专题 “两山”理念

1984–2023 年 9 个母亲湖水体面积时空变化数据集

赵慧子[1,4]，赵凯[2]，曹引[1,3*]，赵红莉[1,3]，蒋云钟[1,3,4]，

张怀文[1]，干雨[1]，华雯静[1]

1. 中国水利水电科学研究院水资源研究所，北京 100038
2. 水利部信息中心，北京 100053
3. 水利部数字孪生流域重点实验室，北京 100038
4. 河海大学，南京 210024

摘要：湖泊作为重要的水资源和生态系统载体，对区域水资源供给、生态环境保障以及区域经济、社会、文化发展都有着重要支撑作用，但受气候变化及人类活动影响，部分湖泊面临水体面积萎缩的威胁。本研究以水利部印发《母亲河复苏行动方案（2022–2025 年）》中 9 个母亲湖（白洋淀、衡水湖、大黄堡洼、团泊洼、七里海、南大港湿地、岱海、沙湖、长湖）作为研究对象，利用 Landsat-5 TM、Landsat-8/9 OLI 和 Sentinel-2 MSI 等多种卫星数据源协同，提取了 9 个母亲湖 1984-2023 年 40 年水体面积，共计 4791 期监测结果。本研究以 9 个母亲湖共 55 期同步 1 m 空间分辨率高分 2 号卫星影像提取水体面积作为相对真值，利用广泛使用的混淆矩阵精度评价方法验证数据集精度。经验证，本数据集总精度为 94.7%，Kappa 系数为 0.86。本数据集可为母亲湖复苏成效定量评估提供数据支撑。

关键词：母亲湖；水体面积；Landsat-5；Landsat-8/9；Sentinel-2

数据库（集）基本信息简介

数据库（集）名称	1984–2023 年 9 个母亲湖水体面积时空变化数据集
数据通信作者	曹引（caoyin@iwhr.com）
数据作者	赵慧子、赵凯、曹引、赵红莉、蒋云钟、张怀文、干雨、华雯静
数据时间范围	1984-2023年
地理区域	白洋淀（38°46’N –38°59’N，115°49’E –116°6’E）、衡水湖（37°32’N –37°39’N，115°33’E –115°39’E）、大黄堡洼（39°24’N –39°32’N，117°10’E –117°17’E）、团泊洼（38°50’N –38°58’N，117°3’E –117°9’E）、七里海（39°16’N –39°19’N，117°27’E –117°38’E）、南大港湿地（38°27’N –38°36’N，117°25’E –117°35’E）、岱海（40°30’N –40°37’N，112°35’E –112°46’E）、沙湖（38°44’N –38°49’N，106°18’E –106°25’E）、长湖（30°20’N –30°32’N，112°10’E –112°32’E）
空间分辨率	Sentinel-2：10 m；Landsat-5/8/9：30 m
数据量	501 MB（解压后）
数据格式	*.shp，*.xlsx

www.sciengine.com/CSD
ISSN 2096-2223
CN 11-6035/N

文献 CSTR：
32001.14.11-6035.noda.2024.0164.zh

文献 DOI：
10.11922/11-6035.noda.2024.0164.zh
数据 DOI：
10.57760/sciencedb.13319

文献分类：地球科学

收稿日期：2025-01-28
录用日期：2025-04-10
发表日期：2025-09-18

* 论文通信作者
曹引：caoyin@iwhr.com

数据服务系统网址	https://doi.org/10.57760/sciencedb.13319
基金项目	国家自然科学基金（52309040、52279031）；国家重点研发计划（2023YFF1304202）。
数据库（集）组成	数据集包含白洋淀、衡水湖、大黄堡洼、团泊洼、七里海、南大港湿地、岱海、沙湖、长湖等9个母亲湖1984–2023年水体面积矢量数据，以及9个母亲湖1984-2023年水体面积统计表。

引　言

湖泊作为重要的水资源和生态系统，承载着水资源供给调节、区域气候调节、生物多样性保护等功能，同时其提供了重要的生态系统服务功能[1]，此外，湖泊对地方经济社会的可持续发展同样至关重要[2]。然而，由于人类活动和气候变化等因素影响加剧，许多湖泊面临着严重的面积减少和生态退化的挑战[3-4]。2022 年起，水利部开展了全面排查确定断流河流、萎缩干涸湖泊修复名录的工作，随后水利部印发《母亲河复苏行动方案（2022—2025 年）》，2023 年确定了 88 条（个）河湖名单，其中包含 79 条母亲河及 9 个母亲湖，并逐步开展对其综合治理、系统治理、源头治理。列入名单的 9 个母亲湖对地区经济、社会、文化发展有着突出作用，也因靠近城市或人类聚集区，受人类活动影响尤为剧烈，历史上出现了严重的水质恶化或水体面积萎缩的问题，在实施了一系列合理的修复措施后，母亲湖均取得了一定的复苏成效，但如何评价母亲湖复苏效果成为行业关注的问题。母亲湖面积变化一定程度上能够直观地表现母亲湖复苏的效果，获取母亲湖长时序水体面积变化趋势对分析母亲湖复苏成效，指导母亲湖复苏行动持续开展具有重要意义。

