A new global dataset of mountain glacier centerlines and lengths

Dahong Zhang [1, 2], Gang Zhou [1, 2], Wen Li [1, 2], Shiqiang Zhang [1, 2], Xiaojun Yao [3], Shimei Wei [3]

[1] College of Urban and Environmental Science, Northwest University, Xi'an 710127, PR China

[2] Shaanxi Key Laboratory of Earth Surface System and Environmental Carrying Capacity, Northwest University, Xi'an 710127, PR China

[3] College of Geography and Environment Sciences, Northwest Normal University, Lanzhou 730070, PR China

Correspondence: Shiqiang Zhang (zhangsq@lzb.ac.cn)

Abstract. The ~~Length~~ length of a glacier is ~~one of the~~a key determinant~~s~~ of ~~glacier~~ its geometry and is an important parameter ~~of~~ in glacier ~~inventory~~ inventories and modeling; glacier centerlines are the lines along which the main flow of glaciers takes place and, thus, are crucial inputs for many glaciological applications. In this study, the centerlines and maximum lengths of global glaciers were extracted using a~~n~~ self-designed automatic extraction algorithm based on the latest global glacier inventory data, digital elevation ~~data~~ model (DEM), and European allocation theory. ~~The glacier polygons were reconstructed according to the geometric principle and an automatic checking algorithm for the global glacier outlines was designed to filter erroneous or unsupported glacier outlines. The DEMs of global glacier-covered regions were compiled using available DEMs. An updated automatic extraction tool was designed independently, and a parameterization scheme with empirical thresholds was applied for data production.~~ The accuracy of the dataset was evaluated ~~using~~ through random visual assessments ~~with visible interpretation~~ and comparisons ~~comparative analysis~~ with the Randolph Glacier Inventory (RGI) version 6.0~~another dataset~~. A total of 8.25% of the outlines of the RGI were excluded, including~~The~~ 10,764 erroneous glacier polygons, 7,174 ice caps, and 419 nominal glaciers.~~ from the Randolph Glacier Inventory (RGI) version 6.0 were identified and excluded, accounting for 8.25% of the total. In~~ A total of~~,~~ 198,137 glacier centerlines were generated, accounting for 99.74% of the ~~total~~ input glaciers~~ and 91.52% of the RGI v6.0~~. The accuracy of glacier centerlines was 89.68%. ~~The~~ A comparison between the dataset and the previous dataset~~s~~ suggested that ~~the majority of~~most glacier centerlines were slightly longer than those in RGI v6.0~~. The extraction method of this study has a strong ability to obtain the maximum length of glaciers~~, meaning that the maximum lengths of some glaciers ~~were~~ had been likely underestimated in the past. The ~~dataset~~ constructed dataset ~~includes~~ comprises 17 sub-datasets, ~~such as the~~including global glacier centerline~~ dataset~~s, ~~global glacier~~ maximum length~~ dataset~~s, and ~~global glacier~~ DEM ~~dataset~~s, all of which can be found at~~ link~~: https://doi.org/10.11922/sciencedb.01643 (Zhang and Zhang, 2022).

1 Introduction

Mountain glaciers~~ which are distinct from the Greenland and Antarctic ice sheets,~~ are ~~also~~ shrinking rapidly (Hugonnet et al., 2021)~~. They are~~, altering regional hydrology (Pritchard, 2019), raising global sea levels (Cazenave, 2018), and ~~elevating~~ increasing the risk of natural hazards (Shukla and Sen, 2021; Zheng et al., 2021)~~. These glaciers~~, and are among the most climate-sensitive constituents of the world's natural water towers (Immerzeel et al., 2019). Under the influence of global climate change, studies on glacier area changes (Sommer et al., 2020; Li et al., 2021), ice thickness (Farinotti et al., 2019), mass balance (Zemp et al., 2019; Vargo et al., 2020; Wu et al., 2021; Mankoff et al., 2021), ice velocity field (Thogersen et al., 2019; Solgaard et al., 2021; Franke et al., 2022), the impact of debris-cover (Scherler et al., 2018; Shukla et al., 2020; Herreid and Pellicciotti, 2020), glacier meltwater (Noel et al., 2020), sediment release (Aciego et al., 2015; Li et al., 2019), and related hazards (Zhou et al., 2021b; Stuart-Smith et al., 2021; Kääb et al., 2021)~~ in glacier-covered regions~~ are essential for global water resources supply assessment and disaster prevention and reduction.

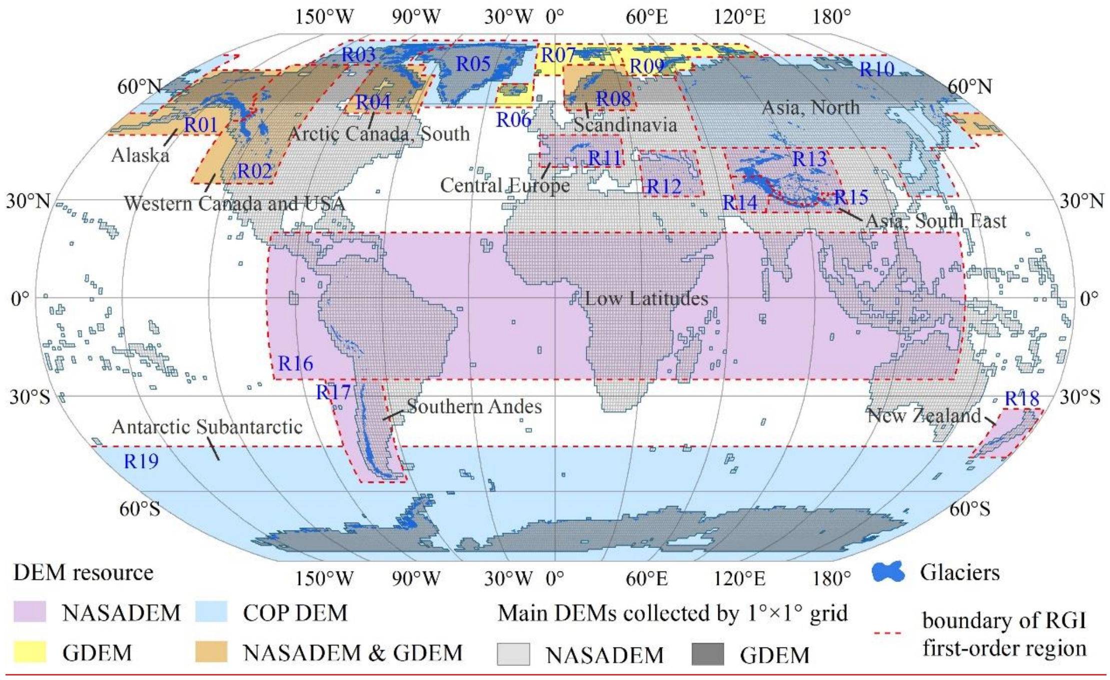
The most ~~obvious~~ noticeable distinction between glaciers and other natural ice bodies is their property ~~to~~ of moving~~e~~ towards lower altitudes under the influence of gravity. Glacier flow lines ~~are~~ correspond to a glacier's~~the~~ motion trajectories, ~~of a glacier~~ and the main flow line is the key trajectory. Due to the lack of glacier velocity field data, the main flow lines cannot be obtained on a large scale.~~The main flow line cannot be obtained on a large scale owing to the lack of glacier velocity field data.~~ The glacier centerline, generated via the axis line method (Le Bris and Paul, 2013; Machguth and Huss, 2014; Kienholz et al., 2014; Zhang et al., 2021), is typically used to represent the main flow line. The glacier centerline is a critical parameter for analyzing the ice velocity field (Heid and Kääb, 2012; Melkonian et al., 2017), estimating ~~the~~ glacier volumes (Li et al., 2012; Gao et al., 2018), and developing glacier models (Oerlemans, 1997; Sugiyama et al., 2007; Maussion et al., 2019).

Glacier length, usually ~~referring~~ refers to the maximum length of a glacier centerline (main flow line)~~,~~ and represents the longest motion trajectory of a glacier, which is ~~one of~~among the key determinants of glacier geometry and a basic parameter of glacier inventories (RGI Consortium, 2017) and modeling (Maussion et al., 2019). Glacier length fluctuations can be used to quantify glacier changes (Zhou et al., 2021a), such as by identifying glacier advancement, surge, or retreat. Glacier length fluctuations (e.g., Leclercq et al., 2014) have also been used to study the relationships with changes in glacier area (Winsvold et al., 2014) and the geometric structure of a glacier (Herla et al., 2017), estimate glacier volume in combination with the glacier area (Lüthi et al., 2010), and reconstruct annual averaged surface temperatures over the past 400 years on hemispherical and global scales (Leclercq and Oerlemans, 2011).

~~The~~ A complete global ~~complete~~ inventory (RGI Consortium, 2017) of glacier outlines (RGI Consortium, 2017) was created following the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC AR5). Three types of ~~To meet the demand for large-scale acquisition of glacier length,~~ automatic and semi-automatic methods have been proposed to meet the demand for large-scale acquisitions of glacier lengths.~~There are three types of methods: first~~First, ~~the~~ there are typical hydrological analysis methods (Schiefer et al., 2008), but they ~~the~~ result in lengths that are longer than equivalent maximum distances taken along typical longitudinal centerline profiles~~;~~. The second type is~~,~~ a simplified algorithm based on the skeleton theory (Le Moine and Gsell, 2015), but ~~this method~~ it has not been widely used~~;~~. ~~third~~Third, there are centerline methods based on the axis concept~~,~~ proposed by Le Bris and Paul (2013)~~,~~ and first applied to ~~the calculation~~ calculating~~of~~ global glacier length ~~for the first time~~ by Machguth and Huss (2014). However, with this type of algorithm, the glacier centerlines tend to be noticeably deflected by their tributaries ~~it is difficult to extract complex glaciers using the method~~ (Le Bris and Paul, 2013). The cost grid-least-cost route approach of Kienholz et al. (2014) based on the axis concept ~~has higher~~is more ~~accuracy~~accurate~~,~~ but also~~it requires~~ more labor and time-consuming, which limits its application to global glaciers. The trade-off function approach of Machguth and Huss (2014), which is ~~was~~ based on the axis concept, ~~their results cover~~has been applied to almost all global mountain glaciers ~~in the world~~ but excludes the centerlines of the branches of glaciers. Despite many attempts~~Therefore, researchers have been trying~~ to overcome these ~~difficulties~~ limitations in recent years (Yao et al., 2015; Yang et al., 2016; Ji et al., 2017; Hansen et al., 2020; Xia, 2020; Zhang et al., 2021)~~.~~, ~~To~~ to date, global datasets of the centerline and length of mountain glaciers are rare~~, including that of glacier branches~~. Based on our recent study ~~(Zhang et al., 2021)~~ on successfully extracting the glacier centerline using the Euclidean allocation method (Zhang et al., 2021), ~~in this study,~~ we aim to combine ~~free,~~publicly available digital elevation data into one global digital elevation model (DEM) ~~with~~ at ~~30~~ 30-m resolution and extending from 90°N to 90°S, to check and correct the global glacier outlines~~,~~ and obtain a new graphic dataset of the centerline and length ~~for~~ of global mountain glaciers ~~based on the updated DEM and outlines~~.

2 Study region and data

The glacier dataset used in this study was the Randolph Glacier Inventory version 6.0 (RGI v6.0) ~~RGI version 6.0~~ (http://www.glims.org/RGI/randolph.html, last accessed: 15 November 2021) released via the Global Land Ice Measurements from Space initiative (GLIMS), which is a globally complete collection of digital glacier outlines ~~of glaciers~~, excluding ice sheets (Pfeffer et al., 2014). RGI v6.0 includes 216,502 global glaciers (215,547 glaciers described in the product handbook) ~~worldwide~~, with a total area of 705,738.793 km^2 (RGI Consortium, 2017). All glaciers can be divided into 19 first-order glacier regions (Radić and Hock, 2010), ~~and these regions~~which were used in our study (Fig. 1).



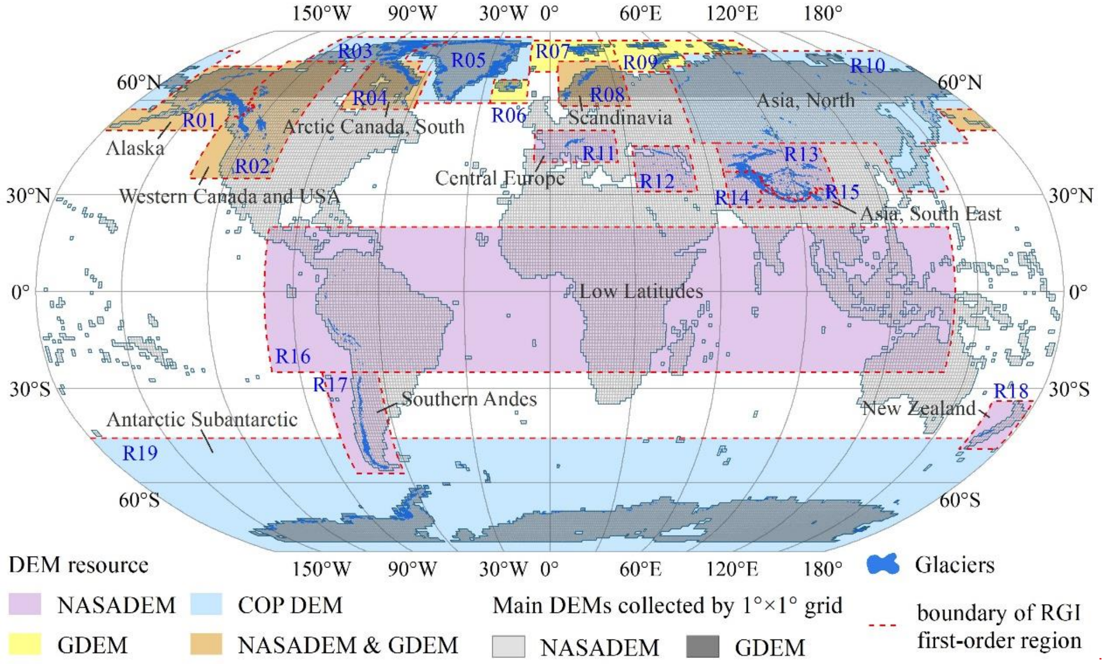


Figure 1. Distribution of global glaciers, first-order glacier regions, and Digital Elevation Models (DEMs) used. The background is the global DEM grid (1°×1°) covered by NASADEM and GDEM. GDEM and COP DEM represent the ASTER GDEM v3 and the Copernicus DEM, respectively. **Notes:** R01: Alaska; R02: Western Canada and USA; R03: Arctic Canada, North; R04: Arctic Canada, South; R05: Greenland Periphery; R06: Iceland; R07: Svalbard and Jan Mayen; R08: Scandinavia; R09: Russian Arctic; R10: North Asia; R11: Central Europe; R12: Caucasus and Middle East; R13: Asia, Central; R14: Asia, South West; R15: South Asia East; R16: Low Latitudes; R17: Southern Andes; R18: New Zealand; R19: Antarctic Subantarctic.

Five DEM products (Table 1) were ~~collected~~ used in ~~preliminary studies~~this study. The National Aeronautics and Space Administration (NASA) DEM (NASADEM) (https://lpdaac.usgs.gov/news/release-nasadem-data-products/, last accessed: November 17, 2021) was released by the Land Processes Distributed Active Archive Center (LP DAAC) in January 2020. NASADEM is the reprocessed version of the NASA~~As a modernization of the DEM and associated products generated from the~~ Shuttle Radar Topography Mission (SRTM) data (Farr et al., 2007)~~, the NASADEM,~~ with a low mean absolute error (MAE) (Carrera-Hernández, 2021)~~, is the successor of the NASA SRTM V3. The~~ and improved root mean square error (RMSE)~~ of NASADEM is smaller than that of SRTM~~ (Uuemaa et al., 2020). Serving the zonal extent of (56°S, 61°N), NASADEM was used as the preferred DEM in this study because of its superior performance. The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) is a 14-channel imaging instrument operating on the Terra satellite of NASA since 1999. The ASTER Global Digital Elevation Model (GDEM) version 3 (https://lpdaac.usgs.gov/news/nasa-and-meti-release-aster-global-dem-version-3/, last accessed: November 17, 2021) (Abrams et al., 2020) was released by Japan's Ministry of Economy, Trade, and Industry (METI) and NASA in July 2019. Using ICESat data, Carabajal and Boy (2016) found that ASTER GDEM v3 displayed smaller means, similar medians, and less scatter than the ASTER GDEM v2 in Greenland and Antarctica.~~ To determine the zonal extent of (56°S, 83°S) and (61°N, 83°N),~~ ASTER GDEM v3 was used as the second priority

DEM to cover the zonal extents of (56°S, 83°S) and (61°N, 83°N)~~in this study~~.

