

基于数值模拟与机器学习融合的CO₂封存研究

刘 阳, 黎 智

安徽理工大学地球与环境学院, 安徽 淮南

收稿日期: 2025年7月6日; 录用日期: 2025年8月8日; 发布日期: 2025年8月14日

摘 要

随着全球碳排放持续增长加剧气候变化, 二氧化碳地质封存(CCUS)技术的重要性日益凸显, 其中深部咸水层和煤层因其巨大潜力成为主要封存场所。融合数值模拟与机器学习技术成为突破瓶颈的关键。本文综述了该融合方法在碳封存领域的研究进展与应用。数值模拟方面, 以COMSOL为代表的多物理场耦合软件能够精准模拟CO₂在多孔介质中的运移如两相流、溶解扩散、岩石变形等。煤层吸附膨胀效应, 其数学模型构建、关键参数设定与边界条件处理是可靠模拟的基础。机器学习方面, 支持向量机、随机森林和神经网络等算法, 利用数值模拟生成的数据集, 可高效构建算法模型, 显著加速参数优化、不确定性分析、封存量预测及稳定性评估。工程实例验证了该融合框架的显著优势, 通过代理模型将历史拟合计算量降低数个数量级。研究表明, 数值模拟与机器学习的协同大幅提升了碳封存预测的效率和精度, 解决了强非均质性和超长期预测难题, 为CCUS工程的高效决策与优化提供了强有力的技术支撑, 有力推动了碳中和目标下油气开发与碳封存技术的融合发展。

关键词

CCUS, 机器学习, 数值模拟, 代理模型

Research on CO₂ Sequestration Based on the Integration of Numerical Simulation and Machine Learning

Yang Liu, Zhi Li

School of Earth and Environment, Anhui University of Science and Technology, Huainan Anhui

Received: Jul. 6th, 2025; accepted: Aug. 8th, 2025; published: Aug. 14th, 2025

Abstract

As global carbon emissions continue to increase and climate change intensifies, the importance of

carbon dioxide geological storage (CCUS) technology is becoming increasingly prominent, with deep saline and coal seams becoming the main storage sites due to their enormous potential. Integrating numerical simulation and machine learning technology has become the key to breaking through bottlenecks. This article reviews the research progress and application of this fusion method in the field of carbon sequestration. In terms of numerical simulation, multi physics coupling software represented by COMSOL can accurately simulate the transport of CO₂ in porous media, such as two-phase flow, dissolution diffusion, rock deformation, etc. The mathematical model construction, key parameter setting, and boundary condition treatment of coal seam adsorption expansion effect are the basis for reliable simulation. In terms of machine learning, algorithms such as support vector machines, random forests, and neural networks can efficiently construct algorithm models using datasets generated from numerical simulations, significantly accelerating parameter optimization, uncertainty analysis, storage prediction, and stability evaluation. Engineering examples have verified the significant advantages of this fusion framework, reducing the computational complexity of historical fitting by several orders of magnitude through proxy models. Research has shown that the synergy between numerical simulation and machine learning has significantly improved the efficiency and accuracy of carbon sequestration prediction, solved the problems of strong heterogeneity and ultra long term prediction, provided strong technical support for efficient decision-making and optimization of CCUS engineering, and effectively promoted the integrated development of oil and gas development and carbon sequestration technology under the goal of carbon neutrality.

Keywords

CCUS, Machine Learning, Numerical Simulation, Surrogate Model

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着 CO₂ 的过度排放, 温室效应明显加剧, 气候变化已经成为全球最严峻的挑战之一, 全球碳排放量持续增长使二氧化碳地质封存技术的重要性日益凸显[1]-[3]。《京都议定书》将 CO₂ 列为温室气体之首, 其封存需求迫在眉睫。深部咸水层因其分布广泛和巨大封存潜力成为首选场所, 煤层则兼具提高甲烷采收率(ECBM)与封存的双重效益。由于地质构造的复杂多变、多相流耦合作用以及长期封存安全性等难题使封存量预测面临极大挑战。传统的数值模拟可以演化物理过程, 但是计算成本较高, 数据驱动模型缺乏物理约束。COMSOL 与机器学习的融合是解决这一问题的关键, 前者通过多物理场耦合精准模拟出 CO₂ 运移的物理化学过程, 后者则通过代理模型加速参数优化与不确定性分析, 形成“物理机理 + 数据驱动”的双引擎预测框架。

