

基于声发射参数的随机森林模型在阀门内漏检测的应用

朱治衡

重庆科技大学石油与天然气工程学院, 重庆

收稿日期: 2025年5月9日; 录用日期: 2025年6月19日; 发布日期: 2025年6月27日

摘要

针对天然气站场阀门内漏检测效率低、环境干扰强的问题等问题, 本研究提出基于声发射参数与随机森林算法的阀门内漏检测方法。通过构建气体阀门内漏模拟实验平台, 在0.3~0.8 MPa压力范围内采集DN100闸阀的声发射信号, 采用小波包分解对18.75~75.0 kHz特征频段进行降噪处理, 提取压力、幅度、能量、RMS、ASL等5个特征参数构建特征矩阵。通过随机森林分类模型对阀门正常关闭(标签-1)与内漏状态(标签1)进行分类建模, 结果表明: 在测试集上, 模型准确率可达93.33%, 其中阀门压差与能量参数对阀门内漏判别具有较高贡献度。该方法有效解决了传统检测方法对微小泄漏响应迟滞的问题, 为工业阀门状态监测提供了新的技术路径。

关键词

声发射检测, 随机森林, 阀门内漏检测

The Application of Random Forest Model Based on Acoustic Emission Parameters in Valve Internal Leakage Detection

Zhiheng Zhu

College of Petroleum and Natural Gas Engineering, Chongqing University of Science and Technology, Chongqing

Received: May 9th, 2025; accepted: Jun. 19th, 2025; published: Jun. 27th, 2025

Abstract

Aiming at the problems of low detection efficiency and strong environmental interference of valve

internal leakage in natural gas station, this study proposes a valve internal leakage detection method based on acoustic emission parameters and random forest algorithm. By constructing a gas valve internal leakage simulation experiment platform, the acoustic emission signals of DN100 gate valve were collected in the pressure range of 0.3~0.8 MPa. The wavelet packet decomposition was used to denoise the characteristic frequency band of 18.75~75.0 kHz, and five characteristic parameters such as pressure, amplitude, energy, RMS and ASL were extracted to construct the characteristic matrix. The random forest classification model is used to classify and model the valve normal closing (label-1) and internal leakage state (label 1). The results show that the accuracy of the model can reach 93.33% on the test set, and the valve pressure difference and energy parameters have a high contribution to the valve internal leakage discrimination. This method effectively solves the problem of hysteresis response of traditional detection methods to small leakage, and provides a new technical path for industrial valve condition monitoring.

Keywords

Acoustic Emission Detection, Random Forest, Valve Internal Leakage Detection

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

阀门在制造业和工程建设中扮演着至关重要的角色。然而，阀门密封性的优劣直接关系到生产系统的安全性和有效操作。阀门泄漏可能带来多种严重后果，涵盖了环境污染、火灾、爆炸、中毒事件，甚至危及人员的生命安全[1]。这不仅对环境产生负面影响，还可能引发法律诉讼和巨大的经济损失。因此，进行有效的阀门泄漏检测研究对于确保安全性和可持续性发展至关重要。

近年来，声发射检测作为一种新兴的无损检测技术，在工程和工业领域得到广泛应用。其具有诸如准确、快速、无损、经济等特点，使其备受青睐[2]。该技术利用物体产生的声波信号来检测材料的内部状态，特别适用于阀门内漏的检测。声发射技术不仅能够识别微小声音变化，还能提供有关泄漏源的位置和程度的信息。与传统的检测方法相比，它具有更高的准确性和灵敏度，能够快速捕捉潜在问题，从而降低漏报率和误报率。

