第 28 卷　第 1 期
2026 年　2 月

古　地　理　学　报
JOURNAL OF PALAEOGEOGRAPHY（Chinese Edition）

Vol. 28　No. 1
Feb. 2026

文章编号：1671-1505(2026)01-0381-17　　**DOI**:10.7605/gdlxb.2026.048

基于多梯度提升器算法的四川盆地茅口组一段含海泡石层系分布预测*

宋金民[1]　任　杉[1]　刘树根[1,2]　文　龙[3]　李智武[1]　罗　冰[3]
李柯然[1]　杨　迪[1]　汪　华[3]　叶玥豪[1]　金　鑫[1]
张钊益[1]　郭嘉欣[1]　邵兴鹏[1]　张祖兵[1]

1 成都理工大学油气藏地质及开发工程全国重点实验室，四川成都 610059
2 西华大学，四川成都 610039
3 中国石油西南油气田公司，四川成都 610051

摘　要　四川盆地中二叠统茅口组一段含海泡石层系具有自生自储、源内成藏的特点，为碳酸盐岩非常规气藏勘探的新领域，目前对其的识别和分布预测较为薄弱。基于岩心、薄片、XRD、测井与录井资料，通过敏感性分析选取 CNL、DEN、GR、RT、RXO 和 AC 6 条测井曲线，利用 SMOTE 算法处理特征平衡性差异，提出含海泡石层系基于多梯度提升器算法的综合预测流程。具体如下：(1) CatBoost 算法以滑石作为预测分析载体，对是否含海泡石作二分类判断；(2) CatBoost算法对含海泡石层系进行滑石产状的多分类判断；(3) XGBoost算法对含海泡石层系的滑石含量进行回归预测；(4) 识别不同产状含海泡石层系的有效厚度。茅一段含海泡石层系(滑石) 主要发育斑点状、透镜状和层状 3 种产状。预测结果显示，斑点状滑石发育在川北和川中—蜀南地区，厚呈向北东增大趋势；透镜状滑石主要分布在川西地区；层状滑石见于川西北、川中和蜀南地区。四川盆地茅一段含海泡石层系整体具有“北东厚、南西薄”的分布特征，通江—长寿凹陷地区为含海泡石层系的沉积中心，合川—威远—泸州地区为含海泡石层系优势发育区，这为下一步的勘探部署提供了依据。

关键词　梯度提升器算法　机器学习　含海泡石层系　茅一段　二叠系　四川盆地

第一作者简介　宋金民，男，1983 年生，博士，成都理工大学教授，主要从事油气储层地质的研究和教学工作。E-mail：songjinmin2012@cdut.edu.cn。

中图分类号：P588.24^{+}5　　**文献标志码**：A

Distribution prediction of the sepiolite-containing succession in the Member 1 of Maokou Formation in Sichuan Basin based on multi-gradient boosting algorithm

SONG Jinmin[1]　REN Shan[1]　LIU Shugen[1,2]　WEN Long[3]　LI Zhiwu[1]　LUO Bing[3]
LI Keran[1]　YANG Di[1]　WANG Hua[3]　YE Yuehao[1]　JIN Xin[1]
ZHANG Zhaoyi[1]　GUO Jiaxin[1]　SHAO Xingpeng[1]　ZHANG Zubing[1]

1 *State Key Laboratory of Oil and Gas Reservoir Geology and Exploitation, Chengdu*

* 国家自然科学基金项目（编号：42572132，U24B6001）和国家科技重大专项（编号：2025ZD1400403）联合资助。[Co-funded by the National Naturral Science Foundation of China（Nos. 42572132，U24B6001）and National Science and Technology Major Project（No. 2025ZD1400403）]
收稿日期：2025-08-29　　改回日期：2025-10-23

University of Technology, *Chengdu* 610059, *China*
2 *Xihua University*, *Chengdu* 610039, *China*
3 *PetroChina Southwest Oil & Gas Field Company*, *Chengdu* 610051, *China*

Abstract The sepiolite-containing succession in the Member 1 of the Middle Permian Maokou Formation in the Sichuan Basin exhibit self-generation and self-accumulation characteristics, positioning them as a promising new frontier for the exploration of unconventional gas reservoirs. However, the identification and predictive distribution of these strata are still underdeveloped. This study utilizes core samples, thin sections, X-ray diffraction (XRD) analysis, well logging, and mud logging data, performing sensitivity analysis to select six key well log curves: CNL, DEN, GR, RT, RXO, and AC. The SMOTE algorithm is employed to address feature imbalance. The workflow for predicting sepiolite-containing succession based on multi-gradient boosting algorithms is as follows: (1) The CatBoost algorithm is used for binary classification to determine the presence of sepiolite; (2) CatBoost performs multi-class classification to categorize the morphology of sepiolite-containing succession; (3) XGBoost is applied for regression analysis to predict the talc content in the sepiolite-containing succession; (4) The effective thickness of sepiolite-containing succession is identified based on talc content. The sepiolite-containing succession in the Member 1 of Maokou Formation primarily exhibit three morphologies: spotty, lenticular, and layered. Prediction results indicate that spotty talc is mainly developed in the northern and central southern regions of Sichuan Basin, with thickness increasing toward the northeast; lenticular talc is predominantly distributed in the western Sichuan region; and layered talc is mainly found in north western, central and southern Sichuan Basin. Overall, the sepiolite-containing succession in the Member 1 of Maokou Formation show a distribution pattern of 'thicker in the northeast, thinner in the southwest.' The Tongjiang-Changshou intracratonic sag area and the Hechuan-Weiyuan-Luzhou region serve as the sedimentary centers for the effective thickness of sepiolite-containing succession, providing a basis for future exploration and deployment in these areas.

Key words gradient boosting machines, machine-learning, sepiolite containing succession, Member 1 of Maokou Formation, Permian, Sichuan Basin

About the first author SONG Jinmin, born in 1983, Ph. D., is a professor at Chengdu University of Technology. He is specialises in research and teaching within the field of petroleum reservoir geology. E-mail: songjinmin2012@cdut.edu.cn.

0 引言

近年来，四川盆地中二叠统茅口组一段（以下简称茅一段）泥质灰岩储集层的天然气勘探取得系列突破，共 19 口探井钻获天然气（夏文谦等，2023；宋金民等，2024a；王斌等，2024；翟常博等，2024），其中 10 口井测试日产气量超过 2×10^4 m^3（宋金民等，2024a），TT1 和 DB1H 井测试日产气量达 31.00×10^4 m^3 和 55.6×10^4 m^3（范建平等，2022；梁兴等，2024），揭示出其巨大的勘探潜力（梁兴等，2024）。

四川盆地茅一段为碳酸盐岩缓坡沉积（宋金民等，2024b），为泥岩—泥质灰岩—灰岩韵律层构成的含海泡石层系（宋金民等，2024b）。茅一段泥岩—泥质灰岩韵律结构内富含海泡石（滑石），其具有极强的吸附性，能够促进有机质富集，提升烃源岩品质（宋金民等，2024b），同时层内发育大量的有机质孔、裂缝、溶孔（缝）、粒缘孔（缝）和滑石收缩孔（缝）（范建平等，2022），形成自生自储、源内成藏、大面积连续分布的天然气藏（张培先等，2021；张宇等，2022）。前期针对含海泡石层系的研究主要集中在海泡石沉积模式（宋金民等，2024a）、黏土矿物类型及产

状（宋金民等，2024b）和源储组合模式（宋金民等，2022）等方面。四川盆地茅一段海泡石产状主要有透镜状、斑点状、层状和生屑状，不同岩性内产状不一，石灰岩内以透镜状、斑点状为主，泥岩—泥质灰岩则可见层状海泡石（宋金民等，2022）；含海泡石层系内主要自生黏土矿物类型为海泡石和滑石（最终演化结果）；形成自生自储和旁（下）生侧（上）储型的高效源储配置模式，其最有利勘探区主要在“C”形的通江—长寿凹陷内（刘树根等，2022）。

随着人工智能、机器学习及大数据运算技术的快速发展，油气勘探领域正在经历由经验驱动向数据驱动的转型。这些前沿技术能够在大规模、多维度的地质与测井资料中高效提取特征与规律为复杂地质体系的精细刻画和有利区预测提供了新的途径。Wang 等（2024）以测井数据为基础，采用反向传播神经网络（BPNN）预测区块内裂缝发育度，建立了地应力控制下的裂缝发育分布模型。Koray 等（2024）结合无监督聚类与多种监督回归算法改进机器学习流程，有效提升了孔隙度与渗透率建模精度。前人已将机器学习算法应用到茅一段的研究中，包括含滑石多矿物优化模型（殷树军等，2023）、BP 神经网络算法（王心乾等，2024）、多元回归预测方法（蒋文博等，2018）、基于极限学习机算法的神经网络模型（白烨等，2021），主要是利用多维测井数据进行黏土矿物含量的预测。目前针对茅一段海泡石—滑石类黏土矿物产状的预测则相对欠缺，且海泡石（滑石）含量相对较低（刘树根等，2022），直接运用“端到端”预测模型可能出现偏差。因此需要先利用机器学习模型判定海泡石是否发育，然后增加海泡石层系类型的分类任务，目前这方面的研究尚显薄弱。作者将滑石视为海泡石演化的最终产物，利用梯度提升器算法（Friedman，1999）中的 CatBoost（Prokhorenkova *et al.*，2018）及 XGBoost（Chen and Guestrin，2016）模型重新审视“端到端”预测流程，以海泡石（滑石）作为预测分析载体对含海泡石层系存在性（二分类问题）、海泡石层系的滑石产状（多分类问题）和滑石含量（回归问题）进行综合分析，建立适用于含海泡石层系的综合预测的多梯度提升器算法工作流程，对四川盆地茅一段含海泡石层系不同滑石产状的分布进行预测，为后续的勘探提供参考。

1　区域地质背景

中二叠世茅口组沉积时期，四川盆地总体处于稳定的克拉通演化阶段（宋金民等，2024a），茅一段沉积时期，继承栖霞组的沉积格局，为碳酸盐岩缓坡沉积环境。受控于洋盆扩张和台地内部的正断层活动（刘昇等，2023），栖霞期形成的通江—长寿凹陷进一步扩大，凹陷内沉积大量含海泡石层系，呈现“两台一凹”的构造-沉积格局（宋金民等，2023，2024b）（图 1-A）。

茅口组自下而上划分为茅一段、茅二段、茅三段和茅四段。茅一段为碳酸盐岩中缓坡—外缓坡沉积（杨跃明等，2020；范建平等，2023），底部发育硅质层和泥岩，向上以“眼皮眼球”状灰岩为主，眼球灰岩为泥晶生屑灰岩和生屑泥晶灰岩，眼皮灰岩为泥质生屑灰岩；茅二段—茅三段水体持续变浅，由内缓坡过渡为弱镶边型台地沉积，发育生屑泥晶灰岩、泥晶生屑灰岩和白云岩（蒋航等，2023；梁兴等，2023）；茅四段受东吴运动影响，遭受了不同程度的剥蚀，在盆地内发育不完整，以泥晶灰岩和生屑灰岩为主（宋金民等，2024a）。

