

doi:10.6041/j.issn.1000-1298.2018.01.022

基于离散粒子群和偏最小二乘的水源地浊度高光谱反演

曹引¹ 冶运涛² 赵红莉² 蒋云钟² 王浩^{1,2} 严登明¹

(1. 东华大学国家环境保护纺织工业污染防治工程技术中心, 上海 201620;

2. 中国水利水电科学研究院流域水循环模拟与调控国家重点实验室, 北京 100038)

摘要: 随着面源污染的加剧, 导致水源地水体富营养化程度日趋严重, 浊度作为衡量水体富营养化的一项重要指标, 是水质评价的重要参数。为降低浊度偏最小二乘(Partial least squares, PLS)反演模型建模的不确定性, 提高模型反演精度, 提出了基于离散粒子群(Discrete binary particle swarm optimization, DBPSO)和偏最小二乘的水体浊度反演模型。以2015年10月在南水北调东线重要水源地微山湖获取的水体浊度和准同步的HJ-1A HSI高光谱数据为例, 利用HJ-1A HSI B26-B105(中心波长: 518~870 nm)全谱段光谱反射率(Original spectral reflectance, OSR)和归一化光谱反射率(Normalized spectral reflectance, NSR)直接构建浊度OSR-PLS和NSR-PLS反演模型, 同时利用离散粒子群算法优选输入浊度PLS反演模型的最佳原始波段反射率和归一化光谱反射率, 在此基础上提出并构建浊度OSR-DBPSO-PLS和NSR-DBPSO-PLS反演模型; 然后对上述模型进行精度评价, 分析光谱归一化处理 and 特征波段优选对PLS模型反演精度的影响, 选择精度最高的模型反演微山湖水体浊度分布。结果表明: NSR-PLS模型精度($R^2 = 0.91$)高于OSR-PLS模型($R^2 = 0.50$), 对波段进行归一化处理能提高浊度PLS反演模型精度; DBPSO能够识别浊度PLS反演的最佳波段, 浊度PLS建模所需的波段数由80个分别减少为44个(OSR波段)和36个(NSR波段), 在此基础上构建的OSR-DBPSO-PLS模型($R^2 = 0.96$)和NSR-DBPSO-PLS模型($R^2 = 0.97$)均具有较高精度, 显著高于直接利用全谱波段构建的浊度PLS模型反演精度; 选择综合误差最小的NSR-DBPSO-PLS模型反演微山湖水体浊度, 反演结果符合实际, 该模型适用于HJ-1A HSI数据和内陆水体浊度反演。

关键词: 水源地; 浊度; 高光谱反演; 离散粒子群; 偏最小二乘; HJ-1A HSI

中图分类号: X832 文献标识码: A 文章编号: 1000-1298(2018)01-0173-10

Satellite Hyperspectral Retrieval of Turbidity for Water Source Based on Discrete Particle Swarm and Partial Least Squares

CAO Yin¹ YE Yuntao² ZHAO Hongli² JIANG Yunzhong² WANG Hao^{1,2} YAN Dengming¹

(1. State Environmental Protection Engineering Center for Pollution Control in Textile Industry, Donghua University, Shanghai 201620, China

2. State Key Laboratory of Simulation and Regulation of Water Cycle in River Basin, China Institute of Water Resources and Hydropower Research, Beijing 100038, China)

Abstract: With the increase of non-point source pollution emissions, the degree of eutrophication in water source is becoming seriously increase. Turbidity is an important parameter of water quality assessment, as an indicator of eutrophication. A discrete binary particle swarm optimization-partial least squares (PLS) model was proposed to reduce modeling uncertainty of turbidity retrieval using PLS and improve retrieval accuracy. A discrete binary particle swarm optimization was used to select original spectral reflectance and normalized spectral reflectance of concurrent HJ-1A HSI hyperspectral data with the turbidity obtained from October, 2015 in Weishan Lake as the input of partial least squares model. OSR-PLS and NSR-PLS model retrieving turbidity were developed by using original spectral reflectance

收稿日期: 2017-04-13 修回日期: 2017-05-27

基金项目: 国家重点研发计划项目(2017YFC0405801、2017YFC0405804)、国家自然科学基金项目(51309254)、高分辨率对地观测系统重大专项(08-Y30B07-9001-13/15-01)、中国水利水电科学研究院科研专项“十三五”重点科研项目(WR0145B272016)和中国水利水电科学研究院流域水循环模拟与调控国家重点实验室开放基金项目(IWHR-SKL-201517)

作者简介: 曹引(1991—), 男, 博士生, 主要从事流域水资源遥感及水循环模型数据同化研究, E-mail: yinc0416@163.com

通信作者: 冶运涛(1983—), 男, 高级工程师, 博士, 主要从事水信息学与智慧流域、流域水循环模拟研究, E-mail: yeyuntao@iwahr.com

and normalized spectral reflectance in full spectrum (HJ-1A HSI B26-B105 with 518 nm to 870 nm (central wavelength)). Meanwhile, the OSR-DBPSO-PLS and NSR-DBPSO-PLS models were developed to retrieve turbidity by using the selected original spectral reflectance and normalized spectral reflectance. The influence of spectral normalized and the characteristic band optimized on PLS model retrieval accuracy were analyzed based on the four models' elevation. Finally, the highest accuracy model was used to retrieve the turbidity distribution in Weishan Lake. The results indicated that the accuracy of NSR-PLS ($R^2 = 0.91$) model was better than that of OSR-PLS model ($R^2 = 0.50$). The normalization of reflectance can improve PLS accuracy of turbidity retrieval. DBPSO can identify the optimal original and normalize spectral reflectance. The number of bands required for turbidity PLS modelling was reduced from 80 to 44 (OSR) and 36 (NSR), respectively. The OSR-DBPSO-PLS ($R^2 = 0.96$) and NSR-DBPSO-PLS ($R^2 = 0.97$) modelling based on 44 OSR and 36 NSR, respectively, had high accuracies, which were significantly higher than the OSR-PLS and NSR-PLS modelling by full spectrum. The NSR-DBPSO-PLS model with minimal comprehensive error was selected to retrieve turbidity in Weishan Lake, which was suitable for inland water turbidity retrieval based on HJ-1A HSI data.

Key words: water source; turbidity; hyperspectral retrieval; discrete particle swarm; partial least squares; HJ-1A HSI

0 引言

湖泊作为人类重要的水源地,具有灌溉、防洪、供水和养殖以及重要的生态系统功能^[1]。随着社会经济的高速发展,农业生产逐渐由传统农业向具有集约化、专业化、智能化等特征现代农业发展^[2-3],以高投入、高产出、高排放为特征的现代农业带来严重的面源污染,主要表现为农业生产中化肥低利用率导致大量无法被植物吸收利用的氮、磷、钾等养分直接随着灌溉水、地表径流等进入水源地^[3],加上生活污水直接排放带来的营养物质,导致水源地水体富营养化日趋严重^[4]。浊度阻碍光在水体中的传输,对水中浮游植物的生长具有重要影响,是表征水体富营养化状况的一项重要指标^[5]。当水生植物分布较少时,水体浊度由总悬浮浓度主导,其变化可以识别水体悬浮物的输移、沉淀以及再悬浮过程^[6]。浊度作为一种重要的水质参数,可有效表征水质污染状况,遥感技术作为一种区域性监测手段,越来越多地应用于水质监测和水环境管理,为实现水体浊度高频次、大范围的监测提供了可能。