遥感影像具有长时序、大范围、动态监测等优点，目前已成为湖泊水体面积长时序变化监测的有效手段[5]，内陆湖泊精细监测往往需要采用中高分辨率卫星，然而，基于中高分辨率卫星的长时序、高频次内陆水体监测仍是挑战[6]。一方面受限于轨道、飞行高度、在轨时间等，与低空间分辨率但高时间分辨率的 MODIS 相比，Landsat-8/9、Sentinel-2 等中高分辨率卫星重复监测能力仍有待提高；另一方面光学遥感易受到云覆盖、阴影等影响[7-8]，进一步影响了实际有效监测频次。因此，仅采用单一卫星数据源难以同时获取长时序、高时间和高空间分辨率影像。当前我国湖泊水体面积数据集生产多以 MODIS、Landsat 系列、Sentinel-2 等单一卫星或卫星系列为数据源，基于 MODIS 生产的湖泊水体面积数据集具有较高的时间分辨率[9-10]，但受空间分辨率所限，数据集水体提取精度略有劣势；以 Landsat 系列、Sentinel-2 等中高空间分辨率卫星为数据源生产的水体面积数据集监测频次整体偏低[11-13]。利用多源中高空间分辨率卫星协同开展湖泊水体面积监测，能够有效提高湖泊水体面积监测频次和监测精度，同时延长监测时间范围[14]。近年来，部分学者尝试利用多源卫星协同开展湖泊水体面积监测研究。陈楚楚等[15]基于 Landsat TM、OLI 数据，采用 MNDWI 提取了天津市团泊洼 1995-2020 年间六个年份的水体面积，并分析了团泊洼的水体面积变化；张戈等[16]比较了 Sentinel-2 和 Landsat 两种数据源水面信息提取差异和协同适用性；Nicholas 等[17]使用协同的 Landsat-8/9 和 Sentinel-2（Harmonized Landsat-8/9 Sentinel-2，HLS）数据生产了 2023 年至今的全球范围动态地表水范围产品，永久地表水和季节性地表水产品的平均准确率为 96.1%。

本数据集以《母亲河复苏行动河湖名单（2022—2025 年）》中 9 个母亲湖为研究对象，基于 1984 年-2023 年 40 年来 Landsat-5 TM、Landsat-8/9 OLI、Sentinel-2 MSI 等多源遥感影像，采用基于面向对象和双峰自动阈值的湖泊水体面积提取方法提取湖泊水体，并通过目视解译对提取结果进行修正，生产了 1984 年—2023 年 9 个母亲湖水体面积数据集，共计 4791 期监测结果。本数据集可为母亲湖

复苏成效定量评估以及后续母亲河湖复苏行动开展提供支撑。

1 数据采集和处理方法

1.1 数据采集

本研究整合了国际上广泛应用的 Landsat-5 TM、Landsat-8/9 OLI、Sentinel-2 MSI 等中高空间分辨率卫星传感器地表反射率产品开展 9 个母亲湖水体提取，利用多源卫星协同监测以提升母亲湖水体提取的时空分辨率和可靠性。此外，为验证数据集水体提取精度，利用和数据集时间匹配的 1 m 空间分辨率的高分 2 号卫星提取水体面积作为相对真值，验证数据集水体面积精度。不同卫星信息详细介绍如下。

1.1.1 Sentinel-2

Sentinel-2 卫星是一组多光谱成像卫星，由 2A 和 2B 两颗星组成，全球重访周期为 5 天，Sentinel-2 卫星各搭载一枚多光谱成像仪（MultiSpectral Instrument，MSI），幅宽 290km，蓝、绿、红、近红波段影像空间分辨率为 10 m。本数据集使用 Sentinel-2 L2A 级影像，该影像为地表反射率产品，已经完成正射校正、辐射定标和大气校正，产品获取时间为 2017 年至今。

1.1.2 Landsat-5

Landsat-5 为美国陆地卫星系列的第五颗卫星，搭载多光谱扫描仪（Multispectral Scanner System，MSS）和专题制图仪（Thematic Mapper，TM）两枚传感器，影像全球重访周期为 16 天，幅宽 185km。本数据集使用 Landsat-5 TM Collection2 影像，该影像为地表反射率产品，已经完成正射校正、辐射定标和大气校正，空间分辨率为 30m，产品获取时间为 1984–2012 年。

1.1.3 Landsat-8/9

Landsat-8/9 分别为美国陆地卫星系列的第八、第九颗卫星，均搭载陆地成像仪（Operational Land Imager，OLI）和热红外传感器（Thermal Infrared Sensor，TIRS）两枚传感器，但 Landsat-9 搭载的 OLI-2 和 TIRS-2 传感器成像精度上要优于 Landsat-8，Landsat-8/9 组网后的全球重访周期为 8 天，幅宽 185km。本数据集使用地表反射率产品 Landsat-8 OLI/TIRS Collection 2 和 Landsat-9 OLI-2/TIRS-2 Collection 2，产品已完成正射校正、辐射定标和大气校正，空间分辨率为 30m，Landsat-8 产品获取时间为 2013 年至今，Landsat-9 产品获取时间为 2021 年 10 月至今。

