

水体多环芳烃组分识别小样本分析方法研究

祝 玮^{1,2}, 杨瑞芳^{1*}, 赵南京^{1*}, 殷高方¹, 肖 雪¹, 刘建国¹, 刘文清¹

1. 中国科学技术大学环境科学与光电技术学院, 安徽 合肥 230026

2. 中国科学院合肥物质科学研究院, 安徽光学精密机械研究所, 环境光学与技术重点实验室, 安徽 合肥 230031

摘 要 多环芳烃(PAHs)是一类在自然环境中常见且广泛存在的有毒有害有机物。其主要来源有自然界的各种微生物以及植物的生物合成, 富含植被区域的天然火灾, 火山的喷发物, 化石燃料以及人为工业碳氢化合物的不完全燃烧和运输过程中的石油泄漏等。多环芳烃的毒性较为强烈, 具有生物致癌性, 遗传毒性和致突变性。它对于人体呼吸系统, 循环系统, 神经系统有着多方面的危害, 是一种重要的有机污染物, 因此有必要对多环芳烃的现场监测和分析方法进行研究。目前对于多环芳烃的分析方法主要有化学分析法和光谱分析法。化学分析法包含有前处理的化学滴定法, 液相色谱法(LC), 高效液相色谱法(HPLC), 气相色谱质谱法(GC-MS); 光谱学分析法涉及紫外吸收光谱, 荧光光谱和三维荧光光谱等。三维荧光光谱同时获得激发波长和发射波长的信息, 因而包含的光学信息十分丰富, 灵敏度高, 光谱特征显著, 在实际水体的现场检测和水体样本混合组分的快速研究有明显的优势。常见的三维荧光光谱解析方法有平行因子分析法(PARAFAC), 多维偏最小二乘法(N-PLS)等。平行因子分析是分析多环芳烃重叠三维荧光光谱的一种有效方法。但有时由于多种组分的荧光较弱, 它对三维荧光光谱的欠定分析并不能得到令人满意的结果。为了从两个样品中提取更多的成分, 提出一种基于奇异值分解(SVD)和 PARAFAC 的方法。首先对每个观测样本进行奇异值分解, 根据累积贡献率选取合适的奇异值, 构造新的伪样本来突出微弱的荧光信号。然后, 将两个观测样品及其对应的伪样品输入 PARAFAC, 恢复组分光谱。为验证所提方法的有效性, 对三组不同荧光强度的多环芳烃重叠三维荧光光谱进行了分析。结果表明, 从两个混合样品中提取并识别出 6 个多环芳烃的纯组分光谱, 其分辨发射和激发光谱与标准光谱的相似性均在 0.80 以上。

关键词 欠定分析; 弱荧光信号; 平行因子分析; 奇异值分解; 多环芳烃

中图分类号: O433.4 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)11-3494-07

引 言

来自自然和人为来源的多环芳烃(PAHs)几乎存在于所有自然水环境中^[1-2]。它们是有毒有害的有机污染物, 对动物和人具有引发突变和致癌的威胁^[3-4]。因此对实际水体中多环芳烃的快速现场监测十分重要。化学色谱法是一种理想的测量多环芳烃的方法, 但是有比较耗费时间和耗费金钱的样品前处理, 包括分析物的提纯和预先浓缩^[5-6]。目前化学色谱法的替代方法是三维荧光光谱技术(TFST)^[7-8]。TFST具有较高的灵敏度, 是一种很有前途的快速、无损地现场监测水生环境中荧光多环芳烃的技术^[9-10]。它同时测量多环芳

烃的激发和发射矩阵来获取各多环芳烃的荧光光谱信息, 可作为区分不同种类多环芳烃^[11]的特征指纹。然而由于部分多环芳烃的结构相近和自然水样的成分复杂, 以致造成荧光峰重叠和未知荧光组分的干扰导致三维荧光光谱缺乏一定的选择性, 阻碍了其直接识别水中多环芳烃污染的实用性。因此, 成分光谱提取成为多环芳烃三维荧光光谱分析的主要研究热点之一, 并发展了包括平行因子分析(PARAFAC)在内的多种多维分辨率方法^[12-14]。

平行因子分析是当前应用最广泛的分析方法之一, 它将三维荧光光谱数据组直接分解为相对浓度分数, 激发光谱和发射光谱, 从而获得底层荧光组分^[15-17]。通过对多环芳烃^[18]的激发和发射光谱与标准荧光光谱的比对, 可以实现组分的

收稿日期: 2022-04-15, 修订日期: 2022-09-19

基金项目: 国家自然科学基金项目(61805255), 中国科学院仪器设备功能开发技术创新项目(E03HBF11291), 中国科学院合肥物质科学研究院院长基金“融合专项”重点支持项目(YZJJZX202009)资助

