



文献 CSTR:

32001.14.11-6035.noda.2024.0165.zh



文献 DOI:

10.11922/11-6035.noda.2024.0165.zh

数据 DOI:

10.57760/sciencedb.j00001.01041

文献分类: 地球科学

收稿日期: 2025-03-21

录用日期: 2025-03-25

发表日期: 2025-12-29

基于多源数据的青海省夏季地表水分布数据集 (2019–2023)

冯心熠^{1,2,3}, 肖正洁^{1,2,3}, 夏兴生^{1,2,3*}, 周强^{1,2,3}, 陈琼^{1,2,3}, 潘耀忠^{3,4}

1. 青海师范大学, 青藏高原地表过程与生态保育教育部重点实验室, 西宁 810016

2. 青海师范大学, 地理科学学院, 青海省自然地理与环境过程重点实验室, 西宁 810016

3. 青海师范大学, 高原科学与可持续发展研究院, 西宁 810016

4. 北京师范大学, 地理科学学部遥感科学国家重点实验室, 北京 100875

摘要: 青海省位于青藏高原东北部, 素有“中华水塔”的美誉。构建青海省地表水的分布数据集对中华水塔资源、生态安全、灾害风险监测及应对具有重要的支撑作用。本研究基于 GEE 平台, 利用哨兵一号和哨兵二号数据计算得到双极化水体指数 SDWI、归一化水体指数 NDWI、归一化积雪指数 NDSI 三个指数, 结合 SRTM 数据并运用阈值分割的方法对青海省夏季地表水进行了提取, 得到了空间分辨率为 10 m 的青海省夏季地表水分布数据集 (2019–2023), 并对每年地表水面积进行统计。通过构建混淆矩阵对数据集进行精度验证, 结果显示: 数据集平均总体精度达到 90%, 平均 Kappa 系数为 0.85。本数据集能够为青海省地表水变化监测, 气候变化研究以及相关政策制定提供数据支撑。

关键词: 地表水; 水体指数; 阈值分割; 多源数据; 青藏高原

数据库(集)基本信息简介

数据库(集)名称	基于多源数据的青海省夏季地表水分布数据集(2019–2023)
数据通信作者	夏兴生 (xxs@qhnu.edu.cn)
数据作者	夏兴生、冯心熠、肖正洁、周强、陈琼、潘耀忠
数据时间范围	2019–2023年
地理区域	地理范围为31°9′–39°19′N, 89°35′–103°04′E, 其中包括了青藏高原的东北部。涉及的地区有西宁市、海东市、海北藏族自治州、黄南藏族自治州、海南藏族自治州、果洛藏族自治州、玉树藏族自治州、海西蒙古族藏族自治州。
空间分辨率	10 m
数据量	245 MB
数据格式	*.tif, *.xlsx
数据服务系统网址	https://doi.org/10.57760/sciencedb.j00001.01041

* 论文通信作者

夏兴生: xxs@qhnu.edu.cn

基金项目	第二次青藏高原综合科学考察研究（2019QZKK0906、2019QZKK0603）；国家自然科学基金项目（42201027）。
数据库（集）组成	本数据集由两部分组成：（1）2019–2023年每年青海省夏季地表水分布数据，tif格式；（2）2019–2023年青海省夏季地表水面积统计结果，xlsx格式。

引言

地表水作为地球水体的重要组成部分，包括陆地表面可见的河流、湖泊、沼泽、湿地、冰川、雪山、雪地等水体，在维系陆地生态系统、满足生产生活用水方面扮演着不可或缺的角色^[1]。地表水的时空变化不仅仅是指示一般的自然规律，也可以从中窥探全球变化和人类活动影响背景下的波动或者异常，从而服务于气候、灾害的预报，甚至生产活动的安排、水资源调度策略的制定等事关生态系统维护和人类发展的工作。

青藏高原分布着全球范围最广、数量最多、海拔最高的高原湖泊群，同时拥有大量冰川和湿地，且长期被积雪覆盖，是长江、黄河、澜沧江、雅鲁藏布江等众多大江大河的发源地，所以其地表水的分布及变化对东亚、南亚水资源调配和全球变化的监测至关重要^[2]。众多专家学者也以高原湖泊、冰川、积雪、湿地为对象，从多方面进行了研究探索并形成了丰富的成果，不仅为青藏高原的生态修复和环境保护做出了贡献，也为青藏高原及周边地区的发展提供了有力的数据支撑。但是，青藏高原独特的地貌气候特征使得更多研究主要集中在单个要素的提取方面^[3]，针对青藏高原的冰川积雪和水体作为整体进行提取并对其进行时间序列分析研究则较少。

遥感技术因具有省时、高效的优点，不仅为大范围地表水的监测提供了丰富的数据源，也提供了便捷有效的手段，而现阶段地表水遥感提取的主流数据主要集中在 Landsat 系列、MODIS、Sentinel-2 等光学遥感数据和 Sentinel-1 雷达遥感数据，提取方法主要有面向对象的方法、水体指数方法、机器学习、深度学习等^[4]。这些都为后续地表水的持续监测提供了数据和方法支撑。

本研究以青藏高原东北部的青海省行政区划范围为研究区，以时间序列的地表水提取为目的，利用双极化水体指数（SDWI）、归一化水体指数（NDWI）、归一化积雪指数（NDSI）构建青海省夏季地表水（水体和冰川积雪）分布数据集，期望能够为大范围、长时间序列地表水监测提供技术方案，也为青藏高原地表水的变化研究提供数据支撑。

