引文格式：敦力民,尹海丹,马嵩. 基于哨兵 2 号时序影像的水华提取[J]. 地理空间信息, 2024, 22(1): 44-48. doi:10. 3969/j.issn. 1672-4623.2024.01.010

基于哨兵2号时序影像的水华提取

敦力民[1]，尹海丹[2]，马　嵩[3]
（1. 沈阳市勘察测绘研究院有限公司，辽宁 沈阳 110004；2. 四川省国土空间规划研究院，四川 成都 610081；
3. 广东省深圳生态环境监测中心站，广东 深圳 518038）

摘　要：水华现象给人类生活和社会发展造成了负面影响，而目前设计的水华指数较难应用于哨兵2号影像。以哨兵2号遥感影像为数据源，通过分析水华、水体、水田等地物的时序光谱特征，提出从时序特征的角度区分水体与水田，再利用多元逐步线性回归的思想进行水华提取的方法；并与常用的浮游藻类指数（FAI）进行了对比实验。结果表明，该方法的提取精度优于FAI，总体精度达到90%以上，Kappa系数达到0.8以上，并能有效减少水田、水生植物、薄云等虚警，具有良好的提取精度与普适性。

关键词：哨兵2号；水华提取；水华指数
中图分类号：P237　　**文献标志码：**B　　**文章编号：**1672-4623（2024）01-0044-05

Algae Bloom Extraction Based on Sentinel-2 Time-series Images

DUN Limin[1], YIN Haidan[2], MA Song[3]
(1. Shenyang Geotechnical Investigation & Surveying Research Institute Co., Ltd., Shenyang 110004, China;
2. Sichuan Institute of Territorial and Spatial Planning, Chengdu 610081, China;
3. Guangdong Shenzhen Ecological Environment Monitoring Center Station, Shenzhen 518038, China)

Abstract: The phenomenon of algae bloom has caused negative impact on human life and social development, and the currently designed algae bloom indexes are difficult to be applied to Sentinel-2 images. Based on Sentinel-2 images, we analyzed the time-series spectral features of algae bloom and other features, and proposed a new algae bloom extraction method, which distinguished water bodies and water fields from the perspective of time-series features, and then used the idea of multivariate stepwise linear regression to extract algae bloom. We conducted a comparison experiment between this method and FAI in the study area. The results show that the extraction accuracy of method in this paper is better than FAI, the overall accuracy reaches more than 90%, and the Kappa coefficient reaches more than 0.8, which can prove the good extraction accuracy and universality of this method.

Key words: Sentinel-2, algae bloom extraction, algae bloom index

湖泊作为水资源的一部分，既能调节气候、维持生态多样性，又能涵养水源，对内陆生态系统的重要性不言而喻[1]。然而，全球变暖、人为活动的不断加剧以及社会经济的快速发展导致我国乃至世界范围内的生态环境恶化，尤其是富营养化引起的水华问题，已成为社会广泛关注的焦点[2]。遥感技术可迅速、高效获得多方位、多源化的对地观测资料，可监测水体中水华现象的发生和发展情况，还可获取水华的空间分布范围和爆发强度等信息，对于掌握内陆水体的水华现象具有重要意义[3-4]。目前利用遥感监测水华现象的方法主要包括影像分类法[5]、水质参数反演法[6]和多波段光谱指数模型[7]，其中多波段光谱指数模型具有耗时少、易应用的特点，被广泛应用于水华提取中。水华与植被的光谱特性相近，因此很多学者利用植被指数建立水华提取模型，如Rouse J W[8]提出的归一化植被指数（NDVI）；但已有研究显示，NDVI易受大气、薄云等干扰，且对浑浊水质也十分敏感[9]。Hu C M[7]提出了浮游藻类指数（FAI）模型用以检测漂浮藻类。Xing Q G[10]等提出了一种基于虚拟基线高度的漂浮藻类遥感识别指数VB_FAH，并应用于Landsat、HJ-1等卫星影像中。Fang C[11]等对FAI进行了改进，建立了AFAI，并采用该指数对1983—2016年的呼伦湖藻华进行了监测。然而，上述水华提取指数通常是针对MODIS、Landsat遥感影像设计的，无法精确监测中小型水体的水华现象，且由于各卫星影像设置的波段中心波长不同，针对其他卫星影像的水华提取指数

