Improving Semi-Automated Glacial Mapping with a Multi-Method Approach: Applications in Central Asia

Taylor Smith[1], Bodo Bookhagen[1], and Forest Cannon[2]
[1]Institute for Earth and Environmental Sciences, Universität Potsdam, Germany
[2]Geography Department, University of California, Santa Barbara, USA

Corresponding authors:
Taylor Smith and Bodo Bookhagen
Institute for Earth and Environmental Sciences
Universität Potsdam
Potsdam-Golm 14476, Germany
Email: tsmith@uni-potsdam.de, bodo@geo.uni-potsdam.de

Abstract.

There exist several global glacier databases, but they are often limited to a single glacier-area measurement in time. This limits longitudinal studies of glacier change, particularly on hard to map debris-covered glaciers, which require extensive manual digitization.

To minimize time spent on manual digitization, we designed and implemented a glacier mapping algorithm in the Tien Shan which delineates both clean glacier ice – methods which are well documented – and glacier debris tongues, which have distinct spectral signatures from clean glacier ice, and often require extensive manual intervention. This research improves upon methods developed to automatically delineate glacier areas using spectral, topographic, velocity, and spatial relationships. We found that the algorithm misclassifies between 2 and 10% of glacier areas, as compared to a $\sim$750 glacier control dataset.

The algorithm does not completely solve the difficulties inherent in classifying glacier areas from remotely sensed imagery, but does represent a significant improvement over purely spectral-based classification schemes, such as the relationship between Landsat bands one, three and five or the Normalized Difference Snow Index.

1 Introduction

Changes in glacier area have long been considered one of the best indications of climate change (i.e. Oerlemans, 2005), although more recent studies have also assessed glacier volume changes with the advent of new remote sensing techniques (e.g., Berthier et al., 2007; Aizen et al., 2007; Gardelle et al., 2012, 2013; Bolch et al., 2012; Kääb et al., 2012; Stocker, 2013). This study focuses on mapping glaciers over a large spatial scale using publicly available remotely sensed data. Several high-resolution glacier outline databases have been produced, most notably the Global Land Ice Measurements from Space (GLIMS) project (Armstrong et al., 2005; Raup et al., 2007, 2014), and the recently produced supplemental GLIMS dataset known as the Randolph Glacial Inventory (RGI) v4.0 (Arendt et al., 2012; Pfeffer et al., 2014). A coherent, complete, and accurate global glacier database is important for several reasons, including monitoring global glacier changes driven by climate change, natural hazard detection and assessment, and analysis of the role of glaciers in natural and built environments, including glacier contributions to regional water budgets and hydrologic cycles (Racoviteanu et al., 2009). Precision in glacier outlines is of utmost importance for monitoring changes in glaciers, which may change less than 15-30 m/yr ($\sim$1-2 pixels of Landsat Enhanced Thematic Mapper (ETM+)/yr). Thus, spatially accurate glacier outlines are imperative for precise glacier change detection (Paul et al., 2004, 2013).

Several methods have been developed to delineate clean glacier ice (i.e. Hall et al., 1987; Paul, 2002; Paul et al., 2002; Racoviteanu et al., 2008a,b; Hanshaw and Bookhagen, 2014), relying primarily on spectral data available on satellites such as Landsat and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). Although significant progress has been made towards automated glacier outline retrieval using satellite imagery, these methods struggle to accurately map debris-covered glaciers, or other glaciers with irregular spectral profiles (Paul et al., 2004; Bolch et al., 2007; Racoviteanu et al., 2008b; Scherler et al., 2011a). Much of this difficultly stems from the similarities in spectral profiles of debris located on top of a glacier tongue and the surrounding hillslopes where the debris is sourced from. The majority of studies examining debris-covered glaciers employ extensive manual digitization in a Geographic Information System (GIS), which is very time consuming, and can introduce significant user-generated errors (Paul et al., 2013; Pfeffer et al., 2014; Raup et al., 2014). Building on the multi-spectral, topographic, and spatially-weighted methods developed by Paul et al. (2004), we present a refined rules-based classification algorithm based on spectral, topographic, land-cover, velocity, and spatial relationships between glacier areas

and the surrounding environment. The algorithm has been designed to be user-friendly, globally applicable, and built upon open-source tools.

2 Study Area and Data Sources

2.1 Study Area

In this study we use a suite of 62 Landsat Thematic Mapper (TM), Enhanced Thematic Mapper+ (ETM+) and Optical Land Imager (OLI) images across a spatially and topographically diverse set of study sites comprising eight Landsat footprints (Path/Row combinations: 144/30, 145/30, 147/31, 148/31, 149/31, 151/33, 152/32, 153/33) along a $\sim$1,500km profile from the Central Pamir to the Central and Central-Eastern Tien Shan (Figure 1) to analyze the results of our classification algorithm.

The study area contains a wide range of glacier types and elevations, with both small and clean-ice dominated glaciers, as well as large, low-slope, and debris-covered glaciers. The diversity in glacier types in the region provides and ideal test area, particularly in regards to mapping glaciers with long and irregular debris tongues, such as the Inylchek and Tomur glaciers in the Central Tien Shan.

The wintertime climate of the study area is controlled by both the Winter Westerly Disturbances (WWDs) and the Siberian High, which dominate regional circulation and create strong precipitation gradients throughout the range, which extends from Uzbekistan in the west through China in the east (Figure 1) (Lioubimtseva and Henebry, 2009; Narama et al., 2010; Bolch et al., 2011; Sorg et al., 2012; Cannon et al., 2014). The western edges of the region tend to receive more winter precipitation in the form of snow, with precipitation concentrated in the spring and summer in the central and eastern reaches of the range (Narama et al., 2010).

Sorg et al. (2012) note that due to the interaction of the Siberian High – a semi-permanent thermal high-pressure system which extends from Siberia towards the Tien Shan – and more western continental atmospheric patterns such as WWDs, the study area has a distinct precipitation gradient with decreasing precipitation from the west to east. There is also a strong gradient between outer and inner reaches of the Tien Shan, as moisture is precipitated on the windward (northwest) side of topography, moving eastwards from the Pamirs (Sorg et al., 2012). These climate zones contribute to the diversity of glaciers in the study region.

2.2 Data Sources

Our glacier mapping algorithm is based on several datasets. The Landsat 5 (TM), 7 (ETM+), and 8 (OLI) platforms were chosen as the primary spectral data sources, as they provide spatially and temporally extensive coverage of the study area (Table 1). ASTER can also be used as a source of spectral information, but here we chose to focus on the larger footprint and longer timeseries available through the Landsat archive. In addition to spectral data, the 2000 Shuttle Radar Topography Mission (SRTM) Digital Elevation Model (DEM) ($\sim$90m) was leveraged to provide elevation, slope, and hillshade information (Farr and Kobrick, 2000). The SRTM data and its derivatives were downsampled to 30m to match the resolution of the Landsat images using bilinear resampling. The USGS Hydrosheds river network (15 second resolution, $\sim$500m) was also used as an input dataset (Lehner et al., 2008).

3 Methods

Our glacier classification algorithm uses several sequential thresholding steps to delineate glacier outlines. The scripts used in this study are available in the Data Repository, with updates posted to http://github.com/ttsmith89/GlacierExtraction/. It is noted if the step requires manual processing or is part of a script.

1. Data Preparation
    (a) Landsat TM/ETM+ data is aligned to a single base image (Manual)
    (b) Velocity fields are calculated with Normalized Cross Correlation (Manual)
    (c) The Hydrosheds river network is buffered to a polygon feature (Manual)
    (d) Training lakes and any Manual Debris points are created (Manual)
    (e) SRTM data is used to create Slope and Hillshade images (Python Script)
    (f) All input datasets are matched to a single extent and spatial resolution (30m) (Python Script)
2. Glacier Classification Steps
    (a) Lakes are classified using the Normalized Difference Water Index (NDWI) (Python Script)
    (b) Clean-ice glacier outlines are created using Landsat Bands 1,3, and 5 (Matlab Script)

(c) 'Potential debris areas' are generated from low-slope and low slope-variability areas (Matlab Script)

(d) Low elevation areas are removed (Matlab Script)

(e) Low- and high-velocity areas are removed (Matlab Script)

(f) Distance weighting metrics are used to remove areas distant from river networks or clean glacier ice (Matlab Script)

(g) Distance weighting metrics are used to remove areas distant from glacier lakes and manual seed points (Matlab Script)

(h) The resulting glacier outlines are cleaned with statistical filtering (Matlab Script)

3. Post-processing

(a) Glacier outlines are exported to ESRI shapefile format for use in a GIS (Python Script)

3.1 Data Preparation

For accurate glacier delineation, we primarily used Landsat images which were free of new snow, and had less than 10% cloud cover. However, we have also included scenes with limited snowcover and cloud cover in our analysis to understand the impacts of snow and clouds on our classification algorithm. We find that the presence of fresh snow in images tends to overclassify glacier areas and classify non-permanent snow as glaciers. Additionally, cloud covered glaciers cannot be correctly mapped by the algorithm (Paul et al., 2004; Hanshaw and Bookhagen, 2014).

After selecting a series of Landsat images, we co-registered each image to a 'master' image specific to each Landsat Path/Row combination, using the Automated Registration and Orthorectification Package (AROP) (Gao et al., 2009) (Step 1(a)). Master images are denoted in Table 1 with an asterisk. This ensures that glacier outlines are properly matched, and that glacier outlines derived from subsequent Landsat scenes are consistent in space. This step is considered optional for Landsat TM and ETM+, as the referencing only moves the images on the order of 1-2 pixels, and is not needed for Landsat OLI, as high-quality georeferencing is already included in the L1T product. Once the data were georeferenced and registered, a pair of scripts perform the glacier classification. The algorithm uses Landsat imagery, a hydrologically corrected DEM, a velocity surface derived from image cross-correlation, and the Hydrosheds 15s river network (buffered by 200m and converted to a raster) as the primary inputs (Steps 1(b) and 1(c)). The algorithm generates slope and hillshade images from the DEM according to image date, time, and location, and rectifies additional

input datasets described below for processing by resampling and reprojecting each dataset to the same spatial extent and resolution (30m to match the Landsat data) (Steps 1(e) and 1(f)). Although the current algorithm leverages a few proprietary Matlab commands, we will continue to update the code with the goal of using only open-source tools and libraries in the future.

