

doi: 10.3788/gzxb20164505.0530002

变压器油中颗粒污染物的中红外光谱检测

陈彬, 韩超, 刘阁

(重庆工商大学 废油资源化技术与装备工程研究中心, 重庆 400067)

摘 要: 配制了含不同颗粒污染等级的变压器油样, 利用中红外光谱扫描获得油样的红外光谱数据, 再采用连续投影算法提取油样红外光谱的有效波长变量, 分别应用偏最小二乘法和支持向量机法方法建立了颗粒污染等级与中红外光谱有效波长的模型. 所配置的油样红外光谱经过连续投影算法提取的有效波长具有特定颗粒污染物特征波长的特点, 所建两种模型的预测效果均优于全谱的偏最小二乘法和支持向量机法模型, 对验证集样本数据预测的决定系数分别为 0.892 9、0.934 3, 均方根误差为 6.372×10^{-3} 、 3.07×10^{-3} , 获得了较好的预测效果, 为变压器油中颗粒物的检测提供了借鉴.

关键词: 污染物检测; 光谱分析; 连续投影算法; 偏最小二乘法; 支持向量机; 变压器油; 颗粒污染物

中图分类号: O657.33

文献标识码: A

文章编号: 1004-4213(2016)05-0530002-5

Detection on Particulate Pollutant in Transformer oil Based on the Mid-Infrared Spectrum

CHEN Bin, HAN Chao, LIU Ge

(Engineering Research Centre for Waste Oil Recovery Technology and Equipment, Chongqing Technology and Business University, Chongqing, 400067, China)

Abstract: The different particle pollution degree in transformer oil samples were made up, the infrared spectrum data of the oil samples were acquired by using the infrared spectrum scanning. Using the successive projections algorithm, the effective wavelength variables of the oil samples were extracted. Based on the extracted wavelength variables, two models of both the effective wavelength of the infrared spectrum and the particle contamination pollution degree were established by using partial least squares and support vector machine method. The effective wavelength of successive projections algorithm extracted from the infrared spectrum of transformer oil samples has the characteristics of the wavelength of specific particle contamination, and the prediction effects of the models are better than partial least squares model and support vector machine model using the full infrared spectrum data of the oil samples. Besides, the determination coefficient of the prediction set of oil samples are 0.892 9, 0.934 3 respectively with the two models, and the root mean square error are 6.372×10^{-3} , 3.07×10^{-3} respectively, thereby the satisfactory prediction results has achieved, these provide a reference for the detection of the particle contamination in transformer oil.

Key words: Pollutant detection; Spectrum analysis; Successive Projections Algorithm; Partial Least Squares; Support Vector Machine; Transformer oil; Particulate pollutant

OCIS Codes: 300.6340; 040.3060; 070.4790; 110.3080

0 引言

变压器油在变压器设备运行中具有冷却、绝缘以

及灭弧的作用, 其质量的好坏对变压器设备是否能够安全运行有着至关重要的作用. 变压器油设备在生产过程或运行中会产生金属粉末或其他颗粒污染物, 如

基金项目: 国家自然科学基金(No. 51375516)、重庆基础与前沿研究项目(Nos. cstc2014jcyjA90015, cstc2013jcyjA90021)资助

第一作者: 陈彬(1972—), 男, 教授, 博士, 主要研究方向为油液污染控制技术方面的研究. Email: hustchb@163.com

导师(通讯作者): 刘阁(1973—), 女, 副教授, 主要研究方向为油水分离技术研究, Email: 397289154@qq.com

收稿日期: 2015-11-05; 录用日期: 2015-12-24

http://www.photon.ac.cn

金属铜、铁等金属氧化物以及胶质物,从而改变油液的粘度、表面张力等性能,降低变压器油的绝缘强度,极大地影响变压器设备的安全运行,因此对变压器油中颗粒污染物的检测一直受到研究人员的重视^[1-3]。

