

# 近红外光谱分析技术在油品质量分析中的应用

姚成, 高俊

(南京工业大学理学院, 江苏南京 210009)

[摘要] 分析了近红外光谱在油品质量分析中应用的工作原理, 介绍了常用的几种近红外光谱定量、定性分析方法. 定量分析方法主要有多元线性回归(MLR)、主成分回归(PCR)、偏最小二乘法(PLS)与人工神经网络法(ANN)等; 定性分析方法包括峰位鉴别法、模式识别法等, 并对这几种方法进行了比较. 进一步介绍了近红外光谱技术在柴油、汽油等油品质量分析中的应用成果. 结果表明, 近红外光谱分析技术是分析油品质量的一种有效而便利的方法, 具有相当的工业应用价值.

[关键词] 近红外光谱, 油品, 定性分析, 定量分析

[中图分类号] O657.33, [文献标识码] A, [文章编号] 1672-1292(2004)02-0001-04

在炼油厂的油品分析中, 测定油品的许多物性参数, 如辛烷值、馏程、密度、烃类组成、折光率等, 应用常规的分析方法速度慢, 并且当质检部门得到数据时物流的性质可能已发生了变化. 近红外光谱(NIR)在线分析方法是一种间接测定方法, 使用先进的光谱统计技术和光缆技术, 能快速、准确地测定油品的辛烷值等物性参数, 连续、真实地提供油品的性质, 使装置生产得到优化. 现代近红外光谱技术是光谱测量技术、化学计量学技术和基础测量技术的有机结合, 可以快速、高效地对样品进行定性和定量分析<sup>[1-5]</sup>. 随着计算机技术的快速发展和化学计量学研究的日趋深入, 近红外光谱分析已成为20世纪90年代以来发展最快的分析测试技术之一. 近年来, 近红外光谱(NIR)作为快速分析技术已成功应用于许多领域, 其优越性是快速、低耗、无污染、能在线分析, 更适合于生产和控制的需要, 是典型的绿色分析技术. 在石油化工领域, 近红外光谱的最早也是最成功的应用是测定汽油的辛烷值和族的组成<sup>[6,7]</sup>. 此后多年中, 近红外光谱几乎用于石油化工的各个环节, 并取得了良好的应用效果和经济效益.

## 1 工作原理

近红外光谱(波长范围700~2500 nm)的产生主要是样品分子中含X-H键(X为C, O, N, S等)基团的化合物在中红外区域基频振动的倍频及合频吸收, 与中红外光谱相比, 谱带较宽且强度较弱, 因此在近红外光谱范围内样品不经稀释可直接在常规样品池中进行测定. 由于不同基团产生的光谱在吸收峰位和强度上有所不同, 随着样品组成的变

化其光谱特征也将发生变化, 这就为近红外光谱的定性和定量分析奠定了理论基础.

对油品而言, 石油产品主要为含有各种不同的C-H基团信息的烃类化合物, 例如甲基C-H(913 nm)、烯烃C=C-H(895 nm)、亚甲基C-H(934 nm)和芳烃C-H(875 nm)等基团信息<sup>[7]</sup>. 油品中不同结构的烃类化合物含量变化都会导致近红外光谱的变化, 尽管这种变化非常细微, 但通过化学计量学方法对光谱数据的处理, 便能得到样品组成变化的信息. 因此可将油品的近红外光谱和组成数据, 通过合理的谱图预处理(如平滑、微分)和化学计量学校正方法(如人工神经网络法)建立起近红外光谱与组成间的校正模型<sup>[8]</sup>, 通过建立的校正模型便可快速得到组成结果.

## 2 定量分析方法

近红外光谱分析方法由两个要素组成: 一是准确、稳定地测定样品的吸收或漫反射光谱图谱的硬件技术(即光谱仪器), 主要要求是保持长时间的稳定性; 另一个是利用多元校正方法计算测定结果的软件技术. 随着仪器制造技术与计算机技术的发展完善, 两个要素得以实现, 软件校正方法的研究也不断趋于成熟. 目前较常用的校正方法有多元线性回归、主成分回归、偏最小二乘法与人工神经网络等等.

