

呼伦湖 2002–2021 年悬浮物浓度数据集

李含含^{1,2}, 许金朵^{2,3}, 隗晓琪^{2,4}, 黄泽晖^{2,4}, 马荣华^{2,3,5*}

ISSN 2096-2223

CN 11-6035/N



文献 CSTR:

32001.14.11-6035.csd.2023.0093.zh



文献 DOI:

10.11922/11-6035.csd.2023.0093.zh

数据 DOI:

10.57760/sciencedb.07504

文献分类: 地球科学

收稿日期: 2023-04-20

开放同评: 2023-05-05

录用日期: 2023-12-01

发表日期: 2023-12-26

1. 南京信息工程大学, 遥感与测绘工程学院, 南京 210044
2. 中国科学院南京地理与湖泊研究所, 湖泊与环境国家重点实验室, 南京 210008
3. 国家地球系统科学数据中心, 湖泊-流域分中心, 南京 210008
4. 中国科学院大学, 北京 100049
5. 草原生态安全省部共建协同创新中心, 呼和浩特 010021

摘要: 悬浮物浓度 (Suspended particular matter, SPM) 作为评价水质参数之一, 监测湖泊水体悬浮物长期变化对湖泊管理和生态恢复具有重要价值。呼伦湖作为我国北方的重要湖泊, 是“一湖两海”的重要部分, 监测其水质变化对北方生态安全具有重要意义。通过对 2002–2021 年 MODIS Aqua 数据进行处理, 去除水汽和臭氧吸收及瑞利散射的影响, 获得瑞利校正反射率 (Rayleigh-corrected reflectance, R_{rc}), 利用构建的机器学习模型获得了呼伦湖 2002–2021 年的月均和年均悬浮物浓度数据集。经验证, 该模型平均百分比误差为 16.5%, 满足湖泊悬浮物长期观测的实际需求。本数据集可为呼伦湖的生态环境变化监测和水质环境管理提供理论支撑和实践参考。

关键词: MODIS; 悬浮物浓度; 呼伦湖

数据库 (集) 基本信息简介

数据库 (集) 名称	呼伦湖 2002–2021 年悬浮物浓度数据集
数据通讯作者	马荣华 (rhma@niglas.ac.cn)
数据作者	李含含, 许金朵, 隗晓琪, 黄泽晖, 马荣华
数据时间范围	2002–2021
地理区域	呼伦湖
空间分辨率	250 m
数据量	17.78 MB
数据格式	*.tif
数据服务系统网址	https://doi.org/10.57760/sciencedb.07504
基金项目	中国科学院“十四五”网络安全和信息化专项 (CAS-WX2022SDC-SJ05, CAS-WX2021SF-0306)。
数据库 (集) 组成	该数据集由 2 部分组成, 数据格式为 TIFF 格式, 分别为年均和月均悬浮物浓度影像。第 1 部分是年均悬浮物产品, 共 20 个数据, 为呼伦湖 2002–2021 年的年均悬浮物浓度影像, 数据大小为 2.62 MB; 第二部分为逐年月均产品, 共 118 个数据, 为呼伦湖 2002 年 7 月至 2021 年 10 月的呼伦湖月均悬浮物浓度影像, 数据大小为 15.16 MB。

* 论文通信作者

马荣华: rhma@niglas.ac.cn

引言

湖泊是地表极其重要的水资源，是人类活动与发展的热点区域，为流域工农业发展和人类生活提供稳定而清洁的水源^[1]，监测湖泊水质变化尤为重要。习近平总书记十分关心内蒙古“一湖两海”（呼伦湖、乌梁素海、岱海）的生态综合治理，而湖泊水质的变化监测为解决“一湖两海”的水质改善和环境质量提高等问题提供数据支撑和科学依据。呼伦湖是内蒙古第一大湖泊，湖泊面积约为 2339 km²，同时也是内陆同纬度地带最大的草原型湖泊^[2]，作为我国北方寒冷干旱地区内陆湖泊的典型代表，在涵养水源、调节气候、防风固沙及维系呼伦贝尔草原生态系统平衡乃至我国北方生态安全屏障等方面发挥着不可替代的作用^[3]。

悬浮物浓度作为水色参数三要素之一，影响着水体的光学特性和物质交换，也是环境监测的一项重要指标，在一定程度上能综合反映水体的水质特征和水体化学元素迁移、转化和归宿的特征和规律^[4]。高悬浮物浓度会降低水体透明度，限制光在水体中的传输，限制水体浮游生物和沉水植被的生长，影响初级生产力，从而影响湖泊生态环境的变化^[5]。因此，对悬浮物浓度进行动态监测具有十分重要的意义。

