

基于拉曼光谱评估变压器油 - 屏障式绝缘老化状态

汪帮瑞, 薛建侠, 左小玉, 傅雪平, 孙一叶*

温州大学电气与电子工程学院, 浙江 温州

收稿日期: 2023年3月6日; 录用日期: 2023年5月30日; 发布日期: 2023年6月7日

摘要

变压器油 - 屏障式绝缘老化状态的准确评估对电力系统的安全可靠运行至关重要。本工作基于拉曼光谱检测平台及自行搭建的老化设备, 对45#绝缘油与牛皮纸组成的绝缘系统进行加速热老化实验, 依据对热老化样品的取样时间, 热老化样品分为6类, 采用支持向量机对原始数据进行分类, 并用网格法、粒子群算法、遗传算法优化支持向量机参数。由于样本数量多以及光谱维度高, 优化支持向量机参数时间较长, 为进一步提升分类准确率以及分类速度, 对原始光谱数据使用非对称重新加权惩罚最小二乘法、Savitzky-Golay平滑、主成分分析等预处理方法, 消除基线漂移与环境噪声的影响, 并且降低分类时间, 提高分类准确率, 该工作为变压器油 - 屏障式绝缘老化状态的有效评估提供技术支持和理论依据。

关键词

拉曼光谱, 油 - 屏障式绝缘, 基线校正, 支持向量机

Evaluation of Transformer Oil-Barrier Insulation Aging State Based on Raman Spectroscopy

Bangrui Wang, Jianxia Xue, Xiaoyu Zuo, Xueping Fu, Yiye Sun*

College of Electrical and Electronic Engineering, Wenzhou University, Wenzhou Zhejiang

Received: Mar. 6th, 2023; accepted: May 30th, 2023; published: Jun. 7th, 2023

Abstract

It is critical to accurate assessment of the aging state of transformer oil-barrier insulation for the

*通讯作者。

文章引用: 汪帮瑞, 薛建侠, 左小玉, 傅雪平, 孙一叶. 基于拉曼光谱评估变压器油-屏障式绝缘老化状态[J]. 电气工程, 2023, 11(2): 64-73. DOI: 10.12677/jee.2023.112008

safe and reliable operation of power systems. In this study, the accelerated thermal aging experiment was carried out on the insulation system composed of 45# insulating oil and kraft paper based on the Raman spectrum detection platform and self-built aging equipment. Based on the sampling time of thermal aging samples, the thermal aging samples were divided into 6 categories, and the support vector machine was used to classify the original data, and the parameters of the support vector machine were optimized by Gridsearch, Particle Swarm Optimization (PSO), and Genetic Algorithm (GA). It takes a long time to optimize the parameters of the support vector machine due to the large number of samples and the high spectral dimension. In order to further improve the classification accuracy and speed, asymmetric reweighting penalty least squares, SG smoothing, PCA and other preprocessing methods are used for the original spectral data, which eliminate the influence of baseline shift and environmental noise, and reduce the classification time, and improve the classification accuracy. The research paves a way to effectively evaluate the aging state of transformer oil-barrier insulation.

Keywords

Raman Spectroscopy, Oil-Barrier Insulation, Baseline Correction, Support Vector Machine

Copyright © 2023 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

变压器是输配电系统的主要设备,其状态对电力系统的安全性和可靠性至关重要。变压器发生事故的可能性随着绝缘的恶化而增加。油-屏障式绝缘是油浸式变压器的主要绝缘类型,变压器在长期运行中,由于热老化的不可逆消极影响[1],导致其油-屏障式绝缘系统的机械性能发生下降,影响它的安全稳定运行。因此,准确识别变压器油-屏障式绝缘的不同老化阶段,特别是对于运行时间越久的变压器,油-屏障式绝缘老化状态监测的研究对于变压器运行寿命的预测具有重要的意义。

目前,对于油-屏障式绝缘老化检测的方法有绝缘纸的聚合度(DP)检测,糠醛浓度检测,油中溶解气体分析(DGA),然而,这些方法都有各自的缺点。DP在采样和测量时需要切断电源并且需要对变压器进行吊芯取样,因此,DP测量在现场应用仍有诸多不便。对于糠醛浓度检测,有高效液相色谱(HPLC)、紫外(UV)分光光度法和比色法等方法。HPLC有操作复杂、洗脱困难等缺点;紫外分光光度法稳定性差,易受变压器油中除糠醛外的其他有机物的影响;由于比色法需要目视比色管,这导致比色法精度较低,易受不同操作人员的影响。溶解气体分析(DGA)常使用气相色谱(GC)来进行,然而,色谱柱的性能会随着时间的推移而下降,GC监测系统需要由专业人员在实验室中操作。