茅一段自上而下细分为 a、b 和 c 3 个亚段，茅一 a 亚段为厚层状泥晶生屑灰岩、生屑泥晶灰岩夹少量泥质生屑灰岩，测井曲线上表现为自然伽马（GR）较高，呈箱状，声波时差（AC）向上减小，补偿中子（CNL）呈锯齿状变化；茅一 b 亚段主要为泥质生屑灰岩、泥晶生屑灰岩，自然伽马（GR）相对茅一 a 亚段略微升高，中下部为箱状，顶界处呈锯齿状高值，声波时差（AC）、补偿中子（CNL）值呈箱形降低；茅一 c 亚段主要为泥质生屑灰岩和泥晶生屑灰岩的韵律层，见少量泥岩和硅质岩，电性表现为底部锯齿状自然伽马（GR）高值，向上降低呈箱状高值，声波时差（AC）呈锯齿状高值（图 1-B）。含海泡石（滑石）层系的自然伽马（GR）测井曲线形态较为平缓，判别标准为 28~65 API；电阻率（RT）测井曲线呈锯齿状，其值小于 100 Ω·m；声波时差（AC）测井曲线平缓或呈钟形，分布于 50~77 μs/m 之间；补偿密度（DEN）测井曲线平缓；补偿中子测井（CNL）曲线呈平缓状或钟形，判别标准值介于 0.75%~11.5%之间（图 1）。

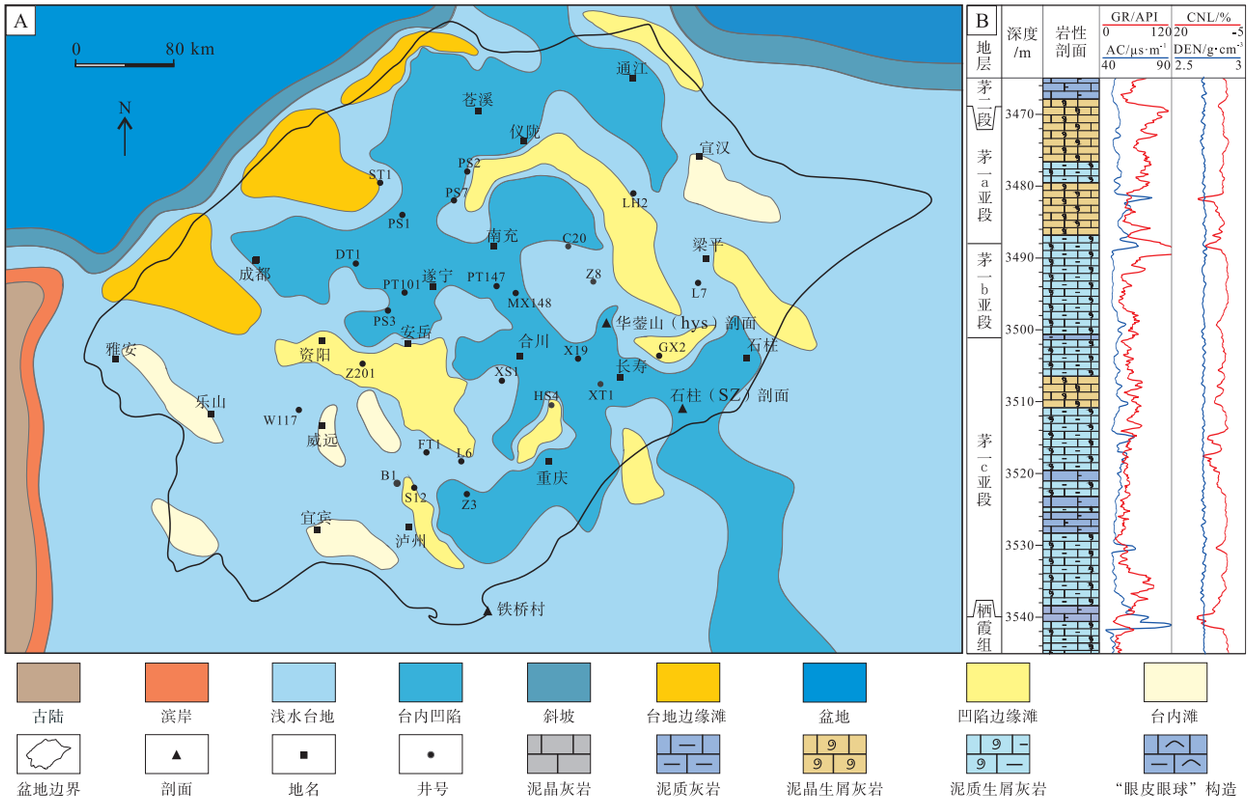


图 1 四川盆地茅一段沉积相图（A）与综合柱状图（B）（修改自宋金民等，2024b）

Fig. 1 Sedimentary facies map（A）and integrated stratigraphic column（B）of the Member 1 of Maokou Formation of Sichuan Basin（modified from Song *et al.*, 2024b）

2 含海泡石层系岩石学特征

茅一段含海泡石层系主要赋存于泥岩—泥质灰岩—灰岩韵律层中（宋金民等，2022）。含海泡石层系在宏观露头剖面上以层状或眼皮状产出（宋金民等，2024a），在岩心上呈黑色斑点状、透镜状及层状滑石与灰岩不规则状交互产出（图 2）。含海泡石层系中斑点状滑石赋存于含生屑泥晶灰岩或泥晶生屑灰岩中，集中发育在茅一 a 亚段，多沿层理分布，斑点直径为 2～5 mm，单层厚度范围 0.25～0.76 m，整体占比 55%～66%，表现出多层叠置和韵律性发育的特征（图 2-A）。透镜状滑石多赋存在泥晶生屑灰岩中，多见于茅一 b 亚段及茅一 c 亚段，以黑色透镜状状沿层理嵌套产出，单层厚度 0.13～0.66 m，整体占比 15%～27%，部分可见定向排列现象，与灰岩呈穿插共生状态（图 2-B）。层状滑石以黑色层状产出于泥岩或泥质灰岩，主要发育在茅一 b 亚段及茅一 c 亚段，单层厚度 0.64～0.72 m，整体占比 0%～21%，厚度较均一、连续性好（图 2-C）。

茅一段含海泡石层系主要由泥质灰岩、生屑泥晶灰岩和泥岩组成。泥质灰岩镜下呈颗粒结构，灰泥基质含量为 20%～30%，黏土矿物含量为 20%～35%，呈波状起伏；颗粒类型以生物碎屑为主，含量 30%～60%，包括有孔虫、腕足类和双壳类（图 3-A），生物壳体易发生硅化（图 3-B），整体较破碎并呈定向分布的特征（图 3-C）。石灰岩镜下呈泥晶结构，生屑含量为 10%～30%，生屑保存较完整，可见斑点状滑石（图 3-D）、透镜状滑石（图 3-E），滑石含量为 5%～15%，生屑间为灰泥基质。泥岩主要成分为滑石和石英，镜下见大量滑石聚集（图 3-F），滑石含量为 30.0%～75.4%，石英含量为 19.7%～24.6%。

含海泡石层系中的滑石在单偏光下呈浅褐色，正交偏光下最高干涉色可以达到Ⅲ级橙色。滑石切面为片状、鳞片状和纤维状结构，常与有机质共生（胡东风等，2020）。斑点状滑石零星分布于灰泥