1 数据采集和处理方法

1.1 数据采集

本数据集生产过程采用的数据源为已经过辐射校正、地形校正的哨兵一号遥感数据产品、哨兵二号 L2A 级数据产品（已经过辐射校正、大气校正、系统几何校正）以及 SRTM 数据。以上产品基于 Google Earth Engine 云平台（<https://earthengine.google.com>）获取，主要筛选时间为 2019 年至 2023 年每年的 6 月 1 日至 8 月 31 日，且哨兵二号遥感影像云量小于 10%。对于 SRTM 数据根据研究区进行裁剪后计算得到研究区坡度和高程数据。

1.2 数据处理方法

1.2.1 双极化水体指数

双极化水体指数（Sentinel-1 Dual-Polarized Water Index, SDWI）是由贾诗超^[5]等根据水体在 Sentinel-1 双极化（VV 和 VH 极化）影像上灰度值不同的特性提出的一种水体指数，以达到增强水体特征的效果，并同时消除植被等其他地物的干扰，因此，主要用于提取地表水。其模型表达式如下：

$$K_{SDWI} = \ln(10 \cdot VV \cdot VH) \quad (1)$$

其中， K_{SDWI} 为水体指数， VV 和 VH 分别为 SAR 影像中 VV 极化和 VH 极化的像素值。

1.2.2 归一化差异水体指数

归一化差异水体指数（Normalized Difference Water Index NDWI）因可以快速、准确地提取水体范围，被广泛用于水资源管理、洪涝监测、环境保护等领域^[6]。其计算公式如下：

$$NDWI = \frac{(Green - NIR)}{(Green + NIR)} \quad (2)$$

其中， $Green$ 表示绿波段反射率， NIR 表示近红外波段反射率。

1.2.3 归一化差异积雪指数

归一化差异积雪指数（Normalized Difference Snow Index NDSI）^[7]通过计算绿光波段和近红外波段之间反射率的差异得到，能够检测冰川积雪的存在。计算公式如下：

$$NDSI = \frac{Green - SWIR}{Green + SWIR} \quad (3)$$

其中， $Green$ 表示绿光波段反射率， $SWIR$ 表示短波红外波段反射率。

1.2.4 大津算法

大津算法（OTSU），也称为最大类间方差法，是一种确定影像分割阈值的算法，且不受图像亮度和对比度的影响^[8]。OTSU 将图像分成背景和前景两部分，以前景和背景的方差最大来动态确定图像的阈值，通过获得的阈值进行二值化分割，最终形成二值图像。当背景和前景之间的方差越大，说明这两部分差别越大、越易区分，因此，当两部分之间的方差最大时意味着错分的概率最小^[9]。具体算法如下：

$$w_1 = \frac{n_1}{h \cdot w} \quad w_2 = \frac{n_2}{h \cdot w} \quad (4)$$

$$u_1 = \frac{N_1}{n_1} \quad u_2 = \frac{N_2}{n_2} \quad (5)$$

$$T = \arg\left\{\max\left[w_1 \cdot w_2 (u_1 - u_2)^2\right]\right\} \quad (6)$$

其中， h 为影像高度， w 为影像宽度； n_1 为大于阈值的像素，即前景； n_2 为小于等于阈值的像素，即背景； N_1 为 n_1 灰度值的和， N_2 为 n_2 灰度值的和； w_1 、 w_2 为影像前景和背景的像元概率； u_1 、 u_2 为影像前景和背景的平均灰度值， T 为计算得到的阈值。

1.2.5 数据处理流程

本数据集制作技术路线如图 1 所示。为了规避分割尺度对地表水提取的影响，本研究通过预实验发现对像元数小于等于 50702×28513 的区域进行阈值分割的结果是最优的。因此，在后续的工作中将研究区划分为 9 个子区进行地表水的逐区提取。具体如下：

