

LIF技术与ELM算法的电力变压器故障诊断研究

闫鹏程^{1,2}, 张超银^{2*}, 孙全胜², 尚松行², 尹妮妮¹, 张孝飞²

1. 安徽理工大学, 深部煤矿采动响应与灾害防控国家重点实验室, 安徽 淮南 232001
2. 安徽理工大学电气与信息工程学院, 安徽 淮南 232001

摘要 电力变压器油的检测分析是电力变压器故障诊断的有效方法, 快速识别电力变压器油的油样对电力变压器故障诊断工作有重大意义。常规的电力变压器油的检测技术主要是气相色谱法, 此方法操作比较复杂, 且不适合在线检测, 不能及时发现变压器的故障隐患。提出一种激光诱导荧光光谱(LIF)技术与极限学习机(ELM)算法的电力变压器故障诊断研究的方法。实验采集四种油样, 分别为热性故障油、电性故障油、局部受潮油以及原油。使用激光发生器激发油样而发射荧光, 获取不同油样光谱数据, 采用MSC、SNV预处理算法对光谱数据进行处理, 防止噪声等因素干扰。随后, 利用KPCA和PCA降维, 主成分个数皆取5, KPCA处理后显示MSC预处理的累计贡献率最高, 为99%, 经MSC预处理的PCA模型累计贡献率依然达到95%以上, Original-KPCA与Original-PCA模型的累计贡献率均达到65%以下, 可以发现, 采用预处理的模型, 累计贡献率均有上升。最后, 分别对两种降维后的数据利用ELM进行回归拟合。实验表明, KPCA、PCA两种降维方式, KPCA算法表现性能较好, 处理数据时间更短, 提高了模型的可靠性和效率。同KPCA降维方式下, MSC-ELM模型的拟合优度 R^2 为0.999 41, 均方误差MSE为0.074%; SNV-ELM拟合优度 R^2 为0.999 08, 均方误差MSE为0.129%; Original-ELM拟合优度 R^2 为0.996 95, 均方误差MSE为0.399%; 对比可以发现MSC比SNV处理后的效果更好, MSC-KPCA-ELM模型表现效果最佳, 预测值与真实值更为接近, 均方根误差最小。结果证明, MSC-KPCA-ELM模型结合激光诱导荧光光谱技术更加适用于对电力变压器是否发生故障的快速诊断, 精确判断为哪种故障类型, 保障电力设备的运行安全。

关键词 激光诱导荧光光谱; 极限学习机; 变压器油; KPCA; PCA

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)05-1459-06

引言

自20世纪70年代以来, 我国的经济迅速地发展, 电力工业也相应地大规模壮大, 导致电力变压器的使用日益增多^[1-2]。根据国家电力联合会的相关统计, 2021年1月—2月份, 全社会用电量为12 588亿kW时, 同比增长22.2%, 全国工业用电量比制造业用电量增加1 846亿kW时。发电机组保持绿色发展, 核能、风能、水能等能源的发电量均有提升, 其中燃气发电机组容量超过一亿千瓦, 装机容量的增加, 对电网中的关键设备要求更高^[3-4]。煤炭发电是电力生产的主要能源, 煤矿中的电力变压器更是在电能传输中起关键作用。因此, 定期对煤矿变压器检测维护显得十分重要。

电力变压器常见的故障为电性故障、局部受潮以及热性故障。电性故障主要表现为短路故障, 比如单相对地短路、两相对地短路、三相之间的短路, 出口处短路最为常见。局部受潮表现为套管部位密封不严谨, 储油柜内有积水等, 电力变压器油中掺杂水分会影响其绝缘特性^[5-8]。异常电流过热会导致电力变压器出现热性故障, 比如涡流、环流等因素, 都会引起电流的瞬间增加, 引起电力变压器出现故障^[9-10]。

目前对电力变压器油检测的方法大多是气相色谱法, 这种方法操作比较复杂, 不适合在线检测, 不能快速地检查出故障原因^[11-14]。本文提出激光诱导荧光光谱技术结合ELM算法识别出电力变压器故障类型。首先, 通过激光打入未知油样, 得到荧光光谱, 再通过MSC、SNV预处理算法进行噪声处理, 采用KPCA与PCA算法对得到的光谱数据进行降