3.2 Lake Delineation

Glacial debris tongues tend to host supra-glacier lakes, particularly during the summer months when snow cover is lowest (Quincey et al., 2007; Gardelle et al., 2011). These lakes are used as seed points for distance-weighting mechanisms to more accurately delineate glacier debris tongues. Lakes are delineated using the NDWI (Gao, 1996), after which misclassified areas are removed by masking out high slopes and shadowed areas (Step 2(a)) (Huggel et al., 2002; Worni et al., 2013; Nie et al., 2013; Hanshaw and Bookhagen, 2014). To increase the modularity of the algorithm, we did not rely on a fixed NDWI value across all datasets, but instead used a manually generated set of index lakes that exhibit the spectral properties desired. Several lakes in each Landsat scene are selected, with a range of geographic settings (e.g., far removed from any glacier, on top of a debris tongue, at the base of a glacier). A small spectral buffer (0.025) is then added to the by-scene NDWI value to account for lakes with differing spectral signatures to those of the index lakes (generally below -0.05). This small manual step greatly increases the number of correctly classified lakes by providing scene-specific NDWI thresholds for each processed image.

3.3 Clean Ice Delineation

Calculations are performed on rasterized versions of each input dataset, which have been standardized to the same matrix size. The first step in the classification process leverages Landsat Bands 1, 3, and 5 (Step 2(b)). For Landsat 8 OLI images, a slightly different set of bands is used to conform to OLI's modified spectral range. For simplicity, bands referenced in this publication refer to Landsat 7 ETM+ spectral ranges. The ratio of TM3/TM5 (value $\geq$2), with additional spectral information from TM1 (value $>$250) has been used in previous research as an effective means of delineating glacier areas (e.g., Hall et al., 1987; Hanshaw and Bookhagen, 2014), but is not effective in delineating debris-covered glacier areas (Figure 2). In our algorithm, we use a threshold of TM3/TM5 $\geq$ 2 and TM1 $>$60 to map clean glacier ice. The end result of this step is the spectrally-derived glacier outlines, which are later integrated back into the workflow before statistical filtering (Figure 2). Here

we choose the fairly conservative threshold values to ensure that we do not remove clean glacier ice. We find that increasing the TM1 threshold results in tighter classification of glacier debris tongues, but also removes some areas properly classified as glacier, particularly at high elevations. Thus, we err on the side of overclassification with our delineation of clean glacier ice.

3.4 Debris-covered Ice Delineation

3.4.1 Topographic Filtering

Building on the work of Paul et al. (2004), low slope areas (between 1 and 24°) are isolated as areas where debris-covered glaciers are likely to exist (Step 2(c)). As glacier surfaces tend to be rougher than the surrounding areas, a standard deviation filter (3x3 kernel size) is also applied to the slope and used to mask out areas of low slope variability. Low elevation areas (defined on a scene-by-scene basis, generally below 2500-3000m in our study area) are then masked out to decrease processing time (Step 2(d)). These thresholding steps are performed independent of the previous, spectrally delineated, glacier outlines. In essence, this step identifies areas where there is the potential for a debris covered glacier to exist. Additional thresholding is then performed on this 'potential debris area' subset to identify debris-covered glacier areas (Figure 3).

As can be seen in Figure 3, extensive areas which are not glacier or glacier debris tongue are identified in this step. However, this step generally removes all pixels outside of the main glacierized areas of any scene, and allows the algorithm to work on a subset of the image. The next step uses a generalized velocity surface to subset the 'potential debris area'.

3.4.2 Velocity Filtering

The Correlation Image Analysis Software (CIAS) (Kääb, 2002) tool, which uses a method of statistical image cross-correlation, is used to derive glacier velocities from Landsat Band 8 panchromatic images. This method functions by tracking individual pixels across space and time, and provides a velocity surface at the same resolution as the input datasets (15m) (Step 1(b)). The velocity surface is then upsampled using bilinear resampling to provide a consistent velocity estimate across the entire Landsat scene. We then standardized the velocity measurements to m/yr using the capture dates of the two Landsat images. As glacier velocity can change significantly throughout the year, and clean images were not available for at exactly the same intervals for each Path/Row combination, there is some error in our velocity fields. We mitigate this range in velocities by using Path/Row-

specific velocity thresholds during our classification process. It is important to note that cloud-free and snow-free images are essential for this step, as the presence of snow or cloud cover can disrupt the correlation process, resulting in anomalous velocity measurements. An example binary velocity mask is shown in Figure 4 (Step 2(e)). This mask shows areas which are removed from the 'potential debris areas' in red, as they fall outside of the expected range of debris-tongue velocities.

We only used one multi-year velocity measurement for each path/row combination to derive general areas of movement/stability for glacier classification, as using stepped velocity measurements over smaller time increments did not show a noticeable improvement in glacier classification. These velocities ranged generally from 4-30 m/yr across the different scenes, with very few glacier pixels falling outside of this range. Scene-specific thresholds were chosen based on both manual inspection of the velocity surfaces, and their impact on glacier classification on a scene-specific test dataset. Furthermore, a method of frequential cross-correlation using the co-registration of optically sensed images and correlation (COSI-Corr) tool (Leprince et al., 2007; Scherler et al., 2011b) was tested and did not show any appreciable improvement in velocity measurements (Heid and Kääb, 2012).

The velocity step is most important for removing hard-to-classify pixels along the edges of glaciers, and wet sands in riverbeds. These regions are often spectrally indistinguishable from debris tongues, but have very different velocity profiles. It is important to note, however, that this step also removes some glacier area, as not all parts of a glacier are moving at the same speed. This can result in small holes in the delineated glaciers, which the algorithm attempts to rectify using statistical filtering.

3.4.3 Spatial Weighting

After topographic and velocity filtering, a set of spatially weighted filters was constructed. The first filtering step uses the Hydrosheds river network to remove 'potential debris areas' which are distant from the center of a given glacier valley (Figure 5, Step 2(f)). As glaciers occur along the flowlines of rivers, and the Hydrosheds river network generally delineates flowlines all the way to the peaks of mountains, the river network provides an ideal set of seed points with which to remove misclassified pixels outside of river valleys. A second distance weighting is then performed using both the supra-glacier lakes detected during the lake delineation step, and any manual seed points provided (Step 2(g)). As debris tongues must occur in proximity to one or more of: (1) glacier areas, (2) the centerlines of valleys, or (3) supra-glacier lakes, these two steps are effective in removing overclassified areas (Figure 6). At this step, it is possible to add manual seed points, which may

be necessary for some longer debris tongues. We note that these are optional, and the majority of glaciers do not need the addition of manual seed points. However, for certain glaciers which do not have many lakes, or do not have lakes large enough to be delineated by Landsat, the addition of manual control points has been observed to increase the efficacy of the algorithm.

The spatial weighting step is essential for removing pixels spatially distant from any glacier or lake area. In many cases, large numbers of river pixels, and in some cases dry sand pixels, have similar spectral and topographic profiles to debris covered glaciers. This step effectively removes the majority of pixels outside the general glaciated area(s) of a Landsat scene, as can be seen in Figure 6, which shows the results of the velocity filtering (black), underneath the results of the spatial filtering (blue).

3.4.4 Statistical Filtering

Once the spatial weighting steps are completed, the glacier outlines are generally accurate. A set of three filters are then applied, in order to remove isolated pixels, bridge gaps between isolated glacier areas, and fill holes in large contiguous areas (Step 2(h)). First, a 3x3 median filter is applied, followed by an 'area opening' filter, which fills holes in contiguous glacier areas. Finally an 'image bridging' filter is applied to connect disjointed areas, and fill holes missed by the area opening filter.

This step is essential for filling holes and reconnecting separated glacier areas. As our initial filtering methods are based on a fixed set of threshold values, there are often glacier pixels which are removed. For example, some pixels in the middle of a debris tongue may be moving more slowly than the provided velocity threshold, and are thus removed. This problem is somewhat, but not completely, mitigated by the statistical filtering (Figure 7).

When the final filtered outlines are compared to the original spectral outlines, there is a clear improvement in classification of debris tongues (Figure 8).

3.5 Creation of Manual Control Datasets

Manual control datasets encompassing $\sim$750 glaciers ($\sim$3000 km^2) were created to test the efficacy of the glacier mapping algorithm. These datasets were digitized off of Landsat imagery in a GIS, and then corrected with higher resolution imagery in Google Earth. The datasets are coherent in space, but cover two different times ($\sim$2000 and $\sim$2011, depending on the dates of the available Landsat scenes). The bulk of the manually digitized glaciers fall within the boundary of Landsat Path/Row

combination 147/031, as this is the most heavily glacierized sub-region of our study area. However, we have digitized glaciers throughout the eight Path/Row combinations to avoid biasing our statistics and algorithm to one specific scene extent. We have also considered a wide range of size classes in our manual dataset ($<$0.5km^2 to 500+ km^2), as well as both clean and debris-covered glaciers. We note that although the manual datasets here are considered 'perfect', there is inherent error in any manual digitization in a GIS (e.g., Paul et al., 2013). Due to the lack of ground truth information, we have estimated the overall uncertainty of the manual dataset to be 2% based on previous experiments (Paul et al., 2002, 2013).

Before any comparisons between glaciers can be performed, glacier complexes must be split into component parts. A set of manually edited watershed boundaries, derived from the SRTM DEM, were used to split both the manual and algorithm datasets into individual glacier areas for analysis. In this way, the diverse datasets and classified glacier areas can be split into the same subset areas for statistical comparison.