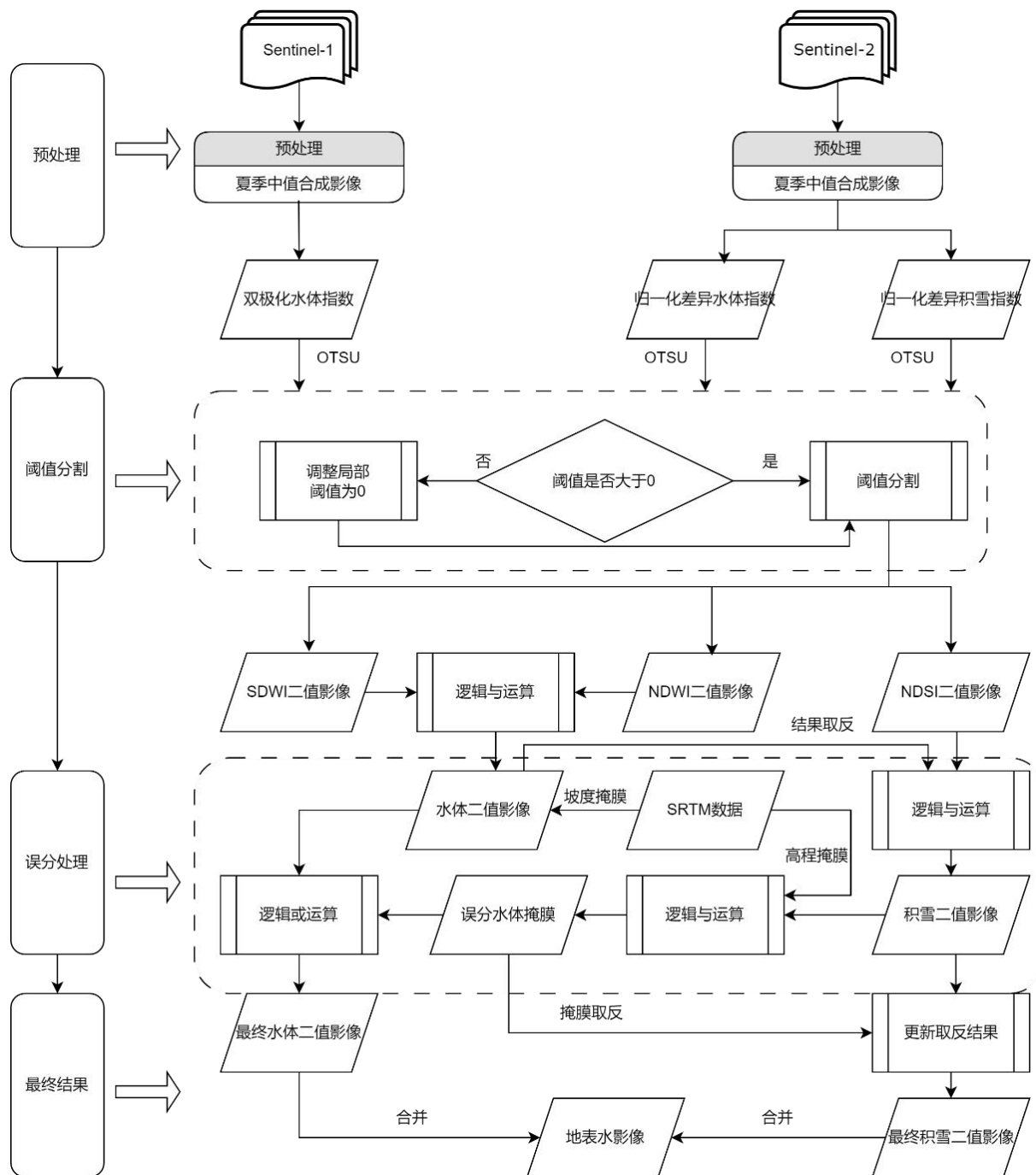


图 1 数据集制作技术路线图

Figure 1 Flowchart of the dataset generation

(1) 基于经过预处理的哨兵一号、哨兵二号数据产品计算得到 SDWI、NDWI、NDSI，并通过大津算法分别计算各指数局部分割阈值。当阈值小于 0 时，重新调整局部阈值为 0，最终确定分割阈值^[10-11]，并获得初步的地表水分布。随后在 GEE 平台上逐区域对水体范围和冰川积雪范围进行下

一步处理。

（2）水体范围的确定。为减小山体阴影的影响^[12]，通过 SRTM 计算得到坡度，结合目视解译统计局部有水体存在且坡度最小的坡度阈值，得到坡度掩膜，然后将 NDWI 和 SDWI 阈值分割结果中的重叠部分结合坡度掩膜^[13]，得到水体的二值结果 S_{w1} 。为减少水体与积雪的误分，下一步将结合 SRTM 高程数据^[14]，对局部区域进行目视解译，确定积雪覆盖的最低高程阈值，并将低于该阈值的积雪区域识别为水体 S_{w2} 。使用 $\text{lt}()$ 函数，输入高程数据得到高程掩膜，再对 NDSI 分割结果使用 $\text{updateMask}()$ 函数^[15]更新高程掩膜得到被误分的水体结果 S_{w2} ，再与之前的水体结果进行合并得到最终子区的水体二值影像。过程如下：

$$T_{\text{slope}} < T_{\text{坡度阈值}}, T_{\text{NDWI}} > 0 \text{ 且 } T_{\text{SDWI}} > 0 \quad (7)$$

$$T_{\text{NDSI}} > 0 \text{ 且 } T_{\text{elevation}} < T_{\text{高程阈值}} \quad (8)$$

$$S_w = S_{w1} + S_{w2} \quad (9)$$

式中， T_{slope} 为坡度， T_{NDWI} 为 NDWI 指数， T_{SDWI} 为 SDWI 指数， T_{NDSI} 为 NDSI 指数， $T_{\text{elevation}}$ 为高程， S_{w1}, S_{w2} 分别为第一步提取出的水体结果和第二步提取出的水体结果。

（3）冰川积雪范围的确定。对第一步提取出来的水体二值结果取反，使用 $\text{and}()$ 函数和 NDSI 的阈值分割结果取交集，得到积雪二值结果。然后使用 $\text{updateMask}()$ 函数更新前一步误分水体二值结果取反的掩膜，得到最终子区的冰川积雪二值影像。

$$R_{\text{NDWI}} = 0 \text{ 且 } T_{\text{NDSI}} > 0 \quad (10)$$

$$R_{\text{SI}} = 1 \text{ 且 } R_{\text{W2}} = 0 \quad (11)$$

式中， R_{NDWI} 为子区水体 S_{w1} 的最终二值结果， T_{NDSI} 为 NDSI 指数， R_{SI} 为第一步确定的冰川积雪二值结果， R_{W2} 为第二步确定的水体二值结果。