传统的人工巡测由于时空尺度差且耗费巨大，难以在悬浮物浓度时空监测上发挥有效作用。卫星遥感具有高频率、大范围观测等特点，能够提供同一时刻区域面状水域的遥感影像，因此在水环境监测中有其特有的优势^[6]。目前遥感估算悬浮物浓度的算法主要分为两类：经验模型与半分析模型，但针对不同水体类型的湖泊尚未有统一的可用模型^[5]。曹志刚等（2016）^[7]利用多波段组合建立模型，获得了洪泽湖悬浮物浓度的长期变化；王卷乐等（2016）^[8]利用 MODIS（Moderate-resolution Imaging Spectroradiometer）数据的多种光谱指数，逐季建立光谱指数与悬浮物浓度的回归关系，继而选择悬浮物浓度的反演模型，反演了鄱阳湖连续多年、季相尺度的悬浮物浓度。近年来，随着计算机科学和人工智能技术的发展，机器学习被广泛应用于悬浮物浓度的遥感反演中^[9]。然而，多数研究集中于模型算法的研究，而很少形成长期可供科学界参考和使用的公开数据集。

呼伦湖悬浮物浓度与气候变化、人类活动、水体富营养化等因素相关，悬浮物浓度数据集对呼伦湖大范围水监测与评估、区域研究与生态保护等具有重要意义。本研究利用机器学习模型实现呼伦湖悬浮物浓度的高精度反演，以此为基础，构建长时间、逐年月相尺度的悬浮物数据集。本数据集为掌握呼伦湖悬浮物浓度的总体时空分布、演变特征和驱动力因素提供基础，为我国生态环境调查评估、区域环境管理等提供基础数据。

1 数据采集和处理方法

1.1 实测数据采集和处理

呼伦湖地理坐标介于 116° 58′ –117° 48′ E, 48° 33′ –49° 20′ N 之间，是内蒙古第一大湖，湖面呈不规则斜长方形，其湖水主要来源于海拉尔河、克鲁伦河、东部乌尔逊河和东北部新开河。本研究于 2016 年 8 月采集了均匀分布在呼伦湖的 29 个样点（图 1），样点位置及实测 SPM 如表 1。每个样点使用棕色采样瓶（2 L）采集表层至 50 cm 水柱内的混合水样，放到装有冰块的保温箱遮光保存，并于当天做过滤处理，后置于冰箱中保存；同时记录测量时的天气、水体透明度、风速、风向、水质状况、采样时间等辅助数据，并使用 GPS 记录采样位置^[5]。悬浮物浓度的测量使用“称重法”，

利用 450℃ 烘烤过后的 47 mm 的 GF/F 膜过滤水样,过滤后的膜放至烘箱 105°烘烤 4 小时,用 0.0001 g 精度的天平称重计计算得到总悬浮物浓度^[7]。

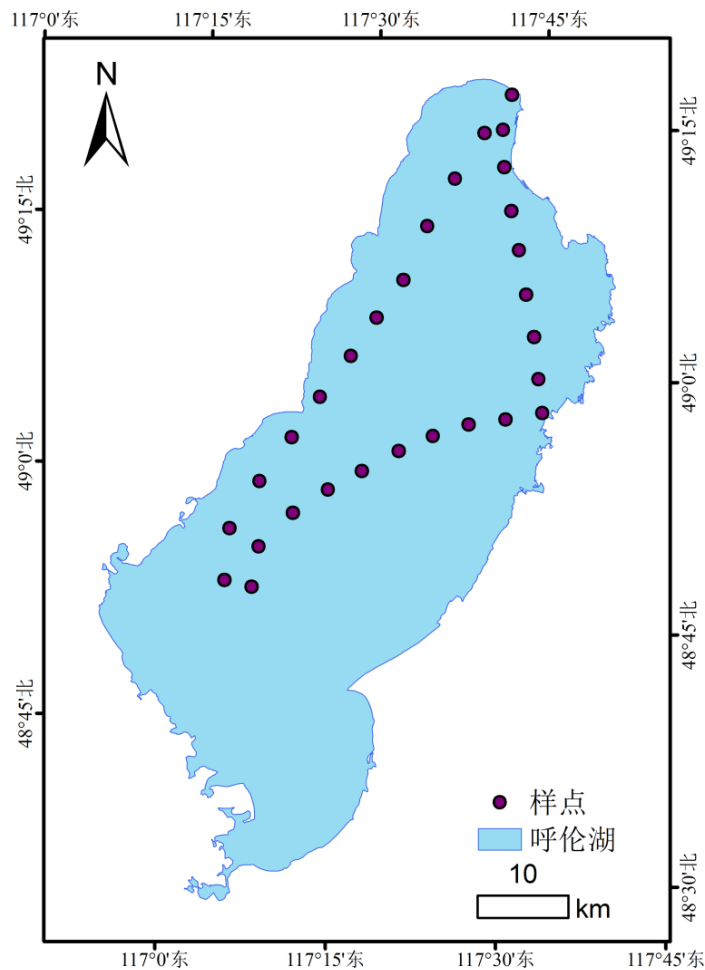


图 1 呼伦湖采样点分布

Figure 1 Distribution of sampling sites in Hulun Lake

表 1 实测点位置及 SPM 浓度

Table 1 The location of the measured points and SPM concentrations

点位号	点位位置		实测 SPM(mg/L)
	经度 (东经)	纬度 (北纬)	
HL01	117.683924	49.306249	28.00
HL02	117.663827	49.27241	28.60
HL03	117.658466	49.235129	29.00
HL04	117.660617	49.190701	29.43
HL05	117.664523	49.150874	30.67
HL06	117.66679	49.105926	31.20
HL07	117.67025	49.06270	32.67
HL08	117.669031	49.02033	33.24
HL09	117.668386	48.986535	34.20

