



A new method for deriving glacier centerlines applied to glaciers in Alaska and northwest Canada

C. Kienholz¹, J. L. Rich¹, A. A. Arendt¹, and R. Hock^{1,2}

¹Geophysical Institute, University of Alaska Fairbanks, USA

²Department of Earth Sciences, Uppsala University, Uppsala, Sweden

Correspondence to: C. Kienholz (christian.kienholz@gi.alaska.edu)

Received: 6 October 2013 – Published in The Cryosphere Discuss.: 25 October 2013

Revised: 28 January 2014 – Accepted: 10 February 2014 – Published: 25 March 2014

Abstract. This study presents a new method to derive centerlines for the main branches and major tributaries of a set of glaciers, requiring glacier outlines and a digital elevation model (DEM) as input. The method relies on a “cost grid–least-cost route approach” that comprises three main steps. First, termini and heads are identified for every glacier. Second, centerlines are derived by calculating the least-cost route on a previously established cost grid. Third, the centerlines are split into branches and a branch order is allocated. Application to 21 720 glaciers in Alaska and northwest Canada (Yukon, British Columbia) yields 41 860 centerlines. The algorithm performs robustly, requiring no manual adjustments for 87.8 % of the glaciers. Manual adjustments are required primarily to correct the locations of glacier heads (7.0 % corrected) and termini (3.5 % corrected). With corrected heads and termini, only 1.4 % of the derived centerlines need edits. A comparison of the lengths from a hydrological approach to the lengths from our longest centerlines reveals considerable variation. Although the average length ratio is close to unity, only ~ 50 % of the 21 720 glaciers have the two lengths within 10 % of each other. A second comparison shows that our centerline lengths between lowest and highest glacier elevations compare well to our longest centerline lengths. For > 70 % of the 4350 glaciers with two or more branches, the two lengths are within 5 % of each other. Our final product can be used for calculating glacier length, conducting length change analyses, topological analyses, or flowline modeling.

1 Introduction

Glacier centerlines are a crucial input for many glaciological applications. For example, centerlines are important for determining glacier length changes over time (Leclercq et al., 2012; Nuth et al., 2013), analyzing velocity fields (Heid and Kääb, 2012; Melkonian et al., 2013), estimating glacier volumes (Li et al., 2012; Linsbauer et al., 2012), and one-dimensional modeling of glaciers (Oerlemans, 1997a; Sugiyama et al., 2007). Also, glacier length, derived from centerlines, is an important parameter for glacier inventories (Paul et al., 2009).

So far, most of the above applications have relied on labor-intensive manual digitization of centerlines. The few studies that derive centerlines fully automatically use a hydrological approach and/or derive only one centerline per glacier (Schiefer et al., 2008; Le Bris and Paul, 2013). Here, we present a new algorithm that allows for deriving multiple centerlines per glacier based on a digital elevation model (DEM) and outlines of individual glaciers. Moreover, this algorithm splits the centerlines into branches and classifies them according to a geometry order. The approach is tested on all glaciers in Alaska and adjacent Canada with an area > 0.1 km², corresponding to 21 720 out of 26 950 glaciers. We carry out a quality analysis by visual inspection of the centerlines, and compare the derived lengths to the lengths obtained from alternative approaches.

2 Previous work

The automatic derivation of glacier centerlines and the resulting glacier length is challenging (Paul et al., 2009; Le Bris and Paul, 2013). Consequently, only few automated approaches have been proposed so far (Schiefer et al., 2008; Le Bris and Paul, 2013) and often centerlines have been digitized manually even for large-scale studies (e.g., Burgess et al., 2013b).

In their large-scale study, Schiefer et al. (2008) applied an automated approach to all British Columbia glaciers that is based on hydrology tools. For each glacier, this approach derives one line that represents the maximum flow path that water would take over the glacier surface. Schiefer et al. (2008) find that these lengths are 10–15 % longer than distances measured along actual centerlines. Because lower glacier areas (\sim ablation areas) are typically convex in cross section, their flow paths are deflected toward the glacier margins, which leads to length measurements that are systematically too long in these cases. While systematically biased lengths in a glacier inventory can be corrected for, the actual lines need major manual corrections before they can be used in glaciological applications such as flowline modeling. Therefore, Paul et al. (2009) suggest the use of the above automated algorithm in the concave (i.e., higher) part of the glacier, combined with manual digitization in the convex (lower) part of the glacier.

Le Bris and Paul (2013) present an alternative method for calculating glacier centerlines based on a so-called “glacier axis” concept, which derives one centerline per glacier between the highest and the lowest glacier elevation. Le Bris and Paul (2013) first establish a line (the “axis”) between the highest and the lowest glacier point, which is then used to compute center points for the glacier branches. These center points are connected, starting from the the highest glacier elevation and following certain rules (e.g., “always go downward”, “do not cross outlines”). Smoothing of the resulting curve leads to the final centerline. The algorithm is applicable to most glacier geometries, yielding results similar to manual approaches (Le Bris and Paul, 2013). A limitation of their approach is the fact that the derived centerline between the highest and lowest glacier elevation does not necessarily represent the longest glacier centerline or the centerline of the main branch, both of which are frequently required in glaciological applications. For example, for glacier inventories, it is recommended to measure glacier length along the longest centerline, or, alternatively, to average the lengths of all glacier branches (Paul et al., 2009).

3 Test site and data

Our algorithm is tested on glaciers located in Alaska and adjacent Canada (Fig. 1a). For brevity, we hereafter refer to these glaciers as Alaska glaciers. From the complete Alaska

glacier inventory (Arendt et al., 2013), we extract all glaciers with a minimal area threshold of $> 0.1 \text{ km}^2$, thus eliminating small glacierets and possible perennial snowfields. This results in 21 720 glaciers with $86\,400 \text{ km}^2$ of ice total, which accounts for more than 99 % of the area of the complete Alaska inventory.

The Alaska glacier inventory comprises glacier outlines derived from satellite imagery taken between ~ 2000 and 2012. The outlines used herein either stem from manual digitization at the University of Alaska Fairbanks or from automated band ratioing followed by visual quality checks and manual corrections (Bolch et al., 2010; Le Bris et al., 2011, Fig. 1a).

Four DEM products are combined to create a continuous 60 m DEM consistent with the time span covered by the glacier outlines (Fig. 1b). South of 60° N , we rely on the Shuttle Radar Topography Mission (SRTM) DEM (<http://www2.jpl.nasa.gov/srtm>, last access: 25 July 2013), taken in February 2000 (Farr et al., 2007). Over Alaska, the SRTM DEM has a native spatial resolution of 30 m, while the resolution is 90 m over Canada. North of 60° N , we use a high-quality DEM derived from airborne interferometric synthetic aperture radar (IFSAR) data obtained in 2010 (Geographic Information Network of Alaska, GINA; <http://ifsar.gina.alaska.edu>; last access: 25 July 2013). For areas not covered by the IFSAR DEM, we use DEMs derived from data from the high-resolution stereo (HRS) imaging instrument onboard the Système Pour l’Observation de la Terre (SPOT) satellite taken within the scope of the SPIRIT program (time span 2007–2008; Korona et al., 2009) and the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) instrument onboard the Terra satellite (ASTER GDEM2, time span 2000–2011; Tachikawa et al., 2011; <http://asterweb.jpl.nasa.gov/gdem>; last access: 25 July 2013). The individual DEM tiles are merged into one data set giving priority to the highest-quality DEM available for each area. While the used DEMs represent roughly the same time span, the goals and scopes of the individual campaigns are different, which becomes apparent in the contrasting quality of the DEMs. For example, the GDEM has nearly global coverage, but limited quality, while the IFSAR DEM is of high quality, but only available for parts of Alaska.

4 Method

Our goal is to design an algorithm that (1) creates centerlines for the main glacier branches as well as major tributaries, (2) yields a quality comparable to a manual approach, and (3) requires minimal data and manual intervention. It would be desirable to derive centerlines that represent actual flowlines (i.e., ice trajectories); however, this would require coherent velocity fields without gaps. While corresponding algorithms are applied for single glaciers (e.g., McNabb et al., 2012), stringent velocity data requirements make large-scale

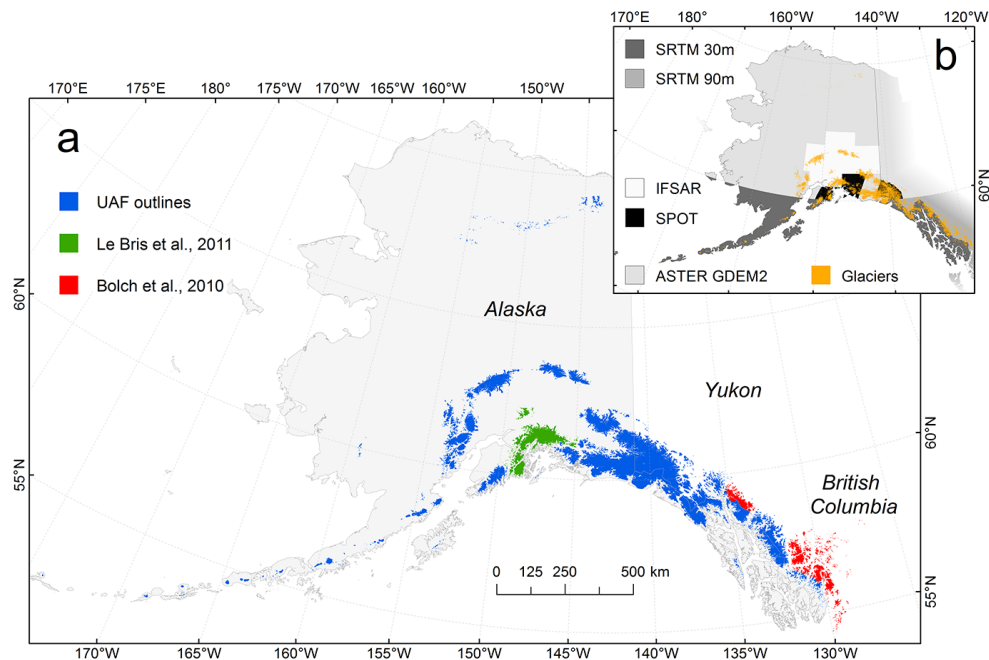


Fig. 1. The study area comprising glaciers of Alaska and adjacent Canada. **(a)** The three main outline sources marked in colors. **(b)** The DEM products used within the study.

applications difficult. Here, we aim at obtaining centerlines that are close to flowlines by only using glacier outlines and a DEM as input. Such centerlines differ from flowlines mostly in areas where centerlines from different glacier branches converge. Actual flowlines would not converge completely, but run side by side to the glacier terminus, as illustrated, for example, by Farinotti et al. (2009). Although not entirely consistent with real flowlines, centerlines are used for modeling purposes (e.g., Oerlemans, 1997b; Leclercq et al., 2012). They can also be used for applications that generally have lower requirements than does modeling (e.g., determining length changes, deriving glacier length for inventories).

