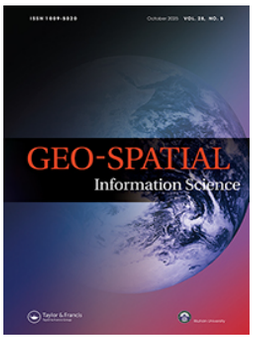


Geo-spatial Information Science

ISSN: 1009-5020 (Print) 1993-5153 (Online) Journal homepage: www.tandfonline.com/journals/tgsi20

A 10-m glacial lake dataset along the Sichuan-Tibet Railway using Sentinel-2 images

Qian Deng, Yong Nie, Jonathan L. Carrivick, Gordon G. D. Zhou, Qiyuan Lyu & Jida Wang

To cite this article: Qian Deng, Yong Nie, Jonathan L. Carrivick, Gordon G. D. Zhou, Qiyuan Lyu & Jida Wang (23 Dec 2025): A 10-m glacial lake dataset along the Sichuan-Tibet Railway using Sentinel-2 images, Geo-spatial Information Science, DOI: 10.1080/10095020.2025.2601921

To link to this article: https://doi.org/10.1080/10095020.2025.2601921

Published online: 23 Dec 2025.

GEO-SPATIAL INFORMATION SCIENCE
https://doi.org/10.1080/10095020.2025.2601921

Taylor & Francis
Taylor & Francis Group

A 10-m glacial lake dataset along the Sichuan-Tibet Railway using Sentinel-2 images

Qian Deng [a,b], Yong Nie [a,b], Jonathan L. Carrivick [c], Gordon G. D. Zhou [a], Qiyuan Lyu [a] and Jida Wang [d,e]

[a]Key Laboratory of Mountain Hazards and Engineering Resilience, Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu, China; [b]College of Resource and Environment, University of Chinese Academy of Sciences, Beijing, China; [c]Water@leeds, School of Geography and, University of Leeds, Leeds, UK; [d]Department of Geography and Geographic Information Science, University of Illinois Urbana-Champaign, Urbana, IL, USA; [e]Department of Geography and Geospatial Sciences, Kansas State University, Manhattan, KS, USA

ABSTRACT

The loss of glacier mass and the increase in glacial lakes in High Mountain Asia due to climate change led to water-related crises and hazards, such as glacial lake outburst floods (GLOFs). The Sichuan-Tibet Railway (STR), which is crucial for the development of West China, is facing challenges due to these water-related crises. Therefore, there is an urgent need for a high-quality dataset of glacial lakes along the STR corridor, which remains lacking. This study employed a semi-automated lake mapping method, accompanied by rigorous quality assurance and quality control, to create a comprehensive, high-resolution dataset of glacial lakes in approximately 2021, covering the entire STR corridor. The 10-m resolution glacial lake dataset comprises 17,566 glacial lakes, covering a total area of 582.89 ± 31.45 km^2. This dataset features a larger number of lakes yet exhibits a lower mapping error compared to existing inventories in this region. The high-quality glacial lake dataset has the potential to benefit various applications, including assessing lake mapping accuracy, training deep learning algorithms, evaluating water sources, assessing risks related to glacial lakes, and understanding the interactions between glaciers and lakes.

ARTICLE HISTORY
Received 3 October 2024
Accepted 5 December 2025

KEYWORDS
Glacial lake outburst floods; lake mapping; tibetan plateau; remote sensing; GIS

1. Introduction

Glaciers and glacial lakes are sensitive indicators of climate change (Nie et al. 2017; Shugar et al. 2020; Zheng et al. 2021). Climate change is causing glaciers to lose mass at an accelerating rate (Brun et al. 2017; Lee et al. 2021; Shean et al. 2020) and leading to the expansion of glacier lakes in High Mountain Asia (HMA) in terms of both numbers and areas (Chen et al. 2021; Furian, Maussion, and Schneider 2022; Wang et al. 2020). These rapid changes affect water resources and related hazards (Nie et al. 2021; Pan et al. 2025). For instance, there has been an increasing trend in outburst floods from glacial lakes in HMA, posing a potential threat to downstream populations and infrastructure (Nie et al. 2023; Zheng et al. 2021). One of the significant infrastructure projects currently under construction is the Sichuan-Tibet Railway (STR), which will span 1580 km from east to west (Lu and Cai 2019; Xue et al. 2021). The STR aims to connect Chengdu and Lhasa (Figure 1) with a total investment exceeding $44 billion. Once completed, the STR will be crucial for the economic development of Western China. However, it is also expected to face challenges related to water resources and water-related hazards.

Glacial lakes were primarily formed due to deglaciation (Carrivick and Tweed 2013). They are dynamic, usually expanding as the glaciers melt, but some lakes can drain to form glacial lake outburst floods (GLOFs) (Aggarwal et al. 2017; Carrivick and Quincey 2014; Carrivick and Tweed 2019; Carrivick, Sutherland, et al. 2022). Glacial lake datasets are essential for understanding this dynamism and assessing water resources and GLOF hazards (Lesi et al. 2022a; Sahu et al. 2023). In alpine environments where field investigations are challenging, satellite remote-sensing images are the most viable method for producing glacial lake datasets (Li et al. 2025; Nie et al. 2017; Zhang et al. 2025). Most open-access glacial lake datasets have been generated using Landsat images with a spatial resolution of 30 m at regional scales (Chen et al. 2021; Wang et al. 2020; Zhang et al. 2015). The freely available Sentinel-2 imagery, which provides a finer spatial resolution, enables the identification of more glacial lakes with more precise boundaries compared to Landsat-derived glacial lake outlines (Lesi et al. 2022a). Despite the availability of 10-m Sentinel-2 images since late 2015 (Drusch et al. 2012), a 10-m glacial lake dataset is still absent in the STR region.

CONTACT Yong Nie nieyong@imde.ac.cn

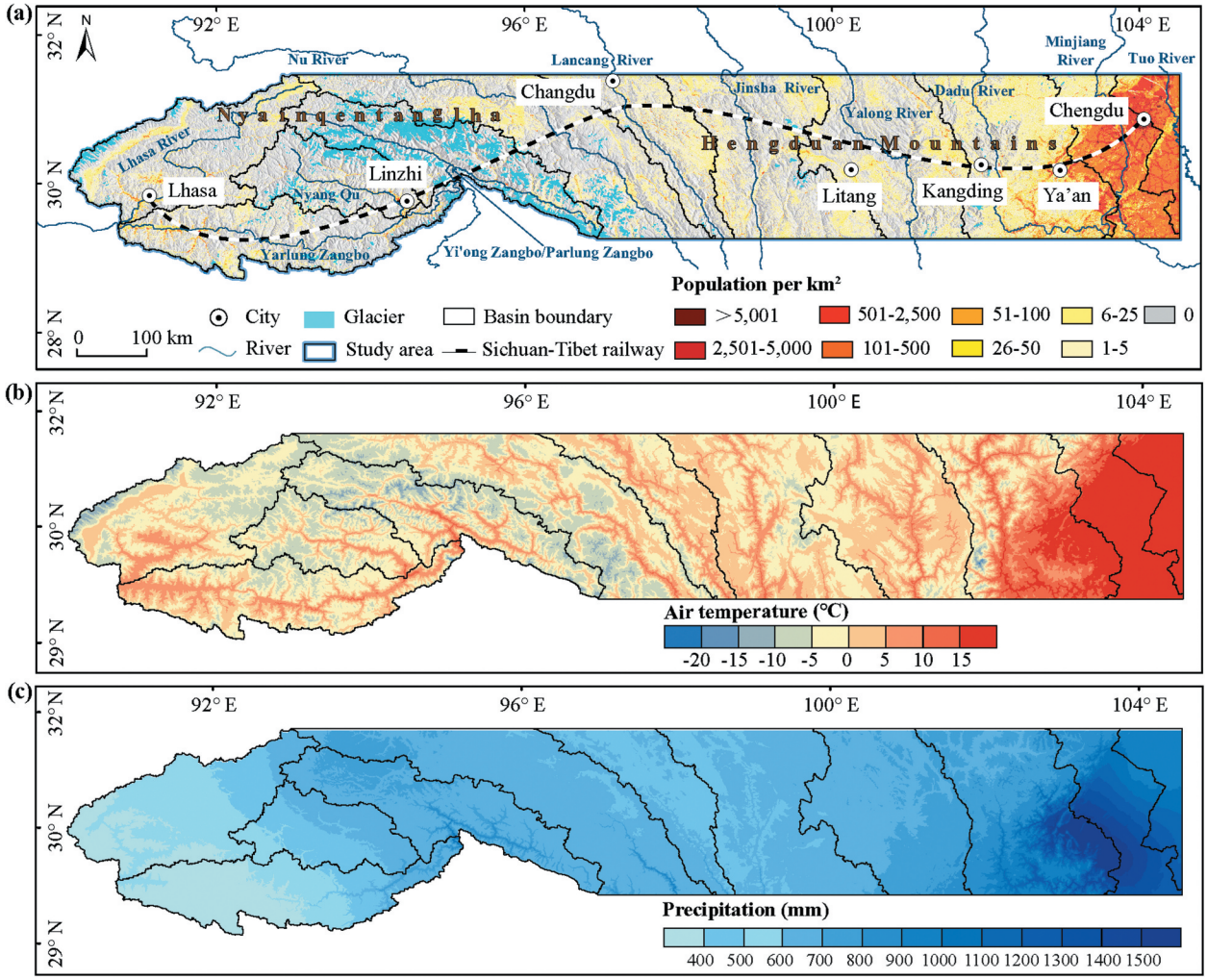


Figure 1. Location of the study area associated with the population distribution (Rose et al. 2020), rivers, glaciers (RGI Consortium 2017), mountains (a), air temperature (b), and precipitation (c) along the sichuan-Tibet Railway.

