

1967–2014 年科西河流域面积 $\geq 0.05 \text{ km}^2$ 冰湖轮廓

数据集

ISSN 2096-2223

CN 11-6035/N

张大弘¹, 宫鹏¹, 姚晓军^{1*}, 刘娟¹



文献 DOI:

10.11922/csdata.2019.0064.zh

数据 DOI:

10.11922/sciencedb.885

文献分类: 地球科学

收稿日期: 2019-09-05

开放同评: 2019-11-28

录用日期: 2020-05-14

发表日期: 2020-06-28

1. 西北师范大学地理与环境科学学院, 兰州 730070

摘要: 科西河流域发育有众多的冰川和冰湖, 冰川覆盖率达 9.78%, 冰湖数量 221 个, 冰湖分布及变化不仅是预测该区域冰湖溃决洪水灾害的基础数据, 而且可以在一定程度上反映高山寒区气候及冰川水资源的变化。本数据集应用 1960s–2010s 多源遥感影像数据 (Corona、Landsat MSS/TM/ETM+/OLI) 及地形图进行目视解译获得了科西河流域 1960s、1970s、1980s、1990s、2000s 和 2010s 的 6 期冰湖数据, 解译精度控制在一个像元以内。研究区空间范围为 26.86°–29.14°N, 85.39°–88.95°E, 覆盖印德拉瓦迪、孙科西、绒辖曲、利库科拉、牛奶河、阿润和塔木尔等子流域。本数据集包括科西河流域内面积 $\geq 0.05 \text{ km}^2$ 的冰湖, 数据集时序较长且包含冰湖类型及其他几何属性信息, 可作为科西河流域冰湖时空变化、冰湖水量变化评估的基础数据, 也可作为喜马拉雅山地区冰湖溃决危险性评估、水资源变化、气候变化及冰湖与冰川间变化响应关系等研究的重要依据。

关键词: 冰湖; 科西河; 目视解译; 青藏高原

数据库 (集) 基本信息简介

数据库 (集) 名称	1967–2014 年科西河流域面积 $\geq 0.05 \text{ km}^2$ 冰湖轮廓数据集
数据作者	张大弘, 宫鹏, 姚晓军, 刘娟
数据通信作者	姚晓军 (yaoxj_nwnu@163.com)
数据时间范围	1967–2014 年
地理区域	地理范围包括 26.86°–29.14°N, 85.39°–88.95°E, 北至雅鲁藏布江南部, 南到尼泊尔与印度交界处, 西临加德满都, 东达中、尼、印三国接壤地区。
空间分辨率	全色影像 1.8 m (Corona) / 15 m (Landsat ETM+/OLI), 多光谱影像 78 m (Landsat MSS) / 30 m (Landsat TM/ETM+/OLI); 地理坐标系为 WGS1984, 投影坐标系为 UTM。
数据量	143.94 MB
数据格式	ESRI Shapefile 文件 (压缩为 .zip 格式)
数据服务系统网址	http://www.sciencedb.cn/dataSet/handle/885
基金项目	国家自然科学基金 (41561016、41861013、41801052); 西北师范大学青年教师科研能力提升计划 (NWNLUKQN-14-4)。
数据库 (集) 组成	数据集共包括 2 个数据文件: (1) Basin_Koshi.zip 是科西河流域边界数据, 数据量 0.1 MB; (2) Glacial_lake_Koshi.zip 是 1967–2014 年科西河流域面积 $\geq 0.05 \text{ km}^2$ 的冰湖轮廓数据, 数据量 143.84 MB。

* 论文通信作者

姚晓军: yaoxj_nwnu@163.com

引言

冰湖溃决洪水 (Glacial Lake Outburst Flood, GLOF) 是高山冰川作用区常见自然灾害之一^[1-2], 主要由冰川阻塞湖和冰碛阻塞湖溃决所引发。就中国西藏自治区而言, 20 世纪以来发生的冰湖溃决洪水 (泥石流) 事件超过 30 次^[3]; 欧洲的阿尔卑斯山、南美洲的安第斯山、中亚的天山以及青藏高原周边高山地区也是冰川溃决洪水的主要分布区^[4]。冰湖溃决洪水有可能造成人员伤亡、牲畜死亡、淹没草场或农作物以及毁坏道路桥梁等损失, 因此防治冰湖溃决洪水灾害是当前这些区域的重要任务之一, 而冰湖分布及变化正是预测冰湖溃决洪水灾害的最基础数据。冰湖扩张或萎缩可以灵敏地反映气候变化^[5], 其面积或水量的变化能从一定程度上反映区域水资源分布的变化。因此, 构建较长时间序列的冰湖数据集对于研究冰湖分布与变化、预测冰湖溃决洪水灾害、合理利用水资源及探讨气候-冰川-冰湖间耦合关系具有重要意义。

