Responses to Editor Decision to manuscript TC-2020-294

Dear Editor,

Thank you very much for your valuable comments to improve this manuscript. We responded point by point to each comment as listed below, along with a clear indication of the location of the revision.

If you have any queries, please don't hesitate to contact us at the address below. Looking forward to hearing from you.

Thank you and best regards.
Sincerely,

Dahong Zhang
Email: zhangdh_yx@163.com

Please Notes: Text in BLACK is the editor's comments and our responses are marked in BLUE. In addition, the notation used to locate the changes first defines the page number, then the line number(s). For example, **P4L15** means that the described modification to the manuscript can be found on the 15th line on the 4th page in the track-changes file.

Comments:

L20: were more superior => were superior

Response: It has been modified. **(P1L20)**

L20: provided => provides

Response: It has been modified. **(P1L20)**

L24: Glacier is => Glaciers are

Response: It has been modified. **(P2L24)**

L35: glacier centerline => the glacier centerline

Response: It has been modified. **(P2L36)**

L37: one-dimensional glacier model => one-dimensional glacier models

Response: It has been modified. **(P2L38)**

L103: was given => is given

Response: It has been modified. **(P7L108)**

L181: simple glacier => simple glaciers

Response: It has been modified. **(P16L194)**

L181: compound glacier => compound glaciers

Response: It has been modified. **(P16L194)**

In addition, we found and then changed two mistakes: (i) the wrong Eq.1 has been corrected **(P7L109)** and (ii) the default value of P_9 in Table 1 was not up-to-date and has been updated **(P8L114)**.

Responses to Reviewer #1 to manuscript TC-2020-294

Thanks for your helpful comments to improve this manuscript.

Please Notes: Text in BLACK is the reviewer' comments and our responses are in BLUE. In addition, the notation used to locate the changes first defines the page number, then the line number(s). For example, **P4L15** means that the described modification to the manuscript can be found on the 15th line of the 4th page in the track-changes file.

Specific Comments:

Title: I think the title should point out that the author's approach is different from other studies, for example "base on …".

Response: We changed the title to "A new automatic approach for extracting glacier centerlines based on Euclidean allocation", which can reflect that our approach is different from other studies. **(P1L1)**

*(P1L20) "the largest length" -> "the longest length" or "the maximum length".

Response: We revised it to "the longest length". **(P1L21)**

*(P2L30) "Alternatively" might be "Therefore".

Response: It has been modified. **(P2L31)**

*(P2L31) Please add a sentence to explain the role of the two concepts of glacier axis and glacier centerline and their relationship with glacier flowline.

Response: We have further explained the related concepts involved in the question:

Glacier centerline is a central line close to the main flowline of glacier, which can be acquired base on glacier axis and be used to simulate the glacier flowline. **(P2L32)**

In addition, explanations of the relationship of some related concepts are shown in Figure A1.

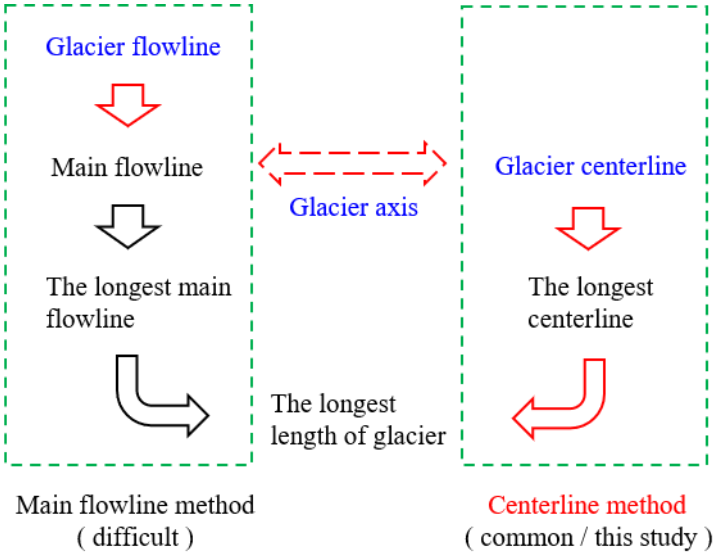


Figure A1: The schematic of the relationship of some related concepts.

*(P2L45) Delete "automatic". It is too early to mention the importance of automatic extraction algorithm because it cannot be illustrated above.

Response: This word was deleted. **(P3L48)**

*(P2L46- P3L60) This section seems not make clear the challenge of current glacier centerlines extraction.

Response: So far, the biggest challenge for glacier centerline extraction is still automation. In the past, glacier length was determined manually in a laborious way. In recent years, several authors mentioned in the section have tried to extract the centerlines in batches, however, the results are not satisfactory. In this regard, we added the following summary:

So, the current biggest challenge is still the implementation of automation extraction of glacier centerline and the acquirement of more information about glacier length. **(P3L61)**

*(P4L80) The provincial boundary is not obvious to see in Figure 1, and the number of map's scale is best such as 100, 200, 500, 1000 km.

Response: The Figure 1 was remapped. **(P5L84)**

*(P5L85) "arcpy" -> "ArcPy"

Response: It has been modified. **(P5L92, P17L209)**

*(P5L95) Make some parameters clear, for example, P_G, A, P, A_G.

Response: We rewrote acronym of each parameter to clarify their meanings, listed in the Appendix A. The relevant parameters are explained as follows:

Table A1 The list of main acronyms in this study. (P38L454)

Acronyms	Description
A_t	The given area of an equilateral triangle
A_g	The polygon's area of the glacier's outer boundary
A_l	The final auxiliary line
A_r	The ridgelines of the glacier surface
G_{br}	The bare rock in glacier
G_{fcl}	The final glacier centerline
G_{fl}	The feature lines of glacier surface
G_{cl}	The original glacier centerline
G_{Labl}	The length in the ablation region of the glacier
G_{Lacc}	The length in the accumulation region of the glacier
G_{Lmax}	The longest length of the glacier
G_{Lmean}	The average length of the glacier
G_{pl}	The polyline of the outer boundary of the glacier
G_{po}	The polygon of the outer boundary of the glacier
L_{max}	The longest glacier length of RGI v6.0
D_L	The difference between G_{Lmax} and L_{max}
P_t	The given perimeter of an equilateral triangle
P_g	The perimeter of the glacier's outer boundary
P_{max}	The local highest point of glacier outline
P_{min}	The lowest point of glacier outline
RGI	The Randolph Glacier Inventory
SCGI	The Second Chinese Glacier Inventory
Z_{med}	The median elevation of the glacier

Please note that in Table A1, the parameters A_t, P_t, A_g and P_g correspond to A, P, A_G and P_G in the manuscript, respectively. The four parameters involved in the comment are explained as follows:

A_t *(A)*: The given area of an equilateral triangle;

P_t *(P)*: The given perimeter of an equilateral triangle; **(P7L103)**

A_g *(A_G)*: The polygon's area of the glacier's outer boundary;

P_g *(P_G)*: The perimeter of the glacier's outer boundary. **(P7L104)**

*(P5L101) Author should explain where the formula 1-3 comes from?

Response: Formula 1-3 are proposed in this study. Formula 1 expresses the relationship between the perimeter and the area of an equilateral triangle. Formula 2 represents the method for determining the glacier grade in this study. Formula 3 expresses the proportional coefficients for determining the relevant parameters of different levels based on the aspect ratio of the equilateral triangle corresponding to the area of glacier's outer outline.

The main basis is the classification of glacier scale and the scale of glaciers is divided into 12 levels in the SCGI. The values of classification intervals are 0.1, 0.5, 1, 2, 5, 10, 20, 50, 100, 200 and 300 km^2. Combined with the sensitivity of the algorithm to each grade of glaciers during the experiment, this research divides the glaciers into 5 grades (interval value: 1, 5, 20 and 50 km^2). In the experiment, we also found that when the outer perimeter of glaciers (P_g) of the same scale differs greatly, the extraction results of glacier centerlines differ greatly. In addition, the shape of alpine glacier resembles a triangle. Therefore, the P_g was considered in the glaciers' classification in this study, and the classification results were fine-tuned according to the above three formulas with reference to the values of the SCGI's grading intervals.

*(P7L124) Some word's fonts in Figure 2 are not uniform. Please check. In addition, I have a question, did DEM need preprocessing? Such as filling.

Response: It has been checked that Figure 2 includes two fonts. The main body of the flow chart uses the Times New Roman (nine pounds) and the module name uses the Microsoft Elegant Black (10 pounds).

Figure 2 briefly shows the processing for DEM. The actual processing includes a series of preprocessing such as clipping, filling, condition selection, focus statistics, and inverse terrain calculations.

*(P8L134) median elevation Z_{min} -> median elevation Z_{med}. Please check the full text.

Response: It has been modified. **(P11L141, P34L401, P34L402)**

*(P9L144) "the material flow" -> "the mass flow".

Response: It has been modified. **(P12L152)**

*(P9L147) As for post-processing, please introduce in more detail.

Response: Firstly, the ridgelines of the glacier surface (A_r) were obtained by clipping the ridge lines using G_{po}. The set of all possible starting points of auxiliary lines was gained by intersecting A_r with G_{pl}. Then, the ridgeline clusters connected to each starting point were achieved and marked by traversing the point set. The number of auxiliary lines was initially determined. Finally, the longest length of each auxiliary line was calculated by adopting the critical path algorithm. The final auxiliary lines (A_l) were obtained by screening all auxiliary lines using the three parameters of P_4, P_5 and P_{11}.

The related processing methods are explained in the P10L153- P11L159 of the manuscript. The processing objects (the disconnected lines and the abnormal lines) of steps i and ii are shown in the discussion section (Figure 15). The post-processing of steps iii, iv and v are shown in Figure A2 in more detail.

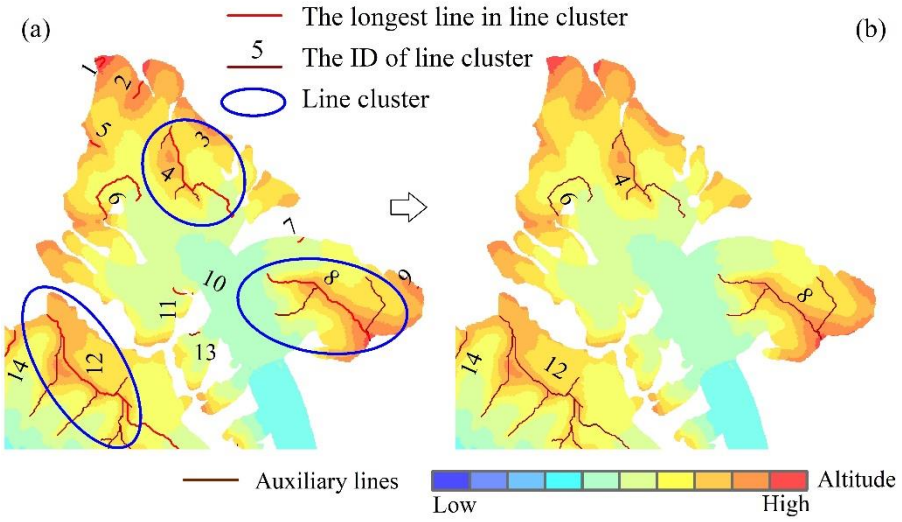


Figure A2: The schematic of post-processing. (a) Before pre-processing; (b) After pre-processing. A total of nine line-clusters are removed by screening.

*(P12L198) How exactly did the authors get the final glacier centerlines?

Response: Firstly, the feature polylines (G_{fl}) after automatically deriving by the program are input, and the function of Euclidean allocation in ArcPy is called to generate the division glacier surface. Then the common edges between regions on the dividing glacier surface are identified. Finally, the common edges are automatically checked and processed (including smoothing process) to obtain the corresponding vector data. This study regards them as the final glacier centerlines.

*(P14L234) How exactly did the authors visual inspection? Some glacier centerlines may be visually indistinguishable.

Response: The method of visual inspection is detailed in section 4.2 (Sample selection and assessment criteria). Indeed, we also found this problem, however, it is hard to avoid. This research is based on a 2D algorithm. Theoretically, the extraction result of the glacier centerline is correct as long as it meets its definition. Nevertheless, we still loaded it on Google Earth for inspection. In addition, we compared it with the glacier length in the RGI v6.0, and further evaluated the extraction results of glacier centerlines.

*(P19L280) Is the DEM used for maximum length calculation in RGI6.0 same with the author's?

Response: We all used SRTM DEM to calculate the longest length of the glaciers. The difference is the spatial resolution of SRTM DEM (this study: 30 m; RGI v6.0: 90 m).

*(P24L364) Maybe I missed some details. How did the authors get ELA through Z_{min}? Maybe the author meant Z_{med}?

Response: We thank the reviewer for the comment. ELA is estimated by Z_{med}, and the relevant content has been corrected above.

*(P26L409) When the article was accepted, I requested the authors to consider making the source code or tool available on Github or some elsewhere.

Response: We agree to you. We will provide an executable file and test results if the paper can be published.

Responses to Reviewer #2 to manuscript TC-2020-294

Thank you very much for your helpful comments to improve this manuscript.

Please Notes: Text in BLACK is the reviewer' comments and our responses are in BLUE. In addition, the notation used to locate the changes first defines the page number, then the line number(s). For example, **P4L15** means that the described modification to the manuscript can be found on the 15th line of the 4th page in the track-changes file.

Specific Comments:

1. The author uses numerous abbreviations. It would be easier for readers to follow if the author could apply a list of these abbreviations.

Response: Thanks for your suggestion. We added a list (Appendix A: Table A1) of main acronyms at the end of this paper.

Table A1 The list of main acronyms in this study. (P38L454)

Acronyms	Description
A_t	The given area of an equilateral triangle
A_g	The polygon's area of the glacier's outer boundary
A_l	The final auxiliary line
A_r	The ridgelines of the glacier surface
G_{br}	The bare rock in glacier
G_{fcl}	The final glacier centerline
G_{fl}	The feature lines of glacier surface
G_{cl}	The original glacier centerline
G_{Labl}	The length in the ablation region of the glacier
G_{Lacc}	The length in the accumulation region of the glacier
G_{Lmax}	The longest length of the glacier
G_{Lmean}	The average length of the glacier
G_{pl}	The polyline of the outer boundary of the glacier
G_{po}	The polygon of the outer boundary of the glacier
L_{max}	The longest glacier length of RGI v6.0
D_L	The difference between G_{Lmax} and L_{max}
P_t	The given perimeter of an equilateral triangle
P_g	The perimeter of the glacier's outer boundary
P_{max}	The local highest point of glacier outline
P_{min}	The lowest point of glacier outline
RGI	The Randolph Glacier Inventory
SCGI	The Second Chinese Glacier Inventory
Z_{med}	The median elevation of the glacier

2. It would be better if the author could provide more detailed, necessary explanations in the figure captions. Not all the figures are self-explainable. For instance, in figure 4, why the background elevation maps look differently in the first and second columns? For the DEM in the third column, some areas are masked out. It would be better if the author could explain why and how these areas are masked out.

Response: Thanks a lot for your comments. We renamed some figures in the manuscript.