随着激光技术和电荷耦合元件检测技术的发展,拉曼光谱已广泛应用于食品检测[2]、材料分析[3]等领域的定性或定量分析。近些年来,有学者使用拉曼光谱进行糠醛浓度检测,检测限达到 14.4 mg/L [4],表明拉曼光谱在变压器的早期故障诊断中显示出一定的潜力。油-屏障式绝缘是变压器的重要组成部分,根据绝缘油的型号不同,可分为 10#、25#、45#三种,其中 45#油-屏障式绝缘常用于低温、超高压环境,然而,很少有报道将拉曼光谱应用于 45#绝缘油与绝缘纸组成的绝缘类型以间接评估变压器的老化状况。因此,本文研究的主要目的是探究 45#油-屏障式绝缘的拉曼光谱特性,并探索其应用于油浸式变压器老化阶段评估的可行性。

在本文研究中,使用自行搭建的热老化设备,制备不同老化油样样品,使用拉曼光谱仪测得各个时

间段老化样品的拉曼谱线。根据不同的老化时间，模拟不同的老化程度，对不同老化程度的样品数据进行分类，探究热老化对变压器油 - 屏障式绝缘类型的影响。

2. 实验部分

2.1. 拉曼光谱检测装置

本文根据激光拉曼原理搭建了如图 1 所示的检测油样实验装置。实验时，785 nm 的激光器发射的激光经探头穿过盛放样品的压盖玻璃瓶，照射在油样表面，并发生拉曼散射，散射光经光谱仪采集后通过信号传输线将采集的数据传给微型计算机，进而完成对油样的检测。该实验装置采用厦门奥谱天成光电有限公司生产的拉曼光谱仪，型号为 ATR3110-785。为了提高信噪比，减小宇宙射线与荧光的干扰，获得稳定的油样拉曼光谱图，将积分时间设置为 5 s，激光功率设置为 100 MW，平均次数设置为 1 次。

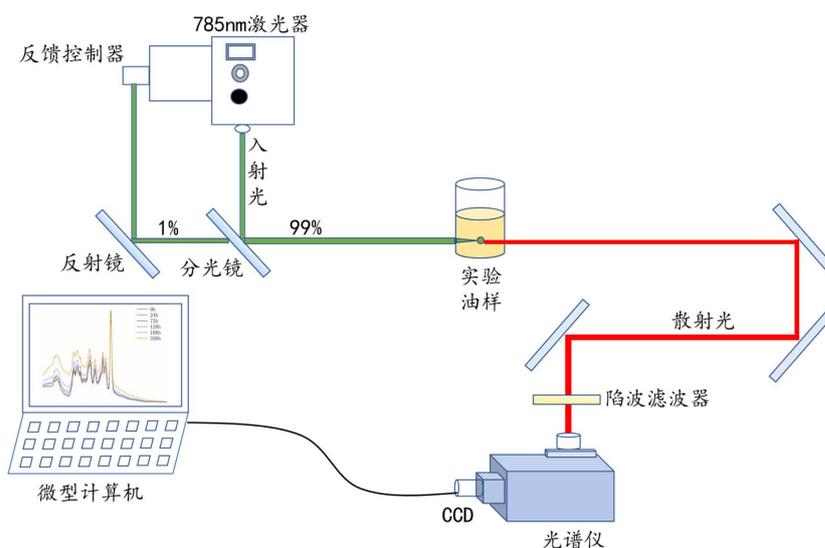


Figure 1. Diagram of oil sample testing device
图 1. 油样检测装置图

2.2. 油 - 屏障式绝缘老化样品的制备

本采用 45 号矿物绝缘油与 0.2 mm 厚的牛皮绝缘纸，先将绝缘油在 80℃ 下干燥脱水 72 h，再将绝缘纸裁剪成碎末状后干燥 48 h，绝缘油与绝缘纸干燥完成后，将绝缘纸与绝缘油以 1:8 质量比混合装瓶，在 40℃ 下充分混合 24 h，然后开始热老化加速实验。

