

基于不同光谱主导因子的内陆湖泊水体叶绿素浓度三波段反演模型研究

施坤¹, 李云梅^{1*}, 刘忠华¹, 徐祎凡¹, 徐昕¹, 吴传庆², 朱利²

(1. 南京师范大学教育部虚拟地理环境教育部重点实验室, 南京 210046; 2. 环境保护部卫星环境应用中心, 北京 100029)

摘要: 利用分层聚类法, 对实测的遥感反射率进行分类, 结合固有光学特征和实测的水色要素, 确定每种类别遥感反射率光谱变化的主导影响因素。根据太湖、滇池、三峡库区及巢湖的地面遥感实验, 将其遥感反射率分为 3 类, 类型一的光学特性由总悬浮物和浮游藻类共同主导, 类型二的光学特性由浮游藻类主导, 类型三的光学特性由悬浮物主导。根据每类水体光学特征的差异, 分别建立叶绿素浓度的三波段反演模型, 反演的平均相对误差为 23.8%, 均方根误差为 8.5 mg/m³, 其反演精度要高于未经分类而直接建立的三波段模型。

关键词: 三波段模型; 浮游藻类; 总悬浮物; 叶绿素浓度; 光谱主导因子; 太湖

中图分类号: X87 文献标识码: A 文章编号: 0250-3301(2010)12-2873-09

Three Wavelengths Factor Model of Chlorophyll Concentrations Inversion Based on Different Spectral Dominating Factors for Inland Lake

SHI Kun¹, LI Yun-mei¹, LIU Zhong-hua¹, XU Yi-fan¹, XU Xin¹, WU Chuan-qing², ZHU Li²

(1. Key Laboratory of Virtual Geographic Environment, Ministry of Education, Nanjing Normal University, Nanjing 210046, China;

2. Satellite Environment Application Center, Ministry of Environmental Protection, Beijing 100029, China)

Abstract: A hierarchical cluster analysis is applied to the data set of R_{rs} spectra. Combining the inherent optical and measured water color characteristics, we determine factors of the variability for each class of R_{rs} spectra. The R_{rs} measured from Lake Taihu, Chaohu, Dianchi and Three Gorges Reservoir can be divided into three spectrally distinct classes: class 1, class 2 and class 3. Class 1 is associated with water optical property dominated by total suspended matter and phytoplankton. Class 2 is associated with water optical property dominated by phytoplankton. And Class 3 is associated with water optical property dominated by total suspended matter. The three wavelengths factor Chlorophyll a concentration inversion models were developed according to different class water optical characteristics, which means relative error and RMSE are 23.8% and 8.5 mg/m³, and have higher accuracy than the model developed without classification.

Key words: Three wavelengths factor model; phytoplankton; total suspended matter; chlorophyll concentration; spectra dominating factors; Lake Taihu

叶绿素浓度反演模型构建的关键是反演因子的确定, 也就是模型变量的确定。较好的反演因子能够尽可能多地反映叶绿素浓度的信息, 而较少反映其他物质的信息。确定叶绿素反演因子的方法主要有: 波段比值^[1~4]、反射峰位置^[5]和三波长因子法^[6~11, 4]等。三波长因子法^[8, 11]是目前反演叶绿素浓度较为有效的方法之一。

利用三波长因子法反演叶绿素浓度的主要模型包括: Dall' Olmo 等^[6]提出的利用 $[R_{rs}^{-1}(671) - R_{rs}^{-1}(710)] R_{rs}(740)$ 因子建立的估算模型 (Dall' Olmo 模型); Gitelson 等^[7]提出的含因子 $[R_{rs}^{-1}(675) - R_{rs}^{-1}(695)] R_{rs}(730)$ 的估算模型 (Gitelson 模型); 周冠华等^[9]利用三波长因子 $[R_{rs}^{-1}(666) - R_{rs}^{-1}(688)] R_{rs}(725)$ 建立的模型 (周冠华模型); 徐京萍等^[8]的 $[R_{rs}^{-1}(689.89) - R_{rs}^{-1}(692.77)] R_{rs}(799.18)$ 三波长

因子模型 (徐京萍模型); Sun 等^[11] $[R_{rs}^{-1}(661) - R_{rs}^{-1}(691)] R_{rs}(727)$ 三波长因子模型 (Sun 模型)。三波长因子法以生物光学模型为理论基础, 通过 3 个波长反射率的组合, 不但保证了光谱因子对浮游植物色素 (以 Chla 为主) 变化的敏感性, 而且还减少了水体中其他组分的干扰, 因此, 一定程度上能提高内陆混浊湖泊水体中叶绿素 a 浓度的反演精度。但是, 三波长光谱因子形式在不同水体或同一水体的不同季相具有差异性, 同时线性统计回归模型中的

收稿日期: 2010-01-06; 修订日期: 2010-02-26

基金项目: 国家自然科学基金项目 (40971215); 国家水体污染控制与治理科技重大专项 (2009ZX07527-006); 江苏省研究生培养创新工程项目 (CX10B_392Z); 南京师范大学优秀博士研究生学位论文培育项目 (2010BS0036)

作者简介: 施坤 (1985 ~), 男, 博士研究生, 主要研究方向为环境遥感, E-mail: ahuashikun@yahoo.com.cn

* 通讯联系人, E-mail: liyunmei@njnu.edu.cn

参数受样本变化的影响较大,这都限制了模型的通用性^[11].其差异产生的主要原因是由于不同水体或者不同时间的同一水体具有不同的光学特性.由于在不同水体或在不同时间同一水体中,叶绿素、总悬浮物、黄色物质的浓度差异较大,导致主导水体光学特性的因子不同,进而造成其固有光学特性的差异.遥感反射率的变化特征主要是受水体固有光学特性的控制^[12].水体的固有光学特性参数主要包括吸收和散射,在不同的水体,吸收和散射对遥感反射率大小的影响是不同的.在叶绿素浓度较高而总悬浮物浓度相对较低的水体,叶绿素的吸收可能是遥感反射率变化的主导因子;在总悬浮物浓度较高而叶绿素浓度相对较低的水体,总悬浮物的散射可能是遥感反射率变化的主导因子;而在总悬浮物浓度和叶绿素浓度相对较为均衡的水体,遥感反射率的变化可能是由两者共同作用的结果.遥感反射率光谱变化特征较为一致的水体,其光学特性的主导因子应该是一致的.

因此,本研究根据实测的太湖、滇池、巢湖以及三峡库区等内陆湖泊水体的遥感反射率,首先利用分层聚类法对遥感反射率进行分层分类,然后对每种类别的遥感反射率及其对应的固有光学特征进行分析,以确定每类水体光学特性的主导影响因子,进而确定每种类别水体叶绿素反演的三波段因子.根据每种类别的三波段因子,分别建立叶绿素浓度反演模型.

