

# 基于偏最小二乘回归的藻类荧光光谱特征波长选取

余晓娅 张玉钧 殷高方 肖 雪 赵南京 段静波 石朝毅 方 丽

(中国科学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** 针对藻类荧光光谱解析中常见的信息冗余和光谱相关性问题, 基于偏最小二乘(PLS)的方法, 提出了区间蒙特卡罗偏最小二乘(IMC-PLS)方法, 有效地解决了特征波长的选取问题。根据特征色素荧光峰位置预选出特征区域, 综合利用了此特征区域内单个波段的信息和不同的随机波段组合对于模型的贡献, 基于荧光光谱的三线性特点, 联合了发射波长和激发波长的信息。研究表明, 与无信息变量消除算法(UVE)相比, IMC-PLS 反演 4 种藻类浓度得到的平均相对标准偏差分别降低了 0%、34.3%、55.9%、30.5%, 选择出的特征波长数和运算时间分别减少了 80.1%、81.3%, IMC-PLS 方法有效地解决了实时监测问题, 也为离散三维荧光光谱仪器的研制提供了理论支持。

**关键词** 光谱学; 特征波长; 区间蒙特卡罗偏最小二乘回归; 无信息变量消除; 荧光光谱; 藻类

**中图分类号** O433.4 **文献标识码** A **doi**: 10.3788/AOS201434.0930002

## Feature Wavelength Selection of Phytoplankton Fluorescence Spectra Based on Partial Least Squares

Yu Xiaoya Zhang Yujun Yin Gaofang Xiao Xue Zhao Nanjing Duan Jingbo  
Shi Chaoyi Fang Li

(Key Laboratory of Environmental Optics and Technology, Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei, Anhui 230031, China)

**Abstract** For spectral information redundancy and correlation in phytoplankton spectral analysis, interval Monte Carlo partial least squares (IMC-PLS) which effectively solves the problem of feature wavelength selection is presented based on partial least squares (PLS). Feature region is preselected according to the location of the pigment fluorescence peaks, the internal informations of a single band and the contributions of different random band combinations to the model are plenary used. Based on three-linear feature of fluorescence spectra, emission wavelength band and excitation wavelength band are considered as a unit. The result shows that comparing with the uninformative variable elimination (UVE), feature wavelength points and computation time obtained by IMC-PLS decrease by 80.1% and 81.3% and average relative tolerances (ARTs) by inversion of four algae concentrations decrease by 0%, 34.3%, 55.9%, 30.5%. IMC-PLS algorithm effectively solves the problem of real-time monitoring, and provides theoretical support for the development of a discrete three-dimensional fluorescence spectrometer meanwhile.

**Key words** spectroscopy; feature wavelength; interval Monte Carlo partial least squares; uninformative variable elimination; fluorescence spectra; phytoplankton

**OCIS codes** 300.6170; 300.6280; 260.2510

收稿日期: 2014-03-20; 收到修改稿日期: 2014-05-04

基金项目: 国家自然科学基金(61378041)、国家 863 计划(2014AA06A509, 2013AA065502)、安徽省杰出青年科学基金(1108085J19)、安徽省自然科学基金(1408085MD72)、中国科学院仪器设备功能开发技术创新项目(yg2012071)

作者简介: 余晓娅(1985—), 女, 博士研究生, 主要从事三维荧光光谱数据处理方面的研究。E-mail: xyu@aiofm.ac.cn

导师简介: 张玉钧(1964—), 男, 博士, 研究员, 主要从事光学信息处理和环境监测技术等方面的研究。

E-mail: yjzhang@aiofm.ac.cn(通信联系人)

本文电子版彩色效果请详见中国光学期刊网 www.opticsjournal.net

# 1 引言

近年来河流湖泊污染日益严重,给生产生活和生态系统的平衡带来了恶劣影响,亟需建立水质监测体系。藻类是湖泊中的主要污染物,识别不同藻类能够有效预警水质的变化<sup>[1-2]</sup>。目前常用的荧光光谱方法通常只能识别蓝藻、绿藻和硅藻三种藻<sup>[3-4]</sup>,但是,其他藻类在水体污染中也扮演者一定的角色,比如隐藻水华在淡水渔场经常出现<sup>[5]</sup>等,因此需要建立多种藻的识别体系。