4 Results

4.1 Statistical Analysis of Algorithm Errors

A subset of 215 glaciers from the manual control datasets of varying size and topographic setting was chosen for more detailed analysis. The unedited, algorithm-generated, glacier outlines were compared against both spectral outlines, which only classify the glacier areas via commonly used spectral subsetting (using TM1, TM3, and TM5, produced in Step 2(b)) and the manual control datasets. Figure 9 shows the bulk elevation distributions across 215 glaciers for each dataset in 10m elevation bins.

There is some apparent bias in our algorithm towards low elevation areas, which represent the debris-covered portions of glaciers and are the most difficult areas to classify. There is also a bias in our control dataset towards underclassifying the high elevation areas, which we attribute to user bias in removing isolated rock outcrops within glaciers, as opposed to simply defining accumulation areas as a single polygon. In general, the algorithm and the control dataset are well matched below 4000 meters; above this the spectral dataset and the algorithm dataset begin to align closely. This threshold represents the general transition from debris-covered glaciers to clean glacier ice in the study area.

In order to examine persistent bias throughout the algorithm classification, under- and over-classified areas were examined. To determine areas of overclassification (underclassification), the manually (algorithm) generated dataset was subtracted from the algorithm (manual) dataset, leaving only pixels which are overclassified (underclassified). Figure 10 shows the elevation distributions of under and over classified areas. The algorithm tends to consistently overclassify areas across the range of glacier elevations, which we attribute here to differences in manual and algorithm treatment of steep and de-glaciated areas within glacier accumulation zones. Importantly, the algorithm underclassifies a much smaller number of pixels, generally corresponding to areas below 4000m, where debris tongues are dominant. The majority of these pixels are along the edges of glacier debris tongues, which are removed by the algorithm due to their low relative velocity. It is also possible that some of these pixels are 'dead ice', which is difficult to differentiate from debris tongues by visual inspection. The total misclassification of algorithm-derived outlines against two independent manual control datasets are 2% and 10% respectively, which represents a significant improvement from a pure spectral delineation approach.

4.2 Vertex Distance Matching

To capture changes in the shape of the glacier outlines between the initial spectral classification and the final algorithm output, we computed the distance between pairs of glacier vertices. We first reduced our manual control dataset to component vertices, which were then matched to the closest vertex in the spectral and final algorithm results polygons, respectively. The results of this distance matching can be seen in Figure 11.

The distance distribution for the algorithm dataset shows generally close agreement between the algorithm and manual control datasets. The spectral dataset also contains a large percentage of vertices close to a 1:1 agreement with the manual control dataset, which are primarily those vertices at the upper edges of glaciers, or vertices from small, debris-free glaciers. The difference in these two distributions is attributed to the increased precision with which the algorithm maps debris-covered glacier outlines.

4.3 Comparison to a Random Sampling of Glacial Areas

In order to examine sampling bias in our analysis, the researchers used 465 GLIMS glacier identification numbers (centroids, point features) which overlapped with the manual control datasets. A

random subset of 100 of these points was chosen for this analysis. As can be seen in Figure 12, similar patterns emerge between the randomly sampled glaciers and the sampling used in other sections of this manuscript. There is evidence of more noise in the random sample, as some glaciers which we avoided due to closeness to wet sand or other hard-to-classify areas were chosen during the random sampling. However, the relationship between the algorithm and the manual datasets remains significant (ks-test passes at 99% confidence interval).

5 Discussion

5.1 Unused Filtering Steps

Two additional topographic indices – spatial Fast Fourier Transforms (FFTs), also known as 2D FFTs, and ASTER surface roughness measurements – were tested during the development of the algorithm, although neither provided significant improvement. We attempted to derive frequential information from several Landsat and ASTER bands, with limited success. Some glaciers exhibit a unique frequency signature when analyzed using spatial FFTs, although these were not consistent across multiple debris-covered glaciers with differing surface characteristics. Additionally, the FFT approach was tested against a principal component analysis (PCA) image derived from all Landsat bands, without significant improvement to the algorithm.

We also attempted to integrate surface roughness measurements using the ASTER satellite, which contains both forward looking (3N - nadir) and backwards looking (3B - backwards) images, primarily intended for the generation of stereoscopic DEMs. The difference in imaging angle provides the opportunity to examine surface roughness by examining changes in shadowed areas (Mushkin et al., 2006; Mushkin and Gillespie, 2011). We found that there are slight surface roughness differences between glaciated and non-glaciated areas on some debris tongues, but that these differences are not significant enough to use as a thresholding metric. Furthermore, the nature of the steep topography limits the efficacy of this method, as valleys which lie parallel the satellite flight path and those which lie perpendicular to the flight path show different results. Thus, the algorithm relies on the velocity and slope thresholds to characterize the topography of the glacier areas.

5.2 Algorithm Use Cases and Caveats

The glacier outlines provided by the algorithm are a useful first pass analysis of glacier area. It is often more efficient to digitize only misclassified areas, as opposed to digitizing entire glacier areas by hand (Paul et al., 2013). Paul et al. (2013) also note that for clean ice, automatically derived glacier outlines tend to be more accurate, and it is only in the more difficult debris-covered and shadowed areas that manual digitization becomes preferable. In the algorithm presented here, clean ice thresholding was implemented using TM1, TM3, and TM5, but as the algorithm primarily operates on 'potential debris areas', as opposed to whole glaciers, different spectral classification schemes could be easily substituted for regions where other spectral measurements, such as the Normalized Difference Snow Index, perform best.

The algorithm moves a step further than spectral-only classification and attempts to classify glacier areas as accurately as possible, including debris-covered areas. As can be seen in Figure 13, the algorithm compares well with both the control dataset and the RGI v4.0 – a $\sim$198,000 glacier global dataset derived from multiple data sources – across a range of glacier types (Step 2(a)). However, it does not perfectly align with either dataset. In Figure 13, a tendency to remove pixels along the edge of glacier debris tongues can be observed, which we attribute to the fact that the center of debris tongues often move faster than the edges. Furthermore, both the algorithm results and the manual control dataset underestimate glacier area as compared to the RGI, due to the removal of non-clean ice pixels at high altitudes or high slopes, which are generally within the accumulation area of a glacier but rarely are covered by permanent ice. These two types of classification bias are easily rectified with minimal manual intervention. Furthermore, the raw algorithm output identifies the furthest reaches of the glacier tongues effectively in most cases, as can be seen in three long debris tongues shown in Figure 13.

Without post-processing, these raw glacier outlines can be used to analyze regional glacier characteristics, such as slope, aspect, and hypsometry. Even if glacier outlines are not perfectly rectified in space, at the scale of watersheds, satellite image footprints, or mountain ranges, errors of under- and over-classification even out, yielding valuable regional statistics (Figure 9). As the method can be easily modified to fit the topographic and glacier setting of any region, it is a powerful tool for analyzing glacier changes over large scales over the period of Landsat TM, ETM+ and OLI coverage. Small glacier changes are also captured by the algorithm, as can be seen in Figure 14.

Figure 14 also illustrates some potential errors with the algorithm where river sand is sometimes

delineated as glacier area. In many cases, the same areas are captured across different timestamps, as the topographic and velocity data used to define 'potential debris areas' is mostly static in time, excepting the distance weighting steps. However, these areas are generally small and easily removed during manual inspection of results.

The second use case for the algorithm is as a substitute for simple spectral ratios. Particularly in regions with numerous debris-covered glaciers, manual digitization of glacier tongues is time consuming. Our algorithm provides a robust baseline set of glacier outlines which can be corrected manually, with minimal extra processing time. Processing time is also decreased when a large set of Landsat scenes are considered, as generating the input velocity surfaces can take longer than processing glacier outlines from dozens of Landsat scenes.

Although the algorithm represents a step towards improved glacier classification, there are several important caveats to keep in mind: (1) Lack of data density and temporal range limits the efficacy of individual glacier analysis; the algorithm presented in this paper was not designed with individual glacier studies in mind, and in many cases, such as in mass balance studies, more accurate manual glacier outlines are necessary. Furthermore, (2) the algorithm relies on manual intervention to separate individual glaciers which are connected through overlapping classified areas, or which are part of glacier complexes.

Finally, (3) the algorithm relies heavily on the fidelity of the Landsat images provided, in that glacier outlines on images with cloud cover or snow cover are less likely to be well defined. This creates a data limitation, as many glaciated areas are subject to frequent snow and cloud cover, and thus have a limited number of potentially useful Landsat images for the purpose of this algorithm. We have found that OLI images are generally better classified than TM/ETM+ images, potentially due to the increased sensor range onboard OLI reduces saturation in the spectral bands used for glacier ice detection. The difference in the sensors (8-bit on TM/ETM+ and 12-bit on OLI) allows for better radiometric quantization and a higher signal to noise ratio. The change of OLI to a pushbroom sensor also results in higher radiometric sensitivity. Reduced saturation in the spectral domain reduces the impact of atmospheric moisture which is present when there are clouds within the scene. As the glaciers are classified by ratios of spectral bands, having a wider range of possible TM 1, 3, and 5 values reduces errors due to band saturation. However, this conclusion will need to be verified with further data acquisitions, and remains conjecture.

6 Conclusions

This study presents an enhanced glacier classification methodology based on the spectral, topographic, and spatial characteristics of glaciers. Our algorithm represents a step forward toward (semi-) automated glacier classification, in that it outperforms spectral-only algorithms in wide use in the glaciology and remote sensing communities. Although it does not completely solve the difficulties associated with debris-covered glaciers, it can effectively and rapidly characterize glaciers over a wide area. The first steps of this algorithm, which seek to characterize maximum possible glacier area, are lake delineation, spectral delineation, and topographic filtering. Following these steps, a set of velocity, spatial, and statistical filters are applied to accurately delineate glacier outlines, including their debris-covered areas.

