

Remote Estimation of Chlorophyll-a in Xingyun Lake by Sentinel-2 Satellite

Yumeng Zhang, Quan Wang*, Chunyu Duan, Baoli Jiang, Chaojie Yang, Jie Qin, Baoyin Zhang, Yang Jin

College of Chemistry, Biological and Environment, Yuxi Normal University, Yuxi Yunnan
Email: 657118913@qq.com, *hackerquanw@qq.com

Received: Dec. 28th, 2019; accepted: Jan. 10th, 2020; published: Jan. 17th, 2020

Abstract

Sentinel-2A satellite was used to establish chlorophyll-a remote sensing estimation models for the Xingyun Lake in this study. The correlation between the spectral reflectivity of each band and the concentration of chlorophyll-a was analyzed by the concurrent observation. According to DVI, RVI, NDVI, TBM and MCI algorithms, a combination of different sensitive bands and algorithms is used to set up 21 models. Then the modeling dataset is divided into all datasets, high and low concentration datasets. Three data sets are substituted into 21 models to establish 63 regression equations. The modeling and validation effects of these regression equations are analyzed. The remote sensing estimation of chlorophyll-a concentration in Xingyun Lake is carried out by using different models in different chlorophyll-a concentration ranges of Xingyun Lake. The result shows that: 1) The spectral reflectivity of the 4 Red-Edge bands and Near-Infrared band of the Sentinel-2 satellite is strongly positive related to the concentration of chlorophyll-a. When the concentration of chlorophyll-a was higher than that of 0.1 mg/L, the best model is that the rRMSE and NMAE of RVI1H, were 4.01% and 3.95%, respectively. When the concentration of chlorophyll-a is lower than 0.1 mg/L, the best model is that the rRMSE and NMAE of NDVI1L are 25.95% and 19.32%, respectively. TBM1L model is more suitable to calculate the average value of the whole lake, and its MNB is -0.57%. 2) The regression equation of modeling has high determination coefficient, which can only show that the linearity of the modeling data set is better, but whether the model is applicable or not is mainly based on the error of the verification data.

Keywords

Three-Band Model, Normalized Vegetation Index, Ratio Vegetation Index, Difference Vegetation Index, Maximum Chlorophyll Index, Xingyun Lake, Chlorophyll-a, Sentinel-2 Satellite

基于Sentinel-2卫星的星云湖叶绿素a遥感估算研究

张雨萌, 王 泉*, 段春钰, 蒋宝丽, 杨超杰, 秦 洁, 张葆莹, 金 杨

*通讯作者。

玉溪师范学院, 化学生物与环境学院, 云南 玉溪
Email: 657118913@qq.com, hackerquanw@qq.com

收稿日期: 2019年12月28日; 录用日期: 2020年1月10日; 发布日期: 2020年1月17日

摘要

本研究采用了哨兵2A卫星建立针对星云湖的叶绿素a遥感估算模型。通过星地同步观测分析了各波段反射率与叶绿素a浓度的相关性, 依据DVI, RVI, NDVI, TBM, MCI这5种算法采用不同敏感波段的组合建立了21个模型, 再将建模数据集分为全部数据集, 高浓度和低浓度数据集, 将3个数据集代入21个模型建立63个回归方程, 并分析这些回归方程的建模和验证效果, 提出针对星云湖不同叶绿素a浓度范围的湖区采用不同的模型进行叶绿素a浓度的遥感估算。本研究结果表明: 1) 哨兵2卫星4个红边波段和近红外波段反射率均与叶绿素a浓度呈强烈正相关, 叶绿素a浓度高于0.1 mg/L时, 最佳模型是RVI1H, 其rRMSE和NMAE分别为4.01%和3.95%。叶绿素a浓度低于0.1 mg/L时, 最佳模型是NDVI1L, 其rRMSE和NMAE分别为25.95%, 19.32%, 采用TBM1L模型估算比较适合计算全湖的平均值, 其MNB为-0.57%。2) 建模的回归方程决定系数高, 只能说明建模数据集的线性较好, 但是模型是否适用, 主要依据还是验证数据的误差。综上, 本研究建立的星云湖叶绿素a遥感估算模型, 对于星云湖的蓝藻水华遥感监测具有一定的参考价值。

关键词

三波段模型, 归一化植被指数, 比值植被指数, 差值植被指数, 最大叶绿素指数, 星云湖, 叶绿素a, 哨兵2卫星

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

叶绿素 a (Chl.a)浓度是衡量水体富营养化程度的重要水质参数, 由于湖泊藻类爆发和消亡速度很快, 传统的采样监测不能及时反映湖泊藻类水华的发生和发展, 自动监测站可以高频率地获取水质数据, 但是由于费用较高, 往往不能布设很多监测位点, 不能获取湖泊藻类爆发的空间分布情况。遥感估算叶绿素 a 浓度可以实现高时间频率、且空间上连续覆盖整个湖区的监测数据, 可以更好地、详细地掌握湖泊藻类水华的发生和发展情况。

针对内陆水体的叶绿素 a 遥感估算的方法主要有经验模型法、半经验模型法和生物光学模型法[1]。生物光学模型法主要基于水体的固有光学特征来模拟, 虽然该方法具有很强的稳定性, 但是由于生物光学特征参数很难准确获得, 限制了该方法的广泛应用。经验模型纯粹基于叶绿素 a 实测浓度与卫星传感器各波段光谱反射率之间的回归方程建立, 没有考虑光学活性物质的光谱特征。半经验算法基于待测光学活性物质的光谱特征来选择光谱波段用于建模, 建模过程也较为简单, 主要是通过建立采样点的光谱反射率与地面实测值之间的回归方程。根据光学活性物质的浓度变化幅度采用的回归方法有线性回归, 指数回归和多项式回归, 最常用的还是线性回归方法, 按照使用的光学卫星传感器波段数量, 可以分为单波段、双波段、三波段法。

基于叶绿素 a 在红光和绿光波段有 2 个主要的吸收峰, 又大量反射近红外光, 因此, 红光和近红外这 2 个波段常用于植被及浮游植物色素等含有叶绿素的地物遥感反演。最常用的基于红光和近红外的遥感指数有通过波段比值计算得到比值植被指数(RVI) [2], 以及波段差值与比值相结合计算得到的归一化植被指数(NDVI) [2] [3]。