NASADEM and ASTER GDEM v3 do not cover all glacierized regions, missing parts of the polar regions and the Kamchatka Peninsula. Because of their high temporal and spatial resolutions at high latitudes, the ~~reference~~ Reference ~~elevation~~ Elevation ~~model~~ Model of Antarctica (REMA) (Howat et al., 2019) and ArcticDEM (https://www.pgc.umn.edu/data/arcticdem/, last accessed: November 17, 2021) were preferred as the supplementary data of our preliminary studies in these glacier regions. Nevertheless, ArcticDEM and REMA ~~are~~ were found to be inadequate~~not suitable~~ because of insufficient coverage and sporadic data. Some other DEMs~~(Fan et al., 2022; Zhang et al., 2022)~~ in high latitude areas (Fan et al., 2022; Zhang et al., 2022) were also not considered because their spatial resolutions are very different from that required in this study. Therefore, the wide-coverage Copernicus DEM (https://spacedata.copernicus.eu/web/cscda/cop-dem-faq, last accessed: November 17, 2021)~~ with a wide coverage~~ was finally ~~determined~~ selected as the supplementary dataset for glacier regions not entirely covered ~~via~~ by~~the~~ NASADEM and ASTER GDEM v3 ~~completely~~. The Copernicus DEM was~~ recently~~ released ~~(~~in November 2020~~)~~. ~~and t~~The accuracy assessment undertaken by its development team (the product handbook) using TanDEM-X DEM/World DEM ICESat GLAS reference points found an absolute vertical accuracy of approximately 10 m at the periphery of Antarctica and Greenland. ~~Finally~~In summary, NASADEM, ASTER GDEM v3, and Copernicus DEM were compiled to create a ~~30~~ 30-m DEM ~~of the completely~~that covered the study area completely.

Table 1. All DEMs collected in this study

DEM	Extent	Resolution	Access
NASADEM	[56°S, 61°N]	30 m	https://search.earthdata.nasa.gov/search
ASTER GDEM v3	[83°S, 83°N]	30 m	https://gdemdl.aster.jspacesystems.or.jp/
ArcticDEM	[55°N, 90°N]	2 m	https://earthengine.google.com/
REMA	[60°S, 88°S]	2m / 8m	https://earthengine.google.com/
Copernicus DEM	Global	30 m	https://panda.copernicus.eu/web/cds-catalogue/panda

Note: The interval in the 'Extent' column represents all landmasses within ~~a~~ the zonal range, ~~but coverage may~~although not~~ exist for~~ all areas within the range may be covered.

In addition, graphical data (Machguth and Huss, 2014) of glacier length in *. xy format (the UTM projection)~~an unknown projection coordinate system in High Asia were collected~~, which correspond to the attribute of the glacier maximum length (L_{max}) in RGI v6.0, were collected in High Asia. Because these data ~~was~~ were obtained from an unofficial source, we could not access their~~ data description~~ documentations and~~ only~~ recovered only the coordinates ~~matching between these~~of points matching~~and~~ some of the glaciers in RGI v6.0. ~~Registration~~ The registration of the *. xy file depends on~~ the~~ matching~~ between~~ its filename ~~and~~ with the feature identity document (FID) of the glacier polygon of RGI v6.0 in the same glacier area. The glacier lengths (*MHMLDS*) of successful registration were used as graphical validation data for this study.

3 Methods

3.1 Outline of workflow

This study relied on two key input datasets: the global glacier inventory and the compiled global glacier elevation. ~~The goal of this study was to establish a new dataset of global graphic glacier centerlines and lengths.~~ An outline of the workflow for establishing a new dataset of global graphic glacier centerlines and lengths is shown in Figure 2. The process was divided into six parts: (1) design an algorithm to check all glacier outlines~~, marks,~~ and exclude defective glacier polygons; (2) buffer glaciers to produce a mask containing global glaciers and their buffers; (3) mosaic the compiled global DEMs according to the mask~~s~~ in step 2 ~~of different glacier regions~~ to prepare the global glacier elevation data; (4) determine the automatic extraction parameters of global glacier centerlines ~~around the world~~ by repeated testing in each region; (5) input the global DEM ~~and~~, glacier outline dataset, and all parameters into the designed automatic extraction software (Zhang et al., 2021) to generate ~~the~~ global centerlines and length~~s in the global glacier~~; and (6) verify and compare with ~~other~~ existing centerline results ~~obtained via different methods~~ to evaluate the accuracy of the new datasets.

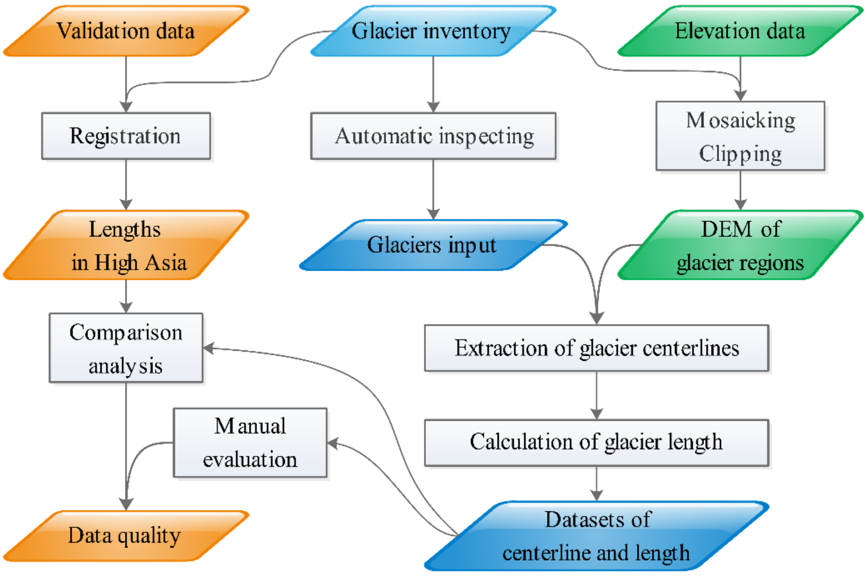


Figure 2. Workflow of the centerline and length dataset production.

3.2 Illustration of key methods

3.2.1 Pre-processing of glacier outlines

This study had strict requirements for glacier outlines~~;~~ and, therefore, all glacier complexes ~~should be~~were divided into individual glaciers ~~before~~ prior the centerline extraction. However, because of the limited semi-automatic glacier segmentation approach (Kienholz et al., 2013) and the high-priority strategy of completeness of coverage adopted by RGI v6.0 (RGI Consortium, 2017), some glaciers were not supported by our algorithm. These unsupported glaciers included three categories: glacier complexes with/without inaccurate segmentation (Fig. 3a-b), erroneous~~incorrect~~ glacier outlines (Fig. 3c) resulting from the vectorization, and flawed glaciers (Fig. 3d-f) generated by the automatic extraction algorithm. For the third category, we designed an identification algorithm ~~(see paragraph 3)~~ to mark and screen them (described in the last paragraph of this section). The flaws in these glacier outlines were mainly caused by ~~topology~~ topological errors of polylines/polygons, such as unclosed, sawtooth, and overlapping polygons. The first two categories ~~do~~ did not affect the ~~program's~~ algorithm's normal operation; however, the extraction accuracy ~~of the extraction results~~ is not always ~~difficult to~~ guaranteed. We ~~cannot~~ could not identify ~~them~~ the source of the problem at the time of the study, ~~present~~ and a solution is needed to improve the quality of the global glacier inventory.

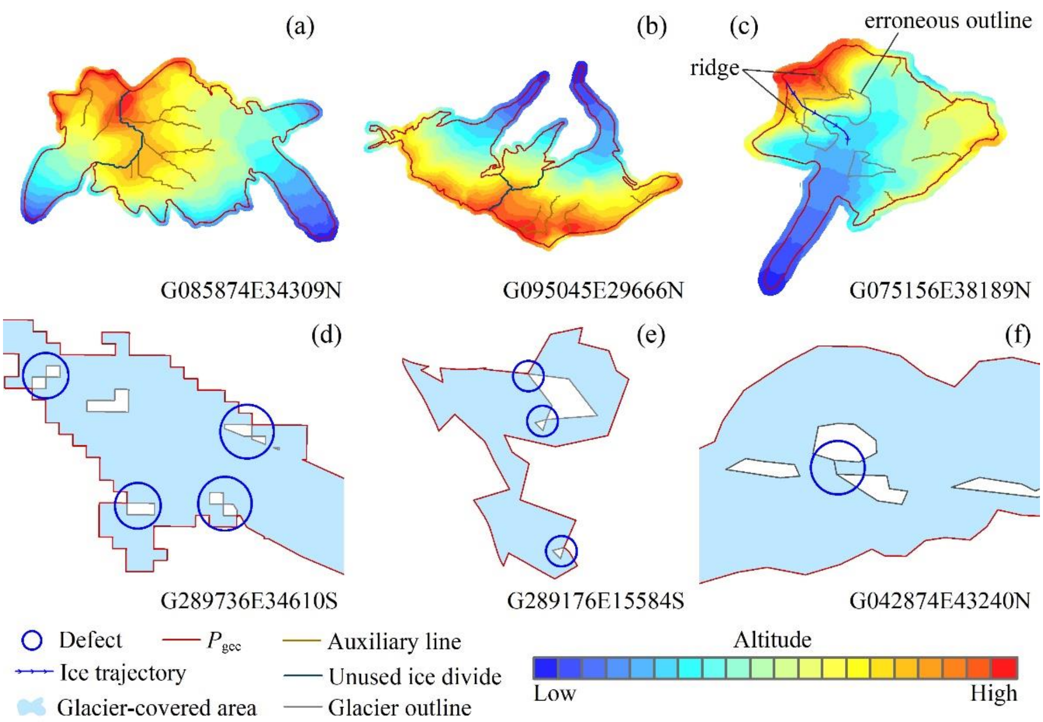


Figure 3. Schematic of three types of flawed glacier outlines. **(a-b)** Glacier complexes with/without inaccurate segmentation. **(c)** ~~Incorrect~~ Erroneous glacier outline caused by vectorization. **(d-f)** ~~Panels represent three c~~Common problems in flawed glaciers: defects in the automatic algorithm, defects in post-processing, and artificial errors. The ~~Auxiliary~~ auxiliary line represents lower-grade ice divides in ~~the~~ an individual glacier, which is part of the ridge lines.

In this study, we defined the external contour of a glacier (P_{gec}), namely, the polygon corresponding to the longest closed polyline of the glacier, to reduce the storage of DEMs and improve the efficiency of batch processing. The buffer masks of all glaciers (buffer distance: approximately 100 m) were generated by their P_{gec} to meet the requirement for the extent of input DEMs to be slightly larger than the P_{gec}. The buffer masks generated initially were partially~~relatively~~ broken because there ~~are~~ were overlaps or gaps between ~~the~~ adjacent polygons of the buffer zone; thus, ~~. We merged small spots to remove~~ polygons with a perimeter of less than 12 times the buffer ~~distance on the glacier buffer~~ mask~~s~~ distances of each region were removed.

The third category of flawed glaciers (Fig. 3d-f) ~~with flaws~~ was identified by obtaining P_{gec}. ~~In the third category, t~~The most common error type ~~is~~ was a glacier outline with two or more closed polylines with the same endpoint ~~in a glacier~~. These flawed glacier outlines were identified by assessing whether there were multiple polylines sharing endpoints after converting the glacier from a polygon to a polyline. However, these outlines did not include the unclosed types: ~~There were also~~ a few glacier ~~polygons~~ outlines that appeared to be closed polylines, but had geometric flaws such as non-coinciding head and tail endpoints of the polylines.~~with false closed polylines, which are the head and tail endpoints of the polylines that do not coincide, but the distance is less than the tolerance. The solution are as follows: flawed glacier outlines were identified by judging whether there were multiple polylines sharing endpoints after converting the glacier from a polygon to polylines, but these outlines do not include the false closed type.~~

3.2.2 Preparation of input datasets

All the data associated with the dataset production were processed in units of first-order glacier regions. The input glacier outlines excluded all the defective glacier outlines. Similarly, the nominal glaciers (represented by an ellipse) and ice caps remarked in RGI v6.0 were also treated, which were distinguished by two attributes: status (nominal glacier) and form (ice cap). The inspection ~~results (Table 2)~~ of glacier outlines show that there are 10,764 defective glacier outlines (*FGODS*) in RGI v6.0, accounting for approximately 4.97% of the total (216,502) (Table 2). After excluding nominal glaciers (461) and ice caps (7,174), 198,646 glaciers remained as input glacier outlines (*IGODS*), accounting for 91.75% of the global mountain glaciers.

Table 2. Preprocessing results of different glacier regions and information of input datasets.

Region	Region Name	Total	Ice Cap	Nominal glacier	Flawed glacier	Glacier input	DEM input
R01	Alaska	27108	0	0	704	26404	NASADEM, GDEM
R02	Western Canada and USA	18855	0	0	1564	17291	NASADEM, GDEM
R03	Arctic Canada, North	4556	650	0	47	3869	COP DEM
R04	Arctic Canada, South	7415	953	0	63	6409	NASADEM, GDEM
R05	Greenland Periphery	20261	1658	0	1547	17247	COP DEM
R06	Iceland	568	133	0	1	435	GDEM
R07	Svalbard	1615	144	0	12	1460	GDEM
R08	Scandinavia	3417	0	4	75	3338	NASADEM, GDEM
R09	Russian Arctic	1069	460	0	0	609	GDEM
R10	North Asia	5151	5	116	136	4899	COP DEM
R11	Central Europe	3927	0	2	76	3849	NASADEM
R12	Caucasus Middle East	1888	0	339	2	1547	NASADEM
R13	Central Asia	54429	1545	0	28	52858	NASADEM
R14	South Asia West	27988	295	0	1946	25792	NASADEM
R15	South Asia East	13119	289	0	4	12826	NASADEM
R16	Low Latitudes	2939	0	0	724	2215	NASADEM
R17	Southern Andes	15908	623	0	3828	11734	NASADEM
R18	New Zealand	3537	0	0	0	3537	NASADEM
R19	Antarctic Subantarctic	2752	419	0	7	2327	COP DEM
--	--	216502	7174	461	10764	198646	--

Note: GDEM and COP DEM ~~represent~~ refer to ASTER GDEM v3 and Copernicus DEM, respectively.

The P_{gec} of all glaciers in RGI v6.0 ~~constitute~~ comprises the global glacier external contour dataset (*GGECDS*), which generated the global mountain glacier's buffer mask dataset (*GGBMDS*) ~~of global mountain glaciers~~. The collected DEMs were extracted using *GGBMDS* and 43,035 DEM tiles were generated. They were then mosaicked according to different first-order glacier regions to generate a global glacier elevation dataset (*GGEDS*). The details of the two input datasets are presented in Table 2.

3.2.3 Generation of centerline and glacier length

Glacier centerlines and length were automatically extracted with the ~~The automatic extraction tool of~~ 'GlacierCenterlines_Py27' (Update to version 5.2.1) tool~~was used~~, which is based on the axis concept and Euclidean allocation (Zhang et al., 2021). The principle is briefly explained as follows: the highest and lowest points of the external outline of a glacier ~~as two endpoints~~ were extracted as two ~~,~~ endpoints that divide the glacier outline into two parts; In the glacier polygon, points that have the equal shortest distances to the two parts were identified as other vertices; The line formed by two endpoints and these other vertices was regarded as the glacier centerline.~~cells with the equal shortest distances from the cell to both sides were identified in a glacier polygon, and the line formed by these cells was regarded as the glacier centerline.~~ The maximum length of glaciers was calculated using an algorithm similar to the critical path. The updated contents focused on formulating the parameterization scheme (Appendix A: Table A1) for extracting global glacier centerlines, as well as repairing some newly discovered bugs, such as a dead cycle in the process of auxiliary line extraction. All glacier outlines included in the *IGODS* were divided into ten levels ~~(Table 3) using~~based on the proportion of cumulative area after ranking the area of all input glacier polygons from small to large (Table 3). The ~~User-defined~~ Albers projection (See Supplement for detailed parameter files) with WGS1984 ~~as the reference ellipsoid were~~ was used as a unified projection coordinate system for each glacier region.~~The central meridian, standard parallel 1, standard parallel 2, and origin latitude of the different glacier regions were determined by their spatial extent.~~ The empirical values of the other parameters were determined in repeated attempts, and their values ~~had a~~were significantly ~~correlation~~ correlated with the glacier scale. The generated glacier centerlines ~~generated~~ were merged according to the glacier regions. ~~and~~ Then, the graphics and attribute information of glacier length were exported as corresponding independent ESRI shapefiles. In addition, ~~some~~other data ~~key~~ associated with the dataset production ~~data~~ were exported, such as the segmentation results of glacier outlines, the lengths in the accumulation and ablation region of each glacier, the lowest points, the local highest points (P_{max}), the ~~failed~~ extracted failed glacier ~~outlines~~outlines ~~dealt~~, and logs.

