}^{reg}) \Delta t \quad (4)$$

$$F^{perf} = \sum_{t=1}^N (K_{WE,t}^{perf} \lambda_{WE,MC,t}^{perf} P_{WE,t}^{reg} R_{WE,t} + K_{T,t}^{perf} \lambda_{T,MC,t}^{perf} P_{T,t}^{reg} R_{T,t}) \Delta t \quad (5)$$

$$F_{reg}^{pen} = \sum_{t=1}^N \rho (|p_{WE,t}^{reg} - P_{WE,MC,t}^{reg}| + |p_{T,t}^{reg} - P_{T,MC,t}^{reg}|) \Delta t \quad (6)$$

式中: F_{reg} 为调频市场总收益; F^{cap} 、 F^{perf} 分别为调频容量收益和调频性能收益; F_{reg}^{pen} 为调频市场偏差惩罚费用; $\lambda_{WE,MC,t}^{cap}$ 、 $\lambda_{WE,MC,t}^{perf}$ 分别为第 t 个调度时段风储

系统中标调频容量价格和中标调频性能价格; $\lambda_{T,MC,t}^{cap}$ 、 $\lambda_{T,MC,t}^{perf}$ 分别为第 t 个调度时段火电厂中标调频容量价格和中标调频性能价格; $K_{WE,t}^{perf}$ 、 $K_{T,t}^{perf}$ 分别为第 t 个调度时段风储系统和火电厂实时调频性能得分; $p_{WE,t}^{reg}$ 、 $p_{T,t}^{reg}$ 、 $R_{WE,t}$ 、 $R_{T,t}$ 分别为第 t 个调度时段风储系统和火电厂在调频市场的实际出力和各自的里程调用率; $P_{WE,MC,t}^{reg}$ 、 $P_{T,MC,t}^{reg}$ 分别为第 t 个调度时段风储系统和火电厂在调频市场的中标容量。

2.2 报价模型

火电厂报价模型反映其参与能量-调频市场的行为决策, 报价目标为火电厂自身效益最大化, 火电厂报价函数如下:

$$F_T = I_T^{en} + I_T^{reg} - C_T(P_{T,B,t}^{en} + P_{T,B,t}^{reg}) \quad (7)$$

$$I_T^{en} = \sum_{t=1}^N \lambda_{T,B,t}^{en} P_{T,B,t}^{en} \Delta t \quad (8)$$

$$I_T^{reg} = \sum_{t=1}^N (\lambda_{T,B,t}^{cap} K_{T,B,t}^{perf} P_{T,B,t}^{reg} + \lambda_{T,B,t}^{perf} K_{T,B,t}^{perf} P_{T,B,t}^{reg} R_{T,t}) \Delta t \quad (9)$$

$$C_T(P) = \sum_{t=1}^N [a(P_t \Delta t)^2 + bP_t \Delta t + c] \quad (10)$$

式中: F_T 为火电厂报价; I_T^{en} 、 I_T^{reg} 分别为火电厂在电能量市场和调频市场的预期收益; $C_T(P)$ 为火电厂运行成本函数; $P_{T,B,t}^{en}$ 、 $P_{T,B,t}^{reg}$ 分别为第 t 个调度时段火电厂能量市场和调频市场竞标容量; $\lambda_{T,B,t}^{en}$ 、 $\lambda_{T,B,t}^{cap}$ 、 $\lambda_{T,B,t}^{perf}$ 分别为第 t 个调度时段火电厂能量市场竞标价格、调频容量竞标价格和调频性能竞标价格; P_t 为第 t 个调度时段火电厂功率; a 、 b 、 c 为火电厂成本系数。

风储系统报价模型反映风储系统参与能量-调频市场的预期收益, 报价目标为风储系统总收益最大化, 其报价函数如下:

$$F_{WE} = I_{WE}^{en} + I_{WE}^{reg} - C_{WE} \quad (11)$$

$$I_{WE}^{en} = \sum_{t=1}^N \lambda_{WE,B,t}^{en} P_{WE,B,t}^{en} \Delta t \quad (12)$$

$$I_{WE}^{reg} = \sum_{t=1}^N \lambda_{WE,B,t}^{cap} K_{WE,t}^{perf} P_{WE,B,t}^{reg} \Delta t + \sum_{t=1}^N \lambda_{WE,B,t}^{perf} K_{WE,t}^{perf} P_{WE,B,t}^{reg} R_{WE,t} \Delta t \quad (13)$$

$$C_{WE} = \sum_{t=1}^N \rho |p_{WE,t} - P_{WE,B,t}| \Delta t + \sum_{t=1}^N \{ \gamma_1 [(1-\beta_t) p_{ESS,t}^{ch} + \beta_t p_{ESS,t}^{dis}] + \gamma_2 E_{ESS,t} \} \Delta t \quad (14)$$

式中: F_{WE} 为风储系统报价; I_{WE}^{en} 、 I_{WE}^{reg} 分别为风储系统在电能量市场和调频市场的预期收益; C_{WE} 为机

组运行成本; $\lambda_{WE,B,t}^{\text{en}}$ 、 $\lambda_{WE,B,t}^{\text{cap}}$ 、 $\lambda_{WE,B,t}^{\text{perf}}$ 分别为第 t 个调度时段风储系统能量市场竞标价格、调频容量竞标价格和调频性能竞标价格; $P_{WE,B,t}^{\text{en}}$ 、 $P_{WE,B,t}^{\text{reg}}$ 分别为第 t 个调度时段风储系统能量市场和调频市场竞标容量; $p_{WE,t}$ 、 $P_{WE,B,t}$ 分别为第 t 个调度时段风储系统实际总出力和总竞标容量; $p_{ESS,t}^{\text{ch}}$ 、 $p_{ESS,t}^{\text{dis}}$ 分别为第 t 个调度时段储能充、放电功率; β_t 为储能充放电控制变量, 为 1 时表示放电, 为 0 时表示充电; $E_{ESS,t}$ 为第 t 个调度时段储能系统剩余电量; γ_1 、 γ_2 分别为储能单位充放电成本和储能单位维护成本。

2.3 实时调频性能得分模型

为充分体现各市场主体在调频市场中的调频能力, 优化调频资源配置, 文中参考美国 PJM 市场的实时调频性能得分模型, 评估市场成员响应调频信号实现动作的准确性, 并分析调频资源对调频收益的影响, 其计算如下:

$$K_t^{\text{perf}} = 1 - \frac{1}{K} \sum_{k=1}^K \frac{|p_{t,k} - P_{MC,t}^{\text{reg}} S_{t,k}|}{P_{MC,t}^{\text{reg}} V_t} \quad (15)$$

式中: K_t^{perf} 为市场成员的调频性能得分; K 为调度时间内的调频信号总数; $p_{t,k}$ 为调频资源在第 t 个调度时段内第 k 个调频出力; $P_{MC,t}^{\text{reg}}$ 为调频资源第 t 个调度时段调频中标容量; $S_{t,k}$ 为第 t 个调度时段第 k 个调频指令信号; V_t 为第 t 个调度时段调频信号绝对值的平均值。原则上 K_t^{perf} 取值范围为 $[0,1]$, K_t^{perf} 越大表明调频资源对调频信号的响应程度越高, 调频性能越好。

此外, 各市场主体的调频性能得分还应综合考虑响应时间、调节速度等调频性能指标, 然而文中所研究的能量-调频市场采用小时级时间尺度, 因此调度中心在评估调频资源时主要关注调节精度, 即文中所提调频性能得分。

2.4 联合出清模型

市场成员参与调频市场进行报价时, 向市场报价决策集合中包含调频容量价格、调频性能价格, 为体现各市场成员作为调频资源在市场出清的综合优势, 调度中心依据历史数据计算调频资源实时调频性能得分, 依据各成员竞价策略排序, 计算如下:

$$\lambda_{Ad,t}^{\text{cap}} = \frac{\lambda_{B,t}^{\text{cap}}}{f_b} \times \frac{c_q}{K_t^{\text{perf}}} \quad (16)$$

$$\lambda_{Ad,t}^{\text{perf}} = \frac{\lambda_{B,t}^{\text{perf}}}{f_b} \times \frac{c_q}{K_t^{\text{perf}}} R_t \quad (17)$$

$$\lambda_{Ad,t}^{\text{reg}} = (\lambda_{Ad,t}^{\text{cap}} + \lambda_{Ad,t}^{\text{perf}}) / c_q \quad (18)$$

