

三维荧光光谱及平行因子分析在石油类污染物检测分析中的应用

潘 钊, 王玉田, 邵小青, 吴希军, 杨丽丽

燕山大学河北省测试计量技术及仪器重点实验室, 河北 秦皇岛 066004

摘要 以三维荧光光谱技术结合平行因子分析算法研究了石油类污染物识别与浓度测量方法, 重点分析了两种以上成品油共存时对测量的影响。以 0# 柴油、97# 汽油与煤油的 CCl_4 溶液为测量样本, 通过两种或三种油不同配比的混合溶液来模拟多种石油类污染物共存的状态, 研究平行因子分析方法在复杂混合物共存体系成分分析时的特点。实验分别针对汽柴油混合溶液、柴煤油混合溶液以及存在少量煤油干扰成分的汽柴油溶液, 分解得到各溶质的激发与发射特征光谱, 实现了各混合样品中主要成分含量的同时测量, 计算了平均回收率。结果表明, 该方法能够实现石油类污染物中共存成分的识别与浓度测量。

关键词 三维荧光光谱; 平行因子分析; 石油类污染物; 成分检测

中图分类号: X830.2, O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2012)03-0714-05

引言

石油类污染物对土壤、水环境等具有广泛的污染影响, 危害植物与动物的生长发育, 并会通过食物链等方式进入人体损害人体健康^[1]。无论原油还是汽油、柴油等石化产品都含有丰富的芳香烃成分^[2], 而芳香烃普遍具有较强的荧光特性^[3], 所以荧光方法成为识别与测量石油类污染物的重要方法。原油或其各种工业产品多含有相同或相似的芳香烃等荧光成分, 其荧光光谱是不同成分荧光叠加的产物。各种石油类物质共存时, 识别其成分种类和测定特定成分的浓度相对困难, 特别是针对环境中的微量石油类污染物, 很难用化学分离的方法将其逐个分离和分别测量。通常的检测方法并不能区别石油类污染物是原油或是某种石油产品, 只是测量其整体含量, 但在某些情况, 如分析污染物来源等则需要对石油类污染物种类进行详细区分。平行因子分析(PARAFAC)是一种针对三线性复杂数据的分析方法, 近年来在化学分析领域有较多应用, 多位研究者在此方面进行了大量的实验工作^[4-8]。本工作尝试使用平行因子分析方法对混合溶液的三维荧光光谱数据进行分析, 来实现以“数学分离”代替“化学分离”对混合溶液中成分进行识别与浓度测量。

1 测量原理

由多个样品的三维荧光光谱构成的数据阵列即是一组典

型的三线性数据。一个三线性数据阵列 \underline{X} 的 PARAFAC 模型由三个负载矩阵 \mathbf{A} , \mathbf{B} , \mathbf{C} 构成, 以数学形式可表示为

$$\underline{X} = \sum_{f=1}^F a_f \otimes b_f \otimes c_f \quad (1)$$

式中, a_f , b_f 和 c_f 分别为 \mathbf{A} , \mathbf{B} , \mathbf{C} 矩阵中的第 f 列, F 为分解运算的因子数, 也就是 \underline{X} 中的主成分数。 \otimes 表示张量积。

计算 PARAFAC 模型的基本方法是使用交替最小二乘法(ALS), 即假设两个负载矩阵已知, 利用它们去估算第三个负载矩阵。算法过程如下:

- ①确定模型的因子数 F , 即混合样品中主成分数;
- ②为负载矩阵 \mathbf{B} , \mathbf{C} 赋予设定的初始值;
- ③以初始化后的 \mathbf{B} 和 \mathbf{C} 为基础, 用式(2)来估计载荷矩阵 \mathbf{A} 的值,

$$\mathbf{A} = \mathbf{XZ}^T(\mathbf{ZZ}^T)^{-1} \quad (2)$$

式中, \mathbf{X} 为 \underline{X} 的展开, \mathbf{Z} 为由列向量 z_f 构成的矩阵, $z_f = b_f \otimes c_f$;