1 数据与方法

1.1 研究区及样点分布

2008 年 11 月、2009 年 6 月、8 月、9 月进行了太湖、巢湖、三峡库区、滇池野外实验.对太湖 72 个样点、巢湖 32 个样点、三峡 25 个样点、滇池 25 个样点进行了观测(图 1、2).测量的参数包括水体中总悬浮物浓度、有机悬浮物浓度、黄色物质的吸收系数、无机悬浮物浓度、叶绿素浓度、色素颗粒物的吸收系数(a_{ph})、散射系数、遥感反射率.其中,遥感反射率

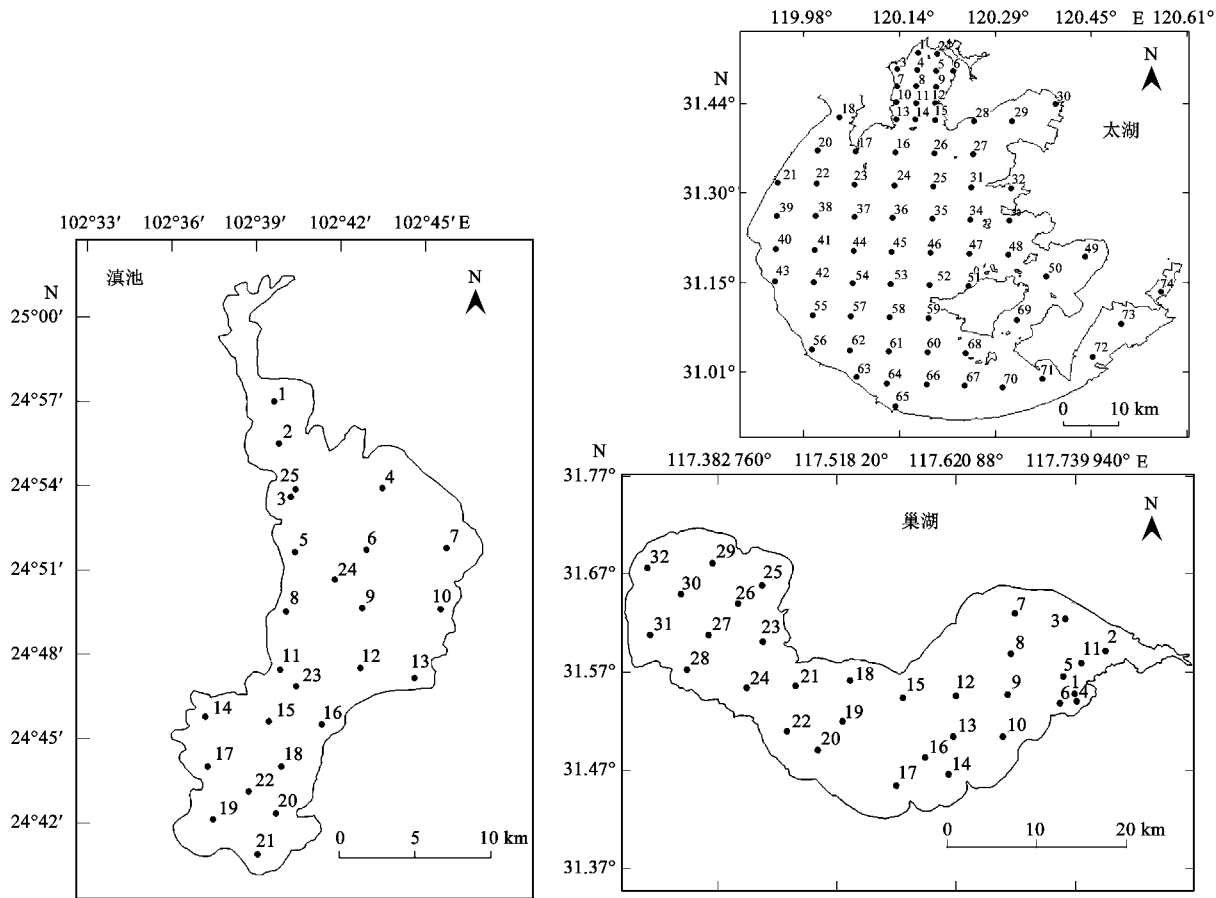


图 1 滇池、太湖、巢湖样点分布示意

Fig. 1 Sketch map shows the distribution of sample stations in Lake Dianchi, Taihu and Chaohu

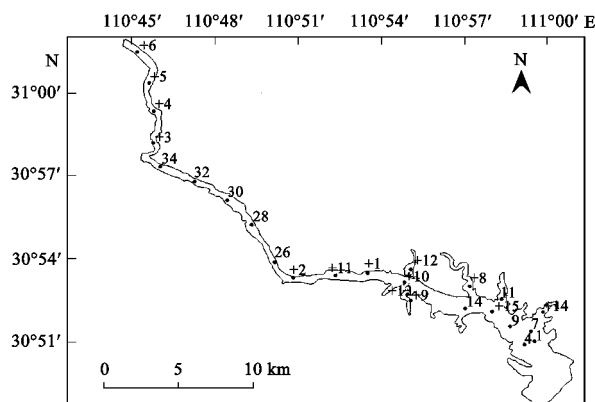


图2 三峡库区样点分布示意

Fig. 2 Sketch map shows the distribution of sample stations in Three Gorges Reservoir

和散射系数在野外直接观测,总悬浮物浓度、有机悬浮物浓度、黄色物质的吸收系数、无机悬浮物浓度、叶绿素浓度、色素颗粒物的吸收系数(a_{ph})、总悬浮物的吸收系数等参数通过室内分析获取。

1.2 参数的测定

运用 Wetlabs 公司的光谱吸收和衰减仪(AC-S)获得太湖水体颗粒物的散射系数,由 AC-S 直接获得的是水体吸收系数 $a(\lambda)$ 和衰减系数 $c(\lambda)$,为了得到更为精确的吸收系数,需对吸收系数作温度和散射纠正,纠正以及颗粒物散射系数的获取方法见文献[13]。

总悬浮物浓度、有机、无机悬浮物浓度采用常规的干燥、烘烧、称重法(GB 11901-89 标准)测定。先利用 $0.45 \mu\text{m}$ 的 GF/F 滤膜过滤一定体积的水样,即可得黄色物质水样,将黄色物质水样放入比色皿中,利用分光光度计测量黄色物质的吸光度,然后获得各个波长处黄色物质的吸收系数 [$a_{CDOM}(\lambda)$],并做散射校正,黄色物质吸收系数获取及散射校正的方法见文献[14]。叶绿素浓度的测量采用常规的化学分析方法,用 $0.45 \mu\text{m}$ 的 GF/F 滤膜过滤,90% 的热乙醇提取,然后利用分光光度计检测,叶绿素浓度获取的详细步骤见文献[15]。

