

平行因子法结合自组织映射神经网络的 三维荧光特征及其与水质的关系

王 娟^{1,2}, 张 飞^{1,2,3}, 王小平^{1,2}, 杨胜天^{1,4}, 陈 芸⁵

¹新疆大学资源与环境科学学院, 新疆 乌鲁木齐 830046;

²新疆大学绿洲生态教育部重点实验室, 新疆 乌鲁木齐 830046;

³新疆智慧城市与环境建模普通高校重点实验室, 新疆 乌鲁木齐 830046;

⁴北京师范大学地理学与遥感科学学院, 北京 100875;

⁵澳大利亚联邦科工组织水土研究所, 澳大利亚 堪培拉 2601

摘要 利用三维荧光技术进行水质监测对于干旱区绿洲河流水质的有效管理具有重要的意义。以三维荧光技术为手段,以艾比湖流域地表水为研究对象,结合平行因子(PARAFAC)法和自组织特征映射神经网络(SOM)方法,探讨了艾比湖流域地表水溶解性有机质的三维荧光特征及其与地表水水质指标之间的关系。通过 PARAFAC 法,有效提取了艾比湖流域地表水样中的 4 种荧光组分,C1 荧光峰对应物质为紫外区类富里酸,C2 荧光峰对应物质为类富里酸,C3 包括 2 个峰 C3(T1)和 C3(T2),其中 C3(T1) 荧光峰对应物质为类蛋白,C3(T2)荧光峰对应物质为类腐殖酸,C4 荧光峰对应物质为类腐殖质。经 SOM 训练,在不同聚类层中探讨水质参数分布情况,水质状况由差到好的顺序依次为博河上游、精河绿洲、乌苏周边农田、艾比湖周边。在艾比湖流域丰水期,酸碱度(pH)、电导率(EC)、溶解氧(DO)、化学需氧量(COD)和五日生化需氧量(BOD₅)与水样的三维荧光峰具有较为显著的相关性,而总磷(TP)、总氮(TN)及氨氮(NH₃⁺-N)与各荧光峰相关性较弱。分别建立 pH、EC、DO、COD 及 BOD₅ 与各荧光组分间的多元线性回归方程,求得相关系数 *R* 分别为 0.579、0.632、0.502、0.762 和 0.785,可以在一定程度上利用各荧光组分模拟水质参数的变化情况。在利用 PARAFAC 探讨地表水荧光特征的基础上,SOM 网络作为一种有效的水体荧光光谱分析工具,可为干旱区水质监测和河流水质污染治理提供科学依据。

关键词 光谱学;水质监测;三维荧光;平行因子分析法;自组织映射神经网络;艾比湖流域

中图分类号 K90 **文献标识码** A

doi: 10.3788/AOS201737.0730003

Three-Dimensional Fluorescence Characteristics by Parallel Factor Method Coupled with Self-Organizing Map and Its Relationship with Water Quality

Wang Juan^{1,2}, Zhang Fei^{1,2,3}, Wang Xiaoping^{1,2}, Yang Shengtian^{1,4}, Chen Yun⁵

¹College of Resources and Environment Science, Xinjiang University, Urumqi, Xinjiang 830046, China;

²Key Laboratory of Oasis Ecology, Ministry of Education, Xinjiang University, Urumqi, Xinjiang 830046, China;

³Key Laboratory of Xinjiang Wisdom City and Environment Modeling, Urumqi, Xinjiang 830046, China;

⁴College of Geography and Remote Sensing Science, Beijing Normal University, Beijing 100875, China;

⁵Land and Water, Commonwealth Scientific and Industrial Research Organization, Canberra 2601, Australia

收稿日期: 2016-12-09; **收到修改稿日期:** 2017-03-02

基金项目: 国家自然科学基金(41361045)、国家自然科学基金(新疆联合基金本地优秀青年人才培养专项)(U1503302)、2016 年新疆大学大学生创新训练计划重点项目(XJGRI2016014)

作者简介: 王 娟(1992—),女,硕士研究生,主要从事干旱区资源与环境遥感方面的研究。

E-mail: wangjuan6231@163.com

导师简介: 张 飞(1980—),男,博士,副教授,主要从事干旱区资源与环境遥感应用方面的研究。

E-mail: zhangfei3s@163.com (通信联系人)

Abstract Using three-dimensional fluorescence technique for water quality monitoring has great significance to manage the river water quality in oases of arid areas. The major objectives of this study are to apply parallel factor (PARAFAC) method and self-organizing map (SOM) to assessing fluorescence properties, and to characterize and quantify the relationship between fluorescence and water quality indexes in the Ebinur Lake watershed. Fluorescence components are identified by the PARAFAC in the Ebinur Lake watershed. C1 corresponds to ultraviolet humic-like fulvic acid and C2 corresponds to humic-like fulvic acid. C3 includes two peaks, C3(T1) and C3(T2), C3(T1) corresponds to protein-like substance and C3(T2) corresponds to humic-like acid. C4 corresponds to humic-like substance. SOM technique is employed for the exploratory analysis of fluorescence components of water samples in the Ebinur Lake watershed. Discussing the distribution of water quality parameters in different clustering layers shows that the water quality tends to be better as a sequence of Bortala River upstream, JingHe Oasis, farmland surrounding Wusu, and Ebinur Lake nearby. There exist certain relationships between water quality indicators and fluorescence peaks. In wet season, the pH, electricity conductivity (EC), dissolved oxygen (DO), chemical oxygen demand (COD) and biological oxygen demand in five days (BOD_5) have significant correlations with fluorescence peaks, while the total phosphor (TP), total nitrogen (TN) and NH_3^+-N have weak correlations with fluorescence peaks. Multiple linear regression equations are established between the pH, EC, DO, COD, BOD_5 and fluorescent components, and correlation coefficients R are 0.579, 0.632, 0.502, 0.762 and 0.785, respectively. In general, based on the investigation of the fluorescence characteristics using PARAFAC, the SOM network can be used as an effective tool for analyzing water fluorescence spectra, and it can help to provide a scientific basis for water quality monitoring and river pollution controlling in arid regions.

