Landsat and Sentinel-derived glacial lake dataset in the China-Pakistan Economic Corridor from 1990 to 2020

Muchu Lesi[1], Yong Nie[1, *], Dan H. Shugar[2], Jida Wang[3], Qian Deng[1, 4], Huayong Chen[1], Jianrong Fan[1]

[1]Institute of Mountain Hazards and Environment, Chinese Academy of Sciences, Chengdu, China.

[2]Water, Sediment, Hazards, and Earth-surface Dynamics (waterSHED) Lab, Department of Geoscience, University of Calgary, Alberta, T2N 1N4, Canada

[3]Department of Geography and Geospatial Sciences, Kansas State University, Manhattan, Kansas 66506, USA

[4]University of Chinese Academy of Sciences, Beijing 100190, China

[*]Corresponding author, nieyong@imde.ac.cn

Abstract. The China-Pakistan Economic Corridor (CPEC) is one of the flagship projects of the One Belt One Road Initiative, which faces threats from water shortage and mountain disasters in the high-elevation region, such as glacial lake outburst floods (GLOFs). An up-to-date high-quality glacial lake dataset with parameters such as lake area, volume and type, which is fundamental to water resource and flood risk assessments, and predicting glacier-lake evolutions, is still largely absent for the entire CPEC. This study describes a glacial lake dataset for the CPEC using a threshold-based mapping method associated with rigorous visual inspection workflows. This dataset includes (1) multi-temporal inventories for 1990, 2000, and 2020 produced from 30 m resolution Landsat images, and (2) a glacial lake inventory for the year 2020 at 10 m resolution produced from Sentinel-2 images. The results show that, in 2020, 2234 lakes were derived from the Landsat images, covering a total area of 86.31±14.98 km^2 with a minimum mapping unit of 5 pixels (4500 m^2), whereas 7560 glacial lakes were derived from the Sentinel-2 images with a total area of 103.70±8.45 km^2 with a minimum mapping unit of 5 pixels (500 m^2). The discrepancy shows that Sentinel-2 is able to detect a significant quantity of smaller lakes than Landsat due to its finer spatial resolution. Glacial lake data in 2020 was validated by Google Earth-derived lake boundaries with a median (±standard deviation) difference of 7.66±4.96 % for Landsat-derived product and 4.46±4.62 % for Sentinel-derived product. The total number and area of glacial lakes from consistent 30 m resolution Landsat images remain relatively stable despite a slight increase from 1990 to 2020. A range of critical attributes have been generated in the dataset, including lake types and mapping uncertainty estimated by an improved Hanshaw's equation. This comprehensive glacial lake dataset has potential to be widely applied in studies on water resource assessment, glacial lake-related hazards, glacier-lake interactions, and is freely available at https://doi.org/10.12380/Glaci.msdc.000001 (Lesi et al., 2022).

1 Introduction

Glaciers in High-mountain Asia (HMA) play a crucial role in regulating climate, supporting ecosystems, modulating the release of freshwater into rivers, and sustaining municipal water supplies (Wang et al., 2019; Viviroli et al., 2020), agricultural irrigation, and hydropower generation (Pritchard, 2019; Nie et al., 2021). Most HMA glaciers are losing mass in the context of climate change (Brun et al., 2017; Maurer et al., 2019; Shean et al., 2020; Bhattacharya et al., 2021), therefore, unsustainable glacier melt and the passing of peak water are reducing the hydrological role of glaciers (Huss and Hock, 2018) and impacting downstream ecosystem services, agriculture, hydropower and other socioeconomic values (Carrivick and Tweed, 2016; Nie et al., 2021). The present and future glacier changes not only impact water supply for downstream area but also alter the frequency and intensity of glacier-related hazards, such as glacier lake outburst floods (GLOFs) (Nie et al., 2018; Rounce et al., 2020; Zheng et al., 2021), and rock and ice avalanches (Shugar et al., 2021). Global glacial lake number and total area both increased between 1990 and 2018 in response to glacier retreat and climate change (Shugar et al., 2020), affecting the allocation of freshwater resource. The Indus is globally the most important and vulnerable water tower unit where glaciers, lakes and reservoir storage contribute about two-thirds of the water supply (Immerzeel et al., 2020). Ice-marginal lakes store ~1% of total ice discharge in Greenland and accelerate lake-terminating ice velocity by ~25% (Mankoff et al., 2020; Carrivick et al., 2022). An increasing frequency and risk of GLOFs (Nie et al., 2021; Zheng et al., 2021) is threatening Asian population and infrastructures in the mountain ranges, such as the China-Pakistan Economic Corridor (CPEC), as a flagship component of One Belt One Road Initiative (Battamo et al., 2021; Li et al., 2021). The northern section of the CPEC passes through Pamir, Karakoram, Hindu Kush and Himalaya mountains where droughts and glacier-related hazards are frequent and severe (Hewitt, 2014; Bhambri et al., 2019; Pritchard, 2019), threatening local people, the existing, under-construction and planned infrastructures, such as highways, hydropower plants and railways. Understanding the risk posed by water shortage and glacier-related hazards is a critical step to sustainable development for the CPEC.

Glacial lake inventories with a range of attributes benefit water resource assessment and disaster risk assessment related to glacial lake (Wang et al., 2020; Carrivick et al., 2022), and contribute to predicting glacier-lake evolution and cryosphere-hydrosphere interactions under climate change (Nie et al., 2017; Brun et al., 2019; Maurer et al., 2019; Carrivick et al., 2020; Liu et al., 2020). Remote sensing is the most viable way to map glacial lakes and detect their spatio-temporal changes in the high-elevation zones where in situ accessibility is extremely low (Huggel et al., 2002; Quincey et al., 2007). Studies in glacial lake inventories using satellite observations have been heavily conducted at regional scales recently, such as in the Tibetan Plateau (Zhang et al., 2015), the Himalaya (Gardelle et al., 2011; Nie et al., 2017), the HMA (Wang et al., 2020; Chen et al., 2021), the Tien Shan (Wang et al., 2013), the Alaska (Rick et al., 2022), the Greenland (How et al., 2021) and the northern Pakistan (Ashraf et al., 2017). However, the latest glacial lake mapping in 2020 is still absent along the CPEC. Among existing studies, Landsat archival images are the most widely used due to their multi-decadal record of earth surface observations, reasonably high spatial resolution (30 m), and publicly available distribution (Roy et al., 2014). Freely available Sentinel-2 satellite images show a better potential than Landsat in glacial lake mapping and inventories due to their higher spatial resolution (10 m) and a global coverage, but have only been available since late 2015 (Williamson et al., 2018; Paul et al., 2020). Glacial lake inventories using Sentinel-2 images are relatively scarce at regional scales, and studies of the latest glacial lake mapping as well as comparisons of glacial lake dataset derived from Sentinel-2 and Landsat observations are still lacking.

Discrepancies between various glacial lake inventories (Zhang et al., 2015; Shugar et al., 2020; Wang et al., 2020; Chen et al., 2021; How et al., 2021) result from differences in mapping methods, minimum mapping units, definition of glacial lakes, time periods, data sources and other factors. For example, manual vectorization method was widely adopted at the earlier stage for its high accuracy. However, it is time-consuming associated with high labor intensity and is only practical at regional scales (Zhang et al., 2015; Wang et al., 2020). Automated and semi-automated lake mapping methods, such as multi-spectral index classification (Gardelle et al., 2011; Nie et al., 2017; Zhang et al., 2018; How et al., 2021), have been developed to improve the efficiency of glacial lake inventories using optical images, although manual modification is often unavoidable to assure the quality of lake data impacted by cloud cover, mountain shadows, seasonal snow cover and frozen lake surfaces (Sheng et al., 2016; Wang et al., 2017, 2018). Backscatter images from Synthetic Aperture Radar (SAR) (Wangchuk and Bolch, 2020; How et al., 2021) were used to remove the impact of cloud cover for lake mapping. Besides, other approaches such as hydrological sink detection using DEM (How et al., 2021) and land surface temperature-based detection method (Zhao et al., 2020) were also used for lake inventories. Different classification methods impact the results of lake mapping and monitoring. So far, we are lacking a unified standard for the classification system of glacial lakes (Yao et al., 2018). Existing classification systems are generally used for their individual research purposes, mainly based on the relative positions of glacial lakes and glaciers, the supply conditions of glaciers, and the attributes of dams. In addition to different classification standards, the same type of glacial lakes may also have different names given by different scholars. For example, ice-marginal (Carrivick and Quincey, 2014; Carrivick et al., 2020), ice-contact (Carrivick and Tweed, 2013) and proglacial (Nie et al., 2017) lakes all represent glacial lakes sharing the boundary with glaciers. Glacier lakes in currently available datasets have been traditionally categorized by their spatial relationship with upstream glaciers (Gardelle et al., 2011; Wang et al., 2020; Chen et al., 2021), and classification attributes considering the formation mechanism and the properties of dams are rare or incomplete in the CPEC (Yao et al., 2018; Li et al., 2020). Dam type classification of glacial lakes provides a crucial attribute for glacier-lake interactions and risk assessment (Emmer and Cuřín, 2021). Therefore, an up-to-date glacial lake dataset with critical, quality-assured parameters (e.g. lake area, volume and type) is necessary.

This study aims to (1) present an up-to-date glacial lake dataset in the CPEC in 2020 using both Landsat 8 and Sentinel-2 images to accurately document its detailed lake distribution; (2) present two historical glacial lake datasets for the CPEC to show extent in 1990 and 2000 using consistent 30-m Landsat images to reveal glacial lake changes at three time periods (1990, 2000 and 2020); and (3) generate a range of critical attributes for glacial lake inventories to benefit studies on water resource evaluation, risk assessment of GLOFs, glacier –lake evolution modeling in the HMA.

2 Study area

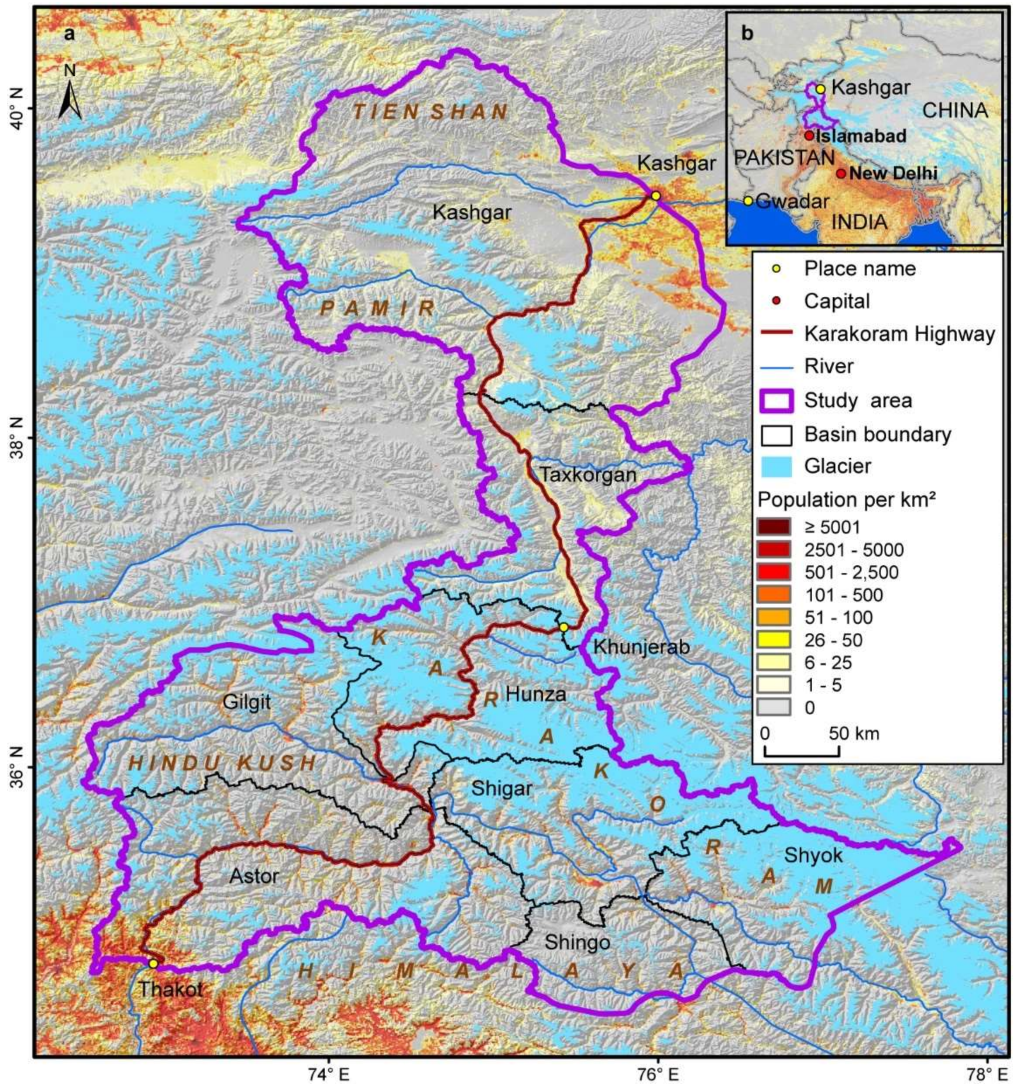


Figure 1. Location of the study area associated with distribution of glaciers (RGI Consortium, 2017), mountains, basins and population (Rose et al., 2021) (a), and its location within the CPCE (b).

The northern part of the CPEC is selected as the study area (Figure 1). The CPCE, originating from Kashgar of the Xinjiang Uygur Autonomous region, China and extending to Gwadar Port, Pakistan (Ullah et al., 2019; Yao et al., 2020), is connecting China and Pakistan via the only Karakoram Highway. The study area covers all the drainage basins along Karakoram Highway starting from Kashgar and ending at Thakot, with a total area of ~125,000 km^2. The upper Indus basins beyond the Pakistani-administrated border are excluded in this study due to spatial coverage of the CPCE. The entire study area is divided into eight sub-basins, covering most of the Karakoram with the highest elevation up to 8611 m, western Himalaya and Tien Shan, eastern Hindu Kush and Pamir Mountains. The 9710 glaciers in the study area cover a total area of 17,447 km^2 and nearly 60% of glaciers are distributed in the Karakoram (5818 glaciers with a total area of 14,067.52 km^2) (RGI Consortium, 2017). Most glaciers in the western Himalaya and eastern Hindu Kush are losing mass in the context of climate change (Kääb et al., 2012; Yao et al., 2012; Brun et al., 2017; Shean et al., 2020; Hugonnet et al., 2021), whereas the glaciers in the eastern Karakoram and Pamir have shown unusually little changes, including unchanged, retreated, advanced and surged glaciers (Hewitt, 2005; Kääb et al., 2012; Bolch et al., 2017; Brun et al., 2017; Shean et al., 2020; Nie et al., 2021). The spatially heterogeneous distribution and changes of glaciers are primarily explained as a result of differences in the dominant precipitation-bearing atmospheric circulation patterns that include the winter westerlies the Indian summer monsoon, their changing trends and their interactions with local extreme topography (Yao et al., 2012; Azam et al., 2021; Nie et al., 2021).