作者简介: 祝 玮, 1992 年生, 中国科学技术大学环境科学与光电技术学院博士研究生 e-mail: 360619273@qq.com

* 通讯作者 e-mail: rfyang@aiofm.ac.cn; hjzhao@aiofm.ac.cn

定性识别。以上所引的研究工作都是试验混合物的数量大于或至少等于纯组分的数量的情况。但在实际应用中, 样品数量过多是一项费时费力的工作。因此, 从较少检测的样品中提取和识别更多的组分, 即多组分三维荧光光谱的欠定分解, 具有实际意义。在光谱分析领域对欠定分解的研究相对较少^[19-20]。PARAFAC 可以从较少的样品中回收更多的成分。但是, 由于一些信号的荧光发射较弱, 对于过少的样品, 它可能会给出不准确的结果。奇异值分解(singular value decomposition, SVD)是一种有效的去噪工具, 可以实现信号与噪声之间的无偏微分^[21-22]。奇异值分解去噪的基本原理是保留奇异值显著的奇异分量。受此启发, 利用 SVD 对微弱荧光信号的高亮进行了研究。

这里提出了一种基于奇异值分解(SVD)和 PARAFAC 的欠定分解方法。该方法通过选取有效奇异值重构观测样品的激发发射矩阵结构, 解决了微弱荧光信号的突出问题。根据奇异值的特点, 确定激励发射矩阵的最优结构作为信号的新的伪样本。将观测样品与新样品结合, 利用 PARAFAC 得到纯组分光谱。最后, 通过回收光谱与标准光谱的比较, 识别出污染物成分。以多环芳烃为例, 从两种多环芳烃混合物的观测光谱中成功地恢复了 6 种纯组分荧光光谱, 并利用相似系数进一步评价其组分识别的效率。

1 方法原理

1.1 三维 PARAFAC 方法

作为一种三线性算法, 三维 PARAFAC 可以对三维阵列进行如下分解(Bro 1997; Nikolajsen 等 2003 年)^[15, 23]

$$x(i, j, k) = \sum_{f=1}^F a_{if} b_{jf} c_{kf} + \epsilon_{ijk}$$

$$i = 1, \dots, I; j = 1, \dots, J; k = 1, \dots, K \quad (1)$$

式(1)中, $x(i, j, k)$ 是 i 样品的一系列尺寸为 $J \times K$ (发射波长 j , 激发波长 k) 的激发-发射矩阵“叠加”而产生的三维矩阵 $x(I \times J \times K)$ 的元素。其中 I 为样本数, J 为发射波长, K 为激发波长, F 为因子数, 每个 F 对应一个平行因子分量。 a_{if} 为第 i 个样品中 f 组分的相对浓度, b_{jf} 和 c_{kf} 分别为 f 组分在 λ_j 波长处的发射和检测波长 λ_k 处的波长激发。将 a_{if} , b_{jf} 和 c_{kf} 的所有元素收集到三个加载矩阵 \mathbf{A} , \mathbf{B} 和 \mathbf{C} 中, 分别对应浓度分数矩阵, 就可以计算出相应的发射光谱和激发光谱。

1.2 奇异值分解

开始的 SVD 过程将数据矩阵 \mathbf{G} 分解为三个矩阵 $\mathbf{U}_{M \times N}$, $\mathbf{S}_{N \times N}$ 和 $\mathbf{V}_{N \times N}$

$$\mathbf{G} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (2)$$

式(2)中, \mathbf{V}^T 是矩阵 \mathbf{V} 的转置, $\mathbf{S}_{N \times N}$ 是一个对角矩阵, 其角元素为奇异值。利用 s 的 $s < N$ 个元素(记为 \mathbf{S})和对应的 \mathbf{U} 的列(记为 \mathbf{U})和 \mathbf{V}^T 的行(记为 \mathbf{V}^T)重构方程(2), 得到新的数据矩阵 \mathbf{G}

$$\mathbf{G} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (3)$$

奇异值分解通过分析有意义或不显著的奇异值/向量, 将有价值的信号从噪声中分离出来。

2 实验和结果讨论

实验以蒽(AN)、菲(PHE)、芘(PY)、芴(FLU)、蒽(ACE)和荧蒽(FLA)这六种多环芳烃(PAHs)为分析对象。不仅因为它们的光谱重叠较广, 而且荧光强度差异较大。在相同浓度下, AN, PY 和 FLU 的荧光强度相对于 PHE, ACE 和 FLA 较强, FLA 为最弱的荧光发射。以去离子水为背景制备样品组, 其中 AN, PHE, PY, FLU, ACE 和 FLA 的浓度分别为

$$2/4/6/10/3/9 \mu\text{g} \cdot \text{L}^{-1} \quad (X1)$$

$$7/3/1/2/5/4 \mu\text{g} \cdot \text{L}^{-1} \quad (X2)$$

$$1/2/3/5/7/10 \mu\text{g} \cdot \text{L}^{-1} \quad (X3)$$

$$8/5/2/7/1/3 \mu\text{g} \cdot \text{L}^{-1} \quad (X4)$$

$$3/6/9/4/10/8 \mu\text{g} \cdot \text{L}^{-1} \quad (X5)$$

$$9/8/7/9/3/2 \mu\text{g} \cdot \text{L}^{-1} \quad (X6)$$

以此来验证算法的有效性。

每次测试对由两个样本组成的一组样本进行带奇异值或不带奇异值的三维 PARAFAC 分解。通过比较 PARAFAC 模型的计算谱与纯组分溶液标准谱的相似程度, 使用相似系数来评价谱图的分辨性能。计算得到的光谱与标准光谱之间的相似系数根据方程(4)(Tucker 1951(A Method for Synthesis of Factor Analysis Studies))计算