收稿日期：2022-08-04。
项目来源：国家自然科学基金资助项目（41971295）。
第一作者简介：敦力民（1970—），硕士研究生，正高级工程师，主要从事城市勘测生产工作，E-mail：dunlmi@sina.com。

波段选取有待进一步探讨。哨兵2号（Sentinel-2）在时间、空间和光谱上都有很大的优势，且设计了3个红边波段，在监测植被信息或水华现象等方面具有很大潜力。目前以Sentinel-2的MSI数据为基础的水华监测研究还较少，因此本文针对Sentinel-2影像，提出了一种内陆水华提取方法，旨在为水华遥感监测提供技术支撑和案例参考。

1　研究区概况

我国拥有湖泊、水库、河流以及人工湖等多种类型的内陆水体，为保证水华提取方法的普适性和稳健性，本文选取太湖、巢湖、滇池、汤逊湖、星云湖、潼湖和华侨城人工湖7个典型内陆水体作为研究对象（表1），上述水体均存在水华等水环境污染问题，必须开展有效的遥感监测，以辅助相关部门的治理[12-15]。

表1　研究区信息

湖泊	地理位置	湖泊大小/km²
太湖	江苏省苏州市	2 250
巢湖	安徽省巢湖市	780
滇池	云南省昆明市	298
汤逊湖	湖北省武汉市	47.6
星云湖	云南省玉溪市	34.71
潼湖	广东省惠州市	17.9
华侨城人工湖	广东省深圳市	4

2　数据来源与研究方法

2.1　数据来源与预处理

本文采用Sentinel-2卫星的Level-1C级数据产品监测水华。该卫星是高分辨率多光谱成像卫星，包含2A和2B两颗卫星，重访周期为5 d，并携带了一枚多光谱仪器，可覆盖13个光谱波段，波长范围400~2 200 nm，空间分辨率分别从10 m、20 m到60 m。本文采用Google Earth Engine平台提供的Sentinel-2 L1C级影像数据，已经过辐射校正和几何校正，因此预处理主要包括大气校正、去云处理和将波段重采样至10 m分辨率。

2.2　研究方法

本文方法包括2个步骤：①通过时序特征，提取水域范围；②通过回归分析，构建水华指数（图1）。

2.2.1　纯净像元选取

本文选取水华、水体、水田的光谱特征和时序特征进行分析，并在太湖、巢湖、滇池、星云湖等影像中，选择其纯净像元。其中，水体与水田像元的样本勾画参考了与影像时间接近的Google Earth高分辨率影像；而水华样本的选取则通过事先查阅相关文献、新闻报道并辅以目视解译的方法，对发生水华现象对应时间点的Sentinel-2影像进行勾画，样本示例见图2。人工选择了3 397个样本，其中水华样本1 080个，水田样本1 260个，水体样本1 057个。由于样本勾画过程中可能受到薄云、雾气等影响，导致部分样本的光谱反射率异常，因此需对初步勾画的样本进行提纯，采用3σ准则剔除异常值。最终得到3 253个样本，其中水华样本1 015个，水田样本1 205个，水体样本1 033个。

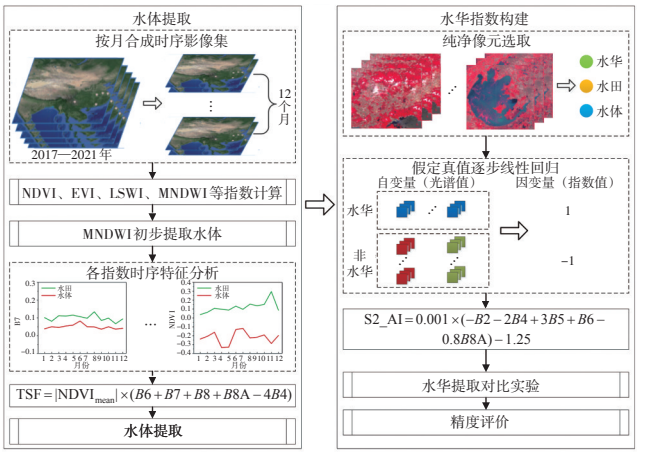


图1　本文方法流程图



a 水华　　b 水体　　c 水田

图2　部分勾画样本示例

2.2.2　时序特征提取水域

水华提取的第一步是提取水域。本文提取的水域是指一段时间内水体出现的范围，同时包括水华的空间范围。根据预处理后的2017—2021年Sentinel-2影像，绘制水体与水田的月平均光谱反射率曲线（图3）发现，在B1~B4波段（中心波长443~665 nm）水体与水田在一年当中的反射率非常接近，且二者光谱曲线随月份的变化不明显，趋势基本一致；在B5~B12波段（中心波长705~2 190 nm），水田的反射率均高于水体，且在7—9月差异最大，水田随月份的变化较大，夏秋季波动较大，而水体的反射率较平稳。另外，计算NDVI的月平均光谱曲线发现，水田的NDVI高于水体。因此，本文采用单因素方差分析解译时序特征，更直观地提取水田与水体地物反射率差异更大的波段，并基于此建立了一种时序特征（TSF）用于剔除被误分为水体的水田像素。