Here, we apply a method that we call a “cost grid–least-cost route approach”. The workflow consists of three main steps that are implemented using the Python™ programming language and ESRI®’s arcpy module. The first step comprises the identification of glacier termini and glacier heads. The second step encompasses the calculation of a cost grid, followed by determining and optimizing the least-cost route to derive the glacier centerlines. In the third step, these centerlines are split into branches and a geometry order is introduced. Because the three steps are associated with different uncertainties, we have separate modules for each step of the algorithm.

4.1 Step 1 – Identification of glacier heads and termini

In the first step, we identify glacier heads and termini by applying a search algorithm based on the DEM and glacier

outline. We aim to derive one terminus per glacier and one head for each major glacier branch. The centerlines eventually run from each glacier head to the terminus.

4.1.1 Glacier terminus

A natural way of identifying the glacier terminus is by extracting the lowest glacier cell (e.g., Le Bris and Paul, 2013). To better constrain our lowest point to the actual terminus, we apply the corresponding query on a low-pass filtered and “filled” DEM. “Filling” refers to removing depressions within the DEM that could hamper the identification of the actual glacier terminus. In this method, filling and filtering are most important for large receding glaciers, which end in flat terrain and are generally characterized by a rough surface with numerous depressions.

Figure 2 shows the glacier terminus as automatically obtained for Gilkey Glacier, an outlet glacier of the Juneau Icefield located in southeast Alaska.

4.1.2 Glacier heads

Since we aim to derive centerlines of all major glacier branches, we need to identify heads for each of these branches. A three-step procedure is adopted to identify these heads (Fig. 3). First we identify local elevation maxima along the glacier outlines. Our algorithm samples the 60 m DEM in predefined steps (100 m) along the glacier outline, including nunataks. The 100 m sampling distance precludes repetitive sampling of the same DEM cell. Each sampled elevation is

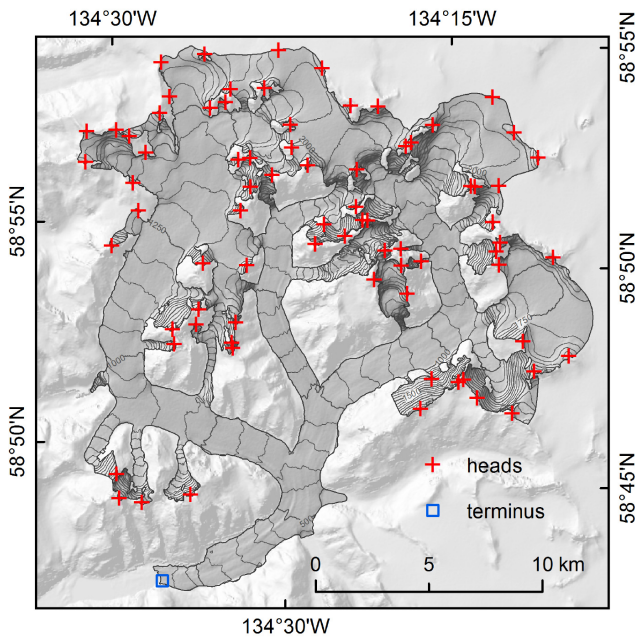


Fig. 2. Automatically derived glacier heads (red crosses) and terminus (blue square) on Gilkey Glacier, Juneau Icefield, southeast Alaska. The 50 m contours and the shaded-relief background are derived from the SRTM DEM.

then compared to its neighboring points along the outline. A possible glacier head is identified if the local point is higher than its neighbors (five neighbors in each direction, Fig. 3a) and if the point is higher than the lowest one-third of the elevation distribution of all the sampled points of the corresponding glacier (Fig. 3b). We apply the second criterion since glacier heads are typically located at higher elevations. We also want to avoid assigning centerlines to low-lying minor tributaries.

Given the irregular shapes of typical glacier outlines, the workflow above can result in multiple heads per glacier branch. To remove multiple heads per branch, we introduce a minimum linear distance r (m) that the derived heads must be apart. Since larger glaciers tend to have wider basins, we define r as a function of glacier area S (m^2):

$$r = \begin{cases} q_1 \cdot S + q_2 & : r \leq r_{\max} \\ r_{\max} & : r > r_{\max} \end{cases} \quad (1)$$

q_1 , q_2 and r_{\max} are constants given in Table 1, constraining r to values between 500 and 1000 m. In case one or more heads are within r , we only retain the highest head and erase all others (Fig. 3c). If two heads are apart by a distance less than r but separated by ice-free terrain, both heads are retained. Ice-free terrain is identified if the circle defined by r splits the glacierized area into two or more parts (Fig. 3d).

In case no head is identified using the above steps, which can be the case for small glaciers, we identify the highest

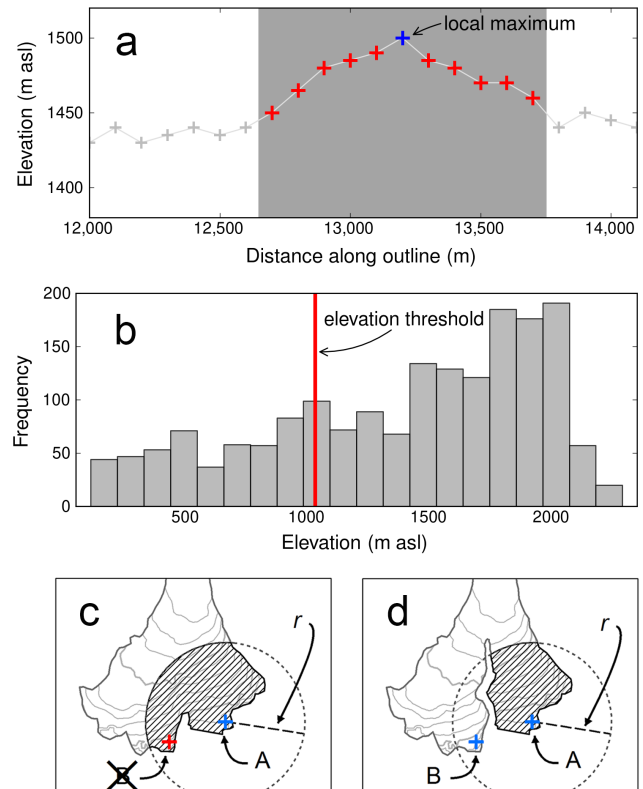


Fig. 3. Three-step procedure to identify the heads of major glacier branches. (a) Identification of local maxima along the glacier outline by comparing each point to its five neighbors in each direction. (b) Histogram of the elevation distribution of all sampled points along the glacier outline. Only points above the elevation threshold are retained. (c, d) Glacier area covered by circle with radius r around identified head A; heads A and B are separated by less than r . In case (c), head B (with lower elevation than head A) is eliminated. In case (d), both heads are retained because the ice-free ridge completely separates the two heads (i.e., splits the circle into two disconnected parts).

glacier elevation as the glacier head. In the case of Gilkey Glacier, 77 heads are identified (Fig. 2).

4.2 Step 2 – Establishment of cost grid and determination of least-cost route

This step establishes a cost or penalty grid (here used as synonyms) with high values at the glacier edge and in upper reaches of the glacier. The penalty values decrease towards the glacier center as well as towards lower elevations. The least-cost route from a glacier head to the glacier terminus yields the centerline.

4.2.1 Cost/penalty grid and route cost

First, we create a penalty grid with $10 \text{ m} \times 10 \text{ m}$ cell size according to Eq. (2). We choose this small cell size to obtain a detailed representation of the glacier outlines in the gridded

Table 1. Used parameter values and units.

Symbol	Value	Unit	Equation
q_1	2×10^{-6}	m^{-1}	1
q_2	500	m	1
r_{\max}	1000	m	1
f_1	1000	–	2
f_2	3000	–	2
a	4.25	–	2
b	3.5–4.0*	–	2
u_1	2×10^{-6}	m^{-1}	4
u_2	200	m	4
l_{\max}	400	m	4
j_1	0.1	–	7
j_2	0.01	m^{-j_3}	7
j_3	0.7	–	7
w_1	1×10^{-6}	m^{-1}	8
w_2	150	m	8
k_{\max}	650	m	8

* Derived from optimization (Sect. 4.2.2).

map. A larger cell size would yield a coarser glacier grid omitting small-scale features such as small nunataks. The penalty value p_i of each grid cell i within the glacier is computed by

$$p_i = \left(\frac{\max(d) - d_i}{\max(d)} \cdot f_1 \right)^a + \left(\frac{z_i - \min(z)}{\max(z) - \min(z)} \cdot f_2 \right)^b, \quad (2)$$

where d_i is the Euclidean distance from cell i to the closest glacier edge and z_i is the cell’s elevation. $\max(d)$ is the glacier’s maximum Euclidean distance from any point to its closest edge, and $\max(z)$ and $\min(z)$ are the glacier’s maximum and minimum elevation. The coefficients f_1 and f_2 and the exponents a and b are given in Table 1.