（4）将所有子区最终确定的水体范围和冰川积雪范围合并，得到青海省夏季地表水分布数据集。

$$\sum_{i=1}^9 S = S_{wi} + S_{si} \quad (12)$$

式中， S_{wi} 为各子区水体提取结果， S_{si} 为各子区积雪提取结果， S_w 为最终地表水结果。

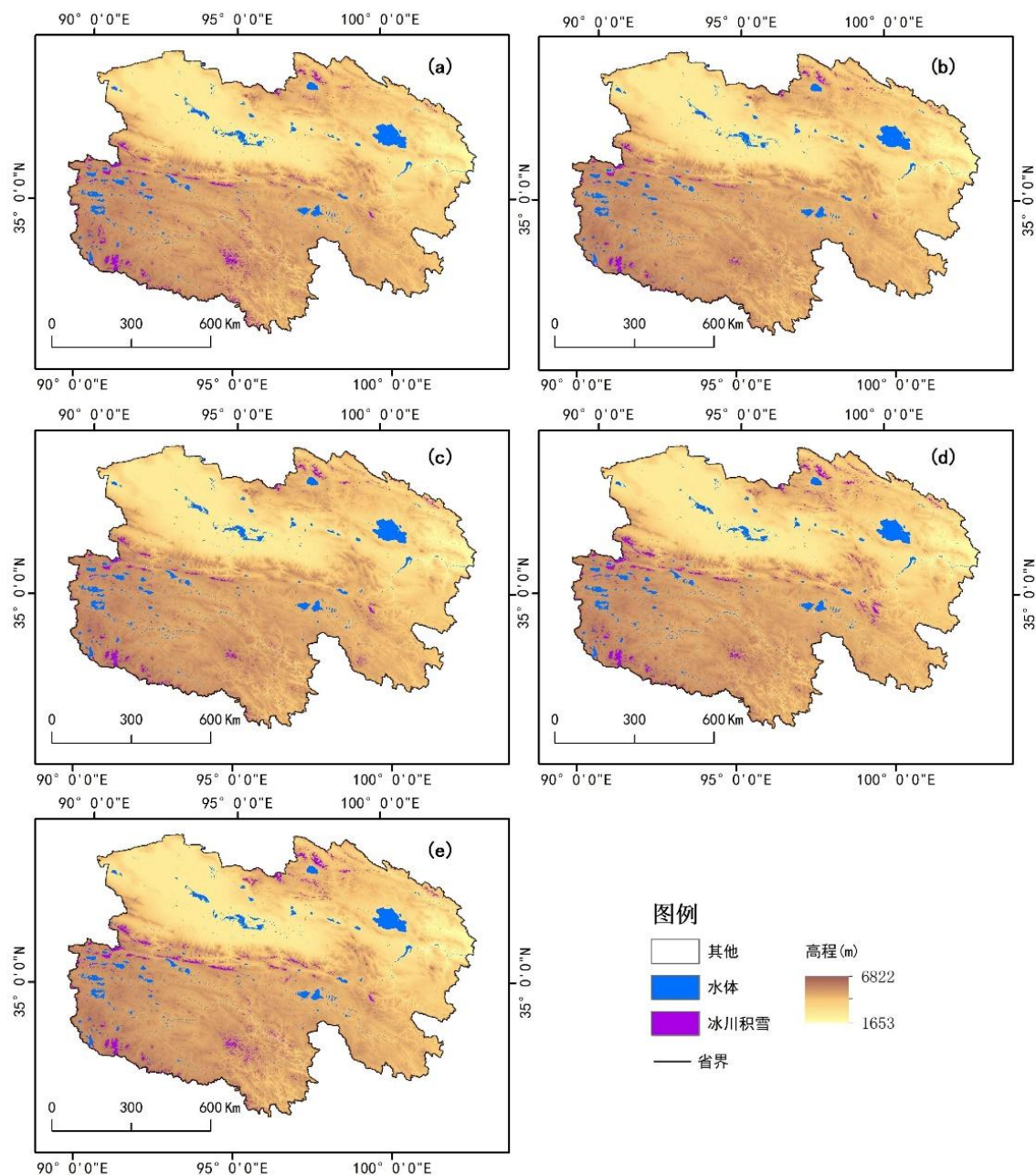
2 数据样本描述

2.1 数据集信息

青海省夏季地表水分布数据集主要包含两部分，分别为：2019–2023 年青海省夏季地表水分布栅格数据（包括水体和冰川积雪）；2019–2023 年夏季地表水面积统计 excel 表格。栅格数据命名格式为 YYYY_qh.tif，其中 YYYY 表示年份，qh 表示研究区简称。数据集存储为 tif、xlsx 格式。数据集由 26 个文件组成，压缩后为 245 MB。

2.2 数据样本

数据集结果为栅格格式和表格格式。栅格数据空间分辨率为 10 m。将最后提取结果分为水体、冰川积雪、其他三类。其中像元值 1 代表水体，像元值 2 代表冰川积雪，像元值 0 的为其他。表格数据为 2019–2023 年青海省夏季地表水面积统计结果。数据样本展示结果如图 2 所示。



(a)为2023年青海省夏季地表水分布结果,(b)为2022年青海省夏季地表水分布结果,(c)为2021年青海省夏季地表水分布结果,(d)为2020年青海省夏季地表水分布结果,(e)为2019年青海省夏季地表水分布结果。(a) Summer surface water distribution in Qinghai Province in 2023, (b) summer surface water distribution in 2022, (c) summer surface water distribution in 2021, (d) summer surface water distribution in 2020, and (e) summer surface water distribution in 2019.