Key words spectroscopy; water quality monitoring; three-dimensional fluorescence; parallel factor analysis; self-organizing map; Ebinur Lake watershed

OCIS codes 300.6280; 120.4630; 120.6650; 240.6490

1 引 言

近年来,水体中有机污染物浓度的在线检测,如生物需氧量(BOD)、化学需氧量(COD)及总有机碳(TOC)等,已经成为监测地表水资源污染源的主要依据^[1]。在干旱地区,绿洲的动植物、地表水资源及地下水通常会受到高盐分物质的影响,在此大环境下,传统的水质参数监测存在一定的局限性^[2]。因此,探讨在干旱区的水质监测对于干旱区绿洲生态环境的管理至关重要。

三维荧光光谱(EEM)技术已被应用于地表水资源检验及水系管理等方面^[3]。由于该技术具有高度的灵敏性及不破坏样品结构等特点,在水体监测和水处理领域引起了学者的广泛关注^[4-5]。在给水系统中,可以用EEM对地下水^[6]、地表水^[7]和循环水^[8]等系统的水质进行评价。在污水处理方面,利用EEM可以有效去除水处理过程中的有机物^[9]。其中,平行因子(PARAFAC)法可以识别不同类型荧光团的独立光谱^[10-11],将重叠的荧光峰分成几个独立的组分,进而分析EEM矩阵以及更微小的变化^[12],故PARAFAC的优势能满足干旱区盐分对水质影响的监测需要。

自组织特征映射(SOM)是人工神经网络算法的一个分支,是一种自组织、自学习网络可视化的方法。它可以将多维空间数据经过非线性映射在低维格点中表达出来^[13],可以很好地将样本和变量连接起来^[14-15]。近年来,因为SOM可以处理非线性关系在环境研究中越来越受欢迎。Kaltch等^[16]和Céréghino等^[17]探讨了SOM在环境科学中的应用,尤其是对水资源的分类。Chon^[18]评论了SOM技术在生态学中的应用。因水质监测数据具有不确定性、高维及非线性等特点,利用现代模式识别与数据挖掘方法进一步对水质监测数据进行分析,可以在一定程度上弥补传统方法的不足^[19]。新疆艾比湖流域是一个涵盖山地、绿洲、荒漠等生态环境的典型干旱区域,多年来在流域内气候和人为因素的双重扰动下,因流域内绿洲农业的发展以及人类生活所产生的污水直接排放或汇入河流所引发的区域性生态环境问题,已经成为关系到新疆社会经济可持续发展的主要问题。本文选取典型的干旱平原区艾比湖流域作为研究对象,利用三维荧光PARAFAC法结合SOM法,研究艾比湖流域水体中相关物质的三维荧光特征,并尝试性地探讨其与水质参数之间的关系,以期控制、治理艾比湖流域河流水质污染提供新思路,为调控绿洲生态环境的实际应用和发展提供依据。

2 实 验

2.1 研究区概况

艾比湖流域地处 $43^{\circ}38' \sim 45^{\circ}52'N$ 和 $79^{\circ}53' \sim 85^{\circ}02'E$ 之间,是国家级荒漠自然生态保护区,主要由湿地、荒漠植被及野生生物等混合组成。研究区面积为 50621 km^2 ,包括博乐市、精河县、乌苏市、温泉县、托里县南部、奎屯市及克拉玛依的独山子区等。区内入湖河流主要有精河、博尔塔拉河和奎屯河等。艾比湖流域年平均气温为 $7.36 \text{ }^{\circ}\text{C}$,多年平均降水量为 $100 \sim 200 \text{ mm}$,多年平均蒸发量为 $1500 \sim 2000 \text{ mm}$,属典型的北温带大陆性干旱气候^[20]。近年来,艾比湖流域水资源退化严重,突出表现为水质咸化、水域面积缩小、天然绿洲面积减少、耕地盐渍化、土地沙化、草场退化严重等^[21]。随着艾比湖湖区面积的减小,湖床裸露面积不断增大,在阿拉山口常年大风的影响下,艾比湖流域成为了主要的沙尘源地,影响着北疆地区的生态环境。研究范围如图 1 所示。

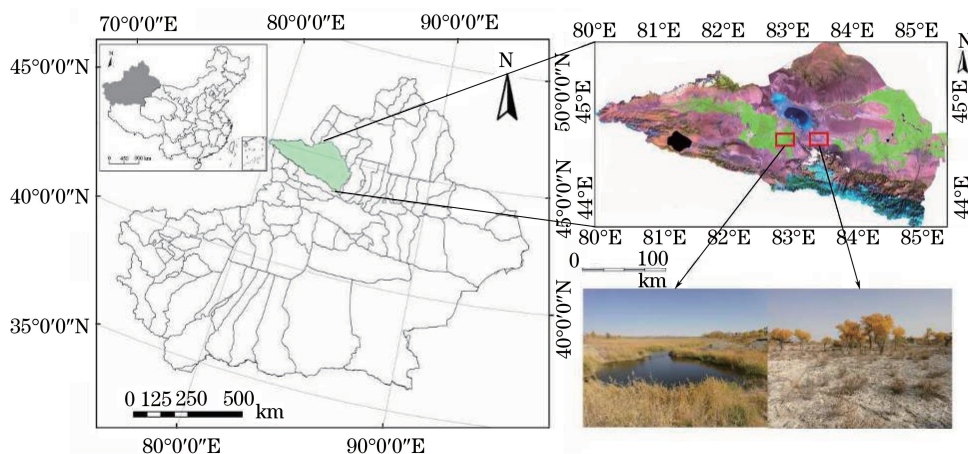


图 1 研究区位置图

Fig. 1 Location map of the studied area

2.2 数据采集与处理

由于艾比湖流域以盐业、卤虫产业为该区域的支柱产业,没有其他重工业的影响,因此在此研究区内不考虑工业废水污染,主要对位于温泉县的博尔塔拉河上游、精河、精河县绿洲、托托乡农用土地、艾比湖周边、艾比湖鸟岛国家生态保护区、甘家湖梭梭林自然保护区、奎屯河、阿克齐苏河以及乌苏地区等区域进行取样。2016年5月共采集56个表层(0.5 m)水样(图 2),同时记录每个采样点周围的景观分布状况。将所采集的水样放置于预先处理过的白色聚乙烯塑料瓶中并放入保温箱。将水样送回实验室过玻璃纤维微孔膜(Whatman GF/F, $450 \text{ }^{\circ}\text{C}$ 马氟弗炉中灼烧 5 h 后使用)后,扫描各水样样品中有色可溶性有机物(CDOM)的