点位号	点位位置		实测 SPM(mg/L)
	经度 (东经)	纬度 (北纬)	
HL10	117.612453	48.984535	34.67
HL11	117.557067	48.984302	35.20
HL12	117.502088	48.977432	36.00
HL13	117.448681	48.966723	38.67
HL14	117.390646	48.951489	40.00
HL15	117.33725	48.937446	43.33
HL16	117.280812	48.918515	50.00
HL17	117.224311	48.889536	53.50
HL18	117.16781	48.860557	57.00
HL19	117.206402	48.850396	58.00
HL20	117.184595	48.911131	60.00
HL21	117.237494	48.954185	62.00
HL22	117.293301	48.993867	64.80
HL23	117.342635	49.030473	66.00
HL24	117.395965	49.067154	68.50
HL25	117.441167	49.101951	70.50
HL26	117.487968	49.136072	72.00
HL27	117.533441	49.186511	77.33
HL28	117.583261	49.230026	78.67
HL29	117.635978	49.271504	81.00

1.2 遥感数据下载与处理

1.2.1 MODIS 数据下载

以 MODIS (Terra: 1999 年至今; Aqua: 2002 年至今) 为代表的中分辨率海洋水色传感器被广泛应用于水质及物候监测^[1]。MODIS 具有 36 个光谱通道, 分布在 0.4–14 μm 的电磁波谱内, 其空间分辨率为 1–2 波段为 250 m、3–7 波段为 500 m 和 8–36 波段为 1000 m, 可以实现一天覆盖全球一次^[10]。MODIS 的陆地波段 (前 7 个波段, 光谱范围为 443–2130 nm) 具有较高的空间分辨率 (250–500 m) 且在浑浊湖泊不饱和, 故使用这 7 个波段提取悬浮物浓度^[1]。本文选择从美国航空航天局海洋生物学处理组 (Ocean Biology Processing Group, OBPG, <https://oceandata.gsfc.nasa.gov>) 的数据网站下载呼伦湖区域 2002–2021 年的 MODIS Aqua Level-1A (L1A) 数据。呼伦湖每年 11 月至来年 4 月为结冰期, 因此只用 5 月至 10 月的遥感影像来进行处理分析。

1.2.2 数据处理

在 SeaDAS 8.1 软件中对下载的 MODIS L1A 数据进行再处理, 生成 Level-1B 数据, 然后进行处理以估算反射率。由于常用的水体大气校正算法难以用于内陆湖泊, 导致遥感反射率 (R_{rs}) 存在较大不确定性, 因此选择从传感器的辐射度中去除水汽和臭氧的吸收以及瑞利散射, 以生成瑞利校正

的反射率 (R_{rc} , 公式 1) [11]。事实上, R_{rc} 已广泛用于监测湖泊悬浮物浓度[12]。

$$R_{rc}(\lambda) = \frac{\pi L'_{t,\lambda}}{F_0 \times \cos\theta_0} - R_{r,\lambda} \quad (1)$$

其中 $L'_{t,\lambda}$ 是除去水汽和臭氧吸收后的大气层顶部 (TOA) 辐射度, $R_{r,\lambda}$ 是瑞利散射, F_0 是地外太阳辐射度, θ_0 是太阳天顶角[13]。

使用 SeaDAS 8.1 中的内部工具, 将具有相对较高空间分辨率 (250–500 m) 的前 7 个波段 (469–2130 nm) 与具有 500 m 空间分辨率的最后 5 个波段重采样到 250 m。采用掩膜处理来去除低质量的 R_{rc} 像素。

(1) 云掩膜: 在海洋水色处理中, 默认的云掩膜标志 ($R_{rc}(859) > 0.027$) 通常会将浑水等像素误判而排除, 造成大量数据丢失, 这对呼伦湖效果尤为显著。根据对过去研究的阈值测试, 发现 $R_{rc}(2130) > 0.021$ 的阈值适合于研究湖泊^{错误:未找到引用源。}, 但这样特定的阈值仍会掩盖一些蓝藻浮渣和极端浑浊的水域。

(2) 陆地掩膜: 呼伦湖水体变动大, 每景 MODIS 数据根据归一化差异水指数 (Normalized Difference Water Index, NDWI) 确定湖泊边界, 然后利用中国湖泊边界^{错误:未找到引用源。}对提取水体进行进一步筛选。此外还采用浮游藻类指数 (floating-algae index FAI) 的阈值 (即 -0.004) 来去除蓝藻水华像元干扰。为去除陆地邻近效应, 采用了 Feng 等 (2018)^{错误:未找到引用源。}的方法移除近岸的 4 个像素。

