第30卷,第3期
2010年3月

光 谱 学 与 光 谱 分 析
Spectroscopy and Spectral Analysis

Vol. 30, No. 3, pp663-666
March, 2010

衰减全反射傅里叶变换红外光谱技术结合模式识别进行油品鉴别

刘 倩[1,2], 孙培艳[2,3], 高振会[2,3], 蔡文生[1], 邵学广[1*]

1. 南开大学化学学院, 天津 300071
2. 国家海洋局海洋溢油鉴别与损害评估技术重点实验室, 山东 青岛 266033
3. 国家海洋局北海环境监测中心, 山东 青岛 266033

摘 要 海洋溢油是主要环境灾害之一, 而且近年来其发生频率呈上升趋势。快速地对油品进行种类鉴别、来源评估有利于及时采取应急措施, 因此具有重要意义。采用衰减全反射傅里叶变换红外光谱法(ATR-FTIR)对25种不同来源的油品进行了检测, 用不同数据预处理方法对原始光谱进行了预处理, 继而用主成分分析(PCA)和系统聚类分析(HCA)方法对光谱进行了分类鉴别。结果表明用多元散射校正(MSC)和连续小波变换(CWT)方法进行数据预处理可以提高分类的准确性, 使分类结果与油样的实际来源一致。该方法对正构烷烃差异较大的油品进行了很好的区分, 但对差异较小的油品其分辨能力仍有一定局限性。因此提供了一种快速的油品鉴别方法, 可用于溢油事件的初步鉴定, 从而为油品的进一步鉴定提供有用信息。

关键词 油品; 分类; 衰减全反射傅里叶变换红外光谱; 主成分分析; 系统聚类分析

中图分类号: O657.3 **文献标识码**: A **DOI**: 10.3964/j.issn.1000-0593(2010)03-0663-04

引 言

在海洋石油勘探与开发工作中, 极易因试油、运输、储存及其他原因造成突发性溢油事件, 不仅对海洋造成严重的污染, 而且会对周围的人类、动物、植物都造成不同程度的危害。因此对石油样品进行快速、准确的分析, 进而对样品进行来源鉴定、评估尤为重要。目前油类分析的主要方法有气相色谱-质谱法[1-3]、气相色谱-火焰离子化检测器[4]、红外光谱法[5]、紫外光谱法、荧光法[6]、同位素质谱法[7]、重量分析法[8]等。

红外光谱法作为一种快速、经济的油品分析方法, 在实验室溢油鉴定中应用较早[9]。随着便携式红外光谱仪的诞生, 其能进行原位分析, 红外光谱技术开始用于石油现场快速排查[10-11]。近年来, 化学计量学发展迅速, 近红外光谱结合模式识别方法在油品鉴别工作中也已有了重要应用。

衰减全反射傅里叶变换红外光谱法(ATR-FTIR)克服了传统的透射方法制样操作, 使样品不需要前处理即可进行测定, 大大地提高了样品分析速度。更重要的是, ATR-FTIR可以提供官能团、化合物类别等结构信息, 信号强, 且结合模式识别方法可以从数据中最大限度地提取信息, 并按照样品的某种性质对样品集进行分类。Andrate课题组曾将此方法应用到油品的类别鉴定中, 对"威望"号(Prestige)油轮燃料油泄漏事件中的相关油品进行了快速鉴定, 并对油品在不同条件的风化进行了监控[12-14]。但是油类样品基质复杂, 光谱背景较高, 且测量条件的变化易使光谱漂移, 会影响分类结果的准确性。因此, 利用ATR-FTIR结合模式识别方法进行油品鉴别时, 进行数据预处理很重要, 但目前有关此类工作的报道还很少。

本文旨在选取合适的数据预处理方法, 减小背景和噪声的影响, 从而提高分类结果的准确性。本工作采用ATR-FTIR对25种不同来源的油样进行检测后, 采用多种数据预处理方法对原始光谱进行背景校正和背景扣除, 然后用主成分分析(PCA)和系统聚类分析(HCA)对样品进行了分类。结果表明多元散射校正(MSC)[15]和连续小波变换(CWT)[16,17]预处理方法可以提高分类的准确性, 且分类结果与油样的实际来源一致。