红外光谱技术是一种简便、快速、无污染、低成本的分析技术,已经广泛应用于食品、化工等领域^[4-6]。近年来,随着红外光谱仪和化学计量学的发展,红外光谱法在润滑油分析中也有了一定的进展,尤其是在酸值、碱值、含水率等测定方面开展了一些有实际应用意义的工作。加拿大 McGill 大学红外光谱小组提出了用红外光谱法测定润滑油酸值和碱值的方法^[7]。在待测润滑油中加入一定量的邻苯二甲酰亚胺钾,通过酸碱反应将邻苯二甲酰亚胺钾转换为邻苯二甲酰亚胺,根据其在红外光谱中的羰基特征吸收(1727cm^{-1} 或 1774cm^{-1}),计算出待测润滑油的酸值。A R Caneca 等采用判别分析法^[8],按照使用里程对润滑油进行了分类,并通过偏最小二乘法(Partial Least Squares, PLS)建立了红外光谱法测定 $40\text{ }^{\circ}\text{C}$ 粘度的校正模型,用于监测柴油机油的使用情况^[9]。本课题组^[10-11]基于油液的红外光谱,采用 PLS 法、T_S 模糊辨识等数学建模方法建立了变压器油中含水量的预测模型,并获得了满意的预测准确度,实现了对油中含水量的监测。张瑜^[12]等采用了最小二乘支持向量机(Least Squares Support Vector Machine, LS-SVM)建立可见-近红外光谱检测模型,并且结合无信息变量消除算法(Uninformative Variable Elimination, UVE)与连续投影算法(Successive Projections Algorithm, SPA)进行光谱有效波长选取,对润滑油酸值进行无损检测。

为了实现变压器油中颗粒污染物含量的准确、快速的检测^[13],本文采用国标 ISO4406 作为油液污染度标准配制 50 组不同污染度的含 Cu、Fe、Si 等氧化物颗粒的油样,对这些油样进行中红外光谱扫描,来探讨含不同颗粒污染等级的油样红外光谱数据与油样中颗粒污染等级之间的关系;运用 SPA 从红外光谱信息中提取有用信息变量,从而提高建模的准确性和效率;再结合 PLS、支持向量机法(Support Vector Machine, SVM)对测试数据建立油样红外光谱数据与油样中颗粒污染物含量的预测模型,实现对变压器油中颗粒污染物含量的预测。

1 实验仪器与方法

1.1 实验样品与仪器

采用 25# 变压器油(中国石油天然气股份有限公司润滑油分公司生产)作为原始油样;采用 HIAC8012 油品污染度测试仪器(HACH Company 生产)对配制的不同污染度的油液进行检测,得到污染度等级在 15/13-22/16 (ISO4406) 范围内的油样;采用 KQ-

400KDB 型高功率数控超声波振荡器(昆山苏美超声仪器有限公司生产,超声功率 400 W,超声频率 40KHz)对不同污染度的油液进行预处理;采用傅里叶变换红外光谱仪 IRAffinity-1(日本岛津生产)对 50 组油样进行红外光谱数据采集,选用 KBr 固定厚度液体池,液体池垫片厚度选取 $100\mu\text{m}$,分辨率 2 cm^{-1} ,扫描次数 45 次,扫描范围设置为 $4700\sim 340\text{ cm}^{-1}$ (共 4523 个变量),此波段范围内光谱用作分析。

1.2 实验方法

1.2.1 配制油样

分别称取 Cu、Fe、Si 等氧化物颗粒粉末各 0.005g,将称好的固体粉末与 1L 25# 变压器油混合,经 8 h 超声波振荡器充分振荡。为保证实验油样与实际应用中油样的一致性,振荡过程存在的超声波加热作用使得油样温度保持在 $40\sim 80\text{ }^{\circ}\text{C}$ 之间。将振荡均匀的油样用 HIAC8012 油品污染度测试仪器进行检测,此时获得油样为污染度等级 22/16(ISO4406)的油样;在此基础上,分别按体积比例兑入原变压器油,并采用 HIAC8012 油品污染度测试仪器进行检测,得到 15/13-22/16 梯度污染度等级的油样(共 50 组)分别盛入 500 mL 锥形瓶中,待实验。