收稿日期: 2021-04-21, 修订日期: 2021-06-08

基金项目: 国家重点研发计划重点专项(2018YFC0604503), 安徽省自然科学基金青年项目(1808085QE157), 安徽省博士后科研经费项目(2019B350), 中国煤炭工业协会2018年度科学技术研究指导性计划项目(MTKJ2018-258)资助

作者简介: 闫鹏程, 1988年生, 安徽理工大学电气与信息工程学院讲师 e-mail: pcyan1988@126.com

* 通讯作者 e-mail: 517806552@qq.com

维,再经过 ELM 算法识别出电力变压器的故障类型。

1 实验部分

1.1 材料

实验样本是从黄山市和宿州市的国家电网收集的电力变压器油样,分别为热性故障油、电性故障油、局部受潮油以及原油,故障油均为单状态下的油样。每种油样分别采集 50 个,共 200 个,存放在密闭无光的容器中。

1.2 荧光光谱的采集

采用的仪器为美国 Ocean optics 公司生产的激光诱导荧光光谱仪,型号为 USB2000+,光谱仪探头为可浸入式,型号为 FPB-405-V3。整个油样的光谱采集流程如图 1 所示。

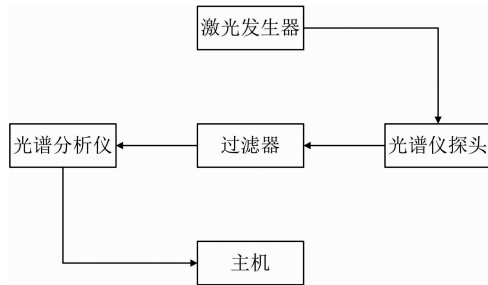


图 1 油样的光谱采集流程图

Fig. 1 Flow chart of oil sample spectrum collection

为了防止其他光线的干扰,整个实验均在暗室中进行。由激光发生器发射出波长为 405 nm 的激光,经过光纤传输,到达微型光谱仪探头,探头把激光直接输出到被测油样中,而后接收被测油样受激发而产生的荧光,经过过滤器,筛选出所需要的光谱波段,再由光纤传输到光谱仪进行光电信号转换等,最后经数据线传送到主机。主机上由 OceanView 软件显示荧光光谱,波长范围 370~1 050 nm,积分时间 100 ms,多次扫描平均为 3,滑动平均宽度为 1。

1.3 光谱预处理

由于在光谱获取的过程中,难免会有一些不确定因素影响光谱图的质量。第一,一些干扰噪声掺杂在生成的光谱图像中,会增大光谱数据的错误概率。第二,由于数据量大,数据的品质不统一。为了提高高品质的数据,以便得到更高质量的结果,因此,我们需要对获得的光谱数据采用两种预处理方法,分别为多元散射校正法(MSC)、标准矢量归一化法(SNV)。

光谱预处理目前常用的一种方法是多元散射校正算法,降低散射的影响,增加光谱数据的准确性。由于难以获取理想光谱数据,因此,假设所有光谱数据的平均值为“理想光谱”。其次,将采集的油样光谱与理想光谱进行一元线性回归,得到每个油样的基线平移量和偏移量。最后,校正每个油样的光谱。

$$A = \frac{1}{n} \sum_{i=1}^n A_i \quad (1)$$

$$A_i = k_i \bar{A} + b_i \quad (2)$$

$$A_{i(MSC)} = \frac{A_i - b_i}{k_i} \quad (3)$$

标准矢量归一化算法的主要作用是削减光程变化、表面散射等因素的影响。SNV 算法是对光谱阵的行运算的,处理的对象为某一条光谱。SNV 归一化的公式

$$SNV_i = \frac{x_{i,k} - \bar{x}_i}{\sqrt{\sum_{k=1}^n \frac{(x_{i,k} - \bar{x}_i)^2}{m} - 1}} \quad (4)$$

