

基于DeepSets-TD3的多障碍井约束井眼轨迹智能控制方法

马伟民, 黄新武*, 王 瑜

(中国地质大学(北京)工程技术学院, 北京, 100083)

摘要: 针对密集井网条件下井眼轨迹控制面临的多邻井约束、障碍物输入顺序敏感和连续动作稳定控制问题, 本文提出DeepSets-TD3多障碍井约束井眼轨迹智能控制方法。该方法以最小曲率法递推井眼轨迹, 将井斜角和方位角增量作为连续控制动作; 以靶点相对位移、井斜角及方位角正余弦构成基础状态, 以障碍井表面净距及相对方位正余弦构成集合元素, 并通过共享编码网络与均值-最大值对称池化形成置换不变表征。在固定四障碍布置测试中, DeepSets-TD3的平均目标误差为7.52 m, 物理碰撞率为0。随机四障碍测试中, DeepSets-TD3的平均目标误差为11.2 m, 物理碰撞率为1.2%, 近失误缓冲区率为1.9%, 狗腿严重度合规率为99.6%, 且四项指标均优于普通TD3。结果表明, 集合表征能够减弱障碍井排列变化引起的策略退化, 提高复杂邻井约束下轨迹控制的精度、安全性与泛化能力。

关键词: 井眼轨迹控制; 多障碍井; 邻井防碰; 深度强化学习; TD3; DeepSets; 最小曲率法; 智能控制

中图分类号: P634.7; TE243 **文献标识码:** A **文章编号:** 2096-9686(2026)S1-0102-10

Intelligent wellbore trajectory control method for multiple obstacle-well constraints based on DeepSets-TD3

MA Weimin, HUANG Xinwu*, WANG Yu

(School of Engineering and Technology, China University of Geosciences, Beijing 100083, China)

Abstract: Aiming at the challenges in borehole trajectory control under dense well patterns, including constraints from multiple adjacent wells, sensitivity to the input order of obstacles, and stable control of continuous actions, this paper proposes an intelligent control method for multi-obstacle well constrained borehole trajectory named DeepSets-TD3. The method recursively calculates the borehole trajectory using the minimum curvature method, and takes the increments of inclination and azimuth as continuous control actions. The basic state consists of the relative displacement of the target, inclination angle, as well as the sine and cosine values of azimuth angle. Each element in the set is constructed by the surface clear distance of obstacle wells and the sine and cosine values of relative azimuth. A permutation-invariant representation is formed via a shared encoder network and mean-max symmetric pooling. In the test with four fixed obstacles, the DeepSets-TD3 achieves an average target error of 7.52 m and zero physical collision rate. In the random four-obstacle test, the average target error is 11.2 m, the physical collision rate is 1.2%, the near-miss buffer rate is 1.9%, and the compliance rate of dogleg severity reaches 99.6%. All four indicators outperform the conventional TD3. The results demonstrate that the set representation can mitigate policy degradation caused by

收稿日期: 2026-07-31; 修回日期: 2026-09-15 DOI: 10.12143/j.ztgc.2026.S1.011

基金项目: 地球深部探测与矿产资源勘查国家科技重大专项“绿色高效精准智能化钻探技术与装备”所属专题课题“高效精准智能化钻探技术集成研究”(编号: 2024ZD1003105.4)

第一作者: 马伟民, 男, 汉族, 2003年生, 硕士研究生, 地质工程专业, 主要从事钻探自动化与智能化控制等方面的研究工作, 北京市海淀区学院路29号, 15646501952@163.com。

通信作者: 黄新武, 男, 汉族, 1972年生, 副教授, 地质工程专业, 博士, 主要研究方向为钻探自动化与智能化、工程物探、资源与环境勘查, 北京市海淀区学院路29号, hwx666@hotmail.com。

引用格式: 马伟民, 黄新武, 王瑜. 基于DeepSets-TD3的多障碍井约束井眼轨迹智能控制方法[J]. 钻探工程, 2026, 53(S1): 102-111.

MA Weimin, HUANG Xinwu, WANG Yu. Intelligent wellbore trajectory control method for multiple obstacle-well constraints based on DeepSets-TD3[J]. Drilling Engineering, 2026, 53(S1): 102-111.

changes in the arrangement of obstacle wells, and improve the accuracy, safety and generalization capacity of trajectory control under complex constraints of adjacent wells.

Key words: wellbore trajectory control; multiple obstacle wells; anti-collision of adjacent wells; deep reinforcement learning; TD3; DeepSets; minimum curvature method; intelligent control

0 引言

水平井、分支井和工厂化井组提高了储层动用程度,也显著增加了井眼轨迹设计与随钻控制的复杂度。智能钻井技术正由单点自动化向感知、决策与执行闭环演进,井眼轨迹控制需要同时处理靶点命中、井身质量、狗腿度限制、邻井防碰和导向工具可执行性等约束^[1]。

近年来,非直井轨迹研究逐步从固定井型参数优化转向多目标、实时和自适应控制。顾海鹏^[2]围绕多目标轨迹优化与随钻自适应校正构建连续控制框架,并采用双延迟深度确定性策略梯度(Twin Delayed Deep Deterministic Policy Gradient, TD3)直接输出井斜角和方位角调整量。

在离线轨迹规划方面,Liu等^[3]利用三维Dubins曲线描述定向井轨迹,以曲率约束和路径长度为核心开展海底井场轨迹规划;D'Angelo等^[4]将方向钻井建议表述为Pareto多目标决策问题,使轨迹长度、井身质量和工程约束能够在同一框架内权衡。

面向密集井网,Wang等^[5]建立多轨迹协同优化模型,重点处理井口布局、轨迹长度、轨迹相似性与碰撞风险之间的耦合;Dou等^[6]进一步将测量误差椭圆与压裂影响域耦合为复合障碍体,用于页岩气三维开发条件下加密井轨迹的抗压裂干扰设计。这类方法增强了复杂约束表达能力,但多依赖离线求解或预先给定的障碍参数,难以直接应对随钻状态变化和邻井集合动态更新。

强化学习为连续闭环控制提供了另一条路径。Zhu等^[7]提出目标感知的迁移强化学习方法,以提高地质扰动条件下的轨迹跟踪与自适应能力。Gu等^[8]将TD3用于井眼轨迹优化与钻进控制,并对状态、动作和奖励函数进行了工程化建模。Li等^[9]在Soft Actor-Critic框架中引入自注意力机制,用于增强复杂地层状态特征提取。