关于冰湖的概念, 学者们给出的定义不尽相同^[6-12], 但都强调冰湖是冰川作用所形成的, 区别主要是冰川作用时期与湖泊物质来源。例如, 在《冰冻圈科学辞典》^[6]中, 冰湖被定义为由冰川作用形成的湖泊; 姚治君等^[12]认为冰湖是末次冰期以来冰川运动或退缩产生的融水在冰川表面、前部或侧部汇集形成的湖泊, 此定义强调冰湖与冰川间的地理位置关系; 姚晓军等^[9]从冰湖编目和冰湖溃决洪水角度提出冰湖是以现代冰川融水为主要补给源或在冰碛垄洼地内积水形成的天然水体, 并给出了冰湖类型及遥感判识指标。考虑冰湖识别的易判别性、可操作性及一致性, 本文采用姚晓军等^[9]提出的冰湖定义及其分类体系。

喜马拉雅山中段是冰湖溃决最频繁的地区之一^[13], 其中科西河流域上游历史上发生此类事件数量最多^[14-15]。作为恒河水系的重要支流, 科西河流经中国、尼泊尔和印度 3 国, 流域内具有世界上最大的高差, 最为完整的气候、土壤和植被带谱。在科西河流域, 分布有多座 8000 m 以上山峰, 现代冰川发育, 冰川覆盖率达 9.78%。据统计, 科西河流域历史上发生过 17 次较大规模的冰湖溃决事件, 其中中国境内高达 12 次^[6]。2015 年尼泊尔博克拉地区发生 8.1 级地震, Kargel 等^[17]对震后灾区分析时指出, 未来当该区域发生更大震级的地震或震源靠近冰湖分布区时, 发生严重冰湖溃决洪水灾害的可能性依然存在。除地震外, 雪 (冰) 崩、强降水、冰川强烈消融、冰碛坝内死冰消融等亦是引发冰湖溃决的诱因^[15,18]。因此, 建立科西河流域长时间序列的冰湖数据集可为分析该地区冰湖时空变化特征、冰湖危险性评估及冰湖溃决洪水 (泥石流) 灾害等研究提供基础数据支撑和科学依据。

1 数据采集和处理方法

1.1 数据来源与预处理

为获取科西河流域不同年代的冰湖数据, 本研究选用的遥感影像数据源主要包括 USGS/NASA (<http://earthexplorer.usgs.gov>) 提供的 14 景 Corona 卫星影像和 56 景 Landsat MSS/TM/ETM+/OLI 遥感影像 (表 1)。其中, Corona 卫星影像成像时间集中在 1960s, 是该时段可获得的质量较好的遥感影像; Landsat MSS/TM/ETM+ 遥感影像时间为 1970s、1980s、1990s、2000s, Landsat OLI 遥感影像则反映了 2013 年以来该流域冰湖状况。以 Landsat MSS/TM/OLI 数据作为主要数据源, 可以免费获取大量的数据, 且回访周期短, 利用成像时间接近的影像通过交叉验证解译方法可以避免云和山体阴影等因素的干扰, 提高冰湖解译精度。为保证数据的准确性, 本研究同时选用了该流域 1966–1980

年的 83 幅地形图，其中包括 4 幅 1:50 万、22 幅 1:10 万、48 幅 1:5 万和 7 幅 1:2.5 万地形图。受遥感影像空间分辨率限制，按 Landsat MSS 影像空间分辨率（78 m）计算， 0.006 km^2 是像元解译的理论阈值，考虑历史上该流域已溃决的冰湖面积普遍较大及冰湖面积扩张时长，本研究仅解译面积 $\geq 0.05 \text{ km}^2$ 的冰湖。此外，冰川轮廓矢量数据可以辅助识别或确定冰湖位置，研究区所用的冰川轮廓数据由科技部科技基础性工作专项“中国冰川资源及其变化调查”项目组和国际山地中心(ICIMOD)提供^[19]。