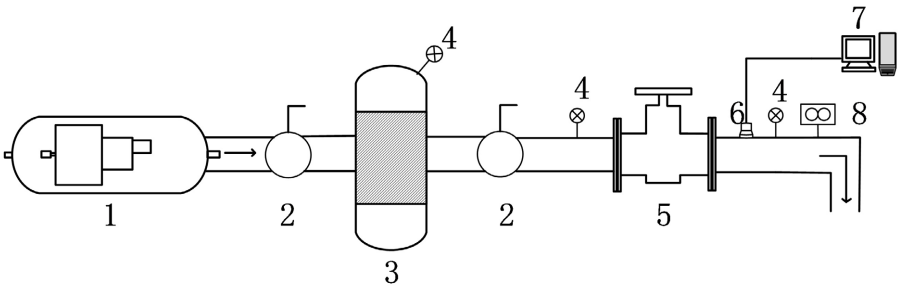
在当前阀门泄漏声发射信号处理领域，主要采用两种方法进行研究。一种是基于回归建模的方法，另一种是基于分类建模的方法。回归建模方法的目标在于建立阀门泄漏声发射信号特征量与泄漏量之间的定量关系模型。这一方法通过对声发射信号的特征参数进行分析，如频率分布、能量分布、时域波形等，结合泄漏量的实际测量数据，利用回归分析建立起声发射信号特征与泄漏量的数学模型，该模型通过分析声发射信号的特征参数，确定其与泄漏量之间的定量关系，以实现阀门泄漏情况的准确量化和预测。国内外学者围绕阀门泄漏声发射检测的核心问题——微弱信号分离、多参数关联建模及工业场景适应性展开持续探索。Seong 团队[3]将声发射频域特征与止回阀泄漏状态关联，构建了基于人工神经网络的泄漏识别框架。戴光等[4]通过 Lighthill 气动声学模型揭示了闸阀泄漏声场分布规律，提出扩压腔区域为最佳检测位点，为实验传感器布设提供了理论支撑。随着机器学习驱动的多特征融合成为研究热点。国外 Kaewwaewnoi 等[5]研究人员提出了基于声发射源几何分布关系的阀门液体泄漏速率预测模型。他们在研究过程中推导出了信号均方根(RMS)与泄漏率 Q 之间的关系，为建立阀门泄漏故障特征库以及进行阀门泄漏率的定量诊断提供了理论基础。Meland [6]提出小波包 - 峭度联合分析法，在宽频信号中提取微泄漏特征。曹国梁[7]基于小波包能量比例特征构建支持向量回归(SVR)模型，实现二分类 90%准确率。

Hoi Yin Sim 等研究人员[8]通过离散小波变换对声发射信号进行时频分析,采用 k 近邻算法(KNN)和支持向量机(SVM)分类算法对阀门状态进行分类。这一研究进一步证明了智能诊断方法在阀门泄漏检测中的有效性。Ye 团队[9]设计了一套压力工况模拟实验系统,通过在 0.4~0.8 MPa 动态压力区间内,分别对通径 40 mm 和 50 mm 的液控阀门进行泄漏状态模拟,采集不同工况下的声发射信号。基于时频联合分析技术提取信号特征参数,构建多特征融合的随机森林分类模型,实现了阀门密封失效状态的有效识别。这一研究为声发射技术在泄漏检测中的应用提供了新的思路。吴凯文[10]通过网格搜索算法对支持向量机进行超参数优化,构建了高精度分类模型,其结果表明阀门泄漏预测模型识别准确率在 93%以上。Aniyom 和 Chikwe [11]提出基于随机森林 + XGBoost 的混合机器学习模型的天然气管道泄漏检测方法,结合温度、压力和流量参数实现泄漏识别准确率达 93%。

2. 阀门内漏声发射信号采集实验

2.1. 实验装置

气体阀门内漏模拟实验平台能够复现工业环境中常见阀门的内部泄漏现象。通过这一设备,模拟各类阀门在实际使用过程中出现的密封失效情况。实验平台示意图如图 1 所示,主要由空气压缩机、控制球阀、稳压罐、压力表、待测阀门 DN100 闸阀、声发射传感器、声发射采集系统和流量计组成。