色素颗粒物的吸收系数 [$a_{ph}(\lambda)$] 利用分光光度计测量,先测量总颗粒物的吸收系数,然后对总颗粒物进行漂白得到非色素颗粒物,再测定其吸收系数 [$a_{nap}(\lambda)$],两者之差即色素颗粒物的吸收系数,具体测量方法见文献[16]。

光谱测量采用美国 ASD 公司生产的 ASD FieldSpec Pro 便携式光谱辐射计,其测量的波段范围在 $350 \sim 1050 \text{ nm}$ 之间。提取遥感反射率必须要观

测的数据有:标准灰板(其反射率 < 0.3)、天空光、水体等的光谱辐亮度信息,每个对象都要采集 10 条以上的光谱数据,剔除异常光谱数据,剩余数据做均值处理。

为了避开水体对太阳的直射反射和船体等阴影对光场的破坏作用,采用一定的观测角度进行光谱测量,仪器观测平面与太阳入射平面的夹角 $90^\circ < \varphi_v < 145^\circ$,仪器与水面法线方向的夹角 $10^\circ < \theta_v < 30^\circ$ 。天空光在水面的反射是不可避免的,因此,在仪器面向水体进行测量后,将仪器在观测平面内向上旋转一个角度,使得观测方向的天顶角与 θ_v 相同,测量天空光的辐射亮度。遥感反射率提取的具体方法见文献[17]。

在获得各个样点的水体遥感反射率后,观察光谱特征的变化,剔除光谱特征变化异常的数据,最后保留 133 个样本数据用于数据的分析。

1.3 三波段叶绿素浓度反演方法

三波段模型最早被应用于植被遥感中,后来 Gitelson 通过改进,将该模型成功地应用于富营养化水体叶绿素浓度的反演。三波段模型的物理推导过程见文献[7]。该模型的概念模式可以表达为:

$$[R_{rs}^{-1}(\lambda_1) - R_{rs}^{-1}(\lambda_2)] R_{rs}(\lambda_3) \propto \text{Chla} \quad (1)$$

式中, $R_{rs}^{-1}(\lambda_1)$ 、 $R_{rs}^{-1}(\lambda_2)$ 分别是 λ_1 、 λ_2 的遥感反射率的倒数,表征了叶绿素 a 在这 2 个波段处的吸收强度; $R_{rs}(\lambda_3)$ 是波段 λ_3 处的遥感反射率, Chla 是叶绿素浓度。

λ_1 的位置应选择在叶绿素 a 浓度对浮游植物色素吸收最敏感的区域,通常选择在 $660 \sim 690 \text{ nm}$ 范围^[6],因为在此范围浮游植物具有相对较高的吸收。为了最大程度上排除黄色物质、非色素颗粒物的吸收和后向散射 [$b_b(\lambda)$] 的干扰,引入波长 λ_2 。 λ_2 的位置选择具有如下特点^[7]: ① $a_{ph}(\lambda_2) \ll a_{ph}(\lambda_1)$; ② $a_{CDOM}(\lambda_2) \approx a_{CDOM}(\lambda_1)$; ③ $a_{nap}(\lambda_2) \approx a_{nap}(\lambda_1)$; ④ $b_b(\lambda_2) \approx b_b(\lambda_1)$ 。由于纯水的吸收系数为常数,所以上式中的 $R_{rs}^{-1}(\lambda_1) - R_{rs}^{-1}(\lambda_2)$ 只受到 a_{ph} 、 b_b 和辅助参数的影响,为进一步消除 b_b 和辅助参数的影响,引入波长 λ_3 。 λ_3 位置具有如下特点: ① $a_{ph}(\lambda_3)$ 、 $a_{CDOM}(\lambda_3)$ 和 $a_{nap}(\lambda_3)$ 都近似为 0,水体的吸收 [$a_w(\lambda)$] 主要由纯水主导; ② $b_b(\lambda_3) \approx b_b(\lambda_2) \approx b_b(\lambda_1)$; ③ $a_w(\lambda_3) \gg b_b(\lambda_3)$ 。

综上所述,通过 3 个波长的组合可以极大地减少水体其他组分对反演叶绿素 a 的干扰,增强叶绿素 a 信息在遥感反射率光谱上的反映,3 个波段的具体位置通常采用迭代法确定。

1.4 遥感反射率的分类

采用分层聚类的分析方法对实测的光谱遥感反射率进行分类. 分层聚类法是非监督分类方法的一种,和快速聚类不同的是在聚类分析之间,不需要事先确定类别数. 分层聚类根 6 类的过程又可以分为分解法和凝聚法. 本研究采用凝聚法对遥感反射率进行聚类,具体的步骤见文献[18],采用不同样本之间的距离作为衡量样本之间相似性的指标. 采用 Ward 方法进行类别间距离的计算,基于方差分析的思想,如果分类结果正确,则同类样本之间的离差平方和应当较小,不同类的样本之间的离差平方和应当较大. 本研究利用 Z^2 检验统计量来确定分类数.

$$Z^2 = 1 - \frac{P_c}{T}$$

(2)

式中, P_c 为分类数为 G 类时的总类内离差平方和; T 为所有变量的总离差平方和. Z^2 越大,说明分为 G 类时每个类内的离差平方和越小,也就是分为 G 类是合适的. Z^2 的值在 $0 \sim 1$ 之间,当 n 个样本分成 n 类时 $Z^2 = 1$;当分成 1 类时 $Z^2 = 0$. 当从 $G + 1$ 类合并成 G 类时,如果 Z^2 值的变化在全局的各次变化较剧烈,那么分成 $G + 1$ 类较适合.

2 结果与讨论

2.1 遥感反射率的分类

图 3 显示了利用层次聚类遥感反射率进行分类的类别数与 Z^2 之间的关系. 由图 3 可知,当遥感反射率的类别由 3 类聚为 1 类时, Z^2 从 0.8 降到 0, Z^2 变化较为剧烈. 因此本研究认为把遥感反射率分成 3 类较为合适. 分别命名为类型一(图 4)、类型二(图 5)和类型三(图 6). 表 1 显示了每类遥感反射率对应的生物-光学特性. 图 7 显示了 3 种类别遥感反射率的平均值. 类型二具有较低的遥感反射率(图 5、7),由藻类色素引起的 2 处反射峰较为明显,由藻类叶绿素和胡萝卜素的弱吸收和细胞的散射作用形成的反射峰出现在 569 nm 左右;由叶绿素 a 的

表 1 不同类型水体的生物光学量

Table 1 Statistical bio-optical parameters in different class waters

光学及生物量	平均值		
	类型一	类型二	类型三
$a_{ph}(675)/m^{-1}$	0.364 02	0.984 672	0.223 64
TSM/ $mg \cdot L^{-1}$	50.013 84	33.683 33	99.934 72
OSM/ $mg \cdot L^{-1}$	10.787 11	16.469 38	14.245 83
ISM/ $mg \cdot L^{-1}$	39.226 73	17.213 95	85.688 89
ISM/TSM	0.757 369	0.505 409	0.821 167
Chla/ $mg \cdot L^{-1}$	18.138 3	48.617 31	15.082 62
$b_p(532)/m^{-1}$	40.582 3	21.088 23	64.286 9
$a_{CDOM}(440)/m^{-1}$	0.847 387	0.549 675	0.448 989

