Improving Semi-Automated Glacier 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 author:
Taylor Smith
Institute for Earth and Environmental Sciences
Universität Potsdam
Potsdam-Golm 14476, Germany
Email: tsmith@uni-potsdam.de

Abstract.

Studies of glaciers generally require precise glacier outlines. Where these are not available, extensive manual digitization in a Geographic Information System (GIS) must be performed, as current algorithms struggle to delineate glacier areas with debris cover or other irregular spectral profiles. Although several approaches have improved upon spectral band ratio delineation of glacier areas, none have entered wide use due to complexity or computational intensity.

In this study, we present and apply a glacier mapping algorithm in Central Asia which delineates both clean glacier ice and debris-covered glacier tongues. The algorithm is built around the unique velocity and topographic characteristics of glaciers, and further leverages spectral and spatial relationship data. We found that the algorithm misclassifies between 2 and 10% of glacier areas, as compared to a $\sim$750 glacier control dataset, and can reliably classify a given Landsat scene in 3-5 minutes.

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 band ratio of Landsat 7 bands three and five or the Normalized Difference Snow Index. The main caveats of the algorithm are (1) classification errors at an individual glacier level, (2) reliance on manual intervention to separate connected glacier areas, and (3) dependence on fidelity of the input Landsat data.

1 Introduction

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 Glacier Inventory (RGI) v4.0 (Arendt et al., 2012; Pfeffer et al., 2014). Smaller-scale glacier databases are also available, such as the Chinese Glacier Inventory (CGI) v2 (Guo et al., 2015). 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; Vaughan et al., 2013). 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+) Panchromatic band/yr).

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 multi-spectral data from 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; Bhambri et al., 2011). Much of this difficulty stems from the similarities in spectral profiles of debris located on top of a glacier tongue and the surrounding landscape. 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, 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 analyze the results of our classification algorithm using a suite of 40 Landsat Thematic Mapper (TM), Enhanced Thematic Mapper+ (ETM+) and Optical Land Imager (OLI) images (1998-2013) 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 ∼1,500 km profile from the Central Pamir to the Central and Central-Eastern Tien Shan (Figure 1, Table 1).

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 an ideal test area – particularly in mapping glaciers with long and irregular debris tongues, such as the Inylchek and Tomur glaciers in the Central Tien Shan (Shangguan et al., 2015).

The wintertime climate of the study area is controlled by both the Winter Westerly Disturbances 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).

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 V4.1 (SRTM) Digital Elevation Model (DEM) (∼90 m, void-filled) was leveraged to provide elevation and slope information (Jarvis et al., 2008). The SRTM data and its derivatives were downsampled to 30 m to match the resolution of the Landsat images using bilinear resampling. The

USGS Hydrosheds river network (15 second resolution, ∼500 m) 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 Pre-processing
    (a) Velocity fields are calculated with Normalized Image Cross Correlation (Manual, can be automatized)
    (b) The Hydrosheds river network is rasterized (Manual, can be automatized)
    (c) Optional manual debris points are created (Manual, optional)
    (d) SRTM data is used to create a hillslope image (Python Script)
    (e) All input datasets are matched to a single extent and spatial resolution (30 m) (Python Script)
2. Glacier Classification Steps
    (a) Clean-ice glacier outlines are created using Landsat Bands 1,3, and 5 (Matlab Script)
    (b) 'Potential debris areas' are generated from low-slope areas (Matlab Script)
    (c) Low-elevation areas are removed (Matlab Script)
    (d) Low-velocity areas are removed (Matlab Script)
    (e) Distance-weighting metrics are used to remove areas distant from river networks or clean glacier ice (Matlab Script)
    (f) Distance-weighting metrics are used to remove areas very distant from clean glacier ice and manual seed points (Matlab Script)
    (g) 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 snow- and cloud-cover in our analysis to understand their impacts 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). We use the USGS Level 1T orthorectified Landsat scenes to ensure that the derived glacier outlines are consistent in space (Hansen and Loveland, 2012; Nuimura et al., 2014).

The algorithm uses Landsat imagery, a void-filled DEM, a velocity surface derived from image cross-correlation, and the Hydrosheds 15s river network (buffered by 200 m and converted to a raster) as the primary inputs (Steps 1(a) and 1(b)). The algorithm generates a slope image from the DEM and rectifies additional input datasets described below for processing by resampling and reprojecting each dataset to the same spatial extent and resolution (30 m to match the Landsat data) (Steps 1(d) and 1(e)). 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 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 7 Bands 1, 3, and 5 (Step 2(a)). 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 $>$25) has been used in previous research as an effective means of delineating glacier areas (e.g., Hall et al., 1987; Paul and Kääb, 2005; Hanshaw and Bookhagen, 2014), but is not effective in delineating debris-covered glacier areas (Figure 2A). In our algorithm, we use a threshold of TM3/TM5 $\geq$ 2 and TM1 $>$60 to map clean glacier ice. While these thresholds perform well over many scenes, using thresholds which are not tuned to specific scene conditions can generate large errors, particularly in shadowed areas (Paul et al., 2015). Here we choose fairly conservative threshold values to ensure that we do not remove clean glacier ice. We find that in-

creasing the TM1 threshold results in tighter classification of debris-covered glacier tongues, but also removes some areas properly classified as glacier, particularly in steep areas of the accumulation zone. Thus, we err on the side of overclassification with our delineation of clean glacier ice. The end result of this step is the spectrally-derived (clean-ice) glacier outlines, which are later integrated back into the workflow before statistical filtering (Figure 2A).

3.3 Debris-covered Ice Delineation

3.3.1 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(b)). We choose the relatively high threshold of 24°, as opposed to the 12° suggested for the Himalaya (Bolch et al., 2007), the 15° suggested for the Himachal Himalaya (Shukla et al., 2010), or the 18° suggested for the Garhwal Himalaya (Bhambri et al., 2011) to ensure that we do not prematurely remove debris-covered areas. As we use the low-slope areas as our initial maximum likely debris extent, a conservative slope threshold helps reduce errors of underclassification. Low-elevation areas (automatically defined on a scene-by-scene basis based on the average elevation of clean-ice areas minus 1750, generally below 2500-3000 m in the study area) are then masked out to decrease processing time (Step 2(c)). These thresholding steps identify areas where there is the potential for a debris-covered glacier to exist, and are performed independently of the previous, psectrally delineated, glacier outlines. Additional thresholding is then performed on this 'potential debris area' subset to identify debris-covered glacier areas (Figure 2B).