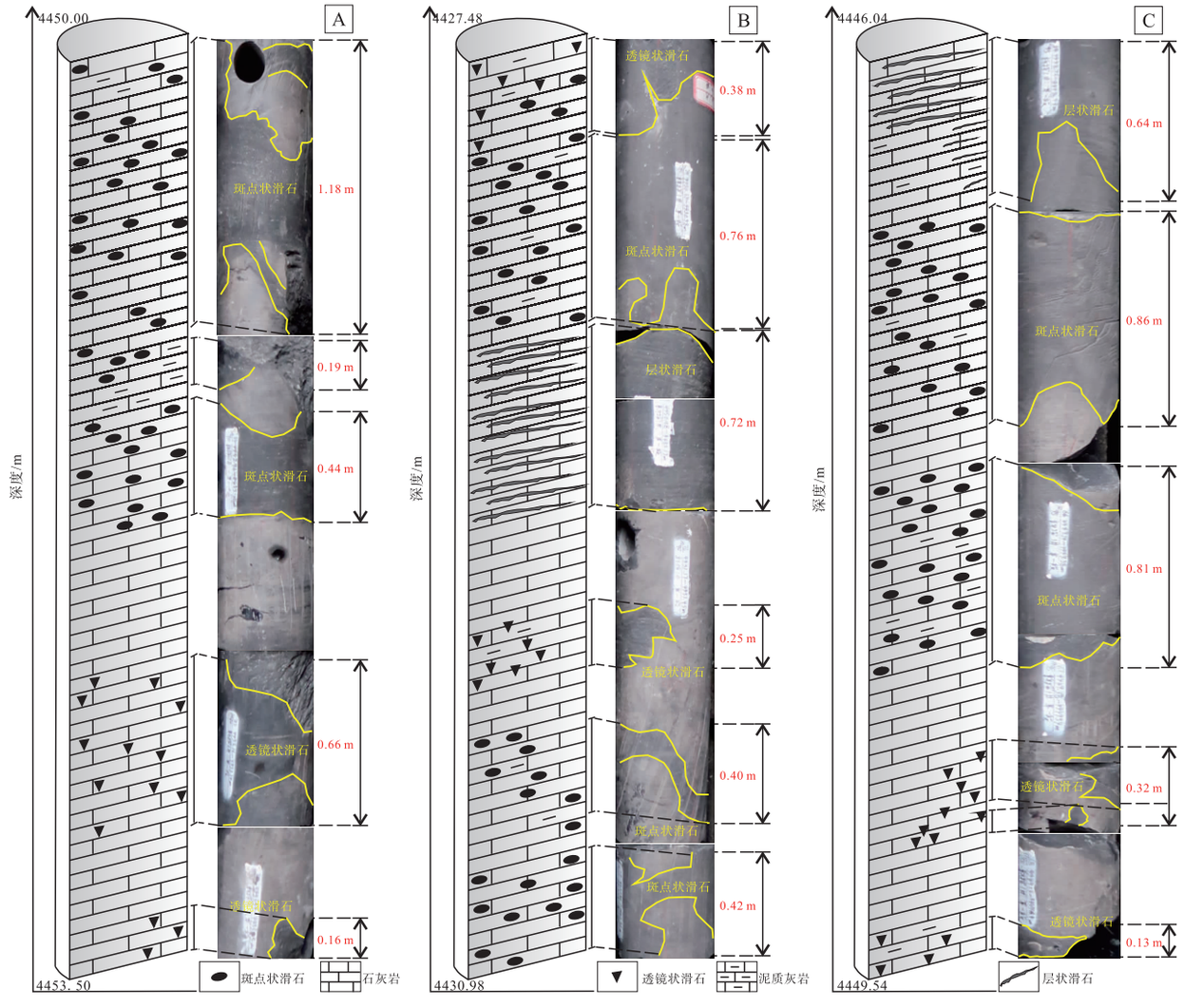


图 2　四川盆地 XT1 井茅一段不同产状的海泡石（滑石）的宏观发育特征

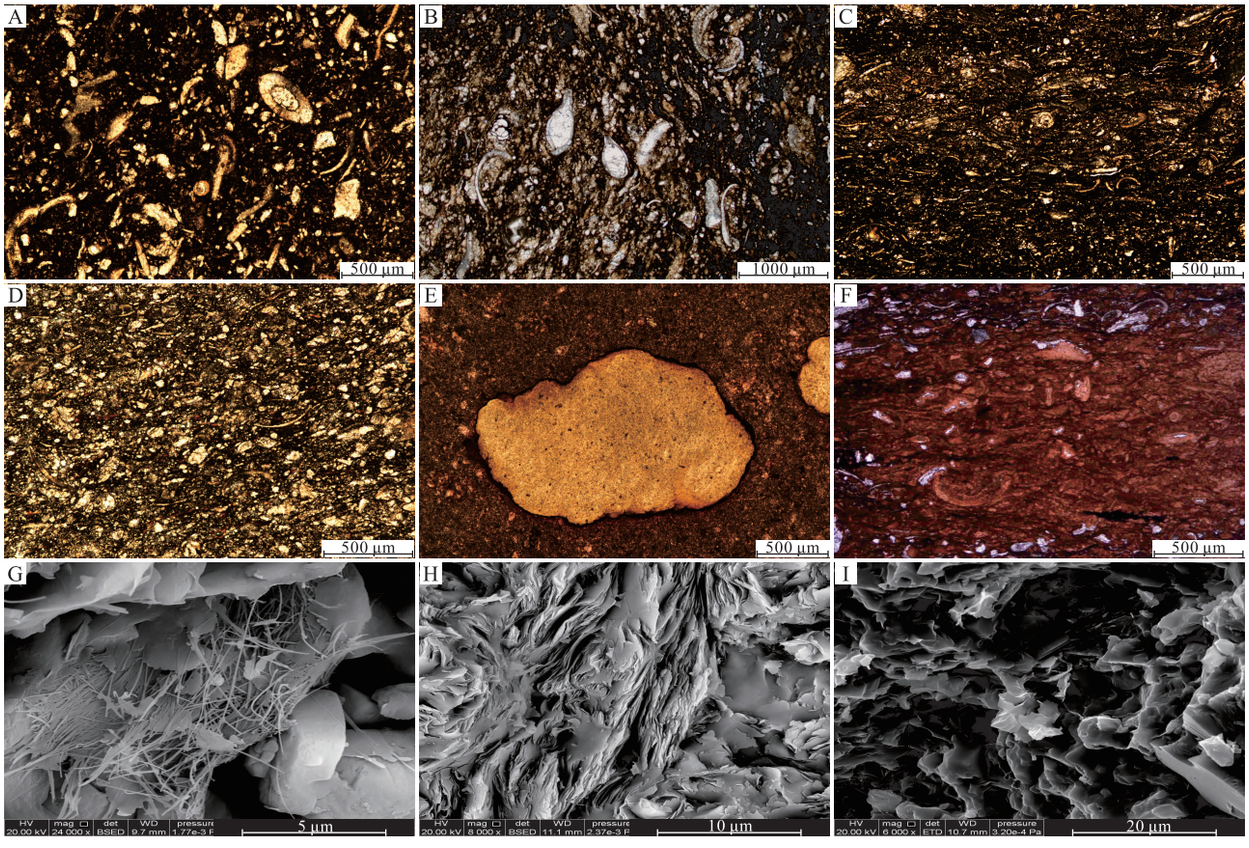
Fig. 2　Macro-developmental characteristics of sepiolite (talc) containing formation in the Member 1 of Maokou Formation of Well XT1 in Sichuan Basin

基质中(图 3-D)，粒径小，分布不均匀。XRD 结果显示滑石含量为 0.15%～0.75%，石英含量为 0.75%～1.72%，方解石含量为 93%～98.64%。透镜状滑石多为片状集合体(图 3-E)，呈不均匀分布于灰泥基质中，滑石间发育明显的收缩缝，XRD 结果显示滑石含量为 0.14%～3.27%，石英含量为 0.36%～5.53%，方解石含量为 93.20%～98.26%。透镜状滑石在扫描电镜下可见海泡石向滑石转化的过渡形态。层状滑石由片状滑石集合体组成(图 3-F)，分布于方解石晶粒间。XRD 结果显示滑石含量为 8.79%～22.17%，石英含量为 0.69%～8.77%，方解石含量为 70.87%～84.50%。

扫描电镜下，海泡石单体为纤维状(图 3-G)，滑石以片状结构为特征(图 3-H，3-I)，呈片状集合体充填于方解石之间，存在“海泡石—斯蒂文石—无序滑石—滑石”的成岩演化序列，指示滑石是海泡石转化而成（宋金民等，2024b），因此文中以滑石的累计厚度表征茅一段含海泡石层系。

3　基于机器学习的测井数据识别

X1 井测井响应特征表明，海泡石（滑石）含量与自然伽马（GR）、电阻率（RT）、冲洗带电阻率（RXO）、声波时差（AC）与补偿中子（CNL）显著正相关(图 4)，可作为海泡石（滑石）判定



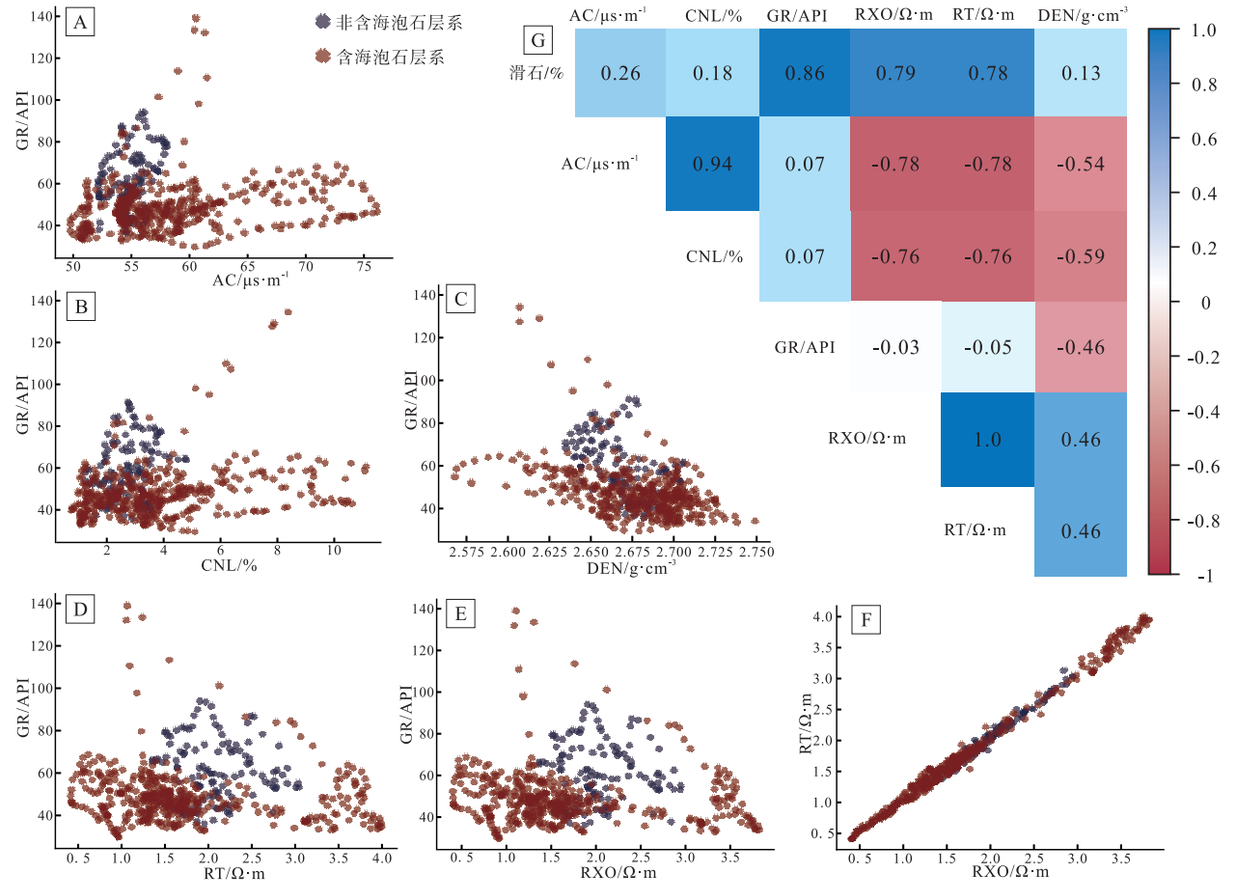
A—泥质灰岩，见腕足、有孔虫、双壳，单偏光，×5，hys 剖面；B—泥质灰岩，生物壳体发生硅化，单偏光，×5，井深 4445.32 m，XT1 井；C—泥质灰岩，生物碎屑破碎、呈定向分布，单偏光，×5，hys 剖面；D—生屑泥晶灰岩，斑点状滑石，×5，单偏光，hys 剖面；E—生屑泥晶灰岩，透镜状滑石，×5，单偏光，hys 剖面；F—泥岩，层状滑石，星点状硅质，×5，单偏光，井深 3425.96 m，CZO 井；G—纤维状海泡石，扫描电镜，SZ 剖面；H—滑石，扫描电镜，井深 4431.31 m，XT1 井；I—滑石，扫描电镜，井深 4433.87 m，XT1 井

图 3 四川盆地茅一段不同产状滑石的中观和微观特征

Fig. 3 Petrological characteristics of the sepiolite containing formation in the Member 1 of Maokou Formation of Sichuan Basin

指标。海泡石（滑石）层系的自然伽马（GR）判别标准为 28~65API（图 4-A 至图 4-E），GR 与含海泡石层系中的滑石含量相关系数为 0.86（图 4-G）。声波时差（AC）分布于 50~77 μs/m 之间（图 4-A），AC 与滑石含量相关系数为 0.26（图 4-G）。补偿中子测井（CNL）判别标准值介于 0.75%~11.5%之间（图 4-B），CNL 与石含量相关系数为 0.18（图 4-G）。补偿密度（DEN）分布于 2.625~2.725 g/cm^3 之间（图 4-C），DEN 与滑石含量相关系数为 0.13（图 4-G）。地层电阻率（RT）与冲洗带电阻率（RXO）均分布于（0.5~2）×10^{5}Ω·M 之间（图 4-D，4-E），RT、RXO 与滑石含量相关系数分别为 0.78、0.79（图 4-G）。同时，声波时差（AC）与补偿中子（CNL）、地层电阻率（RT）、冲洗带电阻率（RXO）及补偿密度（DEN）相关性较强，相关系数分别为 0.94、-0.78、-0.78 及-0.54。补偿中子（CNL）与地层电阻率（RT）、冲洗带电阻率（RXO）及补偿密度（DEN）相关系数分别为-0.76、-0.76、-0.59。测井数据的相关性特征表明，该判别标准能够有效识别含海泡石层系。

梯度提升器算法（Grandient Boosting Machine，GBM）是 Boosting（提升）算法的一种，是基于加法模型（additive model）和梯度下降的集成学习方法（Friedman，1999）。核心思想是串行地生成多个弱学习器，每个弱学习器的目标是拟合先前累加模型的损失函数的负梯度，使累积模型损失往负梯度的方向减少。通过不同的权重将机器学习器进行线性组合，使表现更优秀的学习器得到重用，有效提高预测精度，广泛应用于分类和回归问题。其代



A—自然伽马（GR）与声波时差（AC）交会图；B—自然伽马（GR）与补偿中子（CNL）交会图；C—自然伽马（GR）与补偿密度（DEN）交会图；D—自然伽马（GR）与地层电阻率（RT）交会图；E—自然伽马（GR）与冲洗带地层电阻率（RXO）交会图；F—地层电阻率（RT）与冲洗带地层电阻率（RXO）交会图；G—海泡石（滑石）相关系数矩阵（AC. 声波时差，CNL. 补偿中子，GR. 自然伽马，RXO. 冲洗带地层电阻率，RT. 地层电阻率，DEN. 补偿密度）（刘树根等，2022）

图 4　测井特征与海泡石层系（滑石）关系图

Fig. 4　Plot of logging features versus sepiolite (talc) containing formation

表性的算法包括 CatBoost（Prokhorenkova *et al.*, 2018）、XGBoost（Chen and Guestrin, 2016）等，其在处理类别特征和提高准确性方面表现突出。