1. 压缩机; 2. 控制球阀; 3. 稳压罐; 4. 压力表; 5. 待测阀门组; 6. 传感器; 7. 声发射采集系统; 8. 流量计。

Figure 1. Valve internal leakage data acquisition test platform
图 1. 阀门内漏数据采集试验平台

声发射检测系统参数设置如表 1 所示。

Table 1. Acoustic emission detection system parameter setting
表 1. 声发射检测系统参数设置

采集模式	包络采集
通道	单通道
采样门限	35 db
内置放大器增益	40 db
采样率	600 kHz
波形采集	打开
采样参数	时间、振幅、能量、rms、asl 等

实验前,首先对环境噪声进行采集,根据环境噪声的等级确定声发射实验的门限值,然后在 0.3 Mpa、0.4 Mpa、0.5 Mpa、0.6 Mpa、0.7 Mpa、0.8 Mpa 下对 DN100 的闸阀提取不同泄漏率下声发射信号值。为了确认声发射信号所在的频段,需要在保持声发射传感器与阀门相对位置不变的情况下,仅通过控制阀

门的开度来采集无泄漏情况和发生泄漏情况的声发射信号。

先关闭待测阀门，打开控制球阀，调节阀门进口的压强，当稳压罐上的压力表稳定在 0.3 MPa 时，开始实验。首先记录阀门关闭状态下的信号，接着逐渐打开待测阀门，观察流量计的数值，找到传感器可检测的最小内漏率并记录。这组实验完成后，改变阀门进口的压力，使其分别为 0.4 Mpa、0.5 Mpa、0.6 Mpa、0.7 Mpa、0.8 Mpa 重复上述实验。

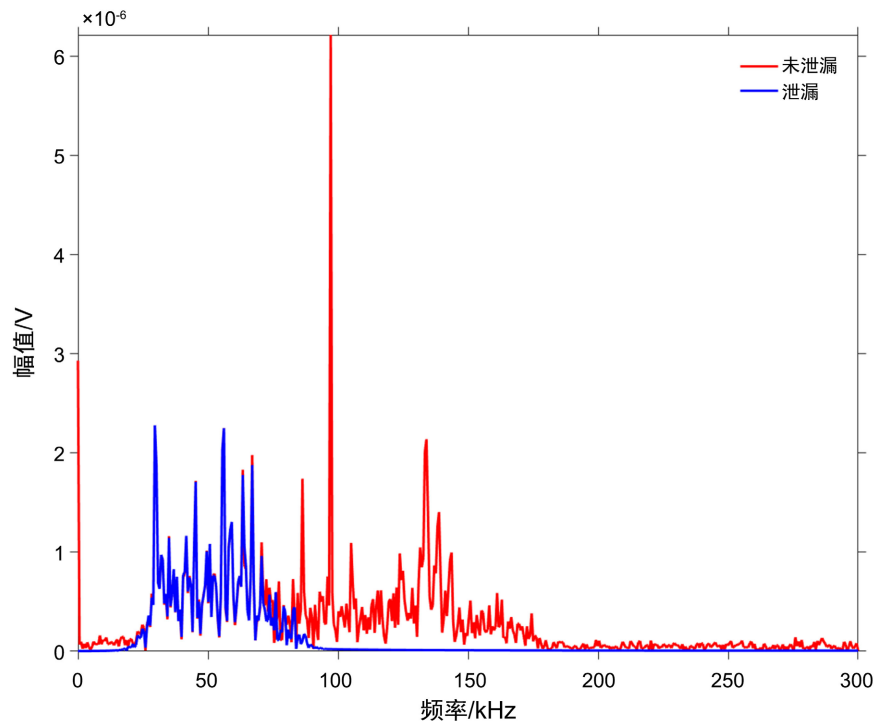
此外，为降低背景噪声的干扰，在传感器和管线之间涂抹了适量厚度的耦合剂，并采用磁性夹具固定。40 d 前置放大器用于放大感测的阀门内漏信号，以降低信号传输过程中的噪声影响，且信号增益变动在 1 dB 以内。数据采集器对经过放大的信号进行 A/D 转换，并将转换后的数字信号传输到计算机系统。设置信号处理终端的采样频率为 600 kHz，这符合香农采样定理，其奈奎斯特频率为 300 kHz。同时，设置记录时长为 5 秒，且每个泄漏率采集 3 组数据。