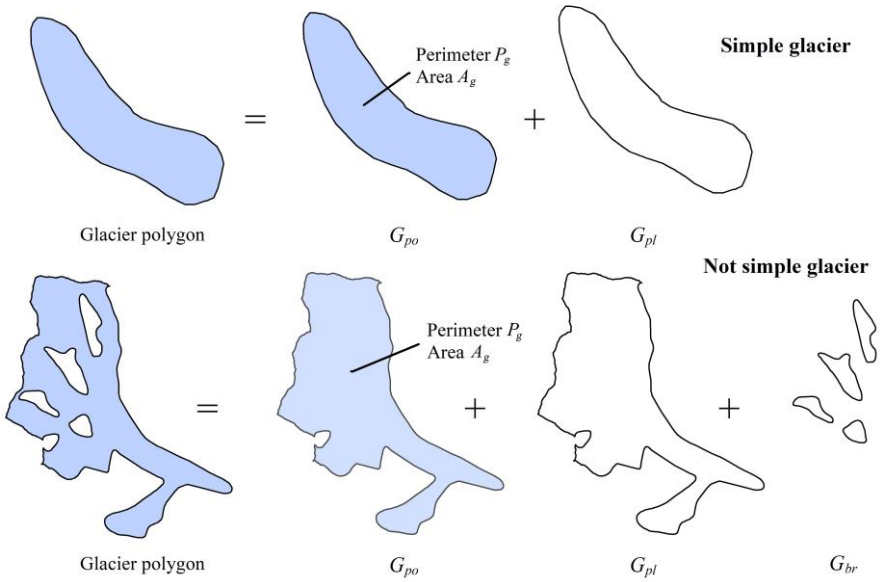


Figure 3: The schematic of processing raw data (G_{po} denotes the polygon of the glacier; G_{pl} denotes the polyline of glacier's outer boundary; and G_{br} denotes the boundary of the bare rock in glacier). (P12L145)

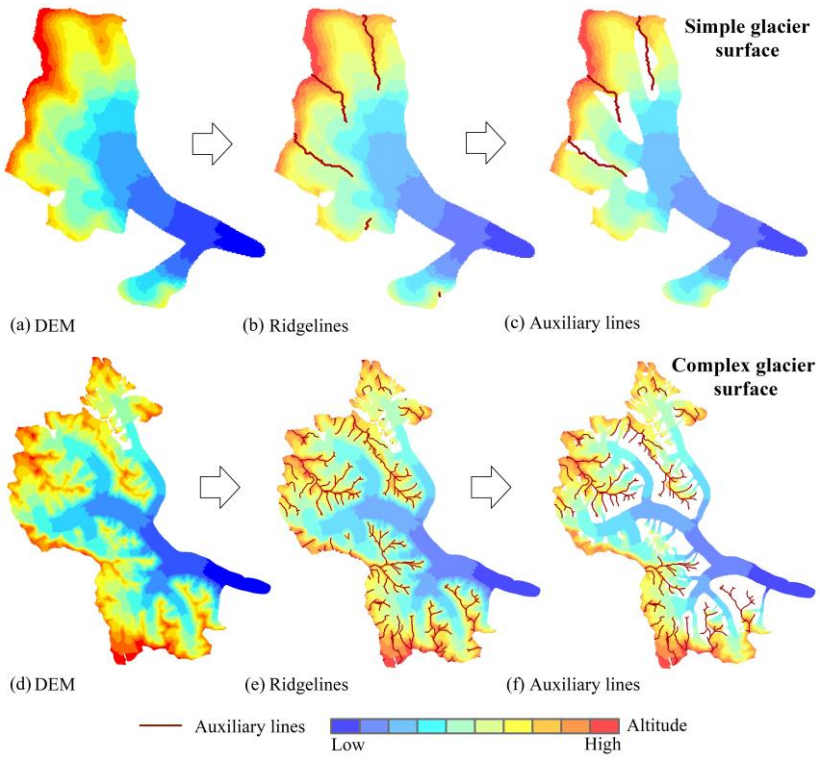


Figure 4: The schematic of extracting auxiliary lines. (a) and (d) demonstrate the digital elevation model (DEM) around the glacier; (b) and (e) show the ridgelines in region covered by DEM; (c) and (f) show the auxiliary lines in glacier. (P14L162)

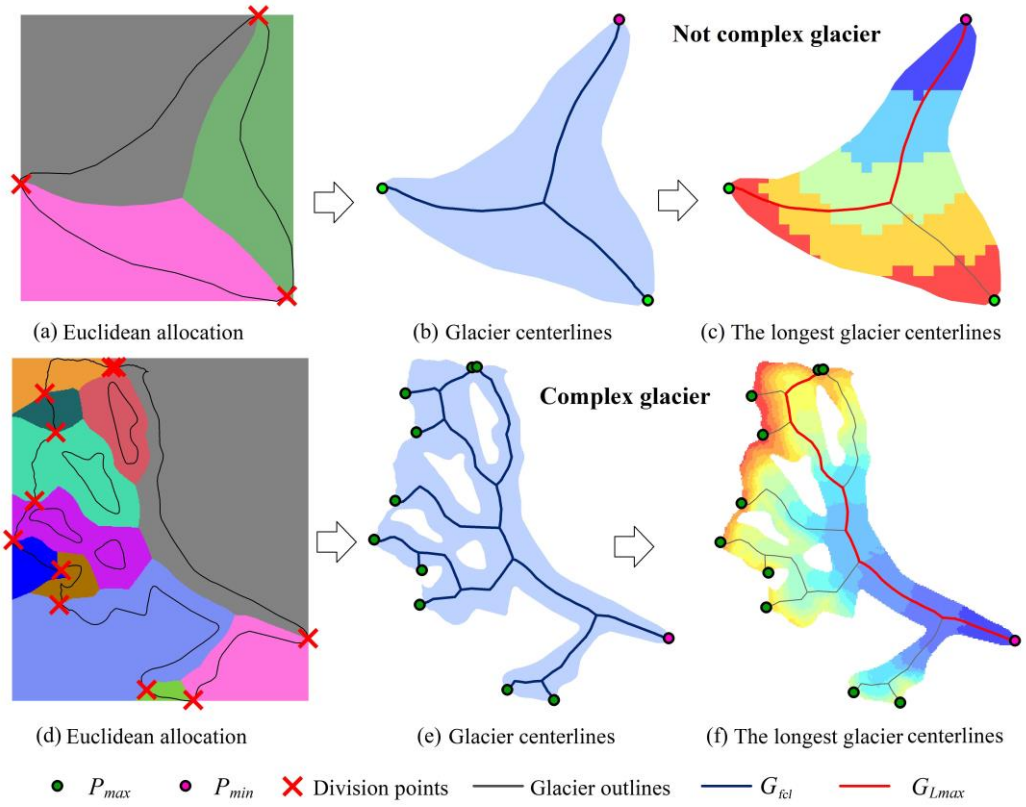


Figure 6: The schematic of extracting centerlines and the longest centerline of the glacier. (a) and (d) show the results after executing the European allocation, and the different colors represent the regions which have the shortest distance to the corresponding edges of the glacier; (b) and (e) represent the centerlines(G_{fcl}), the local highest point (P_{max}) and lowest point (P_{min}) of the glacier; (c) and (f) demonstrate the longest centerline (G_{Lmax}) of the glacier and the background is the digital elevation model with the graduated red (high)– blue (low) color. (P19L218)

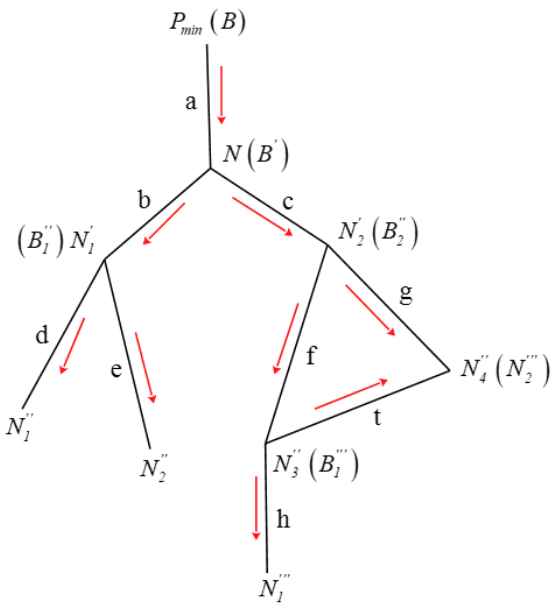


Figure 7: The schematic of calculating glacier length (The red arrow represents the search direction of the branches of glacier centerline). (P20L235)

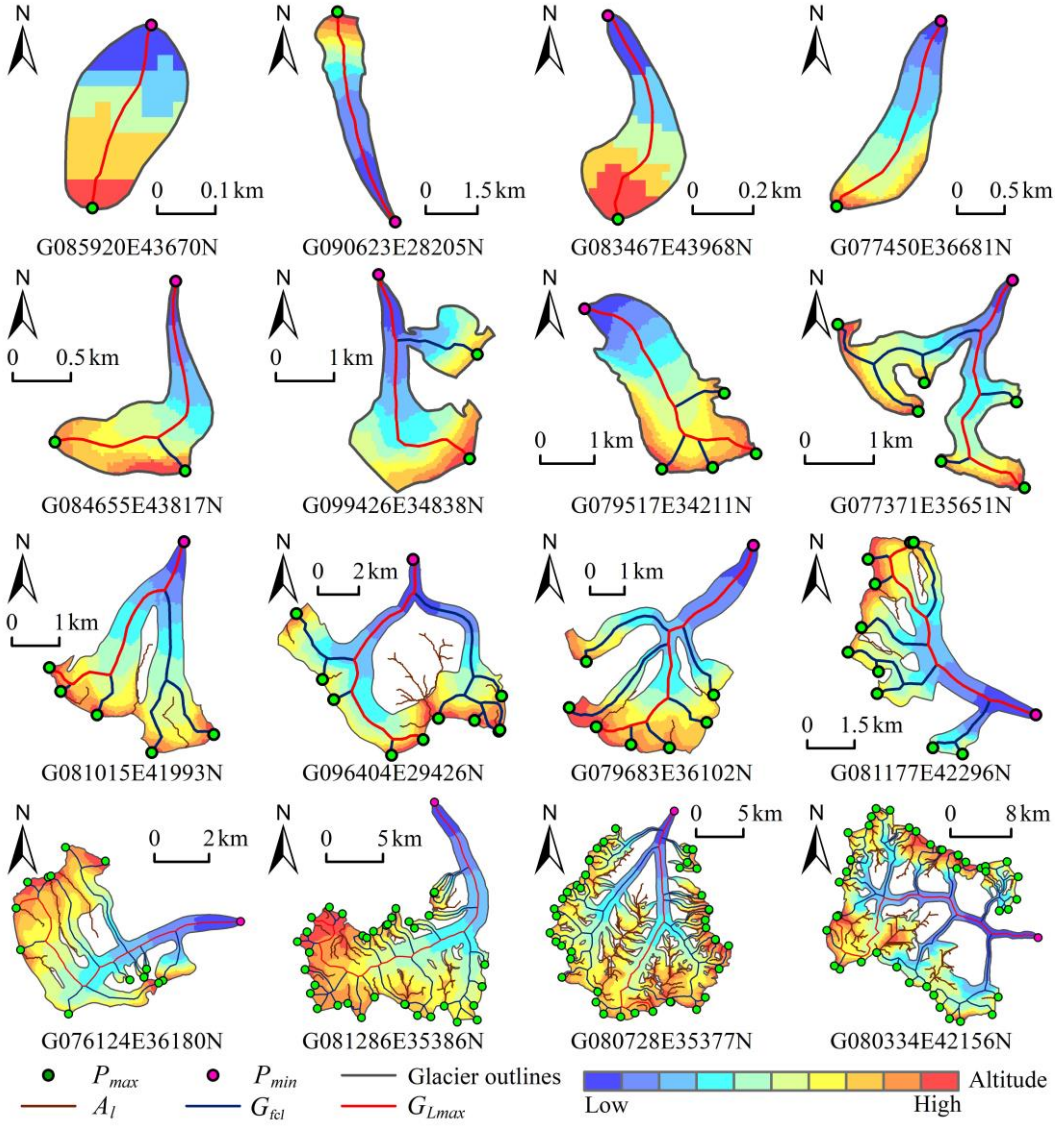


Figure 8: The centerlines for some typical glaciers (P_{max} and P_{min} denote the local highest point and lowest point in the boundary of the glacier, respectively; A_l denotes the auxiliary lines; G_{fcl} and G_{Lmax} denote the centerlines and the longest centerline of the glacier). (P23L259)

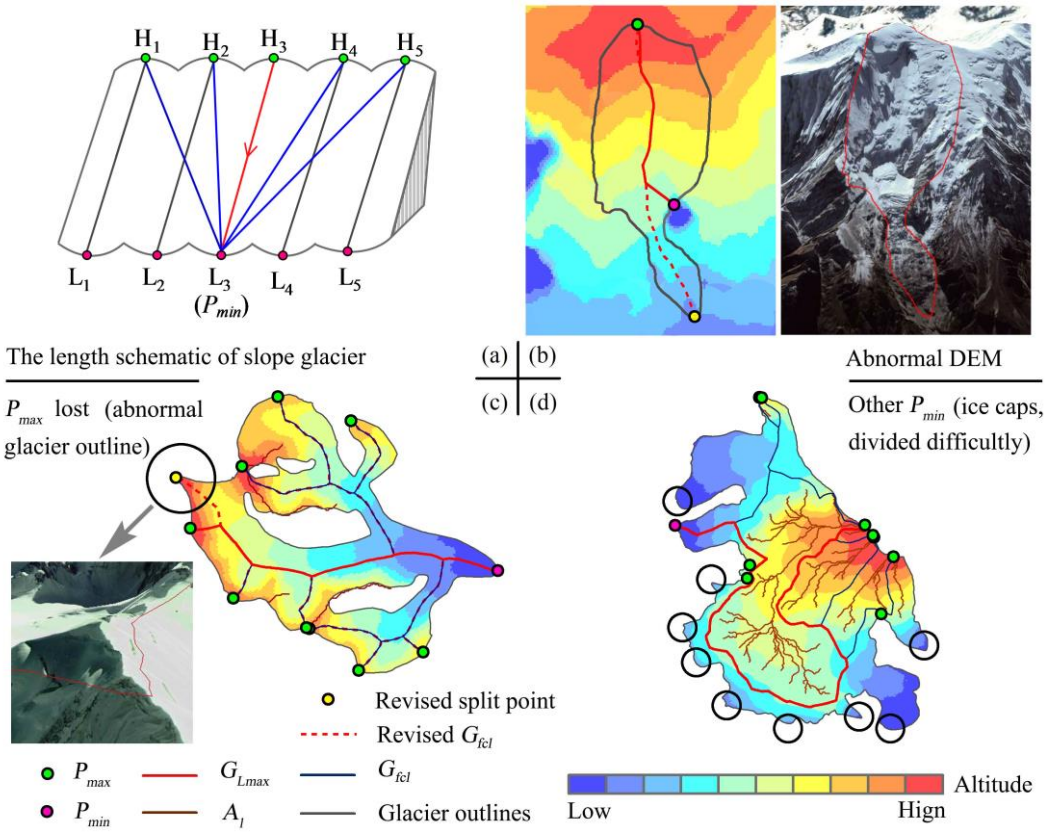


Figure 13: The schematic of probable causes for the abnormal of the longest glacier length. In Figure b, the red dashed line indicates the revised glacier centerline, and the yellow point is the correct lowest point (P_{min}). In Figure c, the red dashed line represents the missing branch, and the yellow point is a local highest point (P_{max}) missed by the algorithm. In Figure d, the black circle indicates some probable exits of the glacier, which needs to be divided into individual glaciers before extracting the centerlines. (P31L338)