$$r = \frac{x \times s^T}{\|x\| \|s\|} \quad (0 \leq r \leq 1) \quad (4)$$

式(4)中, x 为计算得到的光谱向量, s 为标准光谱向量。系数 r 反映了计算得到的光谱与标准光谱的相似程度。 r 值越大, 计算谱 x 与标准谱 s 的相似度越高, 通常认为相似性系数小于 0.8 的成分识别不成功。

首先对第一个实验数据集(X1 和 X2)采用直接三线性分解的三维 PARAFAC 分解进行拟合, 根据式(1)进行初始化。在运行 PARAFAC 之前, 需要确定组分数。所有样品都是在去离子水中加入 6 个不同比例的荧光多环芳烃制备而成。预测组分数需要大于等于添加的多环芳烃的种类数, 所以选择 6, 7, 8 和 9 为可行的组分数。图 1 给出了 6 个分量重建的发射和激发光谱与对应的 6~9 个分量的标准光谱的相似系数。FLA 和 ACE 的激发光谱(Ex)或发射光谱(Em)相似系数均小于 0.80, 成分数为 6; PHE 和 FLA 的相似系数低于 0.80, 成分数为 7; FLA 和 ACE 的相似系数低于 0.80, 成分数为 8; FLA 和 ACE 的相似系数低于 0.80, 成分数为 9。

为了提高 PARAFAC 对两个样本光谱的分辨性能, 可以再构造两个伪样本; 构造伪样本的原则是, 当累积贡献率首次大于 99.9%时, 相对应的奇异值视为强信号, 而将此奇异值前的所有奇异值设为零, 具体方法如下:

(1)对于每个被测样本, 按照式(2)计算奇异值;

(2)奇异值的累积贡献率计算如式(5)

