

四川盆地A地区页岩孔隙度预测及主控因素分析

邢亚冰, 王仲旭, 吴欣雨

重庆科技大学石油与天然气工程学院, 重庆

收稿日期: 2025年3月5日; 录用日期: 2025年4月2日; 发布日期: 2025年4月14日

摘要

孔隙度是深层页岩储层含气性评价及有利区优选的关键参数。为评估LightGBM算法在深层页岩储层孔隙度预测评价中的适用性, 选取四川盆地A地区五峰组-龙马溪组深层页岩储层实测孔隙度为标签以及GR等7个测井参数为特征, 建立基于LightGBM算法的孔隙度预测评价模型。研究表明: LightGBM算法模型的孔隙度评价精度及泛化能力($R^2 = 0.979$, $RMSE = 0.622$, $MAPE = 31.5\%$)优于多元回归算法模型, 能够精确地评价纵向上孔隙度变化规律; 建立的LightGBM算法模型对深层页岩储层中孔隙度评价具有良好的适用性, 为深层页岩含气性预测提供了新的思路, 并明确研究区孔隙度主要受到TOC及黏土含量的共同控制。

关键词

LightGBM算法, 孔隙度, 深层页岩储层, 四川盆地A地区

Shale Porosity Prediction and Analysis of Main Controlling Factors in Area A of Sichuan Basin

Yabing Xing, Zhongxu Wang, Xinyu Wu

School of Petroleum Engineering, Chongqing University of Science & Technology, Chongqing

Received: Mar. 5th, 2025; accepted: Apr. 2nd, 2025; published: Apr. 14th, 2025

Abstract

Porosity is a key parameter for evaluating gas bearing properties of deep shale reservoirs and selecting favorable areas. In order to evaluate the applicability of LightGBM algorithm in the prediction and evaluation of deep shale reservoir porosity, the measured porosity of deep shale reservoir

文章引用: 邢亚冰, 王仲旭, 吴欣雨. 四川盆地 A 地区页岩孔隙度预测及主控因素分析[J]. 地球科学前沿, 2025, 15(4): 339-346. DOI: 10.12677/ag.2025.154034

in Wufeng Formation and Longmaxi Formation in area A of Sichuan Basin was selected as the label and 7 logging parameters such as GR were featured, and a porosity prediction and evaluation model based on LightGBM algorithm was established. The results show that the porosity evaluation accuracy and generalization ability of LightGBM algorithm model ($R^2 = 0.979$, RMSE = 0.622, MAPE = 31.5%) are superior to the multiple regression algorithm model, and can accurately evaluate the longitudinal porosity variation law. The established LightGBM algorithm model has good applicability to the evaluation of porosity in deep shale reservoirs, providing a new idea for the prediction of gas content in deep shale reservoirs, and it is clear that the porosity of the study area is mainly controlled by TOC and clay content.

Keywords

LightGBM Algorithm, Porosity, Deep Shale Reservoir, A Region of Sichuan Basin

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

页岩气作为一种清洁天然气能源，在能源转型过渡期发挥了重要作用[1]。孔隙度作为目前直观反映页岩储集能力以及影响页岩气产能的重要参数，是页岩气储层含气性评价的关键指标[2]。目前，国内外学者主要利用实验测试分析以及多元回归法来测得页岩孔隙度。实验分析法受制于金钱和时间成本，并且由于通过实验测试得到的数据是离散数据难以直观反映其垂向分布规律。孔隙度受多种地质因素的影响而具有较强的储层非均质性，导致利用多元回归法来计算孔隙度时计算精度较差。现今地质领域与人工智能算法的结合越来越密切，前人将深度学习、随机森林、卷积神经网络、梯度决策提升树等算法融入到页岩储层参数预测领域并取得了不错的成果[3]。因此本文引入 LightGBM 算法来计算孔隙度，LightGBM 算法是微软公司于 2017 年提出，通过组合多组弱学习器，减少单一弱学习器预测的误差，具有训练效果好、速度快、不易过拟合的特点，在多学科有着较好的应用效果[4]。因此，本次研究以四川盆地 A 地区五峰组 - 龙马溪组深层页岩储层为例，基于测井数据利用 LightGBM 算法建立孔隙度预测评价模型，随后基于岩心分析实验数据对 LightGBM 算法模型及多元回归法预测结果进行对比验证。通过页岩储层孔隙的相关性分析明确研究区孔隙度变化的主要控制因素。

