

三塘湖盆地条湖凹陷八道湾组煤焦油产率 及富油性预测*

武鸿飞¹, 李鑫^{1†}, 王兴刚², 焦立新², 曹志雄², 李斌², 韦波¹, 冯烁¹

(1. 新疆大学地质与矿业工程学院, 新疆乌鲁木齐 830017; 2. 中国石油吐哈油田分公司勘探开发研究院, 新疆哈密 839000)

摘要: 富油煤集煤、油、气属性为一体, 促进其勘探开发对保障我国油气资源供应、实现煤炭清洁高效利用具有重要战略价值。为此, 以三塘湖盆地条湖凹陷煤样焦油产率、工业组分、元素分析、煤岩分析数据为基础, 结合测井响应, 建立了富油煤焦油产率测井预测模型, 预测研究区八道湾组富油煤资源量。结果表明: 煤焦油产率与挥发分产率、氢元素含量、镜质组含量呈正相关关系, 与灰分产率、惰质组含量呈负相关关系; 挥发分产率、氢元素含量、镜质组含量与声波时差、补偿密度测井值呈较好的负相关关系, 与自然伽马测井值相关性较差。此外, 建立了基于机器学习的煤焦油产率预测模型, 预测值与实际值的相关系数为0.92, 90%煤样的焦油产率预测结果相对误差小于20%, 75%煤样的相对误差小于15%。

关键词: 富油煤; 焦油产率; 预测模型; 条湖凹陷

DOI: 10.13568/j.cnki.651094.651316.2025.02.08.0001

中图分类号: P618.11 **文献标识码:** A **文章编号:** 2096-7675(2025)03-0312-011

引文格式: 武鸿飞, 李鑫, 王兴刚, 焦立新, 曹志雄, 李斌, 韦波, 冯烁. 三塘湖盆地条湖凹陷八道湾组煤焦油产率及富油性预测[J]. 新疆大学学报(自然科学版中英文), 2025, 42(3): 312-322.

英文引文格式: WU Hongfei, LI Xin, WANG Xinggang, JIAO Lixin, CAO Zhixiong, LI Bin, WEI Bo, FENG Shuo. Prediction of coal tar yield and oil-richness in the Badaowan formation of Tiaohu sag, Santanghu Basin[J]. Journal of Xinjiang University(Natural Science Edition in Chinese and English), 2025, 42(3): 312-322.

Prediction of Coal Tar Yield and Oil-Richness in the Badaowan Formation of Tiaohu Sag, Santanghu Basin

WU Hongfei¹, LI Xin¹, WANG Xinggang², JIAO Lixin², CAO Zhixiong²,
LI Bin², WEI Bo¹, FENG Shuo¹

(1. School of Geology and Mining Engineering, Xinjiang University, Urumqi Xinjiang 830017, China; 2. Exploration and Development Research Institution, China Petroleum Tuha Oilfield Branch, Hami Xinjiang 839000, China)

Abstract: Tar-rich coal integrates the properties of coal, oil and gas. Promoting its exploration and development has important strategic value for ensuring the supply of oil and gas resources in China, and realizing the clean and efficient utilization of coal. Therefore, based on the data of tar yield, industrial components, elemental analysis and coal petrography analysis of coal samples in Tiaohu sag of Santanghu Basin, combined with logging response, a logging prediction model of tar-rich coal tar yield is established, and the tar-rich coal resources of Badaowan formation in the study area are predicted. The results show that: the yield of coal tar is positively correlated with volatile yield, hydrogen content and vitrinite content, and negatively correlated with ash yield and inertinite

* 收稿日期: 2025-02-08

基金项目: 新疆维吾尔自治区天山英才计划“新疆富油煤地质成因与薄层富油煤原位热解提取煤基油气基础研究”(2022TSYCLJ0021); 新疆维吾尔自治区科技重大专项“利用二氧化碳提高新疆难开采煤层气采收率技术体系研发”(2022A03015-3)。

作者简介: 武鸿飞(1999—), 男, 硕士生, 从事非常规油气与勘探的研究, E-mail: 1498472737@qq.com.

† 通讯作者: 李鑫(1990—), 男, 博士, 教授, 主要从事非常规油气与勘探的研究, E-mail: lixinwaxj@xju.edu.cn.

content. There is a good negative correlation between volatile yield, hydrogen content, vitrinite content and acoustic time difference, compensated density logging values, and a poor correlation with natural gamma logging values. In addition, the prediction model of coal tar yield based on machine learning is established. The correlation coefficient between the predicted value and the actual value is 0.92. The relative error of tar yield prediction results of 90% coal samples is less than 20%, and the relative error of 75% coal samples is less than 15%.

Key words: tar-rich coal; tar yield; prediction model; Tiaohu sag

0 引言

我国能源结构具有“缺油、少气、相对富煤”的资源禀赋特征^[1-3],煤作为国家能源的主体,其地位短期内不会改变^[4-5].推动煤炭安全绿色开采、清洁高效低碳利用是能源生产和消费革命的要求,也是新时代煤炭工业重要的发展方向.富油煤具有油、气资源属性,可通过中低温(500~700℃)热解产出油气^[6].根据《矿产资源工业要求手册(2014年修订本)》,按照焦油产率($T_{ar,d}$)把煤分为:高油煤, $T_{ar,d}>12\%$;富油煤, $7\%<T_{ar,d}\leq 12\%$;含油煤, $T_{ar,d}\leq 7\%$.从资源评价角度,将焦油产率大于7%的煤统称为富油煤^[7-8].

关于富油煤的研究,目前多聚焦于热解、资源潜力、原位开发、高效提油、主控因素和赋存规律等方面.富油煤主要分布在新疆、内蒙古、云南等地的低中阶煤中^[8].同一成煤时期,煤焦油产率受物质组成、沉积环境、埋藏深度等因素控制,表现出一定差异^[9].焦油产率的预测在富油煤地质储量评价、富油煤高效清洁利用、富油煤资源管理规划等多个领域都具有重要的现实作用.煤焦油产率的预测主要有两种方法:一是基于地球物理测井与煤性质之间的相关性预测,发现富油煤具有“低密度、高中子和高电阻率”的常规测井响应特征,由此建立焦油产率预测模型^[8,10-11];二是基于煤质的预测方法,建立敏感性参数与焦油产率之间的回归预测模型^[10,12].焦油产率与煤中有机质类型和含量、煤中无机物质的组成、煤的变质程度密切相关,焦油产率与煤岩煤质指标呈现复杂的多元非线性关系^[8,13],导致直接使用测井数据构建焦油产率预测模型误差较大.以往研究在预测焦油产率方面主要面临数据限制、预测精度不高、适用范围有限、处理非线性和模糊关系的能力不足、高维度和高计算复杂度等挑战,导致以往的焦油产率预测方法存在预测精度较低、适用范围较窄的问题,同时以往煤质测试数据较少,现阶段补充测试困难.本文基于已有地质勘查钻孔开展的测井工作,建立“测井参数-煤岩煤质-焦油产率”之间的定量关系,并基于此形成焦油产率预测模型.