本文实验温度采用 130℃，不会高于 45 号绝缘油的闪点温度(135℃)保证实验操作的安全性，同时，又可以确保油 - 屏障式绝缘样品的快速热老化。避免了温度设置过低，增加了热老化时间，达不到加速热老化效果；也避免了温度过高，会破坏绝缘油内部的分子结构，使得检测结果不具有真实性的缺点。

本文加速热老化样品制备设备采用鼎鑫宜公司生产的数显恒温油浴锅，设备型号 DXY-12H，该设备广泛用于石油、化工领域，主要用于化学物质的蒸馏、干燥、浓缩以及浸渍化学药品或生物制品。油浴锅有箱主体、油锅、锅盖和温度控制系统等组成，其加温件采用优质不锈钢管式电加热管，直接浸入油浴锅进行热交换，加热效率高、热能损耗小，温度控制采用“工业及专家自整定 PID 技术”控制，与传统的 PID 控制方式相比具有更小的温度超调、更快的稳定时间、更好的控温精度等优点。控温仪还具有定时控制、控温误差修正、偏差报警保护等功能。油浴锅具体参数如表 1 所示。

Table 1. Specific parameters of the oil bath**表 1.** 油浴锅具体参数

型号	电压	功率	温度
DXY-12H	220 V/50 Hz	1200 W	RT + 5°C~300°C

3. 光谱预处理及其分类算法

3.1. Savitzky-Golay 平滑

Savitzky-Golay 滤波是由一种在时域内基于局域多项式最小二乘法拟合的滤波方法[5]。其最大的特点在于在滤除噪声的同时可以确保信号的形状、宽度不变。Savitzky-Golay 原理如下：

设有 $n = 2m + 1$ 个等间距量测点 $x_i (i = -m, -m + 1, \dots, 0, \dots, m - 1, m)$ ，在滤波过程中，令 $x_i = i (i = -m, -m + 1, \dots, 0, \dots, m - 1, m)$ ， n 为滤波窗口的宽度，假设窗口内的数据点可采用 $k - 1$ 次多项式进行拟合。即

$$y_i = a_0 + a_1 i + a_2 i^2 + \dots + a_{k-1} i^{k-1} \quad (1)$$

于是就有了 n 个这样的方程，构成 k 元线性方程组，写成矩阵形式为：

$$Y_{(2m+1) \times 1} = X_{(2m+1) \times k} \times A_{k \times 1} + E_{(2m+1) \times 1} \quad (2)$$

要使上述矩阵有解，必须满足 $2m + 1 > k$ ，系数 A 的最小二乘解为 $\hat{A} = (X^T \cdot X)^{-1} \cdot X^T \cdot Y$ ，滤波完成后的 $\hat{Y} = X \cdot (X^T \cdot X)^{-1} \cdot X^T \cdot Y = B \cdot Y$ ，其中 $B = X \cdot (X^T \cdot X)^{-1} \cdot X^T$ ，称为滤波系数， Y 为原始光谱， \hat{Y} 为经 SG 平滑后的光谱， X 只由滤波带宽 $(2m + 1)$ 和多项式次幂 $(k - 1)$ 决定。

3.2. 非对称重新加权惩罚最小二乘法

为了提高后续数据处理的效果，需要对平滑后的原始光谱数据进行基线校正。本文使用非对称重新加权惩罚最小二乘法(arPLS)来进行基线校正，arPLS 是由韩国全南大学 SungJuneBaek 等人提出的一种基线校正算法[6]，对权重更新的表达式如下：

$$w_i = \begin{cases} \text{logistic}(y_i - z_i, m_{d^-}, \sigma_{d^-}) & y_i \geq z_i \\ 1, & y_i \leq z_i \end{cases} \quad (3)$$

$$\text{logistic}(d, m, \sigma) = 1 / \left\{ 1 + \exp \left[2(d - (-m + 2\sigma)) / \sigma \right] \right\} \quad (4)$$

其中 y 表示原始信号， z 表示拟合信号，下标 i 表示光谱信号的每个点， $d = y - z$ ， d^- 为 $d < 0$ 的部分， m 为均值， σ 为标准差。当信号与基线之间的差异小于估计的噪声平均值时，逻辑函数对低于或高于基线的信号给予几乎相同的权重，避免了基线在无峰区域被低估，在峰值区域被高估，保证所得出的基线更接近于实际基线。