1.3 悬浮物浓度反演模型构建

考虑到 SPM 存在较大的时空变异性, 呼伦湖的实测数据难以独立构建模型, 因此本文使用 Cao 等 (2023) [17] 的机器学习模型从 MODIS Aqua 获取呼伦湖的悬浮物浓度。同时, 我们使用前文所述的 29 个实测点对该模型在呼伦湖反演的结果进行验证 (详见节 3 数据质量控制和评估)。简单而言, Cao 等 (2023) 的模型基于支持向量机所开发, 使用全国 133 个湖泊采集的 903 个样点所构建。模型输入变量包括了 $R_{rc}(469)$, $R_{rc}(555)$, $R_{rc}(645)$, $R_{rc}(859)$, $R_{rc}(1240)$ 和三个波段比率 ($R_{rc}(555) / R_{rc}(469)$, $R_{rc}(645) / R_{rc}(555)$, $R_{rc}(859) / R_{rc}(555)$), 输出变量为 SPM 浓度。在模型训练前, 输入变量进行标准化, 并归一化到 0–1 范围; 悬浮物浓度亦进行标准化并使用 log 函数进行转换。采用网格搜索法来确定支持向量机的参数, 最终获得悬浮物反演模型。模型的主要参数如表 2 所示。更多的细节参考 Cao 等 (2023) [17]。

表 2 用于估算 SPM 的支持向量机回归模型参数

Table 2 Hyperparameters for the support vector regression model to estimate SPM

参数	调整范围	最佳值
kernel	linear, rbf, sigmoid, poly	rbf
epsilon	0.01–1.1 间隔为 0.01	0.03
C	1-10 间隔为 0.1	3.8
gamma	scale, auto	scale
tol	$1e^{-5}$, $1e^{-4}$, $1e^{-3}$, $1e^{-2}$	0.0001

注: kernel 为核函数、epsilon 被用来定义模型对误差的容忍度、C 为错误项的惩罚因子、gamma 为核函数系数、tol 为残差收敛条件。

2 数据样本描述

基于上述 SPM 反演模型完成 2002–2021 年间的呼伦湖悬浮物浓度反演，形成年均浓度分布数据集（图 2）和月均浓度分布数据集（图 3）。总体上，呼伦湖的悬浮物浓度较高（平均值 53.6 mg/L，标准差为 7.95 mg/L），表明呼伦湖水体较为浑浊。从空间分布上看，呼伦湖悬浮物浓度西南部比东北部变化明显，最大值、最小值均在西南部。在人类活动、社会因素和自然因素的影响下，在 2002 年至 2021 年期间，悬浮物浓度整体呈先上升后下降的趋势，其中 2006–2013 年较高，自 2014 年逐步下降。2006 年为补充呼伦湖水量，从海拉尔河引水，外源的输入增加，因此 2006–2013 年悬浮物浓度维持较高状态。2014 年以来悬浮物浓度下降与风速的下降和流域的植被恢复有关。

