



DOI:10.12404/j.issn.1671-1815.2507596

引用格式: 刘宏, 李军, 杨宏伟, 等. 基于录井时序数据与多尺度分解的井漏智能诊断模型: 以伊朗 Azadegan 油田为例[J]. 科学技术与工程, 2026, 26(26): 11297-11307.

Liu Hong, Li Jun, Yang Hongwei, et al. Intelligent lost circulation diagnosis model based on logging time series data and multi-scale decomposition: Azadegan Oilfield in Iran[J]. Science Technology and Engineering, 2026, 26(26): 11297-11307.

基于录井时序数据与多尺度分解的井漏智能诊断模型: 以伊朗 Azadegan 油田为例

刘宏¹, 李军^{1,2*}, 杨宏伟¹, 祁文莉³, 柳贡慧^{1,4}, 杨钰龙⁵

(1. 中国石油大学(北京)石油工程学院, 北京 102249; 2. 中国石油大学(北京)克拉玛依校区石油学院, 克拉玛依 834000;

3. 中国石油玉门油田采油工艺研究院, 酒泉 735211; 4. 北京工业大学机械与能源工程学院, 北京 100124;

5. 中国石油大学(北京)油气资源与工程全国重点实验室, 北京 102249)

摘 要 井漏是钻井作业中常见的一种复杂工况,严重影响钻井的安全性和效率。现有的井漏预警模型大多集中在预测井漏发生时刻,而对井漏的严重程度及其风险评估关注较少。为此,提出了一种融合录井时序数据与多尺度分解的井漏智能诊断模型,该模型结合了改进自适应噪声完备集合经验模态分解(improved complete ensemble empirical mode decomposition with adaptive noise, ICEEMDAN)、本征模态分量(intrinsic mode function, IMF)筛选准则以及粒子群优化算法(particle swarm optimization, PSO)优化支持向量机(support vector machine, SVM)。首先,基于伊朗 Azadegan 油田的钻井完井报告,筛选出对不同漏失程度有明显表征的录井参数。其次,利用 ICEEMDAN 对录井参数数据进行多尺度分解,通过计算皮尔逊相关系数和能量熵,筛选出与井漏特征变化相关性强的 IMF 分量,并重构数据以去除中高频噪声干扰。最后,采用 PSO 优化 SVM 模型构建井漏诊断模型,其中 PSO 算法优化模型的惩罚系数 C 和核参数 γ ,从而提升模型的准确性和泛化能力。研究表明,ICEEMDAN-PSO-SVM 模型的测试集准确率达 98% 以上,相较于 PSO-SVM、随机森林(random forest, RF)和 SVM,在关键性能指标上有显著提升。该研究为降低钻井过程中的安全风险、提高工程效率和经济效益提供了重要支持。

关键词 井漏诊断; 井漏严重程度; ICEEMDAN; IMF; PSO-SVM

中图分类号 TE28;

文献标志码 A

Intelligent Lost Circulation Diagnosis Model Based on Logging Time Series Data and Multi-scale Decomposition: Azadegan Oilfield in Iran

LIU Hong¹, LI Jun^{1,2*}, YANG Hong-wei¹, QI Wen-li³, LIU Gong-hui^{1,4}, YANG Yu-long⁵

(1. College of Petroleum Engineering, China University of Petroleum (Beijing), Beijing 102249, China; 2. College of Petroleum, China University of Petroleum(Beijing) at Karamay, Karamay 834000, China; 3. Research Institute of Oil Production Technology, Yumen Oilfield, Petro China,

Jiuquan 735211, China; 4. College of Mechanical and Energy Engineering, Beijing University of Technology, Beijing 100124, China;

5. National Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing 102249, China)

[Abstract] Lost circulation is a common complex condition in drilling operations that severely compromises safety and efficiency. Most existing early-warning models focus mainly on predicting when lost circulation occurs, with insufficient attention to its severity and risk assessment. To address this gap, an intelligent lost circulation diagnostic model integrating logging time series data and multi-scale decomposition was proposed, combining ICEEMDAN (improved complete ensemble empirical mode decomposition with adaptive noise), IMF (intrinsic mode function) screening criteria, and SVM (support vector machine) optimized by PSO (particle Swarm optimization). First, from drilling and completion reports of Iran's Azadegan Oilfield, logging parameters that effectively characterize different leakage degrees were selected. Second, the logging data undergo multi-scale decomposition via ICEEMDAN. IMF components strongly correlated with leakage characteristics were filtered using Pearson correlation coefficients and energy entropy, and the data are reconstructed to eliminate medium-to-high-frequency noise. Finally, a diagnostic model was built with the PSO-optimized SVM, in which the PSO optimizes the penalty coefficient C and kernel parameter γ , to enhance accuracy and generalization. Results indicate the

收稿日期: 2025-11-02; 修订日期: 2026-04-07

基金项目: 新型油气勘探开发国家科技重大专项(2025ZD1403206, 2025ZD1401900); 国家自然科学基金重大科研仪器研制项目(52227804); 国家自然科学基金联合基金(U22B2072); 国家自然科学基金面上项目(52474018)

第一作者: 刘宏(1999—),男,汉族,山东潍坊人,博士研究生。研究方向:智能钻井与井漏检测。E-mail: hongliu_5695@163.com。

* 通信作者: 李军(1971—),男,汉族,河北沧州人,博士,教授。研究方向:页岩油气井射孔完井、井筒完整性、高温高压井自动压井。E-mail: lijun446@vip.163.com。

投稿网址: www.stae.com.cn

ICEEMDAN-PSO-SVM model's test set accuracy exceeds 98%, outperforming PSO-SVM, RF (random forest) and SVM across key performance metrics. This study offers valuable support for mitigating drilling safety risks and enhancing engineering efficiency and economic benefits.

[**Keywords**] lost circulation diagnosis; severity of lost circulation; ICEEMDAN; IMF; PSO-SVM

全球工业化推进下,能源需求激增带来紧迫挑战。深层、超深层及非常规油气资源的勘探开发是增储关键,其钻井作业的安全与高效直接影响油气顺利开采^[1-2]。钻井过程中,天然裂缝地层或地层与井筒压力不平衡易引发井漏,不仅造成经济损失,还可能污染地层环境。因此,准确识别并预防井漏的发生,对提升作业安全性与效率意义重大^[3]。

当前井漏检测技术主要分为4类:传统监测、地质力学分析、仪器测量和机器学习预测。传统监测依赖人工观察泥浆池液面及井口流体状态,准确性、稳定性不足且成本较高;地质力学分析通过地层力学特性与井内参数变化评估井漏风险^[4];仪器测量借助流量计、温压传感器等实时监测关键参数以确定漏失位置,随钻测井与随钻环空压力监测结合可实时发现异常、识别井漏^[5],但这类技术易受井下极端环境影响,且设备复杂、成本偏高;随着人工智能发展,机器学习在油气行业应用渐广,通过算法模型识别井漏模式与趋势,可预测潜在风险。蒯研锋等^[6]提出基于专家经验的井漏智能预测方法,采用LSTM和Bi-LSTM捕捉综合录井曲线动态自相关特征,提升了预测准确性;张矿生等^[7]提出因果关系约束下的水平井井漏智能预警方法并应用于长庆油田,通过LSTM集成网络实现预警,准确率达95.6%。而现有井漏预警模型多聚焦于预测井漏发生时刻,对漏失严重程度及风险评估关注不足,尤其缺乏对轻微渗漏与严重漏失的针对性预测。此外,现有模型多通过改进结构提升性能,数据处理方法有限,未能充分解决钻井现场高频噪声对综合录井数据的负面影响,可能降低模型准确性。因此,需依托有效的数据处理技术与适配模型,提高井漏程度风险预测的准确性。