由于内陆浑浊水体含有许多影响叶绿素 a 吸收的物质, 例如溶解性有机固体(CDOM), 悬浮颗粒物等, 在较宽的近红外波段范围内容易导致双波段经验模型估算叶绿素 a 浓度的稳定性较差。Dall'Olmo 等[4]提出的三波段算法(TBA), 采用了一个叶绿素吸收峰 λ_1 , 一个狭窄的叶绿素荧光峰 λ_2 和一个纯水吸收峰 λ_3 组合, λ_1 和 λ_2 对黄色物质与无机悬浮物的吸收系数近似相等, 两者相减可以去除这些物质的影响, 三个波段总后向散射系数近似相等, TBA 算法能够部分去除非色素物质的吸收信号, 在轻度和中度浑浊水体中得到广泛应用。宋开山[5]等采用哨兵 3 卫星数据估算了澳洲, 美国和中国的多个湖泊的上千个采样点, 取得了时间跨度长达 5 年的叶绿素 a 浓度数据, 研究认为三波段算法取得了与遗传算法同样的稳定性, 其相对均方根误差 rRMSE 为 33.4%。最大叶绿素指数(MCI)算法是一种基线算法, 其原理是将 λ_1 与 λ_3 的遥感反射率连线作为基线, 然后用 λ_2 处的遥感反射率与该基线的距离来作为自变量与叶绿素 a 浓度进行回归分析, 理论上该算法可以去除水中其它色素物质的干扰, 该算法充分考虑了浮游植物固有光学特征, 具有一定的稳定性[6]。然而, MCI 算法与 NAP 和 CDOM 之间有微弱的相关性, 因此用 MCI 反演叶绿素 a 会产生不可忽略的误差。叶绿素 a 浓度的变化也会影响浮游植物色素吸收峰的数值和位置 [7], 即随着叶绿素 a 浓度增加, 峰值逐渐增大, 并且向长波方向移动, 出现明显的“红移”现象[8], 所以针对不同浓度范围的叶绿素 a 遥感估算, 选择不同的波段来运算效果会更好。Matsushita 等对东亚地区 5 个湖泊水体进行遥感估算研究, 将湖泊区域按照不同的 MCI 范围, 采用不同的回归模型来估算叶绿素 a 浓度, 组合算法得到的归一化平均绝对误差(NMAE)仅有 13.3% [9]。

随着卫星传感器技术的进步, 狭窄的红边波段对叶绿素荧光激发非常敏感, 同时还可以避免其它色素物质的干扰, 非常适合用于叶绿素 a 的遥感估算。欧洲空间局(ESA)的哨兵 2 (Sentinel-2)和我国的高分六号卫星[10]都装备了红边波段的传感器, 且这 2 种卫星都具有较高的空间分辨率, 其中哨兵 2 卫星的 MSI 传感器有 4 个红边波段。以往的研究往往采用某一个模型算法就固定采用一种波段组合来进行计算, 灵活性较差。理论上, MSI 传感器 4 个红边波段都是对叶绿素 a 非常敏感的, 可以分别将不同的红边波段作为上述算法中的叶绿素敏感波段 λ_2 , 分别代入模型中计算, 优选出误差最小的模型。

有一些采用经验模型估算叶绿素 a 的研究得到了较好的效果, 误差较小, 但是这些研究采用的建模数据集和验证数据集均来自同一次采样的数据, 无从得知模型的适应性, 采用不同日期的星地同步观测数据分别建模和验证模型才能得知模型的普适性。

杨国范等[11]研究清河水库的叶绿素 a 浓度遥感估算, 建模和验证采用同一天的星地同步观测数据, 其 LS-SVM 模型平均相对误差为 7.21%, 线性回归模型平均相对误差为 16.43%。杨硕等[12]采用实测光谱数据来优选 TBM 模型中三个波段的位置, 建模与验证数据集来自于不同批次的采样, 前后相差一个月, 结果验证数据的平均相对误差在 20%~25%之间。马荣华等采用 MODIS 影像, 基于经验正交函数 EOF 估算太湖叶绿素 a 浓度, 用不同批次的数据集进行建模和验证, 得到无偏均方根误差接近 80% [13]。前述宋开山等[5]的研究采用的 TBM 模型虽然得到了 rRMSE 超过 30%的结果, 但是其采用了时间和地域跨度均很大的数据集来验证, 其验证结果可以证明模型具有很强的普适性。

星云湖位于云南省玉溪市江川区, 属于高原断层淡水湖, 是云南省九大高原湖泊之一, 总面积约 34 km²。[14] 20 世纪 90 年代初期, 星云湖开始出现藻类水华, 营养级别属于中 - 富营养水平, 到了 90 年代末, 藻类水华大面积周期性爆发, 营养级别达到富营养化的程度, 水质下降到劣 V 类[15]。

迄今为止,针对星云湖的叶绿素 a 估算的算法研究都是采用 MODIS 卫星建立经验估算模型[16] [17], 由于 MODIS 卫星对浮游植物敏感的波段空间分辨率为 1 km, 而整个星云湖只相当于 34 个像元, 因此采用 MODIS 估算星云湖的叶绿素浓度的结果会非常粗糙。

Sentinel-2 (哨兵 2) 卫星是欧盟“全球环境与安全监测”计划的第二颗卫星, 由两颗卫星组成(A 星和 B 星), 携带一枚多光谱成像仪 MSI, 拥有 13 个光谱波段, 10 m 空间分辨率, 双星重访周期达到 5d, 目前全世界在轨的可以公开获取数据的多光谱卫星中, Sentinel-2 是空间分辨率和光谱分辨率最高的。采用哨兵 2 对星云湖进行遥感监测, 可以获得 34 万个像元, 空间解析度远比 MODIS 高[10]。

本研究采用 Sentinel-2 卫星进行星地同步观测, 以此建立星云湖叶绿素 a 浓度估算模型, 采用 MCI、TBM、NDVI、RVI、DVI, 5 种算法, 并将多个叶绿素敏感的红边波段应用在上述模型中, 得到多种算法的组合, 优选误差最小的算法用于星云湖叶绿素 a 浓度的遥感估算。本研究建立的星云湖叶绿素 a 估算模型对实现针对星云湖的高时空分辨率的蓝藻水华遥感监测具有重要的参考意义。

