

近红外光谱技术在推进剂质量检测中的应用

王菊香, 瞿 军, 邢志娜, 刘 洁
(海军航空大学岸防兵学院, 山东 烟台 264001)

摘 要: 近红外光谱分析方法具有过程简单、快速、结果准确的特点。介绍了近红外光谱技术在液体推进剂质量检测中的应用情况, 对光谱预处理、波段选择和建模方法进行了分析。通过平滑、微分、正交信号校正、小波变换等光谱预处理方法, 可有效去除高频噪音、基线漂移等对信息提取的影响。近红外光谱是含氢基团的倍频吸收峰, 根据待测组分的基团贡献选择合适的光谱范围建模, 可简化模型, 提高精度。偏最小二乘法是最常用的多元校正方法, 对多数分析项目可建立准确的校正模型。对含量低, 信息不丰富或受其他成分干扰严重的组分或性质可采用间隔组分偏最小二乘法和 BP-神经网络法, 以提高模型的准确度和预测能力。

关键词: 近红外光谱, 液体推进剂, 偏最小二乘法, 间隔组合偏最小二乘法, BP-神经网络法
中图分类号: V511-34 **文献标识码:** A **文章编号:** 1672-9374(2018)02-0082-06

Application of NIRS analysis technology in liquid propellant quality detection

WANG Juxiang, QU Jun, XING Zhina, LIU Jie
(Coast Defense Army College, Naval Aviation University, Yantai 264001, China)

Abstract: Application of near-infrared spectrum (NIRS) spectroscopy in liquid propellant quality determination is introduced in this paper. The spectrum preprocessing, waveband selection and modeling method of NIR are analyzed. The influence of high frequency noise and base line drift on information extraction can be effectively eliminated by spectral pretreatment methods, such as smoothing, differential coefficient, orthogonal signal correction or wavelet transform. NIRS is absorption peak of frequency multiplication of hydrogenous radical. Modeling in right wavelength range which is selected according to radical contribution of component under determination can simplify the model and improve accuracy of the analytical result. Partial least square (PLS), a common multivariate calibration method, is often used to set up the accurate calibration model for most analysis items. The interval partial least squares (iPLS) algorithm and BP artificial neural network (ANN) algorithm can be adopted for some components with low content and not abundant information or the components which are seriously disturbed by other components to improve the accuracy and predictive ability of calibration model.

Keywords: NIRS; liquid propellant; PLS; iPLS; BP-ANN

0 引言

推进剂是火箭和导弹的能源和工作介质,性能高低影响飞行速度或目标的命中率。液体推进剂在运输、贮存、转注过程中,由于挥发、吸湿、空气氧化及设备腐蚀等因素的影响,成分极易发生改变。因此使用单位对贮存过程中和加注前的液体推进剂都按技术要求进行分析化验,以保证质量。传统化验分析采用化学分析法和简单的仪器分析法,存在分析时间长、过程复杂、化学试剂使用多等缺点,且许多液体推进剂毒性大,化验过程中易造成人身伤害和环境污染。近红外光谱技术是样品不经处理直接测量、操作方便、无污染的一种快速分析测试技术^[1],在石油、农业、医药等领域应用广泛^[2-3]。采用近红外光谱分析技术不仅可提高液体推进剂质量检测的效率,还可实现现场和在线分析。近红外光谱定量分析是依靠所建分析模型实现的,建立准确性高、适用性强的分析模型是实现液体推进剂组分或性质快速准确测定的前提。文中介绍的分析方法对液体推进剂的红外光谱分析模型的建立进行了大量的研究工作,并建立了多种液体推进剂相关分析项目的准确的分析模型。

1 近红外光谱分析原理

近红外光谱一般分短波区(780~1 100 nm)和长波区(1 100~2 500 nm),主要是被测样品的含氢基团(C—H, O—H, N—H, S—H)的基频振动的倍频和合频吸收。液体推进剂一般都是含碳、含氮的有机化合物,含有丰富的含氢基团,在近红外光谱区有特征吸收。而液体推进剂的密度、运动粘度、闪点等物理性质与组成密切相关,随着组成的变化而变化。因此,采用近红外光谱可实现组成和理化性质的同时测定。具体研究过程包括常规方法测定一系列样品的相关分析项目,得到基础数据;进行光谱扫描得到光谱数据;进行光谱预处理、波段选择,用多元校正方法处理数据建立校正模型;模型验证,形成分析模型。