针对现有研究局限,从数据处理与模型优化两方面展开针对性优化。在数据处理方面,钻井现场中高频噪声复杂,传统单一小波变换去噪效果不足^[8],而多尺度信号分解算法(如EMD)可将信号分解为不同尺度子信号,精准分离噪声、保留有效信息^[9]。在模型优化方面,井漏程度诊断属于分类任务,SVM在相关领域应用广泛。李凯阳等^[10]提出一种基于变分模态分解结合粒子群优化支持向量机模型进行手势分类,模型识别6种手势动作时识别准确率可达97.30%;邢昭等^[11]联合支持向量机

和梯度提升决策树构建模型以用于研究区的滑坡危险性评价,有效预测出研究区域的极高危险区、高危险区、中危险区和低危险区的占比情况;Deng等^[12]的IFOA-SVM井漏预测模型,能以更少迭代次数实现更高预测准确性;白凯等^[13]用ISSA优化IRF模型参数,性能显著优于经典模型。通过调研可知,合理应用多尺度分解数据处理方法与群智优化算法优化模型参数,对提升井漏诊断模型性能至关重要。

基于上述分析,现提出一种基于ICEEMDAN-SVM的井漏程度诊断方法。在数据处理阶段,采用ICEEMDAN算法对原始钻井信号进行多尺度分解,有效分离高频噪声与有效信息;在模型构建阶段,利用粒子群算法(particle swarm optimization, PSO)对SVM的惩罚系数 C 和核参数 γ 进行协同优化,构建PSO-SVM诊断模型。最后进行井漏程度分类准确率、泛化性能等方面的对比实验验证。研究成果可为复杂工况下井漏风险评估及现场防控措施制定提供依据,对保障钻井安全、降低井漏损失具有重要意义。

1 井漏机理与类型分析

1.1 井漏机理

井漏的发生与地层特性、井眼结构、钻井工艺参数和钻井液性能密切相关。高渗透性或裂缝发育地层易使钻井液渗入形成漏失通道。大尺寸或不规则井眼会致钻井液流速不均、局部压力过高,引发井漏。钻压过大、转速过高、排量过大等工艺参数问题,可能破坏地层结构或导致井底压力失衡。钻井液密度若过高,可能超地层破裂压力,压裂地层,形成漏失通道^[14-15]。井漏发生时,钻井工程与钻井液参数等均会发生显著变化。然而,仅依据单一参数识别井漏易出现较大误差,因此需综合多种参数变化进行判断^[16]。

1.2 井漏案例数据集

研究选取了伊朗Azadegan油田的数据以训练模型,用于预测钻井作业中的井漏程度。Azadegan油田是世界上最大的未开发油田之一,其内部包含I、II和III 3个不同的构造块。Mardanirad等^[17]对Azadegan油田中20口井的每日钻井、泥浆和完井报告进行了深入分析,并调研了不同区块的日常钻井和录井数据,结果表明大多数井漏事件发生在III

区。因此,从该区域中选取了 5 口具有代表性的邻井,这些井涵盖了不同的漏失情况,并利用其数据构建了井漏识别数据集。依据井漏机理,选取了井眼结构、钻井工程和钻井液等多项参数,具体参数信息如表 1 所示。此外,还从每日钻井报告中提取了每米钻井深度的钻井液漏失率,并根据漏失量(桶/h)将漏井程度划分为 5 个等级^[18],从“无漏失”到“完全漏失”依次编码为 0~4,具体划分标准如表 2 所示。

表 1 井漏数据集参数信息

Table 1 Information on the parameters of the lost circulation data set

参数	不同值数量	零值数量	平均值	标准差
井眼尺寸/in	5	0	11.90	4.61
钻深/m	4 082	0	2 077.40	1 002.70
钻速/(m·h ⁻¹)	2 928	0	8.96	12.30
钻压/klbf	5 060	0	13.16	8.14
转速/(r·min ⁻¹)	4 266	0	137.90	43.49
扭矩/(klbf·ft ⁻¹)	619	1 906	3.45	2.07
立管压力/psi	4 301	0	1 880.30	648.20
出口流量/%	4 888	0	58.05	12.43
泵冲次数/(次·min ⁻¹)	2 096	0	108.87	36.19
泥浆密度/pcf	91	0	74.95	4.66

注:1 in = 25.4 mm;1 klbf = 4.448 kN;1 klbf/ft = 1.356 kN·m;
1 psi = 6.895 × 10⁻³ MPa;1 pcf = 16.018 kg/m³。

表 2 漏井程度及划分标准

Table 2 Degree of lost circulation and classification criteria

井漏程度	漏失量/(桶·h ⁻¹)	标签编码
无漏失	0	0
轻微渗漏	(0,10]	1
中等漏失	(10,100]	2
严重漏失	>100	3
完全漏失	无返回	4

2 相关原理概述

2.1 改进自适应噪声完备集合经验模态分解

ICEEMDAN 作为自适应分解算法,在每步迭代中向待分解信号添加经 EMD 分解白噪声后得到的第 k 阶 IMF 分量,从而有效抑制了传统 EMD 算法中的模态混叠现象^[19],其流程如下。

(1)在待分解信号 x 中加入第 1 白噪声的模态分量 $E_1(w^i)$,得到加噪信号 x^i 为

$$x^i = x + \beta_0 E_1(w^i) \quad (1)$$

式(1)中: β_0 为第 1 阶段的信噪比; w^i 为高斯白噪声。

(2)利用 EMD 计算加噪信号 x^i 的局部均值,取均值得到第一残差 R_1 ,进而得到第一模态分量 IMF₁。

$$R_1 = \frac{1}{I} \sum_{i=1}^I L(x^i) \quad (2)$$

$$\text{IMF}_1 = x - R_1 = x - \frac{1}{I} \sum_{i=1}^I L(x^i) \quad (3)$$

式中: $L(\cdot)$ 为局部均值算子; I 为模态分解的次数。

(3)通过类推,将白噪声添加到 $k-1$ 残差中,得到 k 残差,最后将得到 k 模态分量。

$$\text{IMF}_k = R_{k-1} - R_k \quad (4)$$

$$\text{IMF}_k = \frac{1}{I} \sum_{i=1}^I L[R_{k-1} + \beta_{k-1} E_k(w^i)] \quad (5)$$

2.2 本征模态分量筛选准则

在 ICEEMDAN 分解后,并非所有 IMF 包含显著有效信息。为了去除噪声并保留原始信号主体,采用皮尔逊相关系数对 IMF 分量进行筛选,仅保留与原信号相关性高的分量用于重构。相关系数 P_k 计算式^[20]为

$$P_k = \frac{\sum_{j=1}^N x \text{IMF}_k(j)}{\sqrt{\sum_{j=1}^N x^2 \sum_{j=1}^N \text{IMF}_k^2(j)}} \quad (6)$$