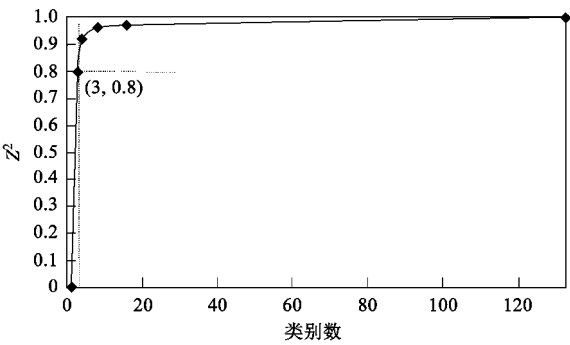


图 3 Z^2 与分类数之间的关系

Fig. 3 Relationship between Z^2 and class numbers

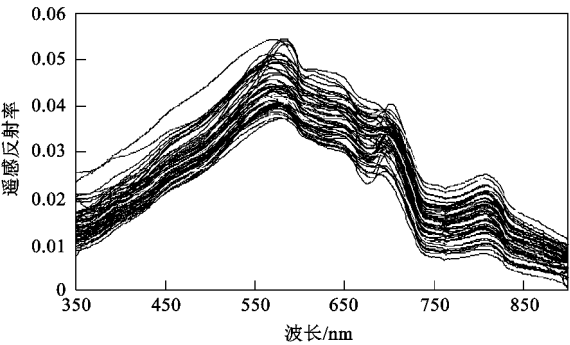


图 4 类型一反射率光谱曲线

Fig. 4 Spectrum of the remote sensing reflectance in class 1 water

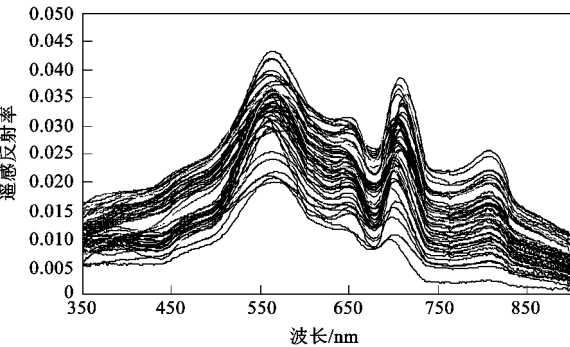


图 5 类型二遥感反射率光谱变化

Fig. 5 Spectrum of the remote sensing reflectance in class 2 water

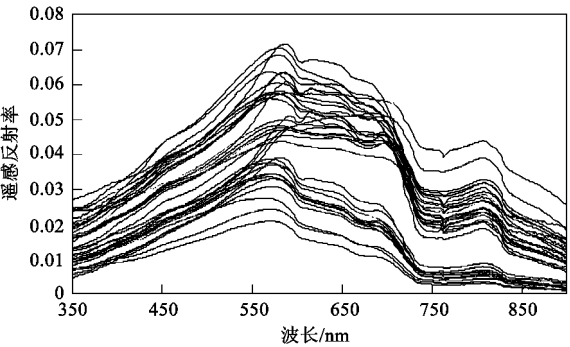


图 6 类型三遥感反射率光谱变化

Fig. 6 Spectrum of the remote sensing reflectance in class 3 water

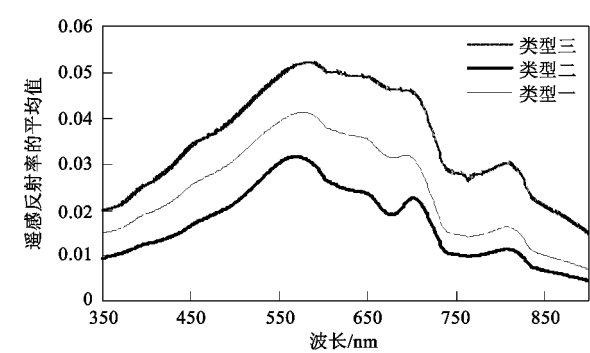


图 7 三类遥感反射率的平均值
Fig. 7 Mean spectrum of the remote sensing reflectance in 3 class waters

荧光作用而产生的荧光峰出现在 700 nm 左右;由藻类叶绿素在红光波段处的强吸收而形成的吸收峰也较为明显,出现在 685 nm 左右;由表 1 可知类型二的色素颗粒物在 675 nm 处的吸收系数、叶绿素浓度、黄色物质的吸收以及有机悬浮物占总悬浮物的比例的平均值分别为 0.98 m^{-1} 、 48.36 mg/m^3 、 0.84 m^{-1} 、 0.5 ,要大于类型一和类型三的平均值.类型二水体的总悬浮物浓度、无机悬浮物浓度、532 nm 处的散射系数的平均值分别为 33.6 mg/L 、 17.2 mg/L 、 21 m^{-1} ,要小于类型一、类型三的平均值.类型二的样点共计 52 个,占总样点的 39% 左右,主要分布在滇池、太湖梅梁湾、巢湖的西半湖、三峡支流等叶绿素浓度较高而总悬浮物浓度相对较低的水体.可见,类型二水体的光学特性主要是由浮游藻类主导的.类型三具有较高的遥感反射率(图 6、7),由藻类色素引起的 2 处反射峰不明显,由藻类叶绿素和胡萝卜素弱吸收和细胞的散射作用形成的反射峰出现在 583 nm 左右;由于叶绿素 a 的荧光作用而产生的荧光峰出现在 697 nm 左右;由藻类叶绿素在红光波段处的强吸收而形成的吸收峰出现在 683 nm 左右;由表 1 可知类型三的色素颗粒物在 675 nm 的吸收系数、叶绿素浓度、黄色物质的吸收以及有机悬浮物占总悬浮物的比例的平均值分别为 0.22 m^{-1} 、 15 mg/m^3 、 0.44 m^{-1} 、 0.2 ,要小于类型一和类型二的平均值.而类型三水体的总悬浮物浓度、无机悬浮物浓度、532 nm 处的散射系数的平均值分别为 99.9 mg/L 、 85.68 mg/L 、 64 m^{-1} ,要大于类型一、类型二的平均值.类型三的样点共计 33 个,占总样点数的 24% 左右,主要分布在太湖湖心区域、三峡库区等叶绿素浓度较低而总悬浮物浓度相对较高的水体.可知,类型三水体的光学特性主要是由悬浮物特别是

无机悬浮物主导的.类型一的遥感反射率的大小介于类型二与类型三之间(图 4、7),由藻类色素引起的 2 处反射峰不如类型二的明显,但比类型三的明显.由藻类叶绿素和胡萝卜素弱吸收和细胞的散射作用形成的反射峰出现在 577 nm 左右;由叶绿素 a 的荧光作用而产生的荧光峰出现在 697 nm 左右;由藻类叶绿素红光波段处的强吸收而形成的吸收峰出现在 683 nm 左右;由表 1 可知类型一的色素颗粒物在 675 nm 处的吸收系数、叶绿素浓度、黄色物质的吸收以及有机悬浮物占总悬浮物的比例的平均值分别为 0.36 m^{-1} 、 18 mg/m^3 、 0.54 m^{-1} 、 0.25 ,介于类型三和类型二之间.类型一水体的总悬浮物浓度、无机悬浮物浓度、532 nm 处的散射系数的平均值分别为 50 mg/L 、 39 mg/L 、 40 m^{-1} ,也介于类型三和类型二之间.类型三共计 48 个样点,占总样点的 37% 左右.可见,类型一水体的光学特性主要是由悬浮物和浮游藻类共同主导的.