The first term, $\frac{\max(d) - d_i}{\max(d)}$, normalizes d_i to the maximum Euclidean distance found on the glacier, and leads to a Euclidean penalty contribution that ranges between zero and one (i.e., zero at the cell(s) with $\max(d)$ and one along the glacier edge). By multiplying the term with f_1 , we scale the normalized values to a range between zero and f_1 . The second term, $\frac{z_i - \min(z)}{\max(z) - \min(z)}$, normalizes the elevation of each grid cell to the elevation range of the entire glacier, and yields an elevation penalty contribution that is zero at the glacier terminus, where $z = \min(z)$, and one at the highest glacier point, where $z = \max(z)$. f_2 scales these values to a range between zero and f_2 . While the first term tends to force the least-cost route to the glacier center, the second term tends to force it downslope. The exponents a and b control the weight each of these terms have. The normalization in Eq. (2) is implemented to make the same a and b exponents better transferable to glaciers of different size and geometry.

f_1 and f_2 scale the normalized values back to actual glacier dimensions. Both coefficients are derived from medium-sized glaciers located in the Alaska Range that were

used to calibrate the initial a and b exponents. As we applied an unnormalized version of Eq. (2) to calibrate a and b , that is, $p_i = (\max(d) - d_i)^a + (z_i - \min(z))^b$, f_1 (1000, ~ equivalent to $\max(d)_{\text{calib.}}$) and f_2 (3000, ~ equivalent to $\max(z)_{\text{calib.}} - \min(z)_{\text{calib.}}$) are necessary to use these initially calibrated a and b values in the normalized Eq. (2).

To obtain plausible centerlines, a strong increase in the penalty values is required close to the glacier boundary and at higher glacier elevations. High Euclidean distance-induced penalty values at the glacier boundary are crucial to prevent centerlines from reaching too close to the glacier edge, which would not match their expected course. A strong elevation-induced penalty gradient at higher glacier altitudes is important to ensure that centerlines choose the correct branch from the start. By using a and b as exponents and not as coefficients, we do obtain the highest penalty gradients close to the glacier edges and at high glacier elevations.

Figure 4a shows the initial cost grid obtained for Gilkey Glacier. The first part of Eq. (2) is dominating (penalties decrease strongly towards the branch centers), while the second part of the equation has a lower effect.

Using the cost grid obtained from Eq. (2), we calculate the least-cost route from each head to the glacier terminus, which corresponds to the path with the minimum route cost. The route cost c is defined by the sum of the penalty values p_i between the glacier head and the terminus,

$$c = \sum_{i=1}^I p_i, \quad (3)$$

where I is the total number of cells crossed from the glacier head to the glacier terminus. The path with the lowest c is not necessarily the shortest path (minimal I), because the penalty values p_i are not constant. For example, given a meandering glacier, the shortest route is expensive, because it crosses cells with very high penalty values at the edge of the glacier. Instead, it is cheapest to stay near the center of the glacier. Although this route crosses more cells (higher I), the resulting sum of penalties is smaller as the penalty values (p_i) are considerably smaller near the glacier center.

In a next step, we convert the above least-cost route, which is obtained as a raster data set, to a vector format. We then smooth the corresponding curve using a standard polynomial approximation with exponential kernel (PAEK) algorithm (Bodansky et al., 2002). This algorithm calculates a smoothed centerline by applying a weighted average on the vertices of a moving line subsegment. A longer subsegment leads to more smoothing. We define the length of the subsegment, l (m), for every glacier individually, by

$$l = \begin{cases} u_1 \cdot S + u_2 & : l \leq l_{\max} \\ l_{\max} & : l > l_{\max} \end{cases}, \quad (4)$$

where S is the glacier area in m^2 . The constants u_1 , u_2 and l_{\max} are given in Table 1. l is increased as a function of the

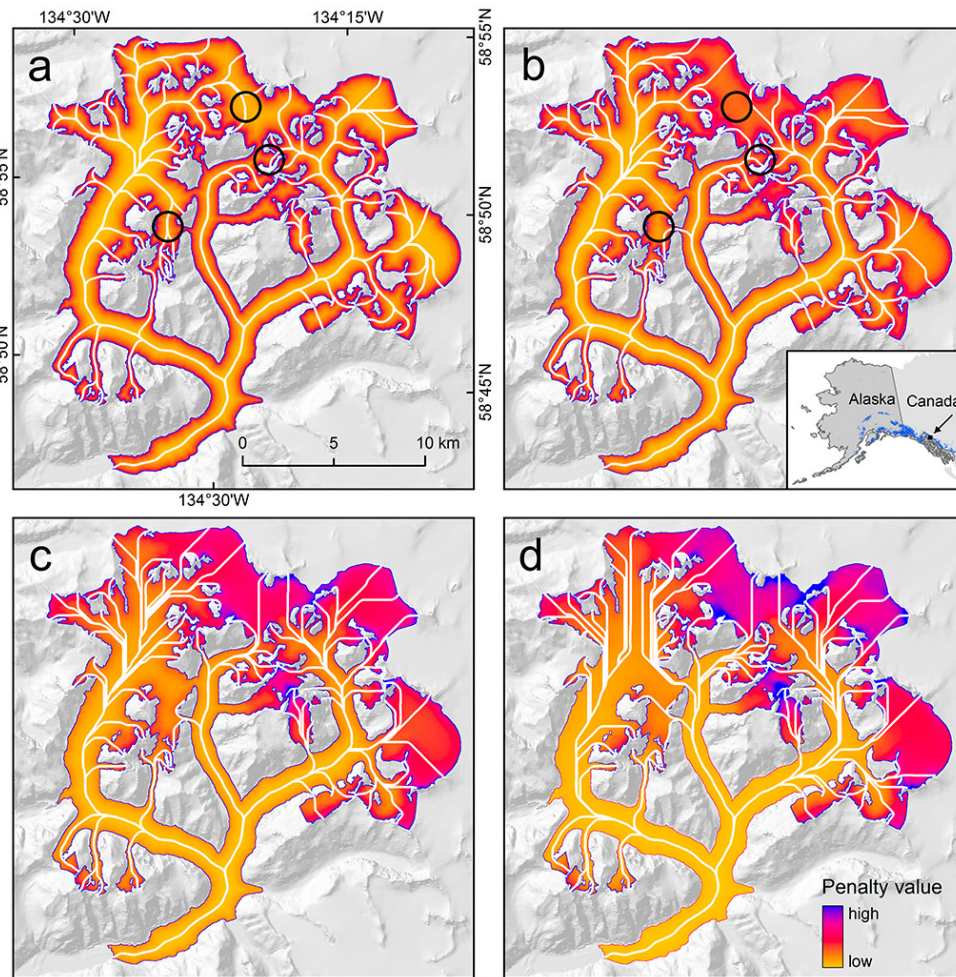


Fig. 4. Selected penalty grids of Gilkey Glacier and corresponding centerlines, using a constant a value of 4.25, but varying b . **(a)** Initial cost grid with $b = 3.5$. The penalty values strongly decrease towards the center of the branches. There is a subsidiary decline from higher to lower elevations. The black circles indicate centerlines that take implausible routes with significant upslope flow. **(b)** The penalty grid with $b = 3.7$ (i.e., $\Delta b = 0.2$). The initially wrong centerlines now take the correct routes; however, other centerlines have partially deviated from their expected courses. **(c)** $b = 3.9$ ($\Delta b = 0.4$). **(d)** $b = 4.1$ ($\Delta b = 0.6$). Many centerlines are cutting corners, especially in higher glacier reaches. Δb is limited to a maximum of 0.5; thus b values of 4.1 are not allowed during the optimization.

glacier area to account for the wider branches and the smooth course of the centerline typical for larger glaciers.

For simple glacier geometries, the first term alone in Eq. (2) already creates plausible centerlines; however, for more complex geometries, the resulting centerlines can “flow” unreasonably upslope and choose a wrong route. Because elevation (or slope) is neglected in the first term, the centerlines stick to the glacier center regardless of the topography. To remedy this problem, the elevation-dependent term is essential in Eq. (2). The elevation-dependent term also forms the basis for the optimization step introduced next.

4.2.2 Optimization

During the optimization we aim to find a combination of a and b values (Eq. 2) that provides the most plausible solution

for each centerline. Our approach is based on the following considerations: if a narrow and a wide basin are connected in their upper reaches, centerlines will typically flow through the wide basin, because the penalty values are smaller in the center of the wide basin compared to the center of the narrow basin. This is illustrated in Fig. 4a, where wide basins have lower penalty values in the center than narrow basins. In some cases, the centerlines may flow a significant distance upslope and make a major detour to reach a wider basin, and still have minimum route cost. In these cases, the low penalty values in the center of the wide basin overcompensate for the additional penalties due to the detour. This implies that the weight of the second term of Eq. (2) is too low compared to the weight of the first term. Incrementally increasing the second term of Eq. (2) will eventually force the centerline to

take a route that is shorter and typically characterized by less upslope flow. Ideally, this is the correct centerline.

To obtain the initial centerlines, we apply a and b values of 4.25 and 3.5, respectively (Table 1), as derived from tests in the Alaska Range. During the optimization, we keep a constant and raise b in discrete steps, thereby increasing the weight of the elevation-dependent term in Eq. (2).

Not all centerlines require an optimization, and in case centerlines require optimization, b cannot be increased infinitely, as this leads to a loss of the expected course of the centerline. Figure 4a illustrates this on Gilkey Glacier, where most initial centerlines have plausible routes. Only the centerlines marked by circles have implausible routes with major upslope flow and thus need optimization. Figure 4b–d show that an increase of b improves the lines in need of optimization (i.e., the implausible centerlines circled in Fig. 4a take the correct route in Fig. 4b), while it may diminish the quality of the remaining centerlines due to the higher weight of the elevation-dependent term, which forces centerlines to “cut corners” instead of sticking to the glacier center. Accordingly, a criterion is needed to determine whether optimization is necessary and, in case it is, to decide when to terminate the optimization. For this purpose, we sample the DEM along each centerline and determine the total elevation increase in meters (Δz_{up} , Fig. 5) and the maximum number of samples with continuously increasing elevation (n_{up} , Fig. 5) by

$$\Delta z_{up} = \sum_{i=1}^I \Delta z_{up,i} \quad (5)$$

$$n_{up} = \max(n_{up,i}), \quad (6)$$

where I is the total number of individual centerline sections with upslope flow. Δz_{up} and n_{up} are then used to compute the threshold m :

$$m = j_1 \cdot n_{up} + j_2 \cdot \Delta z_{up}^{j_3}, \quad (7)$$

where j_1 , j_2 and j_3 are parameters given in Table 1. Function type and parameters in Eq. (7) are empirically chosen so that m yields the upper limit of Δb , the latter defined as the increase of b relative to its initial value of 3.5. Beyond m , solutions are no longer expected to improve and thus are not computed during the optimization. Δb , and hence b , are incrementally raised until Δb equals or exceeds m , at which point the optimization terminates. Figure 6 illustrates this workflow.