表 1 数据集所采用的遥感影像列表

影像编号/轨道号	卫星/传感器	空间分辨率/m	影像获取时间
DS1037-1039DA026	Corona KH-4	1.8 (全色)	1967-11-11
DS1037-1039DA027	Corona KH-4	1.8 (全色)	1967-11-11
DS1037-1039DA028	Corona KH-4	1.8 (全色)	1967-11-11
DS1044-1070DA095	Corona KH-4	1.8 (全色)	1967-11-11
DS1044-1070DA096	Corona KH-4	1.8 (全色)	1967-11-11
DS1044-1070DA097	Corona KH-4	1.8 (全色)	1967-11-11
DS1108-2217DA063	Corona KH-4	1.8 (全色)	1969-12-18
DS1108-2217DA064	Corona KH-4	1.8 (全色)	1969-12-18
DS1108-2217DA065	Corona KH-4	1.8 (全色)	1969-12-18
DS1108-2217DA066	Corona KH-4	1.8 (全色)	1969-12-18
DS1049-2136DA148	Corona KH-4	1.8 (全色)	1968-12-21
DS1108-2233DA055	Corona KH-4	1.8 (全色)	1968-12-21
DS1014-2218DA192	Corona KH-4	1.8 (全色)	1968-12-21
DS1014-2218DA068	Corona KH-4	1.8 (全色)	1972-04-21
139/040	Landsat MSS	78	1976-12-18
	Landsat TM	30	1989-12-28; 1990-01-13; 2000-10-07; 2000-11-08; 2011-08-27
	Landsat OLI	30	2013-06-13; 2014-01-07
139/041	Landsat MSS	78	1976-12-18
	Landsat TM	30	1989-10-25; 1989-11-10; 1999-09-19; 2000-12-26; 2011-08-27
	Landsat OLI	30	2013-06-13; 2013-11-20; 2013-12-06
140/040	Landsat MSS	78	1976-12-19
	Landsat TM	30	1989-01-17; 1992-09-30; 1999-11-29; 2000-09-12; 2000-09-28; 2010-10-26; 2010-11-11; 2010-04-09; 2010-04-25
	Landsat OLI	30	2013-11-11; 2014-05-22

影像编号/轨道号	卫星/传感器	空间分辨率/m	影像获取时间
140/041	Landsat MSS	78	1976-12-19; 1977-03-19
	Landsat TM	30	1989-01-17; 1992-09-22; 1999-10-28; 2000-10-30; 2009-10-15
	Landsat OLI	30	2014-01-05; 2014-05-22
141/040	Landsat MSS	78	1976-12-02
	Landsat TM	30	1991-11-30; 1993-03-08; 1999-12-06; 2000-10-05; 2000-11-22; 2009-06-08; 2009-10-30; 2010-10-01
	Landsat OLI	30	2013-12-04
141/041	Landsat MSS	78	1976-12-02
	Landsat TM	30	1991-11-30; 2001-02-26; 2001-12-27; 2009-05-07; 2010-11-02
	Landsat OLI	30	2014-04-11; 2014-05-13

基于光学遥感影像提取冰湖轮廓时易受云、雪及山体阴影的影响，因此在选取遥感影像时遵循少云多期原则，以达到多时段多角度的交叉解译目的，从而尽量避免漏画、错画问题。除 Corona 遥感影像需进行几何校正和图幅拼接外，获取的 Landsat 卫星系列遥感影像在数据生产时已经过初步预处理（系统辐射校正、地面控制点几何校正及基于 DEM 的地形校正），经检查同一目标地物在不同影像中无明显错位、变形的情况，而目视解译又属于定性判别地类，故这些遥感影像未做其他预处理。