$$TSF = \left|NDVI_{mean}\right| \times (B6 + B7 + B8 + B8A - 4B4) \quad (1)$$

式中，$|NDVI_{mean}|$为12景影像中NDVI平均值的绝对值；*B*4、*B*6、*B*7、*B*8、*B*8A分别为基于12景影像均值合成影像的相应波段反射率。

本文选取适当的TSF阈值区分水田与水体，并基于最大类间方差法（OTSU）的自适应阈值比较仅使用MNDWI与加入TSF特征后的水体提取效果（图4），初步实验表明，后者能更完整地提取水体区域，较好地区分靠近水体的水田、农田等区域，保留了河道。

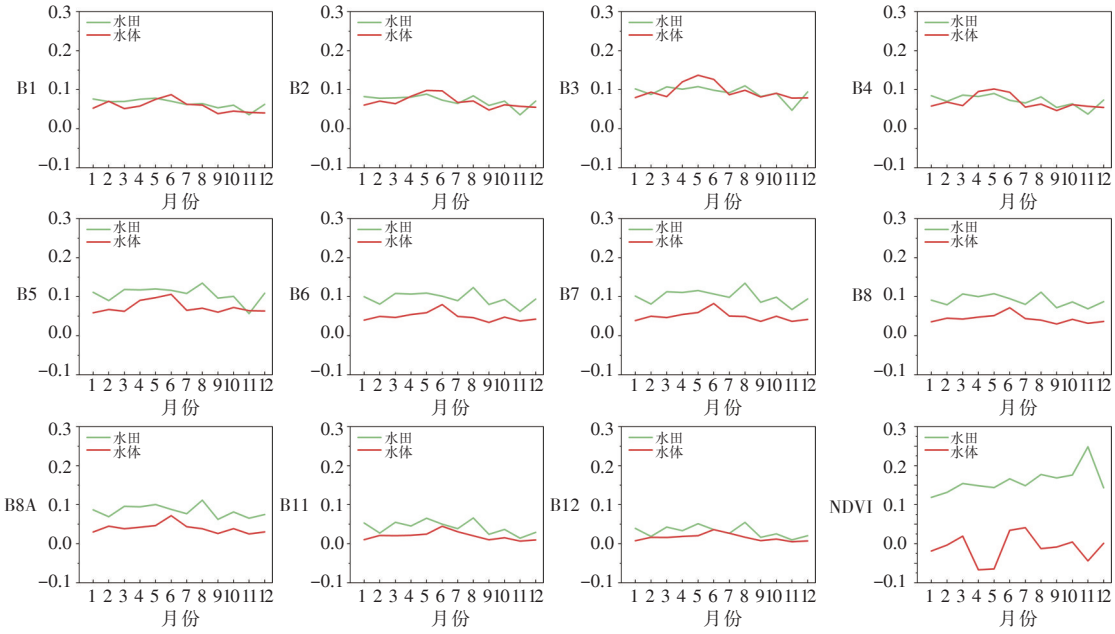
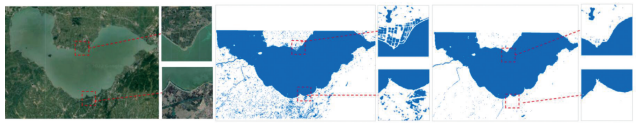