1.1.4 高分 2 号

高分 2 号（以下称 GF-2）卫星为我国高分辨率对地观测系统卫星之一，2014 年投入使用，重访周期小于 5 天（卫星侧摆±23°），GF-2 卫星搭载 1 m 分辨率全色和 4 m 分辨率多光谱相机[18]。GF-2 卫星可从中国资源卫星应用中心（https://data.cresda.cn/#/home）下载 LEVEL 1A 影像，下载后需要进行正射校正、影像融合与裁剪等预处理，可获得 1m 空间分辨率多光谱影像，优于 Landsat-5 TM、Landsat-8/9 OLI 30 m 分辨率和 Sentinel-2 MSI 10 m 分辨率，因此，选择同日 GF-2 卫星水体面积监测值作为相对真值，用于验证数据集水体提取精度。

1.2 数据处理

利用 GEE 和湖泊缓冲区矢量裁剪获取 1984–2023 年覆盖 9 个母亲湖的逐期 Landsat-5 TM、Landsat-8/9 OLI 和 Sentinel-2 MSI 影像，云量设置为 30%，利用基于面向对象和双峰自动阈值的水体提取方法自动提取 9 个母亲湖水体面积，并对自动提取的水体面积矢量进行人工目视解译修正，最终获得 9 个母亲湖逐期水体面积矢量，具体流程如图 1 所示。

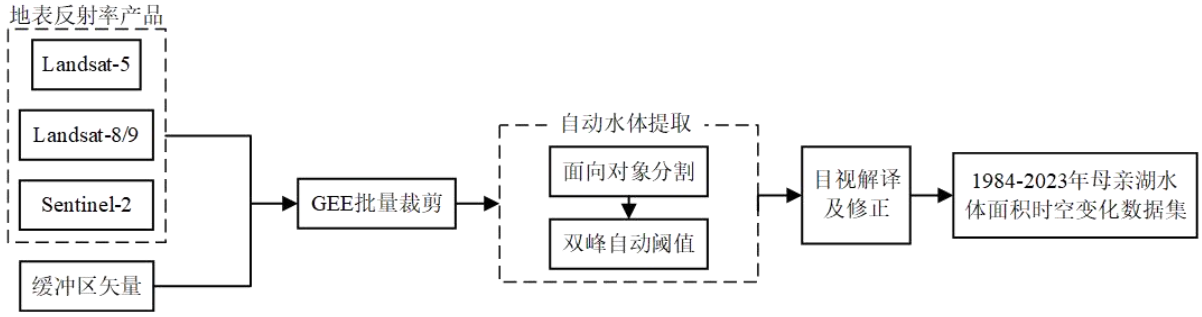


图 1　1984–2023 年母亲湖水体面积时空变化数据集生产流程

Figure 1　Production process of the dataset of spatial-temporal variations in water areas of 9 mother lakes (1984-2023)

1.2.1 影像裁剪

基于影像和湖泊先验范围绘制湖泊缓冲区矢量，缓冲区面积约为湖泊面积的 1.5-2 倍。使用 Google Earth Engine 的代码编辑器功能在线裁剪 Landsat-5 和 Landsat-8/9 和 Sentinel-2 L2A 级影像，云覆盖值设置为小于 30%。

1.2.2 自动水体提取

（1）面向对象分割

利用母亲湖缓冲区矢量裁剪影像生成湖泊缓冲区遥感影像，记为 IoL（Imageries of Lakes）。利用面向对象分割的方法将裁剪后的 IoL 进行多尺度分割，即基于像元尺度合并光谱、纹理特征相似的邻近像元，获得若干多边形对象。纹理特征包括不同波段的角二阶矩、对比度、熵、均匀性、相关性等。

（2）双峰自动阈值

基于面向对象分割得到的多边形尺度，计算逐个多边形归一化差异水体指数（NDWI）均值（式（1）），统计得到 IoL 的 NDWI 直方图，具体例子见图 2。由于水体与陆地的 NDWI 存在显著差异，IoL 的 NDWI 会呈现明显的双峰特征[19]，通过计算机自动识别双峰之间的最小值对应的 NDWI 值并将其作为水体提取的最优阈值，利用阈值分割法获得湖泊水体面积矢量，从而实现湖泊水体的自动提取。双峰自动阈值法可以根据每一期缓冲区影像 NDWI 的双峰特征自动确定最佳水体提取阈值，降低不同季节水体光谱特征变化对水体提取阈值的影响。

$$NDWI_{ploygon} = \frac{\sum_{i=1}^{n} \frac{B_{2_i} - B_{4_i}}{B_{2_i} + B_{4_i}}}{n} \quad (1)$$

式中：$NDWI_{ploygon}$代表某面向对象分割多边形的 NDWI 均值，n代表多边形对象个数，B_{2_i}、B_{4_i}分别代表多边形中第i个对象绿波段、近红波段的地表反射率值。

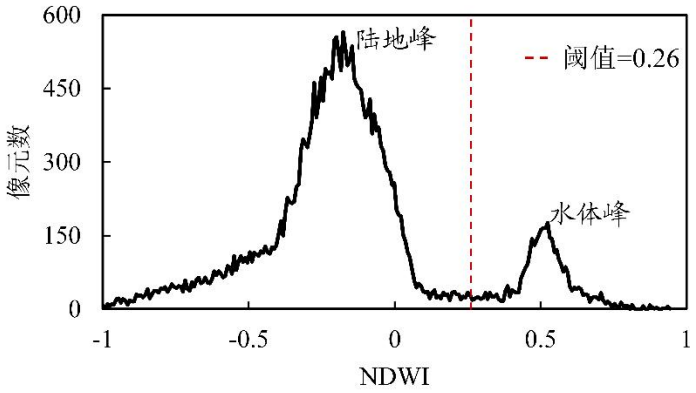


图 2　双峰阈值法 NDWI 直方图

Figure 2　NDWI histogram using the bimodal thresholding method

1.2.3　目视解译及修正

通过人工目视解译对自动提取的水体进行逐期审核，剔除云覆盖、光谱异常等导致的无效提取结果，针对最优阈值确定有误的自动提取结果，通过人工修改阈值对结果进行修正，对修改阈值后仍存在错提、漏提的部分进行手动修正。