Table 3. ~~Statistics of g~~Global glaciers stratified by area~~by different levels~~.

Level	Count	Area/km^2	Acc. area/km^2	Percent	Interval/km^2
L1	165593	1.00	41313.79	10%	[0.01, 1.00]
L2	22833	3.57	82629.47	20%	(1.00, 3.57]
L3	6906	11.39	123947.69	30%	(3.57, 11.39]
L4	2149	35.51	165282.14	40%	(11.39, 35.51]
L5	698	103.10	206631.32	50%	(35.51, 103.10]
L6	262	248.26	247917.55	60%	(103.10, 248.26]
L7	113	521.40	289227.71	70%	(248.26, 521.40]
L8	55	1087.47	330595.34	80%	(521.40, 1087.47]
L9	27	2657.74	374312.14	90%	(1087.47, 2657.74]
L10	10	6004.85	413136.71	100%	(2657.74, 6004.85]
Total	198646	--	--	--	--

3.2.4 Accuracy assessment

A random assessment was prioritized to assess the accuracy of the extracted centerlines~~in this study~~. We randomly selected 100 glaciers in each of the 19 glacier regions, ~~and obtained 19 samples with~~obtaining a total of 1,900 glacier centerlines (N_G). These glacier centerlines were divided ~~by artificial inspection~~ into three first-level categories (Zhang et al., 2021): correct (I), inaccurate (II), and incorrect (III). Type II mostly ~~contains~~ contained glaciers with accurate glacier maximum lengths but missing, redundant, or unreasonable branches of glacier centerlines. When calculating the dataset~~verification~~ accuracy, Types I and II were regarded as correct, and only Type III was considered incorrect. Finally, the ~~glacier~~ proportion of Type III glaciers in the sample was counted, and the valuation result (R) was calculated using Eq. (1)~~.~~:

$$R = \sum_{i=1}^{19} \frac{S_i \times N_{T_i}}{N_G} \quad , (1)$$

where N_G is the total ~~quantity~~ number of glaciers centerlines and N_{Ti} and S_i are the verification accuracy and the number of glaciers in ~~the corresponding glacier region of~~ the *i* th ~~sample~~ glacier region (i = 1, 2, 3, ..., 18,19), respectively.

This study's ~~All glacier~~ maximum glacier lengths (G_{Lmax}) ~~in this study~~ were compared with the L_{max} (Machguth and Huss, 2014) in RGI v6.0 using linear correlation and ratio analysis. ~~Here, we took L_4 – L_{10} at the glacier level as the same grade for statistics.~~ The correlations between G_{Lmax} and L_{max} were established according to different glacier regions, and glacier levels, and the length ratio~~,~~ R_r (Eq. 2)~~,~~ was calculated~~.~~:

$$R_r = \frac{G_{L\max}}{L_{\max}} \qquad (2)$$

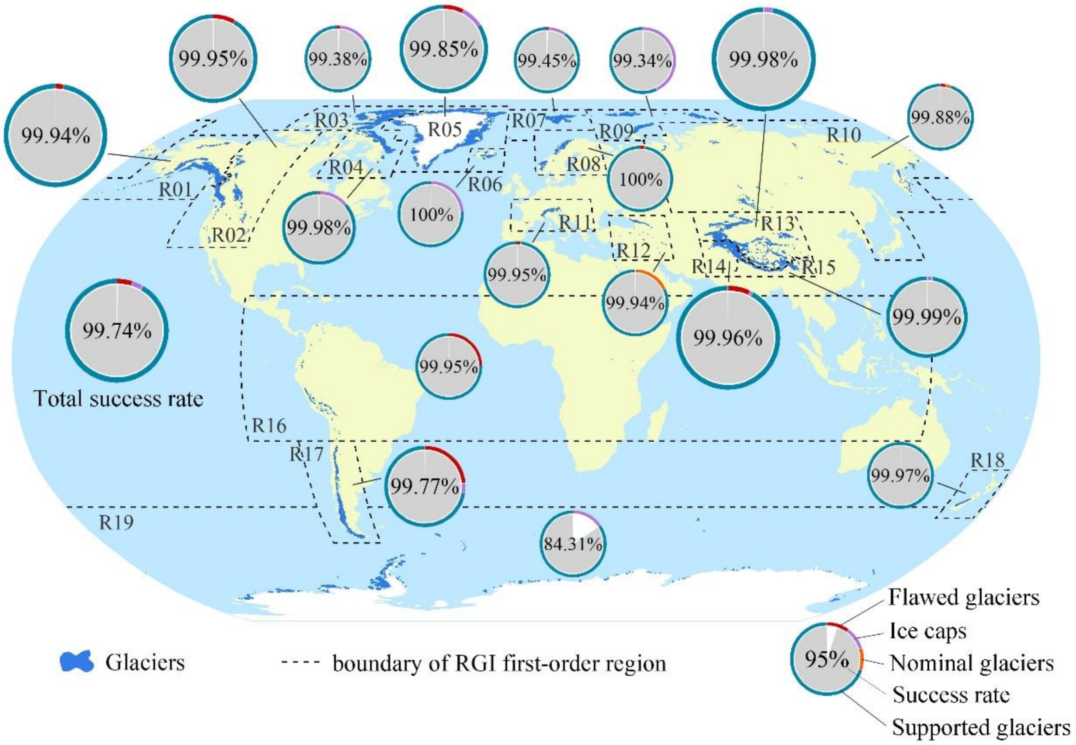
In addition, considering the differences between the graphics, we also collected graph data of glacier length extracted by Machguth and Huss (2014). Considering the limited availability of the data (obtained: R13–R15), we only compared two glacier-covered regions in the Himalayas: Mount Qomolangma and Kangchenjunga (the world's third-highest mountain) and their surrounding areas.

~~$$R_r = \frac{G_{L\max}}{L_{\max}} \qquad (2)$$~~

4 Results

4.1 Centerline and length of glaciers

Taking the *IGODS*, *GGEDS*, and other model parameters (Appendix A: Table A1) as input data, 198,137 glacier centerlines were automatically generated using the centerline extraction tool of 'GlacierCenterlines_Py27 v5.2.1', with an overall success rate of 99.74%. The number and proportion of flawed glacier outlines, nominal glaciers, ice caps, input glacier outlines, and extraction results for distinct glacier regions are shown in Fig. 4.



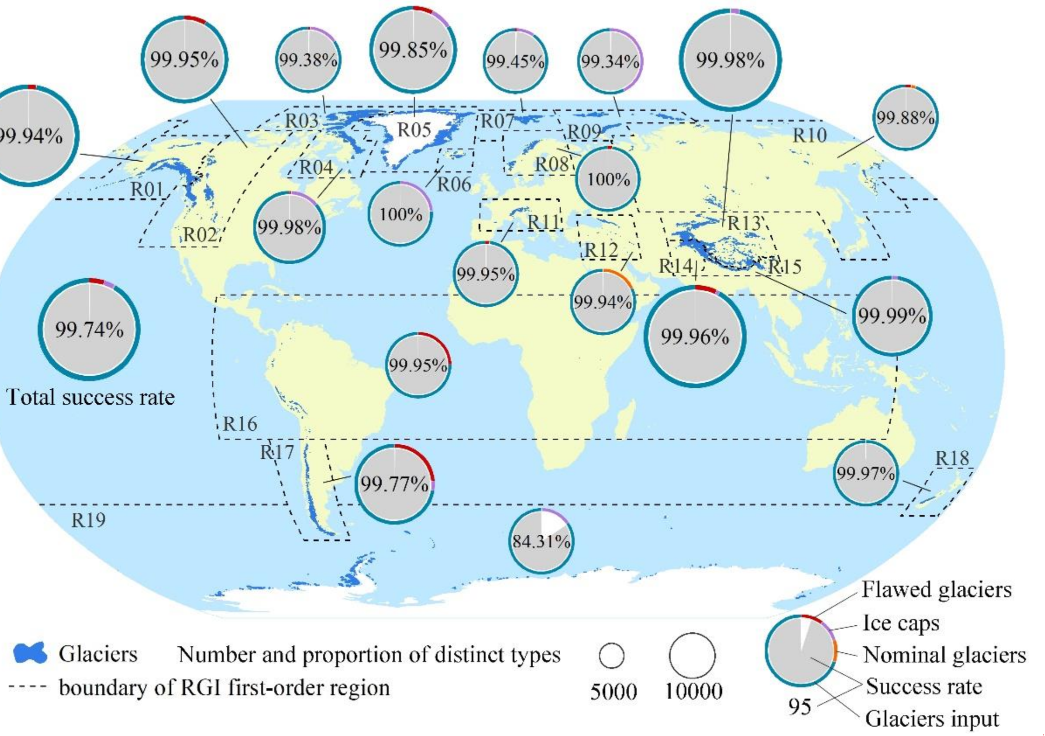


Figure 4. Extraction results of glacier centerlines in different glacier regions. The ring in the pie chart rings represents indicate the proportions of input glaciers number and of the number of excluded three glacier types of excluded glaciers with total number of glaciers in the region. The Pie pie chart represents the correct rate, which is the proportion of correctly extracted glaciersthe extraction result number with input glacier quantity. The size of the pie/ring represents the grade of the glacier numbers in the region.

Except for Antarctica and Subantarctica (R19), the success rates of extracting glacier centerlines in other glacier regions were was greater than 99%, which indicates that the automatic extraction algorithm for glacier centerlines is robust. A small number of glacier outlines with falsely closed boundariesproblems and unidentified ice caps were the main reasons for the failure of the automatic extraction of glacier centerlines; however, it is difficult to establish rules for accurately identifying these glacier polygons. In total, 510 unsuccessful glacier outlines were identified, of which Antarctic-Subantarctic (R19) accounted for 71.57%; Southern Andes (R17) and Greenland Periphery (R05) for 5.29% and 5.1%, respectively; Arctic Canada North (R03) and Alaska (R01) for 4.71% and 2.94%, respectively; and other glacier regions for less than 2%.

Overall, the global glacier centerline dataset (*GGCLDS*) constructed in this study contained 91.52% of the total glaciers in RGI v6.0. The lengths of each branch of the glacier centerline were derived and the longest branch lengths of the glacier centerline were defined as the glacier maximum length (G_{Lmax}), which were used to form the global glacier maximum length dataset (*GGMLDS*). The average centerline length of all branches of a glacier is called the glacier mean length (G_{Lmean}). In addition, the median glacier altitude was regarded as the equilibrium line altitude (ELA) (Machguth and Huss, 2014);. the The part with G_{Lmax} higher than ELA was regarded as the length of the glacier accumulation zone (G_{Lacc}), and the part lower than ELA was regarded as the length of the glacier ablation zone (G_{Labl}), which formed the glacier accumulation zonal length dataset (*GACLDS*) and glacier ablation zone length dataset (*GABLDS*). The key process data corresponding to *GGCLDS* were also output, to form the glacier outline segmentation results (*GOSRDS*), lowest points (*GGLPDS*), local highest points (*GLHPDS*), and unsuccessful glacier outlines (*GUGODS*). The fields involved in all datasets are listedexplained in Table 4.

Table 4. Description of the attributes contained in all datasets.

Name	Data type	Char. length	Description
GLIMS_ID	Char.	14	Unique code of a glacier
Type	Long int.	4	Glacier grade in this study
MaxL	Float	8	Glacier maximum length (Unit: m)
MeanL	Float	8	Glacier average length (Unit: m)
ELA	Long int.	4	Equilibrium line altitude (Unit: m)
AccL	Float	8	Length in the accumulation region (Unit: m)
AblationL	Float	8	Length in the ablation region (Unit: m)
Id	Long int.	8	Data code of the same glacier
BS	Long int.	8	Tag of the same segment in a glacier
RASTERVALU	Long int.	4	Altitude of a P_{max} (Unit: m)

Note: Char.: and int. represent Character; and int.: integer, respectively.

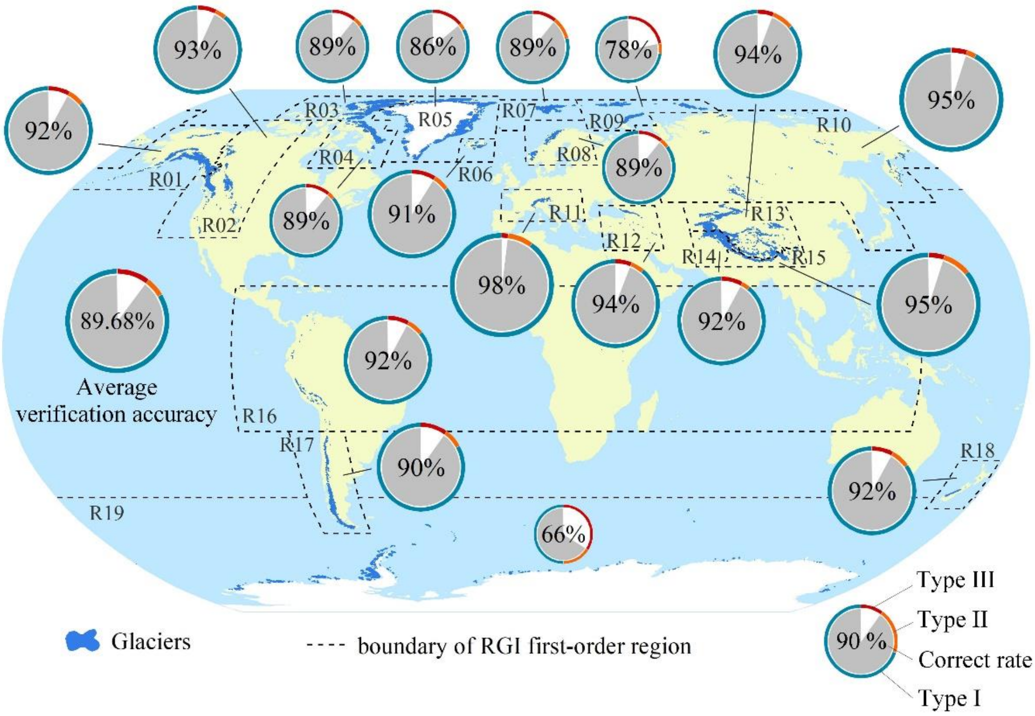
The glacier outlines of RGI v6.0 without centerlines results obtained in this study were limited by the quality of the glacier polygons, which mainly correspond to the flawed glacier outlines (*FGODS*), and the identified ice caps in RGI v6.0 (Table 2). Among the *FGODS* (10,764), Southern Andes (R17) had the most, followed by Southwest Asia (R14); Western Canada and USA (R02) and Greenland Periphery (R05), with slightly more than 1,500; and Low Latitudes (R16) and Alaska (R01), with slightly more than 700. There were 451 in other glacier regions, including two regions with 0 defective glacier outline, the Russian Arctic (R09) and New Zealand (R18). Among the ice caps (7174) identified by RGI v6.0, slightly more than 1,500 were in R05 and Central Asia (R13), between 500 and 1,000 in the Arctic Canada South (R04), Arctic Canada North (R03), and R17, and less than 500 in other glacier regions. Nominal glaciers (461) existed in three glacial regions:

Caucasus Middle East (R12), North Asia (R10), and Scandinavia (R08).

4.2 Data validation

4.2.1 Random ~~self-~~assessment results

The evaluation results using random samples from the glacier centerline dataset suggested that the average verification accuracy of the glacier centerline dataset was 89.68%. There were significant differences ~~in~~ across the ~~accuracy~~ accuracies of the 19 glacier regions around the world (Fig. 5). Among them, R11, R15 and R10, R09, and R19 ~~were~~ had the highest (98%), second highest (95%), second lowest ~~lower~~ (78%), and lowest (50%) accuracies, respectively. In terms of types, the average proportions of Types I and II were 83.53% and 6.16%, respectively. The proportions of Type I in R07 and R09 ~~was~~ were relatively low, at 79% and 73%, respectively, and the lowest in R19 was only 50%. Type II had the highest proportion in R19 at 16%, followed by R07 (10%). Moreover, Type II accounted for more than 5% in seven regions: R11, R13, R17, R18, R16, R01, and R06.