2. 区域地质背景

四川盆地 A 地区构造上位于四川盆地川中平缓构造带与川南低陡构造带交界处，整体呈北东 - 南西向展布[5]。目前研究区深层页岩气勘探开发的主力层系为五峰组和龙马溪组底部地层，目的层发育一套深水 - 浅水陆棚沉积背景的富有机质海相页岩。五峰组地层岩性主要为黑色含硅质页岩，富含笔石化石[6]；龙马溪组主要为泥页岩，钙质、碳质含量高，自下而上粉砂含量逐渐增多[7]。五峰组厚度为 5~11 m，龙马溪组地层厚度介于 200~530 m，地层整体横向展布较稳定[8]。五峰组上部以及龙马溪组底部地层有机质含量高，是页岩气开采的主力层段。

3. 数据及方法

3.1. 数据准备及预处理

本文测井数据来自于四川盆地 A 地区 12 口页岩气井，本次样品均取于四川盆地 A 地区五峰组 - 龙

马溪组深层页岩储层，在进行了充分的岩心观察后设计合理的取样方案以保证实验结果的代表性。其中 A12 井的数据不参与模型训练，用以检验所建立的孔隙度评价模型的泛用性。样本数据标签(因变量)为孔隙度通过实验测试获得，之后选取与孔隙度数据对应深度的 GR (自然伽马)、KTh (无铀伽马)、U (铀)、DEN (密度)、CNL (补偿中子)、AC (声波时差)及 LLD (深侧向电阻率) 7 个测井参数作为样本数据特征(自变量)。针对孔隙度及其所对应的测井参数展开岩石物理分析，对 GR、KTh、U、AC、DEN、CNL 及 Rt 测井曲线参数及孔隙度进行相关性拟合，在这 7 个测井参数中筛选出与孔隙度相关性较好的 AC 和 DEN 测井曲线参数(图 1)基于此利用多元回归法进行拟合得到多元回归孔隙度测井评价模型。