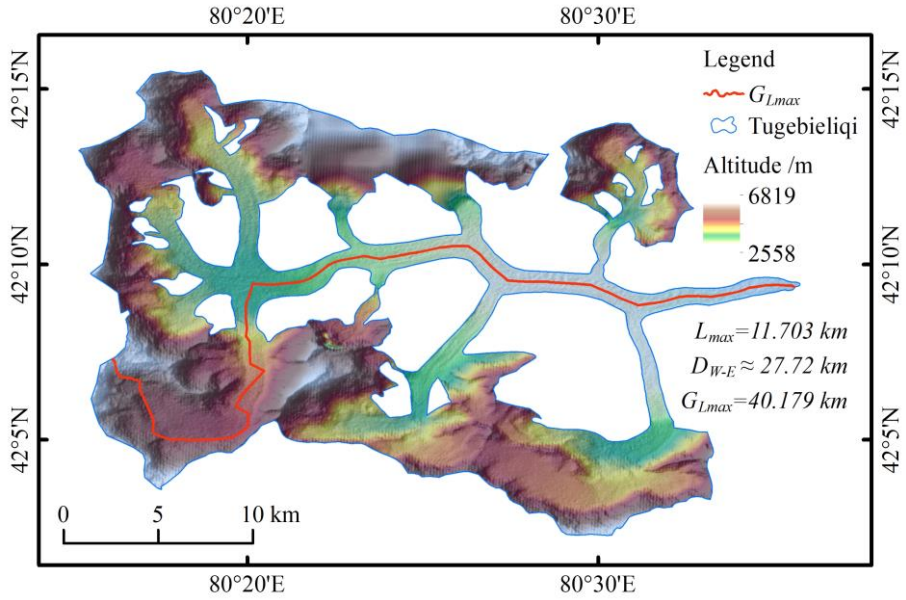


Figure 14: The schematic of the longest centerline of the Tugebieliqi Glacier (L_{max}: the corresponding length of this glacier in the RGI v6.0; $D_{W\text{-}E}$: the distance from west to east of this glacier; G_{Lmax}: the length calculated by our method). (P19L218)

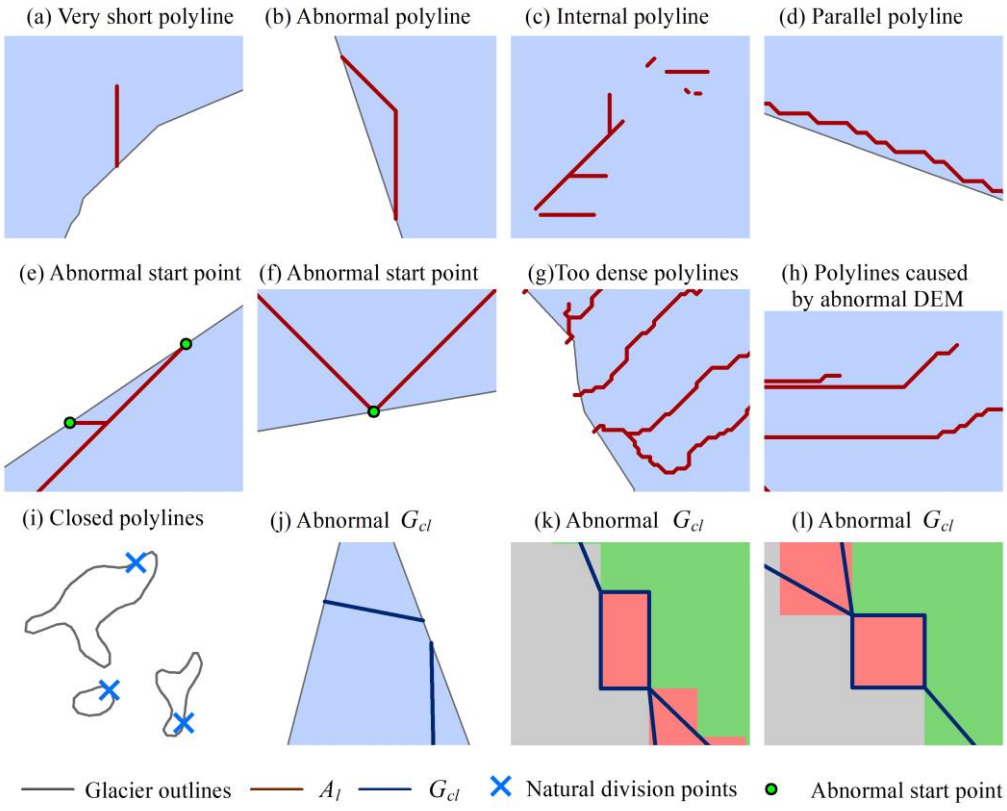


Figure 15: The schematic of discontinuous short polylines. Subgraphs a-h represent type (i), i represents type (ii), j represents type (iii) and k-l represent type (iv). The background in subgraphs a-h and j represent glacier-covered areas. Subgraph i shows several closed polylines, which does not fill background color. The different background colors in subgraphs k-l represent different areas of the glacier surface after the European allocation. (P36L417)

3. Page 6, Line 118: it would be better if the author could explain more about each rule. For instance, (1) why the local highest points must be higher than ELA? (2) Why a glacier has only one exit? The author also mentioned that this single exit could cause problems (See Figure 13d).

Response: This paper takes the four rules as the preconditions for the implementation of the algorithm, and clarifies that the processing unit of the algorithm is an individual glacier polygon instead of a no divided glacier such as the ice sheet ice cap. The detailed explanation is as follows:

(i) As glacier heads, the local highest points are typically located at higher elevations. It is generally considered to be higher than the altitude of 1/3 (Kienholz et al., 2014), or 1/2 glacier area, and Median area altitude (latter: Z_{med}) can be approximated as ELA (Machguth and Huss, 2014).

(ii) This study assumes that all glacier polygons are correctly divided into single glaciers, that is, there is only one glacier terminus (exit). It is generally considered to be the lowest point of the polyline of the outer boundary of a glacier.

(iii) The auxiliary polyline is used to intervene in the generation of centerline for the upper part of a glacier, so it only acts on the accumulation region of glaciers.

(iv) The feature polylines of the glacier surface are composed of the polylines of the outer boundary of the glacier, auxiliary polylines, and the boundary of the bare rock area, which together determine the flow direction of a glacier centerline.

In addition, the reason for this problem (Figure 13d) is that the glacier was not divided into a single glacier in the Second Chinese Glacier Inventory.

4. Figure 2: (1) Could the author explain about extracting DEM and buffering DEM?

Response: In the flow chart of this research, boxes represent the process and the parallelograms represent the results. Therefore, as show as in the flow chart, “buffering DEM” is obtained in the process of “clipping DEM”. Specifically, "extracting DEM" refers to the clipping of the DEM, which appeared twice: one is to use the buffering polygon of the outer boundary of the glacier to clip DEM to obtain the "buffering DEM". Its purpose is to extract the feature information of the glacier such as the lowest point, the local highest points and the auxiliary lines; the other is to use the glacier polygon to extract DEM to estimate the ELA.

5. Figure 3: what is the difference between G_{po} and G_{pl}.

Response: We added the more complete captions in Figure 3:

Figure 3: The schematic of processing raw data (G_{po} denotes the polygon of the glacier; G_{pl} denotes the polyline of glacier’s outer boundary; and G_{br} denotes the boundary of the bare rock in glacier). **(P12L145)**

G_{po} denotes one glacier in 2D geometry (i.e., polygon), and G_{pl} denotes one glacier in 1D geometry (i.e., polyline). Specifically, G_{po} represents the polygon of the outer boundary of the glacier, and G_{pl} refers to the polyline of the outer boundary of the glacier. To identify them more clearly, we collected them in the list (Appendix A: Table A1), as shown as the response to Comment 1.

6. Page 8, Line 142: it would be better if the author could provide more information about hydrologic analysis.

Response: For a more detailed presentation, we added the workflow of hydrological analysis (the shaded region in Figure A2) and changed the relevant description in the manuscript as follows:

Based on the inverse terrain method, the extraction of ridgelines was easily accomplished by the workflow of hydrologic analysis.

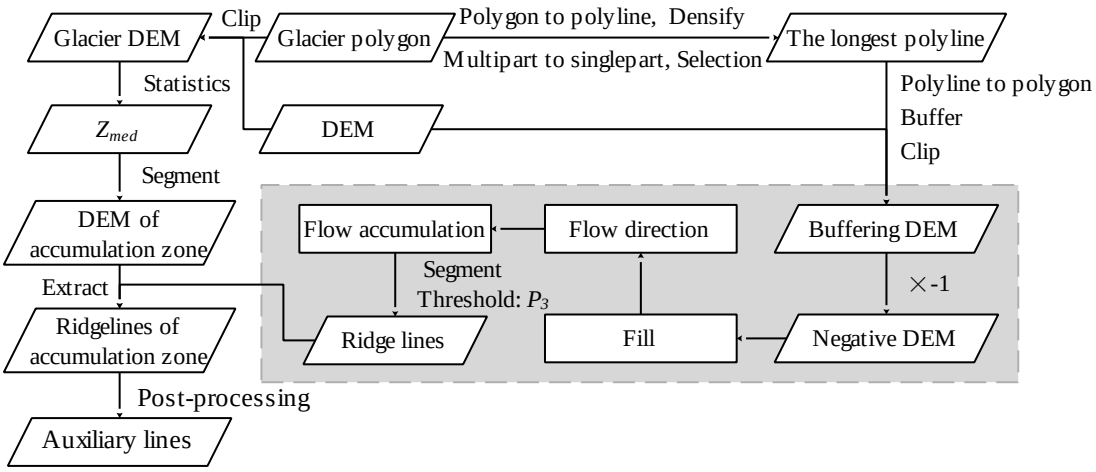


Figure A2: The flow chart of the auxiliary line. The workflow of hydrological analysis is shown in the shaded region.

7. Page 9, Line 148: About identifying abnormal lines, was it done automatically or manually?

Response: It is identified automatically by the program. The whole process of the glacier centerlines extraction is no one intervened from data input to results generation.

8. Page 9, Line 154: Could the author provide more information about the ergodic algorithms?

Response: This part is a detailed explanation of the five steps of post-processing the ridgelines. The ergodic algorithms are shown in Figure A3, which specifically reflects in the following two aspects:

(i) Given a starting point from the set of all possible starting points of auxiliary lines, all the ridgelines of corresponding glaciers are traversed to determine the line cluster composed of the polylines directly or indirectly connected to it.

(ii) Given a line cluster, all the polylines that make up the line cluster are traversed to find the longest ridgeline starting from the starting point.

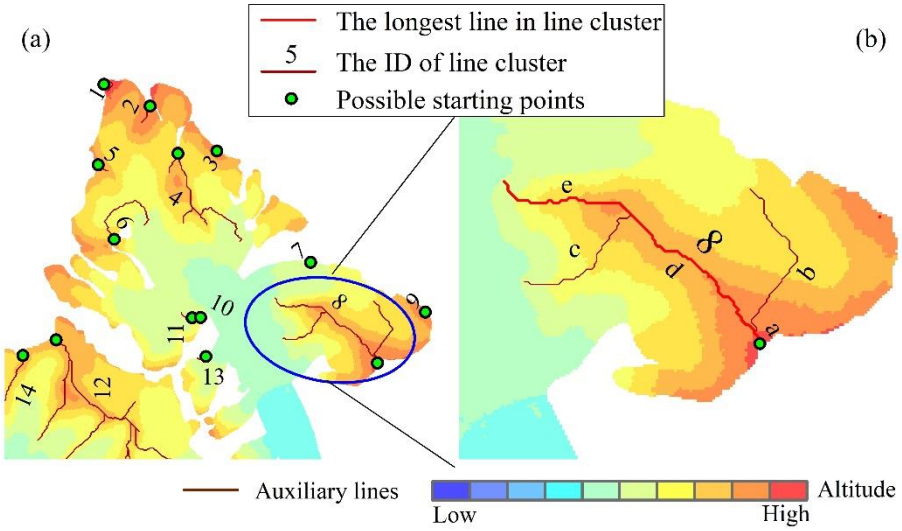


Figure A3: The diagram of the application of the traversal algorithms in the part. (a) 14 line clusters in the figure are identified from all polylines; (b) Line cluster eight consists of five polylines, the longest one is [a, d, e].

9. Page 9, Line 158: Could the author illustrate more about how exactly they screen auxiliary lines using P_4, P_5, and P_{11}?

Response: P_4 is used to control the shortest auxiliary line, filtering some extremely short auxiliary lines. Only the longest length of the line cluster is less than P_4 can be retained.

P_5 is used to filter some extremely long auxiliary lines, and the given threshold is relatively large. It has a particularity and is mainly aimed at some narrow and long glaciers.

P_{11} acts as a switch. When the perimeter of the glacier's outer boundary is greater than the value, parameter P_5 will be used.

10. Figure 5: the definition of Natural division point is missing. It would be better if the author could provide an example showing the natural division point.

Response: In order to make readers better understand the natural division point in this study, we added a schematic of the natural division point, which is shown in Figure A4. It is determined by the storage structure of the closed polyline. It is assumed that there is a set of coordinates [a, b, c, d, e, f, g], which represents the vertex set (V) of a polyline (L). If L is a closed polyline, then g's coordinate of the last member of V is equal to a's coordinate. Although these two coordinates represent the same position and L is also closed visually, a (the polyline head) and g (the polyline end) are separated in data storage. A breakpoint is formed between a and g, which is the natural division point in this paper.

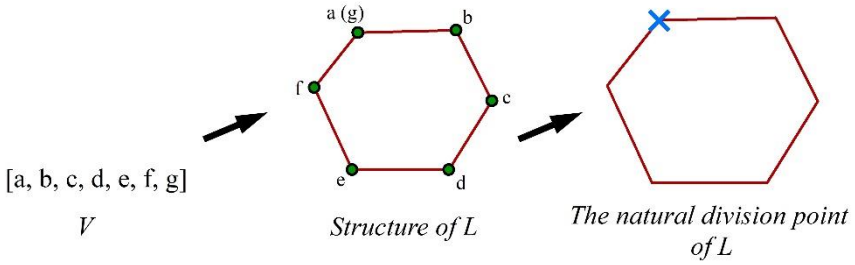


Figure A4: The schematic of the natural division point.

11. Page 12, Line 196: It would be better if the author could explain more about how Euclidean allocation could get glacier centerlines from G_{fl}. To me, the Euclidean is the key part for extracting the glacier centerlines. So I think it is worthwhile to illustrate more about it.

Response: We added more related descriptions, as following:

Original glacier centerlines (G_{cl}) were achieved with the function of Euclidean allocation in ArcPy, which needed the input of G_{fl} and setting the value of P_8. Firstly, the feature lines (G_{fl}) after automatically deriving by the program are input, and the function of Euclidean allocation in ArcPy is called to generate the division glacier surface. Then the common edges between regions on the dividing glacier surface are identified. Finally, the common edges are automatically checked and processed to obtain G_{cl}. **(P17L209)**

The function of Euclidean allocation in ArcPy is used to calculate the nearest source for each cell based on Euclidean distance. It can be divided into three steps:

(i) As the input source locations, G_{fl} is converted to the grid format with a spatial resolution of P_8 according to the ID of the polyline clusters;

(ii) The last step also generated a grid data with an extent of the bounding box of G_{fl} and a spatial resolution of P_8, which is equivalent to an equidistant scatter array, and can be used as the output source locations;

(iii) By calculating the Euclidean distance between each output source location and the all input source locations one by one, the closest input source (polyline cluster) is determined, and its ID is assigned as the value of the output source location. The raster consisting of the updated output source locations is then exported, that is, the raster of glacier surface after segmented by the function of Euclidean allocation.

12. Page12, Line 197: The author uses the Peak algorithm to get the final glacier centerlines from

the glacier centerline. What is the purpose of this step? Figure 2 shows that the Peak algorithm is to smooth the polyline. Why do we need to smooth the polyline?

Response: Glacier centerline (G_{fcl}) represents the main flow line of a glacier. The smoothing algorithm can eliminate the zigzag pattern or irregular polylines in the result to make it closer to the actual main flow line of a glacier. This is consistent with the processing methods of the other two related papers (Kienholz et al., 2014, Machguth and Huss, 2014). The Peak algorithm selected in this study is as same as that adopted by Kienholz et al., which is relatively simple and has a better smoothing effect. The difference is that in this study, the zigzag pattern or irregular polylines are caused by the lower spatial resolution of Euclidean allocation (depending on P_8: to trade-off the extraction efficiency and the accuracy of the results), while in their study, those are caused by the low-quality DEM.