1.2 研究区概况

科西河流域位于青藏高原南部、喜马拉雅山中段（ 26.86° – 29.14°N ， 85.39° – 88.95°E ），北抵雅鲁藏布江，南至尼泊尔与印度交界处，西临加德满都附近，东达中、尼、印三国接壤地区。流域内地势北高南低，海拔落差极大，主河全长 255 km，上游分 Arun（中国境内称为朋曲）、Sun Koshi（中国境内称为波曲）和 Tamor 三个支流（图 1）。流域东西长 340 km，南北宽 230 km，总面积约 $5.5 \times 10^4 \text{ km}^2$ ，位于中国和尼泊尔境内的面积分别占流域总面积的 51.96% 和 48.04%^[20]。流域内分布有 5 座海拔 8000 m 以上山峰，围绕这些山峰发育了众多的冰川及冰湖。受地形影响，研究区气候类型差异明显，自北向南依次分布着喜马拉雅山过渡带、高山带、亚高山带、温带、亚热带和热带等，后三种气候类型分布在河流中下游，比较适宜农业发展，分布着大量人口。据统计，尼泊尔境内总人口 434.59 万人，占流域总人口的 96.57%^[18]。

1.3 数据处理过程

冰湖作为一类自然水体，目前已有一些比较成熟的基于遥感影像的自动提取算法，如水体指数法、波段比值法、监督分类法及面向对象分割法等。由于本数据集采用的本底数据存在全色影像、地形图等多源数据，加之自动提取算法对影像质量要求较高，为保证数据采集精度，本数据集统一采用目视解译方法，具体步骤如下：

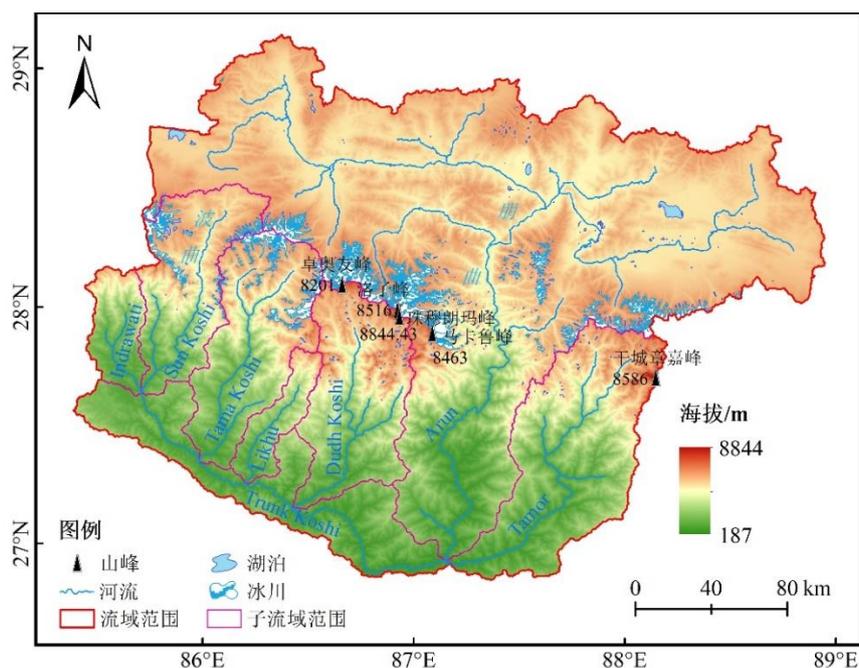


图 1 研究区概况图

(1) 获取解译底图：下载、收集并整理科西河流域的遥感影像、地形图等可提供解译对象信息的资料，按时间、类型、位置等归类，确保每个时段每个位置至少有一期影像覆盖。

(2) 确定解译方式：因冰湖具有季节性变化特征，一年之内最稳定的时段为 9–12 月，且少量影像（如 Landsat ETM+）局部存在坏带，故以其他相邻时相的影像作为补充，通过交叉检验解译方式获取精度较高的冰湖轮廓信息。

(3) 冰湖边界目视解译：以交叉检验解译的方式逐时期进行，显示比例尺控制在 1:2000 左右，沿水陆、冰陆或者冰雪界线像元中心数字化冰湖边界，精度控制在一个像元以内。

(4) 确定冰湖与非冰湖：采用姚晓军等^[4]提出的冰湖定义、冰湖与非冰湖的界定方法，逐一对不同时期提取到的研究区湖泊数据进行判定，剔除由大气降水、地下水等形式补给的非冰湖，包括研究区内大量存在的古冰蚀湖，最终得到研究区不同时期的冰湖数据集。