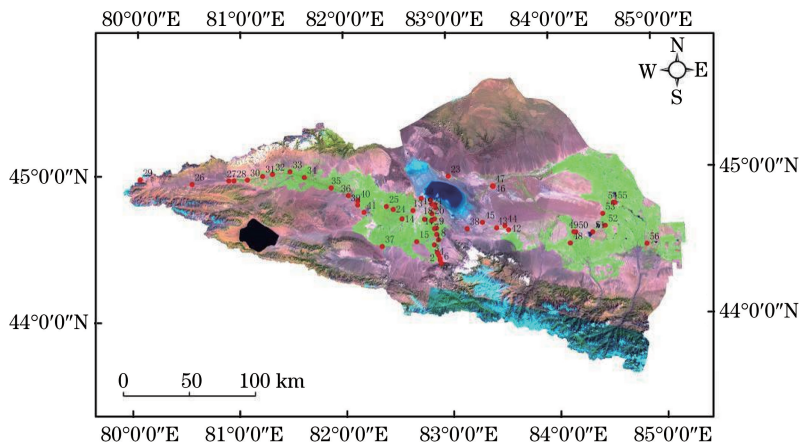


图 2 采样点分布图

Fig. 2 Distribution map of sampling points

三维荧光光谱,同时测定的水质指标有酸碱度(pH)、电导率(EC)、溶解氧(DO)、化学需氧量(COD)、五日生化需氧量(BOD₅)、总磷(TP)、总氮(TN)和氨氮(NH₃⁺-N)等,具体测定方法如表 1 所示。

表 1 水质指标及实验方法

Table 1 Water indexes and experimental methods

Parameter	Experimental method
pH	pH-40A portable pH acidity meter
DO	Iodine quantity method (GB/T 7489-87) using visible light spectrophotometer 722 N test instrument for determining the content of DO
COD	Dichromate method (GB 11914-89) using standard COD digestion apparatus (K-100) for determining the content of COD
BOD ₅	Dilution and inoculation method (HJ 505-2009) with constant temperature incubator (HWS-150 type) for determining the content of BOD ₅
TP	Ammonium molybdate spectrophotometric method (HJ 636-2012) using visible light spectrophotometer 722 N for determining the content of TP
TN	Ultraviolet spectrophotometry (HJ 535-2009) by ultraviolet visible light spectrophotometer, UV-6100 for determining the content of TN
NH ₃ ⁺ -N	Nessler's reagent spectrophotometry with the amount of visible light spectrophotometer 722 N determination the content of NH ₃ ⁺ -N
EC	Handheld conductivity meter

荧光光谱分析法:采用 Hitachi F-7000 型高灵敏度荧光光谱分析仪进行荧光测定。激发光源为 150 W 氙弧灯;光电倍增管(PMT)电压为 700 V;信噪比大于 110;带通:激发波长 $\lambda_{ex} = 10$ nm,发射波长 $\lambda_{em} = 10$ nm;响应时间设为自动。三维荧光光谱扫描速率为 2400 nm/min,激发扫描波长范围为 200~450 nm,发射扫描波长范围为 260~550 nm,同步扫描激发光谱扫描波长范围为 260~550 nm, $\Delta\lambda = \lambda_{em} - \lambda_{ex} = 30$ nm^[22];扫描光谱进行仪器自动校正后,将所得到的光谱扣除超纯水空白,可减少拉曼散射及仪器条件对荧光光谱的影响。

2.3 研究方法

2.3.1 PARAFAC 法

PARAFAC 法可以将一个由多个 EEM 数据构成的三维阵列 X 利用交替最小二乘原理分解为 3 个载荷矩阵,从而对三维荧光光谱进行解析。通过裂半分析和残差分析来检验 PARAFAC 模型的有效性,从而确定最优可溶性有机物(DOM)组分的数目^[23]。确立最优分组的数目后,扣除 $\lambda_{em} < \lambda_{ex} + 20$ nm 的区域消除一级瑞利散射对 EEM 的影响,扣除部分用零替换。计算平台为 Matlab 2013a。

2.3.2 荧光强度和荧光组分占总组分比例的计算

通过 PARAFAC 模型得出每种组分的相对荧光强度,每种组分的荧光强度 I_i 计算公式为^[24]

$$I_i = S_i \times \lambda_{ex_max}(i) \times \lambda_{em_max}(i), \quad (1)$$

式中 S_i 代表第 i 种组分的相对荧光强度; $\lambda_{ex_max}(i)$ 代表第 i 种组分激发负载的最大值; $\lambda_{em_max}(i)$ 代表第 i 种组分发射负载的最大值。总荧光强度 I_T 和荧光组分占总组分比例 P_i 的计算公式为

$$I_T = \sum_1^n I_n, \quad (2)$$

$$P_i = I_i / I_T, \quad (3)$$

式中 n 为组分总数。

2.3.3 基于 SOM 的荧光组分特征解析

SOM 神经网络不需要数据符合一定的分布,结构简单且不受局部过小的影响,对细节信息处理能力较好,其结果保持输入模式的分布特性和拓扑结构不变^[25]。K-means 算法基于迭代原理将数据对象划分到不同的簇中,使得相同簇内部各对象间具有较高的相似程度,而不同簇之间的对象相似程度则较低。

选用基于无分层 K-means 分类的 SOM 法进行研究区三维荧光组分的研究,分为以下步骤:

1) 将待聚类 PARAFAC 组分输入 SOM 网络,采用计算网络大小的拓扑值来选择神经元数量,输出结

果用量化误差(QE)和拓扑图形误差(TE)的最小值确定。其中,QE 具有可以确定构建的神经网络分辨原始输入数据的能力^[13];TE 可以测量神经网络质量,即评估网络是否已适合训练。确定网络大小后,进行网络训练,进而得到一组权值。

2) 以 SOM 的聚类结果得到的权值作为初始聚类中心,对 K-means 算法进行初始化,再执行 K-means 算法对其进行聚类。无分层 K-means 分类的 SOM 法弥补了 SOM 网络本身收敛时间较长和 K-means 算法对初始聚类中心聚类效果较差的不足,可以既高效又较准确地完成聚类。计算平台为 Matlab 2013a。

3 结果与讨论

3.1 艾比湖流域水体的荧光组分特征研究

利用三维荧光结合 PARAFAC 模型得到艾比湖流域水样共有 4 个荧光组分,如图 3 所示。在此基础上,根据图 3 对其荧光峰位置进行了描述,结果见表 2。分析可得,5 类荧光峰的中心位置分别为:组分 C1 荧光峰对应物质为紫外区类富里酸($\lambda_{ex}/\lambda_{em} = 260 \text{ nm}/420 \text{ nm}$);组分 C2 荧光峰对应物质表现为类富里酸