$$k = 1, 2, \dots, m; i = 1, 2, \dots, n$$

式(4)中, m 为波长点数, n 为校正集样品数, \bar{x}_i 为第 i 样品的光谱数据平均值。

两种预处理算法所生成的图像和原始图像进行对比,如图 2 所示,波段在 370~550 和 900~1 050 nm 基本趋于一致,

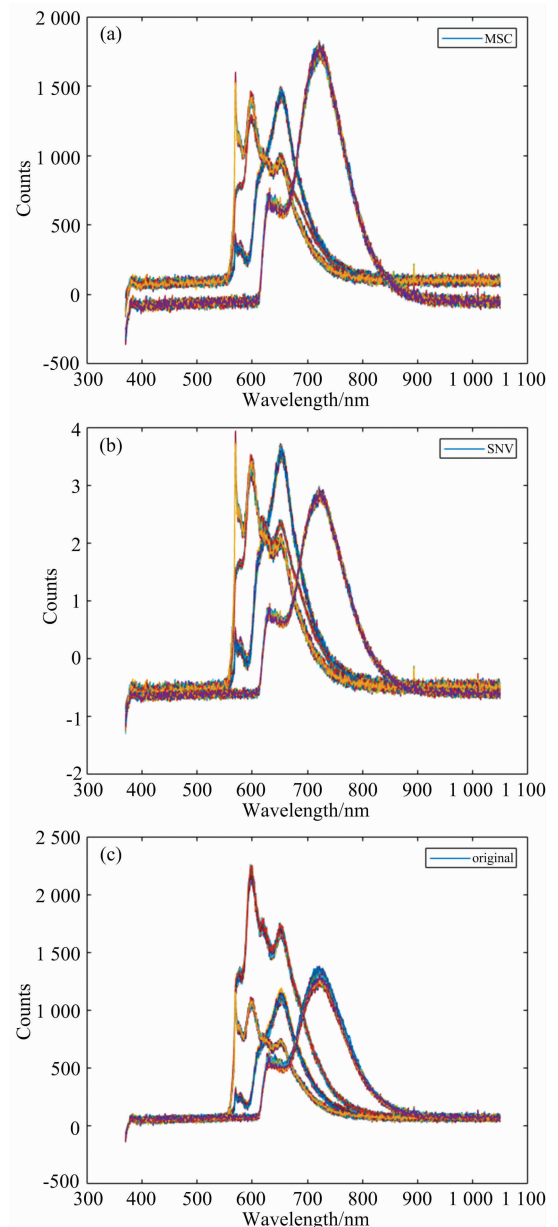


图 2 两种预处理及原始图像

Fig. 2 Two kinds of preprocessing and original image processing

主要不同集中在 550~900 nm, 四种波的波峰也主要集中在此, 可以看出有四类不同的波形, 分别代表四种不同的油样。

1.4 PCA 和 KPCA 降维

每个样本生成的光谱数据为 2 048 个, 一共 200 个样本, 由于数据维度过大, 会造成模型的复杂性, 因此, 采用 PCA 与 KPCA 算法降低数据的维度, 减少整个模型的训练时间。

主成分分析(PCA)是把原来的变量映射到新的变量空间, 在新的变量空间可以用若干个变量替代原来的变量, 并尽可能保留原变量的数据内容, 新变量之间相互正交, 消除原变量的共线性问题。

核主成分分析(KPCA)是通过核函数把原始数据映射到高维空间, 再利用更高的维度空间进行线性降维。

对预处理后的数据和原始数据进行降维, 如图 3 和图 4 所示, 当主成分数为 1 时, 只有 MSC-PCA 和 MSC-KPCA 的累计贡献度达到 85% 以上, 其他模型的累计贡献度均低于此值。随着主成分数的增加, 累计贡献度也随之上升, 因此, 选择主成分数为 5, MSC-PCA 累计贡献度为 95.68%, Original-PCA 与 SNV-PCA 累计贡献度分别为 61.74% 和 78.58%, MSC-KPCA 累计贡献度达到最高, 为 99.00%, Original-KPCA 与 SNV-KPCA 均低于 MSC-KPCA 的累计贡