2.2. 实验数据预处理及特征提取

由于声发射传感器采集的阀门内漏声发射信号具有随机信号的特点，还含有电机噪声和其它环境噪声，因此，除了在硬件上要对信号进行放大和抗混叠低通滤波之外，还要对转换成的数字信号进行预处理，从而确定信号的特征频带。

以 DN100 闸阀 0.5 Mpa 下的泄漏数据与未泄漏数据为例，对采集到的声发射信号进行频谱分析，如图 2 所示。当阀门未泄漏时，信号在 0~200 kHz 的频带内的幅值分布较为平均。当阀门发生泄漏时，泄漏频率在 18.75~75.0 kHz 区间段，随着泄漏量的增加，信号的幅值随着泄漏量的增加逐渐抬升。因此，在预处理时，保留该频段的小波包节点系数，将其他频段噪声信号的小波包节点系数降为 0，重构声发射信号。

原始泄漏信号与降噪后信号对比如图 3 所示，相比于未处理的图像，重构后的图像整体幅值有较大回落，这是因为环境噪声所在频段被滤波后，只保留了幅值较低的声发射信号。



(a) 阀门泄漏与未泄漏

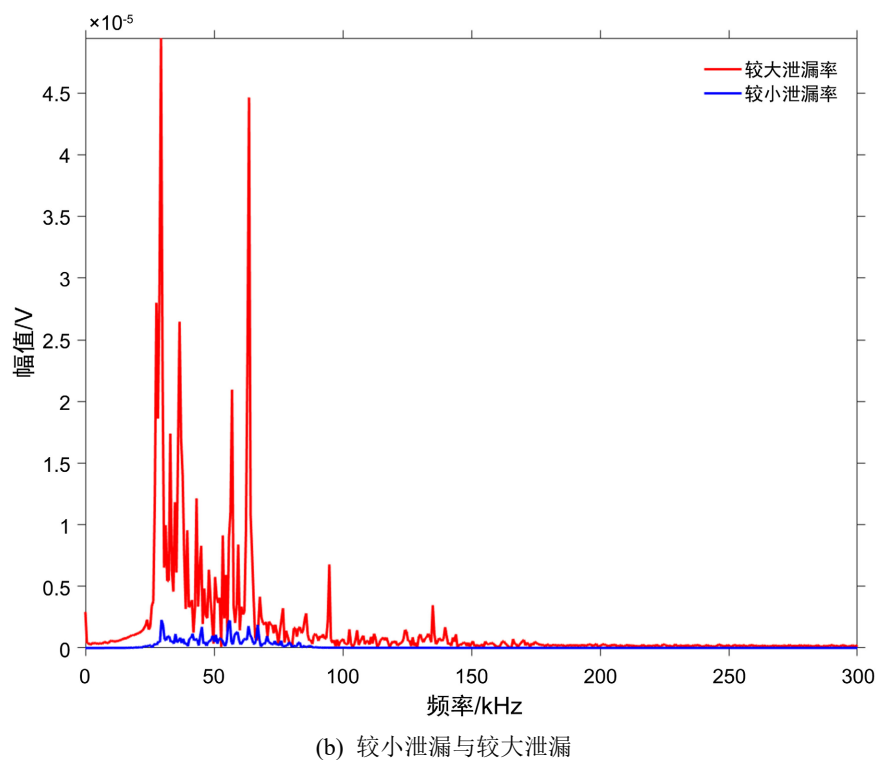


Figure 2. DN100 gate valve leakage acoustic emission signal spectrum diagram
图 2. DN100 闸阀泄漏声发射信号频谱图