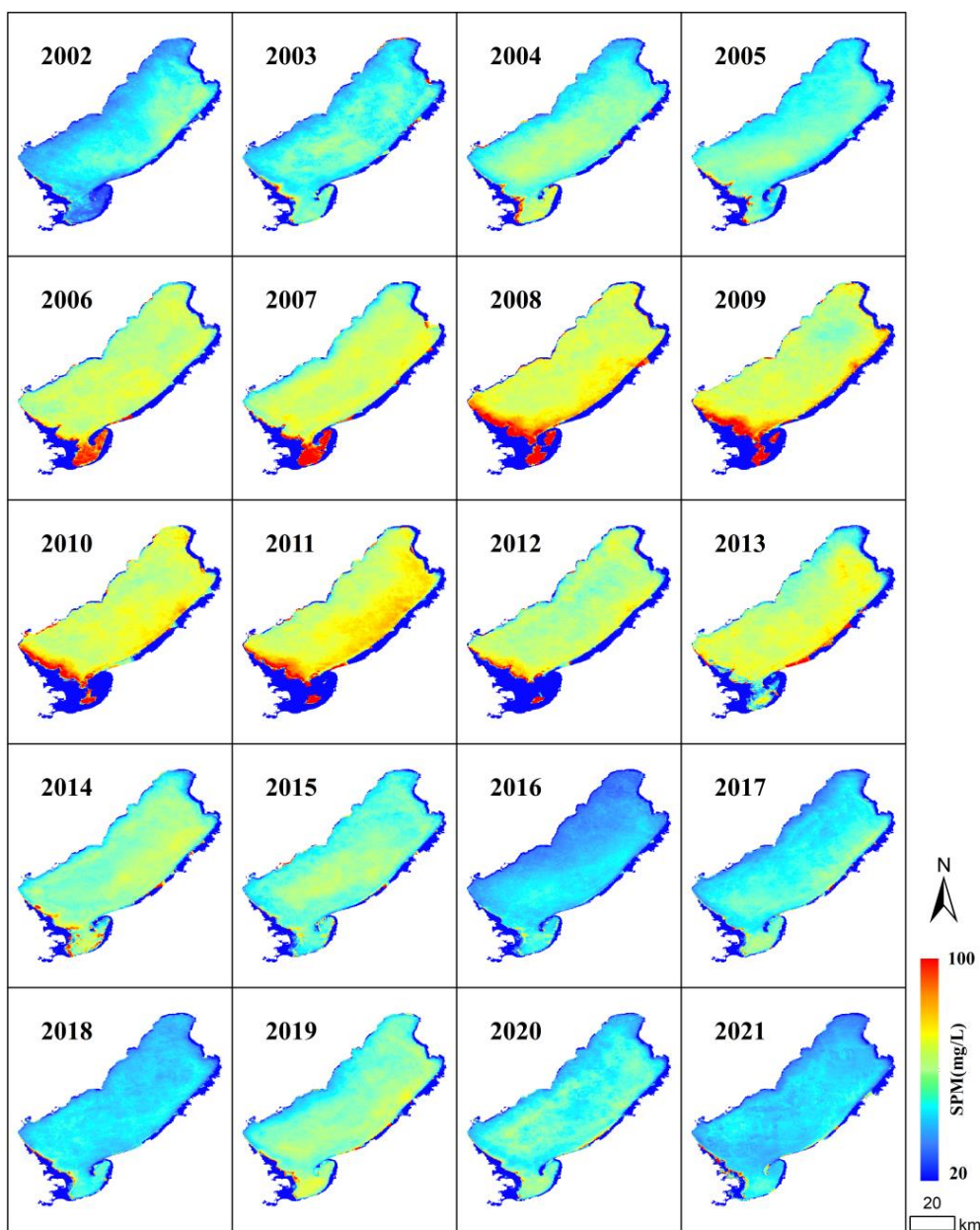


图 2 2002–2021 年呼伦湖年均悬浮物浓度空间分布

Figure 2 Spatial distribution of annual mean suspended matter concentration in Hulun Lake from 2002 to 2021

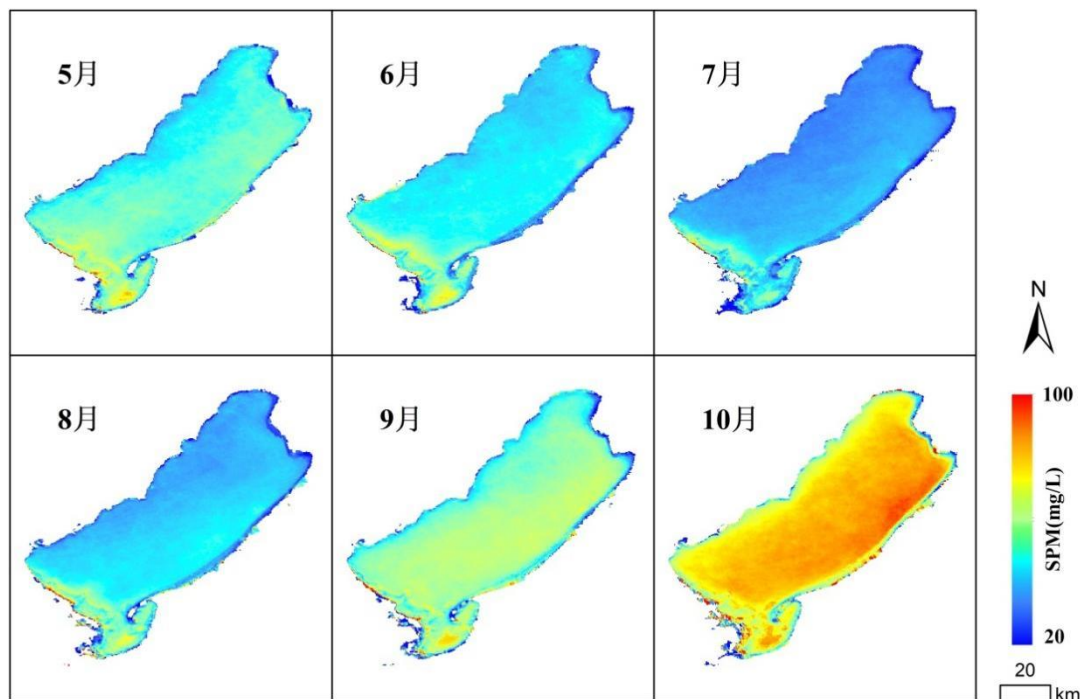


图 3 2002–2021 年呼伦湖月均悬浮物浓度空间分布

Figure 3 Spatial distribution of monthly mean suspended matter concentration in Hulun Lake from 2002 to 2021

在季节上，呼伦湖的悬浮物浓度 7 月最低（平均值 37.7 mg/L，标准差为 6.95 mg/L），10 月最高（平均值 74.9 mg/L，标准差为 13.20 mg/L），这是因为呼伦湖为浅水湖泊（平均水深为 5.7m），风速引起的底泥再悬浮作用是悬浮物浓度季节变化的主要原因^[7]。呼伦湖秋季风速较高，因此在 9 月和 10 月具备较高的浓度。

