Topography and Climate in the Upper Indus Basin: Mapping Elevation-Snow Cover Relationships

T. Smith*[1], A. Rheinwalt[1], and B. Bookhagen[1]

[1]Institute of Geosciences, Universität Potsdam, Germany

Abstract

The Upper Indus Basin (UIB), which covers a wide range of climatic and topographic settings, provides an ideal venue to explore the relationship between climate and topography. While the distribution of snow and glaciers is spatially and temporally heterogeneous, there exist regions with similar elevation-snow relationships. In this work, we construct elevation-binned snow-cover statistics to analyze 3,415 watersheds and 7,357 glaciers in the UIB region. We group both glaciers and watersheds using a hierarchical clustering approach and find that (1) watershed clusters mirror large-scale moisture transport patterns and (2) are highly dependent on median watershed elevation. (3) Glacier clusters are spatially heterogeneous and are less strongly controlled by elevation, but rather by local topographic parameters that modify solar insolation. Our clustering approach allows us to clearly define self-similar snow-topographic regions. Eastern watersheds in the UIB show a steep snow cover-elevation relationship whereas watersheds in the central and western UIB have moderately sloped relationships, but cluster in distinct groups. We highlight this snow-cover-topographic transition zone and argue that these watersheds have different hydrologic responses than other regions. Our hierarchical clustering approach provides a potential new framework to use in defining climatic zones in the cyrosphere based on empirical data.

Keywords

Snow-Cover, Hierarchical Clustering, Glaciers, Upper Indus Basin

Highlights

1. Watersheds (24 - 1,410 km^2) and glaciers have distinctive elevation-snow cover relationships in the UIB.
2. Elevation-binned snow-cover statistics can be used to group self-similar regions.
3. Clusters of glaciers and watersheds provide a novel way of defining empirical topographic-climatic zones.

1 Introduction

The Upper Indus Basin (UIB) is a key source of water for millions of people across India, Pakistan, China, and Afghanistan (Vaughan et al., 2013; W. W. Immerzeel et al., 2010; Bolch et al., 2012). Water stored in snow and ice is responsible for more than 50% of the downstream yearly discharge in the Indus; seasonal snow-water contributions to the water budget are higher for many sub-catchments (Bookhagen & Burbank, 2010; W. W. Immerzeel et al., 2010; Tahir et al., 2011; Huss et al., 2017). The UIB is also highly dependent on the consistency of snowfall and snowmelt; there is a lack of reservoir capacity to buffer seasonal water shortages (Barnett et al., 2005; W. Immerzeel & Bierkens, 2012; Smith et al., 2017; Athar et al., 2019), especially as regional glaciers shrink (Gardelle et al., 2012; Kapnick et al., 2014; Bookhagen, 2016; Treichler et al., 2019; Shean et al., 2020; Farinotti et al., 2020). Changes in high-elevation snow and snowmelt will also be felt downstream in natural, agricultural, and urban settings (e.g., W. W. Immerzeel et al., 2010; Lutz et al., 2016; Bookhagen, 2017).

The significant snow-water resources of the Indus are not evenly distributed – there are strong topographic and structural controls on where, when, and how much precipitation is deposited as snow and rain (Cannon et al., 2015; Smith & Bookhagen, 2019). Topography also plays a role in long-term changes in snow-water storage (Smith & Bookhagen, 2018, 2020a; Huning & AghaKouchak, 2020), snowmelt (Smith et al., 2017; Lund et al., 2019), and regional glacier stability (Kapnick et al., 2014; Treichler et al., 2019; Shean et al., 2020; Farinotti et al., 2020; Abdullah et al., 2020); some glaciers in the UIB are growing in opposition to general regional trends (Hewitt, 2005; Treichler et al., 2019; Bolch et al., 2019; Shean et al., 2020). While the water budget of the UIB is increasingly well understood, high-elevation snow and glacier dynamics remain uncertain (W. Immerzeel et al., 2015; Bolch et al., 2012) due to a distinct lack of in-situ measurements at high elevations. Studies using empirical (Smith & Bookhagen, 2018; Kääb et al., 2015; Shean et al., 2020) and modeling (Palazzi et al., 2013; Maussion et al., 2014) data have attempted to estimate snow and glacier dynamics, but are limited by the low spatial resolutions of key datasets. In particular, modeling approaches depend strongly on the spatial resolution of the terrain (e.g., Cannon et al., 2017; Norris et al., 2015, 2020) and climate (Yoon et al., 2019) data used to force the models.

Precise hydro-meteorologic in-situ measurements are difficult to obtain over complex terrain and often face high uncertainties when they are used to characterize large regions (Liu et al., 2018; Pellicciotti et al., 2012; Fowler & Archer, 2006; Baudouin et al., 2020). Remotely-sensed data have the advantage of providing spatially extensive measurements over long time spans. There remain, however, uncertainties in these data as well – in particular, cloud cover limits the utility of optical data in many seasons, and the spatial resolutions of other climate data remain low (Smith & Bookhagen, 2018). It is thus important to emphasize that insights obtained from remote-sensing data should be validated with further local-scale and in-situ studies.

Monitoring regional cryospheric trends – via in-situ, modeled, or remotely-sensed data – thus often requires gathering sparse data into self-similar groups. In particular, the mass balance signatures of single glaciers are typically noisy and uncertain; grouping glaciers before regional analysis is essential for removing outliers, extending the length of measured time series, and increasing the number of useful observations. Previous work and approaches have attempted to group and classify regions of High Mountain Asia into distinctive glacio-climatic domains (e.g., Bolch et al., 2019; Shean et al., 2020; Scherler et al., 2011; Kääb et al., 2015; Smith et al., 2017), but the complex topographic and climate setting of the UIB makes defining self-similar regions difficult.

In this study, we use high-resolution topography (1 arc second Shuttle Radar Topographic Mission (SRTM) data (JPL, 2020)) and snow-cover data to examine topographic controls on the distribution and character of snow and glaciers in the UIB region. Using a consistent binning approach, we first quantify the relationship between elevation and snow-cover for a number of watersheds and glaciers; we hypothesize that these relationships are diagnostic of climatic and topographic zones. We further test whether the elevation-binned data can be leveraged in a hierarchical clustering framework to coherently group and aggregate diverse watersheds and glaciers in the UIB region.

1.1 Study Area

The UIB (area: ∼425,000 km^2 (Lutz et al., 2016)) covers a wide range of topographic and cryospheric settings from the warmer, low-elevation foreland, across the Karakoram and into the dry Tibetan interior (Figure 1). The lower reaches of the Indus basin in the northwestern Himalaya are located at the end of the monsoonal conveyor belt stretching from the Bay of Bengal to the northwest and receive moderate (<1 m/yr) amounts of monsoonal moisture during the summer season (e.g., Bookhagen & Burbank, 2006, 2010; Malik et al., 2016) (cf. Figure 1). In contrast, the UIB and adjacent regions are strongly influenced by Westerly Disturbances (e.g., Cannon et al., 2015; Dimri et al., 2015) leading to significant snowcover (>80%) and snow-water storage (>75 mm) at high elevations (e.g., Wulf et al., 2016; Smith & Bookhagen, 2019, 2020a; Fowler & Archer, 2006; Norris et al., 2015; Bonekamp et al., 2019). The hydrologic budget of the UIB is dominated by snowmelt, but rainfall during the monsoon season is an important factor in lower-elevation areas (e.g., Bookhagen & Burbank, 2010; Smith & Bookhagen, 2018; Tahir et al., 2011; W. W. Immerzeel et al., 2009; Huss et al., 2017; Wulf et al., 2016).

While snow-covered area (SCA) is often used as a first-order proxy of snow-water storage, the relationship between SCA and snow-water storage is non-linear (Supplemental Figure 1) and has a complex spatio-temporal pattern. In the more westerly reaches of the UIB, SCA and snow-water equivalent (SWE) generally mirror topography – the higher reaches of the UIB have more snow. In the eastern areas, however, the drier Tibetan interior has much lower SCA and SWE at similarly high elevations. This reflects the topographic shielding of Westerly Disturbances that mostly impact the western Pamir and UIB regions (Dimri et al., 2015; Cannon et al., 2015). The majority of snow-cover and snow-water storage is found in the high north and north-central reaches of the UIB (Figure 1). Throughout the UIB, there is a marked disconnect between SCA and SWE, where areas with high SWE volumes are not always fully snow-covered (Figure 1, Supplemental Figure 1).