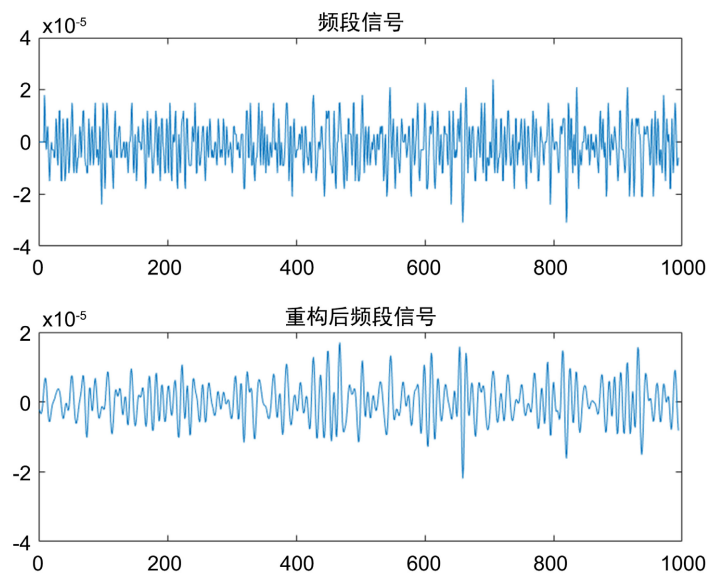


Figure 3. The comparison diagram of the original leakage signal and the signal after noise reduction reconstruction
图 3. 原始泄漏信号与降噪重构后信号对比图

3. 随机森林分类模型

3.1. 随机森林原理

随机森林(Random Forest)是一种强大且灵活的机器学习算法，常用于数据分类任务。它由多个决策

树组成，每个决策树都独立地对输入数据进行分类[12]。最终，通过对它们的结果进行平均处理，获得了更加可靠的结果，如图 4 所示。

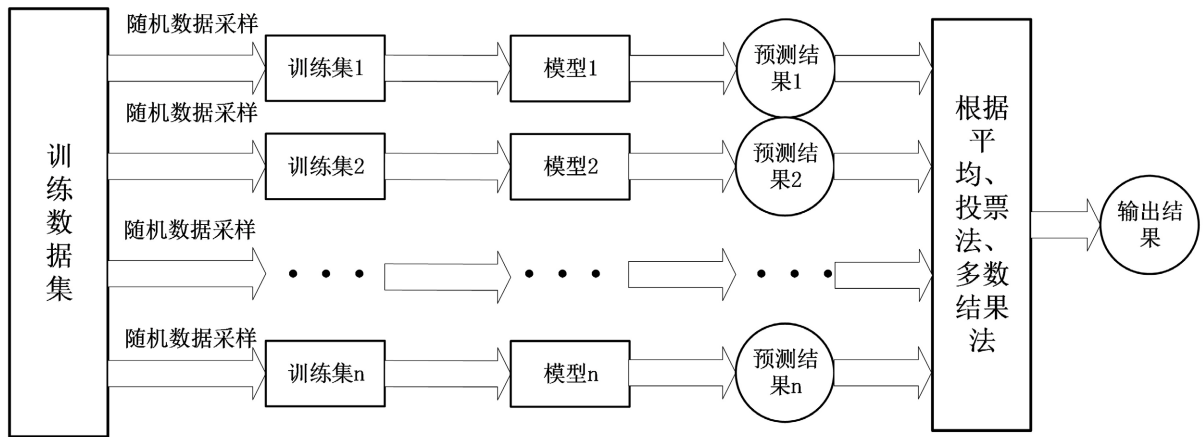


Figure 4. Random forest structure diagram
图 4. 随机森林结构图

随机森林不容易过拟合，这是因为它由多个决策树组成，每棵树都是基于随机子集和特征子集构建的。这种随机性减少了模型对训练数据的过度依赖，提高了模型的泛化能力。对于不平衡的数据集，随机森林可以平衡误差，即使某些类别的样本比其他类别更多或更少，它也能够有效地处理这种情况，不至于使得模型偏向于数量更多的类别[13]。此外，即使有很大一部分的特征遗失，随机森林仍可以维持较高的准确度。这是因为每棵树都是基于不同的随机子集构建的，因此即使部分特征丢失，仍然有足够的信息来进行分类。