2 样品集的选择和数据测定

2.1 样品集选择

样品集是用于建立和验证分析模型的一系列

样品组合,理想的样品集要求如下:

- 1) 样品集的组成应与待测未知样品的组成基本一致;
- 2) 样品集浓度变化范围包含以后要测定的样品的浓度范围;
- 3) 组分浓度呈均匀分布;
- 4) 具有足够的样品数以便建立准确的定量关系。

为了获得满足上述要求的样品,可从生产厂家和使用单位贮存库分批次进行样品采集,然后按不同的比例混合。以保证得到校正集浓度范围宽,分布均匀,适应各使用单位的分析要求。

建立校正模型时,一般将样品集按K-S算法按约3:1的比例分成校正集和验证集,分别用于提高校正模型和验证模型的准确度。

2.2 基础数据的测定

采用传统的标准分析方法,对被测组分和性质进行化验,得到组分和性质的测定结果作为基础数据。为了保证测定结果的准确性,要求有化验分析经验的熟练化验员进行测定,并取满足重复要求的多次测定结果的平均值作为最终结果。

2.3 近红外光谱数据采集

为了保证在不同时间采集的近红外光谱具有一致性,要求光谱仪器稳定、可靠,环境温度和湿度变化小。一般样品池采用恒温控制,光谱采集过程温度变化不超过 $\pm 0.5\text{ }^{\circ}\text{C}$ 。连续采集10张光谱取平均值。

3 光谱预处理

近红外光谱仪所采集的光谱包含了除样品自身信息外的其他无关信息和噪音,如电噪音、样品背景和杂散光等。光谱预处理可消除与待测信息无关的信息和噪音,提高校正模型的准确度。

3.1 平滑处理

平滑处理可消除光谱的高频噪声,目前最常用的是S-G卷积平滑法。平滑处理的一个非常重要的内容就是窗口宽度的选择。一般情况下,高频噪声对近红外吸收光谱的影响较小,根据组成情况可适当进行5点或9点平滑。但对含量较低(1%以下)的组分,为了保留有用信息可不进行平滑处理。

3.2 微分处理

微分处理一般用于消除近红外光谱的基线漂移和其他背景干扰,分辨重叠峰,提高分辨率和灵敏度^[4-6]。在进行光谱的微分处理时,要合理选择微分级数和窗口宽度。一般情况下,处理效果无明显差别的情况下选择一阶微分。但对混胺燃料等颜色较深的样品,在近红外区受颜色影响较重(见如图 1),采用二阶微分可在一定程度上消除颜色的影响和混胺谱图基线的整体漂移。微分窗口宽度可通过校正模型标准偏差(SEC)的变化进行选择,图 2 为混胺水分的校正偏差随窗口宽度的变化曲线,当微分宽度为 4 nm 时,校正集标准偏差最低。

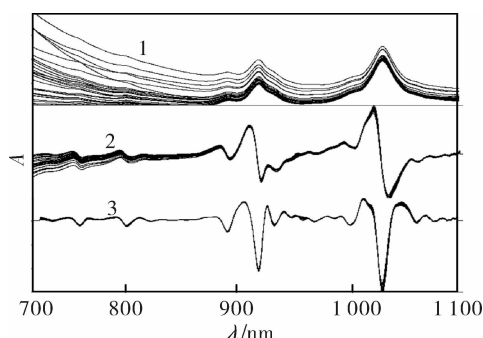


图 1 混胺的近红外光谱(1)和一阶(2)、二阶(3)微分谱图

Fig. 1 NIRS (1) of mixed amine and its 1st order (2) and 2nd order (3) derivative spectra

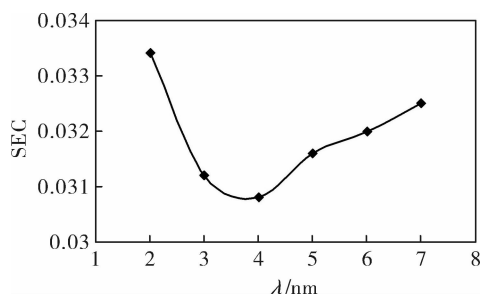


图 2 二阶微分宽度对混胺中水分 SEC 的影响
Fig. 2 Influence of 2nd - order derivative width on SEC in mixed amine