梯度提升器算法中，CatBoost 使用“有序提升”（Ordered Boosting）和随机排列技术，直接处理类别特征，减少了目标泄漏（target leakage）问题，能够更好地捕捉数据中的复杂关系，适用于分类任务（Prokhorenkova *et al.*, 2018）。XGBoost 是一种高效的梯度提升算法，其分布式计算能力和优化大规模数据集能力能够高效处理高维数据并且能有效减少过拟合（Chen and Guestrin, 2016）。因此，针对四川盆地含海泡石层系存在性二分类问题以及滑石产状多分类问题选取 CatBoost 模型，滑石含量回归预测任务选取 XGBoost 模型。

3.1　数据处理

收集四川盆地 127 口含海泡石层系数据，采用 Z-score 方法对每一道曲线数据进行标准化处理，标注出—含海泡石层系存在性—滑石产状分类标签。对于测井数据 X_k 及其对应标签 Y，构成集合 D，$D=\{X_k,Y_k\}\,k=1,2,\cdots,n$，其中，$X_k$ 表示第 k 道测井曲线数据，对于实际测井曲线构成的特征数据集，X_k 表示为：

$$X_k=(X_1^T,\cdots X_n^T)=\begin{pmatrix}(x_1^1,\cdots,x_1^m)^T\\ \vdots\\ (x_k^1,\cdots,x_k^m)^T\\ \vdots\\ (x_n^1,\cdots,x_n^m)^T\end{pmatrix}\tag{1}$$

式中 m 表示纵向采样数目，$m=\frac{\text{地层总厚度}}{\text{采样间隔}}$。含海泡石层系预测以滑石作为预测分析载体。滑石标签 Y_k 包含地层是否存在滑石、滑石产状类型及滑石含量，表示为：

$$Y_k=(Y_1^T,\ Y_2^T,\ Y_3^T)=\begin{pmatrix}(y_1^1,\ \cdots,\ y_1^m)^T\\(y_2^1,\ \cdots,\ y_2^m)^T\\(y_3^1,\ \cdots,\ y_3^m)^T\end{pmatrix} \quad (2)$$

式中 Y_1^T 表示地层是否存在滑石，$Y_1^T\in[0,\ 1]$，0 表示不发育滑石，1 表示发育滑石。Y_2^T 表示滑石含量。Y_3^T 表示滑石产状，$Y_3^T\in[0,\ 1,\ 2]$，0 表示斑状，1 表示透镜状，2 表示层状。

利用机器学习算法对滑石各类标签预测问题，表示为：

$$\forall\, y_i^j\in Y_k,\ \exists\, F_i^j(x)\,|\,R^m\rightarrow R\Rightarrow \min(Loss(F_i^j(x),\ y_i^j)) \quad (3)$$

数据预处理包含四川盆地待预测测井数据类型统计和完整性检验。在保证数据集和预测集数据特征一致的前提下，将测井曲线最低频率和完整性下限制设置为 0.9，统计结果见表 1。引入随机森林参数打分法进一步检验与滑石发育的相关性（Hastie *et al.*, 2009）（表 2），在确保测井曲线完整度前提下，综合结果选择自然伽马（GR）、声波时差（AC）、补偿中子（CNL）、密度（DEN）、地层真电阻率（RT）、冲洗带地层电阻率（RXO）曲线作为数据特征，采用 MinMaxScaler 方法（Pedregosa *et al.*, 2011）进行数据归一化。

表 1 四川盆地茅一段样本井与待预测井不同测井曲线完整度统计

Table 1 Statistics on completeness of different logging curves of sample wells and wells to be predicted in the Member 1 of Maokou Formation of Sichuan Basin

类型	不同测井曲线完整度									
	GR	**AC**	**CNL**	**DEN**	**RT**	**RXO**	SP	PE	MSFL	CAL
样本井	**0.99**	**0.98**	**0.97**	**0.92**	**0.91**	**0.91**	0.85	0.64	0.43	0.30
待预测井	**0.99**	**0.97**	**0.96**	**0.93**	**0.90**	**0.90**	0.82	0.70	0.38	0.28

然而，含海泡石—滑石层系二分类问题和多分类问题中，存在滑石空间局部分布不均、二分类和多分类任务过程中的焦点差异导致样本标签分布不均衡，造成模型过采样/欠采样（Cao *et al.*, 2019）引入 SMOTE（Synthetic Minority Oversampling Technique）算法对少数样本进行补充（Chawla *et al.*, 2002）。SMOTE 通过在少数类样本的特征空间中构造合成样本来扩充数据集（公式 4），共获得测井数据—滑石/非滑石、测井—滑石产状标签 50800 条（图 5）。

表 2 四川盆地茅一段随机森林打分器对不同测井曲线打分结果统计

Table 2 Statistics of scoring results of different logging curves by random forest scorer in the Member 1 of Maokou Formation of Sichuan Basin

测井类型	**GR**	**AC**	CAL	**DEN**	**RT**	**RXO**	**SP**	**PE**	MSFL	**CNL**
分数	**0.54**	**0.21**	−0.034	**0.24**	**0.21**	**0.21**	**0.34**	**0.42**	−0.17	**0.28**

注：加粗数据表示测井数据得分大于 0.2。

$$x_{new}=X_i+\delta\cdot(x_{nn}-x_i),\ \delta\in[0,\ 1] \quad (4)$$

其中，x_i 为少数类样本，x_{nn} 为其在 k 近邻中的任意一个邻居样本，δ 为 [0, 1] 区间的随机数。通过该方式在原有少数类样本之间插值生成新的样本点。

3.2 模型建立

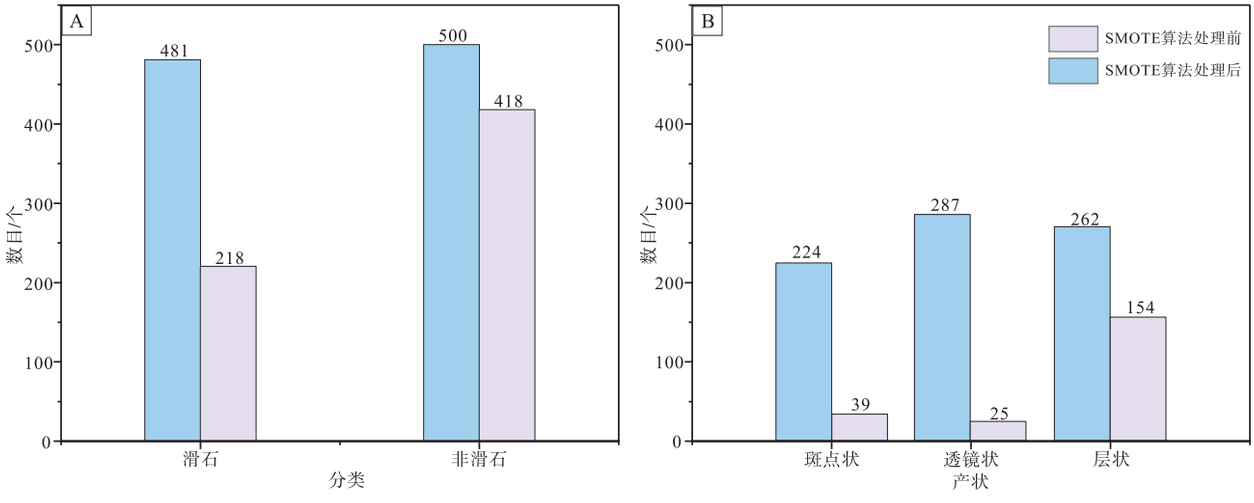
CatBoost 是基于对称决策树的梯度提升模型（Gradient Boosting Decision Tree，GBDT），将类别特征（Categorical Features）与梯度提升（Gradient Boosting）组合，能够评估变量重要性（Friedman, 1999）（图 6-A）。基于预测值更改（Prediction Values Change）的特征值 F 是模型中决策树（trees）上所有叶节点（leafs）的加权和，用下式表示：

$$I_F=\sum_{trees,\ leafs_F}c_1(v_1-avr)^2+c_2(v_2-avr)^2 \quad (5)$$

式中，$avr=\frac{c_1v_1+c_2v_2}{c_1+c_2}$，$c_1$ 和 c_2 分别表示左叶节点和右叶节点中对象的总权重，v_1 和 v_2 分别表示左叶节点和右叶节点的值。

XGBoost（eXtreme Gradient Boosting）极致梯度提升（Prokhorenkova *et al.*, 2018），也是基于 GBDT 的一种算法。给定 1 个包含 n 个样本和 m 个特征的数据集 $D=\{(x_i,\ y_i)\}$（$|D|=n$ $|D|=n$，$x_i\in R^m$，$y_i\in R$），梯度提升树模型使用 K 个加性函数来预测输出结果（图 6-B）。

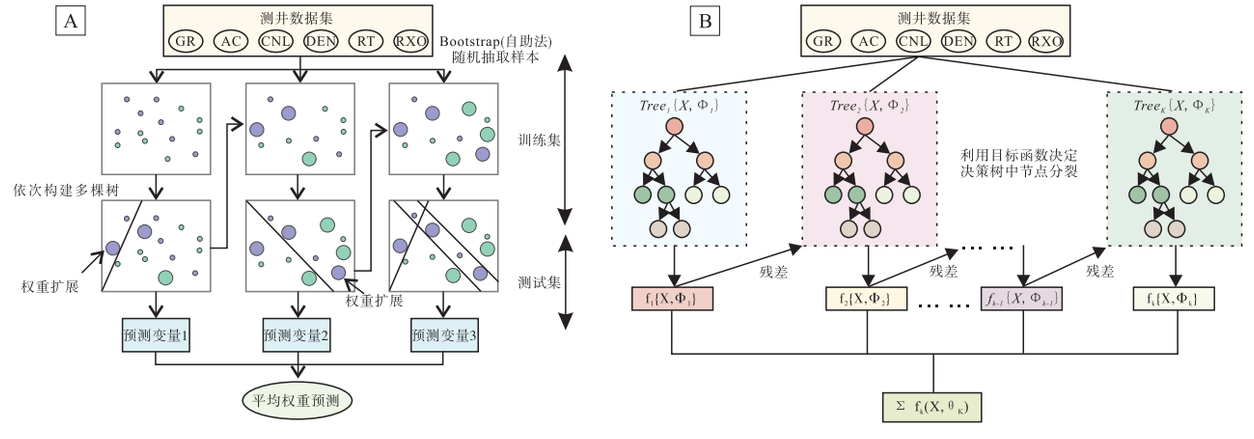
$$\hat{y}_i=\varphi(x_i)=\sum_{k=1}^{K}f_k(x_i),\ f_k\in F \quad (6)$$



A—滑石及非滑石分类数据引入 SMOTE 算法前后样本分布；B—引入 SMOTE 算法滑石产状分类前后样本分布

图 5　引入 SMOTE 算法前后样本数据分布

Fig. 5　Distribution of sample data before and after the introduction of SMOTE algorithm



A—CatBoost 模型结构；B—XGBoost 模型结构

图 6　模型结构示意图

Fig. 6　Schematic diagram of the model structure

式中，$F=\{f(x)=wq(x)\}(q:R^m\rightarrow T, w\in RT)$ 是回归树（CART）的空间；q 表示每棵树的结构；T 是树中叶子的数量。每个 f_k 对应于一个独立的树结构 q 和叶子权重 w；w_i 表示第 i 个叶子上的分数；使用树中的决策规则（q）将其分类到叶子中，并通过对应叶子中的分数求和，用于计算最终预测值。