图3　水体与水田各波段月平均光谱反射率曲线



a 真彩色影像　b MNDWI水体提取图　c MNDWI+TSF水体提取图

图4　水体提取效果对比

2.2.3　Sentinel-2水华指数

目前主流的FAI大多是基于MODIS、Landsat的光谱波段特征而设计的，很难适用于其他卫星影像。Sentinel-2包含3个红边波段，中心波长分别为705 nm、740 nm和783 nm，在监测植被信息或水华现象等方面具有很大潜力。本文根据勾画的纯净像元，绘制了3类样本的平均光谱反射率（图5），可以看出，水华的光谱特征曲线与典型植被较相似，在近红外波段的反射率急剧升高，因此可基于该特征较好地区分水华与周围的纯净水体。具体表现为：在绿光波段处有一个小的反射峰，在红光波段附近有一个吸收谷，在红边1波段后，反射率陡然升高，在红边至近红外范围内形成高反射峰，水华现象越严重，该处的峰值越高，从短波红外波段开始，反射率下降，并趋近于0；水田的趋势虽与水华类似，在绿光、红边波段和近红外波段存在反射峰，但在红边和近红外波段的反射率峰值远低于水华，相对于水体而言，水田在红光、近红外和红边波段都表现出较高的光谱反射率；水体光谱曲线的反射峰处于绿光波段，之后反射率随波长的增加而下降，在短波红外波段基本为0，相对于水田和水华，水体在红边—近红外—短波红外波段的反射率较小。

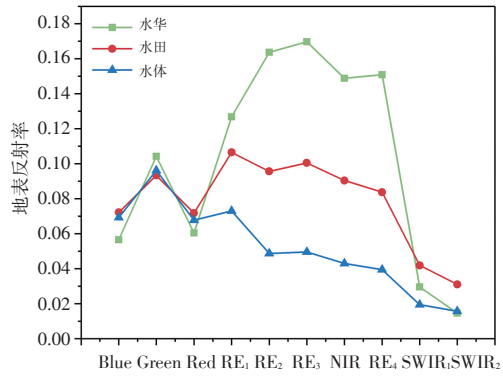


图5　水华、水田、水体的光谱反射曲线

根据水华在Sentinel-2的独特光谱特征，采用多元逐步线性回归的思想筛选合适的波段，构建Sentinel-2水华指数。模型的自变量为水体、水田和水华样本的光谱反射率，因变量为需要得到的水华指数。假设水华像元的指数值为1，而非水华像元的指数值为-1，采用多元逐步线性回归分析结果以及水华与非水华的光谱特征差异进行微调，保留水华反射率较低的B2和B4波段以及反射率较高的B5、B6和B8A波

段。在TSF提取的水域结果上，进行水华提取。水华指数的计算公式为：

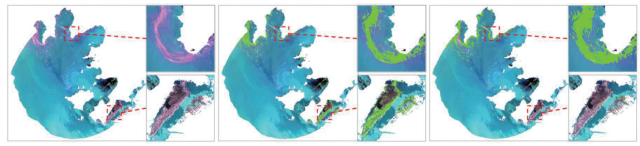
$$S2_AI = 0.001 \times (-B2 - 2B4 + 3B5 + B6 - 0.8B8A) - 1.25 \quad (2)$$

3　结果分析与应用

3.1　基于OTSU自适应阈值的水华提取

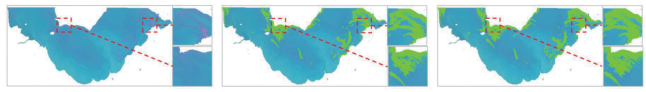
本文采用FAI和上述水华提取方法在太湖、巢湖、滇池3个水体中进行应用和比较。将FAI模型应用于Sentinel-2影像，R_{NIR}参数有5种可能性，R_{SWIR}参数有两种可能性，因此FAI有10种组合方式。由于水华遥感监测是一个长期的过程，需要针对不同时间的多幅Sentinel-2影像进行提取，采用FAI模型或S2_AI模型提取水体水华时，阈值的稳定性对提取精度的影响很大，因此本文利用OTSU方法选取水华的分割阈值，并依据一年内阈值变异系数大小选取了太湖、巢湖和滇池的最优FAI模型，即太湖（B4/B8/B11）、巢湖（B4/B5/B12）、滇池（B4/B6/B11）。水华提取结果见图6~8，可以发现，在太湖的提取结果中，FAI（B4/B8/B11）指数和本文方法（TSF+S2_AI）均能较好地识别水华，在空间分布上较相似，但通过放大对比局部细节可知，FAI（B4/B8/B11）指数不能较好地提取水华的边缘信息，且在太湖的东南部沿岸区域有大量水田和水生植物，其光谱特征与水华相似，因此以往研究中常被误分为水华，而本文方法可以很好地规避这一错误；在巢湖的提取结果中，FAI（B4/B5/B12）指数和本文方法均能较好地识别水华，且与目视解译结果较一致，但通过放大对比局部细节可知，两种方法的主要区别在于对发生水华现象的边缘像素的判定，FAI（B4/B5/B12）指数提取的水华会忽视掉一些边缘信息，而本文方法提取的水华信息包含的范围更广；在滇池的提取结果中，FAI（B4/B6/B11）指数和本文方法的差异较大，具体表现在FAI（B4/B6/B11）指数容易漏分，仅能提取水华现象较明显的像素，而本文方法提取结果与目视解译结果基本一致，导致FAI漏检的原因可能是该指数最初是针对沿海区域提出的，不适用于滇池这样的高原湖泊。相较而言，本文提出的S2_AI指数是综合多个水体发生水华现象的样本而设计的，更适用于滇池、星云湖等高原湖泊的水华遥感监测。