Meanwhile, the risk of filtering is also very little, because the filtering result (shorter part) is always consistent with the forward trend of glacier centerline.

13. Page 19, Line 287: It would be better if the author could provide examples showing that their results are more consistent with the actual conditions of glaciers comparing with RGI v6.0.

Response: There are two reasons why the results of this study are more consistent with the actual conditions of glaciers: (i) For some glaciers with large differences in the longest length of glaciers extracted by the two algorithms, visual inspection can reveal this conclusion. (ii) The spatial resolution of DEM used in this research is better than the RGI v6.0. Correspondingly, the results are more consistent with the actual conditions of glaciers. As an example, figure 14 can reflect this result to some extent.

In the past two years, we have been looking for a set of existing graphical data of glacier centerlines that can be used for verification for this study. We also tried to ask the authors of related papers for help in the form of E-mail. However, we did not get any available information.

Fortunately, RGI v6.0 provides the numerical data of the longest length of glaciers, and most of the corresponding glacier polygons are derived from SCGI. We used the field of GLIMS_ID shared by the two sets of data for matching, and finally obtained the set of numerical data that was used to verify the results of this research.

14. Page 19, Line 291: The tolerance here is 90 meters (3 pixels of DEM). It would be better if the author could explain why they choose this value.

Response: The main reason for choosing 3 DEM pixels (90 m) as the tolerance is that the spatial resolution of DEM used to calculate the longest glacier length in RGI v6.0 is 90 meters. Another reason is that in this study, whether the selection of the local highest points or the process of the Euclidean allocation (given maximum P8: 30 m), at least 3 pixels are needed to determine a local highest point or an effective vertex of glacier centerline.

Therefore, this is based on the theoretical maximum error of this study and the minimum error of the longest glacier length in RGI v6.0 as the tolerance of statistics.

15. Page 19, Line 292: If I understand it correctly, I suggest the author rephrase the sentence as "There were 22017 glaciers within the tolerance, 925 glaciers with negative D_L and 15111 glaciers with positive D_L that are out of the tolerance.

Response: It has been modified. **(P28L318)**

16. Figure 13: For 13b, c, and d, where is the correct glacier centerline? Also, in figure 13d, what do these black circles mean? Please add more information in the figure caption (See comment 2).

Response: We added more information to the caption of Figure 13:

Figure 13: The schematic of probable causes for the abnormal of the longest glacier length. In Figure b, the red dashed line indicates the revised glacier centerline, and the yellow point is the correct lowest point (P_{min}). In Figure c, the red dashed line represents the missing branch, and the yellow point is a local highest point (P_{max}) missed by the algorithm. In Figure d, the black circle indicates some probable exits of the glacier, which needs to be divided into individual glaciers before extracting the centerlines. **(P31L338)**

The revised Figure 13 is shown in Figure A5. We added the correct centerline and the lowest point in subgraph b, and added the local highest points missed by the algorithm and the revised glacier centerline to subgraph c. At the same time, we updated the legend in the new figure.

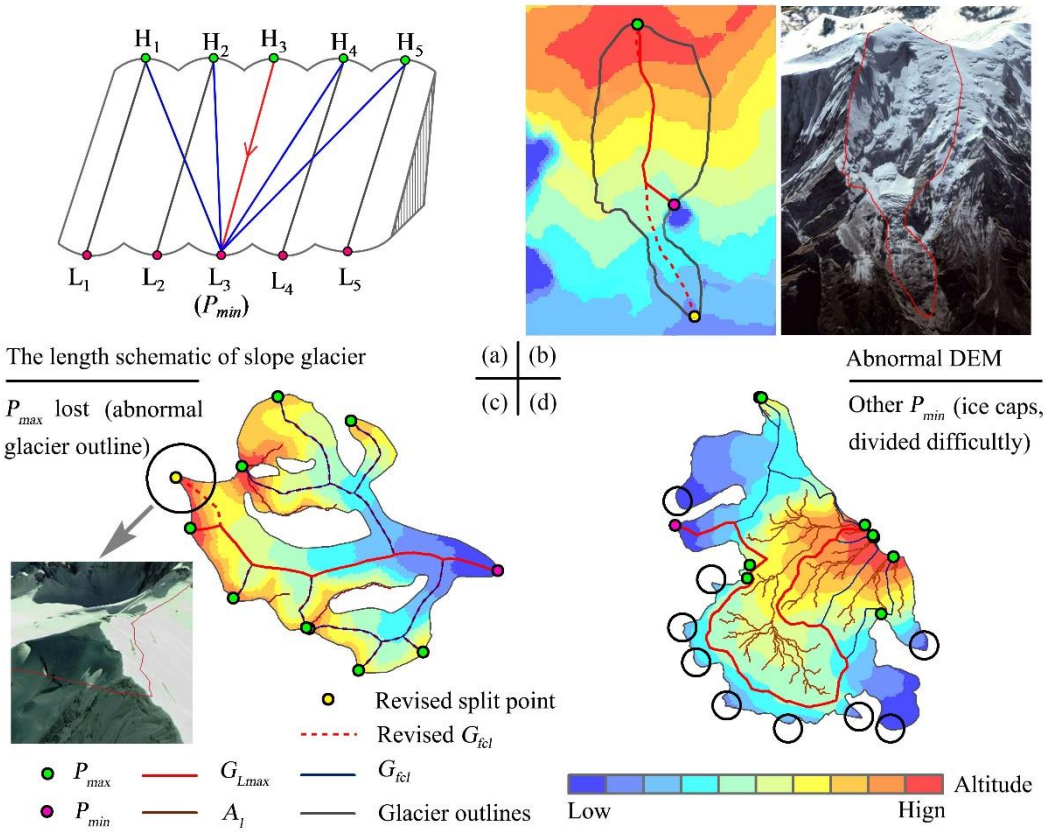


Figure A5: The revised Figure 13. (P30L337)

17. Figure 15: Please add legends of regions with different colors or illustrate them in the figure caption. Please consider numbering each subfigure (See comment 2).

Response: The revised Figure 15 is shown in Figure A6. We ranked the 12 sub-graphs from a to l in the revised Figure 15 and added more detailed caption:

Figure 15: The schematic of discontinuous short polylines. Subgraphs a-h represent type (i), i represents type (ii), j represents type (iii) and k-l represent type (iv). The background in subgraphs a-h and j represent glacier-covered areas. Subgraph i shows several closed polylines, which does not fill background color. The different background colors in subgraphs k-l represent different areas of the glacier surface after the European allocation. **(P36L417)**

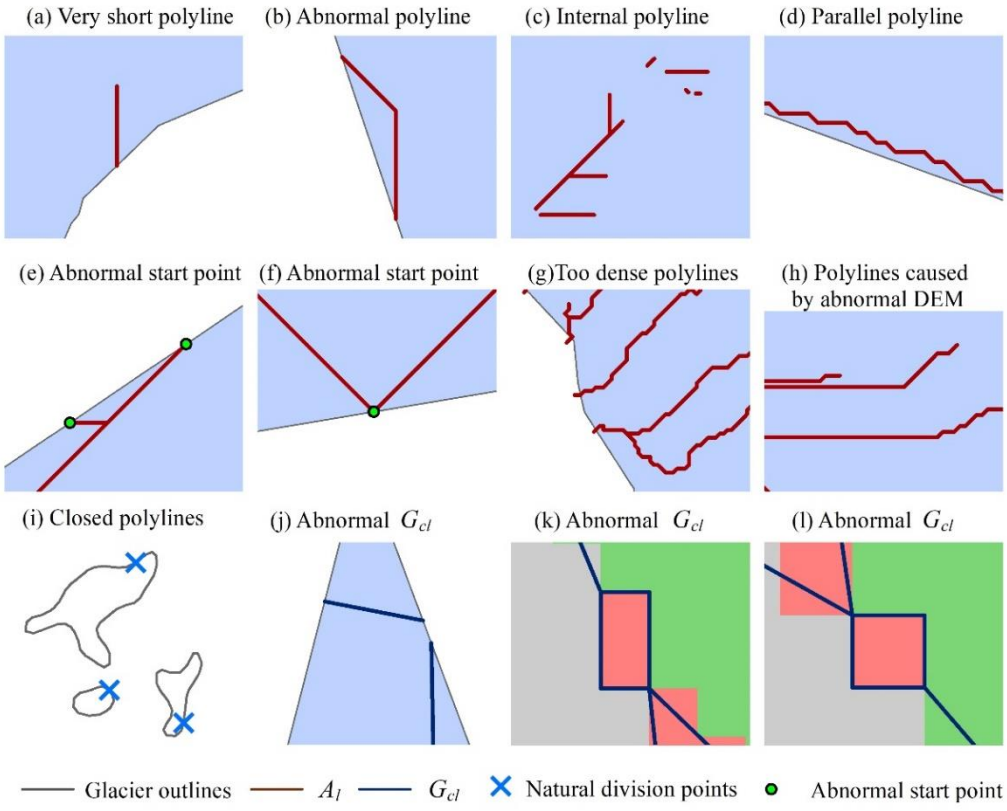


Figure A6: The revised Figure 15. (P35L415)

18. Figure 15: In the first two figures of the second row, where is the abnormal start point? In the fourth figure of the second row, why it is due to abnormal DEM?

Response: The gray polylines in Figure 15 represent glacier outlines. Intersection points of glacier outline and A_l is the anomaly starting point in the first two figures of the second row (subgraphs e and f). We added these intersection points to the figure and the corresponding legend (See figure A6).

In the figure (subgraph h) in the second row and fourth column, A_l is a straight line. There are generally two reasons for this situation: (i) the corresponding area is a flat surface with a slope of almost zero; (ii) the topography of the corresponding region is extremely complex, and the quality of DEM is too poor. The elevation values in a region are almost same because they are derived by interpolation. In the accumulation region of a glacier, the latter accounts for the vast majority. Thus, it is believed that this situation is caused by the abnormal DEM.

Technical Corrections:

Page 5, Line 85: Consider to change "arcpy" to "ArcPy".

Response: It has been modified. **(P5L92, P17L209)**

Figure 2: In the part of the extraction of centerlines, it seems that G_{cline} and G_{gcf} should switch their position according to the author's definition.

Response: We rewrote acronym of each parameter to clarify their meanings, listed in the Appendix A (Table A1) and can also be found in the response to Comment 1.

Page 9, Line 148: In the third part of the post-processing, is it "numbers" or "members"?

Response: It should be "members", and refers to the elements that make up a line cluster.

Figure 9: Consider changing "inexact" to "inaccurate" for consistency.

Response: It has been modified. **(P26L309)**

Figure 12: Consider changing "*DL*" to "D_L" in the figure caption. For the figure on the right-hand side, the blue color should represent the number of $+D_L$, if I understand it correctly.

Response: It has been modified. **(P28L314, P29L323)**

References:

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, doi: 10.5194/tc-8-1741-2014, 2014.

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, doi: 10.5194/tc-8-503-2014, 2014.

A new automatic approach for extracting glacier centerlines based on Euclidean allocation

Dahong Zhang[1,2], Xiaojun Yao[1,3], Hongyu Duan[1], Shiyin Liu[4], Wanqin Guo[5], Meiping Sun[1], Dazhi Li[1]

[1]College of Geography and Environment Sciences, Northwest Normal University, Lanzhou, China

[2]College of Urban and Environmental Sciences, Northwest University, Xi'an, China

[3]National Cryosphere Desert Data Center, Lanzhou, China

[4]Institute of International Rivers and Eco-security, Yunnan University, Kunming, China

[5]State Key Laboratory of Cryospheric Science, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou, China

Correspondence to: Xiaojun Yao (yaoxj_nwnu@163.com)

Abstract. Glacier centerlines are crucial input for many glaciological applications. From the morphological perspective, we proposed a new automatic method to derive glacier centerlines, which is based on the Euclidean allocation and the terrain characteristics of glacier surface. In the algorithm, all glaciers are logically classified as three types including simple glacier, simple compound glacier and complex glacier, with corresponding process ranges from simple to complex. The process for extracting centerlines of glaciers introduces auxiliary reference lines, and follows the setting of not passing through bare rock. The program of automatic extraction of glacier centerlines was implemented in Python and only required glacier boundary and digital elevation model (DEM) as input. Application of this method to 48571 glaciers in the second Chinese glacier inventory automatically yielded the corresponding glacier centerlines with an average computing time of 20.96 s, a success rate of 100% and a comprehensive accuracy of 94.34%. A comparison of the longest length of glaciers to the corresponding glaciers in the Randolph Glacier Inventory v6.0 revealed that our results were ~~more~~ superior. Meanwhile, our final product ~~provided~~ provides more information about glacier length, such as the average length, the ~~largest~~ longest length, the lengths in the accumulation and ablation regions of each glacier.

1 Introduction

Glaciers are an important freshwater resource on earth and a vital part of the cryosphere (Muhuri et al., 2015). According to the Fifth Assessment Report (AR5, https://www.ipcc.ch/) published by the Intergovernmental Panel on Climate Change (IPCC), there are 168331 glaciers (including ice caps) in the world, with a total area of 726258 km^2 apart from ice sheets. Glaciers move towards lower altitude by gravity, which is the most obvious distinction between glacier and other natural ice bodies. The glacier flowlines are the motion trajectories of a glacier, and the main flowline is the longest flow trajectory of glacier ice. Due to the differences in the speed and moving direction of any point at the surface or inside the glacier, the calculation of the main flowline of glaciers requires a coherent velocity field data, which is difficult to obtain on the global or regional scale (McNabb et al., 2017). Therefore, some concepts such as the glacier axis and the glacier centerlines were proposed (Le Bris and Paul, 2013; Kienholz et al., 2014; Machguth and Huss, 2014). Glacier centerlines are the central lines close to main flowlines of glaciers, which can be acquired based on glacier axis and be used to simulate glacier flowline.

As an important model parameter, the glacier centerline can be used to determine the change of glacier length (Leclercq et al., 2012a; Nuth et al., 2013), analyze the velocity field (Heid and Kääb, 2012; Melkonian et al., 2017), estimate the glacier ice volume (Li et al., 2012; Linsbauer et al., 2012), and develop one-dimensional glacier models (Oerlemans, 1997; Sugiyama et al., 2007). Meanwhile, the length of the longest glacier centerline is one of the key determinants of glacier geometry and an important parameter of glacier inventory (Paul et al., 2009; Leclercq et al., 2012b). The length and area of glacier can be also used to estimate the large-scale glacier ice volume (Zhang and Han, 2016; Gao et al., 2018). The length change at the terminus of a glacier can directly reflect the state of motion, e.g., glacier recession, glacier advance or surging (Gao et al., 2019). Winsvold et al. (2014) analyzed the changes of glacier area and length in Norway, using glacier inventories derived from Landsat TM/ETM+ images and digital topographic maps. Herla et al. (2017) explored the relationship between the geometry and length of glaciers in the Austrian Alps based on a third-order linear glacier length model. Leclercq et al. (2012, 2014) reconstructed annual averaged surface temperatures in the past 400 years on hemispherical and global scale from glacier length fluctuations. These studies indicated that both the extraction of contemporary glacier

length and the reconstruction of historical glacier length require more accurate extraction methods of glacier flowlines.