When compared visually and statistically against a control dataset, or a wide-area glacier database such as the Randolph Glacier Inventory v4.0, our algorithm remains robust across the wide range of glacier sizes and types found in Central Asia. The algorithm developed here will be applicable to a wide range of glaciated regions, particularly in those regions where debris-covered glaciers are dominant, and extensive manual digitization of glacier areas is required. The raw output is usable for rough statistical queries on glacier area, hypsometry, slope, and aspect; however, manual intervention should be applied before using algorithm glacier outlines for more in-depth area change or mass balance studies.

Acknowledgements. This work was supported through the Earth Research Institute (UCSB) through a Natural Hazards Research Fellowship, as well as the NSF grant AGS-1116105. We would like to thank Frank Paul and Wanqin Guo for their detailed and helpful reviews, as well as Tobias Bolch for his contribution to the development of the manuscript.

References

Aizen, V., Aizen, E., and Kuzmichonok, V.: Glaciers and hydrological changes in the Tien Shan: simulation and prediction, Environmental Research Letters, 2, 045 019, 2007.

Arendt, A., Bolch, T., Cogley, J., Gardner, A., Hagen, J., Hock, R., Kaser, G., Pfeffer, W., Moholdt, G., Paul, F., et al.: Randolph Glacier Inventory [v2. 0]: A Dataset of Global Glacier Outlines. Global Land Ice Measurements from Space, Boulder Colorado, USA, Digital Media, 2012.

Armstrong, R., Raup, B., Khalsa, S., Barry, R., Kargel, J., Helm, C., and Kieffer, H.: GLIMS glacier database, National Snow and Ice Data Center, Boulder, Colorado, USA, 2005.

Berthier, E., Arnaud, Y., Kumar, R., Ahmad, S., Wagnon, P., and Chevallier, P.: Remote sensing estimates of glacier mass balances in the Himachal Pradesh (Western Himalaya, India), Remote Sensing of Environment, 108, 327–338, 2007.

Bolch, T., Buchroithner, M. F., Kunert, A., and Kamp, U.: Automated delineation of debris-covered glaciers based on ASTER data, in: Geoinformation in Europe (Proc. of 27th EARSel Symposium, 04-07 June 2007), Bozen, Italy, pp. 403–410, 2007.

Bolch, T., Peters, J., Yegorov, A., Pradhan, B., Buchroithner, M., and Blagoveshchensky, V.: Identification of potentially dangerous glacial lakes in the northern Tien Shan, Natural Hazards, 59, 1691–1714, 2011.

Bolch, T., Kulkarni, A., Kääb, A., Huggel, C., Paul, F., Cogley, J., Frey, H., Kargel, J., Fujita, K., Scheel, M., et al.: The state and fate of Himalayan glaciers, Science, 336, 310–314, 2012.

Cannon, F., Carvalho, L., Jones, C., and Bookhagen, B.: Multi-annual variations in winter westerly disturbance activity affecting the Himalaya, Climate Dynamics, pp. 1–15, 2014.

Farr, T. G. and Kobrick, M.: Shuttle Radar Topography Mission produces a wealth of data, Eos, Transactions American Geophysical Union, 81, 583–585, 2000.

Gao, B.-C.: NDWI - a normalized difference water index for remote sensing of vegetation liquid water from space, Remote sensing of environment, 58, 257–266, 1996.

Gao, F., Masek, J., and Wolfe, R. E.: Automated registration and orthorectification package for Landsat and Landsat-like data processing, Journal of Applied Remote Sensing, 3, 033 515–033 515, 2009.

Gardelle, J., Arnaud, Y., and Berthier, E.: Contrasted evolution of glacial lakes along the Hindu Kush Himalaya mountain range between 1990 and 2009, Global and Planetary Change, 75, 47–55, 2011.

Gardelle, J., Berthier, E., and Arnaud, Y.: Slight mass gain of Karakoram glaciers in the early twenty-first century, Nature geoscience, 5, 322–325, 2012.

Gardelle, J., Berthier, E., Arnaud, Y., and Kääb, A.: Region-wide glacier mass balances over the Pamir-Karakoram-Himalaya during 1999-2011., Cryosphere, 7, 2013.

Hall, D., Ormsby, J., Bindschadler, R., and Siddalingaiah, H.: Characterization of snow and ice reflectance zones on glaciers using Landsat Thematic Mapper data, Ann. Glaciol, 9, 1–5, 1987.

Hanshaw, M. N. and Bookhagen, B.: Glacial areas, lake areas, and snowlines from 1975 to 2012: status of the Cordillera Vilcanota, including the Quelccaya Ice Cap, northern central Andes, Peru., Cryosphere Discussions, 7, 2014.

Heid, T. and Kääb, A.: Evaluation of existing image matching methods for deriving glacier surface displacements globally from optical satellite imagery, Remote Sensing of Environment, 118, 339–355, 2012.

Huggel, C., Kääb, A., Haeberli, W., Teysseire, P., and Paul, F.: Remote sensing based assessment of hazards from glacier lake outbursts: a case study in the Swiss Alps, Canadian Geotechnical Journal, 39, 316–330, 2002.

Kääb, A.: Monitoring high-mountain terrain deformation from repeated air-and spaceborne optical data: examples using digital aerial imagery and ASTER data, ISPRS Journal of Photogrammetry and remote sensing, 57, 39–52, 2002.

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

Lehner, B., Verdin, K., and Jarvis, A.: New global hydrography derived from spaceborne elevation data, EOS, Transactions American Geophysical Union, 89, 93–94, 2008.

Leprince, S., Ayoub, F., Klingert, Y., and Avouac, J.-P.: Co-registration of optically sensed images and correlation (COSI-Corr): An operational methodology for ground deformation measurements, in: Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International, pp. 1943–1946, IEEE, 2007.

Lioubimtseva, E. and Henebry, G. M.: Climate and environmental change in arid Central Asia: Impacts, vulnerability, and adaptations, Journal of Arid Environments, 73, 963–977, 2009.

Mushkin, A. and Gillespie, A.: Using ASTER Stereo Images to Quantify Surface Roughness, in: Land Remote Sensing and Global Environmental Change, pp. 463–481, Springer, 2011.

Mushkin, A., Gillespie, A., Danilina, I., O'Neal, M., Pietro, L., Abbott, E., and Balick, L.: Using sub-pixel roughness estimates from ASTER stereo images to compensate for roughness effects in the thermal infrared, in: RAQRS II: 2nd International Symposium on Recent Advances in Quantitative Remote Sensing, 2006.

Narama, C., Kääb, A., Duishonakunov, M., and Abdrakhmatov, K.: Spatial variability of recent glacier area changes in the Tien Shan Mountains, Central Asia, using Corona (˜ 1970), Landsat (˜ 2000), and ALOS (˜ 2007) satellite data, Global and Planetary Change, 71, 42–54, 2010.

Nie, Y., Liu, Q., and Liu, S.: Glacial Lake Expansion in the Central Himalayas by Landsat Images, 1990–2010, PloS one, 8, e83 973, 2013.

Oerlemans, J.: Extracting a climate signal from 169 glacier records, Science, 308, 675–677, 2005.

Paul, F.: Changes in glacier area in Tyrol, Austria, between 1969 and 1992 derived from Landsat 5 Thematic Mapper and Austrian Glacier Inventory data, International Journal of Remote Sensing, 23, 787–799, 2002.

Paul, F., Kaab, A., Maisch, M., Kellenberger, T., and Haeberli, W.: The new remote-sensing-derived Swiss glacier inventory: I. Methods, Annals of Glaciology, 34, 355–361, 2002.

Paul, F., Huggel, C., and Kääb, A.: Combining satellite multispectral image data and a digital elevation model for mapping debris-covered glaciers, Remote Sensing of Environment, 89, 510–518, 2004.

Paul, F., Barrand, N., Baumann, S., Berthier, E., Bolch, T., Casey, K., Frey, H., Joshi, S., Konovalov, V., Bris, R. L., et al.: On the accuracy of glacier outlines derived from remote-sensing data, Annals of Glaciology, 54, 171–182, 2013.

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

Quincey, D., Richardson, S., Luckman, A., Lucas, R., Reynolds, J., Hambrey, M., and Glasser, N.: Early recognition of glacial lake hazards in the Himalaya using remote sensing datasets, Global and Planetary Change, 56, 137–152, 2007.

Racoviteanu, A. E., Arnaud, Y., Williams, M. W., and Ordonez, J.: Decadal changes in glacier parameters in the Cordillera Blanca, Peru, derived from remote sensing, Journal of Glaciology, 54, 499–510, 2008a.

Racoviteanu, A. E., Williams, M. W., and Barry, R. G.: Optical remote sensing of glacier characteristics: a review with focus on the Himalaya, Sensors, 8, 3355–3383, 2008b.

Racoviteanu, A. E., Paul, F., Raup, B., Khalsa, S. J. S., and Armstrong, R.: Challenges and recommendations in mapping of glacier parameters from space: results of the 2008 Global Land Ice Measurements from Space (GLIMS) workshop, Boulder, Colorado, USA, Annals of Glaciology, 50, 53–69, 2009.

Raup, B., Kääb, A., Kargel, J. S., Bishop, M. P., Hamilton, G., Lee, E., Paul, F., Rau, F., Soltesz, D., Khalsa, S. J. S., et al.: Remote sensing and GIS technology in the Global Land Ice Measurements from Space (GLIMS) project, Computers & Geosciences, 33, 104–125, 2007.

Raup, B. H., Khalsa, S. J. S., Armstrong, R. L., Sneed, W. A., Hamilton, G. S., Paul, F., Cawkwell, F., Beedle, M. J., Menounos, B. P., Wheate, R. D., et al.: Quality in the GLIMS Glacier Database, in: Global Land Ice Measurements from Space, pp. 163–182, Springer, 2014.

Scherler, D., Bookhagen, B., and Strecker, M. R.: Spatially variable response of Himalayan glaciers to climate change affected by debris cover, Nature Geoscience, 4, 156–159, 2011a.

Scherler, D., Bookhagen, B., and Strecker, M. R.: Hillslope-glacier coupling: The interplay of topography and glacial dynamics in High Asia, Journal of Geophysical Research: Earth Surface, 116, 2011b.