在防碰与泛化方面,Zhu等^[10]构建深度循环Q网络轨迹设计框架,将碰撞规避、狗腿度和目标进入参数纳入奖励函数,并与粒子群优化(Particle Swarm Optimization, PSO)方法进行比较;许秋皓^[11]

通过域随机化和渐进式学习提高井眼轨迹控制对地层变化的适应性;樊宸^[12]利用空间注意力机制提取三维障碍环境特征,为复杂障碍场的风险感知提供了另一种思路。

现有研究仍存在一个容易被忽略的建模问题:多障碍井在数学上属于无序集合,而常规神经网络往往按障碍井编号直接拼接特征。井号或输入槽位并不具有稳定的物理语义,同一组障碍井仅因排列顺序改变就可能引起状态语义错位。DeepSets通过共享元素编码器和对称聚合函数获得置换不变表征,为多邻井风险信息建模提供了直接方法^[13]。

基于此,本文将DeepSets集合编码器嵌入TD3的策略网络(Actor)和双价值网络(Critic),构建DeepSets-TD3多障碍井约束轨迹控制方法。TD3通过双价值网络、目标策略平滑和延迟策略更新缓解连续控制中的价值过估计问题^[14]。

试验以深度确定性策略梯度(Deep Deterministic Policy Gradient, DDPG)、普通TD3和PSO作为参照,其中DDPG是确定性连续动作强化学习的典型基线^[15]。本文设置无障碍、单障碍、双障碍、四障碍固定布置、随机四障碍和困难四障碍等递进试验,重点检验集合表征能否削弱障碍井输入顺序敏感性,并在障碍物位置和半径变化时保持到靶精度、避碰可靠性与井身质量。

1 实验方案与评价指标

本文以四项指标作为随机四障碍场景的核心评价依据(表1)。物理碰撞率表示轨迹进入障碍井实体范围的回合比例;平均目标误差表示未发生物理碰撞回合终点到目标点三维距离的算术平均值;近失误缓冲区率表示未发生实体碰撞但最小净距落入0~20 m缓冲区的回合比例;狗腿严重度(Dogleg Severity, DLS)合规率表示满足DLS上限的轨迹点比例。平均最小净距和安全违规率作为补充诊断指标,用于分析风险和安全裕度。

DeepSets-TD3和普通TD3均采用5个训练随

表1 四项核心评价指标

指 标	定 义	工 程 含 义
物理碰撞率 P_{col}	进入障碍井实体范围的回合比例	邻井防碰可靠性
平均目标误差 d_{mean}	终点到目标点三维距离的算术平均值	总体到靶精度
近失误缓冲区率 P_{near}	$0 < \text{最小净距} \leq 20 \text{ m}$ 的回合比例	近碰风险与安全裕度
DLS 合规率 P_{DLS}	满足 DLS 上限的轨迹点比例	井身质量与可施工性

机种子,即1、42、100、500和1024,每个随机种子下模型均训练10000回合,后续评价回合模型参数不再更新。每个种子在每种固定布置下评价100回合,两种方法各形成15组场景统计和1500个评价回合。随机四障碍测试中,两种方法每个种子均评价1000回合,即每种方法共5000回合;困难四障碍测试仅对DeepSets-TD3进行5000回合压力评价。各场景保持起点、着陆点、靶点、造斜点(kickoff point, KOP)、步长、动作边界和DLS约束一致,以减弱非算法因素对比较结果的影响。

2 多障碍井约束轨迹控制模型

2.1 最小曲率法轨迹递推

井眼轨迹由测深(Measured Depth, MD)、井斜角 I 、方位角 A 、北向位移 N 、东向位移 E 和真垂深(True Vertical Depth, TVD)描述。相邻测点之间采用最小曲率法递推,使曲线段近似为空间圆弧。

$$\cos\beta = \cos I_k \cos I_{k+1} + \sin I_k \sin I_{k+1} \cos \Delta A \quad (1)$$

$$RF = \begin{cases} \frac{2}{\beta} \tan\left(\frac{\beta}{2}\right), & \beta > 0, \\ 1, & \beta \approx 0, \end{cases} \quad (2)$$

$$\Delta N = \frac{\Delta MD}{2} (\sin I_k \cos A_k + \sin I_{k+1} \cos A_{k+1}) RF \quad (3)$$

$$\Delta E = \frac{\Delta MD}{2} (\sin I_k \sin A_k + \sin I_{k+1} \sin A_{k+1}) RF \quad (4)$$

$$\Delta TVD = \frac{\Delta MD}{2} (\cos I_k + \cos I_{k+1}) RF \quad (5)$$

$$DLS = \beta \frac{180}{\pi} \cdot \frac{30}{\Delta MD} \quad (6)$$

式中: β ——相邻测点的空间狗腿角; ΔA ——方位角增量; RF ——曲率半径校正因子; ΔMD ——相邻测点的测深增量; I_k, I_{k+1} ——第 $k, k+1$ 个测点的井斜角; ΔN ——相邻测点间的北向位移增量; ΔE ——相邻测点间的东向位移增量; ΔTVD ——相邻测点间的垂深增量。

仿真采用三段式水平井:直井段保持井斜角接

近 0° ,造斜段由 KOP 逐步增斜并趋近着陆点,水平段保持井斜角接近 90° 并沿目标方位到达终点。起点为 $(0, 0, 0) \text{ m}$,着陆点为 $(495, 495, 1000) \text{ m}$,靶点为 $(1500, 1500, 1000) \text{ m}$,KOP 为 300 m ,目标方位为 45° 。垂直段步长为 20 m ,造斜点之后步长为 10 m ,最终接近目标时,最后一步缩短为剩余距离。造斜段和水平段采用 10 m 离散步长,是为了近似一根钻杆单根的长度,并在具有实际工程意义的空间尺度上解析曲率控制和避障决策。初始垂直段采用较粗的 20 m 步长,是因为造斜点之前尚未开始主动转向,且通常不会发生障碍物交互,因此可以减少不必要的离散化,最大测深 5000 m ,DLS 上限为 $6^\circ/30 \text{ m}$ 。控制动作为 $a = [\Delta I, \Delta A]$,其中 $\Delta I \in [-1.2^\circ, 1.2^\circ]$, $\Delta A \in [-1.8^\circ, 1.8^\circ]$ 。