In order to obtain the length of glaciers, some automatic or semi-automatic methods were proposed in recent years. Schiefer et al. (2008) extracted the longest flow path on the ice surface based on a hydrological model, which was generally 10% to 15% larger than the glacier length. Le Bris et al. (2013) accomplished the automatic extraction of flow lines from the highest point to the terminus of a glacier based on the concept of glacier axis, with a verification accuracy of 85%. Unfortunately, the branches of glacier centerlines have not been extracted and the length is not necessarily the maximum for huge or complex glaciers (Paul et al., 2009). Machguth et al. (2014) proposed an extraction method of glacier length based on the slope and width of glacier with a success rate of 95-98%, however the branches of glacier centerlines could not be extracted either. Kienholzs et al. (2014) applied the grid–least-cost route approach to the automatic extraction of glacier flow lines, having an automation degree of 87.8% with additional manual intervention. Yao et al. (2015) proposed the semi-automatic method of extraction glacier centerlines based on Euclidean allocation theory, which required the expertise and experiences for composite valley glaciers and ice caps. So, the current biggest challenge is still the implementation of automation extraction of glacier centerline and the acquirement of more information about glacier length. The aims of this study are to design an algorithm to: (i) automatically generate centerlines for the main body of each glacier and its branches; (ii) automatically calculate the longest length, average length, the length of accumulation region, and the length of ablation region of each glacier, along with corresponding polylines; and (iii) improve the degree of automation as much as possible on the premise of ensuring the accuracy of glacier centerlines.

2 Input data and test region

The glacier dataset used in this study is the Second Chinese Glacier Inventory (SCGI) released by National Tibetan Plateau Data Center (http://westdc.westgis.ac.cn/data), which has been approved by some organizations (e.g., WGMS, GLIMS, NSIDC, etc.) and adopted in the Randolph Glacier Inventory (RGI) v6.0 (Guo et al., 2017). According to the SCGI (Fig.1), there were 48571 glaciers in China, with a total area of 51766.08 km^2, accounting for 7.1% of the glacier area in the world

except for the Antarctic and Greenland ice sheets (Liu et al., 2015). Due to the lack of automatic method to calculate glacier's length, there was no length property in the SCGI, and some subsequent studies haven't made great breakthroughs (Yang et al., 2016; Ji et al., 2017).

The SCGI was produced based on Landsat TM/ETM+ images and ASTER images in the period of 2004-2011 and SRTM v4.1 with a spatial resolution of 90 m (Liu et al., 2015). In this study, we selected SRTM1 DEM v3.0 (http://www2.jpl.nasa.gov/srtm, last accessed on March 2, 2013, with a spatial resolution of 30 m) (Farr et al., 2007) in consideration of its free access and higher data quality, which was used to identify division points on the glacier outlines, extract ridge lines in the coverage region of glaciers, and generate the glacier centerlines. Additionally, we extracted glaciers in China from the RGI v6.0 provided by GLIMS (http://www.glims.org/RGI/). There are 38053 glaciers matching the graphic position of the SCGI. The field of L_{max} of RGI v6.0 provides the length of the longest flowlines on the glacier surface, which was calculated with the algorithm proposed by Machguth et al. (2014). For verifying the validity and accuracy of glacier centerlines, we compared the extracted longest length of glaciers with the value of L_{max} in the RGI v6.0.

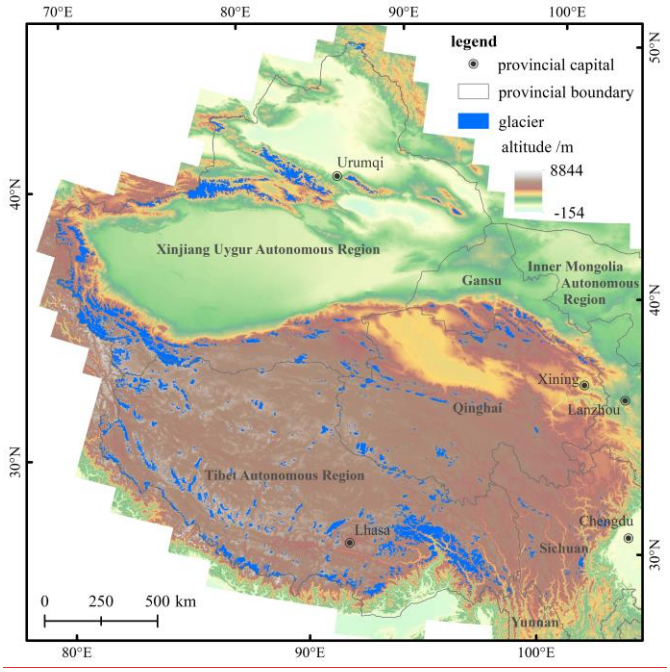


Figure 1: The distribution of glaciers in China.

3 Principles and algorithm of glacier centerline extraction

In order to implement the automatic extraction of glacier centerlines, we have designed a new set of algorithms. Relevant parameters and processing procedures are introduced as follows.

3.1 Model parameters

The code was written in Python and partially invoked the site package of ArcPy. The calculation of the glacier centerlines relies on two basic inputs: (i) glacier in the form of polygon with a unique value field and a projection coordinate system (unit: m), (ii) DEM data having the spatial resolution and acquisition time close to the images used for glacier inventory. We defined

11 adjustable parameters named P_i (i=1,…,11) (see Table 1), which were achieved by classifying glacier polygon through a set of reasonable rules. The purpose is to improve the degree of automation and the accuracy. Three key parameters are described as:

—P_3: the threshold of flow accumulation, to control the generation of auxiliary lines.

—P_6: the step size of searching the local highest points, to control the extraction of extremely high points.

—P_8: The grid cell size of Euclidean allocation, to improve the algorithm efficiency.

In the algorithm, the number of the local highest points is affected by the perimeter of the glacier (P_g). We took the given area (A_t) and the perimeter (P_t, Eq.1) of the equilateral triangle corresponding to A_t as the grading threshold. According to the area (A_g) and the perimeter (P_g) of each glacier's outer boundary, all glaciers were divided into five levels (Eq.2), which represented the five levels of glacier polygon with difference in P_g. The built-in parameters were set according to the different levels (Table 1). P_4, P_5 and P_9 were controlled in proportion to the side length of the equilateral triangle corresponding to P_t. The proportional coefficient was T (Eq.3). According to the actual situation of the repeated programing test, the empirical value of each parameter is given in Table 1.

$$P(A_t) = 2 \times 3^{0.75} \times A_t^{0.5} \tag{1}$$

$$L(A_g, i) = \begin{cases} i: A_g \in [A_{i-1}, A_i) \text{ and } P_g \in [P(A_i), +\infty) \text{ and } i \in (1,5] \\ i: A_g \in [A_i, A_{i+1}) \text{ and } P_g \in \left[P(A_i), P(A_{i+1})\right) \text{ and } i \in [1,5] \\ i: A_g \in [A_{i+1}, A_{i+2}) \text{ and } P_g \in \left(0, P(A_{i+1})\right) \text{ and } i \in [1,5) \\ 0: \text{ the above conditions aren't met} \end{cases} \quad A = \{0,1,5,20,50,+\infty\}, \; i = \{1,2,3,4,5\} \tag{2}$$

$$f(T) = \frac{P_g}{3 \times 2 \times T} \tag{3}$$

Table 1 The description of adjustable parameters.

Levels / Par.	1 L(~~A_G~~A_g,1)	2 L(~~A_G~~A_g,2)	3 L(~~A_G~~A_g,3)	4 L(~~A_G~~A_g,4)	5 L(~~A_G~~A_g,5)	Parameter elucidation
*P_1			"10 meters"			Maximum distance between adjacent vertexes of polyline
*P_2			"30 meters"			Buffer distance outside the glacier outline
P_3	500	600	700	800	800	The threshold of accumulative flow
P_4	f(10)	f(11)	f(12)	f(13)	f(15)	The length of the shortest auxiliary line
P_5	f(2)	f(3)	f(4)	f(5)	f(6)	The length of the longest auxiliary line
P_6	50	60	70	80	80	The interval for searching the local highest points
P_7	0.2	0.2	0.5	0.5	1	The matching tolerance of the vertexes of polyline
P_8	1	5	15	15	30	The size of grid cell in Euclidean allocation
P_9	f(~~15~~10)	f(~~30~~15)	f(~~60~~30)	f(~~120~~60)	f (~~240~~120)	Minimum distance between the adjacent local highest points
P_{10}	5	10	15	20	30	The smoothing tolerance of polylines
*P_{11}			P(A_L=5)			Threshold to control the length of the longest auxiliary line

Note: the parameters with "*" are constant.

3.2 Computation flow

In this paper, glaciers were divided into three categories: simple glacier (extremely high point: single, auxiliary line: no, the area of bare rock: no), simple compound glacier (extremely high point: several, auxiliary line: no, the area of bare rock: no), and complex glacier (extremely high point: several, auxiliary line: yes, the area of bare rock: yes). Following the principle from simple to complex, the algorithm was composed of six main steps: data preprocessing, extraction of auxiliary lines, identification of division points, reconstruction of feature lines, extraction of centerlines and the calculation of glacier length. The flow chart of the algorithm is illustrated in Fig.2.

The automatic extraction of glacier centerlines in this study obeys the following rules: (i) the elevation of the local highest points must be higher than the equilibrium line altitude (ELA), (ii) a glacier has only one exit, which is the lowest point of the polyline of the glacier's outer boundary ~~of the glacier~~ (G_{pl}); (iii) the auxiliary line only acts on the accumulation region of

glacier; (iv) the G_{pl}, auxiliary lines, and bare rock region simultaneously serve as barrier lines to restrict the flow direction of the glacier centerlines.

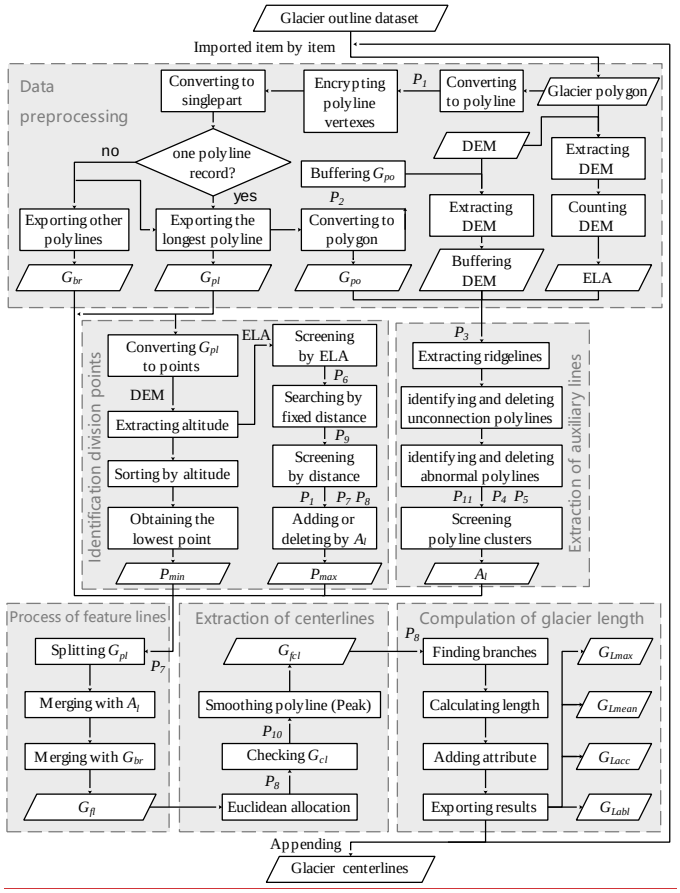


Figure 2: The flow chart of algorithm.

3.3 Critical processes

3.3.1 Data preprocessing

The data preprocessing includes four parts: (i) checking the input data, (ii) pre-processing the glacier outlines, (iii) fine-tuning the built-in parameters, and (iv) calculating the ELA of glaciers. First, the polygon of the glacier's outer boundary (G_{po}), the polyline of glacier's outer boundary (G_{pl}) and the boundary of the bare rock in glacier (G_{br}) were obtained by splitting the glacier outlines in the importing module. These temporary data would be used as the input parameters of other modules in subsequent process. Secondly, the module exported the number of closed lines in glacier outlines, A_g and P_g, which were used to determine the number of bare rocks on the glacier surface, the type and level of glaciers. Thirdly, according to the parameter adjusting rules at the level of glaciers, 11 built-in parameters (see Table 1) were fine-tuned. Finally, the median elevation (Z_{med}) of each glacier aided by its DEM was computed, which was then used to estimate the ELA of each glacier. The schematic diagram of processing glacier outlines is shown in Fig.3.

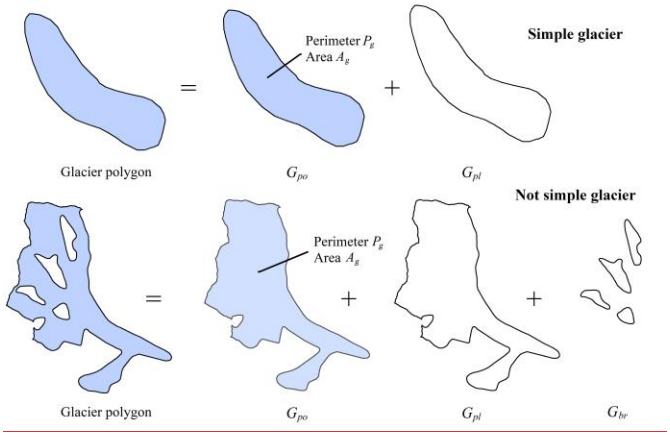


Figure 3: The schematic of processing raw data (G_{po} denotes the polygon of the glacier; G_{pl} denotes the polyline of glacier's outer boundary; and G_{br} denotes the boundary of the bare rock in glacier).

3.3.2 Extraction of auxiliary lines

For making glacier centerlines more reasonable, we introduced the auxiliary lines that represent the internal ridgelines of glaciers to intervene in the generation of centerline for the upper part of a glacier. The extraction of auxiliary lines included the extraction of ridgelines and post-processing. Based on the inverse terrain method, the extraction of ridgelines was easily accomplished by the workflow of hydrologic analysis. The post-processing was relatively complicated. The main reason was that the auxiliary lines were tree-like polylines starting from the upper boundary of the glacier. In principle, the mass flow in the location of the auxiliary lines on the glacier surface could be obviously blocking-up, which was equivalent to the ice divide. The preliminary ridgelines needed to be screened once more combining with DEM by traversal method. Determining the cluster of auxiliary lines was the main problem to be solved by the algorithm of this part. According to the designed algorithm, it could be divided into five parts in post-processing: (i) identifying and deleting the disconnected lines, (ii) identifying and deleting the abnormal lines, (iii) determining the members of line cluster, (iv) determining the longest length of line cluster, and (v) screening the line clusters. The schematic diagram of extracting the auxiliary lines is shown in Fig.4.

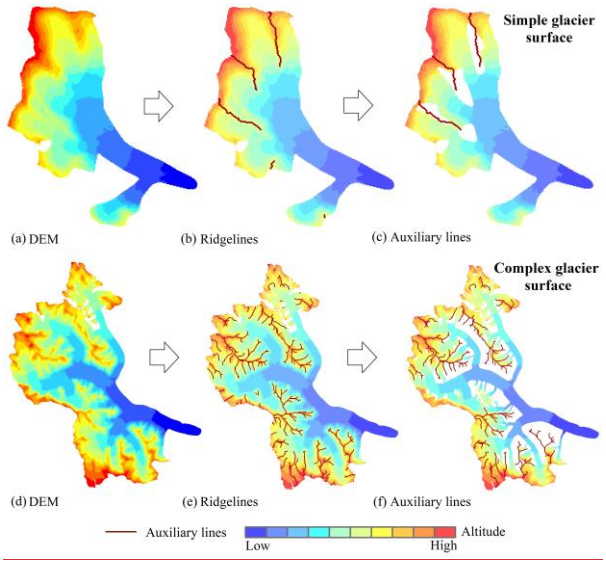


Figure 4: The schematic of extracting auxiliary lines. **(a) and (d) demonstrate the digital elevation model (DEM) around the glacier; (b) and (e) show the ridgelines in region covered by DEM; (c) and (f) show the auxiliary lines in glacier.**

The automatically extracted ridgelines were often disconnected, so it was necessary to remove independent existence or unreasonable ridgelines using the auxiliary data such as DEM, ELA and G_{po} by ergodic algorithms. Firstly, the ridgelines of the glacier surface (A_r) were obtained by clipping the ridge lines using G_{po}. The set of all possible starting points of auxiliary lines was gained by intersecting A_r with G_{pl}. Then, the ridgeline clusters connected to each starting point were achieved and marked by traversing the point set. The number of auxiliary lines was initially determined. Lastly, the longest length of each auxiliary line was calculated by adopting the critical path algorithm. The final auxiliary lines (A_l) were obtained by screening all auxiliary lines using the three parameters of P_4, P_5 and P_{11}.