3 Data sources

Both Landsat and Sentinel-2 images have been employed to map glacial lakes between 1990 and 2020 in the CPEC (Figure 2). A total number of 71 Landsat Thematic Mapper (TM), Thematic Mapper Plus (ETM+) and Landsat 8 Operational Land Imager (OLI) images with a consistent spatial resolution of 30 m were downloaded from the United States Geological Survey Global Visualization Viewer (GloVis, https://glovis.usgs.gov/) to be used to create glacial lake inventories in 1990, 2000 and 2020. High-quality Landsat-5 images around 2010 are insufficient to cover the entire study area, so we were unable to map lakes in 2010 due to Landsat-7's scan-line corrector errors and significant cloud covers. In addition, 39 Sentinel-2 images (23 scenes in 2020) were downloaded from Copernicus Open Access Hub (https://scihub.copernicus.eu/) to produce the 10-m resolution glacial lake inventory in 2020. All images used in this study have been orthorectified before download, but we still find that one Sentinel-2 image was not well matched with Landsat images, leading to the discrepancy between the two glacial lake datasets. We manually georeferenced the shifted image to minimize the difference between Sentinel and Landsat derived glacial lakes.

Cloud and snow covers heavily affect the usability of optical satellite images (Wulder et al., 2019) and their availability in the entire study area, so we took advantage of the images acquired before and after each of the baseline years 1990, 2000 and 2020 to construct the glacial lake inventories. Only 4 images in 1990 (the largest covering the study area), 16 images in 2000 and 23 images in 2020 were used for matching baseline year. Spatially, high-quality images in given baseline years were preferentially chosen, or we selected one or more alternative images acquired in adjacent years to delineate glacial lakes by removing the effect of cloud and snow covers. To minimize the impact of intra-annual changes of glacial lakes, most of used images (82% for Sentinel-2 and 75% for Landsat) were acquired from August to October in the given baseline year with cloud coverage of <20% for each image. For some specific scenes where cloud cover exceeded the threshold of 20%, we selected more than one image to remedy the effect of cloud contamination (Nie et al., 2010, 2017; Jiang et al., 2018).

Other datasets used include the Randolph Glacier Inventory version 6.0 (Pfeffer et al., 2014; RGI Consortium, 2017) and the Glacier Area Mapping for Discharge from the Asian Mountains (GAMDAM) glacier inventory (Sakai, 2019). These two glacier datasets were used to determine glacial lake types, such as ice-contact, ice-dammed and unconnected-glacier-fed lakes. The Shuttle Radar Topography Mission Digital Elevation Model (SRTM DEM) at a 1-arc second (30 m) resolution (Jarvis et al., 2008) was employed to extract the altitudinal characteristics of the glacial lakes. The absolute vertical accuracy of the SRTM DEM is 16 m (90%) (Rabus et al., 2003; Farr et al., 2007). We also applied other published glacial lake datasets for comparative analysis. They include the glacial lake inventories of HMA in 1990 and 2018 downloaded from http://doi.org/10.12072/casnw.064.2019.db (Wang et al., 2020), the Third Pole region in 1990, 2000 and 2010 publicly shared at http://en.tpedatabase.cn/ (Zhang et al., 2015), the Tibet Plateau from 2008 to 2017 accessed at https://doi.org/10.5281/zenodo.3700282 (Chen et al., 2021), and the entire world in 1990, 2000 and 2015 provided at https://nsidc.org/data /HMA_ GLI/versions/1 (Shugar et al., 2020). In addition, field survey data collected between 2017 and 2018 were also used to assist in lake mapping and glacial lake type classification.

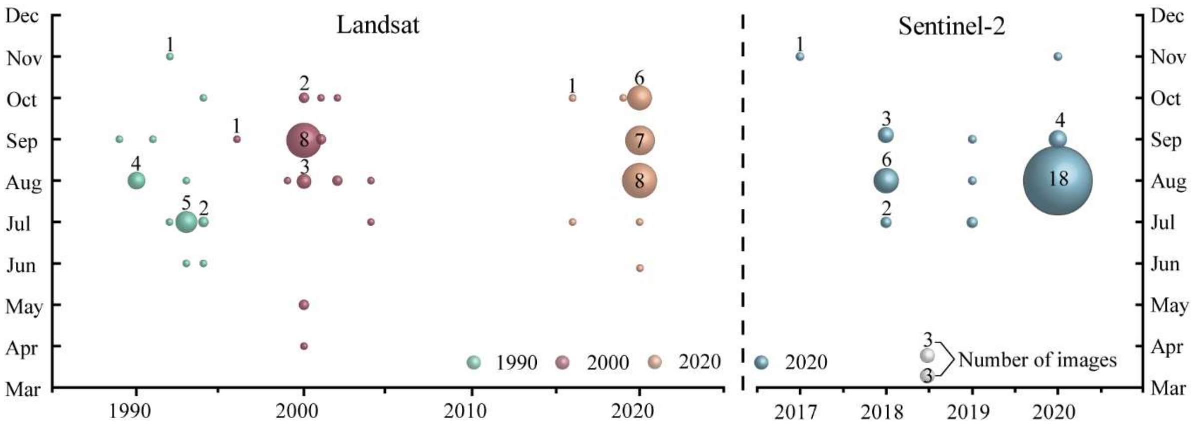


Figure 2. Acquisition years and months of Landsat and Sentinel-2 images selected for glacial lake inventories. The bubble size indicates the available high-quality image number.

4 Glacial lake inventory methods

4.1 Definition of glacial lakes

We consider a glacial lake as one that formed as a result of modern or ancient glaciation. Contemporary glacial lakes are easily recognized using a combination of glacier inventories and remote sensing images. Ancient glacial lakes can be identified from periglacial geomorphological characteristics, including moraine remnants and U-shaped valleys that are discernible from satellite observations (Post and Mayo, 1971; Westoby et al., 2014; Nie et al., 2018; Martín et al., 2021). A 10-km buffering distance of RGI 6.0 glacier boundaries that has been widely used in previous studies (Zhang et al., 2015; Wang et al., 2020), was created to help mapping glacial lakes. A few glacial lakes in the study area (a total of 84 lakes for Sentinel-2 dataset and 55 lakes for Landsat dataset in 2020) beyond the buffering zone, located near buffering boundaries, were intentionally included due to clear evidence of glaciation (**Figure 3**). Landslide-dammed lakes (Chen et al., 2017) in the buffering zone were excluded in our inventories because of their irrelevance to glaciation. All glacial lakes in the study area were mapped according to our definition. We were able to implement this definition by carefully leveraging the spectral properties of glacial lakes and the periglacial geomorphological features that are often evident in remote sensing images (see more in sections 4.3 and 4.4).

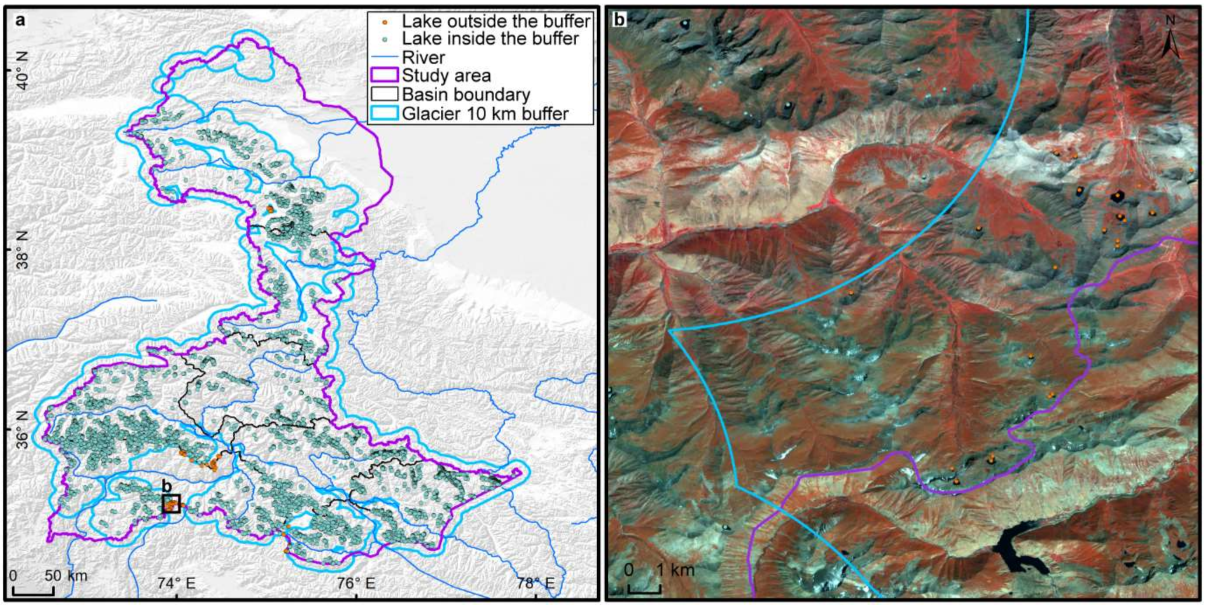


Figure 3. The 10-km buffer zone of RGI 6.0 glacier boundaries (a) and Sentinel-derived glacial lakes located near buffering boundary within the study area (b).

4.2 Interactive lake mapping

A human-interactive and semi-automated lake mapping method (Wang et al., 2014; Nie et al., 2017, 2020) was adopted to accurately extract glacial lake extents using Landsat and Sentinel-2 images, based on the Normalized Difference Water Index (NDWI) (Mcfeeters, 1996). The NDWI uses the green and near infrared bands and is calculated by the following equation:

$$NDWI = \frac{Band_{Green} - Ban\ _{NIR}}{Band_{Green} + Band_{NIR}} \quad (1)$$

where the green band and near infrared band were provided by both Landsat and Sentinel multispectral images.

Specifically, the method calculated the NDWI histogram based on the pixels with each user-defined and manually-drawn region of interest. The NDWI threshold that separates lake surface from land was interactively determined by screening the NDWI histogram against the lake region in the imagery (Wang et al., 2014; Nie et al., 2020). This way, the determined NDWI threshold can be well-tuned to adapt various spectral conditions of the studied glacier lakes. The raster lake extents segmented by the thresholds were then automatically converted to vector polygons. We first completed the glacial lake inventory in 2020 using this interactive mapping method, and the 2020 inventory was then used as a reference to facilitate the lake mapping for other periods.

The minimum mapping unit (MMU) was set to 5 pixels for both Landsat (0.0045 km^2) and Sentinel-2 images (0.0005 km^2) in this study. MMU determines the total number and area of glacial lakes in the dataset, and varies in the previous studies, such as 3 pixels (Zhang et al., 2015), 6 pixels (Wang et al., 2020), or 9 pixels (Chen et al., 2021) for a regional scale, or 55 pixels (Shugar et al., 2020) for a global scale. While a smaller threshold leads to a large quantity of lakes mapped, it also generates larger mapping noises or uncertainties.

Considering this signal-noise balance and our focus on identifying prominent glacier lake dynamics in the study area, we opted to use 5 pixels as the MMU for both Landsat and Sentinel-2 images.

Several procedures were taken to assure the quality assurance and quality control for lake mapping, including 1) visual inspection and modification using the threshold-based mapping method for each lake according to Landsat, Sentinel-2 and Google Earth high-resolution images overlaying preliminarily lake boundary extraction at the given time period; 2) time series check for Landsat-derived glacial lake datasets from 1990 and 2020, and cross-check between Landsat and Sentinel-2-derived lake dataset in 2020 to reduce errors of omission and commission; 3) topological validation of glacial lake mapping, such as repeated removal, elimination of small sliver polygons; and 4) logical check for lake types between two classification systems of glacial lakes. False lake extents resulting from cloud or snow cover, lake ice, and topographic shadows (Nie et al., 2017, 2020) were modified using previous semi-automated mapping method based on alternative images acquired in adjacent years. Those procedures were time-consuming, but helped to minimize the effect of cloud and snow covers, lake mapping errors, and to maximize the quality of the produced lake product and the derived glacial lake changes.

4.3 Classification of glacial lakes

Two glacial lake classification systems (GLCS) have been established based on relationship of interaction between glacial lakes and glaciers as well as lake formation mechanism and dam material properties. In the first GLCS (GLCS1), glacial lakes were classified into four types based on their spatial relationship to upstream glaciers: supraglacial, ice-contact, unconnected-glacier-fed lakes, and non-glacier-fed lakes according to Gardelle et al. (2011) and Carrivick et al. (2013). Alternatively, combining the formation mechanism of glacial lakes and the properties of natural dam features, glacial lakes were classified into five categories (herein named GLCS2) modified from Yao's classification system (2018): supraglacial, end-moraine-dammed, lateral-moraine-dammed, glacial-erosion lakes and ice-dammed lakes. Subglacial lakes were excluded due to the mapping challenge from spectral satellite images alone. Characterization and examples for each type are provided in Table 1 and Table 2. Individual glacial lakes were categorized to the specific types for each GLCS according to available glacier inventory data, geomorphological and spectral characteristics interpreted from Landsat, Sentinel and Google Earth images. The synergy of these two GLCSs is beneficial to predicting glacier-lake evolutions and providing fundamental data for water resource and glacial lake disaster risk assessment.

Table 1. Classification system of glacial lake types according to the relationship between glacial lakes and glaciers (© Google Earth 2019). Glacier outlines are from RGI 6.0 (RGI Consortium, 2017), and the yellow marker represents target lake.