学习模型中使用的函数集合采用 L2 正则化目标（Dorogush *et al.*, 2018）：

$$L(\varphi)=\sum_i l(\hat{y}_i, y_i)+\sum_k \Omega(f_k) \tag{7}$$

$$\Omega(f)=\gamma T+\frac{1}{2}\lambda \| w \|^2 \tag{8}$$

式中，l 表示损失函数；正则化项 Ω（f）2 个参数：γ 为控制树的数量；λ 为控制叶子节点权重的 $L2$ 范数 $\| w \|$。

3.2.1　模型参数

神经网络模型由 AutoKeras ®工具包设计（Jin *et al.*, 2019），保证模型达到最优化。

3.2.1　训练参数

训练过程中首先随机预留部分数据作为检验集，设置数据集比例为训练集：验证集：检验集 = 7:1:2。对于含海泡石（滑石）层系存在性二分类和滑石产状分类问题，模型训练效果均由交叉熵函数计算，对于二分类问题，交叉熵函数为：

$$Cross\_entropy\_binary = \frac{1}{N}\{-[y_i \cdot \log(p_i) + (1-y_i)\cdot\log(1-p_i)]\} \quad (9)$$

式中，N 表示样本总数，y_{ij} 表示第 i 个样本的标签，p_i 表示第 i 个样本的预测值，$p_i \in [0, 1]$。

多分类问题作为二分类问题的拓展，交叉熵计算方法为：

$$Cross\_entropy\_multi = -\frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{class} y_{ij}\cdot\log(p_{ij}) \quad (10)$$

式中，N 表示样本总数，y_{ij} 表示第 i 个样本包含类别 c 的标签，p_{ic} 表示第 i 个样本对于类别 c 的预测值。

回归问题采用均方根误差（Root Mean Square Error，RMSE）（Hyndman and Koehler，2006）进行计算，计算公式如下：

$$RMSE = \sqrt{\sum_{i=1}^{n}(Y_i - f(xi))^2} \quad (11)$$

4 结果

4.1 模型结果

4.1.1 分类问题结果

通过梯度提升器算法中的 CatBoost 及 XGBoost 模型同 LightGBM 模型、随机森林（Random Forest）、人工神经网络（Artificial Neural Network，ANN）、支持向量机（Support Vector Machine，SVM）、KNN（K-Nearest Neighbors）及贝叶斯网络（Naive Bayes）进行对比发现（表 3），相同训练参数（训练轮次为 500）下梯度提升器算法对含海泡石层系的分类问题及回归预测效果均优于其他模型。

综合对比梯度提升器算法中的 CatBoost 及 XGBoost 模型在含海泡石层系存在性二分类问题的表现。训练结果显示，CatBoost、XGBoost 模型在训练误差和验证误差方面表现接近，均为 0.01 与 0.03，训练与验证准确度均达到 98% 与 97%。滑石产状的多分类问题模型表现结果显示，CatBoost 模型保持了较高准确性，CatBoost 模型训练误差与验证误差分别为 0.01 与 0.02，训练准确度为 96%，验证准确度为 92%。XGBoost 模型的训练误差与验证误差分别为 0.07 与 0.11，训练与验证准确度分别为 92% 与 91%，表现出过拟合倾向和泛化能力不足的特征。基于两类任务中的优异表现，选取 CatBoost 模型作为主要分类器，分别用于含海泡石（滑石）层系存在性二分类与滑石产状多分类的预测分析。

4.1.2 回归问题结果

CatBoost、XGBoost 模型均能处理回归问题。对于滑石含量的回归预测任务，在迭代次数为 50 时，2 种模型均表现出一定的拟合能力。对比结果显示 XGBoost 模型的性能更为优越。XGBoost 的训练误差与验证误差分别为 0.07 与 0.05，显著低于 CatBoost 的 0.25 与 0.24。表明在滑石含量的连续变量预测中，XGBoost 在拟合精度与泛化能力上更具优势，更适合作为回归分析的首选模型。

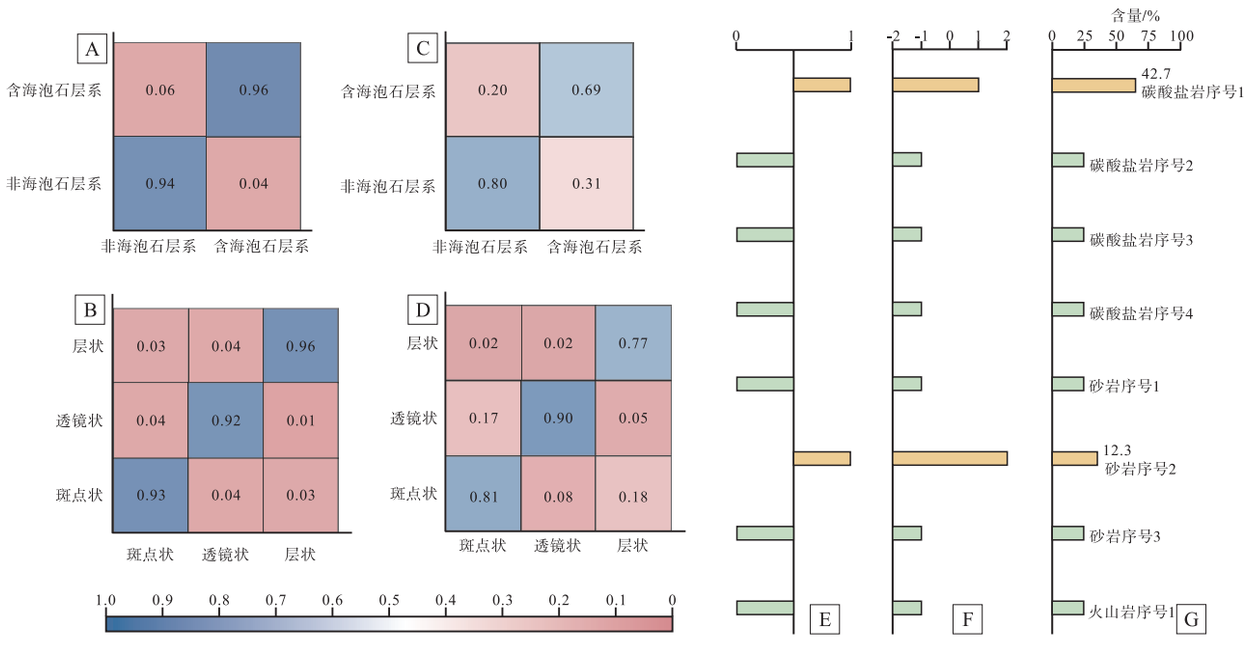
4.2 模型检验

模型泛化能力检验对于模型预测结果的可靠性至关重要。检验集预留数据测试显示（图 7），CatBoost 在含海泡石层系存在性二分类问题中有较好表现，对非海泡石层系预测准确率为 94%，对含海泡石预测准确率为 96%（图 7-A）。XGBoost 对非海泡石层系预测准确率为 80%，对含海泡石层系预测准确率为 69%（图 7-B）。综合对比显示 CatBoost

表 3 不同机器学习表现结果对比

Table 3 Comparison of performance results of different machine learning methods

模型类型	二分类判定滑石存在性问题				滑石产状多分类问题				滑石含量回归问题	
	训练误差	验证误差	训练准确度	验证准确度	训练误差	验证误差	训练准确度	验证准确度	训练误差	验证误差
CatBoost	0.01	0.03	0.98	0.97	0.01	0.02	0.96	0.92	0.25	0.24
XGBoost	0.01	0.03	0.98	0.97	0.07	0.11	0.92	0.91	0.07	0.05
LightGBM	0.01	0.02	0.98	0.98	0.45	0.48	0.74	0.68	0.29	0.27
随机森林	0.18	0.18	0.54	0.54	0.84	0.85	0.32	0.31	0.23	0.21
ANN	0.21	0.14	0.92	0.95	1.01	0.97	0.41	0.41	0.64	0.44
SVM	/	0.02	/	0.95	/	0.60	/	0.62	/	0.38
KNN	/	0.20	/	0.65	/	0.87	/	0.11	/	1.27
贝叶斯	/	0.73	/	0.25	/	0.92	/	0.08	/	/



A—CatBoost 二分类混淆矩阵；B—CatBoost 多分类混淆矩阵；C—XGBoost 二分类混淆矩阵；D—XGBoost 多分类混淆矩阵；
E—二分类预留数据泛化能力检验；F—多分类数据泛化能力检验；G—滑石含量数据泛化能力检验

图 7　检验集预留数据检验结果图（混淆矩阵中横坐标为实际属性，纵坐标为预测属性）

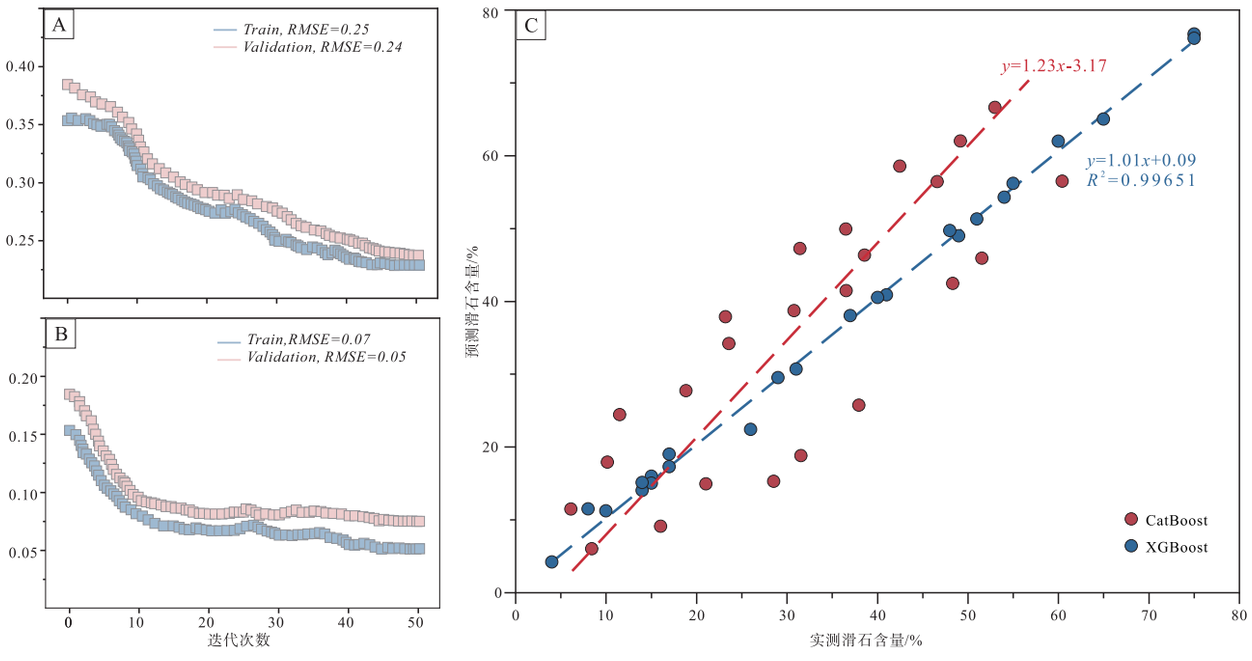
Fig. 7　Validation results with reserved test set data (the abscissa represents actual attributes, and the ordinate represents predicted attributes in the confusion matrix)