不同类型的遥感反射率对应的固有光学-生物特性具有较为明显的差异,类型二遥感反射率往往对应于较高的叶绿素浓度和较低的总悬浮物浓度,色素颗粒物的吸收作用较强,而悬浮物的散射作用相对较弱,往往呈现出较为明显的峰值和谷值.类型三水体的遥感反射率往往对应于较高的总悬浮物浓度和较低的叶绿素浓度,色素颗粒物吸收作用相对较弱,而总悬浮物的散射相对较强.类型一遥感反射率对应于较为适中的叶绿素和总悬浮物浓度,色素颗粒物的吸收作用和总悬浮物的散射作用相当.由于不同类型遥感反射率对应的叶绿素浓度不同,从而造成由于色素的荧光和吸收作用而产生的反射峰和谷值的位置不同,比如在叶绿素浓度较高的类型二水体,由叶绿素 a 的荧光作用而产生的荧光峰在 700 nm 处;而类型三的则出现在 697 nm 处,这与 Gitelson 等^[5]的研究结果一致.

2.2 三波段叶绿素浓度反演模型

2.2.1 三波段因子的选择

波长 λ_1 、 λ_2 和 λ_3 初始位置的选择均具有各自的特点,而 λ_1 和 λ_3 相比于 λ_2 更容易确定,所以本研究先确定 λ_1 和 λ_3 的初始位置,并通过迭代优化确定 λ_2 的位置.

λ_1 初始位置应选择叶绿素 a 浓度对浮游植物色素吸收最敏感的区域.对各波长吸收系数 $a_{ph}(\lambda)$ 与叶绿素 a 浓度进行皮尔逊相关性分析(图 8),比较各波长的 Pearson 相关系数发现,类型一水体最大值为 0.95,出现在 672 nm 处,类型二水体最大值为

0.97, 出现在 675 nm 处, 类型三水体最大值为 0.68, 出现在 685 nm 处, 未分类的水体最大值为 0.95, 出现在 674 nm 处(显著性水平为 0.01), 表明光学特性主导因子不同的水体, 叶绿素 a 对浮游植物色素的吸收较为敏感的波段是不一样的. 因此, 对不同类型的水体, 本研究选择不同的 λ_1 .

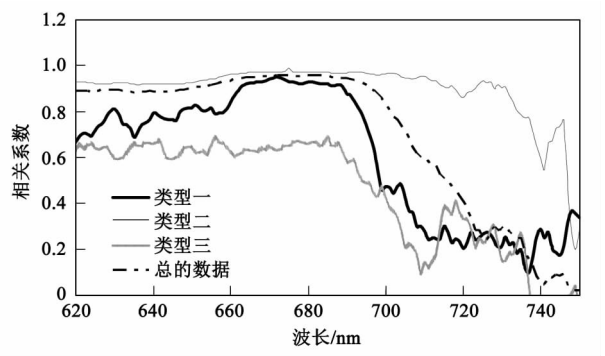


图 8 叶绿素浓度与吸收系数的相关系数
Fig. 8 Pearson correlation coefficient between Chlorophyll a concentration and $a_{ph}(\lambda)$

λ_3 的引入是为消除辅助参数与总后向散射系数 b_b 的影响. 在此波长处, 水体组分(除纯水外)的吸收系数都需接近于 0, 即 $a(\lambda_3) \approx a_w(\lambda_3)$, 且 $a(\lambda_3) \gg b_b(\lambda_3)$; 总后向散射系数 b_b 随波长的变化不大, 可视为常数. 研究表明近红外 730 ~ 800 nm 范围符合这些要求^[6], 而在此范围内 750 nm 附近的纯水吸收系数较大, 所以所有类型水体的 λ_3 初始位置都选择 750 nm 处.

通过对 Pearson 相关系数 R 和线性回归的均方根误差 RMSE 的分析进行波长迭代优化. 首先, 设定 λ_1 初始波长分别为 672 nm(类型一水体)、675 nm(类型二水体)、685 nm(类型三水体)、674 nm(未分类的水体), 设定所有类型水体的 λ_3 初始波长为 750 nm. 在 400 ~ 800 nm 范围对 λ_2 进行优化. 由表 2 可知, 类型一水体 R 的最大值为 0.91, RMSE 的最小值为 3.09 mg/m³, 同时出现在 697 nm 处. 接着, 取 $\lambda_1 = 672$ nm 和 $\lambda_2 = 697$ nm 对 λ_3 进行优化, 同样选择 R 值最大和 RMSE 值最小对应的波长作为 λ_3 此次优化的最佳位置. 按照同样的方法依次对 3 个波长进行迭代优化, 经过 4 次优化后, 三波长的位置已不再发生变化, 说明此时的位置为最佳. λ_1 、 λ_2 和 λ_3 的最佳位置分别为 678、697 和 717 nm, 对应的 R 和 RMSE 分别为 0.97 和 3.07 mg/m³(表 2). 按照同样的方法对其他类型水体做优化迭代运算, 确定每种类型水体的 3 个波段的位置(表 3 ~ 5).