在对藻类进行现场识别时,常用仪器采用二维发射光谱进行识别,存在识别藻种有限和正确率低的问题,本文尝试用离散三维荧光光谱来识别藻类,以期望达到识别的藻种增加、识别率提高的效果。三维荧光光谱可以提供比二维荧光光谱更多的信息<sup>[6]</sup>,但是用所有波段数据进行回归建模时,计算量很大,无法实现现场快速监测,而且一些光谱区域容易受到噪声干扰,或者基本不包含信息,直接把全光谱带入回归建模会影响计算精度。因此,需要对采集到的波长进行优选,以实现实时监测,使得预测结果误差最小。

常用的波长选择方法有遗传算法<sup>[7-8]</sup>,蚁群算法<sup>[9]</sup>,偏最小二乘(PLS)<sup>[10]</sup>和无信息变量消除算法(UVE)<sup>[11]</sup>。遗传算法和蚁群算法在数据点数较多的情况下,计算量大,收敛速度缓慢;PLS通过分别考察单个变量对于建模误差的影响,来决定最终建模时变量的去留;UVE则通过考察所有波长对应的偏最小二乘回归系数的稳定性来选择特征波长,实现简单,计算速度快,但是用PLS和UVE选择特征波长时,没有考虑到波长之间的相互联系以及进行组合建模时对模型的贡献。韩清娟<sup>[12]</sup>提出了用蒙特卡罗偏最小二乘方法(MC-PLS)选择特征波长,增强了特征波长选择结果的稳定性和可靠性,考虑了特征波长之间的联系,但是这种方法对单个波长反复进行蒙特卡罗运算耗费大量时间。受荧光区域索引<sup>[13]</sup>和荧光特征区域<sup>[14]</sup>的启发,本文提出了区间蒙特卡罗偏最小二乘方法(IMC-PLS),先用荧光区域积分选择出特征荧光区域,在此特征区域内,考虑荧光光谱的三线性特点,以每个波段为单位,用蒙特卡罗方法分别从发射波段和激发波段两个方面来随机选择波段组合,进行PLS建模计算预测误差,选择出特征发射波段和特征激发波段后,组合两组特征波段用PLS进行浓度反演。

本文给出特征波长选择方法的理论基础,用实测4种藻类荧光光谱数据检验所提方法,并与常用

的UVE进行对比,分析了方法的有效性。

# 2 方法与原理

## 2.1 PLS

PLS是Word于1983年提出的多元统计数据分析方法<sup>[15]</sup>,考察自变量之间的多重相关性,不是直接用因变量和自变量进行回归建模,而是在系统变量中提取若干对系统具有最佳解释能力的新综合变量,再对新综合变量进行回归建模。它将浓度矩阵 $\mathbf{Y}$ 和相应的量测响应矩阵 $\mathbf{X}$ 同时进行主成分分解,以提取新综合变量,即

$$\mathbf{X} = \mathbf{TP} + \mathbf{E}, \quad (1)$$

$$\mathbf{Y} = \mathbf{UQ} + \mathbf{F}, \quad (2)$$

式中 $\mathbf{T}$ 和 $\mathbf{U}$ 分别为 $\mathbf{X}$ 和 $\mathbf{Y}$ 的得分矩阵, $\mathbf{P}$ 和 $\mathbf{Q}$ 分别为 $\mathbf{X}$ 和 $\mathbf{Y}$ 的载荷矩阵, $\mathbf{E}$ 和 $\mathbf{F}$ 分别为运用PLS去拟合矩阵 $\mathbf{X}$ 和 $\mathbf{Y}$ 时所引进的误差。

PLS就是利用各列向量相互正交的特征响应矩阵 $\mathbf{T}$ 和特征浓度矩阵 $\mathbf{U}$ 进行回归建模,即

$$\mathbf{U} = \mathbf{TB}, \quad (3)$$

得到回归系数矩阵,又称关联矩阵 $\mathbf{B}$ ,表达式为

$$\mathbf{B} = (\mathbf{T}^T \mathbf{T}^{-1}) \mathbf{T}^T \mathbf{U}. \quad (4)$$

因此,PLS的主要步骤包括对浓度矩阵 $\mathbf{Y}$ 和相应的量测响应矩阵 $\mathbf{X}$ 的主成分分解以及对关联矩阵 $\mathbf{B}$ 的计算。