- ④以上一步得到的 \mathbf{A} 的估计值和 \mathbf{C} 为基础, 用同样的方法估计矩阵 \mathbf{B} ; 以 \mathbf{A} 和 \mathbf{B} 为基础估算矩阵 \mathbf{C} ;
- ⑤以各步得到的估计值重复③和④的过程, 逐次迭代直到收敛。

PARAFAC 方法分解结果具有唯一性, 而且具有化学意义, 与实际分析对象相一致。随着计算机运算速度的提高, PARAFAC 方法得以较方便地应用于化学计量学领域。

收稿日期: 2011-06-20, 修订日期: 2011-09-28

基金项目: 国家自然科学基金项目(60974115, 61071202)资助

作者简介: 潘 钊, 1982 年生, 燕山大学电气工程学院博士研究生

e-mail: panzh_zach@hotmail.com, panzh_zz@yahoo.com.cn

2 实验部分

2.1 样品配置

取市场所购某 97# 汽油、某 0# 柴油和普通煤油各 1 g, 分别加少量 CCl_4 溶解后, 分别置于 3 个 100 mL 容量瓶中, 用 CCl_4 定容, 得到 3 种 $0.01 \text{ g} \cdot \text{mL}^{-1}$ 标准溶液, 在常温下保存。

取 24 个 10 mL 容量瓶(编号 1#—24#), 加入不同体积各标准溶液, 加 CCl_4 定容, 得到不同配比石油类 CCl_4 溶液样本, 各样本中溶质浓度如表 1 所示。24 个样本中, 1#—12# 作为校正样, 13#—24# 作为预测样。

Table 1 Concentration of calibration and prediction samples ($\text{g} \cdot \text{mL}^{-1}$)

	校正样品			预测样品			
	柴油	汽油	煤油	柴油	汽油	煤油	
1	0.005	0.001	0	13	0.000 3	0.000 8	0
2	0.002	0.002	0	14	0.000 2	0.000 6	0
3	0.001	0.005	0	15	0.000 4	0.000 3	0
4	0.000 8	0.000 1	0	16	0.001 2	0.000 2	0
5	0.000 5	0.000 5	0	17	0.000 3	0	0.000 8
6	0.000 1	0.000 4	0	18	0.000 2	0	0.000 6
7	0.005	0	0.001	19	0.000 4	0	0.000 3
8	0.002	0	0.002	20	0.001 2	0	0.000 2
9	0.001	0	0.005	21	0.000 3	0.000 8	0.000 1
10	0.000 8	0	0.000 1	22	0.000 2	0.000 6	0.000 1
11	0.000 5	0	0.000 5	23	0.000 4	0.000 3	0.000 1
12	0.000 1	0	0.000 4	24	0.001 2	0.000 2	0.000 1

2.2 仪器

使用英国 Edinburgh Instruments FS920 稳态荧光光谱仪来获得样品荧光光谱。测量时设置激发/发射狭缝宽度均为 2.78 mm, 对应分辨率 5 nm, 积分时间 0.1 s。激发波长范围 300~400 nm, 荧光发射扫描范围 330~500 nm。发射扫描波长起点始终滞后激发波长 20 nm, 以充分避免瑞利散射光谱干扰。激发与发射步长均为 5 nm。

2.3 结果与讨论

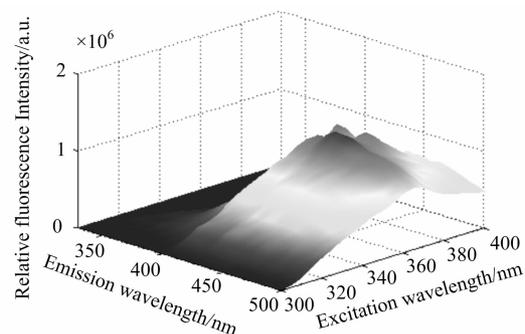
2.3.1 标准溶液的三维荧光光谱

以上述设置分别测量三种标准溶液, 得到荧光数据结果如图 1, 从图中可以看到在扫描范围内三种油都有明显的荧光发射, 虽然荧光强度各不相同, 但荧光峰具有明显的重叠区域。因此, 当两种或三种石油类物质共存于溶液中时, 难以使用光谱方法区分或定量测量其中某一种油的含量。