刘永忠等研究表明, 超临界 CO₂ 在盐水层中的运移涉及两相流驱替、溶解扩散及岩石形变等复杂过程, 需通过多物理场耦合模型精确描述[4] [5]。而煤层封存更需考虑吸附膨胀引起的渗透率动态变化, 丁洋等通过 COMSOL 建立的流固耦合模型揭示了煤基质膨胀对裂隙压力场的显著影响[6]。在过去几十年中, 一些商业规模的项目(如: Sleipner、In Salah 和神华示范工程等)证明了 CCUS 技术的可行性[7] [8]。然而数值模拟每进行一次计算其所需计算资源量非常庞大, 在上述研究中一众学者只进行数值计算未考虑计算资源与参数优化问题, 找出一种能够减少计算且能够优化参数的方法亟需解决。为此笔者在众多学者研究基础上综述一种数值模拟与机器学习融合的高效方法。

2. 数值模拟

众多学者广泛使用的数值模拟计算软件有 ABAQUS、ANSYS、FLAC3D、COMSOL 等。数值模拟一般流程为前处理建立计算几何模型、定义材料属性、确立控制方程与边界条件、多物理场耦合、网格划分与依赖性验证；利用软件内置有限元求解器求解；后处理导出应力应变云图、等值线图、动态变化过程、时间监测曲线等进行分析，流程见图 1。在二氧化碳封存模拟中 COMSOL Multiphysics 数值计算软件凭借其灵活的多物理场耦合功能、模块化建模环境、内置多孔介质本构模型和多相流功能已经成为碳封存模拟的主流工具。

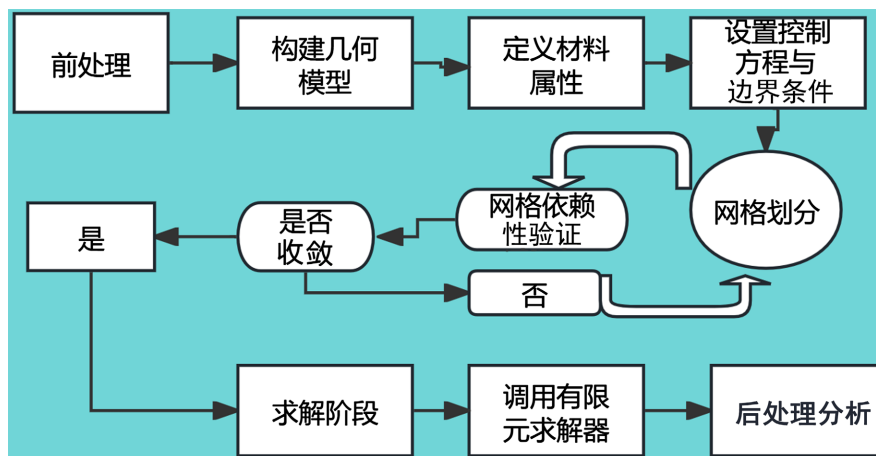


Figure 1. Numerical simulation flowchart

图 1. 数值模拟流程图

该软件能够将流体流动、化学反应与岩土力学等过程进行很好的耦合，通过有限元法求解耦合偏微分方程组(PDEs)。

2.1. 数学模型构建

在 CO₂ 封存应用中，遵循多孔介质不混溶两相流的质量守恒规律。为了描述此种现象，研究者一般使用基于达西定律的多孔介质多相流方程进行表征[4] [5]。COMSOL 通常耦合达西定律用来描述多孔介质渗流、Fick 扩散定律刻画组分迁移及水平集方法进行气液界面追踪。

以刘永忠团队建立的超临界 CO₂-盐水两相流模型为例[4]，其控制方程包括：

$$\frac{\partial(\phi \rho_{\alpha} S_{\alpha})}{\partial t} + \nabla \cdot (\rho_{\alpha} \mathbf{u}_{\alpha}) = Q_{\alpha}, \quad \alpha = \text{CO}_2, \text{brine} \quad (1)$$