As can be seen in Figure 2B, extensive areas that are not glacier or debris-covered glacier tongue are identified in this step. However, this step greatly reduces the processing time of subsequent steps by removing pixels outside of the main glacierized areas of any scene and allowing the algorithm to work on a subset of the image from this point forward.

3.3.2 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(a)). 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 at exactly the same intervals for each Path-Row combination, there is some error in our velocity fields. However, as the velocity surface is used to remove stable ground, which is generally well-defined despite changes in glacier velocities, errors in the velocity surface do not contribute significantly to glacier classification errors, excepting on slower-moving parts of debris-covered glacier tongues. It is important to note that images must be cloud-free over glaciers and snow-free off glaciers for this step, as the presence of snow or cloud cover can disrupt the correlation process, resulting in anomalous velocity measurements. An example velocity surface is shown in Figure 2C (Step 2(d)). Red areas are removed from the 'potential debris areas', 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. This also improved our classification of slow-moving glaciers, which may not change significantly over only a single year. These velocities ranged generally from 4.5-30 m/yr across the different scenes. A single velocity threshold of 5 m/yr was used across all scenes to remove stable ground. 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. Generating a velocity field is the most computationally expensive step of the algorithm.

3.3.3 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 2D, Step 2(e)). As glaciers occur along the flowlines

of rivers, and the Hydrosheds river network generally delineates flowlines nearly 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 the clean-ice outlines generated in Step 2(a), as well as any manual seed points provided (Step 2(f)). As debris tongues must occur in proximity to either glacier areas or the centerlines of valleys, these two steps are effective in removing overclassified areas (Figure 2D). The spatial weighting performed here differs from that proposed by Paul et al. (2004) in that it uses a measure of geodesic distance from given seed points, as opposed to maintaining entire polygons which are connected to clean-ice areas. This difference helps remove non-glacier areas that are distant from clean ice, but still connected by at least a single pixel to clean-ice areas. 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 irregular or cirque glaciers, the addition of manual seed points has been observed to increase the efficacy of the algorithm. In processing the Landsat imagery presented here, we have not used additional manual seed points.

The spatial weighting step is essential for removing pixels spatially distant from any clean-ice 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 glacierized area(s) of a Landsat scene, as can be seen in Figure 2E.

3.3.4 Statistical Filtering

Once the spatial weighting steps are completed, 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(g)). 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 necessary for filling holes and reconnecting separated glacier areas that result from the initial threshold-based filtering steps. For example, slow-moving pixels in the middle of a debris-covered glacier tongue that were removed based on velocity filtering are often restored by the statistical filtering (Figure 3). The improved classification of debris areas between the clean-ice and final

algorithm outputs can clearly be seen in Figure 3.

3.4 Creation of Manual Control Datasets

Manual control datasets encompassing ∼750 glaciers (∼11,000 km^2) were created to test the efficacy of the glacier mapping algorithm. These datasets were digitized from 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 (∼2000 and ∼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 (<0.5 km^2 to 500+ km^2), as well as both clean-ice 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). Figure 4 shows the size-class distribution of the manual control dataset, with logarithmic area scaling.

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

Over the eight Landsat footprints used in this study, we map ∼44,000 km^2 of glaciers over two distinct time slices. Several additional time periods were mapped, but not included in the statistical analysis presented in this manuscript.

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 spectral outlines, which only classify the glacier areas via commonly used spectral subsetting (using TM1, TM3, and TM5, produced in Step 2(b)), the manual control datasets, and the CGI v2. Figure 5A shows the bulk elevation distributions across 215 glaciers for each dataset in 10 m 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. This bias also stems from misclassified areas in shadows, particularly in north-facing glaciers. 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 and generally follow the manually digitized data. This threshold represents the general transition from debris-covered glaciers to clean glacier ice in the study area. Our algorithm output is also well-matched with the CGI v2, except at very high elevations where it overclassifies some areas as compared to the CGI v2.

In order to examine inherent bias throughout the algorithm classification, under- and over-classified areas were examined for a subset of the control dataset. To determine areas of overclassification (underclassification), the manually (algorithm) generated dataset was subtracted from the algorithm (manual) dataset, leaving only pixels that were overclassified (underclassified). Figure 5B 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-glacierized areas within glacier accumulation zones. Importantly, the algorithm underclassifies a much smaller number of pixels, generally corresponding to areas below 4000 m, where debris tongues are dominant. The majority of these pixels are along the edges of debris-covered glacier 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 is 2% and 10% respectively, which represents a significant improvement from a pure spectral delineation approach.

To investigate sampling bias in our analysis, we used 465 GLIMS glacier identification numbers (centroids, point features) that 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 5C, 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 (Kolmogorov–Smirnov test passes at 99% confidence interval).

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 a set of X/Y pairs for each component vertex, which were then matched to the closest vertex in the spectral and final algorithm result's polygons, respectively (Figure 6).

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. Both datasets were normalized by their whole-dataset maximum distances.

5 Discussion

5.1 Comparison with Previous Glacier Mapping Algorithms

Several authors have presented alternative debris-covered glacier classification methods and schemes using thermal and spectral data (Taschner and Ranzi, 2002), topographic and neighborhood analysis (Paul et al., 2004), clustering of optical and thermal data (Bolch et al., 2007), maximum likelihood classification (Shukla et al., 2010), slope and curvature clustering combined with thermal data (Bhambri et al., 2011), decision tree classification and texture analysis, (Racoviteanu and Williams, 2012) and object-based classifications (Rastner et al., 2014). While all of these methods present improvements over basic clean-ice delineation as proposed by Hall et al. (1987), they each have shortcomings that limit their range of use. Table 2 shows a comparison of these different methods

alongside the algorithm presented in this study.