2 Data and Methods

2.1 Topographic and Glacier Data

In this study, we rely on the reprocessed NASA SRTM Digital Elevation Model (DEM) with a nominal resolution of 1 arc second (∼30 m) for topographic information (JPL, 2020). These data have been shown to be reliable elevation indicators in steep terrain (e.g., Purinton & Bookhagen, 2018). To derive watershed units, we first hydrologically correct the DEM by filling all pits, then derive flow direction, calculate flow accumulation, and extract watersheds with stream orders between 3 and 5 using standard GIS approaches (Schwanghart & Scherler, 2014). Our analysis of elevation-snow relationships relies on 3,415 fourth-order watersheds with areas between 24 and 1,410 km^2. Using these watersheds, we subset topographic (elevation, slope, aspect) and climatic (snow-covered area, normalized-difference snow index (NDSI)) data for further analysis. We rely on the RGI v6 (Arendt et al., 2015) for glacier outlines in

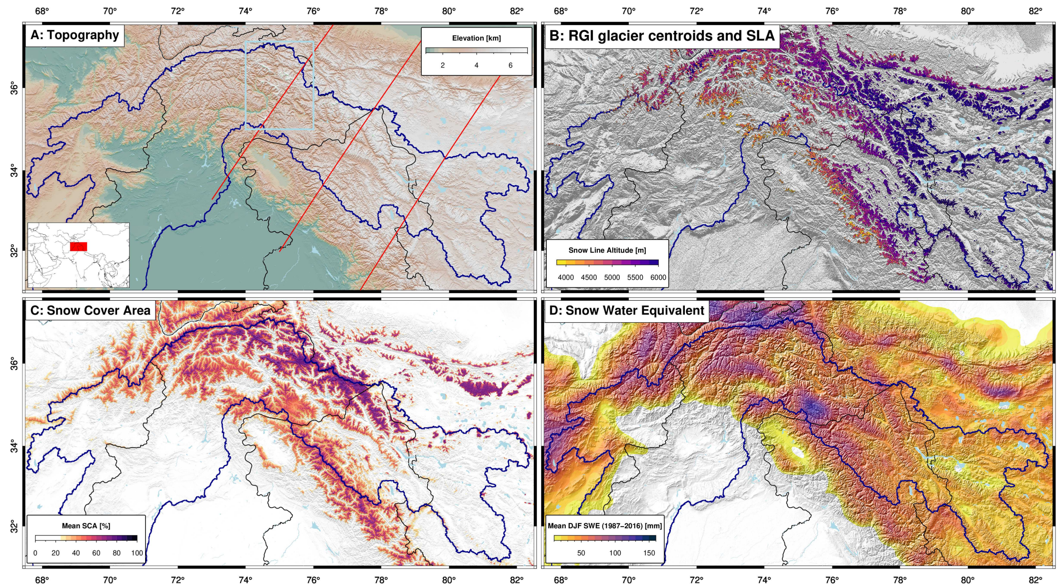

Figure 1. Topographic and climatic setting of the Upper Indus Basin (UIB). (A) Topography (SRTM, (JPL, 2020)), with red profile lines displayed in Supplemental Figure 1. (B) RGI V6 (Arendt et al., 2015) data and snow line altitude approximated by median elevation of glacial outline (e.g., Braithwaite & Raper, 2009; Racoviteanu et al., 2019). Snow line altitudes show a west-to-east and south-to-north gradient also documented in the SCA and SWE profiles (cf. Supplemental Figure 1). (C) Annual snow-covered area (SCA) calculated from MODIS (2001-2020, (Hall et al., 2002)). (D) December-January-February (DJF) average snow-water equivalent from SSM/I passive microwave data (1987-2016, (Smith & Bookhagen, 2020a; M. Brodzik et al., 2016)). There are large topographic variations throughout the UIB, with commensurate differences in snow-cover and snow-water storage.

and around the UIB (Figure 1B), reduced to only those glaciers with areas larger than 1 km^2. We use a subset of 7,357 glaciers (total area: 37,643 km^2) in this analysis, of which 3,830 (total area: 25,093 km^2) fall within the Indus basin. 1,021 of the 3,415 chosen watersheds contain glaciers larger than 1 km^2.

2.2 Snow Data Preparation

Recent advances in passive microwave snow-water equivalent (SWE) estimation provide SWE estimates at $\sim$3 km spatial resolution (M. J. Brodzik et al., 2012; M. Brodzik et al., 2016; Early & Long, 2001; Long & Brodzik, 2016; Chang et al., 1987; Smith & Bookhagen, 2020a), which is a drastic improvement upon previous 0.25 x 0.25° estimates of SWE (Smith & Bookhagen, 2016, 2018). While the spatial resolution of SWE estimates remains too coarse for fine-scale topographic analysis, it can be used to put higher resolution data into context (Figure 1).

In this study, we rely upon two higher-resolution snow-cover datasets: (1) MODIS SCA estimates (MOD10A1, 500m, 2001-2020, (Hall et al., 2002)) and (2) the Landsat 8 archive (2014-2020). MODIS SCA data have been shown to be above 90% accurate across a range of land cover types (Hall & Riggs, 2007; Parajka et al., 2012), and are thus well-suited to the broad delineation of snow-cover across elevations in the UIB. MODIS SCA data is converted to long-term means, December-January-

February (DJF) means, and June-July-August (JJA) means using Google Earth Engine (Gorelick et al., 2017).

We also rely on Google Earth Engine to pre-process and cloud mask the Landsat 8 archive (Gorelick et al., 2017). Using the masked and calibrated Landsat data, we calculate the long-term NDSI second percentile – used here as a proxy for persistent annual snow-cover – and standard deviation at 30 m spatial resolution. Unfortunately, due to the short length of the Landsat 8 time series, we cannot generate reliable seasonal NDSI estimates. We instead rely on the SCA estimates from MODIS to compare seasonal differences in snow character.

2.3 Insolation Estimation

Insolation is a strong driver of local microclimatic variation, which exerts a first-order control on cyrospheric processes (e.g., Smith & Bookhagen, 2020b; Olson et al., 2019; Cuffey & Paterson, 2010; Dozier, 1980). It can have a particularly large impact on glaciers; previous work has found that the main aspect orientation of a glacier plays a large role in its mass balance (e.g., Evans, 1977; Evans & Cox, 2005). At high elevations, sublimation also plays a role in controlling snow persistence (Rupper & Roe, 2008) as well as short- and long-wave radiation balances (Bonekamp et al., 2019). In our analysis, we calculate the total in-plane irradiance as the sum of the beam (I_{beam}), sky diffuse (I_d), and ground reflected components (I_{ground}): $I_{tot} = I_{beam} + I_d + I_{ground}$. We follow the method of Klucher (1979) using the pvlib software package (Holmgren et al., 2018) to measure the diffuse irradiance from the sky on a tilted surface, where we explicitly define surface tilt, surface azimuth angles, and elevation for each DEM grid cell. We integrate solar radiation calculations over an entire year with a time step of four hours and use the annual average as our insolation estimation for each grid cell. From these measurements, we derive by-watershed and by-glacier radiation medians using all pixels of a glacier or watershed polygon.

2.4 Topographic and Climatic Data Analysis

To explore the relationship between topography, snow-cover, and glaciers in the UIB, we use our watershed and glacier polygons to create subsets of topographic (elevation, slope, aspect) and climatic (SCA, NDSI) data. We then subdivide each environmental and topographic variable into 50 m elevation bins, running from 500 to 7000 m asl. For each elevation bin, we capture the number of pixels, both annual and seasonal averages of SCA, and long-term NDSI minimum (2^{nd} percentile) and variability (standard deviation). This processing yields unique elevation-binned statistics for each variable, which we use for further analysis (Figure 2).