其中 ϕ 为孔隙率， ρ 为密度， S 为饱和度， u 为达西流速， Q 为源汇项[4]。该模型结合 Peng-Robinson 状态方程描述超临界 CO₂ 的热力学性质，并通过对数正态随机孔隙率表征地层非均质性，成功重现了 CO₂ 羽流在径向和垂向的迁移规律。煤层封存则需额外引入煤岩吸附变形方程。煤基质膨胀导致的渗透率演化常采用 Shi-Durucan 模型：

$$k = k_0 e^{-\gamma(\sigma - \sigma_0) + \beta \varepsilon_s} \quad (2)$$

其中 γ 为应力敏感系数， ε_s 为吸附应变， β 为耦合系数[4]-[6]。黄陵矿区废弃矿井的模拟显示，CO₂ 在采空区内的分布呈现“近底板富集特性”，且吸附态封存量占比高达 63.52% (注入井深 95 m 时)，验证了煤岩吸附对封存量的主导作用[7]。主要数学物理方程见表 1。

Table 1. Core physical modules for carbon sequestration simulation in COMSOL
表 1. COMSOL 中碳封存模拟的核心物理模块

物理过程	对应模块	关键方程
多孔介质流动	Darcy's Law	$\nabla \cdot (\rho u) = \partial(\phi \rho) / \partial t$
组分运移与扩散	Transport of Diluted Species	$\partial c / \partial t + \nabla \cdot (-D \nabla c) = R$
气-液两相流	Level Set/Phase Field	$\partial \phi / \partial t + u \cdot \nabla \phi = \gamma \nabla \cdot (\nabla \phi)$
岩土力学	Solid Mechanics	$\nabla \cdot \sigma + F = 0$

2.2. 关键参数与边界条件

数值计算模型的可靠性依赖于地质参数的准确性。以咸水层为例, 需输入渗透率($10^{-15} \sim 10^{-12} \text{ m}^2$)、孔隙率(0.05~0.3)、毛管压力曲线如 Brooks-Corey 模型及相对渗透率等参数。煤层还需煤的吸附等温线, 常采用 Langmuir 方程和力学参数如弹性模量、泊松比[4] [5]。

在设置边界条件时需考虑多方面因素, 反映实际工况。通常把注入井设置压力边界或质量、流量边界。在时间尺度上, 一般设置流动仿真时间为 50 年。地层外边界可设为零通量即封闭地层或静水压力开放系统。在前三十年为注入状态, 后二十年为封闭地层状态[9]。

在整个过程中需要重点考虑泄漏风险评估, 引入虚拟井模拟裂隙通道。通过大量研究表明泄漏速率通常低于注入量的 0.3% [10]。

3. 机器学习

机器学习是人工智能领域中常用技术之一, 是一门研究如何通过计算手段, 利用经验(通常以数据形式存在)来自动改善系统自身性能的学科; 其核心在于让计算机系统不依赖于显式的编程指令, 而是通过从积累的经验(数据)中自动分析、归纳并获取蕴含的规律(模式或知识), 进而基于这些学习到的规律对新的未知情况做出有效的预测或决策, 从而在完成特定任务(如分类、回归、聚类、降维等)的能力上实现提升; 简言之, 机器学习致力于构建能够从数据中“学习”并随着经验积累而不断进化和优化的算法与模型, 它既是实现人工智能的关键途径, 也是从数据中挖掘价值、发现知识以解决复杂现实问题的核心技术[11]。

3.1. 支持向量机

支持向量机是一种常用的监督学习算法, 主要用于分类和回归任务, 适用于数据分布复杂、边界难以划分的情境[12]-[14]。在碳封存研究中, 支持向量机常用于预测封存稳定性和样本分类[13]。例如, 在煤层中预测稳定性时, 涉及煤孔隙率、注入量、温度等多个变量, 支持向量机可在多维特征空间中有效划分稳定与不稳定区域[12]-[14]。

3.2. 随机森林

随机森林算法[14]是一种集成学习方法, 通过构建多个决策树并融合其决策结果进行最终预测, 在保留决策树直观优势的同时, 有效缓解了过拟合问题, 适用于数据量大、特征多、对模型泛化能力要求高的场景。在碳封存研究中, 随机森林算法可用于储层特性分析和 CO_2 运移分布预测, 结合多种特征如地质构造、孔隙度和注入速率等, 综合分析海量数据, 提高预测的准确度。