2.2 TD3算法与奖励函数

TD3使用双价值网络、目标策略平滑和延迟策略更新。目标值与策略网络损失(Actor)分别为:

$$y = r + \gamma \min_{j=1,2} Q'_j(s', \pi'(s') + \epsilon) \quad (7)$$

$$L_\pi = -\mathbb{E}_{s \sim D} [Q_1(s, \pi(s))] \quad (8)$$

式中: δ ——终止标志,终止转移取1,否则取0; ϵ ——经幅值裁剪的目标策略噪声; D ——经验回放池; y ——TD3目标值; r ——当前转移获得的即时奖励; Q'_j ——第 j 个目标价值网络; s ——当前状态; s' ——下一状态; L_π ——策略网络损失函数; $\mathbb{E}_{s \sim D}$ ——经验回放池 D 中采样状态的期望; π' ——目标策略网络。

折扣因子 γ 为 0.97 ,软更新系数 τ 为 0.005 ,策略噪声标准差为 0.2 ,噪声裁剪幅值为 0.5 ,策略延迟为 2 ,Actor 和 Critic 学习率均为 3×10^{-4} 。训练回合数为 10000 ,回放池容量为 200000 ,批量大小为 256 ,起始随机探索步数为 5000 。

奖励函数同时考虑向目标推进、着陆点接近、井斜与方位引导、DLS 限制、动作平滑、障碍井安全边界和实体碰撞终止:

$$r = r_{\text{prog}} + r_{\text{target}} + r_{\text{att}} + r_{\text{DLS}} + r_{\text{smooth}} + r_{\text{obs}} + r_{\text{collision}} \quad (9)$$

式中: r_{prog} ——目标推进; r_{target} ——终点接近; r_{att} ——井斜与方位引导; r_{DLS} ——狗腿严重度约束; r_{smooth} ——动作平滑; r_{obs} ——软安全边界; $r_{\text{collision}}$ ——实体碰撞项。

障碍井约束采用硬碰撞与软安全边界结合的方式。当轨迹进入障碍井实体半径时立即判定碰撞并终止;当净距小于软安全裕度但尚未碰撞时,根据接近程度施加连续惩罚,从而在碰撞发生前形成可学习的风险梯度。

2.3 状态空间与障碍井集合表征

基础状态由目标相对位移、当前井斜角和当前方位组成。为避免方位角在 0° 和 360° 附近产生不连续,方位角采用正弦和余弦编码:

$$s_b = \left[\frac{\Delta N}{2000}, \frac{\Delta E}{2000}, \frac{\Delta TVD}{2000}, \frac{I}{90^\circ}, \sin A, \cos A \right] \quad (10)$$

对于第 i 口障碍井,环境计算当前钻头点到障碍井表面的净距 d_i 和相对方位 θ_i 。 d_s 为距离归一化尺度, $\text{clip}(\cdot)$ 用于限制极端距离值。单个障碍井特征定义为:

$$x_i = \left[\text{clip}\left(\frac{d_i}{d_s}, -1, 2\right), \sin\theta_i, \cos\theta_i \right] \quad (11)$$

普通 TD3 将 N 口障碍井特征按输入顺序直接拼接,四障碍场景下观测维度为 18。该结构隐含了不同输入槽位具有固定语义的假设。DeepSets-TD3 则将障碍井表示为集合 $X = \{x_1, x_2, \dots, x_N\}$, 对每个元素使用共享编码器 ϕ , 并通过均值池化与最大池化形成全局上下文:

$$z_i = \phi(x_i), \quad c = \rho \left(\left[\text{mean}_i z_i, \max_i z_i \right] \right) \quad (12)$$

共享编码保证不同障碍井接受同一特征判别规则;均值池化反映整体风险水平,最大值池化突出局部高风险障碍井。二者拼接后供 Actor 和双 Critic 共同使用,使策略关注净距、相对方位和整体障碍场,而不是人为编号顺序,DeepSets-TD3 的轨迹控制流程如图 1 所示。

2.4 域随机化与测量误差

由于已有井轨迹通常大量交错分布,若逐一建模将大幅提升观测空间维度。本文将空间邻近的多条井轨迹聚类为半径 50 m 的竖直圆柱体作为障碍井,用简化后的圆柱集合表征井群危险区域,降低多障碍物场景下轨迹规划的计算负担。在训练阶段对

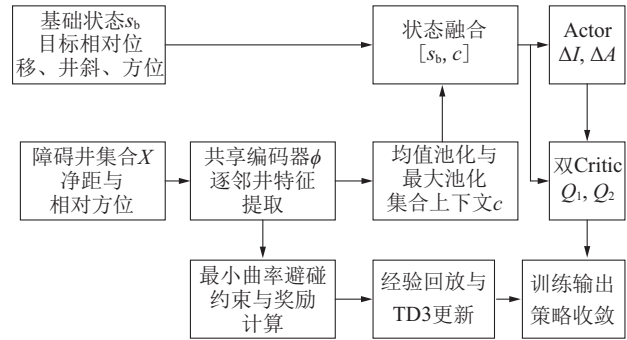


图1 DeepSets-TD3多障碍井约束轨迹控制流程

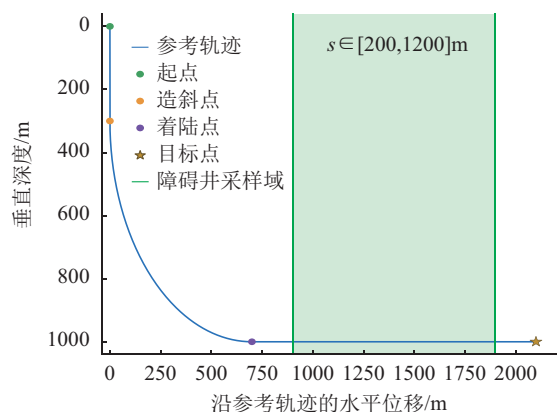
障碍井位置进行域随机化。如图2所示,分别从垂向剖面 and N-E 平面展示参考井眼轨迹与障碍井采样域。以着陆点为纵向坐标 $s=0$, 障碍井沿 45° 名义路径方向在 $s \in [200, 1200]$ m 内采样, 垂直于名义路径的横向偏移 $l \in [-200, 200]$ m; 障碍井半径在训练阶段固定为 50 m, 软安全裕度为 80 m。困难四障碍评价将障碍井半径扩展为 50~80 m, 并设置 30 m 最小安全净距阈值, 用于考察可行空间压缩后的鲁棒性。