2.5 Clustering Approach

As can be seen in Figure 2, the median elevation or elevation range of each watershed is not sufficient to differentiate climatic regions. The elevation-binned snow-cover medians, however, can be separated into groups with similar characteristics. In this study, we rigorously group these binned statistics using hierarchical clustering (Müllner, 2011; Murtagh & Contreras, 2012; Smith et al., 2017; Clubb et al., 2019). In short, hierarchical clustering involves partitioning a set of observations into coherent groups based on an arbitrary distance measure between all pairs of observations. The method is flexible and applicable to a wide range of data types and sizes, with the main constraint being that some measure of distance between each pair of measurements must be defined.

Determining the distance between two sets of observations is the first step for deriving hierarchical clusters. A wide range of methods are commonly used to deter-

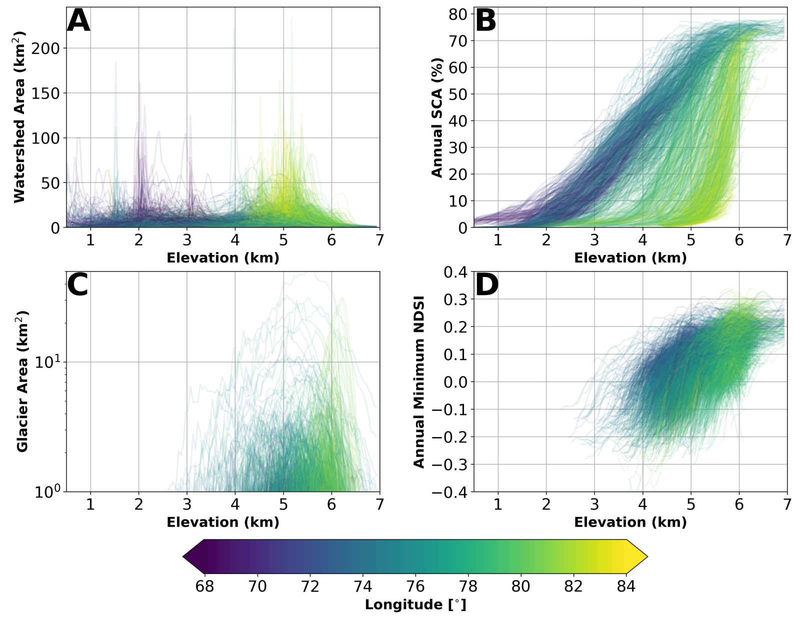

Figure 2. Elevation-binned medians of key variables. (A, B) Fourth-order watersheds and (C,D) glacier polygons showing (A,C) hypsometry and (B,D) elevation-snow cover relationships. Colors scale with longitude from west (dark) to east (light). More easterly watersheds and glaciers generally have steeper snow-cover curves – snow-covered area (SCA) increases rapidly with elevation. We note that the annual SCA cover of each watershed (B) can be visually grouped into two areas: (1) a shallow elevation-SCA relationship for western watersheds and (2) a steep relationship for eastern watersheds.

mine those distances (Murtagh & Contreras, 2012; Deza & Deza, 2009); we define the distance d between two elevation-binned data medians u and v as

$$d = \sum(|(u - v)|)/n \tag{1}$$

where n is the number of elevation bins that the two sets of binned medians share. We choose to normalize the summed distances to account for partially-overlapping sets of binned medians (e.g., watersheds or glaciers that share some, but not all, elevation bands). We also remove from our analysis any watersheds that do not have an annual average DJF SCA of at least 5% to minimize noise from mostly snow-free catchments.

Not all watersheds or glaciers in the UIB overlap in elevation; there exists a subset of binned medians for which the distance d is undefined (e.g., $n = 0$). Simply put, it is not possible to compute a distance between data sets which share no bases. These edge cases cannot simply be removed as outliers – they have well-defined distances to other sets of binned medians. They do, however, pose a problem for hierarchical clustering, which does not support undefined distances between cluster members.

We use two different methods to account for the undefined distance problem: (1) choose a subset of data that has no undefined distances (e.g., all binned medians pass through the same elevation range), and (2) define the distance between non-overlapping clusters as the median of all other distances (Figure 3, Supplemental Figure 2). While the distance between non-overlapping pairs cannot be directly determined, it can be inferred from their relative distances to other data. We choose the median of all other

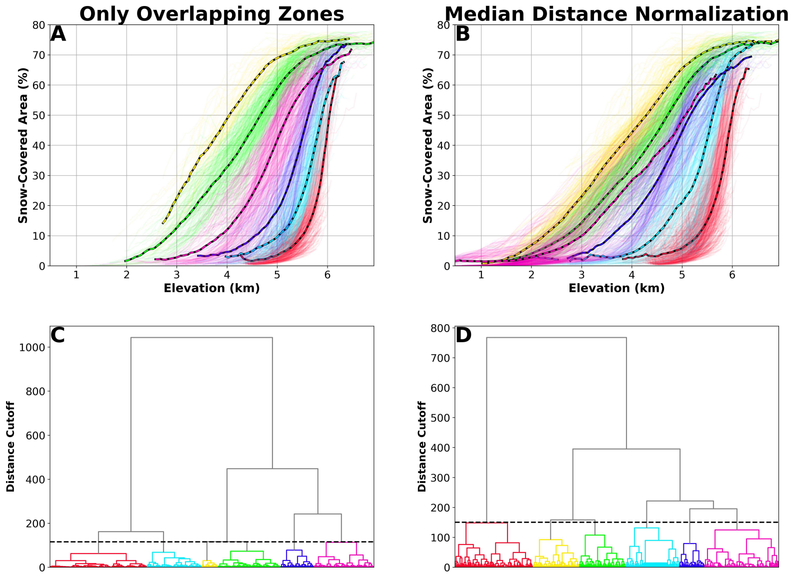

Figure 3. Dendrograms of the two distance calculations explored in this study: (A,C) Only distances between watersheds with overlapping elevation ranges are calculated and used for clustering. (B,D) For watersheds with no elevation overlap, the median distance of all other distances is used. Note the distinctive groups of elevation-snow-cover relations that result from both clustering steps. Clusters for both methods defined on elevation-binned median snow-covered area (SCA) using the watersheds of stream-order 4 (cf. Figure 1).

distances as a reasonable value to characterize those undefined distances. It should be noted that the undefined distance problem is less significant for glacier clustering, as glaciers are much more strongly confined to a limited elevation range.

Both of the proposed methods successfully partition the binned medians (cf. Figure 2). Regardless of which distance measurement we use, we define a linkage matrix from the distances d using Ward's method (Clubb et al., 2019; Müllner, 2011; Murtagh & Contreras, 2012). The cluster dendrograms for both approaches can be seen in Figure 3.

3 Results

3.1 Snow Cover Clusters

Clusters of watersheds based on elevation-binned SCA medians show a coherent spatial pattern and distinct differences in the shape of the elevation-SCA relationships (Figure 4), and are partitioned into pseudo-evenly sized groups (n=194-608). While the elevation-binned SCA medians for each cluster overlap, they do not maintain the same slope – each cluster represents a different relationship between elevation and snow-cover. In the mid-elevations of the central and western UIB, for example, SCA changes relatively slowly with elevation (green lines, n=364, Figure 4B). In contrast, those clusters in the eastern UIB and through the Kunlun Shan exhibit rapidly changing SCA with elevation (red lines, n=608, Figure 4D).