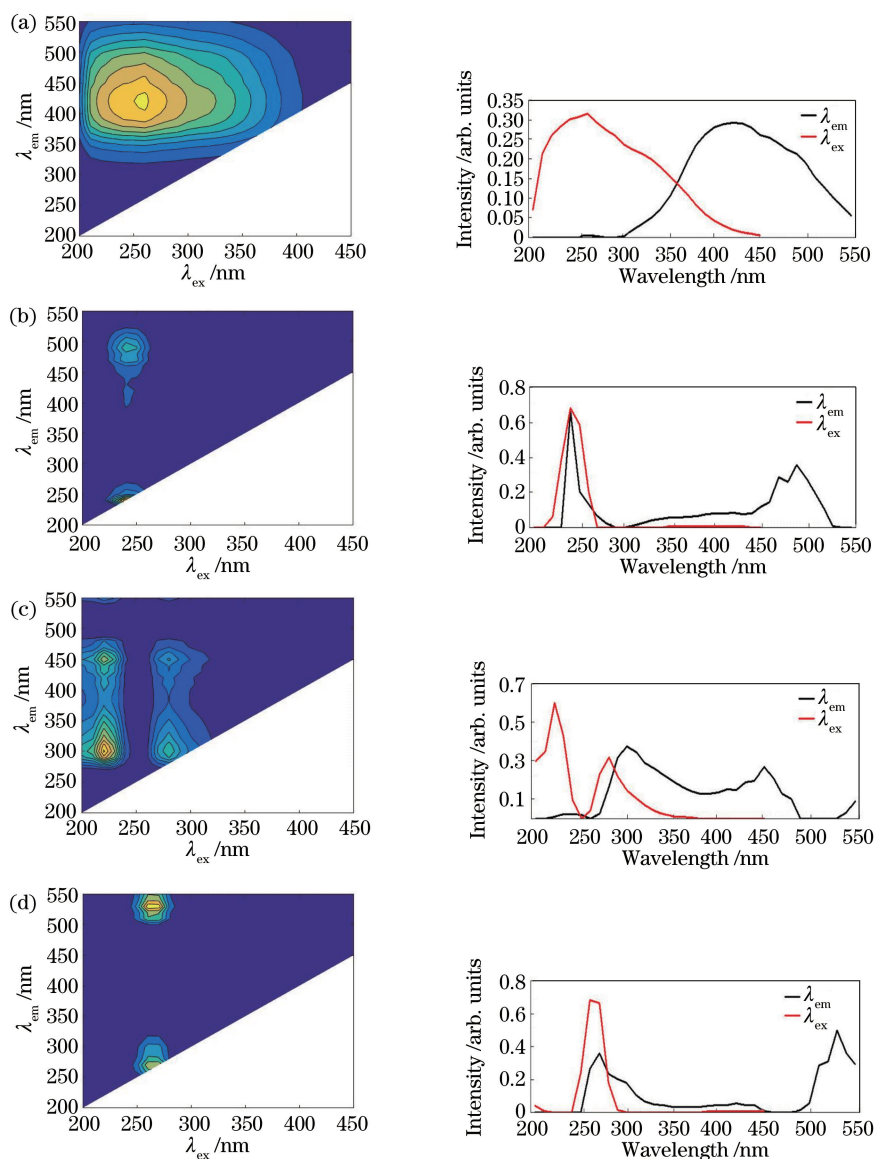


图 3 PARAFAC 法得到的艾比湖流域地表水的 4 个荧光组分。(a)组分 C1;(b)组分 C2;(c)组分 C3;(d)组分 C4

Fig. 3 Spectral characteristics of the four fluorescence components identified by the PARAFAC method.

(a) Component C1; (b) component C2; (c) component C3; (d) component C4

($\lambda_{ex}/\lambda_{em} = 240 \text{ nm}/490 \text{ nm}$);组分 C3 包含两个荧光峰,分别用 C3(T1)和 C3(T2)表示,其中 C3(T1)为类蛋白 B 峰($\lambda_{ex}/\lambda_{em} = 280 \text{ nm}/300 \text{ nm}$),表现为芳香性蛋白类结构,主要由微生物和浮游植物的降解产生,该峰易受农业生产用水和生活污水等外源的影响等;C3(T2)荧光峰对应物质表现为类腐殖酸($\lambda_{ex}/\lambda_{em} = 280 \text{ nm}/450 \text{ nm}$);组分 C4 荧光峰对应物质表现为类腐殖质($\lambda_{ex}/\lambda_{em} = 260 \text{ nm}/530 \text{ nm}$),主要由陆源土壤通过径流、渗流和淹水释放等水文过程输入^[26]。

表 2 PARAFAC 法得到的艾比湖流域地表水 4 个组分的荧光峰位置

Table 2 Fluorescence peak of four components in the surface water of the Ebinur Lake watershed by PARAFAC method

Component	Peak	Excitation wavelength /emission wavelength /nm	Description and probable source	Property	Reference
C1		260/420	Ultraviolet humic-like fulvic acid	Fulvic acid and humic acid fluorescence	[27]
C2		240/490	Humic-like fulvic acid	Contain lots of phenolic hydroxyl and carbonyl groups	[28]
C3	T1	280/300	Protein-like	Microbial degradation, structure of aromaticity protein	[29]
	T2	280/450	Humic-like acid	Easier than protein-like degradation	[30]
C4		260/530	Humic-like	Degradation of humic substances	[30]

3.2 基于 SOM 荧光组分特征的解析

用 SOM 方法结合 PARAFAC 替代传统的寻峰法分析荧光组分信息,在一定程度上优化了分析荧光组分特征的方法。利用 Matlab 2013a 以及 SOM Toolbox2.0,结合 K-means 聚类法,对 PARAFAC 模型鉴别结果进行进一步解析,实现了各荧光组分的初始化、训练和可视化输出。在 SOM 网络的训练过程中,根据输入样本的数量($5\sqrt{n}$)决定网格的大小。不断地修正和调整输入层、输出层神经元的连接权值,可以得到的训练效果为 35(7×5)个神经元,得到的 QE 及 TE 值分别为 0.497 和 0.018。

在 SOM 训练荧光组分的基础上,结合 K-means 算法及 Davies-Bouldin 指数(DBI)确定最佳聚类数^[15]。结果显示,当平均方差在不同聚类数间的值低于 5%时对应的 DBI 最低,此时对应的聚类数为最佳聚类结果。因此,输入训练好的神经元节点的权值,通过 K-means 聚类分析结合 DBI 值为指标选取聚类数,最终进行分类,结果如图 4 所示。从图 4 可以看出,当聚类数为 4 时,DBI 最低,因此荧光光谱可以分成 4 类区域,分别为聚类 I、聚类 II、聚类 III 和聚类 IV,最终结果如图 5 所示。