水质遥感监测的关键在于遥感反演模型的构建。目前基于实测光谱和卫星遥感数据的水体浊度反演模型主要有经验模型^[7-8]、半分析模型^[9-10]和分析模型^[11]。相对于经验和半分析模型,虽然以生物光学模型为代表的分析模型具有明确的物理意义和较好的移植性,但所需的水质参数固有光学特性却难以准确获取,导致生物光学模型在水质监测中应用较少^[12];经验模型和半分析模型仅能利用一个或几个光谱波段信息,限制了模型反演精度的进

一步提高。偏最小二乘法作为一种多元回归方法,可以利用多个波段信息来反演水质参数浓度,在叶绿素 a 浓度和总悬浮物浓度反演中应用较多,在浊度反演中应用较少。有研究表明,当输入偏最小二乘模型的波段数远大于样点数时,模型的不确定性大大增加^[13]。对于偏最小二乘法建模采用波段个数的确定,诸多学者采用了不同的方法^[14-24]。上述学者研究的水质指标主要是叶绿素 a 和总悬浮物浓度,鲜有水体浊度偏最小二乘反演相关研究;所采用的数据源以实测高光谱数据源为主,基于卫星高光谱数据源的偏最小二乘模型在水体浊度反演中的适用性尚缺乏研究。

针对上述不足,本文以典型内陆浅水湖微山湖为研究区,以 HJ-1A HSI 高光谱影像为数据源,首先对 HJ-1A HSI 高光谱反射率进行归一化处理,分别构建基于原始光谱反射率和归一化光谱反射率的浊度偏最小二乘反演模型,同时利用离散粒子群算法(DBPSO)确定浊度偏最小二乘(PLS)模型建模所需的特征波段,然后基于优选的特征波段提出并构建水体浊度 DBPSO-PLS 反演模型,研究光谱归一化处理和特征波段优选对偏最小二乘模型精度的影响;最后选择精度最高的模型反演微山湖水体浊度分布,研究基于卫星高光谱数据源的偏最小二乘模型在水体浊度反演中的适用性。

1 数据

1.1 研究区概况

微山湖地处 116°58' ~ 117°21' E, 34°27' ~ 34°52' N,位于山东省南四湖的最南端,隶属于山东省济宁市微山县,微山湖是南四湖面积最大的一个

湖泊,面积达660 km²,是中国北方最大的淡水湖;其平均水深不足2 m,春夏湖泊中生长有各种水生植物,是典型的内陆浅水草型湖泊;微山湖作为南水北调东线工程重要的湖泊水源地,同时具有防洪、灌溉、养殖和航运等生态系统服务功能,遥感监测微山湖水体浊度对南水北调东线水质的监控与管理具有重要意义。

1.2 数据获取与处理

2015年10月16—17日,在微山湖布设了23个采样点(图1),利用哈希浊度计现场测量浊度。各采样点处水体浊度如图2所示。由图2可以看出,水体浊度较高,且不同采样点差异显著,这是因为10月末,微山湖生长的水生植物腐烂降解形成大量悬浮物质,导致水体十分浑浊。

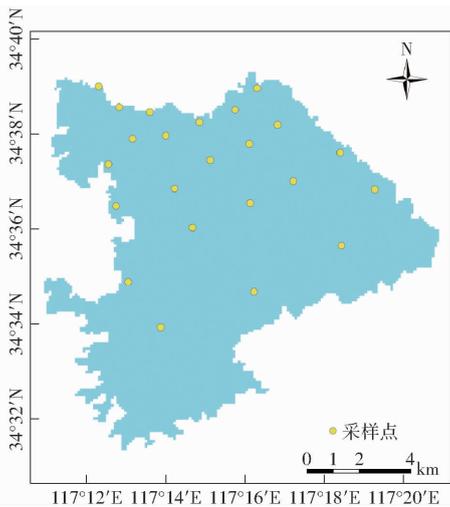


图1 微山湖采样点分布

Fig. 1 Distribution of sampling site in Weishan Lake

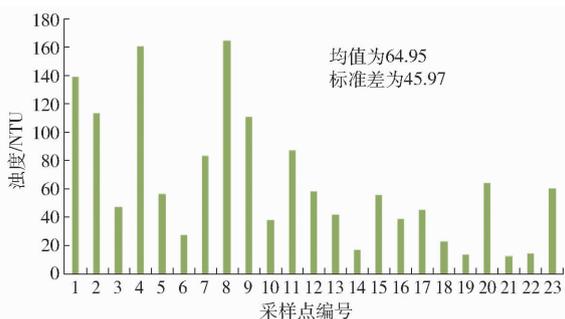


图2 各采样点处浊度

Fig. 2 Turbidity in 23 sampling points

HJ-1A HSI 作为我国第一个高光谱卫星传感器,空间分辨率为100 m,具有115个波段,但由于B1~B25(中心波长:460.04~516.17 nm)和B106~B115(中心波长:877.52~951.54 nm)受噪声干扰,仅B26~B105(中心波长:518.805~870.005 nm)共80个波段适用于水体浊度反演。本研究从中国资源卫星应用中心(<http://www.cresda.com/CN>)获取到一景2015年10月16日的高光谱影像,对HJ-

1A HSI 高光谱影像进行预处理,以经过几何校正的Landsat 8影像作为基准影像进行几何校正,利用FLAASH模型进行大气校正,最后提取微山湖水体区域(图1)。

2 方法

2.1 偏最小二乘法

偏最小二乘法(PLS)是集主成分分析、典型相关分析于一体的一种多元回归方法,其建模原理如下:自变量光谱矩阵 $\mathbf{X}_{n \times p}(x_1, x_2, \dots, x_p)$ 和因变量水质参数矩阵 $\mathbf{Y}_{n \times q}(y_1, y_2, \dots, y_q)$ 分别代表PLS模型的输入和输出。其中 n 代表建模样本数, p, q 分别代表PLS建模输入HJ-1A HSI光谱波段数和输出水质参数个数, $p \leq 80, q = 1$ 。 \mathbf{Y} 即为建模样点处水体浊度,对光谱矩阵 \mathbf{X} 和浊度矩阵 \mathbf{Y} 进行标准化处理得到 \mathbf{E}_0 和 \mathbf{F}_0 ,然后分别从 \mathbf{E}_0 和 \mathbf{F}_0 提取第1个主成分 t_1 和 u_1 ,计算去除主成分信息的残差矩阵 \mathbf{E}_1 和 \mathbf{F}_1 ,然后利用 \mathbf{E}_1 和 \mathbf{F}_1 替换 \mathbf{E}_0 和 \mathbf{F}_0 提取第2个主成分,循环迭代,依次提取主成分 t_1, t_2, \dots, t_h 和 u_1, u_2, \dots, u_h ,最佳主成分数 h 采用“留一”交互验证法来确定,具体过程参照文献[14];最后利用提取的最佳主成分构建偏最小二乘模型。

偏最小二乘法提取的主成分既要携带光谱矩阵的变异信息,即不同样点光谱之间的差异特征,同时也要对浊度矩阵具有很好的解释能力,不同主成分对光谱矩阵中各个波段光谱和浊度矩阵的解释能力计算参照文献[25]。