SCA and NDSI clusters are somewhat controlled by topography – both sets of clusters roughly follow the divide seen in hypsometry clusters between foreland, high mountain, and internal Tibetan Plateau areas (Supplemental Figure 3), and have a

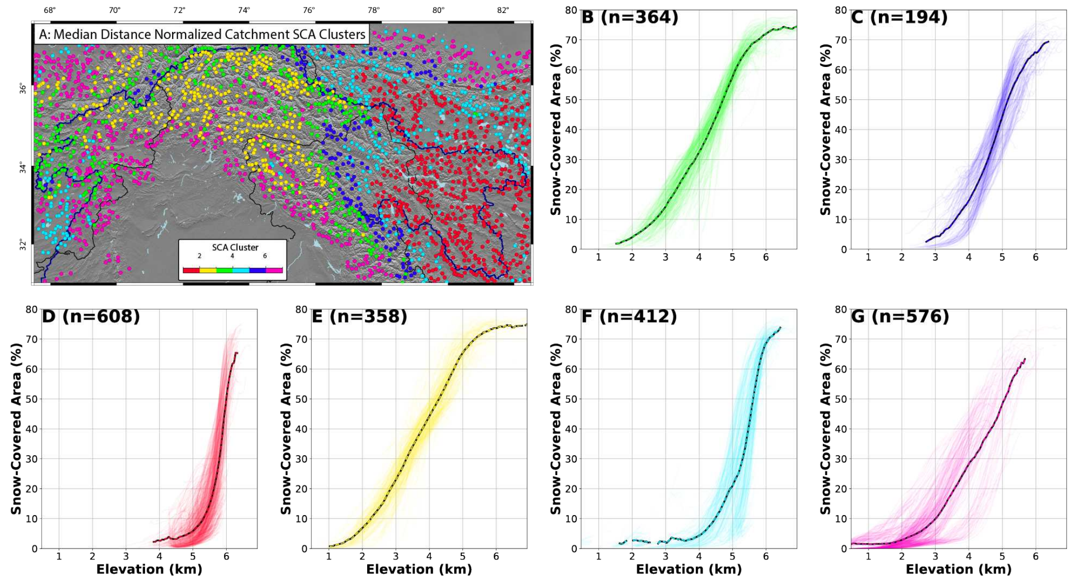

Figure 4. Clusters defined on the relationship between elevation and snow-covered area (SCA) using the watersheds of stream order four. The elevation-binned SCA medians of each cluster are distinct, though the binned medians of many clusters overlap. Number of watersheds in each cluster displayed on each individual chart.

similar distribution of cluster sizes. However, the elevation-binned snow-cover medians of each cluster remain distinct (Figure 4). The split between wetter and drier parts of the Tibetan Plateau region is not as strongly expressed in the simple hypsometry clusters (Supplemental Figure 4), but is apparent in both SCA (Figure 4) and NDSI (Supplemental Figure 5) clusters.

3.2 Glacier Clusters

The clustering approach can be extended to regional glaciers. While SCA provides useful context to glacier character, the spatial resolution (500 m) is too coarse for many small glaciers. We rely instead on the 2014-2020 second percentile NDSI as the basis for our glacier clusters, which serves as a proxy for snow persistence and stability at each elevation bin (Figure 5). Glacier clusters based on second percentile NDSI are fairly well-distributed into six clusters of roughly even sizes, with the exception of cluster 2 (yellow points, n=2965) which is located mainly in the Tibetan interior.

Hypsometry-based clusters (Supplemental Figure 6) are generally broken into two very large clusters defined by the majority of small- to mid-sized glaciers (n=2162, 3200), with larger glaciers comprising several smaller clusters. NDSI clusters, however, are more strongly split by elevation and aspect – those glaciers on opposing sides of a major drainage divide are often grouped into different clusters (Figure 5).

While aspect is a key control on glacier size, shape, and character (Evans, 1977; Evans & Cox, 2005), aspect does not seem to play a dominant role in controlling glacier clusters in our data. While we identify more glaciers with generally north-east aspects than any other direction (Supplemental Figure 7), these glaciers do not group into distinct clusters. The lack of clear aspect-based clusters is likely indicative of (1) the difficulty in assigning a single aspect value to large glacier areas, and (2) the large role of seasonal moisture transport and temperature regimes in controlling glacier size,

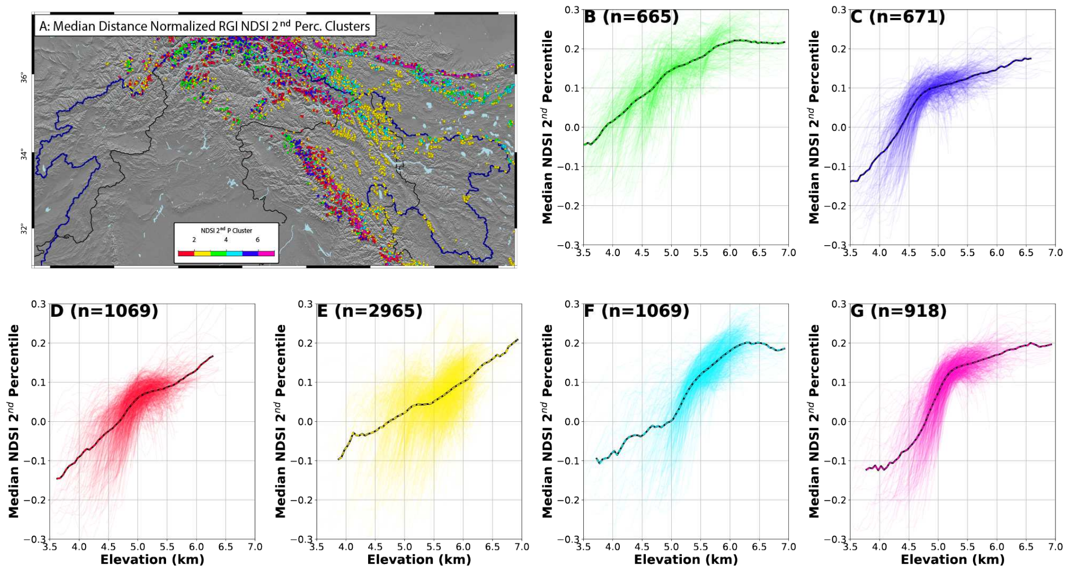

Figure 5. Clusters defined on the relationship between elevation and normalized-difference snow index (NDSI) 2nd percentile using the RGI glacier outlines (Arendt et al., 2015). The elevation-binned NDSI medians of each cluster overlap significantly; however, the slope of the elevation-NDSI relationship is distinctly different between clusters. Number of glaciers in each cluster displayed on each individual chart.

topographic setting, and stability (Fujita, 2008; Kapnick et al., 2014; Shean et al., 2020).

4 Discussion

It is clear that topography – both elevation and aspect – exerts a first-order control on the distribution of snow in the UIB (Figure 1). However, elevation is not enough to differentiate functional regions – climate dynamics also play a major role in controlling snowfall. The direction, timing, and magnitude of moisture transport throughout the region has been shown to strongly influence snow and glacier character (e.g., Fujita, 2008; Kapnick et al., 2014). Using our clusters based upon elevation-snow relationships, we can (1) explore the factors that lead to cluster formation, and (2) examine the relationship between our empirical clusters and the analysis regions commonly used to delineate self-similar regions in and around the UIB.

4.1 Climatic Controls on Watershed Clusters

When we compare the topographic and environmental setting of the different clusters defined by elevation-binned SCA medians, it is clear that watershed median elevation plays a dominant role (Figure 6). Some clusters, for example cluster 6 (low elevation, pink dots) and cluster 1 (high elevation, red dots) are found in distinct and non-overlapping elevation bands. However, many clusters are mixed across similar elevations, indicating that differences in SCA across watersheds are a stronger control on cluster formation.