3.3 均值中心化

一些数据增强算法,如均值中心化、标准化和归一化等,可增强样品之间的差异,提高模型的稳健性和预测能力。一般近红外光谱都采用均值中

心化处理(见图 3),均值中心化处理后的谱图数据分布在零点的两侧,反映的光谱变化信息。

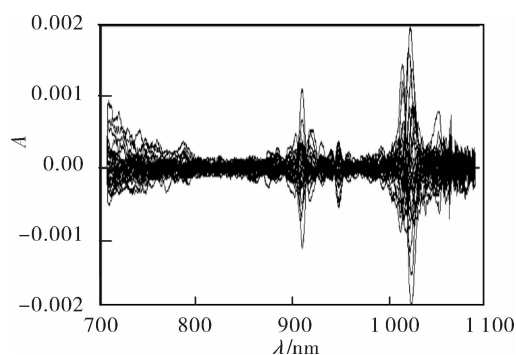


图 3 混胺光谱均值中心化处理

Fig. 3 mixed amine Spectra mean value after centralization preprocessing

3.4 正交信号校正

正交信号校正(OSC)是在建立定量校正模型前,将光谱数据与化学值正交,滤除与化学值无关的光谱信号^[7]。该法能在保持最大的有效光谱信息的前提下简化模型、提高准确度。正交信号校正过程中需要根据校正集或验证集的标准偏差(SEC或SEP)对主成分数进行选择。如 3 号喷气燃料密度正交处理的最佳主成分数为 2(见图 4)。

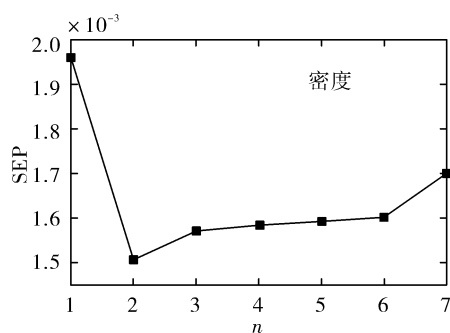
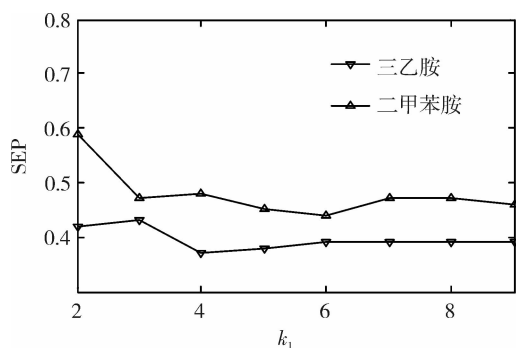


图 4 正交信号主成分数的选择

Fig. 4 Selection of principal component number of orthogonal signal

3.5 小波变换

小波变换(WT)具有多分辨率分析的优点^[8],常用与近红外光谱的去噪及压缩处理^[9-10]。小波变换预处理中需要根据模型的标准偏差选择小波分解层数。如混胺中三乙胺和二甲苯胺的小波分解层数分别为 6 和 4 时趋于稳定(见图 5)。

图5 SEP 随小波分解层数(k_1)的变化图Fig. 5 Variation of SEP with wavelet decomposition order (k_1)

3.6 预处理方法的选择

对光谱进行预处理要综合考虑分析组成或性质的特点,采用合理的预处理组合方法。表1是各预处理方法对混胺中三乙胺和二甲苯胺分析模型的影响。结果表明,三乙胺的预测分析中,只有WT和OSC结合时取得了良好的效果。而二甲苯胺的预测分析中,一阶微分-OSC和WT-OSC两种预处理方法的结合均取得较好的效果。