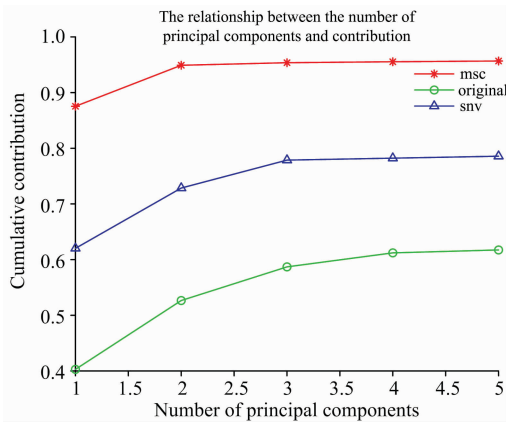


图 3 PCA 累积贡献率

Fig. 3 Cumulative contribution rate of PCA

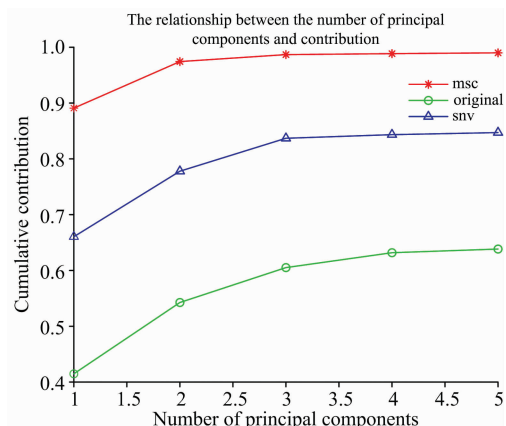


图 4 KPCA 累积贡献率

Fig. 4 Cumulative contribution rate of KPCA

献度, 分别为 63.84% 和 84.72%。结合数据可以看出, MSC 预处理后的数据经两种方式降维效果最好, 累计贡献度最高。同时, 经预处理比未经预处理得到的降维数据更佳。

2 结果与讨论

极限学习机(ELM)是一种前馈神经网络算法, 该算法包含输入层、隐含层、输出层。在创建网络时, 输入层和隐含层的连接权值以及隐含层神经元的阈值均为随机设定, 且设定以后不再调整。隐含层和输出层之间的连接权值 β 可通过以下方程组的最小二乘解获取

$$\min_{\beta} \| H\beta - T' \| \tag{5}$$

其解为

$$\hat{\beta} = H^+ T' \tag{6}$$

式(6)中, H^+ 为隐含层输出矩阵 H 的 Moore-Penrose 广义逆。

2.1 ELM 的预测模型的基本步骤

(1)导入 PCA 与 KPCA 降维后的数据。

(2)训练集/测试集的产生。本实验采用四种油样, 每类油样分别贴入标签 1, 2, 3 和 4, 即标签 1 对应热性故障油, 标签 2 对应电性故障油, 标签 3 对应局部受潮油, 标签 4 对应原油。每一种油样中, 随机选取 40 个作为训练集, 10 个为测试集, 共 160 个训练集, 40 个测试集。

(3)数据归一化, 把不同范围的数据映射到指定的区间, 减小数据差异较大对模型性能的影响, 提高预测模型的收敛速度, 缩短训练时间。

(4)ELM 创建/训练。用 `elmtrain()` 函数创建并训练 ELM, 设置隐含层神经元个数以及激活函数类型, 参数 TYPE 设置为 0。

(5)ELM 仿真测试。训练完成后, 用 `elmpredict()` 函数对测试集进行仿真测试。设置 TF 和 TYPE, 并与训练函数参数保持一致。

(6)绘图与性能评价。直观反映结果, 并作出合理判断。

2.2 结果与性能评价

由于测试集与训练集是随机产生, 因此每次结果都会不同, 某次运行的 ELM 预测模型拟合, 结果如图 5 所示。预测结果最好的是 MSC-KPCA-ELM 模型, 预测值与真实值基本一致。表现效果最差的是未经预处理的 Original-PCA-ELM, 预测值在真实值周边波动。直观的看, 经 MSC、SNV 预处理的模型比未经预处理的模型预测更为准确。