A comprehensive glacial lake dataset requires a thorough quality inspection to ensure accuracy and reduce mapping errors resulting from image quality and processing procedures (Lesi et al. 2022a; Nie et al. 2017). Mapping glacial lakes using optical satellite images presents several challenges, including the effects of mountain shadows, cloud cover, and snow, which need to be addressed (Lesi et al. 2022a; Sahu et al. 2023; Wang et al. 2025; Wangchuk and Bolch 2020). Various mapping techniques, including manual delineation (Wang et al. 2020), multi-spectral index classification (How et al. 2021), and deep learning (Pi et al. 2022; Zhang et al. 2026), have been employed to tackle these issues. Glacial lake products, encompassing attributes like lake area, volume, and type, are valuable for researchers studying glacier-lake interactions (Carrivick, How, et al. 2022; Carrivick, Sutherland, et al. 2022; King et al. 2018; Liu et al. 2020; Maurer et al. 2019; Sutherland et al. 2020), water resources, and risk assessment for GLOFs (Dubey and Goyal 2020). Categorizing glacial lakes based on their spatial relationship with glaciers enables the prediction of glacier evolution under different climate change scenarios (Kraaijenbrink et al. 2017).

This study aims to develop a detailed and accurate dataset of glacial lakes that covers all regions related to STR. The 10-m glacial lake dataset (STR-GL-2021) can be utilized to create advanced deep-learning models for lake mapping and to verify automated lake mapping products. Additionally, this dataset can potentially be used for research on water resource assessment, hazards related to glacial lakes, and modeling the evolution of glacier lakes.

2. Materials and methods

2.1. Study area

Given the importance of the STR and its surrounding resources and environment, we have designed a study area that encompasses the entire STR region. Our study area encompasses 312,000 km^2 and includes all 11 STR-related basins (Lehner and Grill 2013). This comprises three complete drainage basins (Lhasa River, Nyang Qu, and Yi'ong Zangbo/Parlung Zangbo) and eight partial basins (Yarlung Zangbo, Nu River, Lancang River, Jinsha River, Yalong River, Dadu River, Minjiang River, and

Tuo River) (Figure 1). Specifically, this study area was defined by considering the spatial interaction between GLOFs and the STR. We selected the northernmost latitude of the Lhasa River basin (31.29° N) and the southernmost latitude of the Yi'ong Zangbo/Parlung Zangbo basin (29.14° N) as reference points to delineate the eight incomplete north-south river basins.

The STR study area is characterized by complex topography and active tectonic structure, encompassing the Hengduan Mountains, Nyainqentanglha, and Himalaya Mountains (Lu and Cai 2019; Xue et al. 2021). The study area is located at the intersection of the westerlies, Indian monsoon, and southeast monsoon (Nie et al. 2021). The annual mean air temperature across the study area was 1.5°C between 1990 and 2021 (Peng 2024a), and air temperature has been increasing at a rate of 0.2°C per decade over the past four decades (Zhang et al. 2021). The annual precipitation in the study area is 691 mm, with higher values in the eastern section compared to the western section (Peng 2024b). A population of 27 million is distributed unevenly across the study area, with population density decreasing gradually from east to west (Rose et al. 2020).

2.2. Data

We utilized 144 Sentinel-2 images from around 2021 to map glacial lakes. The images were sourced from the Copernicus Open Access Hub (https://scihub.copernicus.eu/., accessed 15 December 2023). Due to heavy cloud cover and snow in the study area, capturing sufficient images to cover the entire region in a single year (e.g. 2021) was not possible. We initially chose Sentinel-2 products from 2021 as a benchmark for each tile region (Figure 2), prioritizing those with the lowest cloud and snow cover, all of which are under 22% (supplementary documents). Seventy-five out of the 144 images (52%) were taken in 2021. Consequently, the glacial lake dataset is labeled "STR-GL-2021." Next, we selected images from 2020 to fill the gap left by cloud and snow cover affecting the area in 2021. We gradually extended the acquisition year of Sentinel-2 products back to 2017 to guarantee complete coverage of the study area with high-quality images. The images from 2020 and 2021 account for 88% of the total count (Figure 3), while the remaining 12% were collected between 2017 and 2019. Most of the images (76%) were acquired in September, October, and November, indicating that our glacial lake data reflect their status in the autumn. This focus on the same season significantly minimized the impact of seasonal changes in glacial lakes on the mapping data. The average cloud coverage percentage of the Sentinel-2 images used in this study is 10% (supplementary documents).

2.3. Methods

This study mapped glacial lakes using an NDWI-threshold-based method and a thorough user-interactive quality assurance process. Creating a dataset for glacial lakes involves three steps (Figure 4): image data processing and calculation, user-interactive inspection and revision, and validating and adding lake attributes.

2.3.1. Step 1: image data processing and calculation

We processed the image data and calculated the masks and waterbodies. We used Interactive Data Language (IDL) codes to unzip all Sentinel-2 raw images automatically. We selected four 10 m bands (blue, green, red, and near-infrared (NIR)) and two 20 m short-wave infrared bands (SWIR1 and SWIR2, resampled to 10 m) to stack and create new named images (Drusch et al. 2012). We made cloud and cloud shadow masks using the UKIS-CSMASK algorithm and stacked images (Wieland, Fichtner, and Martinis 2022). We set the NDWI threshold at 0.1 to extract all waterbodies initially and automatically (Mcfeeters 1996), which included non-glacial lakes as well as incorrect or incomplete glacial lake outlines at this stage. In our study, we set the minimum mapping unit (MMU) to nine pixels (0.0009 km^2) (Chen et al. 2021; Nie et al. 2017). We removed the extracted waterbodies intersecting with cloud and

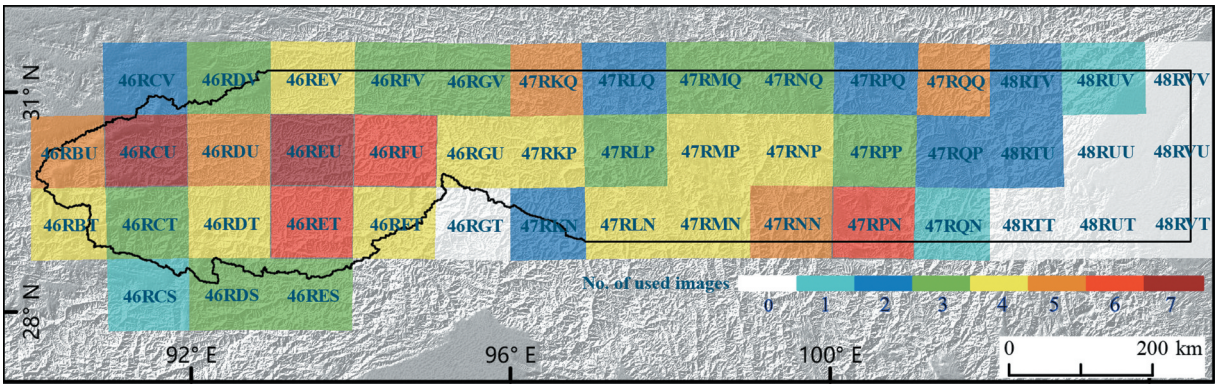


Figure 2. Spatial distribution and number of Sentinel-2 products used in the study, tagged by their tile code in the military grid system. The number of unused images indicates that there is no glacial lake at the tile location.

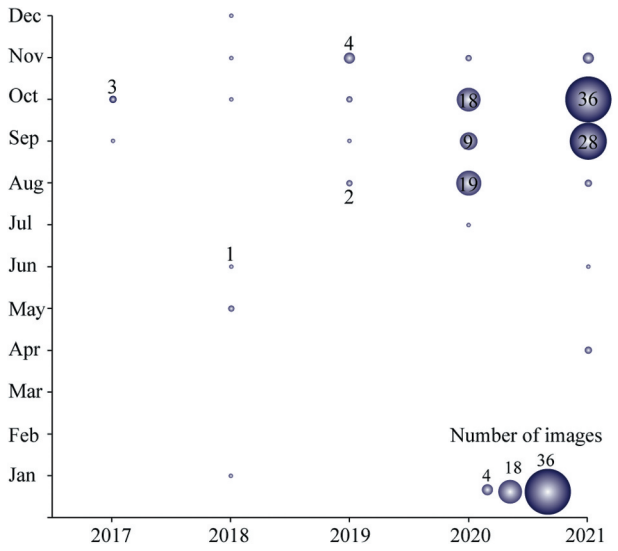


Figure 3. Yearly and monthly distribution of Sentinel-2 images used in the study.

cloud shadow masks and converted the remaining ones into vector files.

2.3.2. Step 2: user-interactive inspection and revision

A thorough quality assurance process was used to identify glacial lakes from vectorized waterbodies. This process involved removing non-glacial lakes and adding any previously omitted glacial lakes that were not included in the initial dataset. Non-glacial lakes (Table 1), including rivers, landslide-dammed lakes, artificial lakes, and reservoirs, were excluded using overlapping images combined with expert experience and field investigations, as they are not related to glaciation (Lesi et al. 2022a). Some glacial lakes were either omitted or had inaccurate outlines due to cloud cover, topographic shadows, and snow, as observed in Sentinel-2 imagery and existing glacial lake datasets (Wang et al. 2020; Wei et al. 2021). To address these issues, the NDWI-threshold-based interactive lake mapping tool (Lesi et al. 2022a; Li and Sheng 2012; Nie et al. 2017, 2020; Wang, Sheng, and Tong 2014) was utilized to recreate the omitted or incorrect outlines of glacial lakes using Sentinel-2 images. One person operated the user-interactive mapping tool while a qualified professional verified the results (Table 2). This method enhances the consistency and accuracy of glacial lake mapping, helping to prevent errors that may occur when multiple operators are involved. This tool involves three steps: (1) calculating NDWI in a region of interest (ROI, the blue rectangle in Figure 4), (2) setting the optimal threshold to extract the raster lake outline based on a bimodal (two-peak) distribution of the pixel NDWI values (Wang, Sheng, and Tong 2014), (3) converting the lake raster to vector for further editing in step 3. The NDWI threshold was determined on a case-by-case basis and visually validated by superimposing the source image. An NDWI threshold above 0.1 is generally applicable to most glacial lakes.