3.2. 声发射参数选取

如表 2，构建特征矩阵，覆盖时域(RMS, ASL)、频域(能量、幅度)及工况(压差)多维信息，满足泄漏状态分类的完备性需求。

Table 2. Acoustic emission parameter selection
表 2. 声发射参数选取

参数名称	物理意义
压力	阀门上下游压差
幅度	声发射信号最大幅值，反映泄漏瞬态冲击强度
能量	信号能量积分值，表征泄漏过程持续释放的机械能总量
RMS	有效值电压均方根，表征信号有效功率
ASL	平均信号电平，反映信号基线稳定性，识别间歇性泄漏引起的电平漂移

为了得到每个泄漏量下声发射信号的 5 个特征量，并将它们组成随机森林算法的特征数据矩阵 x ，即：

$$\begin{cases} x_i = (p_{mpa}(i), d_{db}(i), e_{kpj}(i), r_{rms}(i), a_{asl}(i)) \\ x = [x_1, x_2, \dots, x_l] \end{cases}$$

式中, $p_{mpa}(i)$ 为阀门压力; $d_{db}(i)$ 为幅度; $e_{kpj}(i)$ 为能量; $r_{rms}(i)$ 为有效值电压均方根; $a_{asl}(i)$ 为平均信号电平, 分别为 x 的 5 个特征量的值。

3.3. 随机森林分类建模过程

(1) 添加标签

在分类标签方面, 标签-1 代表阀门正常关闭的状态, 即阀门的密封性能良好, 没有发生泄漏; 而标签 1 则代表阀门存在内漏现象, 即阀门在关闭状态下仍有介质泄漏。通过这样的标签设定, 随机森林分类模型能够对阀门的工作状态进行准确识别。

(2) 选择训练样本

为了划分样本集为测试样本集和训练样本集, 需要按照一定比例将样本集分开。这个过程包括随机打乱样本的顺序, 然后根据需要的比例(70%训练样本和 30%测试样本)将样本分配到不同的集合中。训练样本集的特征矩阵和对应的标签矩阵被用来训练随机森林分类模型。将测试样本集的特征矩阵输入到这个已训练好的随机森林模型中, 以便对测试样本集进行分类。

(3) 随机森林建模过程

在随机森林建模过程中, 首先需要进行数据处理, 以获取随机森林模型所需的输入特征矩阵和标签矩阵。处理过程包括以下步骤:

- 1) 对输入的特征矩阵进行归一化处理, 以确保各个特征的数值范围相似, 避免某些特征对模型的影响过大。
- 2) 将特征矩阵转置, 以适应随机森林模型的输入格式要求。
- 3) 设置随机森林模型的参数, 包括决策树的数目和最小叶子节点数等。
- 4) 评估模型在训练集上的性能表现。