$$R(i) = (A(1)^2 + \dots + A(i)^2) / (A(1)^2 + \dots + A(N)^2)$$

$$i = 1, \dots, N \quad (5)$$

(3)为了提高微弱荧光信号的分辨率, 根据式(3)重新构

造新的样本, 将第 i 奇异值设为 0。

其中 $A(i)$ 为第 i 个奇异值。 N 是奇异值的个数。当 $R(i)$

的幅值首次大于 99.9% 时, 对应的第 i 个奇异值视为强信号。

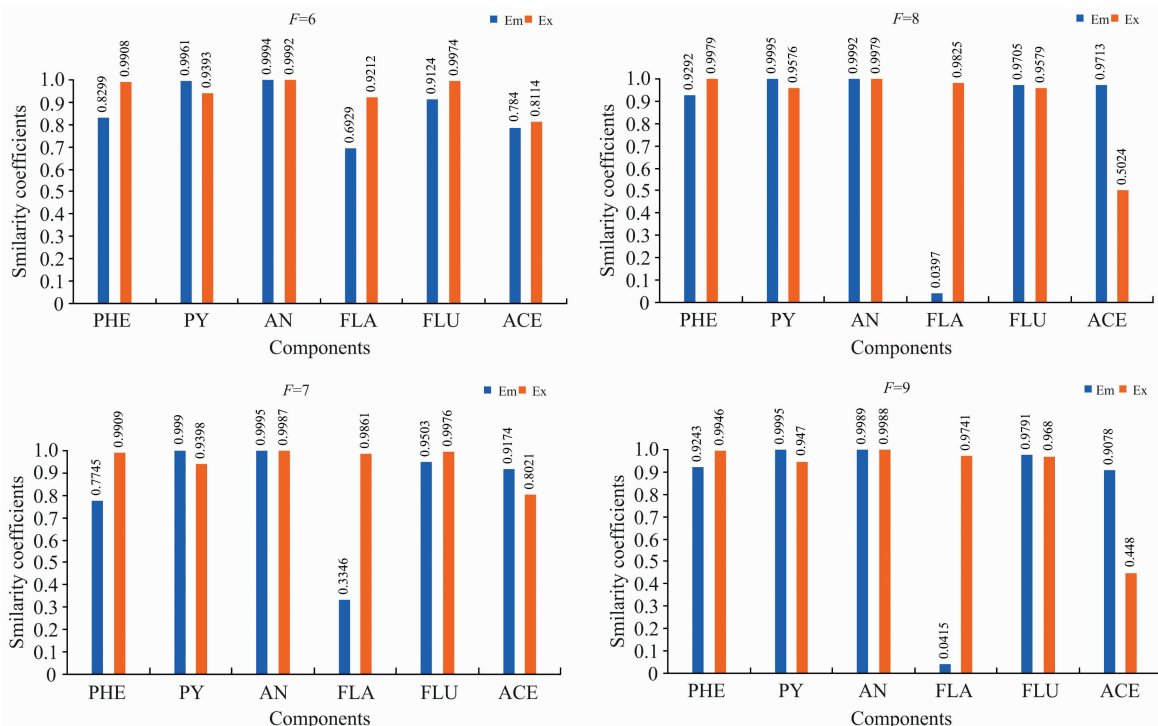


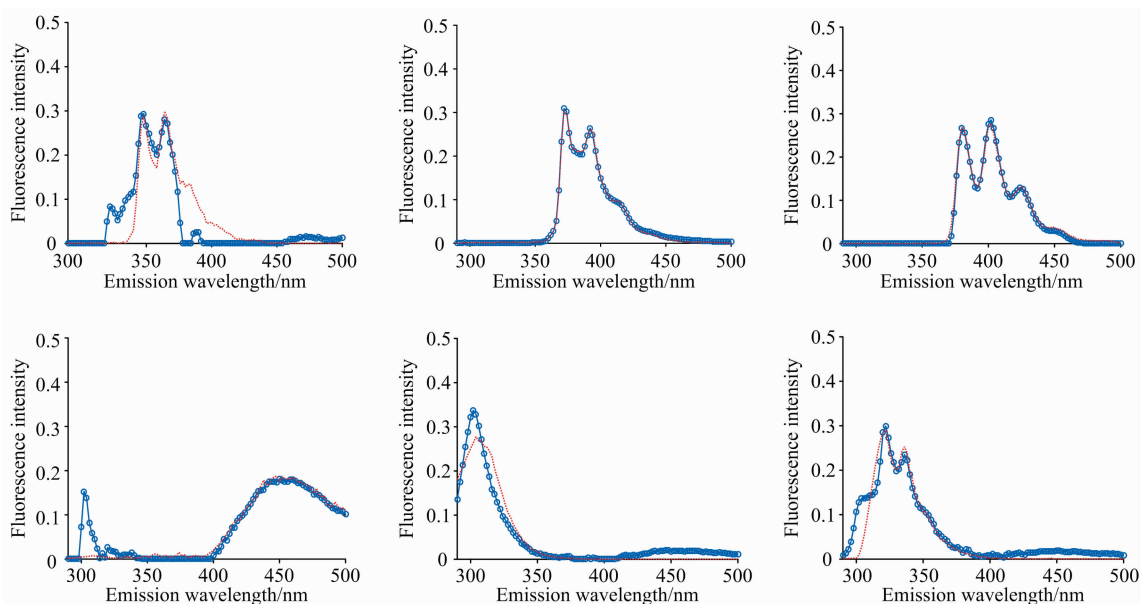
图 1 可行组分数设置从 6 到 9 下样本集 X1 和 X2 中的六种不同组分激发/发射光谱和标准光谱对比的相似系数

Fig. 1 The similarity coefficients between extracted emission and excitation spectra of six components and the corresponding standard spectra with component numbers from 6 to 9 for samples X1 and X2

表 1 样本集 X1 和 X2 的前 8 个奇异值累积贡献率

Table 1 The first 8 cumulative contribution rates of singular values for the first set samples X1 and X2

	1	2	3	4	5	6	7	8
X1	0.836 0	0.980 7	0.991 6	0.996 7	0.998 7	0.999 6	0.999 7	0.999 7
X2	0.932 1	0.986 6	0.996 1	0.998 7	0.999 2	0.999 6	0.999 6	0.999 7