There is a clear split at ~4500 m where watersheds go from seasonally to permanently snow-covered (Figure 6D). Low- to mid-elevation areas all have significantly negative minimum NDSI values, indicating that for at least part of the year they are

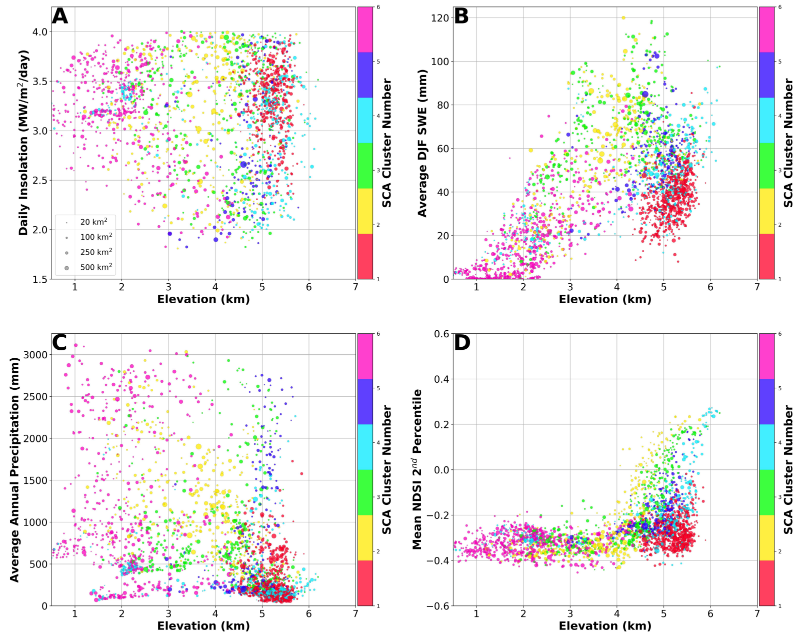

Figure 6. Median elevation of each watershed is compared to (A) annual insolation, (B) average December-January-February (DJF) snow-water equivalent (SWE), (C) annual precipitation (GPM, 2001-2020 (GPM Science Team, 2014)), and (D) annual minimum (2nd percentile) normalized difference snow index (NDSI). Clusters determined from elevation-binned SCA medians. Dots sized by watershed area. Elevation is a first-order control on cluster formation.

snow-free. Clusters 2, 3, and 4 span the range between seasonally snow-free (e.g., NDSI 2nd percentile <0) and permanently snow-covered. These watersheds, however, are not necessarily responsible for significant water storage; clusters 1 and 4 store less water in snow than the lower-elevation clusters 2 and 3. These watersheds are generally in the Tibetan interior (Figure 6C), and receive far less moisture than those watersheds on the exterior of the Plateau.

4.2 Limits to Topographic Controls on Glacier Clusters

Previous studies have noted the influence of received solar radiation on snow-cover (Smith & Bookhagen, 2020b) and glaciers (e.g., Evans, 1977; Evans & Cox, 2005; Olson et al., 2019; Bonekamp et al., 2019). We find that the majority of glaciers in our study region are on average north-facing (Supplemental Figure 7); many large glaciers, however, remain south-facing. It is important to note that the bias in glacier aspects does not mirror large-scale topography; north- and south-facing watersheds are evenly distributed in our study area. East- and west-facing watersheds are less frequent due to the regional tectonic setting which encourages the formation of north- and south-facing valleys (Supplemental Figure 7).

As with the watershed clusters, elevation plays a role in controlling the formation of glacier clusters. However, the role of elevation in controlling glacier clusters is much more subtle; cluster members are well mixed across elevations (Figure 7). Each cluster has rather a slightly different size distribution (Figure 7B); cluster 3, for example, has the largest elevation ranges, but not necessarily the most areally extensive glaciers.

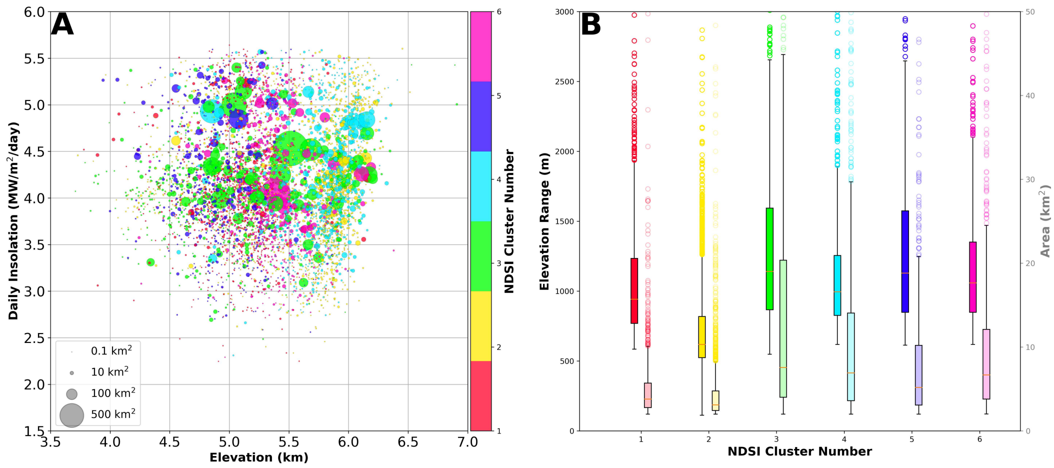

Figure 7. (A) Elevation compared to annual insolation, with dots colored by cluster number and sized by glacier area. Clusters based on NDSI 2$^{\text{nd}}$ percentile (cf. Figure 5). While elevation plays a role in determining cluster formation, it is not as strong of a control as for the watershed clusters (cf. Figure 6). (B) Elevation range (solid boxes, left axis) and area (transparent boxes, right axis) statistics by cluster. Boxes filled between the 25th and 75th percentiles, with thin lines extending to a maximum of 1.5 times the interquartile range. Outliers shown as individual dots above the thin lines. Colors match clusters in panel (A). Elevation ranges and areas are similar across clusters.

The minimal role of elevation in controlling glacier cluster formation can be attributed to a few likely causes. First, glacier clusters are formed over a limited elevation range, which naturally minimizes the difference between elevation-binned medians. Furthermore, snow character will tend to be less variable in the high-elevation zones where glaciers form – especially over glacier accumulation areas. This serves to emphasize differences in the ablation zone of glaciers, where snow tends to vary more across elevations and seasons. From these factors, we can infer that glacier cluster formation is more strongly controlled by topography (e.g., glacier slope and aspect) and climate (e.g., temperature and precipitation) than by elevation, which is a proxy commonly used to group glaciers. Glacier size and elevation range are also poor proxies (Figure 7). Additional high-resolution glacier data distinguishing between snow, ice, and debris could provide alternate and useful parameters for a clustering analysis (Scherler et al., 2011; Smith & Bookhagen, 2016; Racoviteanu et al., 2019). As many glaciers throughout the UIB and the greater High Mountain Asia region have debris-covered tongues, further research to constrain the role of these lower-elevation glacier regions in cluster formation is essential. A detailed analysis of the role of debris-cover is, however, beyond the scope of this study.

4.3 Towards Empirical Clusters for the Cryosphere

It is often important to aggregate data over spatial scales, particularly when the data are uncertain. Some form of regionalization is common in climate (e.g., Vaughan et al., 2013) and glacier (e.g., Kääb et al., 2015; Bolch et al., 2019; Shean et al., 2020) literature. These schemes generally use either (1) regular grids, or (2) semi-objective regions based on climate, topography, and rough mountain range delineations (Bolch et al., 2019). Our approach does not result in spatially contiguous areas, but rather delineates coherent regions from a topographic-climatic perspective. While some similarities between our clusters and previously published regions are apparent, there are also distinct differences (Figure 8).