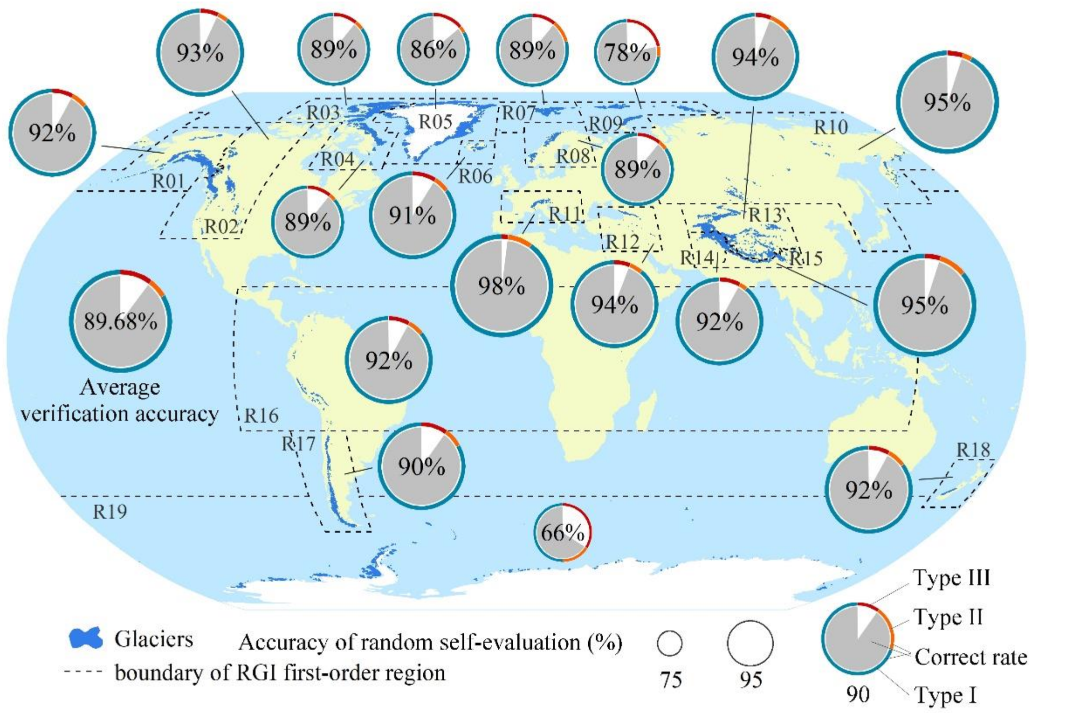


Figure 5. Statistical chart of random evaluation results. The ~~ring in the~~ pie chart rings ~~represents~~ show the proportion of each type with the total number of samples in the region. The ~~Pie~~ pie chart ~~represents~~ shows the correct rate, which is the proportion ~~of the number~~ of Types I and II ~~with~~ in each region ~~sample quantity~~. The size of the pie/ring represents the grade of the correct rate in the region. Types I, II, and III (See Section 3.2.4) ~~represent the centerline of~~are correct, inaccurate, and incorrect centerlines, respectively.

The above results indicate that, in addition to the three glacier regions of R07, R09, and R19, the random samples of the glacier centerline dataset ~~have~~ show excellent performance in terms of accuracy, particularly in R02, R12, and R14. The unmarked ice cap and local low-quality DEM were the main reasons for the poor quality of the glacier centerlines in R07 and R09~~, respectively~~. Owing to glacier complexes and low altitude differences in low-quality DEMs at the glacier tongues, ~~the quality of the~~ glacier centerlines obtained in R19 ~~was~~ were of poor quality~~.~~, but were included for completeness~~However, from the viewpoint of dataset coverage, we provided the extraction results of the glacier centerline in R19~~.

4.2.2 Comparisone with previous results

~~After applying this algorithm to the global glacier inventory RGI v6.0, w~~We compared the glacier lengths (G_{Lmax}) automatically obtained in this study with those (L_{max}) obtained by Machguth and Huss (2014) (Fig. 6). After eliminating 5408 glaciers with L_{max} value of -9 (~~no results~~missing value), the length values of the other 192,728 glaciers in the global glacier length dataset were ~~directly~~ compared directly. The G_{Lmax} and L_{max} were generally comparable (Fig. 6a). The glaciers in grades L_4–L_{10} showed excellent ~~fitting degrees~~agreement, while those of L_1–L_3 determined the linear correlation coefficient owing to their large number. There were approximately 35,000 ~~number of~~ glaciers with a length ratio (R_r) between G_{Lmax} and L_{max} greater than 1.55, ~~(Fig. 6b) was approximately 35,000, which~~ and these were excluded from the histogram in Fig. 6b ~~statistics~~

because there was a high probability that the length of at least one of the two datasets was ~~erroneous~~wrong. The peak value of the histogram~~ (Fig. 6b)~~ of R_r is in the interval 1.05–1.15 and R_r in the interval 0.95–1.25 accounts for 64.55% (Fig. 6b). The glacier length G_{Lmax} determined in this study was generally 10% longer than L_{max}~~ and the average value was approximately 10%~~, which ~~indicates~~ suggests that the glacier centerline lengths were probably underestimated in previous studies. In addition, the ~~abnormal value of the~~ length ratio of glacier L_1 was the highest and the median value was high (Fig. 6c). The R_r values of glaciers L_4–L_{10} fluctuated greatly. The R_r distributions of glaciers L_2 and L_3 were relatively concentrated. The reason for this is that the length of glacier L_1 was affected by the DEM, while glaciers L_4–L_{10} were mainly ~~disturbed~~ impacted by differences in glacier scale and the accuracy of the auxiliary line.

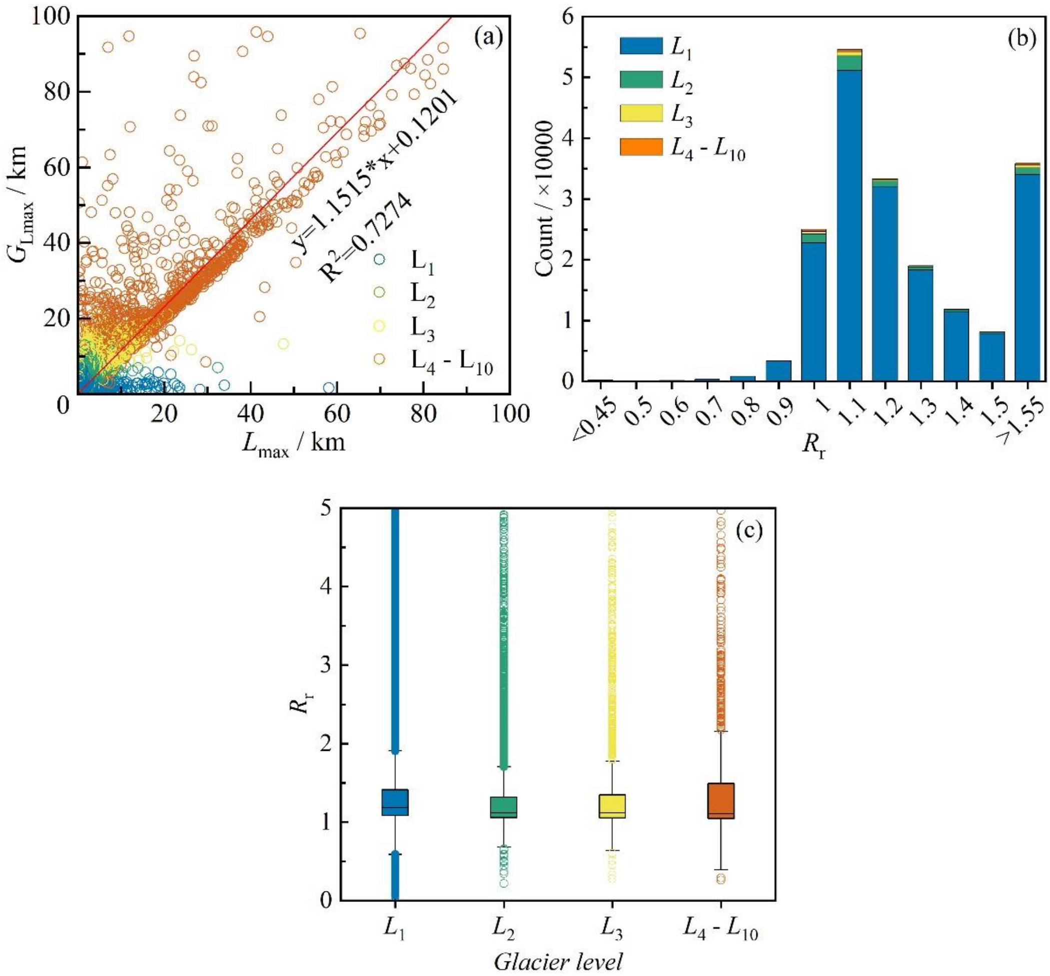


Figure 6. Comparison ~~of~~ between longest centerlines calculated in this study and by Machguth and Huss (2014). **(a)** Linear regression of maximum length for all input glaciers (*IGODS*), determined ~~in the~~as G_{Lmax}~~,~~ calculated in this study and L_{max} obtained in Machguth and Huss (2014). **(b)** Histogram of length ratio (R_r, G_{Lmax}/L_{max}) for distinct grades of glaciers. **(c)** Box plots of length ratio (R_r) for different scales of glaciers.

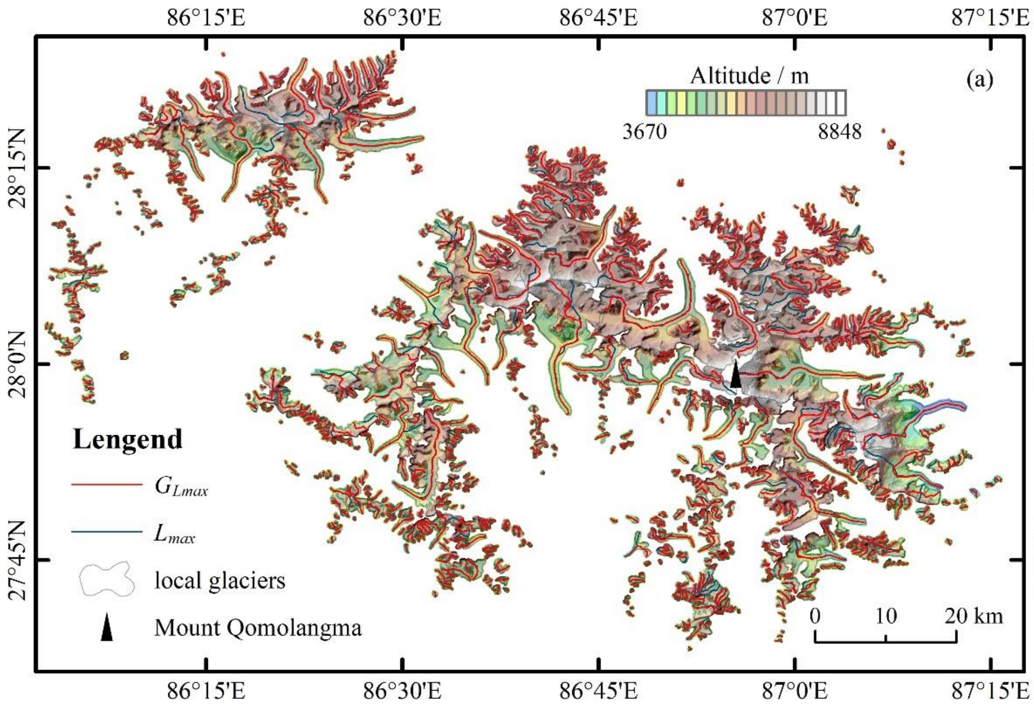
Comparisons between G_{Lmax} and L_{max} for each first-order glacier region and all random samples are shown in Appendix B. The~~re was a preferable~~ fit~~ting degree~~ between G_{Lmax} and L_{max} was better in seven glacier regions including R01, R04, R07, and R12–R15, in which the R^2 was larger than 0.95 (Fig. B1). The R_r in R17 (R^2 = 0.8174), R05 (R^2 = 0. 8136), and R03 (R^2 = 0.6311) were ~~poor~~lower, whereas that in R19 (R^2 = 0.5487) was the worst. The R^2 values of the other eight glacier regions were between 0.85 and 0.95. The histograms (Fig. B2) suggest that G_{Lmax} and L_{max} fitted well in R04, R06, R07, R09, and R12–R15 because they had recognizable single peak values. The peak values of R03, R05, R17, and R19 were not prominent and the proportion of glaciers with $R_r > 1.55$ was extremely high, further increasing the uncertainty in glacier length ~~results~~ estimates in these four regions. R01, R07, R08, R11–R15, and R18 performed well in the box plot (Fig. B3), whereas the results for R09 were not good. Moreover, the fit~~ting degree~~ of all random samples was poor (Fig. B1, R^2 = 0.7547), the peak value was more prominent (Fig. B2), and the length ratio distribution of glaciers of different grades was relatively scattered (Fig. B3). In general, the glacier lengths of R07 and R12–R15 were the closest, while there were significant differences in R03, R05, R17, and R19.

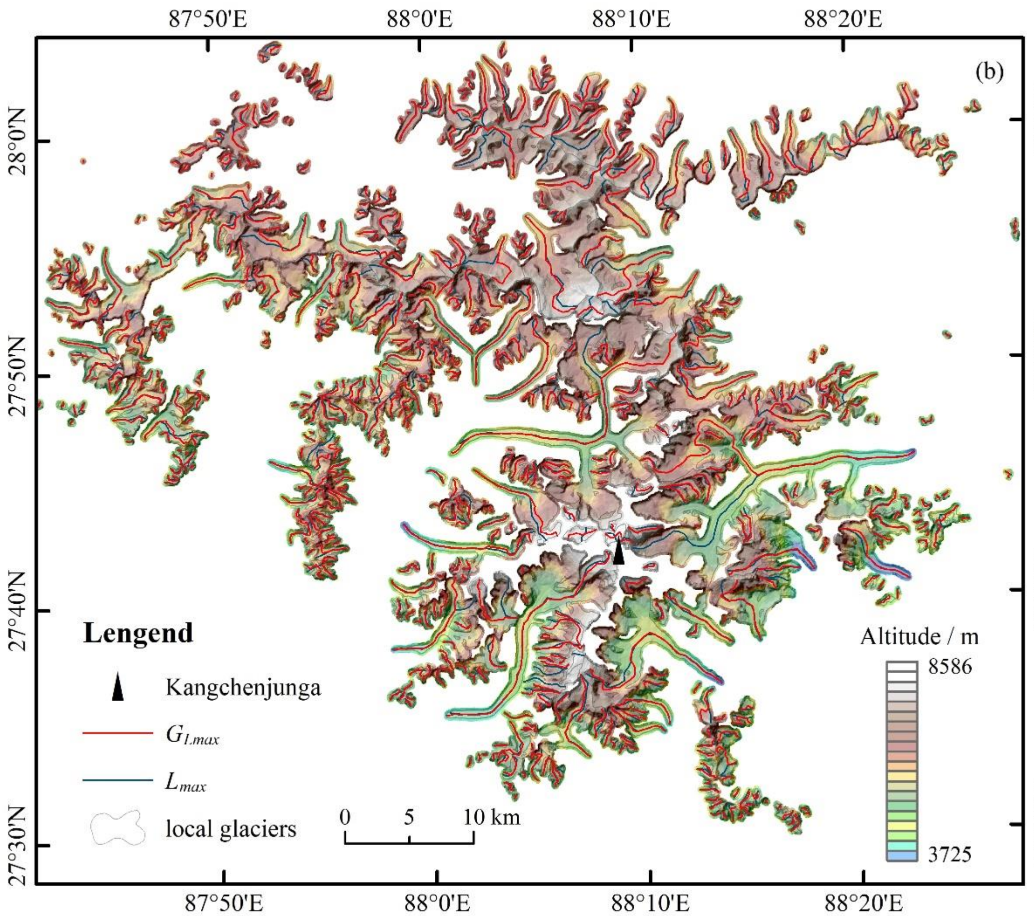
Furthermore, the graphic results, which were collected for the maximum length of glaciers in parts of High Asia (Machguth and Huss, 2014), were used to compare with this study's~~the results~~. ~~In two~~ Two parts of R15~~,~~ were showed, which were Mount Qomolangma and its surrounding area (Fig. 7a) and Kangchenjunga and its surrounding area (Fig. 7b).~~, the glaciers showed a flaky distribution for mapping.~~ ~~Visible~~ Visual comparison ~~was~~ suggested that the extraction ~~method~~ approach used in this study ~~had likely a strong ability to obtain the maximum length of glaciers (Fig. 7a)~~was robust (Fig. 7a) and that its sensitivity to topography was lower than that of Machguth and Huss (2014) (Fig. 7b). ~~Both sets of glacier length extraction schemes were valid and there were large~~ Large differences in glacier length extraction schemes are present only in a few glaciers or in certain types of glaciers, such as slope glaciers and ice caps.