Equation (7) mimics our concept of having a high m (and thus a high b) if a large n_{up} coincides with a high Δz_{up} (“worst case”, e.g., C-1 in Fig. 7a), as this is a very strong indicator for a wrong course of the preliminary centerline. If both n_{up} and Δz_{up} are zero, m is also zero as we assume that the centerline follows the correct branch. If a high Δz_{up} occurs in conjunction with a lower n_{up} or vice versa (e.g., C-2 and C-3 in Fig. 7a), m is reduced compared to the worst-case

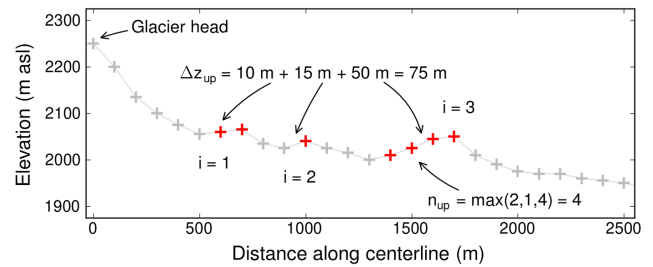


Fig. 5. Elevation profile along a glacier centerline illustrating the definition of Δz_{up} (total elevation increase, Eq. 5) and n_{up} (maximum number of samples with continuous elevation increase, Eq. 6) for $I = 3$ sections.

scenario. In such cases, we are less confident that the lines are actually wrong as such patterns occur occasionally even if the line follows the correct branch. For example, a C-2-like pattern can be caused by blunders in the DEM, while a C-3-like pattern can occur along the centerlines of larger glaciers. By using Eq. (7) we allocate a lower m to these possibly correct cases than to cases where we are more certain that they are actually wrong (C-1).

Tests indicate that many wrong centerlines shift to the correct branch within a Δb of 0.5, that is, with a b value between 3.5 and 4.0. Thus, we limit the maximum Δb (Δb_{max} in Fig. 6) to 0.5, no matter the threshold m derived from Eq. (7) (i.e., no matter the magnitude of n_{up} and Δz_{up}). Not obtaining a better solution within a Δb of 0.5 may indicate wrong divides (i.e., glaciers are not split correctly, and the centerline has to flow over a divide) or problems with the DEM (i.e., blunders within a large area that cannot be bypassed). A narrow branch located next to a very wide branch may also prevent a correct solution. Even if the algorithm found the correct branch with a Δb value larger than 0.5, the resulting line likely had an implausible shape due to the high b value (Fig. 4d).

For our large-scale application, we raise b by increments of 0.1 per iteration (Fig. 6), thus reaching Δb_{max} with five iterations. Smaller increments do not considerably improve the results, but increase the computing time. Fig. 7b shows a typical example of the progression of Δz_{up} during the optimization, in the case of there being a better branch option: Δz_{up} remains high during the first three iterations because the centerline keeps following the wrong branch. In iteration four (i.e., with $\Delta b = 0.4$), the centerline shifts to the correct branch and Δz_{up} decreases greatly. This drop in Δz_{up} is typically associated with a drop in n_{up} (not shown in Fig. 7b), which results in a lower m according to Eq. (7). In this example, m falls below Δb , and thus the optimization terminates automatically as soon as the centerline shifts to the correct branch. Iteration five is omitted, which reduces the computing time.

To obtain Δz_{up} and n_{up} , we sample the DEM only along the uppermost 25 % of the centerlines’ length. Centerlines

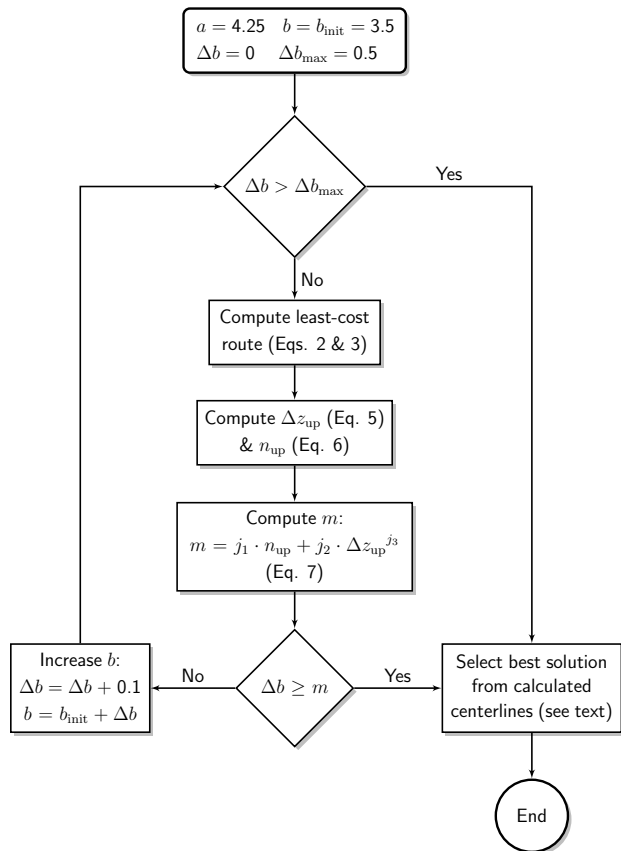


Fig. 6. Flow chart illustrating the optimization steps, including important equations and parameter values. Rectangles show the main operations, while diamonds represent decisions.

most often take the wrong route (i.e., flow upslope) in this uppermost glacier section, where the different glacier branches tend to be interconnected. The same is very unlikely in the lowermost glacier part, as there is typically only one branch left. Even if more than one branch is left, these branches are generally separated by ice-free terrain that cannot be crossed by centerlines. Moreover, upslope flow in the uppermost glacier part is clearer evidence of a wrong route than upslope flow in lower parts, where the surface may be more irregular, for example, due to varying debris coverage. Including lower glacier parts could lead to large Δz_{up} and n_{up} values although the centerline takes the correct route.

The optimization results in a data set that contains one centerline per glacier head and iteration. As the last solution from the optimization is not necessarily the best one, a strategy is needed to select the best solution from the set of centerlines. We choose the solution that has the smallest n_{up} . If two or more solutions have the same n_{up} , we order them according to Δz_{up} and choose the one with the smallest Δz_{up} . If this does not lead to a unique solution either, we take the one solution with the lowest Δb .

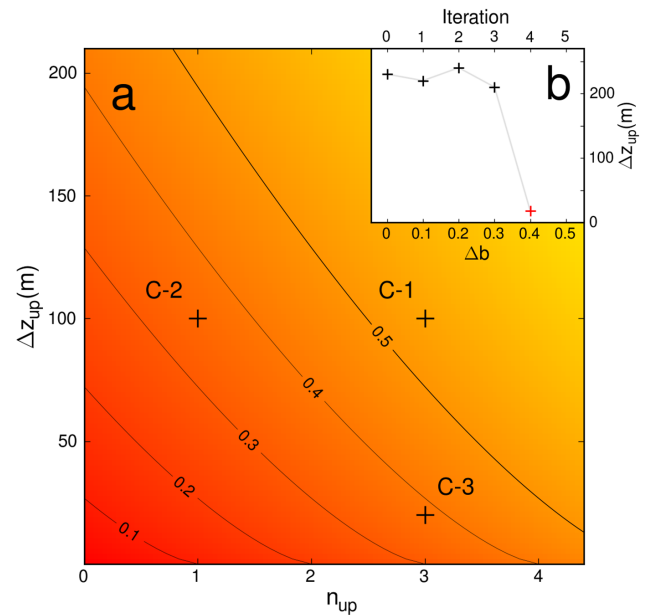


Fig. 7. (a) The threshold m (colored surface and contours) as a function of the number of samples with continuous elevation increase, n_{up} , and the total elevation increase, Δz_{up} . C-1 exemplifies a case with both high n_{up} and Δz_{up} (“worst case”), while C-2 and C-3 mark cases with a high value for one of the parameters and a low value for the other. (b) Typical progression of Δz_{up} during the optimization procedure. Δb is raised by increments of 0.1 per iteration.

Figure 8a (corresponding to Fig. 4a) shows the preliminary set of centerlines as derived for Gilkey Glacier, while Fig. 8b shows the optimized set of glacier centerlines. The three red circles indicate implausible centerlines with significant upslope flow that are successfully adapted during the optimization step. The orange circle shows an implausible centerline that is not improved because there is no upslope flow in this case (the optimization does not respond because Δz_{up} and n_{up} are zero). In Fig. 8b, the blue bold numbers indicate which iterative solution is chosen as the best solution. In 55 cases, this is the initial solution “0” (i.e., using $b = 3.5$) or solution “1” ($b = 3.6$); in 22 cases, higher-order solutions ($3.7 \leq b \leq 4.0$) are picked.

The derived centerlines reach all the way from the glacier heads to the glacier termini. If two or more branches converge, lines start to overlap. As the optimization step can result in different b values for each individual centerline, the derived centerlines do not necessarily overlap perfectly. The green circle in Fig. 8b shows an example of imperfect overlap in the case of Gilkey Glacier.