## 2.2 UVE

UVE是基于PLS的特征波长选择方法<sup>[16-17]</sup>,将随机噪声矩阵加入到光谱矩阵中,通过PLS、留一交叉验证等运算获得一组回归系数,计算各个波长下回归系数的平均值与标准差的比值来表征各个波长的稳定性,小于特定阈值的比值对应的波长变量被认为是无信息变量,予以删除。

UVE具体计算步骤如下:

1) 对光谱矩阵 $\mathbf{X}(n, p)$ ( $n$ 为样本数, $p$ 为波长点数)进行PLS建模,根据预测误差均方根(RMSEP,  $R_{\text{RMSEP}}$ )确定PLS模型的建模因子数 $A_{\text{LV}}$ ,记作 $A_1, A_2 = A_1, R_{\text{RMSEP}}$ 记为 $R_{\text{RMSEP}}^{\text{OLD}}$ 。

$$R_{\text{RMSEP}} = \left[ \sum_{i=1}^n (\hat{y}_i - y_i)^2 / n \right]^{1/2}, \quad (5)$$

式中 $\hat{y}_i$ 是估计值, $y_i$ 是测量值。

2) 在光谱矩阵 $\mathbf{X}$ 中加入人工噪声矩阵,新矩阵变为 $\mathbf{X}(n, 2p)$ 。

3) 将新矩阵按 $A_{\text{LV}} = A_2$ 进行 $n$ 次留一交叉建模,获得系数矩阵 $\mathbf{B}(n, 2p)$ 。

4) 计算每个波长对应的系数的标准差[std( $\cdot$ )]及

平均值 $[\text{mean}(\cdot)]$ ,并定义稳定性系数 $S(j)$ , $S(j) = \text{mean}[B(j)]/\text{std}[B(j)]$ ,其中 $B(j)$ 为第 $j$ 个波长下的系数。

5) 取人工噪声矩阵变量 $[p+1, 2p]$ 所对应的 $S(j)$ 的最大绝对值 $S_{\max}$ 为阈值,在 $[1, p]$ 内去除绝对值小于 $S_{\max}$ 对应的波长。

6) 对去除无信息变量后的新矩阵按 $A_{LV} = A_2$ 建模,获得 $R_{\text{RMSEP}}^{\text{NEW}}$ 。

7) 若 $R_{\text{RMSEP}}^{\text{NEW}} > R_{\text{RMSEP}}^{\text{OLD}}$ ,则算法结束。

8) 若 $R_{\text{RMSEP}}^{\text{NEW}} < R_{\text{RMSEP}}^{\text{OLD}}$ ,说明按 $A_{LV} = A_2$ 建模存在过拟合,则 $A_2 = A_2 - 1$ , $R_{\text{RMSEP}}^{\text{OLD}} = R_{\text{RMSEP}}^{\text{NEW}}$ ,返回步骤3)重新计算。

### 2.3 IMC-PLS

$\mathbf{X}(n, p, q)$  ( $q$ 为发射波长点数)是实测三维荧光光谱,选取特征波长的主要步骤如图1所示,图中EX代表激发波段,EM代表发射波段。