4.2.3 Uncertainties and possibilities for improvement

Although we compared the two current global length datasets, it is still difficult to accurately characterize~~reflect~~ the dataset's quality ~~of the dataset~~ in this study. For ~~some~~ glaciers ~~that~~ for which ~~are not provided~~ centerlines were not provided in this dataset, ~~data~~ users need to update the corresponding glacier outlines and could use the automatic extraction tool provided in this study to generate their centerlines, ~~which involves~~including the defective glacier outlines (*FGODS*), nominal glaciers and ice caps of the RGI v6.0. Specifically, the centerlines of the *FGODS* rely on the glacier outlines that meet the requirements of this study. These glacier outlines include glacier inventory data from other sources~~,~~ or the *FGODS* that ~~are~~ were repaired by some algorithm~~s~~ or manual process. Nominal glaciers are similar to *FGODS*, and also require users to obtain corresponding glacier outlines. Automatic approaches to dividing ice caps from glacial complexes into individual glaciers are currently limited, and data users can only use their own ~~criterion~~ criteria to ~~divide~~ separate the ice caps and then use our tool to generate the centerlines. In addition, prioritizing the coverage of this dataset, we designed a geometry-based algorithm to repair *FGODS* and provided ~~data~~ users with their centerlines in the form of a supplementary dataset. ~~, and c~~Corresponding codes and results can be seen in sub-datasets *CODES* and *SUP_220707*.

The automatic extraction algorithm in this study is more suitable ~~for application~~ to single-outlet glaciers, particularly valley glaciers; it is not suitable for ice caps, flat-top glaciers, and tidal glaciers, ~~that are~~which tend to be widely distributed in the Antarctic, sub-Antarctic, northern Canadian Arctic, ~~and~~ among other areas. In short, the uncertainties in this dataset come probably from the centerlines of some slope glaciers and the ice caps that are not identified in RGI v6.0, or a few centerlines with unpredictable quality due to the input data such as the incorrect glacier polygons, erroneous DEMs. In future work, ~~better~~ improved glacier ~~inventory~~ inventories and more accurate DEMs ~~are useful for the~~will contribute to ~~improvement~~ improving ~~of~~ centerline quality. Furthermore, ~~On the other hand,~~ optimizing the automatic glacier segmentation approach, the DEM-based extraction algorithm of glacier feature lines and the centerline trade-off algorithm ~~are also~~will also likely ~~probable ways to~~ further improve the accuracy of glacier centerlines. In addition, centerline accuracy ~~it is~~will probably ~~beneficial~~ benefit ~~to~~ from further improving the classification~~clarify the~~ type of each glacier in the glacier inventory ~~for the estimates of centerline accuracy~~.


86°15'E
86°30'E
86°45'E
87°0'E
87°15'E
28°15'N
28°0'N
27°45'N
(a)
Altitude / m
3670
8848
Lengend
G_{Lmax}
L_{max}
local glaciers
Mount Qomolangma
0
10
20 km


87°50'E
88°0'E
88°10'E
88°20'E
28°0'N
27°50'N
27°40'N
27°30'N
(b)
Lengend
Kangchenjunga
G_{Lmax}
L_{max}
local glaciers
0
5
10 km
Altitude / m
8586
3725

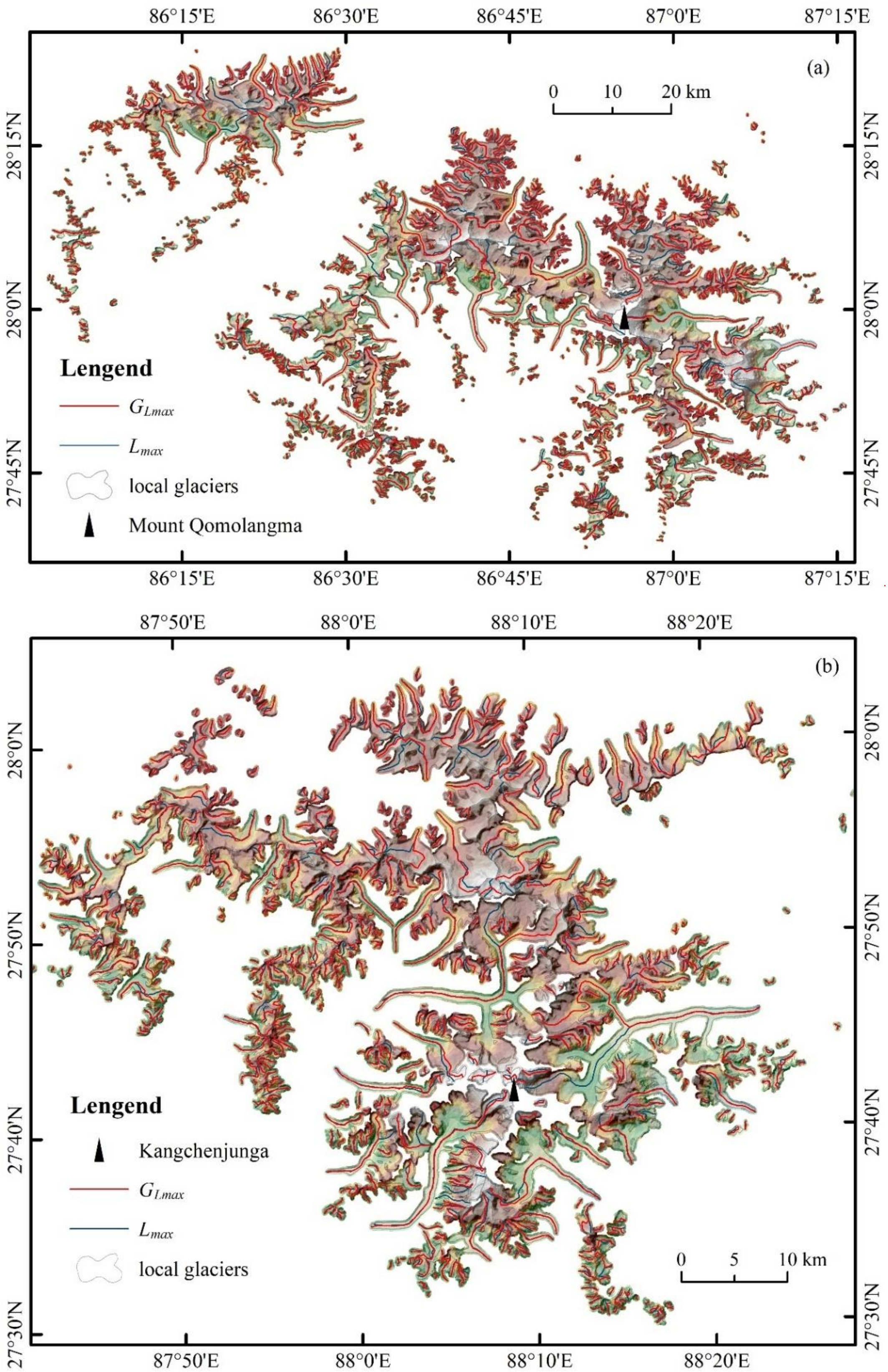


Figure 7. ~~Visible~~ Visual comparison of the longest center lines calculated in this study and by Machguth and Huss (2014) for~~. The figure shows~~ two glacier-covered regions in the Himalayas, covering Mount Qomolangma (**panel a**) and Kangchenjunga (**panel b**, the world's third highest mountain) and their surrounding areas. In ~~The~~ the background is the DEM used for the calculation.

5 Data availability

The ~~Global~~ global glacier centerline dataset (*GGCLDS*), global glacier maximum length dataset (*GGMLDS*), and other relevant datasets are available at https://doi.org/10.11922/sciencedb.01643 (Zhang and Zhang, 2022). All 17 sub-datasets of this dataset are listed in Table 5.

Table 5. Description of the sub-datasets ~~members~~ contained in this dataset.

Acronym	Data format	Data volume	Description
IGODS	*.shp	316 MB	Input glacier outline dataset
GGEDS	*.tif	3.70 GB	Global glacier elevation dataset
GGCLDS		838 MB	Global glacier centerline dataset
GGMLDS		616 MB	Global glacier maximum length dataset
GACLDS		302 MB	Global glacier accumulation region length dataset
GABLDS		358 MB	Global glacier ablation region length dataset
GOSRDS		1.16 GB	Global glacier outline segmentation result dataset
GLHPDS		11 MB	Global glacial local highest point dataset
GLPDS	*.shp	6.25 MB	Global glacial lowest point dataset
GUGODS		3.95 MB	Unsuccessful global glacier outline dataset
FGODS		119 MB	Global flawed glacier outline dataset
GGECDS		334 MB	Global glacier external contour dataset
GGBMDS		374 MB	Global glacier buffer mask dataset
MHMLDS		8.32 MB	The maximum length of Machguth and Huss in High Asia
SUP_220707		681 MB	Updated the centerlines of the repaired *FGODS*
CODES	*.py	40 KB	Related codes of data process in bulk
LOGS	*.txt	1.27 MB	Related logfiles of data process in bulk

6 Conclusions

In this study, a new global dataset ~~on the~~ of glacier centerline~~s of global glaciers~~ was constructed, and the maximum length was calculated based on the global glacier inventory (RGI v6.0) and global glacier region DEM (*GGEDS*, composed of NASADEM, ASTER GDEM v3, and Copernicus DEM). ~~In~~ A total of~~,~~ 198,137 glacier centerlines were generated, accounting for 99.74% of the ~~total number of~~ imported glaciers (*IGODS*) and 91.52% of ~~the total number of~~ the global glacier inventory. The ~~comprehensive~~ overall extraction accuracy of these glacier centerlines (*GGCLDS*) used in random self-assessment was 89.68%. The glacier length (G_{Lmax}) obtained in this study was, on average, ~~generally~~ approximately 10% longer than that of L_{max} ~~on average~~. Nevertheless, our method showed an ~~stronger~~ improved ability to obtain the maximum length, and we believe that the resulting errors were controllable. Furthermore, the preprocessing algorithm we designed accurately identified 10,764 erroneous glacier polygons from RGI v6.0, which formed the defective glacier dataset (*FGODS*).

~~A~~ The global dataset ~~containing~~ contains 17 sub-datasets ~~was generated through the above work~~, including two basic input datasets (*IGODS* and *GGEDS*), two key result datasets (*GGCLDS* and *GGMLDS*), four process datasets, six derived result datasets, and three supplementary datasets. Ice caps, nominal glaciers, and erroneous glacier polygons were eliminated from most sub-datasets ~~in this study~~, accounting for approximately 8.25% of the total RGI v6.0. The poor status of these glacier polygons ~~was not sufficient to~~ did not support the automatic extraction of glacier centerlines, which needs to be improved in future work. Inevitably, ~~there were~~ some defects in the algorithm or datasets ~~that~~ will also need to be ~~considered~~ addressed in future research. For instance, the glacial regions (R19 and R03) ~~with~~ had the worst results but were nevertheless added to the dataset to prioritize data coverage integrity. ~~It is worth noting that t~~The global glacier DEM dataset (*GGEDS*), global glacier external outline dataset (*GGECDS*), and global glacier buffer mask datasets (*GGBMDS*) cover all glaciers in RGI v6.0. Accordingly, they will help ~~researchers~~ design more efficient automated extraction algorithms to produce datasets containing all types of glacier centerlines and lengths worldwide, which is our next goal.

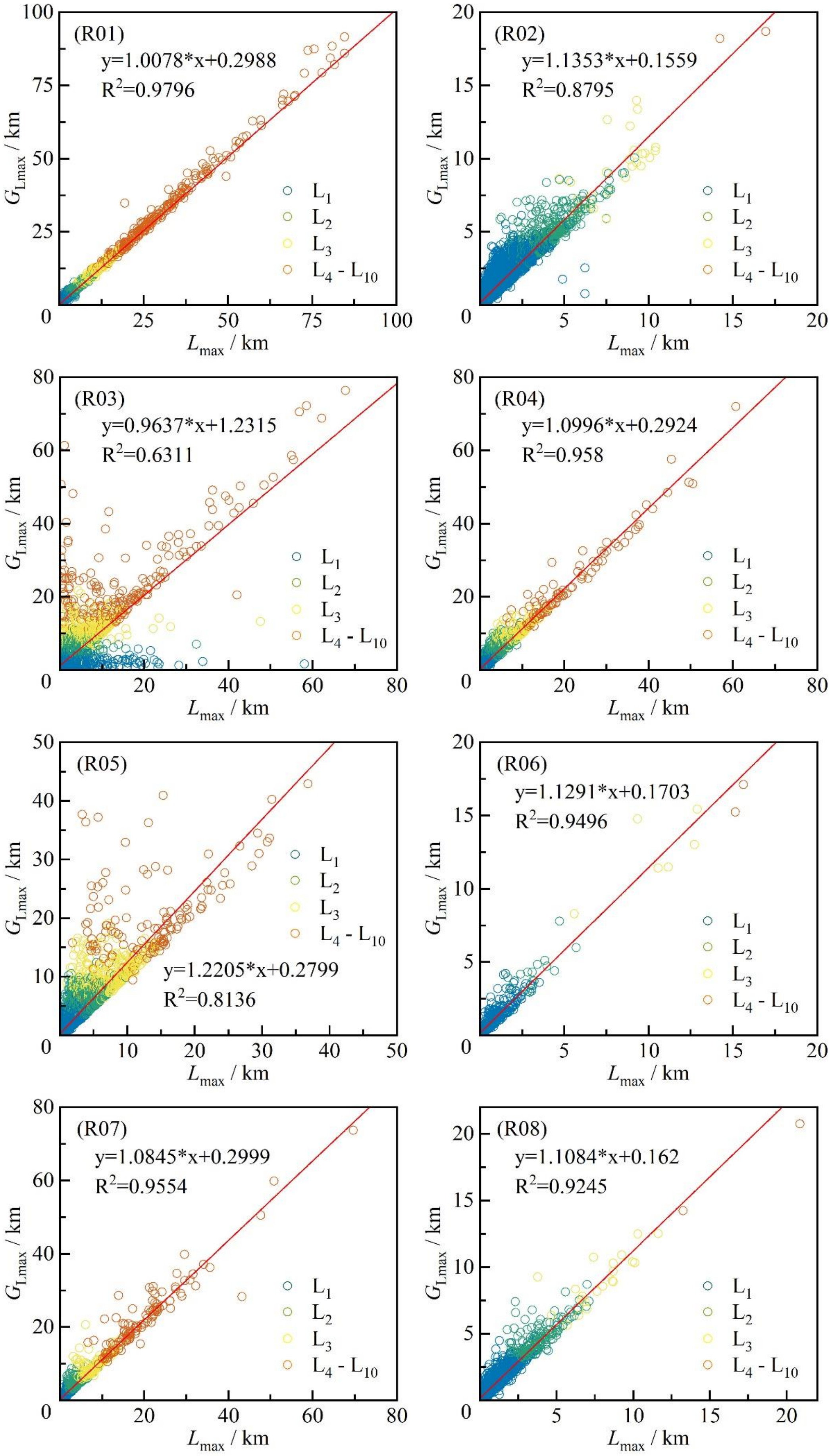
Appendix A: Model parameters resulting from the Central Asia Glacier and extended to worldwide calculations are listed in Table A1.