在现场, 控制器接收到的钻头姿态会受到随钻测量(MWD)系统测量误差的污染。本文通过向观测到的井斜、方位角和测深加入零均值高斯噪声来建模这一影响: 智能体完全在无噪声条件下训练($\sigma_0 = \sigma_\varphi = \sigma_{\text{md}} = 0$), 但在 MWD 的噪声下评估, 其中 $\sigma_0 = 0.1^\circ$, $\sigma_\varphi = 0.2^\circ$, $\sigma_{\text{md}} = 0.5$ m。TD3 既面对未见过的障碍井布局, 又面对训练期间从未遇到过的测量扰动。因此, 在该条件下的稳健表现提供了更严格的仿真鲁棒性测试。

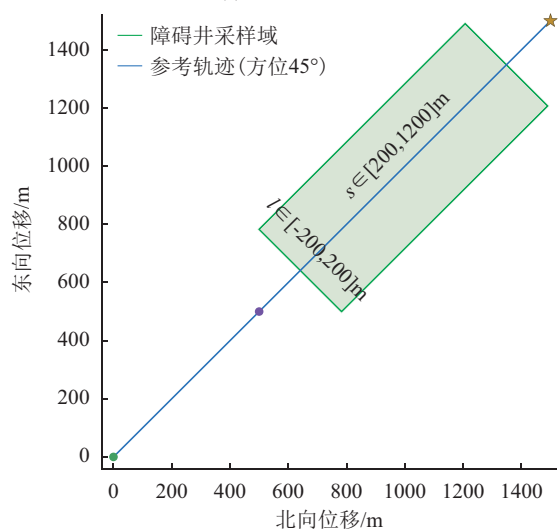
3 实验结果与分析

3.1 基础实验: 无障碍、单障碍与双障碍

无障碍实验用于验证连续轨迹控制算法在不受障碍约束时的基础目标命中能力。图3给出了DDPG与普通TD3在3个随机种子(42、100、1024)下的训练奖励和历史最优目标误差收敛过程, 半透明区域表示跨种子波动区间。总体来看, 两种算法均能逐步提高回合奖励并降低目标误差。相比DDPG, TD3的奖励提升过程相对平缓, 但其历史最优目标误差在训练后期波动更小, 表现出更稳定的目标逼近特征。评价结果显示, TD3的平均目标误差为 10.70 m, 低于DDPG的 18.31 m。考虑到井眼轨迹控制更关注目标命中精度、跨随机种子的稳



(a)侧视图(垂向剖面)



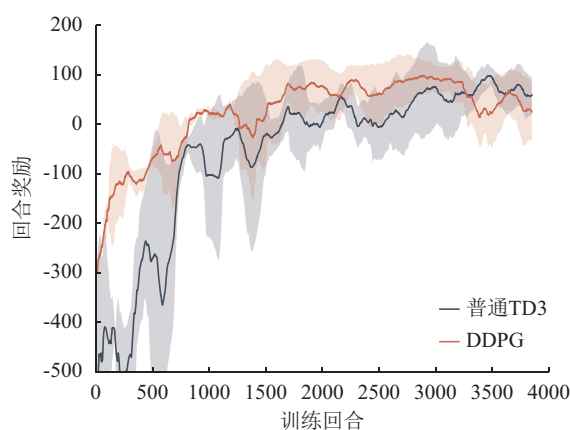
(b)俯视图(N-E平面)

图2 环境建模中的障碍井域随机化范围示意

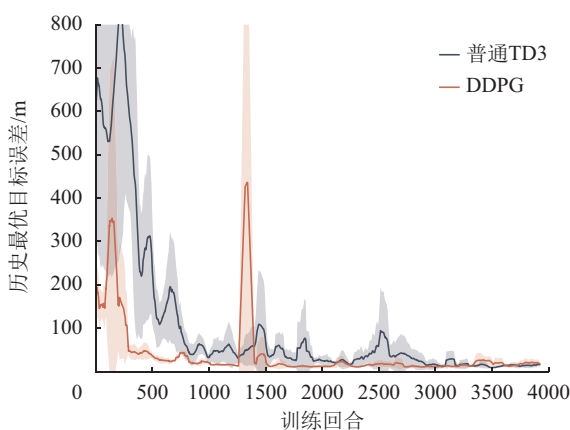
定性以及后续障碍约束场景中的工程可靠性,后续实验选取TD3作为基础连续控制算法。在算法层面,TD3的双Q保守估计和延迟策略更新有助于减弱连续动作空间中的价值过估计,从而提高策略训练的稳定性。

单障碍实验用于检验障碍井软惩罚项与碰撞终止机制对避障行为的约束效果。图4展示了障碍物分别位于水平段前部、中部和末端时的代表性轨迹。评价结果显示,单障碍TD3在三种障碍位置下均未发生物理碰撞;其中,障碍物位于水平段中部的平均目标误差为4.07 m,平均最小净距为28.16 m。上述结果表明,在障碍数量较少且输入维度固定的条件下,普通TD3能够在保持目标逼近精度的同时学习基本避让动作,为后续多障碍场景实验提供了基线参照。

双障碍随机实验用于评估算法从单障碍避让向



(a)训练奖励变化



(b)历史最优目标误差变化

图3 无障碍物场景下训练奖励与历史最优靶距变化

多障碍随机场景扩展时的稳定性。图5给出了普通TD3与PSO在随机双障碍场景下的物理碰撞率、目标误差、近失误缓冲区率和DLS合规率对比。评价结果显示,普通TD3在3个随机种子(42、100、1024)、共3000回合评价中的目标误差为17.0 m,物理碰撞率为1.0%;相比之下,PSO的目标误差为68.7 m,物理碰撞率为6.0%。同时,PSO的近失误缓冲区率为7.2%,高于普通TD3的1.8%,表明其在随机双障碍场景中更容易出现贴近障碍的轨迹。综合目标误差、碰撞率和近失误缓冲区率结果,普通TD3相比PSO表现出更好的目标逼近能力和避碰稳定性。

3.2 四障碍固定布置:测试输入顺序敏感性

为直观说明两种方法在固定布置下的行为差异,固定四障碍测试包括S1_BAC、S2_CBA和S3_CAB三组布置。为了使定性比较具有诊断性而不是仅展示个例,本文根据三类障碍物角色构造了