3.3. 神经网络

神经网络是一种模拟人脑神经元网络结构与工作方式的人工智能算法。根据结构和学习方式的不同,神经网络算法[15]可进一步细分为多种类型,其中前馈神经网络(FNN)适用于一般性任务,卷积神经网络(CNN)在图像处理领域表现卓越,循环神经网络(RNN)因其独特的循环结构,适合处理具有时间依赖性的序列数据,能通过共享权重捕捉长期的数据特征关联[16]。在研究碳封存时 CO_2 的运移规律尤为重要,需要其随着时间尺度的泄露状态。RNN 算法与长短期记忆网络(LSTM)算法能够充分发挥处理时间序列数据的优势,挖掘数据中的潜在信息,因而被广泛应用。

4. 模型融合与工程实例

4.1. 数据集构建与特征工程

构建准确机器学习模型的核心在于数据集。通过数值模拟技术可以减少数据集生成成本。运用 COMSOL 生成训练数据集以减少现场原位实验与相似模拟实验。以深层咸水层封存效率预测为例,考虑地质参数如孔隙度、渗透率、地层倾角;注入参数如速率、压力、温度及流体特性如盐度、 CO_2 纯度等多维度变量[17]。通过实验设计如拉丁超立方采样在参数空间均匀取样、K 折交叉验证等[14]。通过编写自动化脚本与 COMSOL 求解器耦合自动生成在改变不同参数变量下的数值,对应输出数据集包括束缚空间封存、溶解封存及累计注入量等参数指标。运用相互耦合的代理模型可以极大减少数据集生成的高昂时间与计算成本。

特征工程环节需筛选物理意义明确的输入参数。例如,咸水层封存中无量纲组合(如毛细管数 Ca 、重力数 Ng)可替代原始变量;煤层封存则需突出吸附势与应力敏感系数的影响[5]-[7]。

在构建出准确的机器学习代理模型后,后续工程中可通过在机器学习模型输入对应参数预测出对应结果;其次可根据机器学习的初步结果减少 COMSOL 数值模拟参数调整计算成本,更加高效通过求解器得出计算结果;最后可根据现场实际情况进行相似模拟实验,通过等比例缩放尺寸、时间与参数大小得出实验结果。将机器学习、数值模拟与相似模拟实验三种手段得出结果相互印证,通过多尺度数据融合更加高效精准表征 CO_2 地质封存的运移-解吸规律。同时三种方法相互对照可以有效解决来自数值模型与机器学习代理模型自身的双重不确定性,增强整个系统的鲁棒性,对应流程见图 2。