3.3.3 Identification of division points

The division points include the lowest point (P_{min}) and the local highest point (P_{max}). The ordered point set (h) was obtained

after converting G_{pl} from a polyline to a point set and extracting the elevation for the point set. The method for obtaining P_{min} was relatively simple, as showed in Eq. (4).

$$P_{min} = Min(h_1, h_2, \cdots, h_n) \quad (4)$$

In comparison, the extraction of P_{max} was more complicated. It was necessary to ensure the extraction of all possible branches of the centerlines and avoid the redundancy of branches. The algorithm could be divided into four steps: (i) obtaining the local highest point set (M″) by filtering h (Eq.5, Eq.6) according to P_6, (ii) removing the elements (Eq.7) at an altitude lower than ELA from M", (iii) removing the elements (Eq.8) of adjacent distance less than P_9 from M', and (vi) checking, deleting or adding some local highest points (Eq.9) using the auxiliary lines to ensure that there was at least one local highest point among adjacent auxiliary lines.

$$H_i = \left\{h_{i-\frac{P_6}{2}}, \cdots, h_{i-1}, h_i, h_{i+1}, \cdots, h_{i+\frac{P_6}{2}}\right\}, i \in \left[\frac{P_6}{2}, n - \frac{P_6}{2}\right] \quad (5)$$

$$M'' = \{h_i | h_i \geq Max(H_i)\} \quad (6)$$

$$M' = \left\{M_j'' \middle| M_j'' \geq ELA\right\}, j \in \left[0, card(M'')\right) \quad (7)$$

$$M = \{M_k' | d(M_{k-1}', M_k') \geq P_9, and\ d(M_k', M_{k+1}') \geq P_9\}, k \in \left[0, card(M')\right) \quad (8)$$

$$P_{max} = M \cup \left\{l_j \middle| l_j \geq Max(L_i)\right\} \quad (9)$$

3.3.4 Reconstruction of feature lines

Feature lines of glacier surface were used to express G_{pl}, G_{br}, A_l, P_{max}, P_{min}, and the intersection of A_l and G_{pl}. The schematic diagram of merging the glacier surface features is illustrated in Fig.5.

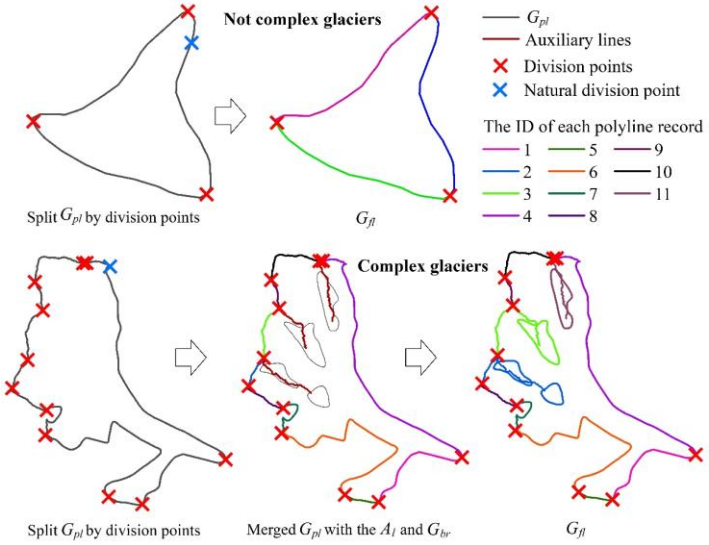


Figure 5: The schematic of extracting the polyline features of glacier surface.

For simple glaciers and simple compound glaciers, it was only necessary to merge P_{max} and P_{min} into a vector file, then split G_{pl}, and allocate one unique code for each polyline after converting it from multipart to singlepart. For complex glacier, the processing method was composed of several steps. First of all, the G_{pl} split by division points needed to be combined with G_{br} (if any) and A_l (if any) into a vector file. After converting it from multipart to singlepart, program would allocate again code for each polyline and remark it as G_{sp1}. Secondly, polyline records in G_{sp1} were selected one by one with A_l, and then the polyline records belonging to the same part in G_{sp1} were merged, which was recorded as G_{sp2}. Thirdly, G_{edge} was exported by selecting G_{sp2} using G_{pl}, and G_{alone} was exported after switching selection, which represented the bare rock region that still existed independently after merging the glacier outlines with the auxiliary lines. Finally, adopting the proximity algorithm, each element (if any) in G_{alone} was processed in turn with G_{edge}. Specifically, it needed three steps: (i) The vertex set E (Eq.10) of G_{edge} and the vertex set U (Eq.11) of G_{alone} were obtained. (ii) The pairs of polylines (Eq.12) matched by serial number were calculated and made the corresponding marks in G_{sp2}; (iii) The feature lines (G_{fl}) of glacier surface were reconstructed by

merging the same marks in G_{sp2}.

$$E_i = \{E_{ij} | j \in [0, card(E)\} \quad (10)$$

$$U_p = \{U_{pq} | q \in [0, card(U)\} \quad (11)$$

$$D = \left\{(p,i) \middle| Min\left(d(U_p, E_i)\right)\right\} \quad (12)$$

3.3.5 Extraction of glacier centerlines

Original glacier centerlines (G_{cl}) were achieved with the function of Euclidean allocation in ArcPy, which needed the input of G_{fl} and set the value of P_8. Firstly, the feature lines (G_{fl}) after automatically deriving by the program are input, and the function of Euclidean allocation in ArcPy is called to generate the division glacier surface. Then the common edges between regions on the dividing glacier surface are identified. Finally, the common edges are automatically checked and processed to obtain G_{cl}. The final glacier centerlines (G_{fcl}) were obtained by processing G_{cl} with Peak algorithm, after setting the tolerance for smoothing polylines (P_{10}). The schematic diagram of extracting G_{fcl} and the longest length of glaciers (G_{Lmax}) is shown in Fig.6.

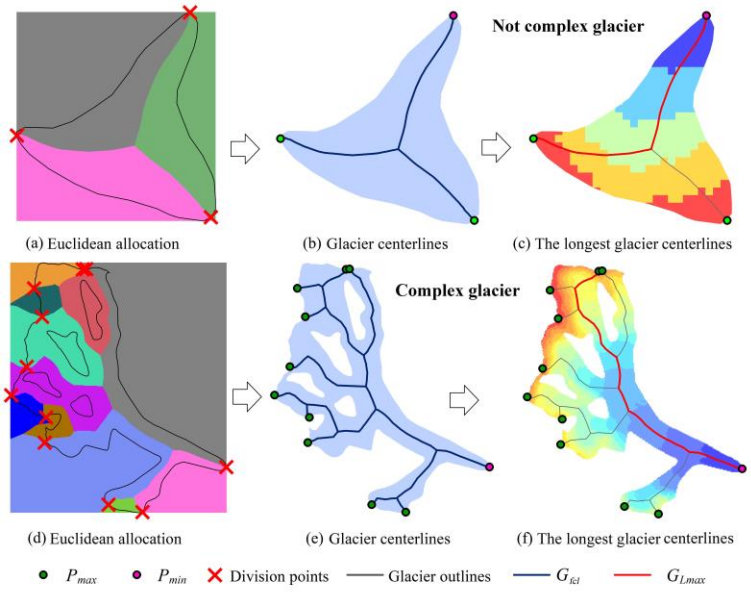


Figure 6: The schematic of extracting centerlines and the longest centerline of the glacier. (a) and (d) show the results after executing the European allocation, and the different colors represent the regions which have the shortest distance to the corresponding edges of the glacier; (b) and (e) represent the centerlines(G_{fcl}), the local highest point (P_{max}) and lowest point (P_{min}) of the glacier; (c) and (f) demonstrate the longest centerline (G_{Lmax}) of the glacier and the background is the digital elevation model with the graduated red (high)– blue (low) color.

3.3.6 Calculation of glacier length

The final code of the G_{fcl} was determined by P_{min} after G_{fcl} being converted from multipart to singlepart and was given in a unified format. Then all branches of glacier centerlines and glacier length were achieved using algorithm (Fig.7) similar to the critical path. This work consisted of four steps: (i) the polyline set of G_{fcl} was recorded as C (Eq.13), then the sets of polyline length (L) and polyline endpoint (S) (Eq.13) were obtained; (ii) the initial search point (B) (Eq.14), the end of glacier centerline, was determined by the coordinates of P_{min} based on the above steps.

The common endpoint set (N) (Eq.14) with the next parts of glacier centerlines was obtained, and then the polyline code corresponding to B was recorded; (iii) each element in N was used as a new starting point for search respectively (B') (Eq.15), which was used to get the common endpoint set (N') (Eq.15) with the next parts of glacier centerlines. The coding of the corresponding polyline set of each glacier branch was recorded separately and (vi) the above process continued until all branches of glacier centerline trace back to its corresponding P_{max} (Eq.16).

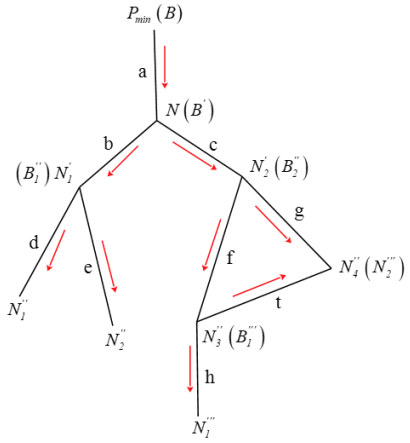


Figure 7: The schematic of calculating glacier length (The red arrow represents the search direction of the branches of glacier centerline).

$$C_i = \{C_{ij} | j \in [0, card(C)\}$$
$$S = \{(s_i, e_i) | s_i = C_{[i][0]}, and\ e_i = C_{[i][card(C_i)-1]}\} \quad (13)$$

$$B = \{k | P_{min} \in S_k\}, k \in [0, card(S)\big)$$
$$N = \{P | P \neq P_{min}, and\ P \in S_B\} \quad (14)$$

$$B' = \{k | N \in S_k, and\ k \neq B\}, k \in \big[0, card(S)\big)$$
$$N' = \left\{P \middle| P \neq N, and\ P \in S_{B_m'}\right\}, m \in \left[0, card\left(B'\right)\right) \quad (15)$$

$$res = \{\{a, b, d\}, \{a, b, e\}, \{a, c, f, h\}, \{a, c, g\}, \{a, c, f, t\}\} \quad (16)$$

The length of each branch of glacier centerlines was counted. The average length (Eq.17) of all branches was named as the average length of a glacier (G_{Lmean}). The longest length (Eq.18) of all branches was named as the longest length of a

glacier (G_{Lmax}). In addition, the part above ELA in G_{Lmax} was regarded as the accumulation region length (G_{Lacc}) of a glacier, and the part of G_{Lmax} with altitude lower than ELA was regarded as the ablation region length (G_{Labl}) of a glacier. Finally, the corresponding vector data were generated and some attributes including the corresponding polyline code, glacier code, the value of glacier length were added.

$$L_{mean} = \frac{SUM\left(L_{res_i}\right)}{card(res)} \quad (17)$$

$$L_{max} = Max\left(L_{res_i}\right) \quad (18)$$

4 Accuracy evaluation and the results

4.1 Methods of quality analysis

Here, we used the SCGI as the test data to run the designed program, including 48571 glaciers. The extraction results of some typical examples of glaciers (from simple to complex) are presented in Fig.8. The accuracy of glacier centerlines was evaluated based on a random verification method in this study. All glaciers (total quantity: N_G) corresponding to the samples were obtained and arranged in ascending order of the area. Specifically, 100 random integers were generated in the set of [0, N_G). Glaciers with corresponding serial number were exported as samples. After the visual inspection, the accuracy evaluation was conducted based on the following statistical analysis.

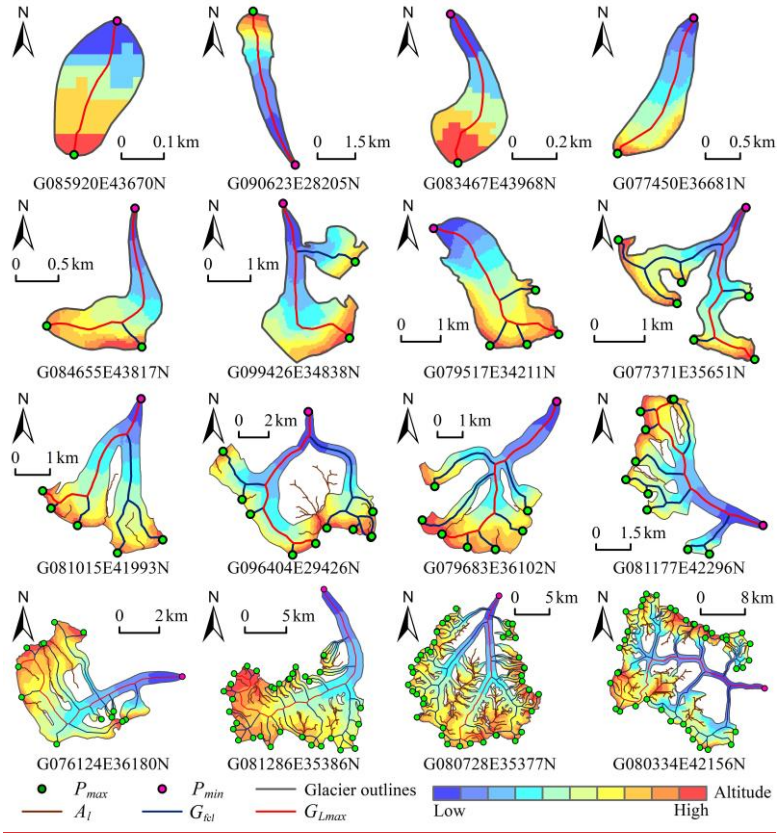


Figure 8: The centerlines for some typical glaciers **(P_{max} and P_{min} denote the local highest point and lowest point in the boundary of the glacier, respectively; A_l denotes the auxiliary lines; G_{fcl} and G_{Lmax} denote the centerlines and the longest centerline of the glacier).**

Firstly, 100 glaciers were randomly selected from the glacier dataset as samples to obtain a verification accuracy (R_1) (Eq.19). Secondly, each level of glaciers was separately taken as the total (N_T), and 100 glaciers were randomly selected. There were 5 samples for 5 levels, which were used to calculate a verification accuracy (R_2) (Eq.20) by taking the number proportion of each glacier level as the weight. Then, 100 glaciers with the largest, middle and smallest areas were selected separately as

samples. The verification accuracy (R_3) (Eq.21) was derived using 1:2:1 as the allocation proportion of weight. Finally, the average value of R_1, R_2 and R_3 was used as the comprehensive accuracy (R) (Eq.22). Among them, S_i represented the verification accuracy of the *ith* sample (i = {1,2,3,4,5,6,7,8,9}).

$$R_1 = S_1 \tag{19}$$

$$R_2 = \sum_{i=5}^{9} \frac{S_i \times N_{T_i}}{N_G} \tag{20}$$

$$R_3 = 0.25 \times S_2 + 0.5 \times S_3 + 0.25 \times S_4 \tag{21}$$

$$R = \frac{R_1 + R_2 + R_3}{3} \tag{22}$$

4.2 Sample selection and assessment criteria

Visual inspection in combination with satellite images and topographic maps is the most direct evaluation method for extraction results. Using 48571 glaciers in China as the test data, nine samples of 900 glaciers were selected for three verifications according to the evaluation method defined in section 4.1. The samples used for verification and relative information are given in Table 2.