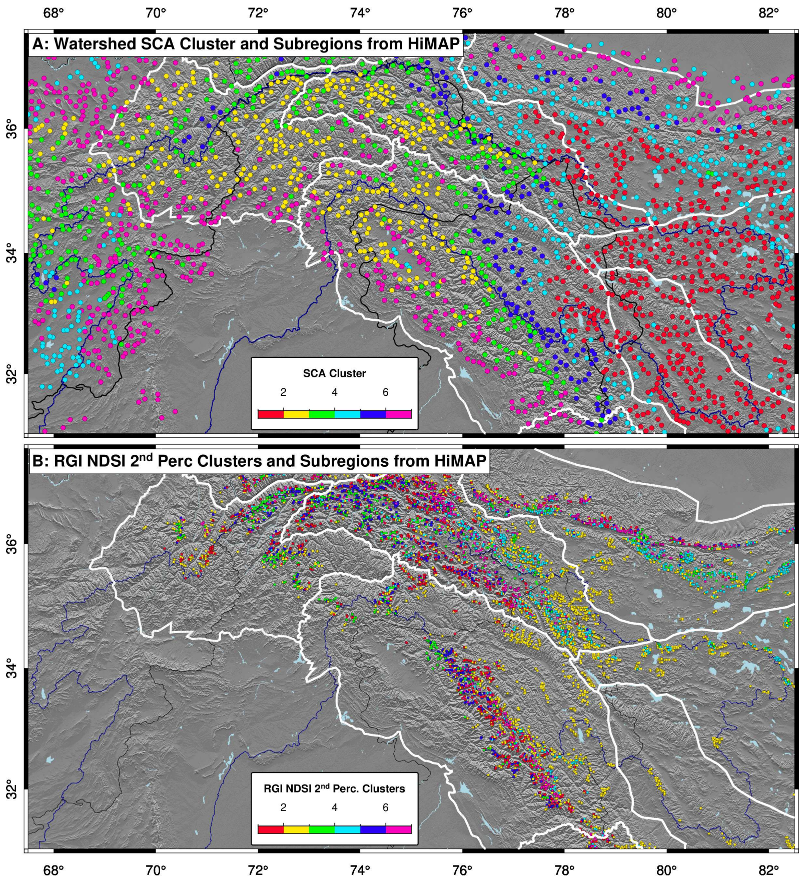

Figure 8. HiMAP zones (Bolch et al., 2019; Shean et al., 2020) compared to (A) watersheds of stream order four clustered based on elevation-binned snow-covered area (SCA) (cf. Figure 4), and (B) glacier clusters based on the elevation- normalized-difference snow-index (NDSI) relationship (cf. Figure 5). Rough spatial zones do a poor job of grouping self-similar glaciers.

Watershed clusters are relatively spatially coherent and generally represent large-scale climatic-topographic zones (Figure 8A). These zones could be useful for future studies as coherent analysis regions based on a defined climatic-topographic gradient with self-similar behavior. It is important to note that within each major watershed (e.g., the UIB) there are multiple watershed clusters. As each cluster has a unique elevation-snow relationship, their snow-water storage regimes will differ under current and future climate scenarios. These differences have implications for the timing and volume of snow-water storage and snowmelt in major watersheds across the region.

Glacier clusters are much more spatially heterogeneous than watershed clusters, and do not smoothly conform to glacier regions delineated in previous work, which generally uses elevation, climate, and spatially coherent mountain ranges to analyze glacier regions (e.g. Scherler et al., 2011; Kääb et al., 2015; Bolch et al., 2019; Shean et al., 2020). This is unsurprising given the strong role of aspect (Evans & Cox, 2005) and precipitation seasonality (Fujita, 2008) in controlling glacier mass balance. There remain, however, some similarities between our clusters and the glacier regions of Bolch et al. (2019). In particular, glacier cluster 4 is largely confined to the Western Kunlun

Shan and Tibet interior. The Karakoram region is, however, fairly evenly split between all six clusters defined in this study, indicating that aggregated glacier statistics in this area may be more uncertain than in more homogeneous regions.

5 Conclusion

The distribution of snow and glaciers in the UIB is spatially heterogeneous and highly dependent on regional precipitation patterns and topography; elevation-binned medians of snow-cover are also strongly influenced by latitude, with western areas having relatively shallow elevation-snow relationships, and eastern areas having steep elevation-snow relationships. Based on high-resolution topography (SRTM 1-arcsec) and snow-cover data from MODIS and Landsat, we propose a novel method of delineating climatic-topographic zones using a hierarchical clustering approach. We find that both watersheds and glaciers can be clustered into self-similar groups based on the empirical relationship between snow-cover and elevation. Watershed clusters are strongly influenced by median watershed elevation; glacier clusters are less so, and reflect stronger climatic and insolation-derived control on cluster formation. We emphasize that the topo-climatic clusters that we derive are different from those based solely on topography – the spatial distribution of snow-cover plays a large role in cluster formation, particularly for glaciers. We propose that clustering glaciers and watersheds using empirical data could provide a novel way of aggregating data in the complex environment of the UIB. Analyses based on empirical clusters could help reduce the errors propagated from uncertain in-situ, remotely sensed, and modeled data into regional hydrologic and climate analyses.

Acknowledgments

The State of Brandenburg (Germany) through the Ministry of Science and Education and the NEXUS project supported T.S. for part of this study (grant to. B.B.). We also acknowledge support from the BMBF ORYCS project.

Author contributions

T.S. and B.B. designed the study and prepared and analyzed the data. A.R. and B.B. contributed to the development of the methodology. All authors wrote the manuscript led by T.S.

Competing financial interests

The authors declare no competing financial interests.

Data Availability

SCA data is publicly available (Hall et al., 2002). Clustering results for both watersheds and glaciers are available on Zenodo (https://doi.org/10.5281/zenodo.4469473). NDSI data accessed via Google Earth Engine (Gorelick et al., 2017). SWE data can be found here: https://doi.org/10.5281/zenodo.3898517.

References

Abdullah, T., Romshoo, S. A., & Rashid, I. (2020). The satellite observed glacier mass changes over the upper indus basin during 2000–2012. *Scientific reports*, *10*(1), 1–9.

Arendt, A., Bolch, T., Cogley, J., Gardner, A., Hagen, J., Hock, R., … others (2015). Randolph glacier inventory [v5.0]: A dataset of global glacier outlines. global land ice measurements from space, boulder colorado, usa. *Digital Media*.

Athar, H., et al. (2019). Contribution of changing precipitation and climatic oscillations in explaining variability of water extents of large reservoirs in pakistan. *Scientific reports*, *9*(1), 1–14.

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. (2005). Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, *438*(7066), 303–309.

Baudouin, J.-P., Herzog, M., & Petrie, C. A. (2020). Cross-validating precipitation datasets in the indus river basin. *Hydrology and Earth System Sciences*, *24*(1), 427–450. Retrieved from https://hess.copernicus.org/articles/24/427/2020/ doi: 10.5194/hess-24-427-2020

Bolch, T., Kulkarni, A., Kääb, A., Huggel, C., Paul, F., Cogley, J., … others (2012). The state and fate of himalayan glaciers. *Science*, *336*(6079), 310–314.

Bolch, T., Shea, J. M., Liu, S., Azam, F. M., Gao, Y., Gruber, S., … others (2019). Status and change of the cryosphere in the extended hindu kush himalaya region. In *The hindu kush himalaya assessment* (pp. 209–255). Springer.

Bonekamp, P. N. J., de Kok, R. J., Collier, E., & Immerzeel, W. W. (2019). Contrasting meteorological drivers of the glacier mass balance between the karakoram and central himalaya. *Frontiers in Earth Science*, *7*, 107. Retrieved from https://www.frontiersin.org/article/10.3389/feart.2019.00107 doi: 10.3389/feart.2019.00107

Bookhagen, B. (2016). Glaciers and monsoon systems. In L. Carvalho & C. Jones (Eds.), *Monsoons and climate change: Observations and modeling* (p. 63-174). Springer. doi: 10.1007/978-3-319-17220-0_15

Bookhagen, B. (2017). The influence of hydrology and glaciology on wetlands in the himalayas. In H. Prins & T. Namgail (Eds.), *Bird migration across the himalayas: Wetland functioning amidst mountains and glaciers* (p. 175-188). doi: 10.1017/9781316335420.014

Bookhagen, B., & Burbank, D. W. (2006). Topography, relief, and trmm-derived rainfall variations along the himalaya. *Geophysical Research Letters*, *33*(8). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006GL026037 doi: https://doi.org/10.1029/2006GL026037

Bookhagen, B., & Burbank, D. W. (2010). Toward a complete himalayan hydrological budget: Spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge. *Journal of Geophysical Research: Earth Surface (2003–2012)*, *115*(F3).

Braithwaite, R., & Raper, S. (2009). Estimating equilibrium-line altitude (ela) from glacier inventory data. *Annals of Glaciology*, *50*(53), 127–132.