2.2 离散粒子群算法

偏最小二乘模型若输入波段过多,模型的复杂性和建模不确定性将会增加^[13],此外输入的光谱波段并非全部对浊度反演敏感,因此需要对PLS建模输入波段进行优化,确定水体浊度PLS反演的最佳波段。离散粒子群算法(DBPSO)是KENNEDY和EBERHART^[26]在粒子群算法基础上提出的一种离散群体智能全局优化算法,具有计算效率高、收敛速度快和鲁棒性强等优势^[27],广泛应用于各种离散优化求解。DBPSO算法搜索步骤如下:

(1)初始化粒子位置 $\mathbf{X}_k = (x_{k,1}, x_{k,2}, \dots, x_{k,m})$ 和速度 $\mathbf{V}_k = (v_{k,1}, v_{k,2}, \dots, v_{k,m})$,每个粒子位置代表一组HJ-1A HSI高光谱反射率组合,包含 p 个0或1,0代表对应的波段未被选中,1代表对应的波段被选中。其中 p 为总波段数, k 为迭代次数,初始取 $k=0$;粒子速度用于更新粒子位置, m 为粒子个数,取 $m=20$,下同。

(2)结合PLS模型,计算粒子适应度,适应度利用浊度PLS模型验证均方根误差(Root mean square

error, RMSE) 和建模决定系数 R^2 比值表示为

$$F_{fitness} = R_{MSE} / R^2 \quad (1)$$

其中

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2)$$

$$R_{MSE} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (3)$$

式中 y_i ——样本点 i 浊度实测值

\bar{y} ——实测浊度的平均值

\hat{y}_i ——样本点 i 浊度反演值

n ——建模或验证样本点个数

(3) 确定个体极值 $P_k = (p_{k,1}, p_{k,2}, \dots, p_{k,m})$ 和全局极值 $P_{g,k} = \max(p_{k,1}, p_{k,2}, \dots, p_{k,m})$, 个体极值代表使水体浊度 PLS 反演模型适应度最小的各个粒子, 全局极值代表所有粒子中使水体浊度 PLS 反演模型适应度最小的粒子, 根据个体极值和全局极值计算粒子搜索速度, 然后根据粒子速度更新粒子位置。

$$v_{k+1,i} = wv_{k,i} + c_1 R_{and1} (p_{k,i} - x_{k,i}) + c_2 R_{and2} (P_{g,k} - x_{k,i}) \quad (4)$$

$$S_{ig}(v_{k+1,i}) = 1 / (1 + \exp(-v_{k,i})) \quad (5)$$

$$x_{k+1,i} = \begin{cases} 1 & (R_{and_{k+1,i}} < S_{ig}(v_{k+1,i})) \\ 0 & (R_{and_{k+1,i}} \geq S_{ig}(v_{k+1,i})) \end{cases} \quad (6)$$

式中 $v_{k+1,i}$ ——第 $k+1$ 次迭代粒子 i 的搜索速度

$v_{k,i}$ ——第 k 次迭代粒子 i 的搜索速度

$x_{k+1,i}$ ——第 $k+1$ 次迭代时粒子 i 的位置

c_1 ——搜索控制因子, 控制算法的局部搜索能力

c_2 ——搜索控制因子, 控制算法的全局搜索能力

w ——惯性权重

$R_{and_{k+1,i}}$ ——0~1 之间的随机数

$S_{ig}(v_{k+1,i})$ ——粒子第 i 维取 1 的概率

由式(4)、(5)可以看出粒子越接近最优解, $v_{k+1,i}$ 越趋向 0, 此时维变化率 $S_{ig}(v_{k+1,i})$ 接近 0.5, 随机性很高, 全局搜索能力强, 但局部搜索能力差, 难以收敛, 为克服这一不足, 改变粒子位置更新方式^[28]

$$S_{ig}(v_{k+1,i}) = \begin{cases} 1 - \frac{2}{1 + \exp(-v_{k,i})} & (v_{k,i} \leq 0) \\ \frac{2}{1 + \exp(-v_{k,i})} - 1 & (v_{k,i} > 0) \end{cases} \quad (7)$$

$$x_{k+1,i} = \begin{cases} 0 & (R_{and_{k+1,i}} \leq S_{ig}(v_{k+1,i}), v_{k,i} \leq 0) \\ x_{k,i} & (R_{and_{k+1,i}} > S_{ig}(v_{k+1,i}), v_{k,i} \leq 0) \\ 1 & (R_{and_{k+1,i}} \leq S_{ig}(v_{k+1,i}), v_{k,i} > 0) \\ x_{k,i} & (R_{and_{k+1,i}} > S_{ig}(v_{k+1,i}), v_{k,i} > 0) \end{cases} \quad (8)$$

此时, 当 $v_{k+1,i}$ 越趋向 0, 粒子接近最优解时快速实现收敛, 具有很好的局部搜索能力, 但全局搜索能力较差。为了保证算法在迭代前中期具有很好的全局搜索能力, 在迭代后期具有很好的局部搜索能力, 将两种粒子位置更新方式进行组合, 在迭代前中期(迭代次数小于等于最大迭代次数的 70%) 利用式(5)和式(6)对粒子位置进行更新, 该种更新方式记为 style1, 在迭代后期(迭代次数大于最大迭代次数的 70%) 时利用式(7)、(8)对粒子位置进行更新, 该种更新方式记为 style2。此时离散粒子群算法在迭代前期具有很强的全局搜索能力, 在迭代后期具有很强的局部搜索能力, 可最终实现算法的收敛。

2.3 离散粒子群-偏最小二乘法

将上述离散粒子群和偏最小二乘算法相结合, 利用离散粒子群优化算法确定输入偏最小二乘模型中的最佳光谱波段, 构建水体浊度 DBPSO-PLS 反演模型, DBPSO-PLS 模型构建详细步骤如下:

(1) 设置参数值(惯性权重 w , 搜索控制因子 c_1, c_2), 初始化粒子位置和速度。由于光谱波段反射率为离散数据, 粒子位置中用 0、1 表征对应波段是否被选中, 其中 0 代表未选中, 1 代表选中。

(2) 结合 PLS 模型计算各个粒子 PLS 浊度反演的适应度, 确定个体极值和全局极值。根据各个粒子位置在建模数据集中分别选择粒子选中的波段或特征变量作为输入, 水体浊度作为输出构建偏最小二乘模型, 计算粒子适应度, 确定个体极值和全局极值。

(3) 更新粒子速度, 根据迭代次数选择粒子位置更新方式。

(4) 循环迭代步骤(2)、(3), 直至达到最大迭代次数。迭代次数小于等于最大迭代次数的 70% 时根据 style1 进行位置更新, 迭代次数大于最大迭代次数的 70% 时根据 style2 进行位置更新。