1.2.4　精度验证方法

以各湖泊同日匹配 GF-2 水体提取结果作为相对真值，采用广泛使用的混淆矩阵精度评价方法验证数据集水体面积精度[20-21]，评价指标包括用户精度、生产者精度、总体精度和 Kappa 系数。混淆矩阵计算方法如表 1 所示。

表 1　混淆矩阵

Table 1　Confusion matrix

	GF-2 提取结果		错分误差	用户精度
	水体	非水体		
水体	正确分类为水体的样本数（TP）	错误分类非水体为水体的样本数（FP）	$\frac{FP}{\mathrm{TP}+\mathrm{FP}}$	$1-\frac{FP}{\mathrm{TP}+\mathrm{FP}}$
非水体	错误分类水体为非水体的样本数（FN）	正确分类为非水体的样本数（TN）	$\frac{FN}{\mathrm{FN}+\mathrm{TN}}$	$1-\frac{FN}{\mathrm{FN}+\mathrm{TN}}$
漏分误差	$\frac{FN}{\mathrm{TP}+\mathrm{FN}}$	$\frac{FP}{\mathrm{FP}+\mathrm{TN}}$		
生产者精度	$1-\frac{FN}{\mathrm{TP}+\mathrm{FN}}$	$1-\frac{FP}{\mathrm{FP}+\mathrm{TN}}$		
总精度	$P_0=(TP+TN)/(TP+FP+FN+TN)$			
Kappa 系数	$k=\frac{P_0-P_e}{1-P_e}$，$P_e=(TP+FN)(TP+FP)+(FP+TN)(FN+TN)/n^2$，$n$为样本点总数			

2 数据样本描述

本数据集包括：（1）白洋淀、衡水湖、大黄堡洼、团泊洼、七里海、南大港湿地、岱海、沙湖、长湖共 9 个母亲湖 1984-2023 年逐期水体面积矢量数据，以 shp 格式存储，命名分别为“传感器名称_湖泊编码_YYYYMMDD.shp”（传感器名称分别为 S2X_MSI/LT05/LC08/LC09，YYYY 为年份，MM 为月份，DD 为日），以岱海 2023 年 10 月 27 日监测结果为例，矢量数据展示如图 3 所示，母亲湖对应湖库编码如表 2 所示；（2）母亲湖 1984–2023 年水体面积统计表，以 xlxs 格式存储，记录 9 个母亲湖 1984-2024 年逐期水体面积值，以岱海为例，表格示例如表 3 所示。

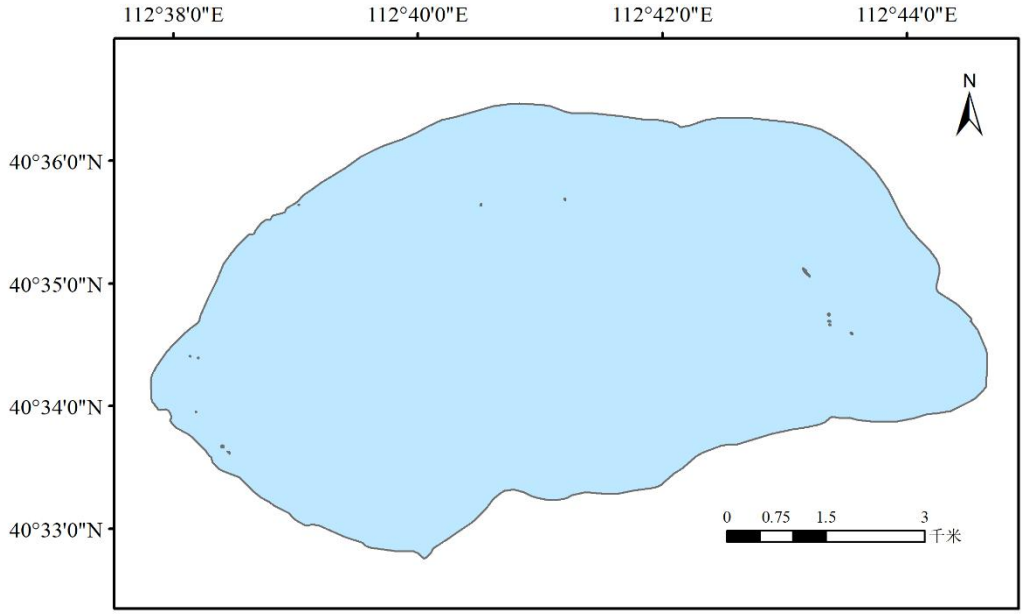


图 3　2023 年 10 月 27 日岱海水体面积

Figure 3　Daihai Lake water area on October 27, 2023

表 2　9 个母亲湖湖泊编码对照表

Table 2　Lake codes for the 9 mother lakes

序号	湖库名称	湖库编码
1	白洋淀	CD001
2	衡水湖	CF001
3	大黄堡洼	ZB002
4	团泊洼	ZB001
5	七里海	ZB003
6	南大港湿地	CF002
7	岱海	KA165
8	沙湖	D3064
9	长湖	F6006