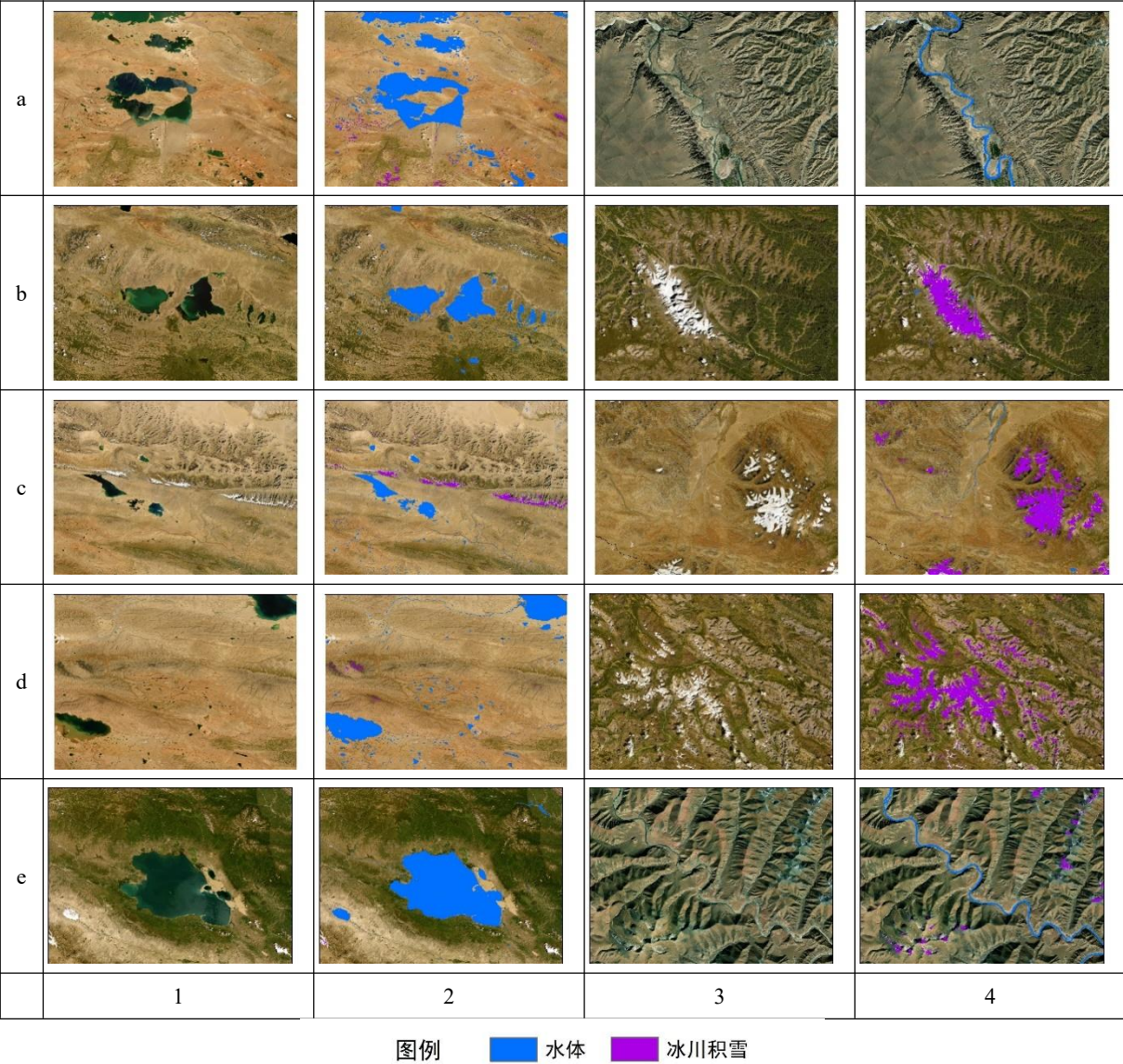
图2 青海省夏季地表水数据集（2019-2023）

Figure 2 Summer surface water dataset of Qinghai Province (2019-2023)

3 数据质量控制和评估

3.1 地表水提取结果分析

图 3 为 2019–2023 年部分地区地表水提取效果图。结果显示，湖泊、河流、坑塘以及冰川积雪提取效果较好（图 a2、b2、c2、e2）。在山区，地表水提取能很好地避免山体阴影带来的影响（图 a4、e4），面积较小的坑塘也能被准确提取（图 d2）。提取结果里存在部分积雪水体误分，主要表现在地形复杂高海拔地区，积雪和水体存在异物同谱现象^[16]。尽管已经通过局部地区高程和坡度的掩膜以及基于指数的逻辑运算进行了区分，但是并没有完全消除，积雪和水体误分的现象依然个别存在（图 c4）。



(a)为 2023 年地表水提取结果对比,(b)为 2022 年地表水提取结果对比,(c)为 2021 年地表水提取结果对比,(d)为 2020 年地表水提取结果对比,(e)为 2019 年地表水提取结果对比。(a) Comparison of surface water extraction results in 2023, (b) comparison of surface water extraction results in 2022, (c) comparison of surface water extraction results in 2021, (d) comparison of surface water extraction results in 2020, and (e) comparison of surface water extraction results in 2019.