各预测模型的性能指标如表 1 所示, MSC-KPCA-ELM 模型最佳, 平均相对误差 MRE 为 0.96%, 均方根误差 RMSE 为 0.027 1, 均接近于 0。SNV-KPCA-ELM 模型的平均相对误差 MRE 为 1.74%, 均方根误差 RMSE 为 0.036, Original-KPCA-ELM 模型平均相对误差 MRE 为 2.41%, 均方根误差 RMSE 为 0.063 2。可以看出, 同为 KPCA 降维的模型, 经 MSC 预处理的效果要好。MSC-PCA-ELM 模型的平均相对误差 MRE 为 1.18%, 均方根误差 RMSE 为 0.033 9, 对比后发现, 同为 MSC 预处理, KPCA-ELM 模型比 PCA-ELM 模型效果更佳。效果最差的模型是 Original-PCA-ELM,

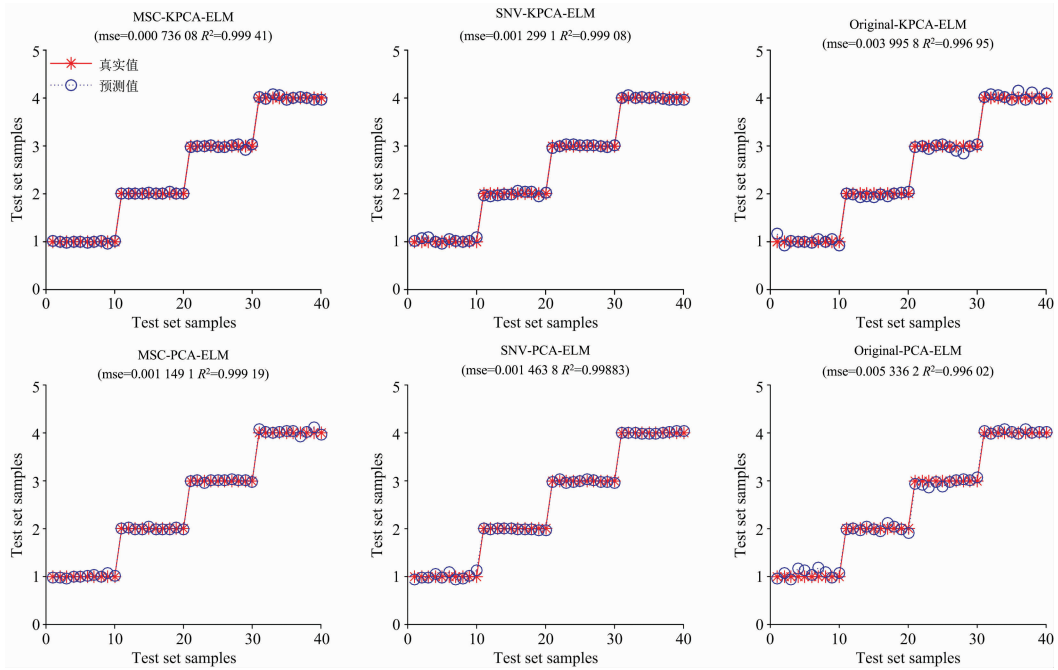


图 5 ELM 测试集输出模型预测结果对比

Fig. 5 Comparison of prediction results of ELM test set output models

平均相对误差最佳的模型高 2.54%，均方根误差 RMSE 高 0.045 9。

0.12%~3.68%，验证出预测结果更为准确。误差幅度变化最大的为 Original-PCA 模型，最高值为 18.74%，较 MSC-PCA 模型高出了 12.02%，可以看出，此模型预测结果相对不佳。

表 1 ELM 测试集输出模型性能指标对比

Table 1 Performance comparison of ELM test set output models

预处理	性能评价	KPCA-ELM	PCA-ELM
MSC	运行时间/s	0.5	3.296 9
	平均相对误差(MRE)	0.009 6	0.011 8
	均方误差(MSE)	0.000 7	0.001 1
	均方根误差(RMSE)	0.027 1	0.033 9
	拟合优度(R^2)	0.999 41	0.999 19
SNV	运行时间/s	0.5175	3.343 8
	平均相对误差(MRE)	0.017 4	0.018 5
	均方误差(MSE)	0.001 3	0.001 5
	均方根误差(RMSE)	0.036 0	0.038 3
	拟合优度(R^2)	0.999 08	0.998 83
Original	运行时间/s	0.578 1	3.484 4
	平均相对误差(MRE)	0.024 1	0.035 0
	均方误差(MSE)	0.004 0	0.005 3
	均方根误差(RMSE)	0.063 2	0.073 0
	拟合优度(R^2)	0.996 95	0.996 02