式(6)中: N 为序列长度值;IMF _{k} 为每阶 IMF 分量, $k=1,2,\dots,n$ 。

井漏发生时,钻井参数中的有效信息在频域与能量域均会显著改变。因此在 ICEEMDAN 分解的基础上,观察每阶 IMF 的能量熵的动态差异,可精准区分有效信号与噪声分量,能量熵 H_{EN} 的计算式^[21]为

$$H_{\text{EN}} = - \sum_{k=1}^n P_k \log_2 P_k \quad (7)$$

式(7)中: P_k 为第 k 个 IMF 与总能量的比值。

2.3 群智优化诊断模型

SVM 是一种机器学习算法,旨在高维空间中找到一个最优超平面以实现数据的分类,并使得距离该平面最近的数据点到该平面距离最大^[22]。SVM 的惩罚系数 C 和核参数 γ 是影响模型性能的关键因素。传统方法常依赖人为经验设置参数,易导致模型性能欠佳。粒子群优化算法(PSO)通过模拟鸟群觅食行为,具有全局搜索能力强、收敛速度快的优点,可有效解决 SVM 参数优化问题^[23]。根据式(8),粒子群中的所有粒子会根据自己找到的当前个体极值和整个粒子群共享的当前全局最优解来调整自己的速度和位置,从而优化 SVM 的 C 和 γ ,提升模型在分类和回归任务中的性能,PSO 优化 SVM 的流程如图 1 所示。

$$\begin{cases} V_{\text{id}}^{m+1} = \omega V_{\text{id}}^m + c_1 r_1 (P_{\text{id}}^m - X_{\text{id}}^m) + c_2 r_2 (P_{\text{gd}}^m - X_{\text{id}}^m) \\ X_{\text{id}}^{m+1} = X_{\text{id}}^m + V_{\text{id}}^{m+1} \end{cases} \quad (8)$$

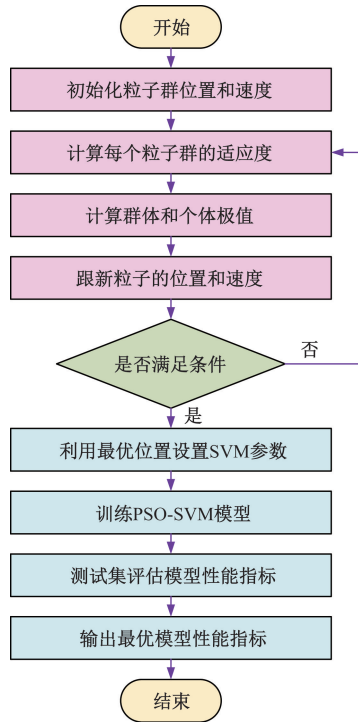


图1 粒子群优化诊断模型流程图

Fig. 1 Flowchart of particle swarm optimization diagnostic model

式(8)中; m 为迭代次数; ω 为惯性权重; P_{id}^m 和 P_{gd}^m 分别为粒子个体最优值和种群最优值; X_{id}^m 和 V_{id}^m 分别为粒子 s 次迭代中的位置和速度; c_1 和 c_2 均为非负常数的学习因子; r_1 和 r_2 均为分布于 $[0,1]$ 的随机数。

2.4 基于 ICEEMDAN-PSO-SVM 的井漏智能诊断模型框架

如图2所示,提出的基于 ICEEMDAN-PSO-SVM

的井漏智能诊断模型框架主要分为3个阶段。

阶段1:输入井漏表征参数,针对含高频噪声较多的数据,采用 ICEEMDAN 算法进行多尺度分解,结合多指标 IMF 筛选准则完成去噪处理,通过信号重构提升数据集质量。

阶段2:构建基于 PSO 优化 SVM 的井漏诊断模型,利用 PSO 算法对模型的惩罚系数 C 与核参数 γ 进行全局优化,进而提升模型的识别准确性与泛化能力。

阶段3:将优化后的 ICEEMDAN-PSO-SVM 模型分别与不同模型进行性能对比,验证其综合诊断性能;此外,通过对比不同数据处理方法对模型性能的提升效果,进一步验证 ICEEMDAN 算法在综合录井数据去噪中的优势。

3 井漏诊断识别案例

3.1 数据处理过程

研究以伊朗 Azadegan 油田 5 口邻井的漏失信息为基础,构建井漏数据集(表1和表2),并将其用于井漏诊断模型的训练与测试。首先对数据集进行实时数据质量控制,过滤掉离群值数据记录。这些异常数据可能源于泵送系统故障或数据采集系统安装不当。其次,在钻井施工现场,由于设备、环境等因素的影响,综合录井数据通常会受到噪声干扰,过多的噪声会严重影响数据集的质量,进而降低模型的性能。具体分析表1中的井漏数据集可知:井眼尺寸和钻深属于井眼结构范畴,数据采集过程不受噪声影响,因此无须额外的去噪处理;出口流量、泵冲次数和泥浆密度,这些数据

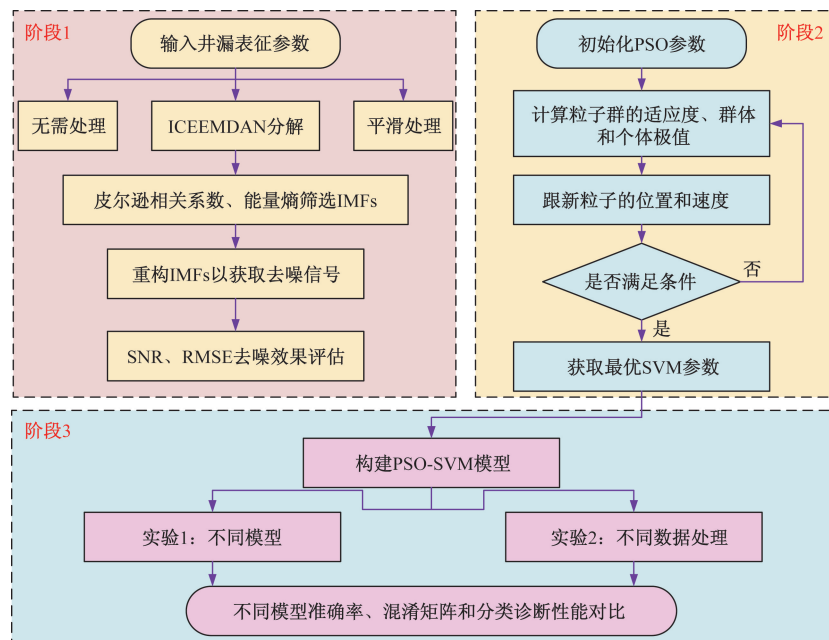


图2 基于 ICEEMDAN-PSO-SVM 的井漏智能诊断模型框架

Fig. 2 Framework for an intelligent lost circulation diagnosis model based on ICEEMDAN-PSO-SVM