1.2.2 采集中红外光谱数据

采用傅里叶变换红外光谱仪 IRAffinity-1 对预处理好的 50 组不同污染度等级的油样进行中红外光谱数据采集,采用 KBr 固定厚度液体池,液体池垫片厚度选取 $100\mu\text{m}$,分辨率 2 cm^{-1} ,扫描次数 45 次,扫描范围设置为 $4700\sim 340\text{ cm}^{-1}$ (共 4523 个变量),将此波段范围内光谱用作分析。其中随机选择 40 组样品为模型校正集,10 组作为预测集。光谱数据处理采用 Matlab7.6 进行。

2 油中颗粒污染物中红外光谱建模

2.1 建模数据预处理

原始光谱数据采集时受到实验温度、样本不均匀、高频随机噪音以及基线漂移等不利因素的影响,因此需采取一定的预处理方法来实现有效地建模。对 50 组油样的原始光谱信号进行平滑预处理,处理后的光谱信号与原始光谱信号之间的误差为 0.1407,较好地保留了原始光谱的基本特征,提高了光谱曲线的平滑性。预处理前后的光谱曲线如图 1。

光谱经过平滑预处理后,由于光谱信息变量多,全波段共有 4523 个变量,计算量大,以全谱数据建立模型易产生冗余信息,因此需要提取有效波长,从而提高建模的效率和准确度。

选用连续投影算法(SPA)对光谱信息变量进行筛选,该方法能够消除波长变量间的共线性影响,使向量间的共线性降到最小,从而降低模型的复杂度,提高建

模的效率,且可以使检测模型达到更高的准确度,是一种有效的波长选取方法^[14-15].采用 SPA 法在 340~4 700 nm 全谱范围内筛选有效波长变量,共提取变量 23 个,如表 1.可以看出筛选出的前 5 个波长其中 4 个波长主要集中在 Fe-O 键伸缩振动宽谱带区 370~390 nm 范围内,其中选择的第 3 个特征波长 983.738 nm 是 Si-O 键特征吸收峰位置,第 7~8 个波长 342.379 nm、358.775 nm 以及第 10 个波长之后的数据在 Cu-O 键伸缩振动区 340 nm 附近,表明经过 SPA 选择的波长具有油中颗粒物光谱的特征意义,能有效提高模型的准确性.

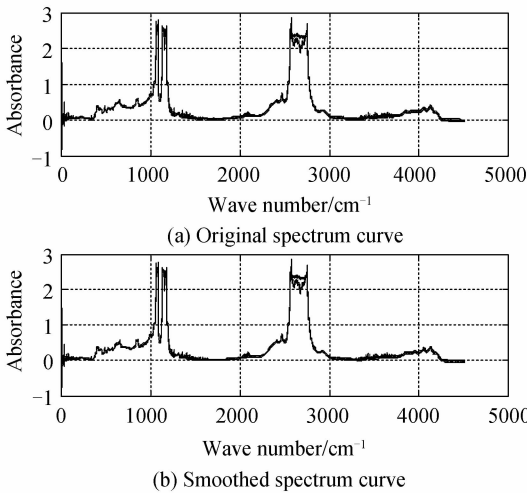


图 1 油样光谱曲线

Fig. 1 The oil samples spectral curve

表 1 SPA 提取的有效波长

Table 1 The effective wavelength of SPA extraction

Method	Wavelength/nm
SPA	2 923.246, 384.815, 376.135, 983.738,
	375.170, 2 844.161, 342.379, 358.775,
	2 820.05, 344.308, 346.237, 343.343,
	341.415, 362.63296, 350.095, 353.953,
	370.348, 340.451, 348.166, 2 996.544,
	354.917,364.561,355.882

油中颗粒物污染度数据范围在 229~22 771 之间,进行归一化处理,把数据映射到 0~1 范围之内,使得程序运行时收敛加快,提高模型的精确性.