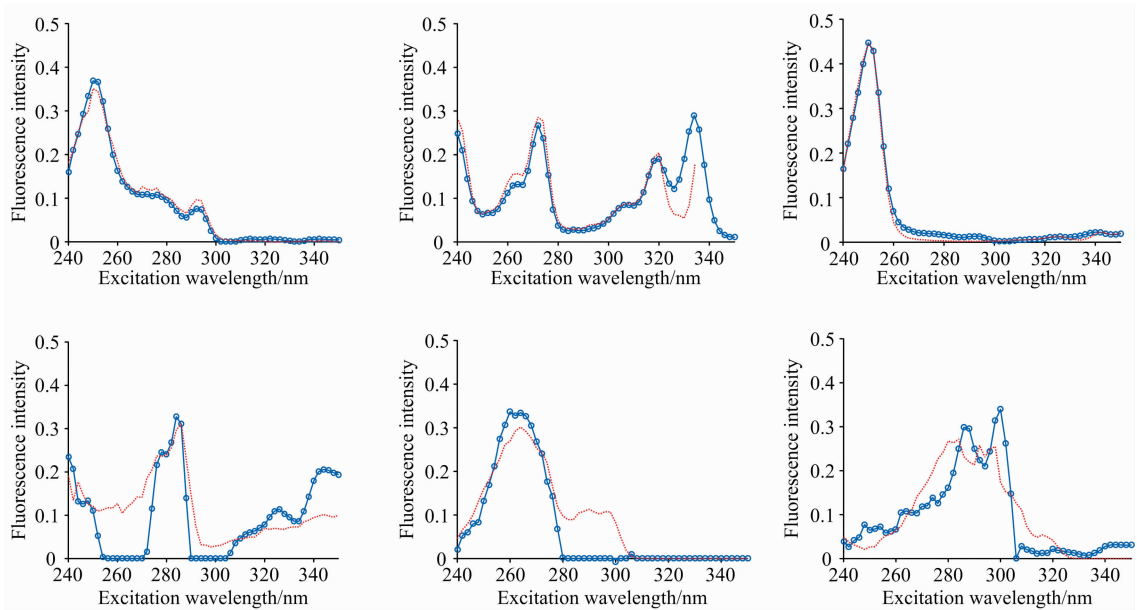


图 2 样本集 X1, X2 和两个伪样本提取出的六组分 (PHE/PY/AN/FLA/FLU/ACE) 发射光谱/激发光谱 (蓝线) 以及相应的组分数为 8 的标准光谱 (红线)

Fig. 2 The extracted emission (PHE/PY/AN/FLA/FLU/ACE) and excitation (PHE/PY/AN/FLA/FLU/ACE) spectra of six components (blue lines) and the corresponding standard spectra (red lines) with component number 8 for samples X1, X2 and two pseudo samples

第一个样本集的奇异值累积贡献率列于表 1。将样本 X1 的前 6 个奇异值设为零, 样本 X2 的前 5 个奇异值设为零,

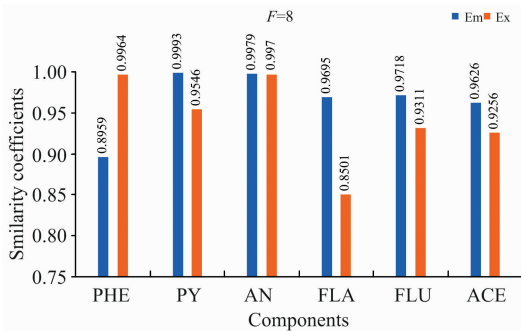


图 3 可行组分数设置从 6 到 9 下样本集 X1, X2 和两个伪样本中的六种不同组分激发/发射光谱和标准光谱对比的相似系数

Fig. 3 The similarity coefficients between extracted emission and excitation spectra of six components and the corresponding standard spectra with component numbers from 6 to 9 for samples X1, X2 and two pseudo samples

其余奇异值分别构成两个新的对角矩阵。根据式(3)重构两个伪样本。在每次测试中, 将两个观察样本和两个新的伪样本一起作为 PARAFAC 算法的输入。图 2 给出组分数为 8 的 PARAFAC 双向提取的发射和激发光谱。可以看出, 双向提取的发射光谱和激发光谱与相应的标准光谱吻合得很好, 特别是 PY, AN, FLA, FLU 和 ACE 的发射光谱, 以及 PHE, PY, AN, FLU 和 ACE 的激发光谱。计算提取的发射光谱和激发光谱与 6 个组分对应的标准光谱的相似系数, 均在 0.85 以上(图 3)。可见, 构建新的奇异值伪样本提高了弱荧光多环芳烃组分光谱的分辨能力和微弱信号的分解与放大特性。

然后根据 Eq. (1)对第二个样本集 X3 和 X4 的实验数据集进行三维 PARAFAC 分解拟合。图 4 给出了提取的 6 个组分的发射和激发光谱与对应的 6~9 个组分设置下的标准光谱的相似系数, 其相似系数均小于 0.80。根据 Eq. (3)构造两个新的伪样本。通过设置 X3 的前 6 个奇异值和样本 X4 的前 5 个奇异值为 0; 将四个样品(X3, X4 和两个新的伪样本)输入到 PARAFAC 算法和设置组分数为 9。提取的发射和激发光谱与对应的标准光谱的相似系数均在 0.81 以上(图 5)。表 2 为样本集 X3 和 X4 的前 8 个奇异值累积贡献率。

表 2 样本集 X3 和 X4 的前 8 个奇异值累积贡献率

Table 2 The first 8 cumulative contribution rates of singular values for the first set samples X3 and X4

	1	2	3	4	5	6	7	8
X3	0.8318	0.9602	0.9848	0.9928	0.9979	0.9991	0.9994	0.9995
X4	0.8448	0.9810	0.9955	0.9979	0.9990	0.9993	0.9996	0.9997