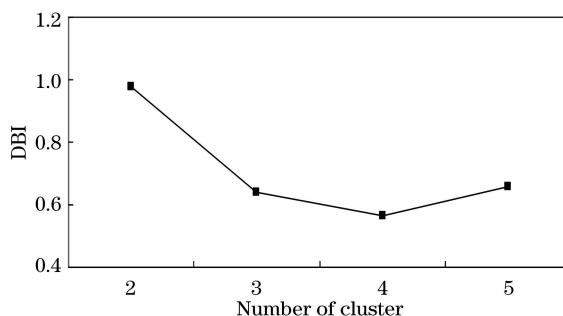


图 4 基于 K-means 不同聚类层的 DBI

Fig. 4 DBI for different clusters based on K-means

图 5 中包含 56 个采样点,结合实际采样点的分布状况,可以得到聚类 I 主要包括博尔塔拉河上游至精河绿洲农田附近的采样点,聚类 II 主要包括奎屯河、阿克齐苏河乃至乌苏地区的采样点,聚类 III 主要包括精河及精河绿洲周围的采样点,聚类 IV 则主要为艾比湖周边水产养殖基地附近的采样点。进一步探讨艾比湖流域荧光组分特征,各组分变量可视化如图 6 所示,图中 d 表示荧光强度的大小。

图 6 为不同水样荧光光谱的 4 个组分在 SOM 神经网络下可视化的输出结果。从总体上看,由于博尔

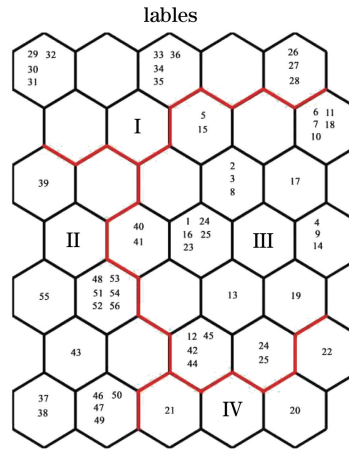


图 5 基于 SOM 的聚类样点分布图

Fig. 5 Sample distribution map obtained via SOM

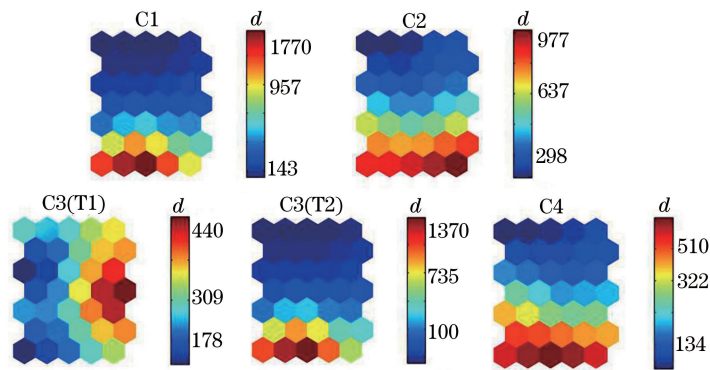


图 6 基于 SOM 的 4 个荧光峰组分的可视化图

Fig. 6 Visualization of four fluorescence peak components based on SOM

塔拉河上游位于温泉县,在艾比湖流域内相对海拔较高,该区域人口稀少,水体中有机物主要来源于藻类、动植物死亡后腐败溶解及微生物的胞外分泌物等,故聚类 I 内包含的采样点在各组分内荧光强度均较弱。从图 6 可以看出,除 C3(T1)外,其他组分在相应的神经元之间颜色分布较相似,表明各组分之间具有一定的关联性。结合表 2 可知,组分 C1 荧光峰对应的主要为紫外区类富里酸物质,在乌苏地区及艾比湖周边水产养殖基地采样点的荧光强度较强;组分 C2 荧光峰对应的主要为类富里酸物质,呈现为精河绿洲-奎屯河、阿克齐苏河-艾比湖周边荧光强度逐渐增强的趋势;组分 C3(T1)荧光峰对应的主要为类蛋白质物质,主要表现为集中在精河绿洲内部采样点的荧光强度较强;组分 C3(T2) 荧光峰对应的主要为类腐殖酸物质,与组分 C1 的荧光强度趋势大致一致;组分 C4 荧光峰对应的则主要为类腐殖质物质,且与组分 C2 的荧光强度表现出相似分布趋势。水体中的有机物质除了来源于浮游植物外^[31],最主要的来源就是陆源输入^[32]。总体而言,研究区分布在耕地周边及艾比湖周边的采样点较易受到影响,是因为研究区内主要以农业为主,畜牧业为辅,且艾比湖区位于精河绿洲附近,受人为干扰较大,周围分布着大量入湖河流及人工养殖基地(如鱼塘、蟹池等),故类腐殖质等有机物主要来源于浮游生物、绿洲内部的土壤、腐烂植物和农田堆肥等,同时有机物也会随灌溉水和雨水等进入艾比湖湖区附近。为了更直观地观测各聚类层水质参数的分布情况,统计出不同层中各水质参数值,结果如图 7 所示。

从图 7 可以看出,在不同聚类层,水质参数分布情况各不相同。从总体上看,pH 在各层中差异较小,范围为 8.5~9.0;而 EC 变化范围则较大,尤其在艾比湖周边受人工养殖水产等因素的干扰,采样点 EC 含量较高;DO 浓度在 4 个聚类层中差异不大,对评判各聚类层水质好坏不具有代表性。在此基础上,BOD₅ 和 COD 的质量浓度从大到小的顺序依次为博河上游、精河绿洲、乌苏周边农田、艾比湖周边。TN 和 TP 的质量浓度在精河绿洲附近较高,同时,由于 5 月艾比湖流域地表径流因冰雪融水等流量加大,在注入艾比湖过