表 3　1984–2023 年岱海逐期水体面积统计表（以 2023 年 10 月 27 日示例）

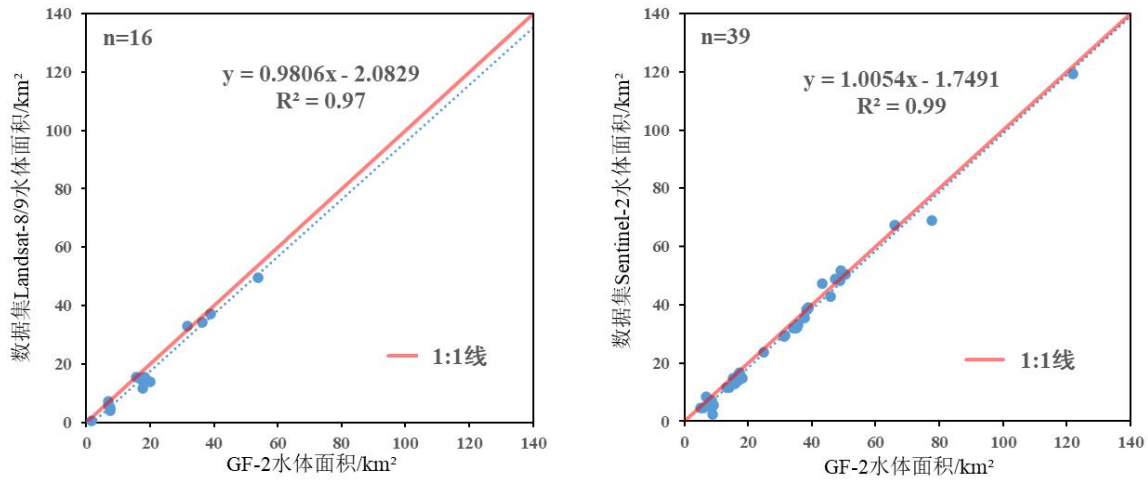
Table 3　Daihai lake water area statistics (1984–2023) with 27 October 2023 as an example）

WaterCode	Year	Month	Day	Sensor	FileName	Area
KA165	2023	10	27	S2X_MSI	S2X_MSI_KA165_20231027	43.837

3　数据质量控制和评估

以每个湖泊缓冲区矢量为范围，分别生成 500 个随机样本点，计算混淆矩阵（表 4）。9 个母亲湖共有 55 期 GF-2 有效同步监测结果，限于篇幅，表 4 仅列出 9 个母亲湖整体水体面积提取精度。由表 4 可知，9 个母亲湖水体提取总精度为 95.5%。9 个母亲湖中，长湖水体面积提取总精度最高为 99%，大黄堡洼、白洋淀水体较为零散，对水体面积提取精度影响较大，总体精度分别为 91.4%、93.1%。总体来看，9 个湖泊水体的错分误差为 6.5%，漏分误差为 13.9%，总精度为 94.7%，Kappa 系数为 0.861。

为进一步验证 30 m 空间分辨率 Landsat 系列卫星和 10 m 空间分辨率卫星监测水体面积精度和一致性，通过时间匹配获取了和 1 m 高空间分辨率 GF-2 卫星同天的 16 期 Landsat-8/9 影像与 39 期 Sentinel-2 影像，以 GF-2 卫星水体提取结果为相对真值，验证 Landsat-8/9 和 Sentinel-2 卫星水体提取精度和一致性（图 4a 和图 4b）。验证结果显示，Landsat-8/9 卫星监测水体面积与 GF-2 卫星监测水体面积的决定系数（R^2）达 0.98，略低于 Sentinel-2 卫星（R^2=0.99），跟 Landsat-8/9 卫星空间分辨率低于 Sentinel-2 卫星空间分辨率的特征一致。由图 4a 和图 4b 可以看出，Landsat-8/9 卫星监测水体面积和 GF-2 卫星监测水体面积散点大都十分靠近 1:1 线，均与 GF-2 监测水体面积高度一致（R^2>0.98），证明 30 m 空间分辨率 Landsat 系列卫星与 10 m Sentinel-2 卫星监测水体面积具有一致性，可利用 Landsat 系列卫星和 Sentinel-2 卫星协同开展母亲湖水体监测。



（a）　（b）

图 4　Landsat-8/9、Sentinel-2 监测水体面积与 GF-2 监测水体面积散点图

Figure 4　Scatter pots of water surface areas from Landsat-8/9 and Sentinel-2 versus GF-2

表 4　55 期同步影像水体面积混淆矩阵

Table 4　Confusion matrix of water areas for 55 periods of synchronized images

	GF-2 提取结果		错分误差	用户精度
	水体	非水体		
水体	6346	439	6.5%	93.5%
非水体	1025	19690	4.9%	95.1%
漏分误差	13.9%	2.2%		
生产者精度	86.1%	97.8%		
总精度	94.7%			
Kappa 系数	0.861			