3.3. 主成分分析

由于实验过程中采集的光谱数据维度为 2048 维的高维数据，不仅增加了存储空间，而且影响了分类效果，所以在使用分类算法前需要对原始数据进行降维处理。主成分分析(PCA)是一种应用广泛的数据降维方法，将原变量进行转换[7]，使得少数几个新变量是原变量的线性组合，同时，这些新变量要尽可能表达原变量的数据特征而不丢失信息。

3.4. 支持向量机

支持向量机(Support Vector Machine, SVM)是一种用于模式分类的有监督学习方法，由 Vapnik 提出，

SVM 的基本思想来源于线性判别的最优超平面, 所谓最优分类超平面就是要求分类面不但能将两类样本无错误的分开, 而且要使分类间隔最大。通过构造多个一对一分类器, SVM 可实现多类样本的分类判别。

4. 结果与讨论

4.1. 油 - 屏障式绝缘样品谱图

为方便查看不同老化时间油 - 屏障式绝缘样品的拉曼光谱图变化, 将每个老化时间数据的平均光谱绘制在图中, 对 45# 绝缘油样进行分析。原始光谱如图 2 所示。

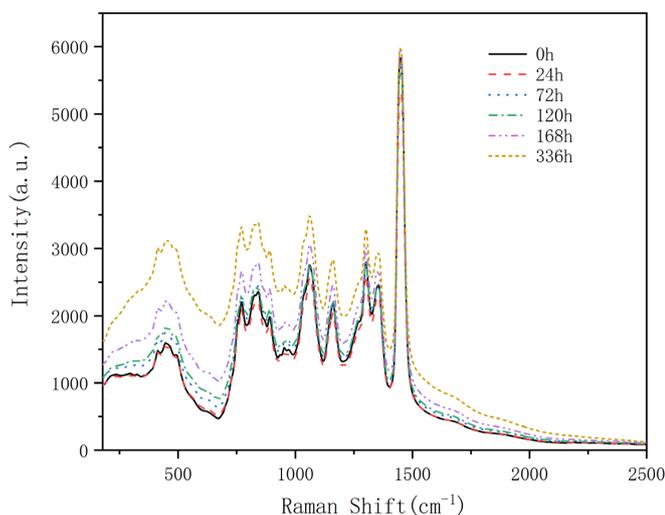
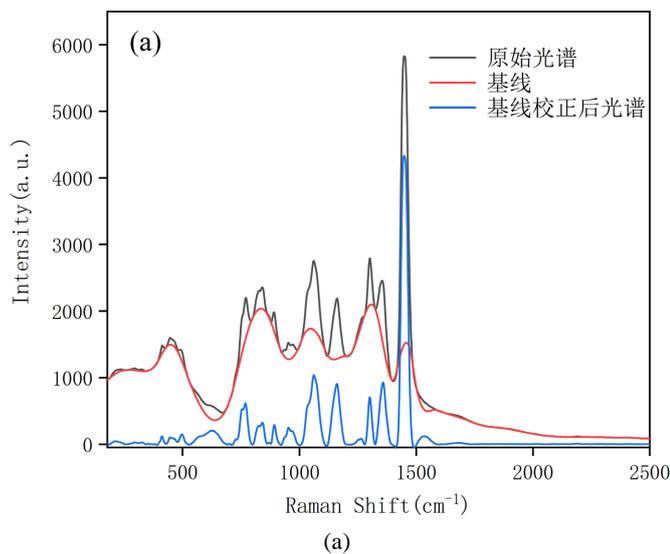


Figure 2. Major raw spectrum of oil samples

图 2. 油样原始光谱图

观察原始光谱图, 可以看出, 由于荧光干扰引起的基线漂移, 使得老化油样的原始光谱处在不同的基准值之下, 不能很好的区别拉曼谱峰。为此, 使用 arPLS 方法对原始光谱进行基线校正处理, 以老化时间 0 h 为例, 展示原始光谱、基线、基线校正后的光谱三者之间的关系, 如下图 3(a)所示, 基线校正后, 对谱线作 SG 平滑处理, 如图 3(b)所示, 通过 SG 平滑, 能够有效消除基线校正无法消除的噪声干扰。