4. 随机森林分类结果

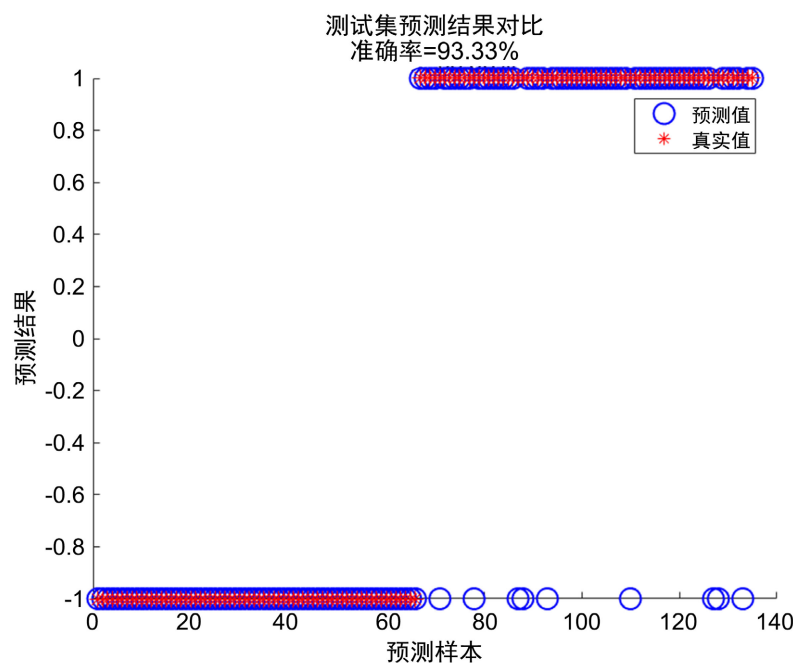


Figure 5. Random forest classification results

图 5. 随机森林分类结果

在压力 0.3 Mpa~0.8 Mpa 下, 采集了 DN100 闸阀 150 组阀门正常关闭与泄漏数据。随机森林模型经贝叶斯超参数优化后, 其决策树的数量 68; 每个节点划分时考虑的最大特征数为 3, 增加模型的随机性, 降低树之间的相关性; 叶节点最小样本数为 5。基于图 5 的结果, 可以看出构建的随机森林分类模型在测试集上表现出优异的性能, 其预测准确率为 93.33%, 提供了强有力的数据分类能力, 包括其对高维度数据的适应性、抗过拟合能力、对不平衡数据的平衡处理以及对特征缺失的鲁棒性。

特征 1~5 分别对应压力、幅度、能量、RMS、ASL 的特征量。对每个阀门的数据绘制其特征重要性。得出相同的结果, 在这五个特征值中, 影响阀门泄漏等级的主要特征是阀门前后端压差以及能量, 其中最重要的是阀门前后端压差, 如图 6 所示。

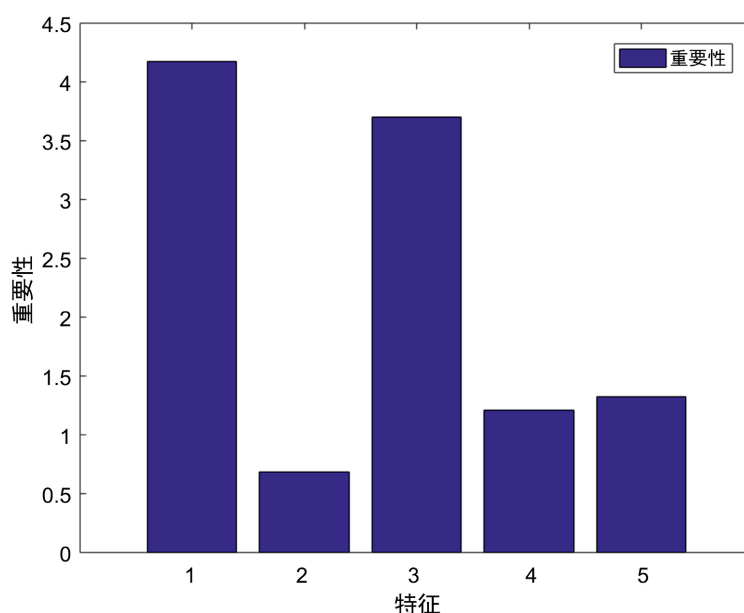


Figure 6. Importance of leak identification features

图 6. 泄漏识别特征的重要性

从特征贡献度来看, 阀门上下游压差与能量参数的累计重要性最大, 其物理机制可解释为: ① 压差直接影响流体通过阀座缝隙的流速与湍流强度, 进而决定声发射信号的主频分布; ② 能量参数反映泄漏持续时间的累积效应, 能够捕捉传统时域参数(如幅度)易忽略的微弱信号变化。