Sorg, A., Bolch, T., Stoffel, M., Solomina, O., and Beniston, M.: Climate change impacts on glaciers and runoff in Tien Shan (Central Asia), Nature Climate Change, 2, 725–731, 2012.

Stocker, D. Q.: Climate change 2013: The physical science basis, Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Summary for Policymakers, IPCC, 2013.

Worni, R., Huggel, C., and Stoffel, M.: Glacial lakes in the Indian Himalayas - From an area-wide glacial lake inventory to on-site and modeling based risk assessment of critical glacial lakes, Science of the Total Environment, 468, S71–S84, 2013.

Table 1. Data table listing Landsat capture dates used in this study. Organized by WRS2 Path/Row combinations. Starred dates denote 'Master' images to which others were rectified. Bold dates indicate use for Velocity profiles.

	144/030	145/030	147/031
Number of Images	11	10	12
Date Range of Images	2002-2013		2000-2013
LT5 Capture Dates	Jul 31, 2006	Sep 2, 1998*	Aug 19, 2011
	Aug 8, 2009	Oct 4, 1998	Oct 2, 1998
	Sep 27, 1998	Jul 22, 2006	Sep 6, 2006
	Jul 13, 2011	**Aug 10, 2007**	Aug 24, 2007
		Sep 11, 2007	Oct 3, 2010
		Oct 2, 2009	Aug 3, 2011
		Aug 2, 2010	
		Jul 4, 2011	
		Sep 6, 2011	
LE7 Capture Dates	**Sep 14, 2002***		Sep 14, 2000*
	Jul 7, 2000		**Oct 5, 2002**
	Aug 8, 2000		Aug 18, 2002
	Jun 7, 2001		
LC8 Capture Dates	**Oct 22, 2013**	**Sep 27, 2013**	**Sep 25, 2013**
	Aug 19, 2013		Sep 9, 2013
	Sep 4, 2013		May 7, 2014
Projection	WGS 1984 45N		WGS 1984 44N
Comments	Eastern Edge of Study Area		Vicinity of Inylchek Glacier
	148/031	149/031	150/032
Number of Images	13	3	5
Date Range of Images	2002-2013	1999-2013	1998-2013
LT5 Capture Dates	Sep 16, 2007	Sep 7, 2007	Oct 23, 1998
	Sep 11, 2011		Jul 1, 2009
	Aug 22, 1998		
	Aug 12, 2006		
	Sep 13, 2006		
	Jul 30, 2007		
LE7 Capture Dates	**Jul 24, 2002***	**Sep 9, 1999***	**Aug 20, 2001***
	Jul 16, 1999		Sep 24, 2002
	Sep 18, 1999		
	Aug 25, 2002		
LC8 Capture Dates	**Jul 30, 2013**	**Oct 9, 2013**	**Jun 10, 2013**
	Oct 2, 2013		
	May 14, 2014		
Projection	WGS 1984 44N	WGS 1984 43N	WGS 1984 43N
Comments			
	151/033	153/033	
Number of Images	5	3	
Date Range of Images	1998-2013	1998-2013	
LT5 Capture Dates	Sep 28, 1998	Sep 26, 1998	
	Sep 10, 2009		
LE7 Capture Dates	Aug 24, 2000*	**Sep 29, 2002***	
	Sep 28, 2001		
LC8 Capture Dates	**Oct 7, 2013**	**Oct 5, 2013**	
Projection	WGS 1984 43N	WGS 1984 42N	
Comments		Towards Pamir Knot	

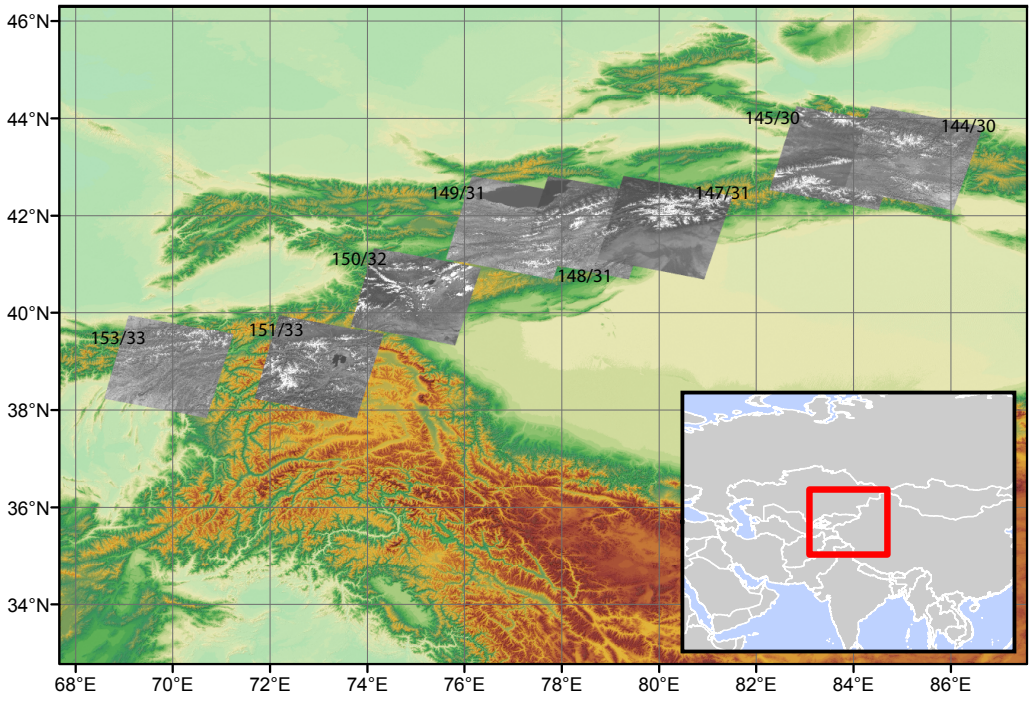

Fig. 1. Greater study area of the Tien Shan, showing SRTM v4.1 topography (Farr and Kobrick, 2000) and locations of eight Landsat scenes outlines (grayscale) used in the study, as well as their Path/Row combinations.

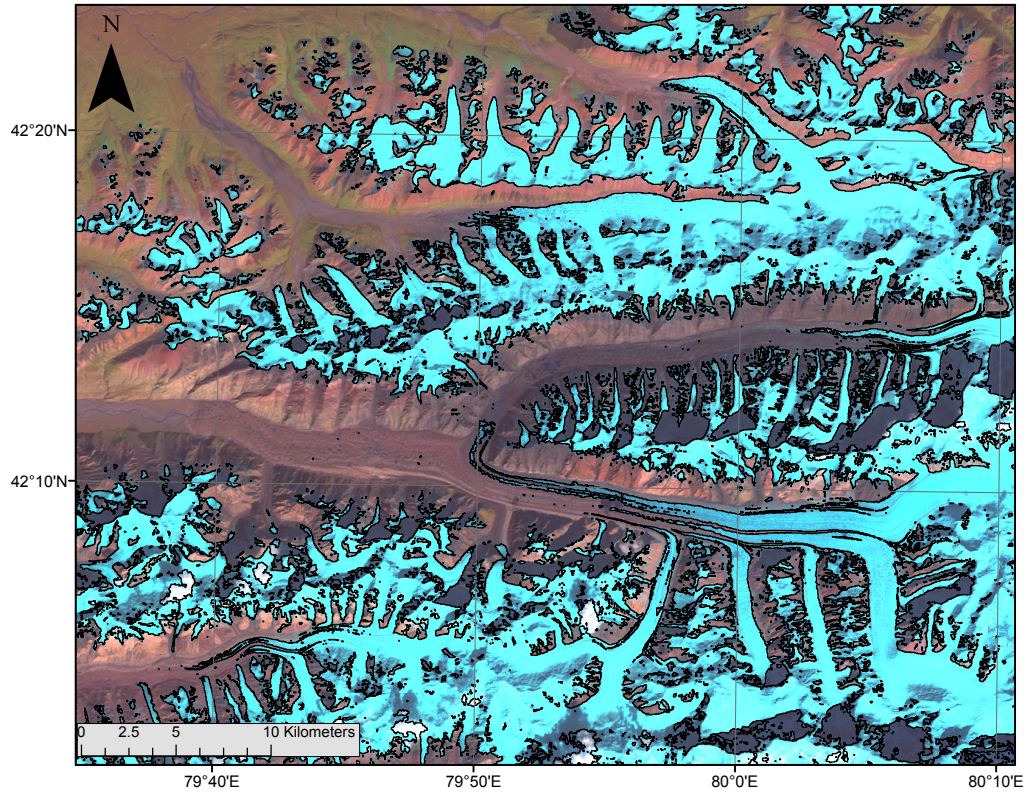

Fig. 2. Characteristic example of a debris-covered glacier tongue (Inylchek Glacier). Spectrally-delineated glacier outlines (black), over Landsat bands B7/B5/B3 loaded as Red/Green/Blue, from image LC81470312013268LGN00. Showing generally well-mapped clean ice, but poor treatment of debris-covered tongues.