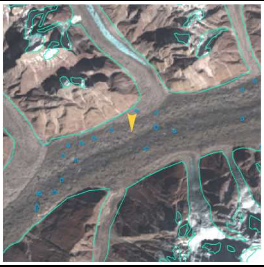
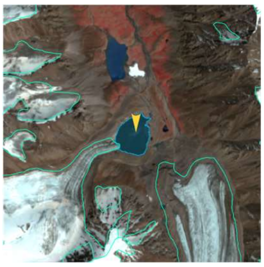
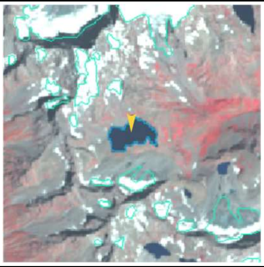
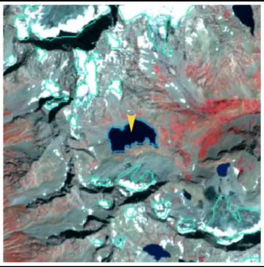
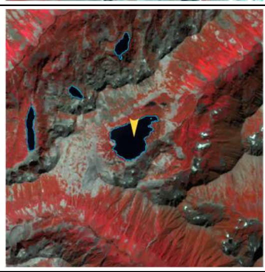
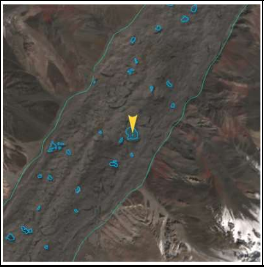
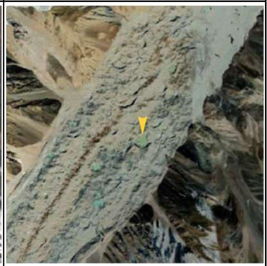
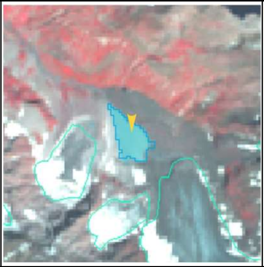
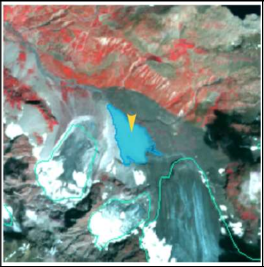
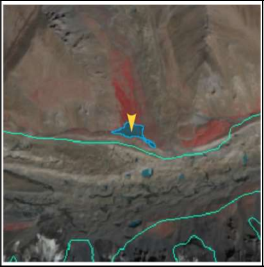
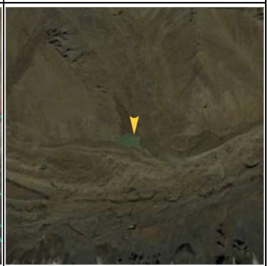
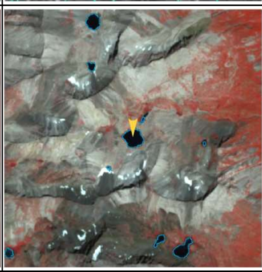
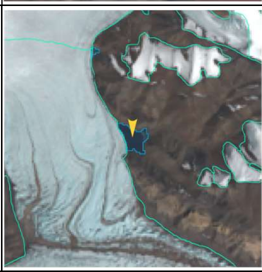
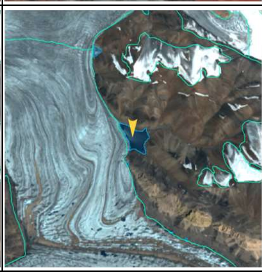
Lake types	Characteristics	Landsat	Sentinel-2	Google Earth
Supraglacial	Lakes formed on the surface of glaciers, generally dammed by ice and thin debris. Case location: 35°43'49.74" N 76°13'53.88" E			
Ice-contact	Lakes dammed by moraine, ice or bedrock, supplied by glacial meltwater and shared boundary with glaciers. Case location: 39°09'32.40" N 73°43'12.00" E			
Unconnected-glacier-fed	Lakes currently supplied by upstream glacial meltwater but disconnected with glaciers. Case location: 35°47'60.00" N 72°55'15.60" E			
Non-glacier-fed	Lakes formed by glaciology, dammed by moraine or bed rock, and currently not supplied by glacial meltwater. Case location: 34°50'39.99" N 74°48'29.31" E			

Table 2. Classification system of glacial lake types according to the formation mechanism of glacial lakes and dam material properties (© Google Earth 2019). Glacier outlines from RGI 6.0 (RGI Consortium, 2017), and the yellow marker represents target lake.

Lake types	Characteristics	Landsat	Sentinel-2	Google Earth
Supraglacial	Lakes formed on the surface of glaciers, generally dammed by ice and thin debris. Case location: 36°46'7.39" N 74°20'7.59" E			
End-moraine-dammed	Lakes formed behind moraines as a result of glacier retreat and downwasting. Case location: 35°42'50.40" N 73°09'57.60" E			
Lateral-moraine-dammed	Lakes formed behind lateral glacial moraine ridges and dammed by debris, different from ice-dammed glacial lake. Case location: 38°28'45.62" N 75°20'52.30" E			
Glacial-erosion	Lakes formed in depressions created by glacial over-deepening. Bedrock dam dominates, partially superimposed by top moraine in rugged terrain. Dams are unclear in the satellite images. Case location: 35°55'55.56" N 73°38'20.13" E			
Ice-dammed	Lakes formed behind glaciers, dammed by glacier ices (partially covered by debris on the top). Case location: 35°28'31.32" N 77°30'46.81" E			

4.4 Attributes of glacial lake data

A total of 18 attribute fields were input into our glacial lake datasets (Table 3). They include lake location (longitude and latitude), lake elevation (centroid elevation), orbital number of the image source, image acquisition date, lake area, lake perimeter, lake types of the two GLCSs, mapping uncertainty, lake water volume and the country, sub-basin, and mountain range associated with the lake. Amongst the attributes, lake location was calculated based on the centroid of each glacial lake polygon associated with the DEM, N represents northing and E represents easting. Orbital number of the image source was filled with the corresponding satellite image, with the codes expressed as "PxxxRxxx" or "Txxxxx", where P and R indicate the path and row for Landsat image and T represents the tile of Sentinel-2 image associated with 5 digit code of military grid reference system. Area and perimeter were automatically calculated based on glacial lake extents. Lake water volume was estimated by area-volume empirical equation (Cook and Quincey, 2015). Lake types were attributed using the characterization and interpretation marks described in Section 4.3. Mapping uncertainty was estimated using our modified equation which will be introduced in section 4.5 and appendix tutorial. Located country, sub-basin and mountain range of each glacial lake was identified by overlapping the geographic boundaries of countries, basins and mountain ranges.

Table 3. Attributes of glacial lake dataset.

Field Name	Type	Description	Note
FID or OBJECTID	Object ID	Unique code of glacial lake	Number
Shape	Geometry	Feature type of glacial lake	Polygon
Latitude	String	Latitude of the centroid of glacial lake polygon	Degree minute second
Longitude	String	Longitude of the centroid of glacial lake polygon	Degree minute second
Elevation	Double	Elevation of the centroid of glacial lake polygon	Unit: meter above sea level
IMGSOURCE	String	Path and row numbers for Landsat image based on World Reference System 2 or Tile number for Sentinel image based on military grid reference system	PxxxRxxx or Txxxxx
ACQDATE	String	Acquisition date of source image	YYYYMMDD
GLCS1	String	The first classification system of glacial lakes based on relationship of interaction between glacial lakes and glaciers	Supraglacial, Ice-contact, Unconnected-glacier-fed, None-glacier-fed
GLCS2	String	The second classification system of glacial lakes based on lake formation mechanism	Supraglacial, End-moraine-dammed, Lateral-

Field Name	Type	Description	Note
		and dam material properties	moraine-dammed, Glacial-erosion and Ice-dammed
Basin	String	Basin name where glacial lake locates in	
Mountain	String	Mountain name where glacial lake locates in	
Country	String	Country name where glacial lake locates in	
Perimeter	Double	Perimeter of glacial lake boundary	Unit: meter
Area	Double	Area of glacial lake coverage	Unit: square meter
Uncertainty	Double	Uncertainty of glacial lake mapping estimated based on modified Hanshaw's equation (2014)	Unit: square meter
Volume	Double	Water volume of glacial lake estimated by area-volume empirical equation	Unit: square meter
Operator	String	Operator of glacial lake dataset	Muchu, Lesi
Examiner	String	Examiner of glacial lake dataset	Yong, Nie

4.5 Error and uncertainty assessment

4.5.1 Improved uncertainty estimating method

We modified Hanshaw's (2014) equation that had been used to calculate lake-area mapping uncertainty. Lake perimeter and displacement error are widely used to estimate the uncertainty of glacier and lake mapping from satellite observation (Carrivick and Quincey, 2014; Hanshaw and Bookhagen, 2014; Wang et al., 2020). Hanshaw and Bookhagen (2014) proposed an equation to calculate the error of area measurement by the number of edge pixels of the lake boundary multiplied by half of a single pixel area. The number of edge pixels is simply calculated by the perimeter divided by the grid size. The equation is expressed as below:

$$Error(1\sigma) = \frac{P}{G} \times 0.6872 \times \frac{G^2}{2} \quad (2)$$

$$D = \frac{Error(1\sigma)}{A} \times 100\% \quad (3)$$

Where G is the cell size of the remote sensing imagery (10 m for Sentinel-2 image and 30 m for Landsat image). P is the perimeter of individual glacial lake (m), and the coefficient of 0.6872 (1σ), which means nearly 69% of the edge pixels are subject to errors (Hanshaw and

Bookhagen, 2014), was chosen assuming that area measurement errors follow a Gaussian distribution. Relative error (D) was calculated by equation 3, in which A is the area of an individual glacial lake.

In the original equation 2, the number of edge pixels varies by the shape of lake and is indicated by $\frac{P}{G}$. However, the pixels in the corner are double counted (Figure 4). The total number of repeatedly calculated edge pixels equals the number of inner nodes. Therefore, we adjusted the calculation of the actual number of edge pixels as the maximum of edge pixels ($\frac{P}{G}$) subtracting the number of inner nodes. Accordingly, the equation of uncertainty estimation for lake mapping is modified as below:

$$Error(1\sigma) = \left(\frac{P}{G} - N_{Inner}\right) \times 0.6872 \times \frac{G^2}{2} \quad (4)$$

Where N_{Inner} is the number of inner nodes (inflection points) of each lake. The modified equation is also suitable for lakes with islands (as illustrated in Figure 4b).

For polygons without islands (Figure 4a), use the following equation:

$$N_{Inner} = \left(\frac{N_{Total} - 4 - 1}{2}\right) \quad (5)$$

N_{Total} is the total number of nodes, including both the outer and inner. N_{Total} is calculated by the "Field Calculator" in ArcGIS, in some cases, it is necessary to remove the redundant nodes before calculating the total number of nodes (See the Appendix for more details). An inner node is a polygon vertex where the interior angle surrounding it is greater than 180 degrees. An outer node is the opposite of the inner node, where the interior angle is less than 180 degrees. We found that the outer nodes are usually four more than the inner nodes in our glacial lake dataset. The total nodes in ArcGIS contain one overlapping node to close the polygon, meaning the endpoint is also the startpoint. This extra count was deleted in the calculation (equation 5).

For polygons with island (Figure 4b) use the following equation:

$$N_{Inner} = \left(\frac{N_{Total} - (N_{Island} + 1) \times 5}{2}\right) \quad (6)$$

N_{Island} is the number of islands within each polygon. A calculation method of N_{Island} is given in the Appendix.

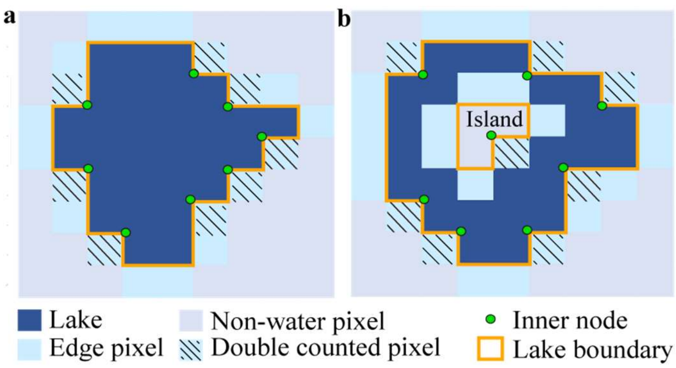


Figure 4. Sketch of estimating the actual edge pixels for uncertainty calculation of individual glacial lake (with (a) and without islands (b)).

4.5.2 Validation of glacial lake mapping

A total of 89 glacial lakes were selected by stratified random sampling and manually digitized based on the Google Earth high resolution images to further validate the absolute error of the glacial lake mapping in 2020 due to lacking of field measurements for glacial lakes in the study area. During the sampling, we set a regulation of minimum lake area greater than 4500 m^2 and relative differing between Landsat- and Sentinel-derived lake areas less than 18% (nearly equaling to the average relative error of ±17.36% for Landsat lake mapping) to minimize the effect of lake changes from multi-temporal satellite observations in circa 2020. The 89 sample lakes range from 0.005 km^2 to 0.802 km^2 with a median (standard deviation) size of 0.047±0.134 km^2 and total area of 8.033 km^2 for Landsat-derived dataset, whereas ranging from 0.005 km^2 to 0.849 km^2 with a median (standard deviation) size of 0.045±0.144 km^2 and total area of 8.447 km^2 for Sentinel-derived dataset.

5 Results

5.1 Glacier lake distribution and changes observed from Landsat

We mapped 2,234 glacial lakes for 2020 across the studied CPEC from Landsat-8 images, with a total area of 86.31±14.98 km^2 (Figure 5a and b). Unconnected-glacier-fed lakes are dominant in the first classification system, followed by non-glacier-fed lakes (Figure 6) whereas glacial-erosion lakes dominate at both number (1478) and area (57.02 km^2) in the second classification system (Figure 7), followed by end-moraine-dammed lakes and supraglacial lakes. Among the classified lakes, 137 are ice-contact lakes and cover an area of 5.56 km^2, implying a higher mean size of ice-contact lakes than supraglacial lakes.

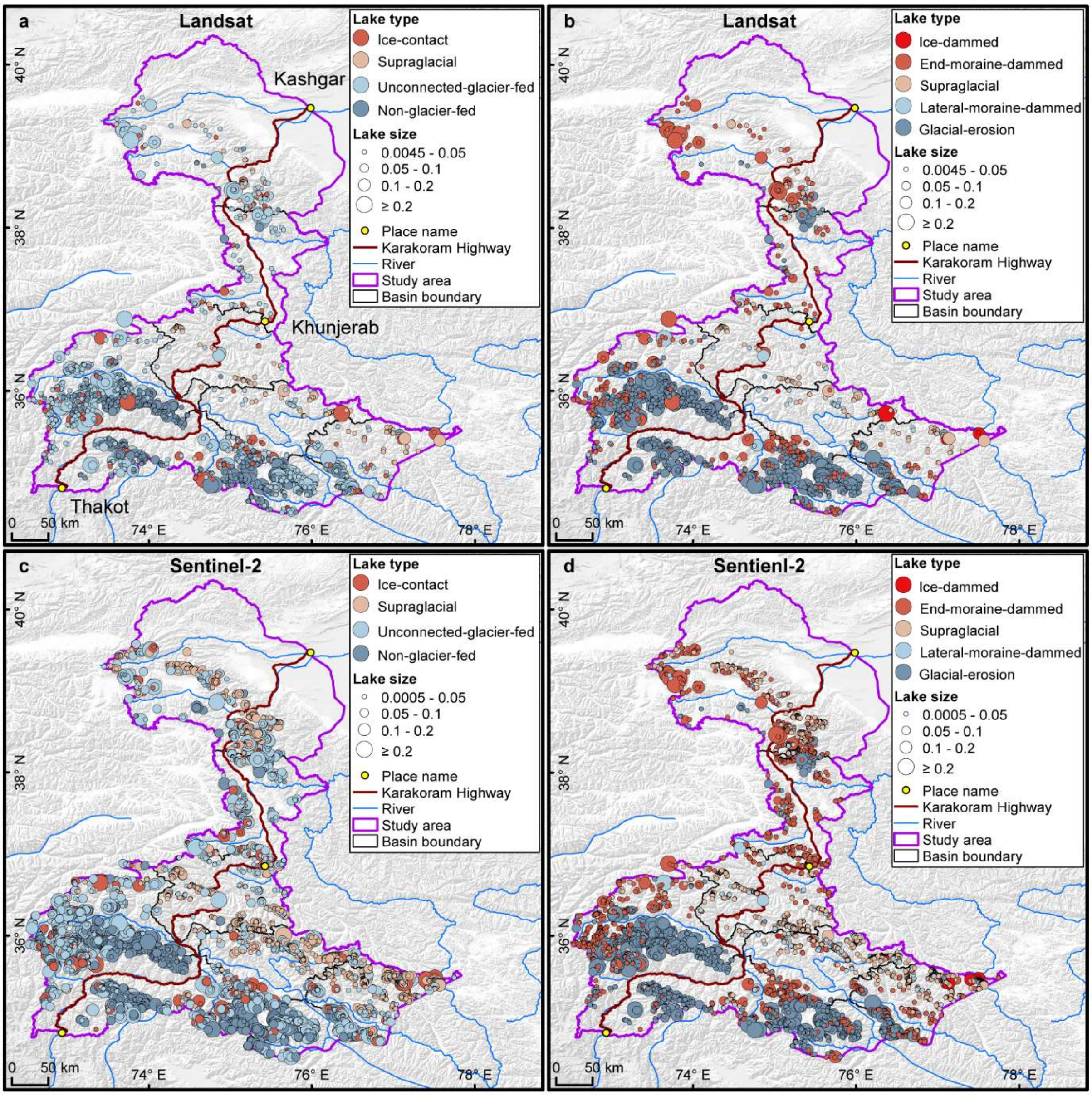


Figure 5. Distribution of glacial lakes in 2020 extracted from Landsat (a, b) and Sentinel-2 (c, d) images. Panels a and c are classified by GLCS1, and GLCS2 for sub-graph b and d.