Our study improves on previous work in three main ways: (1) reduced computational intensity, (2) greater diversity of study area, and (3) increased temporal range of our dataset. The methods proposed in this study, excepting the generation of a velocity field, require very little processing power. Once initial input datasets (velocity surface, rasterized river network) have been created, a Landsat scene can be processed in 3-5 minutes (Ubuntu 14.04, 8 cores (3.6GhZ), 16 GB RAM). When this is compared with the training dataset creation, computationally expensive classification schemes, and neighborhood analyses employed by other studies, there is a clear improvement in efficiency. Secondly, we analyze a significantly larger glacier area than any of the previous studies, which has helped us generalize our algorithm and methods to a wide range of topographic and landcover settings. Finally, we process a multi-year dataset, encompassing 40 Landsat scenes with varying landcover and meteorological settings. This has allowed us to further generalize our algorithm to be effective beyond a single scene or small set of scenes, and to remain effective across a wide spatial and temporal range. The time-dynamic aspect of our algorithm can also complement time-static wide-area datasets, such as the RGI v4.0, the CGI v2, and the forthcoming GAMDAM datasets (Arendt et al., 2012; Guo et al., 2015; Nuimura et al., 2014). While these datasets may provide higher-quality manually digitized outlines for specific glaciers, they only provide a single snapshot in time, and are limited to a specific area of coverage.

5.2 Additional Tested 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 terrain on and off glaciers, 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.

5.3 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. However, because the algorithm operates primarily on 'potential debris areas', any clean ice classification scheme could be used. For example, in other study regions, or for different satellite sensors, other schemes, such as the Normalized Difference Snow Index (Dozier, 1989), may outperform clean ice classification as implemented in this study.

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 7, the algorithm compares well with both the control dataset and the CGI v2 – a high-fidelity, manually edited, dataset – across a range of glacier types (Step 2(a)) (Guo et al., 2015). However, the algorithm outlines do not perfectly align with either dataset. In Figure 7, a tendency to remove pixels along the edge of debris-covered glacier 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 CGI v2, 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 are not always covered by permanent ice. These two types of classification bias are easily rectified with minimal manual intervention. Some bias between the manual or algorithm datasets and the CGI v2 can also be attributed to the difference in time; while the manual and

algorithm datasets share an image date, the CGI v2 was digitized on top of multiple images that may not match up perfectly in time with our datasets. Despite these misclassified areas, the raw algorithm output effectively identifies the furthest reaches of the glacier tongues in most cases, as can be seen in three long debris tongues shown in Figure 7.

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 5A). As the method can be easily modified to fit the topographic and glacier setting of any region, it is a powerful tool for analysing glacier changes over large scales for the period of Landsat TM, ETM+ and OLI coverage. While the algorithm has yet to be applied to large and slow-moving debris-covered glaciers in the Himalaya, a wide range of glacier size classes, speeds, and topographic settings are well classified by the algorithm. For example, even small glacier changes are captured by the algorithm, as can be seen in Figure 8.

Figure 8 also illustrates some potential errors in 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 easily removed during manual inspection of results.

The second use case for the algorithm is as a substitute for simple spectral ratios. Manual digitization of glacier tongues is time consuming, particularly in regions with numerous debris-covered glaciers. Our algorithm provides a robust baseline set of glacier outlines that can be corrected manually, with minimal extra processing time. As generating the input velocity surfaces can take longer than processing glacier outlines from dozens of Landsat scenes, efficiencies are gained when Landsat scenes are processed in bulk. The algorithm as presented in this manuscript takes $\sim$3-5 minutes of actual processing time once the base datasets have been created. For a single Path-Row combination, the time to set up the input datasets (velocity surface, manual debris points) is $\sim$4 hours. Once the initial setup has been completed for a given Path-Row combination, any number of Landsat scenes can be processed very quickly.

Although the algorithm represents a step forward in semi-automated 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- or snow-cover are less likely to be well defined. This creates a data limitation, as many glacierized areas are subject to frequent cloud- and snow-cover, and thus have a limited number of potentially useful Landsat images for the purpose of this algorithm.

6 Conclusions

This study presents an enhanced glacier classification methodology based on the spectral, topographic, and spatial characteristics of glaciers. We present a new method of (semi-) automated glacier classification, which is built upon, but unique from, the work of previous authors. Although it does not completely solve the difficulties associated with debris-covered glaciers, it can effectively and rapidly characterize glaciers over a wide area. Following an initial delineation of clean glacier ice, 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 manually digitized control dataset and the high-fidelity CGI v2, our algorithm remains robust across the wide range of glacier sizes and types found in Central Asia. The algorithm developed here is applicable to a wide range of glacierized regions, particularly in those regions where debris-covered glaciers are dominant, and extensive manual digitization of glacier areas has previously been required. The raw algorithm output is usable for rough statistical queries on glacier area, hypsometry, slope, and aspect; however, manual inspection of algorithm output is necessary before using the generated glacier outlines for more in-depth area change or mass balance studies.

Acknowledgements. Part of 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, Wanqin Guo, and one anonymous reviewer for their detailed and helpful reviews, as well as Tobias Bolch for his contribution to the development of the manuscript.

References

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.

Bhambri, R., Bolch, T., and Chaujar, R.: Mapping of debris-covered glaciers in the Garhwal Himalayas using ASTER DEMs and thermal data, International Journal of Remote Sensing, 32, 8095–8119, 2011.

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.

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.

Dozier, J.: Spectral signature of alpine snow cover from the Landsat Thematic Mapper, Remote sensing of Environment, 28, 9–22, 1989.

Guo, W., Xu, J., Liu, S., Shangguan, D., Yao, X., Wei, J., Bao, W., Yu, P., Liu, Q., and Jiang, Z.: The second Chinese glacier inventory: data, methods and results, Journal of Glaciology, 226, 957–969, doi: 10.3189/2015JoG14J209, 2015.

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.

Hansen, M. C. and Loveland, T. R.: A review of large area monitoring of land cover change using Landsat data, Remote sensing of Environment, 122, 66–74, 2012.

Hanshaw, M. and Bookhagen, B.: Glacial areas, lake areas, and snow lines from 1975 to 2012: status of the Cordillera Vilcanota, including the Quelccaya Ice Cap, northern central Andes, Peru, The Cryosphere, 8, 359, 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.

Jarvis, A., Reuter, H. I., Nelson, A., Guevara, E., et al.: Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database (http://srtm.csi.cgiar.org), 2008.

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.

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.