式中: $\lambda_{Ad,t}^{\text{cap}}$ 、 $\lambda_{Ad,t}^{\text{perf}}$ 分别为第 t 个调度时段市场成员经过调度中心调频性能评估并调整后的容量价格和

性能价格; $\lambda_{Ad,t}^{\text{reg}}$ 为第 t 个调度时段各市场成员参与竞标的排序价格; $\lambda_{B,t}^{\text{cap}}$ 、 $\lambda_{B,t}^{\text{perf}}$ 分别为第 t 个调度时段市场成员的调频容量竞标价格和调频性能竞标价格; c_q 为市场成员的有效申报调频容量; f_b 为调频资源收益因子, 文中研究风储系统属于快速调频资源^[33], 其最大值不超过 2.9, 而火电厂类的传统调频资源收益因子取 1; R_t 为第 t 个调度时段的里程调用率, 由各主体历史数据计算求得。

以系统总用能成本最小化为目标函数, 进行能量-调频市场联合出清:

$$\min (F_{MC}^{\text{en}} + F_{MC}^{\text{cap}} + F_{MC}^{\text{perf}}) \quad (19)$$

$$F_{MC}^{\text{en}} = \sum_{t=1}^N \lambda_{MC,t}^{\text{en}} (P_{WE,MC,t}^{\text{en}} + P_{T,MC,t}^{\text{en}}) \Delta t \quad (20)$$

$$F_{MC}^{\text{cap}} = \sum_{t=1}^N K_{WE,t}^{\text{perf}} \lambda_{WE,MC,t}^{\text{cap}} P_{WE,MC,t}^{\text{reg}} \Delta t + \sum_{t=1}^N K_{T,t}^{\text{perf}} \lambda_{T,MC,t}^{\text{cap}} P_{T,MC,t}^{\text{reg}} \Delta t \quad (21)$$

$$F_{MC}^{\text{perf}} = \sum_{t=1}^N \lambda_{WE,MC,t}^{\text{perf}} K_{WE,t}^{\text{perf}} P_{WE,MC,t}^{\text{reg}} R_{WE,t} \Delta t + \sum_{t=1}^N \lambda_{T,MC,t}^{\text{perf}} K_{T,t}^{\text{perf}} P_{T,MC,t}^{\text{reg}} R_{T,t} \Delta t \quad (22)$$

式中: F_{MC}^{en} 为中标结算时能量市场总用能成本; F_{MC}^{cap} 、 F_{MC}^{perf} 分别为中标结算时调频市场调频容量成本和调频性能成本。

双市场联合出清模型约束条件如下:

$$P_T^{\text{min}} \leq P_{T,MC,t}^{\text{en}} + P_{T,MC,t}^{\text{reg}} \leq P_T^{\text{max}} \quad (23)$$

$$-P_{\text{down}} \leq P_{T,MC,t}^{\text{en}} + P_{T,MC,t}^{\text{reg}} - P_{T,MC,t-1}^{\text{en}} - P_{T,MC,t-1}^{\text{reg}} \leq P_{\text{up}} \quad (24)$$

$$P_{WE}^{\text{min}} \leq P_{WE,MC,t}^{\text{en}} + P_{WE,MC,t}^{\text{reg}} \leq P_{WE}^{\text{max}} \quad (25)$$

$$\mu_{ESS}^{\text{min}} E^{\text{cap}} \leq E_{ESS,t} \leq \mu_{ESS}^{\text{max}} E^{\text{cap}} \quad (26)$$

$$E_{ESS,t} = E_{ESS,t-1} + (1 - \beta_t) p_{ESS,t}^{\text{ch}} - \beta_t p_{ESS,t}^{\text{dis}} \quad (27)$$

式中: P_T^{max} 、 P_T^{min} 和 P_{WE}^{max} 、 P_{WE}^{min} 分别为火电厂和风储系统出力上、下限; P_{up} 、 P_{down} 分别为火电厂向上、向下爬坡约束; μ_{ESS}^{max} 、 μ_{ESS}^{min} 分别为储能系统 SOC 上、下限; E^{cap} 为储能最大容量。

风储系统和多个火电厂处于平等地位自由竞价, 依据自身报价模型向能量-调频市场报价报量, 市场联合出清后返回中标结果以计算各主体市场收益, 如图 2 所示, 其中 L 为智能体数量。多市场主体在参与能量-调频市场相互竞价博弈过程中, 所有市场成员之间的报价决策信息未知。这种情况与多智能体深度强化学习算法的应用场景高度契合, 该算法通过最大化策略熵扩展智能体探索范

围,有效应对不完全信息环境下多市场主体竞价博弈问题。由于该问题变量设置复杂,各市场主体报价决策既相互独立又彼此影响,面临着大量变量交互与频繁的相互调用情形,传统数学优化算法求解难度高。因此文中充分发挥人工智能算法无模型、自学习等优势,采用多智能体深度强化学习算法进行求解。

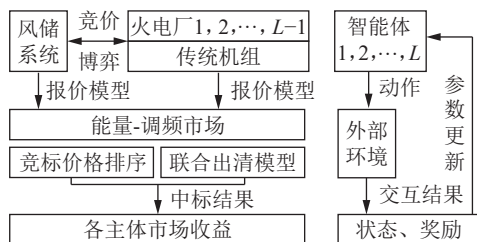


图2 各主体市场收益过程示意

Fig.2 Schematic diagram of the market revenue process for each entity

3 模型求解方法

利用深度强化学习 MASAC 方法求解文中所提模型,该方法与风储系统参与能量-调频市场竞价策略制定问题适配程度高。首先,风储系统多维连续的市场报价行为与多智能体系统中序列决策过程契合程度高;其次,多智能体深度强化学习方法可以有效反映出风电和储能在能量-调频市场中的合作博弈关系。因此,MASAC 方法是应对不完全信息市场环境下多主体竞价博弈问题的有效手段之一。

3.1 深度强化学习 MASAC 算法理论

多个智能体同时与外部环境交互,每个智能体的决策和行为不仅影响自身的状态和回报,还可能影响其他智能体的状态和回报,因此若在单个智能体训练过程中将其他智能体当做外部环境模块,容易使得模型训练时外部环境变化情形过于丰富,外部环境对应的奖励值不确定性加重,算法收敛难度大幅提高。集中式训练分布式执行(centralized training with decentralized execution, CTDE)范式是实现多智能体协作的有效框架之一,在集中式训练阶段,智能体可以利用全局信息来优化联合策略和价值函数,智能体间存在内部交互,以学习到更好的策略,有效提高模型训练的稳定性;在分布式执行阶段,每个智能体根据自身策略做出决策,保证算法应用的灵活性与可行性,执行阶段仅利用局部信息使得算法更具扩展性和实用性,有效应对不完全信息市场带来的局部信息约束。

MASAC 算法是 SAC 算法的拓展,其创新点在

于引入多智能体系统,采用 CTDE 框架,在文中多主体竞价博弈场景中算法学习目标是最大化所有智能体的累积回报,即各市场主体收益最大化。同时多智能体系统特有的 CTDE 范式中包含分布式执行训练模式,智能体决策过程中仅参照自身观测信息,与文中不完全信息市场环境下风储系统竞标过程中对其他市场主体竞标信息未知的场景适配性极高。MASAC 算法沿用 SAC 算法的演员(actor)网络和评论家(critic)网络,核心机制是每个智能体都有自己的 AC 网络,其策略函数 π_{θ_i} 定义了智能体 i 在不同状态下的行为策略,其期望回报梯度可以表示为:

$$\nabla_{\theta_i} J(\theta_i) = E_{s \sim D, a \sim \pi_{\theta_i}} (\nabla_{\theta_i} \ln \pi_{\theta_i}(a_i | s_i) Q^{\pi_{\theta_i}}(s_i, a_i)) \quad (28)$$