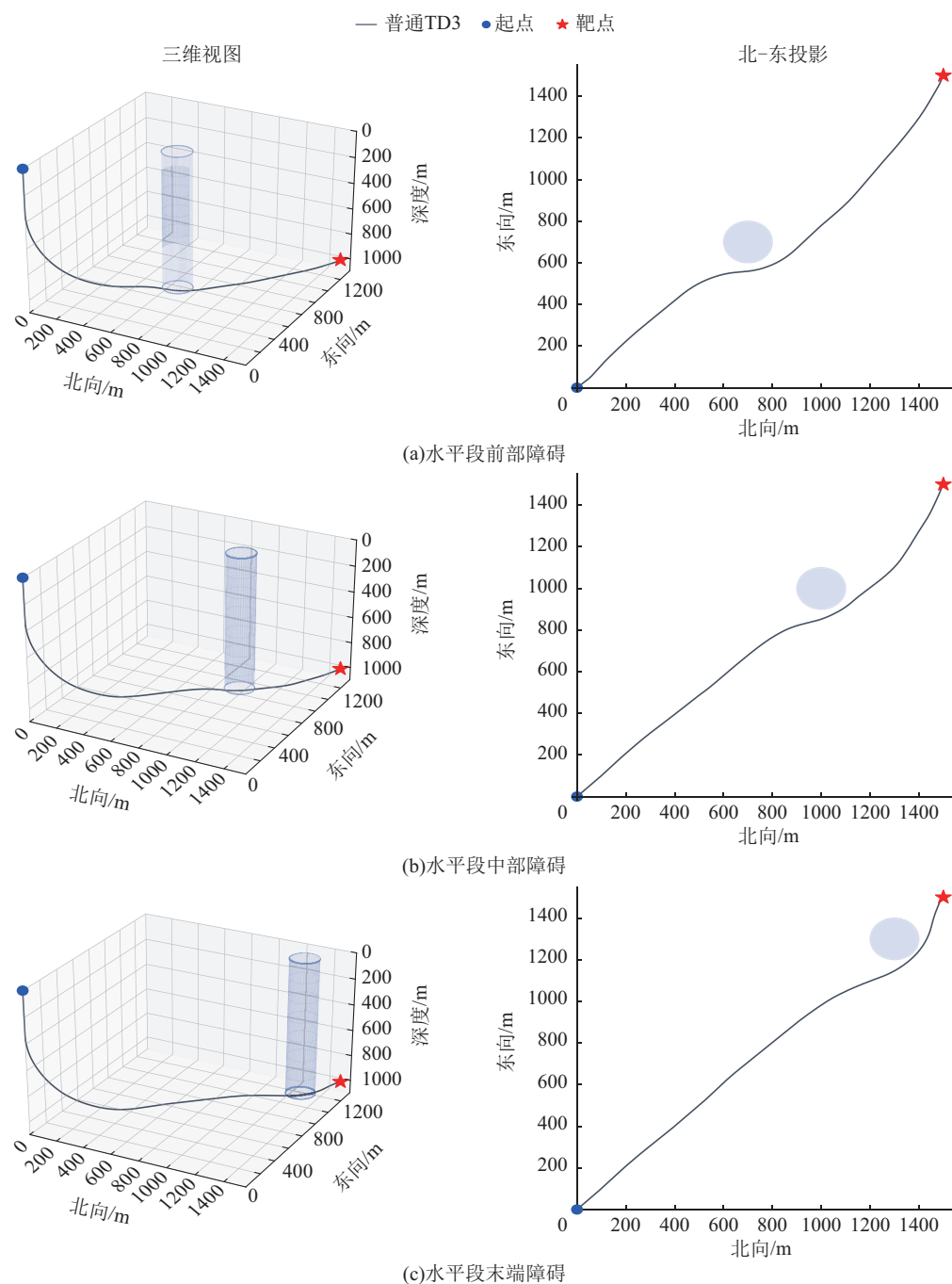


图4 单障碍物不同位置下的代表性轨迹智能控制结果

三组固定布置。在该记号中,A表示偏离名义最短路径的障碍物,B表示位于名义最短路径上的障碍物,C表示由两个障碍物形成的中间狭窄通道。S1_BAC、S2_CBA和S3_CAB通过排列这些角色,测试策略是否能够找到安全且高效的通道,而不是简单沿起点至目标对角线前进。

图6给出了三种场景的代表性轨迹智能控制结果。在该轨迹可视化中,每个障碍物以100 m可视

化包络绘制,该包络由50 m物理障碍物半径和50 m可视化安全缓冲区组成;碰撞判据和压力测试安全间距阈值仍按上文定义分别计算。可以看到,DeepSets-TD3通常会在进入障碍密集区之前完成横向调整,随后绕开位于名义最短路径上的障碍物并选择更安全的通道;普通TD3则更容易沿起点—目标对角线继续前进,在S2_CBA和S3_CAB中表现出明显的迟滞避障或提前终止现象。