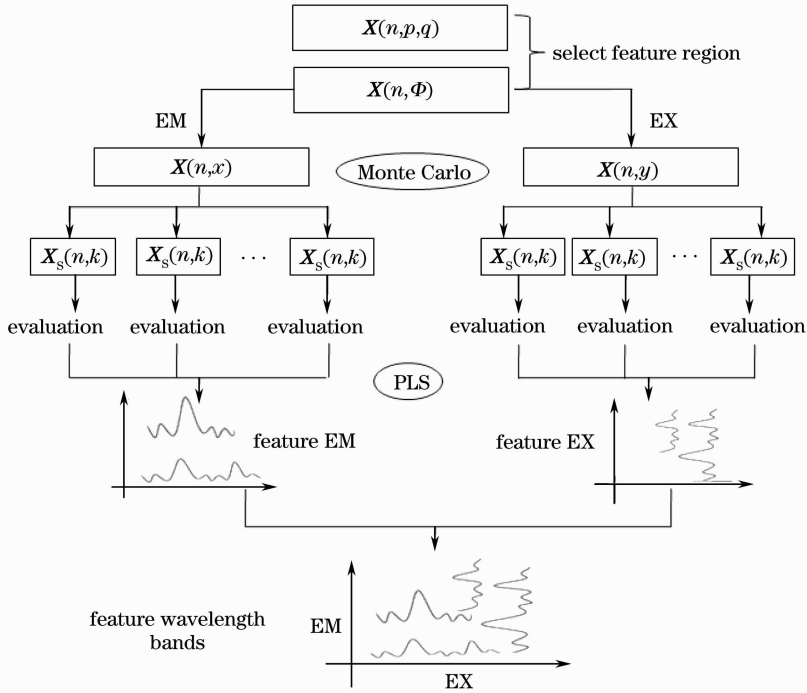


图1 光特征波长选择流程图

Fig.1 Flowchart of feature wavelength selection

1) 预选出特征区域 $\Phi$ 。在整个荧光区域( $p \times q$ )内单位区域的体积分 $\Delta\Phi$ 为

$$\Delta\Phi = \sum \sum I(\lambda_{\text{em}}, \lambda_{\text{ex}}) \Delta\lambda_{\text{em}} \Delta\lambda_{\text{ex}}, \quad (6)$$

式中 $I(\lambda_{\text{em}}, \lambda_{\text{ex}})$ 代表在发射波长 $\lambda_{\text{em}}$ 、激发波长 $\lambda_{\text{ex}}$ 处的荧光强度, $\Delta\lambda_{\text{em}}$ 为单位发射波长, $\Delta\lambda_{\text{ex}}$ 为单位激发波长。

特定区域( $3 \times 3$ )内总体积积分为

$$\Phi_i = \sum \Delta\Phi, \quad (7)$$

$$\Phi_i = \begin{cases} 0, & \Phi_i < \Phi_0 \\ \Phi_i, & \Phi_i \geq \Phi_0 \end{cases}, \quad (8)$$

式中 $\Phi_0$ 为阈值,设 $\Phi_0$ 为荧光强度最大值的 $1/3$ 。

特征区域为特定区域的并集,即

$$\Phi = \bigcup \Phi_i. \quad (9)$$

2) 在特征区域 $\Phi$ 内,分别以一个激发波段和一个发射波段为单位,进行多波段蒙特卡罗随机组合,

以发射波段为例,每次从 $x$ 个发射波段中,选择 $k$  ( $k \ll x$ )个发射波段组成 $\mathbf{X}_s(n, k)$ ,进行偏最小二乘交叉验证,计算不同的随机波段组合所建立模型的预测误差,设定组合内的每个波段都对应此误差。

3) 重复进行蒙特卡罗运算共 $M$  ( $M=100$ )次,以保证方法的稳定性和可靠性,分别计算每次运算涉及到的波段所对应的预测误差,找出预测误差总和最小的波段为特征波段。

4) 把特征激发波段和特征发射波段代入运算进行PLS,并计算浓度的平均相对误差(ART,  $R_{\text{ART}}$ )。

ART是指同一种藻在不同混合样品中回归得到的浓度的相对误差的平均值,其计算公式如下:

$$R_{\text{ART}} = \frac{1}{l} \sum_{j=1}^l R_{\text{RT}}^j, \quad (10)$$

式中 $R_{\text{RT}}^j$ 是第 $j$ 个样品中的相对误差, $l$ 为总样品数。

### 3 实验及其结果分析

#### 3.1 光谱测量

实验中选择 4 种我国主要淡水富营养化优势藻种在实验室进行纯种培养,其中蓝藻门、绿藻门、硅藻门和隐藻门各 1 种,分别为铜绿微囊藻、蛋白核小球藻、脆杆藻和卵形隐藻,所有藻类纯种培养体均由中国科学院水生生物研究所淡水藻种库提供。纯种

培养在光照培养箱中进行,在培养周期 15 d 结束后,取 4 种藻的液态活体培养体,用超纯水稀释至一定浓度,共配制 22 组不同浓度的混合样本。各样本暗适应 2 h 后,采用 F-7000 荧光分光光度计进行测量,扫描范围:激发波长为 350~700 nm,发射波长为 560~750 nm,步长为 5 nm,狭缝宽度为 5 nm,得到的三维荧光光谱如图 2 所示。