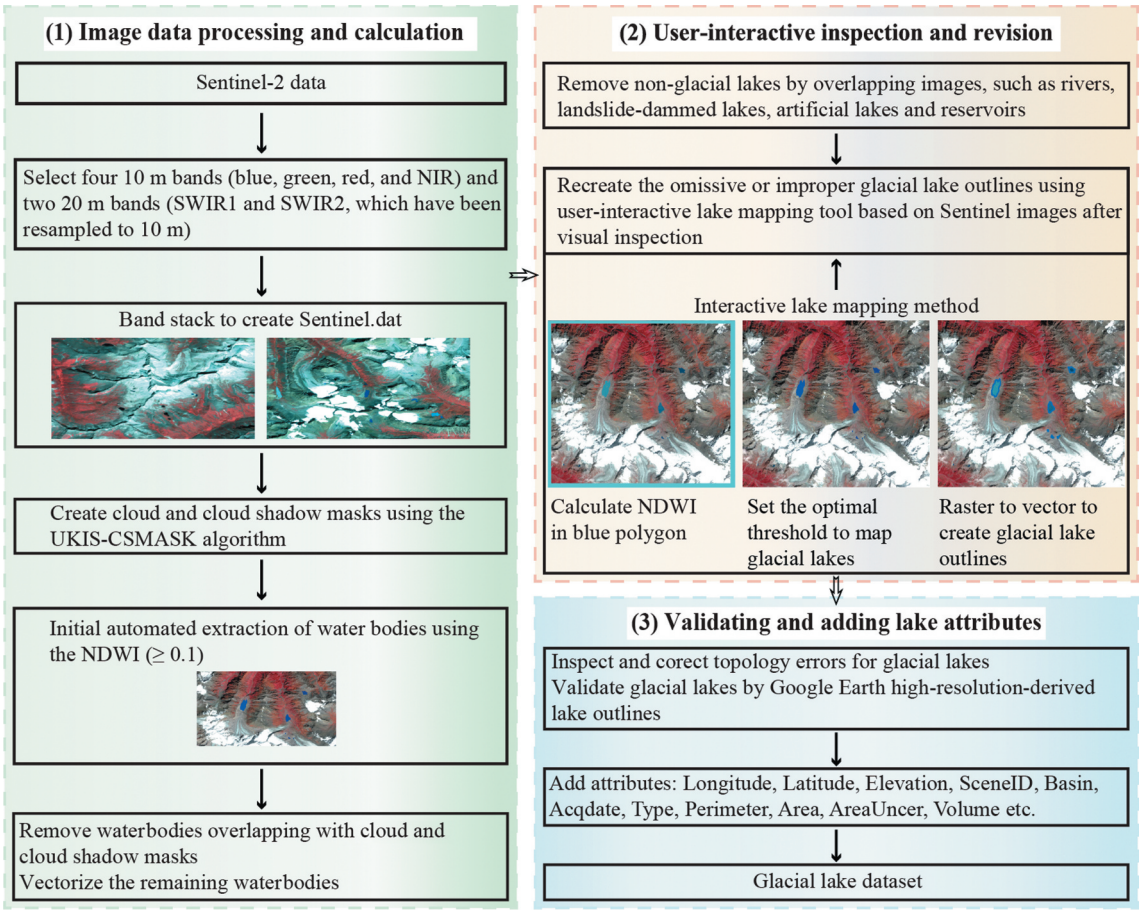


Figure 4. Workflow of the glacial lake inventory in this study.

Table 1. Our glacial lake dataset excluded typical landslide-dammed lakes and artificial lakes/reservoirs. Images are from Sentinel-2 and Google Earth.

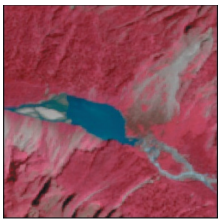
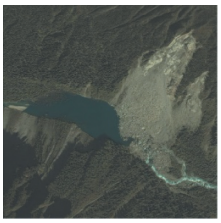
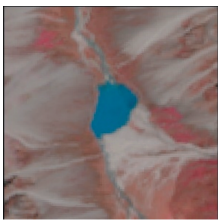
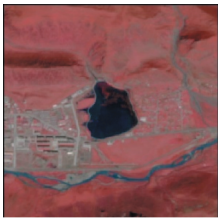
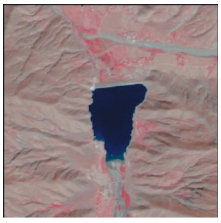
Lake type	Description	Sentinel-2	Google Earth
Landslide-dammed lake	Lakes formed behind a collapsing rock or sliding land dam. Case location: 30°05′53.12″ N 94°29′07.22″ E		
	Lakes formed behind debris-flow-induced dams. Case location: 29°22′41.95″ N 97°30′15.73″ E		
Artificial lake or reservoir	Lakes are heavily affected by human activities. Case location: 30°38′35.52″ N 93°14′43.65″ E		
	Lakes are artificially constructed, such as hydropower reservoirs. Case location: 29°07′46.63″ N 91°06′20.95″ E		

Table 2. The attributes of our glacial lake dataset in circa 2021 were derived from Sentinel-2 images.

Field name	Type	Description	Note
FID	Object ID	Unique code of the glacial lake	Number
Shape	Geometry	The type of glacial lake	Polygon
Longitude	Double	Longitude at the centroid of a glacial lake	Decimal degree
Latitude	Double	Latitude at the centroid of a glacial lake	Decimal degree
Elevation	Long	Elevation at the centroid of a glacial lake	Unit: meters above sea level
SceneID	String	Tile code and date of a source image	Txxxxx_YYYYMMDD
Acqdate	String	The acquisition date of the source image	YYYYMMDD
Basin	String	The basin name where the glacial lake is in	
Type	String	Glacial lake type	Supraglacial (SG), ice-contact (IC), unconnected-glacier-fed (UGF), non-glacier-fed (NGF)
Perimeter	Double	The perimeter of a glacial lake boundary	Unit: m
Area	Double	The area of a glacial lake	Unit: m^2
AreaUncer	Double	Area uncertainty of a glacial lake	Unit: m^2
RAreaUncer	Double	Relative area uncertainty of a glacial lake	Unit: %
Volume	Double	The water volume of a glacial lake	Unit: m^3
Operator	String	The operator of glacial lake mapping	Qian Deng
Examiner	String	The examiner of glacial lake mapping	Yong Nie

2.3.3. Step 3: validating and adding lake attributes

This stage involved inspecting and correcting topology errors in glacial lake outlines in GIS, including removing repeated polygons and eliminating small, sliver-shaped polygons.

We validated our glacial lake data by randomly selecting 165 lakes, stratified by size, reflecting the decreasing frequency of larger glacial lakes (Figure 5(b)). The outlines were manually created using high-resolution (~3 m) images from Google Earth at a similar age. Sixteen attribute fields (Table 2) were populated for each glacial lake in our dataset. A glacial lake's centroid (longitude and latitude) was automatically calculated using the spatial analysis function of GIS. Lake elevation was determined by the centroid of each glacial lake associated with a 1 arc-second SRTM DEM (Nasa

2013). The image source (SceneID) was derived from an image file name consisting of the Sentinel-2 product tile code in the military grid system and the image acquisition date. Acqdate is the acquisition date of Sentinel-2 source images. The classification of glacial lakes (Type) was established based on the spatial relationship between glaciers and glacial lakes (Lesi et al. 2022a). Glacial lakes were grouped into four categories using the characterization and examples provided in Table 3, interpreted from Sentinel-2 and Google Earth images, with the aid of the RGI 6.0 glacier outlines (RGI Consortium 2017). The perimeter and area of the lake were calculated automatically. The sub-basin was automatically determined by overlapping the basin boundaries in the study area. Lake water volume (Volume) was calculated using the area-volume empirical function (Cook and Quincey 2015). A glacial lake's absolute and relative area uncertainty was estimated based on the improved Hanshaw's function (Hanshaw and Bookhagen 2014; Lesi et al. 2022a). The equations used are listed below, and the detailed calculation process is available in the literature (Lesi et al. 2022a):

$$\mathrm{Error}(1\sigma) = \left(\frac{P}{G} - N_{\mathrm{Inner}}\right) \times 0.6872 \times \frac{G^2}{2} \quad (1)$$

$$D = \frac{\mathrm{Error}(1\sigma)}{A} \times 100\% \quad (2)$$

where $\mathrm{Error}(1\sigma)$ is the absolute mapping uncertainty, P is the perimeter of a glacial lake (m), and G is the spatial resolution of the remote sensing imagery (10 m for Sentinel-2 image). N_{Inner} is the number of inner nodes (inflection points) of a lake, and the coefficient of 0.6872 (1σ) represents assuming area measurement errors obey a Gaussian distribution. D represents the relative error calculated by Equation (2), in which A is the area of a corresponding glacial lake.