表 2 类型一水体的三波段因子选择的迭代过程
Table 2 Iteration processes of three wavelengths factors choosing for class 1 water

优化次数	λ_1	λ_2	λ_3	R		RMSE/mg·m ⁻³	
				最大值	对应波长	最小值	对应波长
1	672	—	750	0.91	697	3.09	697
2	672	697	—	0.95	717	3.09	717
3	—	697	717	0.97	678	3.07	678
4	672	—	717	0.97	697	3.07	697

表 3 类型二水体的三波段因子选择的迭代过程
Table 3 Iteration processes of three wavelengths factors choosing for class 2 water

优化次数	λ_1	λ_2	λ_3	R		RMSE/mg·m ⁻³	
				最大值	对应波长	最小值	对应波长
1	675	—	750	0.91	700	11.2	700
2	675	700	—	0.93	737	11.19	737
3	—	700	737	0.94	678	11.18	678
4	675	—	737	0.94	700	11.18	700

表 4 类型三水体的三波段因子选择的迭代过程
Table 4 Iteration processes of three wavelengths factors choosing for class 3 water

优化次数	λ_1	λ_2	λ_3	R		RMSE/mg·m ⁻³	
				最大值	对应波长	最小值	对应波长
1	685	—	750	0.95	697	2.2	697
2	685	697	—	0.97	818	2.16	818
3	—	697	818	0.97	683	2.16	683
4	685	—	818	0.97	697	2.16	697

表 5 未分类水体的三波段因子选择的迭代过程
Table 5 Iteration processes of three wavelengths factors choosing for unclassified water

优化次数	λ_1	λ_2	λ_3	R		RMSE/mg·m ⁻³	
				最大值	对应波长	最小值	对应波长
1	674	—	750	0.91	711	11.3	711
2	674	711	—	0.92	749	11.02	749
3	—	711	749	0.92	663	11.02	663
4	674	—	749	0.92	711	11.02	711

类型一、二、三以及未分类水体的三波段因子分别为(表 6): $[R_{rs}^{-1}(678) - R_{rs}^{-1}(697)]R_{rs}(717)$ 、 $[R_{rs}^{-1}(678) - R_{rs}^{-1}(700)]R_{rs}(737)$ 、 $[R_{rs}^{-1}(683) - R_{rs}^{-1}(697)]R_{rs}(818)$ 、 $[R_{rs}^{-1}(663) - R_{rs}^{-1}(711)]R_{rs}(749)$, 观察这些三波段因子可以发现: λ_1 基本上位于由藻类叶绿素在红光波段处的强吸收而形成的吸收峰处, 此波段浮游藻类的吸收较强; λ_2 位于由叶绿素 a 的荧光而产生的荧光峰处, 该波段色素颗粒物的吸收相对 λ_1 处的要小的多; λ_3 位于散射较小的波段处. 由于藻类叶绿素在红光波段处的强吸收而形成的吸收峰随着叶绿素浓度的增加而向短波长

向移动,因此 λ_1 随着叶绿素浓度的升高而往短波处移动;由叶绿素 a 的荧光而产生的荧光峰,随着叶绿素浓度的增加而往长波方向移动,因此, λ_2 随着叶绿素浓度的升高往长波方向移动; λ_3 与水体中总悬浮物的浓度有关,随着总悬浮物浓度的增加往长波方向移动.

表 6 3 种类型水体的三波段因子及对应的 R 和 RMSE
Table 6 Correlation coefficient and RMSE between three wavelengths factors and Chlorophyll a concentrations in three class waters

水体类型	三波段因子	R	RMSE / $\text{mg}\cdot\text{m}^{-3}$
类型一	$[R_{rs}^{-1}(678) - R_{rs}^{-1}(697)]R_{rs}(717)$	0.97	3.07
类型二	$[R_{rs}^{-1}(678) - R_{rs}^{-1}(700)]R_{rs}(737)$	0.94	11.18
类型三	$[R_{rs}^{-1}(683) - R_{rs}^{-1}(697)]R_{rs}(818)$	0.97	2.16
未分类	$[R_{rs}^{-1}(663) - R_{rs}^{-1}(711)]R_{rs}(749)$	0.92	11.02

2.2.2 反演模型的构建

考虑到线性模型的稳定性较高,因此,根据每种类型水体的遥感反射率分别建立了不同类型的叶绿素浓度反演的线性模型.图 9~12 分别显示了类型一、二、三以及未分类水体叶绿素浓度与三波段因子之间的关系,每种类别叶绿素浓度的反演模型见式(3)~(6).

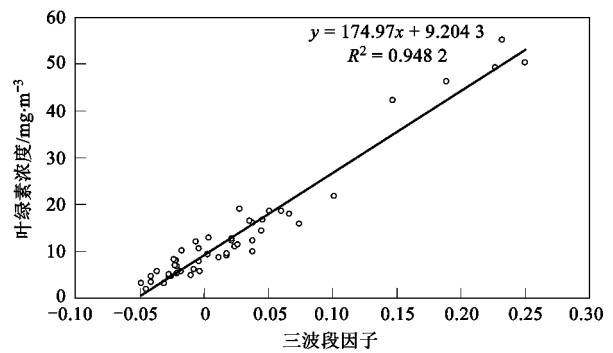


图 9 类型一三波段因子与叶绿素浓度之间的关系
Fig.9 Relationship between three wavelengths factor and Chlorophyll a concentrations for class 1 water

类型一模型:

$$\begin{aligned} \text{Chla} = & 174.97 [R_{rs}^{-1}(678) - R_{rs}^{-1}(697)] \\ & R_{rs}^{-1}(717) + 9.2 \end{aligned} \quad (3)$$

$(R^2 = 0.94, N = 48, p < 0.001)$

类型二模型:

$$\begin{aligned} \text{Chla} = & 148.92 [R_{rs}^{-1}(678) - R_{rs}^{-1}(700)] \\ & R_{rs}^{-1}(737) + 15.7 \end{aligned} \quad (4)$$

$(R^2 = 0.90, N = 52, p < 0.001)$

类型三模型:

$$\text{Chla} = 321.9 [R_{rs}^{-1}(683) - R_{rs}^{-1}(697)]$$

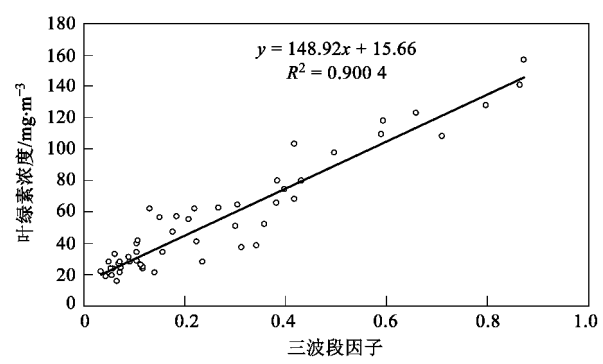


图 10 类型二三波段因子与叶绿素浓度之间的关系
Fig.10 Relationship between three wavelengths factor and Chlorophyll a concentrations for class 2 water

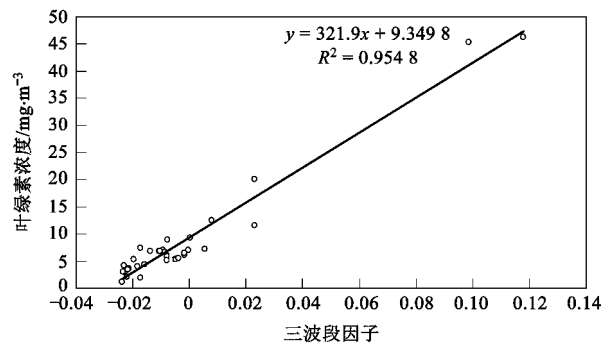


图 11 类型三三波段因子与叶绿素浓度关系
Fig.11 Relationship between three wavelengths factor and Chlorophyll a concentrations for class 3 water