4.3 Step 3 – Derivation of branches and branch order

The ultimate goal of this step is to remove the overlapping sections of the centerlines and to arrive at individual branches that are classified according to a geometric order. We keep

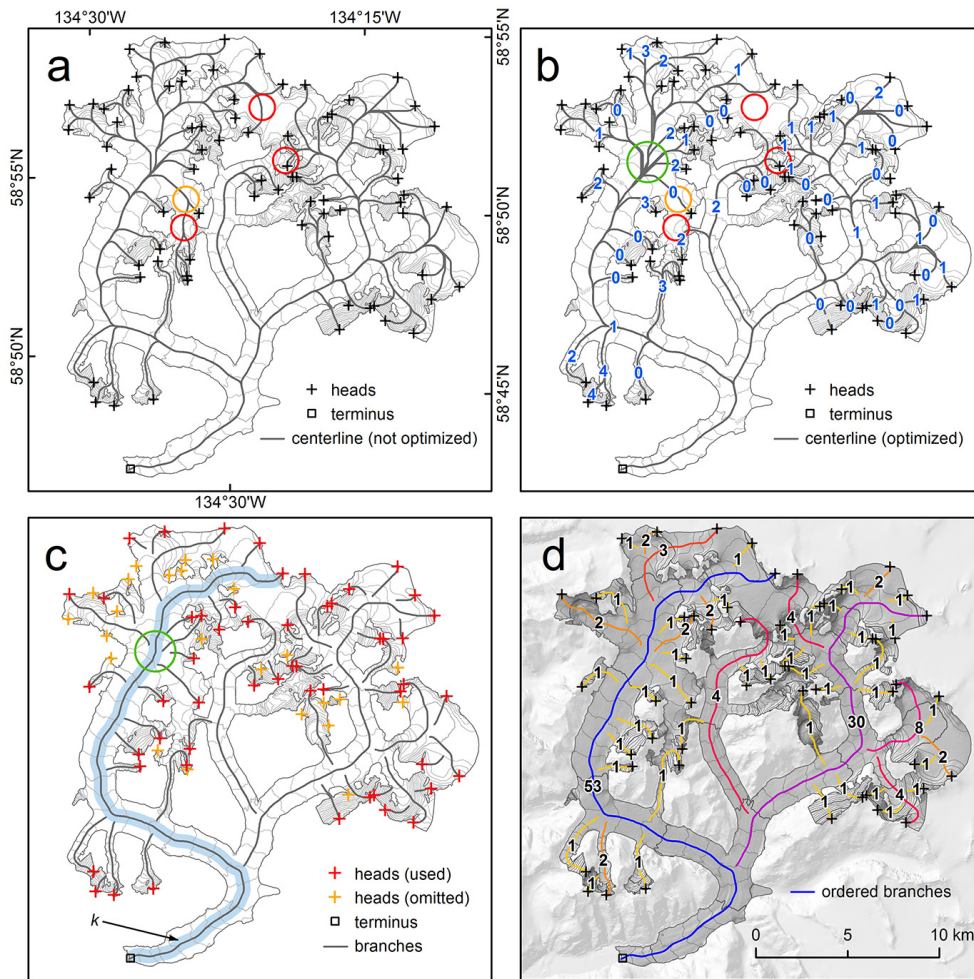


Fig. 8. The main processing steps using the example of Gilkey Glacier, Juneau Icefield. **(a)** The grey lines show the centerlines derived without the optimization, using $a = 4.25$ and $b = 3.5$ (Eq. 2). Red and orange circles mark implausible centerlines. **(b)** Centerlines after the optimization. The lines indicated by red circles are improved, while the line indicated by the orange circle remains unchanged. Blue bold numbers show the final iterative solutions for selected centerlines (e.g., “3” means solution of iteration three) given a Δb increase of 0.1 per iteration. The green circle marks an area of imperfect overlap after optimization of the centerlines, which is due to the different applied b values. **(c)** Branches after splitting the centerlines. The width of the blue area illustrates k (Eq. 8). Overlapping parts (green circle) are eliminated and short segments (belonging to the orange heads) are omitted. **(d)** Branches after allocation of geometric order. The order number indicates the number of first-order branches flowing into the corresponding branch.

the longest centerline that reaches from the head to the terminus. The remaining centerlines are trimmed so that they reach from their head to the next-larger branch. If a trimmed centerline falls below a length threshold, it is deleted.

4.3.1 Branches

We consider the longest centerline as the main branch, following previous studies (Bahr and Peckham, 1996; Paul et al., 2009). This main branch is exported into a separate file. Then, from the initial file, we remove the main branch, including line segments within a distance k (m) of the main branch (Fig. 8c). We apply this minimum distance k because different b values, employed during the optimization step,

can yield centerlines that do not overlap perfectly. Centerlines may run in parallel in the same branch before they converge, or they may diverge again, after having converged higher on the glacier. k , defined according to Eq. (8), allows for eliminating such cases of imperfect overlap:

$$k = \begin{cases} w_1 \cdot S + w_2 & : k \leq k_{\max} \\ k_{\max} & : k > k_{\max} \end{cases} \quad (8)$$

The constants w_1 , w_2 and k_{\max} given in Table 1 constrain k to values between 150 and 650 m. Larger glaciers tend to have wider branches, and parallel running centerlines in the same branch may be farther apart. To account for this, we increase k as a function of S . k is not the actual branch width,

but rather a minimum distance that is required to eliminate cases of imperfect overlap.

The application of k may yield lines that are split into multiple parts (one segment from the head to the first convergence point, the second segment from the first to the second convergence point, etc.). As the segment from the first convergence point to the terminus is already covered by the main branch, we keep only the one segment in contact with the glacier head and remove all the other segments.

The step above is applied iteratively to the longest centerline of the updated initial file until either the number of remaining centerlines reaches zero or their length falls below a certain length. As a length threshold, we reapply r defined in Eq. (1), designating lines shorter than the threshold as noise rather than branches. In cases where r would remove all centerlines (which may occur for small glaciers), Eq. (1) is not applied. Instead, r adopts the length of the longest centerline.

Merging of the identified individual branches yields the final set of branches. In the case of Gilkey Glacier, 53 branches are obtained; 24 out of 77 lines are omitted because their length is below r . In Fig. 8c, the heads of omitted branches are marked in orange. The green circle in Fig. 8c shows an area of initially imperfect overlap after employment of step 3.

4.3.2 Branch order

The above step yields a set of branches for every glacier and also establishes a branch order using the branch length as a criterion: while the longest branch is the main branch (highest order), the shortest branch is the lowest-order branch. Next, we evaluate the number of side branches that contribute to each individual branch. In this case, the branch order increases with the number of contributing branches. This results in main branches that have the highest numbers allocated. The numbers decrease as the branches fork into smaller branches. “1” stands for the lowest-order branch, meaning that there are no other branches flowing into this corresponding branch. The applied branch order is derived from the stream order proposed in Shreve (1966) and gives some first-order information about the branch topology of the glacier.

The implementation consists of a proximity analysis that is iteratively applied to each individual branch. We start by allocating an order of one (lowest order) to every branch. Next, we iterate through the branches in the order of increasing length and flag the branches that are within a distance of k (Eq. 8) from the one branch selected at that iteration step (the reference branch). The proximity analysis is applied only within the glacierized terrain. That is, a branch separated by ice-free terrain is not flagged unless the distance is less than k at a point without ice-free terrain between the branch and the reference branch. By summing the individual orders of the flagged branches, we arrive at the true order of the selected reference branch. This true order is updated

instantaneously because its updated value is required for the next iterations.

Figure 8d shows the result for Gilkey Glacier. The “53” of the main branch indicates that 53 first-order branches converge to make up the main branch.

4.4 Quality analysis and manual adjustments

The quality analysis consists of a visual check, conducted throughout the domain by evaluating the derived centerlines in conjunction with contours (50 m contour spacing), shaded-relief DEMs, and satellite imagery (mostly Landsat). A 25 km × 25 km grid, covering the entire study area, is used for guidance and keeping track of checked regions. We assess visually whether heads and termini are located correctly. For the termini, this means approximately at the center of the tongue; for the heads, at the beginning of a branch. The actual centerlines should flow roughly orthogonal to contour lines and parallel to medial moraines.

Glacier termini are manually moved to the center of the tongues, and glacier heads are added, deleted or moved as needed. To determine the number of moved termini, we compare the coordinates of the initial, automatically derived termini to the coordinates of the checked termini and sum the number of cases with changed coordinates. A similar analysis allows us to distinguish and quantify the two categories “added” and “deleted” heads.

Instead of editing the actual centerlines manually, we establish a new set of lines that we call “breaklines”. The idea is to treat these breaklines like nunataks upon rerunning of step 2 of our workflow. This implies that centerlines may not cross breaklines; moreover, breaklines change the derived penalty raster. Such an approach allows for efficient correction of wrong centerlines. For example, it is useful to adapt a centerline that did not shift into the correct branch despite the optimization procedure. A simple breakline that blocks access to the wrong branch is sufficient to reroute the wrong centerline, which is much faster than manually editing the actual centerline. In the context of the quality analysis, counting the number of set breaklines allows for quantifying wrong centerlines.

To obtain the final set of centerlines, steps 2 and 3 of the workflow are repeated using the adapted heads, termini and breaklines as input, without changing any of the remaining input data or parameters. To allow breaklines as an additional input, an adapted version of the code of step 2 is run.

4.5 Comparison to alternative methods

To identify differences between alternative methods, it would be ideal to compare the actual centerline shapes. However, quantitatively assessing shape agreement is challenging, especially for a large number of glaciers. Here, we use the length as a proxy for agreement and carry out two comparisons. First, we compare the glacier lengths derived from

a hydrological approach to the lengths derived from our longest centerlines. To obtain the hydrological lengths, we run the tool “Flow Length” from the ESRI® ArcGIS software package, which is similar to the tool applied by Schiefer et al. (2008). While this tool yields a length parameter for every glacier, it does not output an actual line that can be checked visually.

In the second comparison, we evaluate agreement between our longest centerlines and our centerlines between highest and lowest glacier elevations. This comparison yields apparent length differences that would occur by applying an approach that considers only the highest glacier elevations as heads (e.g., following Le Bris and Paul, 2013), rather than an algorithm that considers multiple heads. Unlike in the first experiment, we do not compare two approaches that are fundamentally different, but rather the same approach with one different assumption regarding glacier heads. Thus, we expect better agreement in the second experiment.

5 Results

For the 21 720 glaciers with an area $> 0.1 \text{ km}^2$, we obtain 41 860 centerlines, of which 8480 (20.3 %) have a nonzero optimized iterative solution. The centerlines range in length between 0.1 and 195.7 km; the summed length is 87 460 km. Mean glacier length of our sample, as derived from the respective longest centerlines, is 2.0 km. A total of 4350 glaciers (20.0 %) have more than one branch, and the corresponding branch orders reach up to 340. Figure 9 shows the final centerlines for the Delta Range area of the Eastern Alaska Range and Fig. 10 for the Stikine Icefield area, located in the Coast Mountains of southeast Alaska/northwest Canada. The glacier geometries range from large outlet glaciers to medium-sized valley and small cirque glaciers.

Table 2 gives an overview of the manual changes conducted during the quality analysis. Of the 21 720 glaciers, 19 060 (87.8 %) require no manual intervention at all; 2660 glaciers (12.2 %) need at least one manual intervention within the three-step procedure. Most cases of manual intervention are required to adapt the automatically derived glacier heads (1850 deleted, 1070 added) and termini (770 moved), indicating that step 1 or 3 does not yield the expected outcome in these instances. A total of 580 breaklines are used to adjust the course of the actual centerlines, indicating that in these cases, step 2 does not yield the intended result.