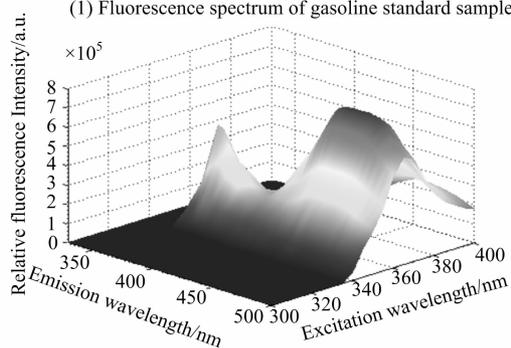
2.3.2 柴油与汽油的混合样品光谱

使用荧光光谱仪以相同的设置扫描柴油与汽油的混合溶液, 得到各混合样品的三维光谱。扣除 CCl_4 本底光谱, 将柴油、汽油标准溶液光谱与混合溶液光谱数据构造 $12 \times 35 \times 21$ 的三线性数据阵列 X_1 。使用核一致诊断方法(core consistency diagnostic)发现, 当因子数为 3 或 4 时, 其核一致系数小于 40%。而使用残差平方和(residual sum of squares)方法,

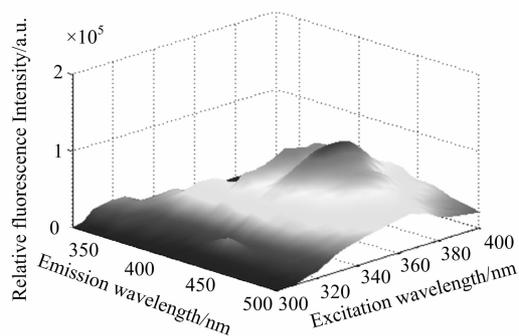
发现当因子数超过 2 时, 其残差平方和会显著下降, 所以选定因子数为 2, 这与实际情况相符。本工作不区分汽油等物质的具体化学成分, 而是将某一种油作为一个整体进行测量。计算 2 因子 PARAFAC 模型, 结果如图 2 所示。将测量得到的柴油、汽油标准溶液实际激发光谱、荧光发射特征光谱表示在对应的图中, 可以看出平行因子分解得到的激发与发射光谱与真实光谱十分吻合。对预测集中的两种油含量的预测结果和回收率如表 2 所示。其中对柴油成分浓度预测的平均回收率为 $97.4\% \pm 2.3\%$, 对汽油成分的为 $101.1\% \pm 0.9\%$ 。



(1) Fluorescence spectrum of gasoline standard sample



(2) Fluorescence spectrum of diesel standard sample

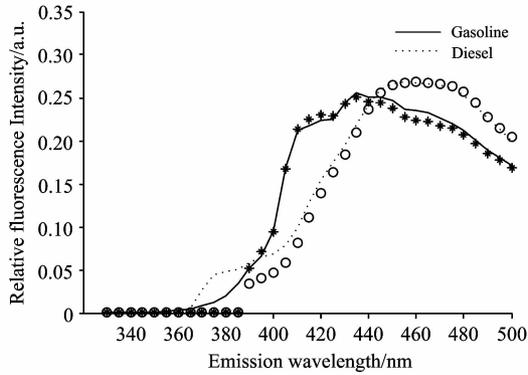


(3) Fluorescence spectrum of kerosene standard sample

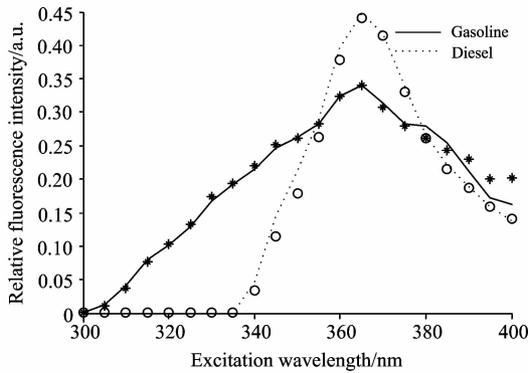
Fig. 1 3-D fluorescence spectra of three standard solutions