(5) 选择全局极值, 构建水体浊度最佳偏最小二乘模型。

水体浊度 DBPSO-PLS 反演模型构建流程如图 3 所示。

2.4 精度评价

利用建模决定系数 R^2 、相对均方根误差 r_{RMSE}

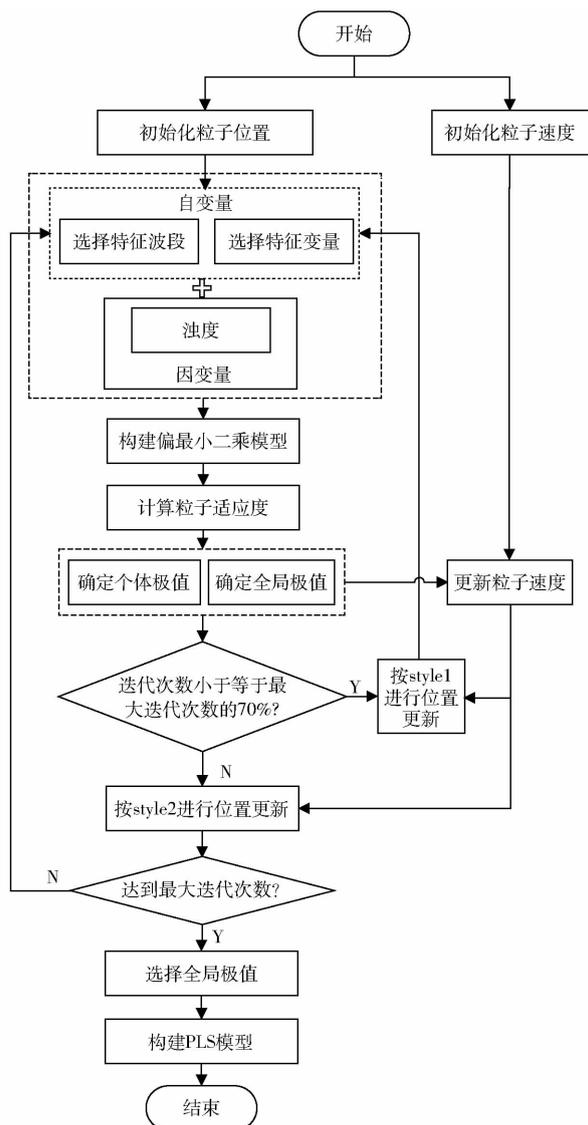


图 3 DBPSO-PLS 反演模型构建流程图

Fig. 3 Flow chart of DBPSO-PLS retrieval model construction

(Relative root mean square error)、平均相对误差 A_{RE} (Average relative error) 和综合误差 C_E (Comprehensive error) 来衡量浊度反演模型精度, 相关计算公式为

$$r_{RMSE} = \frac{R_{MSE}}{y} \quad (9)$$

$$A_{RE} = \sum_{i=1}^n |(\hat{y}_i - y_i)/y_i| / n \times 100\% \quad (10)$$

$$C_E = \frac{C_{E_c} + C_{E_v}}{2} = \frac{r_{RMSE_c} + A_{RE_c} + r_{RMSE_v} + A_{RE_v}}{4} \quad (11)$$

式中下角_c表示建模, _v表示验证。

3 结果与分析

3.1 基于 HJ-1A HSI 的 PLS 模型构建与分析

本次试验共布设 23 个采样点, 其中 15 个采样点

用于建模, 8 个采样点用于验证模型精度, 下同。选择 B26 ~ B105 (中心波长: 518.805 ~ 870.005 nm) 共 80 个波段反射率作为自变量 $X_{n \times p}$, 对应的浊度作为因变量 $Y_{n \times q}$ ($n = 15, p = 80, q = 1$, 下同), 构建基于原始光谱反射率的浊度偏最小二乘反演模型 (OSR-PLS)。有研究认为对光谱进行归一化处理能减少天气条件对不同反射率数值的影响^[29], 因此对光谱进行归一化处理

$$\bar{R}(B_i) = \frac{R(B_i)}{\frac{1}{n} \sum_{i=1}^n R(B_i)} \quad (12)$$

式中 $\bar{R}(B_i)$ —— B_i 波段归一化反射率

$R(B_i)$ —— B_i 波段原始光谱反射率

n —— 波段总数, $n = 80$

利用 B26 ~ B105 归一化波段反射率作为自变量 $\bar{X}_{n \times p}$, 和对应的浊度作为因变量 $Y_{n \times q}$ ($n = 15, p = 80, q = 1$), 构建基于归一化光谱反射率的浊度偏最小二乘反演模型 (NSR-PLS); 浊度 OSR-PLS 和 NSR-PLS 提取的最佳主成分个数以及主成分对光谱矩阵和浊度矩阵的解释能力见表 1, 所有主成分对各个波段光谱的累积解释能力如图 4a 所示。对比发现, NSR-PLS 模型提取的 2 个主成分对光谱 (NSR) 矩阵的累积解释能力要明显低于 OSR-PLS 的 2 个主成分对光谱 (OSR) 矩阵的解释能力, 但对浊度矩阵的累积解释能力 NSR-PLS 提取的 2 个主成分要显著高于 OSR-PLS 模型提取的 2 个主成分。这是因为对光谱进行归一化处理可以降低天气等条件对不同样点光谱的影响, 减小和水体组分无关因素给光谱矩阵带来的变异信息, 保留因水体组分差异导致的光谱变异信息, 提高了 PLS 模型反演精度。由图 4a 可以看出, OSR-PLS 模型提取的 2 个主成分对 520 ~ 870 nm 的光谱均具有较高的累积解释能力, 但对浊度矩阵的解释能力偏低 (表 1), 说明原始光谱变异信息较大, 但仅部分变异信息和水体浊度有关, 导致提取的主成分无法对浊度矩阵进行充分解释; 相比之下, NSR-PLS 提取的 2 个主成分对红波段和近红外波段光谱的解释能力整体高于

表 1 主成分对光谱矩阵和浊度矩阵的累积解释能力

Tab. 1 Explanation ability for spectrum and turbidity of principal components

模型	主成分个数	对光谱矩阵的累积解释能力/%	对浊度矩阵的累积解释能力/%
OSR-PLS	2	82.67	50.21
NSR-PLS	2	42.51	91.19
OSR-DBPSO-PLS	4	90.23	96.44
NSR-DBPSO-PLS	5	58.14	97.35

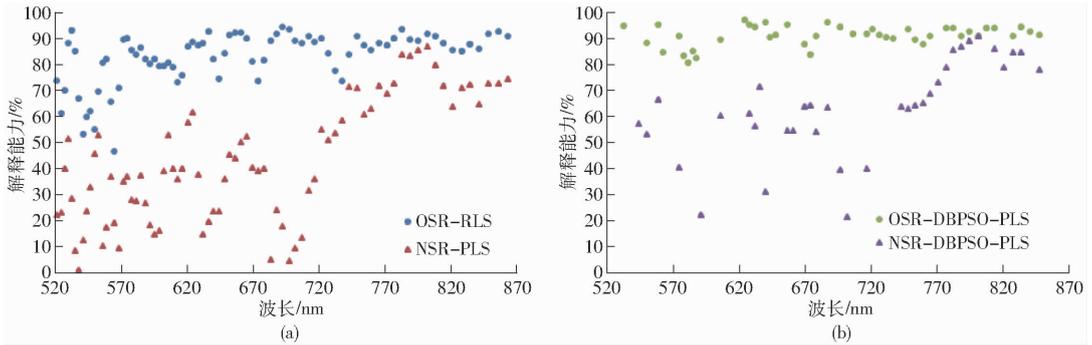


图4 各模型主成分对不同波段光谱的解释能力

Fig. 4 Explanation ability of principles extracted by models for different spectrums