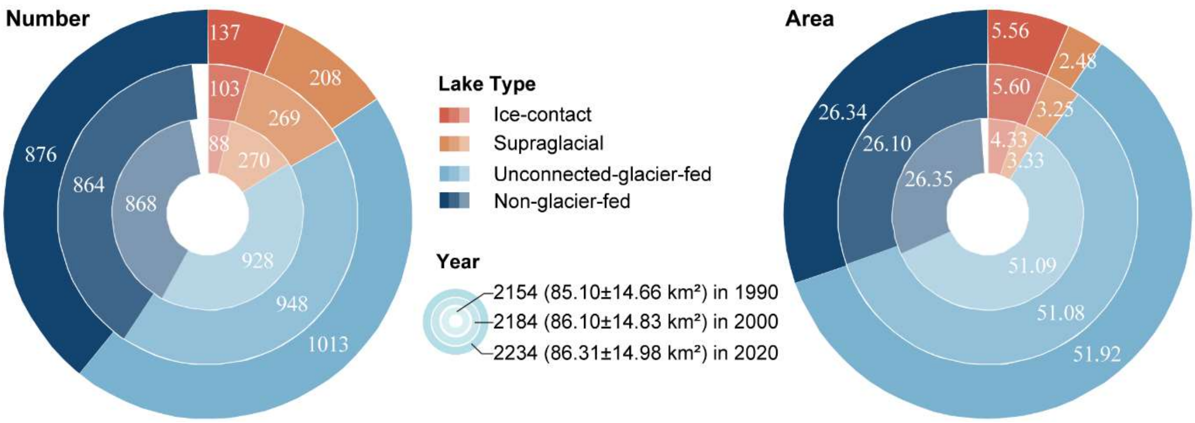


Figure 6. Number and area of different types of glacial lakes classified based on the condition of glacier

supply in the study area (GLCS 1). The outermost ring represents glacial lake data in 2020, middle ring for 2000 and innermost ring for 1990. Lake number and area in 2020 were selected as reference, meaning a concept of "100 %" for a complete ring. Labeled values are scaled in degrees rather the radius of rings.

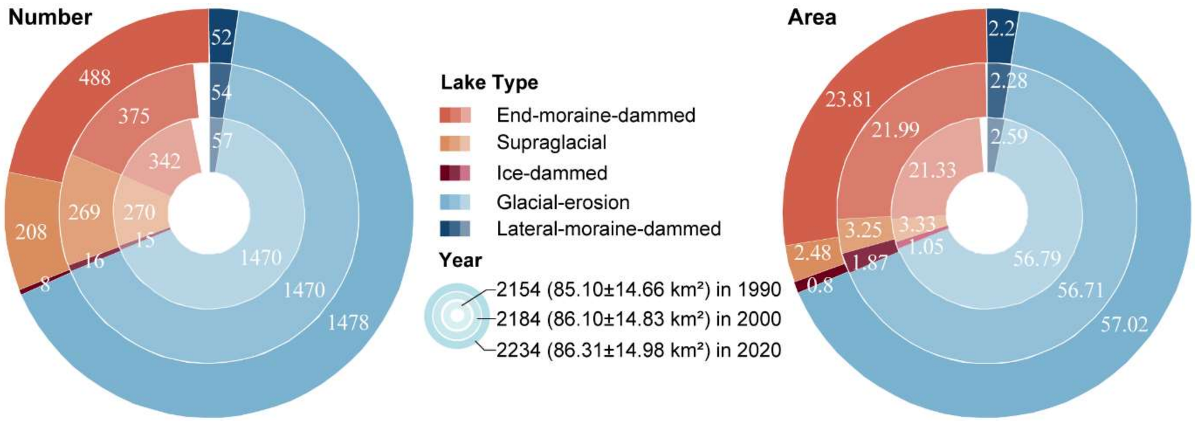


Figure 7. Number and area of different types of glacial lakes classified based on glaciation and nature of dam in the study area (GLCS 2). The outermost ring represents glacial lake data in 2020, middle ring for 2000 and innermost ring for 1990. Lake number and area in 2020 were selected as reference, meaning a concept of "100 %" for a complete ring. Labeled values are scaled in degrees rather the radius of rings.

The total number and area of glacial lakes in the study remain relatively stable with a slight increase between 1990 and 2020, and the changes in count and area among various types of glacial lakes vary substantially (Figure 6 and Figure 7). From 1990 to 2020, the total number of glacial lakes increased by 80 or 3.70%, while the area grew by 1.21 km^2 (or 1.42%). In GLCS1, unconnected-glacier-fed lakes have the largest increase in number, followed by ice-contact and non-glacier-fed lakes, whereas supraglacial lakes decreased by 62 in count. Ice-contact lakes expanded by 1.24 km^2 (equaling an increase of 26% in ice-contact lakes), contributed one third of the total area increase. Supraglacial lakes decreased by 0.85 km^2 in area whereas the areas of unconnected-glacier-fed and non-glacier-fed lakes remained stable as a result of disconnections from glaciers (Figure 6). In GLCS2, end-moraine-dammed lakes increased by 2.48 km^2 and contributed most of the glacier lake area expansion, whereas supraglacial, ice-dammed and lateral-moraine-dammed lakes decreased slightly in both number and area. Glacial-erosion lakes accounted for the maximum percentage (about 66% for both count and area) in each time period and remained stable (Figure 7).

5.2 Glacier lake distribution observed from Sentinel-2

Sentinel-derived results shows that there are 7,560 glacial lakes (103.70±8.45 km^2) in 2020 across the entire CPEC (Table 4) under a MMU of 5 pixels (500 m^2). Compared with Landsat-derived product, glacial lakes from Sentinel-2 have similar spatial distribution characteristics (Figure 5); meanwhile, a larger quantity of glacier lakes, with more accurate boundaries and a greater total lake area, were generated from Sentinel-2 images. The smallest size class (0.0005-0.0045 km^2) contains the maximum lake number (4,969) but the least lake area (7.73±2.62 km^2) (Table 4), which is not available in the Landsat-derived lake data due to a coarser spatial resolution. In each size class, there are also a higher number of larger glacial lakes from Sentinel than that from Landsat images. The discrepancy is mainly attributed to the inconsistency of spatial resolutions and image acquisition dates.

Table 4. Count and area of glacial lakes mapped from Sentinel-2 and Landsat images in 2020 between various size classes.

Lake size km^2	Glacial lakes from Sentinel-2 count (km^2)	Glacial lakes from Landsat count (km^2)	Overlap % (%)
0.0005-0.0045	4969（7.73±2.62）	—	—
0.0045-0.05	2182（35.52±3.72）	1870（31.47±9.57）	85.70 (88.60)
0.05-0.1	237（16.37±0.89）	204（14.07±2.18）	86.08 (85.95)
0.1-0.2	122（16.88±0.68）	115（15.91±1.83）	94.26 (94.25)
⩾0.2	50（27.20±0.54）	45（24.86±1.40）	90.00 (91.40)
Total	7560（103.70±8.45）	2234（86.31±14.98）	—

Note: Overlap % (%) represent the rates in count and area calculated by dividing Landsat-derived lake data by Sentinel-derived data in the same size class respectively.

6 Discussions

6.1 Error and uncertainty of lake mapping

The uncertainty estimated from our improved equation shows that the relative error of individual glacial lake decreases when lake size increases or cell size of remote sensing images reduces (Lyons et al., 2013; Carrivick and Quincey, 2014) (Figure 8). Total area error of glacial lakes in study area is approximate ±14.98 km^2 and ±8.45 km^2 in 2020 for Landsat and Sentinel-2 dataset, respectively, and the average relative error is ±17.36% and ±8.15%. Generally, small lakes have greater relative errors. For example, the mean relative error is 35.38% for Landsat derived glacial lakes between 0.0045 and 0.1 km^2 and 10.63% for glacial lakes greater than 0.1 km^2. The mean area error of Sentinel-derived glacial lakes is almost one third of that extracted from Landsat images for glacial lakes of all or specific size group. Because the relative error was estimated as a function of satellite image spatial resolution and lake perimeter, the calculated error for large lake is proportionally smaller than that of small lake (Salerno et al., 2012) and the error for Landsat-derived lake is naturally greater than that of Sentinel-derived lake at the same size group.

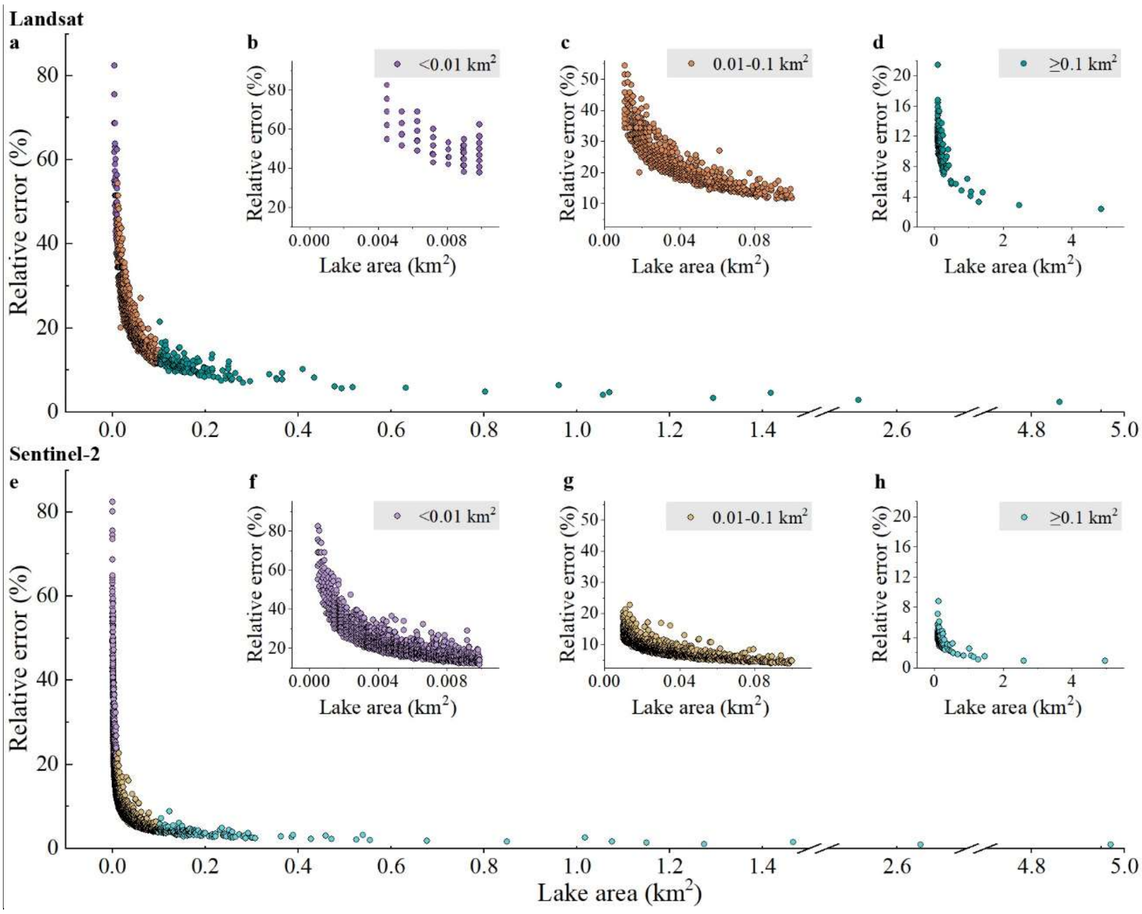


Figure 8. Estimated relative error for glacial lakes of all or specific size ranges in study area. Error estimation is based on the modified equation and lake data extracted from Landsat (a-d) and Sentinel-2 images (e-h).

Our Landsat- and Sentinel-derived glacial lake dataset match well lake boundaries in Google Earth high resolution images (Figure 9). A dense cluster of validation samples along the 1:1 line indicates a high accuracy in lake mapping (Figure 9c and d). The error of 89 sample lakes is 5.48% in total area between Landsat- and Google Earth-derived data, whereas 0.61% for Sentinel- and Google Earth-derived data. The median (±standard deviation) in discrepancy of individual lake area is 7.66±4.96 % for Landsat- and Google Earth-derived data, whereas 4.46±4.62 % for Sentinel- and Google Earth-derived data. Our glacial lake dataset shows a satisfactory mapping accuracy, and of which Sentinel-derived lake data performs more accurate than those from Landsat images.