式中: $\nabla_{\theta_i} J(\theta_i)$ 为关于智能体 i 网络参数 θ_i 的期望回报函数 $J(\theta_i)$ 的梯度; $\pi_{\theta_i}(a_i | s_i)$ 为以 θ_i 为网络参数的策略函数,指智能体 i 在给定状态 s_i 下采取动作 a_i 的概率,表示状态空间到动作空间的映射关系; $D = \{s, a, r, s'\}$ 为经验回放池,其中 s 为当前状态, a 为动作, r 为奖励, s' 为下一状态,用以存储智能体与环境交互过程中产生的经验(即状态、动作、奖励和下一个状态的四元组); $E_{s \sim D, a \sim \pi_{\theta_i}}$ 为状态 s 和动作 a 联合分布下的期望,其中状态 s 从经验回放池 D 中采样,动作 a 从策略 π_{θ_i} 中采样; $Q^{\pi_{\theta_i}}(s_i, a_i)$ 为智能体 i 遵循策略 π_{θ_i} 的期望回报。

MASAC 算法中通过最大化策略的熵,鼓励智能体探索更广泛的策略空间,增强鲁棒性和探索能力。其参数更新主要涉及 actor 网络和 critic 网络的更新。actor 网络的参数更新方法是基于软策略梯度,引入熵正则化项来增强探索能力,而 critic 网络的参数更新方法是基于值函数的近似进行优化,与传统 SAC 算法不同,如式(29)、式(31)所示, MASAC 算法中目标 Q 值和最小化损失函数更新均采用所有智能体的联合动作,以保证智能体后续能够有效利用全局信息来优化联合策略和价值函数,参数更新方法如下:

$$y_i = r_i + \gamma (Q'_i(s'_i, \{a'_1, \dots, a'_L\}) - \alpha \ln \pi'_{\theta_i}(a'_i | s'_i)) \quad (29)$$

$$L(\alpha) = -\alpha (\ln \pi_{\theta_i}(a_i | s_i) + H) \quad (30)$$

$$L(\phi_i) = (Q_i(s_i, \{a_1, \dots, a_L\} | \phi_i) - y_i)^2 \quad (31)$$

$$\nabla_{\theta_i} J(\theta_i) = \nabla_{\theta_i} (\alpha \ln \pi_{\theta_i}(a_i | s_i) - Q^{\pi_{\theta_i}}(s_i, a_i)) \nabla_{\theta_i} \pi_{\theta_i}(s_i | \theta_i) \quad (32)$$

$$\theta'_i = \tau \theta_i + (1 - \tau) \theta'_i \quad (33)$$

$$\phi'_i = \tau \phi_i + (1 - \tau) \phi'_i \quad (34)$$

式中: y_i 为目标 Q 值; r_i 为智能体 i 的即时奖励; γ 为折扣因子; $Q'_i(s'_i, \{a'_1, \dots, a'_L\})$ 为目标 critic 网络在状

态 s'_i 和 L 个智能体动作 $\{a'_1, \dots, a'_L\}$ 下的输出; $\pi_{\theta_i}(a'_i|s'_i)$ 为智能体 i 的目标 actor 网络在状态 s'_i 和动作 a'_i 以及网络参数 θ_i 下的策略函数输出; α 为温度参数, 用于控制策略熵, 在训练过程中自适应更新, 影响智能体探索和利用的平衡, 通常初始化为一个较小的正数; $L(\alpha)$ 为温度参数 α 的最小化损失函数; H 为目标熵, 用于调整温度参数; ϕ_i 为 critic 网络参数; $Q_i(s_i, \{a_1, \dots, a_L\}|\phi_i)$ 为当前 critic 网络在状态 s_i 和 L 个智能体动作 $\{a_1, \dots, a_L\}$ 以及网络参数 ϕ_i 下的输出; $L(\phi_i)$ 为 critic 网络的最小化损失函数; $\nabla_{\theta_i} \pi_{\theta_i}(a_i|s_i)$ 为关于 actor 网络参数 θ_i 的策略函数 π_{θ_i} 的梯度, 用于衡量策略函数 π_{θ_i} 随参数 θ_i 变化的敏感性; τ 为软更新参数, 通常初始化为一个较小的正数; θ'_i 、 ϕ'_i 分别为目标 actor 网络和目标 critic 网络参数。

3.2 模型求解

文中所提模型求解方法如下, MASAC 算法中每一个智能体对应能量-调频市场中的一个市场主体, 算法外部环境即为电力市场, 智能体动作对应为各市场主体的报价决策行为, 智能体观测信息为双市场出清结果和各市场主体运行状态, 算法奖励为各市场主体利润, 风储系统在能量-调频市场中竞价策略不断优化直至风储系统收益最大化。图 3 为 MASAC 算法的理论框架, 其中, π 、 π^* 分别为 actor 网络的策略函数和目标策略函数; Q_{s_i} 、 V_{s_i} 分别为 critic 网络 t 时刻的动作价值函数和状态价值函数输出。

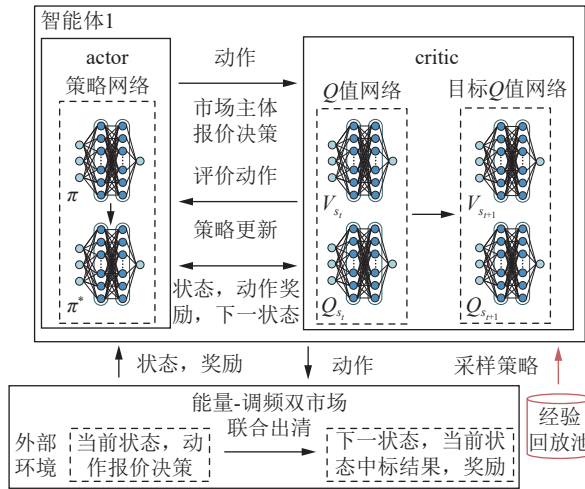


图 3 MASAC 算法的理论框架

Fig.3 Theoretical framework of the MASAC algorithm

(1) 状态空间。

第 t 个调度时段状态空间 s_t 包含所有智能体的观测信息, 即能量-调频市场出清价格和各市场主体的运行状态, 中标容量等信息。

$$s_t = \{\lambda_{MC,t}^{\text{en}}, \lambda_{WE,MC,t}^{\text{cap}}, \lambda_{T,MC,t}^{\text{cap}}, \lambda_{WE,MC,t}^{\text{perf}}, \lambda_{T,MC,t}^{\text{perf}}, s_{WE}, s_T\} \quad (35)$$

$$s_{WE} = \{P_{WE,MC,t}^{\text{en}}, P_{WE,MC,t}^{\text{reg}}, P_{WE,t}, E_{ESS,t}, \beta_t, K_{WE,t}^{\text{perf}}\} \quad (36)$$

$$s_T = \{P_{T,MC,t,j}^{\text{en}}, P_{T,MC,t,j}^{\text{reg}}, P_{T,t,j}, K_{T,t,j}^{\text{perf}}\} \Big|_{j=1,2,\dots,n} \quad (37)$$

式中: s_{WE} 、 s_T 分别为风储系统和火电厂状态观测信息, 包含其各项运行数据; $P_{T,MC,t,j}^{\text{en}}$ 、 $P_{T,MC,t,j}^{\text{reg}}$ 、 $P_{T,t,j}$ 、 $K_{T,t,j}^{\text{perf}}$ 分别为第 t 个调度时段第 j 个火电厂的能量市场中标容量、调频市场中标容量、实际总出力和调频性能得分; n 为火电厂数量。