其他波段的光谱(图4a),这是因为红波段和近红外波段光谱的变异信息主要由总悬浮物浓度和浊度主导,对浊度反演较为敏感^[30],所以NSR-PLS模型提取的主成分对浊度矩阵具有较高的解释能力。

浊度OSR-PLS和NSR-PLS反演模型散点图如图5所示。对浊度OSR-PLS和NSR-PLS反演模型进行精度评价,评价结果见表2。图5可以直观显示两种模型建模和验证样点的分布以及拟合效果,表2则是对两种模型建模和验证精度的定量描述。由图5和表2可以看出,基于原始光谱反射率的PLS模型建模拟合效果较差,建模决定系数 R^2 只有0.50(即主成分对浊度的累积解释能力),模型反演值和实测值分散在1:1线两侧,部分样点模型浊度反演值和实测值的散点偏离1:1线幅度较大,模型建模综合误差为46.36%,验证综合误差为54.57%,整个模型的综合误差达到50.46%;基于

归一化光谱反射率的PLS模型较基于原始光谱反射率的PLS模型建模精度有明显提高,建模 R^2 达到0.91,建模和验证散点均靠近1:1线,建模综合误差为21.32%,验证综合误差为44.14%,模型综合误差为32.73%。对比分析可以发现,对光谱进行归一化处理可以提高偏最小二乘模型反演精度。

3.2 基于HJ-1A HSI的DBPSO-PLS模型构建与分析

偏最小二乘模型虽能综合利用全谱波段信息,但并非全部波段反射率对水体浊度反演都敏感,直接利用全谱波段构建水体浊度反演模型,会因部分不敏感波段的输入限制模型精度。利用DBPSO算法对输入PLS模型的B26~B115共80个波段进行优选,确定反演浊度的最佳波段,利用确定的最佳波段构建浊度DBPSO-PLS模型;同时基于归一化光谱反射率构建水体浊度DBPSO-PLS模型,研究DBPSO优选特征波段对PLS模型反演精度的影响。

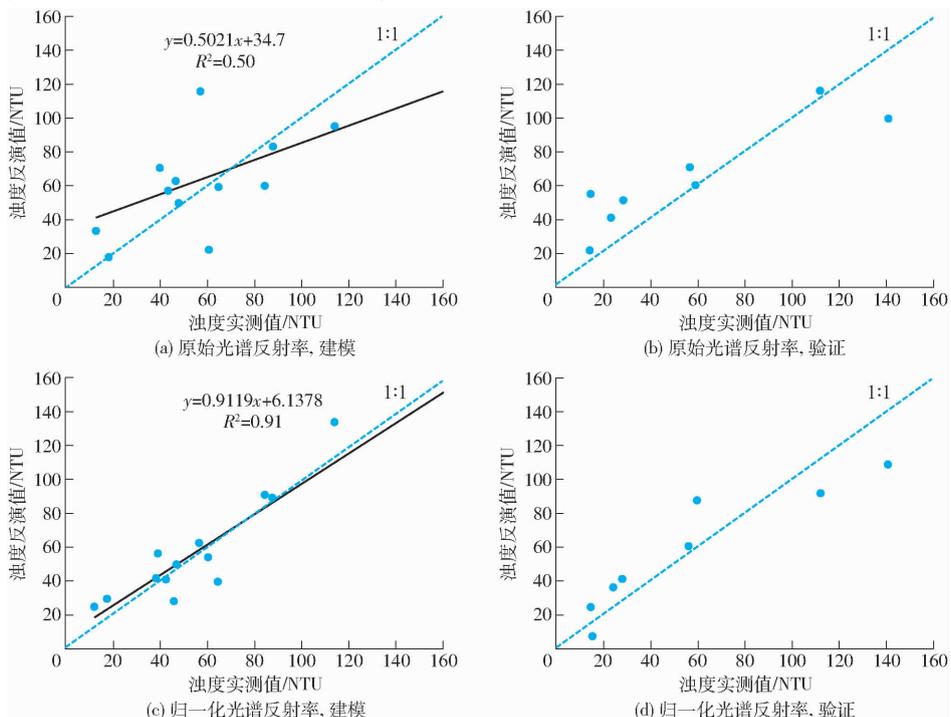


图5 浊度PLS建模和验证散点图

Fig. 5 Modeling and validation scatter plots of PLS retrieving turbidity

表 2 浊度偏最小二乘模型精度评价表

Tab. 2 Accuracy evaluation of PLS model retrieving turbidity based on original spectral reflectance

模型	波段数	$r_{RMSE_c}/\%$	$A_{RE_c}/\%$	$C_{E_c}/\%$	$r_{RMSE_v}/\%$	$A_{RE_v}/\%$	$C_{E_v}/\%$	$C_E/\%$
OSR-PLS	80	45.34	47.38	46.36	41.93	67.20	54.57	50.46
NSR-PLS	80	19.07	23.58	21.32	35.56	52.71	44.14	32.73
OSR-DBPSO-PLS	44	12.13	19.67	15.90	6.93	7.93	7.43	11.66
NSR-DBPSO-PLS	36	10.47	13.23	11.85	5.00	6.43	5.71	8.78

利用 DBPSO 算法分别优选水体浊度反演的最佳原始光谱反射率波段和归一化反射率波段。DBPSO 算法中, $w = 1, c_1 = c_2 = 2$, 初始化 20 个粒子, 速度限制为 $[-4, 4]$, 设置迭代次数为 300 次。DBPSO 算法优选浊度 PLS 反演模型所需最佳光谱反射率波段和归一化光谱反射率波段的适应度函数变化如图 6 所示; 统计最终确定用于浊度 PLS 反演模型构建的最佳原始光谱反射率和归一化光谱反射率波段分布如图 7 所示。分别利用优选的最佳原始光谱反射率和归一化光谱反射率波段构建浊度 PLS 反演模型 (OSR-DBPSO-PLS 和 NSR-DBPSO-PLS 模型), 模型散点图如图 8 所示。浊度 OSR-DBPSO-PLS 和 NSR-DBPSO-PLS 提取的最佳主成分个数分别为 4 个和 5 个

(表 1), 主成分对 DBPSO 确定的光谱矩阵中各个波段光谱的累积解释能力如图 4b 所示。由表 1 可以看出, OSR-DBPSO-PLS 模型提取的 4 个主成分对光谱矩阵和浊度矩阵的累积解释能力分别为 90.23% 和 96.44%; NSR-DBPSO-PLS 模型提取的 5 个主成分对光谱矩阵和浊度的累积解释能力分别为 58.14% 和 97.35%, 与 OSR-PLS 模型、NSR-PLS 模型主成分对光谱矩阵中的各个波段光谱的解释能力类似 (图 4a, 4b), 但 OSR-DBPSO-PLS 模型和 NSR-DBPSO-PLS 模型提取的主成分对浊度矩阵的累积解释能力要高于 OSR-PLS 和 NSR-PLS 模型提取的主成分, 说明 DBPSO 优选后的特征波段光谱包含更多对浊度反演敏感的变异信息。