4　数据使用方法和建议

本数据集为水利部《母亲河复苏行动方案（2022—2025 年）》中提及 9 个母亲湖 1984-2023 年 40 年遥感监测水体面积数据，以.shp 文件储存，可使用 ArcGIS、ENVI 等常用 GIS 与遥感影像处理软件读取使用。此外，本数据集还包含逐期水体面积统计表格，可直接用于数据分析。

数据作者分工职责

赵慧子（1997—），女，辽宁省盘锦市人，博士研究生，研究方向为水资源遥感。主要承担工作：监测方案规划、数据检索与收集、数据生产、数据集整理、论文撰写。

赵凯（1978—），男，山西省忻州市人，学士，高级工程师，研究方向为水利信息化。主要承担工作：数据收集。

曹引（1991—），男，安徽省滁州市人，博士，高级工程师，研究方向为水资源遥感。主要承担工作：方法指导、数据核验、数据论文修改。

赵红莉（1973—），女，安徽省蚌埠市人，博士，教授级高工，研究方向为水利信息化、水资源遥感与立体监测、数字孪生流域。主要承担工作：方法指导、数据论文修改。

蒋云钟（1969—），男，浙江省东阳市人，博士，教授级高工，研究方向为水利信息化、水资源管理与调度、数字孪生流域。主要承担工作：数据论文修改。

张怀文（1997—），男，安徽省合肥市人，博士研究生，研究方向为水资源遥感。主要承担工作：数据复核。

干雨（2000—），男，北京市人，硕士研究生，研究方向为水资源遥感。主要承担工作：目视解译及修改。

华雯静（2001—），女，江苏省无锡市人，硕士研究生，研究方向为水资源遥感。主要承担工作：目视解译及修改。

参考文献

[1]　李冰，杨桂山，万荣荣. 湖泊生态系统健康评价方法研究进展[J]. 水利水电科技进展，

2014,34(6):98-106. DOI:10.3880/j.issn.10067647.2014.06.019. [LI B, YANG G S, WAN R R. Progress on evaluation methods of lake ecosystem health[J]. Advances in Science and Technology of Water Resources, 2014,34(6):98-106. DOI:10.3880/j.issn.10067647.2014.06.019.]

[2] 马元植, 覃小林, 凌红波, 等. 1991—2020 年新疆中小湖泊面积变化时空特征及趋势分析[J]. 干旱区研究, 2024,41(6):905-916. DOI:10.13866/j.azr.2024.06.01. [MA Y Z, QIN X L, LING H B, et al. Spatio-temporal characteristics and trends of area changes in the small and medium-sized lakes in Xinjiang, China, from 1991 to 2020[J]. Arid Zone Research, 2024,41(6):905-916. DOI:10.13866/j.azr.2024.06.01.]

[3] WOOLWAY R I, KRAEMER B M, LENTERS J D, et al. Global lake responses to climate change[J]. Nature Reviews Earth & Environment, 2020, 1(8): 388-403, DOI:10.1038/s43017-020-0067-5.

[4] VOROSMARTY C J, GREEN P, SALISBURY J, et al. Global water resources: vulnerability from climate change and population growth[J]. science, 2000, 289(5477): 284-288, DOI:10.1126/science.289.5477.284.

[5] 张兵, 李俊生, 申茜, 等. 长时序大范围内陆水体光学遥感研究进展[J]. 遥感学报, 2021,25(01):37-52. DOI:10.11834/jrs.20210570. [ZHANG B, LI J S, SHEN Q, et al. Recent research progress on long time series and large scale optical remote sensing of inland water[J]. National Remote Sensing Bulletin, 2021,25(01):37-52. DOI:10.11834/jrs.20210570.]

[6] PEKEL J F, COTTAM A, GORELICK N, et al. High-resolution mapping of global surface water and its long-term changes[J]. Nature, 2016, 540(7633): 418-422, DOI:10.1038/nature20584.

[7] 陶亮亮, 王雨琦. 基于雷达和光学遥感数据的云污染区域光谱重建算法[J]. 农业工程学报, 2024, 40(17):216-223. DOI: 10.11975/j.issn.1002-6819.202405098. [TAO L L, WANG Y Q. Reconstructing spectral information in cloud contaminated regions using object-oriented approach[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(17):216-223. DOI: 10.11975/j.issn.1002-6819.202405098.]

[8] 岳照溪, 张永军, 段延松, 等. DEM 辅助的卫星光学遥感影像山体阴影检测与地形辐射校正[J]. 测绘学报, 2018,47(1):113-122. DOI:10.11947/j.AGCS.2018.20170194. [YUE Z X, ZHANG Y J, DUAN Y S, et al. DEM Assisted Shadow Detection and Topography Correction of Satellite Remote Sensing Images in Mountainous Area[J]. Acta Geodaetica et Cartographica Sinica, 2018,47(1):113-122. DOI:10.11947/j.AGCS.2018.20170194.]

[9] 卢善龙, 金继明, 贾立, 等. 基于 MODIS MOD09Q1 的青海、西藏湖泊水面数据集（2000～2012 年）[J/OL].中国科学数据, 2017, 2(2): 42-50. DOI:10.11922/csdata.170.2016.0113.[LU S L, JIN J M, JIA L, et al. MODIS (MOD09Q1)-derived lake water surface dataset for the Tibetan Plateau (2000–2012)[J/OL]. China Scientific Data, 2017, 2(2): 42-50. DOI: 10.11922/csdata.170.2016.0113.]