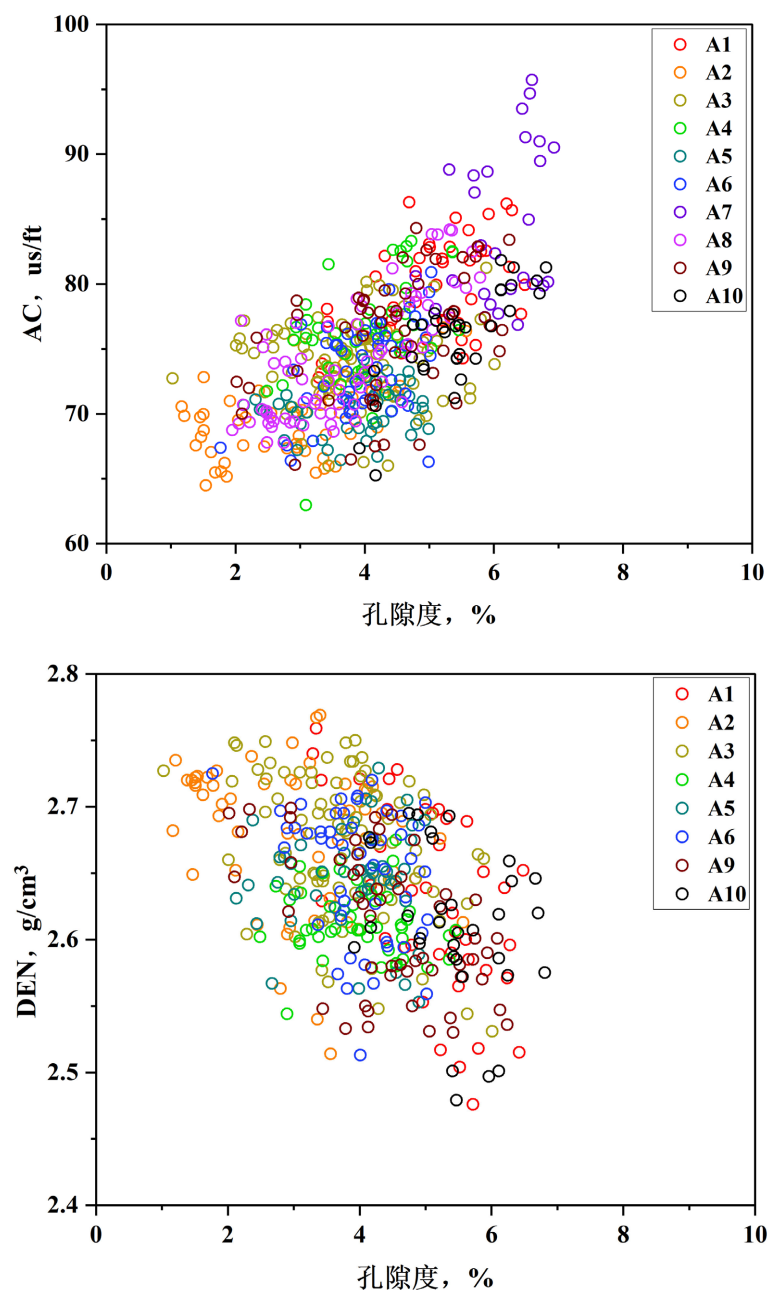


Figure 1. Correlation analysis of pore of target layer and log parameters in study area
图 1. 研究区目的层孔隙及测井曲线参数相关性分析

3.2. LightGBM 算法

LightGBM 是由微软公司提出的一种基于梯度提升决策树框架的集成学习算法[4] [9]，其核心思想是采用前向分布算法生成新的弱学习器(即决策树)，每个弱学习器的构建都是基于上一个弱学习器和标签值之间的差值，模型的最终输出为所有弱学习器输出结果之和(图 2) [10]。为进一步减少计算内存占用、提升运行效率及防止过拟合，LightGBM 算法在确保模型性能的同时，提出了多项改进措施：包括基于直方图的决策树算法、带深度限制的 leaf-wise 的叶子生长策略、基于梯度的单边采样及互斥特征捆绑策略。

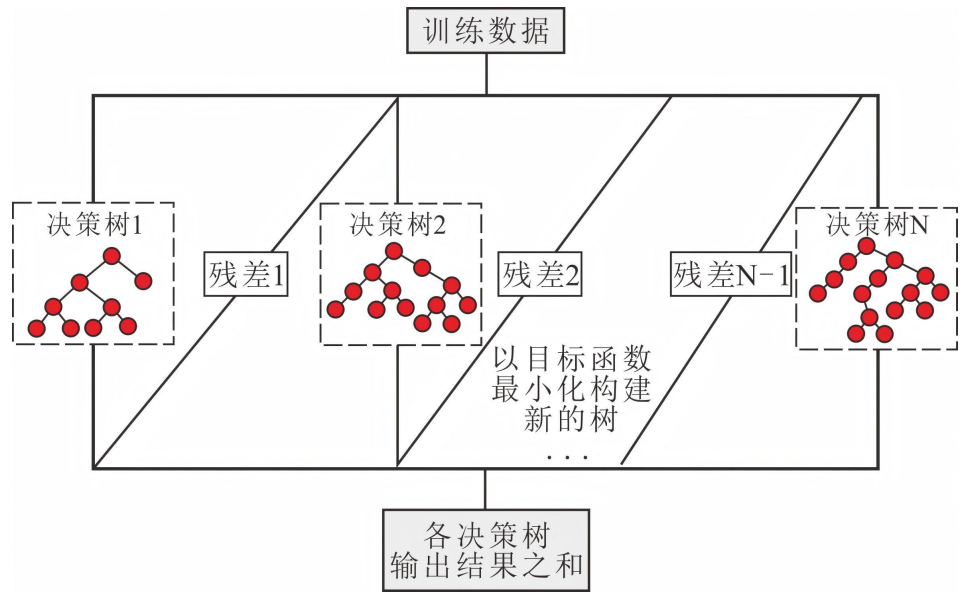


Figure 2. Schematic diagram of gradient lifting decision tree framework (Modified based on reference [11])
图 2. 梯度提升决策树框架原理示意图(基于文献[11]修改)

利用 LightGBM 算法对研究区孔隙度开展预测工作时，以研究区 387 个样本数据，样本标签为孔隙度，样本特征为对应深度的 GR (自然伽马)、KTh (无铀伽马)、U (铀)、DEN (密度)、CNL (补偿中子)、AC (声波时差)及 LLD (深侧向电阻率) 7 个测井参数。通过利用 GridSearchCV (带交叉验证的网格搜索)这种超参数优化方法穷举出最优的超参数组合，并使用 5 折交叉验证来评估每个超参数组合的性能，确保发现全局最优解。本研究基于 LightGBM 算法的孔隙度评价模型超参数取值范围、步长及最优取值为 n_estimators: 456、learning_rate: 0.20、num_leaves: 8。最终建立研究区孔隙度预测模型。