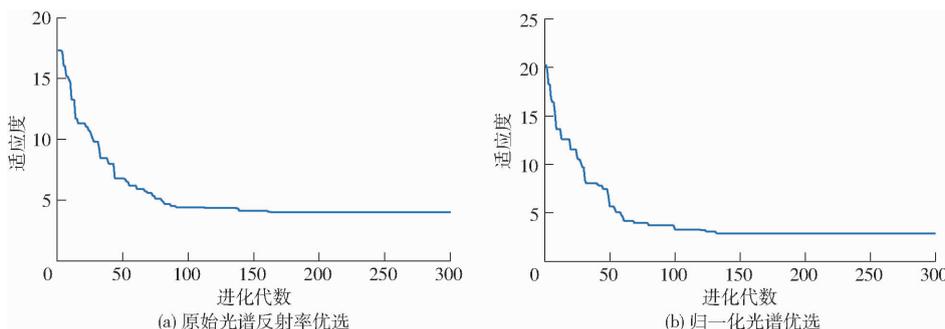


图 6 适应度随迭代次数变化曲线

Fig. 6 Changing curves of fitness value with iteration

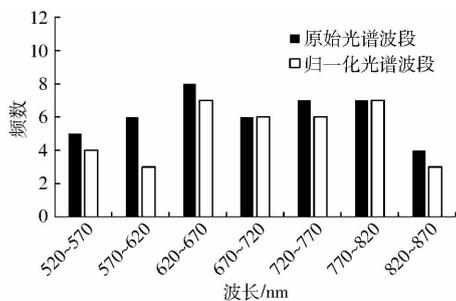


图 7 浊度 DBPSO-PLS 反演模型最佳波段分布

Fig. 7 Optimal bands distribution as turbidity DBPSO-PLS models

由图 6 可以看出, DBPSO 算法优选浊度反演的最佳波段的适应度值在迭代前期迅速减小, 在迭代后期快速收敛, 证明 DBPSO 算法在前期具有很好的全局搜索能力, 在后期具有较好的局部搜索能力, 并能在迭代后期快速收敛, 其中优选归一化光谱反

射率波段的最终适应度值要低于优选原始光谱反射率波段的适应度值。由图 7 可以看出, DBPSO 优选的浊度 PLS 反演模型最佳反演波段主要分布于对浊度反演敏感的红波段和近红波段^[30], DBPSO 确定的最佳归一化光谱反射率波段数为 36 个, 相比于确定的原始光谱反射率波段数有所降低; 由图 8 和表 2 可以看出, 浊度 OSR-DBPSO-PLS 和 NSR-DBPSO-PLS 模型均具有较好的拟合效果, 模型浊度反演值和观测值散点均靠近 1:1 线两侧, 其中浊度 OSR-DBPSO-PLS 模型 R^2 为 0.96, 模型综合误差为 11.66%; 浊度 NSR-DBPSO-PLS 模型 R^2 为 0.97, 模型综合误差为 8.78%。

3.3 DBPSO-PLS 和 PLS 模型对比分析

由图 8 可以看出, 浊度 OSR-DBPSO-PLS 模型精度 ($R^2 = 0.96$) 显著高于 OSR-PLS 模型 ($R^2 =$

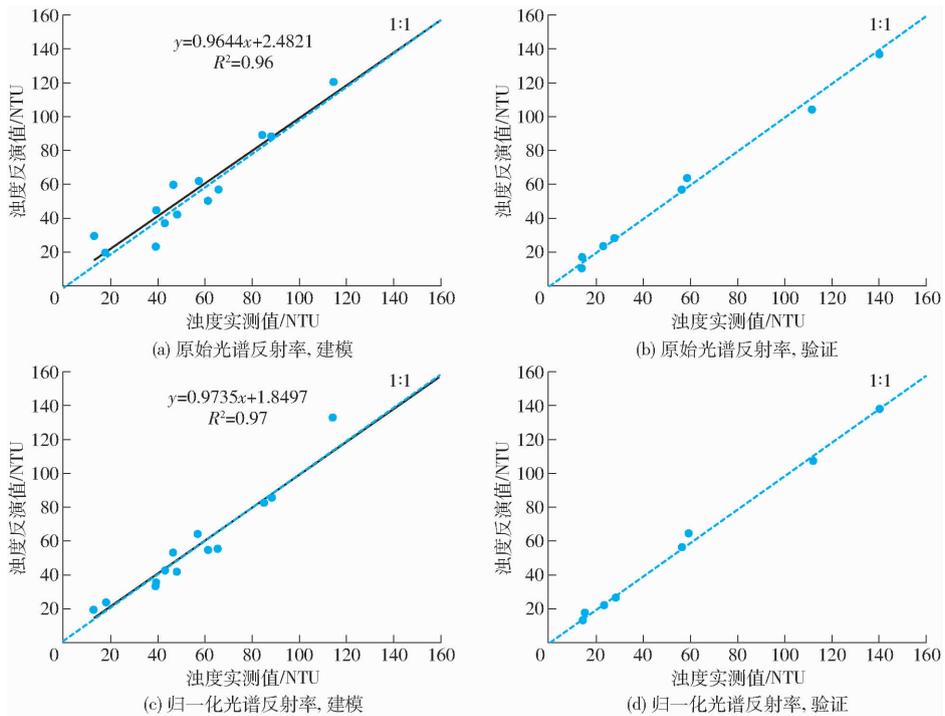


图 8 浊度 DBPSO - PLS 建模和验证散点图

Fig. 8 Modeling and validation scatter plots of DBPSO - PLS retrieving turbidity

0.50), 且浊度 PLS 建模波段数由 80 个减少至 44 个; 浊度 NSR - DBPSO - PLS 模型精度 ($R^2 = 0.97$) 同样显著高于 NSR - PLS 模型 ($R^2 = 0.91$), PLS 建模波段个数由 80 个减少至 36 个。利用 DBPSO 算法结合 PLS 模型优选反演浊度的最佳波段, 利用最佳波段和浊度建立的 PLS 模型 (DBPSO - PLS) 在减少 PLS 建模所需的波段数的同时, 可以提高模型反演精度。

3.4 微山湖水体浊度遥感反演

基于预处理的 HJ-1A HSI 影像, 选择具有最小模型综合误差的 NSR - DBPSO - PLS 模型作为浊度的最佳反演模型, 利用 NSR - DBPSO - PLS 模型反演得到微山湖水体浊度分布如图 9 所示。由图 9 可以看出, 2015 年 10 月微山湖水体浊度空间变异性显著, 微山湖东北方向湖区水体浊度显著高于微山湖西南方向湖区。

微山湖水生植物种类繁多, 其中光叶眼子菜和菹草是其中分布最多的两种植物, 除此之外还分布有较多的篦齿眼子菜和穗花狐尾藻等, 其中菹草在初春时快速繁殖生长, 在初夏之前迅速枯黄腐烂沉入湖底, 而光叶眼子菜、篦齿眼子菜和穗花狐尾藻初春时开始复苏, 在夏末生长得最为茂盛, 此后开始逐渐衰败; 在此次试验之前, 课题组曾于 2014 年 7 月 22—23 日、2014 年 8 月 29 日、2014 年 11 月 17 日、2015 年 4 月 6—9 日、2015 年 5 月 15—16 日和 2016 年 6 月 13 日在微山湖进行过 6 次实地采样, 采样足迹