3 数据质量控制和评估

将 MODIS 遥感数据反演的悬浮物浓度与样点实测值进行结果验证，以保证各期模型在遥感数据上应用的有效性。主要采用决定系数 (R^2)、均方根误差 (RMSE) 和平均绝对百分比误差 (MAPE) 验证，计算方法如下。

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - X_i)^2} \quad (2)$$

$$MAPE = 100 \times \frac{1}{N} \sum_{i=1}^N |Y_i - X_i| / X_i \quad (3)$$

其中 N 表示用于计算的数据个数， i 代表数据的下标， X 和 Y 则分别代表用于评价的数据变量。

将 Cao 等 (2023)^[17] 模型应用于呼伦湖获取 SPM 浓度。利用 29 个实测点的位置提取 SPM，与实测 SPM 浓度对比评价模型的精度。结果发现该模型在实测样点的精度较高 ($R^2=0.71$, RMSE 为 9.6 mg/L, MAPE=16.5%) (图 4)，在呼伦湖表现较好。鉴于模型在呼伦湖的相对误差低于 35%，高

于业内对水色产品算法的精度要求^{错误:未找到引用源。}，故可用该模型对呼伦湖悬浮物浓度进行遥感反演估算。虽然本次覆盖了 SPM 浓度 20–80 mg/L 的范围，具有一定代表性；但是，可能未覆盖其他季节的极端情况，因此未来我们也将增加更多呼伦湖实测数据进一步验证或优化模型。

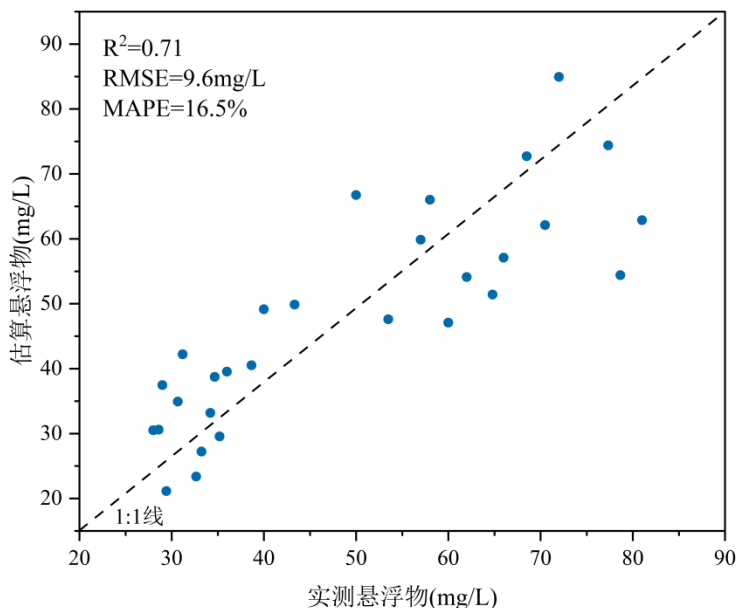


图 4 支持向量机模型精度验证图

Figure 4 Accuracy verification diagram of the support vector machine model

4 数据使用方法和建议

本数据集是基于支持向量机和星地同步数据构建的模型反演结果，其精度满足要求。本数据集实体解压后的 tif 数据可使用专业的 GIS 软件打开和编辑再加工；其时间序列涵盖了 2002–2021 年的呼伦湖悬浮物浓度的时空分布信息，可为呼伦湖大范围水监测和评估、生态保护等提供参考与数据支持。