(5) 冰湖数据质量检查：逐一检查各冰湖轮廓图形数据、对应的影像数据及属性数据等，确保无漏画和错画、各数据项无遗漏和录入正确等。

(6) 冰湖类型识别：将最终获得的冰湖数据导入 Google Earth 软件，根据冰湖分类体系^[4]，逐一识别出所有冰湖的类型并写入属性表。

2 数据样本描述

图 2 展示了科西河流域冰湖整体分布情况。冰湖多边形矢量数据以 `gl_XXXX` 形式命名，其中 `XXXX` 表示年份；冰湖点矢量数据 `gl_LocationCode` 给出了流域内所有的冰湖位置、编码、类型等信息。至 2014 年，流域内共发育冰湖 221 个，其中冰川侵蚀湖 83 个（冰斗湖 10 个，其他冰川侵蚀湖 73 个），冰碛阻塞湖 128 个（终碛阻塞湖 115 个，侧碛阻塞湖 13 个），冰面湖 4 个，其他冰川湖 6 个。以面积持续扩张的冰湖 `GL085838E28322N` 为样本，图 3 展现了该湖泊在研究时段内的边界、周长及面积等的变化情况。图 4 给出了数据集中面积先减小后持续扩张的龙巴萨巴湖不同时期的边界。

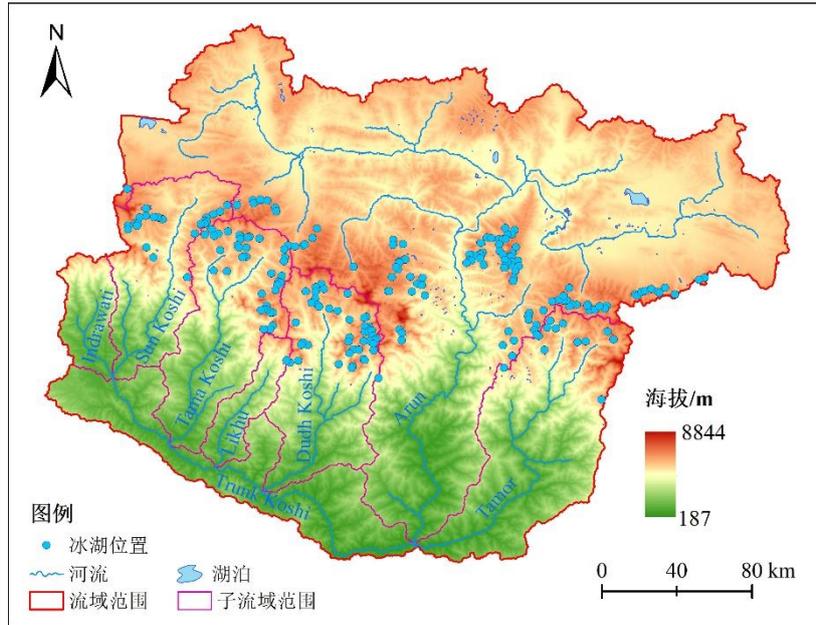


图 2 冰湖分布数据集示意图

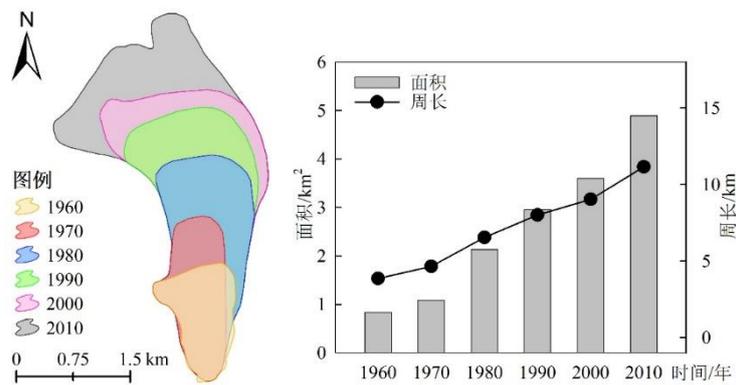


图 3 冰湖 (GL085838E28322N) 轮廓及其面积与周长变化

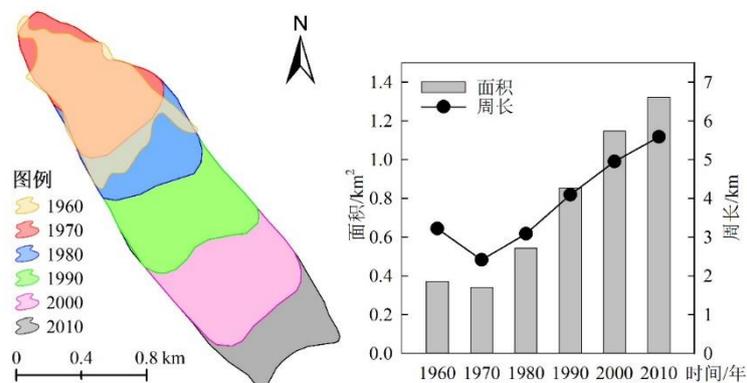


图 4 龙巴萨巴湖 (GL088075E27946N) 轮廓及其面积与周长变化

在本数据集中，每个冰湖对象由图形数据和属性数据两部分组成，其中属性数据包括要素标识码 (FID)、形状类型 (Shape)、冰湖编码 (Code)、冰湖名称 (Name_gl)、冰湖周长 (Perimeter_gl)、

冰湖面积 (Area_gl) 以及冰湖类型 (Type_gl) 等。其中, 冰湖编码结构为 GLnnnnnnEmmmmmN, 编码中的 GL 表示冰湖, n 表示冰湖质心经度, m 表示冰湖质心纬度, E、N 分别代表东经和北纬, n 的前三位和 m 的前两位分别表示经度、纬度的整数部分, 后三位表示经纬度小数前三位。为保证冰川编码的一致性, 以 2010s 时期冰湖质心为准。如冰湖在 2010s 时期不存在, 则以第一次出现年代对应的冰湖质心进行编码。