2.3.3 柴油与煤油混合样品的光谱分析

使用荧光光谱仪以相同的设置扫描柴油与煤油的混合溶液。扣除 CCl_4 本底, 得到混合样品的三维荧光光谱。将煤油与柴油标准溶液光谱与混合溶液光谱构造三线性数据阵列 X_2 。使用核一致方法判断因子数, 假设存在 5 个因子, 其核一致程度始终接近 100%。分析残余方差和发现, 2 因子分析其残余方差和即会显著低于 1 因子分析, 与实际样品构成也

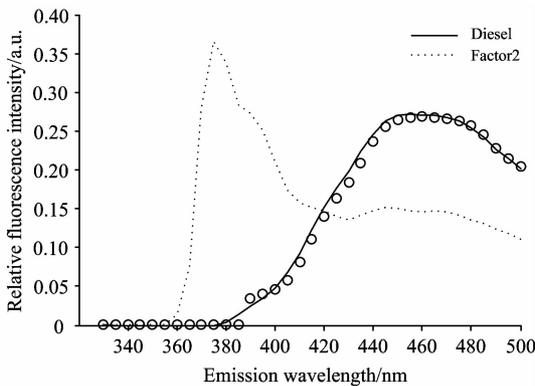


(1) Estimated and true emission spectra

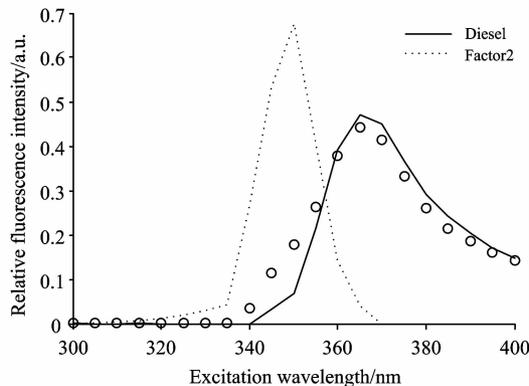


(2) Estimated and true excitation spectra

Fig. 2 The emission spectrum and excitation spectrum estimated by PARARAC



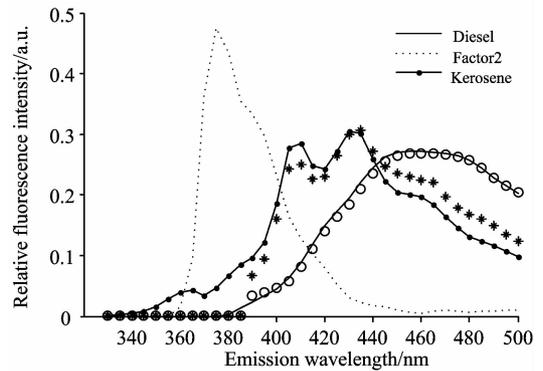
(1) Estimated and true emission spectra



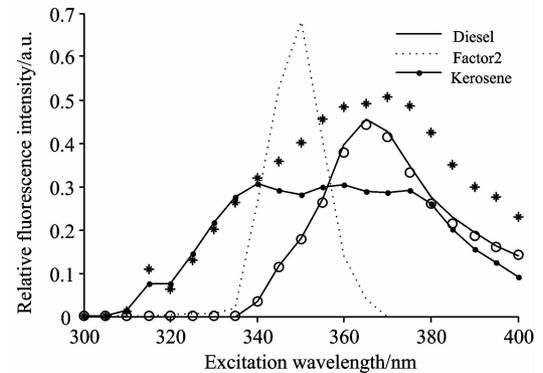
(2) Estimated and true excitation spectra

Fig. 3 The emission spectrum and excitation spectrum estimated by PARARAC(2 factors)