新疆煤炭资源极为丰富^[14],保有资源量为 4.10×10^{11} t,其中富油煤资源量高达 2.05×10^{11} t,占比50%^[15].焦油产率是衡量富油煤勘探开发价值的关键指标,也是实现富油煤“专煤专用”的核心控制因素^[12].为此,本文基于条湖凹陷重点井的煤岩煤质分析数据和测井资料,揭示煤焦油产率的测井响应特征,系统分析富油煤测井响应的敏感因素,利用随机森林算法,建立焦油产率的“煤岩煤质-测井响应”预测模型,旨在为新疆煤基油气资源评价提供参考.

1 地质概况

条湖凹陷位于新疆三塘湖盆地中央凹陷带中部,夹峙于南北两个逆冲推覆体之间(图1).整体为一单斜构造,东部倾角较平缓,一般介于 $5^\circ\sim 15^\circ$;中部倾角较陡,一般介于 $10^\circ\sim 25^\circ$;东部倾角一般介于 $10^\circ\sim 20^\circ$;局部受断层牵引作用,地层倾角较大.

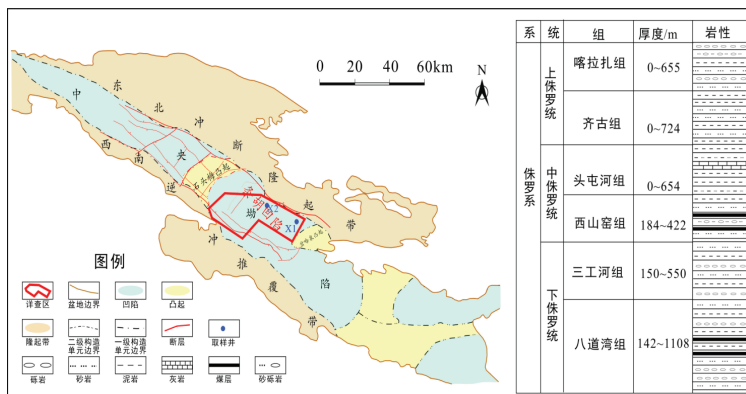


图 1 三塘湖盆地构造格局及条湖凹陷侏罗系综合地层柱状简图

条湖凹陷区含煤地层包括中侏罗统西山窑组和下侏罗统八道湾组,其中:西山窑组含煤1~22层,八道湾组含煤1~12层.西山窑组煤层以多层薄煤和单层厚煤为主,累计厚度在14~38.5 m.八道湾组煤层主要分布在研究区北部、东部,煤层以多层薄煤和单层巨厚煤为主,单层厚度在1~31 m,累计厚度在4~34 m.主力煤层为1号煤层,结构简单,平均厚度超过20 m,最大厚度为31 m.

条湖凹陷八道湾组主力煤层煤样的焦油产率介于2.12%~23.70%,平均值为11.48%,总体为富油煤;镜质组含量介于30%~98%,惰质组含量介于8%~62%,壳质组含量介于0.4%~20%,整体而言煤岩显微组分以镜质组为主,其次为惰质组,壳质组含量最低.镜质组以基质镜质体和碎屑镜质体为主,惰质组以氧化丝质体为主.煤的镜质组随机反射率介于0.47%~0.77%,煤化程度相对较低.总体以高挥发分、低灰分煤层为主.

2 煤样及其测试结果

依据《煤层煤样采取方法(GB/T 482—2008)》,在条湖凹陷北缘八道湾组主力煤层1号煤层采集40件煤样,自上而下系统采集X1井779~817 m深度区间的22件煤样与X2井634~662 m深度区间的18件煤样.为防止氧化,将采集的样品及时用保鲜膜密封,然后进行压碎、筛分,并对样品进行细分,开展煤焦油产率、煤工业分析、煤元素分析、煤岩显微组分与成熟度4个方面煤岩煤质测试分析.收集取样段处伽马测井、密度测井、声波测井值数据,利用煤岩显著测井响应特征进行岩心深度归位,取单件样品段测井值的平均值作为单件样品的测井值.

2.1 煤焦油产率

依据《煤的格金低温干馏试验方法(GB/T 1341—2007)》开展煤样的格金低温干馏试验分析.结果显示,煤焦油产率为3.20%~23.70%,平均值为13.48%,其中:9件煤样焦油产率小于7%,为含油煤;6件煤样焦油产率大于7%、小于12%,为富油煤;25件煤样焦油产率大于12%,为高油煤(表1).

2.2 煤工业分析

依据《煤的工业分析方法 仪器法(GB/T 30732—2014)》开展煤样的工业分析.结果显示,煤样水分含量为3.96%~11.34%,平均值为6.10%;灰分产率为1.53%~23.03%,平均值为5.20%,主要为特低灰煤;挥发分产率为28.63%~57.45%,平均值为47.06%,以高挥发分煤为主(表1).

表 1 煤工业分析、低温干馏焦油产率部分数据

样品编号	工业分析/%			低温干馏试验/%			煤的焦油产率分级
	M_{ad}	A_d	V_{daf}	CR_d	$T_{ar,d}$	$Water_{ad}$	
X01	5.31	9.68	52.40	65.2	10.5	15.5	富油煤
X02	8.24	9.30	32.64	76.7	3.2	13.5	含油煤
X03	7.07	4.24	45.69	70.0	9.5	15.0	富油煤
X04	4.08	17.71	55.85	54.4	21.2	11.5	高油煤
X05	5.75	4.98	49.34	62.5	14.1	14.5	高油煤
X06	8.59	3.47	28.68	69.4	3.9	15.0	含油煤
X07	5.31	2.55	52.99	56.7	19.6	13.3	高油煤
X08	5.56	1.91	49.46	59.9	17.7	13.7	高油煤
X09	7.84	10.85	34.06	66.4	5.5	16.3	含油煤
X10	5.82	1.60	49.25	60.7	17.7	11.5	高油煤
X11	5.02	1.53	53.35	73.0	9.7	8.3	富油煤
X12	5.89	4.08	50.22	56.3	20.4	12.3	高油煤
X13	5.34	1.92	49.97	56.8	19.9	11.3	高油煤
X14	6.24	1.94	49.27	66.4	10.5	12.2	富油煤
X15	5.85	7.58	48.53	64.6	13.0	12.2	高油煤

2.3 煤元素分析

依据《煤的元素分析(GB/T 31391—2015)》开展煤样元素分析.结果显示,煤样氧和硫含量为13.63%~24.64%,平均值为16.72%;碳含量为69.69%~82.14%,平均值为76.62%;氢含量为3.04%~6.46%,平均值为5.36%;氮含量为0.84%~1.25%,平均值为1.01%(表2).