模型在二分类问题中具有更强的泛化能力。多分类任务中，CatBoost 模型对斑点状滑石、透镜状滑石、层状滑石预测准确率分别为 93%、92%、96%（图 7-C）。XGBoost 的预测准确率分别为 81%、90%、77%（图 7-D）。因此，CatBoost 在多分类问题中具有更高的可靠性。

将模型引入检验集测试，通过模型对人工加入噪声的分类、回归表现，进一步评价模型泛化能力。预留噪声数据设置碳酸盐岩（n = 80）、砂岩（n = 60）、火山岩（n = 20）均不含滑石。预留噪声二分类结果设定为 0（不存在）或 1（存在），多分类设置 -1 以指示异常或无关类别。碳酸盐岩（不含滑石）序列中，序号 1 被误判为含海泡石层系；砂岩（不含滑石）序列中，序号 2 被误判为含海泡石层系列。图 7-E 至图 7-G 表明，独立训练的模型结果的相关关系较强，二分类预测出含海泡石层系的样本在滑石产状预测和含量预测中均有明确的对应关系，未出现二分类判断为非含海泡石层系的样本判定出滑石产状及滑石含量的情况。

RMSE（Root Mean Squared Error，均方根误差）是一种常用的衡量回归模型预测误差的指标，尤其适用于连续数值预测任务。RMSE 值越低表示模型预测与实际值之间差异较小，模型表现好。迭代次数为 50 时，2 个模型的损失函数值变化曲线均趋于稳定，CatBoost 的训练误差与验证误差分别为 0.25 与 0.24（图 8-A），XGBoost 的训练误差与验证误差分别为 0.07 与 0.05（图 8-B），表明 XGBoost 模型在回归任务中拟合效果更好。模型在回归问题上的泛化能力检验效果如图 8-C，R^2 是决定系数，表示曲线对散点的拟合能力，用于衡量模型对数据的拟合程度。斜率（Slope）反映每个自变量对目标变量的具体贡献决定着预测和实测的一致性。XGBoost 模型的 R^2 为 0.99651，K 为 1.01，模型泛化能力较强，实测滑石含量与预测滑石含量基本一致，表明 XGBoost 模型在滑石的含量预测上具有较强的可靠性。

综合岩石显微图像资料、CIFLog 软件计算结果与 XRD 实验滑石含量结果，与模型输出的海泡石层系存在性—滑石产状分类—滑石含量预测结果对比，计算 12 口验证井预测结果准确率。以岩石学分析为依据，滑石存在性二分类验证准确率平均为 87.4%，产状多分类准确率平均为 86.16%。以



A—CatBoost 回归问题 RMSE 曲线；B—CatBoost 回归问题 RMSE 曲线；C—回归问题泛化能力检验

图 8 各回归模型在滑石含量预测中的泛化能力对比

Fig. 8 Comparison of generalization capabilities of various regression models in predicting talc content

表 4 验证井对比准确率

Table 4 Verification well comparison accuracy

序号	井代号	岩石学分析验证/%		滑石含量 XRD 结果验证/%	总准确率/%
		滑石存在性二分类	滑石产状多分类		
1	YB7	91.24	90.54	89.41	90.1
2	HB11	78.56	77.98	81.72	79.4
3	LX2	84.17	82.34	76.56	81.0
4	QL16	98.62	97.53	92.14	96.1
5	W118	79.43	78.21	84.99	80.9
6	JT1	87.91	85.67	80.37	84.7
7	CT1	92.08	91.56	93.65	92.4
8	LT1	76.34	75.64	77.89	76.3
9	NC1	93.77	91.12	88.23	91.0
10	LS1	81.56	80.43	79.12	80.0
11	W117	95.43	94.21	91.38	93.0
12	FS1	89.72	88.65	78.77	85.4

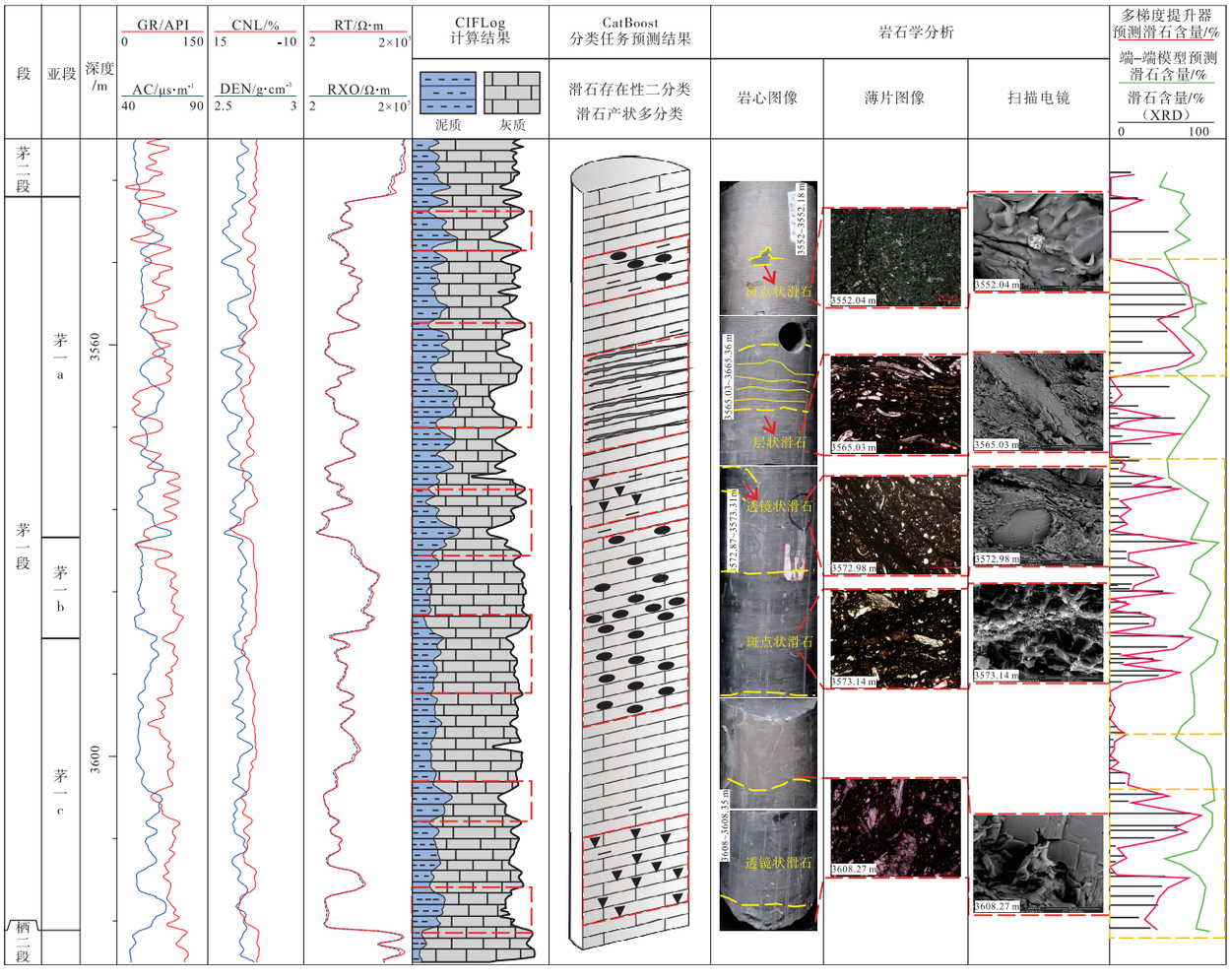
XRD 实验结果验证滑石含量回归结果准确率平均为 84.52%。预测的总准确率范围 79.4%～96.1%，平均值为 81.6%（表 4）。

CIFLog 软件可计算示测井数据约束下的矿物含量，以泥质、灰质单因素含量侧面预测含海泡石层系。对比端—端模型直接预测滑石含量（图 9绿色折线），模型能更准确对海泡石层系及滑石产状进行二分类判别及多分类预测，并捕捉滑石含量的变化趋势。HB1 井茅一 c 亚段中下段发育透镜状滑石，且滑石含量较高，3604～3610 m 井段滑石含量大于 65%，3615～3620 m 井段滑石含量分布较为均匀，平均含量为 60%。茅一 b 亚段整体靠近茅一 c 亚段及茅一 a 亚段滑石含量较高，下部靠近茅一 c 亚段发育斑点状滑石，滑石含量为 81%；上部靠近茅一 a 亚段发育透镜状，滑石含量为 85%。茅一 a 亚段中段发育层状滑石，滑石含量预测结果较高，3550～3563 m 井段的滑石含量平均 73%；上段发育斑点状滑石，滑石含量为 31%（图 9）。滑石含量预测值与 XRD 实验结果值吻合率保持在 89%。

5 海泡石层系预测结果

以四川盆地 254 口海泡石层系归属尚未明确的钻井为研究对象，将其测井数据导入多机器学习模型方案进行预测。基于四川盆地茅一段井资料，对茅一段含海泡石层系共 468 口井有效厚度进行计算，计算方法为：

$$\text{有效厚度} = \sum_{i=1}^{n} \text{单层厚度} \times \text{滑石厚度} \quad (12)$$



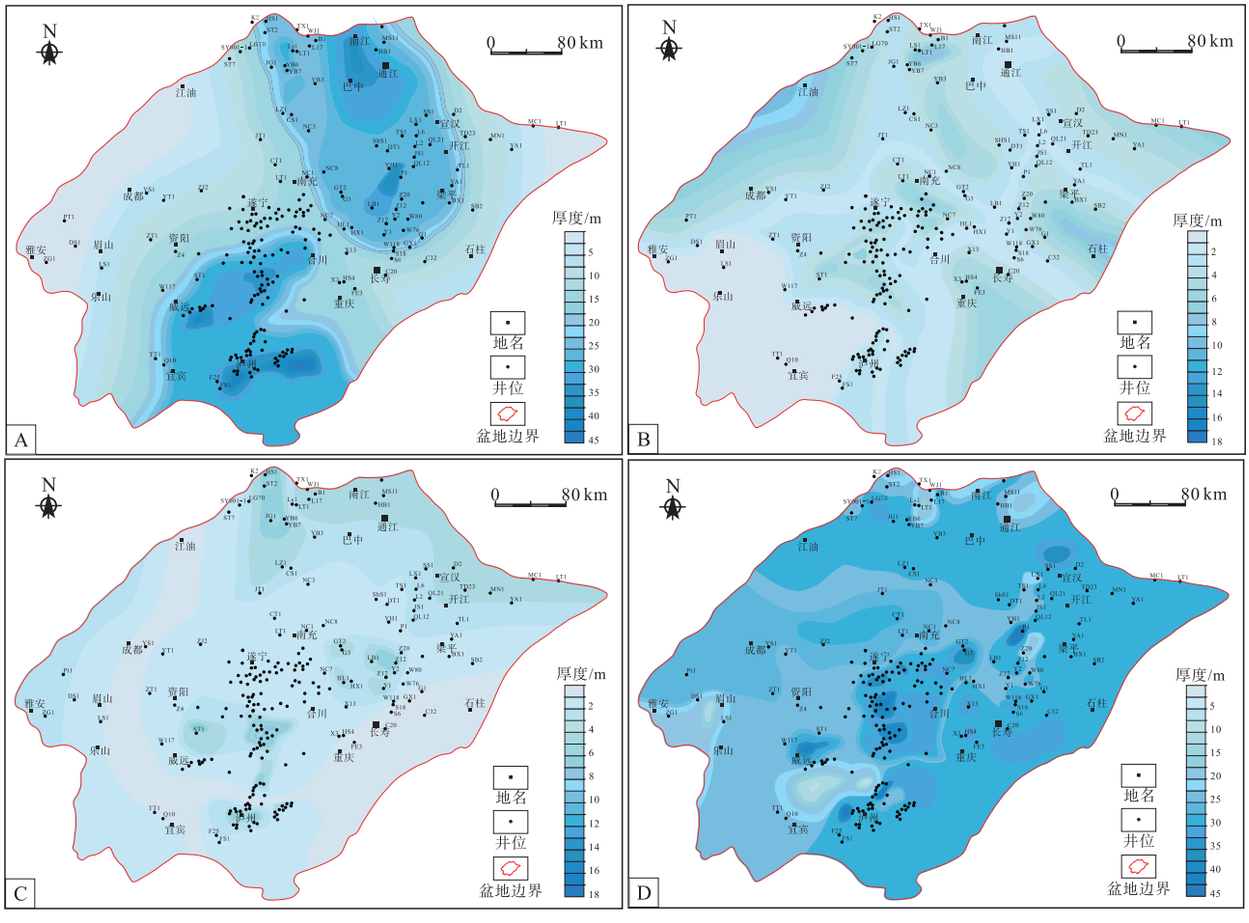
注：CatBoost 分类任务预测结果道红色虚线框表示预测滑石段。滑石含量段黄色虚线框表示滑石含量预测结果高含量段