6 Discussion

6.1 Algorithm

Despite its empirical nature, our approach yields plausible results for a wide range of glacier sizes and shapes, provided both good-quality DEMs and glacier outlines are available.

Table 2. Manual changes attributed to individual error categories. Percentages are relative to the total of each category (e.g., deleted heads vs. total heads).

Category	Number	%
Termini	21 720	
Moved termini	770	3.5
Heads	41 860	
Deleted heads	1850	4.4
Added heads	1070	2.6
Centerlines	41 860	
Breaklines	580	1.4
Glaciers	21 720	
Glaciers with changes	2660	12.2

However, the algorithm has limitations, as shown by the quality analysis. The main challenge is to automatically derive glacier heads and termini, due to the large natural variability inherent in the glacier sample with respect to size, shape and hypsometry. The actual derivation of the centerlines is less error-prone.

6.1.1 Termini

We use the lowest glacier points to automatically identify glacier termini, which can lead to problems described in Le Bris and Paul (2013). Especially if glacier tongues reach low-slope terrain, which is typical for expanded-foot and piedmont glaciers, the lowest glacier cell may not be located in the center of the glacier tongue but rather along the side of the glacier. This “pulls” the line away from the glacier center and leads to centerlines that are not realistic. Although this inconsistency only affects the immediate tongue area, it may interfere with certain applications and thus requires a manual shift of the terminus. In our study area, shifting of misplaced termini accounts for a considerable number of cases requiring manual input (Table 2). Currently, we do not have a reliable automatic approach to detect and adapt such misclassified termini.

In many glacierized areas, there are glaciers that drain into multiple tongues. Our algorithm identifies the lowest point as the only terminus and therefore does not account for multiple termini. To address this problem, we have to manually split these glaciers into separate catchments (i.e., one catchment per tongue), followed by treating the catchments like separate glaciers. In Alaska and northwest Canada, only a handful of glaciers drain into multiple tongues; hence the amount of manual intervention required is relatively small.

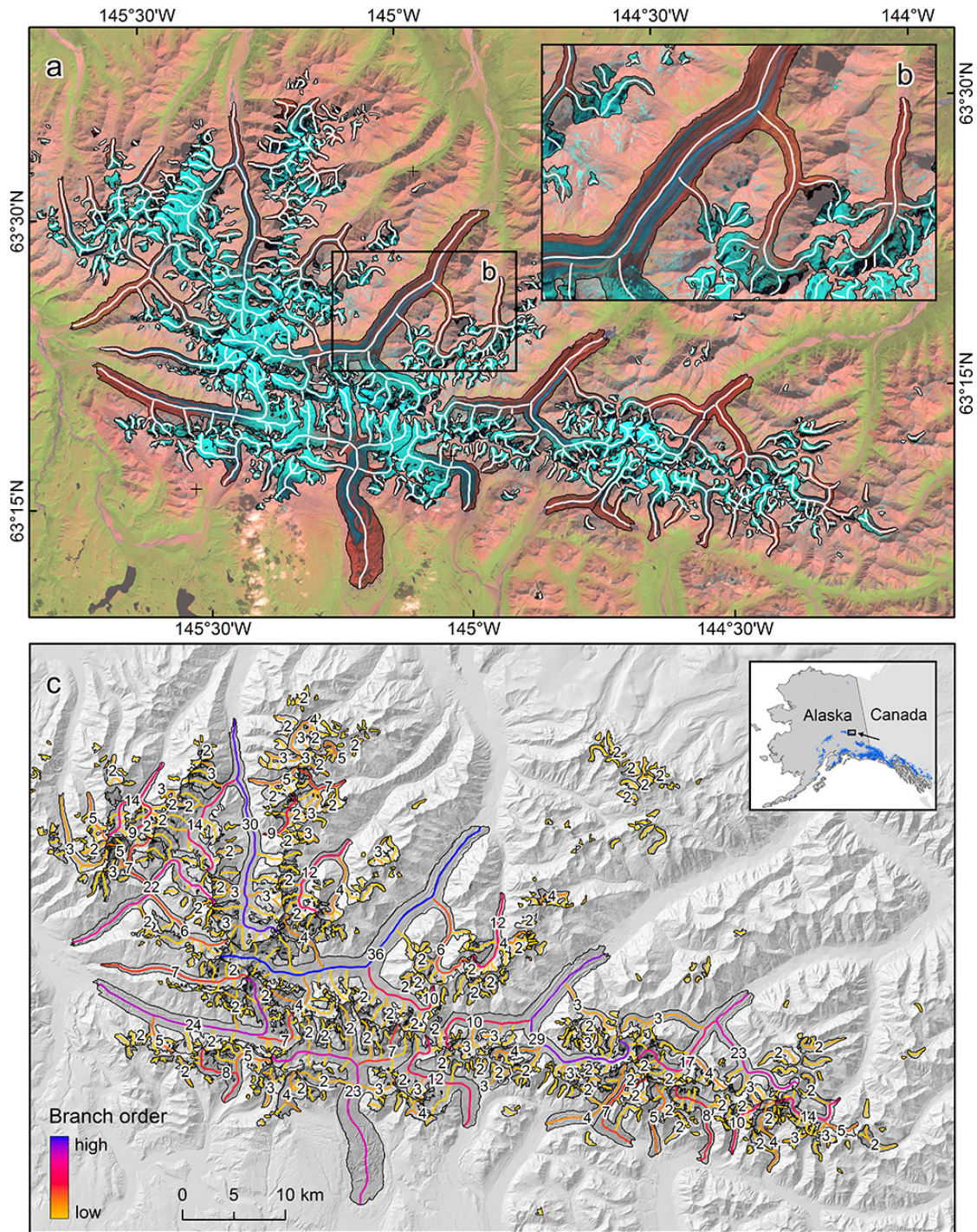


Fig. 9. Derived centerlines for the Delta Range area. The white lines in panel (a) and (b) show the centerlines overlaid on a Landsat 5 Thematic Mapper false-color composite (bands 7-4-2) from 10 August 2005 (scene LT50660162005222PAC00). Moraines are shown in brown, snow and ice in light to dark blue. Panel (c) shows the same centerlines, but with colors indicating the branch order. Branch orders greater than one are labeled. The shaded-relief background is derived from the IFSAR DEM.

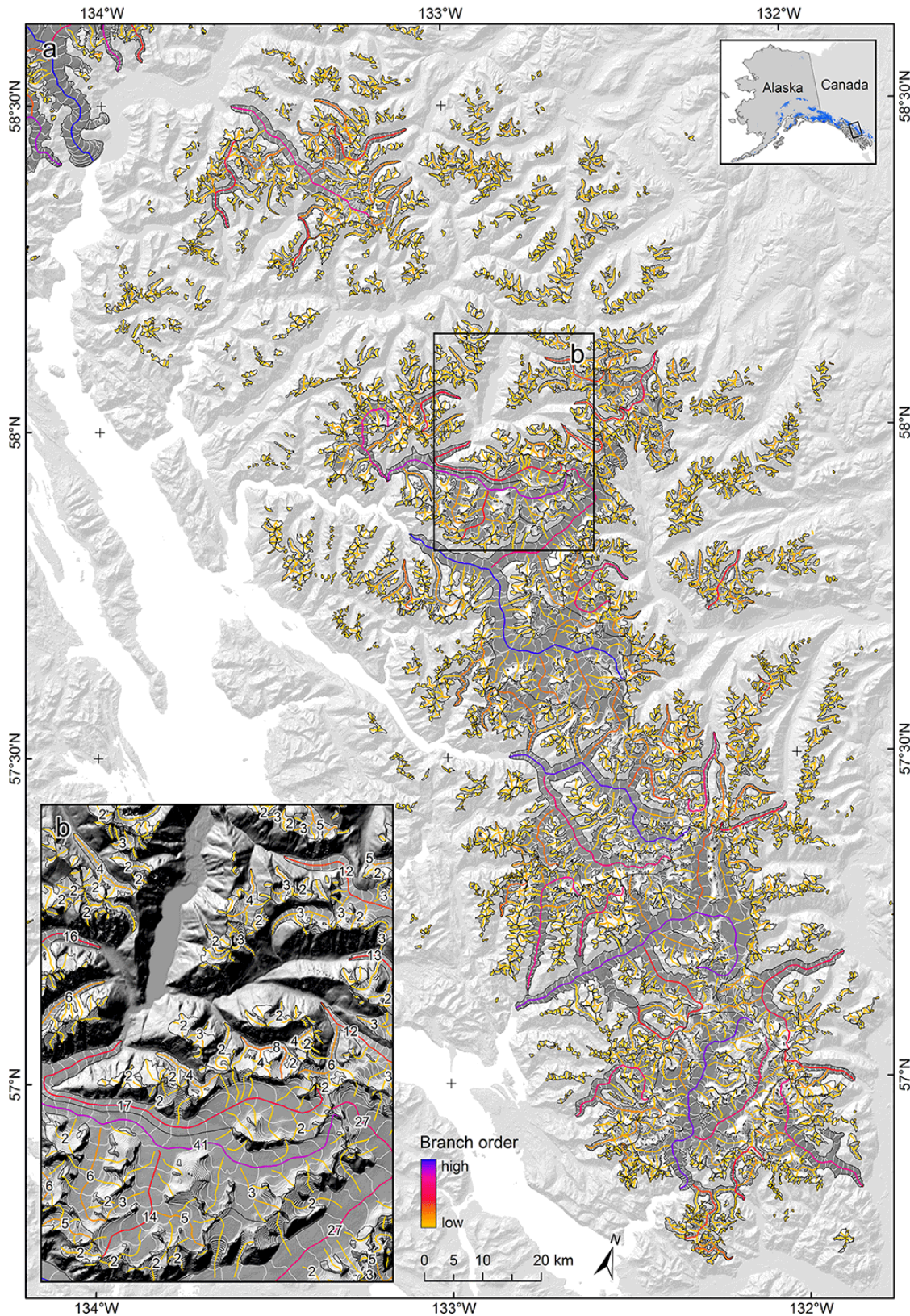


Fig. 10. (a) Derived centerlines for glaciers in the Stikine Icefield area. The line colors indicate branch order. The shaded-relief background is derived from the SRTM DEM. 50 m contours are shown as white lines. Inset (b) shows a subarea with labels quantifying the branch orders greater than one.