不同。若按照实际情况使用 $F=2$ 的平行因子分析,可以得到图 3 的激发和发射光谱。通过对比柴油和煤油标准溶液的三维荧光光谱,可以发现图中左侧的荧光峰其实来自柴油光谱,即柴油三维光谱中靠右侧的荧光峰,而不是煤油的荧光光谱。经分析认为此种现象是由于煤油溶液的整体光谱强度水平显著低于相同浓度的柴油溶液,甚至低于柴油中某些次要成分的荧光强度,以致算法将柴油光谱中强度较弱的次要荧光峰作为一个独立的成分分解出来。增加一个因子,对 X_2 做 3 因子平行因子分析得到激发发射光谱如图 4 所示。



(1) Estimated and true emission spectra



(2) Estimated and true excitation spectra

Fig. 4 The emission spectrum and excitation spectrum estimated by PARARAC (3 factors)

分解得到的估计光谱与柴油真实光谱十分吻合,煤油的估计光谱与真实光谱形状相似,但在和柴油光谱重叠的部分,强度较真实值偏低。对样品中的柴油和煤油浓度预测见表 2, 平均回收率分别为 $108.1\% \pm 7.3\%$ 和 $98.7\% \pm 2.7\%$ 。

2.3.4 存在煤油干扰成分的汽、柴油混合溶液光谱
在汽、柴油混合溶液中加入煤油溶液作为干扰成分。使用荧光光谱仪以相同的设置扫描预测集,扣除 CCl_4 本底,得到预测集的三维荧光光谱。将汽、柴油标准溶液光谱和校正集、预测集光谱共同构成三线性数据阵列 X_3 。对 X_3 做 2 因子平行因子分析,分析所得发射与激发光谱与 2.3.2 节分解结果基本相同。对其中柴、汽油含量预测如表 2 所示,平均回收率为 $96.4\% \pm 3.1\%$ 和 $105.6\% \pm 4.4\%$ 。从结果中可以看到,因为同等浓度的柴、汽油溶液其荧光强度水平相近,且显著高于干扰物煤油在同等浓度下的荧光强度,故在少量煤油干扰存在(不高于校正集中柴油或汽油最低含量)的条件

Table 2 Concentration and recovery of prediction samples

	预测浓度/(g·mL ⁻¹)			回收率/%			平均回收率/%		
	柴油	汽油	煤油	柴油	汽油	煤油	柴油	汽油	煤油
13	0.000 287	0.000 806	—	95.67	100.75	—	97.4±2.3	101.1±0.9	—
14	0.000 191	0.000 604	—	95.50	100.66	—			
15	0.000 395	0.000 303	—	98.75	101.00	—			
16	0.001 197	0.000 204	—	99.75	102.00	—			
17	0.000 339	—	0.000 798	113.00	—	99.75	108.2±7.3	—	98.7±2.7
18	0.000 229	—	0.000 599	115.00	—	99.83			
19	0.000 415	—	0.000 297	103.75	—	99.00			
20	0.001 21	—	0.000 192	100.83	—	96.00			
21	0.000 284	0.000 821	—	94.67	102.63	—	96.4±3.1	105.6±4.4	—
22	0.000 187	0.000 62	—	93.50	103.30	—			
23	0.000 392	0.000 319	—	98.00	106.30	—			
24	0.001 194	0.000 22	—	99.50	110.00	—			

下,可以比较准确的同时预测柴油与汽油在溶液中的浓度。

3 结论

使用荧光光谱仪测量获得某汽油、柴油和煤油样品的 CCl₄ 溶液及不同成分、含量的混合溶液三维荧光光谱,三种石油类物质由于成分相似,所以光谱既存在形状、强度的不同,同时在某些位置存在严重的重叠。应用平行因子模型分析了柴油与汽油、柴油与煤油以及有少量煤油干扰存在的柴

油与汽油混合溶液的荧光光谱数据,可实现以数学分离代替化学分离,同时识别和检测混合溶液中不同石油类成分。使用核一致方法与残余方差和分析结合样品的实际情况确定因子数,分解得到的激发/发射特征光谱与实测光谱吻合,说明选用的因子数适当。煤油成分因光谱强度较低,在与柴油共存时,其测量需要考虑柴油中具有较强荧光强度次要成分的干扰,需要结合激发、发射光谱判断因子的来源。在少量煤油作为干扰存在的条件下,平行因子方法可以比较准确的预测汽油、柴油在混合溶液中的含量。