图 9　HB1 井茅一段滑石含量预测单井验证图（滑石产状图例同图 2）

Fig. 9　Single-well verification chart of sepiolite content prediction in the Member 1 of Maokou Formation of Well HB1

斑点状滑石发育于川中、蜀南及川北地区，平均厚度为 15~18 m，通江—长寿、合川—威远及宜宾—泸州地区发育较厚，最高达 19 m，整体呈向北东方向含量升高的趋势（图 10-A）。透镜状滑石主要分布在川西地区，川西、川中、川北及川东北地区，呈块状分布，厚度为 5~10 m，南充—遂宁、长寿—重庆块状分布特征明显，厚度平均 8 m（图 10-B）。层状滑石整体较薄，厚度普遍小于 5 m，主要发育在川西北、川中和蜀南地区，南充—梁平—长寿地区、合川—资阳—威远地区及泸州地区也较厚，平均厚度为 8 m（图 10-C）。

四川盆地茅一段海泡石层系发育具有北东厚、南西薄的特征（图 10-D），同茅一段烃源岩分布范围一致（江青春等，2021；宋金民等，2024a）。含海泡石层系沉积中心位于通江—长寿凹陷地区，与“两台一凹”的沉积格局一致（宋金民等，2024a）。川中—蜀南地区茅一段含海泡石层系整体厚度大，斑点状、透镜状、层状纵向叠置，海泡石总厚度范围为 20~40 m；川西北剑阁地区和川东重庆—开县地区也发育 3 类滑石，厚度范围分别为 15~30 m 及 20~30 m，与茅一段生烃中心的范围一致（江青春等，2021；宋金民等，2024a），证明茅一段含海泡石层系与烃源岩展布存在耦合性。含海泡石（滑石）层系可形成两大类油气藏：凹陷边缘发育的旁（下）生—侧（上）储型层状白云岩气藏及凹陷内部发育的自生自储型泥质灰岩非常规气藏（刘树根等，2022；宋金民等，2025）。含海泡石（滑石）层系的分布不仅是预测优质储集层的关键，也是揭示茅一段油气成



A—斑点状滑石累计厚度图；B—透镜状滑石累计厚度图；C—层状滑石累计厚度图；D—四川盆地含海泡石层系累计厚度图

图 10 四川盆地茅一段不同产状的含海泡石层系厚度等值线

Fig. 10 Isopach of sepiolite-containing strata with different occurrences of the Member 1 of Maokou Formation in Sichuan Basin

藏规律的重要地质依据。

同时预测结果细致刻画了四川盆地茅一段含海泡石层系分布，其主要发育在川中、川北地区，平均厚度为 27~30 m，通江—长寿凹陷地区及合川—威远—泸州地区有效厚度可达 30~45 m，为海泡石层系优势发育区。进一步揭示了茅一段含海泡石（滑石）层系的空间展布特征，为有利区带优选与油气勘探部署提供了重要参考。

6 结论

1）四川盆地茅一段含海泡石层系主要由泥质灰岩、生屑泥晶灰岩和泥岩组成。含海泡石层系内的滑石是海泡石转化而成，以滑石的累计厚度表征茅一段含海泡石层系。含海泡石层系中滑石含量与自然伽马（GR）、电阻率（RT）、冲洗带电阻率（RXO）、声波时差（AC）与补偿中子（CNL）显著正相关，基于此的判别标准能够有效识别含海泡石层系。

2）综合多种算法对比显示，CatBoost 在含海泡石层系存在性二分类问题中对非海泡石层系预测准确率为 94%，对含海泡石预测准确率为 96%，在多分类任务中对斑点状滑石、透镜状滑石、层状滑石预测准确率分别为 93%、92%、96%；XGBoost 滑石含量预测结果保持 89%的准确率。选取 CatBoost 算法以滑石作为预测分析载体对是否含海泡石作二分类判断及含海泡石层系产状的多分类判断；选取 XGBoost 算法对含海泡石层系内的滑石含量进行回归预测，建立基于多梯度提升器算法的茅一段含海泡石层系分布预测流程。

3）茅一段含海泡石层系主要呈现斑点状、透镜状和层状 3 种产状。斑点状滑石具有川北、川中—

蜀南两大发育中心；透镜状滑石则主要分布在川西地区；层状滑石见于川西北、川中和蜀南地区。四川盆地茅一段含海泡石层系主要发育在川中、川北地区，平均厚度为 27~30 m，通江—长寿凹陷地区及合川—威远—泸州地区有效厚度可达 30~45 m，为海泡石层系优势发育区。

参考文献(References)

白烨，曹乃文，邱庆良. 2021. 极限学习机在黏土矿物分析中的应用. 地质与资源，30(4)：505-511. [Bai Y，Cao N W，Qiu Q L. 2021. Application of extreme learning machine in analysis of clay minerals. Geology and Resources，30(4)：505-511]

范建平，宋金民，江青春，刘树根，叶玥豪，黄士鹏，王佳蕊，苏旺，李立基，金鑫，冯宇翔. 2022. 川东地区中二叠统茅口组一段储层特征与形成模式. 石油与天然气地质，43(6)：1413-1430. [Fan J P，Song J M，Jiang Q C，Liu S G，Ye Y H，Huang S P，Wang J R，Su W，Li L J，Jin X，Feng Y X. 2022. Reservoir characteristics and development model of the Middle Permian Mao-1 Member in eastern Sichuan Basin. Oil & Gas Geology，43(6)：1413-1430]

范建平，宋金民，刘树根，江青春，李智武，杨迪，金鑫，苏旺，叶玥豪，黄士鹏，王佳蕊，姜华，罗平. 2023. 四川盆地中二叠统茅一段灰岩—泥质灰岩韵律层古温度演化及驱动机制. 石油实验地质，45(4)：726-738. [Fan J P，Song J M，Liu S G，Jiang Q C，Li Z W，Yang D，Jin X，Su W，Ye Y H，Huang S P，Wang J R，Jiang H，Luo P. 2023. Paleotemperature evolution and its driving mechanism during the formation of limestone-marl alternations in the first Member of Middle Permian Maokou Formation in Sichuan Basin. Petroleum Geology & Experiment，45(4)：726-738]

胡东风，王良军，张汉荣，段金宝，夏文谦，刘珠江，魏全超，王昆，潘磊. 2020. 碳酸盐岩烃源岩气藏的发现及其油气地质意义：以四川盆地涪陵地区中二叠统茅口组一段气藏为例. 天然气工业，40(7)：23-33. [Hu D F，Wang L J，Zhang H R，Duan J B，Xia W Q，Liu Z J，Wei Q C，Wang K，Pan L. 2020. Discovery of carbonate source rock gas reservoir and its petroleum geological implications：a case study of the gas reservoir in the first Member of Middle Permian Maokou Formation in the Fuling area，Sichuan Basin. Natural Gas Industry，40(7)：23-33]

江青春，汪泽成，苏旺，黄士鹏，曾富英，冯周，毕赫，鲁卫华. 2021. 四川盆地中二叠统茅口组一段泥灰岩源内非常规天然气成藏条件及有利勘探方向. 中国石油勘探，26(6)：82-97. [Jiang Q C，Wang Z C，Su W，Huang S P，Zeng F Y，Feng Z，Bi H，Lu W H. 2021. Accumulation conditions and favorable exploration orientation of unconventional natural gas in the marl source rock of the first Member of the Middle Permian Maokou Formation，Sichuan Basin. China Petroleum Exploration，26(6)：82-97]

蒋航，张玺华，陈聪，高兆龙，山述娇，罗平，陈延贵，李天军，胡罗嘉. 2023. 川北地区茅口中晚期斜坡相沉积特征及其油气地质意义. 海相油气地质，28(4)：381-390. [Jiang H，Zhang X H，Chen C，Gao Z L，Shan S J，Luo P，Chen T J，Li T J，Hu L J. 2023. Sedimentary characteristics of slope facies in mid-late stage of Maokou Formation in northern Sichuan Basin and its petroleum geological significances. Marine Origin Petroleum Geology，28(4)：381-390]

蒋文博，丁晓军，郭华黡，许姗姗，黄斌，李信. 2018. 黏土矿物预测方程的建立及在东营油田的应用. 石油地质与工程，32(3)：95-97，125. [Jiang W B，Ding X J，Guo H N，Xu S S，Huang B，Li X. 2018. Establishment of prediction equations for clay minerals and their application in the Dongying Oilfield. Petroleum Geology and Engineering，32(3)：95-97，125]

梁兴，徐政语，栗维民，马立桥，计玉冰，罗瑀峰，丁邦春. 2023. 蜀南—渝西地区中二叠统茅一段灰质源岩气储层特征及主控因素：以 DB1 井为例. 石油实验地质，45(4)：714-725. [Liang X，Xu Z Y，Li W M，Ma L Q，Ji Y B，Luo Y F，Ding B C. 2023. Characteristics and main controlling factors of the limy source rock gas reservoir in the first Member of the Middle Permian Maokou Formation in the southern Sichuan and western Chongqing area：a case study of well DB 1. Petroleum Geology & Experiment，45(4)：714-725]

梁兴，张磊，蒋立伟，单长安，潘峰，罗瑀峰，孟阳，王高成，舒红林，朱斗星，李博硕，汪振东，姚秋昌，徐政语. 2024. 四川盆地渝西大安地区灰质源岩气勘探突破与资源潜力. 中国石油勘探，29(2)：30-45. [Liang X，Zhang L，Jiang L W，Shan C A，Pan F，Luo Y F，Meng Y，Wang G C，Shu H L，Zhu D X，Li B S，Wang Z D，Yao Q C，Xu Z Y. 2024. Exploration breakthrough and resource potential of limy source rock gas in Da'an area in the western Chongqing，Sichuan Basin. China Petroleum Exploration，29(2)：30-45]

刘昇，范存辉，张本健，张亚，王尉，罗冰，白晓亮. 2023. 四川盆地东部中二叠统茅口组孤峰段展布特征及其油气地质意义. 石油与天然气地质，44(4)：993-1008. [Liu S，Fan C H，Zhang B J，Zhang Y，Wang W，Luo B，Bai X L. 2023. Distribution characteristics of Gufeng Member of the Middle Permian Maokou Formation，eastern Sichuan Basin and its petrogeological significance. Oil & Gas Geology，44(4)：993-1008]