Nuimura, T., Sakai, A., Taniguchi, K., Nagai, H., Lamsal, D., Tsutaki, S., Kozawa, A., Hoshina, Y., Takenaka, S., Omiya, S., et al.: The GAMDAM Glacier Inventory: a quality controlled inventory of Asian glaciers, The Cryosphere Discussions, 8, 2799–2829, 2014.

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. and Kääb, A.: Perspectives on the production of a glacier inventory from multispectral satellite data in Arctic Canada: Cumberland Peninsula, Baffin Island, Annals of Glaciology, 42, 59–66, 2005.

Paul, F., Kääb, 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., Le Bris, R., et al.: On the accuracy of glacier outlines derived from remote-sensing data, Annals of Glaciology, 54, 171–182, 2013.

Paul, F., Bolch, T., Kääb, A., Nagler, T., Nuth, C., Scharrer, K., Shepherd, A., Strozzi, T., Ticconi, F., Bhambri, R., et al.: The glaciers climate change initiative: Methods for creating glacier area, elevation change and velocity products, Remote Sensing of Environment, 162, 408–426, 2015.

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.

Racoviteanu, A. and Williams, M. W.: Decision tree and texture analysis for mapping debris-covered glaciers in the Kangchenjunga area, Eastern Himalaya, Remote Sensing, 4, 3078–3109, 2012.

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.

Rastner, P., Bolch, T., Notarnicola, C., and Paul, F.: A comparison of pixel-and object-based glacier classification with optical satellite images, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7, 853–862, 2014.

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.

Shangguan, D., Bolch, T., Ding, Y., Kröhnert, M., Pieczonka, T., Wetzel, H., and Liu, S.: Mass changes of Southern and Northern Inylchek Glacier, Central Tian Shan, Kyrgyzstan, during 1975 and 2007 derived from remote sensing data, The Cryosphere, 9, 703–717, 2015.

Shukla, A., Arora, M., and Gupta, R.: Synergistic approach for mapping debris-covered glaciers using optical–thermal remote sensing data with inputs from geomorphometric parameters, Remote Sensing of Environment, 114, 1378–1387, 2010.

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.

Taschner, S. and Ranzi, R.: Comparing the opportunities of Landsat-TM and Aster data for monitoring a debris covered glacier in the Italian Alps within the GLIMS project, in: Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International, vol. 2, pp. 1044–1046, IEEE, 2002.

Vaughan, D., Comiso, J., Allison, I., Carrasco, J., Kaser, G., Kwok, R., Mote, P., Murray, T., Paul, F., Ren, J., Rignot, E., Solomina, O., Steffen, K., and Zhang, T.: Observations: Cryosphere. In: Climate Change 2013: The Physical Science Basis., Contribution of Working Group I to the Fifth Assessment Report of the IPCC., 2013.

Table 1. Data table listing Landsat acquisition dates used in this study. Organized by WRS2 Path/Row combinations. Bold dates indicate images used for velocity profiles.

	144/030	145/030	147/031
Number of Images	5	4	12
Date Range of Images	2002-2013	1998-2013	2000-2013
LT5 Acquisition Dates	Sep 27, 1998	Sep 2, 1998	Aug 19, 2011
		Sep 6, 2011	Oct 2, 1998
		Aug 10, 2007	Sep 6, 2006
			Aug 24, 2007
			Oct 3, 2010
			Aug 3, 2011
LE7 Acquisition Dates	**Sep 14, 2002**		Sep 14, 2000
	Aug 7, 2000		**Oct 5, 2002**
			Aug 18, 2002
LC8 Acquisition Dates	**Oct 22, 2013**	**Sep 27, 2013**	**Sep 25, 2013**
	Sep 4, 2013		Sep 9, 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	5	3	4
Date Range of Images	2002-2013	1999-2013	1998-2013
LT5 Acquisition Dates	Aug 22, 1998	Sep 7, 2007	Oct 23, 1998
	Sep 11, 2011		
LE7 Acquisition Dates	**Jul 24, 2002**	**Sep 9, 1999**	**Aug 20, 2001**
	Sep 18, 1999		
LC8 Acquisition Dates	**Jul 30, 2013**	**Oct 9, 2013**	**Jun 10, 2013**
Projection	WGS 1984 44N	WGS 1984 43N	WGS 1984 43N
Comments			
	151/033	153/033	
Number of Images	4	3	
Date Range of Images	1998-2013	1998-2013	
LT5 Acquisition Dates	Sep 28, 1998	Sep 26, 1998	
LE7 Capture Dates	Aug 24, 2000	**Sep 29, 2002**	
	Sep 28, 2001		
LC8 Acquisition Dates	**Oct 7, 2013**	**Oct 5, 2013**	
Projection	WGS 1984 43N	WGS 1984 42N	
Comments		Towards Pamir Knot	

Table 2. Comparison of methods between previous debris-covered glacier mapping studies.

Method	Short Description	Data Inputs	Processing Intensive Steps	Area Covered in Study	Reported Accuracy
Taschner and Ranzi (2002)	Clean-ice detection using Landsat, coupled with ASTER thermal data	Landsat, ASTER	Data resampling, pixel clustering	5.58 km^2, Italian Alps	Not Reported
Paul et al. (2004)	Clean-ice detection using Landsat, coupled topographic analysis and neighborhood analysis	Landsat, ASTER-DEM	Image Polygon Growing neighborhood analysis	23 km^2, Swiss Alps	21% of debris misclassified
Bolch et al. (2007)	A set of training areas based on spectral and topographic information is used to determine classification thresholds	ASTER, ASTER-DEM	Creation and tuning of training dataset	Not reported, Mt. Everest Region	5% total area misclassified
Shukla et al. (2010)	Multiple landcover types mapped using spectral and thermal imagery combined with a DEM	ASTER, AWiFS, DEM	Data conversion and registration, solar illumination analysis, training dataset creation, Maximum Likelihood Classifier	200 km^2, Samudra Tapu glacier, Himachal Pradesh, India	8-14% debris misclassified
Bhambri et al. (2011)	Combination of slope and curvature data analyzed with a clustering algorithm coupled with thermal band thresholding	ASTER, DEM, Landsat, IRS PAN	Manual decisions on glacier slope and curvature clusters	232 km^2, Gongotri Glacier, Garhwal Himalaya, India	0.5-11% debris misclassified
Racoviteanu and Williams (2012)	(1) Decision tree classification with ASTER and topographic data, and (2) texture analysis exploiting surface roughness	ASTER, DEM, Quickbird, Worldview2	Training dataset creation, decision tree set-up, principal component analysis	576.4 km^2, Sikkim Himalaya, NE India	(1) 25%, (2) 31% debris misclassified
Rastner et al. (2014)	Comparison of object- and pixel-based methods of glacier mapping. Both methods use spectral and topographic information as inputs	ASTER, Landsat, DEM	Manual threshold definitions, segmentation processing, iterative thresholding	Not reported, three distinct test regions	11.5% (object-based) and 23.4% (pixel-based) misclassified areas for Himalaya region
This Study	Clean-ice detection coupled with topographic, velocity, and distance weighting thresholds	Landsat, SRTM DEM, River Network	Velocity field calculation, optional debris seed point selection	$\sim$44,000 km^2, Pamir-Tien Shan	2-10% total area misclassified