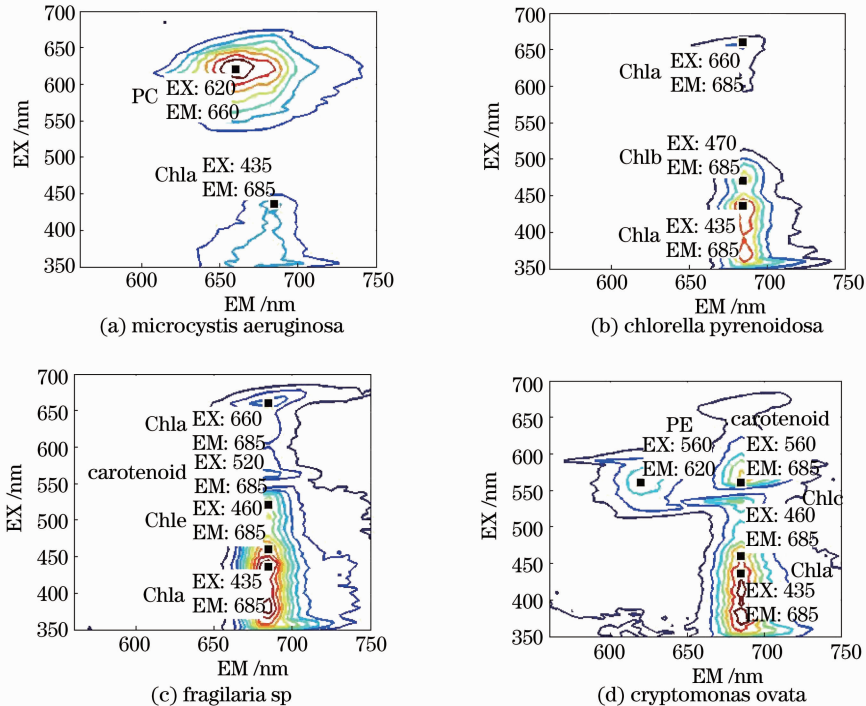


图 2 4 种藻荧光光谱

Fig. 2 Fluorescence spectra of 4 phytoplankton

不同的藻类含有不同的色素,因而具有不同的荧光特征。蓝藻中含有藻蓝蛋白(PC)和叶绿素 a(Chla),峰值分别在 620 nm/660 nm(激发波长/发射波长)和 435 nm/684 nm 处。绿藻中含有 Chla 和叶绿素 b(Chlb),峰值在 435 nm/684 nm 和 470 nm/684 nm 处,其中 660 nm/684 nm 处有 Chla 次吸收峰。硅藻含有 Chla、叶绿素 c(Chlc)和类胡萝卜素,峰值分别在 435 nm/684 nm、460 nm/684 nm 和 480~580 nm/684 nm 处,如图 2 所示,实测的硅藻荧光光谱在 520 nm/685 nm 处有峰值。隐藻中含有的色素较多,有 Chla、藻红蛋白(PE)、Chlc 和类胡萝卜素,峰值分别在 435 nm/684 nm、560 nm/620 nm、460 nm/684 nm 和 560 nm/684 nm 处。实测的 4 种藻类特征色素的峰值位置与 Beutler<sup>[18]</sup>分析的一致。

#### 3.2 预选出特征波长区域

用荧光区域积分提取特征区域,得到具有 529

个波长点的特征区域,分布如图 3 所示,图中方格部分即为所选择的特征区域,可以看出选出的特征区域包含了 4 种藻类的色素峰值点位置,验证了此特征区域选择方法的有效性。

#### 3.3 选择特征波长

把特征区域内的波长分别按行和按列进行排列,舍去长度小于 5 的波段,得到发射波段(图 1 中  $x$ )总共有 23 个和激发波段(图 1 中  $y$ )总共有 42 个,如图 4 所示。如图 4 中虚线所示,用蒙特卡罗方法每次选 3 个(图 1 中  $k$ )波段进行 PLS 建模,共进行 100 次蒙特卡罗计算和对应的建模,计算每次建模时不同波段组合所对应的预测误差,根据误差最小原则,选出了 3 个特征发射波段(分别为 630、685、695 nm)和 3 个特征激发波段(分别为 440、590、610 nm)。特征发射波段和特征激发波段包括的波长点共有 115 个(IMC-PLS1),考虑到波长间隔较小,将间隔扩大为原来的 2 倍,可以将波长点减