2.4 煤岩显微组分与成熟度

依据《煤岩分析样品制备方法(GB/T 16773—2008)》与《煤的显微组分组和矿物测定方法(GB/T 8899—2013)》,制备煤砖光片,进行显微煤岩组分鉴定.鉴定结果以去矿物基的形式表示,结果显示,煤样镜质组含量为60.90%~97.92%,平均值为85.19%;惰质组含量为0.42%~38.40%,平均值为8.25%;壳质组含量为0.70%~18.63%,平均值为6.56%(表2).

2.5 煤层地球物理测井

通过岩心深度归位,取样段处声波时差测井值为406.81~446.70 $\mu\text{s}/\text{m}$,平均值为426.12 $\mu\text{s}/\text{m}$;自然伽马测井值为3.65~16.95 API,平均值为8.76 API;补偿密度测井值为1.14~1.54 g/cm^3 ,平均值为1.26 g/cm^3 (表2).

表2 煤元素分析、显微组分、测井值部分数据

样品 编号	元素分析/%				显微组分/%			平均随机 反射率/%	声波时差 /($\mu\text{s}/\text{m}$)	自然伽马 /API	补偿密度 /(g/cm^3)
	$O_{\text{daf}} + S_{\text{daf}}$	C_{daf}	H_{daf}	N_{daf}	镜质组	惰质组	壳质组				
X01	17.39	76.12	5.24	1.25	85.17	5.93	8.90	0.48	426.35	3.66	1.26
X02	19.12	77.84	3.04	-	67.30	31.80	0.90	0.55	421.83	5.98	1.44
X03	17.11	75.93	5.66	1.24	90.11	0.52	9.37	0.45	406.81	14.94	1.28
X04	15.58	77.05	6.40	0.97	79.10	2.27	18.63	0.41	420.85	8.84	1.33
X05	17.53	75.99	5.52	0.95	94.02	0.92	5.06	0.40	419.48	8.41	1.27
X06	14.02	81.88	4.10	-	61.00	38.20	0.80	0.56	444.45	5.20	1.32
X07	16.14	76.93	5.97	0.96	85.78	0.92	13.30	0.40	425.27	9.27	1.25
X08	16.40	77.06	5.59	0.95	92.32	0.96	6.72	0.46	424.81	7.84	1.24
X09	19.17	73.79	5.88	1.16	67.30	31.80	0.90	0.33	433.90	9.94	1.26
X10	16.00	77.36	5.73	0.90	92.04	1.32	6.64	0.43	423.21	8.52	1.24
X11	24.64	69.69	4.81	0.84	94.50	0.84	4.66	0.45	428.50	9.10	1.22
X12	15.15	77.82	6.06	0.96	90.86	0.92	8.22	0.44	426.77	7.33	1.26
X13	14.95	77.95	6.19	0.90	92.14	1.75	6.11	0.42	427.44	8.38	1.28
X14	19.73	74.02	5.22	1.02	97.92	1.04	1.04	0.45	428.57	9.04	1.14
X15	18.92	74.75	5.36	0.96	87.56	1.00	11.44	0.44	421.05	11.12	1.25

3 煤焦油产率的测井响应影响因素分析

3.1 煤焦油产率的测井响应

采用单参数交会的方法,对区内40件煤样测试分析,结果显示(图2):在参数置信度为95%的情况下,煤焦油产率与声波时差、补偿密度测井值呈负相关关系,与自然伽马测井值相关性较低.如图3所示,煤焦油产率与灰分产率、惰质组含量、碳含量呈负相关关系,与挥发分产率、氢含量、镜质组含量呈正相关关系.由于煤焦油产率与挥发分产率、氢含量、镜质组含量、惰质组含量的相关性更明显,故这4个参数是可用于焦油产率预测建模的敏感性参数.由于镜质组含量与惰质组含量之间为负相关关系,测井响应互为包容,故选择其中1个参数作为敏感性参数即可满足建模需求.