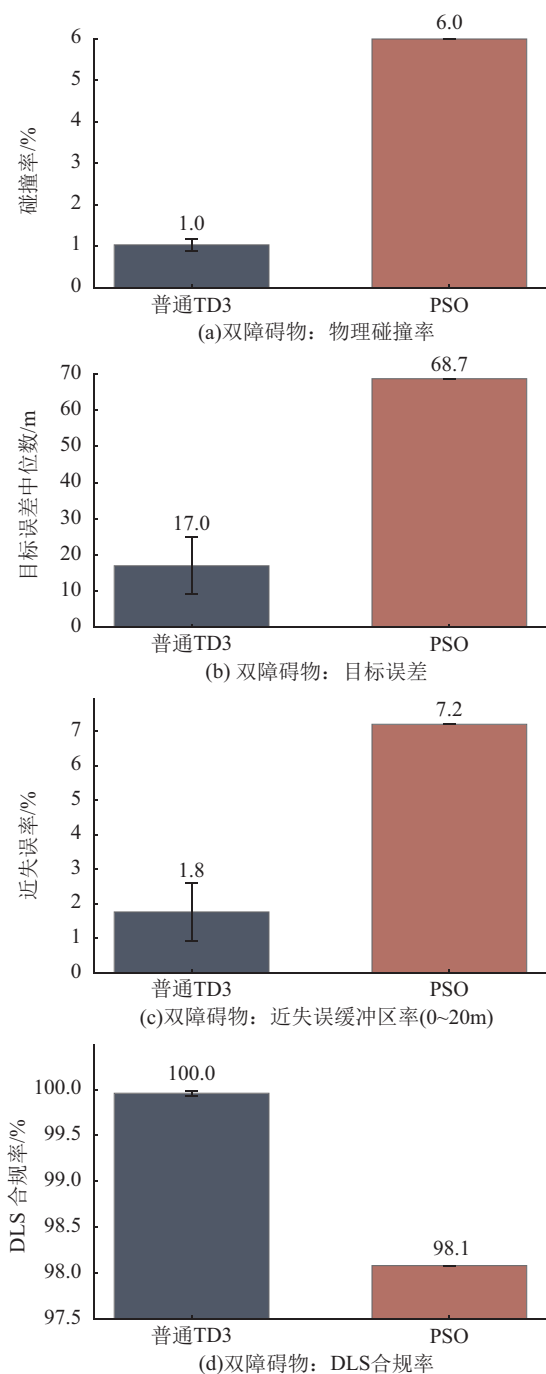


图5 双障碍场景下普通TD3与PSO的核心指标对比

DeepSets-TD3在种子1、42、100、500和1024下的三组布置平均目标误差分别为8.67、7.66、6.09、7.25和7.91 m,所有种子的物理碰撞率均为0;最大单场景平均目标误差为10.98 m,出现在S1_BAC且种子为1时,最小值为5.67 m,出现在S3_CAB且种子为100时。相比之下,普通TD3在

固定四障碍布置下的平均目标误差为370.02 m,物理碰撞率为26.67%,表现出明显策略退化。

固定布置评价中,单个场景内的障碍物几何位置和输入顺序保持不变,但模型参数在评价阶段不再更新,因而普通TD3只能调用训练阶段形成的状态-动作映射。由于普通TD3将障碍井特征按顺序直接拼接,其Actor和Critic中的多层感知机(MLP)易将输入槽位与特定障碍角色绑定;当测试场景中的障碍角色组合与训练过程中形成的槽位语义不一致时,就会出现输入特征错配,进而导致迟滞避障、绕行过度或提前终止。DeepSets-TD3采用共享元素编码和对称池化,避免将障碍物风险与人为输入编号绑定,因此在不同固定障碍角色排列下保持了更稳定的靶点逼近和避障表现。

3.3 随机四障碍测试与困难四障碍泛化评价

随机四障碍测试与训练阶段的障碍采样分布一致,但每个测试回合重新生成障碍井位置。图7给出了普通TD3与DeepSets-TD3在物理碰撞率、平均目标误差、近失误差缓冲区率和DLS合规率四项核心指标上的对比。评价结果显示,普通TD3的物理碰撞率、平均目标误差、近失误差缓冲区率和DLS合规率分别为22.1%、106.4 m、10.1%和99.0%;DeepSets-TD3对应指标分别为1.2%、11.2 m、1.9%和99.6%。同时,DeepSets-TD3的平均最小净距为67.6 m,高于普通TD3的55.9 m。结果表明,DeepSets-TD3在降低实体碰撞和近失误差风险的同时,提高了有效轨迹的到靶精度,并保持了较高的井身质量约束满足程度。

困难四障碍测试进一步提高障碍井半径,并加入30 m最小安全净距阈值,以检验可行空间收缩条件下的鲁棒性。在该条件下,DeepSets-TD3的物理碰撞率为2.4%,平均目标误差为13.3 m,近失误差缓冲区率为3.1%,DLS合规率为99.6%,平均最小净距为61.0 m。与随机四障碍测试相比,碰撞率和近失误差缓冲区率有所增加,平均目标误差略有上升,平均最小净距有所下降,说明障碍半径增大和安全边界收紧会提高避障难度;但DeepSets-TD3仍能维持较低的目标误差、较高的DLS合规率和一定的平均安全净距,表明其在可行空间压缩后仍具有一定鲁棒性。