MSC 预处理后的 KPCA-ELM 模型运行时间为 0.5 s，PCA-ELM 模型运行时间为 3.296 9 s，其中 KPCA 降维的时间为 0.437 5 s，PCA 降维的时间为 3.234 4 s，可以看出，由 KPCA 降维比 PCA 更快，同一状态下的效果会更好。

相对误差能够反映预测值偏离真实值的实际大小，如图 6 所示为 ELM 测试集输出结果的相对误差曲线，误差幅度变化较大主要集中在标签 1 所对应的油样，标签 2, 3 和 4 基本趋于平稳。MSC-KPCA 模型的相对误差波动最小，为

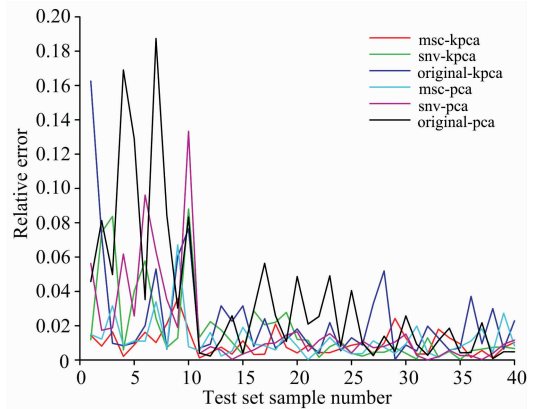


图 6 预测值与真实值的相对误差曲线

Fig. 6 The relative error curve between predicted value and real value

3 结 论

实验对三种电力变压器故障油以及原油利用激光诱导荧光光谱分析技术进行数据分析，采用 MSC、SNV 两种预处理方式，结果发现，MSC 预处理后的数据经 KPCA 降维，其主成分贡献率达到最高，而未经预处理的数据经两种降维算法达到的效果均较差。实验可见，经 KPCA 降维的模型运行时间小于 PCA，验证出使用 LIF 技术获取电力变压器油荧光光谱，结合 MSC-KPCA 模型经 ELM 拟合性能最佳，对检测电力变压器故障类型具有一定的可靠性。

References

- [1] WANG Wei-wang, LIU Ying, HE Jie-feng, et al(王威望, 刘莹, 何杰峰, 等). High Voltage Engineering(高电压技术), 2020, (10): 3362.
- [2] HUANG Da-rong, CHEN Chang-sha, SUN Guo-xi, et al(黄大荣, 陈长沙, 孙国玺, 等). Automation of Electric Power Systems(电力系统自动化), 2017, 41(12): 206.
- [3] HE Zheng-you, LI Bo, LIAO Kai, et al(何正友, 李波, 廖凯, 等). Proceedings of the CSEE(中国电机工程学报), 2020, 40(19): 6193.
- [4] PU Tian-jiao, QIAO Ji, HAN Xiao, et al(蒲天骄, 乔骥, 韩笑, 等). High Voltage Engineering(高电压技术), 2020, 46(2): 369.
- [5] LIU Yun-peng, XU Zi-qiang, LI Gang, et al(刘云鹏, 许自强, 李刚, 等). High Voltage Engineering(高电压技术), 2019, 45(2): 337.
- [6] DAI Quan-min, QI Bo, LI Cheng-rong, et al(戴俊民, 齐波, 李成榕, 等). Transactions of China Electrotechnical Society(电工技术学报), 2018, 33(11): 2613.
- [7] LEI Lin, LAI Zhen-liang, ZHU Jia-long, et al(雷霖, 赖真良, 朱家龙, 等). Power System Protection and Control(电力系统保护与控制), 2017, (17): 26.
- [8] OUYANG Sen, MA Wen-jie(欧阳森, 马文杰). Electric Power Automation Equipment(电力自动化设备), 2018, 38(9): 21.
- [9] MA Hong-zhong, GENG Zhi-hui, CHEN Kai, et al(马宏忠, 耿志慧, 陈楷, 等). Automation of Electric Power Systems(电力系统自动化), 2013, 37(8): 89.
- [10] XIAO Yan-wei, LOU Ping, JIANG Shui-zhong, et al(肖艳伟, 楼平, 蒋水中, 等). Transformer(变压器), 2019, 56(4): 46.
- [11] GONG Mao-fa, ZHANG Yan-pan, LIU Yan-ni, et al(公茂法, 张言攀, 柳岩妮, 等). Power System Protection and Control(电力系统保护与控制), 2015, 43(3): 113.
- [12] LI En-wen, WANG Li-nong, SONG Bin, et al(李恩文, 王力农, 宋斌, 等). Transactions of China Electrotechnical Society(电工技术学报), 2018, 33(19): 4594.
- [13] WEN Quan, ZHOU Lu-bo, WANG Dong, et al(温泉, 周录波, 王栋, 等). Electrical Measurement & Instrumentation(电测与仪表), 2020, 57(13): 23.
- [14] ZHOU Meng-ran, WANG Jin-guo, SONG Hong-ping, et al(周孟然, 王锦国, 宋红萍, 等). Laser & Optoelectronics Progress(激光与光电子学进展), 2020, 57(20): 203001.