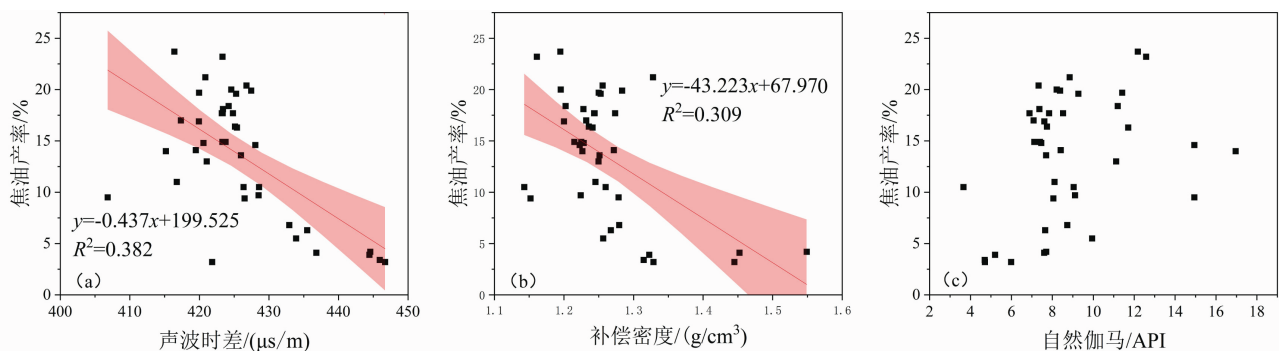


图2 煤样焦油产率与测井响应参数交会图

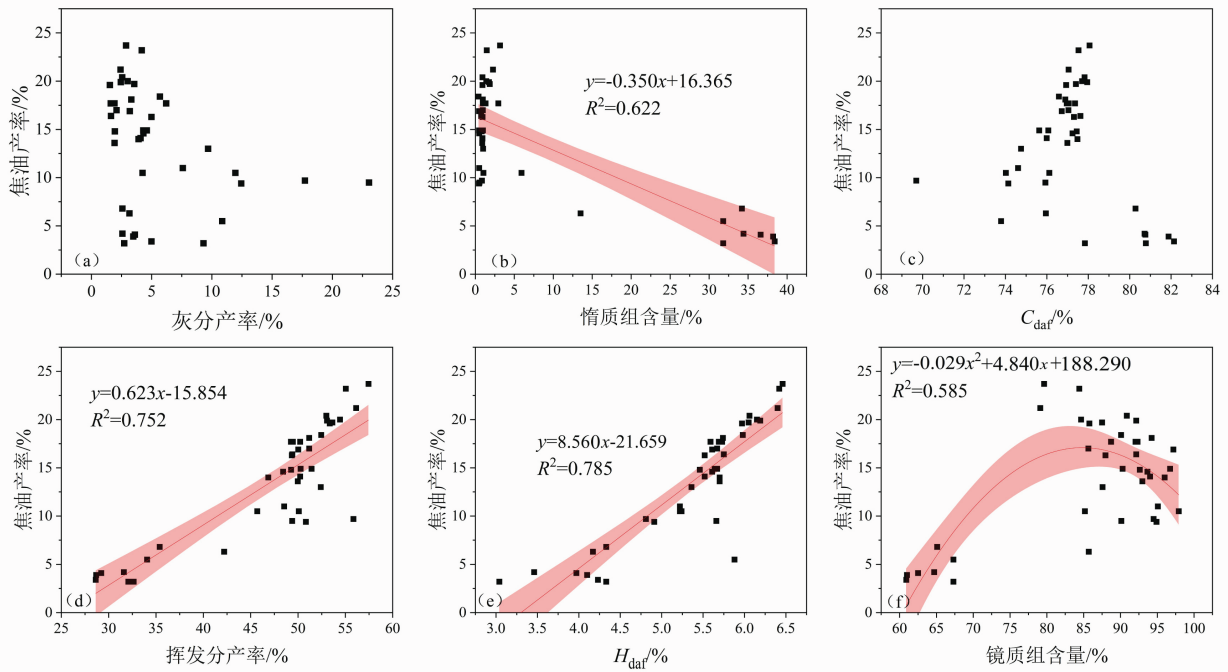


图3 煤样焦油产率与煤岩煤质测试结果相关性分析

3.2 富油煤测井响应的敏感因素分析

前人研究发现,成煤环境、成煤物质条件、煤化程度以及煤岩显微组分组成是控制富油煤发育的主要地质因素,表现为高H/C原子比、低灰分产率、高挥发分产率、高镜质组含量的特点^[6]。在覆水较深、还原程度高的滨湖、浅湖、三角洲及沼泽环境中,植物和浮游生物发育较好,煤的富油性偏高。在温暖且潮湿的还原环境中,成煤植物遗体主要通过凝胶化作用进行转化,使得煤中活性组分(壳质组和镜质组)含量增加,活性组分是煤热解产油的主要组分^[6,16-18]。

煤层测井响应呈“三高三低”的典型特征:电阻率、中子孔隙度和声波时差较高,自然伽马、密度和光电吸收截面指数较低^[19]。不同区域的成煤环境不同,同一煤层的结构、顶底板岩性、物理性质、工业组分等会存在差异,测井曲线形态、特征、幅值等也会不一致,这是利用测井数据预测煤质的物理基础^[20]。

采用单测井参数交会分析方法,对煤的敏感性参数进行测井敏感性分析。如图4所示,在参数置信度为95%的情况下,挥发分产率与声波时差、补偿密度测井值具有较高的负相关关系,与自然伽马测井值相关性较低。上述现象产生的原因是煤中挥发分物质以化合物形式赋存,挥发分产率越高,这意味着煤中含有越多的可燃物质,导致声波在煤层中的传播速度降低。测井响应表现为高声波时差、低密度的特性(图4)。煤中放射性元素主要呈无机态赋存,而伽马值与放射性矿物含量呈正相关^[21]。煤灰中所含的矿物通常含有放射性元素^[22],自然伽马测井值与灰分产率呈正相关,而灰分含量通常与焦油产率呈负相关,导致伽马值与焦油产率呈负相关,间接体现了富油煤低伽马值的特性。

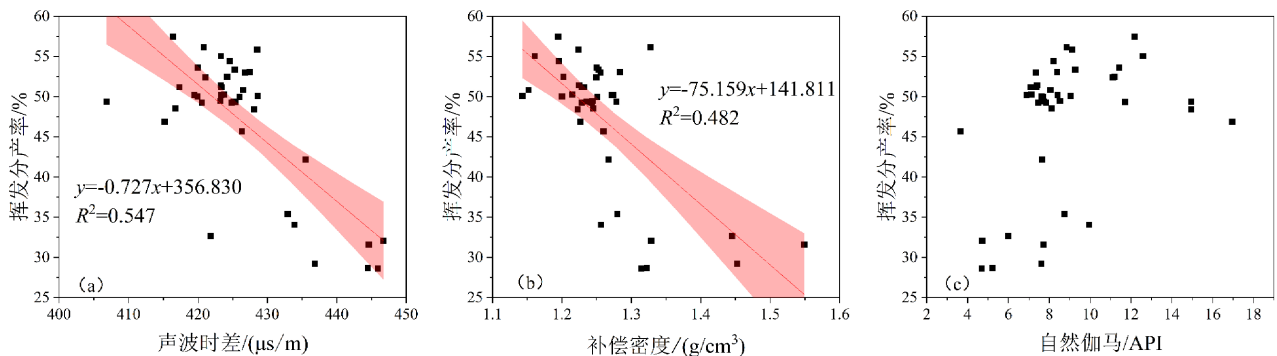


图4 煤样挥发分产率与测井响应参数交会图

煤样氢含量与补偿密度测井值负相关性显著,与声波时差测井值相关性一般,与自然伽马测井值相关性较低(图5),这种趋势在0.05显著水平下是显著的.产生上述现象的原因是煤中氢含量的变化对分子结构产生影响,使声波在煤层中传播速度较慢,高油富氢煤分子结构排列相对松散,加之煤层内部为条带状、片状及层状等界面结构,导致高油煤声波时差测井响应表现为高声波时差、低密度的特性.