2. 研究方法

2.1. 采样点布设

本次研究对星云湖进行了 2 次采样, 采样时间分别为 2018 年 10 月 18 日和 2018 年 11 月 18 日, 采样点的布设如图 1 所示, 采用思拓力 S7-D 型 GPS 定位, 定位误差在 1 米以内, 10 月份有 13 个采样点, 11 月份有 9 个采样点, 编号及位置如图 1 所示。

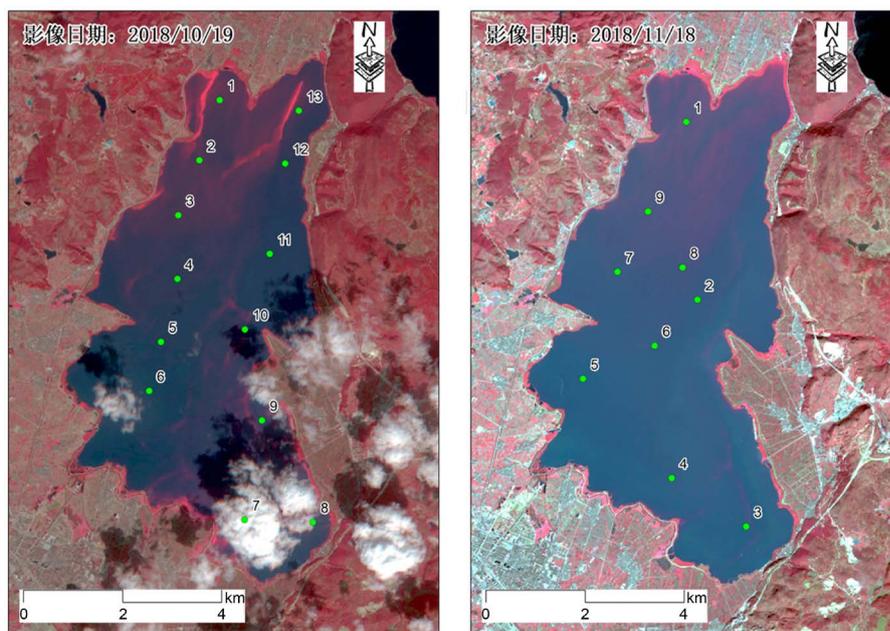


Figure 1. Sampling point layout and the Sentinel-2A satellite image at pseudo-color

图 1. 采样点布设及哨兵 2A 卫星影像假彩色合成图

2.2. 水样叶绿素 a 的测定方法

采用热乙醇萃取分光光度法对水样叶绿素 a 进行测定[18] [19], 即将采集的水样用 0.45 μm 的水系滤膜抽滤后放入冰箱 -20°C 冷冻 24 h。然后取出来将滤膜剪碎放入离心管, 用 80°C 水浴 3 min 的 90% 乙醇倒入离心管, 避光浸泡 4 小时, 然后 3000 转离心 10 min 后再用 90% 乙醇作为参比液进行比色, 先后在

665nm 和 750 nm 波长测消光率 E_{665} 、 E_{750} ，然后在样品比色皿中加 1 滴 1 mol/L 盐酸进行酸化，加盖摇匀，1min 后重新在 665 nm、750 nm 波长测消光率 A_{665} 、 A_{750} ，再按公式(1)进行叶绿素 a 的浓度计算。叶绿素 a 计算公式如下：

$$\text{Chl.a} = 27.9V_{\text{乙醇}} \left[(E_{665} - E_{750}) - (A_{665} - A_{750}) \right] / V_{\text{样品}} \quad (1)$$

其中 Chl.a 是叶绿素的浓度(mg/L)， $V_{\text{乙醇}}$ 是乙醇萃取液定容的体积(ml)， $V_{\text{样品}}$ 是过滤水样的体积(ml)。

2.3. 卫星数据

用于星地同步观测的 Sentinel-2A 卫星数据通过 USGS (<http://glovis.usgs.gov>)获取，原始卫星影像为 L1C 级别，采用官方的预处理工具 SNAP 和 Sen2cor 进行大气校正预处理为 L2A 级别。本研究两次采样的日期和时间分别为 2018 年 10 月 18 日早上 10~12 点，2018 年 11 月 18 日早上 7~9 点(图 1)。Sentinel-2A 卫星过境星龙湖的时间分别为 2018 年 10 月 19 日 11 点 37 分，2018 年 11 月 18 日 11 点 40 分。因此，2018 年 10 月 18 日采样，19 日卫星过境的数据属于准同步数据，相差一天。2018 年 11 月 18 日采样与卫星过境的时间相差在 4 个小时以内，属于同步数据。

2.4. 波段敏感性分析及叶绿素 a 估算模型的建立

本研究首先分析 MSI 各波段与建模数据集叶绿素 a 浓度的相关关系，采用 SPSS 计算皮尔逊相关系数，然后选择具有显著相关性，相关系数高的波段，基于比值植被指数(RVI) [2]、差值植被指数(DVI)、归一化差分植被指数(NDVI) [2]、三波段(TBM) [5]、最大叶绿素指数(MCI) [9]作为自变量，建立相应的线性回归模型。上述算法的表达式如下：

$$\text{RVI} = R_{rs}(\lambda_2) / R_{rs}(\lambda_1) H \quad (2)$$

$$\text{DVI} = R_{rs}(\lambda_2) - R_{rs}(\lambda_1) \quad (3)$$

$$\text{NDVI} = \frac{R_{rs}(\lambda_2) - R_{rs}(\lambda_1)}{R_{rs}(\lambda_2) + R_{rs}(\lambda_1)} \quad (4)$$

$$\text{TBM} = \left[R_{rs}(\lambda_1)^{-1} - R_{rs}(\lambda_2)^{-1} \right] R_{rs}(\lambda_3) \quad (5)$$

$$\text{MCI} = R_{rs}(\lambda_2) - R_{rs}(\lambda_1) - \left[\frac{(\lambda_2 - \lambda_1)}{(\lambda_3 - \lambda_1)} (R_{rs}(\lambda_3) - R_{rs}(\lambda_1)) \right] \quad (6)$$