图 3 2019–2023 年青海省部分地区地表水提取结果图

Figure 3 Surface water extraction results for selected areas in Qinghai Province (2019–2023)

3.2 夏季地表水面积统计分析

本研究通过对 2019–2023 年青海省夏季地表水分布数据集进行面积统计可知（图 4），2019–2023 年青海省夏季地表水整体呈现先下降再回升的趋势，其中，冰川积雪的消融是地表水呈下降趋势的直接原因，水体则整体保持平稳态势。2022 年为 5 年内地表水最少的年份，面积仅为 222.01 万公顷，其中积雪面积为 47.24 万公顷。青海省夏季地表水整体变化趋势符合现阶段青藏高原气候变暖、变湿的实际变化规律^[17]。此外，2021 年青海省夏季地表水统计面积约为 244.71 万公顷，与青海省第三次全国国土调查中的水域调查结果 244.66 万公顷基本相符合，证明本数据集具有较高的可靠性。

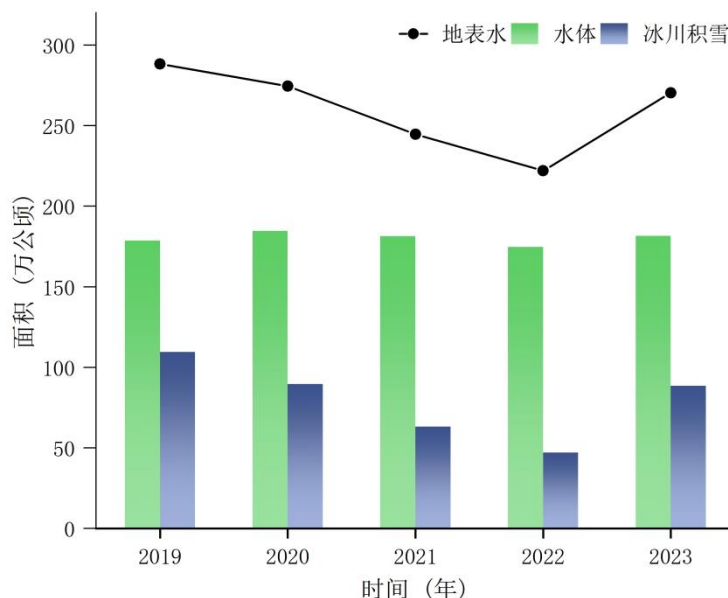


图 4 青海省夏季地表水面积（2019–2023）

Figure 4 Summer surface water area in Qinghai Province (2019–2023)

3.3 数据质量评价

本研究在历史高清遥感影像中选取样本点 1405 个，通过目视解译将样本点分为水体、冰川积雪、其他三类，以此来构建混淆矩阵对青海省夏季地表水分布数据集进行定量化精度验证。使用总体精度、Kappa 系数、生产者精度、用户精度 4 个指标进行精度评价，结果如表 1 所示。青海省夏季地表水分布数据集五年平均精度达到 90%，平均 Kappa 系数为 0.85，最高总体精度为 2020 年的 95%，Kappa 系数为 0.93；最低总体精度为 2021 年的 85%，Kappa 系数为 0.77。

表 1 青海省夏季地表水分布数据集精度表

Table 1 Accuracy assessment of the summer surface water distribution dataset in Qinghai Province

时间（年）	总体精度	Kappa 系数	用户精度	生产者精度
2019	94%	0.91	[87%,98%,100%]	[98%,92%,91%]
2020	95%	0.93	[89%,100%,100%]	[100%,91%,94%]
2021	85%	0.77	[72%,95%,100%]	[98%,95%,65%]
2022	87%	0.81	[76%,97%,97%]	[99%,81%,80%]
2023	89%	0.81	[79%,95%,97%]	[98%,92%,73%]

注：用户精度和生产者精度数值分别代表其他、水体、冰川积雪。

4 数据使用方法和建议

本数据集为 2019–2023 年青海省夏季地表水分布数据集, 格式为 tif 和 xlsx, 可以在 ENVI, ArcGIS 等遥感数据处理或地理信息系统软件中直接打开使用。基于本数据可以进行气候变化、水资源监测、生态安全和自然灾害风险的研究, 也可以为相关政策的制定提供基础数据服务。

作者分工职责

冯心熠 (2001—), 男, 重庆人, 在读硕士研究生, 研究方向为空间数据处理与挖掘。主要承担工作: 遥感数据收集处理, 数据集的开发, 精度验证, 数据论文撰写等。

肖正洁 (2000—), 男, 湖北省襄阳人, 本科, 助理工程师, 研究方向为空间数据处理与挖掘。主要承担工作: 数据的收集及处理、预实验, 论文初稿的撰写等。

夏兴生 (1989—), 男, 甘肃省庆阳人, 博士研究生, 副教授, 研究方向为遥感和地理信息技术的教学与应用。主要承担工作: 提出数据集开发思路、框架, 实施指导及论文的修改定稿等。

周强 (1971—), 男, 陕西省临潼人, 硕士研究生, 教授, 研究方向为土地资源与环境变化。主要承担工作: 提供数据集研发的项目支持, 指导数据集的开发等。

陈琼 (1975—), 女, 浙江省诸暨人, 博士研究生, 教授, 研究方向为土地科学及人地关系。主要承担工作: 数据集的开发指导、审核, 数据论文的修改等。

潘耀忠 (1965—), 男, 山西省大同人, 博士研究生, 教授, 研究方向为资源环境定量遥感应用, 主要承担工作: 数据集的审核, 数据论文科学合理性审查等。