少为 57 个 (IMC-PLS), 两种情况下反演浓度得到的平均相对误差相差不大, 最终选择出的特征波长分布如图 5 所示。

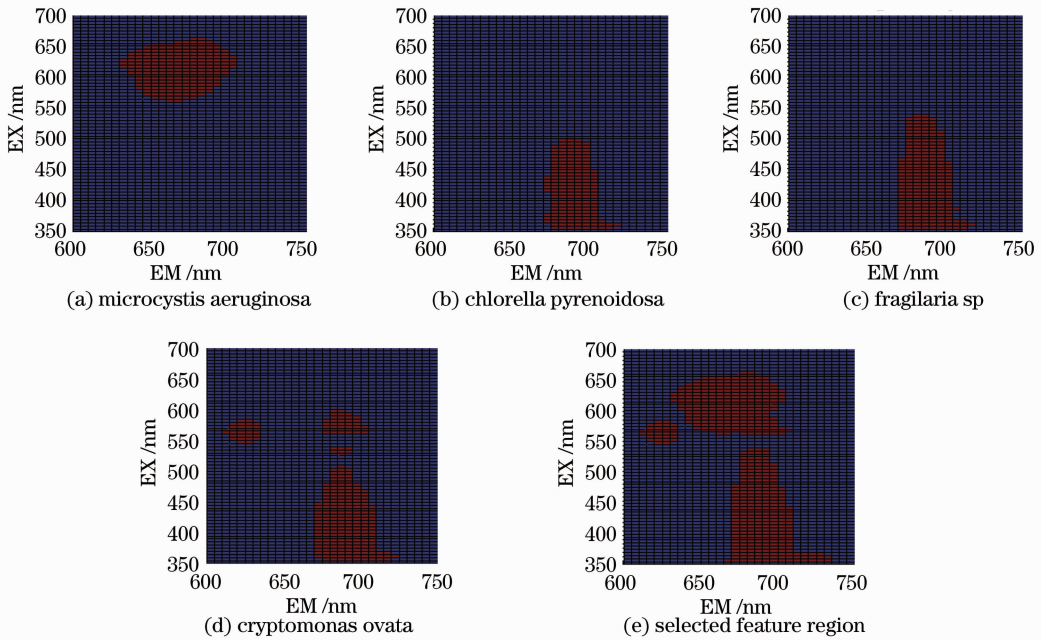


图 3 荧光区域积分选出的特征区域

Fig. 3 Feature regions selected by fluorescence regional integration

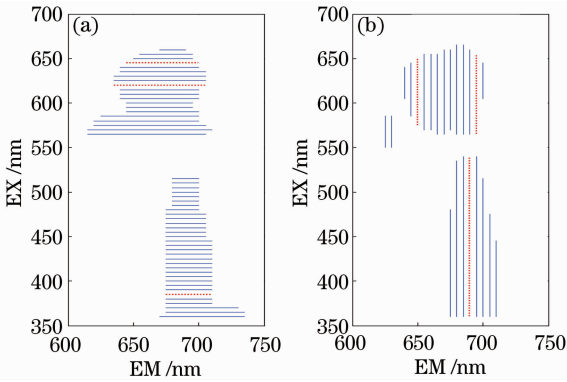


图 4 特征波段选择

Fig. 4 Feature wavelength band selections

为了检验所提方法的优越性, 把它与全波长、特征区域内波长和 UVE 提取的波长进行对比, 把 4 种情况下包含的波长点代入 PLS 对 4 种藻类进行

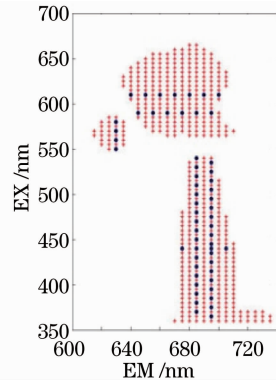


图 5 选出的特征波长

Fig. 5 Selected feature wavelengths

浓度反演, 从计算时间和 ART 两个方面进行评价, 结果如表 1 所示。

表 1 4 种方法的计算时间和 ART

Table 1 Computing time and ART of 4 methods

	Wavelength points	Computing time /s	ART			
			Microcystis aeruginosa	Chlorella pyrenoidosa	Fragilaria sp	Cryptomonas ovata
Total-PLS	3159	5	0.1864	0.1819	0.0777	0.1310
Feature region-PLS	529	1.4	0.0971	0.1218	0.0957	0.1267
UVE-PLS	287	0.75	0.1063	0.1837	0.1416	0.1872
IMC-PLS1	115	0.29	0.1045	0.1365	0.1093	0.1455
IMC-PLS	57	0.14	0.1064	0.1207	0.0625	0.1301