Table 2 The information about validation samples.

Verification identifier	1-whole		2-area		3-levels				
Sample identifier	a	b	c	d	e	f	g	h	i
Selection conditions	Random	Max.	Central	Min.	Random				
Sample number	100	100	100	100	100	100	100	100	100
Total amount	48571		48571		38463	7341	2061	501	205
Proportion of sample (%)	0.21		0.62		0.26	1.36	4.85	19.96	48.78
Proportion of total (%)	100		100		79.19	15.11	4.24	1.03	0.42

Considering the possible defaults of the input data, we set some standards of accuracy evaluation (Table 3). The first level includes three categories: correct (I), inaccurate (II) and incorrect (III). The secondary categories were divided into 11 categories according to probable causes, among which the inaccurate causes and incorrect causes were subclassified as 6 types and 4 types, respectively. Type II involves mostly glaciers with accurate G_{Lmax} but missing, redundant or unreasonable branches of glacier centerlines. When calculating the comprehensive accuracy, category I and II were regarded as correct, and only III

was considered incorrect.

Table 3 The rules of accuracy assessment.

1st-level categories		2nd-level categories	
Code	Descriptions	Code	Descriptions
I	Correct	11	Correct
II	Inaccurate	21	Inaccurate glacier outlines
		22	Inaccurate identification of extreme points
		23	Inaccurate proximity algorithm for bare rock regions
		24	The influence of shunt or convergence in the glacier centerlines
		25	Inaccurate ridgelines
		26	Others (issues that are unknown by the algorithm itself, glaciers or DEM data)
III	Incorrect	31	Undivided glaciers
		32	Ice caps
		33	Slope glacier, i.e., ~~Slope~~ glaciers with little change in slope
		34	Others (unknown issues by the algorithm itself, issues with glaciers and DEM data, indistinguishable glacier types, etc.)

4.3 Statistics of different samples

According to the standards in Table 3, the selected samples were conducted with visual investigation. The results of nine samples were displayed in Fig.9. The statistical results showed that the accuracy of verification-2 was the highest (95.25%), followed by the ~~accuracy of~~ verification-3 ~~was the second~~ (94.76%) and the ~~the accuracy of~~ verification-1 ~~was~~ (93%). The comprehensive accuracy of glacier centerlines was 94.34%, of which category-I and category-II accounted for 86.06% and 8.28%, respectively. Meanwhile, we summarized the frequency of each type in each sample basing on 2nd-level categories. As seen in Fig.10, the problems of centerlines of small glaciers were mainly caused by the inaccurate selection of division points due to the insufficient accuracy of DEM (code: 22) and incorrect calculation results of some ~~slope~~ glaciers with little change in slope (code: 33). The problems of centerlines of large glaciers were mainly concentrated in some types coded in 31 and 32, which needed to be repartitioned and recalculated. In addition, a few problems were found in samples: the upper outlines of glacier were across the ridgeline; a small number of glaciers were not correctly segmented; the altitude in glaciers' DEM was abnormal. It implied that the reasonable glacier outlines and accurate DEM data were the prerequisite for extracting

glacier centerlines and calculating glacier length.

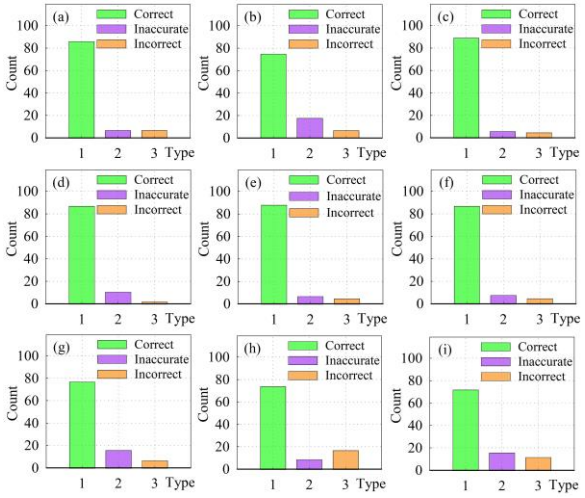
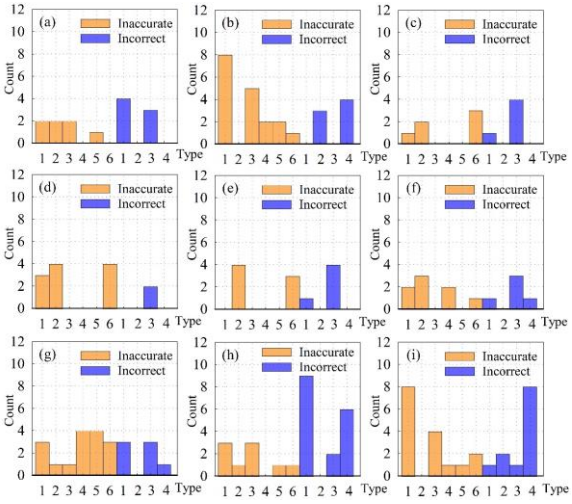


Figure 9: The statistical chart of evaluating results according to the 1st-level categories.



295

Figure 10: The statistical chart of evaluating results according to the 2nd-level categories.

4.4 Comparison to glaciers' maximum length from the RGI v6.0

4.4.1 The statistic of bit order and D_L

In the RGI v6.0, 38053 glaciers in the SCGI were adopted and accounted for 78.35% of the total glaciers in China, by checking the GLIMS_ID in both glacier datasets. As mentioned above, the field L_{max}, the longest glacier length, was contained in the RGI v6.0. In order to further verify the accuracy of glacier length calculated by this method, we calculated the difference (D_L) between G_{Lmax} and L_{max}, and then arranged them in ascending order to generate the distribution diagram of sequence-D_L (Fig.11). If D_L was negative, it meant that the G_{Lmax} of glaciers with the corresponding serial number was smaller than L_{max} and vice versa. Overall, there were only a small part of glaciers with extremely large $|D_L|$ at both ends (Fig.11). After visual inspection, G_{Lmax} was more consistent with the actual status of glaciers.

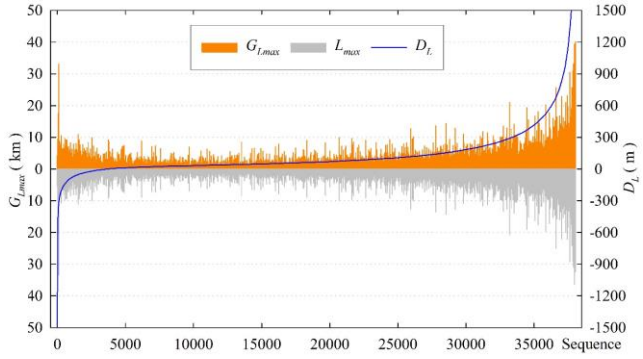


Figure 11: The statistical chart of the difference (D_L) of the longest glacier length between this dataset (G_{Lmax}) and the RGI v6.0 (L_{max}).

In addition, the average value of positive D_L, the average value of negative D_L and the number of glaciers in different levels were calculated (Fig.12). The size of three pixels for DEM was used as the statistical tolerance, which means glaciers within the tolerance range were regarded as consistent extraction results. Statistically, there were 22017 glaciers within tolerance, 925 glaciers with negative D_L and 15111 glaciers with positive D_L that are greater than the tolerance. In terms

of numerical comparison, G_{Lmax} obtained by our method was slightly larger than L_{max} in RGI v6.0.

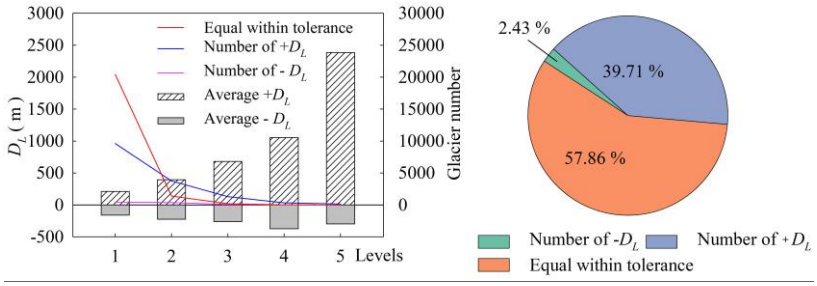


Figure 12: The statistical charts of the difference (D_L) of the longest length of glaciers by two methods in different glacier sizes.

设置了格式：字体：倾斜

设置了格式：字体：倾斜，下标

4.4.2 Analysis of abnormal D_L

Combining the designed algorithm with visual inspection, the preliminary analysis showed that the local abnormal DEM, inaccurate glacier outlines and some glacier types (such as ice cap, slope glacier, etc.) were the main causes of abnormal D_L (Fig.13). Slope glacier is typical multi-origin and multi-exit glacier with almost the same number of local highest points and local lowest points, which often exist in pairs (Fig.13-a). If the local highest point did not match the local lowest point, a value of positive D_L would occur (Fig.13-a, blue polyline). Local abnormalities in DEM generally resulted in a shorter G_{Lmax} (negative D_L), as showed in Fig.13-b. Some key local highest points could not be identified because of the inaccurate outlines, resulting in a large negative D_L (Fig.13-c). For non-single glacier, this algorithm could only identify a lowest point, and all branches of glacier centerlines converge to this point, which would increase the length of most branches and make G_{Lmax} to be too large or even wrong (Fig.13-d).

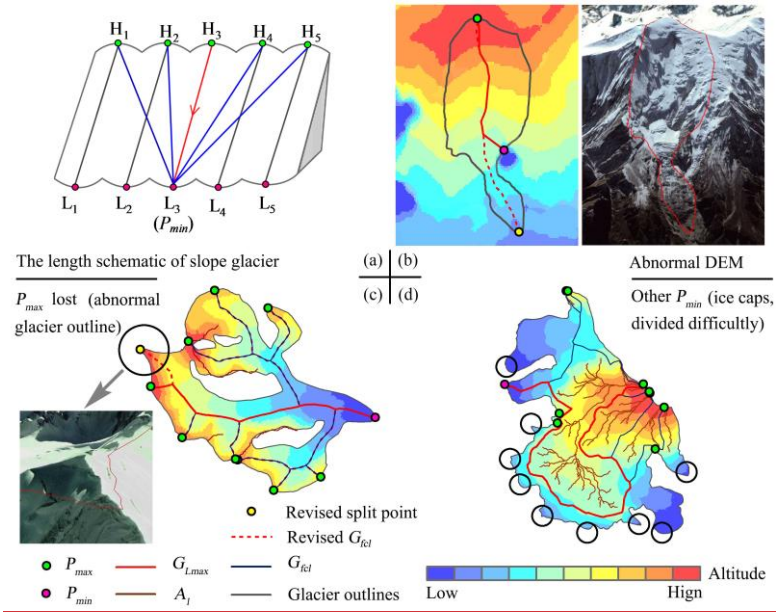


Figure 13: The schematic of probable causes for the abnormal of the longest glacier length. In Figure b, the red dashed line indicates the revised glacier centerline, and the yellow point is the correct lowest point (P_{min}). In Figure c, the red dashed line represents the missing branch, and the yellow point is a local highest point (P_{max}) missed by the algorithm. In Figure d, the black circle indicates some probable exits of the glacier, which needs to be divided into individual glaciers before extracting the centerlines.

The small or abnormal L_{max} of some glaciers was also the main reason of abnormal D_L. An abnormal example is shown in Fig.14. The Tugebieliqi Glacier (GLIMS_ID: G080334E42156N) with the maximum $|D_L|$ is the third largest glacier in China, behind the Sugatyanatjilga Glacier and the Tuomuer Glacier. Its G_{Lmax} was 40.179 km, but its L_{max} in the RGI v6.0 was only 11.703 km. The further measurement by Google Earth showed that the west-east length ($D_{W\text{-}E}$) of the glacier was about 27.72 km, which meant that our result was more conformable to reality.

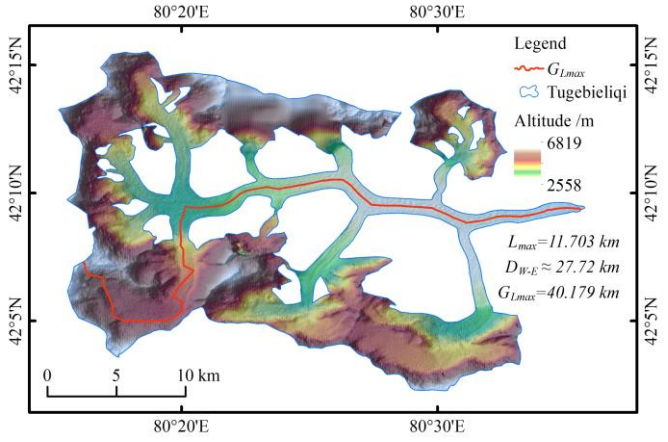


Figure 14: The schematic of the longest centerline of Tugebieliqi Glacier (L_{max}: the corresponding length of this glacier in the RGI v6.0; $D_{W\text{-}E}$: the distance from west to east of this glacier; G_{Lmax}: the length calculated by our method).

5 Discussion

5.1 Performance of the algorithm

In the process of extracting centerlines of glaciers in China, all glaciers were equally divided into eight tasks according to the number and considering the running efficiency of the algorithm. Based on the actual extraction results, five glaciers that failed to execute were added as the ninth task. Tasks coded T1~T9 were executed in the working environment of ArcGIS 10.4 software. Except for T7 and T9 using a Lenovo G410 (processors: Intel(R) Core(TM) i5-4210M CPU @ 2.60 GHz; memories: 4GB DDR3L 1600 MHz; video card: AMD Radeon R5 M230 2GB Discrete graphics) of home laptops, the other seven tasks used seven Dell OptiPlex 7040 (processors：Intel(R) Core(TM) i7-6700 CPU @ 3.40 GHz; memories: 8GB DDR4 2633 MHz; video card: AMD Radeon(TM) R5 340X 2GB Integrated graphics) of the tower server with the same configuration. The task distribution and execution results of the tests are given in Table 4.

Table 4 The statistics of assigning tasks and results of execution in tests.

Task ID	Assigned amount	Completed amount	Completion rate (%)	Total ~~T~~time (h)	Average time (s)

T1	6000	6000	100	31.00	18.60
T2	6000	6000	100	29.75	17.85
T3	6000	5999	99.98	30.53	18.32
T4	6000	6000	100	29.34	17.61
T5	6000	6000	100	33.54	20.12
T6	6000	5999	99.98	31.62	18.97
T7	6000	5999	99.98	58.63	35.18
T8	6571	6569	99.97	38.27	20.97
T9	5	5	100	0.12	86.26
Total	48571	48571	100	282.81	20.96

The results of the tests showed that the program took an average of 20.96 s to ~~extract~~ process an individual glacier, ~~and~~ whereas it spent 86.26 s or even longer for some complex glaciers. Among the first eight processing tasks, T4 took the least time. The main reason was that the assigned glaciers in this task were mostly small and complex glaciers were less, except for the higher machine configuration. T7 took the longest time, and the cause was the lower machine configuration. The results of all tasks were merged to obtain the centerline dataset of the SCGI. It contained seven vector files (56 items) and nine logs, which took up about 912 MB in the storage.

5.2 Influence of glacier outline quality and DEM

The extraction method of glacier centerlines belongs to geometric graphic algorithm and depends on glacier outlines. Natively, comparing with the previous studies, our method has similar problems: (i) the delayed shunt and early convergence of the branches and (ii) the centerlines of same glacier in different periods, which is not geometrically comparable for some glaciers in drastic changes of outlines. The extraction results also showed that the branches of some glacier centerlines did have delayed diversion or early convergence, while the impact on the simulation of glacier's main flowline was limited. Considering that the results of extracting glacier centerlines change with the changes of glacier outlines, the measurement of the length change of glaciers in different periods will be the focus of our future work. We may further design a new algorithm to automatically supplement, extend, delete or modify the benchmarking glacier centerlines, so as to measure the changes of centerlines and length of glaciers in different periods.