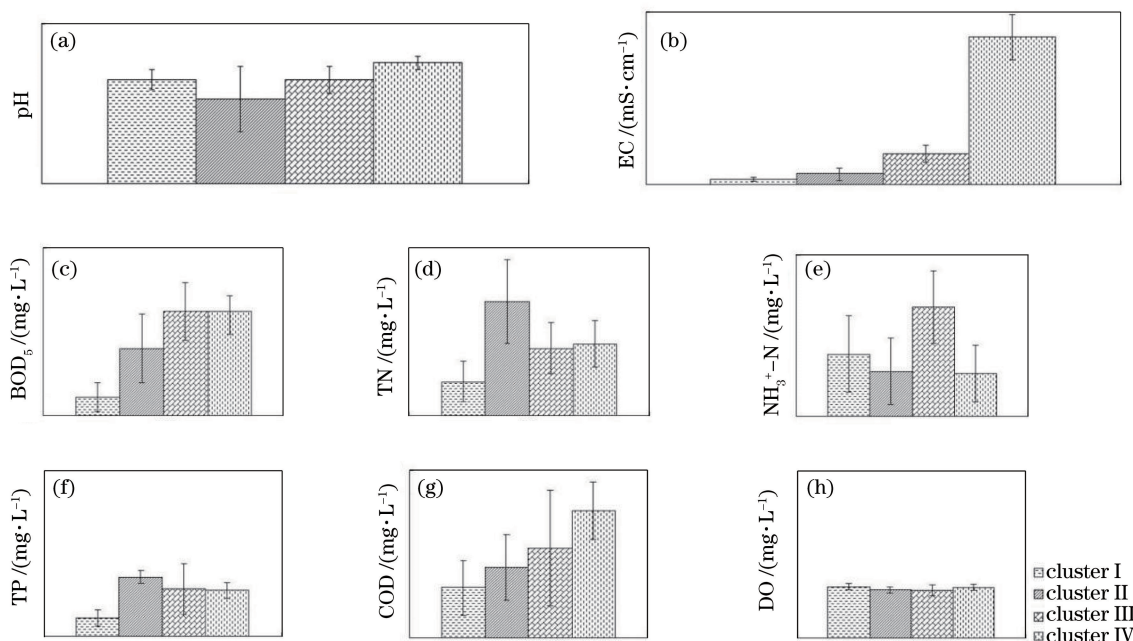


图 7 4 个聚类层中各水质参数的平均值。(a) pH; (b) EC; (c) BOD₅; (d) TN; (e) NH₃⁺-N; (f) TP; (g) COD; (h) DO

Fig. 7 Average values of the water quality parameters in the four clusters.

(a) pH; (b) EC; (c) BOD₅; (d) TN; (e) NH₃⁺-N; (f) TP; (g) COD; (h) DO

程中携带 N 及 P 元素,导致在聚类 IV 中 TN 和 TP 的质量浓度相对较高。乌苏周边的农田附近 NH₃⁺-N 的质量浓度较高。为进一步探讨各水质参数与 PARAFAC 荧光组分之间的关系,对其进行相关分析,结果见表 3。

表 3 荧光光谱与水质因子的相关系数

Table 3 Correlation coefficients between fluorescence spectra and water quality factors

Parameter	C1	C2	C3(T1)	C3(T2)	C4
pH	0.517**	0.325*	-0.340*	0.516**	0.426**
EC	0.545**	0.243	-0.154	0.562**	0.337*
DO	-0.363**	-0.276*	0.097	-0.384**	-0.276*
COD	0.679**	0.349**	-0.003	0.693**	0.383**
BOD ₅	0.711**	0.337*	-0.099	0.726**	0.419**
TP	-0.042	-0.047	-0.013	-0.033	-0.013
TN	-0.177	-0.095	0.319*	-0.179	-0.185
NH ₃ ⁺ -N	-0.102	-0.060	-0.301*	-0.091	-0.060

* $p < 0.05$ (1-tailed); ** $p < 0.01$ (2-tailed)

从表 3 可以看出,在 $p=0.01$ 置信水平下,pH 与 C1、C3(T2)和 C4 呈正相关,相关系数分别为 0.517、0.516和 0.426;在 $p=0.05$ 置信水平下,pH 与 C2 呈正相关,而与 C3(T1)呈负相关,相关系数分别为 0.325 和 -0.340。在 $p=0.01$ 置信水平下,EC 与 C1 及 C3(T2)呈正相关,相关系数分别为 0.545 和 0.562;EC 与 C4 在 $p=0.05$ 置信水平下呈正相关,相关系数为 0.337。DO 与 C1 和 C3(T2)在 $p=0.05$ 置信水平下呈负相关,相关系数分别为 -0.363 和 -0.384;与 C2 和 C4 在 $p=0.01$ 置信水平下呈负相关,相关系数均为 -0.276。在 $p=0.01$ 置信水平下,COD 与 C1、C2、C3(T2)及 C4 均呈正相关,相关系数分别为 0.679、0.349、0.693 及 0.383。在 $p=0.01$ 置信水平下,BOD₅ 与 C1、C3(T2)及 C4 呈较为显著的正相关,相关系数分别为 0.711、0.726 及 0.419;而在 $p=0.05$ 置信水平下 BOD₅ 与 C2 呈正相关,相关系数为 0.337。在 $p=0.05$ 置信水平下,TN 与 C3(T1)呈正相关,而 NH₃⁺-N 与 C3(T1)呈负相关,相关系数分别为 0.319和 -0.301。除此之外,TP、TN 及 NH₃⁺-N 与其他成分间并无显著相关性。艾比湖流域内降水量由于季节性差异分配不均匀,

从而导致流域内主要河流流量也呈现明显的季节性差异,显著影响艾比湖流域的水质状况。在丰水期,流域内冰川融水较多,地表径流较为丰富,导致精河、博尔塔拉河及奎屯河等河流流量大且流速快。艾比湖流域虽以农业为主,但在丰水期内,流域农作物还未开始大面积耕种,农药、化肥等使用较少,因此 TP、TN 及 NH_3^+-N 仅在绿洲部分区域含量较高,对艾比湖流域整体水质的影响并不突出。整体上看,丰水期内 TP、TN 及 NH_3^+-N 不是影响荧光光谱的主要因素,与各荧光组分间相关性较小。因此,为进一步研究水质参数与各荧光组分间的综合相关关系,以 C_1 、 C_2 、 $C_3(T_1)$ 、 $C_3(T_2)$ 、 C_4 为自变量,以与荧光组分较为相关的水质参数 pH、EC、DO、COD 及 BOD_5 为因变量,分别建立水质参数与各荧光组分间的多元线性回归方程,如表 4 所示。