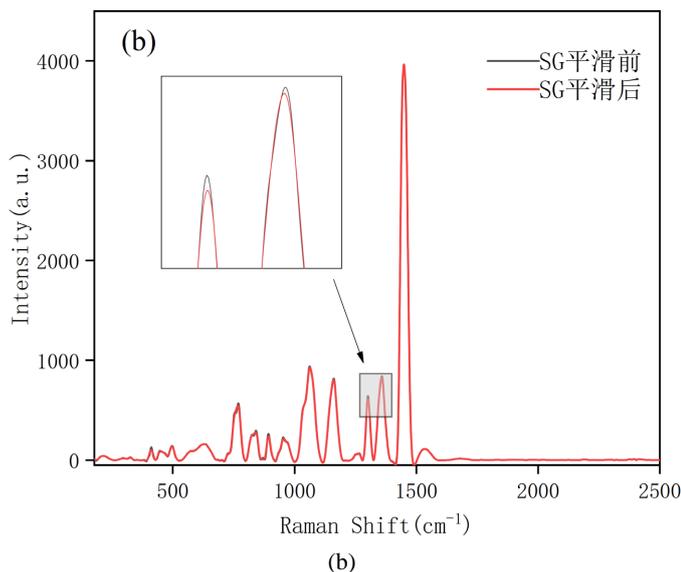


Figure 3. (a) Baseline correction; (b) SG smoothing

图 3. (a) 基线校正; (b) SG 平滑处理

由图 4 可以看出, 经过基线校正、SG 平滑等预处理之后, 各类老化油样的拉曼光谱处在同一个基准值之上。绝缘纸主要由纤维素构成, 而纤维素是由葡萄糖单体经糖苷键构成的高聚合碳氢化合物, 在老化过程中, 由于温度的影响, 连接葡萄糖单体的糖苷键断裂, 生成糠醛等老化物质溶解在油中, 且老化时间越长, 糖苷键断裂的越多[8]。在油中老化物质与绝缘油自身的共同作用下, 测量油样的拉曼光谱强度越来越强。