2.2 油液中颗粒污染物污染浓度模型的构建

基于平滑预处理后的红外光谱数据(全谱),以及经过 SPA 对原始光谱提取有效波长,分别作为 PLS 模型和 SVM 模型的输入变量,油样的颗粒污染等级作为模型输出变量,建立油中颗粒污染物光谱的全谱模型(PLS 和 SVM),SPA-PLS 模型以及 SPA-SVM 模型.模型的评价指标为决定系数(R^2)和预测均方根误差(Root Mean Square Error, RMSE),模型的相关指数越大,预测均方根误差越小,模型的准确度越高.

PLS 在数学建模中是一种应用非常广泛的方法.

随机选取 50 组油样中的 40 组样品作为校正集,其余 10 组作为模型验证集样本,建立油中颗粒污染物光谱的全谱 PLS 模型以及 SPA-PLS 模型;并利用建立的 PLS 模型对剩余 10 组油样的颗粒含量进行预测.

SVM 是由 Corinna Cortes 等于 1995 年首先提出的一种建模方法,该方法在解决小样本、非线性及高维模式识别问题中具有独特的优势.从 50 组样品中随机选取 40 组作为模型的校正集样本,建立油中颗粒污染物光谱的全谱 SVM 模型以及 SPA-SVM 模型;并利用建立的模型对剩余 10 组油样的颗粒含量进行预测.

建立的油中颗粒污染物光谱全谱模型(PLS 和 SVM),SPA-PLS 模型、SPA-SVM 模型的预测结果见表 2.

表 2 油中颗粒污染物光谱模型的预测结果

Table 2 Prediction results by the models of oil particulate contaminant infrared spectrum

Model	Calibration set		Verification set	
	RMSE	R^2	RMSE	R^2
PLS	9.4998×10^{-2}	0.8589	0.1369	0.7269
SPA-PLS	3.2339×10^{-4}	0.9908	6.372×10^{-3}	0.8929
SVM	9.6008×10^{-4}	0.9675	5.592×10^{-3}	0.8813
SPA-SVM	1.4096×10^{-4}	0.9962	3.070×10^{-3}	0.9343

由表 2 可知,SPA-PLS 模型相对全波段 PLS 模型、SPA-SVM 模型相对全波段 PLS 模型在校正集和验证集的预测相关指数都有所提高,并且预测均方根误差都有所降低,说明 SPA 能有效地提取油样光谱中的有效信息.经过 SPA 对全波段光谱进行有效变量选择后,作为 PLS 模型以及 SVM 模型的输入变量建立的 SPA-PLS 模型和 SPA-SVM 模型能有效地预测中红外油中颗粒物的含量;SPA-PLS 模型和 SPA-SVM 模型对验证集样本进行预测的决定系数分别为 0.892 9、0.934 3,均方根误差分别为 6.372×10^{-3} 、 3.07×10^{-3} ,说明 SPA-PLS 模型和 SPA-SVM 模型能够有效地预测中红外油中颗粒物的含量,获得了满意的预测准确度.

图 2 为建立的油中颗粒污染物红外光谱 4 种预测模型验证集的预测结果与实测值的比较.

图 2(a)、(b)为验证集的全谱 PLS 模型和 SPA-PLS 模型的预测结果与实测值的比较,可以明显看出经过 SPA 有效选择变量后,建立的 PLS 模型的输出能更好地跟踪颗粒物含量的变化,建立的全波段 PLS 模糊模型的预测值与实测值的误差变化范围为 0.32,而 SPA-PLS 模糊模型的预测值与实测值基本吻合,其误差变化范围为 0.21.