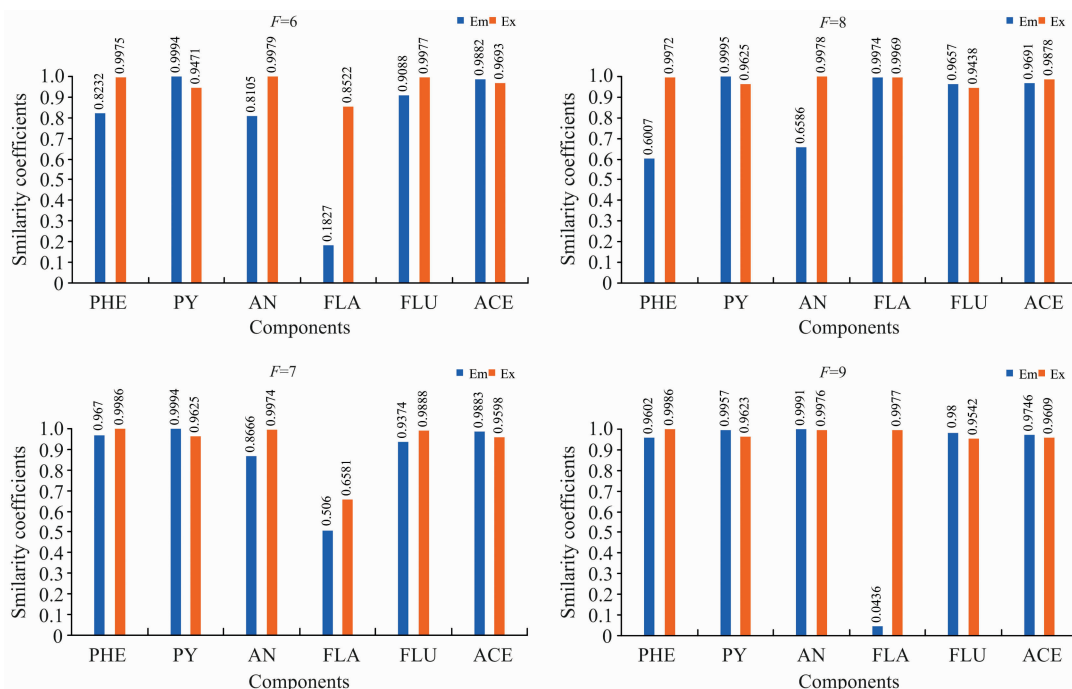


图 4 可行组分数设置从 6 到 9 下样本集 X3, X4 中的六种不同组分激发/发射光谱和标准光谱比对的相似系数

Fig. 4 The similarity coefficients between extracted emission and excitation spectra of six components and the corresponding standard spectra with component numbers from 6 to 9 for samples X3 and X4

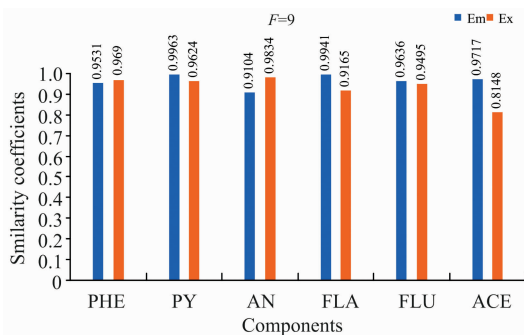


图 5 可行组分数设置从 6 到 9 下样本集 X3, X4 和两个伪样本中的六种不同组分激发/发射光谱和标准光谱比对的相似系数

Fig. 5 The similarity coefficients between extracted emission and excitation spectra of six components and the corresponding standard spectra with component numbers from 6 to 9 for samples X3, X4 and two pseudo samples

图 6 为最后一组样品 X5 和 X6, 组分数为 6~9, 采用三维 PARAFAC 分解提取的六组分 PAHs 的结果。PHE, FLA

和 ACE 的相似系数均小于 0.80。以同样的方式构造两个新样品通过设置样本 X5 的前 6 个和 X6 的前 5 个奇异值为 0 (表 3), PARAFAC 的分析结果由图 7 中给出。提取的发射光谱和激发光谱及其对应的标准光谱的相似性系数均在 0.81 以上, 表明 PARAFAC 耦合奇异值分解可以提高两个观测样品对荧光强度较弱的多环芳烃的检测能力。表 3 为样本集 X5 和 X6 的前 8 个奇异值累积贡献率。

3 结论

为了从两个样品中识别出更多的组分, 提出了一种基于 SVD 和三维 PARAFAC 的方法。通过奇异值分解对混合样品的每一个三维荧光光谱进行分析, 并根据奇异值的累积贡献率构建新的伪本来突出微弱的荧光信号。结合伪样品, 采用三维 PARAFAC 实现了两个样品水中多环芳烃重叠荧光光谱的欠定盲分离。从两个样品中成功提取了 6 种 (PHE/PY/AN/FLA/FLU/ACE) 多环芳烃的源光谱, 通过比较提取的发射/激发光谱与对应的标准光谱的相似性系数实现成分识别。解析光谱与标准发射/激发光谱的相似性均在 0.80 以上。本研究为促进激发发射荧光光谱技术在水体多环芳烃监测中的应用提供了潜在的科学价值。

表 3 样本集 X5 和 X6 的前 8 个奇异值累积贡献率

Table 3 The first 8 cumulative contribution rates of singular values for the first set samples X5 and X6