(2) 动作空间。

第 t 个调度时段动作空间 a_t 包含所有智能体的动作信息, 为各市场主体的报价决策集合, 即风储系统能量竞标价格、调频容量竞标价格、调频性能竞标价格和能量市场、调频市场的竞标容量, 全部火电厂各自的能量竞标价格、调频容量竞标价格、调频性能竞标价格和能量市场、调频市场的竞标容量。

$$a_t = \{a_{WE,B,t}, a_{T,B,t}\} \quad (38)$$

$$a_{WE,B,t} = \{\lambda_{WE,B,t}^{\text{en}}, \lambda_{WE,B,t}^{\text{cap}}, \lambda_{WE,B,t}^{\text{perf}}, P_{WE,B,t}^{\text{en}}, P_{WE,B,t}^{\text{reg}}\} \quad (39)$$

$$a_{T,B,t} = \{\lambda_{T,B,t,j}^{\text{en}}, \lambda_{T,B,t,j}^{\text{cap}}, \lambda_{T,B,t,j}^{\text{perf}}, P_{T,B,t,j}^{\text{en}}, P_{T,B,t,j}^{\text{reg}}\} \Big|_{j=1,2,\dots,n} \quad (40)$$

式中: $a_{WE,B,t}$ 、 $a_{T,B,t}$ 分别为第 t 个调度时段风储系统和火电厂动作空间; $\lambda_{T,B,t,j}^{\text{en}}$ 、 $\lambda_{T,B,t,j}^{\text{cap}}$ 、 $\lambda_{T,B,t,j}^{\text{perf}}$ 分别为第 t 个调度时段第 j 个火电厂能量竞标价格、调频容量竞标价格、调频性能竞标价格; $P_{T,B,t,j}^{\text{en}}$ 、 $P_{T,B,t,j}^{\text{reg}}$ 分别为第 t 个调度时段第 j 个火电厂的能量市场竞标容量和调频市场竞标容量。

(3) 奖励。

第 t 个调度时段奖励集合 r_t 包含所有市场主体的收益。

$$r_t = \{F_{WE}, F_T\} \quad (41)$$

$$F_T = \{F_{T,j}\} \Big|_{j=1,2,\dots,n} \quad (42)$$

式中: F_{WE} 、 F_T 分别为风储系统和火电厂的收益; $F_{T,j}$ 为第 j 个火电厂的收益, 详细计算见 2.1 节。

将 MASAC 算法中各关键参数与文中所提风储系统竞价模型一一对应, 基于 MASAC 算法模型求解流程如图 4 所示, 用以介绍该算法求解文中风储系统竞价模型的具体求解流程。其中, m 、 M 分别为回合数和最大回合数; T 为单回合的最大时间步。

4 算例分析

4.1 基础数据

文中采用西南地区某风电场站实测数据进行分析, 采样时间间隔设置为 1 h, 一天内进行 24 次采

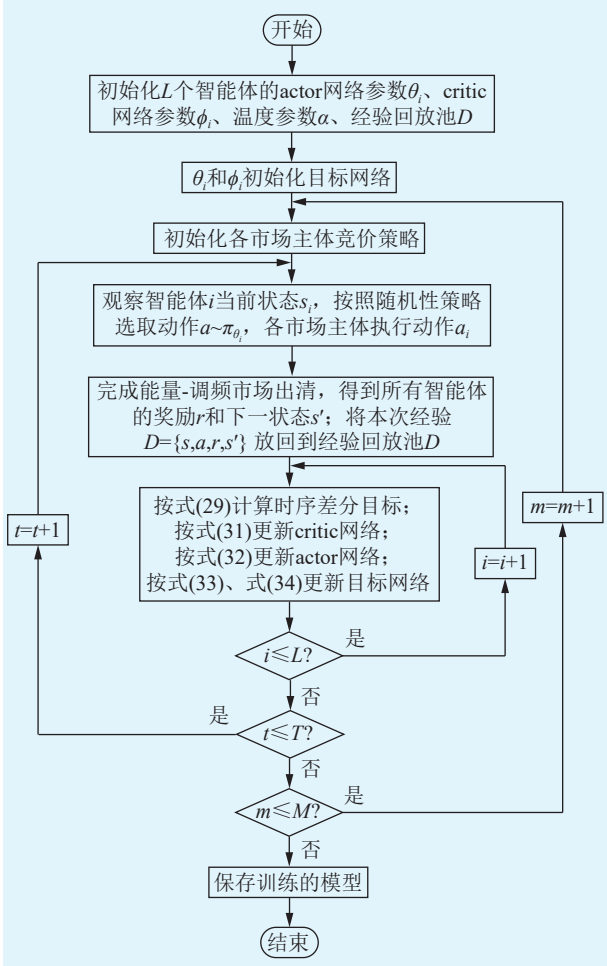


图 4 基于 MASAC 算法模型求解流程
Fig.4 Model solving flow chart based on MASAC algorithms

样,该风电场实际容量为 160 MW·h,储能系统容量按 30% 风电场容量配置,其余相关参数设置如表 1 所示,二者组成风储系统作为市场成员参与能量-调频市场。参考文献[32,34]将常规火电厂设置为 3 个,火电厂参数设置如表 2 所示。为充分体现市场成员竞争性,各时刻市场需求均小于所有市场主体可发电量总和,基于 MASAC 的竞价模型超参数如表 3 所示。

表 1 储能参数设置
Table 1 Parameters setting for energy storage system

参数	数值
容量/(MW·h)	48
功率/MW	14.5
$\mu_{ESS}^{\min}$	0.1
$\mu_{ESS}^{\max}$	0.9
γ_1 /(美元·MW ⁻¹)	25
γ_2 /(美元·MW ⁻¹ ·h ⁻¹)	0.4

4.2 算法性能分析

对比对抗双深度 Q 网络(dueling double deep Q-

表 2 火电厂参数设置
Table 2 Parameters setting for thermal power plant

参数	火电厂1	火电厂2	火电厂3
出力上限/MW	200	200	400
出力下限/MW	30	30	60
爬坡率/(MW·h ⁻¹)	70	70	90
综合调频性能指标	0.7	0.75	0.8
成本一次项系数/(美元·MW ⁻¹ ·h ⁻¹)	5.3	5.8	6.5
成本二次项系数/(美元·MW ⁻² ·h ⁻²)	0.016	0.014	0.018

表 3 MASAC 模型超参数设置
Table 3 Hyperparameters setting for MASAC model

超参数	数值
actor网络学习率	0.01
critic网络学习率	0.01
批尺寸	96
折扣率	0.99
软更新参数	0.005
温度参数初值	0.02

network, D3QN)、SAC 和 MASAC 算法在相同训练时段内应用文中所提风储系统竞价模型的收敛步数和收敛时间,结果如表 4 所示。

表 4 不同算法下收敛步数和收敛时间对比
Table 4 Comparison of convergence steps and convergence time under different algorithms

竞价算法	收敛步数	收敛时间/s
D3QN	135	293.6
SAC	120	262.3
MASAC	50	121.2

图 5 为全部市场主体总回报随训练步数反复迭代的收益曲线。从收敛速度来看, MASAC 算法的收敛步数约为 50, 而 SAC 算法和 D3QN 算法收敛步数约为 120 和 135, 3 种算法的收敛时间分别为 121.2、262.3、293.6 s, MASAC 算法针对文中所提模型的收敛速度明显更快, 相较于其他两种算法而言分别提升约 53.8%、58.7%。从收敛稳定性方面来看, D3QN 算法作为对抗深度 Q 网络(dueling deep Q-network, DQN)算法的改进版本, 其动作空间仍为离散值, 算法收敛后总回报仍处于较高波动水平, 而 MASAC 算法和 SAC 算法得益于其连续动作空间, 收敛后总回报波动区间明显减小, 总波动维持在 2×10⁴ 美元范围内, 收敛稳定性表现明显更优。从算法收益表现来看, MASAC 算法相较于 SAC 和 D3QN 算法而言, 采用了多智能体系统, 与文中所提竞价模型所涉及的多市场主体自由竞价适配程度高, MASAC 算法总收益为 57.2×10⁴ 美元,

相较于 SAC 算法和 D3QN 算法总收益分别提高了 31.8% 和 40.9%, 由此可见 MASAC 算法可以更好地匹配这类不完全信息环境下的多主体竞价博弈问题。