上式中， R_{rs} 是光谱反射率， λ_2 是近红外波段(叶绿素强烈反射波段)， λ_1 是红光波段(叶绿素强烈吸收波段)， λ_3 是 750 nm 附近的波段。本研究将根据 MSI 传感器的波段设置，基于各波段反射率与叶绿素浓度之间的相关分析结果，采取不同的波段组合构成自变量，将与叶绿素浓度呈显著正相关的波段作为上述遥感指数的 λ_2 ，与叶绿素浓度呈显著负相关的波段作为 λ_1 ，建立回归模型。模型建立和模型验证均采用缩减主轴回归分析(RMA) [20]来计算斜率、截距和决定系数 R^2 。

关于建模和验证数据集的选取，由于 10 月 18 日采样的 7, 8, 9, 10 号位点在 10 月 19 日的卫星影像里面有云和云下阴影遮盖，所以将其剔除；1, 2 号位点由于叶绿素 a 浓度太高，将其作为异常值剔除。最终用于验证的有效位点为编号 3, 4, 5, 6, 11, 12, 13 这 7 个点。11 月 18 日采样的 4 号位点数据异常，引起建模误差较大，将其剔除。最终，有 8 个 2018 年 11 月 18 日的采样点的数据用于建模，7 个同年 10 月 18 日采样的数据用于验证。

陈宇炜[19]等认为一般富营养化水体的叶绿素 a 浓度在 0.1 mg/L 以下，因此建模实测数据以 0.1 mg/L 为界限，分为高浓度和低浓度样点，分别建立 All, High 和 Low 三个数据集建模，其中 All 数据集采用

全部 8 个采样点建模, High 数据集采用高浓度的 4 个采样点数据建模, Low 数据集采用低浓度的 4 个采样点数据建模。验证数据也分为 All, High 和 Low 三个数据集, 其中 High 数据集有 4 个采样点, Low 数据集有 3 个采样点(见表 1)。

Table 1. Monitoring data of chlorophyll-a in Xingyun Lake (mg/L)

表 1. 星云湖叶绿素 a 监测数据(mg/L)

日期	最大值	最小值	平均值	样点数	数据集类型
	0.25575	0.02325	0.10296	7	All ^①
2018/10/18	0.25575	0.11625	0.17825	4	High ^②
	0.06975	0.02325	0.0465	3	Low ^③
	0.20088	0.01116	0.09796	8	All ^①
2018/11/18	0.20088	0.10044	0.14787	4	High ^②
	0.03348	0.01116	0.02511	4	Low ^③

① All dataset; ② Data with a concentration greater than 0.1 mg/L of chlorophyll-a is a high concentration dataset; ③ Data with a concentration lower than 0.1 mg/L of chlorophyll-a is a low concentration dataset.

① 全部数据集; ② 叶绿素 a 浓度大于 0.1 mg/L 的数据为高浓度数据集; ③ 叶绿素 a 浓度小于 0.1 mg/L 的数据集为低浓度数据集。

2.5. 误差的估算

本研究采用多个误差评价指标, 包括均方根误差(RMSE), 相对均方根误差(rRMSE), 归一化均方根误差(NRMS), 平均归一化偏倚(MNB), 归一化平均绝对误差(NMAE) [9]。其中 RMSE 是指估计值和观测值偏差的平方与观测次数 n 比值的平方根, 能很好地反映出模型的误差, RMSE 计算公式为:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (X_{\text{esti},i} - X_{\text{meas},i})^2}{N-1}} \quad (7)$$

式中, $X_{\text{esti},i}$ 指估算值, $X_{\text{meas},i}$ 指的是实测值, N 为总的的数据量。

相对均方根误差(rRMSE)的计算公式为:

$$\text{rRMSE} = \text{RMSE} / \text{Mean}_{\text{meas}} \quad (8)$$

其中 $\text{Mean}_{\text{meas}}$ 为实测叶绿素 a 的浓度平均值。

NRMS 表示结果的相对随机不确定性, 它代表 ε_i 的标准差, 计算公式为:

$$\text{NRMS} = \text{stdev}(\varepsilon_i) \% \quad (9)$$

ε_i 是测量值与估算值之差所占实测值的百分比, 计算公式如下:

$$\varepsilon_i = \left((X_{\text{esti},i} - X_{\text{meas},i}) / X_{\text{meas},i} \right) \times 100 \quad (10)$$

MNB 表示估算的平均偏差, 它代表 ε_i 的平均值; NMAE 表示估算的平均绝对误差, 它代表 ε_i 的绝对值的平均值, 计算公式如下:

$$\text{MNB}(\%) = \text{mean}(\varepsilon_i) \quad (11)$$

$$\text{NMAE}(\%) = \text{mean}(|\varepsilon_i|) \quad (12)$$

3. 研究结果与讨论

3.1. 叶绿素 a 浓度测定结果与波段敏感性分析

从现场实测数据来看(表 1), 10 月份监测数据叶绿素 a 浓度较高, 到了 11 月份, 叶绿素 a 浓度有所降低, 从图 1 呈粉红色的区域分布也可以看出 10 月 19 日星云湖南北湖区的蓝藻水华分布比较集中, 颜色较深, 与实地监测叶绿素浓度的分布情况较为一致。2018 年 11 月 19 日的叶绿素 a 浓度高值区域位于星云湖北部湖区, 并且颜色要比 10 月 18 日的浅, 推测 11 月的富营养化程度没有 10 月份的严重。

波段敏感性分析表明在使用建模数据中的全部数据集的时候, 红光波段(Red)反射率值与实测叶绿素 a 浓度呈现显著的负相关, 而 4 个红边波段(VRE)和近红外波段(NIR)反射率与实测叶绿素 a 浓度呈现显著正相关。而高浓度和低浓度数据集与红光波段都是负相关, 但是显著性不强。在红边和近红外波段, 除了低浓度数据集与 B8a 红边波段反射率相关系数很高, 但是显著性不强以外, 其余波段均与相应数据集有显著正相关, 见表 2。本研究将选择 B4 作为 λ_1 , B5, B6, B7, B8, B8a 波段作为 λ_2 代入上述 5 个模型中分析。

Table 2. Pearson correlation analysis result

表 2. 皮尔逊相关性分析结果

数据集	B4(Red)	B5 (VRE)	B6 (VRE)	B7 (VRE)	B8 (NIR)	B8a (VRE)
ALL	-0.741*	0.979**	0.973**	0.967**	0.974**	0.963**
HIGH	-0.305	0.997**	0.995**	0.996**	0.997**	0.985*
LOW	-0.915	0.981*	0.994**	0.978*	0.995**	0.933

*At level 0.05 (double tail), **At level 0.01 (double tail), the correlation was significant.