4. 实例应用

4.1. 孔隙度评价结果

基于岩石物理分析结果挑选出的与孔隙度相关性较好的 AC 和 DEN 测试集数据利用多元回归法拟合出孔隙度预测评价模型： $\Phi = 0.125 * AC - 7.829 * DEN + 15.494$ ，之后利用测试集数据验证所建立的孔隙度预测评价模型。利用 LightGBM 算法应用训练集数据建立孔隙度预测评价模型，然后利用测试集数据对所建立模型的性能进行验证。为客观评价 LightGBM 算法模型的性能，将其与多元回归法孔隙度评价模型进行对比(图 3 和表 1)。两个模型使用相同的数据划分方案，即训练集和测试集数据一致。

Table 1. Comparison of porosity prediction and evaluation effects of different models
表 1. 不同模型孔隙度预测评价效果对比

孔隙度评价方法	特征	训练集			测试集		
		R ²	RMSE	MAPE	R ²	RMSE	MAPE
LightGBM 算法	GR, KTH, U, DEN, CNL, AC, LLD	0.979	0.622	31.5%	0.975	0.684	35.1%
多元回归法	AC, DEN	0.893	0.865	72.5%	0.885	0.789	65.2%

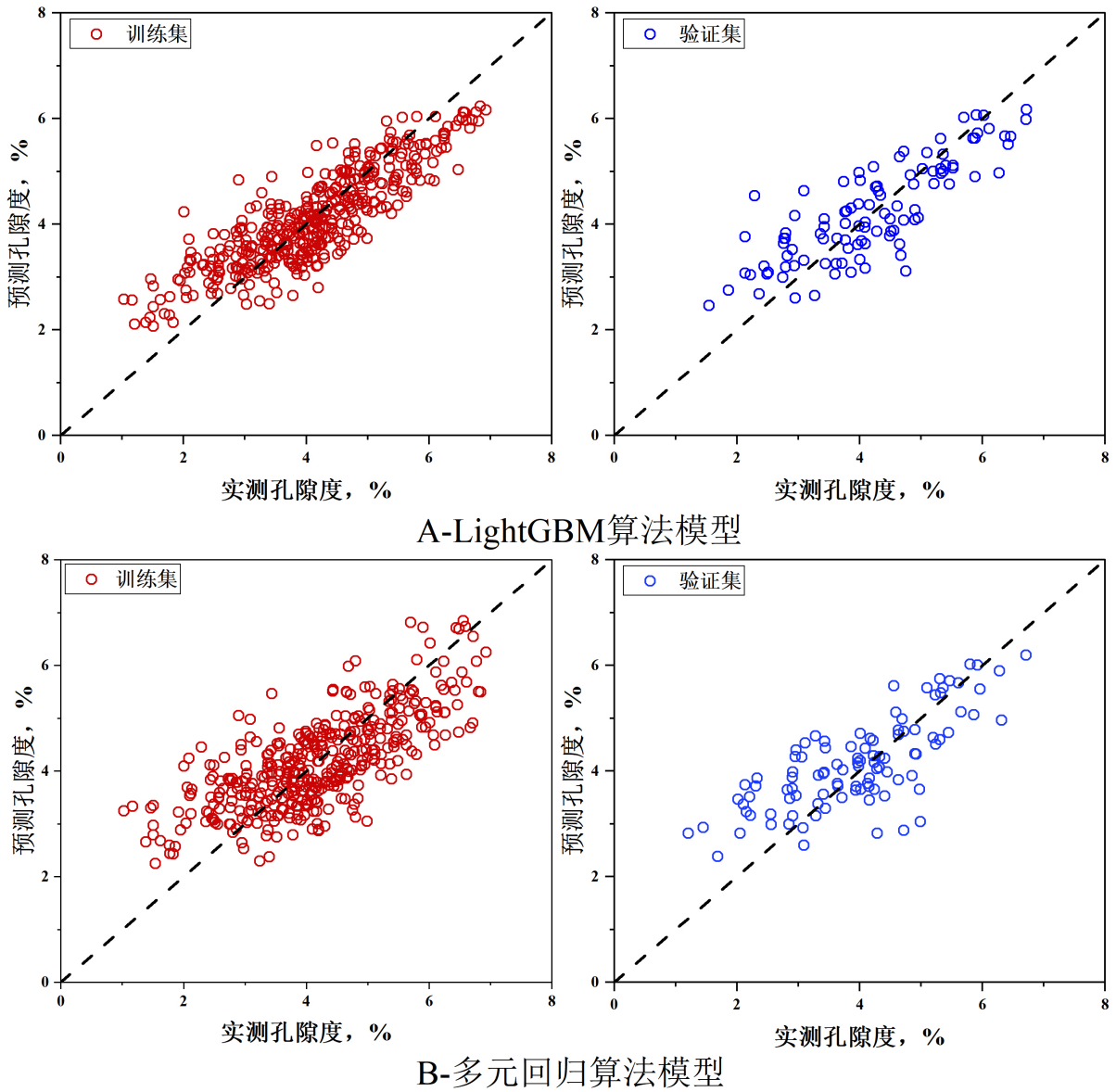


Figure 3. Scatter plot of predicted and true values of porosity of different models
图 3. 不同模型孔隙度预测值与真实值散点图