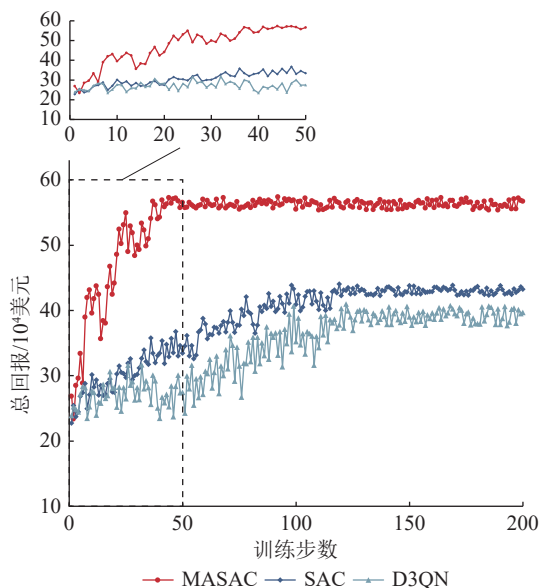


图5 不同算法总回报对比

Fig.5 Comparison of total return of different algorithms

图6和图7为典型日下各市场成员参与能量-调频市场的中标情况。

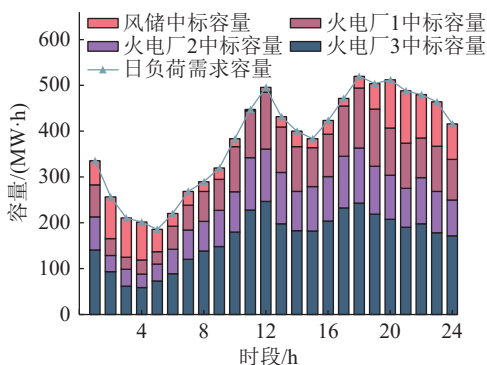


图6 各主体能量市场中标容量及日负荷需求

Fig.6 Winning bid capacity by each entity and daily load demand in the energy market

从图6中可以看出, 在电能量市场中, 仍是火电厂承担了绝大部分负荷需求, 3个火电厂在电能量市场各调度时段累计中标容量共占总负荷需求的 86.9%, 由此可见火电厂这类传统发电资源在电能量市场中仍占据主导地位。相应地, 风储系统在电能量市场中出力占比较低, 仅占 13.1%, 这与风力发电特性相关, 在白天负荷高峰期间风力发电量均较低, 而储能系统受限于自身 SOC、充放电功率特性等因素限制, 无法为风储系统在电能量市场中占据较大份额。

从图7中可以看出, 风储系统在调频市场上中

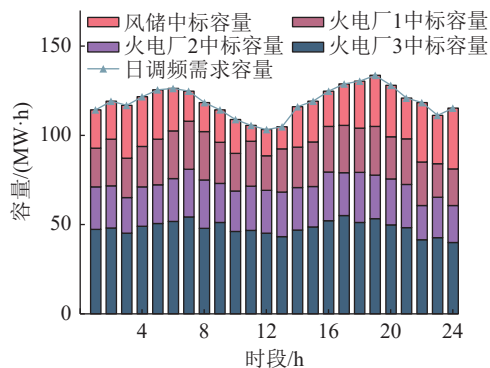


图7 各主体调频市场中标容量及日调频需求

Fig.7 Winning bid capacity by each entity and daily frequency regulation demand in the frequency regulation market

标容量较为稳定, 各调度时段累计中标容量共占总调频需求的 19.2%, 相较于电能量市场, 风储系统在调频市场中能够占据更高份额, 造成该结果的原因是风储系统通过自身配套储能系统, 能够在调频市场考核中获取更高的实时调频性能分数, 使其在竞价排序中能够更容易中标, 在保证风储系统整体经济性的同时, 尽可能参与调频市场, 进一步验证了文中所提风储系统竞价模型的有效性。

图8为风力波动程度高的典型日下风储系统在能量-调频市场中标结果, 可以看出, 即使风电出力波动程度较大, 但风储系统在调频市场中标容量仍较为稳定, 其目的是期望风储系统的实时调频性能分数始终保持在较高水平。风力波动程度高的典型日下的某些时段风力发电量较小, 例如在 11 h 调度时段, 风储系统在电能量市场上的投标量极低, 以保证调频市场中标容量仍能维持原有水平。即使最后实际发电量比原有期望值高, 储能系统也能够及时充电, 保证风电消纳率。而在风力发电量高峰期间, 例如在 1 h 和 21 h 调度时段, 风储系统竞价策略过于保守, 导致中标容量明显低于该时段风电实际发电量。此时, 储能系统通过充电消纳多余风电, 有效避免了弃风惩罚。该风储系统中标结果表明文中所提方法可以在有效提高风储系统的投标收益的同时减少风电弃风现象, 验证了文中所提方法在高风电出力波动情况下的鲁棒性良好, 也体现了风电和储能组成风储系统共同参与电力市场的必要性。

4.3 不同市场竞价模式下风储系统效益分析

为验证文中所提风储系统竞价模型有效性, 设置以下 2 种场景进行比较分析, 市场主体参数及市场需求与 4.1 节和 4.2 节相同。

场景 1: 风储系统联合竞价(文中所提模型)。

场景 2: 风电场和储能系统独立参与竞价。

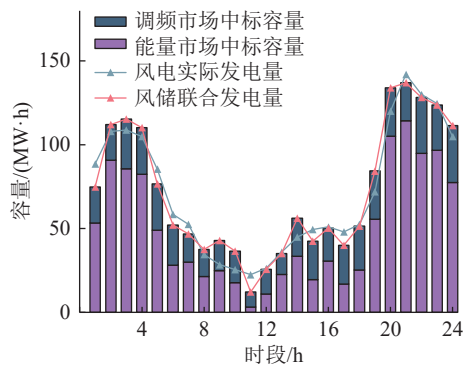


图 8 风储系统在能量-调频市场中标结果
Fig.8 Winning bid results of wind-storage systems in the energy and frequency regulation market

表 5 为不同市场竞价模式下风储系统收益对比,可以看出,场景 1 风储系统收益明显高于场景 2,其原因是风储系统联合竞价可以利用风电场和储能系统的互补特性,在保证市场主体申报体量的同时,平抑风电波动性。此外储能系统与风电场共同参与市场竞价,可以有效提高风储系统调频性能,进而帮助风储系统在价格排序机制中处于优势地位。而在场景 2 中,风电场和储能系统独立参与竞价,风电场受其自身不确定性影响,实际申报过程中竞价策略制定较为保守,而储能系统借其调频性能可以在调频市场中占据优势地位,但受自身 SOC 限制,无法源源不断地进行大体量的调频容量申报,因此不得不舍弃某些时段的竞价收益以维持储能系统 SOC 平衡。综上所述,风储系统联合竞价的参与模式能够有效提高风储系统市场经济性。

表 5 不同场景下风储系统收益对比

Table 5 Comparison of wind-storage system returns under different scenarios

单位: 美元

项目	场景1	场景2	
		风电场	储能
能量市场收益	57 517.345	29 151.688	6 023.301
调频容量收益	37 454.464	18 996.100	6 125.953
调频性能收益	5 349.879	2 931.483	942.658
成本	4 197.842	0	3 561.298
总收益	96 123.846	51 079.271	9 530.614