	1	2	3	4	5	6	7	8
X5	0.842 2	0.962 5	0.988 4	0.997 9	0.999 2	0.999 5	0.999 6	0.999 7
X6	0.870 5	0.969 0	0.996 5	0.998 8	0.999 4	0.999 7	0.999 8	0.999 8

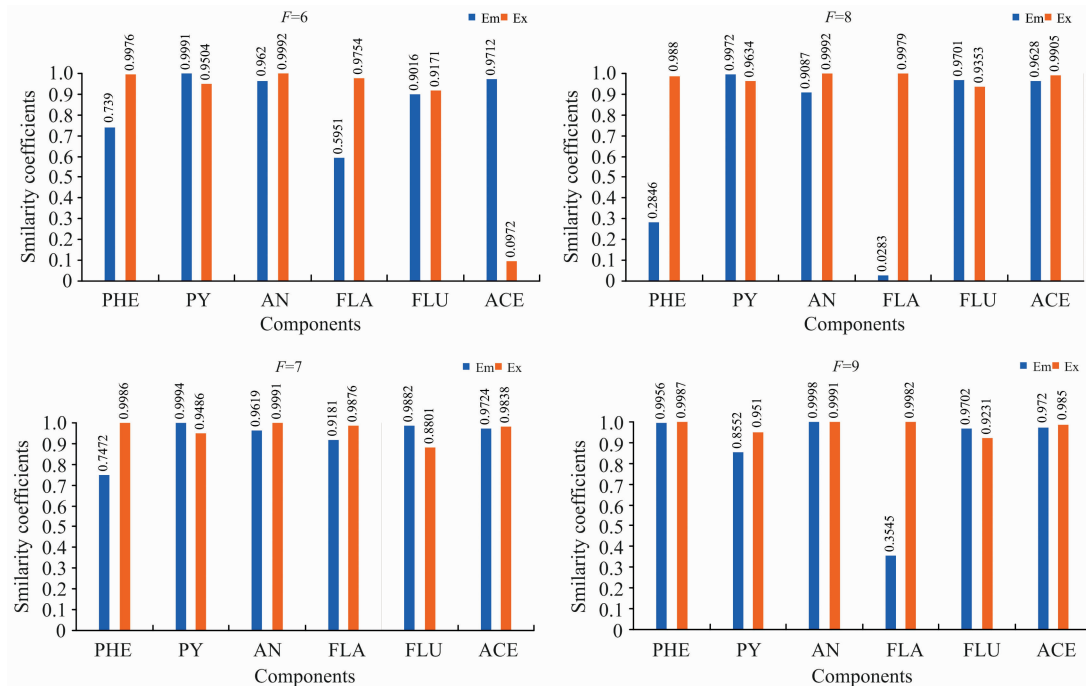


图 6 可行组分数设置从 6 到 9 下样本集 X5, X6 中的六种不同组分激发/发射光谱和标准光谱对比的相似系数

Fig. 6 The similarity coefficients between extracted emission and excitation spectra of six components and the corresponding standard spectra with component numbers from 6 to 9 for samples X5 and X6

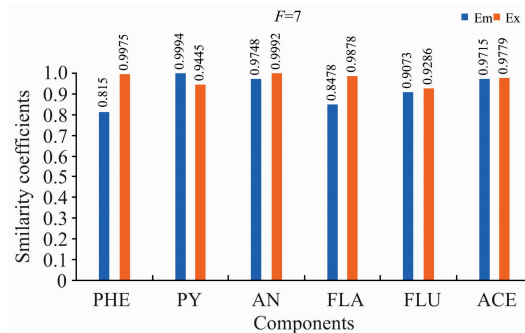


图 7 可行组分数设置从 6 到 9 下样本集 X5, X6 和两个伪样本中的六种不同组分激发/发射光谱和标准光谱对比的相似系数

Fig. 7 The similarity coefficients between extracted emission and excitation spectra of six components and the corresponding standard spectra with component numbers from 6 to 9 for samples X5, X6 and two pseudo samples

References

- [1] Brandt H C A, Watson W P. *Annals of Occupational Hygiene*, 2003, 47(5): 349.
- [2] Kaushik C P, Haritash A K. *Our Earth*, 2006, 3(3): 1.
- [3] Marchand C, St-Arnaud M, Hogland W, et al. *International Biodeterioration and Biodegradation*, 2017, 116: 48.
- [4] Williams E S, Mahler B J, Van Metre P C. *Environmental Science and Technology*, 2013, 47: 1101.
- [5] Wise S A, Sander L C, May W E. *Journal of Chromatography A*, 1993, 642(1-2): 329.
- [6] Walter B W, Udienza H, Mattheu M, et al. *J. Chromatogr. A*, 2014, 1345: 1.
- [7] Taylor T A, Jarvis G B, Xu H, et al. *Instrumentation Science & Technology*, 1993, 21(3): 141.
- [8] Lakowicz J R, Masters B R. *Principles of Fluorescence Spectroscopy*. Third Edition. Plenum Press, 1983.
- [9] Eleoroaristizabal S, Juan A D, Garcia J A, et al. *Chemometrics and Intelligent Laboratory Systems*, 2014, 132: 63.
- [10] Booksh K S, Muroski A R, Myrick M L. *Analytical Chemistry*, 1996, 68(20): 3539.
- [11] Nahorniak M L, Booksh K S. *Analyst*, 2006, 131: 1308.
- [12] Wu H L, Nie J F, Yu Y J, et al. *Analytica Chimica Acta*, 2009, 650(1): 131.