Bare rock region refers to the non-glacial component that is within the outer boundary of the glacier outlines but is not covered by snow or ice. It can be divided into two types: one is the exposed rock protruding on the glacier surface; the other is the cliff generally existing between the upper part of the glacier and the firn basin. The snow or ice on the upper part of the glacier enters the firn basin through the cliffs. And the snow or ice on the cliffs are also important sources of replenishment for firn basin. So the cliffs are theoretically considered to be part of the glacier. However, the cliffs may be similar to the bare rock area during the ablation season, and the cliffs are often accompanied by the presence of image shadows, which will easily cause misjudgments of glacier outlines in interpretation.

Determining the ownership of bare rock regions in G_{fl} will improve the quality of glacier centerlines. In this study, all bare rock regions were considered to be the first category, and such cases were handled accordingly. The first category was divided into two types: (i) the bare rock area on the upper part of the glacier being equivalent to the ice divide and (ii) the bare rock area near the end of the glacier. The attribution of most bare rock areas in the upper part of the glacier can be determined by the intersection point of A_l, G_{pl} with G_{br}. Only a few bare rock areas still exist alone, Eq. (12) was required to determine the segments of the G_{fl} to which they belong. Some bare rock areas located in the ablation area were allowed to exist alone in the G_{fl}, and the probability of their existence was extremely low.

The determination of glacier's ELA is difficult. Some scholars believed that each glacier has its own ELA (Cui and Wang, 2013; Sagredo et al., 2016), but other scholars argued that the ELA of all glaciers in a certain region is the same (Sagredo et al., 2014; Jiang et al., 2018). The measurement of ELA requires continuous and long-term observation data, so it is very difficult to determine the ELA of the glaciers in large-scale. In this study, the ELA used to distinguish between the accumulation area and the ablation area of the glacier was estimated by calculating the median of elevation (Z_{med}). For some glaciers (such as calving glaciers), the Z_{med} is above the actual ELA, which has been reasonably explained by scholars (Braithwaite and Raper, 2009). And it was considered that this overestimation is unlikely to affect the automatic calculation of glacier length (Machguth and Huss, 2014).

5.3 Some other factors influencing centerline of glaciers

Automatic extraction of glacier centerlines was basically carried out during the processing of polylines, so the processing algorithm of polylines in the program occupied a considerable part of codes. Among them, several common problems of disconnected polylines are shown in Fig.15. The following four types are important, which have a great influence on the accuracy and extraction automation of glacier centerlines.

(i) During the post-processing of the auxiliary lines, due to the inaccuracy of ice divide or the problems of DEM, the ridgelines in the edge of the ice divide of some glaciers start at the G_{pl} and end up with the G_{pl} or in parallel along the G_{pl}, which are unreasonable. In response to this problem, the algorithm set corresponding rules for screening in the processing of auxiliary lines, reducing the impact of such problems as much as possible.

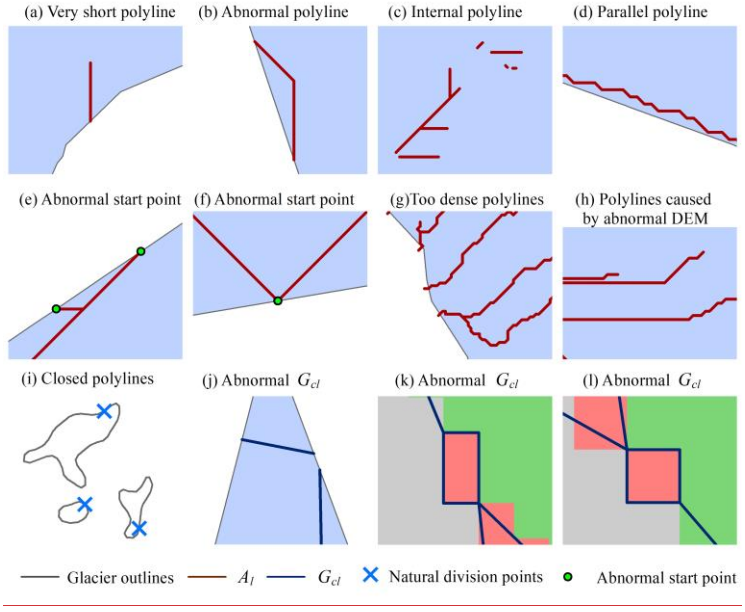


Figure 15: The schematic of discontinuous short polylines. Subgraphs a-h represent type (i), i represents type (ii), j represents type (iii) and k-l represent type (iv). The background in subgraphs a-h and j represent glacier-covered areas. Subgraph i shows several closed polylines, which does not fill background color. The different background colors in subgraphs k-l represent different areas of the glacier surface after the European allocation.

(ii) The visually closed vector polyline is not completely closed. Its start and end are at the same point, which is equal to a natural division point. Unless the natural division point of G_{pl} completely coincides with a certain division point, the number of polyline records in the G_{pl} after division will be one more than we expected. Therefore, it is necessary to identify the natural division point during processing and merge the two disconnected polyline records.

(iii) The algorithm of Euclidean allocation is accomplished based on raster operation, which is equivalent to the equidistant scatter operation with the interval of P_8 on the glacier surface. For some glaciers with horizontal or vertical distribution of the G_{pl}, the extraction will continue after the centerlines overlaps with the G_{pl}. We only need to design the corresponding functions to detect and delete this redundancy of the disconnected polylines.

(iv) In the process of calling the module of Euclidean allocation to generate the centerlines, there is a slight probability that pixels with strictly equal distances will appear. The central axis will generate a regular rectangle based on the raster pixel corresponding to the central point, which will affect the calculation of the G_{Lmax}. In the algorithm, a function to identify and deal with such problems was added after the Euclidean allocation, then the polylines on one side of the diagonal of a rectangle were randomly retained.

6 Conclusions

An automatic method for extracting glacier centerlines based on Euclidean allocation in two-dimensional space was designed and implemented in this study. It only needs the glacier outlines and the corresponding DEM to automatically generate the vector data of glacier centerlines, and provides different properties including the longest length, the average length, the length in the ablation region, the length in the accumulation region of the glacier. The standardized and automatic extraction of

glacier centerlines requires no manual intervention. Meanwhile, we used the SCGI as the test data to run the program and verify its efficiency. The success rate of extracting glacier centerlines was very close to 100% and the comprehensive extraction accuracy reached 94.34%, which reflected the robustness and simplicity of this method.

The automatic extraction algorithm proposed has three advantages: (i) introducing the auxiliary reference lines which ensure the validity of the upper glacier centerlines; (ii) success in automatically obtain the longest centerline of each glacier and the branches of glacier centerlines; (iii) providing more information of glacier lengths than other methods proposed by some scholars. Compared with the longest length of each glacier in the RGI v6.0, the length of the corresponding glacier calculated by our algorithm is in better agreement with the actual length of the glacier. We also identified the possible causes affecting the accuracy of glacier centerlines. In the future, we will focus on improving the time efficiency of the algorithm, providing the updated datasets of glacier centerlines with higher-quality, and identifying the abnormal glacier phenology such as glacier surging rapidly.

Appendix A

The paper uses numerous abbreviations. Explanations of main acronyms are listed in Table A1.

Table A1 The list of main acronyms in this study.

Acronyms	Description
A_t	The given area of an equilateral triangle
A_g	The polygon's area of the glacier's outer boundary
A_l	The final auxiliary line
A_r	The ridgelines of the glacier surface
G_{br}	The bare rock in glacier
G_{fcl}	The final glacier centerline
G_{fl}	The feature lines of glacier surface
G_{cl}	Glacier centerline
G_{Labl}	The length in the ablation region of the glacier
G_{Lacc}	The length in the accumulation region of the glacier
G_{Lmax}	The longest length of the glacier

G_{Lmean}	The average length of the glacier
G_{pl}	The polyline of the outer boundary of the glacier
G_{po}	The polygon of the outer boundary of the glacier
L_{max}	The longest glacier length of RGI v6.0
D_L	The difference between G_{Lmax} and L_{max}
P_t	The given perimeter of an equilateral triangle
P_g	The perimeter of the glacier's outer boundary
P_{max}	The local highest point of glacier outline
P_{min}	The lowest point of glacier outline
RGI	The Randolph Glacier Inventory
SCGI	The Second Chinese Glacier Inventory
Z_{med}	The median elevation of the glacier

Code availability

The code used to support the findings of this study are available in the Supplement~~from the corresponding author upon request~~.

Data availability

The datasets including the SCGI, RGI v6.0 and SRTM1 DEM v3.0 used in this study are freely available. The database of glacier centerlines in the SCGI produced in this study are available from the corresponding author upon request.

Supplement

The Supplement consists of three parts. They are (i) automatic extraction software of glacier centerlines, (ii) software instructions, and (iii) some test results.

Author contribution

Xiaojun Yao designed this algorithm of extracting glacier centerlines and edited the manuscript. Dahong Zhang implemented the program and wrote the first draft of the manuscript. Hongyu Duan tested the program and checked the quality of glacier

centerlines. Shiyin Liu, Wanqin Guo, Meiping Sun and Dazhi Li reviewed and edited the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We thank editors and two reviewers for their valuable comments that improved the manuscript. Dahong Zhang thanks his current supervisor, Professor Shiqiang Zhang of Northwest University in Xi'an, China for valuable ideas and suggestions. This research was funded by the National Natural Science Foundation of China (No.41861013, No.42071089, No.41801052), the Open Research Fund of National Earth Observation Data Center (No. NODAOP2020007) and the Open Research Fund of National Cryosphere Desert Data Center (No. 20D02).

References

Braithwaite, R. J., and Raper, S. C. B.: Estimating equilibrium-line altitude (ELA) from glacier inventory data, Annals of Glaciology, 50, 127-132, doi: 10.3189/172756410790595930, 2009.

Cui, H., and Wang, J.: The methods for estimating the equilibrium line altitudes of a glacier, Journal of Glaciology and Geocryology, 35, 345-354, doi: 10.7522/j.issn.1000-0240.2013.0041, 2013.

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, doi: 10.1029/2005rg000183, 2007.

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, doi: 10. 12118 /j. issn.1000- 6060.2018.06. 08, 2018.

Gao, Y. P., Yao, X. J., Liu, S. Y., Qi, M. M., Duan, H. Y., Liu, J., and Zhang, D. H.: Remote sensing monitoring of advancing glaciers in the Bukatage Mountains from 1973 to 2018, Journal of Natural Resources, 34, 1666-1681, doi: 10.31497/zrzyxb.20190808, 2019.

Guo, W., Liu, S., Xu, J., Wu, L., Shangguan, D., Yao, X., Wei, J., Bao, W., Yu, P., Liu, Q., and Jiang, Z.: The second Chinese glacier inventory: data, methods and results, Journal of Glaciology, 61, 357-372, doi: 10.3189/2015JoG14J209, 2017.

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, doi: 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, doi: 10.1017/aog.2017.15, 2017.

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, doi: 10.1007/s11629-016-4243-5, 2017.

Jiang, D. B., Liu, Y. Y., and Lang, X. M.: A multi-model analysis of glacier equilibrium line altitudes in western China during the Last Glacial Maximum, Science China Earth Sciences, 49, 1231-1245, doi: 10.1360/n072018-00058, 2018.

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, doi: 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, doi: 10.1016/j.cageo.2012.10.014, 2013.

Leclercq, P. W., and Oerlemans, J.: Global and hemispheric temperature reconstruction from glacier length fluctuations, Climate Dynamics, 38, 1065-1079, doi: 10.1007/s00382-011-1145-7, 2012.

Leclercq, P. W., Pitte, P., Giesen, R. H., Masiokas, M. H., and Oerlemans, J.: Modelling and climatic interpretation of the length fluctuations of Glaciar Frías (north Patagonian Andes, Argentina) 1639–2009 AD, Climate of the Past, 8, 1385-1402, doi: 10.5194/cp-8-1385-2012, 2012a.

Leclercq, P. W., Weidick, A., Paul, F., Bolch, T., Citterio, M., and Oerlemans, J.: Brief communication "Historical glacier length changes in West Greenland", The Cryosphere, 6, 1339-1343, doi: 10.5194/tc-6-1339-2012, 2012b.

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, doi: 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, 1020-1030, doi: 10.1029/2011jf002104, 2012.

Linsbauer, A., Paul, F., and Haeberli, W.: Modeling glacier thickness distribution and bed topography over entire mountain ranges with GlabTop: Application of a fast and robust approach, Journal of Geophysical Research: Earth Surface, 117, 3007-3024, doi: 10.1029/2011jf002313, 2012.

Liu, S. Y., Yao, X. J., Guo, W. Q., Xu, J. L., Shang Guan, D. H., Wei, J. F., Bao, W. J., and Wu, L. Z.: The contemporary glaciers in China based on the Second Chinese Glacier Inventory, Acta Geographica Sinica, 70, 3-16, doi: 10.11821/dlxb201501001, 2015.

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, doi: 10.5194/tc-8-1741-2014, 2014.

McNabb, R. W., Hock, R., O'Neel, S., Rasmussen, L. A., Ahn, Y., Braun, M., Conway, H., Herreid, S., Joughin, I., Pfeffer, W. T., Smith, B. E., and Truffer, M.: Using surface velocities to calculate ice thickness and bed topography: a case study at Columbia Glacier, Alaska, USA, Journal of Glaciology, 58, 1151-1164, doi: 10.3189/2012JoG11J249, 2017.

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, doi: 10.3189/2014JoG13J181, 2017.

Muhuri, A., Natsuaki, R., Bhattacharya, A., and Hirose, A.: Glacier Surface Velocity Estimation Using Stokes Vector Correlation, Synthetic Aperture Radar, Singapore, 2015, 9781467372961, 606-609,2015.

Nuth, C., Kohler, J., König, M., von Deschwanden, A., Hagen, J. O., Kääb, A., Moholdt, G., and Pettersson, R.: Decadal changes from a multi-temporal glacier inventory of Svalbard, The Cryosphere, 7, 1603-1621, doi: 10.5194/tc-7-1603-2013, 2013.

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, doi: 10.1017/S0260305500012489 1997.

Paul, F., Barry, R. G., Cogley, J. G., Frey, H., Haeberli, W., Ohmura, A., Ommanney, C. S. L., Raup, B., Rivera, A., and Zemp, M.: Recommendations for the compilation of glacier inventory data from digital sources, Annals of Glaciology, 50, 119-126, doi: 10.3189/172756410790595778, 2009.

Sagredo, E. A., Rupper, S., and Lowell, T. V.: Sensitivities of the equilibrium line altitude to temperature and precipitation changes along the Andes, Quaternary Research, 81, 355-366, doi: 10.1016/j.yqres.2014.01.008, 2014.

Sagredo, E. A., Lowell, T. V., Kelly, M. A., Rupper, S., Aravena, J. C., Ward, D. J., and Malone, A. G. O.: Equilibrium line altitudes along the Andes during the Last millennium: Paleoclimatic implications, The Holocene, 27, 1019-1033, doi: 10.1177/0959683616678458, 2016.

Schiefer, E., Menounos, B., and Wheate, R.: An inventory and morphometric analysis of British Columbia glaciers, Canada, Journal of Glaciology, 54, 551-560, doi: 10.3189/002214308785836995 2008.

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, doi: 10.3189/172756407782871143 2007.

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

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, doi: 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, doi: 10.7522/j. issn.1000-0240.2015. 0173, 2015.

Zhang, W., and Han, H. D.: A review of the ice volume estimation of mountain glaciers, Journal of Glaciology and Geocryology, 38, 1630-1643, doi: 10.7522 /j. issn.1000-0240.2016.0191, 2016.