表1 光谱预处理方法的对校正模型的影响

Tab. 1 Effect of spectral pretreatment method on calibration model

预处理方法	三乙胺			二甲苯胺		
	最佳主因子数	SEC	SEP	最佳主因子数	SEC	SEP
一阶微分	8	0.49	0.51	8	0.45	0.52
平滑	10	0.43	0.50	10	0.51	0.56
WT	9	0.35	0.37	8	0.23	0.44
OSC	6	0.34	0.36	6	0.21	0.43
一阶微分-平滑	8	0.43	0.50	7	0.35	0.51
一阶微分-WT	9	0.36	0.44	8	0.35	0.41
一阶微分-OSC	5	0.36	0.43	4	0.21	0.33
WT-OSC	5	0.33	0.35	4	0.23	0.33

4 波段选择

近红外光谱波段选择主要剔除不相关或非线性变量信息,简化建模计算量,提高所建模型的预

测能力和稳健性^[11-12]。波段的选择一般依据分析项目的吸收信息以及与光谱吸收的相关性。如混胺主要是由三乙胺和二甲苯胺各占约50%组成的混合物,其近红外光谱主要是甲基、亚甲基和氨基等含氢基团的贡献。而混胺中会混入的少量水分主要是O—H键的二级倍频吸收,在940~965 nm处(见图6),选择此特征波段建模其他组分干扰较少。

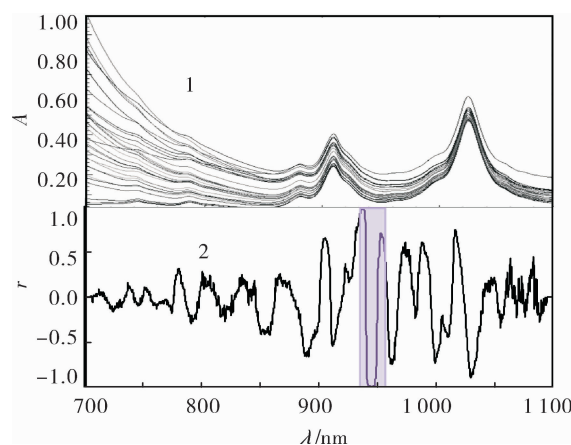


图6 混胺水分与光谱的相关图

Fig. 6 Correlation diagram of water in mixed-amine and NIRS

5 模型建立

化学计量学多元校正建模方法有多元线性回归(MLR)、主成分回归(PCR)、偏最小二乘(PLS)等线性方法和以局部权重回归(LWR)、人工神经网络(ANN)等非线性方法^[13]。其中PLS算法适用于多数分析项目的模型建立。而ANN作为非线性校正方法的代表,也越来越多的应用于近红外光谱分析。在液体推进剂的各项指标的近红外光谱分析中,对建模方法进行了研究,有偏最小二乘法(PLS)、改进的偏最小二乘法和针对微量组分测定的人工神经网络法等。

5.1 偏最小二乘法

在近红外光谱进行预处理和波段选择后,采用偏最小二乘法(PLS)建立校正模型时,必须确定建模的最佳主因子数,并通过校正集标准偏差(SEC)、验证集标准偏差(SEP)和相关系数(R^2)等参数对模型进行评价^[14-16]。一般采用交互校验法^[8]得到

的预测残差平方和(PRESS)最小确定最佳主因子。如图 7,三乙胺的最佳主因子数是 5。

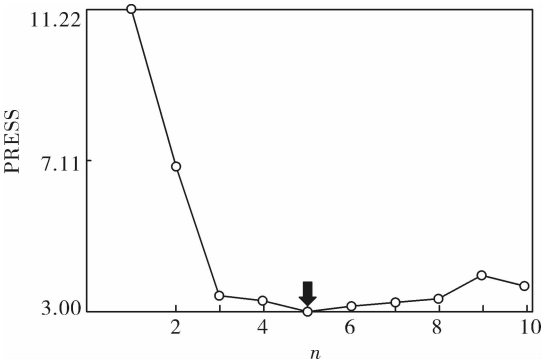


图 7 交互验证偏差平方和随主因子数的变化
Fig. 7 Variation of PRESS obtained by interaction validation with main factor number