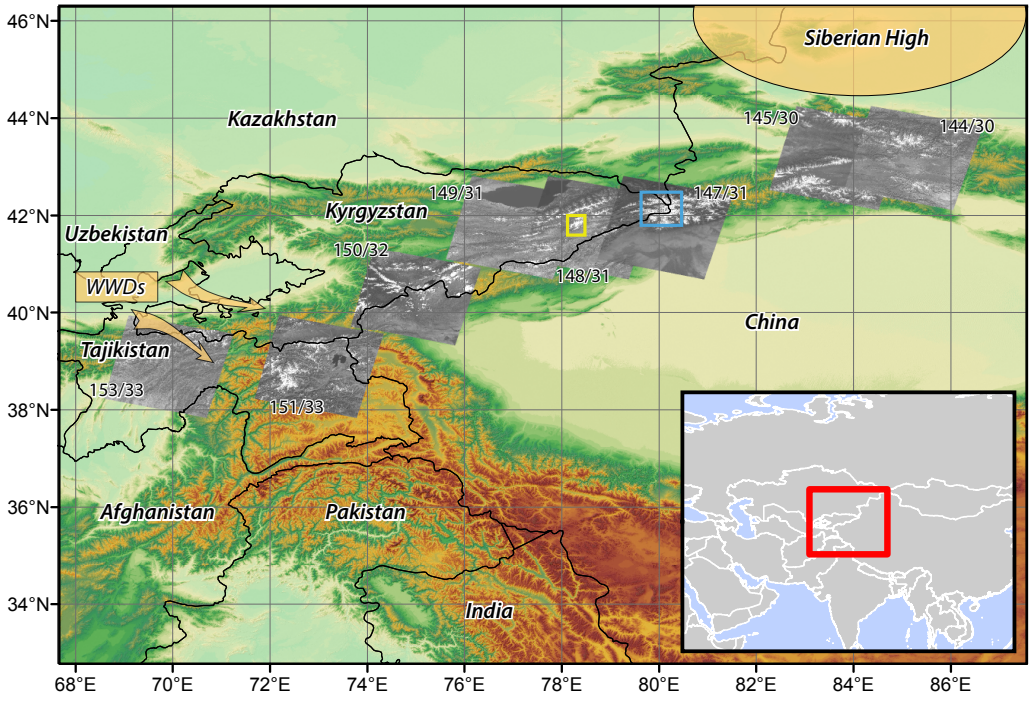

Fig. 1. Greater study area of the Tien Shan, showing SRTM v4.1 topography (Jarvis et al., 2008) and location of eight Landsat image footprints (grayscale) used in the study, along with their Path/Row combinations. Blue box delineates Figures 2-3 and 7, yellow box delineates Figure 8. Winter Westerly Disturbances and Siberian High highlighted in orange.