致 谢

感谢美国航空航天局（NASA）提供 MODIS 数据和处理支持，感谢湖泊–流域数据中心提供实测数据构建模型。

数据作者分工职责

李含含（1997—），女，河南开封市人，硕士生，研究方向为悬浮物浓度遥感监测。主要承担工作：方案设计与实施，论文撰写，数据处理。

许金朵（1982—），女，江苏省睢宁县人，硕士，工程师，研究方向为数据库建设、数据共享与地理信息系统。主要承担工作：数据准备，数据整理与上传。

隗晓琪（2000—），女，陕西省西安市人，硕士生，研究方向为湖泊固有光学特性的遥感反演。主要承担工作：遥感数据预处理，数据产品质量控制。

黄泽晖（1999—），男，河南省洛阳市人，硕士生，研究方向为湖泊蓝藻水华的遥感监测。主要承担工作：实测数据采集与分析处理。

马荣华（1972—），男，山东临沂市人，博士，研究员，研究领域为湖泊环境遥感、湖泊-流域数据共享。主要承担工作：论文修改完善，数据方法指导。

参考文献

- [1] 袁俊, 曹志刚, 马金戈, 等. 1980s 以来巢湖藻华物候时空变化遥感分析[J]. 湖泊科学, 2023, 35(1): 57–72. DOI: 10.18307/2023.0103. [YUAN J, CAO Z G, MA J G, et al. Remote sensed analysis of spatial and temporal variation in phenology of algal blooms in Lake Chaohu since 1980s[J]. Journal of Lake Sciences, 2023, 35(1): 57–72. DOI: 10.18307/2023.0103.]
- [2] 王苏民, 窦鸿身. 中国湖泊志[M]. 北京: 科学出版社, 1998.[WANG S M, DOU H S. China lakes record [M]. Beijing: Science Press Ltd., 1998.]
- [3] 曹秉帅, 徐德琳, 窦华山, 等. 北方寒冷干旱地区内陆湖泊生态安全评价指标体系研究——以呼伦湖为例[J].生态学报, 2021, 41(08): 2996–3006. DOI: 10.5846/stxb202004301059. [CAO B S, XU D L, DOU H S, et al. Index system of ecological security of inland lakes in cold arid region: a case study of Hulun Lake, China[J]. Acta Ecologica Sinica, 2021, 41(8): 2996–3006. DOI: 10.5846/stxb202004301059.]
- [4] 张运林, 秦伯强, 陈伟民, 等. 悬浮物浓度对水下光照和初级生产力的影响[J]. 水科学进展, 2004(05):615–620. DOI: 10.3321/j.issn:1001-6791.2004.05.014. [ZHANG Y L, QIN B Q, CHEN W M, et al. Experimental study on underwater light intensity and primary productivity caused by variation of total suspended matter[J]. Advances in Water Science, 2004(05):615–620. DOI: 10.3321/j.issn:1001-6791.2004.05.014.]
- [5] 曹志刚. 基于 MODIS/Aqua 和 VIIRS/NPP 卫星的悬浮物浓度遥感估算研究——以洪泽湖为例[D]. 北京: 中国科学院大学, 2017.[CAO Z G. Remote sensing of the Suspended Particulate Matter in Lake Hongze: Observation from VIIRS NPP and MODIS Aqua[D]. Beijing: University of Chinese Academy of Sciences, 2017.]
- [6] 李云梅, 黄家柱, 陆皖宁, 等. 基于分析模型的太湖悬浮物浓度遥感监测[J]. 海洋与湖沼, 2006, 37(2): 171–177. DOI: 10.3321/j.issn: 0029-814X.2006.02.011. [LI Y M, HUANG J Z, LU W N, et al. Model-based remote sensing on the concentration of suspended sediments in Taihu Lake[J]. Oceanologia et Limnologia Sinica, 2006, 37(2): 171–177. DOI: 10.3321/j.issn: 0029-814X.2006.02.011.]
- [7] 曹志刚, 段洪涛, 崔海山, 等. 基于 VIIRS 数据洪泽湖悬浮物浓度估算算法研究[J]. 红外与毫米波学报, 2016, 35(4): 462–469.. DOI: 10.11972/j.issn.1001-9014.2016.04.014. [CAO Z G, DUAN H T, CUI H S, et al. Remote estimation of suspended matters concentrations using VIIRS in Lake Hongze, China[J]. Journal of Infrared and Millimeter Waves, 2016, 35(4): 462–469.. DOI: 10.11972/j.issn.1001-9014.2016.04.014.]
- [8] 王卷乐, 陈二洋, 祝俊祥, 等. 鄱阳湖悬浮物浓度 2000~2013 年反演数据集[J/OL].中国科学数据(中英文网络版), 2016(01):29-36. DOI:10.11922/csdata.170.2015.0001. [WANG J L, CHEN E Y, ZHU J X, et al. A dataset of suspended solids concentration inversion for Poyang Lake, China 2000 – 2013[J/OL]. China Scientific Data, 2016(01):29-36. DOI:10.11922/csdata.170.2015.0001.]

- [9] 陈点点, 陈芸芝, 冯险峰, 等. 基于超参数优化 CatBoost 算法的河流悬浮物浓度遥感反演[J]. 地球信息科学学报, 2022, 24(4): 780-791. [CHEN D D, CHEN Y Z, FENG X F, et al. Retrieving suspended matter concentration in rivers based on hyperparameter optimized CatBoost algorithm[J]. Journal of Geo-information Science, 2022, 24(4):780-791.]
- [10] 吕恒, 黄家柱, 江南. 太湖水质参数 MODIS 的遥感定量提取方法[J]. 地球信息科学学报, 2009, 11(01): 104-110. DOI: 10.3969/j.issn.1560-8999.2009.01.016. [LU H, HUANG J Z, JIANG N. Quantitative retrieval of water quality by remote sensing in Taihu Lake based on MODIS data[J]. Journal of Geo-information Science, 2009, 11(01): 104-110. DOI: 10.3969/j.issn.1560-8999.2009.01.016.]
- [11] HU C M, CHEN Z Q, CLAYTON T D, et al. Assessment of estuarine water-quality indicators using MODIS medium-resolution bands: initial results from *Tampa Bay*, FL[J]. Remote Sensing of Environment, 2004, 93(3): 423–441. DOI: 10.1016/j.rse.2004.08.007.
- [12] CAO Z G, DUAN H T, FENG L, et al. Climate- and human-induced changes in suspended particulate matter over Lake Hongze on short and long timescales[J]. Remote Sensing of Environment, 2017, 192: 98–113. DOI: 10.1016/j.rse.2017.02.007.
- [13] SEEGER S B N, WERDELL P J, VANDERMEULEN R A, et al. Satellites for long-term monitoring of inland U.S. lakes: the *MERIS time* series and application for chlorophyll-A[J]. Remote Sensing of Environment, 2021, 266: 112685. DOI: 10.1016/j.rse.2021.112685.
- [14] AURIN D, MANNINO A, FRANZ B. Spatially resolving ocean color and sediment dispersion in river plumes, coastal systems, and continental shelf waters[J]. Remote Sensing of Environment, 2013, 137: 212–225. DOI: 10.1016/j.rse.2013.06.018.
- [15] MA R H, YANG G S, DUAN H T, et al. China's Lakes at present: number, area and spatial distribution[J]. Science China Earth Sciences, 2011, 54(2): 283–289. DOI: 10.1007/s11430-010-4052-6.
- [16] FENG L, HU C M. Land adjacency effects on MODIS Aqua top-of-atmosphere radiance in the shortwave infrared: statistical assessment and correction[J]. Journal of Geophysical Research: Oceans, 2017, 122(6): 4802–4818. DOI: 10.1002/2017jc012874.
- [17] CAO Z G, HU C M, MA R H, et al. MODIS observations reveal decrease in lake suspended particulate matter across China over the past two decades[J]. Remote Sensing of Environment, 2023, 295: 113724. DOI: 10.1016/j.rse.2023.113724.
- [18] Hooker S B. SeaWiFS technical report series: An overview of SeaWiFS and ocean color[R]. 1992.