收稿日期: 2009-03-18, **修订日期**: 2009-06-22

基金项目: 国家自然科学基金项目(20775036)和国家海洋局海洋溢油鉴别与损害评估技术重点实验室课题项目(200816)资助

作者简介: 刘 倩, 女, 1983年生, 南开大学化学学院在读硕士研究生 e-mail: liuq321@mail.nankai.edu.cn

*通讯联系人 e-mail: xshao@nankai.edu.cn

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

1 实验部分

1.1 仪器

Vertex 70 型 FTIR 光谱仪(德国 Bruker 公司),DTGS 检测器,KBr 分光器;P025-1151 型单反射 ATR 附件(德国 Bruker 公司),ZnSe 晶体;光谱数据处理采用 OPUS 软件(德国 Bruker 公司)。

1.2 油品

本文使用了 25 个油样,其中 18 个样品(S1-S11,D1-D7)为第一海域 A 油田的原油,样品 S1-S11 正构烷烃不丰富,样品 D1-D7 正构烷烃丰富;B 为第二海域 B 油田的原油,正构烷烃丰富;C1 和 C2 为第二海域 C 油田的原油,正构烷烃丰富;R1～R4 为正构烷烃丰富的燃料油,其中 R1 为柴油,R2 为机油,R3 和 R4 为柴油机油。所有的原油和燃料油未经过任何处理,直接用于测定。

1.3 光谱数据的采集

本实验设置的仪器工作参数为:光谱扫描范围 4 000～650 cm^{-1},扫描 64 次,分辨率 4 cm^{-1}。以空气为背景,进行光谱采集,图 1 为代表性光谱图(样品 D5)。每条光谱由 1 736个数据点组成。

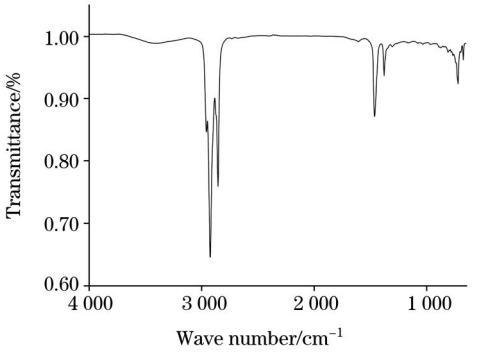

Fig. 1 An example of ATR-FTIR spectra of oil samples

2.4 光谱波段范围的选择

由图 1 可以看出,谱峰为 2 855,2 925,2 957 cm^{-1}的吸收带为脂肪族甲基和亚甲基(CH_2 + CH_3)的伸缩振动。谱峰为 1 460 和 1 380 cm^{-1}吸收带为脂肪族甲基和亚甲基(CH_2 + CH_3)的弯曲振动。谱峰 720 cm^{-1}的吸收带为长链的饱和烷烃。以上所述峰的强度代表了石油中饱和烷烃的丰富度,因此仅选择 2 975～2 685 cm^{-1}和 1 805～650 cm^{-1}的波段进行分类。

2 结果与讨论

2.1 主成分分析(PCA)

PCA 属于无监督式模式识别方法,是经典的特征抽取和降维技术之一,在不具备任何相关知识背景的情况下可以对未知样品进行类别归属。本文利用 PCA 对光谱进行分析,得到前 3 个主成分的累计贡献率大于 99.9%,因此选取前 3 个主成分绘制主成分得分三维谱图,见图 2。从图 2 可以看出,正构烷烃不丰富的 S 类与其他正构烷烃丰富类的分类不是很明显,且 D 类、B 类、C 类、R 类油样重叠也很严重。因此,原始光谱不经过预处理而直接进行聚类不能清楚地将正构烷烃差异较大的油品分开。

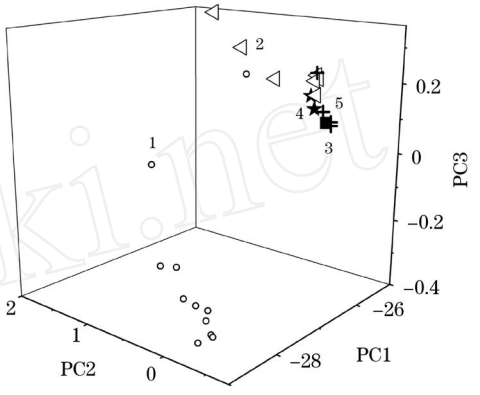

Fig. 2 PC scores scatterplots

1: S; 2: D; 3: B; 4: C; 5: R

2.2 数据预处理方法对分类结果的影响

数据归一化(normalization)、中心化(centering)和标准化(autoscaling)等都是数据增强的方法,能去除量纲影响、消除多余信息、增加样品之间的差异;MSC 和标准归一化(SNV)可以使信号的强度和尺度得到调整,有效地提高光谱的信噪比,从而达到背景校正的目的[18]。本文首先采用以上方法及不同方法的组合对原始光谱进行了预处理,得到的分类结果都不理想,即正构烷烃不丰富的 S 类与其他正构烷烃丰富类的分类仍未得到改善(S7,S11,D1 3 个样品分类不准确)。