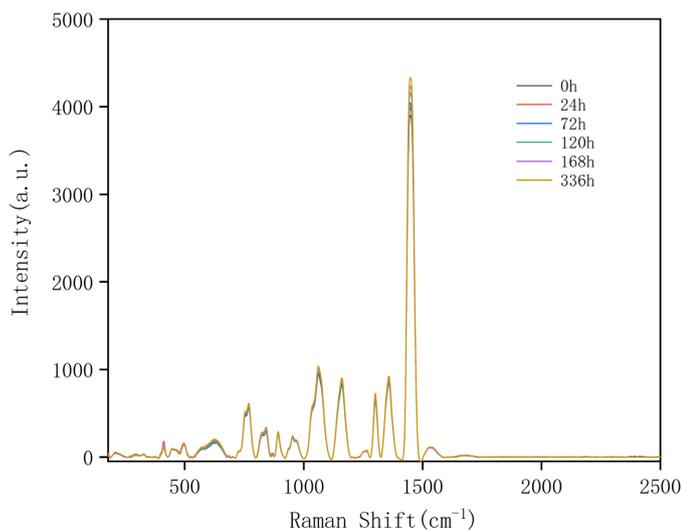


Figure 4. Thermal aging spectra of oil samples at various stages

图 4. 油样各阶段热老化光谱图

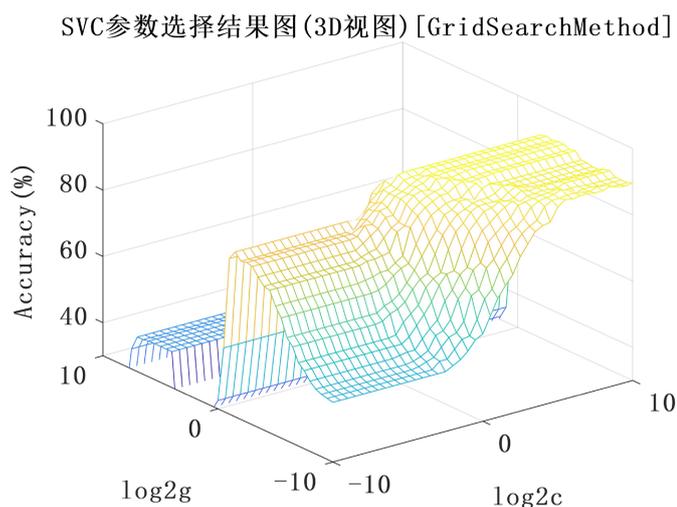
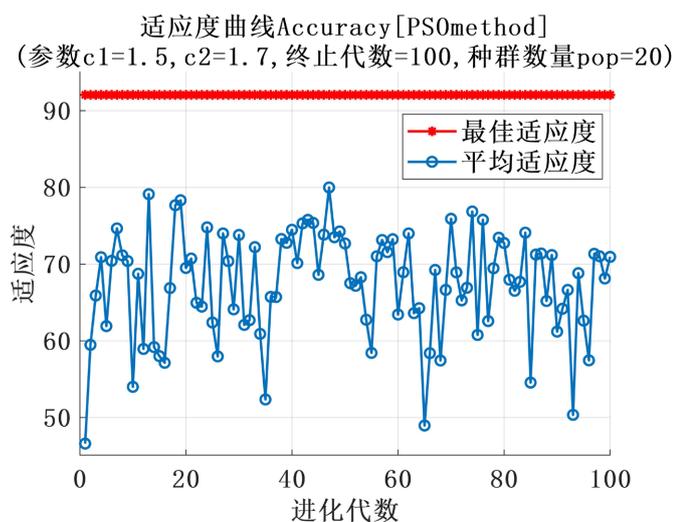
4.2. 分类结果分析

如表 2 所示, 根据不同老化时间段取样将热老化油样分为 6 类, 每个时间段得到 63 条谱线, 其中 42 条用于训练模型, 其余 21 条用于测试模型效果, 对拉曼原始数据进行 SVM 分类。

Table 2. Correspondence between sampling time and categories**表 2.** 取样时间与类别对应关系

取样时间	0 h	24 h	72 h	120 h	168 h	336 h
类别	1	2	3	4	5	6

采用网格法、粒子群、遗传算法寻求 SVM 模型的最佳参数值, 并使用 5 折交叉验证, 避免模型过拟合。其中网格法(Gridsearch)是在给定的参数空间内搜索最优参数, 将参数空间划成长短相同的网格, 其中每一个网格点代表一组参数, 将所有网格点带入 SVM 中验证分类效果, 暴力枚举出最优参数; 粒子群(PSO)优化的基本思想是通过群体中个体之间的协作和信息共享来寻找最优解; 遗传算法(GA)在寻优过程中随机产生多个起始点并同时开始搜索, 由适应度函数来指导搜索方向, 能够在复杂搜索空间寻求全局优化解。三种方法寻优结果可视化如图 5~7 所示。

**Figure 5.** Grid search**图 5.** 网格法**Figure 6.** PSO**图 6.** 粒子群算法

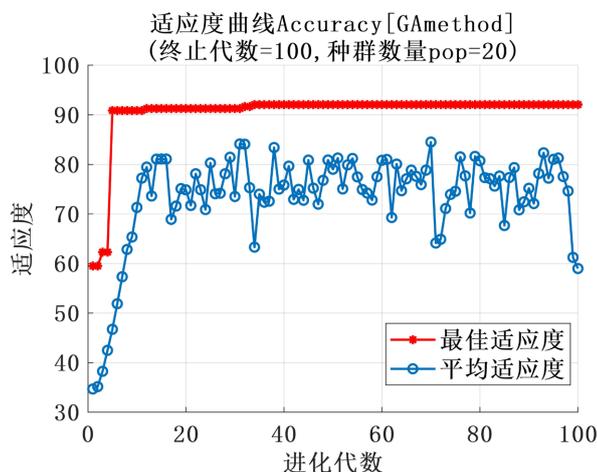


Figure 7. GA
图 7. 遗传算法

表 3 为三种寻优方法性能的对比分析。与网格法相比，使用粒子群算法与遗传算法寻优的交叉验证准确率相同，均为 92.06%，但 PSO、GA 寻优时间较长，这是由于在 PSO 中，粒子的收敛是以轨道的形式实现的，并且粒子的速度总是有限的，导致搜索全部空间耗时较长；而单一的 GA 编码不能全面将优化问题的约束表示出来，需要借助罚函数法将约束问题转为非约束问题，因此大大增加了工作量和求解时间。由表 3 知，三种寻优方法对测试集的效果一致，测试集共有 126 个样本，三种方法均错分 20 个样本，故选择网格法寻优效果最好。