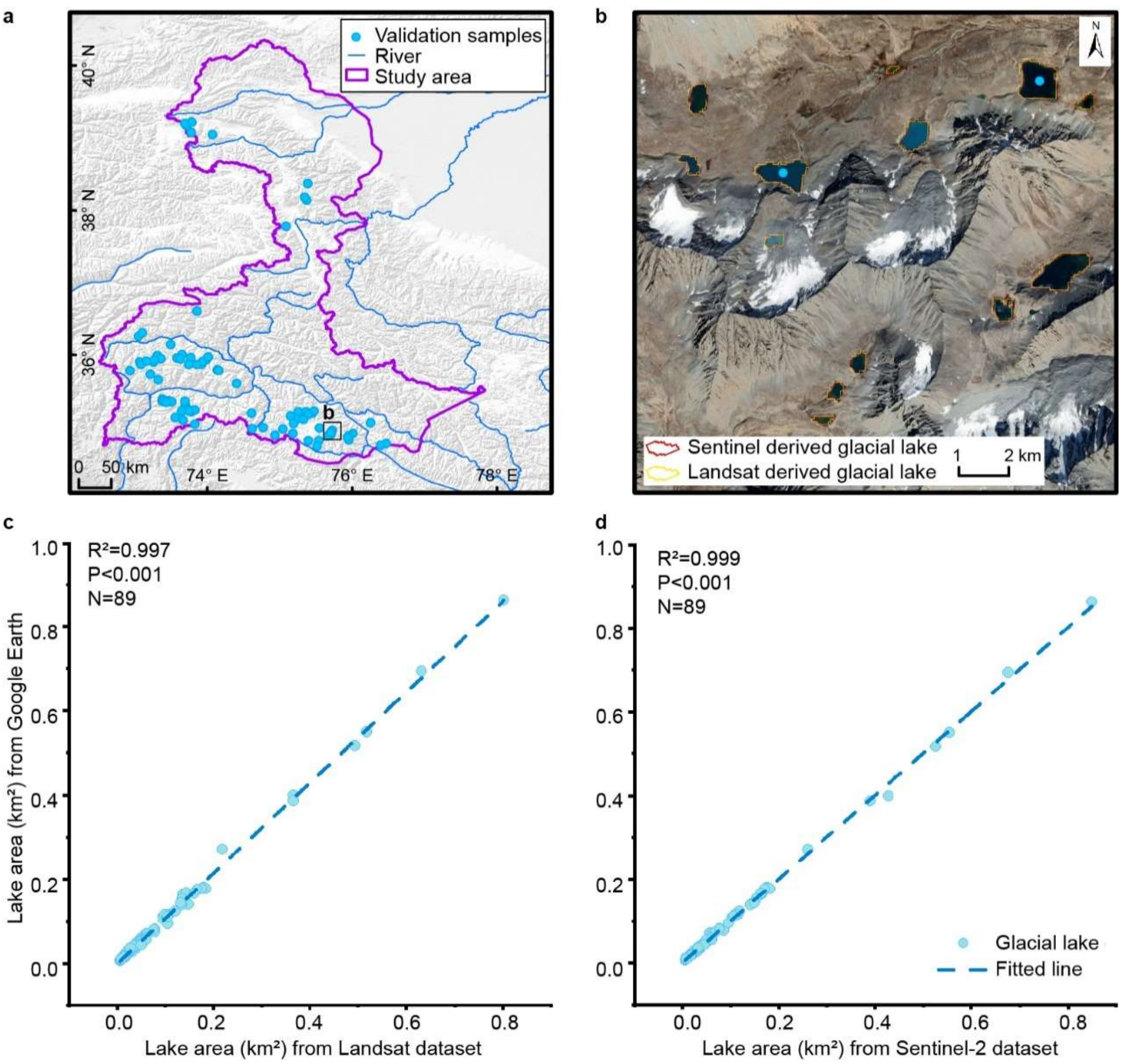


Figure 9. Distribution of validation samples (a), comparison of glacial lakes derived from Landsat and Sentinel-2 overlaying Google Earth image (© Google Earth 2019) in a zoomed site (b), and glacial lake product validated by Google Earth derived lake boundaries (c and d).

6.2 Comparison of Sentinel-2 and Landsat derived products

Glacial lakes from Landsat and Sentinel-2 images have a high consistency in number and area with overlap rates from approximately 86% to 94% for all lakes greater than 0.0045 km^2 (Table 4), indicating a good potential for coordinated utility with Landsat archived observation (Figure 10). Lake extents extracted from Landsat and Sentinel images match well for various types and sizes (Figure 10 and Figure 11, Table 4). The best consistency rate reaches 94% for the glacial lakes between 0.1 km^2 and 0.2 km^2. The difference in area of glacial lakes extracted from Landsat and Sentinel-2 images generally lies within the uncertainty ranges.

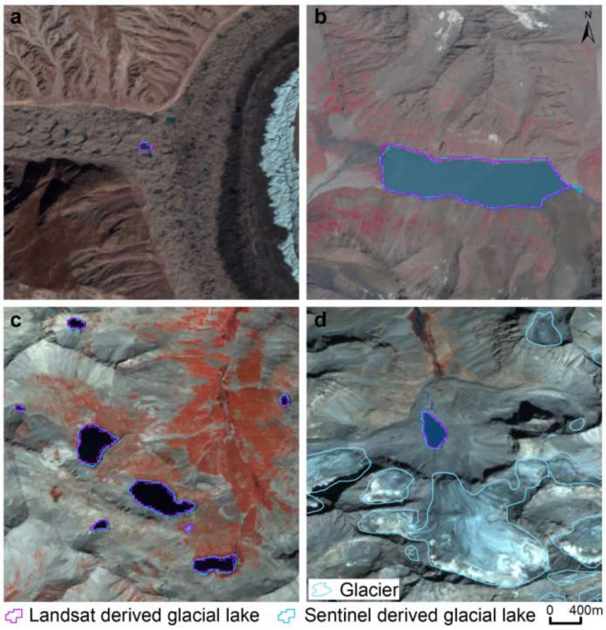


Figure 10. High consistency of lake extents extracted from Landsat and Sentinel-2 images. Lake types shown include supraglacial (a), glacier-fed moraine-dammed (b), unconnected glacial-erosion lake without glacier melt supply (c) and glacier-fed moraine-dammed lakes (d).

Spatial resolution of satellite images plays a primary role in the discrepancies in count and area of glacial lakes extracted from Landsat (30 m) and Sentinel-2 (10 m) observations. Due to a finer spatial resolution, Sentinel-2 images can extract more glacial lakes and more accurate extents than those from Landsat images. We set the same 5 pixels as the MMU for both Landsat and Sentinel-2 images, which corresponds to a minimum area of 0.0045 km^2 and 0.0005 km^2, respectively. The minimum mapping area results in generating nearly 5000 more lakes from Sentinel-2 images than from Landsat images, causing the greatest discrepancy in number (Table 4), such as Figure 11a. Small lakes such as supraglacial lakes play an important role in analyzing glacier evolution and supraglacial drainage systems (Liu and Mayer, 2015; Miles et al., 2018), implying a potential of our dataset to be applied in studies of glacier-lake evolutions. Meanwhile, Sentinel-2 images are able to depict boundaries of glacial lake with a lower uncertainty, as for some small islands and narrow channels (Figure 11b and c) were mapped from Sentinel-2 imagery that were unable to be detected in Landsat imagery.

Different acquisition dates between Sentinel-2 and Landsat images also contribute to the discrepancy of those two glacial lake data. Acquiring same-day images from the two sensors were not always possible due to the impacts of cloud contaminations, topographic shadows, snow cover and revisit periods (Williamson et al., 2018; Paul et al., 2020). Glacial lakes are changing temporally in the context of climate and glacier changes, taking supraglacial lakes for example that evolve dramatically in a short period observed between Landsat and Sentinel-2 images (Figure 11d). Despite our efforts of leveraging all available high-quality images, the overlap of acquisition dates between Landsat and Sentinel-2 images for the same location is relatively low (only 7 scenes of Sentinel-2 images or 112 glacial lakes in 2020) in this study area, and the consequential temporal gaps led to a difference in the number and area of the derived glacial lakes.

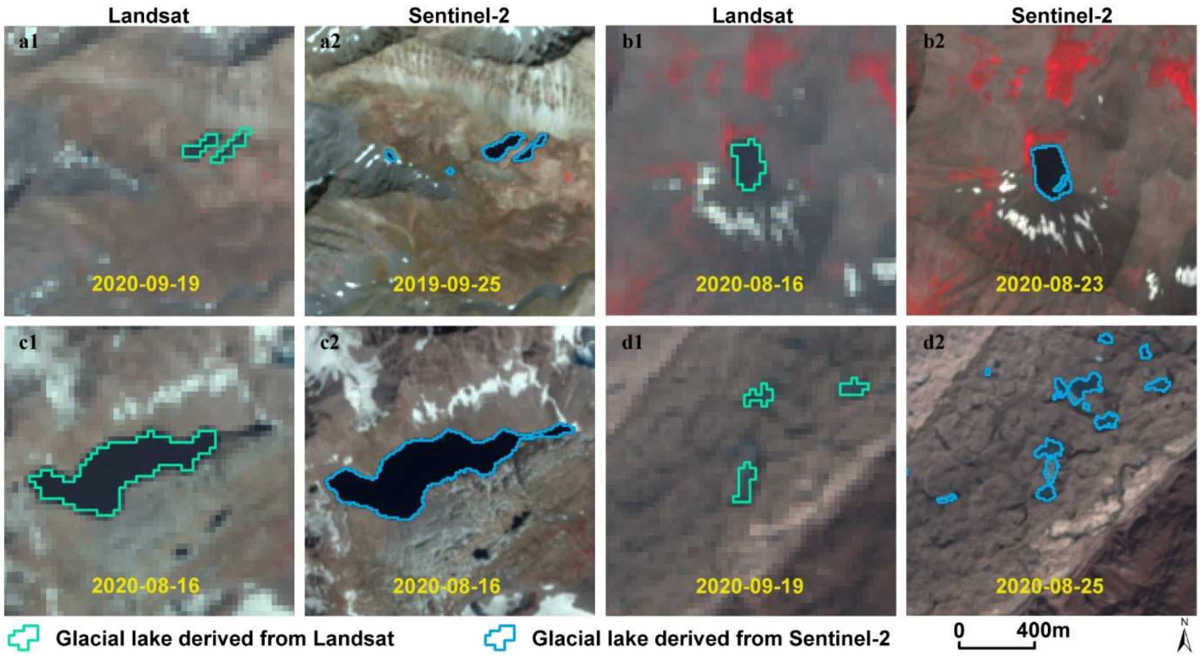


Figure 11. Discrepancy of lake extents extracted from Landsat and Sentinel-2 images.

6.3 Limitation and updating plan

We would like to acknowledge several limitations of our glacier lake dataset, largely due the availability of high quality satellite images in the study area and inadequate field survey data (Wang et al., 2020; Chen et al., 2021). First, it is unlikely to collect enough good-quality images within one calendar year for the entire study area due to high possibility of cloud or snow covers. Even though the capacity of repeat observations for Landsat-8 OLI and Sentinel-2 increased (Roy et al., 2014; Williamson et al., 2018; Wulder et al., 2019; Paul et al., 2020), the 2020 glacial lake dataset has to employ images acquired in adjacent years besides 2020. Most images used from Landsat and Sentinel-2 platforms were imaged in autumn, and some images taken between April and July and in November also were employed. Distribution and changes in glacial lakes primarily represent the characteristics between August and October. Glacial lakes evolve with time and space (Nie et al., 2017), and subtle inter- and intra-annual changes (Liu et al., 2020) for each time period were ignored. Second, field investigation data are limited due to low accessibility of high mountain environment in the study area, which restrained the accuracy in classifying the glacial lake types. Although very high-resolution Google Earth images were utilized to assist in lake type interpretation, occasional misclassification was unavoidable. We implemented two types of classification systems based on a careful utilization of glacier data, DEM, geomorphological features and expert knowledge. However, the lack of in situ survey prohibited a thorough validation of the glacial lake types. Third, the rigorous quality assurance and cross check after semi-automated lake mapping assure the quality of our lake dataset but are still time and cost prohibitive. State-of-the-art mapping methods, such as deep learning method (Wu et al., 2020), Google Earth Engine cloud-computing (Chen et al., 2021) and synergy of SAR and optical images (Wangchuk and Bolch, 2020; How et al., 2021), would be used in the future to balance product accuracy and time cost.

The glacial lake dataset will be updated using newly collected Landsat and Sentinel images at a five-year interval or modified according to user feedbacks. The updated glacial lake dataset will continue to be released freely and publicly on the Mountain Science Data Center sharing platform.

7 Data availability

Our glacial lake dataset extracted from Sentinel-2 images in 2020 and Landsat observation between 1990 and 2020 are available online via the Mountain Science Data Center, the Institute of Mountain Hazards and Environment, the Chinese Academy of Sciences at https://doi.org/10.12380/Glaci.msdc.000001 (Lesi et al., 2022). The glacial lake dataset is provided in both ESRI shapefile format (total size of 22.6 MB) and the Geopackage format (version 1.2.1) with a total size of 9.2MB, which can be opened and further processed by open-source geographic information system software such as QGIS.

8 Conclusions

Glacial lake inventories of the entire China-Pakistan Economic Corridor in 2020 were provided based on Landsat and Sentinel-2 images using a threshold-based semi-automated mapping method. Both Landsat and Sentinel-2 derived glacial lake dataset show similar characteristics in spatial distribution and in the statistics of count and area. By contrast, glacial lake dataset derived from Sentinel-2 images with a spatial resolution of 10 m has a lower mapping error and more accurate lake boundary than those from 30 m spatial resolution Landsat images whereas Landsat imagery is more suitable to analyze spatial-temporal changes at a longer time scale due to its long-term archived observations at a consistent 30 m spatial resolution starting from the late 1980s.

Glacial lakes in the study area remain relatively stable with a slight increase in number and area between 1990 and 2020 according to Landsat observations. Our dataset reveals that 2154 glacial lakes in 1990 covering 85.1 ± 14.66 km^2 increased to 2234 lakes with a total area of 86.31 ± 14.98 km^2. The same mapping method and rigorous workflow of quality assurance and quality control used in this study reduced the error in multi-temporal changes of glacial lakes.

The Hanshaw's error estimation method for pixel-based lake mapping was improved by removing repeatedly calculated edge pixels that vary with lake shape. Therefore, the newly proposed method reduces the estimated value of uncertainty from satellite observations. The average relative error is ±17.36% for Landsat-derived product and ±8.15% for product from Sentinel-2.

Our glacial lake dataset contains a range of critical parameters that maximize their potential utility for water resource and GLOFs risk evaluation, cryosphere-hydrological and glacier-lake evolution projection. The dual classification systems of glacial lake types were developed and are very likely to attract broader researchers and scientists to use our datasets. In comparison with other existing glacial lake datasets, our products were created through a thorough consideration of lake types, cross checks and rigorous quality assurance, and will be updated and released continuously in the Mountain Science Data Center. As such, we expect that our glacial lake dataset will have significant value to cryospheric-hydrology research, the assessment of water resource and glacier-related hazards in the CPEC.

Appendix. The appendix related to this article is available online.

Author contributions. ML and YN conceived the study, ML, YN and XD performed data processing and analysis of the glacial lake inventory data, JW contributed to tool development and mapping methods, ML and YN wrote the manuscript. All authors reviewed and edited the manuscript before submission.

Competing interests. The authors declare no conflict of interest.