刘树根，文龙，宋金民，孙玮，汪华，金鑫，郭海洋，邓宾，江青春，李智武，丁一，叶玥豪，王瀚，范建平，杨迪，李柯然，田小彬，罗平. 2022. 四川盆地中二叠统构造-沉积分异与油气勘探. 成都理工大学学报(自然科学版)，49(4)：385-413. [Liu S G，Wen L，Song J M，Sun W，Wang H，Jin X，Guo H Y，Deng B，Jiang Q C，Li Z W，Ding Y，Ye Y H，Wang H，Fan J P，Yang D，Li K R，Tian X B，Luo P. 2022. Sedimentary topography and tectonic differentiation on the Middle Permian platform and hydrocarbon exploration in Sichuan Basin，SW China. Journal of Chengdu University of Technology(Science & Technology Edition)，49(4)：385-413]

宋金民，刘树根，金鑫，汪华，范建平，江青春，田小彬，叶玥豪，李亚，李智武，杨迪，王佳蕊，罗平. 2022. 四川盆地中二叠统含海泡石层系源储组合新模式及其油气勘探新领域. 成都理工大学学报(自然科学版)，49(2)：129-149. [Song J M，Liu S G，Jin X，Wang H，Fan J P，Jiang Q C，Tian X B，Ye Y H，Li Y，Li Z W，Yang

D, Wang J R, Luo P. 2022. A new model of source-reservoir association and new fields of oil and gas exploration in the Middle Permian sepiolite-bearing strata, Sichuan Basin, China. Journal of Chengdu University of Technology (Science & Technology Edition), 49(2): 129-149]

宋金民, 刘树根, 李智武, 张本健, 邓宾, 江青春, 金鑫, 孙玮, 汪华, 郭海洋, 陈骁, 王佳蕊, 田小彬, 张玺华, 叶玥豪, 杨迪, 范建平, 罗平. 2023. 四川盆地中二叠统油气成藏模式与有利勘探区分布. 天然气工业, 43(11): 54-71. [Song J M, Liu S G, Li Z W, Zhang B J, Deng B, Jiang Q C, Jin X, Sun W, Wang H, Guo H Y, Chen X, Wang J R, Tian X B, Zhang X H, Ye Y H, Yang D, Fan J P, Luo P. 2023. Accumulation model and favorable exploration area distribution of the Middle Permian oil and gas in the Sichuan Basin. Natural Gas Industry, 43(11): 54-71]

宋金民, 江青春, 刘树根, 金鑫, 范建平, 李智武, 黄士鹏, 苏旺, 杨迪, 姜华, 叶玥豪, 王佳蕊, 王俊轲, 任杉. 2024a. 四川盆地中二叠统茅口组一段含海泡石层系古环境与沉积格局指示意义. 石油学报, 45(6): 914-931. [Song J M, Jiang Q C, Liu S G, Jin X, Fan J P, Li Z W, Huang S P, Su W, Yang D, Jiang H, Ye Y H, Wang J L, Wang J K, Ren S. 2024a. Paleoenvironment and sedimentary significances of sepiolite-containing succession in the first Member of Middle Permian Maokou Formation, Sichuan Basin. Acta Petrolei Sinica, 45(6): 914-931]

宋金民, 王佳蕊, 刘树根, 李智武, 罗平, 江青春, 金鑫, 杨迪, 黄士鹏, 范建平, 叶玥豪, 王俊轲, 邓豪爽, 王斌, 郭嘉欣. 2024b. 含海泡石层系泥质灰岩中自生黏土矿物的类型、组成与成岩演化过程: 以川东地区中二叠统茅口组茅一段为例. 石油勘探与开发, 51(2): 311-322. [Song J M, Wang J R, Liu S G, Li Z W, Luo P, Jiang Q C, Jin X, Yang D, Huang S P, Fan J P, Ye Y H, Wang J K, Deng H S, Wang B, Guo J X. 2024b. Types, composition and diagenetic evolution of authigenic clay minerals in argillaceous limestone of sepiolite-bearing strata: a case study of Mao-1 Member of Middle Permian Maokou Formation, eastern Sichuan Basin, SW China. Petroleum Exploration and Development, 51(2): 311-322]

宋金民, 王俊轲, 刘树根, 文龙, 叶玥豪, 罗冰, 李智武, 张本健, 金鑫, 杨迪, 张玺华, 王佳蕊, 周刚, 郭嘉欣, 张钊益, 罗平. 2025. 含海泡石层系有机质赋存类型与富集模式: 以川东地区茅口组一段为例. 石油勘探与开发, 52(5): 1054-1067. [Song J M, Wang J K, Liu S G, Wen L, Ye Y H, Luo B, Li Z W, Zhang B J, Jin X, Yang D, Zhang X H, Wang J R, Zhou G, Guo J X, Zhang Z Y, Luo P. 2025. Occurrence types and enrichment model of organic matter in the sepiolite-containing successions: a case study of the Mao-1 Member of Mid-Permian Maokou Formation, Eastern Sichuan Basin, SW China. Petroleum Exploration and Development, 52(5): 1054-1067]

王斌, 宋金民, 刘树根, 金鑫, 李智武, 叶玥豪, 杨迪, 王佳蕊, 王俊轲, 邓豪爽, 郭嘉欣, 罗平. 2024. 川西地区中二叠统茅口组二段白云岩类型与成因模式. 石油学报, 45(12): 1743-1760. [Wang B, Song J M, Liu S G, Jin X, Li Z W, Ye Y H, Yang D, Wang J R, Wang J K, Deng H S, Guo J X, Luo P. 2024. Types and genetic modes of dolomite in the Member 2 of Middle Permian Maokou Formation in western Sichuan Basin. Acta Petrolei Sinica, 45(12): 1743-1760]

王心乾, 余文端, 马晓东, 周韬, 郜浩, 崔钦宇, 邓空, 陆永潮, 刘占红. 2024. 基于常规测井曲线的页岩岩相识别与应用: 以苏北盆地溱潼凹陷阜宁组二段为例. 油气藏评价与开发, 14(5): 699-706, 713. [Wang X Q, Yu W D, Ma X D, Zhou T, Tai H, Cui Q Y, Deng K, Lu Y C, Liu Z H. 2024. Identification and application of shale lithofacies based on conventional logging curves: a case study of the second Member of Funing Formation in Qintong sag, Subei Basin. Petroleum Reservoir Evaluation and Development, 14(5): 699-706, 713]

夏文谦, 刘瑾, 俞凌杰, 张汉荣, 汪建辉, 魏富彬, 章思遥. 2023. 川东南涪陵地区二叠系茅口组一段泥质灰岩储层特征及主控因素. 矿物岩石, 43(4): 100-108. [Xia W Q, Liu J, Yu L J, Zhang H R, Wang J H, Wei F B, Zhang S Y. 2023. Characteristics and main controlling factors of argillaceous limestone reservoir in the first Member of Permian Maokou Formation in Fuling area, southeast Sichuan Basin. Mineralogy and Petrology, 43(4): 100-108]

杨跃明, 杨雨, 文龙, 张玺华, 陈聪, 陈康, 张亚, 狄贵东, 汪华, 谢忱. 2020. 四川盆地中二叠统天然气勘探新进展与前景展望. 天然气工业, 40(7): 10-22. [Yang Y M, Yang Y, Wen L, Zhang X H, Chen C, Chen K, Zhang Y, Di G D, Wang H, Xie C. 2020. New exploration progress and prospect of Middle Permian natural gas in the Sichuan Basin. Natural Gas Industry, 40(7): 10-22]

殷树军, 文政, 金雪英, 闫伟林, 王春燕, 师江波, 夏文豪, 李金奉, 聂妮, 郑盛敦. 2023. 泥灰岩非常规储层特征及测井表征方法: 以四川盆地合川—潼南地区茅口组一段为例. 石油学报, 44(7): 1105-1117. [Yin S J, Wen Z, Jin X Y, Yan W L, Wang C Y, Shi J B, Xia W H, Li J F, Nie N, Zheng S D. 2023. Characteristics and logging characterization method of unconventional marl reservoir: a case study of the Member 1 of Maokou Formation in Hechuan-Tongnan area, Sichuan Basin. Acta Petrolei Sinica, 44(7): 1105-1117]

翟常博, 林良彪, 尤东华, 刘冯斌, 刘思雨. 2024. 川西南地区中二叠统茅口组一段沉积微相特征及有机质富集模式. 石油与天然气地质, 45(2): 440-456. [Zhai C B, Lin L B, You D H, Liu F B, Liu S Y. 2024. Sedimentary microfacies characteristics and organic matter enrichment pattern of the first Member of the Middle Permian Maokou Formation, southwestern Sichuan Basin. Oil & Gas Geology, 45(2): 440-456]

张培先, 何希鹏, 高全芳, 高玉巧, 孙斌, 蔡潇, 何贵松, 张志萍, 刘娜娜. 2021. 四川盆地东南缘二叠系茅口组一段页岩气藏地质特征及富集模式. 石油与天然气地质, 42(1): 146-157. [Zhang P X, He X P, Gao Q F, Gao Y Q, Sun B, Cai X, He G S, Zhang Z P, Liu N N. 2021. Geological characteristics and enrichment pattern of Permian Mao-1 Member shale gas reservoirs at the southeastern margin of Sichuan Basin. Oil & Gas Geology, 42(1): 146-157]

张宇, 曹清古, 罗开平, 李龙龙, 刘金连. 2022. 四川盆地二叠系茅口组油气藏勘探发现与启示. 石油与天然气地质, 43(3): 610-620. [Zhang Y, Cao Q G, Luo K P, Li L L, Liu J L. 2022. Reservoir

exploration of the Permian Maokou Formation in the Sichuan Basin and enlightenment obtained. Oil & Gas Geology,43(3): 610-620]

Cao K, Wei C, Gaidon A, Arechiga N, Ma, T Y. 2019. Learning imbalanced datasets with label-distribution-aware margin loss. Advances in Neural Information Processing Systems, 32(NeurIPS 2019).

Chawla N V, Bowter K W, Hall L O, Kegelmeyer W P. 2002. SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research,16: 321-357.

Chen T, Guestrin C. 2016. Xgboost: a scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: 785-794.

Dorogush A V, Ershov V, Gulin A. 2018. The CatBoost algorithm. Microsoft Research Technical Report.

Friedman J H. 1999. Greedy function approximation: a gradient boosting machine. Annals of Statistics,29(5): 1189-1232.

Hastie T, Tibshirani R, Friedman J. 2009. The Elements of Statistical Learning(2nd ed). New York: Springer.

Hyndman R J, Koehler A B. 2006. Another look at measures of forecast accuracy. International Journal of Forecasting,22(4): 679-688.

Jin H, Song Q, Hu X. 2019. Auto-Keras: an efficient neural architecture search system. 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining: 1946-1956.

Koray A-M, Bui D, Kubi E, Ampomah W, Amosu A. 2024. Machine learning based reservoir characterization and numerical modeling from integrated well log and core data. Geoenergy Science & Engineering,243: 213296.

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay Edouard. 2011. Scikit-learn: machine learning in Python. Journal of Machine Learning Research,12: 2825-2830.

Prokhorenkova L, Gusev G, Vorobev A, Dorogush A V, Gulin A. 2018. CatBoost: unbiased boosting with categorical features. Advances in Neural Information Processing Systems,31.

Wang Z H, Cai Y D, Liu D M, Lu J, Qiu F, Sun F R, Hu J H, Li Z T. 2024. Characterization of natural fracture development in coal reservoirs using logging machine learning inversion, well test data and simulated geostress analyses. Engineering Geology: 107696.

（责任编辑　郑秀娟）