含有较多的低频噪声分量,且变化较为缓慢。对于这类数据,仅需进行数据平滑处理即可有效去除噪声;钻速、钻压、转速、扭矩和立管压力数据,这些数据含有较多的高频噪声分量,简单的平滑处理无法有效去除其中的高频噪声。因此,需要采用 ICEEMDAN 算法对这些数据进行多尺度分解,并结合多指标 IMF 筛选准则进行去噪处理,以提高数据集的质量。

以钻速数据为例进行数据去噪处理(钻压、转速、扭矩及立管压力数据均采用相同处理流程,故本文中未单独展示其处理过程),如图 3 所示。从钻速数据的时域图中可以观察到,数据中存在大量“毛刺”等高频噪声干扰。为了去除这些噪声,采用 ICEEMDAN 算法对钻速数据进行分解。其中设置添加噪声的标准差(N_{std})为 0.2,白色噪声组数(N_R)为 100,迭代最大次数(Max_{Iter})为 1 000,此时分解效果最佳。从图 3 的分解结果可以看出,ICEEMDAN 算法分解产生的 IMF 数量较少,且残余噪声和虚假分量基本被消除。

如图 4 所示,钻速数据经过多尺度分解后,依次从高频 IMF1 分解到低频 IMF8。其中,高频 IMF 中主要包含大量高频噪声,这些噪声需要进行筛选后去除;而中低频 IMF 则主要包含有效信息分量和数

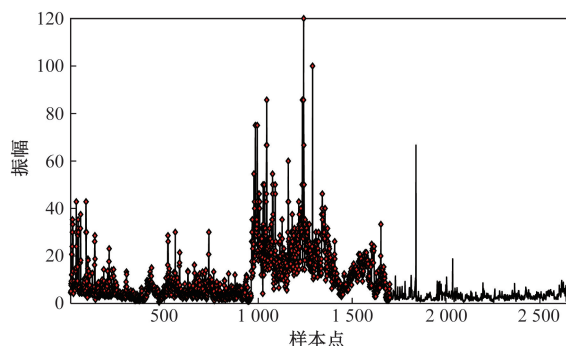


图 3 钻速数据的时域图

Fig. 3 The time-domain graph of the drilling rate data

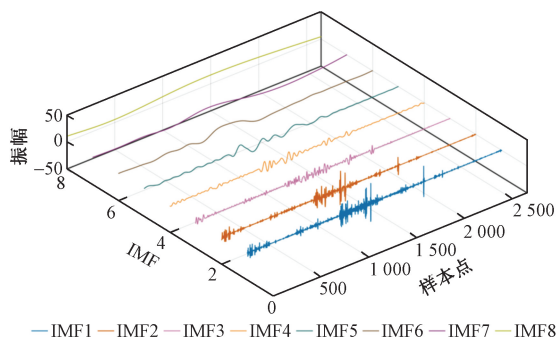


图 4 钻速数据的 ICEEMDAN 分解时域图

Fig. 4 ICEEMDAN decomposition time-domain graph of drilling rate data

据的趋势变化,这些分量需要进行筛选后保留,以确保数据的完整性和准确性。因此,采用皮尔逊相关系数和能量熵筛选相关 IMFs,计算结果如图 5 所示。从图 5 可知,IMF1 和 IMF2 的相关系数较低且能量熵较高,为高频噪声主导的分量,需要去除;IMF3 ~ IMF5 的相关系数较高且能量熵最低,为非平稳信号主成分,进行保留;IMF6 ~ IMF8 与原始信号程度相关,为信号趋势项,且能量在时频域分布较分散,进行保留。利用 IMF3-8 进行信号重构即可得到去噪处理后的钻速数据,如图 6 所示,去噪后的数据噪声明显减少,时域波形趋势更加明显和光滑。为了验证 ICEEMDAN 分解效果的优势,研究分别对比了不同分解算法结合去噪处理的效果,如图 7 所示。可发现采用 ICEEMDAN 分解的去噪数据,其信噪比(signal-to-noise ratio, SNR)最高可达 15.80 dB,且均方误差值(root mean square error, RMSE)最低可至 2.53,去噪效果最优;而采用 EMD 分解,其 SNR 和 RMSE 分别为 8.50 和 4.96,数据质量较低,且重构误差较大。这表明 ICEEMDAN 算法在去除噪声的同时,能够更好地保留信号的有效信息,明显优于传统的 EMD 分解方法。

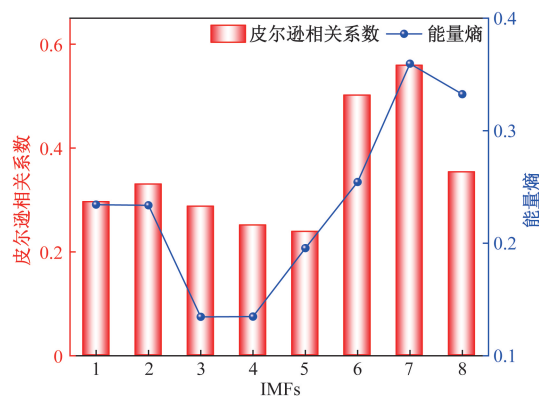


图 5 IMFs 的皮尔逊相关系数和能量熵计算值

Fig. 5 The Pearson correlation coefficient and energy entropy calculation values of IMFs

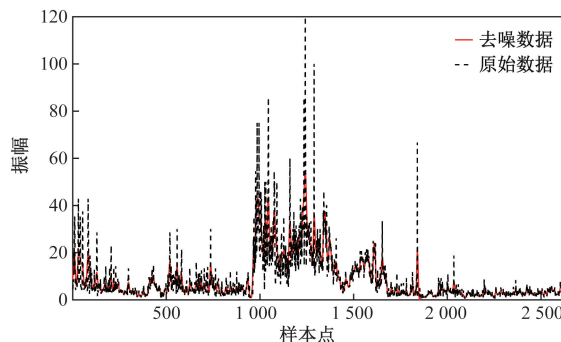


图 6 钻速数据的去噪前后对比结果

Fig. 6 The comparison results of the drilling rate data before and after denoising

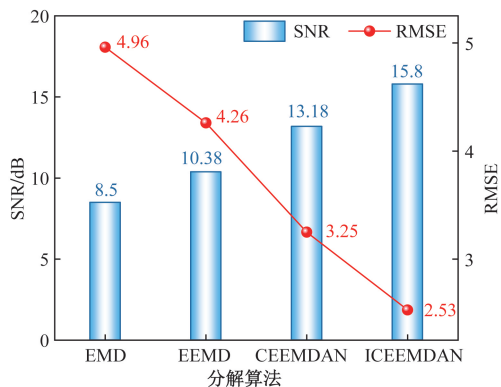


图7 不同分解算法的去噪效果对比

Fig. 7 Comparison of denoising effects of different decomposition algorithms

3.2 诊断模型搭建与优化

在对钻速、钻压、转速、扭矩和立管压力等数据进行去噪处理后,为了解决数据集中多参数特征尺度和量纲不一致的问题,采用了最大-最小归一化方法来处理特征数据,以消除不同量纲和单位的影响,其计算公式^[24]为

$$\tilde{X}_i = \frac{X_i - \min X_i}{\max X_i - \min X_i} \quad (9)$$

式(9)中: $\tilde{X}_i$ 为归一化后的数值; X_i 为原始第*i*个变量; $\min X_i$ 为原始第*i*个变量的最小值; $\max X_i$ 为原始第*i*个变量的最大值。