Acknowledgements.
We are grateful for the editor Kenneth Mankoff and three anonymous referees for their constructive comments that greatly help us to improve this manuscript. This study was supported by the second Tibetan Plateau Scientific Expedition and Research Program (grant 2019QZKK0603), the National Natural Science Foundation of China (Grant Nos. 42171086, 41971153), the International Science & Technology Cooperation Program of China (No. 2018YFE0100100), the Chinese Academy of Sciences "Light of West China" and Natural Sciences and Engineering Research Council of Canada (Grant No. DG-2020-04207).

References

Ashraf, A., Naz, R., Iqbal, M.B.: Altitudinal dynamics of glacial lakes under changing climate in the Hindu Kush, Karakoram, and Himalaya ranges. Geomorphology, 283: 72-79, https://doi.org/10.1016/j.geomorph.2017.01.033, 2017.

Azam, M.F., Kargel, J.S., Shea, J.M., Nepal, S., Haritashya, U.K., Srivastava, S., Maussion, F., Qazi, N., Chevallier, P., Dimri, A.P., Kulkarni, A.V., Cogley, J.G., Bahuguna, I.: Glaciohydrology of the Himalaya-Karakoram. Science, 373: eabf3668, https://doi.org/10.1126/science.abf3668, 2021.

Battamo, A.Y., Varis, O., Sun, P., Yang, Y., Oba, B.T., Zhao, L.: Mapping socio-ecological resilience along the seven economic corridors of the Belt and Road Initiative. J. Clean. Prod., 309: 127341, https://doi.org/10.1016/j.jclepro.2021.127341, 2021.

Bhambri, R., Hewitt, K., Kawishwar, P., Kumar, A., Verma, A., Snehmani, Tiwari, S., Misra, A.: Ice-dams, outburst floods, and movement heterogeneity of glaciers, Karakoram. Global Planet. Change, 180: 100-116, https://doi.org/10.1016/j.gloplacha.2019.05.004, 2019.

Bhattacharya, A., Bolch, T., Mukherjee, K., King, O., Menounos, B., Kapitsa, V., Neckel, N., Yang, W., Yao, T.: High Mountain Asian glacier response to climate revealed by multi-temporal satellite observations since the 1960s. Nat. Commun., 12: 4133, https://doi.org/10.1038/s41467-021-24180-y, 2021.

Bolch, T., Pieczonka, T., Mukherjee, K., Shea, J.: Brief communication: Glaciers in the Hunza catchment (Karakoram) have been nearly in balance since the 1970s. The Cryosphere, 11: 531-539, https://doi.org/10.5194/tc-11-531-2017, 2017.

Brun, F., Berthier, E., Wagnon, P., Kääb, A., Treichler, D.: A spatially resolved estimate of High Mountain Asia glacier mass balances from 2000 to 2016. Nat. Geosci., 10: 668-673, https://doi.org/10.1038/ngeo2999, 2017.

Brun, F., Wagnon, P., Berthier, E., Jomelli, V., Maharjan, S.B., Shrestha, F., Kraaijenbrink, P.D.A.: Heterogeneous Influence of Glacier Morphology on the Mass Balance Variability in High Mountain Asia. J. Geophys. Res-Earth, 124: 1331-1345, https://doi.org/10.1029/2018JF004838, 2019.

Carrivick, J.L., Tweed, F.S.: Proglacial lakes: character, behaviour and geological importance. Quaternary Sci. Rev., 78: 34-52, https://doi.org/10.1016/j.quascirev.2013.07.028, 2013.

Carrivick, J.L., Quincey, D.J.: Progressive increase in number and volume of ice-marginal lakes on the western margin of the Greenland Ice Sheet. Global Planet. Change, 116: 156-163, https://doi.org/10.1016/j.gloplacha.2014.02.009, 2014.

Carrivick, J.L., Tweed, F.S.: A global assessment of the societal impacts of glacier outburst floods. Global Planet. Change, 144: 1-16, https://doi.org/10.1016/j.gloplacha.2016.07.001, 2016.

Carrivick, J.L., Tweed, F.S., Sutherland, J.L., Mallalieu, J.: Toward Numerical Modeling of Interactions Between Ice-Marginal Proglacial Lakes and Glaciers. Front. Earth Sci, 8, https://doi.org/10.3389/feart.2020.577068, 2020.

Carrivick, J.L., How, P., Lea, J.M., Sutherland, J.L., Grimes, M., Tweed, F.S., Cornford, S., Quincey, D.J., Mallalieu, J.: Ice-Marginal Proglacial Lakes Across Greenland: Present Status and a Possible Future. Geophys. Res. Lett., 49: e2022GL099276, https://doi.org/https://doi.org/10.1029/2022GL099276, 2022.

Chen, F., Zhang, M., Guo, H., Allen, S., Kargel, J.S., Haritashya, U.K., Watson, C.S.: Annual 30 m dataset for glacial lakes in High Mountain Asia from 2008 to 2017. Earth System Science Data, 13: 741-766, https://doi.org/10.5194/essd-13-741-2021, 2021.

Chen, X., Cui, P., You, Y., Cheng, Z., Khan, A., Ye, C., Zhang, S.: Dam-break risk analysis of the Attabad landslide dam in Pakistan and emergency countermeasures. Landslides, 14: 675-683, https://doi.org/10.1007/s10346-016-0721-7, 2017.

Cook, S.J., Quincey, D.J.: Estimating the volume of Alpine glacial lakes. Earth Surf. Dynam., 3: 559-575, https://doi.org/10.5194/esurf-3-559-2015, 2015.

Emmer, A., Cuřín, V.: Can a dam type of an alpine lake be derived from lake geometry? A negative result. J. Mt. Sci.-Engl., 18: 614-621, https://doi.org/10.1007/s11629-020-6003-9, 2021.

Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank, D., Alsdorf, D.: The Shuttle Radar Topography Mission. Rev. Geophys., 45: RG2004, https://doi.org/10.1029/2005RG000183, 2007.

Gardelle, J., Arnaud, Y., Berthier, E.: Contrasted evolution of glacial lakes along the Hindu Kush Himalaya mountain range between 1990 and 2009. Global Planet. Change, 75: 47-55, https://doi.org/10.1016/j.gloplacha.2010.10.003, 2011.

Hanshaw, M.N., Bookhagen, B.: 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: 359-

376, https://doi.org/10.5194/tc-8-359-2014, 2014.

Hewitt, K.: The Karakoram Anomaly? Glacier Expansion and the ‘Elevation Effect,’ Karakoram Himalaya. Mt. Res. Dev., 25: 332-340, https://doi.org/10.1659/0276-4741(2005)025[0332:TKAGEA]2.0.CO;2, 2005.

Hewitt, K., 2014. Glaciers of the Karakoram Himalaya： Glacial Environments, Processes, Hazards and Resources. Springer, Dordrecht.

How, P., Messerli, A., Mätzler, E., Santoro, M., Wiesmann, A., Caduff, R., Langley, K., Bojesen, M.H., Paul, F., Kääb, A., Carrivick, J.L.: Greenland-wide inventory of ice marginal lakes using a multi-method approach. Sci. Rep.-UK, 11: 4481, https://doi.org/10.1038/s41598-021-83509-1, 2021.

Huggel, C., Kääb, A., Haeberli, W., Teysseire, P., Paul, F.: Remote sensing based assessment of hazards from glacier lake outbursts: a case study in the Swiss Alps. Can. Geotech. J., 39: 316-330, https://doi.org/10.1139/t01-099, 2002.

Hugonnet, R., Mcnabb, R., Berthier, E., Menounos, B., Nuth, C., Girod, L., Farinotti, D., Huss, M., Dussaillant, I., Brun, F., Kääb, A.: Accelerated global glacier mass loss in the early twenty-first century. Nature, 592: 726-731, https://doi.org/10.1038/s41586-021-03436-z, 2021.

Huss, M., Hock, R.: Global-scale hydrological response to future glacier mass loss. Nat. Clim. Change, 8: 135-140, https://doi.org/10.1038/s41558-017-0049-x, 2018.

Immerzeel, W.W., Lutz, A.F., Andrade, M., Bahl, A., Biemans, H., Bolch, T., Hyde, S., Brumby, S., Davies, B.J., Elmore, A.C., Emmer, A., Feng, M., Fernández, A., Haritashya, U., Kargel, J.S., Koppes, M., Kraaijenbrink, P.D.A., Kulkarni, A.V., Mayewski, P.A., Nepal, S., Pacheco, P., Painter, T.H., Pellicciotti, F., Rajaram, H., Rupper, S., Sinisalo, A., Shrestha, A.B., Viviroli, D., Wada, Y., Xiao, C., Yao, T., Baillie, J.E.M.: Importance and vulnerability of the world’s water towers. Nature, 577: 364-369, https://doi.org/10.1038/s41586-019-1822-y, 2020.

Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E., 2008. Hole-filled seamless SRTM data V4. 2008, International Centre for Tropical Agriculture (CIAT), available from http://srtm.csi.cgiar.org.

Jiang, S., Nie, Y., Liu, Q., Wang, J., Liu, L., Hassan, J., Liu, X., Xu, X.: Glacier Change, Supraglacial Debris Expansion and Glacial Lake Evolution in the Gyirong River Basin, Central Himalayas, between 1988 and 2015. Remote Sens.-Basel, 10: 986, https://doi.org/10.3390/rs10070986, 2018.

Kääb, A., Berthier, E., Nuth, C., Gardelle, J., Arnaud, Y.: Contrasting patterns of early twenty-first-century glacier mass change in the Himalayas. Nature, 488: 495-498, https://doi.org/10.1038/nature11324, 2012.

Lesi, M., Nie, Y., Shugar, D.H., Wang, J., Deng, Q., Chen, H.: 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 CSTR:1a006.11.Glaci.msdc.000001, 2022.

Li, D., Shangguan, D., Anjum, M.N.: Glacial Lake Inventory Derived from Landsat 8 OLI in 2016–2018 in China–Pakistan Economic Corridor. ISPRS international journal of geo-information, 9: 294, https://doi.org/10.3390/ijgi9050294, 2020.

Li, Z., Deng, X., Zhang, Y.: Evaluation and convergence analysis of socio-economic vulnerability to natural hazards of Belt and Road Initiative countries. J. Clean. Prod., 282: 125406, https://doi.org/10.1016/j.jclepro.2020.125406, 2021.

Liu, Q., Mayer, C.: Distribution and interannual variability of supraglacial lakes on debris-covered glaciers in the Khan Tengri-Tumor Mountains, Central Asia. Environ. Res. Lett., 10: 014014 2015.

Liu, Q., Mayer, C., Wang, X., Nie, Y., Wu, K., Wei, J., Liu, S.: Interannual flow dynamics driven by frontal retreat of a lake-terminating glacier in the Chinese Central Himalaya. Earth Planet. Sc. Lett., 546: 116450, https://doi.org/10.1016/j.epsl.2020.116450, 2020.

Lyons, E.A., Sheng, Y., Smith, L.C., Li, J., Hinkel, K.M., Lenters, J.D., Wang, J.: Quantifying sources of error in multitemporal multisensor lake mapping. Int. J. Remote Sens., 34: 7887-7905, https://doi.org/10.1080/01431161.2013.827343, 2013.

Mankoff, K.D., Noël, B., Fettweis, X., Ahlstrøm, A.P., Colgan, W., Kondo, K., Langley, K., Sugiyama, S., van As, D., Fausto, R.S.: Greenland liquid water discharge from 1958 through 2019. Earth Syst. Sci. Data, 12: 2811-2841, https://doi.org/10.5194/essd-12-2811-2020, 2020.

Martín, C.N.S., Ponce, J.F., Montes, A., Balocchi, L.D., Gorza, C., Andrea, C.: Proglacial landform assemblage in a rapidly retreating cirque glacier due to temperature increase since 1970, Fuegian Andes, Argentina. Geomorphology, 390: 107861, https://doi.org/10.1016/j.geomorph.2021.107861, 2021.

Maurer, J.M., Schaefer, J.M., Rupper, S., Corley, A.: Acceleration of ice loss across the Himalayas over the past 40 years. Sci. Adv., 5: eaav7266, https://doi.org/10.1126/sciadv.aav7266, 2019.

Mcfeeters, S.K.: The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int. J. Remote Sens., 17: 1425 - 1432 1996.

Miles, E.S., Watson, C.S., Brun, F., Berthier, E., Esteves, M., Quincey, D.J., Miles, K.E., Hubbard, B., Wagnon, P.: Glacial and geomorphic effects of a supraglacial lake drainage and outburst event, Everest region, Nepal Himalaya. The Cryosphere, 12: 3891-3905, https://doi.org/10.5194/tc-12-3891-2018, 2018.

Nie, Y., Zhang, Y., Liu, L., Zhang, J.: Glacial change in the vicinity of Mt. Qomolangma (Everest), central high Himalayas since 1976. J. Geogr. Sci., 20: 667-686, https://doi.org/10.1007/s11442-010-0803-8, 2010.

Nie, Y., Sheng, Y., Liu, Q., Liu, L., Liu, S., Zhang, Y., Song, C.: A regional-scale assessment of Himalayan glacial lake changes using satellite observations from 1990 to 2015. Remote Sens. Environ., 189: 1-13, https://doi.org/10.1016/j.rse.2016.11.008, 2017.

Nie, Y., Liu, Q., Wang, J., Zhang, Y., Sheng, Y., Liu, S.: An inventory of historical glacial lake outburst floods in the Himalayas based on remote sensing observations and geomorphological analysis. Geomorphology, 308:

91-106, https://doi.org/10.1016/j.geomorph.2018.02.002, 2018.

Nie, Y., Liu, W., Liu, Q., Hu, X., Westoby, M.J.: Reconstructing the Chongbaxia Tsho glacial lake outburst flood in the Eastern Himalaya: Evolution, process and impacts. Geomorphology, 370: 107393, https://doi.org/10.1016/j.geomorph.2020.107393, 2020.

Nie, Y., Pritchard, H.D., Liu, Q., Hennig, T., Wang, W., Wang, X., Liu, S., Nepal, S., Samyn, D., Hewitt, K., Chen, X.: Glacial change and hydrological implications in the Himalaya and Karakoram. Nature Reviews Earth & Environment, 2: 91-106, https://doi.org/10.1038/s43017-020-00124-w, 2021.

Paul, F., Rastner, P., Azzoni, R.S., Diolaiuti, G., Fugazza, D., Le Bris, R., Nemec, J., Rabatel, A., Ramusovic, M., Schwaizer, G., Smiraglia, C.: Glacier shrinkage in the Alps continues unabated as revealed by a new glacier inventory from Sentinel-2. Earth System Science Data, 12: 1805-1821, https://doi.org/10.5194/essd-12-1805-2020, 2020.