Table 3. Comparison of the results of different optimization methods

表 3. 不同寻优方法结果对比

寻优方法	运行时间	交叉验证准确率	测试集准确率
Gridsearch	1137 s	91.67%	84.13%
PSO	1842 s	92.06%	84.13%
GA	2475 s	92.06%	84.13%

从图 8 中看出未经过任何处理的光谱数据分类的效果并不理想，网格法 SVM 测试集分类准确率为 84.13%。长时间老化与短时间老化对比鲜明，类别为 5, 6 的样品实际类别与预测类别完全相同，类别为 1, 2, 3, 4 的样品总计有 20 个被错误归类。短时间不同老化时间区分的效果不是太好，其中实际老化类别为 1 的，在预测时，有 4 个样本被归于类别 2，归于类别 3、4 的各有两个样本；实际类别为 3 的，有 7 个样本被归于类别 4；实际类别为 4 的，有 4 个样本被归于类别 3。导致这种现象原因可能有如下两方面，第一，短时间老化时，绝缘油萃取绝缘纸中的老化物质比较少，第二，由于荧光干扰造成的基线漂移与测量环境产生的噪声影响了拉曼峰的辨别。

为了去除噪声、荧光等信号的干扰，提高分类准确率，对原始数据进行基线校正、平滑预处理。此外，每个油样的拉曼光谱含有 2048 个数据点，高维数据会造成巨大计算量，很大程度上增加了参数寻优的时间，因此采用主成分分析对数据进行特征提取，使得最大程度上保证样本信息的同时减小数据维度。对于本文中的实验数据，前 2 个主成分的累计贡献率超过 95%，可以充分表示原有光谱的光谱信息，故将原始光谱维度降至 2 维。对预处理后的光谱数据进行网格法参数寻优，结果如图 9 所示。网格法寻优运行时间进一步缩短，且交叉验证准确率上升至 94.05%。

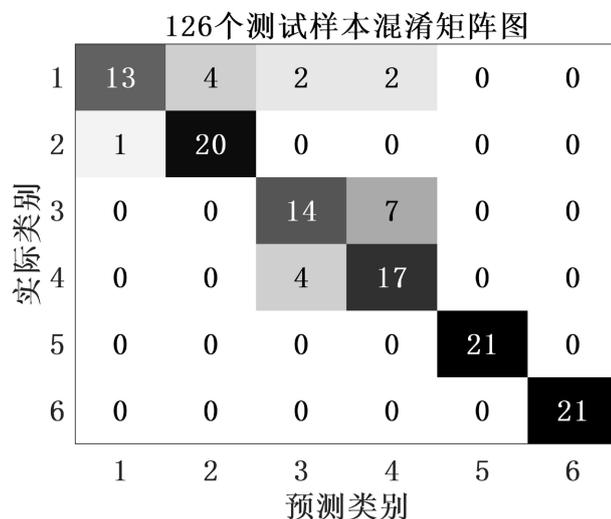


Figure 8. Classification result of oil sample raw data

图 8. 油样原始数据分类结果图

SVC参数选择结果图(3D视图) [GridSearchMethod]