*在 0.05 级别(双尾), **在 0.01 级别(双尾), 相关性显著。

3.2. 模型的构建和验证

本研究采用 2018 年 11 月 18 日的星地同步监测数据进行建模, 然后用 2018 年 10 月 18 日采样, 第二天卫星观测的准同步数据进行验证。本研究中模型的命名规则如下, 例如: MCI1H, 其中的 MCI 为自变量 x 的算法名称, 数字 1 为编号, H 为高浓度数据集, 建模分别采用 2018 年 11 月 18 日的 A, H, L 数据集, 验证数据也对应采用同年 10 月 18 日的 A, H, L 数据集。建模和验证结果如表 3 所示。

建模结果表明, 全部数据集建模误差均较大(见表 3), 误差最小的模型是 DV11A, 其 rRMSE 为 32.77%, NMAE 为 28.44%, MNB 为 2.94%。高浓度数据集建模误差最小的模型是 RV11H (见表 4), 其 rRMSE 和 NMAE 分别为 4.01% 和 3.95%。表现较好的模型有 MCI2H, DV11H 以及所有的 NDVI 各波段组合, 其 NMAE 均小于 10%, rRMSE 均小于 11%。低浓度数据集表现最好的模型是 NDVI1L 和 DV12L (见表 5), 前者 rRMSE 为 25.95%, NMAE 为 19.32%, MNB 为 -13.3%, 后者的 rRMSE 为 25.71%, NMAE 为 24.86%, MNB 为 -17.32%。低浓度数据集 MNB 最低的模型是 TBM1L, 为 -0.57%。MNB 较低的模型模拟值与实测值之间正负偏离抵消较多, 比较适用于计算全湖的平均值。

本研究还发现, 三个数据集建模最佳的模型选择的 λ_2 除了低浓度数据集的 DV12L 选择红边波段 B6 (中心波长 740 nm) 以外, 其余都是红边波段 B5 (705 nm), 与近红外波段 B8 的 842 nm 相距较远, 与红光波段的中心波长 665 nm 距离较近。因此, 全部数据集建模, 用差值植被指数作为自变量, 误差最小, 如果估算全湖平均值可能得到更低的误差。高浓度数据集则用 RVI 指数作为自变量最适合, 无论是估算单个点的还是全湖平均值, 误差均是最低的。此外, 高浓度数据集建模, 自变量可以选用的指数很多, 包括 NDVI, DVI, MCI 都表现不错。而低浓度数据集建模, 可以选用 NDVI 或者 DVI 作为自变量, 它的单点估算误差最低, 如果要做全湖平均值的估算, 选用 TBM 作为自变量是最适合的。

Table 3. Estimation and Verification results of chlorophyll-a model (All dataset)**表 3.** 叶绿素 a 遥感估算模型建立和验证结果(全部数据集)

模型名称	λ_1	λ_2	λ_3	斜率	截距	R ²	RMSE	rRMSE (%)	NRMS (%)	MNB (%)	NMAE (%)
MCI1A	B4	B5	B6	12.618	-0.2925	0.847	0.0862	83.69	89.50	-11.72	70.65
MCI2A	B4	B5	B7	8.059	-0.164	0.915	0.0506	49.18	53.01	-3.12	36.93
MCI1L	B4	B5	B6	1.428	-0.008483	0.983	0.0242	52.05	33.80	-43.04	43.04
MCI2L	B4	B5	B7	1.064	0.002668	0.958	0.0261	56.04	29.11	-50.05	50.05
TBM1A	B4	B5	B6	0.1563	-0.01611	0.963	0.0614	59.62	61.29	18.29	45.33
TBM2A	B4	B7	B6	0.1269	0.06769	0.946	0.0624	60.58	61.93	19.56	45.75
TBM3A	B4	B8	B6	0.132	0.0802	0.961	0.0711	69.06	71.94	18.25	52.98
TBM4A	B4	B8a	B6	0.1321	0.164	0.933	0.0712	69.13	64.52	34.80	53.62
RVI1A	B4	B5		0.1341	-0.2015	0.944	0.0435	42.26	44.42	9.71	33.48
RVI2A	B4	B6		0.1218	-0.05622	0.964	0.0589	57.20	58.36	18.76	43.36
RVI3A	B4	B7		0.1242	-0.0568	0.958	0.0640	62.16	63.13	21.18	46.07
RVI4A	B4	B8		0.1513	-0.06809	0.966	0.0740	71.85	73.82	22.17	53.45
RVI5A	B4	B8a		0.2322	-0.06655	0.958	0.0708	68.80	66.20	31.26	50.65
NDVI1A	B4	B5		0.6619	-0.1441	0.901	0.0484	46.98	50.68	2.35	35.33
NDVI2A	B4	B6		0.2998	0.08199	0.932	0.0450	43.74	45.53	11.70	31.86
NDVI3A	B4	B7		0.3084	0.0836	0.942	0.0507	49.28	51.29	13.18	33.34
NDVI4A	B4	B8		0.3342	0.09956	0.963	0.0583	56.62	59.84	11.67	38.15
NDVI5A	B4	B8a		0.3378	0.1694	0.941	0.0531	51.62	49.87	23.09	36.03
DVI1A	B4	B5		5.055	-0.07714	0.954	0.0337	32.77	35.25	2.94	28.44
DVI2A	B4	B6		4.278	0.06789	0.964	0.0500	48.60	51.01	11.46	38.70
DVI3A	B4	B7		4.35	0.06965	0.956	0.0504	48.97	51.47	11.28	39.56
DVI4A	B4	B8		5.27	0.08582	0.966	0.0583	56.60	60.25	9.60	47.24
DVI5A	B4	B8a		7.953	0.1681	0.946	0.0703	68.30	69.88	21.88	53.23