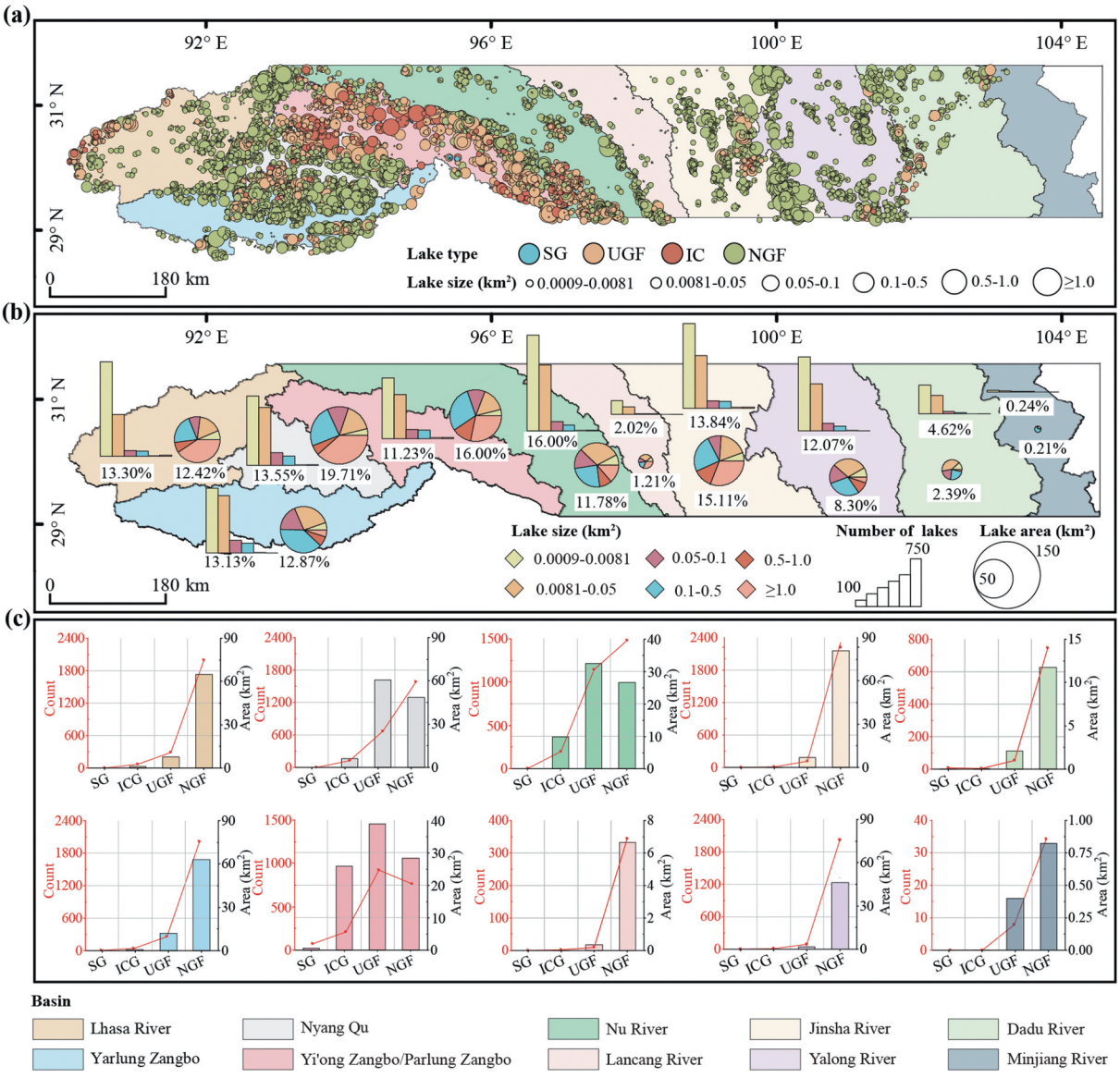
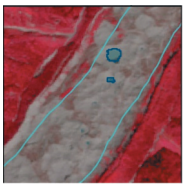
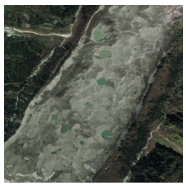
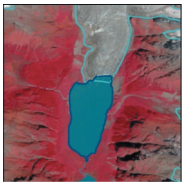
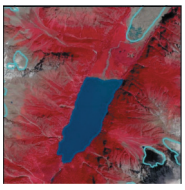
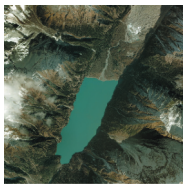
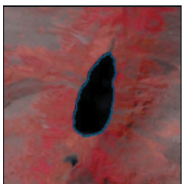


Figure 5. Distribution of STR-GL-2021 across basins. (a) The glacial lake dataset is presented using different size groups and types. (b) Bar and pie charts show the number and area percentage of glacial lakes in each watershed, categorized by size groups. (c) The bar chart illustrates the total number and area of glacial lakes in each basin, categorized by lake type.

Table 3. A classification system of glacial lake types according to the relationship between glacial lakes and glaciers (Lesi et al. 2022a). Glacier outlines are the RGI 6.0 (RGI Consortium 2017.).

Lake types	Characteristics	Sentinel-2	Google Earth
Supraglacial	Lakes are situated on the surface of glaciers. Case location: 30°00′30.10″ N 95°32′00.14″ E		
Ice-contact	Lakes are connected by a glacier terminus and supplied by meltwater. Case location: 30°14′31.40″ N 93°38′15.12″ E		
Unconnected-glacier-fed	Lakes are located downstream of a glacier but are not directly connected to it. Case location: 30°07′32.84″ N 94°05′22.35″ E		
Non-glacier-fed	Lakes were formed by glaciology and are not supplied by glacial meltwater. Case location: 29°17′45.74″ N 93°25′23.76″ E		

3. Data description

The glacial lake dataset (Deng 2022) can be accessed online through the Mountain Science Data Center (http://www.msdc.ac.cn/#/datadetails?id=93.) of the Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, at https://doi.org/10.12380/Glaci.msdc.000002. The dataset is available in Esri shapefile format (21.9 MB) and Geopackage format (18.5 MB). It was reprojected to be the World Geodetic System (WGS) 84 or the Asia North Albers Equal Area Conic coordinate system. There are 16 attribute fields for the glacial lakes, and you can find the details of these attributes in Table 2. Additionally, we provide the outline data of river basins in Esri shapefile format, generated by clipping HydroBASINS data (Lehner and Grill 2013). The study area and statistics of glacial lakes are provided in an Excel spreadsheet format.

The glacial lake dataset includes 17,566 glacial lakes, each with a size of ≥0.0009 km^2, totaling an area of 582.89 ± 31.45 km^2 and a water volume of 16.67 km^3. The altitudes of all glacial lakes varied from 2110 m above sea level (a.s.l.) to 5960 m a.s.l., with 73% located between 4400 m a.s.l. and 5300 m a.s.l., and a mean altitude of 4870 m a.s.l.

The distribution of glacial lakes varies across basins, lake size groups, and lake types along the STR (Figure 5(a)). For example, the Nu River basin has the largest number of glacial lakes (2,810 lakes, accounting for 16.00% of the total), while the Minjiang River basin has the smallest number (43 lakes, accounting for 0.24%). The Nyang Qu basin has the largest proportion (114.86 ± 4.86 km^2, 19.71%), and the Minjiang River basin has the smallest proportion (1.24 ± 0.07 km^2, 0.21%).

Glacial lakes have been categorized into six size categories (Figure 5(b)). There are 9333 glacial lakes, ranging from 0.0009 km^2 to 0.0081 km^2, which make up the majority. As the lake size increases, the total number decreases. For example, there are only 60 glacial lakes larger than 1 km^2. Glacial lakes between 0.1 km^2 and 0.5 km^2 cover the largest percentage, totaling 157.32 ± 5.54 km^2, while glacial lakes ranging from 0.0009 km^2 to 0.0081 km^2 cover the smallest percentage, totaling 32.98 ± 7.22 km^2.

Regarding lake types (Figure 5(c)), NGF glacial lakes are the most common, with a total of 13,275 covering an area of 377.78 ± 22.73 km^2. Next up are the UGF glacial lakes, with 3553 covering an area of 161.75 ± 7.02 km^2. Then, there are 655 IC glacial lakes, covering an area of 42.91 ± 1.62 km^2, primarily distributed in the Nu River and Yi'ong Zangbo/Parlung Zangbo. Lastly, 83 SG lakes cover a total area of 0.44 ± 0.08 km^2, primarily in Yi'ong Zangbo and Parlung Zangbo.

4. Discussion

4.1. Validation of the glacial lake dataset

We randomly selected 165 glacial lakes from our dataset to create reference lakes using high-resolution images from Google Earth around 2021. This selection aimed to verify the accuracy of our glacial lake dataset, given the limited field measurements along the STR. The 165 lakes in our dataset range from 0.0012 km^2 to 5.33 km^2, with a median size of 0.07 km^2, covering all river basins in the study area (Figure 6).

Our glacial lake dataset and the Google Earth-derived lake data showed high consistency in both individual lakes and groups of different sizes, indicating the reliability of our dataset (Figure 6 and Table 4). The mean area difference between lake samples is 0.004 km^2, with most validation samples closely distributed along the 1:1 line, demonstrating high mapping accuracy. The area error for the 165 sample lakes is 1.21%, with a median deviation of 5.26%. The overlap rate exceeds 95% across different size groups (Table 4). Overall, our STR-GL-2021 dataset exhibits satisfactory accuracy.

4.2. Comparison with other published datasets

We compared our glacial lake dataset with two other datasets: the 2018 Wang's glacial lake dataset, digitized from 30 m Landsat images (Wang et al. 2021), and the 2020 Wei's thermokarst lake dataset (Wei 2021).