归一化后的数据集按照 7:3 的比例分割,其中 70% 的数据被分配给训练集,剩余的 30% 数据被分配给测试集,并独立于模型训练过程进行保存。使用井漏数据集在 ICEEMDAN-PSO-SVM、PSO-SVM、随机森林(random forest, RF)和 SVM 4 种模型中进行训练和测试,以比较不同模型的性能。其中,SVM 的 *C* 初始值设置为 0.1,最大值为 100; γ 初始值设置为 1.0,最大值为 100;采用 RBF 核函数。PSO 的主要参数设置如表 3 所示。群智优化算法的适应度曲线变化反映了算法的收敛速度和精度。图 8 展示了 PSO 优化 SVM 的适应度曲线。从图 8 可以看出,适应度值在前 4 次迭代中迅速下降,这表明算法在初期阶段快速找到了较优解。随后,适应度值在大约第 30 次迭代后趋于稳定,这表明算法在此时已经找到了一个最优解,适应度最终达到最低值约 0.022 9。

表3 PSO 的主要参数设置

Table 3 The main parameter settings of PSO

参数	数值
种群规模 <i>N</i>	30
惯性权重 ω	1.0
学习因子 <i>C</i> ₁	1.5
学习因子 <i>C</i> ₂	1.7
最大迭代次数	50

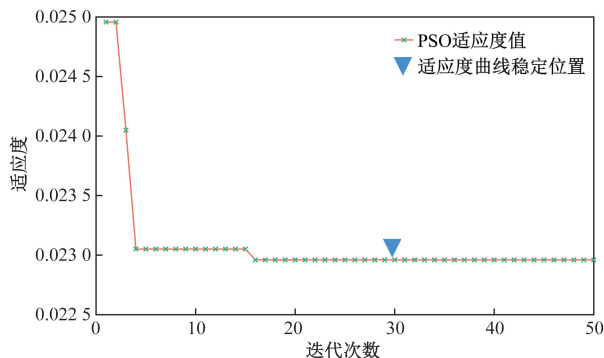


图8 PSO 优化 SVM 的适应度曲线

Fig. 8 PSO optimizes the fitness curve of SVM

3.3 实验结果分析

3.3.1 模型性能评估

采用准确率($A_{accuracy}$)、精确率($P_{recision}$)、召回率(R_{ecall})和 F_1 -Score 指标对诊断模型性能进行评估。此外,虚警率 R_{FA} 和漏警率(R_{MA})对于评估井漏诊断模型的实际应用效果至关重要,能够帮助更全面地评价模型的性能^[25-26]。

$$A_{accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$P_{recision} = \frac{TP}{TP + FP} \quad (11)$$

$$R_{ecall} = \frac{TP}{TP + FN} \quad (12)$$

$$F_1\text{-Score} = \frac{2 \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

$$R_{FA} = \frac{FP}{FP + TN} \quad (14)$$

$$R_{MA} = \frac{FN}{FN + TP} \quad (15)$$

式中:TP 为真阳性;TN 为真阴性;FP 为假阳性;FN 为假阴性。

3.3.2 实验 1:不同模型的性能对比

为了验证 ICEEMDAN-PSO-SVM 在井漏诊断模型的可行性,分别与 PSO-SVM、RF 和 SVM 的性能进行比较,每种模型的训练集准确率如图 9 所示。所提出的 ICEEMDAN-PSO-SVM 模型在训练集上达到了 99.691 4% 的高准确率,而 PSO-SVM 和 RF 的准确率分别为 95.834 5% 和 92.703 5%,表现良好。相比之下,SVM 的准确率仅为 74.816 2%,表现不佳。为了全面评估模型的分类性能,进一步比较了各模型在测试集上的混淆矩阵,如图 10 所示。其中,垂直轴表示真实标签,水平轴表示预测标签。结果显示,SVM 的测试集准确率最低,仅为 75.312 3%,而 ICEEMDAN-PSO-SVM 的测试集准确率最高,达到了 98.644 9%,明显优于其他模型。此外,计算了

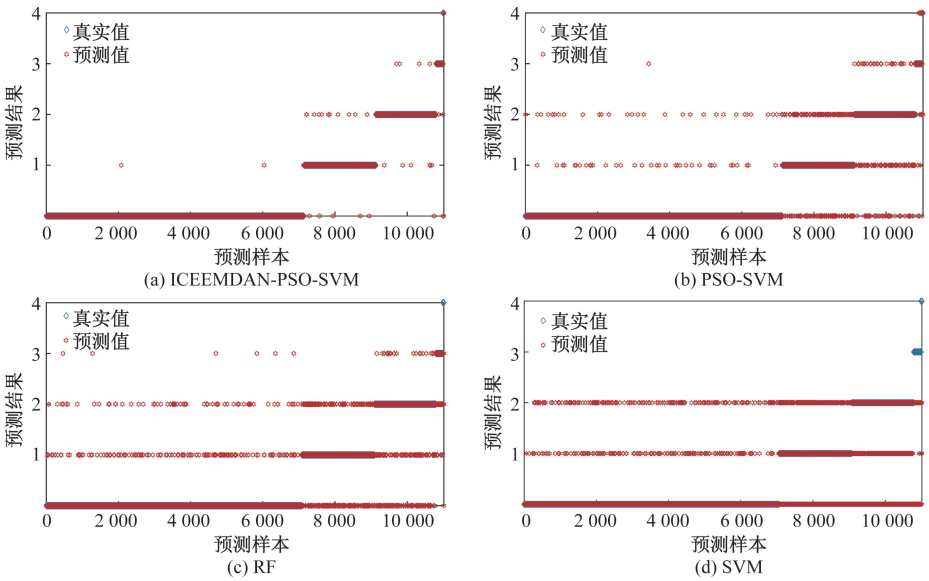


图 9 不同模型的训练集准确率(实验 1)

Fig. 9 Accuracy rates of training sets for different models (Experiment 1)