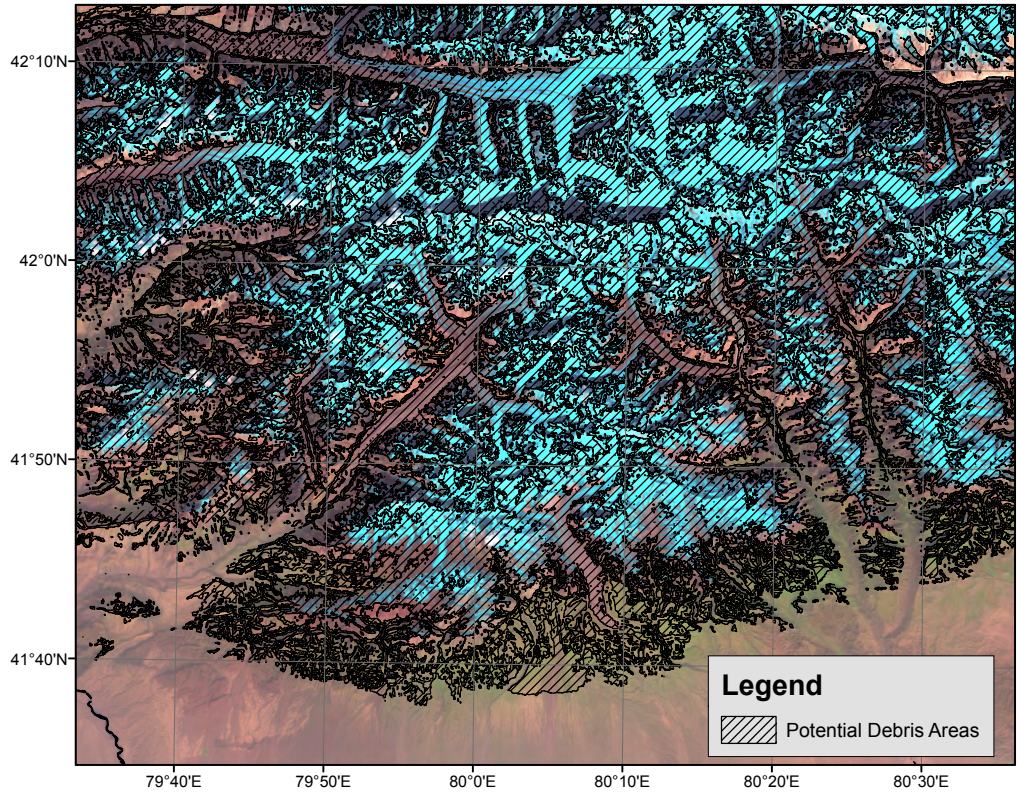

Fig. 3. Black lines show 'Potential Debris Areas', as delineated by slopes between 1-24 degrees, and slope variability (via 3x3 standard deviation filter) above 2, with elevations below 2500m removed. Only these areas are examined in the subsequent thresholding steps, to reduce processing time and misclassification errors.

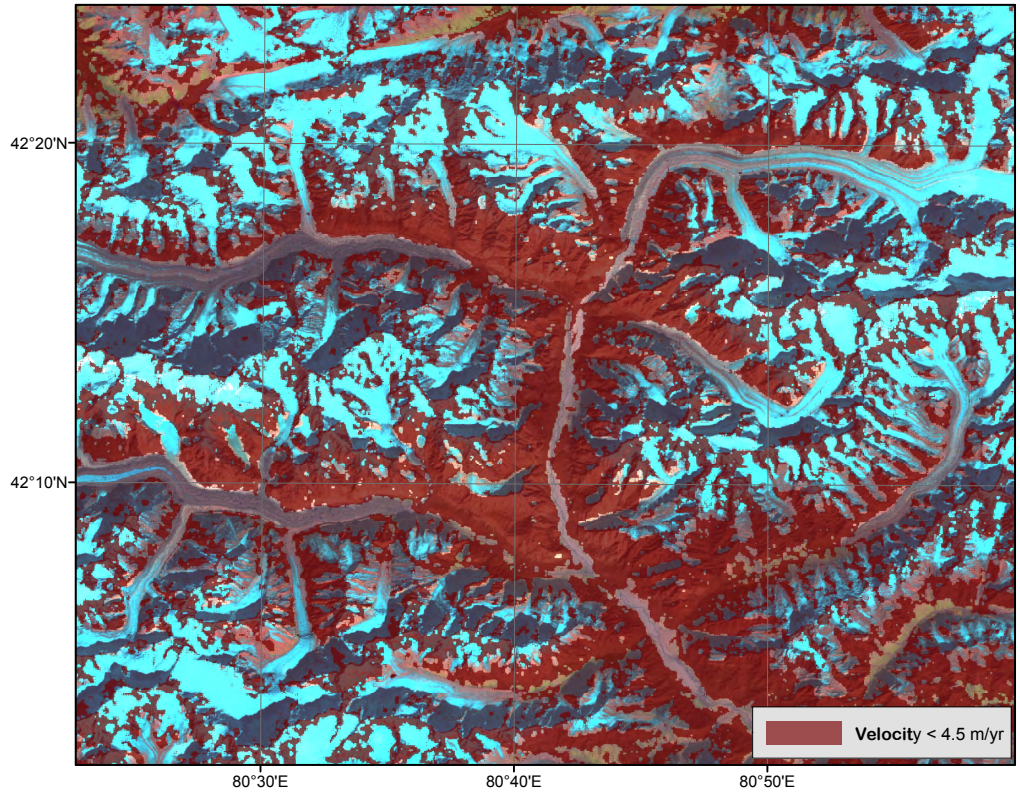

Fig. 4. Example binary velocity mask, generated using Normalized Image Cross Correlation. Areas in red are removed from the 'potential debris areas', because they do not show strong surface movements. Debris-tongues and clean-ice areas visible as B7/5/3 image LC81470312013268LGN00.

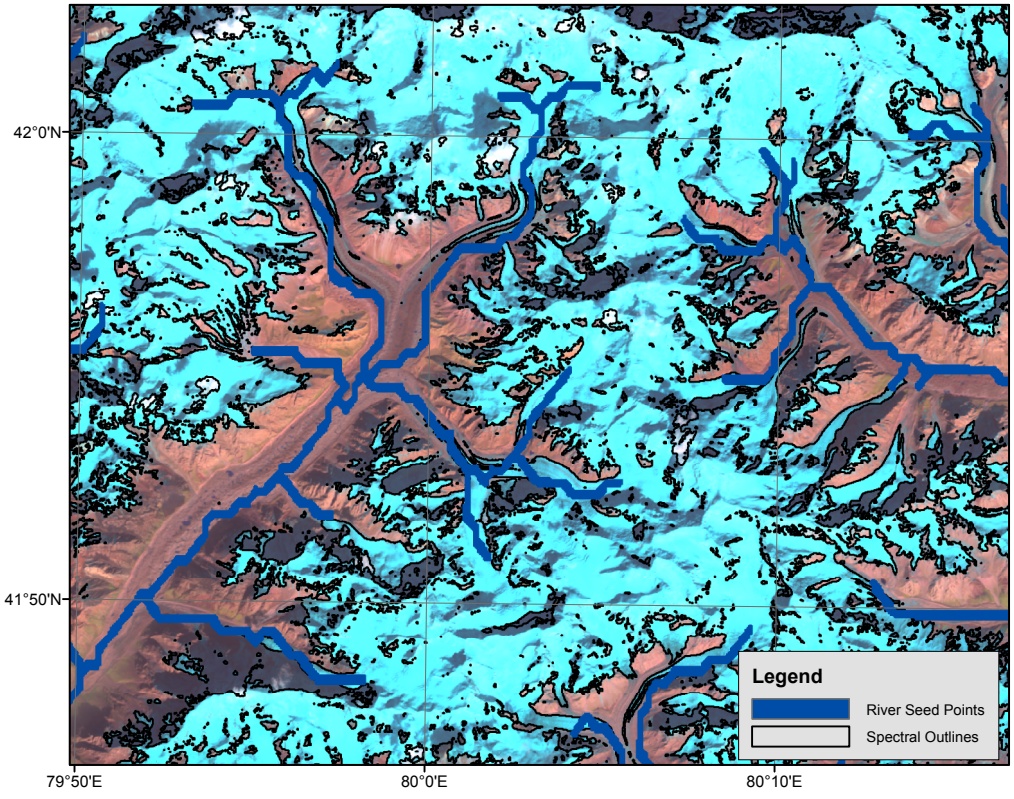

Fig. 5. Example of distance weighting seed points used to remove pixels from the 'potential debris areas' which are distant from either a river valley or classified glacier ice. Rivers in blue, spectrally-delineated glaciers in black. Illustrates the presence of river networks along debris-covered tongues of glaciers where there is little clean glacier ice.

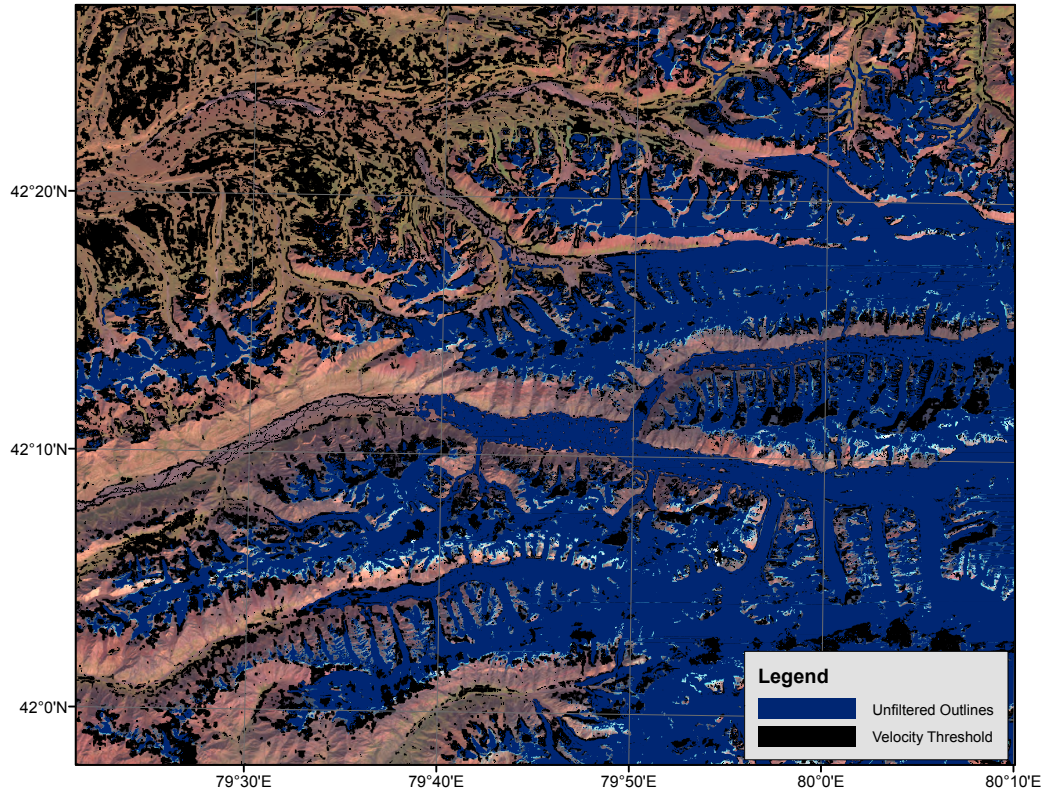

Fig. 6. Example distance weighting classification to remove areas distant from valleys and glacier areas. Results of the velocity filtering in black underneath the results of the distance filtering in blue, to illustrate the difference in classification after the distance threshold step. The distance weighting primarily removes pixels distant from glacier areas, such as the large areas of black in the top left.