4.2.1. Comparison with Wang's Landsat-derived dataset

The overall overlap ratio between Wang's 2018 dataset (Wang et al. 2021) and our 2021 dataset is 75.41% in terms of count and 88.99% in terms of area (Table 5), both at the same MMU of 5400 m^2 and spatial coverage. Our dataset contains more lakes than Wang's dataset, associated with a greater total area. In Wang's dataset, 4394 glacial lakes (297.18 km^2) were manually digitized,

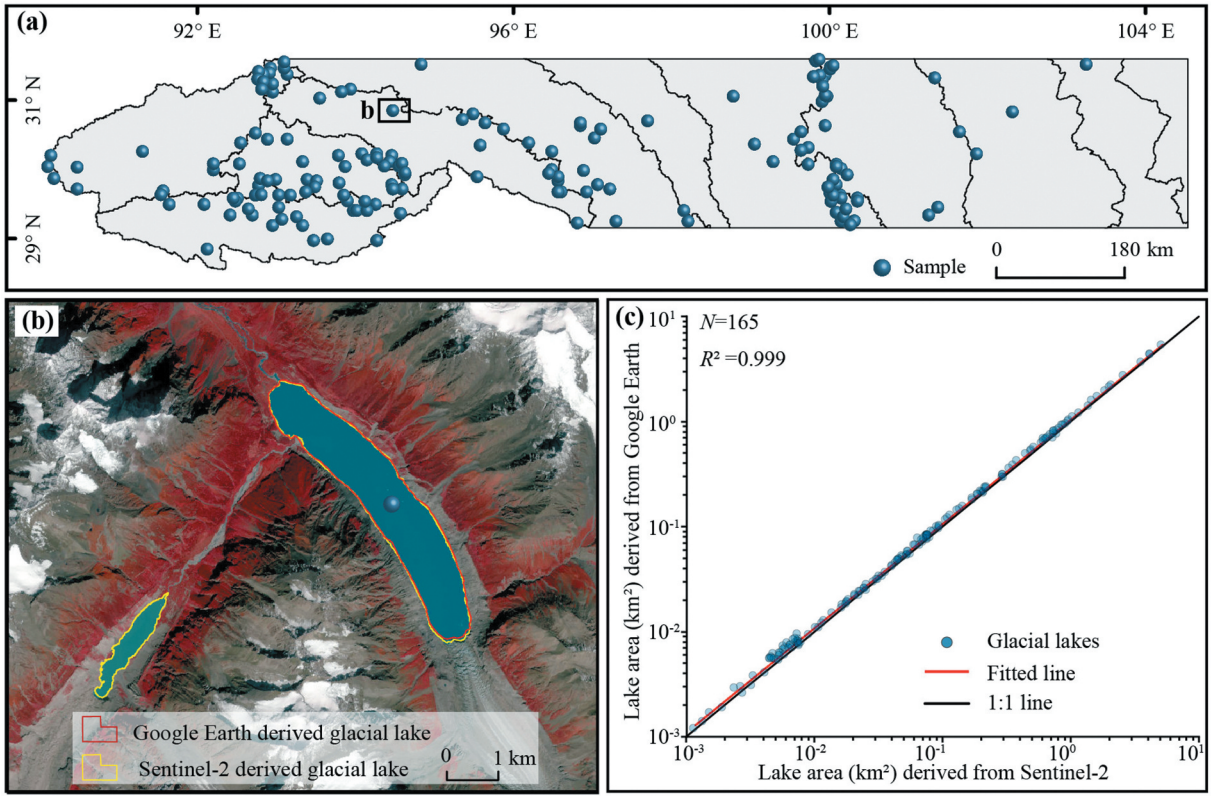


Figure 6. Validation of glacial lakes is conducted using data derived from Google Earth. (a) The spatial distribution of the validation samples is shown, (b) a comparison of a reference glacial lake with our lake data is presented, and (c) consistency between STR-GL -2021 and the reference lakes is demonstrated.

Table 4. Distribution of validation samples in different size groups.

Lake size (km^2)	Number of lakes	Total area from Sentinel-2 (km^2)	Total area from Google Earth (km^2)	Overlap rate (%)
0.0009–0.0081	40	0.22	0.23	95.65
0.0081–0.05	35	0.93	0.93	100.00
0.05–0.1	30	2.23	2.28	97.81
0.1–0.5	25	5.89	6.03	97.68
0.5–1.0	20	14.53	14.80	98.18
≥1.0	15	35.61	35.87	99.28
Total	165	59.41	60.14	98.79

whereas our product mapped 5827 glacial lakes (333.97 km^2). Our product has more lakes in three size groups than Wang's dataset, except for the 0.1–0.5 km^2 and 0.5–1.0 km^2 groups, where the total areas of Wang's dataset are greater than ours. Our dataset, as of 2021, contains 1400 more lakes than Wang's dataset from 2018 for class 0.0054–0.05 km^2. We attributed the discrepancy to the detection of small lakes that the 10 m Sentinel-2 images can extract, whereas the 30 m Landsat images cannot be manually mapped (Figure 7). Our glacial lake dataset has been validated with higher spatial resolution images in Google Earth, showing an improvement over the published dataset.

We compared the same glacial lakes from both datasets and found a high level of consistency, with a total area error of 3.77% (Table 6), ranging from 2.52% to 4.40%. Wang's dataset shows slightly larger lake areas within each size group compared to ours. However, the errors fall within the mapping uncertainty. Variations in image acquisition dates, image quality, and spatial resolutions (Lesi et al. 2022a) are likely responsible for the differences between the two datasets. We are sharing an updated glacial dataset created from Sentinel-2 images, which provides users with more options for studying changes in water resources, evaluating the risk of glacial lake outburst floods, and modeling the changes in glaciers and lakes under different climate change scenarios.

Table 5. Comparing the number and area of glacial lakes between Wang's and our datasets at different size classes.

	Wang's glacial lakes dataset		Our glacial lake dataset		Overlap
Lake size (km^2)	Count	Area (km^2)	Count	Area (km^2)	percentage (%)
0.0054–0.05	3244	65.68	4672	79.20	69.43 (82.93)
0.05–0.1	585	40.92	593	41.26	98.65 (99.18)
0.1–0.5	487	96.00	483	94.24	100.83 (101.87)
0.5–1.0	43	30.71	43	30.40	100.00 (101.02)
≥1.0	35	63.87	36	88.87	97.22 (71.87)
Total	4394	297.18	5827	333.97	75.41 (88.99)

Note: Wang's dataset (Wang et al. 2020) has an MMU of 6 pixels (0.0054 km^2). To ensure comparability, we excluded 3912 glacial lakes (10.76 ± 2.70 km^2) from our dataset, which ranged in size from 0.0009 km^2 to 0.0054 km^2. The overlap percentage (%) indicates the ratios between Wang's dataset and ours in count and area.

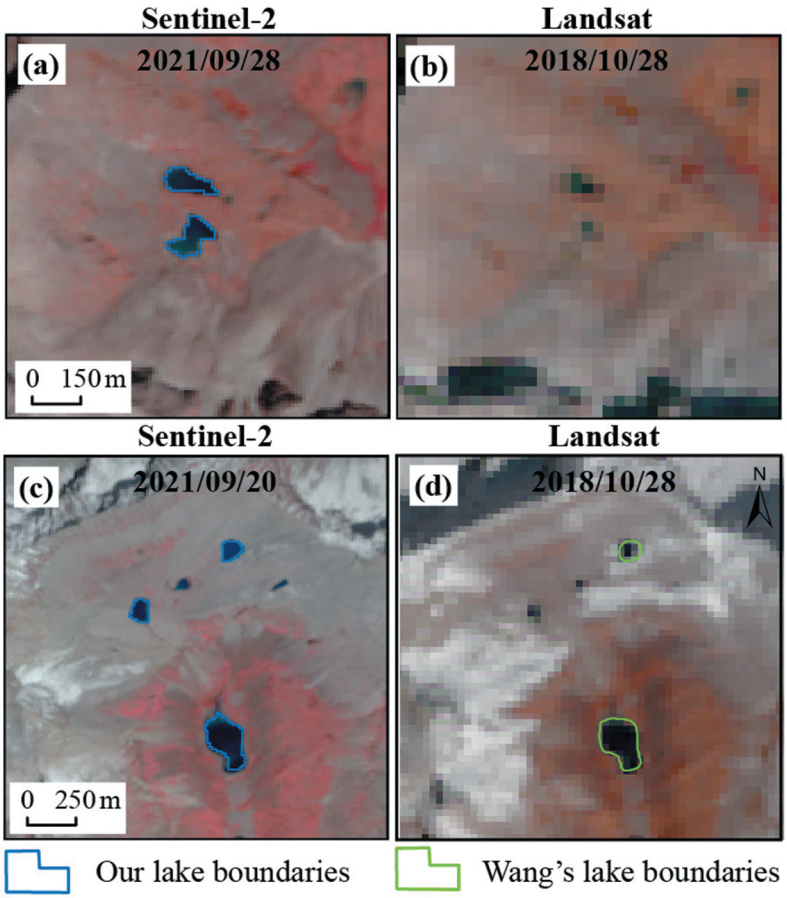


Figure 7. Comparison of Wang's glacial lake dataset and ours for the size of 0.0054–0.05 km^2. (a) Glacial lakes can be mapped from a Sentinel-2 image, but cannot be extracted from a Landsat image (b). We were able to extract more glacial lakes from a Sentinel-2 image than Wang's data from a Landsat image (c, d).

4.2.2. Comparison with Wei's Sentinel-2-derived dataset

The area difference between Wei's 2020 dataset and our 2021 dataset is 11.48% (Table 7) at the MMU of 900 m^2 (9 pixels). Wei's dataset (Wei 2021) was initially intended to study thermokarst lakes and ponds on the Tibetan Plateau, defining a study area outside a 10 km buffer zone of glaciers (Figure 8(a)). We selected 2690 glacial lakes from Wei's dataset by overlapping with our dataset for comparison. Count and area errors ranged from −12.45% to 18.89% across different size groups. Wei's dataset (Wei 2021) contains more lakes in size groups ranging from 0.0081 km^2 to 1.0 km^2 and has a larger total area in all size groups except the 0.0009–0.0081 km^2 group.

Wei's data had a larger lake area due to topographic shadows and seasonal changes, and their effects were removed from our dataset. This difference is evident in the Sentinel-2 images across various lake sizes (Figure 8(b-g)). Another factor contributing to errors

Table 6. Comparing Wang's dataset with our dataset in various size groups.

Lake size (km^2)	Wang's glacial lakes dataset		Our glacial lake dataset		Difference percentage (%)
	Count	Area (km^2)	Count	Area (km^2)	
0.0054–0.05	2932	61.66	2970	59.47	−1.28 (3.68)
0.05–0.1	570	39.80	547	38.17	4.20 (4.27)
0.1–0.5	472	92.82	460	89.10	2.61 (4.18)
0.5–1.0	42	29.89	40	28.63	5.00 (4.40)
≥1.0	32	55.42	31	54.06	3.23 (2.52)
Total	4048	279.59	4048	269.43	0.00 (3.77)

Note: The MMU in the table is 6 pixels (0.0054 km^2), and glacial lakes that overlap between Wang's and our datasets have been reselected for comparison. The difference percentage (%) represents the error between Wang's dataset and ours in count and area, respectively.

Table 7. Comparing the count and area of glacial lakes between Wei's and our datasets across different size categories.