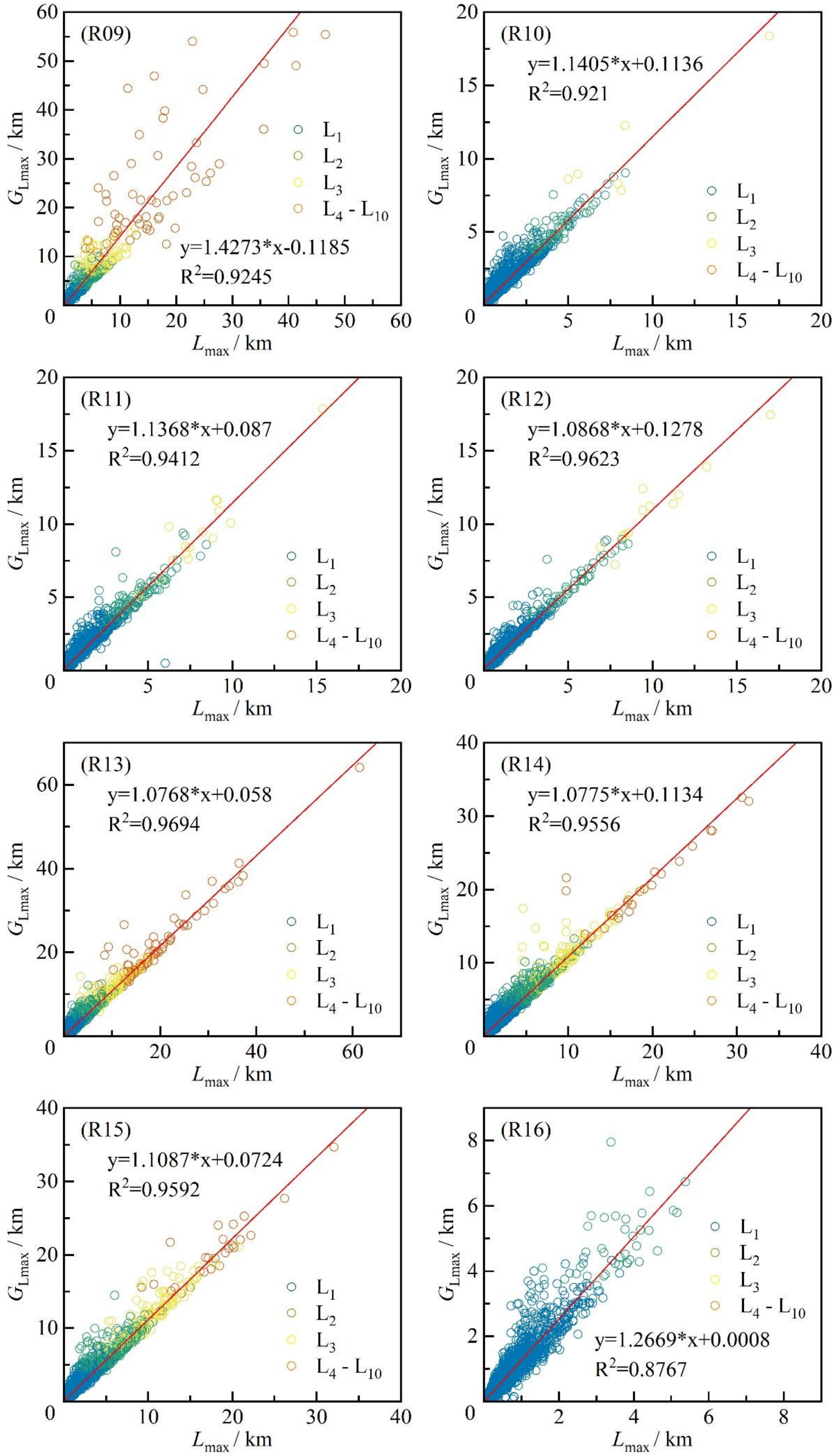
Table A1. Parameterization scheme for extracting global glacier centerlines.

Par.	Description	Value (Levels 1-10)	Unit
P_1	Maximum distance between adjacent vertexes	10	m
P_2	Buffer distance outside the glacier outline	30	m
P_3	Threshold of accumulative flow	5 - 8, 10, 20, 30, 50, 100, 200	int × 10^3
P_4	Length of the shortest auxiliary line	10 - 19	int
P_5	Length of the longest auxiliary line	2 - 11	int
P_6	Interval for searching the local highest points	50, 60, 70, 80, 90,100, 200, 300, 400, 500	count
P_7	Matching tolerance of the vertexes of polyline	0.2 ,0.2 ,0.5 ,0.5 ,1 (L_5 - L_{10})	m
P_8	Size of grid cell in Euclidean allocation	1, 5, 15, 15, 30 (L_5 - L_{10})	m
P_9	Minimum distance between the adjacent $P_{\max}$	10, 15, 30, 60, 120, 150, 200, 300, 400, 500	count
P_{10}	Smoothing tolerance of polylines	5, 10, 15, 20, 30 (L_5 - L_{10})	m
P_{11}	Length threshold of the longest auxiliary line	10190	km^2

Notes: The calculation method for each parameter is detailed in Zhang et al. (2021). $P_{\max}$ and L refer to the local highest points and grades of the glacier, respectively.

Appendix B: Comparison ~~of~~ between longest centerlines calculated in this study and by Machguth and Huss (2014) for all samples and ~~the~~ different first-order glacier regions of RGI v6.0. Linear regression of the two lengths, histogram of length ratio (R_r), and box plots of R_r for glaciers of different grades in these regions were in Figure B1, B2, and B3, respectively.


(R01)
y=1.0078*x+0.2988
R2=0.9796
(R02)
y=1.1353*x+0.1559
R2=0.8795
(R03)
y=0.9637*x+1.2315
R2=0.6311
(R04)
y=1.0996*x+0.2924
R2=0.958
(R05)
y=1.2205*x+0.2799
R2=0.8136
(R06)
y=1.1291*x+0.1703
R2=0.9496
(R07)
y=1.0845*x+0.2999
R2=0.9554
(R08)
y=1.1084*x+0.162
R2=0.9245
L1
L2
L3
L4 - L10
GL,max / km
Lmax / km


(R09)
y=1.4273*x-0.1185
R²=0.9245
(R10)
y=1.1405*x+0.1136
R²=0.921
(R11)
y=1.1368*x+0.087
R²=0.9412
(R12)
y=1.0868*x+0.1278
R²=0.9623
(R13)
y=1.0768*x+0.058
R²=0.9694
(R14)
y=1.0775*x+0.1134
R²=0.9556
(R15)
y=1.1087*x+0.0724
R²=0.9592
(R16)
y=1.2669*x+0.0008
R²=0.8767
L1
L2
L3
L4 - L10
G_Lmax / km
L_max / km

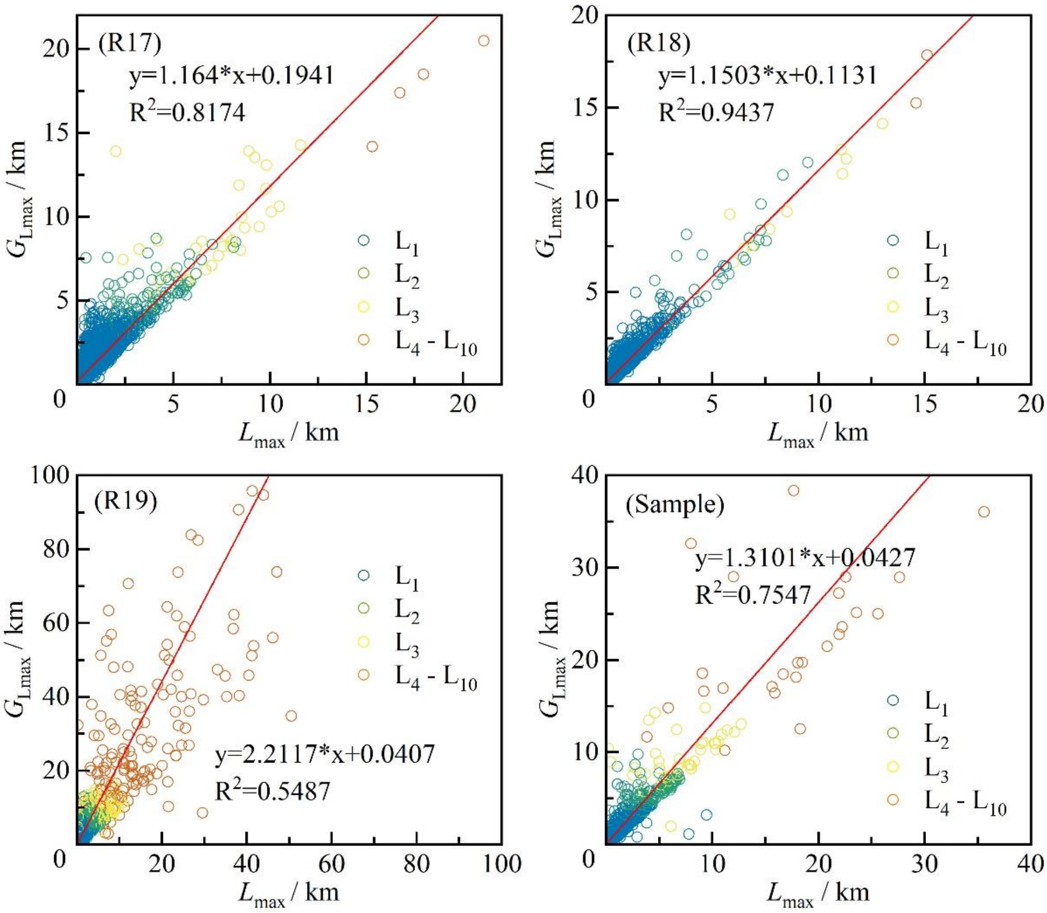
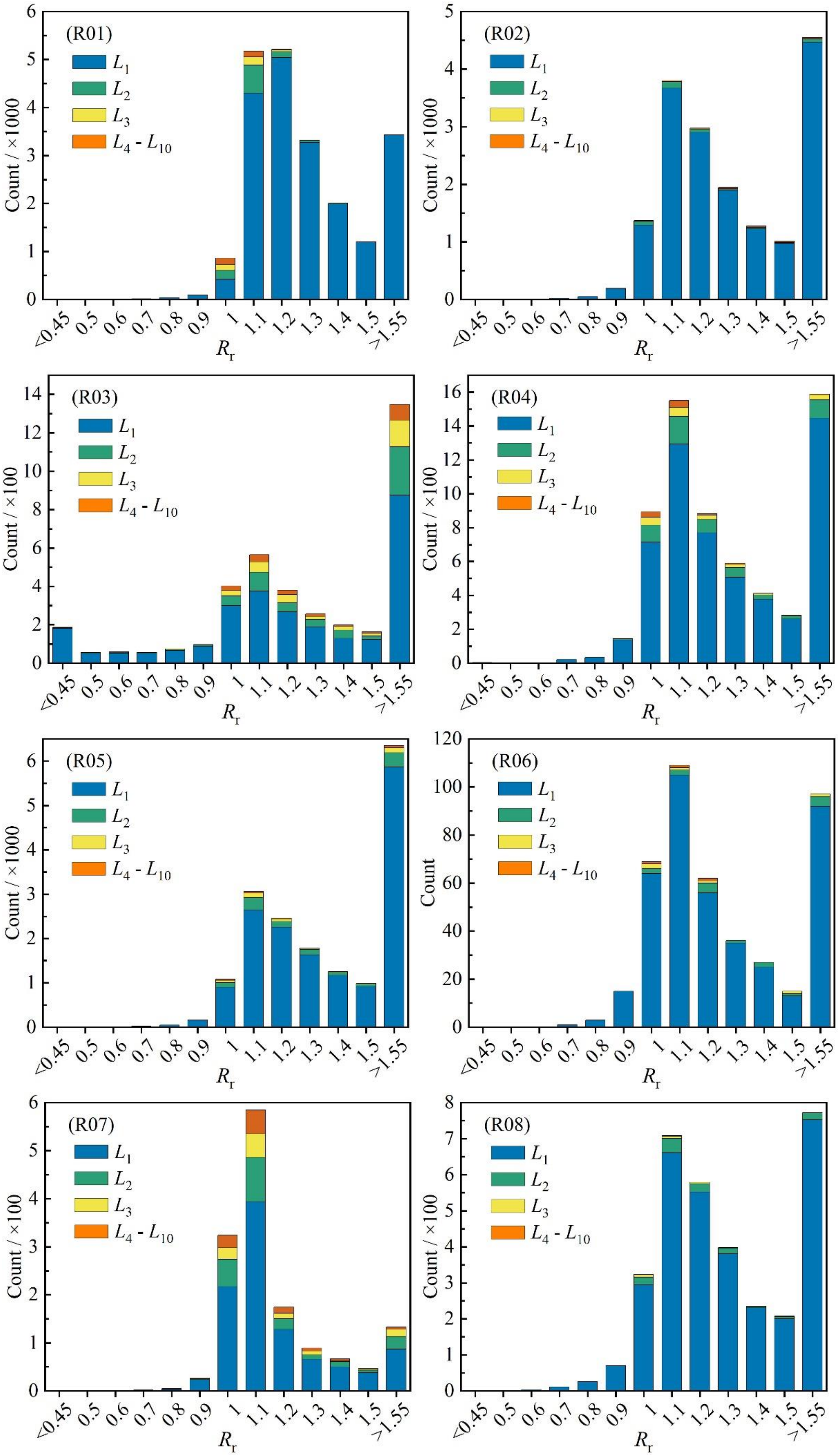
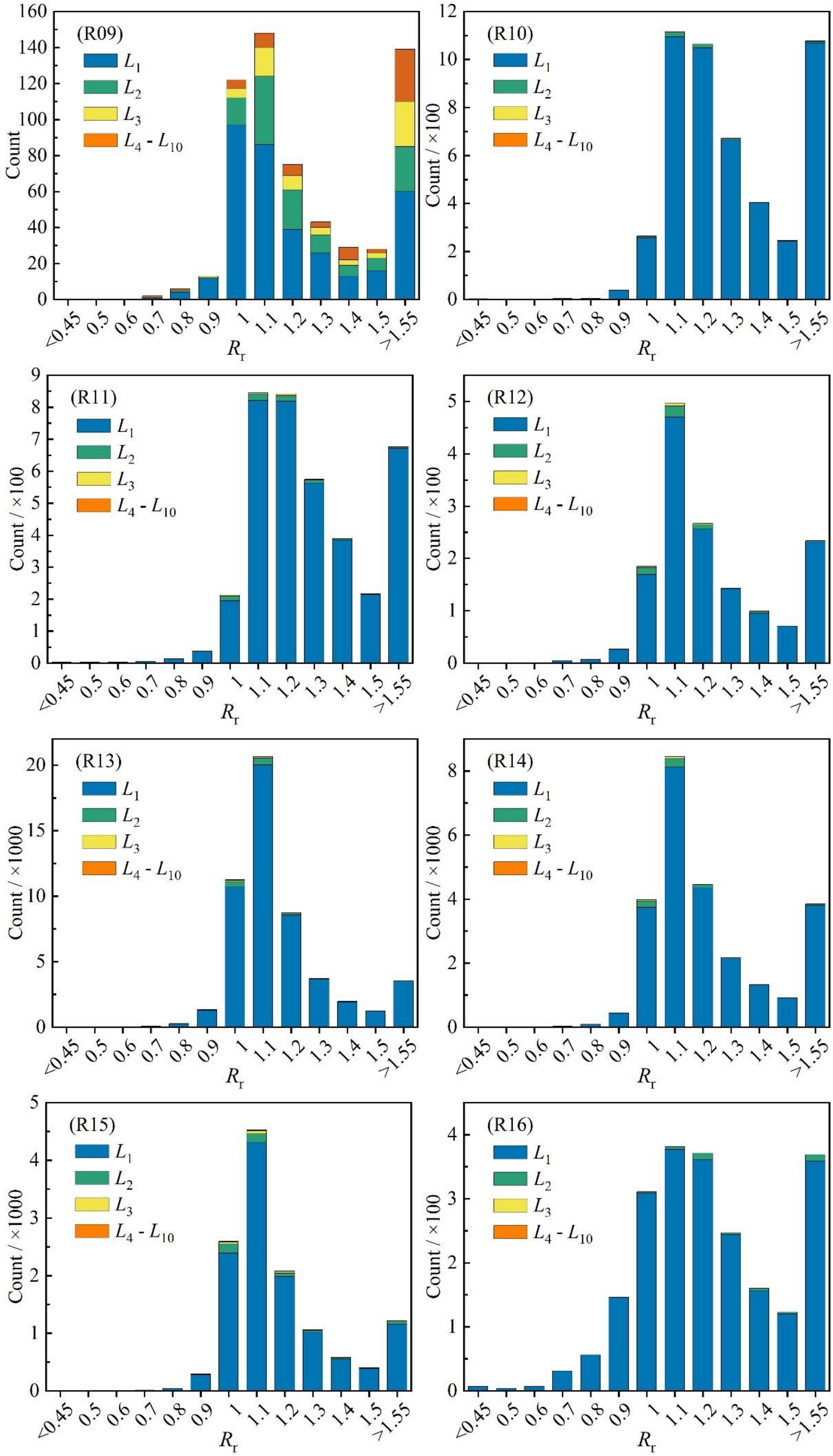


Figure B1. Linear regressions ~~in~~ for different glacier regions between glacier length (G_{Lmax}) calculated in this study and glacier length (L_{max}) calculated by Machguth and Huss (2014).