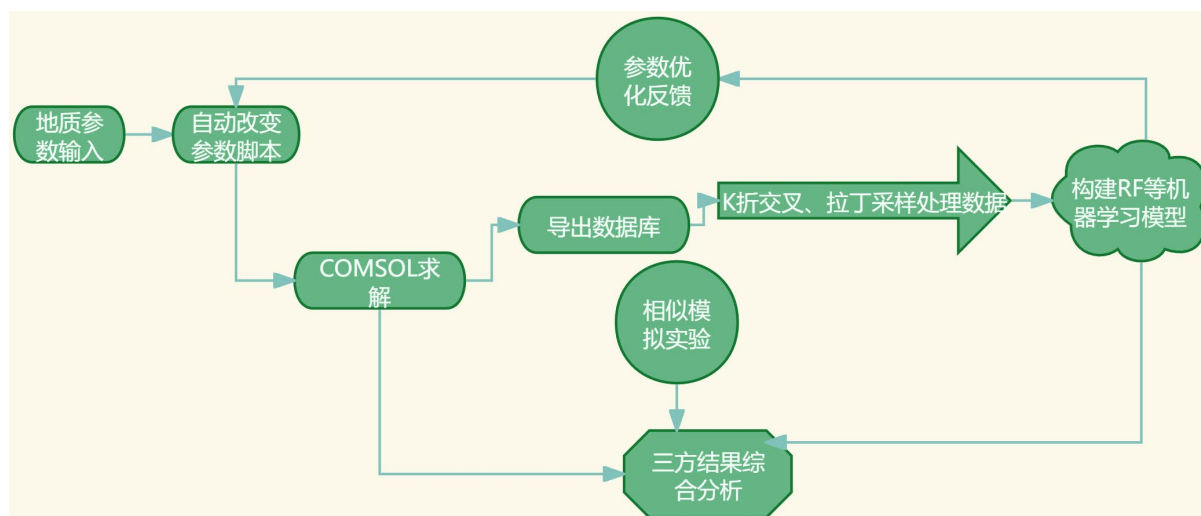


Figure 2. "Physics-Data-Experiment" three-driven framework diagram

图 2. “物理 - 数据 - 实验”三驱动框架图

4.2. 工程实例

在碳捕集、利用与封存(CCUS)工程实践中,游君昱等以美国 Farnsworth Unit (FWU)油田项目为背景,展示了数值模拟与机器学习融合方法的显著成效[18]。该油田自 1955 年开发以来积累近 70 年生产数据,储层非均质性极强,CO₂-水交替驱(CO₂-WAG)进一步加剧了流体分布复杂性,导致传统数值模拟面临历史拟合计算成本高、多目标协同优化困难等挑战。游君昱等研究团队提出混合方法框架:首先构建精细数值模型生成训练数据库,采用人工神经网络(ANN)与高斯核支持向量回归(Gaussian-SVR)建立替代模型,显著降低计算负荷。通过粒子群优化(PSO)算法驱动替代模型,以 62 项关键参数(如各向渗透率)优化历史拟合,将总误差降至 20.2%,计算效率提升显著。60,000 次代理模型运算仅需 300 秒,而同等规模数值模拟需数月[18]。

针对 CCUS 项目多目标优化需求,研究进一步整合多目标粒子群算法(MOPSO)与 SVR 替代模型,以 20 年预测期内累计产油量(f_1)、CO₂埋存量(f_2)和净现值(f_3)为优化目标。基于帕累托最优理论生成三维解集前沿,揭示目标间内在权衡关系:例如产油量与 CO₂埋存量、NPV 呈负相关性。优化后的帕累托解集显示,相较基础方案(产油 1320 万桶、CO₂埋存 106 万吨、NPV 1.83 亿美元),最优方案可实现原油增产 158 万桶(总量达 1478 万桶)、CO₂埋存增加 137 万吨(总量 243 万吨),NPV 最高提升至 2.05 亿美元,CO₂埋存率峰值达 92.9%。该方法为决策者提供灵活方案选择,例如优先满足 90%埋存率约束时仍可筛选 NPV 最大化策略[18]。

该工程实例验证了混合方法在复杂 CCUS 工程中的双重优势:其一,代理模型将历史拟合计算量压缩数个数量级,解决超长期生产数据与强非均质性储层的拟合难题;其二,帕累托前沿量化了“增产-埋存-经济性”的协同与矛盾,突破传统单目标优化的局限性。

5. 总结与展望

5.1. 总结

面对碳封存预测中的地质复杂性和高昂计算成本难题,数值模拟与机器学习的结合提供了有效途径。数值模拟动态刻画了 CO₂运移、溶解、吸附及岩石响应的物理过程,但耗时巨大。机器学习模型则擅长从海量数据中快速学习规律。两者的融合,利用数值模拟生成高质量数据训练机器学习模型,构建出兼顾物理机理与计算效率的快速预测工具。

这种结合在实践中展现出强大优势。它显著缩短了如历史拟合、长期预测等耗时任务的计算时间,从数月压缩到分钟,解决了强非均质储层和长期动态模拟的瓶颈。同时,结合物理基础的机器学习模型,也提升了封存量预测和风险评估的准确性。更重要的是,它能够高效处理多目标优化问题,例如在增产原油、增加 CO₂埋存量和提高经济效益之间找到最佳平衡点,为实际工程决策提供清晰依据,美国 Farnsworth 油田等项目已成功验证了这一点。

总之,数值模拟与机器学习的协同应用,极大地提升了我们理解和优化碳封存过程的能力。它降低了评估成本,缩短了研究周期,为在全球范围内更高效、更可靠地部署碳捕集与封存技术,实现碳中和目标,提供了坚实的技术支撑。这一融合方向将继续推动碳封存技术的进步。

5.2. 挑战与展望

碳中和是我国目前众多产业升级的必经之路,想要达成“碳中和-碳达峰”这一目标二氧化碳地质封存技术是关键一环。然而这一实现过程当前面临着众多挑战如 CO₂在深部地层中会与油气、煤矿及其他封存地层之间存在着无法探查的相互作用,地下封存岩层地质结构、理化性质、渗透性、含水率及温

度等各种因素对 CO₂ 封存时的作用及运移机制也存在着极其复杂影响, 超临界态的 CO₂ 与地层水结合后所形成的 HCO₃⁻ 时刻侵蚀破坏着封存地层的岩腔(如含碳酸盐岩胶结物盖层), 造成潜在的逃逸风险[19]。