综上所述，基于OTSU自适应阈值的FAI水华提取，在太湖、巢湖的提取结果与目视解译结果基本一致，但存在易漏检发生水华现象的边缘像素和易将湖区的水田误分为水华两个问题，而本文方法的提取结果明显更贴近真实情况，且水华边界与水体的区分都较好，边缘信息较丰富；而在滇池，FAI与真实水华情况存在较大差异，仅能提取在假彩色影像上表现为亮红色的水华像素，而本文方法提取结果较全面。



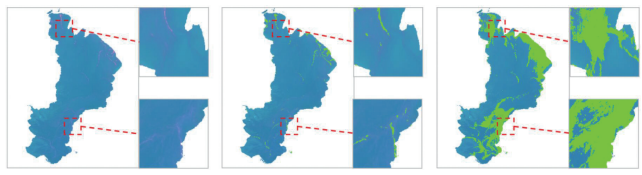
a 假彩色影像　b FAI　c TSF+S2_AI

图6　太湖水华提取结果对比



a 假彩色影像　b FAI　c TSF+S2_AI

图7　巢湖水华提取结果对比



a 假彩色影像　b FAI　c TSF+S2_AI

图8　滇池水华提取结果对比

3.2　定量精度评价

为了更好地定量比较两种水华提取方法，本文选取太湖、巢湖、滇池、汤逊湖、星云湖、滇湖6个水体的遥感影像，以目视解译勾画验证样本，通过总体精度、生产者精度、用户精度和Kappa系数等指标进行精度评估。6个研究区的精度评价结果见表2，可以看出，本文方法在使用OTSU自适应阈值的情况下，总体精度均达到了90%以上，Kappa系数均达到了0.8以上；FAI在使用OTSU自适应阈值的情况下，在巢湖、汤逊湖、星云湖、滇湖的总体精度达到了90%以上，Kappa系数达到了0.7以上，但在太湖、滇池的总体精度、Kappa系数偏低；本文方法的生产者精度均在85%以上，漏检现象较少，用户精度也均在90%以上，说明误分现象也较少；FAI在滇池的生产者精度仅为10.7%，结合图8可知，其漏检现象较严重；FAI的用户精度在太湖、汤逊湖较低，误检率较高，结合图6可知，导致该结果的原因是FAI将太湖东南部的大量水田和水生植物误分为水华，将汤逊湖影像中的薄云误分为水华，说明在复杂的水域环境中，采用传统的水华提取方法难以准确提取水体中的水华。总的来说，在所有研究区中，本文方法的提取结果在各方面均优于FAI，说明本文方法可以更有效、更准确地提取水华。

表2 精度评价结果

研究水体	提取方法	总体精度/%	生产者精度/%	用户精度/%	Kappa系数
太湖	FAI(B4/B8/B11)	84.8	90.3	81.6	0.696
2021-06-22	TSF+S2_AI	95.7	90.5	92.7	0.912
巢湖	FAI(B4/B5/B12)	95.6	94.2	97.0	0.913
2021-08-19	TSF+S2_AI	94.7	91.3	97.9	0.893
滇池	FAI(B4/B6/B11)	46.8	10.7	98.3	0.088
2018-11-28	TSF+S2_AI	92.9	98.7	90.2	0.848
汤逊湖	FAI(B4/B8A/B11)	92.5	98.5	87.8	0.849
2018-08-28	TSF+S2_AI	98.5	97.0	95.6	0.970
星云湖	FAI(B4/B8A/B11)	90.0	94.3	90.4	0.777
2018-11-23	TSF+S2_AI	91.7	87.1	98.7	0.829
潼湖	FAI(B4/B8A/B11)	90.0	93.8	88.2	0.798
2020-03-15	TSF+S2_AI	92.9	93.8	93.8	0.854