(R01)
(R02)
(R03)
(R04)
(R05)
(R06)
(R07)
(R08)
L_1
L_2
L_3
L_4 - L_{10}
Count / ×1000
Count / ×100
Count
R_r
<0.45 0.5 0.6 0.7 0.8 0.9 1 1.1 1.2 1.3 1.4 1.5 >1.55



(R09)
(R10)
(R11)
(R12)
(R13)
(R14)
(R15)
(R16)
L_1
L_2
L_3
L_4 - L_{10}
Count
Count / ×100
Count / ×1000
R_r
<0.45
0.5
0.6
0.7
0.8
0.9
1
1.1
1.2
1.3
1.4
1.5
>1.55

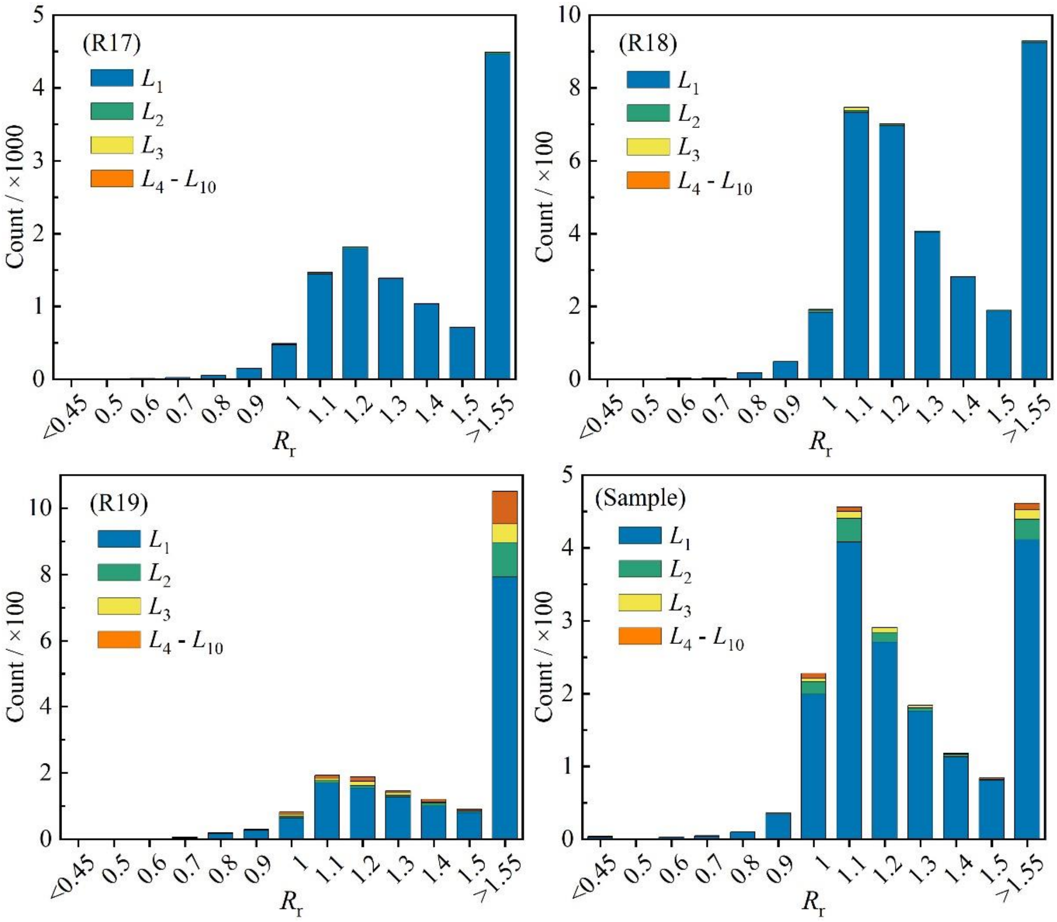
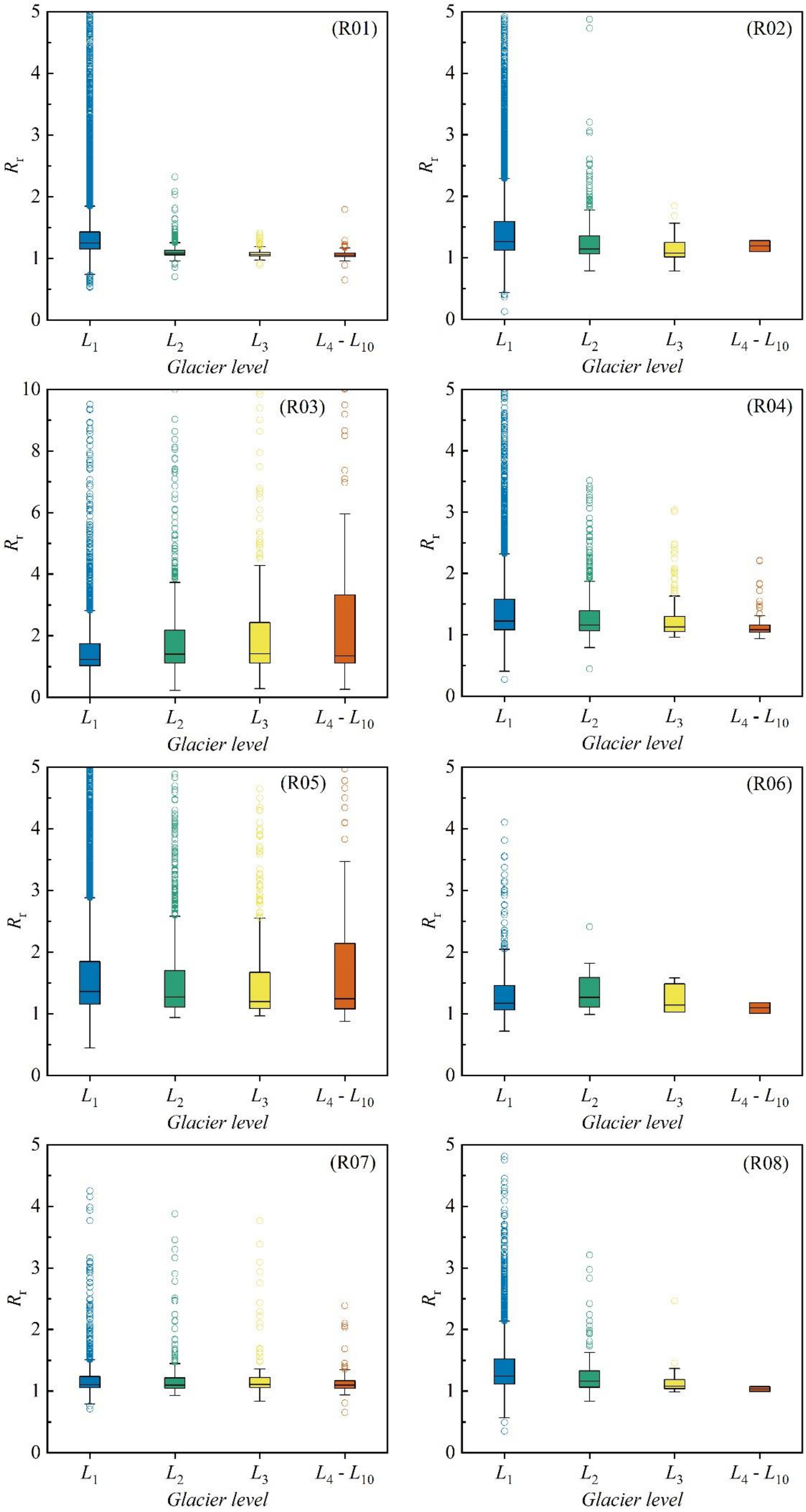
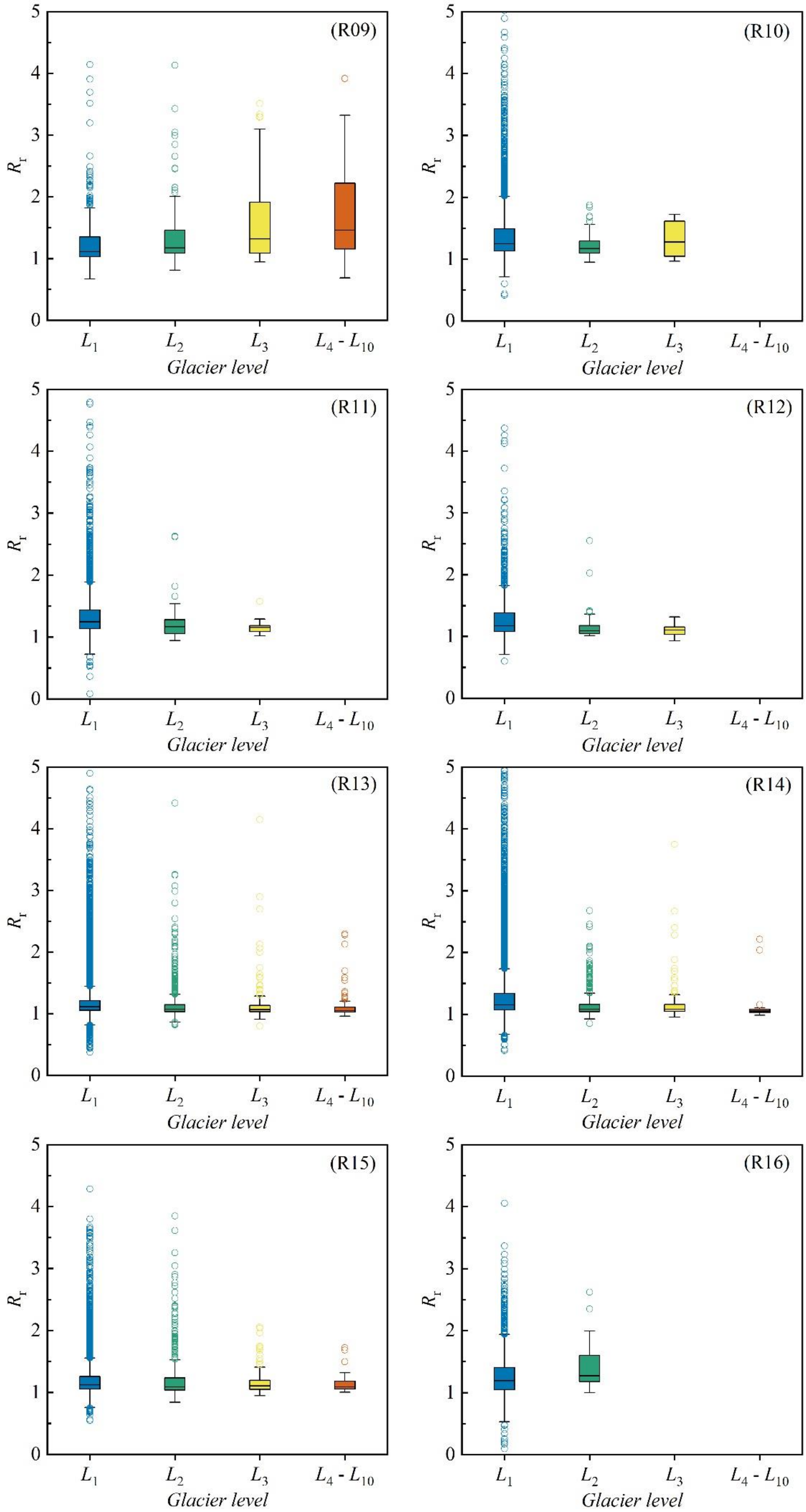


Figure B2. Histograms of the length ratio (R_r, G_{Lmax}/L_{max}) of distinct glacier grades in glacier-covered regions and all samples.


(R01)
(R02)
(R03)
(R04)
(R05)
(R06)
(R07)
(R08)
R_r
L_1
L_2
L_3
L_4 - L_{10}
Glacier level


(R09)
(R10)
(R11)
(R12)
(R13)
(R14)
(R15)
(R16)
R_r
L_1
L_2
L_3
L_4 - L_{10}
Glacier level

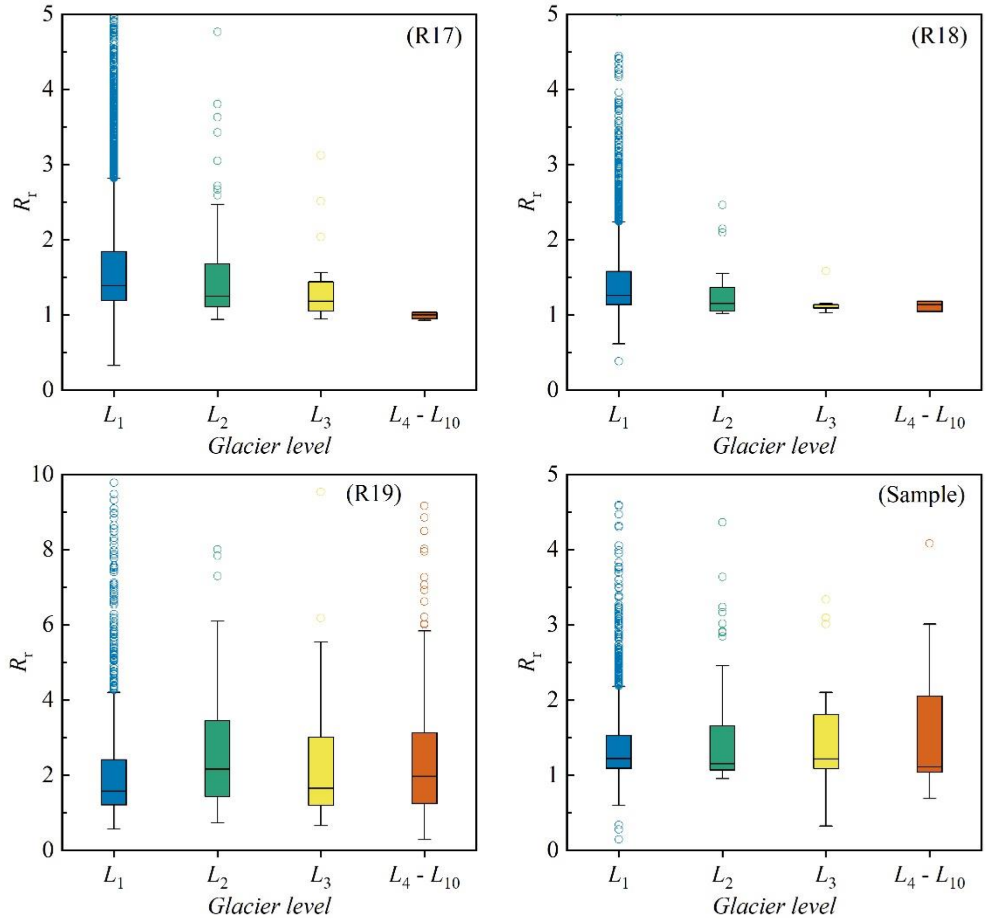


Figure B3. Box plots of length ratio (R_r, G_{Lmax}/L_{max}) of glaciers of distinct grades in every glacier-covered region and whole sample.

Supplement.

The Supplement consists of ~~two~~three parts: (1) ‘GlacierCenterlines_Py27’ (version 5.2.1), the updated automatic extraction tool of glacier centerlines in this study, which fixed some defects compared with version 5.2.0 (https://doi.org/10.5194/tc-151955-2021-supplement). (2) ‘Other_parameters_T1.txt’ is the parameter file for extracting the global glacier centerlines. (3) The parameter files (format: *.prj) of the projected coordinate systems (Albers) of 19 glacier regions used in this study, which can be viewed with a text viewer and directly imported into ArcGIS software for further analysis.

Author contributions.

All authors contributed to writing and editing the manuscript. DZ processed the data, performed all calculations, created all figures, and wrote most of the manuscript. SZ contributed significantly to the development of the analyses, figures, and writing. XY contributed to the development of the data production strategy and writing. GZ and WL contributed to the initial data production. SW participated in writing Chapter 4.

Competing interests.

The authors declare that they have no conflict of interest.

Acknowledgments.

We thank editors, two reviewers and community scholars for their valuable comments that improved the manuscript. The authors would especially like to thank GLIMS for releasing the RGI v6.0 (http://www.glims.org/RGI/randolph.html, last accessed: November 15, 2021), LP DAAC for releasing the NASADEM (https://lpdaac.usgs.gov/news/release-nasadem-data-products/, last accessed: November 17, 2021), METI and NASA for jointly releasing the ASTER GDEM v3 (https://lpdaac.usgs.gov/news/nasa-and-meti-release-aster-global-dem-version-3/, last accessed: November 17, 2021), and the European Space Agency (ESA) for providing the Copernicus DEM (https://spacedata.copernicus.eu/web/cscda/cop-dem-faq, last accessed: November 17, 2021). This work ~~is~~ would not have been possible without the support of open-access data.

Financial support.

This research was funded by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP) (grant number: 2019QZKK020109) and China National Natural Science Foundation (grant numbers: 41730751, 42171124).