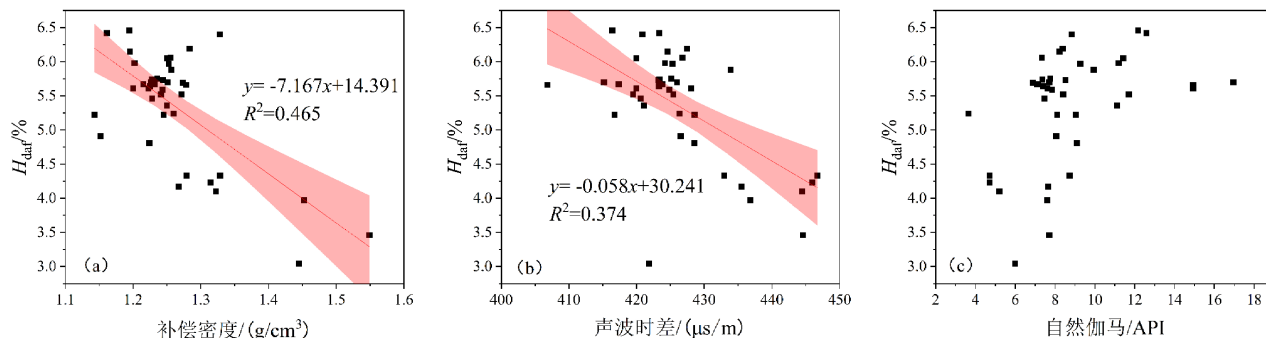


图5 煤样氢含量与测井响应交会图

如图6所示,煤样镜质组含量与声波时差、补偿密度测井值具有较高的负相关关系,与自然伽马测井值相关性较低,这种趋势在0.05显著水平下是显著的.镜质组是煤热解产油的主要组分,密度相对较低.声波在煤层中传播速度越大,声波时差越小,镜质组含量越高.

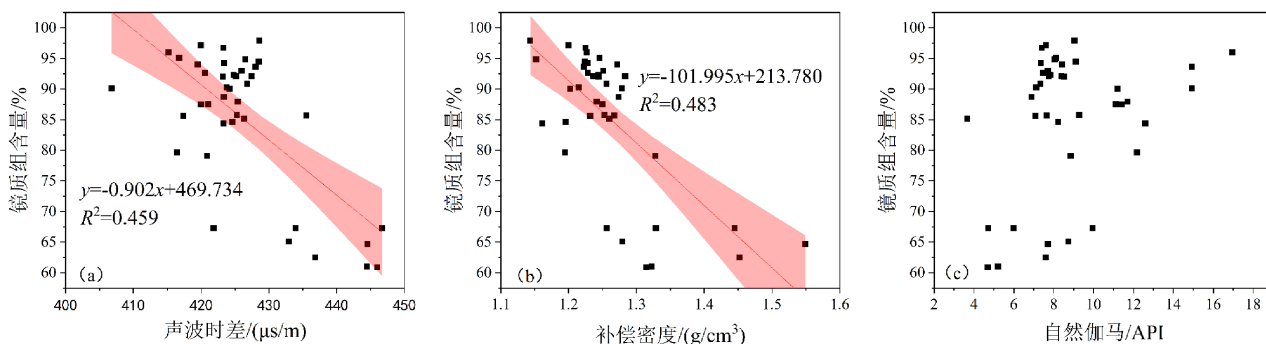


图6 煤样镜质组含量与测井响应交会图

煤样敏感性参数与测井响应之间的离散性较大.测井响应是地层、钻孔、仪器等多种因素叠加的综合反映,测井值随空间位置和时间发生变化,即使同一设备在不同时间的测井值也可能产生误差.因此,笔者建议针对不同的测井值,优先采用同一单位、同一测井设备、同一测井标准所测得的测井值.

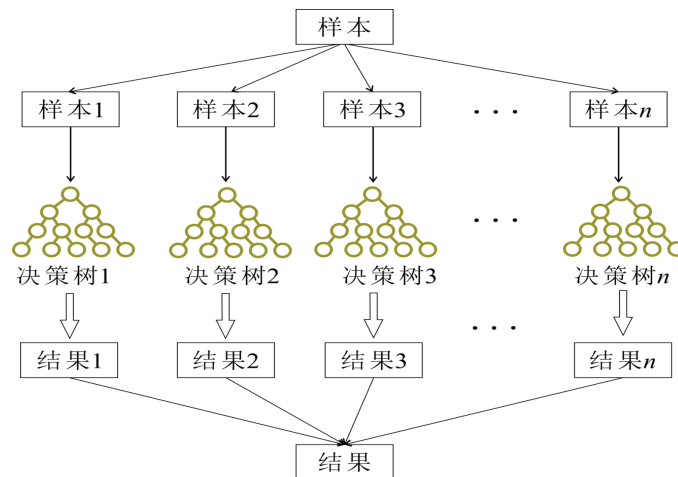
4 煤焦油产率测井预测模型

上述相关性分析显示,挥发分产率、氢含量、镜质组含量增加,焦油产率随之增大,且具有较高的正相关关系,而这3个参数与声波时差、补偿密度测井值具有较高的负相关关系,与自然伽马测井值相关性一般.依据机器学习方法,构建随机森林预测模型.随机森林预测模型的关键在于输入变量的选取,输入变量应相互独立以避免多重共线性,且与输出变量具有较好的相关性.因收集到的资料有限,基于上述研究,以声波时差、自然伽马、补偿密度3类相互独立的测井数值作为“测井数据-煤岩煤质特征”模型的输入层变量,分别以煤样挥发分产率、氢含量和镜质组含量作为输出变量的期望输出值;以挥发分产率、氢含量、镜质组含量3类相互独立的敏感性参数作为“煤岩煤质特征-焦油产率”模型的输入层变量,煤样焦油产率作为输出变量的期望输出值,基于输入参数的特征构建决策树,通过集成所有决策树的预测结果得到最终的预测值,形成“测井数据-煤岩煤质特征-焦油产率”预测模型.

4.1 方法与数据预处理

4.1.1 随机森林

随机森林是一种基于决策树的集成学习算法(图7),其通过自助法重采样技术,从训练样本集 N 中重复随机抽取 m 个样本;然后,随机选取 t 个特征,按以上步骤生成 n 棵决策树,共同组成随机森林^[24-25];对多个决策树的预测结果进行平均或加权平均,确定为最终结果^[26-27].相比多元线性回归,该方法可以更好解决预测问题.