小波变换是近年来发展起来的信号处理技术,由于小波基函数可以认为是某平滑函数的导数,因此 CWT 可以认为是导数和平滑技术的结合[16],CWT 用于导数的计算可以方便地解决高阶导数计算时的噪声问题,为分析信号、特别是信噪比较低的信号的导数计算提供了一种有效工具,而且它还可以扣除变动的背景[17,18]。因此,本文还使用了 CWT 作为数据预处理方法。

使用 CWT 方法进行数据预处理后,正构烷烃不丰富的 S 类与其他正构烷烃丰富类的分类较清楚,但是 S7 和 S11 样品与 S 类其他样品的距离还是较远。然后采用 MSC 和 CWT 预处理方法,经主成分分析得到前 3 个主成分的累计贡献率为 99.5%,选取前 3 个主成分绘制主成分得分三维谱图,如图 3 所示。由图 3 可以看出,25 个油样清楚地分为两大类,一类为正构烷烃不丰富的 S 类,另一类为正构烷烃丰富的其他油样。因此,正构烷烃含量差异较大的两组油品得到了完全正确的分离。图 3 与图 2 相比还可以看出,正构烷烃丰富的 D 类、B 类、C 类、R 类样品的分离也有所改善,但类别之间仍存在交叉。

为了更好地观察此结果,取前 3 个主成分,选用欧式距离和类平均法准则进行系统聚类分析,结果如图 4 所示。从图 4 也可以看出 25 个样品分成了两类,一类为正构烷烃不

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

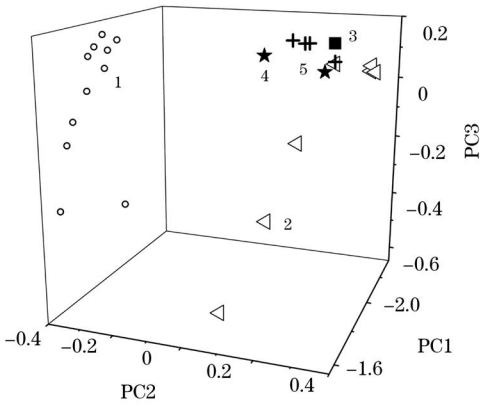

Fig 3　PC scores scatterplots with MSC and CWT

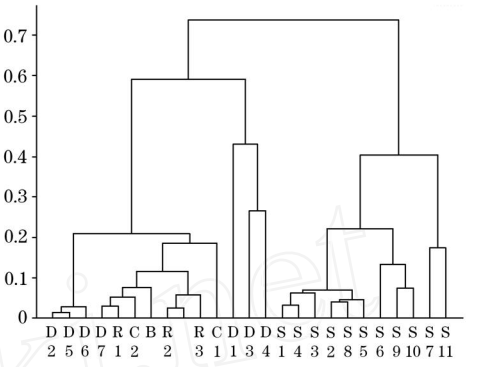

Fig 4　Dendrograms of the oil samples using the first three principal components with MSC and CWT

1: S; 2: D; 3: B; 4: C; 5: R

丰富的 S 类样品，另一类为正构烷烃丰富的其他样品。除了 R1，B，C1，C2 和 D7 外，正构烷烃丰富的 D 类、B 类、C 类和 R 类油样的分离也较好。由此可以说明 MSC 进行背景校正结合 CWT 进行背景扣除，不仅能将正构烷烃含量差异较大的两组油品完全正确分离，而且可以改善正构烷烃差异不大的样品的分离。但此结果也说明，该方法的分辨能力对差异较小的油品有一定局限性。

3　结　论

本研究采用衰减全反射傅里叶变换红外光谱结合主成分分析和系统聚类分析方法对石油样品进行分类鉴别，结果表明通过选择适当的光谱预处理方法进行背景校正和背景扣除，可以消除干扰信息，提高分类结果的准确性。此方法对正构烷烃差异较大的油品进行很好的区分，为油品的快速分类鉴别提供了一种可行的方法，但此方法对差异较小的油品其分辨能力仍有一定局限性，需要进一步的研究。