由表 1 可以看出，三个模型中，LightGBM 算法模型在训练集和测试集数据上的表现都是最优的，R² 分别达到了 0.979 和 0.975，RMSE 分别为 0.62 和 0.68，MAPE 分别为 31.5%和 35.1%。从图 3A 也可以看出，LightGBM 算法模型预测的孔隙度数据点均匀分布于 y = x 线两侧，并且接近于 y = x 线，表明该

模型对孔隙度的评价精度较高,且模型泛化能力强。多元回归法模型在孔隙度评价中的表现中等, R^2 分别为 0.893 和 0.885, RMSE 分别为 0.87 和 0.79, MAPE 分别为 72.5% 和 65.2%。虽然多元回归法模型预测的孔隙度数据点大致分布于 $y=x$ 线两侧,但有一定程度的偏离(图 3B),表明该模型对孔隙度评价精度及模型泛化能力均弱于 LightGBM 算法模型。

为进一步检验 LightGBM 算法模型的泛化能力,将该模型应用于研究区未参与模型训练的 A12 井。计算结果显示(图 4), LightGBM 算法模型对该井目的层段页岩气储层孔隙度预测评价精度较高,效果明显优于多元回归算法模型,进一步证明了 LightGBM 算法模型具有较强的泛化能力,在深层页岩气储层孔隙度评价中具有较好的适用性。综上, LightGBM 算法在深层页岩气储层孔隙度预测评价具有较高的精度和较强的泛化能力。

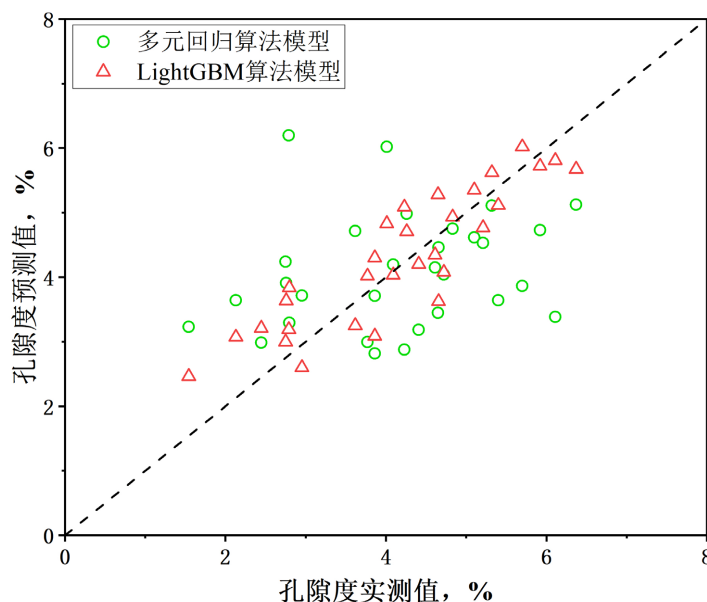


Figure 4. Comparison of the results of different porosity evaluation models for well A12 in Area A, Sichuan Basin
图 4. 四川盆地 A 地区 A12 井不同孔隙度评价模型结果对比