由图 2(c)以及图 2(d)两模型的预测值与实测值的关系图中也可以看出,SPA-PLS 模型除了验证集的第 9 组数据与实测值误差较大外,其它几组的验证集数据都较均匀地分布在回归直线附近,且第 2、6 组数

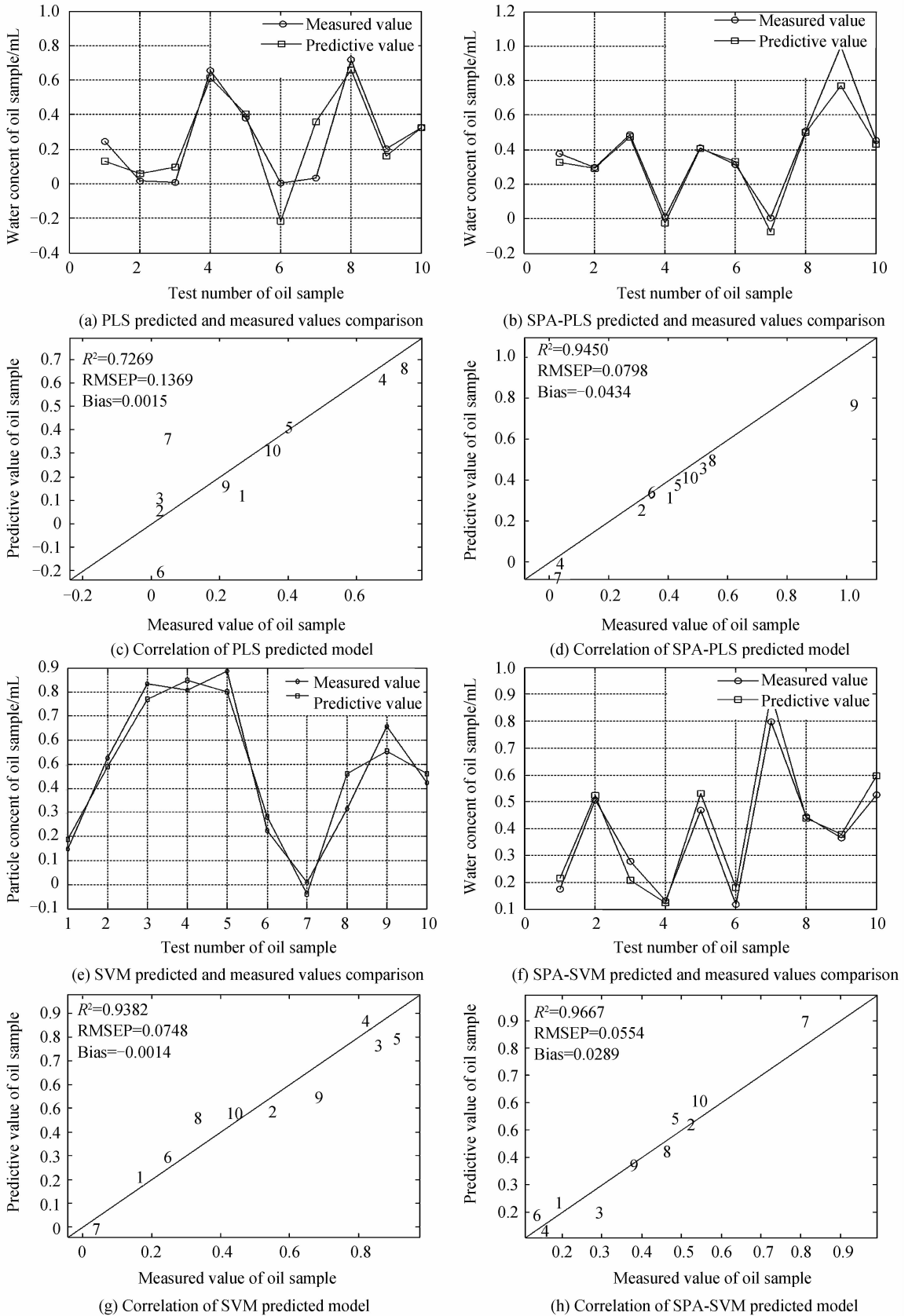


图2 预测结果与实测值的比较

Fig. 2 Predicted and measured values comparison

据分布在回归直线上,而全波段 PLS 模型验证集第 7 组数据偏离回归直线较远,影响了模型预测的精确性,

因此 SAP-PLS 模型相对全波段 PLS 模型预测准确性更高;其中全波段 PLS 模糊模型决定系数 $R^2 =$

0.726 9,均方根误差 RMSEP=0.136 9,SPA-PLS 模型决定系数 $R^2=0.945 0$,均方根误差 RMSEP=0.079 8,可以看出 SPA-PLS 模型的预测准确度更高。