图 7 随机森林模型示意图^[23]

4.1.2 数据预处理

数据预处理是回归预测和机器学习模型构建中极为关键的环节. 开展数据分析前, 通常需要收集大量不同类型的相关指标. 这些指标在性质、量纲、数量级和可用性等方面存在显著差异, 难以直接用于分析, 尤其是指标量级相差较大时, 直接使用会削弱数值较低指标的影响. 因此, 必须对原始指标数据进行变换处理, 以使不同特征具有统一的尺度. 本文使用标准化方法对测井数据进行预处理, 找出每个属性的最小值和最大值(去除异常值), 将原始值 X 通过标准化映射成为区间 $[0, 1]$ 中的值 X' :

$$X' = (X - X_{\min}) / (X_{\max} - X_{\min}). \quad (1)$$

4.2 敏感因素测井响应预测

利用敏感性参数分析数据和测井数据, 采用随机森林回归算法构建研究区挥发分产率、氢含量、镜质组含量-测井预测模型. 以声波时差、自然伽马、补偿密度3类测井数值作为输入层变量, 分别以煤样挥发分产率、氢含量和镜质组含量作为输出层的期望值, 利用随机森林回归算法, 得到挥发分产率、氢含量、镜质组含量预测模型. 相关参数设置如表3所示.

表 3 煤岩煤质敏感性参数预测模型参数设置

参数名	V_{daf}	H_{daf}	镜质组含量
训练用时	0.188 s	0.147 s	0.224 s
数据切分	0.7	0.8	0.7
数据洗牌	是	是	是
交叉验证	否	否	否
节点分裂评价准则	MAE	MAE	MAE
划分时考虑的最大特征比例	NONE	NONE	NONE
树的最大深度	10	10	11
叶子节点的最大数量	50	50	50
决策树数量	140	102	170
有放回采样	TRUE	TRUE	TRUE
袋外数据测试	FALSE	FALSE	FALSE

根据训练集数据预测结果可知, 基于声波时差、自然伽马和补偿密度3类测井数值构建的随机森林模型, 预测的挥发分产率平均绝对误差为1.165, 与实际值的相关系数(R^2)为0.940; 预测的氢含量平均绝对误差为0.201, 预测值与实际值的相关系数为0.890; 预测的镜质组含量平均绝对误差为2.427, 预测值与实际值的相关系数为0.932(表4).

在测试集数据预测结果中,基于声波时差、自然伽马和补偿密度3类测井数值构建的随机森林模型,预测的挥发分产率平均绝对误差为2.425,预测值与实际值的相关系数为0.883;预测的氢含量平均绝对误差为0.128,预测值与实际值的相关系数为0.895;预测的镜质组含量平均绝对误差为2.654,预测值与实际值的相关系数为0.853(表4)。

表4 煤岩煤质敏感性参数预测模型评估结果

敏感地质因素		<i>MSE</i>	<i>RMSE</i>	<i>MAE</i>	<i>MAPE</i>	R^2
V_{daf}	训练集	3.592	1.895	1.165	2.680	0.940
	测试集	10.176	3.190	2.425	5.791	0.883
H_{daf}	训练集	0.083	0.288	0.201	4.006	0.890
	测试集	0.024	0.154	0.128	2.291	0.895
镜质组	训练集	9.775	3.126	2.427	3.043	0.932
	测试集	9.344	3.057	2.654	2.981	0.853

注:*MSE*为均方误差,*RMSE*为均方根误差,*MAE*为平均绝对误差,*MAPE*为平均绝对百分比误差, R^2 为相关系数;下同

4.3 基于随机森林的煤焦油产率预测

基于煤样的敏感性参数测试分析数据和测井数据,采用随机森林回归算法构建镜质组含量、挥发分产率、氢含量测井计算模型;基于煤样的焦油产率和镜质组含量、挥发分产率、氢含量关系,采用随机森林回归算法构建焦油产率计算模型,形成“测井数据-煤岩煤质特征-焦油产率”预测模型。采用挥发分产率、氢含量、镜质组含量作为输入层的变量,以煤焦油产率作为输出层的期望值,利用随机森林回归算法得到焦油产率预测结果。相关参数设置见表5。

表5 煤焦油产率预测模型参数

参数名	参数值	参数名	参数值
训练用时	0.136 s	树的最大深度	11
数据切分	0.7	叶子节点的最大数量	50
数据洗牌	是	节点划分不纯度的阈值	0
交叉验证	否	决策树数量	104
节点分裂评价准则	MAE	有放回采样	TRUE
划分时考虑的最大特征比例	NONE	袋外数据测试	FALSE

以3个敏感性参数为自变量,采用随机森林模型预测煤的焦油产率(表6)。在参数置信度为95%的情况下,训练集的预测结果表明,煤焦油产率预测值平均绝对误差为0.722,预测值与实际值的相关系数为0.973;在测试集的结果中,煤焦油产率预测值平均绝对误差为1.298,预测值与实际值的相关系数为0.919。

表6 煤焦油产率预测模型评估结果

	<i>MSE</i>	<i>RMSE</i>	<i>MAE</i>	<i>MAPE</i>	R^2
训练集	0.985	0.992	0.722	7.126	0.973
测试集	2.711	1.647	1.298	9.964	0.919

4.4 预测模型评价

利用测井信息对煤的3个敏感性参数进行预测,再利用3个参数预测值对焦油产率进行预测。如图8所示,氢含量、挥发分产率、镜质组含量预测值与实际值的相关系数依次为0.917 1、0.922 5和0.940 8,平均绝对误差依次为0.186 7、1.543 0和2.498 4;煤焦油产率预测值与实际值的相关系数为0.915 8,平均绝对误差为1.508 7。预测值与实测值的拟合效果较好,相关系数均高于0.9,同时平均绝对误差也处于较低水平,均小于2.5。

最终预测结果显示,焦油产率预测值在90%的数据上相对误差小于20%。40个样本中仅有4个相对误差大于20%,其预测值均大于实际值(其中2个样本实际焦油产率小于7%)。造成此结果的原因为研究区建模数据中含油煤数据较少,富油煤、高油煤数据较多,导致部分实际值较低的点预测值偏高。焦油产率预测值的平均值为13.299%,实际值的平均值为13.475%,相对误差为1.31%。

上述结果表明基于测井的“测井数据-煤岩煤质特征-焦油产率”模型预测焦油产率可行,具有一定的实际应用前景,同时后续应尽可能测试和收集更多数据资料,以保证模型更加精准。