跨随机种子的误差线表示不同训练随机种子所得模型在随机四障碍测试中的评价波动。尽管

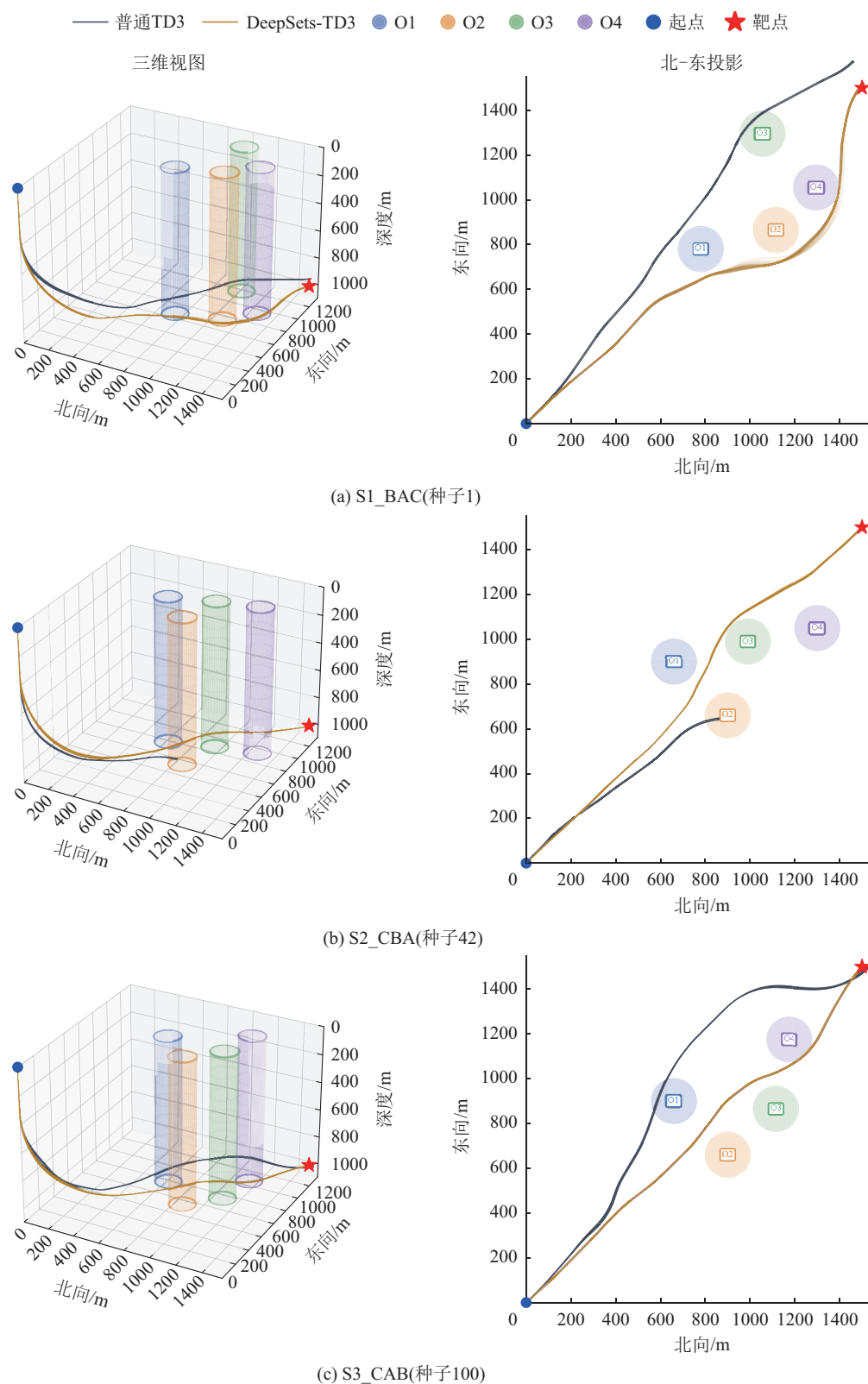


图6 固定四障碍布置下的代表性轨迹智能控制结果

不同随机初始化和训练采样会导致模型性能存在差异,DeepSets-TD3在物理碰撞率、平均目标误差

和近失误缓冲区率上的优势仍保持稳定,说明其性能提升并非依赖某一特定随机初始化。对于井眼

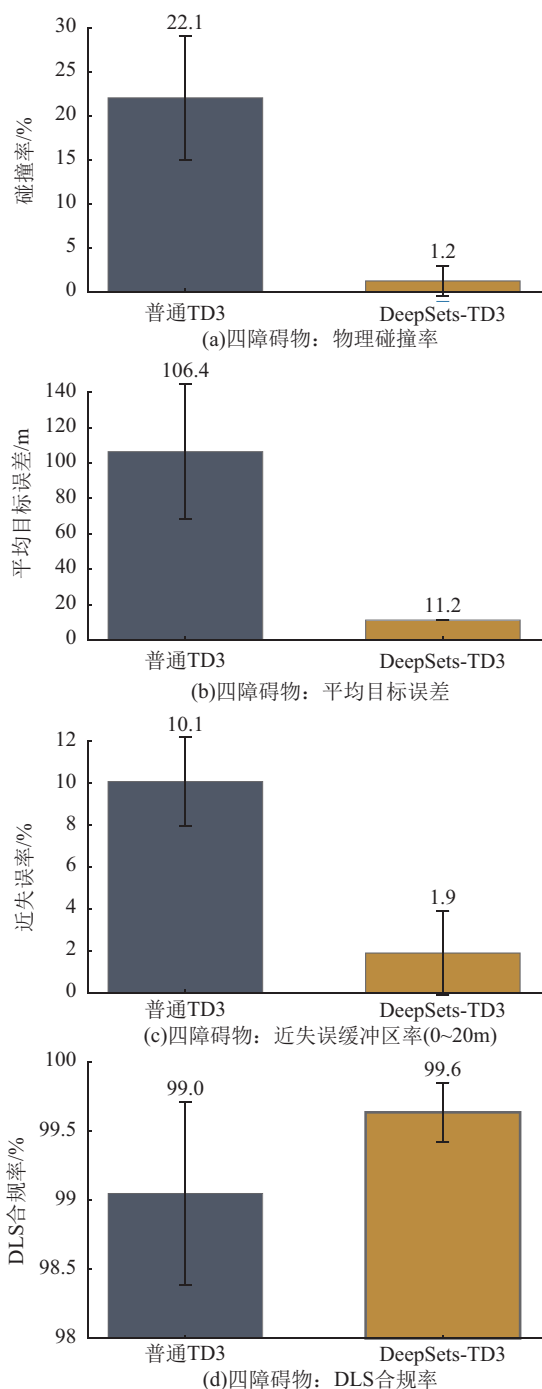


图7 随机四障碍场景下普通TD3与DeepSets-TD3的核心指标对比

轨迹控制任务,这意味着该策略在不同训练条件下仍能形成稳定的避障和到靶行为,有助于降低训练偶然性对轨迹控制精度与邻井防碰可靠性的影响。

3.4 最佳轨迹终点与训练收敛

最佳奖励轨迹的终点MD主要位于2820~2850 m,终点井斜接近90°,终点方位接近目标方位

45°。种子分别取1、42、100、500和1024时的终点统计结果见表2。其中种子取500的终点坐标为(1499.23,1500.72,1000.94) m,与目标点三维误差为1.41 m,种子取1024时误差为2.64 m,其余种子时误差均小于6 m,说明策略在完成绕障后能够回归目标水平段。种子100对应轨迹的终点方位角为53.93°,相较于45°目标方位存在一定偏差,但其终点三维误差仍为5.81 m。该结果说明,在当前奖励权重下,策略更优先保证终点位置逼近、避障安全和轨迹平滑性;终点方位角虽受到目标方位引导项约束,但在个别随机种子下仍可能存在一定残余偏差。因此,在后续优化中可将终端方位误差作为更明确的约束项,以进一步提高轨迹末端姿态控制精度。

图8为四障碍物场景下的训练奖励和历史最优靶距。实线表示5个随机种子的汇总趋势,半透明区域表示跨种子波动区间,可同时反映收敛速度和训练稳定性。

图8(a)显示DeepSets-TD3的训练奖励提升更快,后期稳定值高于普通TD3;图8(b)显示其历史最优靶距更早进入低值区间,且后期波动较小。该趋势与固定布置和随机泛化结果一致,表明集合表征不仅改善最终控制性能,也提高了训练过程的稳定性和样本效率。

4 结论

(1)建立了基于DeepSets-TD3的多障碍井约束并眼轨迹智能控制方法。方法以最小曲率法递推轨迹,以井斜角和方位角增量为连续动作,并用DeepSets形成障碍井集合表征。