图 2(e)、(f)为验证集全波段 SVM 模糊模型和 SPA-SVM 模糊模型的预测结果与实测值的比较,可以看出两个模型的输出都能较好地跟踪油中颗粒物含量的变化,建立的全波段 SVM 模糊模型的预测值与实测值的误差变化范围为 0.17,而 SPA-SVM 模糊模型的预测值与实测值的误差变化范围为 0.1,相对全谱 PLS 模型误差范围缩小了 0.07。

由图 2(g)、(h)两模型的预测值与实测值的关系图中也可以看出,SPA-SVM 模型除了验证集的第 7 组数据与实测值误差较大外,其它数据都较均匀地分布在回归直线附近,且第 2、8、9 组数据分布在回归直线上,而全波段 SVM 模型验证集数据离回归直线距离较远,影响了模型的预测精确性,因此 SPA-SVM 模型相对全波段 SVM 模型预测精确性较高;其中全波段 SVM 模糊模型决定系数 $R^2=0.938 2$,RMSEP=0.074 8,SPA-SVM 模型决定系数 $R^2=0.966 7$,均方根误差 RMSEP=0.055 4,说明应用 SPA-SVM 模型能够充分利用油液中颗粒物含量信息,从而有效实施油中颗粒物的红外光谱特征的建模和预测。

3 结论

运用 SPA 对含颗粒物油样的中红外光谱进行有效波长筛选,并结合偏小二乘法、支持向量机法建立了油中颗粒污染等级与中红外光谱的预测模型,结论表明:

1)含颗粒污染的油样原始光谱数据经平滑预处理后,采用连续投影算法提取了 23 个有效波长;选择的波长主要集中在 Fe-O 键伸缩振动宽谱带区 370~390nm 和 Cu-O 键伸缩振动区 340nm 附近,选择的特征波长 983.738nm 是 Si-O 键特征吸收峰位置,具有油中颗粒物特征光谱的意义;

2)建立了油中颗粒物的红外光谱的 SPA-PLS 模型和 SPA-SVM 模型,SPA-PLS 模型中决定系数 R^2 达到 0.9450,均方根误差 RMSEP 为 0.0798,SPA-SVM 模型中决定系数 R^2 达到 0.9667,均方根误差 RMSEP 为 0.0554,可见 SPA-SVM 模型预测准确度较高;

3)建立的油中颗粒物的红外光谱模型中 SPA-PLS、SPA-SVM 模型相对全谱 PLS、SVM 模型都有较好的预测效果,对验证集样本数据预测的决定系数 R^2 分别为 0.8929、0.9343,均方根误差 RMSEP 为 6.372×10^{-3} 、 3.07×10^{-3} 。

参考文献

[1] RODRIGUEZ D M, AGUILAR FJA, WROBEL K, et al.

Spectrophotometric assay for copper and iron in transformer oil using partial least squares regression (PLS2) [J]. *IEEE Transactions on Dielectrics & Electrical Insulation*, 2007, **13** (6): 1272-1277.

[2] YANG D W, WU J J, YUAN D, et al. Determination of micro copper in transformer oil by FAAS [J]. *Chinese Journal of Spectroscopy Laboratory*, 2013, **30**(1): 216-219.

[3] AKSHATHAA, KUMAR R, SUNDARA R J, et al. Copper corrosion phenomenon in transformers due to DBDS in mineral transformer oil [C]. *Electrical Insulation and Dielectric Phenomena (CEIDP) 2013 IEEE Conference*, 2013: 826-829.