Table 4. Estimation and Verification results of chlorophyll-a model (High dataset)**表 4.** 叶绿素 a 遥感估算模型建立和验证结果(高浓度数据集)

模型名称	λ_1	λ_2	λ_3	斜率	截距	R ²	RMSE	rRMSE (%)	NRMS (%)	MNB (%)	NMAE (%)
MCI1H	B4	B5	B6	9.643	-0.1733	0.675	0.0563	31.60	36.85	9.67	31.55
MCI2H	B4	B5	B7	8.086	-0.1623	0.919	0.0182	10.18	7.46	8.16	8.16
TBM1H	B4	B5	B6	0.1133	0.03592	0.977	0.0416	23.33	27.89	5.06	19.78
TBM2H	B4	B7	B6	0.08882	0.09726	0.991	0.0373	20.92	25.25	3.52	18.25
TBM3H	B4	B8	B6	0.09377	0.1078	0.991	0.0431	24.19	29.51	2.15	22.00
TBM4H	B4	B8a	B6	0.09973	0.1683	0.984	0.0478	26.79	30.82	9.19	20.63
RVI1H	B4	B5		0.09619	-0.09147	0.806	0.0071	4.01	4.41	1.77	3.95
RVI2H	B4	B6		0.08953	0.004734	0.983	0.0385	21.59	26.22	2.80	19.25
RVI3H	B4	B7		0.09011	0.00768	0.997	0.0365	20.47	24.69	3.53	17.82
RVI4H	B4	B8		0.111	-0.003111	0.994	0.0421	23.59	28.81	1.81	21.57
RVI5H	B4	B8a		0.1755	-0.006778	0.991	0.0470	26.34	30.67	8.19	20.87
NDVI1H	B4	B5		0.6836	-0.149	0.866	0.0150	8.44	7.42	-5.88	7.17

Continued

NDVI2H	B4	B6	0.2828	0.09277	0.92	0.0151	8.47	4.18	-7.75	7.75
NDVI3H	B4	B7	0.2814	0.09631	0.952	0.0129	7.21	3.35	-6.67	6.67
NDVI4H	B4	B8	0.2908	0.1129	0.941	0.0195	10.92	7.84	-8.85	8.85
NDVI5H	B4	B8a	0.3078	0.1719	0.984	0.0157	8.81	10.00	-3.31	8.72
DVI1H	B4	B5	4.405	-0.04396	0.973	0.0178	10.00	9.50	6.31	7.58
DVI2H	B4	B6	3.208	0.09494	0.991	0.0375	21.06	25.53	3.05	18.63
DVI3H	B4	B7	3.249	0.09849	0.999	0.0362	20.34	24.39	4.10	17.40
DVI4H	B4	B8	3.964	0.1099	0.994	0.0416	23.35	28.49	2.05	21.23
DVI5H	B4	B8a	6.128	0.1687	0.994	0.0453	25.42	30.39	5.53	21.50

Table 5. Estimation and Verification results of chlorophyll-a model (Low dataset)**表 5.** 叶绿素 a 遥感估算模型建立和验证结果(低浓度数据集)

模型名称	λ_1	λ_2	λ_3	斜率	截距	R ²	RMSE	rRMSE (%)	NRMS (%)	MNB (%)	NMAE (%)
MC11L	B4	B5	B6	1.428	-0.008483	0.983	0.0242	52.05	33.80	-43.04	43.04
MC12L	B4	B5	B7	1.064	0.002668	0.958	0.0261	56.04	29.11	-50.05	50.05
TBM1L	B4	B5	B6	0.09088	0.00304	0.989	0.0211	45.34	52.35	-0.57	39.44
TBM2L	B4	B7	B6	0.06677	0.04338	0.947	0.0239	51.30	59.17	2.34	42.96
TBM3L	B4	B8	B6	0.1057	0.0632	0.971	0.0439	94.41	104.78	26.08	64.90
TBM4L	B4	B8a	B6	0.04402	0.06858	0.232	0.0174	37.44	43.11	-2.87	33.52
RVI1L	B4	B5		0.0358	-0.03484	0.96	0.0158	33.94	29.05	-22.78	31.19
RVI2L	B4	B6		0.06501	-0.02218	0.994	0.0193	41.44	47.81	1.80	34.48
RVI3L	B4	B7		0.06557	-0.02251	0.977	0.0256	55.08	63.36	4.78	44.73
RVI4L	B4	B8		0.09272	-0.03642	0.997	0.0383	82.45	92.32	20.14	59.06
RVI5L	B4	B8a		0.09941	-0.01754	0.909	0.0211	45.34	52.05	4.88	35.16
NDVI1L	B4	B5		0.1024	0.008346	0.927	0.0121	25.95	25.71	-13.33	19.32
NDVI2L	B4	B6		0.09206	0.0404	0.992	0.0142	30.49	32.07	-12.58	28.14
NDVI3L	B4	B7		0.0948	0.0409	0.989	0.0168	36.09	39.30	-12.00	33.91
NDVI4L	B4	B8		0.1243	0.05079	0.999	0.0220	47.30	54.48	-3.36	42.21
NDVI5L	B4	B8a		0.1008	0.06598	0.923	0.0156	33.47	37.96	-6.25	28.21
DVI1L	B4	B5		1.349	-0.002367	0.975	0.0185	39.80	18.39	-36.48	36.48
DVI2L	B4	B6		1.963	0.04154	0.999	0.0120	25.71	21.95	-17.32	24.86
DVI3L	B4	B7		2.078	0.04231	0.988	0.0147	31.61	29.96	-18.06	31.43
DVI4L	B4	B8		2.832	0.05378	0.998	0.0205	44.15	48.84	-12.66	42.72
DVI5L	B4	B8a		2.942	0.07549	0.961	0.0211	45.43	50.60	-12.00	43.80

3.3. 模型应用

根据前述建模和验证的效果优选出来的模型,在实际应用中,本研究提出基于浓度分区的星云湖叶绿素 a 遥感估算方法,既首先选用 DVI1A 模型进行估算,将估算结果按照叶绿素 a 浓度 0.1 mg/L 作为阈值,将湖区分为高浓度区域和低浓度区域,然后分别针对高浓度和低浓度区域使用对应的最适用模型再