Lake size (km^2)	Wei's dataset		Our dataset		Difference percentage (%)
	Count	Area (km^2)	Count	Area (km^2)	
0.0009–0.0081	1153	4.40	1317	4.73	−12.45 (−6.98)
0.0081–0.05	1083	22.68	975	19.96	11.08 (13.63)
0.05–0.1	248	17.67	215	15.10	15.35 (17.02)
0.1–0.5	183	33.93	162	29.88	12.96 (13.55)
0.5–1.0	13	9.00	11	7.57	18.18 (18.89)
≥1.0	10	18.33	10	17.85	0.00 (2.69)
Total	2690	106.01	2690	95.09	0.00 (11.48)

Note: The MMU in Wei's dataset (Wei 2021) is 5 pixels (0.0005 km^2), which is smaller than the 9 pixels in our dataset. Therefore, a 9-pixel threshold was utilized to select glacial lakes for comparison. The difference percentage (%) represents the errors between Wei's dataset and ours in count and area, respectively.

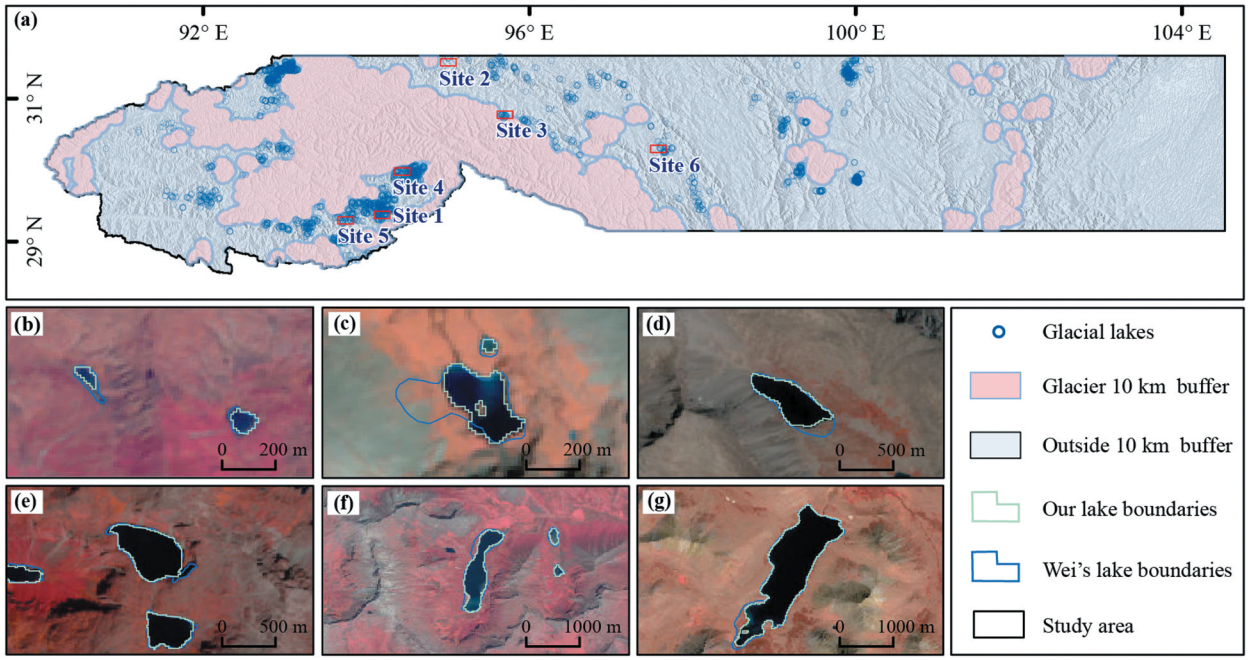


Figure 8. Comparison between Wei's dataset and ours. (a) The spatial distribution of the selected glacial lakes from Wei's dataset and enlarged sub-graph locations. (b)-(g) Examples of glacial lake boundaries in Wei's dataset and ours, shown at the numbering sites in (a).

is the difference in acquisition dates between the Sentinel-2 source images. In our study, we dedicated time to selecting the best image for each tile and inspecting and validating lake data to minimize the impact of clouds, topographical shadows, and other factors, thereby producing a high-quality dataset of glacial lakes.

4.3. Accuracy and limitations in the glacial lake dataset

Several factors, including pixel grid size, georeferencing errors, MMU, and the effects of cloud and snow cover, as well as key parameters of mapping methods and operator experience, influence the accuracy of the glacial lake dataset. In this study, we calculate area uncertainty using an improved equation that incorporates pixel grid size and lake perimeter (Lesi et al. 2022a). This equation considers only the displacement caused by georeferencing satellite imagery. Other challenging factors that are difficult to quantify have been excluded. Additionally, the MMU has a significant influence on the discrepancies in both the number and area of the glacial lake datasets. A smaller MMU typically results in a higher number of detected glacial lakes (Nie et al. 2017). Pixel sizes ranging from 3 to 55 pixels have been used in glacial lake mapping with satellite images (Chen et al. 2021; Lesi et al. 2022a; Nie et al. 2017; Shugar et al. 2020; Wang et al. 2020; Zhang et al. 2015). The research objectives and mapping techniques determine the MMU threshold. In this study, we set the MMU at 9 pixels, a threshold widely accepted in the previous research (Chen et al. 2021; Nie et al. 2017). The MMU impacts uncertainty estimation due to the measurement approach related to pixel grid size and lake perimeter (Hanshaw and Bookhagen 2014; Lesi et al. 2022a). A smaller lake may show less absolute error but a higher relative error according to the uncertainty calculation equation, and vice versa. Examples from our dataset illustrate that the outlines of small glacial lakes produced using our mapping method are reliable (Figures 7 and 8). Our study demonstrates that 10 m Sentinel-2-derived glacial lakes provide more accurate boundaries with less mapping uncertainty compared to 30 m Landsat-derived glacial lake outlines (Figure 7), consistent with earlier research (Lesi et al. 2022a).

A series of mapping procedures and workflows enhances the reliability of our glacial lake dataset. (1) We selected a total of 165 glacial lakes through stratified random sampling and manually digitized them using high-resolution images from Google Earth. This sample size exceeds that of a previous study, which utilized 89 samples (Lesi et al. 2022a), demonstrating the reliability of the lake dataset produced by our mapping workflow. (2) To minimize the effects of clouds, cloud shadows, and snow, we select autumn-dominated Sentinel-2 images for mapping glacial lakes (Figure 3), thereby reducing the impact of seasonal variations and snow cover. We also reference an existing glacial lake dataset (Wang et al. 2020; Wei et al. 2021) to select source images for our dataset, which further minimizes the effects of clouds and cloud shadows. (3) Integrating a user-interactive mapping tool and a data verification process carried out by an assigned individual enhances the consistency and accuracy of glacial lake mapping, reducing potential errors from multiple operators. All these mapping strategies improve the quality of our glacial lake dataset by minimizing mapping errors. Finally, we present the first 10 m glacial lake dataset covering the entire STR-related region.

Our glacial lake dataset has several limitations. (1) The global area-volume empirical function (Cook and Quincey 2015), typically used to estimate glacial lake volumes, may cause regional-scale inaccuracies compared to area-volume equations derived from the Tibetan Plateau (Qi et al. 2025; Zhang et al. 2023). Limited bathymetric measurements in the study area restrict the creation of a localized model that could enhance the accuracy of volume estimates. (2) Temporal discrepancies between the RGI 6.0 glacier outlines and Sentinel-2 imagery employed for delineating glacial lakes may result in minor errors in identifying lake types. Although RGI 6.0 has been combined with Sentinel-2 images to address these issues, the lack of an updated and synchronized glacier dataset limits the precision of lake type classification. (3) There is a seasonal bias in our glacial lake dataset, but its impact on analyzing intra- and inter-annual changes is minimal. Most Sentinel-2 images (76%) were captured in September, October, and November, when the Indian summer monsoon weakens (Nie et al. 2021), resulting in minimal water-level changes in glacial lakes during autumn (Wang et al. 2025). Optical satellite images are inherently affected by snow and cloud cover (Wang et al. 2021; Wei et al. 2021), which means that we cannot gather enough high-quality Sentinel-2 images in a single month to eliminate seasonal bias entirely. The bias caused by the timing difference of satellite image acquisition is unavoidable in similar studies (Lesi et al. 2022b; Wang et al. 2021; Wei et al. 2021) that rely on images from a single satellite platform to create a glacial lake dataset. This issue is common in remote sensing-based glacial lake mapping. Prior research indicates that water levels in glacial lakes on the Tibetan Plateau and nearby regions tend to remain relatively stable (Wang et al. 2025), with less than 1 m of intra-annual variation in 86% of these lakes. While seasonal changes may slightly affect year-to-year comparisons, selecting datasets from comparable seasons helps reduce this effect. (4) Temporal mismatches between our dataset (Deng 2022) and existing datasets (Wang

et al. 2020; Wei 2021) introduce potential biases in comparative studies that are challenging to quantify. Considering the increasing risk of GLOFs (Zheng et al. 2021) due to climate warming, our 10-m resolution glacial lake dataset in 2021 is highly valuable for hazard risk assessments.

Author contributions

CRediT: **Qian Deng:** Data curation, Methodology, Writing – original draft, Writing – review & editing; **Yong Nie:** Conceptualization, Methodology, Project administration, Writing – original draft, Writing – review & editing; **Jonathan L. Carrivick:** Writing – review & editing; **Gordon G. D. Zhou:** Writing – review & editing; **Qiyuan Lyu:** Methodology, Software, Writing – review & editing; **Jida Wang:** Methodology, Software, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Natural Science Foundation of China [Grant numbers 42171086, 41971153, and 41941017], the Science and Technology Department of Tibet Program [Grant number XZ202301ZY0016G], and the Key Laboratory of Mountain Hazards and Engineering Resilience, Chinese Academy of Sciences [Grant number KLMHER-T09].