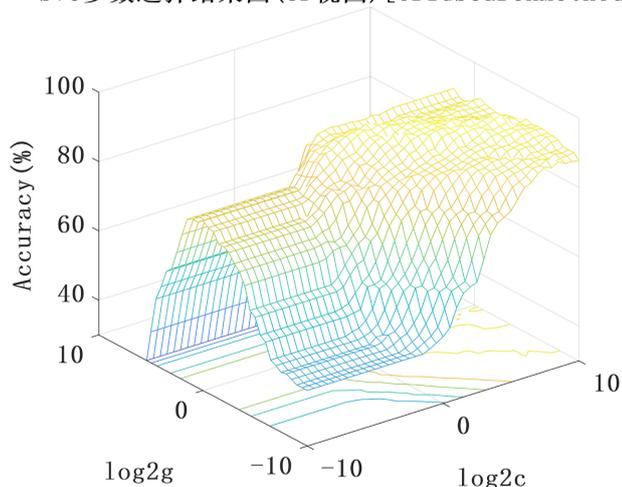


Figure 9. Grid method optimization after data preprocessing

图 9. 数据预处理后网格法寻优

对进行预处理后的数据进行分类, 测试集准确率由 84.13% 提升到 92.86%, 结果如图 10 所示。对短时间热老化样本的错分率已经大大降低, 总错分样本为 7 个。其中实际类别为 1 的样本, 只有 2 个样本被归类为类别 2, 没有归类类别 3、4 的情况, 实际类别为 3 的只有一个样本被归类为类别 4, 实际类别为 4, 5, 6 的分类结果全部正确, 没有错分的情况, 这说明荧光干扰造成的基线漂移影响得到大幅度抑制。但是, 实际类别为 2 的样本错分数增加了 3 个, 对于老化类别为 2 的样本来说, 其分类准确率由 95.23% 下降到 80.95%, 这说明对于老化时间较短的油样, 通过一系列方法对拉曼光谱数据进行预处理, 也可能对某一类的分类结果产生消极影响。通过预处理方法提高了所有类别的总的分类准确率, 特别是对于老化时间较长的样本, 此时绝缘纸溶解在绝缘油中的老化物质较多, 分类准确率高达 100%, 说明此方法能够很好的对运行时间较长的变压器的油 - 屏障式绝缘老化状态作出正确的评估。由此可见, 使用网格法优化 SVM 参数对变压器油 - 屏障式绝缘老化状态的评估具有良好的可行性。

126个测试样本混淆矩阵图

1	19	2	0	0	0	0
2	1	17	2	1	0	0
3	0	0	20	1	0	0
4	0	0	0	21	0	0
5	0	0	0	0	21	0
6	0	0	0	0	0	21
	1	2	3	4	5	6

实际类别
预测类别

Figure 10. Classification result after data preprocessing
图 10. 数据预处理后分类结果图

5. 结论

通过室内老化设备进行加速热老化实验, 获得不同老化时间的油-屏障式绝缘样品, 并对其进行拉曼光谱检测。对原始数据进行 SVM 分类, 通过 Gridsearch、PSO、GA 等方法对 SVM 参数进行寻优, 其中, Gridsearch 在保持较好的寻优效果时, 运行时间较短。为提高 SVM 的分类准确率, 对拉曼光谱进行预处理操作, 包括采用 arPLS 方法消除拉曼光谱的基线漂移, 采用 SG 平滑去除噪声信号, 通过 PCA 有效降低拉曼光谱的数据维度, 经预处理后拉曼数据 SVM 的分类准确率提高至 92.86%, 为进一步探究变压器油-屏障式绝缘老化状态评估提供参考。

基金项目

温州大学大学生创新创业项目, 浙江省大学生新苗人才项目。

参考文献

- [1] 王健一, 刘雪丽, 孙建涛, 等. 油纸绝缘新型老化表征物研究进展与展望[J]. 中国电机工程学报, 2021, 41(21): 7517-7529.
- [2] 邓素梅, 刘厦, 康凯, 等. 表面增强拉曼光谱技术在食品检测中的应用研究进展[J]. 分析试验室, 2022, 41(2): 232-239.
- [3] 吴娟霞, 谢黎明. 二维材料的拉曼光谱研究进展[J]. 科学通报, 2018, 63(35): 3727-3746.
- [4] 顾朝亮, 陈伟根, 邹经鑫, 等. 变压器油中溶解微量糠醛的激光拉曼光谱检测方法[J]. 电工技术学报, 2016, 31(20): 219-227.
- [5] 康明, 韩森坪, 杨洪杰, 等. 基于天然气组分红外光谱图的数据预处理方法研究[J]. 红外技术, 2021, 43(8): 804-808.
- [6] Sung, J.B., Aaron, P., Young, Y.J., et al. (2015) Baseline Correction Using Asymmetrically Reweighted Penalized Least Squares Smoothing. *Analyst*, **140**, 250-257. <https://doi.org/10.1039/C4AN01061B>
- [7] 魏松丽, 黄纪念, 张丽霞, 等. 基于主成分分析的红外预处理温度对花生油香气成分的影响[J]. 食品科技, 2022, 47(12): 149-155.
- [8] 王伟, 马志青, 李成榕, 等. 纤维素老化对油纸绝缘水分平衡的影响[J]. 中国电机工程学报, 2012, 32(31): 100-105+224.