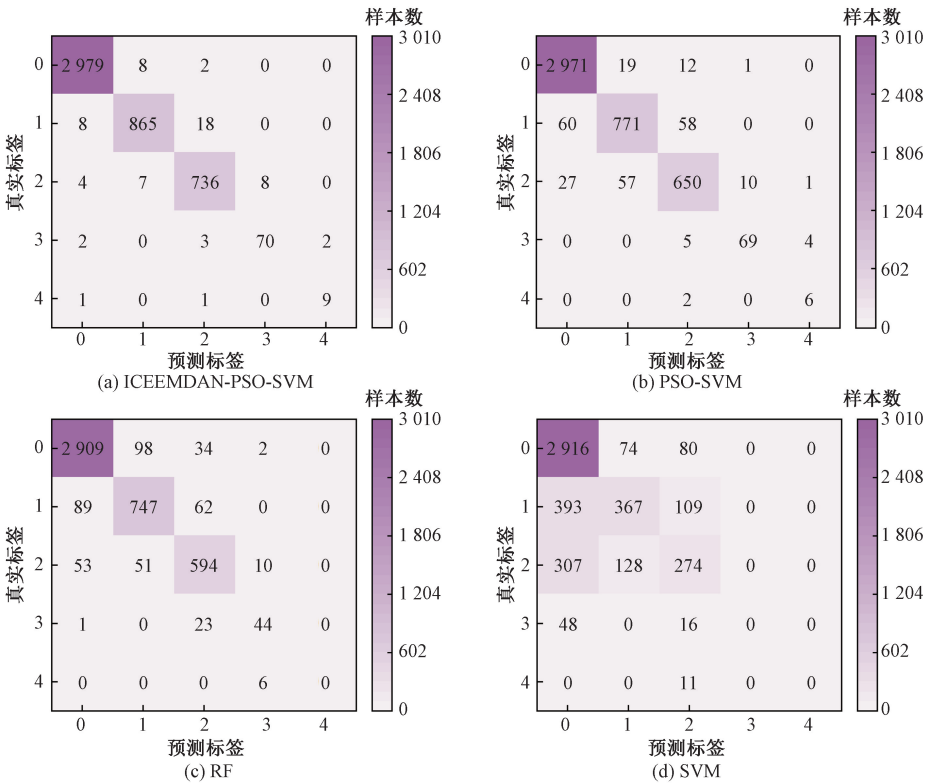


图 10 不同模型的测试集混淆矩阵(实验 1)

Fig. 10 Confusion matrices of test sets for different models (Experiment 1)

4 种模型的精确率、召回率、 F_1 -Score、虚警率和漏警率,以评估其对不同漏失情况的分类诊断能力,结果如图 11 所示。ICEEMDAN-PSO-SVM 在轻微渗漏、中等渗漏和严重渗漏(标签 1 ~ 标签 3)的分类中,精确率、召回率和 F_1 -Score 均超过 94%,虚警率和漏警率均低于 6%。对于完全漏失(标签 4),其精确率和召回率分别为 81.82% 和 75.00%,尽管虚警率和漏警率相对较高,但整体性能仍然最佳。PSO-SVM 和 RF 模型在标签 0 和 1 的分类上表现较好,但在标签 3 和标签 4 上的精确率和召回率较低,

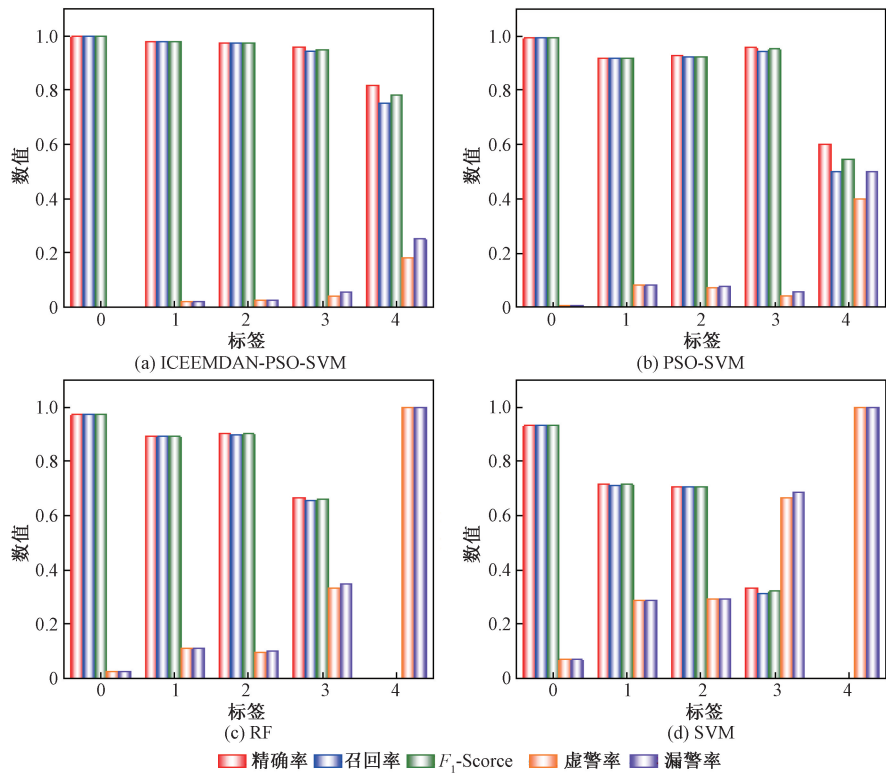


图 11 不同模型的分类诊断性能(实验 1)

Fig. 11 The classification and diagnosis performance of different models (Experiment 1)

虚警率和漏警率相对较高。SVM 仅在标签 0 上表现良好,而在标签 1~标签 4 上的分类效果不理想。上述实验结果表明,研究提出的 ICEEMDAN-PSO-SVM 模型在井漏诊断识别中展现出更优的分类性能。

3.3.3 实验 2:不同数据处理的模型性能对比

为了进一步验证 ICEEMDAN 对综合录井数据去噪效果的优势性,研究对比了 EMD、EEMD、CEEMDAN 和 ICEEMDAN 分解算法对 PSO-SVM 模型性能的影响。不同分解算法的设置信息如表 4 所示。不同数据处理的模型在测试集上的混淆矩阵,如图 12 所示。4 种模型测试集的准确率均可保持在 96%,展示出了较好的性能。然而,评估每个模型对不同漏失情况的分类诊断能力,如图 13 所示。基于 EMD 的模型对于标签 1、标签 2 和标签 4 的精确率和召回率仅维持在 71% 左右,虚警率和漏警率最高可达 28.70%。而随着分解算法的优化,基于 CEEMDAN 的模型对于标签 0~标签 3 的精确率和召回率均可达到 91% 以上,除了标签 4 的性能指标偏低,模型的整体性能接近 ICEEMDAN-PSO-SVM。实验结果进一步表明,不同的数据处理方法有利于改善数据集质量,进而提升模型的诊断性能。

表 4 不同分解算法的设置信息

Table 4 Setting information for different decomposition calculations

分解算法	参数设置	分解层数
ICEEMDAN	$N_{std} = 0.2, N_E = 100, \text{Max}_{iter} = 1\ 000$	8
CEEMDAN	$N_{std} = 0.2, N_E = 100, \text{Max}_{iter} = 1\ 000$	11
EEMD	$N_{std} = 0.2, N_E = 100$	9
EMD	—	9

4 结论

为了准确识别钻井过程中的不同程度漏失的发生,提出了一种融合录井时序数据与多尺度分解的井漏智能诊断模型,得出如下结论。

(1) 采用 ICEEMDAN 算法结合多指标 IMF 筛选准则能有效去除综合录井数据中的噪声干扰。实验证明,重构后的数据 SNR 可高达 15.80 dB, RMSE 可低至 2.53,显示出较理想的去噪效果。

(2) 研究提出的 ICEEMDAN-PSO-SVM 模型准确率可高达 98% 以上,相比于 PSO-SVM、RF 和 SVM 模型,在精确率、召回率、 F_1 -Score、虚警率和漏警率性能指标上均有明显的提升,展示出了良好的井漏诊断性能。