参　考　文　献

[1] Wang Z, Yang C, Hollebone B, et al. Environmental Science & Technology, 2006, 40 (18): 5636.
[2] Hupp A M, Marshall L J, Campbell D I, et al. Analytica Chimica Acta, 2008, 606(2): 159.
[3] Sun P, Bao M, Li G, et al. Journal of Chromatography A, 2009, 1216 (5): 830.
[4] Fernandez-Varela R, Andrade J M, Muniategui S, et al. Marine Pollution Bulletin, 2008, 56(2): 335.
[5] Mattson J S, Mark H B, Kolpack R L, et al. Analytical Chemisty, 1970, 42 (2): 234.
[6] Christensen J H, Hansen A B, Mortensen J. Analytical Chemisty, 2005, 77(7): 2210.
[7] Mansuy L, Philp R P, Allen J. Environmental Science & Technology, 1997, 31 (12): 3417.
[8] Villalobos M, Avila-Forcada A P, Gutierrez-Ruiz M E. Water Air Soil Pollution, 2008, 194(1-4): 151.
[9] Kawahara F K. Environmental Science & Technology, 1969, 3(2): 150.
[10] Permanyer A, Douifi L, Dupuy N, et al. Fuel, 2005, 84(2-3): 159.
[11] Permanyer A, Rebufa C, Kister J. Journal of Petroleum Science and Engineering, 2007, 58(3-4): 464.
[12] Fernandez-Varela R, Suarez-Rodriguez D, Gomez-Carracedo M P, et al. Talanta, 2005, 68(1): 116.
[13] Fernandez-Varela R, Gomez-Carracedo M P, Fresco-Rivera P, et al. Talanta, 2006, 69(2): 409.
[14] Fresco-Rivera P, Fernandez-Varela R, Gomez-Carracedo M P, et al. Talanta, 2007, 74(2): 163.
[15] Geladi P, MacDougall D, Martens H. Applied Spectroscopy, 1985, 39(3): 491.
[16] Shao X G, Leung A K M, Chau F T. Accounts of Chemical Research, 2003, 36(4): 276.
[17] Ma C X, Shao X G. Journal of Chemical Information and Computer Science, 2004, 44(3): 907.
[18] CHU Xiao-li, YUAN Hong-fu, LU Wan-zhen(褚小立，袁洪福，陆婉珍). Progress in Chemistry(化学进展), 2004, 16(4): 528.

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

Classification of Oils by Attenuated Total Reflectance-Fourier Transform Infrared Spectrometry Combined with Pattern Recognition Techniques

LIU Qian[1,2], SUN Pei-yan[2,3], GAO Zhen-hui[2,3], CAI Wen-sheng[1], SHAO Xue-guang[1*]

1. College of Chemistry, Nankai University, Tianjin 300071, China
2. Key Laboratory of Marine Spill Oil Identification and Damage Assessment Technology, State Oceanic Administration, Qingdao 266033, China
3. North China Sea Marine Monitoring Center, State Oceanic Administration, Qingdao 266033, China

Abstract In the present work, the combination of attenuated total reflectance-Fourier transform infrared spectrometry (ATR-FTIR) and pattern recognition, including principal components analysis (PCA) and hierarchical cluster analysis (HCA), is used as a fast and convenient analytical tool to classify oil samples. Twenty five samples including crude oils and fuel oils with different total contents of n-alkanes were analyzed. It was found that multiplicative scatter correction (MSC) and continuous wavelet transform (CWT) as a pretreatment method could improve the classification results of pattern recognition. The classification results were proved to be in agreement with the origin of the oil samples. The oils with high content of n-alkanes and those with low content were classified clearly by this developed method, but it still had some constraint to differentiating oils with little difference. The present work provides a feasible method for quick classification of oils, which can be used for the initial identification of spill oils and afford useful information for the further identification of the oils.

Keywords Oils; Classification; Attenuated total reflectance-Fourier transform infrared spectrometry (ATR-FTIR); Principal components analysis (PCA); Hierarchical cluster analysis (HCA)

(Received Mar. 18, 2009; accepted Jun. 22, 2009)

* Corresponding author

© 1994-2010 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net