最终选择的特征波长点数越少,则需要的计算时间越短,如表 1 所示,从计算时间上看,IMC-PLS 需时最少,最为优越;从平均相对误差上看,预选区域得到的平均相对偏差最小,进一步验证了所用的提取特征区域方法的有效性。因为全波长中往往包含冗余信息,对于回归建模是不利的,全波长的 ART 不是最低的,这也说明了提取特征波长的必要性。UVE 方法得到的波长相对于全波长,波长点数和计算时间都减少了,蓝藻的 ART(0.1063)有了改善,但是对于其他三种藻则没有优越性,这是因为 UVE 考虑的是单个波长对于回归建模的贡献,没有考虑到多个波长对于回归建模的综合作用,此外,蓝藻的主要信息集中在藻蓝蛋白产生的荧光峰里,而其他三种藻包含的荧光谱峰较多,UVE 在单峰情况下有效,在多峰情况下则失去优越性。IMC-PLS 的 ART 和预选区域相比,总体上略有增加[分别为 0.0971、0.1064(蓝藻)和 0.1267、0.1301(隐藻)],这是因为 IMC-PLS 在提取特征波长的过程中会损失小部分信息,在一定程度上,波长点数和误差是一对不可调和的矛盾,波长点数的减少会增大产生误差的可能性,但是保证在误差允许的范围内尽量减少特征波长点数是必要的。

IMC-PLS 方法在两种指标上都最为优越,因为它以波段为单位,考虑了单个波段内所有波长点对于模型的贡献,而且运用蒙特卡罗选择不同波段组合,考虑了不同的随机波段组合对于模型的贡献。

## 4 结 论

提出了一种基于 PLS 的特征波长选择方法——区间蒙特卡罗偏最小二乘回归法。根据藻类荧光光谱的区域积分预选出光谱特征区域,缩小有效数据范围;利用三维荧光光谱的三线性特征,从发射波长波段和激发波长波段两方面分析;用蒙特卡罗随机选取波段组合,综合考虑了不同的随机波段组合对于模型的贡献;用偏最小二乘进行回归建模计算预测误差,进而决定波段的去留;用发射特征波段和激发特征波段进行藻类荧光光谱的浓度反演。虽然选择特征波长会花费一些时间,但是在藻种不变的情况下,只需在开始时进行特征波长选取,之后就可以在选取的特征波长基础上进行回归建模,缩短了运算时间,为研制离散三维荧光光谱仪器提供了理论支持。利用了三维荧光光谱的三线性特点,下一步计划基于立体空间的数据关系挖掘特征信息。