参考文献

- [1] 李景刚, 李纪人, 黄诗峰, 等. *Terra*/MODIS 时间序列数据在湖泊水域面积动态监测中的应用研究: 以洞庭湖地区为例[J]. 自然资源学报, 2009, 24(5): 923–933. [LI J G, LI J R, HUANG S F, et al. Application of *Terra*/MODIS time series data in dynamic monitoring of lake water area variations: a case study in Dongting Lake region, China[J]. Journal of Natural Resources, 2009, 24(5): 923–933.]
- [2] 刘佳, 陈超, 秦宁生, 等. 青藏高原若尔盖生态区水资源对气候变化的响应[J]. 冰川冻土, 2016, 38(2): 498–508. DOI:10.7522/j.issn.1000-0240.2016.0056. [LIU J, CHEN C, QIN N S, et al. Response of water resources to climate change in Zoige, Tibetan Plateau[J]. Journal of Glaciology and Geocryology, 2016, 38(2): 498–508. DOI:10.7522/j.issn.1000-0240.2016.0056.]
- [3] 常亚茹, 张继贤, 韩文立, 等. 利用 Sentinel-1A 数据的纳木错湖面月际变化监测[J]. 遥感信息, 2022, 37(2): 45–52. DOI:10.3969/j.issn.1000-3177.2022.02.006. [CHANG Y R, ZHANG J X, HAN W L, et al. Monitoring inter-monthly change of namtso lake surface using sentinel-1A data[J]. Remote Sensing Information, 2022, 37(2): 45–52. DOI:10.3969/j.issn.1000-3177.2022.02.006.]
- [4] 苏龙飞, 李振轩, 高飞, 等. 遥感影像水体提取研究综述[J]. 国土资源遥感, 2021, 33(1): 9–19. DOI:10.6046/gtzyyg.2020170. [SU L F, LI Z X, GAO F, et al. A review of remote sensing image water extraction[J]. Remote Sensing for Land & Resources, 2021, 33(1): 9–19. DOI:10.6046/gtzyyg.2020170.]
- [5] 贾诗超, 薛东剑, 李成绕, 等. 基于 Sentinel-1 数据的水体信息提取方法研究[J]. 人民长江, 2019, 50(2): 213–217. DOI:10.16232/j.cnki.1001-4179.2019.02.038. [JIA S C, XUE D J, LI C R, et al.

- Study on new method for water area information extraction based on Sentinel-1 data[J]. Yangtze River, 2019, 50(2): 213–217. DOI:10.16232/j.cnki.1001-4179.2019.02.038.]
- [6] 熊珂, 常云鹏, 关键. 基于归一化差异水体指数的水体面积变化监测[J]. 中南林业调查规划, 2024, 43(2): 54–57. DOI:10.16166/j.cnki.cn43-1095.2024.02.011. [XIONG K, CHANG Y P, GUAN J. Monitoring of water area changes based on the normalized difference water index[J]. Central South Forest Inventory and Planning, 2024, 43(2): 54–57. DOI:10.16166/j.cnki.cn43-1095.2024.02.011.]
- [7] 侯靖琪, 王宁练, 石晨烈, 等. 2000—2021 年蒙古高原湖冰物候变化研究[J]. 冰川冻土, 2024, 46(5): 1537–1550. DOI:10.7522/j.issn.1000-0240.2024.0120. [HOU J Q, WANG N L, SHI C L, et al. Study on phenological changes of lake ice in Mongolian Plateau from 2000 to 2021[J]. Journal of Glaciology and Geocryology, 2024, 46(5): 1537–1550. DOI:10.7522/j.issn.1000-0240.2024.0120.]
- [8] OTSU N. A threshold selection method from gray-level histograms[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9(1): 62–66. DOI:10.1109/TSMC.1979.4310076.
- [9] 冯崎, 王琦, 黄海兰, 等. 基于 Sentinel-1 双极化数据改进水体提取的 Otsu 算法[J]. 遥感技术与应用, 2022, 37(4): 1003–1011. DOI:10.11873/j.issn.1004-0323.2022.4.1003. [FENG Q, WANG Q, HUANG H L, et al. Improved otsu algorithm for water extraction based on sentinel-1 dual-polarization data[J]. Remote Sensing Technology and Application, 2022, 37(4): 1003–1011. DOI:10.11873/j.issn.1004-0323.2022.4.1003.]
- [10] ZHANG H B, ZHANG F, ZHANG G Q, et al. Ground-based evaluation of MODIS snow cover product V6 across China: Implications for the selection of NDSI threshold[J]. Science of the Total Environment, 2019, 651: 2712–2726. DOI:10.1016/j.scitotenv.2018.10.128.
- [11] 徐涵秋. 利用改进的归一化差异水体指数(MNDWI)提取水体信息的研究[J]. 遥感学报, 2005, 9(5): 589–595. [XU H Q. A study on information extraction of water body with the modified normalized difference water index (MNDWI)[J]. Journal of Remote Sensing, 2005, 9(5): 589–595.]
- [12] 王西营, 张帅娟. 融合光谱和时相特征的地表水遥感信息提取方法[J]. 内蒙古师范大学学报(自然科学版), 2024, 53(5): 488–496. [WANG X Y, ZHANG S J. A method for extracting surface water remote sensing information by integrating spectral and temporal characteristics[J]. Journal of Inner Mongolia Normal University (Natural Science Edition), 2024, 53(5): 488–496.]
- [13] 何丽. 基于 Google Earth Engine 的湄公河流域资源环境关键要素动态监测与时空分析[D]. 昆明: 云南师范大学, 2021. [HE L. Dynamic monitoring and spatiotemporal analysis of key elements of resources and environment in Mekong Basin based on Google Earth Engine platform[D]. Kun Ming: YUNNAN NORMAL UNIVERSITY, 2021.]
- [14] LI Y, NIU Z G, XU Z Y, et al. Construction of high spatial-temporal water body dataset in China based on sentinel-1 archives and GEE[J]. Remote Sensing, 2020, 12(15): 2413. DOI:10.3390/rs12152413.
- [15] 沈梦瑶. 基于 GEE 平台的海水水体提取算法研究[D]. 湘潭: 湘潭大学, 2024. [SHENG M Y. Research on seawater extraction algorithm based on GEE platform[D], XiangTan: Xiangtan University, 2024.]
- [16] 张永合, 范城城, 陈明, 等. 联合多源卫星遥感影像自动提取高原冰冻淡水湖泊[J]. 地理空间信息, 2025, 23(2): 78–81. DOI:10.3969/j.issn.1672-4623.2025.02.017. [ZHANG Y H, FAN C C, CHEN M, et al. Automatic extraction of plateau frozen freshwater lake using multiple satellite images[J]. Geospatial Information, 2025, 23(2): 78–81. DOI:10.3969/j.issn.1672-4623.2025.02.017.]