5.2 间隔组合偏最小二乘法

间隔组合偏最小二乘法(icPLS)是通过近红外光分段建模来选择光谱区间的一种建模算法,以解决变量多、光谱复共线性多样化等问题对 PLS 建模效果的影响^[17]。图 8 是 3 号喷气燃料闪点的 ic-PLS 建模区间选择,将全谱按 40 个光谱点为单位分为 39 个区间,对每个区间分别进行 PLS 处理,根据 SEC 和 R^2 的大小进行组合,确定最佳建模区间。与全谱 PLS 模型相比,icPLS 对信息区间筛选和优化效果明显。

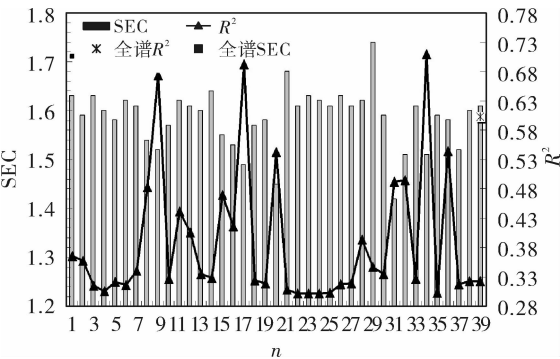


图 8 每个区间的最佳 PLS 模型 SEC 和 R^2
Fig. 8 SEC and R^2 of optimal PLS model in each spectra section

5.3 BP-神经网络法

对含量较低的理化指标由于受其他组分干扰严重,可考虑采用非线性的 BP-神经网络法(BP-

ANN)提高模型的预测能力^[18-19]。3 号喷气燃料总酸值较小,特征信息量易被干扰,BP-ANN 分析模型就取得较好的效果,常采用 3 层 BP 神经网络模型^[20]。影响 BP-ANN 预测结果的参数主要有隐含层神经元个数、学习速率、动量因子和学习次数。神经元个数与隐含层节点数和训练集样本数有关,可通过 SEC 和 SEP 来确定(见图 9),综合考虑总酸值模型信息处理能力和预测精度选择 10。

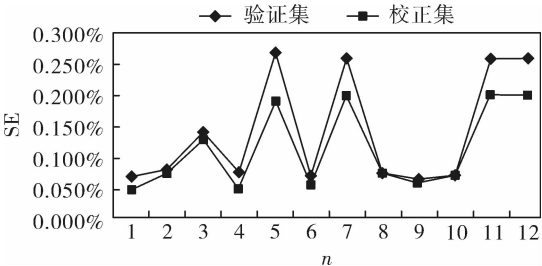


图 9 隐含层数对总酸值模型影响效果图
Fig. 9 Effect of number of hidden neurons on acid value model

学习速率影响网络收敛和网络性能,过大导致系统不稳定,过小导致训练时间长,网络收敛慢;动量因子影响网络建模能力,太小网络处理速度太慢,预测误差大,太大网络极易出现振荡陷入局部最小。如表 2,综合分析 SEC,SEP 和 R^2 确定最优学习效率和动量因子分别为 0.2 和 0.4。总酸值的 BP-ANN 和 PLS 模型验证集的 SEP 分别为 0.000 66 和 0.000 8,BP-ANN 模型预测精度和稳定性都好于 PLS 模型。

表 2 总酸值的学习速率和动量因子优化选择表
Tab. 2 Optimal selection of learning rate and momentum factors of total acid value

参数	学习速率				动量因子		
	0.1	0.2	0.3	0.4	0.3	0.4	0.5
SEP	0.18%	0.065%	0.063%	0.073%	0.069%	0.065%	0.047%
SEC	0.27%	0.061%	0.067%	0.072%	0.061%	0.061%	0.062%
R^2	0.193	0.979	0.977	0.977	0.977	0.979	0.977

6 结论

近红外光谱分析借助化学计量学方法软件实

现液体推进剂的快速分析。为了保证近红外光谱分析的准确性,应选用合适的预处理方法,选择最佳的建模波段,采用合适的多元校正方法,并变换参数进行反复建模,以确定每个分析项目的最佳校正模型。最后要通过对大量的未知样品的预测来确定模型的适用性和可靠性。由于近红外光谱分析是利用标准方法所测实验数据和光谱数据建立关联关系的二次分析方法,在实际使用中要经常对分析模型进行维护,通过仪器校准、补充编外样品或特殊样品来提高模型的准确性和稳健性。

参考文献:

- [1] 陆婉珍. 现代近红外光谱分析技术[M]. 北京:中国石油出版, 2007.
- [2] 刘福莉, 陈华才. 近红外透射光谱法检测三组分食用调和油含量的研究[J]. 光谱学与光谱分析, 2009, 29(8): 2099-2102.
- [3] 褚小立, 袁洪福, 陆婉珍. 近红外分析中光谱预处理及波长选择方法进展与应用[J]. 化学进展, 2004, 16(4): 528-542.
- [4] 王菊香, 邢志娜, 申刚, 等. 光谱预处理和波长选择对混胺燃料各指标近红外光谱定量分析结果的影响比较[J]. 计算机与应用化学, 2013, 30(1): 39-42.
- [5] 王菊香, 邢志娜, 刘洁, 等. 近红外波长选择结合偏最小二乘法测定混胺中微量水分[J]. 分析试验室, 2011, 30(11): 52-55.
- [6] 郑永梅, 张铁强, 张军, 等. 平滑、导数、基线校正对近红外光谱 PLS 定量分析的影响研究[J]. 光谱学与光谱分析, 2004, 24(16): 1546-1548.
- [7] 李华, 王菊香, 邢志娜, 等. 基于正交信号校正的喷气燃料近红外光谱去噪研究[J]. 计算机与应用化学, 2011, 28(1): 97-102.
- [8] 田高友, 褚小立, 袁洪福. 小波变换-偏最小二乘法用于柴油近红外光谱分析[J]. 计算机与应用化学, 2006, 23(10): 971-974.
- [9] 孙百红, 宋少伟. 基于小波分析的发动机转动惯量测量信号特征提取[J]. 火箭推进, 2010, 36(6): 52-55.
- SUN Baihong, SONG Shaowei. Feature extraction for measurement signal of engine rotary inertia moment based on wavelet analysis [J]. Journal of rocket propulsion, 2010, 36(6): 52-55.
- [10] 李华, 王菊香. 光谱预处理方法对混胺近红外定量模型影响的研究[J]. 分析科学学报, 2010, 26(5): 551-554.
- [11] 谷筱玉, 徐可欣, 汪曦. 波长选择算法在近红外光谱法中药有效成分测量中的应用[J]. 光谱学与光谱分析, 2006, 26(9): 1618.
- [12] 王菊香, 邢志娜, 叶勇, 等. 近红外光谱法快速分析液体推进剂组成和性质[J]. 理化检验化学分册, 2009, 45(7): 787-790.
- [13] 褚小立. 化学计量学方法与分析光谱分子技术[M]. 北京:化学工业出版社, 2011.
- [14] XING Zhina, WANG Juxiang, YE Yong, et al. Rapid quantification of kinematical viscosity in aviation kerosene by near-infrared spectroscopy [J]. Energy & fuels, 2006, 20: 2486-2488.
- [15] 王菊香, 申刚, 邢志娜. 近红外光谱快速测定混胺组分含量[J]. 分析化学, 2004, 32(4): 459-463.
- [16] 邢志娜, 王菊香, 申刚. CCD 近红外光谱分析技术在测定红烟硝酸中的应用[J]. 分析科学学报, 2004, 20(3): 278-282.
- [17] 邢志娜, 王菊香, 申刚, 等. 改进偏最小二乘法在航空煤油的近红外光谱分析中的应用[J]. 兵工学报, 2010, 31(8): 1106-1109.
- [18] 丁伟程, 吴建军, 刘洪刚. 基于神经网络算法的液体火箭发动机实时故障检测方法研究[J]. 火箭推进, 2005, 31(5): 5-10.
- DING Weicheng, WU Jianjun, LIU Honggang. Real-time fault detection method based on neural network algorithm for LRE [J]. Journal of rocket propulsion, 2005, 31(5): 5-10.
- [19] 韩晓, 王菊香, 刘洁, 等. 遗传算法结合神经网络用于傅里叶变换红外光谱法测定航空润滑油中水分[J]. 理化检验化学分册, 2012, 48(4): 388-391.
- [20] 韩晓, 王菊香, 刘洁. 基于 BP-神经网络的航空煤油总酸值近红外光谱快速检测[J]. 分析科学学报, 2011, 27(6): 751-75.