- [13] Escandar G M, Goicoechea, Héctor C, et al. *Analytica Chimica Acta*, 2014, 806; 8.
- [14] María del Carmen Hurtado-Sánchez, Lozano V A, Rodríguez-Cáceres M I, et al. *Talanta*, 2015, 134(1): 215.
- [15] Bro R. *Chemometrics and Intelligent Laboratory Systems*, 1997, 38; 149.
- [16] Ferretto N, Tedetti M, Guigue C, et al. *Chemosphere*, 2014, 107; 344.
- [17] Wang H B, Zhang Y J, Xiao X, et al. *Analytical Methods*, 2011, 3(3): 688.
- [18] Yu X Y, Zhang Y J, Yin G F, et al. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2014, 124; 52.
- [19] Kopriva I, Jeric I. *J. Mass Spectrom.*, 2009, 44(9): 1378.
- [20] Kopriva I, Jeric I, Smrecki V. *Anal. Chim. Acta*, 2009, 653(2): 143.
- [21] Schmidt M, Rajagopal S, Ren Z. *Biophysical Journal*, 2002, 84(3): 2112.
- [22] Zhao M, Jia X. *Mechanical Systems and Signal Processing*, 2017, 94; 129.
- [23] Nikolajsem R P H, Booksh K S, Hansen A M, et al. *Anal. Chim. Acta*, 2003, 475; 137.

Study on Small Sample Analysis Method for Identification of Polycyclic Aromatic Hydrocarbons in Water

ZHU Wei^{1,2}, YANG Rui-fang^{1*}, ZHAO Nan-jing^{1*}, YIN Gao-fang¹, XIAO Xue¹, LIU Jian-guo¹, LIU Wen-qing¹

1. School of Environmental Science and Optoelectronic Technology, University of Science and Technology of China, Hefei 230026, China
2. Key Laboratory of Environmental Optics and Technology, Anhui Institute of Optics and Fine Mechanics, Hefei Institute of Physical Sciences, Chinese Academy of Sciences, Hefei 230031, China

Abstract Polycyclic aromatic hydrocarbons (PAHs) are a toxic and harmful organic compound that widely exists in the natural environment. Its main sources are various microorganisms in nature and plant biosynthesis, natural fires in vegetation-rich areas, volcanic eruptions, fossil fuels, incomplete combustion of artificial industrial hydrocarbons and oil leakage during transportation. Polycyclic aromatic hydrocarbons are toxic, with biological carcinogenicity, heritable toxicity and mutagenicity. It harms human respiratory, circulatory and nervous systems in many aspects. Therefore, it is necessary to study the on-site monitoring and analysis methods of polycyclic aromatic hydrocarbons. Chemical analysis methods include chemical titration with pretreatment, liquid chromatography (LC), high-performance liquid chromatography (HPLC), and gas chromatography-mass spectrometry (GC-MS); Spectroscopic analysis studies UV absorption spectra, fluorescence spectra and three-dimensional fluorescence spectra. The three-dimensional fluorescence spectrum obtains the information of excitation and emission wavelength simultaneously, so it contains more optical information, high sensitivity and remarkable spectral characteristics. Therefore, it gains obvious advantages in the field detection of actual water bodies and the rapid study of mixed components of water samples. Common three-dimensional fluorescence spectrum analysis methods include parallel factor analysis (PARAFAC), multidimensional partial least squares (N-PLS), etc. Parallel factor analysis is an effective method to analyze polycyclic aromatic hydrocarbons' overlapping three-dimensional fluorescence spectra. However, sometimes, due to the weak fluorescence signal, the underdetermined three-dimensional fluorescence spectrum analysis cannot get satisfactory results. In order to extract more components from two sample sets, a method based on singular value decomposition (SVD) and PARAFAC is proposed. First, singular value decomposition is used on each observed sample, the appropriate singular value is selected according to the cumulative contribution rate, and a new pseudo sample is constructed to highlight the weak fluorescence signal. They were then inputting two observed samples and their corresponding pseudo samples into PARAFAC to recover the component spectrum. Three groups' three-dimensional fluorescence spectra of polycyclic aromatic hydrocarbons with different fluorescence intensities are analyzed. The pure component spectra of six polycyclic aromatic hydrocarbons were extracted and identified from two mixed samples and the result shows that the similarity between the emission/excitation spectra, and the standard spectra is more than 0.80.

Keywords Underdetermined analysis; Weak fluorescence signal; Parallel factor analysis; Singular value decomposition; Polycyclic aromatic hydrocarbons

* Corresponding authors

(Received Apr. 15, 2022; accepted Sep. 19, 2022)