LIF Technology and ELM Algorithm Power Transformer Fault Diagnosis Research

YAN Peng-cheng^{1,2}, ZHANG Chao-yin^{2*}, SUN Quan-sheng², SHANG Song-hang², YIN Ni-ni¹, ZHANG Xiao-fei²

1. State Key Laboratory of Mining Response and Disaster Prevention and Control in Deep Coal Mine, Anhui University of Science and Technology, Huainan 232001, China

2. College of Electrical and Information Engineering, Anhui University of Technology, Huainan 232001, China

Abstract The detection and analysis of power transformer oil is an effective method for power transformer fault diagnosis, and it is of great significance to quickly identify the oil sample of power transformer oil for power transformer fault diagnosis. The detection technology of conventional power transformer oil is mainly gas chromatography, which is complicated and not suitable for on-line detection, and it can't find the fault hidden danger of transformer in time. A method for transformer fault diagnosis of laser-induced fluorescence spectroscopy (LIF) technology and extreme learning machine (ELM) algorithm is proposed. Four oil samples were collected: thermal fault oil, electrical fault oil, local moisture-affected oil and crude oil. The spectral data of different oil samples are obtained by using a laser generator to emit fluorescence. The spectral data are processed by MSC and SNV preprocessing algorithms to prevent noise and other factors. Subsequently, the use of KPCA and PCA dimension reduction, the main components are taken 5, KPCA processing shows that MSC pretreatment of the cumulative contribution rate of the highest, 99%, MSC pre-processed PCA model cumulative contribution rate is still more than 95%, Original-KPCA and Origin-PCA model cumulative contribution rate of less than 65%, you can find that the use of pretreatment model, cumulative contribution rate has increased. Finally, the data after the two dimensions are regression fit by ELM. Experiments show that KPCA and PCA are two kinds of dimensional reduction methods. The KPCA algorithm performs best, the processing time is

short, and the reliability and efficiency of the model are improved. In the same KPCA dimension reduction mode, the fitting excellence R^2 of the MSC-ELM model is 0.999 41, the mean square error MSE is 0.074%, and the SNV-ELM fit is 0.999 08, the mean square error MSE is 0.129%, The Original-ELM fitting excellence R^2 is 0.996 95, the mean square error MSE is 0.399%, and the comparison can be found that MSC is better than SNV after processing. The MSC-KPCA-ELM model performs best, the prediction value is closer to the real value. The mean square root error is the smallest. The results show that the MSC-KPCA-ELM model, combined with laser-induced fluorescence spectroscopy technology, is more suitable for the rapid diagnosis of whether or not the power transformer has failed, which type of fault is accurately determined, and the operation safety of power equipment is guaranteed.

Keywords Laser induced fluorescence spectroscopy; Extreme learning machine; Transformer oil; KPCA; PCA

(Received Apr. 21, 2021; accepted Jun. 8, 2021)

* Corresponding author