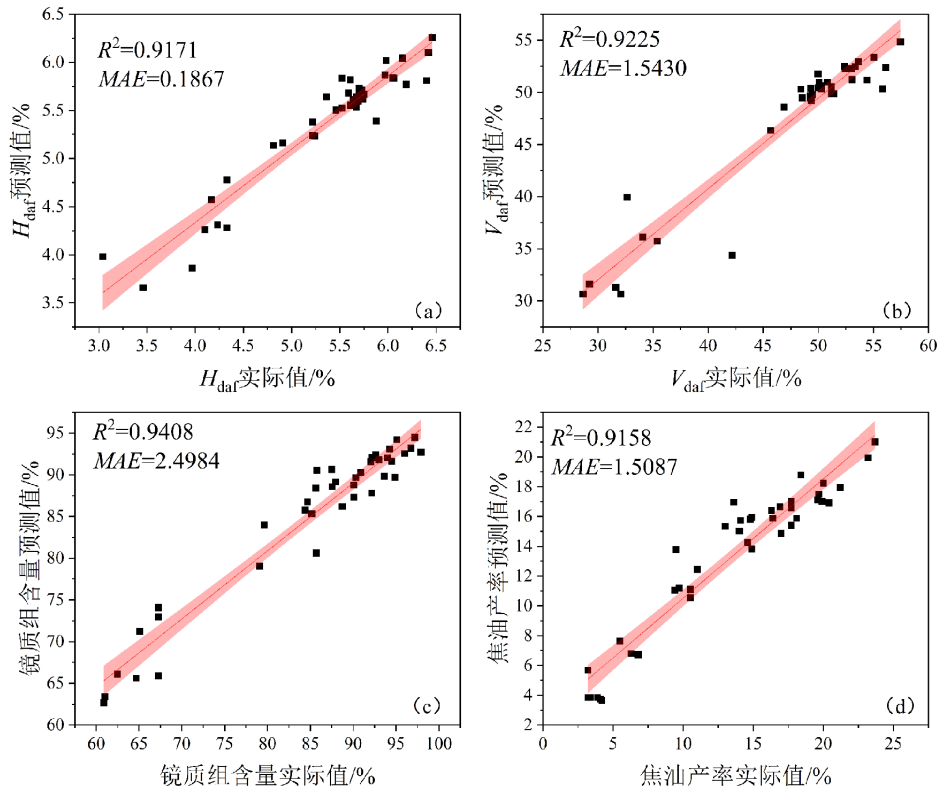


图8 煤样实际值和模型预测值交会图

5 结论

1) 两口取心井煤样的焦油产率为3.20%~23.70%,平均值为13.48%,焦油产率较高,为富-高油煤;煤样水分含量较低,灰分产率低,挥发分产率高;碳元素含量为69.69%~82.14%,氢元素含量为3.04%~6.46%,氮元素含量为0.84%~1.25%;煤样镜质组分占比为60.90%~97.92%,以基质镜质体(胶结半丝质体、丝质体碎片、碎屑惰质体、壳质组等)为主。

2) 煤焦油产率与灰分产率、惰质组含量呈负相关关系,与挥发分产率、氢元素含量、镜质组含量呈正相关关系。煤的挥发分产率、氢元素含量和镜质组含量与声波时差、补偿密度测井值呈较高的负相关性,而与自然伽马测井值的相关性较弱。

3) 利用测井数据、煤岩煤质数据建立的“测井数据-煤岩煤质特征”预测模型,以及焦油产率、煤岩煤质数据建立的“煤岩煤质特征-焦油产率”回归预测模型,预测值与实际值相关系数为0.92,在90%的数据上相对误差小于20%,表明基于机器学习的“测井数据-煤岩煤质特征-焦油产率”模型预测焦油产率可行,具有一定的实际应用前景。

参考文献:

- [1] 周忠科,王立杰. 我国煤基清洁能源发展潜力及趋势[J]. 中国煤炭, 2011, 37(5): 24-27+36.
ZHOU Z K, WANG L J. Development potentiality and tendency of coal-derived clean energy in China[J]. China Coal, 2011, 37(5): 24-27+36. (in Chinese)