[10] 孙芳蒂, 熊立. 2000-2015 年咸海边界数据集[J/OL].中国科学数据, 2019, 4(4): 56-61. DOI: 10.11922/csdata.2019.0033.zh.[SUN F D, XIONG L. A dataset of the Aral Sea periphery during 2000 – 2015[J/OL]. China Scientific Data, 2019, 4(4): 56-61. DOI: 10.11922/csdata.2019.0033.zh.]

[11] 张鸿芳, 姚晓军, 肖建设, 等. 2020 年青海省面积≥1.0 km^{2}湖泊边界数据集[J/OL]. 中国科学数据, 2023, 8(1): 315-327. DOI: 10.11922/11-6035.csd.2021.0075.zh. [ZHANG H F, YAO X J, XIAO J S, et al. A dataset of boundaries data of the lakes (≥ 1.0 km^2) in Qinghai Province in 2020[J/OL]. China Scientific Data, 2023, 8(1): 315-327. DOI: 10.11922/11-6035.csd.2021.0075.zh.]

[12] 张大弘, 李晓锋, 姚晓军. 2000-2014 年中国西北地区面积≥10 km² 主要湖泊边界数据集[J/OL]. 中国科学数据, 2018, 3(4):55-63. DOI: 10.11922/csdata.2018.0041.zh. [ZHANG D H, LI X F, YAO X J. A dataset of main lakes in northwest China with an area above 10 km2 (2000 – 2014) [J/OL]. China Scientific Data, 2018, 3(4): 55-63. DOI: 10.11922/csdata.2018.0041.zh.]

[13] 德吉央宗, 边巴次仁, 曾林, 等. 1975–2021 年西藏南部佩枯错湖泊面积数据集[J/OL].中国科学数据, 2024, 9(4): 352-359. DOI: 10.11922/11-6035.csd.2023.0126.zh. [DEJI Y Z, BIANBA C R, ZENG L, et al. A dataset of lake areas in Paiku Tso, southern Tibet (1975–2021) [J/OL]. China Scientific Data, 2024,9(04):352-359. DOI: 10.11922/11-6035.csd.2023.0126.zh.]

[14] 李树涛, 李聪妤, 康旭东. 多源遥感图像融合发展现状与未来展望[J]. 遥感学报, 2021,25(1):148-166. DOI: 10.11834/jrs.20210259. [LI S T, LI C Y, KANG X D. Development status and future prospects of multi-source remote sensing image fusion[J]. National Remote Sensing Bulletin, 2021,25(1):148-166. DOI: 10.11834/jrs.20210259.]

[15] 陈楚楚, 刘浩, 彭丽瑶, 等. 基于多源遥感数据的湖泊水体演变研究[J]. 测绘与空间地理信息, 2024,47(3):32-35. [CHEN C C, LIU H, PENG L Y, et al. Study on Lake Water Body Evolution Based on Multi-source Remote Sensing Data[J]. Geomatics & Spatial Information Technology, 2024,47(3):32-35.]

[16] 张戈, 夏建新, 杜蕾, 等. Landsat8 和 Sentinel-2A 的水体指数比较与协同应用——以黑河流域典型区为例[J]. 应用基础与工程科学学报, 2020,28(06):1361-1368. DOI:10.16058/j.issn.1005-0930.2020.06.009. [ZHANG G, XIA J X, DU L, et al. Comparison and Collaborative Application of Water Index Based on Landsat8 and Sentinel-2A Data:A Case Study in Heihe River Basin,Northwest China[J]. Journal of Basic Science and Engineering, 2020,28(06):1361-1368. DOI:10.16058/j.issn.1005-0930.2020.06.009.]

[17] ARENA N, BATO G, BEKAERT D, et al. Opera Dynamic Surface Water Extents for Harmonized Landsat Sentinel-2 (DSWX-HLS) Validation Activities[C]. IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, California, USA, 2023.

[18] 廖静娟, 朱彬, 常云蕾, 等. 基于高分 2 号数据的海南岛红树林变化数据集(2015–2019)[J/OL]. 中国科学数据, 2022,7(4):7-17. DOI: 10.11922/11-6035.noda.2021.0016.zh. [LIAO J J, ZHU B, CHANG Y L, et al. A dataset of mangrove forest changes on Hainan Island based on GF-2 data during 2015–2019[J/OL]. China Scientific Data, 2022, 7(4):7-17. DOI: 10.11922/11-6035.noda.2021.0016.zh.]

[19] 张怀文, 曹引, 赵红莉, 等. 2018–2022 年青海省主要湖泊蓄水量时空变化数据集[J/OL]. 中国科学数据, 2024,9(1): 319-332. DOI: 10.11922/11-6035.noda.2023.0009.zh. [ZHANG H W, CAO Y, ZHAO H L, et al. A dataset of spatial-temporal variation of major lake water storage in Qinghai Province during 2018–2022[J/OL]. China Scientific Data, 2024, 9(1): 319-332. DOI: 10.11922/11-6035.noda.2023.0009.zh.]

[20] 郑晨键, 刘炜, 薛永福, 等. 基于 CART 算法的纳木措湖泊面积精确提取[J]. 哈尔滨商业大学学报(自然科学版), 2024,40(03):288-295. DOI:10.19492/j.cnki.1672-0946.2024.03.016.[ZHENG C J, LIU W, XUE Y F,et al. Accurate extraction of Namtso Lake area based on decision tree CART algorithm[J]. Journal of Harbin University of Commerce(Natural Sciences Edition), 2024,40(03):288-295. DOI:10.19492/j.cnki.1672-0946.2024.03.016.]