References

- [1] ZHANG Xue-jia, JI Wei, KANG Zhi-jun(张学佳, 纪巍, 康志军). Science and Technology in Chemical Industry(化工科技), 2008, 16(6): 60.
- [2] NI Yong-nian(倪永年). Application of Chemometrics in Analytical Chemistry(化学计量学在分析化学中的应用). Beijing: Science Press(北京: 科学出版社), 2004. 178.
- [3] XU Bing-jiu(徐秉久). Instrument Analysis(仪器分析). Beijing: Peking University Medical Press(北京: 北京大学医学出版社), 2005. 198.
- [4] Rasmus Bro. Chemometrics and Intelligent Laboratory Systems, 1997, 38(2): 149.
- [5] LIU Hai-long, WU Xi-jun, TIAN Guang-jun(刘海龙, 吴希军, 田广军). Chinese Journal of Lasers(中国激光), 2008, 35(5): 685.
- [6] HUANG Lei, WU Hai-long, NIE Jin-fang, et al(黄蕾, 吴海龙, 聂瑾芳, 等). Computers and Applied Chemistry(计算机与应用化学), 2010, 27(7): 893.
- [7] GUO Wei-dong, YANG Li-yang, WANG Fu-li, et al(郭卫东, 杨丽阳, 王福利, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2011, 31(2): 427.
- [8] LÜ Gui-cai, ZHAO Wei-hong, WANG Hong-tao(吕桂才, 赵卫红, 王红涛). Chinese Journal of Analytical Chemistry(分析化学), 2010, 38(8): 1144.

Application of PARAFAC Method and 3-D Fluorescence Spectra in Petroleum Pollutant Measurement and Analysis

PAN Zhao, WANG Yu-tian, SHAO Xiao-qing, WU Xi-jun, YANG Li-li

Measurement Technology and Instrumentation Key Lab of Hebei Province, Yanshan University, Qinhuangdao 066004, China

Abstract A method for identification and concentration measurement of petroleum pollutant by combining three-dimensional (3-D) fluorescence spectra with parallel factor analysis (PARAFAC) was proposed. The main emphasis of research was the measurement of coexisting different kinds of petroleum. The CCl_4 solutions of a 0[#] diesel sample, a 97[#] gasoline sample, and a kerosene sample were used as measurement objects. The condition of multiple petroleum coexistence was simulated by petroleum solutions with different mixed ratios. The character of PARAFAC in complex mixture coexisting system analysis was studied. The spectra of three kinds of solutions and the spectra of gasoline-diesel mixed samples, diesel-kerosene mixed samples, and gasoline-diesel mixed with small counts of kerosene interference samples were analyzed respectively. The core consistency diagnostic method and residual sum of squares method were applied to calculate the number of factors in PARAFAC. In gasoline-diesel experiment, gasoline or diesel can be identified and measured as a whole respectively by 2-factors parallel factors analysis. In diesel-kerosene experiment, 2-factors parallel factors analysis can only obtain the characters of diesel, and the 3rd factor is needed to separate the kerosene spectral character from the mixture spectrum. When small counts of kerosene exist in gasoline-diesel solution, gasoline and diesel still can be identified and measured as principal components by a 2-factors parallel factor analysis, and the effect of interference on qualitative analysis is not significant. The experiment verified that the PARAFAC method can obtain characteristic spectrum of each kind of petroleum, and the concentration of petroleum in solutions can be predicted simultaneously, with recoveries shown in the paper. The results showed the possibility of petroleum pollutant identification and concentration measurement based on the 3-D fluorescence spectra and PARAFAC.

Keywords Three-dimensional fluorescence spectra; Parallel factor analysis; Petroleum pollutant; Concentration measurement

(Received Jun. 20, 2011; accepted Sep. 28, 2011)