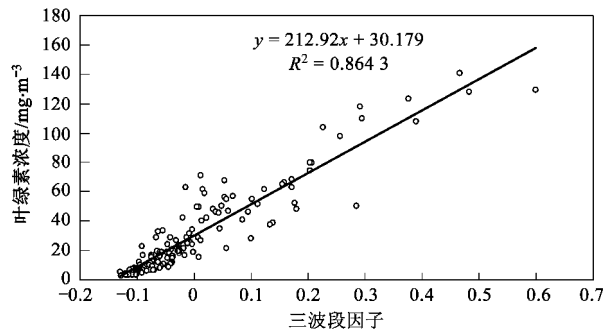


图 12 未分类三波段因子与叶绿素浓度之间的关系
Fig.12 Relationship between three wavelengths factor and Chlorophyll a concentrations for unclassified water

$$\begin{aligned} & R_{rs}^{-1}(818) + 9.3 \end{aligned} \quad (5)$$

$(R^2 = 0.95, N = 33, p < 0.001)$

未分类数据的模型:

$$\begin{aligned} \text{Chla} = & 212.92 [R_{rs}^{-1}(663) - R_{rs}^{-1}(711)] \\ & R_{rs}^{-1}(749) + 30.2 \end{aligned} \quad (6)$$

$(R^2 = 0.86, N = 133, p < 0.001)$

表 7 显示了每种类别三波段因子反演叶绿素浓度的相对误差和均方根误差,类型一反演的平均相对误差为 23%,RMSE 为 3.07 mg/m³;类型二反演的平均相对误差为 19.9%,RMSE 为 11.18 mg/m³;类型三反演的平均相对误差为 28.2%,RMSE 为 2.16 mg/m³;未分类的反演平均误差为 36.5%,RMSE 为 11.02 mg/m³,图 13、14 分别显示了分类以及未分类水体叶绿素浓度的反演值与实测值之间的关系.图 15、16 显示了分类后以及未分类水体反演的相对误差的误差分布情况.可知,通过分类后建立的三波段叶绿素浓度的反演模型的相对误差为 0~0.2 之间的占整个样本数的 50% 以上,相对误差在 0~0.3 之间的占整个样本数的 70% 以上,相对误差基本上集中在 30% 以内;未分类的三波段叶绿素浓度反演模型相对误差在 0~0.2 之间的占整个样本的 40% 左右,相对误差在 0~0.5 之间的占整个样本的 70% 左右.通过分类后的三波段叶绿素浓度反演模型的相对误差的平均值为 0.238,RMSE 为 8.5;未分类的三波段叶绿素浓度反演模型的相对误差的平均值为 0.365,RMSE 为 11.2.利用分类后建立的三波段模型能够很好地提高叶绿素浓度反演的精度.

表 7 不同类型水体三波段叶绿素反演模型的精度分析
Table 7 Accuracy analysis of three wavelengths factor Chlorophyll a concentrations inversion model for each class

水体类型	最大相对 误差/%	最小相对 误差/%	MAPE/%	RMSE /mg·m ⁻³
类型一模型	80	1	23	3.07
类型二模型	80	0.2	19.9	11.18
类型三模型	90	0.3	28.2	2.16
未分类模型	175	0.4	36.5	11.02

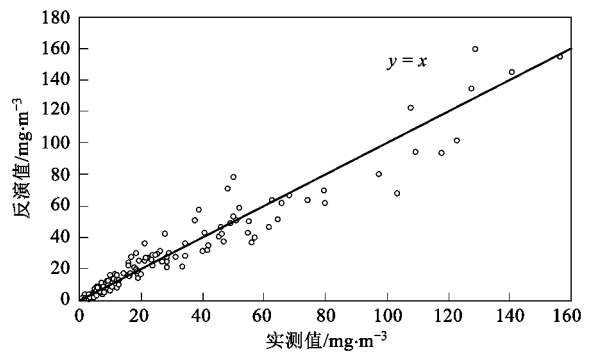


图 13 分类三波段叶绿素浓度反演模型的反演值与实测值的关系
Fig. 13 Relationship between measured and inversion values by classified three wavelengths factor model

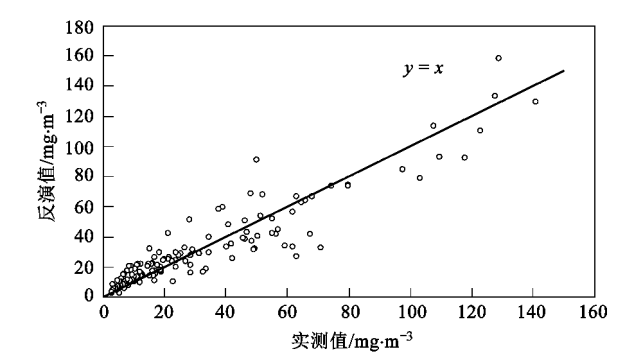


图 14 未分类三波段叶绿素浓度反演模型反演值和实测值的关系
Fig. 14 Relationship between measured and inversion values by unclassified three wavelengths factor model

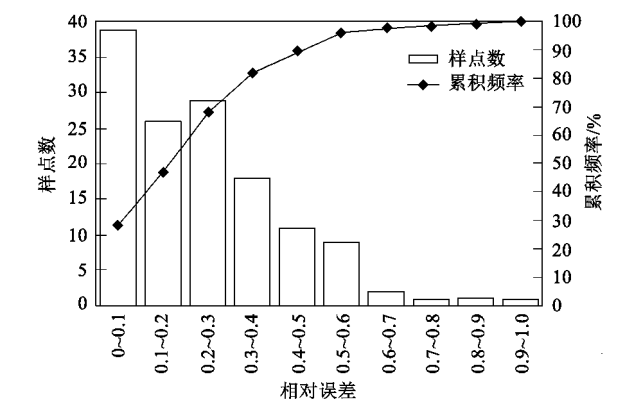


图 15 分类的三波段叶绿素反演模型的误差分布图
Fig. 15 Error distribution of classified three wavelengths factor Chlorophyll a concentrations inversion model

利用相对误差和 RMSE 对以往研究者提出的三波长估算模型在 3 类水体中的应用精度进行分析,三波长估算模型主要包括: Dall’ Olmo 模型^[6]、Gitelson 模型^[7]、周冠华模型^[9]、徐京萍模型^[8]、Sun 模型^[11]等.这些模型并非利用各学者获得的原始参数,而是经过本研究的建模数据重新拟合得到的模型参数,在此不作一一列举.然后,利用验证样本计算各模型的预测误差.由表 8 可知,诸学者提出的三波长因子模型在不同类型水体中的应用精度不同.有的在类型一水体中预测精度较高,如 Sun 模型^[11],其预测平均相对误差为 0.443,均低于其他模型;有的在类型二水体中预测精度较高,比如 Dall’ Olmo 模型^[6],平均相对误差为 0.3,均小于其他水体,其研究的水体的光学特性可能接近类型二;在类型三的水体中,徐京萍模型^[8]应用精度最高,其研究的水体的光学特性可能接近类型三.然而这些模型预测的整体精度都低于本研究所提出的 3 种类型的三波段叶绿素浓度的反演模型.