Data availability statement

The dataset for glacial lakes can be accessed online from the Mountain Science Data Center of the Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, at https://doi.org/10.12380/Glaci.msdc.000002. The Sentinel-2 images were downloaded from the Copernicus Open Access Hub at https://scihub.copernicus.eu/. The 1 arc-second SRTM DEM was obtained from Earthdata at https://www.earthdata.nasa.gov/. The Randolph Glacier Inventory (RGI) 6.0 was downloaded from Global Land Ice Measurements from Space at https://www.glims.org/.

Notes on contributors

Qian Deng received the Master's degree from the Institute of Mountain Hazards and Environment, Chinese Academy of Sciences. Her research interests include remote sensing, scientific data, and hazard risk assessment.

Yong Nie is a professor in the Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, and received the PhD degree from the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences. His research interests include glacial lake outburst floods, remote sensing mapping, scientific data management, and natural hazard risk assessment.

Jonathan L. Carrivick is a professor at the University of Leeds and received a PhD degree from Durham University. His research interests include earth surface processes and landforms in polar, arctic, and alpine environments.

Gordon G. D. Zhou is a professor in the Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, and received the PhD degree from Hong Kong University of Science and Technology. His research interests include the triggering and flow mechanisms of debris flows, granular flows, and multiphase flows, and mechanisms of the hazard chain process.

Qiyuan Lyu received the Master's degree from the University of Electronic Science and Technology of China. His research interests include the development of new algorithms for remote sensing images, scientific data management, and the construction of software platforms.

Jida Wang is an associate professor at the University of Illinois Urbana-Champaign and received a PhD degree from the University of California in 2013. His research interests include the abundance and dynamics of surface water, particularly in lakes, reservoirs, and wetlands.

ORCID

Qian Deng http://orcid.org/0009-0009-1386-3087
Yong Nie http://orcid.org/0000-0002-6075-8564
Jonathan L. Carrivick http://orcid.org/0000-0002-9286-5348
Gordon G. D. Zhou http://orcid.org/0000-0001-9014-2931
Qiyuan Lyu http://orcid.org/0000-0002-8100-0859
Jida Wang http://orcid.org/0000-0003-3548-8918

References

Aggarwal, S., S. C. Rai, P. K. Thakur, and A. Emmer. 2017. "Inventory and Recently Increasing GLOF Susceptibility of Glacial Lakes in Sikkim, Eastern Himalaya." *Geomorphology* 295:39–54. https://doi.org/10.1016/j.geomorph.2017.06.014

Brun, F., E. Berthier, P. Wagnon, A. Kääb, and D. Treichler. 2017. "A Spatially Resolved Estimate of High Mountain Asia Glacier Mass Balances from 2000 to 2016." *Nature Geoscience* 10 (9): 668–673. https://doi.org/10.1038/ngeo2999

Carrivick, J. L., P. How, J. M. Lea, J. L. Sutherland, M. Grimes, F. S. Tweed, S. Cornford, D. J. Quincey, and J. Mallalieu. 2022. "Ice-Marginal Proglacial Lakes Across Greenland: Present Status and a Possible Future." *Geophysical Research Letters* 49 (12): e2022GL099276. https://doi.org/10.1029/2022GL099276

Carrivick, J. L., and D. J. Quincey. 2014. "Progressive Increase in Number and Volume of Ice-Marginal Lakes on the Western Margin of the Greenland Ice Sheet." *Global and Planetary Change* 116:156–163. https://doi.org/10.1016/j.gloplacha.2014.02.009

Carrivick, J. L., J. L. Sutherland, M. Huss, H. Purdie, C. D. Stringer, M. Grimes, W. H. M. James, and A. M. Lorrey. 2022. "Coincident Evolution of Glaciers and Ice-Marginal Proglacial Lakes Across the Southern Alps, New Zealand: Past, Present and Future." *Global and Planetary Change* 211:103792. https://doi.org/10.1016/j.gloplacha.2022.103792

Carrivick, J. L., and F. S. Tweed. 2013. "Proglacial Lakes: Character, Behaviour and Geological Importance." *Quaternary Science Reviews* 78:34–52. https://doi.org/10.1016/j.quascirev.2013.07.028

Carrivick, J. L., and F. S. Tweed. 2019. "A Review of Glacier Outburst Floods in Iceland and Greenland with a Megafloods Perspective." *Earth-Science Reviews* 196:102876. https://doi.org/10.1016/j.earscirev.2019.102876

Chen, F., M. Zhang, H. Guo, S. Allen, J. S. Kargel, U. K. Haritashya, and C. S. Watson. 2021. "Annual 30 m Dataset for Glacial Lakes in High Mountain Asia from 2008 to 2017." *Earth System Science Data* 13 (2): 741–766. https://doi.org/10.5194/essd-13-741-2021

Cook, S. J., and D. J. Quincey. 2015. "Estimating the Volume of Alpine Glacial Lakes." *Earth Surface Dynamics* 3 (4): 559–575. https://doi.org/10.5194/esurf-3-559-2015

Deng, Q. 2022. "A glacial lake dataset along the SichuanTibet Railway." Mountain Science Data Center. https://doi.org/10.12380/Glaci.msdc.000002

Drusch, M., U. Del Bello, S. Carlier, O. Colin, V. Fernandez, F. Gascon, B. Hoersch, et al. 2012. "Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services." *Remote Sensing of Environment* 120:25–36. https://doi.org/10.1016/j.rse.2011.11.026

Dubey, S., and M. K. Goyal. 2020. "Glacial Lake Outburst Flood Hazard, Downstream Impact, and Risk Over the Indian Himalayas." *Water Resources Research* 56 (4): e2019WR026533. https://doi.org/10.1029/2019WR026533

Furian, W., F. Maussion, and C. Schneider. 2022. "Projected 21st-Century Glacial Lake Evolution in High Mountain Asia." *Frontiers in Earth Science* 10:821798. https://doi.org/10.3389/feart.2022.821798

Hanshaw, M. N., and B. Bookhagen. 2014. "Glacial Areas, Lake Areas, and Snow Lines from 1975 to 2012: Status of the Cordillera Vilcanota, Including the Quelccaya Ice Cap, Northern Central Andes, Peru." *The Cryosphere* 8 (2): 359–376. https://doi.org/10.5194/tc-8-359-2014

How, P., A. Messerli, E. Mätzler, M. Santoro, A. Wiesmann, R. Caduff, K. Langley, et al. 2021. "Greenland-Wide Inventory of Ice Marginal Lakes Using a Multi-Method Approach." *Scientific Reports* 11 (1): 4481. https://doi.org/10.1038/s41598-021-83509-1

King, O., A. Dehecq, D. Quincey, and J. Carrivick. 2018. "Contrasting Geometric and Dynamic Evolution of Lake and Land-Terminating Glaciers in the Central Himalaya." *Global and Planetary Change* 167:46–60. https://doi.org/10.1016/j.gloplacha.2018.05.006

Kraaijenbrink, P. D. A., M. F. P. Bierkens, A. F. Lutz, and W. W. Immerzeel. 2017. "Impact of a Global Temperature Rise of 1.5 Degrees Celsius on Asia's Glaciers." *Nature* 549 (7671): 257–260. https://doi.org/10.1038/nature23878

Lee, E., J. L. Carrivick, D. J. Quincey, S. J. Cook, W. H. M. James, and L. E. Brown. 2021. "Accelerated Mass Loss of Himalayan Glaciers Since the Little Ice Age." *Scientific Reports* 11 (1): 24284. https://doi.org/10.1038/s41598-021-03805-8

Lehner, B., and G. Grill. 2013. "Global River Hydrography and Network Routing: Baseline Data and New Approaches to Study the World's Large River Systems." *Hydrological Processes* 27 (15): 2171–2186. https://doi.org/10.1002/hyp.9740

Lesi, M., Y. Nie, D. H. Shugar, J. Wang, Q. Deng, H. Chen, and J. Fan. 2022a. "Landsat- and Sentinel-Derived Glacial Lake Dataset in the China–Pakistan Economic Corridor from 1990 to 2020." *Earth System Science Data* 14 (12): 5489–5512. https://doi.org/10.5194/essd-14-5489-2022

Lesi, M., Y. Nie, D. H. Shugar, J. Wang, Q. Deng, H. Chen, and J. Fan. 2022b. "Landsat- and Sentinel-derived glacial lake dataset in the China–Pakistan Economic Corridor from 1990 to 2020." Mountain Science Data Center. https://doi.org/10.12380/Glaci.msdc.000001

Li, D., M. Wang, H. Guo, and W. Jin. 2025. "On China's Earth Observation System: Mission, Vision and Application." *Geo-Spatial Information Science* 28 (2): 303–321. https://doi.org/10.1080/10095020.2024.2328100

Li, J., and Y. Sheng. 2012. "An Automated Scheme for Glacial Lake Dynamics Mapping Using Landsat Imagery and Digital Elevation Models: A Case Study in the Himalayas." *International Journal of Remote Sensing* 33 (16): 5194–5213. https://doi.org/10.1080/01431161.2012.657370

Liu, Q., C. Mayer, X. Wang, Y. Nie, K. Wu, J. Wei, and S. Liu. 2020. "Interannual Flow Dynamics Driven by Frontal Retreat of a Lake-Terminating Glacier in the Chinese Central Himalaya." *Earth and Planetary Science Letters* 546:116450. https://doi.org/10.1016/j.epsl.2020.116450

Lu, C., and C. Cai. 2019. "Challenges and Countermeasures for Construction Safety During the sichuan–Tibet Railway Project." *Engineering* 5 (5): 833–838. https://doi.org/10.1016/j.eng.2019.06.007

Maurer, J. M., J. M. Schaefer, S. Rupper, and A. Corley. 2019. "Acceleration of Ice Loss Across the Himalayas Over the Past 40 Years." *Science Advances* 5 (6): eaav7266. https://doi.org/10.1126/sciadv.aav7266

Mcfeeters, S. K. 1996. "The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features." *International Journal of Remote Sensing* 17 (7): 1425–1432.