Brodzik, M., Long, D., Hardman, M., Paget, A., & Armstrong, R. (2016). Measures calibrated enhanced-resolution passive microwave daily ease-grid 2.0 brightness temperature esdr. *National Snow and Ice Data Center, Boulder, CO USA*.

Brodzik, M. J., Billingsley, B., Haran, T., Raup, B., & Savoie, M. H. (2012). Ease-grid 2.0: Incremental but significant improvements for earth-gridded data sets. *ISPRS International Journal of Geo-Information*, *1*(1), 32–45.

Cannon, F., Carvalho, L. M., Jones, C., & Norris, J. (2015). Winter westerly disturbance dynamics and precipitation in the western himalaya and karakoram: a wave-tracking approach. *Theoretical and Applied Climatology*, 1–18.

Cannon, F., Carvalho, L. M. V., Jones, C., Norris, J., Bookhagen, B., & Kiladis, G. N. (2017). Effects of topographic smoothing on the simulation of winter precipitation in high mountain asia. *Journal of Geophysical Research: Atmospheres*, *122*(3), 1456-1474. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD026038` doi: https://doi.org/10.1002/2016JD026038

Chang, A., Foster, J., & Hall, D. (1987). Nimbus-7 smmr derived global snow cover parameters. *Annals of glaciology*, *9*(9), 39–44.

Clubb, F. J., Bookhagen, B., & Rheinwalt, A. (2019). Clustering river profiles to classify geomorphic domains. *Journal of Geophysical Research: Earth Surface*, *124*(6), 1417–1439.

Cuffey, K. M., & Paterson, W. S. B. (2010). *The physics of glaciers*. Academic Press.

Deza, M. M., & Deza, E. (2009). Encyclopedia of distances. In *Encyclopedia of distances* (pp. 1–583). Springer.

Dimri, A. P., Niyogi, D., Barros, A. P., Ridley, J., Mohanty, U. C., Yasunari, T., & Sikka, D. R. (2015). Western disturbances: A review. *Reviews of Geophysics*, *53*(2), 225-246. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014RG000460` doi: https://doi.org/10.1002/2014RG000460

Dozier, J. (1980). A clear-sky spectral solar radiation model for snow-covered mountainous terrain. *Water Resources Research*, *16*(4), 709-718. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/WR016i004p00709` doi: https://doi.org/10.1029/WR016i004p00709

Early, D. S., & Long, D. G. (2001). Image reconstruction and enhanced resolution imaging from irregular samples. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(2), 291–302.

Evans, I. S. (1977). World-wide variations in the direction and concentration of cirque and glacier aspects. *Geografiska Annaler: Series A, Physical Geography*, *59*(3-4), 151–175.

Evans, I. S., & Cox, N. J. (2005). Global variations of local asymmetry in glacier altitude: separation of north–south and east–west components. *Journal of glaciology*, *51*(174), 469–482.

Farinotti, D., Immerzeel, W. W., de Kok, R. J., Quincey, D. J., & Dehecq, A. (2020). Manifestations and mechanisms of the karakoram glacier anomaly. *Nature Geoscience*, *13*(1), 8–16.

Fowler, H., & Archer, D. (2006). Conflicting signals of climatic change in the upper indus basin. *Journal of Climate*, *19*(17), 4276–4293.

Fujita, K. (2008). Effect of precipitation seasonality on climatic sensitivity of glacier mass balance. *Earth and Planetary Science Letters*, *276*(1), 14–19.

Gardelle, J., Berthier, E., & Arnaud, Y. (2012). Slight mass gain of karakoram glaciers in the early twenty-first century. *Nature geoscience*, *5*(5), 322–325.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, *202*, 18–27.

GPM Science Team. (2014). Gpm gmi level 1b brightness temperatures, version 03. *Greenbelt, MD, USA: NASA Goddard Earth Science Data and Information Services Center (GES DISC)*.

Hall, D. K., & Riggs, G. A. (2007). Accuracy assessment of the modis snow products. *Hydrological Processes: An International Journal*, *21*(12), 1534–1547.

Hall, D. K., Riggs, G. A., Salomonson, V. V., DiGirolamo, N. E., & Bayr, K. J. (2002). Modis snow-cover products. *Remote sensing of Environment*, *83*(1), 181–194.

Hewitt, K. (2005). The karakoram anomaly? glacier expansion and the'elevation effect,'karakoram himalaya. *Mountain Research and Development*, *25*(4), 332–

340.
Holmgren, W. F., Hansen, C. W., & Mikofski, M. A. (2018). pvlib python: a python package for modeling solar energy systems. *Journal of Open Source Software*, *3*(29), 884. Retrieved from `https://doi.org/10.21105/joss.00884` doi: 10.21105/joss.00884

Huning, L. S., & AghaKouchak, A. (2020). Global snow drought hot spots and characteristics. *Proceedings of the National Academy of Sciences*, *117*(33), 19753–19759.

Huss, M., Bookhagen, B., Huggel, C., Jacobsen, D., Bradley, R., Clague, J., ... Winder, M. (2017). Towards mountains without permanent snow and ice. *Earth's Future*, n/a–n/a. Retrieved from `http://dx.doi.org/10.1002/2016EF000514` (2016EF000514) doi: 10.1002/2016EF000514

Immerzeel, W., & Bierkens, M. (2012). Asia's water balance. *Nature Geoscience*, *5*(12), 841–842.

Immerzeel, W., Wanders, N., Lutz, A., Shea, J., & Bierkens, M. (2015). Reconciling high-altitude precipitation in the upper indus basin with glacier mass balances and runoff. *Hydrology and Earth System Sciences*, *19*(11), 4673–4687.

Immerzeel, W. W., Droogers, P., De Jong, S., & Bierkens, M. (2009). Large-scale monitoring of snow cover and runoff simulation in himalayan river basins using remote sensing. *Remote sensing of Environment*, *113*(1), 40–49.

Immerzeel, W. W., Van Beek, L. P., & Bierkens, M. F. (2010). Climate change will affect the asian water towers. *Science*, *328*(5984), 1382–1385.

JPL, N. (2020). *Nasadem merged dem global 1 arc second v001 [data set]. nasa eosdis land processes daac.* Retrieved 2020-12-10, from `https://doi.org/10.5067/MEaSUREs/NASADEM/NASADEM_HGT.001`

Kääb, A., Treichler, D., Nuth, C., & Berthier, E. (2015). Brief communication: Contending estimates of 2003–2008 glacier mass balance over the pamir–karakoram–himalaya. *The Cryosphere*, *9*(2), 557–564.

Kapnick, S. B., Delworth, T. L., Ashfaq, M., Malyshev, S., & Milly, P. (2014). Snowfall less sensitive to warming in karakoram than in himalayas due to a unique seasonal cycle. *Nature Geoscience*, *7*(11), 834–840.

Klucher, T. (1979). Evaluation of models to predict insolation on tilted surfaces. *Solar Energy*, *23*(2), 111 - 114. Retrieved from `http://www.sciencedirect.com/science/article/pii/0038092X79901105` doi: https://doi.org/10.1016/0038-092X(79)90110-5

Liu, J., Kang, S., Hewitt, K., Hu, L., & Xianyu, L. (2018). Large observational bias on discharge in the indus river since 1970s. *Scientific Reports*, *8*(1), 17291. Retrieved from `https://doi.org/10.1038/s41598-018-35600-3` doi: 10.1038/s41598-018-35600-3

Long, D. G., & Brodzik, M. J. (2016). Optimum image formation for spaceborne microwave radiometer products. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(5), 2763–2779.

Lund, J., Forster, R. R., Rupper, S. B., Marshall, H., Deeb, E. J., & Hashmi, M. Z. U. R. (2019). Mapping snowmelt progression in the upper indus basin with synthetic aperture radar. *Frontiers in Earth Science*, *7*, 318.

Lutz, A. F., Immerzeel, W. W., Kraaijenbrink, P. D., Shrestha, A. B., & Bierkens, M. F. (2016). Climate change impacts on the upper indus hydrology: Sources, shifts and extremes. *PloS one*, *11*(11), e0165630.