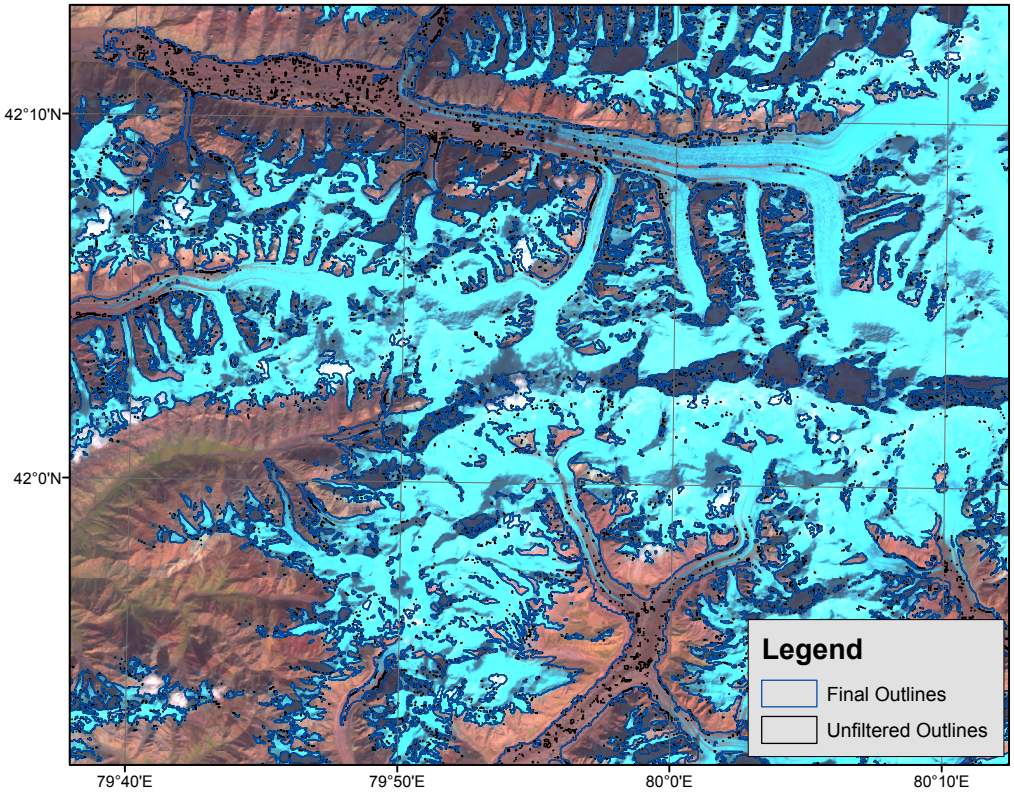

Fig. 7. Impacts of statistical filtering on glacier outlines, with areas in black removed during the filtering process. Primarily small holes in large debris-tongues are removed, while the glacier outlines remain intact during this step. This can be seen on both the Inylchek and Tomur glaciers in this figure.

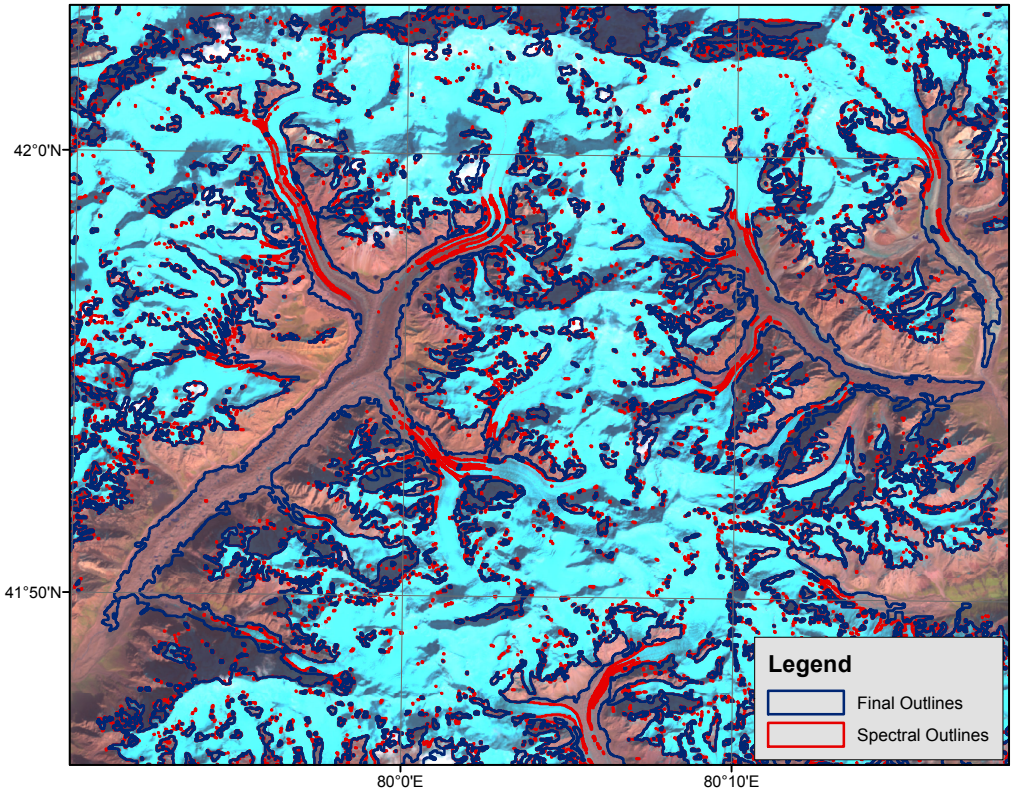

Fig. 8. Final algorithm results (blue), with spectral outlines (red). Illustrates the improved classification by the algorithm across several large debris tongues. Vicinity of the Tomur glacier, OLI image captured Sept 25, 2013 as background.

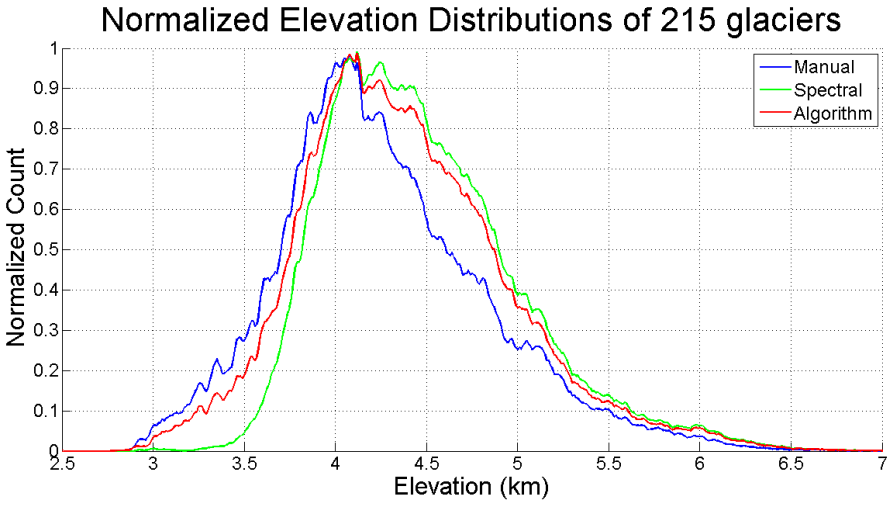

Fig. 9. Bulk elevation distributions of sampled glaciers, with manual delineation (reference dataset, n=215) in blue, algorithm-derived delineation in red, and spectral delineation in green. Illustrates high fidelity algorithm outlines in low-elevation areas, and generally tight agreement between spectral-delineated and algorithm outlines at elevations above 4000m.

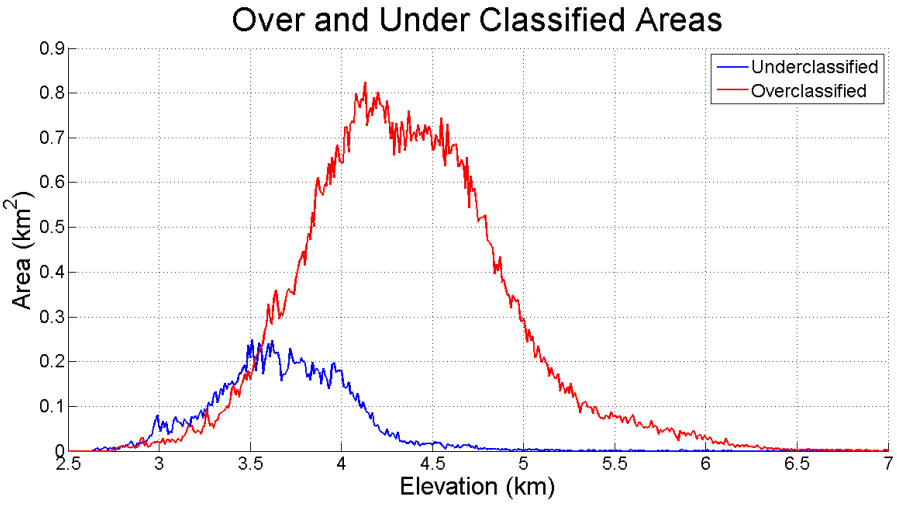

Fig. 10. Elevation distributions of over- and under-classified glacier areas, as compared to a manual control dataset (n=215). Overclassified areas show that the algorithm is weighted towards overclassifying high elevation areas, and does not remove large portions of the accumulation area. Underclassified areas indicate that the algorithm does not perfectly classify low-elevation areas, but that it misclassifies a relatively small area percentage.