5. 结论

本研究通过构建随机森林分类模型, 实现了对 DN100 闸阀正常关闭与内漏状态的智能识别。基于小波包分解的 18.75~75.0 kHz 频段降噪方法可有效提取泄漏特征信号, 克服了环境噪声对微弱泄漏信号的干扰。在 0.3~0.8 MPa 压力工况下, 模型整体分类准确率达到 93.33%, 这一性能提升主要归因于随机森林算法对多参数非线性关系的有效捕捉, 以及小波包分解技术对特征频段的精准提取。本方法通过融合压力、能量等多维度特征, 提升了检测准确率, 验证了特征矩阵构建策略的有效性。特征重要性分析表明, 阀门压差与能量参数是泄漏判别的关键指标。

基金项目

重庆科技大学科技创新项目(YKJCX2320139)。

参考文献

- [1] 唐婷婷, 赵博, 刘晓迪, 等. 声发射检测技术在材料断裂损伤中的研究进展与展望[J]. 中国特种设备安全, 2024, 40(2): 1-7.
- [2] 杨磊, 谭开雨, 吴佳俊. 声发射无损检测技术在管道故障检测中的应用[J]. 内燃机与配件, 2023(13): 90-92.
- [3] Seong, S.H., Kim, J.S., Hur, S., *et al.* (2004) The Development of Fusion Sensor Techniques for Condition Monitoring of a Check Valve. *Key Engineering Materials*, **270**, 2220-2225.
- [4] 戴光, 王兵, 张颖, 等. 闸阀气体内漏喷流声场的数值模拟[J]. 流体机械, 2007, 35(3): 29-32.
- [5] Kaewwaewnoi, W., Prateepasen, A. and Kaewtrakulpong, P. (2010) Investigation of the Relationship between Internal Fluid Leakage through a Valve and the Acoustic Emission Generated from the Leakage. *Measurement*, **43**, 274-282. <https://doi.org/10.1016/j.measurement.2009.10.005>
- [6] Meland, E., Thornhill, N.F., Lunde, E. and Rasmussen, M. (2011) Quantification of Valve Leakage Rates. *AIChE Journal*, **58**, 1181-1193. <https://doi.org/10.1002/aic.12630>
- [7] 曹国梁. 阀门内漏状态识别与量化分析的声发射技术研究[D]: [硕士学位论文]. 东营: 中国石油大学(华东), 2016.
- [8] Sim, H.Y., Ramli, R., Saifizul, A. and Soong, M.F. (2020) Detection and Estimation of Valve Leakage Losses in Reciprocating Compressor Using Acoustic Emission Technique. *Measurement*, **152**, Article 107315. <https://doi.org/10.1016/j.measurement.2019.107315>
- [9] Ye, G., Xu, K. and Wu, W. (2021) Multi-Variable Classification Model for Valve Internal Leakage Based on Acoustic Emission Time-Frequency Domain Characteristics and Random Forest. *Review of Scientific Instruments*, **92**, Article 025108. <https://doi.org/10.1063/5.0024611>
- [10] 吴文凯, 徐科军, 叶国阳. 面向阀门内漏声发射检测的支持向量机分类建模[J]. 计量学报, 2021, 42(8): 1018-1025.
- [11] Aniyom, E. and Chikwe, A. (2025) Prediction of Leak on Gas Pipeline Using a Hybrid Machine Learning Model. *Improved Oil and Gas Recovery*, **9**, 1-11.
- [12] 马飞, 邵礼光, 徐君, 等. 基于小波包分解与随机森林的离心泵故障诊断[J]. 工程设计学报, 2024, 31(6): 741-749.
- [13] 吴冬, 阎卫东, 王井利. 基于特征重要性加权的随机森林点云分类研究[J]. 电子测量技术, 2023, 46(20): 120-127.