6.1.2 Heads

By definition, our algorithm obtains exactly one point per local elevation maximum. The corresponding centerline covers one branch; other branches that may originate from the same area remain without centerlines. This is illustrated in Fig. 8b, where not all branches have a centerline allocated. Because our algorithm does not necessarily yield centerlines for each individual glacier branch, additional heads may have to be set manually, often combined with breaklines (required to prevent centerlines from clustered heads taking the same branch). In our test area, this constraint is responsible for most cases in the category “added heads” (Table 2). Glaciers that fork from one into multiple branches, such as glaciers on volcanoes, are most susceptible to the problem.

We further prescribe that centerlines must run from their head to the terminus. While this is appropriate for many glacier geometries, it may not be the case for hanging or apron glaciers. For example, the minimum point of a small, wide apron glacier may be on one side, while a local maximum may be on the other side of the glacier. This results in a centerline that runs almost parallel to the contours, yielding a maximum length that is too long. In such cases, manual intervention is required to remove implausible head–terminus combinations. In our test area, this problem is responsible for the bulk of corrections with regard to deleted heads (Table 2). By applying a higher minimum glacier area threshold (e.g., 1 km² instead of 0.1 km²), the amount of manual corrections could be reduced, as the challenging glacier geometries tend to have small areas.

6.1.3 Cost grid–least-cost route approach

Our algorithm generally yields plausible centerlines if we derive the routes from cost grids established with a and b values of 4.25 and 3.5, respectively (Eq. 2). The Euclidean distance term controls this initial cost grid, reflecting our assumption that the main flow occurs in the glacier center. If this assumption does not hold, the quality of the resulting centerlines may decline, although we consider elevation in Eq. (2) and also conduct an optimization step. Lower-quality centerlines are found, for example, on glaciers that drain very wide, asymmetric basins. In contrast, our quality analysis indicates that the approach works particularly well for valley glaciers. Alaska and northwest Canada comprise many outlet and valley glaciers, which is an important reason for the relatively high success rate in this area (Table 2).

The existence of nunataks tends to improve the derived centerlines. Due to the high penalty values close to the nunataks, centerlines are forced to flow around nunataks, which is typically consistent with their expected course. In our test area, nunataks are abundant, which simplifies the application of the cost grid–least-cost route approach.

Equation (2) is only one way to obtain a functioning cost grid. Other, possibly shorter equations may yield similar

results. For example, it would be possible to use the same values for f_1 and f_2 upon recalibrating a and b , thus reducing the number of variables by one. Instead of exponentials, one could also attempt to obtain a cost grid using logarithms.

6.1.4 Optimization

The optimization is a crucial element of the cost grid–least-cost route approach and works robustly in general. We identify three cases where the optimization either fails or does not respond at all. First, no optimization occurs if the line continuously flows downslope, despite taking the wrong route (m in Eq. 7 is zero in these cases). Second, the algorithm does not optimize if the upslope flow occurs below the upper 25 % of the centerline’s length (Δz_{up} and n_{up} are only determined within the first 25 %). An m that is not high enough (despite detected upslope flow) is a potential third cause for failure of the optimization. It is difficult to attribute wrong branches to individual cases; however, we hypothesize that the first case causes the largest number of errors.

The presented optimization is the result of experimenting with different optimization approaches and break criteria. Intuitive break criteria such as “optimize until Δz_{up} and n_{up} equal zero” or “optimize until improvement of Δz_{up} and n_{up} equals zero” do not work reliably due to the large variability inherent in the glacier sample. For example, a solution may decline temporarily (resulting in higher Δz_{up} or n_{up}) before it improves again to finally yield the best solution. Iteratively calculating all the solutions according to Eq. (7) and then choosing the best solution out of this set of centerlines generally is most successful.

6.1.5 Branches and geometry order

The proposed approach to derive branches from centerlines works satisfactorily in most cases; however, we rely on various simplifications. For example, Eq. (8) defines a constant k for each glacier, which is then used to split the centerlines into branches. Ideally, k would correspond to the actual local branch width and thus evolve along each individual branch. While such a procedure is easily implementable for simple glacier geometries, it is difficult in areas with many interconnected branches, and thus not implemented here. Assuming a constant k is most problematic for very large glaciers, where the branch widths can vary from less than one to several tens of kilometers. In conjunction with the constant minimum length threshold r (Eq. 1), the constant k may lead to branches that are omitted although they should not be, and vice versa. For large glaciers, such errors can be an important contributor to the categories added and deleted heads (Table 2).

6.2 Influence of DEM and outline quality

Our results depend on the quality of DEM and glacier outlines. While systematic elevation biases (e.g., like those

found in the SRTM DEM) have little to no effect, blunders such as bumps (e.g., found in the ASTER GDEM2, due to a lack of contrast in the corresponding optical imagery) are more severe. They lead to elevation maxima that are not real, which interfere with our search for actual glacier heads. Blunders also affect the course of the centerlines. Artificial bumps lead to upslope flow, and the subsequent optimization picks centerlines that flow around these bumps. These solutions are better according to the optimization criteria Δz_{up} and n_{up} , but in reality may be worse than the initial solutions. In areas where we rely on DEMs with blunders, the DEM quality is responsible for most manual corrections in any error category.

If the glacier is part of a larger glacier complex, correct ice divides along the actual drainage divides as retrieved from the DEM are crucial. In case of erroneous divides, centerlines flow over the divide, which may prompt an optimization although the only solution is to adapt the divides. Likewise, it is important to identify correctly location and shape of nunataks. In case of omitted nunataks, the centerlines may cross the nunataks, which is not intended. Identifying nunataks where there are none (e.g., on a medial moraine) leads to centerlines flowing around these apparent nunataks, yielding an implausible curvy shape.

6.3 Quality assessment

For the quality assessment, we use shaded-relief DEMs, contour lines, and satellite imagery. As medial moraines and contour lines are only a proxy for ice flow direction, it would be ideal to consult actual velocity field data to validate the centerlines. However, at the time of the quality analysis, such flow fields were not available on a larger scale. Recently presented data (Burgess et al., 2013a) may be used for future studies.

Visual assessments involve some degree of subjectivity, which we attempt to minimize by checking the results multiple times. Comparison to centerlines exclusively derived by hand could extend the current quality assessment. However, to be meaningful, such tests should comprise multiple glaciers from each glacier type, manually processed by different technicians. This is very time-consuming and thus beyond the scope of this study.

6.4 Comparison to alternative methods

To quantify method-related differences, we compare the glacier lengths derived from alternative algorithms. A meaningful analysis is supported by the large number of length observations available. Calculating the ratios of the lengths obtained from different methods allows a comparison among glaciers. The histogram in Fig. 11a illustrates contrasts in length arising from the concurrent application of our cost grid–least-cost route and a hydrological approach. The distribution of the obtained ratios is shown in the histogram

in Fig. 11a. While the mean and median are very close to unity ($\bar{R} = 0.99$, $R_{50} = 1.02$), the distribution is left-skewed with a maximum between 1.05 and 1.15 and considerable spread. Only $\sim 50\%$ of the 21 720 glaciers have lengths that are within 10% of each other. Assuming that the centerlines from our cost grid–least-cost route approach are the “correct” reference, the distribution peak between 1.05 and 1.15 confirms the finding of Schiefer et al. (2008) that hydrological approaches tend to overestimate glacier length due to the deflection of the hydrological “flowline” to the glacier edge in convex areas. However, in almost 50% of the cases, the lengths from the hydrological approach are shorter than the lengths from the cost grid–least-cost route approach. The pattern is found throughout all size classes and can occur if the hydrological flowline not only gets deflected in areas with convex glacier geometry but actually leaves the glacier before reaching the lowest glacier elevation. In these cases, the hydrological approach may underestimate glacier length. In the cost grid–least-cost route approach, every centerline is forced to reach the lowest glacier elevation, an assumption that may not always hold, which then leads to an overestimation of the length by our approach. Combined, the two tendencies for under- and overestimation may explain the considerable fraction of ratios between 0.7 and 0.9.

The histogram in Fig. 11b shows how different the lengths would be if centerlines were computed between the lowest and highest glacier elevations only (e.g., Le Bris and Paul, 2013), instead of considering multiple branches per glacier. For this comparison, we extract only the 4350 glaciers that have two or more centerlines, as the two lengths must be identical in the case of the remaining 17 370 glaciers. Excluding the glaciers with one centerline tends to exclude the smallest glaciers; thus, nearly all glaciers $< 0.5 \text{ km}^2$ are omitted. More than 70% of the glaciers have a high ratio between 0.95 and unity, meaning that the centerline originating from the highest glacier elevation has a length that is within 5% of the longest centerline’s length. For most glaciers, the longest centerline is actually identical to the line between the lowest and the highest glacier elevation, which is explained by a strong correlation of elevation range and length. Nevertheless, considerable outliers may occur in individual cases, especially for large glaciers that have many branches.