Nasa, J. P. L. 2013. "NASA Shuttle Radar Topography Mission Global 1 Arc Second." *NASA EOSDIS Land Processes DAAC*. https://doi.org/10.5067/MEaSUREs/SRTM/SRTMGL1.003

Nie, Y., Q. Deng, H. D. Pritchard, J. L. Carrivick, F. Ahmed, C. Huggel, L. Liu, et al. 2023. "Glacial Lake Outburst Floods Threaten Asia's Infrastructure." *Science Bulletin* 68 (13): 1361–1365. https://doi.org/10.1016/j.scib.2023.05.035

Nie, Y., W. Liu, Q. Liu, X. Hu, and M. J. Westoby. 2020. "Reconstructing the Chongbaxia Tsho Glacial Lake Outburst Flood in the Eastern Himalaya: Evolution, Process and Impacts." *Geomorphology* 370 (12): 107393. https://doi.org/10.1016/j.geomorph.2020.107393

Nie, Y., H. D. Pritchard, Q. Liu, T. Hennig, W. Wang, X. Wang, S. Liu, et al. 2021. "Glacial Change and Hydrological Implications in the Himalaya and Karakoram." *Nature Reviews Earth & Environment* 2 (2): 91–106. https://doi.org/10.1038/s43017-020-00124-w

Nie, Y., Y. Sheng, Q. Liu, L. Liu, S. Liu, Y. Zhang, and C. Song. 2017. "A Regional-Scale Assessment of Himalayan Glacial Lake Changes Using Satellite Observations from 1990 to 2015." *Remote Sensing of Environment* 189:1–13. https://doi.org/10.1016/j.rse.2016.11.008

Pan, Y., X. Zhang, J. Jiao, and Y. Xiao. 2025. "Spatiotemporal Variability at Seasonal and Interannual Scales of Terrestrial Water Variation Over Tibetan Plateau from Geodetic Observations." *Geo-Spatial Information Science* 28 (3): 1434–1449. https://doi.org/10.1080/10095020.2024.2331553

Peng, S. 2024a. "1-km monthly mean temperature dataset for China (1901-2023)." National Tibetan Plateau Data Center. https://doi.org/10.11888/Meteoro.tpdc.270961

Peng, S. 2024b. "1-km monthly precipitation dataset for China (1901-2023)." National Tibetan Plateau Data Center. https://doi.org/10.5281/zenodo.3114194

Pi, X., Q. Luo, L. Feng, Y. Xu, J. Tang, X. Liang, E. Ma, et al. 2022. "Mapping Global Lake Dynamics Reveals the Emerging Roles of Small Lakes." *Nature Communications* 13 (1): 5777. https://doi.org/10.1038/s41467-022-33239-3

Qi, M., S. Liu, Z. Zhao, Y. Gao, F. Xie, G. Veh, L. Xiao, J. Jing, Y. Zhu, and K. Wu. 2025. "A Mathematical Model to Improve Water Storage of Glacial Lake Prediction towards Addressing Glacial Lake Outburst Floods." *Hydrology and Earth System Sciences* 29 (4): 969–982. https://doi.org/10.5194/hess-29-969-2025

RGI Consortium. 2017. "Randolph Glacier Inventory–A Dataset of Global Glacier Outlines: Version 6." NSIDC: National Snow and Ice Data Center. https://doi.org/10.7265/4m1f-gd79

Rose, A. N., J. J. Mckee, K. M. Sims, E. A. Bright, A. E. Reith, and M. L. Urban. 2020. *LandScan 2019*. Oak Ridge, TN: Oak Ridge National Laboratory.

Sahu, R., R. Ramsankaran, R. Bhambri, P. Verma, and P. Chand. 2023. "Evolution of Supraglacial Lakes from 1990 to 2020 in the Himalaya–Karakoram Region Using Cloud-Based Google Earth Engine Platform." *Journal of the Indian Society of Remote Sensing* 51 (12): 2379–2390. https://doi.org/10.1007/s12524-023-01773-2

Shean, D. E., S. Bhushan, P. Montesano, D. R. Rounce, A. Arendt, and B. Osmanoglu. 2020. "A Systematic, Regional Assessment of High Mountain Asia Glacier Mass Balance." *Frontiers in Earth Science* 7:363. https://doi.org/10.3389/feart.2019.00363

Shugar, D. H., A. Burr, U. K. Haritashya, J. S. Kargel, C. S. Watson, M. C. Kennedy, A. R. Bevington, R. A. Betts, S. Harrison, and K. Strattman. 2020. "Rapid Worldwide Growth of Glacial Lakes Since 1990." *Nature Climate Change* 10 (10): 939–945. https://doi.org/10.1038/s41558-020-0855-4

Sutherland, J. L., J. L. Carrivick, N. Gandy, J. Shulmeister, D. J. Quincey, and S. L. Cornford. 2020. "Proglacial Lakes Control Glacier Geometry and Behavior During Recession." *Geophysical Research Letters* 47 (19): e2020GL088865. https://doi.org/10.1029/2020GL088865

Wang, J., Y. Sheng, and T. S. D. Tong. 2014. "Monitoring Decadal Lake Dynamics Across the Yangtze Basin Downstream of Three Gorges Dam." *Remote Sensing of Environment* 152:251–269. https://doi.org/10.1016/j.rse.2014.06.004

Wang, W., K. Xu, A. Wang, Y. Chen, X. Deng, and T. Wang. 2025. "ADAC: An Active Domain Adaptive Network With Progressive Learning Strategy for Cloud Detection of Remote Sensing Imagery." *Geo-Spatial Information Science* 28 (4): 1597–1614. https://doi.org/10.1080/10095020.2024.2389958

Wang, X., X. Guo, C. Yang, Q. Liu, J. Wei, Y. Zhang, S. Liu, Y. Zhang, Z. Jiang, and Z. Tang. 2020. "Glacial Lake Inventory of High-Mountain Asia in 1990 and 2018 Derived from Landsat Images." *Earth System Science Data* 12 (3): 2169–2182. https://doi.org/10.5194/essd-12-2169-2020

Wang, X., X. Guo, C. Yang, Q. Liu, J. Wei, Y. Zhang, S. Liu, Y. Zhang, Z. Jiang, and Z. Tang. 2021. "Cataloging Data Set of High Asian Ice Lakes." *National Cryosphere Desert Data Center*. https://doi.org/10.12072/casnw.064.2019.db

Wang, Y., D. Zheng, G. Zhang, J. L. Carrivick, T. Bolch, W. Ren, L. Guo, J. Su, S. Yuan, and X. Li. 2025. "Patterns and Change Rates of Glacial Lake Water Levels Across High Mountain Asia." *National Science Review* 12 (3): nwaf041. https://doi.org/10.1093/nsr/nwaf041

Wangchuk, S., and T. Bolch. 2020. "Mapping of Glacial Lakes Using Sentinel-1 and Sentinel-2 Data and a Random Forest Classifier: Strengths and Challenges." *Science of Remote Sensing* 2 (2): 100008. https://doi.org/10.1016/j.srs.2020.100008

Wei, Z. 2021. "Thermokarst Lake and Pond Dataset of the Qinghai–Tibet Plateau (QTP)." *Zenodo*. https://doi.org/10.5281/zenodo.5509325

Wei, Z., Z. Du, L. Wang, J. Lin, Y. Feng, Q. Xu, and C. Xiao. 2021. "Sentinel-Based Inventory of Thermokarst Lakes and Ponds Across Permafrost Landscapes on the Qinghai-Tibet Plateau." *Earth and Space Science* 8 (11): e2021EA001950. https://doi.org/10.1029/2021EA001950

Wieland, M., F. Fichtner, and S. Martinis. 2022. "UKIS-CSMASK: A Python Package for Multi-Sensor Cloud and Cloud Shadow Segmentation." *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLIII-B3-2022:217–222. https://doi.org/10.5194/isprs-archives-XLIII-B3-2022-217-2022

Xue, Y., F. Kong, S. Li, Q. Zhang, D. Qiu, M. Su, and Z. Li. 2021. "China Starts the world's Hardest "Sky-High Road" Project: Challenges and Countermeasures for sichuan–Tibet Railway." *The Innovation* 2 (2): 100105. https://doi.org/10.1016/j.xinn.2021.100105

Zhang, G., T. Bolch, T. Yao, D. R. Rounce, W. Chen, G. Veh, O. King, S. K. Allen, M. Wang, and W. Wang. 2023. "Underestimated Mass Loss from Lake-Terminating Glaciers in the Greater Himalaya." *Nature Geoscience* 16 (4): 333–338. https://doi.org/10.1038/s41561-023-01150-1

Zhang, G., T. Yao, H. Xie, W. Wang, and W. Yang. 2015. "An Inventory of Glacial Lakes in the Third Pole Region and Their Changes in Response to Global Warming." *Global and Planetary Change* 131:148–157. https://doi.org/10.1016/j.gloplacha.2015.05.013

Zhang, H., Y. Nie, L. Liu, L. Han, Q. Lyu, and W. Wang. 2026. "An Advanced Deep Learning Framework for Mapping Glacial Lakes and Its Application in the Hindu Kush-Karakoram-Himalaya Region." *Journal of Hydrology* 664:134507. https://doi.org/10.1016/j.jhydrol.2025.134507

Zhang, J., T. Zhao, Z. Libo, and R. Lingkun. 2021. "Historical Changes and Future Projections of Extreme Temperature and Precipitation Along the Sichuan–Tibet Railway." *Journal of Meteorological Research* 35 (3): 402–415.

Zhang, Z., X. Hu, B. Yang, K. Deng, M. Zhang, and D. Zhu. 2025. "Enhanced Semantic-Positional Feature Fusion Network via Diverse Pre-Trained Encoders for Remote Sensing Image Water-Body Segmentation." *Geo-Spatial Information Science* 28 (4): 1615–1637. https://doi.org/10.1080/10095020.2024.2416898

Zheng, G., S. K. Allen, A. Bao, J. A. Ballesteros-Cánovas, M. Huss, G. Zhang, J. Li, et al. 2021. "Increasing Risk of Glacial Lake Outburst Floods from Future Third Pole Deglaciation." *Nature Climate Change* 11 (5): 411–417. https://doi.org/10.1038/s41558-021-01028-3