近红外光谱应用于定量分析<sup>[9]</sup>的方法是: (1) 选取一组具有代表性的、已知化学测定值的样品作为校准集, 测量出其近红外光谱, 建立化学测定值与近红外光谱之间的定量数学模型(校准模型); (2) 取另一组已知化学测定值的样品作为预测集,

收稿日期: 2003-11-27.

作者简介: 姚成(1961-), 博士, 教授, 博士生导师, 主要从事现代分析及精细有机合成方面的教学与研究. E-mail: yaocheng@njtu.edu.cn

扫描出其近红外光谱,将光谱值代入校准方程得到样品的预测值,用预测值和化学测定值的相关系数和相对标准偏差来衡量所建模型的可靠程度;(3)若所建的校准方程稳定可靠,即可用此模型来对未知样品进行测定。在建立近红外定量校准方程时,通常使用逐步回归分析(SRA)、多元线性回归(MLR)、主成分回归(PCR)、偏最小二乘法(PLS)等来建立样品化学测定值与样品光谱参数之间的线性模型。但近红外吸收光谱中,光谱参数与样品含量化学测定值之间具有一定的非线性,特别是当样品的含量范围较大时,其非线性较明显。因此也可用人工神经网络(ANN)建立非线性模型。

### 3 定性分析方法

近红外光谱技术的定性分析是依据同类样品在不同波长下具有相同的光谱,借助多元分析方法,使各种样品能够进行聚类识别,结合某一类定量模型进而对化合物的各定量参数作出判断。近红外分析技术的定性与定量分析不同之处在于定量分析必须测定训练集的相应基础数据,精度取决于参考方法本身的精度;定性分析依据主要是光谱图谱本身,因为光谱反映了真实的样品组成和结构信息,相同或近似的样品有着相同或接近的光谱。

#### 3.1 近红外光谱峰位鉴别法<sup>[8]</sup>

在近红外光谱图谱上,依据不同种类油品所含化学成分的不同,含氢基团的倍频与合频振动频率的不同,反映的图谱的峰位、峰数及峰强也不同,成分差异越大,图谱的特征性越强。采用简易的峰位鉴别可对其品种进行鉴别。峰位鉴别试验要求对油品图谱进行一、二阶导数处理(常用的光谱预处理方法),共同比较光谱的原谱、一、二阶导数光谱进行分析。根据试验分析可知同种油品的原谱在不同波段可能有较小的差异,但其一、二阶导数光谱应当接近一致;而不同油品因所含化学成分的不同使得光谱图谱曲线完全不同。如柴油与汽油的光谱图谱,虽然这两种油品馏分中含有部分共有成分,但不同组分的差异掩盖了它们的共性,通过对光谱进行预处理,可以更清晰地鉴别出柴油与汽油。

#### 3.2 模式识别法<sup>[10]</sup>

采用峰位鉴别法主要是分析组分相差较远的不同种油品,这种方法直观、简便,但对于相近或同种油品之间进行鉴别就显得有心无力了。因此需要其它方法如模式识别法来判别。定性分析中采用的识别分析方法一般称为模式识别。在近红外光谱定性分析中常用的模式识别方法很多,主要有聚类分析法、主成分分析结合马氏距离法和人工神经网络

方法等。

#### 3.2.1 聚类分析法<sup>[11]</sup>

油样样品的聚类分析是采用光谱图谱的视图降维显示,将光谱数据矩阵(光谱图形)分解得到主成分载荷矩阵和得分矩阵,然后使用样本集的主成分对应其相应的得分作图,所以聚类分析法又被称为作图法。系统聚类法的主要过程是:先将每个样本自成一类,然后最相似的样本聚为一类,逐次比较新类间的距离并合并最相近的类,直到所有的样本聚为一个大类,最后得到按相似性大小构成的谱系图。由于近红外光谱的图谱特征不像红外光谱那样明显,用于表征油品的精细结构并不常用,主要是利用红外光谱特征中的差别,应用化学计量学方法进行归类。一般步骤是利用近红外光谱进行主成分分析判别,通过训练集样品光谱确定主因子数,来计算未知样品光谱的因子得分,并在多维空间确定其所属的类别,图谱分解的算法可以是主成分分析法(PCA),也可以是偏最小二乘法(PLS)。