机遇与挑战往往是并存的, 尽管复杂的地质环境会对二氧化碳封存造成诸多影响, 但是其经济效益必会日益兴盛, 国家也在出台诸多利好政策。随着数值模拟、机器学习与深度神经网络不断发展, 科学第三范式与第四范式高效结合以及地球物理勘探新技术如三维地震反演等方法的出现[19]。地下封存体的结构的勘察将会越来越透明化、精准化、智能化, 其整个体系的稳定性和安全性也会不断提升。

参考文献

- [1] 李阳, 黄文欢, 金勇, 等. 双碳愿景下中国石化不同油藏类型 CO₂ 驱提高采收率技术发展与应用[J]. 油气藏评价与开发, 2021, 11(6): 793-804.
- [2] 张海龙. CO₂ 混相驱提高石油采收率实践与认识[J]. 大庆石油地质与开发, 2020, 39(2): 114-119.
- [3] 尹书郭, 杨国栋, 冯涛, 等. 页岩储层物性参数对 CO₂ 不同封存机制及封存量的影响[J]. 高校地质学报, 2023, 29(1): 37-46.
- [4] 杨勇, 刘永忠, 于博, 等. 封存 CO₂ 裹挟可挥发性污染物的迁移模型分析[J]. 化工学报, 2015, 66(10): 4155-4162.
- [5] 刘永忠, 王乐, 张甲六. 封存 CO₂ 的泄漏过程预测与泄漏速率的影响因素特性[J]. 化工学报, 2012, 63(4): 1226-1233.
- [6] 丁洋, 周煜尧, 李树刚, 等. CO₂ 封存下煤参数对煤层压力与渗透性影响的数值模拟[J]. 西安科技大学学报, 2025, 45(3): 500-511.
- [7] Zhang, K., Lau, H.C. and Chen, Z.X. (2022) Extension of CO₂ Storage Life in the Sleipner CCS Project by Reservoir Pressure Management. *Journal of Natural Gas Science and Engineering*, **108**, Article 104814. <https://doi.org/10.1016/j.jngse.2022.104814>
- [8] 王永胜. CO₂ 咸水层封存砂岩储层孔隙尺度变化规律及可注性研究: 基于神华 CCS 示范工程[D]: [博士学位论文]. 武汉: 中国地质大学(武汉), 2021.
- [9] 武晓菲, 徐同晖, 何畅, 等. 基于有限元数值模拟的地质封存 CO₂ 运移行为预测方法[J]. 计算机测量与控制, 2024, 32(11): 132-139.
- [10] 丁洋, 汤远卓, 李树刚, 等. 老采空区 CO₂ 封存井位部署方案数值模拟研究[J]. 煤炭科学技术, 2024, 52(S2): 131-141.
- [11] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.
- [12] 杨斌, 路游. 基于统计学习理论的支持向量机的分类方法[J]. 计算机技术与发展, 2006, 16(11): 56-58.
- [13] 李召桐. 支持向量机发展历程及其应用[J]. 信息系统工程, 2024(3): 124-126.
- [14] 吕红燕, 冯倩. 随机森林算法研究综述[J]. 河北省科学院学报, 2019, 36(3): 37-41.
- [15] 孙志军, 薛磊, 许阳明, 等. 深度学习研究综述[J]. 计算机应用研究, 2012(8): 2806-2810.
- [16] 杨丽, 吴雨茜, 王俊丽, 等. 循环神经网络研究综述[J]. 计算机应用, 2018, 38(2): 1-6.
- [17] 姚津津, 郭东东, 朱恒银, 等. 机器学习在天然气水合物领域中的应用研究进展[J]. 油气与新能源, 2025, 37(3): 93-100.
- [18] 游君昱, 文书琴, 何宇娇, 等. 利用基于数值模拟和机器学习的混合方法解决 CCUS 项目的实际问题[C]//长安大学, 陕西省石油学会, 西安市科学技术协会, 延安大学. 第八届数字油田国际学术会议论文集. 2023: 165-171.
- [19] 廖志伟, 羊俊敏, 钟翔宇, 等. 二氧化碳地质封存技术研究进展综述[J]. 地下空间与工程学报, 2024, 20(S1): 497-507.