4.4 实时调频性能分数引入有效性验证

为验证文中所提风储系统竞价策略优势与有效性,设置以下 2 种方案进行比较,市场主体参数及市场需求与 4.1 节和 4.2 节相同。

- 方案 1: 实时调频性能分数模型。
- 方案 2: 固定调频性能分数模型。

图 9 为典型日下风储系统的实时调频性能得分曲线,各时段得分在 0.92~0.98 之间波动。图 10、

图 11 和图 12 分别为典型日不同方案下风储系统能量市场出清价格、调频容量出清价格和调频性能出清价格。通过对比可以发现方案 1 的整体调频容量和调频性能出清价格明显高于方案 2,其原因是方案 1 中的调频市场引入实时调频性能分数机制。方案 2 中,所有市场成员采用固定调频性能分数模型,当风储系统调频市场竞标价格偏高时,由于市场竞争性不强而难以中标,只能通过下调竞标价格以保证其竞争力。长期来看,其他市场成员不得不同样下调竞标价格,导致整体调频容量和调频性能出清价格降低,调频资源参与市场积极性大幅下降。而调频市场的失利同样会使得风储系统在能量市场竞价中采取更为保守的竞价策略,进而导致能量市场出清价格有所下降,各市场主体收益均呈现不同程度的下降。在方案 1 中引入实时调频性能分数模型后,使风储系统竞价策略能够向调频市场倾斜,有效提高风储系统的实时调频性能分数,从而在竞价排序机制中获取优势地位,在稳固调频市场收益的同时保证自身竞争性。

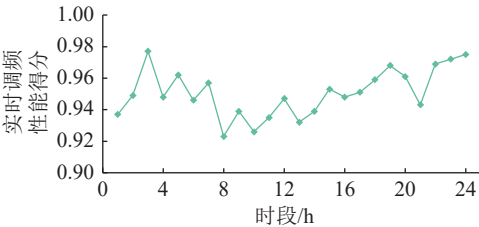


图 9 方案 1 风储系统实时调频性能得分曲线
Fig.9 Real-time frequency regulation performance score curve under scheme 1 for wind-storage system

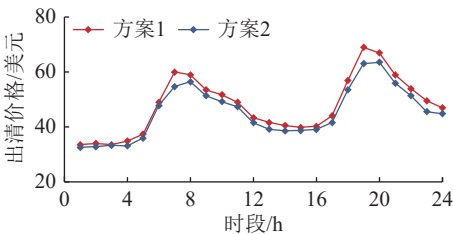


图 10 不同方案下风储系统能量市场出清价格
Fig.10 Clearing prices of energy market under different schemes for wind-storage system

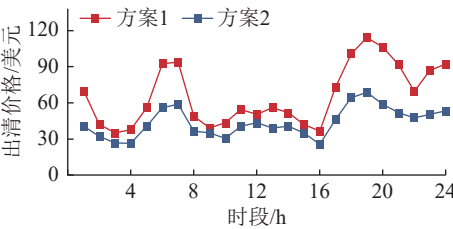


图 11 不同方案下风储系统调频容量出清价格
Fig.11 Clearing prices of frequency regulation capacity under different schemes for wind-storage system

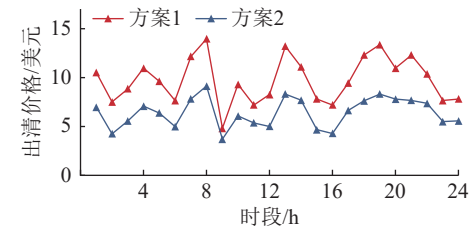


图 12 不同方案下风储系统调频性能出清价格
Fig.12 Clearing prices of frequency regulation performance under different system for wind-storage system

图 13 和图 14 分别为典型日方案 2 能量市场和调频市场中各市场主体中标结果, 图 15 为典型日方案 2 风储系统中中标结果。与方案 1 各市场主体中标情况对比可以发现, 由于方案 2 采取固定调频性能分数机制, 风储系统在调频市场中价格排序机制失去优势地位, 即便风储系统降低报价, 其调频市场中标容量仍不乐观。

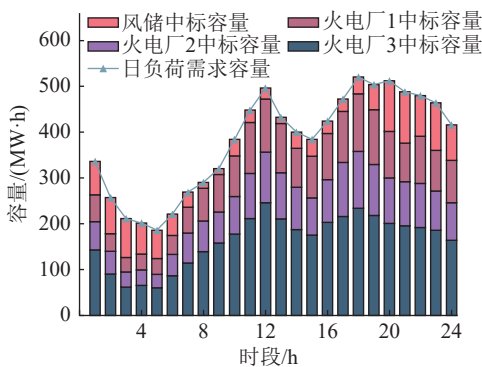


图 13 方案 2 各主体能量市场中标容量及日负荷需求
Fig.13 Winning bid capacity by each entity and daily load demand in the energy market under scheme 2

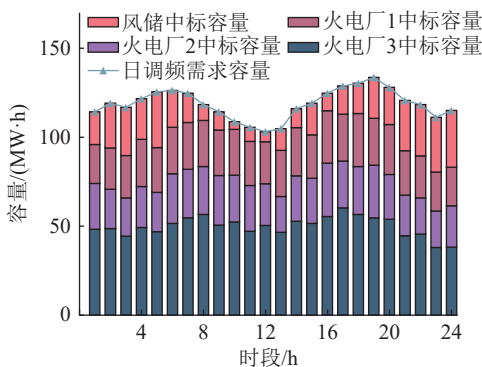


图 14 方案 2 各主体调频市场中标容量及日调频需求
Fig.14 Winning bid capacity by each entity and daily frequency regulation demand in the frequency regulation market under scheme 2

表 6 为不同方案下各市场主体收益对比, 可以看出, 方案 1 各市场主体收益均高于方案 2, 虽然在方案 2 情形下火电厂调频市场中标容量明显高于方案 1, 但由于风储系统被迫调整竞价策略, 导致能

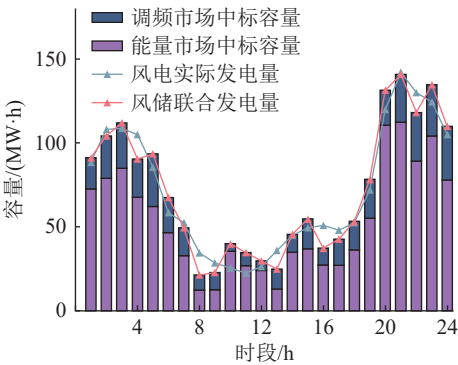


图 15 方案 2 风储系统在能量-调频市场中标结果
Fig.15 Winning bid results of wind-storage systems in the energy and frequency regulation market under scheme 2

量市场价格、调频市场容量价格和性能价格同步下滑, 最终导致方案 2 各市场主体收益均有不同程度下降。这一结果也充分验证了文中所提实时调频性能分数机制的有效性, 通过文中所提模型, 可以使得风储系统总收益和整个系统总收益分别提高 24.6% 和 20.8%, 市场主体总收益的提高不仅能够刺激更多的优质调频资源参与调频市场, 还能促使现有调频资源不断提高自身调频能力以获得更高的调频性能分数, 进而有效提高整个系统调频响应能力。

表 6 不同方案下各主体收益对比

Table 6 Comparison of revenues for each entity under different schemes 单位: 美元

市场主体	方案1收益	方案2收益
风储系统	96 123.846	77 160.852
火电厂1	123 568.017	103 581.180
火电厂2	121 606.332	101 878.673
火电厂3	230 454.624	190 495.086
系统总收益	571 752.819	473 115.791

5 结论

文中研究了基于 MASAC 算法的风储系统参与能量-调频市场竞价模型, 算例结果表明:

(1) 文中构建不完全信息环境下的能量-调频市场交易模型, 用以反映各市场主体报价决策行为与市场出清价格之间的关系, 提出以风储系统取代单独的风电集群参与市场竞价, 该方案能有效改善风储系统调频性能相关指标, 使其在竞价过程中获取优势地位, 从而大幅提高风电消纳率和减少弃风现象。

(2) 文中提出采用 MASAC 算法求解所提竞价模型, 其连续多维的动作空间可以有效避免动

作空间离散化带来的次优解问题,更贴合市场主体的报价决策行为。此外,MASAC算法的多智能系统能够与风储系统、火电厂等多市场主体一一对应,从而有效反映风储系统内部合作关系和风储与火电厂之间的竞争关系,与文中各市场主体间报价决策信息不透明的情形具有高度适配性。通过与SAC算法和D3QN算法进行对比实验,结果表明MASAC算法在求解文中模型上具备独特优势,收敛速度分别提升约53.8%、58.7%,总收益相较于SAC和D3QN两种算法分别提高31.8%和40.9%。