表 4 各水质参数的多元线性回归模型

Table 4 Multiple linear regression models of water quality parameters

Parameter	Multiple linear regression model	R^2	Significance
pH	$4.442 \times 10^{-15} + 1.016 f_{C_1} - 0.387 f_{C_2} - 0.017 f_{C_3(T_1)} - 0.626 f_{C_3(T_2)} + 0.468 f_{C_4}$	0.335	<0.001
EC	$9.143 \times 10^{-16} - 2.694 f_{C_1} - 0.505 f_{C_2} + 0.078 f_{C_3(T_1)} + 3.3470.626 f_{C_3(T_2)} + 0.362 f_{C_4}$	0.399	<0.001
DO	$8.978 + 0.005 f_{C_1} + 0.001 f_{C_2} - 0.0070.626 f_{C_3(T_2)} - 0.002 f_{C_4}$	0.252	<0.001
COD	$6.233 - 0.042 f_{C_1} - 0.006 f_{C_2} + 0.018 f_{C_3(T_1)} + 0.0750.626 f_{C_3(T_2)} - 0.006 f_{C_4}$	0.581	<0.001
BOD_5	$1.058 - 0.010 f_{C_1} - 0.005 f_{C_2} + 0.004 f_{C_3(T_1)} + 0.0180.626 f_{C_3(T_2)} + 0.004 f_{C_4}$	0.617	<0.001

从表 4 可以看出, pH、EC、DO、COD 及 BOD_5 和各荧光组分建立模型所得到的相关系数 R 分别为 0.579、0.632、0.502、0.762 和 0.785。通常当相关系数 $0.5 \leq |R| < 0.8$ 时,被视为中度相关,且结果通过了 $\alpha = 0.05$ 的显著性水平检验,表明模型与数据的拟合程度较好,建立的多元线性回归方程可用,能在一定程度上利用各荧光组分模拟水质参数的变化情况,为干旱区艾比湖流域的水资源保护提供一定的依据。

综上所述,干旱区水质荧光特征与水质指标间的关系呈一定的响应规律,且与 pH、EC 及传统有机质检测指标 DO、COD 和 BOD_5 的关系较为显著。总体而言,艾比湖流域溶解性有机质的组成、特征及来源可以在一定程度上反映水质状况,可为干旱区水质监测和河流水质污染治理提供参考。

4 结 论

利用平行因子法结合 SOM 网络对艾比湖流域水样的三维荧光光谱进行解析,并探讨了水质指标与荧光峰之间的关系,得到以下结论。

1) 通过 PARAFAC 法有效提取艾比湖流域地表水样的 4 种荧光组分, C_1 荧光峰对应物质为紫外区类富里酸; C_2 荧光峰对应物质为类富里酸; $C_3(T_1)$ 荧光峰对应物质为类蛋白,主要具有芳香性蛋白类结构; $C_3(T_2)$ 荧光峰对应物质为类腐殖酸; C_4 荧光峰对应物质为类腐殖质。

2) SOM 训练结果表明,根据各荧光组分对应荧光强度的大小可将样点分为 4 层,不同聚类层中水质参数分布情况各不相同。从总体上看,水质状况由好到差的顺序依次为博河上游、精河绿洲、乌苏周边农田、艾比湖周边。

3) 艾比湖流域水质指标与三维荧光峰存在一定的关系,在艾比湖流域的丰水期,pH、EC 及传统有机质检测指标 DO、COD 和 BOD_5 与水样的三维荧光峰存在较为显著的相关性,TP、TN 及 NH_3^+-N 与各荧光峰的相关性较弱。

4) 建立水质参数与各荧光组分间的多元线性回归方程,pH、EC、DO、COD 及 BOD_5 和各荧光组分建立模型得到的相关系数 R 分别为 0.579、0.632、0.502、0.762 和 0.785,可以在一定程度上利用各荧光组分模拟水质参数的变化情况,为干旱区艾比湖流域的水资源保护提供一定的依据。

参 考 文 献

- [1] Du Shuxin, Du Yangfeng, Wu Xiaoli. Detection of dissolved organic matter based on three-dimensional first-order derivative fluorescence spectrometry[J]. Spectroscopy and Spectral Analysis, 2010, 30(12): 3268-3271.
杜树新, 杜阳锋, 武晓莉. 基于三维荧光导数光谱的水体有机污染物浓度检测[J]. 光谱学与光谱分析, 2010, 30(12): 3268-3271.