References

Abrams, M., Crippen, R., and Fujisada, H.: ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD), Remote Sensing, 12, https://doi.org/10.3390/rs12071156, 2020.

Aciego, S. M., Stevenson, E. I., and Arendt, C. A.: Climate versus geological controls on glacial meltwater micronutrient production in southern Greenland, Earth and Planetary Science Letters, 424, 51-58, https://doi.org/10.1016/j.epsl.2015.05.017, 2015.

Carabajal, C. C. and Boy, J. P.: Evaluation of Aster Gdem V3 Using Icesat Laser Altimetry, ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B4, 117-124, https://doi.org/10.5194/isprsarchives-XLI-B4-117-2016, 2016.

Carrera-Hernández, J. J.: Not all DEMs are equal: An evaluation of six globally available 30 m resolution DEMs with geodetic benchmarks and LiDAR in Mexico, Remote Sensing of Environment, 261, https://doi.org/10.1016/j.rse.2021.112474, 2021.

Cazenave, A.: Global sea-level budget 1993–present, Earth System Science Data, 10, 1551-1590, https://doi.org/10.5194/essd-10-1551-2018, 2018.

Fan, Y., Ke, C.-Q., and Shen, X.: A new Greenland digital elevation model derived from ICESat-2 during 2018–2019, Earth System Science Data, 14, 781-794, https://doi.org/10.5194/essd-14-781-2022, 2022.

Farinotti, D., Huss, M., Fürst, J. J., Landmann, J., Machguth, H., Maussion, F., and Pandit, A.: A consensus estimate for the ice thickness distribution of all glaciers on Earth, Nature Geoscience, 12, 168-173, https://doi.org/10.1038/s41561-019-0300-3, 2019.

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., and Alsdorf, D.: The Shuttle Radar Topography Mission, Reviews of Geophysics, 45, 1-33, https://doi.org/10.1029/2005rg000183, 2007.

Franke, S., Jansen, D., Binder, T., Paden, J. D., Dörr, N., Gerber, T. A., Miller, H., Dahl-Jensen, D., Helm, V., Steinhage, D., Weikusat, I., Wilhelms, F., and Eisen, O.: Airborne ultra-wideband radar sounding over the shear margins and along flow lines at the onset region of the Northeast Greenland Ice Stream, Earth System Science Data, 14, 763-779, https://doi.org/10.5194/essd-14-763-2022, 2022.

Gao, Y. P., Yao, X. J., Liu, S. Y., Qi, M. M., Gong, P., An, L. N., Li, X. F., and Duan, H. Y.: Methods and future trend of ice volume calculation of glacier, Arid Land Geography, 41, 1204-1213, https://doi.org/10.12118/j.issn.1000-6060.2018.06.08, 2018.

Hansen, K., Hasenstab, K., and Schwartzman, A.: Estimating Mountain Glacier Flowlines by Local Linear Regression Gradient Descent, IEEE Transactions on Geoscience and Remote Sensing, 59, 10022-10034, https://doi.org/10.1109/tgrs.2020.3035513, 2020.

Heid, T. and Kääb, A.: Repeat optical satellite images reveal widespread and long term decrease in land-terminating glacier speeds, The Cryosphere, 6, 467-478, https://doi.org/10.5194/tc-6-467-2012, 2012.

Herla, F., Roe, G. H., and Marzeion, B.: Ensemble statistics of a geometric glacier length model, Annals of Glaciology, 58, 130-135, https://doi.org/10.1017/aog.2017.15, 2017.

Herreid, S. and Pellicciotti, F.: The state of rock debris covering Earth's glaciers, Nature Geoscience, 13, 621-627, https://doi.org/10.1038/s41561-020-0615-0, 2020.

Howat, I. M., Porter, C., Smith, B. E., Noh, M.-J., and Morin, P.: The Reference Elevation Model of Antarctica, The Cryosphere, 13, 665-674, https://doi.org/10.5194/tc-13-665-2019, 2019.

Hugonnet, R., McNabb, R., Berthier, E., Menounos, B., Nuth, C., Girod, L., Farinotti, D., Huss, M., Dussaillant, I., Brun, F., and Kaab, 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.

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., and 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, 2019.

Ji, Q., Yang, T.-b., He, Y., Qin, Y., Dong, J., and Hu, F.-s.: A simple method to extract glacier length based on Digital Elevation Model and glacier boundaries for simple basin type glacier, Journal of Mountain Science, 14, 1776-1790, https://doi.org/10.1007/s11629-016-4243-5, 2017.

Kääb, A., Jacquemart, M., Gilbert, A., Leinss, S., Girod, L., Huggel, C., Falaschi, D., Ugalde, F., Petrakov, D., Chernomorets, S., Dokukin, M., Paul, F., Gascoin, S., Berthier, E., and Kargel, J. S.: Sudden large-volume detachments of low-angle mountain glaciers – more frequent than thought?, The Cryosphere, 15, 1751-1785, https://doi.org/10.5194/tc-15-1751-2021, 2021.

Kienholz, C., Hock, R., and Arendt, A. A.: A new semi-automatic approach for dividing glacier complexes into individual glaciers, Journal of Glaciology, 59, 925-937, https://doi.org/10.3189/2013JoG12J138, 2013.

Kienholz, C., Rich, J. L., Arendt, A. A., and Hock, R.: A new method for deriving glacier centerlines applied to glaciers in Alaska and northwest Canada, The Cryosphere, 8, 503-519, https://doi.org/10.5194/tc-8-503-2014, 2014.

Le Bris, R. and Paul, F.: An automatic method to create flow lines for determination of glacier length: A pilot study with Alaskan glaciers, Computers & Geosciences, 52, 234-245, https://doi.org/10.1016/j.cageo.2012.10.014, 2013.

Le Moine, N. and Gsell, P.-S.: A graph-based approach to glacier flowline extraction: An application to glaciers in Switzerland, Computers & Geosciences, 85, 91-101, https://doi.org/10.1016/j.cageo.2015.09.010, 2015.

Leclercq, P. W. and Oerlemans, J.: Global and hemispheric temperature reconstruction from glacier length fluctuations, Climate Dynamics, 38, 1065-1079, https://doi.org/10.1007/s00382-011-1145-7, 2011.

Leclercq, P. W., Oerlemans, J., Basagic, H. J., Bushueva, I., Cook, A. J., and Le Bris, R.: A data set of worldwide glacier length fluctuations, The Cryosphere, 8, 659-672, https://doi.org/10.5194/tc-8-659-2014, 2014.

Li, H., Ng, F., Li, Z., Qin, D., and Cheng, G.: An extended "perfect-plasticity" method for estimating ice thickness along the flow line of mountain glaciers, Journal of Geophysical Research: Earth Surface, 117, n/a-n/a, https://doi.org/10.1029/2011jf002104, 2012.

Li, X., Ding, Y., Hood, E., Raiswell, R., Han, T., He, X., Kang, S., Wu, Q., Yu, Z., Mika, S., Liu, S., and Li, Q.: Dissolved Iron Supply from Asian Glaciers: Local Controls and a Regional Perspective, Global Biogeochemical Cycles, 33, 1223-1237, https://doi.org/10.1029/2018gb006113, 2019.

Li, Y., Li, F., Shangguan, D., and Ding, Y.: A new global gridded glacier dataset based on the Randolph Glacier Inventory version 6.0, Journal of Glaciology, 67, 773-776, https://doi.org/10.1017/jog.2021.28, 2021.

Lüthi, M. P., Bauder, A., and Funk, M.: Volume change reconstruction of Swiss glaciers from length change data, Journal of Geophysical Research, 115, https://doi.org/10.1029/2010jf001695, 2010.

Machguth, H. and Huss, M.: The length of the world's glaciers – a new approach for the global calculation of center lines, The Cryosphere, 8, 1741-1755, https://doi.org/10.5194/tc-8-1741-2014, 2014.

Mankoff, K. D., Fettweis, X., Langen, P. L., Stendel, M., Kjeldsen, K. K., Karlsson, N. B., Noël, B., van den Broeke, M. R., Solgaard, A., Colgan, W., Box, J. E., Simonsen, S. B., King, M. D., Ahlstrøm, A. P., Andersen, S. B., and Fausto, R. S.: Greenland ice sheet mass balance from 1840 through next week, Earth System Science Data, 13, 5001-5025, https://doi.org/10.5194/essd-13-5001-2021, 2021.

Maussion, F., Butenko, A., Champollion, N., Dusch, M., Eis, J., Fourteau, K., Gregor, P., Jarosch, A. H., Landmann, J., Oesterle, F., Recinos, B., Rothenpieler, T., Vlug, A., Wild, C. T., and Marzeion, B.:

The Open Global Glacier Model (OGGM) v1.1, Geoscientific Model Development, 12, 909-931, https://doi.org/10.5194/gmd-12-909-2019, 2019.

Melkonian, A. K., Willis, M. J., and Pritchard, M. E.: Satellite-derived volume loss rates and glacier speeds for the Juneau Icefield, Alaska, Journal of Glaciology, 60, 743-760, https://doi.org/10.3189/2014JoG13J181, 2017.

Noel, B., Jakobs, C. L., van Pelt, W. J. J., Lhermitte, S., Wouters, B., Kohler, J., Hagen, J. O., Luks, B., Reijmer, C. H., van de Berg, W. J., and van den Broeke, M. R.: Low elevation of Svalbard glaciers drives high mass loss variability, Nat Commun, 11, 4597, https://doi.org/10.1038/s41467-020-18356-1, 2020.

Oerlemans, J.: A flowline model for Nigardsbreen, Norway: projection of future glacier length based on dynamic calibration with the historic record, Annals of Glaciology, 24, 382-389, https://doi.org/10.1017/S0260305500012489 1997.

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

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.

Radić, V. and Hock, R.: Regional and global volumes of glaciers derived from statistical upscaling of glacier inventory data, Journal of Geophysical Research, 115, https://doi.org/10.1029/2009jf001373, 2010.

RGI Consortium: Randolph Glacier Inventory – A Dataset of Global Glacier Outlines: Version 6.0: Technical Report, Global Land Ice Measurements from Space, Colorado, USA, 10.7265/N5-RGI-60, 2017.

Scherler, D., Wulf, H., and Gorelick, N.: Global Assessment of Supraglacial Debris-Cover Extents, Geophysical Research Letters, 45, 11,798-711,805, https://doi.org/10.1029/2018gl080158, 2018.

Schiefer, E., Menounos, B., and Wheate, R.: An inventory and morphometric analysis of British Columbia glaciers, Canada, Journal of Glaciology, volume 54, 551-560, 2008.

Shukla, A., Garg, S., Mehta, M., Kumar, V., and Shukla, U. K.: Temporal inventory of glaciers in the Suru sub-basin, western Himalaya: impacts of regional climate variability, Earth System Science Data, 12, 1245-1265, https://doi.org/10.5194/essd-12-1245-2020, 2020.

Shukla, T. and Sen, I. S.: Preparing for floods on the Third Pole, Science, 372, 232-234, https://doi.org/10.1126/science.abh3558, 2021.

Solgaard, A., Kusk, A., Merryman Boncori, J. P., Dall, J., Mankoff, K. D., Ahlstrøm, A. P., Andersen, S. B., Citterio, M., Karlsson, N. B., Kjeldsen, K. K., Korsgaard, N. J., Larsen, S. H., and Fausto, R. S.: Greenland ice velocity maps from the PROMICE project, Earth System Science Data, 13, 3491-3512, https://doi.org/10.5194/essd-13-3491-2021, 2021.

Sommer, C., Malz, P., Seehaus, T. C., Lippl, S., Zemp, M., and Braun, M. H.: Rapid glacier retreat and downwasting throughout the European Alps in the early 21(st) century, Nat Commun, 11, 3209, https://doi.org/10.1038/s41467-020-16818-0, 2020.

Stuart-Smith, R. F., Roe, G. H., Li, S., and Allen, M. R.: Increased outburst flood hazard from Lake Palcacocha due to human-induced glacier retreat, Nature Geoscience, 14, 85-90, https://doi.org/10.1038/s41561-021-00686-4, 2021.

Sugiyama, S., Bauder, A., Zahno, C., and Funk, M.: Evolution of Rhonegletscher, Switzerland, over the past 125 years and in the future : application of an improved flowline model, Annals of Glaciology, 46, 268-274, 2007.

Thogersen, K., Gilbert, A., Schuler, T. V., and Malthe-Sorenssen, A.: Rate-and-state friction explains glacier surge propagation, Nat Commun, 10, 2823, https://doi.org/10.1038/s41467-019-10506-4, 2019.

Uuemaa, E., Ahi, S., Montibeller, B., Muru, M., and Kmoch, A.: Vertical Accuracy of Freely Available Global Digital Elevation Models (ASTER, AW3D30, MERIT, TanDEM-X, SRTM, and NASADEM), Remote Sensing, 12, https://doi.org/10.3390/rs12213482, 2020.

Vargo, L. J., Anderson, B. M., Dadić, R., Horgan, H. J., Mackintosh, A. N., King, A. D., and Lorrey, A. M.: Anthropogenic warming forces extreme annual glacier mass loss, Nature Climate Change, 10, 856-861, https://doi.org/10.1038/s41558-020-0849-2, 2020.

Winsvold, S. H., Andreassen, L. M., and Kienholz, C.: Glacier area and length changes in Norway from repeat inventories, The Cryosphere, 8, 1885-1903, https://doi.org/10.5194/tc-8-1885-2014, 2014.

Wu, K., Liu, S., Jiang, Z., Liu, Q., Zhu, Y., Yi, Y., Xie, F., Ahmad Tahir, A., and Saifullah, M.: Quantification of glacier mass budgets in the Karakoram region of Upper Indus Basin during the early twenty-first century, Journal of Hydrology, 603, https://doi.org/10.1016/j.jhydrol.2021.127095, 2021.

Xia, W.: An Automatic Extraction Method of Glacier Length Based on Voronoi Algorithm - A Pilot Study in the Sanjiangyuan Region, College of Urban and Environmental Science, Northwest University, Xi'an, Shannxi, 2020.

Yang, B. Y., Zhang, L. X., Gao, Y., Xiang, Y., Mou, N. X., and Suo, L. D. B.: An integrated method of glacier length extraction based on Gaofen satellite data, Journal of Glaciology and Geocryology, 38, 1615-1623, https://doi.org/10.7522/j.issn.1000-0240.2016.0189, 2016.

Yao, X. J., Liu, S. Y., Zhu, Y., Gong, P., An, L. N., and Li, X. F.: Design and implementation of an automatic method for deriving glacier centerlines based on GIS, Journal of Glaciology and Geocryology, 37, 1563-1570, https://doi.org/10.7522/j.issn.1000-0240.2015.0173, 2015.

Zemp, M., Huss, M., Thibert, E., Eckert, N., McNabb, R., Huber, J., Barandun, M., Machguth, H., Nussbaumer, S. U., Gartner-Roer, I., Thomson, L., Paul, F., Maussion, F., Kutuzov, S., and Cogley, J. G.: Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016, Nature, 568, 382-386, https://doi.org/10.1038/s41586-019-1071-0, 2019.

Zhang, B., Wang, Z., An, J., Liu, T., and Geng, H.: A 30-year monthly 5 km gridded surface elevation time series for the Greenland Ice Sheet from multiple satellite radar altimeters, Earth System Science Data, 14, 973-989, https://doi.org/10.5194/essd-14-973-2022, 2022.

Zhang, D. and Zhang, S.: A new global dataset of mountain glacier centerline and length, Science Data Bank, https://doi.org/10.11922/sciencedb.01643, 2022.

Zhang, D., Yao, X., Duan, H., Liu, S., Guo, W., Sun, M., and Li, D.: A new automatic approach for extracting glacier centerlines based on Euclidean allocation, The Cryosphere, 15, 1955-1973, https://doi.org/10.5194/tc-15-1955-2021, 2021.

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., and Stoffel, M.: Increasing risk of glacial lake outburst floods from future Third Pole deglaciation, Nature Climate Change, 11, 411-417, https://doi.org/10.1038/s41558-021-01028-3, 2021.

Zhou, S., Yao, X., Zhang, D., Zhang, Y., Liu, S., and Min, Y.: Remote Sensing Monitoring of Advancing and Surging Glaciers in the Tien Shan, 1990–2019, Remote Sensing, 13,

https://doi.org/10.3390/rs13101973, 2021a.
Zhou, Y., Li, X., Zheng, D., Li, Z., An, B., Wang, Y., Jiang, D., Su, J., and Cao, B.: The joint driving effects of climate and weather changes caused the Chamoli glacier-rock avalanche in the high altitudes of the India Himalaya, Science China Earth Sciences, 64, 1909-1921, https://doi.org/10.1007/s11430-021-9844-0, 2021b.