(3)文中引入实时调频性能分数机制并将其应用到调频市场竞标价格排序机制,使得风储系统总收益和整个系统总收益分别提高24.6%和20.8%,该机制使得竞价策略制定更为灵活,既能稳固风储系统作为优质调频资源在调频市场的收益又能保证自身竞争性,从而刺激更多优质调频资源进入市场,提高整个系统的调频响应能力。

文中所提竞价模型用于应对风储系统在不完全信息市场环境下的竞价博弈问题,不仅为风储系统在能量-调频市场中竞价策略的制定提供了参考方案,还对调度中心优化调频资源调用排序和机制起到重要作用,也能优化市场设计,完善市场机制,为企业等其他市场主体提供可持续的市场引导模式。但文中所提模型在高度竞争市场环境下的应用还存在一定局限性,后续研究考虑改进模型求解算法,以提升该模型的工程实用价值。

参考文献:

- [1] 周孝信,赵强,张玉琼,等.“双碳”目标下我国能源电力系统发展趋势分析:绿电替代与绿氢替代[J].中国电机工程学报,2024,44(17):6707-6721.
ZHOU Xiaoxin, ZHAO Qiang, ZHANG Yuqiong, et al. Analysis of the development trend of China's energy and power system under the dual carbon target: green electricity substitution and green hydrogen substitution[J]. Proceedings of the CSEE, 2024, 44(17): 6707-6721.
- [2] 魏旭,刘东,高飞,等.双碳目标下考虑源网荷储协同优化运行的新型电力系统发电规划[J].电网技术,2023,47(9):3648-3661.
WEI Xu, LIU Dong, GAO Fei, et al. Generation expansion planning of new power system considering collaborative optimal operation of source-grid-load-storage under carbon peaking and carbon neutrality[J]. Power System Technology, 2023, 47(9): 3648-3661.
- [3] 黄超,陈湘岳,周琳,等.“双碳”目标下灵活性资源的多维度实时调控模型[J].可再生能源,2023,41(9):1255-1262.
HUANG Chao, CHEN Xiangyue, ZHOU Lin, et al. Multi-dimensional real-time regulation model of flexible resources under 'dual carbon' goals[J]. Renewable Energy Resources, 2023, 41(9): 1255-1262.
- [4] 辛永.“双碳”目标下的电网数字化转型技术研究与应用[J].供用电,2023,40(11):1.
XIN Yong. Research and application of digital transformation technology of power grid under the goal of 'double carbon' [J]. Distribution & Utilization, 2023, 40(11): 1.
- [5] 陈皇森,石立宝.考虑多风电场出力预测误差分布特征的随机机组组合[J].电网技术,2023,47(12):5026-5040.
CHEN Huangsen, SHI Libao. Stochastic unit commitment considering output forecast error distribution characteristics of multiple wind farms[J]. Power System Technology, 2023, 47(12): 5026-5040.
- [6] 王振浩,陈诗伦,葛津铭,等.计及新能源场站调频能力的电力系统最小惯量评估方法[J].太阳能学报,2024,45(8):494-502.
WANG Zhenhao, CHEN Shilun, GE Jinming, et al. Minimum inertia evaluation method of power system considering frequency modulation capability of new energy stations[J]. Acta Energaie Solaris Sinica, 2024, 45(8): 494-502.
- [7] 张龙,孙丹,张旭,等.计及工作点偏移的减载风电系统频率响应模型[J].电力系统自动化,2024,48(16):79-87.
ZHANG Long, SUN Dan, ZHANG Xu, et al. System frequency response model of deloaded wind power system considering working point shifting[J]. Automation of Electric Power Systems, 2024, 48(16): 79-87.
- [8] 负保记,赵文,王建学,等.低碳矿区综合能源系统经济运行优化研究[J].电力系统保护与控制,2024,52(6):177-187.
YUN Baoji, ZHAO Wen, WANG Jianxue, et al. Economic operation optimization of an integrated energy system in low carbon mining area[J]. Power System Protection and Control, 2024, 52(6): 177-187.
- [9] 张大伟,梅晋,王彦洋,等.含新能源场站的变电站无功优化[J].可再生能源,2024,42(10):1416-1420.
ZHANG Dawei, MEI Jin, WANG Yanfeng, et al. Reactive power optimization of substations containing new energy stations[J]. Renewable Energy Resources, 2024, 42(10): 1416-1420.
- [10] 国家市场监督管理总局,国家标准化管理委员会.风电场接入电力系统技术规定第1部分:陆上风电:GB/T 19963.1—2021[S].北京:中国标准出版社,2021.
State Market Regulation Administration, Standardization Administration of the People's Republic of China. Technical specification for connecting wind farm to power system: part 1: on shore wind power: GB/T 19963.1-2021 [S]. Beijing: Standards Press of China, 2021.
- [11] 齐彩娟,车彬,杨燕,等.考虑新能源消纳与储能参与调频的共享储能主从博弈鲁棒定价方法[J].中国电力,2023,56(8):26-39.
QI Caijuan, CHE Bin, YANG Yan, et al. Master-slave game-based robust pricing method of shared energy storage considering renewable energy accommodation and energy storage participating in frequency modulation[J]. Electric Power, 2023,