次估算, 高浓度湖区采用 RVIIH 模型, 低浓度湖区采用 NDVIIL 模型进行估算, 该方法由 ENVI/IDL 代码实现, 计算机逐像元判断第一次估算的结果, 并选择合适的模型进行第二次估算计算, 最终得到星云湖叶绿素 a 浓度模拟结果, 如图 2 和表 6 所示。

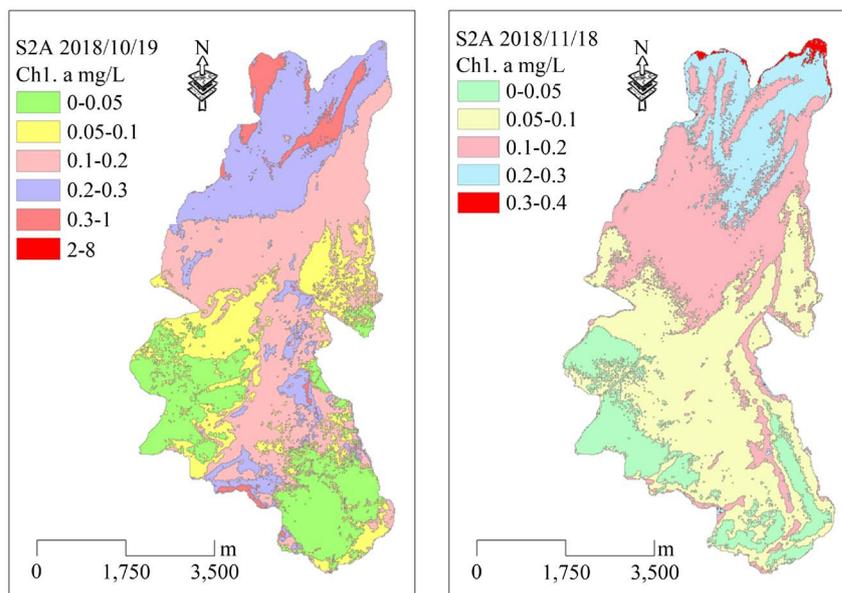


Figure 2. Remote estimation of chlorophyll-a in the Sentinel-2A satellite at the Xingyun Lake
图 2. Sentinel-2A 卫星估算星云湖叶绿素 a 分布图

Table 6. The results of remote estimation chlorophyll-a concentration at Xingyun Lake (mg/L)

表 6. 星云湖叶绿素 a 浓度遥感估算结果(mg/L)

日期	最大值	平均值	平均值 + 2 × 标准差
2018/10/19	7.5112	0.1453	0.3411
2018/11/18	0.4342	0.1146	0.2467

两次遥感估算应用结果表明, 叶绿素 a 高浓度分布区域与卫星影像上的藻类水华大量爆发区域一致。其中 11 月 18 日的藻类水华主要分布在北部湖区, 而 10 月 19 日在南北两边都有大量藻类覆盖湖面, 但是南边部分区域因为有云, 而云团被误识别为低浓度的叶绿素 a, 范围在 0~0.05 mg/L 之间, 见图 2。

基于浓度分区估算星云湖叶绿素 a 的结果最大值分别达到 0.4342 mg/L (11 月 18 日) 和 7.5112 mg/L (10 月 19 日), 估算结果最小值均小于 0, 计算机自动调整为 0, 平均值为 0.1146 mg/L (11 月 18 日), 0.1453 mg/L (10 月 19 日)。2018 年 10 月 19 日估算的结果大于 0.47 mg/L 的像元仅有 8 个, 可以认为是异常值, 其主要的叶绿素 a 浓度分布区间可以用平均值加减 2 倍标准差来表达, 其中平均值减去 2 倍标准差均小于 0, 可调整为 0, 平均值加上两倍标准差的结果如表 6 所示, 10 月 19 日的结果为 0.3411 mg/L, 11 月 18 日估算的结果为 0.2467 mg/L, 10 月份的估算结果明显高于 11 月份, 这与假彩色影像上的主观感受类似。

4. 结论

本研究采用了 Sentinel-2A 卫星对星云湖的叶绿素 a 浓度开展遥感估算, 通过星地同步观测, 建立叶绿素 a 估算模型。首先, 以实测叶绿素 a 浓度 0.1 mg/L 为阈值, 将建模数据分为三个数据集, 分别是高浓度、低浓度和全部数据集; 其次, 将建模数据与哨兵 2A 卫星各波段反射率值做相关分析, 得到 5 个

显著正相关的红边和近红外波段, 根据敏感波段结合 DVI, RVI, NDVI, TBM, MCI 这 5 种算法组合形成 21 种不同波段组合的算法, 将其作为自变量 x , 分别代入 3 个建模数据集, 形成 63 个回归方程, 比较这些回归方程的建模和验证效果。优选出适合不同浓度区间的最佳模型, 提出基于浓度分区的方法对星云湖叶绿素 a 进行遥感估算, 该方法可以灵活地针对不同叶绿素 a 浓度范围的湖区, 应用不同的模型进行运算, 以此来提高叶绿素 a 估算精度。本研究得到如下结论:

1) 哨兵 2 卫星 4 个红边波段和近红外波段反射率均与叶绿素 a 浓度呈强烈正相关, 叶绿素 a 浓度高于 0.1 mg/L 时, 采用 RVI1H 模型, 误差最低, 其 rRMSE 和 NMAE 分别为 4.01% 和 3.95%。叶绿素 a 浓度低于 0.1 mg/L 时, 采用 NDVI1L 模型进行估算, rRMSE 为 25.95%, NMAE 为 19.32%, MNB 为 -13.3%, 采用 TBM1L 模型估算时, MNB 为 -0.57%, 说明模拟值与实测值之间正负偏离抵消较多, 比较适用于计算全湖的平均值。

2) 基于浓度分区对星云湖叶绿素 a 进行估算, 2018 年 10 月 19 日估算的叶绿素 a 浓度大于同年 11 月 18 日的估算结果, 这个结论与标准假彩色影像的感官结果类似。