3 数据质量控制和评估

解译时依据姚晓军等^[4]提出的冰湖定义、冰湖与非冰湖的界定方法, 遵循马荣华^[21]、宫鹏^[20]等所使用的湖泊边界遥感判译原则, 精度控制在一个像元以内。由于参与本数据集解译的长时序多源遥感影像空间分辨率存在差异, 如 Corona KH-4 影像空间分辨率为 1.8 m (全色), Landsat MSS 影像为 78 m (多光谱), Landsat TM 影像为 30 m (多光谱), Landsat ETM+ 和 OLI 影像为 30 m (多光谱) 和 15 m (全色), 因此本数据集的精度虽不及其他高/甚高空间分辨率、单一数据源生产的数据, 整体精度仅能保证在最低空间分辨率 78 m 左右。但从数据时序长度来看, 本数据集可以满足大多数相关领域的研究要求。

4 数据价值

科西河流域分布有数量众多的湖泊, 因此通过区分冰湖与非冰湖以及识别各个冰湖的类型所得到的冰湖数据及湖泊类别属性是本数据集的价值之一, 这可以为当地相关部门以及从事冰湖溃决危险性评价工作的学者们提供宝贵的参考。

本数据集是依据不同时期多源遥感影像数据 (1967–2014 年) 生产的 6 期冰湖边界矢量数据, 特点在于较长时间序列, 可作为科西河流域冰湖时空变化、冰湖水量变化评估等研究的基础数据, 也可作为该区域冰湖溃决危险性评估、水资源变化、气候变化及冰湖与冰川间变化响应关系等研究的重要依据。

5 数据使用方法和建议

本数据集解压后为 ESRI Shapefile 文件, 可使用支持该文件格式的 GIS 与遥感软件读取与编辑。数据集直接反映了 1967–2014 年科西河流域的冰湖分布特征, 进一步分析处理可获得流域内冰湖的时空变化特征等信息, 结合相关资料, 亦可用作冰湖危险性评估及冰湖溃决洪水 (或泥石流) 灾害的基础数据。另外, 如果还有其他来源的数据, 可对本数据做更新修改, 也可用最新的数据更新 2010s 数据, 以延长数据时序。