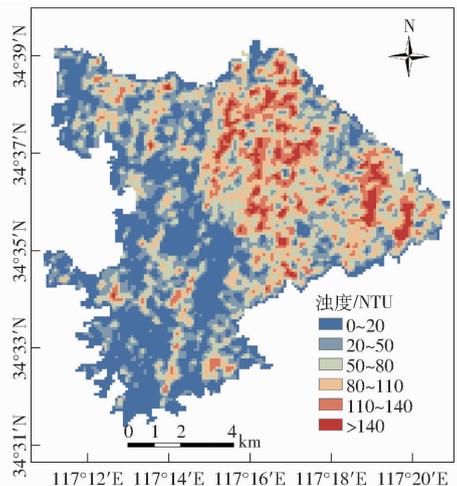


图 9 微山湖水体浊度分布 (2015 年 10 月)

Fig. 9 Distribution of turbidity in Weishan Lake in October, 2015

遍布整个湖区, 采样过程发现穗花狐尾藻主要分布于微山湖东北方向湖区, 而光叶眼子菜、篦齿眼子菜主要生长于西南方向湖区。本次试验时, 穗花狐尾藻已经逐渐凋亡腐烂生成大量残体 (图 10a), 而光叶眼子菜 (图 10b)、篦齿眼子菜 (图 10c) 依旧长势良好, 所以穗花狐尾藻主要生长的东北方向湖区水体浊度偏高, 此外该区域较为开阔, 风浪作用下的底泥再悬浮对浊度也有一定的贡献, 而光叶眼子菜和篦齿眼子菜生长的西南方向湖区由于沉水植物对水体起着过滤、吸附、消浪和抑制底泥上浮的作用^[31], 沉水植物的存在有利于水体浊度的降低^[32], 处于生长期的沉水植物生长区域水质整体较好, 浊度较低

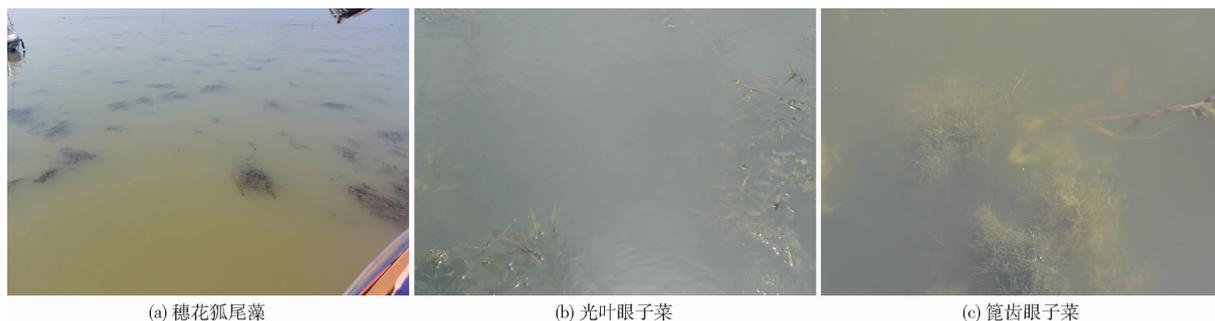


图 10 微山湖 3 种主要水草生长状态(2015 年 10 月)

Fig. 10 Growth of three major aquatic plants in Weishan Lake in October, 2015

(图 10b、10c), 但沉水植物后期死亡分解产生的残体会增加水体浊度(图 10a), NSR-DBPSO-PLS 模型浊度反演结果符合实际。

4 结论

(1) 基于归一化光谱反射率的浊度 PLS 模型精度(NSR-PLS, $R^2 = 0.91$) 显著高于基于原始光谱反射率的浊度 PLS 模型(OSR-PLS, $R^2 = 0.50$), 对波段进行归一化处理能提高水体浊度 PLS 模型反演精度。

(2) 利用 DBPSO 算法优选浊度 PLS 模型反演的特征波段, 在此基础上构建的基于原始光谱反射率的 PLS 模型精度(OSR-DBPSO-PLS, $R^2 = 0.96$) 和基于归一化光谱反射率的 PLS 模型精度

(NSR-DBPSO-PLS, $R^2 = 0.97$) 显著高于直接利用全谱波段构建的 OSR-PLS 和 NSR-PLS 模型, 建模波段数由 80 个分别减少至 44 个和 36 个, DBPSO-PLS 模型在减少 PLS 建模所需的波段数的同时可以提高浊度 PLS 模型反演精度。

(3) 基于归一化光谱反射率的浊度 NSR-DBPSO-PLS 反演模型具有最高精度($R^2 = 0.97$), 利用该模型反演微山湖水体浊度符合实际, 基于卫星高光谱数据的 DBPSO-PLS 模型在水质遥感领域具有应用潜力。

(4) 微山湖中处于生长期的沉水植物有利于降低水体浊度, 其生长区域水质整体较好, 浊度较低, 但沉水植物死亡分解产生的残体会增加水体浊度。

参 考 文 献

- MOSS B. Cogs in the endless machine: lakes, climate change and nutrient cycles: a review [J]. *Science of the Total Environment*, 2012, 434(18):130-142.
- 葛文杰, 赵春江. 农业物联网研究与应用现状及发展对策研究[J/OL]. *农业机械学报*, 2014, 45(7):222-230. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20140735&journal_id=jcsam. DOI:10.6041/j.issn.10001298.2014.07.035.
- GE Wenjie, ZHAO Chunjiang. State of the art and developing strategies of agricultural internet of things [J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2014, 45(7):222-230. (in Chinese)
- 钟珍梅, 黄毅斌, 李艳春, 等. 我国农业面源污染现状及草类植物在污染治理中的应用[J]. *草业科学*, 2017, 34(2):428-435.
- ZHONG Zhenmei, HUANG Yibin, LI Yanchun, et al. Current state of agricultural environmental pollution and herbaceous plants used in controlling pollution in China [J]. *Pratacultural Science*, 2017, 34(2):428-435. (in Chinese)
- PALMER S C J, KUTSER T, HUNTER P D. Remote sensing of inland waters: challenges, progress and future directions [J]. *Remote Sensing of Environment*, 2015, 157:1-8.
- CHEN Z, HU C, MULLER-KARGER F. Monitoring turbidity in Tampa Bay using MODIS/Aqua 250-m imagery [J]. *Remote Sensing of Environment*, 2007, 109(2):207-220.
- GÜTTLER F N, NICULESCU S, GOHIN F. Turbidity retrieval and monitoring of Danube Delta waters using multi-sensor optical remote sensing data: an integrated view from the delta plain lakes to the western-northwestern Black Sea coastal zone [J]. *Remote Sensing of Environment*, 2013, 132(6):86-101.
- 冯奇, 程学军, 沈欣, 等. 利用 Landsat-8 OLI 进行汉江下游水体浊度反演[J]. *武汉大学学报:信息科学版*, 2017, 42(5):643-647.
- FENG Qi, CHENG Xuejun, SHEN Xin, et al. Performance assessment of Landsat-8 OLI image for inland riverine turbidity estimation: a case study on Hanjiang River [J]. *Gemantics and Information Science of Wuhan University*, 2017, 42(5):643-647. (in Chinese)
- BUSTAMANTE J, PACIOS F, DÍAZ-DELGADO R, et al. Predictive models of turbidity and water depth in the Doñana marshes using Landsat TM and ETM+ images [J]. *Journal of Environmental Management*, 2009, 90(7):2219-2225.
- NECHAD B, RUDDICK K G, NEUKERMANS G. Calibration and validation of a generic multisensor algorithm for mapping of turbidity in coastal waters [C]// *Proceedings of SPIE on Remote Sensing of the Ocean, Sea Ice, and Large Water Regions*, 2009: 74730H-1-11.