7 Conclusions

We have developed a three-step algorithm to calculate glacier centerlines in an automated fashion, requiring glacier outlines and a DEM as input. In the first step, the algorithm identifies glacier termini and heads by searching for minima and local elevation maxima along the glacier outlines. The second step comprises a cost grid–least-cost route implementation, which forces the centerlines towards both the central portion and lowest elevations of the glacier. The second step also implements an optimization routine, which obtains the

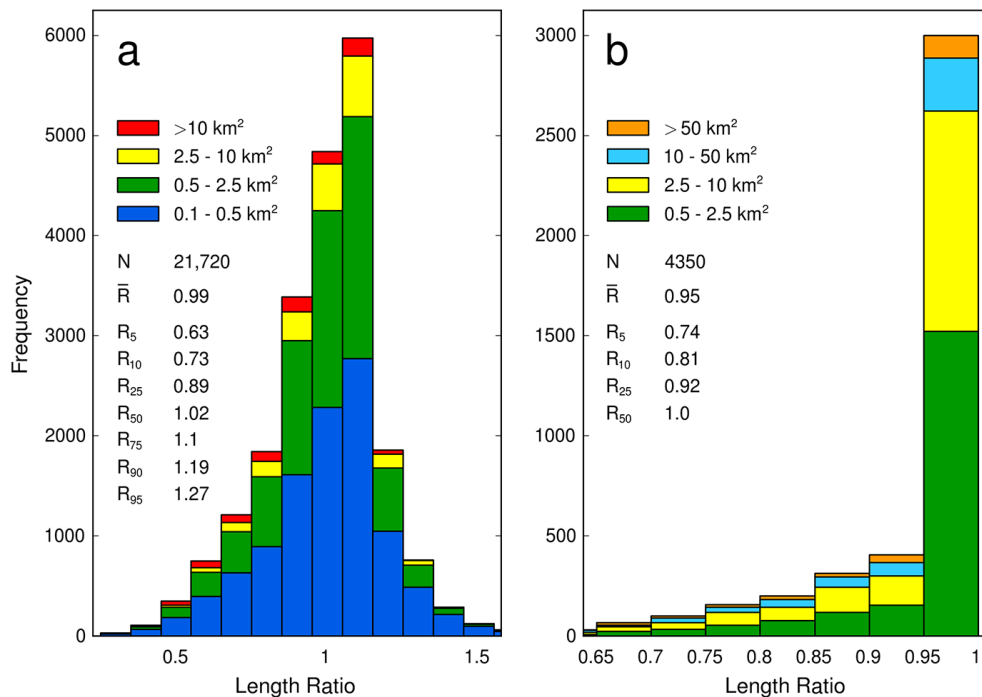


Fig. 11. Histograms illustrating the length ratio distributions of differently derived glacier lengths. Colors distinguish size categories; N indicates the total number of glaciers. Mean (\bar{R}) and selected quantiles (R) are calculated for the complete distributions. Panel (a) shows the ratio of length derived from a hydrological approach to the length derived from our longest centerline. Bar width is 0.1. Panel (b) shows the ratio of length derived from the centerline between the highest and the lowest glacier elevation to the length derived from our longest centerline. Only glaciers with more than one centerline are considered. Bar width is 0.05.

most plausible centerlines by slightly varying the cost grid. In the third step, the algorithm divides the centerlines into individual branches, which are then classified according to branch order (Shreve, 1966).

We have developed and applied our centerline algorithm on a glacier inventory for Alaska and northwest Canada (Arendt et al., 2013). The algorithm is applied to 21 720 glaciers with a minimum area of 0.1 km², yielding 41 860 individual branches ranging in length between 0.1 and 195.7 km. The mean length of the glacier sample is 2.0 km. Our quality analysis shows that the majority (87.8 %) of the glaciers required no manual corrections. The most common errors occurred due to misidentification of either glacier heads or termini (7.0 and 3.5 % of errors, respectively). Once heads and termini are correctly identified, the algorithm determines centerlines for nearly all (98.6 %) cases. The quality analysis further indicates that the algorithm works best for valley glaciers, while apron glaciers tend to be most challenging.

For our sample of 21 720 glaciers, we compare the lengths derived from a hydrological approach (e.g., Schiefer et al., 2008) to the lengths derived from our longest centerlines. We find considerable variation: although the average ratio of the two lengths is close to unity, only ~ 50 % of the glaciers have the two lengths within 10 % of each other. This suggests

that the choice of the applied method may significantly influence the derived glacier lengths. Comparing the lengths from the centerline between highest and lowest glacier elevations to the lengths from the longest centerlines shows that they agree well: > 70 % of the glaciers with two or more branches have the two lengths within 5 % of each other. Agreement is best for small glaciers with few branches. Our results suggest that the centerline between the highest and the lowest glacier point is generally valid to describe the glacier length although considerable outliers may occur in individual cases.

Our derived centerlines do not provide unique, “true” solutions. The results may vary with the parameters chosen in the applied equations; moreover, technician interpretation adds subjectivity within the quality analysis. Nevertheless, the proposed approach contributes towards a standardized derivation of centerlines. The final product may be used, for example, to calculate the glacier length using both suggested methods (Paul et al., 2009) or to conduct topological analyses (Bahr and Peckham, 1996). It may also be used for area-length scaling applications (Schiefer et al., 2008) or as input for flowline modeling (Sugiyama et al., 2007). As soon as good-quality DEMs and glacier outlines are globally available, the scope of the presented project may be expanded from a regional to a global scale.

Acknowledgements. We thank I. Evans and B. Raup for their valuable comments that improved the manuscript. The work was supported by grants from NSF (no. EAR 0943742) and NASA (nos. NNX11AF41G and NNX11AO23G).

Edited by: A. Kääh

References

- Arendt, A., Bolch, T., Cogley, J. G., Gardner, A. S., Hagen, J. O., Hock, R., Kaser, G., Pfeffer, W. T., Moholdt, G., Paul, F., Radi, V., Andreassen, L., Bajracharya, S., Beedle, M., Berthier, E., Bhambri, R., Bliss, A., Brown, I., Burgess, E., Burgess, D., Cawkwell, F., Chinn, T., Copland, L., Davies, B., De Angelis, H., Dolgova, E., Filbert, K., Forster, R. R., Fountain, A., Frey, H., Giffen, B., Glasser, N., Gurney, S., Hagg, W., Hall, D., Haritashya, U. K., Hartmann, G., Helm, C., Herreid, S., Howat, I., Kapustin, G., Khromova, T., Kienholz, C., Koenig, M., Kohler, M., Kriegel, D., Kutuzov, S., Lavrentiev, I., Le Bris, R., Lund, J., Manley, W., Mayer, C., Miles, E., Li, X., Menounos, B., Mercer, A., Moelg, N., Mool, P., Nosenko, G., Negrete, A., Nuth, C., Pettersson, R., Racoviteanu, A., Ranzi, R., Rastner, P., Rau, F., Rich, J., Rott, H., Schneider, C., Seliverstov, Y., Sharp, M., Sigurdsson, O., Stokes, C., Wheate, R., Winsvold, S., Wolken, G., Wyatt, F., and Zheltyhina, N.: Randolph Glacier Inventory [v3.2]: A Dataset of Global Glacier Outlines, Global Land Ice Measurements from Space, Boulder, Colorado, USA, Digital Media, 2013.
- Bahr, D. B. and Peckham, S. D.: Observations and analysis of self-similar branching topology in glacier networks, *J. Geophys. Res.*, 101, 25511–25521, 1996.
- Bodansky, E., Gribov, A., and Pilouk, M.: Smoothing and compression of lines obtained by raster-to-vector conversion, in: *Graphics Recognition Algorithms and Applications*, 256–265, Springer, 2002.
- Bolch, T., Menounos, B., and Wheate, R.: Landsat-based inventory of glaciers in western Canada, 1985–2005, *Remote Sens. Environ.*, 114, 127–137, 2010.
- Burgess, E. W., Forster, R. R., and Larsen, C. F.: Flow velocities of Alaskan glaciers, *Nature Communications*, 4, 2146, doi:10.1038/ncomms3146, 2013a.
- Burgess, E. W., Larsen, C. F., and Forster, R. R.: Summer melt regulates winter glacier flow speeds throughout Alaska, *Geophys. Res. Lett.*, 40, 6160–6164, 2013b.
- Farinotti, D., Huss, M., Bauder, A., Funk, M., and Truffer, M.: A method to estimate the ice volume and ice-thickness distribution of alpine glaciers, *J. Glaciol.*, 55, 422–430, 2009.
- 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, *Rev. Geophys.*, 45, 1–33, 2007.
- Heid, T. and Kääh, 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.
- Korona, J., Berthier, E., Bernard, M., Rémy, F., and Thouvenot, E.: SPIRIT. SPOT 5 stereoscopic survey of Polar Ice: Reference Images and Topographies during the fourth International Polar Year (2007–2009), *ISPRS Journal of Photogrammetry and Remote Sensing*, 64, 204–212, 2009.
- Le Bris, R. and Paul, F.: An automatic method to create flow lines for determination of glacier length: A pilot study with Alaskan glaciers, *Comput. Geosci.*, 52, 234–245, 2013.
- Le Bris, R., Paul, F., Frey, H., and Bolch, T.: A new satellite-derived glacier inventory for western Alaska, *Ann. Glaciol.*, 52, 135–143, 2011.
- 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, *Clim. Past*, 8, 1385–1402, doi:10.5194/cp-8-1385-2012, 2012.
- 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, *J. Geophys. Res.*, 117, F01020, doi:10.1029/2011JF002104, 2012.
- Linsbauer, A., Paul, F., and Haerberli, W.: Modeling glacier thickness distribution and bed topography over entire mountain ranges with GlabTop: Application of a fast and robust approach, *J. Geophys. Res.*, 117, F03007, doi:10.1029/2011JF002313, 2012.
- McNabb, R., Hock, R., O’Neel, S., Rasmussen, L., Ahn, Y., Braun, M., Conway, H., Herreid, S., Joughin, I., Pfeffer, W., Smith, B., and Truffer, M.: Using surface velocities to calculate ice thickness and bed topography: a case study at Columbia Glacier, Alaska, USA, *J. Glaciol.*, 58, 1451–1464, 2012.
- Melkonian, A. K., Willis, M. J., Pritchard, M. E., Rivera, A., Bown, F., and Bernstein, S. A.: Satellite-derived volume loss rates and glacier speeds for the Cordillera Darwin Icefield, Chile, *The Cryosphere*, 7, 823–839, doi:10.5194/tc-7-823-2013, 2013.
- Nuth, C., Kohler, J., König, M., von Deschanden, A., Hagen, J. O., Kääh, 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, *Ann. Glaciol.*, 24, 382–389, 1997a.
- Oerlemans, J.: Climate Sensitivity of Franz Josef Glacier, New Zealand, as Revealed by Numerical Modeling, *Arctic Alpine Res.*, 29, 233–239, 1997b.
- Paul, F., Barry, R. G., Cogley, J. G., Frey, H., Haerberli, 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, *Ann. Glaciol.*, 50, 119–126, 2009.
- Schiefer, E., Menounos, B., and Wheate, R.: An inventory and morphometric analysis of British Columbia glaciers, Canada, *J. Glaciol.*, 54, 551–560, 2008.
- Shreve, R. L.: Statistical Law of Stream Numbers, *J. Geol.*, 74, 17–37, 1966.
- 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, *Ann. Glaciol.*, 46, 268–274, 2007.
- Tachikawa, T., Hato, M., Kaku, M., and Iwasaki, A.: Characteristics of ASTER GDEM version 2, in: *Geoscience and Remote Sensing Symposium (IGARSS)*, 2011 IEEE International, 3657–3660, 2011.