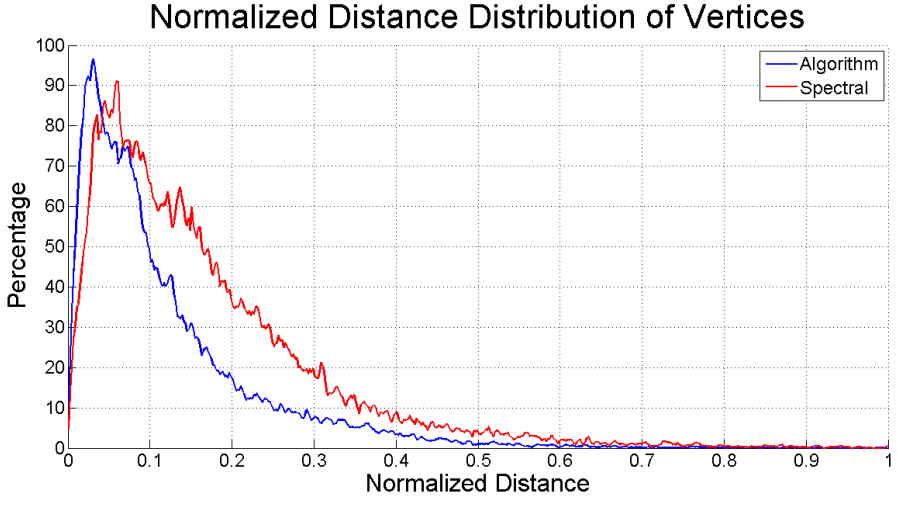

Fig. 11. Normalized vertex distance distributions for algorithm (blue) and spectral (red) vertices, as compared to a manual control dataset. This indicates generally tighter agreement between the algorithm and manual datasets than between the spectral and manual datasets.

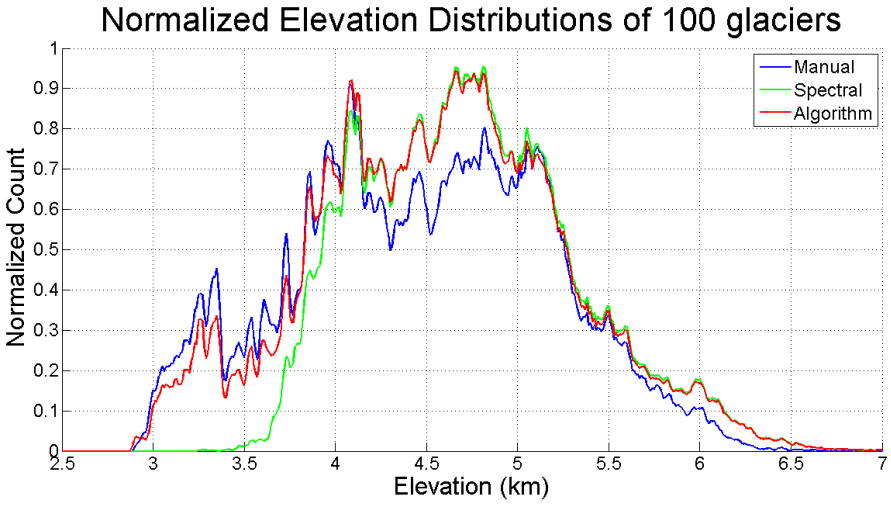

Fig. 12. Averaged elevation differences for a random sample of glaciers overlapping a manual control dataset (n=100). Shows generally tight agreement between the manual glacier dataset and the algorithm dataset below 4000m, with tighter agreement between the spectral and algorithm datasets above 4000m. Indicates improved mapping of debris-tongues by the algorithm, and similar treatment of clean ice by both the algorithm and the spectrally-delineated glaciers.

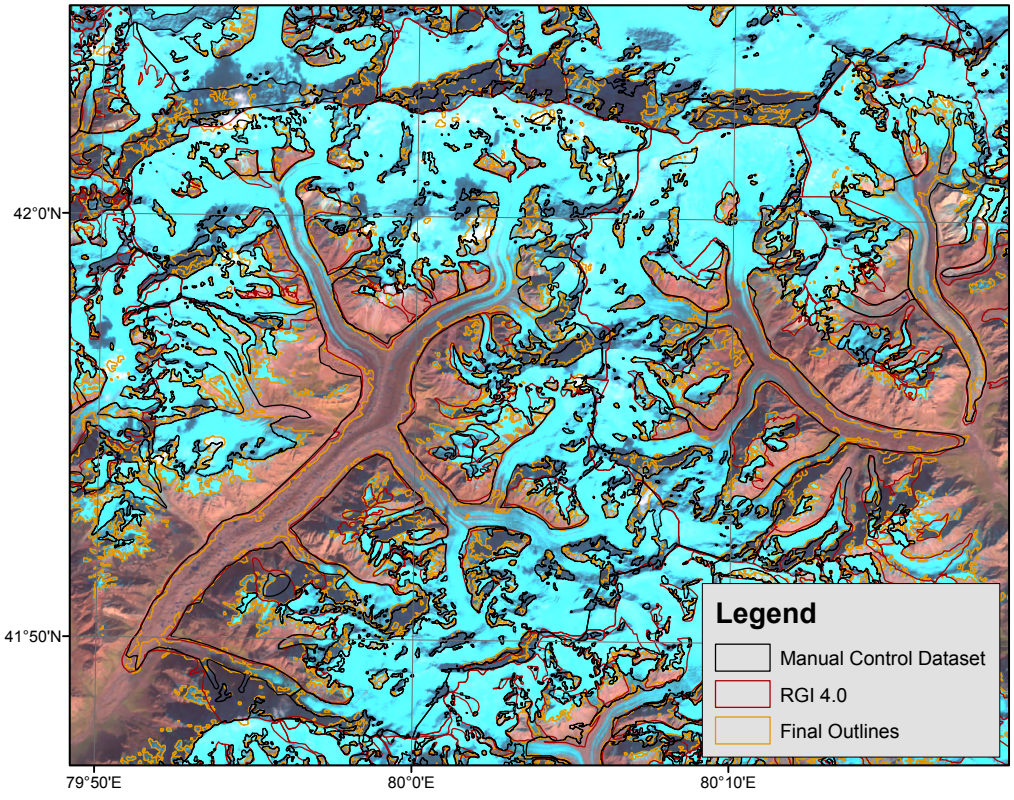

Fig. 13. Algorithm outlines (orange) compared to the control dataset (black) and the RGI v4.0 (red). Illustrates high fidelity in overall debris-tongue length between the three datasets, although the algorithm outlines exhibit noise along the edges of debris tongues.

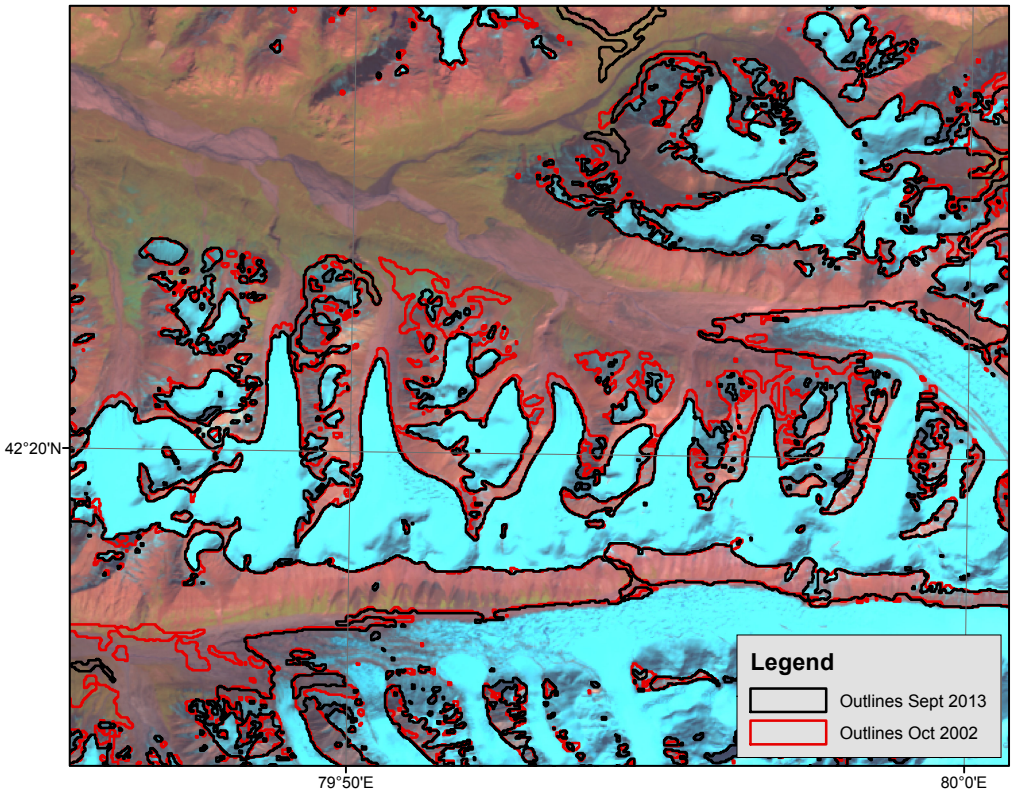

Fig. 14. Algorithm outlines for September 2013 (black) and algorithm outlines for October 2002 (red), showing small retreats in glacier areas, particularly at the debris tongues. Also shows an example of commonly misclassified river sand.