- 10 GOHIN F. Annual cycles of chlorophyll-a, non-algal suspended particulate matter, and turbidity observed from space and in-situ in coastal waters [J]. *Ocean Science*, 2011, 7(5):705 - 732.
- 11 DOGLIOTTI A I, RUDDICK K G, NECHAD B, et al. A single algorithm to retrieve turbidity from remotely-sensed data in all coastal and estuarine waters [J]. *Remote Sensing of Environment*, 2015, 156:157 - 168.
- 12 SONG K, LU D, LI L, et al. Remote sensing of chlorophyll-a, concentration for drinking water source using genetic algorithms (GA)-partial least square (PLS) modeling [J]. *Ecological Informatics*, 2012, 10(7):25 - 36.
- 13 褚小立, 田高友, 袁洪福, 等. 小波变换结合多维偏最小二乘法用于近红外光谱定量分析[J]. *分析化学*, 2006, 34(增刊1):175 - 178.
CHU Xiaoli, TIAN Gaoyou, YUAN Hongfu, et al. Quantitative analysis of near infrared spectroscopy by combination of wavelet analysis and N-way partial least square [J]. *Chinese Journal of Analytical Chemistry*, 2006, 34(Supp. 1):175 - 178. (in Chinese)
- 14 ALI K A, ORTIZ J D. Multivariate approach for chlorophyll-a and suspended matter retrievals in case II type waters using hyperspectral data [J]. *Hydrological Sciences Journal*, 2016, 61(1):200 - 213.
- 15 刘忠华, 李云梅, 吕恒, 等. 基于偏最小二乘法的巢湖悬浮物浓度反演[J]. *湖泊科学*, 2011, 23(3):357 - 365.
LIU Zhonghua, LI Yunmei, LÜ Heng, et al. Inversion of suspended matter concentration in Lake Chaohu based on partial least-squares regression [J]. *Journal of Lake Sciences*, 2011, 23(3):357 - 365. (in Chinese)
- 16 杨燕明, 刘贞文, 陈本清, 等. 用偏最小二乘法反演二类水体的水质要素[J]. *遥感学报*, 2005, 9(2):123 - 130.
YANG Yanming, LIU Zhenwen, CHEN Benqing, et al. Retrieval of oceanic color constituents from case II water reflectance by partial squares regression [J]. *Journal of Remote of Sensing*, 2005, 9(2):123 - 130. (in Chinese)
- 17 徐京萍, 张柏, 宋开山, 等. 用偏最小二乘法提取石头口门水库水质信息[J]. *中国科学院大学学报*, 2007, 24(6):814 - 819.
XU Jingping, ZHANG Bai, SONG Kaishan, et al. Water information extraction of Reservoir Shitoukoumen based on partial least squares regression [J]. *Journal of the Graduate School of the Chinese Academy of Sciences*, 2007, 24(6):814 - 819. (in Chinese)
- 18 ROBERTSON A L. Using a partial least squares (PLS) method for estimating cyanobacterial pigments in eutrophic inland waters [C]//*Proceedings of SPIE on Remote Sensing and Modeling of Ecosystems for Sustainability*, 2009VI:745408 - 1 - 12.
- 19 RYAN K, ALI K. Application of a partial least-squares regression model to retrieve chlorophyll-a [J]. *Ocean Science Journal*, 2016, 51(2):209 - 221.
- 20 XU J, ZHANG B, LI F, et al. Retrieval of total suspended matters using field spectral data in Shitoukoumen Reservoir, Jilin Province, Northeast China [J]. *Chinese Geographical Science*, 2009, 19(1):77 - 82.
- 21 LU D M, SONG K S, LI L, et al. Training a GA - PLS model for chl-a concentration estimation over inland lake in northeast China [J]. *Procedia Environmental Sciences*, 2010, 2(1):842 - 851.
- 22 SONG K, LI L, LI S, et al. Hyperspectral retrieval of phycocyanin in potable water sources using genetic algorithm-partial least squares (GA - PLS) modeling [J]. *International Journal of Applied Earth Observation & Geoinformation*, 2012, 18(18):368 - 385.
- 23 SONG K, LI L, TEDESCO L P, et al. Hyperspectral determination of eutrophication for a water supply source via genetic algorithm-partial least squares (GA - PLS) modeling [J]. *Science of the Total Environment*, 2012, 426(2):220 - 232.
- 24 SONG K, LI L, TEDESCO L P, et al. Remote estimation of chlorophyll-a in turbid inland waters: three-band model versus GA - PLS model [J]. *Remote Sensing of Environment*, 2013, 136:342 - 357.
- 25 葛彦鹏. 基于偏最小二乘法的火电机组关键参数预测模型研究[D]. 北京:华北电力大学, 2013.
GE Yanpeng. Research on the key parameters prediction of thermal power units based on partial least squares methods [D]. Beijing:North China Electric Power University, 2013. (in Chinese)
- 26 KENNEDY J, EBERHART R C. A discrete binary version of the particle swarm algorithm [C]//*Proceedings of the World Multiconference on Systemics, Cybernetics and Informatics*, 1997:4104 - 4108.
- 27 沈林成, 霍霄华, 牛轶峰. 离散粒子群优化算法研究现状综述[J]. *系统工程与电子技术*, 2008, 30(10):1986 - 1990.
SHEN Lincheng, HUO Xiaohua, NIU Yifeng. Survey of discrete particle swarm optimization algorithm [J]. *Systems Engineering and Electronic*, 2008, 30(10):1986 - 1990. (in Chinese)
- 28 刘建华, 杨荣华, 孙水华. 离散二进制粒子群算法分析[J]. *南京大学学报:自然科学版*, 2011, 47(5):504 - 514.
LIU Jianhua, YANG Ronghua, SUN Shuihua. The analysis of binary particle swarm optimization [J]. *Journal of Nanjing University: Natural Sciences*, 2011, 47(5):504 - 514. (in Chinese)
- 29 巩彩兰, 尹球, 匡定波. 黄浦江水质指标与反射光谱特征的关系分析[J]. *遥感学报*, 2006, 10(6):910 - 916.
GONG Cailan, YIN Qiu, KUANG Dingbo. Correlations between water quality indexes and reflectance spectra of Huangpujiang River [J]. *Journal of Remote Sensing*, 2006, 10(6):910 - 916. (in Chinese)
- 30 BUSTAMANTE J, PACIOS F, DÍAZ-DELGADO R, et al. Predictive models of turbidity and water depth in the Doñana marshes using Landsat TM and ETM+ images[J]. *Journal of Environmental Management*, 2009, 90(7):2219 - 2225.
- 31 ZHANG Y, SHI K, LIU X, et al. Lake topography and wind waves determining seasonal-spatial dynamics of total suspended matter in turbid Lake Taihu, China: assessment using long-term high-resolution MERIS Data [J]. *PLOS ONE*, 2014, 9(5):e98055.
- 32 庞翠超, 吴时强, 赖锡军. 沉水植被降低水体浊度的机理研究[J]. *环境科学研究*, 2014, 7(5):498 - 504.
PANG Cuichao, WU Shiqiang, LAI Xijun. Water flow with submerged vegetation and its effect on water turbidity [J]. *Research of Environmental Sciences*, 2014, 7(5):498 - 504. (in Chinese)