- 56(8): 26-39.
- [12] 谢开贵, 赵宇生, 胡博, 等. 考虑风电主动参与频率控制的电力系统运行可靠性评估[J]. 电网技术, 2023, 47(1): 41-54.
XIE Kaigui, ZHAO Yusheng, HU Bo, et al. Operational reliability assessment of power system considering active participation of wind power in frequency control[J]. Power System Technology, 2023, 47(1): 41-54.
- [13] WANG T, JIN M X, LI Y D, et al. Adaptive damping control scheme for wind grid-connected power systems with virtual inertia control[J]. IEEE Transactions on Power Systems, 2022, 37(5): 3902-3912.
- [14] 阮益闽, 宗启航, 姚伟, 等. 计及典型控制的风电场调频能力量化评估及影响因素分析[J]. 电力系统自动化, 2024, 48(8): 42-52.
RUAN Yimin, ZONG Qihang, YAO Wei, et al. Quantitative assessment and analysis of influencing factors on frequency regulation capability of wind farms considering typical control[J]. Automation of Electric Power Systems, 2024, 48(8): 42-52.
- [15] KHESHTI M, LIN S Y, ZHAO X W, et al. Gaussian distribution-based inertial control of wind turbine generators for fast frequency response in low inertia systems[J]. IEEE Transactions on Sustainable Energy, 2022, 13(3): 1641-1653.
- [16] 孙东磊, 王宪, 孙毅, 等. 基于多面体不确定集合的电力系统灵活性量化评估方法[J]. 中国电力, 2024, 57(9): 146-155.
SUN Donglei, WANG Xian, SUN Yi, et al. Polyhedral uncertainty set based power system flexibility quantitative assessment[J]. Electric Power, 2024, 57(9): 146-155.
- [17] 刘俊杰, 林济铿. 考虑风电不确定性的短期合同电量协同分解优化模型及算法[J]. 中国电力, 2023, 56(12): 227-237.
LIU Lingjie, LIN Jikeng. Model and algorithm of cooperative optimization decomposition for short-term contract electricity considering wind power uncertainty[J]. Electric Power, 2023, 56(12): 227-237.
- [18] 郑丁园, 崔双喜, 樊小朝, 等. 计及风电不确定性的综合能源系统多目标分布鲁棒优化调度[J]. 智慧电力, 2024, 52(8): 1-8, 18.
ZHENG Dingyuan, CUI Shuangxi, FAN Xiaochao, et al. Multi-objective distributionally robust optimization scheduling for integrated energy system considering wind power uncertainty[J]. Smart Power, 2024, 52(8): 1-8, 18.
- [19] 王娟娟, 王涛, 刘子菡, 等. 考虑风电和负荷不确定性的输电网多目标柔性规划[J]. 中国电力, 2022, 55(1): 168-177.
WANG Juanjuan, WANG Tao, LIU Zihan, et al. Multi-objective flexible planning of transmission network considering wind power and load uncertainties[J]. Electric Power, 2022, 55(1): 168-177.
- [20] 梅笑妍, 陶苏朦, 张瀛文, 等. 基于四象限储能功率线性化模型的配电网协同优化[J]. 南方电网技术, 2024, 18(6): 112-120.
MEI Xiaoyan, TAO Sumeng, ZHANG Yingwen, et al. Coordinated optimization of distribution network based on linear model of four-quadrant energy storage power[J]. Southern Power System Technology, 2024, 18(6): 112-120.
- [21] 赵熙临, 李品, 付波. 基于频率响应特性的储能辅助电网调频方法[J]. 电力工程技术, 2024, 43(1): 41-49, 67.
ZHAO Xilin, LI Pin, FU Bo. Frequency regulation method assisted by energy storage based on frequency response characteristics[J]. Electric Power Engineering Technology, 2024, 43(1): 41-49, 67.
- [22] 孙鑫, 魏文荣, 李琼林, 等. 计及不确定性的风储联合系统多时间尺度市场参与策略[J]. 电力自动化设备, 2024, 44(2): 18-25.
SUN Xin, WEI Wenrong, LI Qionglin, et al. Multi-time scale market participation strategy of wind-energy storage combined system considering uncertainty[J]. Electric Power Automation Equipment, 2024, 44(2): 18-25.
- [23] 赵宗政, 刘一欣, 郭力, 等. 基于剩余需求曲线的风电-储能一体化电站投标决策方法[J]. 电力系统自动化, 2023, 47(8): 99-108.
ZHAO Zongzheng, LIU Yixin, GUO Li, et al. Bidding decision-making method of wind power-energy storage integrated station based on residual demand curve[J]. Automation of Electric Power Systems, 2023, 47(8): 99-108.
- [24] 何翔路, 娄素华, 吴耀武, 等. 双结算模式下风储一体化电站两阶段市场投标调度策略[J]. 电力系统自动化, 2022, 46(4): 47-55.
HE Xianglu, LOU Suhua, WU Yaowu, et al. Two-stage market bidding and scheduling strategy of integrated wind power and energy storage station in dual-settlement mode[J]. Automation of Electric Power Systems, 2022, 46(4): 47-55.
- [25] 杨楠, 李希喆, 刘毅, 等. 电力市场环境下基于多边不完全信息演化博弈的配电网规划方法研究[J]. 电网技术, 2023, 47(11): 4658-4673.
YANG Nan, LI Xizhe, LIU Yi, et al. Distribution network planning based on multi-lateral incomplete information evolutionary game in power market[J]. Power System Technology, 2023, 47(11): 4658-4673.
- [26] 高冠中, 杨胜春, 郭晓蕊, 等. 深度强化学习在含分布式柔性资源的电网优化调度中的应用研究综述[J]. 中国电机工程学报, 2024, 44(16): 6385-6404.
GAO Guanzhong, YANG Shengchun, GUO Xiaorui, et al. A review of research on the application of deep reinforcement learning in optimization dispatch of power grids with distributed flexible resources[J]. Proceedings of the CSEE, 2024, 44(16): 6385-6404.
- [27] 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.
- [28] 翟苏巍, 李文云, 邱振宇, 等. 基于深度强化学习的风储合作决策方法[J]. 智慧电力, 2023, 51(9): 60-65.
ZHAI Suwei, LI Wenyun, QIU Zhenyu, et al. Decision-making

- algorithm in the cooperation of wind power and energy storage based on deep reinforcement learning[J]. *Smart Power*, 2023, 51(9): 60-65.
- [29] 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: 1801.01290. <https://arxiv.org/abs/1801.01290v2>.
- [30] 傅妍芳, 雷凯麟, 魏佳宁, 等. 基于演员-评论家框架的层次化多智能体协同决策方法[J]. 兵工学报, 2024, 45(10): 3385-3396.
- FU Yanfang, LEI Kailin, WEI Jianing, et al. A hierarchical multi-agent collaborative decision-making method based on the actor-critic framework[J]. *Acta Armamentarii*, 2024, 45(10): 3385-3396.
- [31] 柳丹, 任建宇, 席磊, 等. 基于高维协同软演员-评论家的多智能体自动发电控制[J/OL]. 南方电网技术, 2024: 1-12. (2024-10-28)[2024-12-08]. <https://kns.cnki.net/kcms/detail/44.1643.TK.20241028.0851.002.html>.
- LIU Dan, REN Jianyu, XI Lei, et al. Multi-agent automatic generation control based on high-dimensional collaborative soft actor-critic[J/OL]. *Southern Power System Technology*, 2024: 1-12. (2024-10-28)[2024-12-08]. <https://kns.cnki.net/kcms/detail/44.1643.TK.20241028.0851.002.html>.
- [32] 葛晓琳, 凡婉秋, 符杨, 等. 基于改进柔性策略评价的风火储多主体博弈电能-调频市场联合竞价模型[J]. 电网技术, 2023, 47(5): 1920-1930.
- GE Xiaolin, FAN Wanqiu, FU Yang, et al. Joint bidding model of electricity and frequency regulation market with wind fire storage multi-agent games based on improved soft actor-critic[J]. *Power System Technology*, 2023, 47(5): 1920-1930.
- [33] 陈中飞, 荆朝霞, 陈达鹏, 等. 美国调频辅助服务市场的定价机制分析[J]. 电力系统自动化, 2018, 42(12): 1-10.
- CHEN Zhongfei, JING Zhaoxia, CHEN Dapeng, et al. Analysis on pricing mechanism in frequency regulation ancillary service market of United States[J]. *Automation of Electric Power Systems*, 2018, 42(12): 1-10.
- [34] 赵海岭, 王维庆, 李笑竹, 等. 计及储能参与的电能-调频-备用市场日前联合交易决策模型[J]. 电网技术, 2023, 47(11): 4575-4587.
- ZHAO Hailing, WANG Weiqing, LI Xiaozhu, et al. Day ahead joint trading decision model for electricity, frequency regulation and reserve market considering energy storage participation[J]. *Power System Technology*, 2023, 47(11): 4575-4587.

作者简介:



李钟平

李钟平(2000), 男, 硕士在读, 研究方向为风储系统联合调度(E-mail: zhongpingli1@foxmail.com);

向月(1987), 男, 博士, 教授, 通信作者, 研究方向为电力系统规划与运行优化。

Deep reinforcement learning-driven bidding strategy for wind-storage systems in energy and frequency regulation markets

LI Zhongping, XIANG Yue

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

Abstract: In the power market environment, participation of wind-storage system in both the energy market and the frequency regulation market is essential to enhance economic efficiency and support grid frequency regulation and peak shaving. However, key issues such as formulating bidding strategies for wind-storage systems in energy-frequency regulation dual markets need to be addressed. A bidding model driven by deep reinforcement learning is proposed in this paper to formulate bidding strategies in an incomplete information market environment. Firstly, a framework for wind-storage systems participating in the energy and frequency regulation markets is established to clarify the bidding operation strategies of each market entity. Then, a real-time frequency regulation performance scoring model is introduced to address the differences in frequency regulation response capabilities among various resources. Based on this, a bidding model for wind-storage systems is developed. Finally, a multi-agent deep reinforcement learning method with strong model-free learning capabilities is employed to solve the stochastic game problem in an incomplete information market environment and to handle the multi-agent bidding game relationship. Simulation results indicate that the proposed method can effectively formulate bidding strategies for wind-storage systems participating in the energy and frequency regulation markets. The method achieves high returns while ensuring high convergence stability. As a result, the economic efficiency of wind-storage systems is enhanced, and grid frequency regulation and peak shaving are effectively supported.

Keywords: wind-storage system; energy and frequency regulation market; deep reinforcement learning; real-time frequency regulation performance scoring; actor-critic; multi-entity bidding game

(编辑 吴昊)