(2)固定四障碍布置测试中,DeepSets-TD3平均目标误差为7.52 m,物理碰撞率为0;普通TD3对应指标为370.02 m和26.67%。普通TD3的明显退化表明其将障碍井特征按输入顺序拼接后输入MLP,难以泛化到不同障碍排列和风险组合。而DeepSets-TD3通过置换不变的集合表征减弱了输入顺序敏感性,从而提高了多障碍约束下的靶点逼近和避障稳定性。

(3)随机四障碍测试中,DeepSets-TD3的物理碰撞率为1.2%,平均目标误差为11.2 m,近失误差缓冲区率为1.9%,DLS合规率为99.6%,四项指标均优于普通TD3;提高障碍井半径以及加入最小安全

表 2 DeepSets-TD3 最佳奖励轨迹终点统计								
种子	步数	终点 MD/m	井斜/(°)	方位/(°)	N/m	E/m	TVD/m	误差/m
1	268	2820.0	90.00	45.78	1495.48	1496.82	1001.46	5.72
42	269	2830.0	89.35	46.58	1498.35	1503.16	1003.85	5.25
100	271	2850.0	90.00	53.93	1504.68	1497.38	1002.24	5.81
500	268	2820.0	89.75	42.78	1499.23	1500.72	1000.94	1.41
1024	269	2830.0	90.00	46.92	1500.83	1497.83	1001.25	2.64

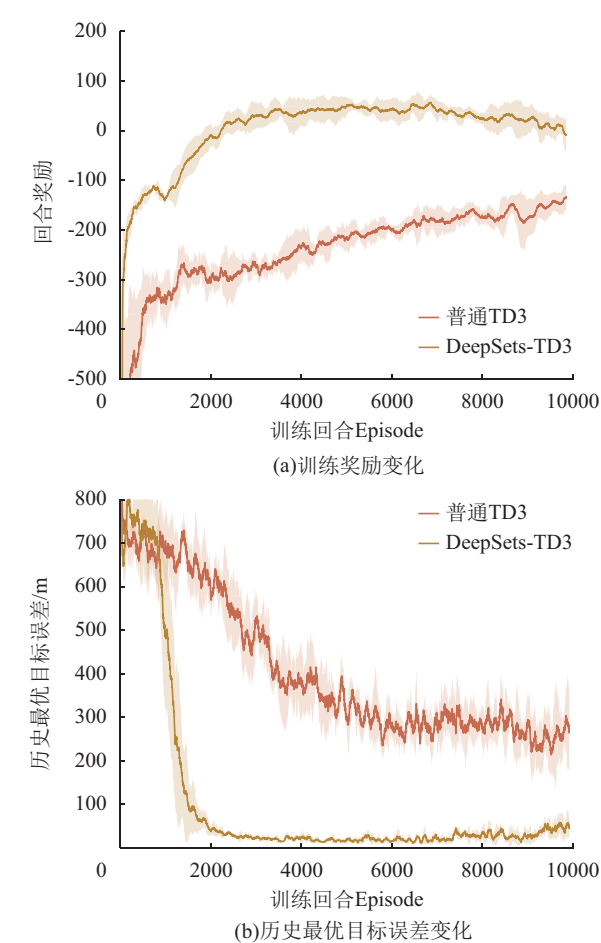


图 8 四障碍物场景下训练奖励与历史最优靶距变化

净距之后,平均目标误差为 13.3 m,物理碰撞率为 2.4%,说明该方法在障碍半径增大与安全边界收紧后仍具有一定鲁棒性。

(4)后续研究应进一步考虑邻井位置不确定性、导向工具响应迟滞和现场闭环约束,并引入控制屏障函数或安全过滤器,以降低困难场景中的安全违规率并提高工程可部署性。

参考文献:

[1] 李根生,宋先知,祝兆鹏,等.智能钻完井技术研究进展与前景

展望[J].石油钻探技术,2023,51(4):35-47.

[2] 顾海鹏.非直井井眼轨迹智能优化与随钻自适应校正策略研究[D].大庆:东北石油大学,2025.

[3] Liu H, Gjersvik T B, Faanes A. Subsea field layout optimization (Part I): Directional well trajectory planning based on 3D Dubins Curve[J]. Journal of Petroleum Science and Engineering, 2022,208:109450.

[4] D'Angelo J, Khaled M S, Ashok P, et al. Pareto optimal directional drilling advisory for improved real-time decision making[J]. Journal of Petroleum Science and Engineering, 2022,210:110031.

[5] Wang G, Huang W, Gao D. Cooperative optimization design method of multitrajectory profile in "well factory" for oil and gas development[J]. SPE Journal, 2025,30(3):1067-1089.

[6] Dou Z, Liu Y, Yang G, et al. Optimization design of infill well trajectory with antifracturing interference for shale gas 3D development[J]. SPE Journal, 2026,31(6):3718-3742.

[7] Zhu D, Xu Q, Wang F, et al. A target-aware well path control method based on transfer reinforcement learning[J]. SPE Journal, 2024,29(4):1730-1741.

[8] Gu H, Wu Y, Li X, et al. Research on wellbore trajectory optimization and drilling control based on the TD3 algorithm[J]. Applied Sciences, 2025,15(13):7258.

[9] Li X, Gu H, Wu Y, et al. Borehole trajectory optimization design based on the SAC algorithm with a self-attention mechanism [J]. Applied Sciences, 2025,15(16):8788.

[10] Zhu W, Liu Y, Wang L, et al. Deep reinforcement learning for well trajectory design with collision avoidance [J]. SPE Journal, 2025,30(8):4545-4560.

[11] 许秋晗.一种地层结构自适应的定向井井眼轨迹控制方法研究[D].北京:中国石油大学(北京),2024.

[12] 樊宸.融合空间注意力机制的智能避障导钻算法研究[D].北京:中国石油大学(北京),2022.

[13] Zaheer M, Kottur S, Ravanbakhsh S, et al. Deep Sets[C]//Advances in Neural Information Processing Systems, 2017: 3391-3401.

[14] Fujimoto S, van Hoof H, Meger D. Addressing function approximation error in actor-critic methods[C]//Proceedings of the 35th International Conference on Machine Learning, 2018: 1587-1596.

[15] Lillicrap T P, Hunt J J, Pritzel A, et al. Continuous control with deep reinforcement learning [C]//International Conference on Learning Representations, 2016.

(编辑 王文)