#### 3.2.2 主成分分析结合马氏距离法<sup>[12]</sup>

由于一般作图分析直观但步骤繁琐,经常在试验报告中给出的是统计结果,实际操作中采用的方法是主成分分析法结合马氏距离法。为了确定未知物是否属于训练集中某种物质,一般应用马氏距离来度量未知样本与特定类之间的距离。由于同类物质组成的差异,未知样本与特定类在二维空间的分布呈现离散性。马氏距离正是应用具有同等大小、形状和取向的椭圆来描述训练集中的每一类物质。马氏距离概念可以使校正集数据的内部变化表现出来,用它来描述每一类样品在二维空间的位置及离散性比较合适。这样也便于计算未知样本与训练集中某一类物质间的距离,以便确定未知物的归属。目前聚类分析、马氏距离等判别方法在近红外定性分析方法中广泛使用,许多商业光谱分析软件也都具有相应的功能。在油品应用分析中可在文献中查到各个地区汽、柴油油品的不同指标十余种。随着大量的不同炼油厂油样样品的搜集,确定油样生产工艺与调和比例关系,可以清楚地分析出某一油品的级别、产地与工艺条件。

#### 3.2.3 人工神经网络法(artificial neural network, ANN)<sup>[13]</sup>

人工神经网络是由大量简单的处理单元广泛互连而形成的复杂网络系统,人工神经网络法是一种应用较多的定性分析方法。较常用于样品定性分析的是反向传播神经网络方法(back propagation BP),它是对大批量样品数据的非线性进行分析,建立模型并分类和预测样品。图1是一个3层BP

网络的结构图,它由3部分组成:输入层、隐含层及输出层,层与层之间采用全互连方式,同一层单元间不存在相互连接,图中圆圈表示神经元.模型是将光谱数据的前四维主成分信息作为神经网络的4个输入层结点,每个主成分都是所有波长点的线性组合.它包括结点,上一层的每一个结点都同下层的每一个结点连接,同外界的联系通过输入、输出层的结点进行,隐含于内部的中间层决定系统的计算能力.在ANN的每条连接线上是权重因子,代表系统记忆能力的重要部分.输入层输入的变量经权重加和后进入隐含层进行非线性变换,隐含层的输出经权重加和后进入输出层,经线性变换得到输出结果.在网络训练过程中,权重的数值可随着正在网络中流通的新信息而改变,输出的预测值( $Y$ )与期望值( $y$ )相比较,所得误差经反向传输,调整权重,使目标变量  $E = (y - Y)^T (y - Y)$  为最小( $T$ 为矩阵的转置),最终得到的权重用于人工神经网络的预测.根据Kolmogorov(映射网络存在)定理,3层BP神经网络可以以任意精度拟合任意连续函数,因而具有强大的非线性建模能力.使用人工神经网络进行建模,不需要事先知道模型的具体形式,因此特别适用于解决复杂的映射问题.

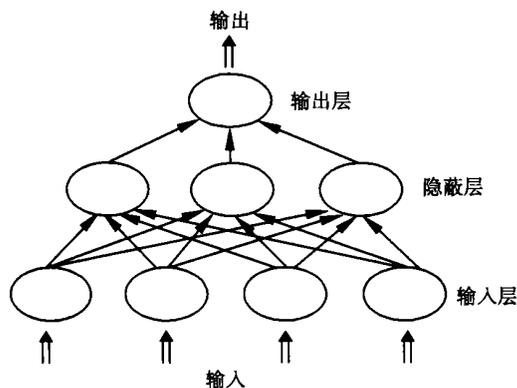


图1 反向传输神经网络的基本结构

### 3.2.4 各种方法的比较

聚类分析法主要应用于定性分析,结果比较直观,但需要作图.由于一般作图分析直观但步骤繁琐,经常在试验报告中给出的是统计结果,实际操作中采用的方法是主成分分析法结合马氏距离法.这种方法分类结果较好而且也能给出保证测量结果优劣的统计结果.人工神经网络的优点是对图形处理有较强的优越性,对于大批量的样品,进行全谱测定,具备较强的抗噪声和抗干扰能力.人工神经网络法是化学计量学中常用于非线性体系的重要校正方法之一.由于人工神经网络法可同时进行定性和定量的测定,这种方法被越来越多的应用于实际生产中.但是ANN应用于定量分析时,其输入

结点不能太多,否则迭代时间过长,得到的结果不理想.