4 结 语

本文根据Sentinel-2影像数据的波段特点，通过分析水华和各类典型地物纯净样本的光谱特征和时序特征差异，提出了一种水华提取方法，并与FAI进行了比较。

1）水体与水田的光谱反射率在红边和近红外波段差异明显，水田的反射率均高于水体，且在夏季差异最大，水田的NDVI在该阶段也高于水体。因此，可从时序特征的角度剔除被误分为水体的水田像素，解决现有水华提取指数易错分水田、水生植物、薄云等地物的问题。

2）本文方法在各研究区的总体精度均达到90%以上，提取结果准确，误检率和漏检率均较低。对比本文方法在太湖、巢湖、滇池、汤逊湖、星云湖和潼湖6个研究区的提取结果发现，总体精度、生产者精度等均较高。

与经典的水华提取指数相比，本文方法（TSF+S2_AI）精度更高、普适性更强，可用于多种类型的内陆水体水华提取研究。不足之处在于，对于某一区域频繁发生水华现象的水体，该方法利用时序特征剔除水田等地物时可能会将部分水体剔除，因此需根据实际情况来调整用于TSF提取的月度影像集合。

参考文献

[1] Guan Q, Feng L, Hou X, et al. Eutrophication Changes in Fifty Large Lakes on the Yangtze Plain of China Derived from MERIS and OLCI Observations[J]. Remote Sensing of Environment,2020,246:111 890

[2] 曹畅,王胜蕾,李俊生,等. 基于MODIS数据的全国144个重点湖库营养状态监测:以2018年夏季为例[J]. 湖泊科学,2021,33(2):405-413

[3] 黄露,王爱华,陈君,等. 国产卫星遥感技术在自然资源调查监测中的应用[J]. 地理空间信息,2020,18(5):73-75

[4] 经波,宁文怡. 基于阈值分割与决策树的SAR影像水体信息提取[J]. 地理空间信息,2021,19(3):46-49

[5] 李瑶,张立福,黄长平,等. 基于MODIS植被指数时间谱的太湖2001—2013年蓝藻爆发监测[J]. 光谱学与光谱分析,2016,36(5):1 406-1 411

[6] Chiswell S M, Safi K A, Sander S G, et al. Exploring Mechanisms for Spring Bloom Evolution: Contrasting 2008 and 2012 Blooms in the Southwest Pacific Ocean[J]. Journal of PlanktonResearch,2019,41:329-348

[7] Hu C M. A Novel Ocean Color Index to Detect Floating Algae in the Global Oceans[J]. Remote Sensing of Environment,2009,113(10):2 118-2 129

[8] Rouse J W. Monitoring Vegetation Systems in the Great Plains with Erts[C]. Third ERTS Symposium,1973

[9] 李旭文,牛志春,姜晟. Landsat5 TM遥感影像上太湖蓝藻水华反射光谱特征研究[J]. 环境监测管理与技术,2010,22(6):25-31

[10] Xing Q G, Hu C M. Mapping Macroalgal Blooms in the Yellow Sea and East China Sea Using HJ-1 and Landsat Data: Application of a Virtual Baseline Reflectance Height Technique[J]. Remote Sensing of Environment,2016,178:113-126

[11] Fang C, Song K S, Shang Y X, et al. Remote Sensing of Harmful Algal Blooms Variability for Lake Hulun Using Adjusted FAI (AFAI) Algorithm [J]. Journal of Environmental Informatics,2018,36:54-68

[12] 洪恬林,李云梅,吕恒,等. 基于MODIS数据的太湖浮游植物物候变化及其对水表温度的响应[J]. 地球信息科学学报,2020,22(10):1 935-1 945

[13] 孙雷. 2014—2018年巢湖水质氮磷含量与蓝藻水华变化情况探究[J]. 环境与发展,2021,33(2):153-156

[14] 谢国清,李蒙,鲁韦坤,等. 滇池蓝藻水华光谱特征、遥感识别及暴发气象条件[J]. 湖泊科学,2010,22(3):327-336

[15] 朱军,李益敏,余艳红. 基于GIS的高原湖泊流域生态安全格局构建及优化研究:以星云湖流域为例[J]. 长江流域资源与环境,2017,26(8):1 237-1 250