[21] 周再文, 何贞铭, 蒋松谕, 等. 多源时序影像洞庭湖流域水体时空演变特征分析[J]. 测绘通报,

2023,(10):20-27. DOI:10.13474/j.cnki.11-2246.2023.0290. [ZHOU Z W, HE Z M, JIANG S Y, et al. Analysis of spatio-temporal evolution characteristics of water bodies in Dongting Lake Basin based on multi-source time series images[J]. Bulletin of Surveying and Mapping, 2023,(10):20-27. DOI:10.13474/j.cnki.11-2246.2023.0290.]

论文引用格式

赵慧子, 赵凯, 曹引, 等. 1984–2023 年 9 个母亲湖水体面积时空变化数据集 [J/OL]. 中国科学数据, 2025, 10(3). (2025-09-18). DOI: 10.11922/11-6035.noda.2024.0164.zh.

数据引用格式

赵慧子, 赵凯, 曹引, 等. 1984–2023 年 9 个母亲湖水体面积时空变化数据集 [DS/OL]. V3. Science Data Bank, 2024. (2025-09-18). DOI:10.57760/sciencedb.13319.

A dataset of spatial-temporal variations in water areas of 9 mother lakes (1984–2023)

ZHAO Huizi[1,4], ZHAO Kai[2], CAO Yin[1,3*], ZHAO Hongli[1,3], JIANG Yunzhong[1,3,4], ZHANG Huaiwen[1], GAN Yu[1], HUA Wenjing[1]

1. Department of Water Resources, China Institute of Water Resources and Hydropower Research, Beijing 100038, P. R. China

2. Information Cener, Ministry of Water Resources, Beijing 100053, P. R. China

3. Key Laboratory of River Basin Digital Twinning of Ministry of Water Resources, Beijing 100038, P. R. China

4. Hohai University, Nanjing 210024, P.R.China

*Email: caoyin@iwhr.com

Abstract: As critical carriers of water resources and ecological systems, lakes play a vital role in regional water supply, ecological security, and the development of regional economy, society, and culture. However, some lakes are increasing threatened by shrinking water surface areas due to the influence of climate change and human activities. This study focuses on the nine mother lakes (Baiyangdian, Hengshui Lake, Dahuangpuwa, Tuanbowa, Qilihai wetland, Nandagang wetland, Daihai Lake, Sand Lake, Long Lake) identified in the *Mother River Revitalization Action Plan (2022–2025)* issued by the Ministry of Water Resources. Using a combination of satellite data sources, including Landsat-5 TM, Landsat-8/9 OLI, and Sentinel-2 MSI, we extracted the water surface areas of these nine mother lakes over a 40-year period from 1984 to 2023, resulting in 4,791 monitoring records. This study employed water surface areas derived from 55 periods of synchronized high-resolution (1 m) GF-2 (Gaofen-2) satellite images of nine mother lakes as reference data. The dataset accuracy was validated using the widely adopted confusion matrix method for accuracy assessment. Validation of this dataset shows an overall accuracy of 94.7%, and a Kappa coefficient of 0.86. This dataset provides valuable data support for the quantitative assessment of the revitalization efforts of the mother lakes.

Keywords: mother lakes; water area; Landsat-5; Landsat-8/9; Sentinel-2

Dataset Profile

Title	A dataset of spatial-temporal variations in water areas of 9 mother lakes (1984-2023)
Data corresponding author	CAO Yin (caoyin@iwhr.com)
Data authors	ZHAO Huizi, ZHAO Kai, CAO Yin, ZHAO Hongli, JIANG Yunzhong, ZHANG Huaiwen, GAN Yu, HUA Wenjing
Time range	1984–2023
Geographical scope	Baiyangdian (38°46'N –38°59'N，115°49'E –116°6'E), Hengshui Lake (37°32'N –37°39'N，115°33'E –115°39'E), Dahuangpuwa (39°24'N –39°32'N，117°10'E –117°17'E), Tuanbowa (38°50'N –38°58'N， 117°3'E –117°9'E), Qilihai Wetland (39°16'N –39°19'N， 117°27'E –117°38'E), Nandagang Wetland (38°27'N –38°36'N，117°25'E –117°35'E), Daihai Lake (40°30'N –40°37'N，112°35'E –112°46'E), Sand Lake (38°44'N –38°49'N，106°18'E –106°25'E), and Long Lake (30°20'N –30°32'N，112°10'E –112°32'E)
Spatial resolution	Sentinel-2 (10 m)；Landsat-5/8/9 (30 m)
Data volume	501 MB (unzipped)
Data format	*.shp, *.xlsx
Data service system	https://doi.org/10.57760/sciencedb.13319
Source(s) of funding	National Natural Science Foundation of China (52309040, 52279031); National Key Research and Development Program of China (2023YFF1304202).
Dataset composition	The dataset consists of water area shapefiles of 9 mother lakes (Baiyangdian, Hengshui Lake, Dahuangpuwa, Tuanbowa, Qilihai Wetland, Nandagang Wetland, Daihai Lake, Sand Lake and Long Lake) from 1984 to 2023, as well as statistical tables of their water areas over the same period.