表 8 其他模型在三类水体中的预测精度比较

Table 8 Comparison of the predictive error from different models for each class water

水体类型	误差	Dall’Olmo 模型	Gitelson 模型	周冠华模型	徐京萍模型	Sun 模型
类型一	MAPE/%	51.7	45.2	87	86.8	44.3
	RMSE/mg·m ⁻³	8.62	8.07	9.91	9.61	8.32
类型二	MAPE/%	30.4	33.5	47.6	34.6	45.5
	RMSE/mg·m ⁻³	14.44	15.24	23.32	14.09	21.47
类型三	MAPE/%	74.5	53.6	81.5	50.2	83.3
	RMSE/mg·m ⁻³	5.03	3.20	7.01	2.69	7.06

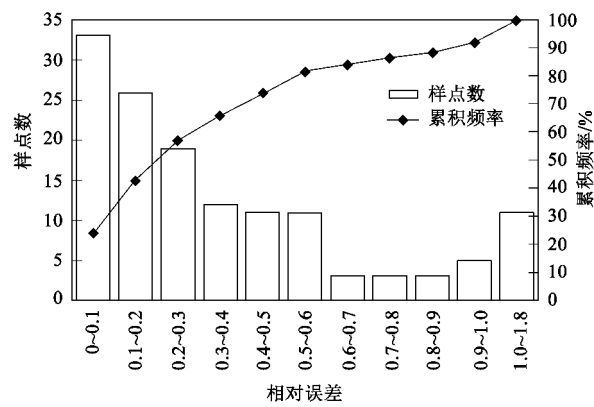


图 16 未分类的三波段叶绿素反演模型的误差分布

Fig. 16 Error distribution of unclassified three wavelengths factor

Chlorophyll a concentrations inversion model

3 结论

(1)利用层次分类法,对实测的太湖、滇池、巢湖、三峡库区等水体的遥感反射率进行分类,并分别针对不同类型水体建立叶绿素浓度的三波段遥感反演模型.研究表明,上述研究区水体可以分成 3 种类型,类型一的光学特性由浮游藻类和总悬浮物共同主导;类型二的光学特性由浮游藻类主导;类型三的光学特性则由总悬浮物主导.

(2)3 种类型水体的三波长因子不同,分别为: $[R_{rs}^{-1}(678) - R_{rs}^{-1}(697)]R_{rs}(717)$ 、 $[R_{rs}^{-1}(678) - R_{rs}^{-1}(700)]R_{rs}(737)$ 、 $[R_{rs}^{-1}(683) - R_{rs}^{-1}(697)]R_{rs}(818)$,与文献中其他三波段模型如 Dall’Olmo 模型、Gitelson 模型、周冠华模型、徐京萍模型等相比,不仅波长因子不同,而且精度也有所不同.结果表明,通过分类再建立模型的方法,能够提高叶绿素浓度反演精度.

参考文献:

[1] Gons H J. Optical teledetection of chlorophyll a in turbid inland waters [J]. Environ Sci Technol, 1999, **33**(7): 1127-1132.

[2] Jiao H B, Zha Y, Gao J, *et al.* Estimation of chlorophyll-a concentration in Lake Tai, China using in situ hyperspectral data [J]. Int J Remote Sens, 2006, **27**(19): 4267-4276.

[3] 段洪涛,张柏,宋开山,等. 查干湖叶绿素 a 浓度高光谱定量

模型研究[J]. 环境科学,2006,**27**(3):503-507.

[4] Zhang Y L, Liu M L, Qin B Q, *et al.* Modeling remote-sensing reflectance and retrieving chlorophyll-a concentration in extremely turbid case-2 waters(Lake Taihu, China) [J]. IEEE Trans Geosci Remote Sens,2009,**47**(7):1937-1948.

[5] Gitelson A A. The peak near 700 nm on radianc spectra of algae and water; relationships of its magnitude and position with chlorophyll concentration [J]. Int J Remote Sens, 1992, **13**: 3367-3373.

[6] Dall’Olmo G, Gitelson A. Effect of bio-optical parameter variability on the remote estimation of chlorophyll a concentration in turbid productive waters; experimental results [J]. Appl Opt, 2005, **44**(3): 412-422.

[7] Gitelson A A, Schalles J F, Hladik C M. Remote chlorophyll-retrieval in turbid, productive estuaries; Cheapeake Bay case study[J]. Remote Sens Environ, 2007, **109**: 464-472.

[8] 徐京萍,张柏,宋开山,等. 近红外波段二类水体悬浮物生物光学反演模型研究[J]. 光谱学与光谱分析, 2008, **28**(10): 2273-2277.

[9] 周冠华,柳钦火,马荣华,等. 基于半分析模型的波段最优化组合反演混浊太湖水体叶绿素 a [J]. 湖泊科学,2008,**20**(2):153-159.

[10] 杨煜,李云梅,王桥,等. 富营养化的太湖水体叶绿素 a 浓度模型反演[J]. 地球信息科学学报,2009,**11**(5):597-603.

[11] Sun D Y, Li Y M, Wang Q. A unified model or remotely estimating chlorophyll a in Lake Taihu, China, based on SVM and in situ hyperspectral[J]. IEEE Trans Geosci Remote Sens, 2009, **47**(8):2957-2965.

[12] Gordon H R, Morel A. Remote assessment of ocean color for interpretation of satellite visible imagery: a review [A]. In: Lecture Notes on Coastal and Estuarine Studies; 4 [M]. New York: Springer Verlag, 1983. 114.

[13] Moore C, Barnard A, Hankins D, *et al.* Spectral Absorption and Attenuation Meter (ac-s) User ‘s Guide, Revision A [M]. America: WET Labs Inc, 2004. 5-20.

[14] 孙德勇,李云梅,黄家柱,等. 太湖梅梁湾夏季水体组分光谱吸收特性[J]. 生态学报,2008,**28**(2):749-760.

[15] Le C F, Li Y M, Zha Y, *et al.* Specific absorption coefficient and the phytoplankton package effect in Lake Taihu, China [J]. Hydrobiologia, 2009, **619**:27-37.

[16] 孙德勇,李云梅,王桥,等. 太湖水体中悬浮颗粒物的比吸收光谱特性及其参数化分析[J]. 环境科学学报, 2009,**29**(9): 1820-1829.

[17] 唐军武,王晓梅,宋庆君,等. 黄、东海二类水体水色要素统计反演模式[J]. 海洋科学进展,2004,**22**(B10):1-7.

[18] 王苏斌. SPSS 统计分析[M]. 北京:电子工业出版社,2007. 290-300.