## 4 近红外光谱分析技术在油品中的应用

近红外谱可用于测量原油中的饱和烃馏分、香烃馏分、松脂馏分及沥青馏分的含量<sup>[14]</sup>,也用于测量原油中的气/油比率(gas/oil ratio)<sup>[15]</sup>.

徐广通,陆婉珍<sup>[16]</sup>等对柴油性质的近红外光谱进行一系列研究,解释了柴油中各种C-H基团的存在对柴油十六烷值、密度、折光、馏程等性质的影响,将近红外光谱作为一种研究油品组成与性质关系的有效手段.可在文献上查到的有运用近红外光谱法测定柴油中的芳烃含量<sup>[17]</sup>、快速测定柴油物理性质<sup>[16]</sup>、测柴油十六烷值<sup>[18]</sup>、测定柴油闪点<sup>[19]</sup>、测定柴油凝点<sup>[20]</sup>等.

近红外光谱分析技术在汽油分析中运用较多,如近红外光谱快速测定汽油详细族组成<sup>[21]</sup>、测定汽油辛烷值<sup>[22]</sup>、测定汽油中乙醇<sup>[23]</sup>、对汽油模糊聚类分析<sup>[24]</sup>、在汽油调和中的运用<sup>[25]</sup>等等.还可查到运用近红外光谱法结合各类化学计量学方法的文献<sup>[26]</sup>.

## 5 结论

综合来讲,近红外光谱分析技术摆脱了传统实验室光谱分析方法拘泥于分析化合物精细化学结构的方式,可用于测定油品的等级、各种物性参数如凝点、辛烷值等,并可用于油品的调和,对分析油品质量提供了有效而便利的方法,具有相当的工业应用价值.因此,在油品分析领域中进行和建立近红外光谱分析方法是十分有益的.

以上简要介绍在油品质量分析中应用近红外光谱常用到的几种定性定量分析方法,实际工作中定性分析与定量分析在软件的实现上关联性较强.近红外光谱技术作为一种间接测量技术,校正集样品基础数据标准方法的测定是决定模型质量好坏的关键因素,但是建立一个好的模型将要耗费巨大的财力、物力和时间.因此在油品分析中,为节省建模费用,可尝试实现模型库共享和校正模型的可移植转换.随着近红外光谱仪器的标准化及各类人们需要的定性分析与定量分析模型的开发,实现模型共享、校正模型的可移植转换及多组分的同时测定,将会使近红外分析技术在油品分析领域得到更为广阔的开发和应用前景.