3) 低浓度叶绿素 a 估算模型误差较高浓度模型的误差大, 主要原因是本研究基于两次不同时间的星地同步和准同步观测来建模和验证, 不同时间的卫星数据, 由于大气状况不同, 大气校正也并不是完美地消除所有大气的影晌, 相同的地物在两次不同时相的影像中的反射率值是有细微的差别的, 而不同批次的采样和实验也会带来一定的误差, 气象因子方面, 风的作用会改变浮藻的位置, 而低浓度范围的叶绿素 a 估算精度更容易受到这些误差的影响。在今后的研究当中需要积累更多的星地同步观测数据, 同时也要加强对星云湖中其它光学敏感物质的监测研究, 排除各种干扰物质的影响, 进一步提高叶绿素 a 的估算精度。

4) 建模的时候得到的回归方程的决定系数高, 只能说明建模数据集的线性较好, 但是模型是否适用, 主要依据还是验证数据的误差。

基金项目

玉溪师范学院大学生创新创业训练计划项目(编号 2018A33), 云南省地方本科高校(部分)基础研究联合专项项目(编号 2017FH001-100, 2018FH001-067, 2018FD094, 2017FD161)和云南省教育科学规划项目(编号 GJZ171813)联合资助。

参考文献

- [1] O'Reilly, J.E. and Werdell, P.J. (2019) Chlorophyll Algorithms for Ocean Color Sensors: OC4, OC5 & OC6. *Remote Sensing of Environment*, **229**, 32-47. <https://doi.org/10.1016/j.rse.2019.04.021>
- [2] Gupta, R.K., Prasad, S., Nadham, T.S.V., et al. (1993) Relative Sensitivity of District Mean RVI and NDVI over an Agrometeorological Zone. *Advances in Space Research*, **13**, 261-264. [https://doi.org/10.1016/0273-1177\(93\)90555-P](https://doi.org/10.1016/0273-1177(93)90555-P)
- [3] Zhang, J., Li, M., Sun, Z., et al. (2018) Chlorophyll Content Detection of Field Maize Using RGB-NIR Camera. *IFAC-PapersOnLine*, **51**, 700-705. <https://doi.org/10.1016/j.ifacol.2018.08.114>
- [4] Giorgio, D.O. and Gitelson, A.A. (2005) Effect of Bio-Optical Parameter Variability on the Remote Estimation of Chlorophyll-a Concentration in Turbid Productive Waters: Experimental Results. *Applied Optics*, **44**, 412-422. <https://doi.org/10.1364/AO.44.000412>
- [5] Song, K., Lin, L., Tedesco, L.P., et al. (2013) Remote Estimation of Chlorophyll-a in Turbid Inland Waters: Three-Band Model versus GA-PLS Model. *Remote Sensing of Environment*, **136**, 342-357. <https://doi.org/10.1016/j.rse.2013.05.017>
- [6] 刘阁, 李云梅, 吕恒, 等. 基于 MERIS 影像的洪泽湖叶绿素 a 浓度时空变化规律分析[J]. 环境科学, 2017, 38(9): 3645-3656.
- [7] 周琳, 马荣华, 段洪涛, 等. 浑浊 II 类水体叶绿素 a 浓度遥感反演(I): 模型的选择[J]. 红外与毫米波学报, 2011, 30(6): 531-536.

-
- [8] Xu, J.P., Fang, L., Bai, Z., *et al.* (2008) Improved Conceptual Three-Band Model for Chlorophyll-a Retrieval in Inland Case-II Waters. *Proceedings of SPIE—The International Society for Optical Engineering*, Volume 7145, 71451L. <https://doi.org/10.1117/12.813039>
- [9] Matsushita, B., *et al.* (2015) A Hybrid Algorithm for Estimating the Chlorophyll-a Concentration across Different Trophic States in Asian Inland Waters. *ISPRS Journal of Photogrammetry and Remote Sensing*, **102**, 28-37. <https://doi.org/10.1016/j.isprsjprs.2014.12.022>
- [10] Christopherson, J.B., Ramaseri Chandra, S.N. and Quanbeck, J.Q. (2019) Joint Agency Commercial Imagery Evaluation—Land Remote Sensing Satellite Compendium. 208. <https://doi.org/10.3133/cir1455>
- [11] 杨国范, 阎孟冬, 殷飞. 清河水库叶绿素 a 浓度反演模型研究[J]. 遥感信息, 2016, 31(5): 74-78.
- [12] 杨硕, 王世新, 周艺, 等. 叶绿素反演三波段模型的多时相应用[J]. 遥感信息, 2010(5): 98-104.
- [13] Qi, L., *et al.* (2014) An EOF-Based Algorithm to Estimate Chlorophyll a Concentrations in Taihu Lake from MODIS Land-Band Measurements: Implications for Near Real-Time Applications and Forecasting Models. *Remote Sensing*, **6**, 10694-10715. <https://doi.org/10.3390/rs61110694>
- [14] 郑田甜, 赵祖军, 赵筱青, 等. 云南星云湖水质变化及其人文因素驱动力分析[J]. 湖泊科学, 2018, 30(1): 79-90.
- [15] 冯梅. 星云湖水质评价及富营养化成因分析[J]. 环境科学导刊, 2005, 24(s2): 96-99.
- [16] 种丹, 李浩杰, 范硕, 等. 基于 MODIS 数据的云南九大高原湖泊叶绿素 a 浓度反演[J]. 生态学杂志, 2017, 36(1): 277-286.
- [17] 张洁, 张志. 基于 modis 数据的云南抚仙湖星云湖水质污染遥感调查方法研究[J]. 水文地质工程地质, 2008, 35(5): 92-96.
- [18] 冯青英, 陈盛, 程麒, 等. 应用热乙醇法提取浮游植物中叶绿素 a 的探讨[J]. 安徽农业科学, 2012, 40(29): 14398-14399, 14413.
- [19] 陈宇炜, 陈开宁, 胡耀辉. 浮游植物叶绿素 a 测定的“热乙醇法”及其测定误差的探讨[J]. 湖泊科学, 2006, 18(5): 550-552.
- [20] Harper, W. (2016) Reduced Major Axis Regression. In: *Wiley StatsRef: Statistics Reference Online*, John Wiley & Sons, Hoboken, 1-6. <https://doi.org/10.1002/9781118445112.stat07912>