Pfeffer, W.T., Arendt, A.A., Bliss, A., Bolch, T., Cogley, J.G., Gardner, A.S., Hagen, J., Hock, R., Kaser, G., Kienholz, C., Miles, E.S., Moholdt, G., Mölg, N., Paul, F., Radić, V., Rastner, P., Raup, B.H., Rich, J., Sharp, M.J.: The Randolph Glacier Inventory: a globally complete inventory of glaciers. J. Glaciol., 60: 537-552, https://doi.org/10.3189/2014JoG13J176, 2014.

Post, A., Mayo, L.R., 1971. Glacier dammed lakes and outburst floods in Alaska: U.S. Geological Survey Hydrologic Investigations Atlas 455, U.S. Geological Survey.

Pritchard, H.D.: Asia’s shrinking glaciers protect large populations from drought stress. Nature, 569: 649-654, https://doi.org/10.1038/s41586-019-1240-1, 2019.

Quincey, D.J., Richardson, S.D., Luckman, A., Lucas, R.M., Reynolds, J.M., Hambrey, M.J., Glasser, N.F.: Early recognition of glacial lake hazards in the Himalaya using remote sensing datasets. Global Planet. Change, 56: 137-152, https://doi.org/10.1016/j.gloplacha.2006.07.013, 2007.

Rabus, B., Eineder, M., Roth, A., Bamler, R.: The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. ISPRS J. Photogramm., 57: 241-262, https://doi.org/10.1016/S0924-2716(02)00124-7, 2003.

RGI Consortium: Randolph Glacier Inventory – A Dataset of Global Glacier Outlines: Version 6.0: Technical Report, https://doi.org/10.7265/N5-RGI-60, 2017.

Rick, B., Mcgrath, D., Armstrong, W., Mccoy, S.W.: Dam type and lake location characterize ice-marginal lake area change in Alaska and NW Canada between 1984 and 2019. The Cryosphere, 16: 297-314, https://doi.org/10.5194/tc-16-297-2022, 2022.

Rose, A., Mckee, J., Sims, K., Bright, E., Reith, A., Urban, M.: LandScan Global 2020, https://doi.org/https://doi.org/10.48690/1523378, 2021.

Rounce, D.R., Hock, R., Shean, D.E.: Glacier Mass Change in High Mountain Asia Through 2100 Using the Open-Source Python Glacier Evolution Model (PyGEM). Front. Earth Sci, 7: 331, https://doi.org/10.3389/feart.2019.00331, 2020.

Roy, D.P., Wulder, M.A., Loveland, T.R., C. E., W., Allen, R.G., Anderson, M.C., Helder, D., Irons, J.R., Johnson, D.M., Kennedy, R., Scambos, T.A., Schaaf, C.B., Schott, J.R., Sheng, Y., Vermote, E.F., Belward, A.S., Bindschadler, R., Cohen, W.B., Gao, F., Hipple, J.D., Hostert, P., Huntington, J., Justice, C.O., Kilic, A., Kovalskyy, V., Lee, Z.P., Lymburner, L., Masek, J.G., Mccorkel, J., Shuai, Y., Trezza, R., Vogelmann, J., Wynne, R.H., Zhu, Z.: Landsat-8: Science and product vision for terrestrial global change research. Remote Sens. Environ., 145: 154-172, https://doi.org/10.1016/j.rse.2014.02.001, 2014.

Sakai, A.: Brief communication: Updated GAMDAM glacier inventory over high-mountain Asia. The Cryosphere, 13: 2043-2049, https://doi.org/10.5194/tc-13-2043-2019, 2019.

Salerno, F., Thakuri, S., D'Agata, C., Smiraglia, C., Manfredi, E.C., Viviano, G., Tartari, G.: Glacial lake distribution in the Mount Everest region: Uncertainty of measurement and conditions of formation. Global Planet. Change, 92-93: 30-39 2012.

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

Sheng, Y., Song, C., Wang, J., Lyons, E.A., Knox, B.R., Cox, J.S., Gao, F.: Representative lake water extent mapping at continental scales using multi-temporal Landsat-8 imagery. Remote Sens. Environ., 185: 129-141, https://doi.org/10.1016/j.rse.2015.12.041, 2016.

Shugar, D.H., Burr, A., Haritashya, U.K., Kargel, J.S., Watson, C.S., Kennedy, M.C., Bevington, A.R., Betts, R.A., Harrison, S., Strattman, K.: Rapid worldwide growth of glacial lakes since 1990. Nat. Clim. Change, 10: 939-945, https://doi.org/10.1038/s41558-020-0855-4, 2020.

Shugar, D.H., Jacquemart, M., Shean, D., Bhushan, S., Upadhyay, K., Sattar, A., Schwanghart, W., Mcbride, S., de Vries, M., Mergili, M., Emmer, A., Deschamps-Berger, C., Mcdonnell, M., Bhambri, R., Allen, S., Berthier, E., Carrivick, J.L., Clague, J.J., Dokukin, M., Dunning, S.A., Frey, H., Gascoin, S., Haritashya, U.K., Huggel, C., Kaab, A., Kargel, J.S., Kavanaugh, J.L., Lacroix, P., Petley, D., Rupper, S., Azam, M.F., Cook, S.J., Dimri, A.P., Eriksson, M., Farinotti, D., Fiddes, J., Gnyawali, K.R., Harrison, S., Jha, M., Koppes, M., Kumar, A., Leinss, S., Majeed, U., Mal, S., Muhuri, A., Noetzli, J., Paul, F., Rashid, I., Sain, K., Steiner, J., Ugalde, F., Watson, C.S., Westoby, M.J.: A massive rock and ice avalanche caused the 2021 disaster at Chamoli, Indian Himalaya. Science, 373: 300-306, https://doi.org/10.1126/science.abh4455, 2021.

Ullah, S., You, Q., Ali, A., Ullah, W., Jan, M.A., Zhang, Y., Xie, W., Xie, X.: Observed changes in maximum and minimum temperatures over China- Pakistan economic corridor during 1980–2016. Atmos. Res., 216: 37-51, https://doi.org/10.1016/j.atmosres.2018.09.020, 2019.

Viviroli, D., Kummu, M., Meybeck, M., Kallio, M., Wada, Y.: Increasing dependence of lowland populations on mountain water resources. Nature Sustainability, 3: 917-928, https://doi.org/10.1038/s41893-020-0559-9, 2020.

Wang, J., Sheng, Y., Tong, T.S.D.: Monitoring decadal lake dynamics across the Yangtze Basin downstream of Three Gorges Dam. Remote Sens. Environ., 152: 251-269, https://doi.org/10.1016/j.rse.2014.06.004, 2014.

Wang, J., Sheng, Y., Wada, Y.: Little impact of the Three Gorges Dam on recent decadal lake decline across China's Yangtze Plain. Water Resour. Res., 53: 3854-3877, https://doi.org/10.1002/2016WR019817, 2017.

Wang, J., Song, C., Reager, J.T., Yao, F., Famiglietti, J.S., Sheng, Y., Macdonald, G.M., Brun, F., Schmied, H.M., Marston, R.A., Wada, Y.: Recent global decline in endorheic basin water storages. Nat. Geosci., 11: 926-932, https://doi.org/10.1038/s41561-018-0265-7, 2018.

Wang, X., Ding, Y., Liu, S., Jiang, L., Wu, K., Jiang, Z., Guo, W.: Changes of glacial lakes and implications in Tian Shan, Central Asia, based on remote sensing data from 1990 to 2010. Environ. Res. Lett., 8: 44052, https://doi.org/10.1088/1748-9326/8/4/044052, 2013.

Wang, X., Liu, S., Zhang, J.: A new look at roles of the cryosphere in sustainable development. Advances in Climate Change Research, 10: 124-131, https://doi.org/10.1016/j.accre.2019.06.005, 2019.

Wang, X., Guo, X., Yang, C., Liu, Q., Wei, J., Zhang, Y., Liu, S., Zhang, Y., Jiang, Z., Tang, Z.: Glacial lake inventory of high-mountain Asia in 1990 and 2018 derived from Landsat images. Earth System Science Data, 12: 2169-2182, https://doi.org/10.5194/essd-12-2169-2020, 2020.

Wangchuk, S., Bolch, T.: Mapping of glacial lakes using Sentinel-1 and Sentinel-2 data and a random forest classifier: Strengths and challenges. Science of Remote Sensing, 2: 100008, https://doi.org/https://doi.org/10.1016/j.srs.2020.100008, 2020.

Westoby, M.J., Glasser, N.F., Brasington, J., Hambrey, M.J., Quincey, D.J., Reynolds, J.M.: Modelling outburst floods from moraine-dammed glacial lakes. Earth-Sci. Rev., 134: 137-159, https://doi.org/10.1016/j.earscirev.2014.03.009, 2014.

Williamson, A.G., Banwell, A.F., Willis, I.C., Arnold, N.S.: Dual-satellite (Sentinel-2 and Landsat 8) remote sensing of supraglacial lakes in Greenland. The Cryosphere, 12: 3045-3065, https://doi.org/10.5194/tc-12-3045-2018, 2018.

Wu, R., Liu, G., Zhang, R., Wang, X., Li, Y., Zhang, B., Cai, J., Xiang, W.: A Deep Learning Method for Mapping Glacial Lakes from the Combined Use of Synthetic-Aperture Radar and Optical Satellite Images. Remote Sens.-Basel, 12: 4020 2020.

Wulder, M.A., Loveland, T.R., Roy, D.P., Crawford, C.J., Masek, J.G., Woodcock, C.E., Allen, R.G., Anderson, M.C., Belward, A.S., Cohen, W.B., Dwyer, J., Erb, A., Gao, F., Griffiths, P., Helder, D., Hermosilla, T., Hipple, J.D., Hostert, P., Hughes, M.J., Huntington, J., Johnson, D.M., Kennedy, R., Kilic, A., Li, Z., Lymburner, L., Mccorkel, J., Pahlevan, N., Scambos, T.A., Schaaf, C., Schott, J.R., Sheng, Y., Storey, J., Vermote, E., Vogelmann, J., White, J.C., Wynne, R.H., Zhu, Z.: Current status of Landsat program, science, and applications. Remote Sens. Environ., 225: 127-147, https://doi.org/https://doi.org/10.1016/j.rse.2019.02.015, 2019.

Yao, C., Wang, X., Zhao, X., Wei, J., Zhang, Y.: Temporal and Spatial Changes of Glacial Lakes in the China-Pakistan Economic Corridor from 1990 to 2018. Journal of Glaciology and Geocryology, 42: 33-42, https://doi.org/https://doi.org/10.7522/j.issn.1000-0240.2020.0009, 2020.

Yao, T., Thompson, L., Yang, W., Yu, W.S., Gao, Y., Guo, X.J., Yang, X.X., Duan, K.Q., Zhao, H.B., Xu, B.Q., Pu, J.C., Lu, A.X., Xiang, Y., Kattel, D.B., Joswiak, D.: Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. Nat. Clim. Change, 2: 663-667, https://doi.org/10.1038/NCLIMATE1580, 2012.

Yao, X., Liu, S., Han, L., Sun, M., Zhao, L.: Definition and classification system of glacial lake for inventory and hazards study. J. Geogr. Sci., 28: 193-205, https://doi.org/10.1007/s11442-018-1467-z, 2018.

Zhang, G., Yao, T., Xie, H., Wang, W., Yang, W.: An inventory of glacial lakes in the Third Pole region and their changes in response to global warming. Global Planet. Change, 131: 148-157, https://doi.org/10.1016/j.gloplacha.2015.05.013, 2015.

Zhang, M., Chen, F., Tian, B.: An automated method for glacial lake mapping in High Mountain Asia using Landsat 8 imagery. J. Mt. Sci.-Engl., 15: 13-24, https://doi.org/10.1007/s11629-017-4518-5, 2018.

Zhao, W., Xiong, D., Wen, F., Wang, X.: Lake area monitoring based on land surface temperature in the Tibetan Plateau from 2000 to 2018. Environ. Res. Lett., 15, https://doi.org/10.1088/1748-9326/ab9b41, 2020.

Zheng, G., Allen, S.K., Bao, A., Ballesteros-Cánovas, J.A., Huss, M., Zhang, G., Li, J., Yuan, Y., Jiang, L., Yu, T., Chen, W., Stoffel, M.: Increasing risk of glacial lake outburst floods from future Third Pole deglaciation. Nat. Clim. Change, 11: 411-417, https://doi.org/10.1038/s41558-021-01028-3, 2021.

Appendix

Tutorial for Improved Uncertainty Estimating Method

The Hanshaw's equation was originally proposed for pixelated polygons (such as a polygon directly extracted from a remote sensing image), and performed more robustly than manually digitized polygons (where vertices do not necessarily follow the pixel edges). Our improved method also performs better for pixelated polygons. This tutorial is dedicated to helping implement our improved uncertainty estimation method.

Procedure of uncertainty estimating method (using ArcGIS (© ESRI) **for example)**

1. Removing redundant nodes (optional)

We found that a small proportion (~1%) of the pixelated lake polygons (directly extracted from satellite images) have redundant nodes, which affects the value of inner nodes. If no redundant nodes exist, this step can be skipped. Or, we recommend using the "Simplify Polygon" tool in ArcGIS to remove those nodes (Figure A1).

In the Simplify Polygon panel

- Input your dataset.
- Set the output path and output file name.
- Choose the simplification algorithm. We recommended "POINT_REMOVE".
- Set the tolerance of simplification algorithm. In this step, we need to ensure that the polygon boundaries remain unchanged after deleting redundant nodes. Generally, a tolerance of 1 meter will suffice, or you can adjust the threshold until your satisfaction.

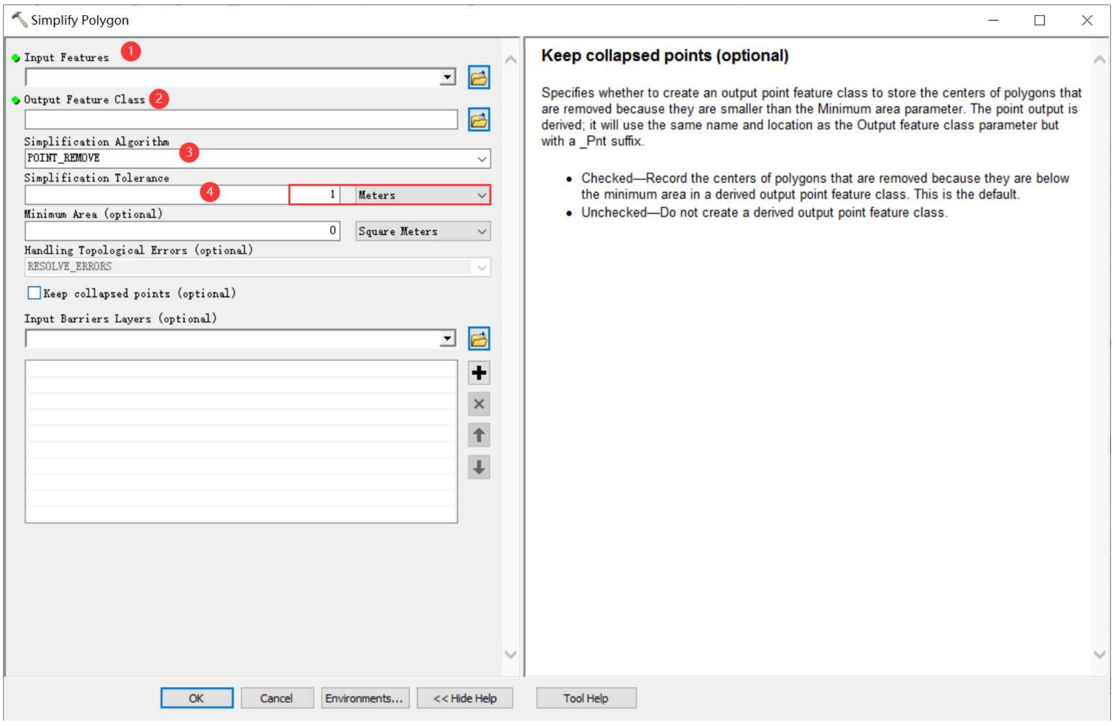


Figure A1. Input and option for Simplify Polygon in ArcGIS.