### [参考文献]

[1] Mc Clure W F. Near infrared spectroscopy: The giant is rurr

- ning strong[J]. Anal Chem, 1994, 66(1): 43A-53A.
- [2] Solomon S. Real time on line analysis of organic and non organic compounds for food, fertilizers and pharmaceutical product[P]. 世界专利: WO 9740361, 1997 10 30.
- [3] Lambert D, Descales B, Bages S, *et al.* Optimize steam cracking with online NIR analysis[J]. Hydrocarbon Processing, 1995, 74(12): 103-108.
- [4] Zilberman I, Bignan J, Sela I, *et al.* Use telecommunications for real-time process control[J]. Hydrocarbon Processing, 1996, 75(5): 91-97.
- [5] Alain E, Michel S, Sebastien O, *et al.* On line NIR analysis and advanced control improve gasoline blending[J]. Oil & Gas J, 1994, 92(42): 49-56.
- [6] Kelly J J, Barlow C H, Jinguji T M, *et al.* Prediction of gasoline octane numbers from near infrared spectral features in range 660~1250 nm[J]. Anal Chem, 1989, 61(4): 313-320.
- [7] Kelly J J, Callis J B. Nondestructive analytical procedure for simultaneous estimation of the major classes of hydrocarbon constituents finished gasoline[J]. Anal Chem, 1990, 62(14): 1444-1451.
- [8] 陆婉珍, 袁洪福, 徐广通. 现代近红外光谱分析技术[M]. 北京: 中国石化出版社, 2000. 146-172.
- [9] 吉海彦, 严衍禄. 主成分-人工神经网络在近红外光谱定量分析中的应用[J]. 分析测试学报, 1999, 18(3): 12-14.
- [10] 许禄. 化学计量学方法[M]. 北京: 科学出版社, 1995. 200-275.
- [11] 俞汝勤. 化学计量学导论[M]. 长沙: 湖南教育出版社, 1991. 158-163.
- [12] 史永刚, 冯新沪, 李子存. 化学计量学[M]. 北京: 中国石化出版社, 2002. 173-175.
- [13] Hagan M T, Demuth H T, Beale M H. 神经网络设计[M]. 戴葵, 宋辉, 谭明峰, 等译. 北京: 机械工业出版社, 2002. 197-257.
- [14] Aske N, Kallevik H, Sjöblom J. Determination of saturate, aromatic, resin, and asphaltenic (SARA) components in crude oils by means of infrared and near infrared spectroscopy[J]. Energy and Fuels, 2001, 15(5): 1304-1312.
- [15] Mullins O C, Terabayashi T, Kegasawa K. Gas/Oil ratio of live crude oils determined by near infrared spectroscopy[J]. Proceedings of SPIE, 2001, 4201(6): 73-81.
- [16] 徐广通, 袁洪福, 陆婉珍. 近红外光谱技术快速测定柴油物理性质[J]. 石油学报, 1999, 15(5): 63-68.
- [17] 徐广通, 陆婉珍, 袁洪福. 近红外光谱法测定柴油中的芳烃含量[J]. 石油化工, 1999, 28(4): 263-265.
- [18] 徐广通, 陆婉珍, 袁洪福. 近红外光谱测定柴油十六烷值[J]. 石油学报, 1999, 15(4): 65-66.
- [19] 王艳斌, 袁洪福, 陆婉珍, 等. 人工神经网络用于近红外光谱测定柴油闪点[J]. 分析化学, 2000, 28(9): 1070-1073.
- [20] 李子存, 史永刚, 冯新沪. 柴油凝点的近红外光谱法预测[J]. 分析仪器, 2000, 18(3): 30-32.
- [21] 褚小立, 袁洪福, 纪长青, 等. 近红外光谱快速测定重整汽油详细族组成[J]. 石油化工, 2001, 30(11): 866-870.
- [22] 康蕴天, 李敬清. 近红外光谱法测定汽油辛烷值[J]. 现代科学仪器, 2002, (1): 50-52.
- [23] 张金生, 李丽华, 曹艳君. 逐步多重线性回归-近红外光谱法测定汽油中乙醇[J]. 抚顺石油学院学报, 1998, 18(2): 1-6.
- [24] 刘海旺, 张光红. 基于近红外光谱的汽油模糊聚类分析[J]. 石油炼制与化工, 1998, 29(6): 59-61.
- [25] 马竞涛, 蔡大雄, 陆文琼, 等. 近红外光谱法在线测定在汽油调合中的应用[J]. 石油炼制与化工, 1998, 29(1): 67-68.
- [26] 史永刚, 冯新沪, 李子存. 化学计量学在近红外光谱定性分析中的应用[J]. 光谱实验室, 1999, 16(3): 237-239.

## Application of Near Infrared Spectroscopy in Oil Quality Analysis

YAO Cheng, GAO Jun

(School of Sciences, Nanjing University of Technology, Nanjing 210009, China)

**Abstract:** The article analyzed the working theory in the application of near infrared spectrum and briefly discussed several common technologies in qualitative and quantitative analysis of near infrared spectroscopy. Quantitative analysis methods include multivariate linear regression, principal component regression, partial least squares method and artificial neural network *etc.*; qualitative analysis methods include NIR spectrum peak position identification, pattern recognition *etc.*, and all these methods are compared. Moreover, the article introduced some achievements in application in gasoline, diesel oil quality analysis with near infrared spectroscopy. Near infrared spectroscopy methods are efficient and convenient for the application of analysis of oil quality, it has considerable applied value in industry.

**Key words:** Near Infrared Spectrum, Oil quality, Qualitative Analysis, Quantitative Analysis

[责任编辑: 严海琳]