- [2] 王双明,师庆民,王生全,等. 富油煤的油气资源属性与绿色低碳开发[J]. 煤炭学报, 2021, 46(5): 1365-1377.
WANG S M, SHI Q M, WANG S Q, et al. Resource property and exploitation concepts with green and low-carbon of tar-rich coal as coal-based oil and gas[J]. Journal of China Coal Society, 2021, 46(5): 1365-1377. (in Chinese)
- [3] 王苗,王长安,宁星,等. 富油煤原位热解技术研究现状及进展[J]. 煤炭学报, 2024, 49(9): 3969-3984.
WANG M, WANG C A, NING X, et al. Research progress of in-situ pyrolysis technology for tar-rich coal[J]. Journal of China Coal Society, 2024, 49(9): 3969-3984. (in Chinese)
- [4] 马丽,王双明,段中会,等. 陕西省富油煤资源潜力及开发建议[J]. 煤田地质与勘探, 2022, 50(2): 1-8.
MA L, WANG S M, DUAN Z H, et al. Potential of oil-rich coal resources in Shaanxi Province and its new development suggestion[J]. Coal Geology & Exploration, 2022, 50(2): 1-8. (in Chinese)
- [5] 吴琰杰,王帅,何磊,等. 吐哈盆地小草湖凹陷东缘侏罗系煤系烃源岩埋藏史、热演化史模拟[J]. 西北地质, 2021, 54(4): 180-191.
WU Y J, WANG S, HE L, et al. Research on the burial history and the thermal evolution history of the Jurassic coal-measure source rocks in the eastern margin of Xiaocaohu sag[J]. Northwestern Geology, 2021, 54(4): 180-191. (in Chinese)
- [6] 王双明,鲍园,郝永辉,等. 富油煤研究进展与趋势[J]. 煤田地质与勘探, 2024, 52(4): 1-11.
WANG S M, BAO Y, HAO Y H, et al. Research on tar-rich coals: Progress and prospects[J]. Coal Geology & Exploration, 2024, 52(4): 1-11. (in Chinese)
- [7] 师庆民,王双明,王生全,等. 神府南部延安组富油煤多源判别规律[J]. 煤炭学报, 2022, 47(5): 2057-2066.
SHI Q M, WANG S M, WANG S Q, et al. Multi-source identification and internal relationship of tar-rich coal of the Yan'an formation in the south of Shenfu[J]. Journal of China Coal Society, 2022, 47(5): 2057-2066. (in Chinese)
- [8] 田瀚,李宁,王双明,等. 富油煤焦油产率测井评价方法研究[J]. 煤田地质与勘探, 2024, 52(7): 97-107.
TIAN H, LI N, WANG S M, et al. A log-based method for evaluating the tar yield of tar-rich coal[J]. Coal Geology & Exploration, 2024, 52(7): 97-107. (in Chinese)
- [9] 马丽,段中会,贺丹,等. 陕北三叠纪煤田中-特高油产率煤分布规律及成因机理[J]. 中国煤炭地质, 2022, 34(1): 32-36.
MA L, DUAN Z H, HE D, et al. Medium-extra high tar yield coal resources distribution pattern and genetic mechanism in northern Shaanxi[J]. Coal Geology of China, 2022, 34(1): 32-36. (in Chinese)
- [10] 闫和平,段中会,王金锋. 黄陵矿区富油煤焦油产率与补偿密度关系模型预测方法研究[J]. 中国煤炭地质, 2022, 34(10): 25-30.
YAN H P, DUAN Z H, WANG J F. Study on the relationship model between oil-rich coal tar yield and compensation density in Huangling mining area[J]. Coal Geology of China, 2022, 34(10): 25-30. (in Chinese)
- [11] 赵军龙,闫和平,王金锋,等. 基于测井信息的煤焦油产率预测方法研究[J]. 地球物理学进展, 2023, 38(4): 1702-1712.
ZHAO J L, YAN H P, WANG J F, et al. Research on coal tar productivity prediction method based on logging information[J]. Progress in Geophysics, 2023, 38(4): 1702-1712. (in Chinese)
- [12] 乔军伟,王昌建,赵泓超,等. 基于煤岩煤质多元指标的BP神经网络焦油产率预测方法研究[J]. 煤田地质与勘探, 2024, 52(7): 108-118.
QIAO J W, WANG C J, ZHAO H C, et al. A method for predicting the tar yield of tar-rich coals based on the BP neural network using multiple indicators of coal petrography and coal quality[J]. Coal Geology & Exploration, 2024, 52(7): 108-118. (in Chinese)
- [13] 孙杰,邹卓,张莉,等. 新疆北部矿区煤中焦油产率分析及热解生油实验研究[J]. 中国煤炭地质, 2023, 35(5): 14-19.
SUN J, ZOU Z, ZHANG L, et al. Analysis of tar yield and experimental study on oil generation from coal pyrolysis in northern Xinjiang[J]. Coal Geology of China, 2023, 35(5): 14-19. (in Chinese)
- [14] 卞正富,张益东,王猛,等. 新疆煤炭资源开发潜力与开发策略[J]. 煤炭学报, 2024, 49(2): 967-977.
BIAN Z F, ZHANG Y D, WANG M, et al. Research on the potential and strategy for coal resources exploitation in Xinjiang[J]. Journal of China Coal Society, 2024, 49(2): 967-977. (in Chinese)
- [15] 宁树正,张莉,徐小涛,等. 新疆北部早、中侏罗世富油煤分布规律及控制因素[J]. 煤炭科学技术, 2024, 52(1): 244-254.
NING S Z, ZHANG L, XU X T, et al. Distribution of Early and Middle Jurassic tar-rich coal and its geological controls in northern Xinjiang[J]. Coal Science and Technology, 2024, 52(1): 244-254. (in Chinese)
- [16] 李盛富,陈洪德,周剑,等. 新疆伊犁盆地南缘中生代以来构造演化与聚煤规律研究[J]. 西北地质, 2016, 49(2): 220-228.
LI S F, CHEN H D, ZHOU J, et al. Tectonic evolution and coal accumulation about the southern margin of Yili Basin in Xinjiang since Middle Cenozoic Era[J]. Northwestern Geology, 2016, 49(2): 220-228. (in Chinese)

- [17] 李洪, 姚亚明, 张俊敏, 等. 焉耆盆地煤系源岩特征、生烃机理及烃源岩倾向性研究[J]. 西北地质, 2015, 48(1): 229-240.
LI H, YAO Y M, ZHANG J M, et al. Research on characteristics and hydrocarbon-generating mechanism of Jurassic coal measures source rock of Yanqi Basin and analysis on hydrocarbon-generating capacity and tendency(liquid and gas) of different source rock[J]. *Northwestern Geology*, 2015, 48(1): 229-240. (in Chinese)
- [18] 孙明波, 樊泽霞, 王书琪, 等. 塔里木盆地煤层坍塌机理研究[J]. 石油大学学报(自然科学版), 2004, 28(2): 49-51.
SUN M B, FAN Z X, WANG S Q, et al. Synthetic sloughing mechanism of coal beds in Tarim Basin[J]. *Journal of the University of Petroleum, China*, 2004, 28(2): 49-51. (in Chinese)
- [19] 钱玉萍, 张国华. 煤层工业组分的测井评价方法研究及应用[J]. 海洋石油, 2021, 41(3): 48-51+71.
QIAN Y P, ZHANG G H. Research and application of logging evaluation method for industrial components of coalbed[J]. *Offshore Oil*, 2021, 41(3): 48-51+71. (in Chinese)
- [20] 余永鹏, 张广兵, 黄自军, 等. 多源测井数据预测煤层工业组分和发热量模型研究[J]. 物探与化探, 2024, 48(1): 185-193.
YU Y P, ZHANG G B, HUANG Z J, et al. A prediction model of the industrial components and calorific values of coal seams based on multi-source log data[J]. *Geophysical and Geochemical Exploration*, 2024, 48(1): 185-193. (in Chinese)
- [21] ITO T, OHBUCHI A, NAKAJIMA T, et al. Identifying the source of natural gamma-rays in shallow-marine siliciclastic strata and their significance for shale evaluation: A case study of the CO₂ storage aquifer at the Nagaoka site, Japan[J]. *Journal of Natural Gas Science and Engineering*, 2017, 46: 782-792.
- [22] MADZIVIRE G, MALEKA P P, VADAPALLI V R K, et al. Fate of the naturally occurring radioactive materials during treatment of acid mine drainage with coal fly ash and aluminium hydroxide[J]. *Journal of Environmental Management*, 2014, 133: 12-17.
- [23] BREIMAN L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5-32.
- [24] KENSERT A, ALVARSSON J, NORINDER U, et al. Evaluating parameters for ligand-based modeling with random forest on sparse data sets[J]. *Journal of Cheminformatics*, 2018, 10(1): 49.
- [25] TEMEL F A, YOLCU O C, TURAN N G. Artificial intelligence and machine learning approaches in composting process: A review[J]. *Bioresource Technology*, 2023, 370: 128539.
- [26] HAN J T, GU Y J. Study on handling range inputs methods on C4.5 algorithm[C]//2009 International Forum on Computer Science-Technology and Applications. December 25-27, 2009. Chongqing, China. IEEE, 2009: 47-49.
- [27] THAKUR D, MARKANDIAH N, RAJ D S. Re optimization of ID3 and C4.5 decision tree[C]//2010 International Conference on Computer and Communication Technology(ICCCCT). September 17-19, 2010. Allahabad, Uttar Pradesh, India. IEEE, 2010: 448-450.

责任编辑: 张自强