4.2. 孔隙度主控因素分析

渝西五峰组-龙一 1 亚段页岩储层孔隙度的控制因素是多方面的,包括沉积环境、岩石矿物成分、有机质含量、成岩作用、构造作用以及热演化程度等因素的控制。这些因素相互作用,共同决定了页岩储层的孔隙发育特征和储气能力。本次研究根据孔隙度与其控制因素的相关性分析,分析结果显示研究区五峰组-龙一 1 亚段页岩储层孔隙度主要受 TOC 含量以及黏土含量控制。

研究区页岩有机质成熟度高,有机质普遍经历热裂解生气阶段,在形成过程中发育大量有机质孔隙,大量发育的有机质孔是研究区五峰组-龙一 1 亚段页岩发育高孔隙度的关键。在 $\text{TOC} < 3\%$ 时,研究区目的层有机质孔发育速度变大,在 $\text{TOC} > 5.0\%$ 的高 TOC 含量情况下,页岩中有机质孔发育减缓。一方面,黏土矿物在页岩中多以黏土复合体的形式出现,呈片状、板状及絮状,内部发育纳米级孔隙,在黏土矿物含量较低时,孔隙度快速增大[12]。另一方面由于黏土属于塑性矿物,受挤压作用而发生变形,黏土质量分数过高导致研究区目的层页岩抗压能力明显减弱,使页岩更容易被压实,孔隙遭受压实破坏,与黏土矿物相关的储层原生孔隙减小,在高黏土含量下页岩中有机质含量低,有机质孔隙发育较少。TOC 含量处于 $1.5\% \sim 5.0\%$ 以及黏土矿物含量在 $20\% \sim 40\%$ 之间时,研究区五峰组-龙一 1 亚段页岩储层发育高