数据作者分工职责

张大弘 (1993—), 男, 甘肃平凉人, 硕士研究生, 主要研究方向为 GIS 设计与开发。主要承担工作: 基础数据收集、处理及论文撰写。

宫鹏 (1990—), 男, 甘肃礼县人, 硕士, 主要研究方向为 GIS 与冰冻圈变化。主要承担工作: 基础数据处理、目视解译。

姚晓军（1980—），男，山西夏县人，博士，副教授，主要研究方向为地理信息技术与冰冻圈变化。主要承担工作：总体方案设计、冰湖类型确定。

刘娟（1993—），女，甘肃定西人，硕士研究生，主要研究方向为环境科学与地理信息系统。主要承担工作：数据集后期处理、冰湖类型识别。

参考文献

- [1] RICHARDSON S D, REYNOLDS J M. An overview of glacial hazards in the Himalayas[J]. *Quaternary International*, 2000, 65/66: 31-47
- [2] OSTI R, EGASHIRA S. Hydrodynamic characteristics of the Tam Pokhari glacial lake outburst flood in the Mt. Everest region, Nepal[J]. *Hydrological Processes*, 2009, 23(20): 2943-2955.
- [3] 姚晓军, 刘时银, 孙美平, 等. 20 世纪以来西藏冰湖溃决灾害事件梳理[J]. *自然资源学报*, 2014, 29(8): 1377-1390.
- [4] CUI P, DANG C, CHENG Z, et al. Debris flows resulting from glacial-lake outburst floods in Tibet, China[J]. *Physical Geography*, 2010, 31(6): 508-527.
- [5] 王欣, 刘时银, 莫宏伟, 等. 我国喜马拉雅山区冰湖扩张特征及其气候意义[J]. *地理学报*, 2011, 66(7):895-904.
- [6] 秦大河, 姚檀栋, 丁永建, 等. 冰冻圈科学辞典[M]. 北京: 气象出版社, 2014: 2-54.
- [7] 中国科学院-水利部成都山地灾害与环境研究所, 西藏自治区交通厅科学研究所. 西藏泥石流与环境[M]. 成都: 成都科技大学出版社, 1999: 79.
- [8] 曹学诚, 刘周周, 李维胜. 尼泊尔 4·25 地震震前冰湖制图与潜在危险性分析[J]. *冰川冻土*, 2016, 38(3): 573-583.
- [9] 姚晓军, 刘时银, 韩磊, 等. 冰湖的界定与分类体系——面向冰湖编目和冰湖灾害研究[J]. *地理学报*, 2017, 72(7): 1173-1183.
- [10] 唐升贵, 刘发祥, 赵振远. 西藏念青唐古拉山东段冰湖的时空分布特征及其影响因素探讨[J]. *科技资讯*, 2014, 12(16): 35-36.
- [11] 陈晨, 郑江华, 刘永强, 等. 近 20 年中国阿尔泰山区冰川湖泊对区域气候变化响应的时空特征[J]. *地理研究*, 2015, 34(2): 270-284.
- [12] 姚治君, 段瑞, 董晓辉, 等. 青藏高原冰湖研究进展及趋势[J]. *地理科学进展*, 2010, 29(1): 10-14.
- [13] REYNOLDS J M. Glacial-lake outburst floods (GLOF) in the Himalayas: an example of hazard mitigation from Nepal[J]. *Geoscience and Development*, 1995, 2: 6-8.
- [14] LIU C H, SHARMA C K. Report on First Expedition to Glaciers and Glacier Lakes in the Pumqu (Arun) and Poqu (Bhote-SunKosi) River basins, XiZang (Tibet), China[M]. Beijing: Science Press, 1988.
- [15] RICHARDSON S D, REYNOLDS J M. An overview of glacial hazards in the Himalayas[J]. *Quaternary International*, 2000, 65/66, 31-47.
- [16] 邓伟, 张懿铨. 气候变化下 Koshi 河流域资源、环境与发展[M]. 成都: 四川科学技术出版社, 2014:82.
- [17] KARGEL J S, LEONARD G J, SHUGAR D H, et al. Geomorphic and geologic controls of geohazards induced by Nepal's 2015 Gorkha earthquake[J]. *Science*, 2016, 351(6269): aac8353.

- [18] YAMADA T, SHARMA C K. Glacier lakes and outburst floods in the Nepal Himalaya[J]. IAHS, 1993, 218: 319-330.
- [19] 刘时银, 姚晓军, 郭万钦, 等. 基于第二次冰川编目的中国冰川现状[J]. 地理学报, 2015, 70(1): 3-16.
- [20] 宫鹏, 姚晓军, 孙美平. 1967–2014 年科西河流域冰湖时空变化[J]. 生态学报, 2017, 37(24): 8422-8432.
- [21] 马荣华, 杨桂山, 段洪涛, 等. 中国湖泊的数量、面积与空间分布[J]. 中国科学: 地球科学, 2011, 41(3): 394-401.