Malik, N., Bookhagen, B., & Mucha, P. J. (2016). Spatiotemporal patterns and trends of indian monsoonal rainfall extremes. *Geophysical Research Letters*, *43*(4), 1710–1717.

Maussion, F., Scherer, D., M?lg, T., Collier, E., Curio, J., & Finkelnburg, R. (2014). Precipitation seasonality and variability over the tibetan plateau as resolved by the high asia reanalysis. *Journal of Climate*, *27*(5), 1910 - 1927. Retrieved from `https://journals.ametsoc.org/view/journals/clim/27/5/`

jcli-d-13-00282.1.xml doi: 10.1175/JCLI-D-13-00282.1

Müllner, D. (2011). Modern hierarchical, agglomerative clustering algorithms. *arXiv preprint arXiv:1109.2378*.

Murtagh, F., & Contreras, P. (2012). Algorithms for hierarchical clustering: an overview. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *2*(1), 86–97.

Norris, J., Carvalho, L. M., Jones, C., & Cannon, F. (2015). Wrf simulations of two extreme snowfall events associated with contrasting extratropical cyclones over the western and central himalaya. *Journal of Geophysical Research: Atmospheres*, *120*(8), 3114–3138.

Norris, J., Carvalho, L. M., Jones, C., & Cannon, F. (2020). Warming and drying over the central himalaya caused by an amplification of local mountain circulation. *npj Climate and Atmospheric Science*, *3*(1), 1–11.

Olson, M., Rupper, S., & Shean, D. E. (2019). Terrain induced biases in clear-sky shortwave radiation due to digital elevation model resolution for glaciers in complex terrain. *Frontiers in Earth Science*, *7*, 216. Retrieved from https://www.frontiersin.org/article/10.3389/feart.2019.00216 doi: 10.3389/feart.2019.00216

Palazzi, E., Hardenberg, J., & Provenzale, A. (2013). Precipitation in the hindukush karakoram himalaya: Observations and future scenarios. *Journal of Geophysical Research: Atmospheres*, *118*(1), 85–100.

Parajka, J., Holko, L., Kostka, Z., & Blöschl, G. (2012). Modis snow cover mapping accuracy in a small mountain catchment–comparison between open and forest sites. *Hydrology and Earth System Sciences*, *16*(7), 2365–2377.

Pellicciotti, F., Buergi, C., Immerzeel, W. W., Konz, M., & Shrestha, A. B. (2012). Challenges and Uncertainties in Hydrological Modeling of Remote Hindu Kush–Karakoram–Himalayan (HKH) Basins: Suggestions for Calibration Strategies. *Mountain Research and Development*, *32*(1), 39 – 50. Retrieved from https://doi.org/10.1659/MRD-JOURNAL-D-11-00092.1 doi: 10.1659/MRD-JOURNAL-D-11-00092.1

Purinton, B., & Bookhagen, B. (2018). Measuring decadal vertical land-level changes from srtm-c (2000) and tandem-x ($\sim$ 2015) in the south-central andes. *Earth Surface Dynamics*, *6*(4), 971–987. Retrieved from https://esurf.copernicus.org/articles/6/971/2018/ doi: 10.5194/esurf-6-971-2018

Racoviteanu, A. E., Rittger, K., & Armstrong, R. (2019). An automated approach for estimating snowline altitudes in the karakoram and eastern himalaya from remote sensing. *Frontiers in Earth Science*, *7*, 220. Retrieved from https://www.frontiersin.org/article/10.3389/feart.2019.00220 doi: 10.3389/feart.2019.00220

Rupper, S., & Roe, G. (2008). Glacier changes and regional climate: A mass and energy balance approach. *Journal of Climate*, *21*(20), 5384 - 5401. Retrieved from https://journals.ametsoc.org/view/journals/clim/21/20/2008jcli2219.1.xml doi: 10.1175/2008JCLI2219.1

Scherler, D., Bookhagen, B., & Strecker, M. R. (2011). Spatially variable response of himalayan glaciers to climate change affected by debris cover. *Nature Geoscience*, *4*(3), 156–159.

Schwanghart, W., & Scherler, D. (2014). Short communication: Topotoolbox 2 – matlab-based software for topographic analysis and modeling in earth surface sciences. *Earth Surface Dynamics*, *2*(1), 1–7. Retrieved from https://esurf.copernicus.org/articles/2/1/2014/ doi: 10.5194/esurf-2-1-2014

Shean, D., Bhushan, S., Montesano, P., Rounce, D., Arendt, A., & Osmanoglu, B. (2020). A systematic, regional assessment of high mountain asia glacier mass balance. front. *Earth Sci*, *7*, 363.

Smith, T., & Bookhagen, B. (2016). Assessing uncertainty and sensor biases in passive microwave data across high mountain asia. *Remote Sensing of Environ-*

ment, *181*, 174–185.

Smith, T., & Bookhagen, B. (2018). Changes in seasonal snow water equivalent distribution in high mountain asia (1987 to 2009). *Science advances*, *4*(1), e1701550.

Smith, T., & Bookhagen, B. (2019). Remotely sensed rain and snowfall in the himalaya. In A. P. Dimri, B. Bookhagen, M. Stoffel, & T. Yasunari (Eds.), *Himalayan weather and climate and their impact on the environment* (p. 225-249). Springer. doi: 10.1007/978-3-319-21650-8_11

Smith, T., & Bookhagen, B. (2020a). Assessing multi-temporal snow-volume trends in high mountain asia from 1987 to 2016 using high-resolution passive microwave data. *Frontiers in Earth Science*, *8*, 392. Retrieved from https://www.frontiersin.org/article/10.3389/feart.2020.559175 doi: 10.3389/feart.2020.559175

Smith, T., & Bookhagen, B. (2020b). Climatic and biotic controls on topographic asymmetry at the global scale. *Journal of Geophysical Research: Earth Surface*, *n/a*(n/a), e2020JF005692. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020JF005692 (e2020JF005692 2020JF005692) doi: https://doi.org/10.1029/2020JF005692

Smith, T., Bookhagen, B., & Rheinwalt, A. (2017). Spatiotemporal patterns of high mountain asia's snowmelt season identified with an automated snowmelt detection algorithm, 1987–2016. *The Cryosphere*, *11*(5), 2329–2343. Retrieved from https://www.the-cryosphere.net/11/2329/2017/ doi: 10.5194/tc-11-2329-2017

Tahir, A. A., Chevallier, P., Arnaud, Y., & Ahmad, B. (2011). Snow cover dynamics and hydrological regime of the hunza river basin, karakoram range, northern pakistan. *Hydrology and Earth System Sciences*, *15*(7), 2275–2290. Retrieved from http://www.hydrol-earth-syst-sci.net/15/2275/2011/ doi: 10.5194/hess-15-2275-2011

Treichler, D., Kääb, A., Salzmann, N., & Xu, C.-Y. (2019). Recent glacier and lake changes in high mountain asia and their relation to precipitation changes. *The Cryosphere*, *13*(11), 2977–3005.

Vaughan, D., Comiso, J., Allison, I., Carrasco, J., Kaser, G., Kwok, R., ... Zhang, T. (2013). Observations: Cryosphere. in: Climate change 2013: The physical science basis. *Contribution of Working Group I to the Fifth Assessment Report of the IPCC.*.

Wulf, H., Bookhagen, B., & Scherler, D. (2016). Differentiating between rain, snow, and glacier contributions to river discharge in the western himalaya using remote-sensing data and distributed hydrological modeling. *Advances in Water Resources*, *88*, 152–169.

Yoon, Y., Kumar, S. V., Forman, B. A., Zaitchik, B. F., Kwon, Y., Qian, Y., ... Mukherjee, A. (2019). Evaluating the uncertainty of terrestrial water budget components over high mountain asia. *Frontiers in Earth Science*, *7*, 120. Retrieved from https://www.frontiersin.org/article/10.3389/feart.2019.00120 doi: 10.3389/feart.2019.00120