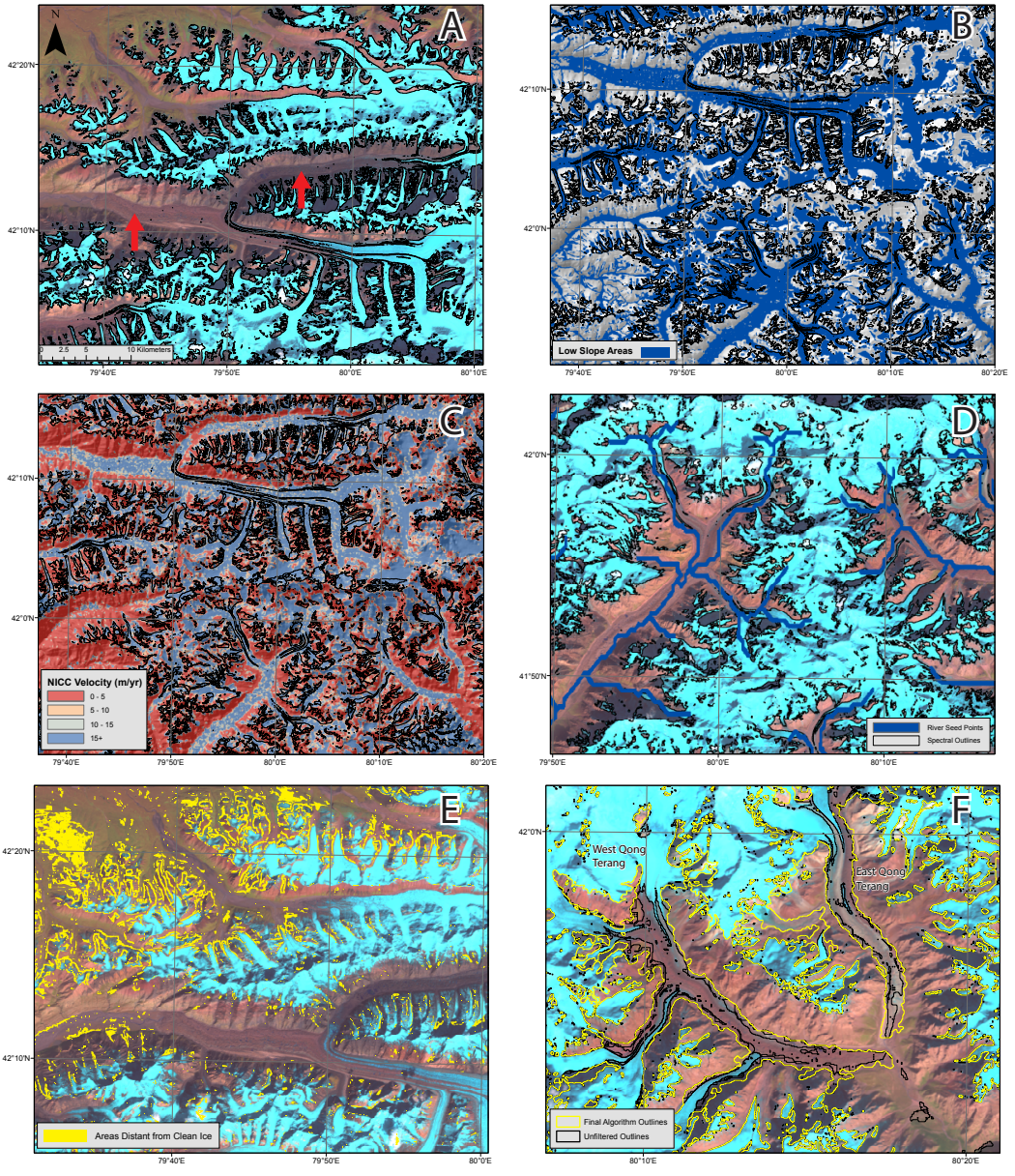

Fig. 2. (A) Characteristic example of a debris-covered glacier tongue (Inylchek Glacier). Spectrally-delineated glacier outlines (black), over Landsat bands B7/B5/B3 (R/G/B), from image LC81470312013268LGN00, with poorly mapped debris-covered tongues (red arrows). (B) Blue areas show 'Potential Debris Areas', as delineated by slopes between 1-24 degrees, with elevations below ∼2500 m removed, SRTM hillshade underneath, clean-ice outlines overlain in black. (C) Example of a glacier velocity surface, generated using Normalized Image Cross Correlation (NICC). Areas in red are slow-moving areas and represent stable ground, clean-ice outlines overlain in black. (D) Example of distance-weighting seed areas used to remove pixels from the 'potential debris areas' which are distant from either a river valley or classified glacier ice. Rivers in blue, clean-ice outlines overlain in black. (E) Areas removed by second distance weighting step (yellow). (F) Impacts of statistical filtering on glacier outlines, with areas in black removed during the filtering process. East and West Qong Terang Glaciers.

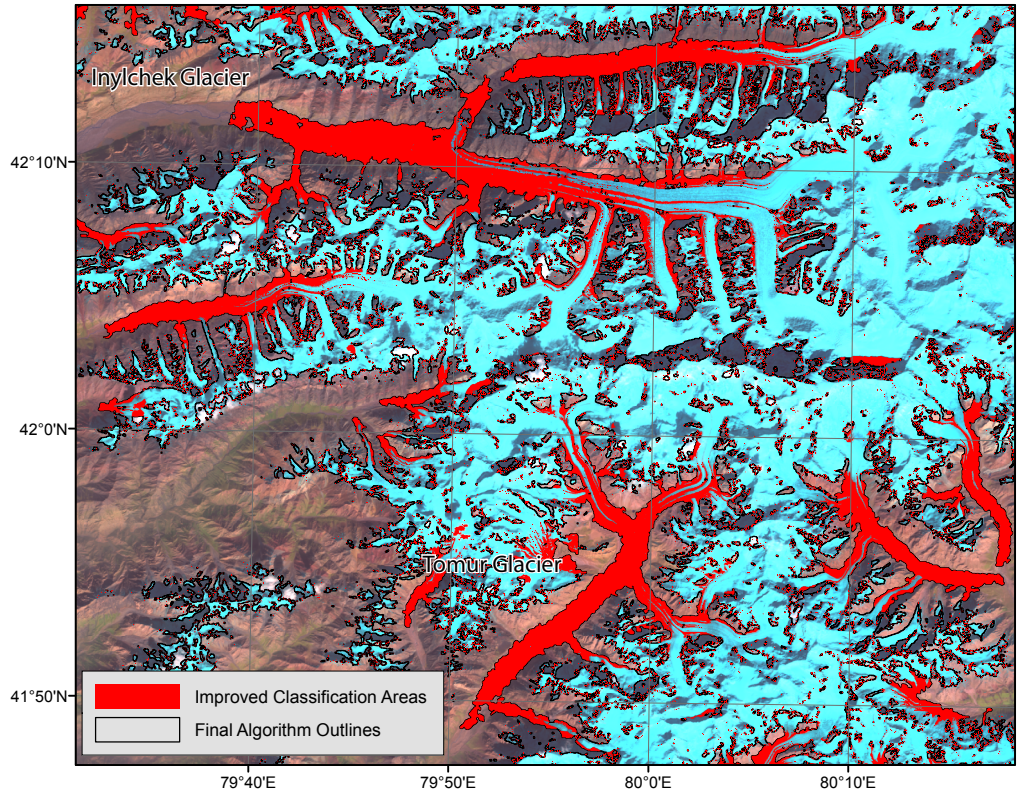

Fig. 3. Final algorithm outlines (black) with areas classified in addition to the clean-ice delineation in red. Landsat OLI image captured Sept 25, 2013 as background.

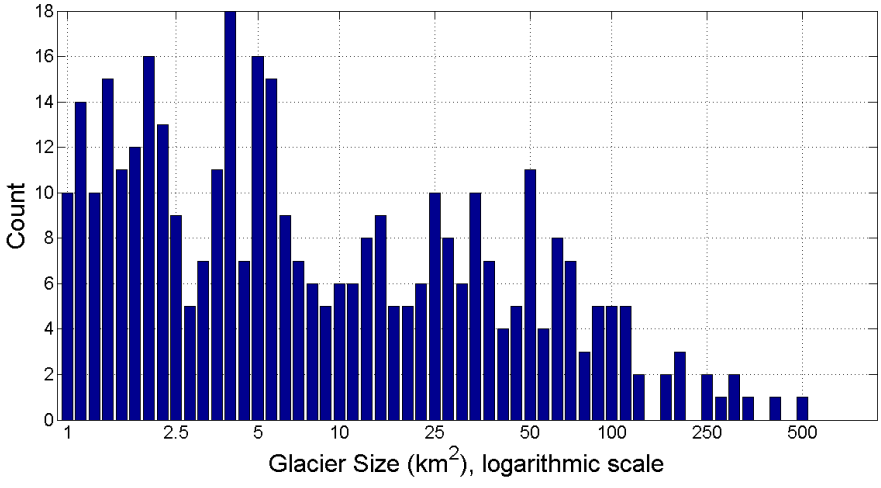

Fig. 4. Glacier size class distribution (n=750) for the manual control dataset. Note the logarithmic x-axis to account for a wide range of glacier sizes.