- [2] Bellos D, Sawidis T, Tsekos I. Nutrient chemistry of River Pinios (Thessalia, Greece)[J]. *Environment International*, 2004, 30(1): 105-115.
- [3] Zhang Y, Liang X, Wang Z, *et al.* A novel approach combining self-organizing map and parallel factor analysis for monitoring water quality of watersheds under non-point source pollution[J]. *Scientific Reports*, 2015, 5(3): 16079.
- [4] Lai Bo, Zhou Yuexi, Dou Lianfeng, *et al.* Variation trend of three-dimensional fluorescence characteristics of aromatic organic pollutants in the ABS wastewater treatment process[J]. *Acta Optica Sinica*, 2011, 31(1): 0130001.
赖波, 周岳溪, 窦连峰, 等. ABS 废水处理过程中芳香类有机污染物三维荧光特性的变化规律[J]. *光学学报*, 2011, 31(1): 0130001.
- [5] Xi Liuhua, Chen Guoqing, Zhu Zhuowei, *et al.* Determination of vintages of sesame flavor liquors by three-dimensional fluorescence spectroscopy[J]. *Laser & Optoelectronics Progress*, 2016, 53(12): 123002.
奚留华, 陈国庆, 朱焯炜, 等. 基于三维荧光光谱测定芝麻香型白酒的年份[J]. *激光与光电子学进展*, 2016, 53(12): 123002.
- [6] Stedmon C A, Sereďyńska-Sobecka B, Boe-Hansen R, *et al.* A potential approach for monitoring drinking water quality from groundwater systems using organic matter fluorescence as an early warning for contamination events[J]. *Water Research*, 2011, 45(18): 6030-6038.
- [7] Bierozna M, Baker A, Bridgeman J. Relating freshwater organic matter fluorescence to organic carbon removal efficiency in drinking water treatment[J]. *Science of the Total Environment*, 2009, 407(5): 1765-1774.
- [8] Murphy K R, Hambly A, Singh S, *et al.* Organic matter fluorescence in municipal water recycling schemes: toward a unified PARAFAC model[J]. *Environmental Science and Technology*, 2011, 45(7): 2909-2916.
- [9] Yu Huibin, Song Yonghui, Tu Xiang, *et al.* Assessing removal efficiency of dissolved organic matter in wastewater treatment using fluorescence excitation emission matrices with parallel factor analysis and second derivative synchronous fluorescence[J]. *Bioresource Technology*, 2013, 144: 595-601.
- [10] Hudson N, Baker A, Ward D, *et al.* Can fluorescence spectrometry be used as a surrogate for the biochemical oxygen demand (BOD) test in water quality assessment? An example from South West England[J]. *Science of the Total Environment*, 2008, 391(1): 149-158.
- [11] Liu Yanfang, Su Rongguo, Zhou Qianqian, *et al.* Rapid modeling offshore eutrophication technique using optical parameters of CDOM[J]. *Chinese J Lasers*, 2014, 41(12): 1215001.
刘艳芳, 苏荣国, 周倩倩, 等. 基于 CDOM 光学参数的近海富营养化快速评价技术[J]. *中国激光*, 2014, 41(12): 1215001.
- [12] Baghoth S A, Sharma S K, Amy G L. Tracking natural organic matter (NOM) in a drinking water treatment plant using fluorescence excitation-emission matrices and PARAFAC[J]. *Water Research*, 2011, 45(2): 797-809.
- [13] Kohonen T. *Self-organizing maps*[M]. 3rd ed. Berlin: Springer, 2001.
- [14] Kohonen T. Essentials of the self-organizing map[J]. *Neural Networks*, 2013, 37: 52-65.
- [15] Zhou P, Huang J, Pontius R G, *et al.* New insight into the correlations between land use and water quality in a coastal watershed of China: Does point source pollution weaken it?[J]. *Science of the Total Environment*, 2016, 543: 591-600.
- [16] Kalthoff A M, Hjorth P, Berndtsson R. Review of the self-organizing map (SOM) approach in water resources: Analysis, modelling and application[J]. *Environmental Modelling and Software*, 2008, 23(7): 835-845.
- [17] Céréghino R, Park Y S. Review of the Self-Organizing Map (SOM) approach in water resources: Commentary[J]. *Environmental Modelling and Software*, 2009, 24(8): 945-947.
- [18] Chon T S. Self-Organizing Maps applied to ecological sciences[J]. *Ecological Informatics*, 2011, 6(1): 50-61.
- [19] Li Wei, Yao Xiaoyan, Liang Zhiwei, *et al.* Assessment of surface water quality using self-organizing map and Hasse diagram technique[J]. *Acta Scientiae Circumstantiae*, 2013, 33(3): 893-903.
李伟, 姚笑颜, 梁志伟, 等. 基于自组织映射与哈斯图方法的地表水水质评价研究[J]. *环境科学学报*, 2013, 33(3): 893-903.
- [20] Zhang F, Tiyyip T, Johnson V C, *et al.* The influence of natural and human factors in the shrinking of the Ebinur Lake, Xinjiang, China, during the 1972-2013 period[J]. *Environmental Monitoring and Assessment*, 2015, 187(1): 4128.
- [21] Jilil Abduvali, Mu Guijin. Analysis on the dust storms and their disasters in the lakebed region of Ebinur Lake, Xinjiang[J]. *Arid Land Geography*, 2002, 25(2): 149-154.
吉力力·阿不都万里, 穆桂金. 艾比湖干涸湖底尘暴及其灾害分析[J]. *干旱区地理(汉文版)*, 2002, 25(2): 149-154.
- [22] Cui Y, Wu Q, Yang M, *et al.* Three-dimensional excitation-emission matrix fluorescence spectroscopy and fractions of dissolved organic matter change in landfill leachate by biological treatment[J]. *Environmental Science and Pollution*

- Research, 2016, 23(1): 793-799.
- [23] Stedmon C A, Bro R. Characterizing dissolved organic matter fluorescence with parallel factor analysis: A tutorial[J]. *Limnology and Oceanography Methods*, 2008, 6(11): 572-579.
- [24] Zhang S, Chen Z, Wen Q, *et al.* Assessing the stability in composting of penicillin mycelial dreg via parallel factor (PARAFAC) analysis of fluorescence excitation-emission matrix (EEM) [J]. *Chemical Engineering Journal*, 2016, 299: 167-176.
- [25] Li Juan, Hao Zhigang. Tourism planning based on Google Earth virtual earth platform[J]. *Remote Sensing for Land and Resources*, 2010, 22(1): 130-133.
李 娟, 郝志刚. 基于 Google Earth 虚拟地球平台的旅游规划研究[J]. *国土资源遥感*, 2010, 22(1): 130-133.
- [26] Gao Jie, Jiang Tao, Li Lulu, *et al.* Ultraviolet-visible (UV-Vis) and fluorescence spectral characteristics of dissolved organic matter (DOM) in soils of water-level fluctuation zones of the Three Gorges Reservoir Region[J]. *Environmental Science*, 2015, 36(1): 151-162.
高 洁, 江 韬, 李璐璐, 等. 三峡库区消落带土壤中溶解性有机质(DOM)吸收及荧光光谱特征[J]. *环境科学*, 2015, 36(1): 151-162.
- [27] Yu G H, He P J, Shao L M. Novel insights into sludge dewater ability by fluorescence excitation-emission matrix combined with parallel factor analysis[J]. *Water Research*, 2010, 44(3): 797-806.
- [28] Liu Xiaohan, Zhang Yunlin, Yin Yan, *et al.* Application of three-dimensional fluorescence spectroscopy and parallel factor analysis in CDOM study[J]. *Transactions of Oceanology and Limnology*, 2012(3): 133-145.
刘笑菡, 张运林, 殷 燕, 等. 三维荧光光谱及平行因子分析法在 CDOM 研究中的应用[J]. *海洋湖沼通报*, 2012(3): 133-145.
- [29] Song Xiaona, Yu Tao, Zhang Yuan, *et al.* Distribution characterization and source analysis of dissolved organic matters in Taihu Lake using a three-dimensional fluorescence excitation-emission matrix[J]. *Acta Scientiae Circumstantiae*, 2010, 30(11): 2321-2331.
宋晓娜, 于 涛, 张 远, 等. 利用三维荧光技术分析太湖水体溶解性有机质的分布特征及来源[J]. *环境科学学报*, 2010, 30(11): 2321-2331.
- [30] Li Haibo, Sun Chen, Liu Xiaoling, *et al.* Assessing organic matter removal from municipal wastewater by excitation-emission matrix fluorescence[J]. *China Environmental Science*, 2016, 36(8): 2371-2379.
李海波, 孙 晨, 刘晓玲, 等. 三维荧光光谱解析城市污水有机物的去除特征[J]. *中国环境科学*, 2016, 36(8): 2371-2379.
- [31] Zhang Y, van Dijk M A, Liu M, *et al.* The contribution of phytoplankton degradation to chromophoric dissolved organic matter (CDOM) in eutrophic shallow lakes: Field and experimental evidence[J]. *Water Research*, 2009, 43(18): 4685-4697.
- [32] Rochelle-Newall E J, Fisher T R. Production of chromophoric dissolved organic matter fluorescence in marine and estuarine environments: an investigation into the role of phytoplankton[J]. *Marine Chemistry*, 2002, 77(1): 7-21.