## 参 考 文 献

- 1 Y P Li, C Y Tang, Z B Yu, *et al.*. Correlations between algae and water quality: factors driving eutrophication in Lake Taihu, China [J]. *Int J Environ Sci Technol*, 2014, 11(1): 169–182.
- 2 Liu Jing, Liu Wenqing, Zhao Nanjing, *et al.*. Phytoplankton chlorophyll fluorescence characteristics excited by various light qualities and intensities [J]. *Acta Optica Sinica*, 2013, 33(9): 0930001.  
刘 晶, 刘文清, 赵南京, 等. 浮游植物在不同光质和光强激发下的叶绿素荧光特性[J]. *光学学报*, 2013, 33(9): 0930001.
- 3 Wang Zhigang, Liu Wenqing, Zhang Yujun, *et al.*. The classified measuring of three dimensional excitation-emission fluorescence matrix technique on phytoplankton concentration in water body [J]. *China Environmental Science*, 2008, 28(2): 136–141.  
王志刚, 刘文清, 张玉钧, 等. 三维荧光光谱法分类测量水体浮游植物浓度[J]. *中国环境科学*, 2008, 28(2): 136–141.
- 4 Wang Zhigang, Liu Wenqing, Zhang Yujun, *et al.*. The phytoplankton classified measure based on excitation fluorescence spectra technique [J]. *China Environmental Science*, 2008, 28(4): 329–333.  
王志刚, 刘文清, 张玉钧, 等. 基于激发荧光光谱的浮游植物分类测量方法[J]. *中国环境科学*, 2008, 28(4): 329–333.
- 5 L Mhlanga, W Mhlanga, P Tendaupenyu. Response of phytoplankton assemblages isolated for short periods of time in a hyper-eutrophic reservoir (Lake Chivero, Zimbabwe) [J]. *Water SA*, 2014, 40(1): 157–164.
- 6 Duan Yali, Su Rongguo, Shi Xiaoyong, *et al.*. Differentiation of phytoplankton populations by in vivo fluorescence based on high-frequency component of wavelet [J]. *Chinese J Lasers*, 2012, 39(7): 0715003.  
段亚丽, 苏荣国, 石晓勇, 等. 基于小波高频分量的浮游植物活体荧光识别技术研究[J]. *中国激光*, 2012, 39(7): 0715003.
- 7 R Leardi. Genetic algorithms in chemometrics and chemistry: a review [J]. *Journal of Chemometrics*, 2001, 15(7): 559–569.
- 8 Zhao Jiewen, Hui Zhe, Huang Lin, *et al.*. Quantitative detection of TVB-N content in chicken meat with hyperspectral imaging technology [J]. *Laser & Optoelectronics Progress*, 2013, 50(7): 073003.
- 9 赵杰文, 惠 喆, 黄 林, 等. 高光谱成像技术检测鸡肉中挥发性盐基氮含量[J]. *激光与光电子学进展*, 2013, 50(7): 073003.
- 10 M Shamsipur, V Zare-Shahabadi, B Hemmateenejad, *et al.*. Ant colony optimisation: a powerful tool for wavelength selection [J]. *Journal of Chemometrics*, 2006, 20(3-4): 146–157.
- 11 M Zhang, S Zhang, J Iqbal. Key wavelengths selection from near infrared spectra using Monte Carlo sampling-recursive partial least squares [J]. *Chemometrics and Intelligent Laboratory Systems*, 2013, 128: 17–24.
- 12 W Cai, Y Li, X Shao. A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra [J]. *Chemometrics and Intelligent Laboratory Systems*, 2008, 90(2): 188–194.
- 13 Han Qingjuan. *New Chemometric Algorithms and Their Application Studies for Multiway Spectroscopy data* [D]. Changsha: Hunan University, 2008.  
韩清娟. 多维光谱数据解析的化学计量学算法及应用研究[D]. 长沙: 湖南大学, 2008.
- 14 J Zhou, J J Wang, A Baudon, *et al.*. Improved fluorescence excitation-emission matrix regional integration to quantify spectra for fluorescent dissolved organic matter [J]. *Journal of Environmental Quality*, 2013, 42(3): 925–930.
- 15 Du Shuxin, Du Yangfeng, Yuan Zhibao. Characteristic region selection methods for three-dimensional fluorescence spectrometry [J]. *Chinese Journal of Luminescence*, 2012, 33(3): 341–345.  
杜树新, 杜阳锋, 袁之报. 三维荧光光谱的特征区域选择方法



- [J]. 发光学报, 2012, 33(3): 341—345.
- 15 V Esposito Vinzi, W W Chin, J Henseler, *et al.*. Handbook of Partial Least Squares: Concepts, Methods and Applications [M]. Heidelberg, Dordrecht, London, New York: Springer, 2010.
- 16 V Centner, D L Massart, O E de Noord, *et al.*. Elimination of uninformative variables for multivariate calibration [J]. Analytical Chemistry, 1996, 68(21): 3851—3858.
- 17 W Cai, Y Li, X Shao. A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra [J]. Chemometrics and Intelligent Laboratory Systems, 2008, 90(2): 188—194.
- 18 M Beutler. Spectral Fluorescence of Chlorophyll and Phycobilins as an *In-Situ* Tool of Phytoplankton Analysis-Models, Algorithms and Instruments [D]. Kile: Kile University, 2003.

栏目编辑：史敏