论文引用格式

张大弘, 宫鹏, 姚晓军, 刘娟. 1967–2014 年科西河流域面积 $\geq 0.05 \text{ km}^2$ 冰湖轮廓数据集[J/OL]. 中国科学数据, 2020, 5(2). (2020-05-21). DOI: 10.11922/csdata.2019.0064.zh.

数据引用格式

张大弘, 宫鹏, 姚晓军, 刘娟. 1967–2014 年科西河流域面积 $\geq 0.05 \text{ km}^2$ 冰湖轮廓数据集[DB/OL]. Science Data Bank, 2019. (2019-09-05). DOI: 10.11922/sciencedb.885.

A dataset of outlines of glacial lakes with an area of no less than 0.05 km^2 in the Koshi River basin from 1967 to 2014

Zhang Dahong¹, Gong Peng¹, Yao Xiaojun^{1*}, Liu Juan¹

1. College of Geography and Environmental Science, Northwest Normal University, Lanzhou 730070, P.R. China

* Email: yaoxj_nwnu@163.com

Abstract: There are numerous glaciers and glacial lakes in the Koshi River basin. The coverage rate of glaciers reaches 9.78% and the number of glacial lakes is 221. The distribution and change of glacial lakes in this region are the most important factors for predicting the disaster of glacial lake outburst floods (GLOFs), and can also reflect the changes of climate and glacial water resources in alpine frigid areas to a certain extent. The dataset includes the vectorized outlines of glacial lakes with an area of no less than 0.05 km^2 during six periods in the Koshi River basin, that is, 1960s, 1970s, 1980s, 1990s, 2000s and 2010s, which were acquired by using multi-source remote sensing image (Corona, Landsat MSS/TM/ETM+/OLI) and topographic maps to conduct a visual interpretation. We controlled the accuracy of interpretation within one pixel. Lying between 26.86° – 29.14°N and 85.39° – 88.95°E , the study area is composed of several sub-basins including Indrawati, Sun Koshi, Tama Koshi, Likhu, Dudh Koshi, Arun and Tamor. The dataset has a long time series, type and geometric properties of glacial lakes, which can be used as the basic data for assessing the temporal-spatial changes of glacial lakes and the changes in water volume of glacial lakes in the Koshi River basin. In addition, it can also be used as an important basis for the risk assessment of GLOFs, the changes of water resources, climate change and the relationship

between glacial lakes and glaciers in this region.

Keywords: glacial lake; Koshi River; visual interpretation; Tibetan Plateau

Dataset Profile

Title	A dataset of outlines of glacial lakes with an area of no less than 0.05 km^2 in the Koshi River basin from 1967 to 2014
Data corresponding author	Yao Xiaojun (yaoxj_nwnu@163.com)
Data authors	Zhang Dahong, Gong Peng, Yao Xiaojun, Liu Juan
Time range	1967 – 2014
Geographical scope	Latitude & longitude ($26.86^\circ - 29.14^\circ\text{N}$, $85.39^\circ - 88.95^\circ\text{E}$); located in the Central Himalayas; The specific areas: north to the south of the Yarlung Zangbo River, south to the border of Nepal and India, West to Kathmandu in Nepal, East to the borders of China, Nepal and India.
Spatial resolution	Panchromatic image: 1.8 m (Corona) / 15 m (Landsat ETM+/OLI); Multispectral image: 78 m (Landsat MSS) / 30m (Landsat TM/ETM+/OLI); GCS: WGS1984, PCS: UTM.
Data volume	143.94 MB
Data format	ESRI Shapefile file (compressed in *.zip format)
Data service system	< http://www.sciencedb.cn/dataSet/handle/885 >
Sources of funding	National Natural Science Foundation of China (41561016, 41861013, 41801052); Youth Scholar Scientific Capability Promoting Project of Northwest Normal University (NWNLU-LKQN-14-4).
Dataset composition	The dataset consists of 2 subsets: (1) Basin_Koshi.zip is made up of boundary data of Koshi River basin, with a data volume of 0.1 MB; (2) Glacial_lake_Koshi.zip is made up of vectorized outlines of glacial lakes with an area of no less than 0.05 km^2 in the Koshi River basin from 1967 to 2014, with a data volume of 143.84 MB.