2. Calculating the total number of nodes using ArcGIS (Figure A2):

- Add a new field in the attribute table of dataset.
- Open Field Calculator.
- Switch the parser to python mode, and enter the following code "!shape.pointcount!" in the blue box to calculate the total number of nodes for each glacial lake boundary.

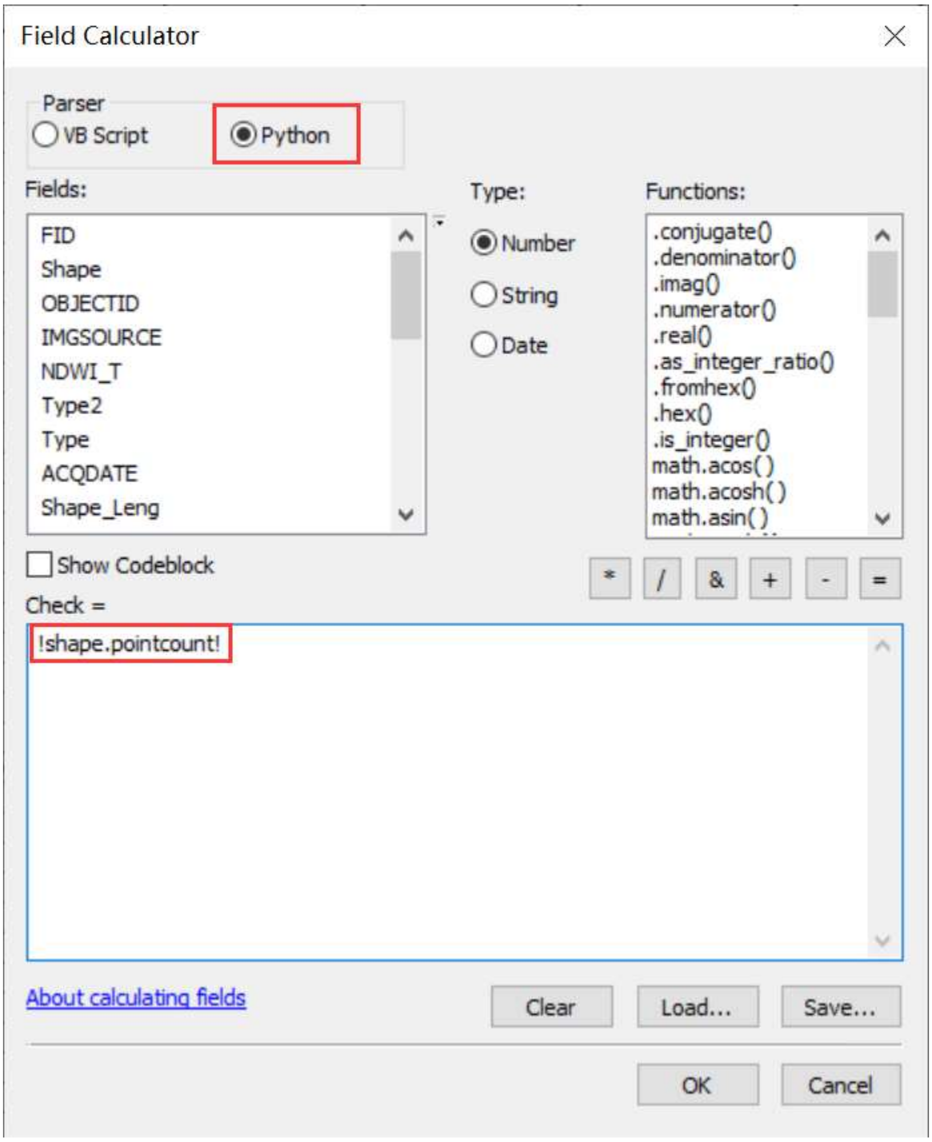


Figure A2. Total node calculation in ArcGIS.

3. Calculating the number of inner nodes:

For polygons without islands (Figure A3), use the equation 5. An inner node is a polygon vertex where the interior angle surrounding it is greater than 180 degrees. An outer node is the opposite of the inner node, where the interior angle is less than 180 degrees. We found that the outer nodes are usually four more than the inner nodes in our glacial lake dataset. The total nodes in ArcGIS contain one overlapping node to close the polygon, meaning the endpoint is also the startpoint. This extra count was deleted in the calculation (equation 5).

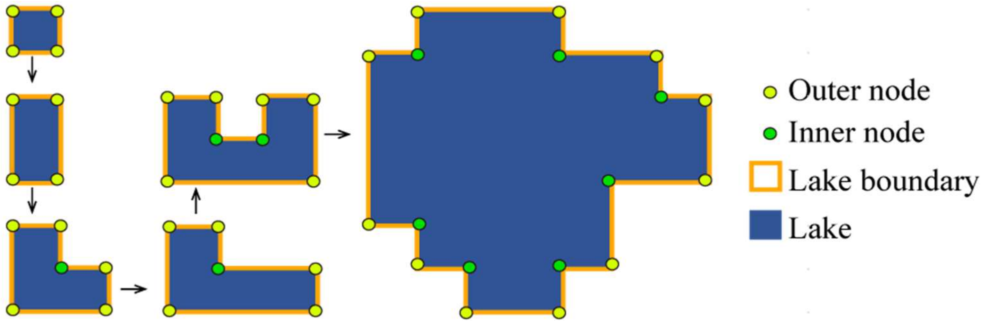


Figure A3. Sketch of outer and inner nodes of various glacial lakes without island.

For polygons with island (Figure A4) use the equation 6.

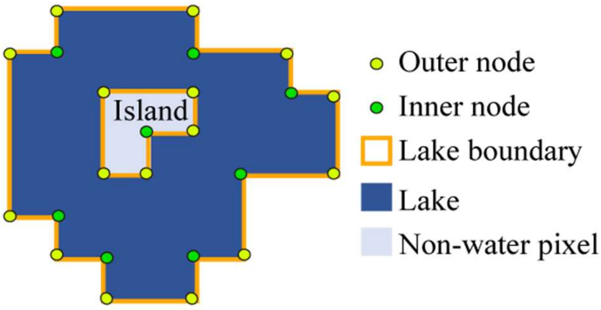


Figure A4. Sketch of outer and inner nodes for glacial lake with island.

We further specify the steps below to help implement equation 6.

Sept 1: detect the number of islands within each polygon.

- Convert the initial lake polygon to polyline using the “Feature To Line” tool (Figure A5).

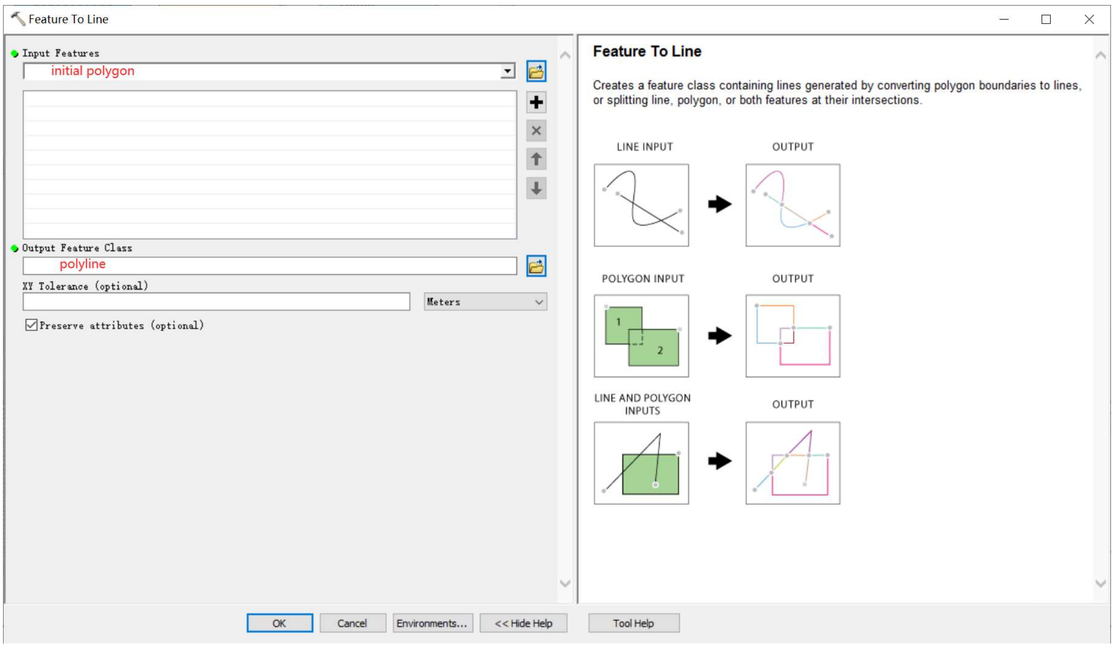


Figure A5. Feature To Line tool in ArcGIS

- Convert the polyline to generate a new polygon (Figure A6).

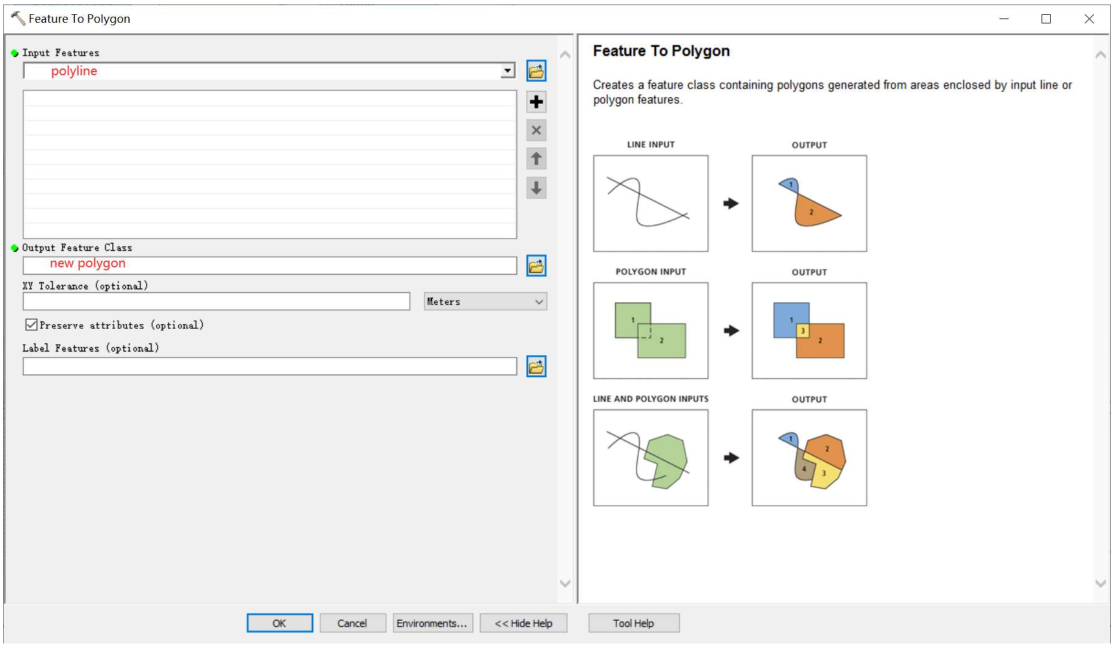


Figure A6. Feature To Polygon tool in ArcGIS

- Erase the new polygon by the initial polygon, which outputs the islands. Then we can count how many islands there are in each lake (Figure A7).

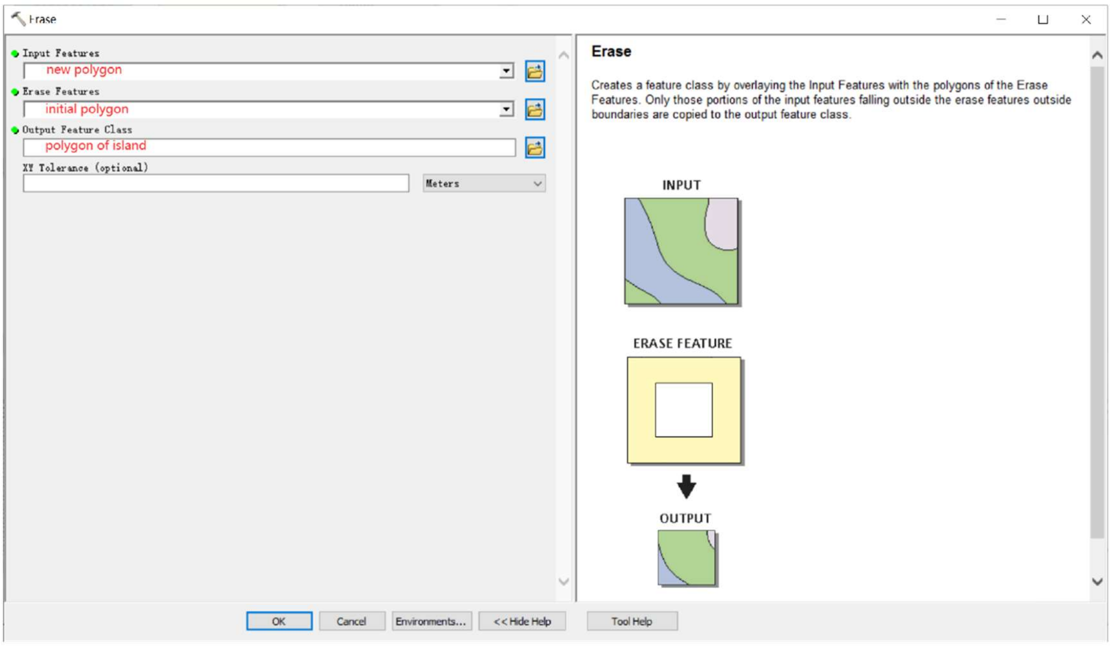


Figure A7. Erase tool in ArcGIS.

Step 2: calculate the number of inner nodes for each polygon with island using equation 6.

4. Calculating the uncertainty of lake mapping using equation 4.