[17] 丁佳, 刘星雨, 郭玉超, 等. 1980—2015 年青藏高原植被变化研究[J]. 生态环境学报, 2021, 30(2): 288–296. DOI:10.16258/j.cnki.1674-5906.2021.02.007. [DING J, LIU X Y, GUO Y C, et al. Study on vegetation change in the Qinghai-Tibet Plateau from 1980 to 2015[J]. Ecology and Environmental Sciences, 2021, 30(2): 288–296. DOI:10.16258/j.cnki.1674-5906.2021.02.007.]

论文引用格式

冯心熠, 肖正洁, 夏兴生, 等. 基于多源数据的青海省夏季地表水分布数据集（2019 – 2023）[J/OL]. 中国科学数据, 2026, 11(1). (2025-12-29). DOI: 10.11922/11-6035.noda.2024.0145.zh.

数据引用格式

夏兴生, 冯心熠, 肖正洁, 等. 基于多源数据的青海省夏季地表水分布数据集（2019–2023）[DS/OL]. V2. Science Data Bank, 2025. (2025-04-07). DOI: 10.57760/sciencedb.j00001.01041.

A dataset of summer surface water distribution in Qinghai Province based on multi-source data (2019–2023)

FENG Xinyi^{1,2,3}, XIAO Zhengjie^{1,2,3}, XIA Xingsheng^{1,2,3*}, ZHOU Qiang^{1,2,3},
CHEN Qiong^{1,2,3}, PAN Yaozhong^{3,4}

1. Key Laboratory of Tibetan Plateau Land Surface Processes and Ecological Conservation (Ministry of Education), Qinghai Normal University, Xining 810016, P.R. China
2. Qinghai Provincial Key Laboratory of Physical Geography and Environmental Process, School of Geographical Science, Qinghai Normal University, Xining 810016, P.R. China
3. Academy of Plateau Science and Sustainability, Qinghai Normal University & Beijing Normal University, Xining 810016, P.R. China
4. State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, P.R. China

*Email: xxs@qhnu.edu.cn (XIA Xingsheng)

Abstract: Qinghai Province, situated in the northeastern region of the Qinghai-Tibet Plateau, is famously known as the “Water Tower of China”. Constructing a surface water distribution dataset for this region is crucial for monitoring water resource, ecological security, and disaster risk, and for supporting response efforts in China. This study, based on the Google Earth Engine (GEE) platform, uses Sentinel-1 and Sentinel-2 to calculate three indices: the Dual-Polarization Water Body Index (SDWI), the Normalized Difference Water Index (NDWI), and the Normalized Difference Snow Index (NDSI). By integrating these indices with SRTM data and applying a threshold segmentation method, the summer surface water distribution in Qinghai Province was extracted, and a dataset with a 10-meter spatial resolution spanning from 2019 to 2023 was obtained. Additionally, the annual surface water area was statistically calculated. A confusion matrix was used to evaluate the accuracy of the dataset. The results show that the average overall

accuracy of the dataset reached 90% and the average Kappa coefficient is 0.85. This dataset can provide data support for monitoring surface water changes, studying climate change impacts, and formulating related policies in Qinghai Province.

Keywords: surface water; water index; threshold segmentation; multi-source data; Qinghai-Tibetan Plateau

Dataset Profile

Title	A dataset of summer surface water distribution in Qinghai Province based on multi-source data (2019 – 2023)
Data corresponding author	XIA Xingsheng (xxs@qhnu.edu.cn)
Data authors	XIA Xingsheng, FENG Xinyi, XIAO Zhengjie, ZHOU Qiang, CHEN Qiong, PAN Yaozhong
Time range	2019–2023
Geographical scope	The northeastern part of the Qinghai-Tibet Plateau (31°9′–39°19′N and 89°35′–103°04′E), covering Xining City, Haidong City, Haibei Tibetan Autonomous Prefecture, Huangnan Tibetan Autonomous Prefecture, Hainan Tibetan Autonomous Prefecture, Golog Tibetan Autonomous Prefecture, Yushu Tibetan Autonomous Prefecture, and Haixi Mongolian and Tibetan Autonomous Prefecture.
Spatial resolution	10 m
Data volume	245 MB
Data format	*.tif, *.xlsx
Data service system	https://doi.org/10.57760/sciencedb.j00001.01041
Source(s) of funding	The Second Tibetan Plateau Scientific Expedition and Research (2019QZKK0906, 2019QZKK0603); the National Natural Science Foundation of China (Grant No. 42201027).
Dataset composition	This dataset consists of two parts: (1) annual summer surface water distribution data for Qinghai Province from 2019 to 2023, stored in tif format; (2) Statistical results of summer surface water area in Qinghai Province from 2019 to 2023, stored in xlsx format.