孔隙区(图 5)。

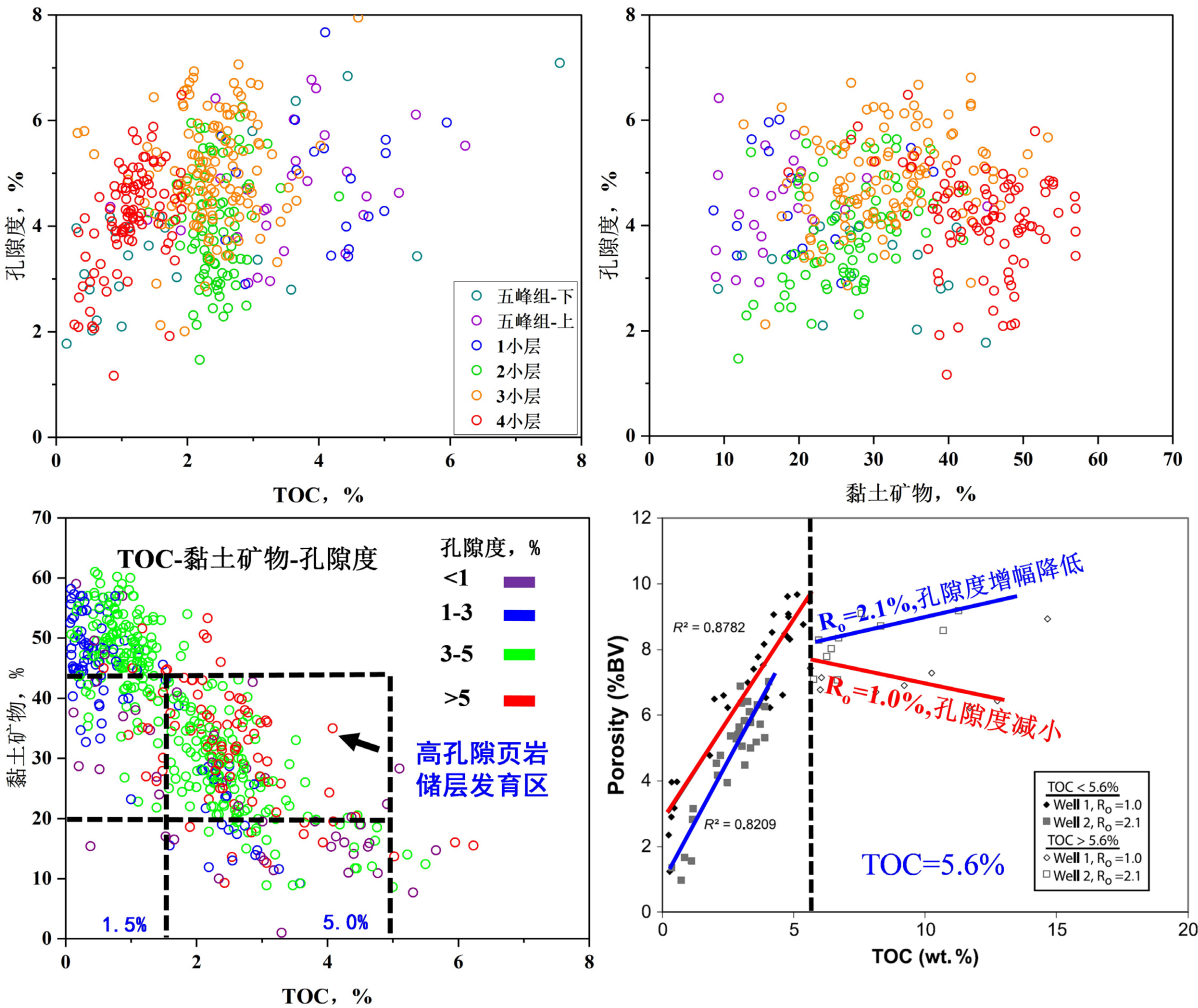


Figure 5. Analysis of the main controlling factors of porosity in the target layer of the study area
图 5. 研究区目的层孔隙度主控因素分析

5. 结论

- (1) 由于 LightGBM 算法更强大的非线性关系处理能力及更多的特征输入，相较于多元回归法，LightGBM 算法模型在深层页岩气储层孔隙度评价中具有更高的精度及更强的泛化能力。
- (2) 明确了研究区目的层孔隙度的主要控制因素为 TOC 和黏土矿物含量，TOC 含量处于 1.5%~5.0% 以及黏土矿物含量在 20%~40% 之间时，研究区五峰组 - 龙一 1 亚段页岩储层发育高孔隙区。

致 谢

感谢各位作者对于本文章提供的建议及帮助，同时也感谢期刊编辑为文章发表给予的帮助。

基金项目

重庆科技大学科技创新基金项目“渝西地区有机质含量测井预测方法优选及主控因素研究”(立项编号: YKJCX2320145)资助

参考文献

- [1] 薛晓辉, 李金龙, 刘红俊, 等. “双碳”背景下我国页岩气产业发展潜力分析[J]. 能源与环保, 2023, 45(9): 169-174.
- [2] 胡曦, 李亚丁, 但霞, 等. 四川盆地泸州地区五峰组——龙一 1 亚段页岩气地质特征与富集高产主控因素[J]. 天然气技术与经济, 2024, 18(4): 15-23.
- [3] 张少龙, 闫建平, 郭伟, 等. 基于岩石物理相的深层页岩气地质——工程甜点参数测井评价方法——以四川盆地 LZ 区块五峰组——龙马溪组为例[J]. 石油地球物理勘探, 2023, 58(1): 214-227.
- [4] Meng, Q. (2017) LightGBM: A Highly Efficient Gradient Boosting Decision Tree: Neural Information Processing Systems. 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, 4-7 December 2017, 569-577.
- [5] 尹兴平, 蒋裕强, 付永红, 等. 渝西地区五峰组——龙马溪组龙一 1 亚段页岩岩相及储层特征[J]. 岩性油气藏, 2021, 33(4): 41-51.
- [6] 韩贵生, 雷治安, 徐剑良, 等. 四川盆地深层页岩气储层特征及高产主控因素——以渝西区块五峰组——龙马溪组为例[J]. 天然气勘探与开发, 2020, 43(4): 112-122.
- [7] 张成林, 张鉴, 李武广, 等. 渝西大足区块五峰组——龙马溪组深层页岩储层特征与勘探前景[J]. 天然气地球科学, 2019, 30(12): 1794-1804.
- [8] 舒红林, 何方雨, 李季林, 等. 四川盆地大安区块五峰组——龙马溪组深层页岩地质特征与勘探有利区[J]. 天然气工业, 2023, 43(6): 30-43.
- [9] 邹辰, 吴永辉, 章超, 等. 渝西地区龙马溪组页岩低阻主控因素及有利区预测[J]. 东北石油大学学报, 2023, 47(2): 81-90.
- [10] Friedman, J.H. (2001) Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*, **29**, 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- [11] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [12] 叶玥豪, 陈伟, 汪华, 等. 四川盆地上二叠统大隆组页岩储层特征及其控制因素[J]. 石油与天然气地质, 2024, 45(4): 979-991.