论文引用格式

李含含, 许金朵, 隗晓琪, 等. 呼伦湖 2002 – 2021 年悬浮物浓度数据集[J/OL]. 中国科学数据, 2023, 8(4). (2023-12-25). DOI: 10.11922/11-6035.csd.2023.0093.zh.

数据引用格式

李含含, 许金朵, 隗晓琪, 等. 呼伦湖 2002 – 2021 年悬浮物浓度数据集[DS/OL]. Science Data Bank, 2023. (2023-05-05). DOI: 10.57760/sciencedb.07504.

A dataset of suspended particulate matter concentration in Hulun Lake from 2002 to 2021

LI Hanhan^{1,2}, XU Jinduo^{2,3}, WEI Xiaoqi^{2,4}, HUANG Zehui^{2,4}, MA Ronghua^{2,3,5*}

1. School of Remote Sensing and Geomatics Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, P.R. China

2. State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, P. R. China

3. Lake-Watershed Science Data Center, National Earth System Science Data Center, Nanjing 210008, P. R. China

4. University of Chinese Academy of Science, Beijing 100049, P.R. China

5. Collaborative Innovation Center for Grassland Ecological Security (Jointly Supported by the Ministry of Education of China and Inner Mongolia Autonomous Region), Hohhot 010021, P. R. China

*Email: rhma@niglas.ac.cn

Abstract: Suspended particular matter (SPM) concentration is one of the parameters in assessing water quality. The monitoring of the long-term variations of suspended particular matter is of great value for lake management and ecological restoration in lake water. As an important lake in the north of China, Hulun Lake is an important part of the "One Lake, Two Seas" initiative. The monitoring of its water quality changes is of great significance to ensuring the ecological security in Northern China. Through the processing of MODIS Aqua data from 2002 to 2021, and by mitigating the impact of water vapor and ozone absorption as well as Rayleigh scattering, we derived Rayleigh-corrected reflectance (R_{rc}), and subsequently generated the dataset of monthly and yearly mean suspended particulate matter concentration in Hulun Lake spanning the years from 2002 to 2021. The average percentage error of the model was verified to be 16.5%, which can meet the practical needs of long-term observation of the suspended particulate matter in lakes. This dataset can provide theoretical support and practical reference for the monitoring of ecological changes and environmental management of water quality in Hulun Lake.

Keywords: MODIS; SPM concentration; Lake Hulun Lake

Dataset Profile

Title	A dataset of suspended particulate matter concentration in Hulun Lake from 2002 to 2021
Data corresponding author	MA Ronghua (rhma@niglas.ac.cn)
Data author	LI Hanhan, XU Jinduo, WEI Xiaoqi, HUANG Zehui, MA Ronghua
Time range	2002–2021
Geographical scope	Hulun Lake
Spatial resolution	250 m
Data volume	17.78 MB
Data format	*.tif
Data service system	< https://doi.org/10.57760/sciencedb.07504 >

Sources of funding	The 14th Five-Year Plan of the Chinese Academy of Sciences for Network Security and Information Technology (CAS-WX2022SDC-SJ05, CAS-WX2021SF-0306).
Dataset composition	The dataset is composed of two parts in TIFF format, which are annual and monthly mean suspended particulate matter concentration images. Part One consists of 20 data entries of the annual mean suspended particulate matter concentration images in Hulun Lake from 2002 to 2021, with a data volume of 2.62MB; Part Two consists of 118 data entries of the monthly mean suspended particulate matter concentration images in Hulun Lake from July 2002 to October 2021, with a data volume of 15.16MB.