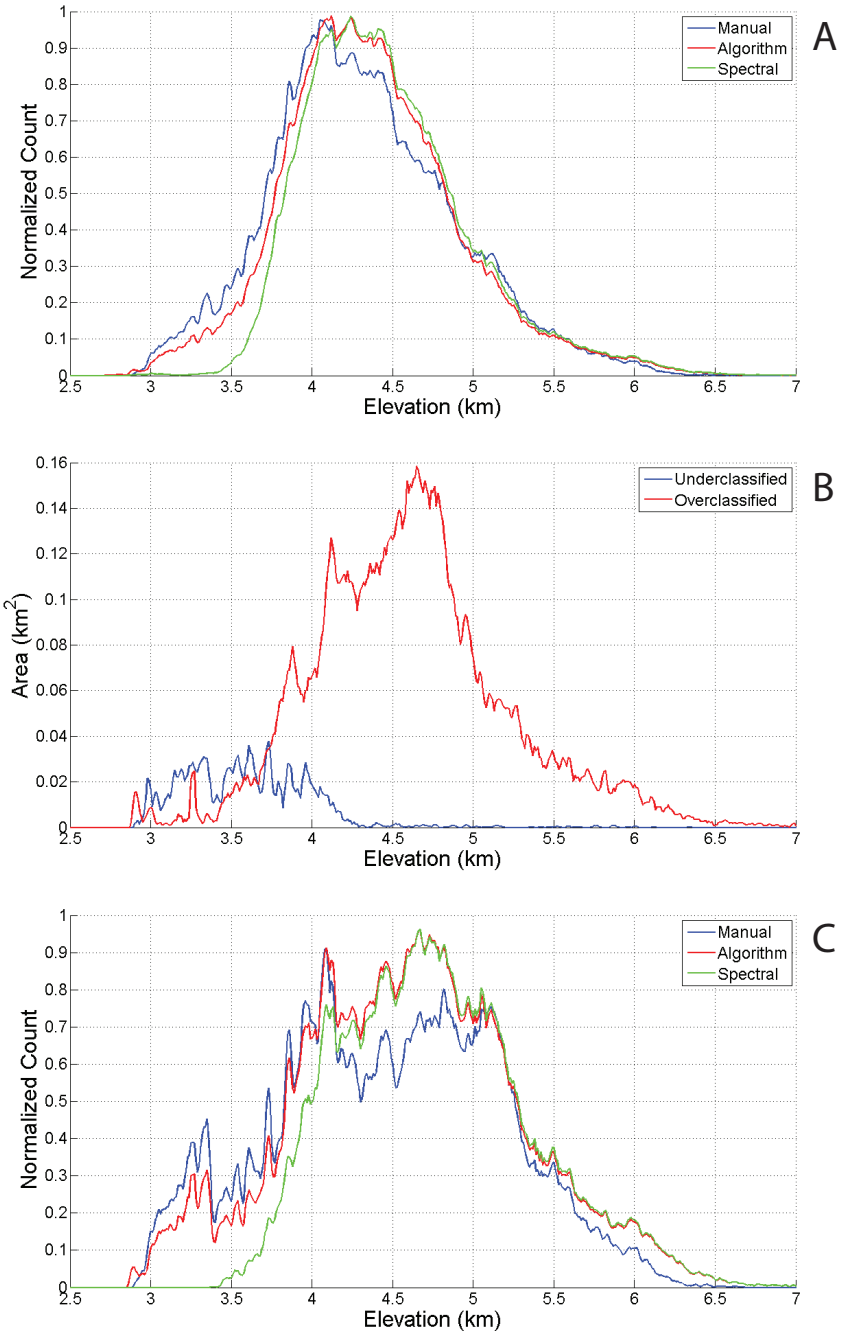

Fig. 5. (A) Bulk elevation distributions of sampled glaciers, with manual delineation (reference dataset, n=215, 4,500 km^2) in blue, algorithm-derived delineation in red, spectral delineation in green, and CGI v2 in black. Values have been normalized to maximum probability. (B) Elevation distributions of over- and under-classified glacier areas, as compared to a manual control dataset (n=75, 330 km^2). 5.5% is overclassified, and 0.8% is underclassified. (C) Averaged elevation differences for a random sample of glaciers overlapping a manual control dataset (n=100, 100 km^2).

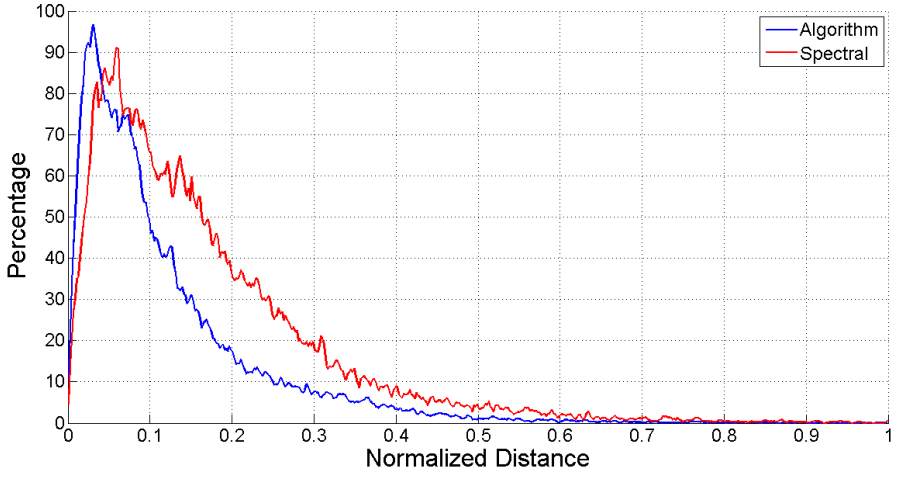

Fig. 6. Vertex distance distributions for algorithm (blue) and spectral (red) vertices, as compared to a manual control dataset, normalized to the maximum distance.