[4] WU D, HE Y, SHI J H, et al. Exploring near and mid-infrared spectroscopy to predict trace iron and zinc contents in powdered milk [J]. *Journal of Agricultural and Food Chemistry*, 2009, **57**(5): 1697-1704.

[5] WU D, CHEN X J, SHI P Y, et al. Determination of alpha-linoleic acid and linoleic acid in edible oils using near-infrared spectroscopy improved by wavelet transform and uninformative variable elimination [J]. *Analytical Chimica Acta*, 2009, **634** (2): 166-171.

[6] YAN Yan-lu, ZHAO Long-lian, HAN Dong-hai, et al. Basic and application of near infrared spectrum analysis [M]. China Light Industry Press, Beijing, 2005

严衍禄,赵龙莲,韩东海,等.近红外光谱分析基础与应用 [M]. 北京:中国轻工业出版社,2005.

[7] VANF R, SEDMAN J, YAYLAYAN V, et al. Determination of acid number and base number in lubricants by Fourier transform infrared spectroscopy. [J]. *Applied Spectroscopy*, 2003, **57**(11): 1425-1431.

[8] CANECA A R, PIMENTEL M F, MATTA C E D, et al. Assessment of infrared spectroscopy and multivariate techniques for monitoring the service condition of diesel-engine lubricating oils. [J]. *Talanta*, 2006, **70**(2): 344-352.

[9] JEREMY A, INGELA P. Quantifying colocalization by correlation: the pearson correlation coefficient is superior to the mander's overlap coefficient [J]. *International Society for Advancement of Cytometry*, 2010, **77a**(8): 733-742.

[10] CHEN Bin, WU Hong-yang, HAN Chao, et al. Viscosity of insulating oil with trace water based on the mid-infrared spectral analysis [J]. *Acta Photonica Sinica*, 2015, **44**(6): 119-124.

陈彬,武宏阳,韩超,等.用中红外光谱法分析含微水绝缘油的粘度 [J]. *光子学报*, 2015, **44**(6): 119-124.

[11] CHEN Bin, LIU Ge. Analysis on near infrared spectroscopy of water content in oil using t-s fuzzy identifying model [J]. *Acta Photonica Sinica*, 2014, **43**(2): 0230001.

陈彬,刘阁.基于近红外光谱的润滑油中含水量 T-S 模糊辨识 [J]. *光子学报*, 2014, **43**(2): 0230001.

[12] ZHANG Yu, WU Di, HE Yong, et al. Non-invasive measurement of acidvalue of lubricant using visible and near infrared spectroscopy [J]. *Infrared*, 2011, **32**(12): 39-44. 张瑜,吴迪,何勇,等.基于可见-近红外光谱技术的润滑油酸值无损检测方法研究 [J]. *红外*, 2011, **32**(12): 39-44.

[13] LANGSHOLT M, ZARRUK G A. Particle transport in semi-dilute turbulent pipe flow [J]. *Journal of Dispersion Science and Technology*, 2015, **36**(10): 1513-1526.

[14] CHEN Bin, LIU Ge, ZHANG Xian-ming. Analysis on near infrared spectroscopy of water content in lubricating oil using successive projections algorithm [J]. *Infrared & Laser Engineering*, 2013, **42**(12): 3168-3174.

陈彬,刘阁,张贤明.连续投影算法的润滑油中含水量的近红外光谱分析 [J]. *红外与激光工程*, 2013, **42**(12): 3168-3174.

[15] CHEN Bin, WU Hong-yang, LIU Ge, et al. Detection on interfacial tension of insulating oil with trace water based on mid-infrared spectroscopy [J]. *Optics and Precision Engineering*, 2015, **23**(9): 2452-2458.

陈彬,武宏阳,刘阁,等.含绝缘油表面张力的中红外光谱检测 [J]. *光学精密工程*, 2015, **23**(9): 2452-2458.