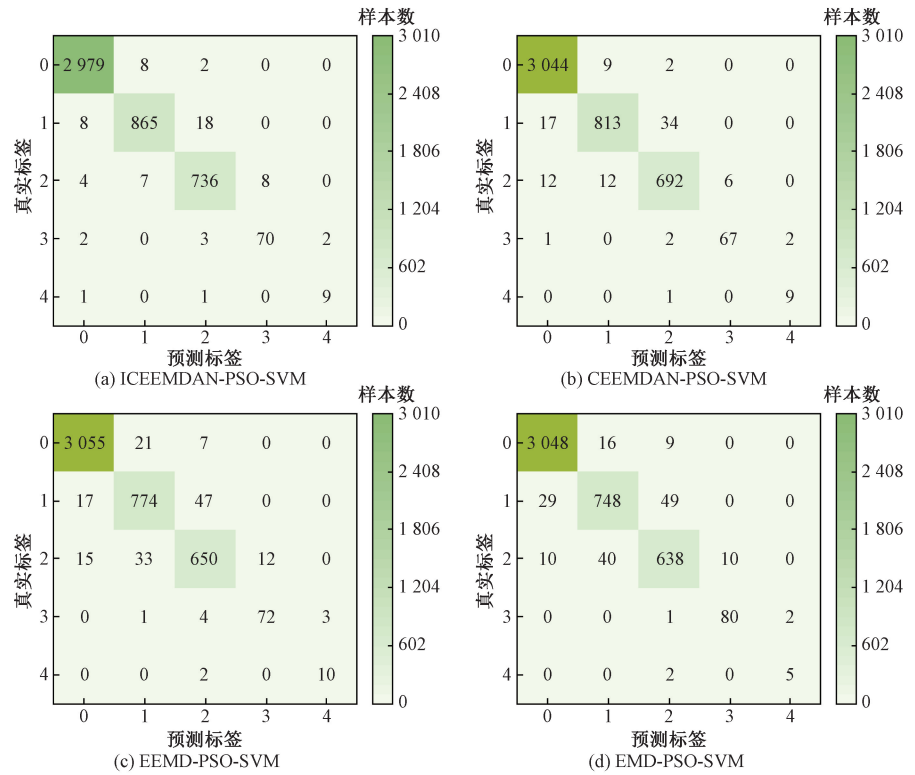


图 12 不同模型的测试集混淆矩阵(实验 2)

Fig. 12 Confusion matrices of test sets for different models (Experiment 2)

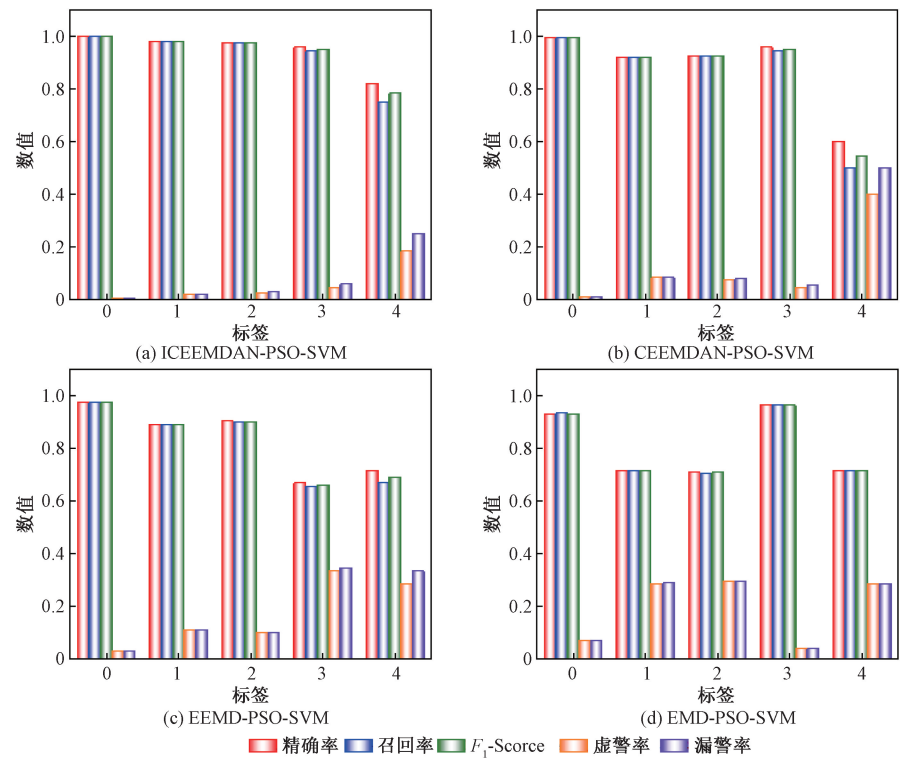


图 13 不同模型的分类诊断性能(实验 2)

Fig. 13 The classification and diagnosis performance of different models (Experiment 2)