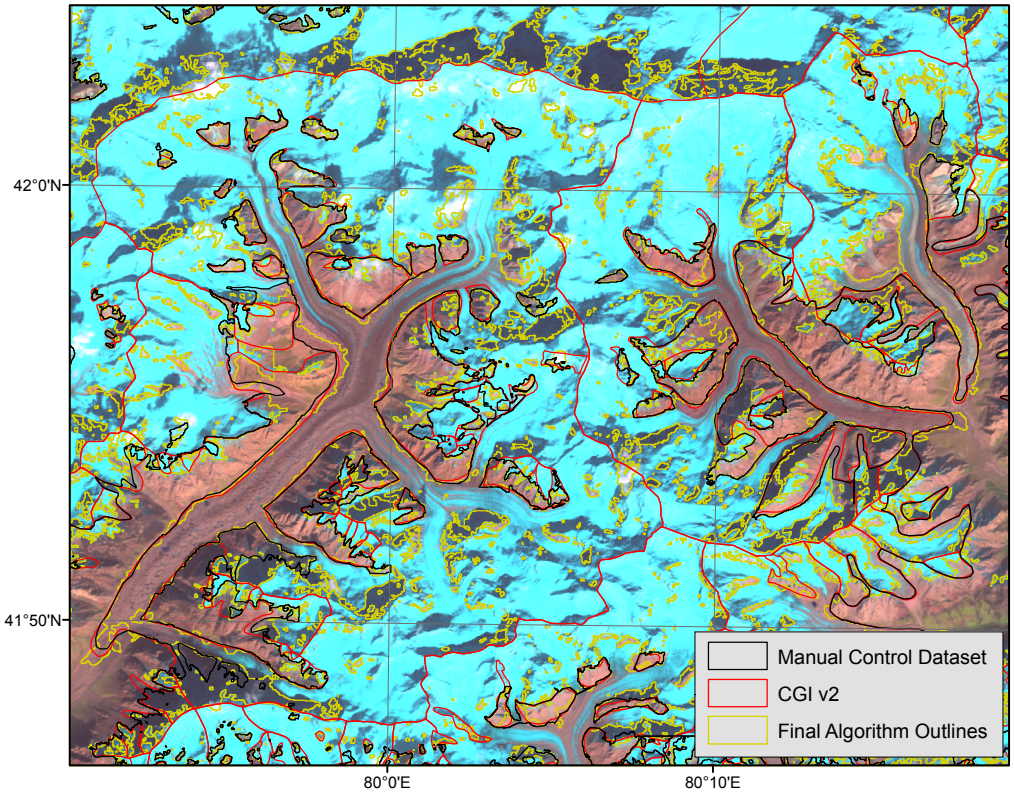

Fig. 7. Algorithm outlines (yellow) compared to the control dataset (black) and the CGI v2 (red). Illustrates high fidelity in overall debris-tongue length between the three datasets, although the algorithm outlines exhibit noise along the edges of debris tongues.

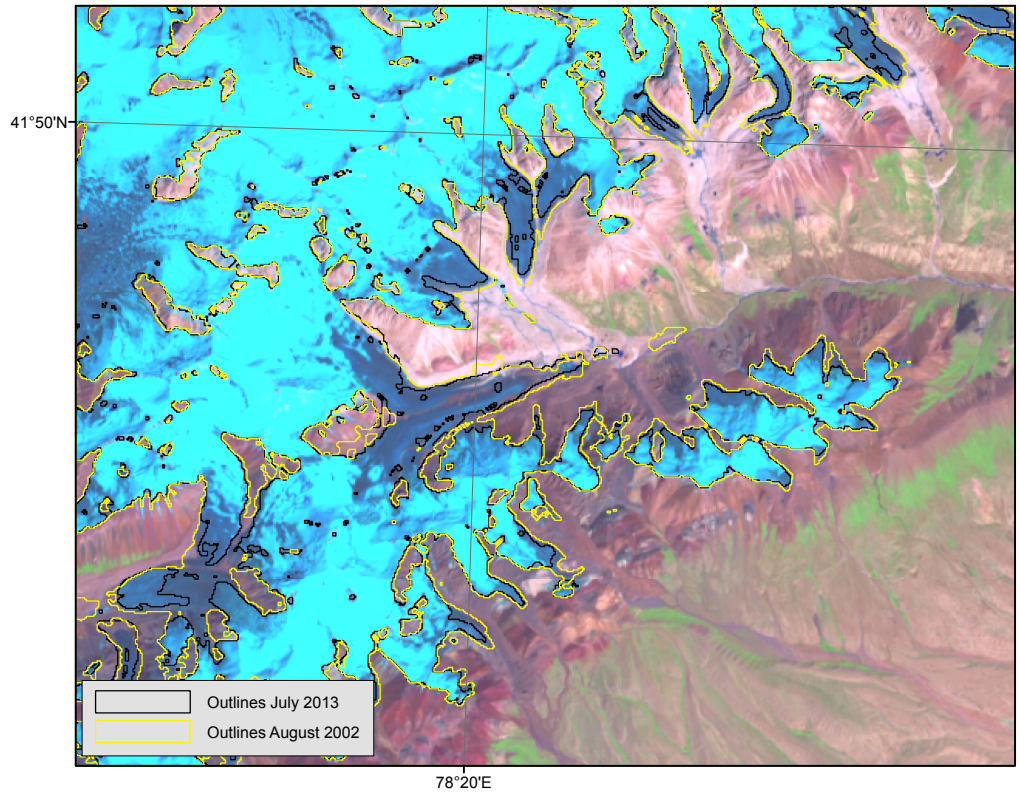

Fig. 8. Algorithm outlines for July 2013 (black) and algorithm outlines for August 2002 (yellow), showing small retreats in glacier areas, particularly at the debris tongues. Vicinity of the Akshiirak glacierized massif, central Tien Shan.