(3) 研究进一步表明,优质的数据处理方法有助于提高数据集的质量,从而提升井漏诊断的准确

率。数据处理与模型优化并非相互排斥,而是相互促进,共同增强模型的整体性能。

参 考 文 献

- [1] Zhang R H, Su N Y, Liao L M, et al. Effects of different thermal insulated drill pipe deployment methods on wellbore temperature control in ultra-deep wells[J]. *Petroleum Science*, 2025, 22(10): 4174-4194.
- [2] 高阳东, 赖维成, 张新涛, 等. 中国海油油气勘探新进展与下一步勘探方向[J]. *中国海上油气*, 2024, 36(3): 1-10.
Gao Yangdong, Lai Weicheng, Zhang Xintao, et al. New progress and future directions of CNOOC oil and gas exploration[J]. *China Offshore Oil and Gas*, 2024, 36(3): 1-10.
- [3] 陈卫婷, 朱忠喜, 吴德胜, 等. 基于瞬态压力波邻域差值计算的井漏识别方法[J]. *科学技术与工程*, 2025, 25(34): 14649-14661.
Chen Weiting, Zhu Zhongxi, Wu Desheng, et al. Well leakage location method based on transient pressure wave neighbourhood difference calculation[J]. *Science Technology and Engineering*, 2025, 25(34): 14649-14661.
- [4] Ao Y, Zhu Z X, Zhang N, et al. Solution and analysis of wellbore temperature and pressure field coupling model under lost circulation[J]. *ACS Omega*, 2022, 7(32): 28675-28684.
- [5] Wang N, Wang J, Sun B, et al. Study of transient responses in the APWD measurements during gas influx[J]. *Journal of Natural Gas Science and Engineering*, 2016, 35: 522-531.
- [6] 蔺研锋, 闵超, 代博仁, 等. 基于动态特征和深度神经网络的钻井井漏失事故预测[J]. *西安石油大学学报(自然科学版)*, 2022, 37(3): 64-69.
Lin Yanfeng, Min Chao, Dai Boren, et al. Prediction of drilling leakage accident based on dynamic features and deep neural network[J]. *Journal of Xi'an Shiyu University(Natural Science Edition)*, 2022, 37(3): 64-69.
- [7] 张矿生, 宫臣兴, 陆红军, 等. 基于集成学习的井漏智能预警模型及智能推理方法[J]. *石油钻采工艺*, 2023, 45(1): 47-54.
Zhang Kuangsheng, Gong Chenxing, Lu Hongjun, et al. Intelligent early warning model and intelligent reasoning method based on integrated learning for lost circulation[J]. *Oil Drilling & Production Technology*, 2023, 45(1): 47-54.
- [8] 段世明, 宋先知, 姚学喆, 等. 压力与规则双约束下的溢流智能预警模型[J]. *石油机械*, 2025, 53(1): 1-9.
Duan Shiming, Song Xianzhi, Yao Xuezhe, et al. Intelligent warning model for overflow under constraints of pressure and rules[J]. *China Petroleum Machinery*, 2025, 53(1): 1-9.
- [9] 杨苗, 任玲, 王双, 等. 基于 CEEMDAN-小波阈值的电容式棉种监测信号降噪方法[J]. *农业机械学报*, 2025, 56(5): 71-81.
Yang Miao, Ren Ling, Wang Shuang, et al. Noise reduction method of capacitive cotton seed monitoring signal based on CEEMDAN-wavelet threshold[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2025, 56(5): 71-81.
- [10] 李凯阳, 韩团军, 刘庆江. 表面肌电信号熵特征和改进 PSO-SVM 的手势识别方法[J]. *科学技术与工程*, 2025, 25(35): 15155-15164.
Li Kaiyang, Han Tuanjun, Liu Qingjiang. Gesture recognition method based on surface electromyography (sEMG) entropy features and improved PSO-SVM[J]. *Science Technology and Engineering*, 2025, 25(35): 15155-15164.
- [11] 邢昭, 孟小军, 袁晶晶, 等. 信息量支持下 SVM-GBDT 模型的滑坡危险性评价[J]. *科学技术与工程*, 2025, 25(7): 2712-2720.
Xing Zhao, Meng Xiaojun, Yuan Jingjing, et al. Landslides hazard assessment based on SVM-GBDT and gradient boosting decision tree model supported by information quantity[J]. *Science Technology and Engineering*, 2025, 25(7): 2712-2720.
- [12] Deng S, Pei C Y, Yan X P, et al. Lost circulation prediction method based on an improved fruit fly algorithm for support vector machine optimization[J]. *ACS Omega*, 2023, 8(36): 32838-32847.
- [13] 白凯, 戴升升, 张照硕, 等. 群智能算法优化改进随机森林算法的井漏预测[J]. *现代电子技术*, 2025, 48(14): 159-168.
Bai Kai, Dai Shengsheng, Zhang Zhaoshuo, et al. Lost circulation prediction based on improved random forest algorithm optimized by swarm intelligence algorithm[J]. *Modern Electronics Technique*, 2025, 48(14): 159-168.
- [14] 孙金声, 杨景斌, 白英睿, 等. 深层超深层钻井液技术研究进展与展望[J]. *石油勘探与开发*, 2024, 51(4): 889-898.
Sun Jinsheng, Yang Jingbin, Bai Yingrui, et al. Research progress and development of deep and ultra-deep drilling fluid technology[J]. *Petroleum Exploration and Development*, 2024, 51(4): 889-898.
- [15] 林英松, 蒋金宝, 秦涛. 井漏处理技术的研究及发展[J]. *断块油气田*, 2005(2): 4-7, 89.
Lin Yingsong, Jiang Jinbao, Qin Tao. The development of well loss processing technology[J]. *Fault-Block Oil & Gas Field*, 2005(2): 4-7, 89.
- [16] 刘颖彪, 徐生江, 田龙, 等. 钻井复杂事故预警技术研究进展[J]. *新疆石油天然气*, 2025, 21(2): 15-23.
Liu Yingbiao, Xu Shengjiang, Tian Long, et al. Research progress on early warning technology of drilling complex accidents[J]. *Xinjiang Oil & Gas*, 2025, 21(2): 15-23.
- [17] Mardanirad S, Wood D A, Zakeri H. The application of deep learning algorithms to classify subsurface drilling lost circulation severity in large oil field datasets[J]. *SN Applied Sciences*, 2021, 3(9): 785.
- [18] 李恒丰, 李琳, 赵志峰. 井漏预测模型的超参数优化与解释: 以伊朗马伦油田为例[J]. *科学技术与工程*, 2024, 24(35): 15006-15013.
Li Hengfeng, Li Lin, Zhao Zhifeng. Drilling fluid lost-circulation prediction model hyperparameter optimization and interpretation: a case study of the Marun Field in Iran[J]. *Science Technology and Engineering*, 2024, 24(35): 15006-15013.
- [19] 李思奇, 夏卯, 鲁思兆, 等. 基于二次分解时频图和 SE-DSMC-BSA 的轻量化有载分接开关机械故障识别方法[J]. *振动与冲击*, 2025, 44(11): 268-279, 308.
Li Siqi, Xia Mao, Lu Sizhao, et al. Lightweight OLTC mechanical fault recognition method based on secondary decomposition time-frequency maps and SE-DSMC-BSA[J]. *Journal of Vibration and Shock*, 2025, 44(11): 268-279, 308.
- [20] 陈永高, 钟振宇. 基于改进 EEMD 算法的桥梁结构响应信号模态分解研究[J]. *振动与冲击*, 2019, 38(10): 23-30.
Chen Yonggao, Zhong Zhenyu. Research on modal modal decom-

- position of response signals for a bridge structure based on the improved EEMD [J]. *Journal of Vibration and Shock*, 2019, 38 (10): 23-30.
- [21] 张龙, 宋成洋, 邹友军, 等. 基于 VMD 多特征融合与 PSO-SVM 的滚动轴承故障诊断[J]. *机械设计与研究*, 2019, 35 (6): 96-104.
- Zhang Long, Song Chengyang, Zou Youjun, et al. A rolling bearing fault diagnosis method based on VMD multi-feature fusion and PSO-SVM [J]. *Machine Design & Research*, 2019, 35 (6): 96-104.
- [22] Vladimir N. Vapnik. *The nature of statistical learning theory* [M]. New York: Springer, 1995.
- [23] Zhu Z X, Liu H, Wu D S, et al. A diagnosis method for loss of circulation based on transient-pressure wave analysis and particle swarm optimization[J]. *Physics of Fluids*, 2024, 36(7): 077121.
- [24] 盛武, 樊斌斌. EMD-GRU 组合模型在煤矿瓦斯体积分数预测中的应用[J]. *安全与环境学报*, 2025, 25(4): 1339-1348.
- Sheng Wu, Fan Binbin. Application of the EMD-GRU combined model for predicting gas volume fraction in coal mines [J]. *Journal of Safety and Environment*, 2025, 25(4): 1339-1348.
- [25] 陈凯枫, 杨学文, 宋先知, 等. 基于工程录井数据的井漏智能诊断方法[J]. *石油机械*, 2022, 50(11): 16-22.
- Chen Kaifeng, Yang Xuwen, Song Xianzhi, et al. An intelligent diagnosis method for lost circulation based on engineering logging data [J]. *China Petroleum Machinery*, 2022, 50(11): 16-22.
- [26] 刘浩男, 文晓涛, 何健, 等. 基于随机森林算法的 AVO 类型判别[J]. *中国海上油气*, 2020, 32(5): 73-81.
- Liu Haonan, Wen Xiaotao, He Jian, et al. AVO type discrimination based on random forest algorithm [J]. *China Offshore Oil and Gas*, 2020, 32(5): 73-81.