Journal of Himalayan Earth Sciences Volume 58, No. 1, 2025 pp. 01-20

Extracting clear ice surface of mountainous glaciers of Karakoram Range using Machine Learning for different Band Ratio compositions of OLI: Case Study of Hunza Sub-Basin

***Syed Najam ul Hassan[1, 2], Mohd Nadzri Md. Reba[2], and Aftab Ahmed[2]**

[1]Geoscience & Digital Earth Centre (INSTeG), Research Institute for Sustainability & Environment (RISE), Universiti Teknologi Malaysia (UTM), Johar Bahru, 83110 Malaysia.

[2]Departmenty of Computer Sciences, Karakoram International University, Gilgit, 15100, Pakistan

**Corresponding author: syed.najam@kiu.edu.pk*

Submitted date: 16/05/2024 Accepted date: 25/10/2024 Published online: 31/03/2025

Abstract

Glaciers in the Hindu Kush-Karakoram-Himalaya region impact Earth's climate, contribute freshwater downstream, and influence weather patterns of precipitation and temperature. However, the region needs more detailed information about its glaciers. Specifically, the stability of glaciers in the Karakoram range of the Hunza sub-basin is a well-known anomaly. Therefore, monitoring its glaciers is needed to understand the dynamics of climate change in HKH. Glacier inventory is baseline data for monitoring, and the clear-ice surface is a quantifying parameter of glacier changes. Recently, Operational Land Imager (OLI), exploited with machine learning (ML), is highly recommended for glacier monitoring due to improved accuracy. So, it is necessary to update the current status of glaciers in sub-basin using OLI and ML. Therefore, the study aims a) to evaluate the current extent of clear ice in the sub-basin to examine stability and b) to exploit the application of ML for extracting clear ice from OLI and assess accuracy. Google Earth environment is used to derive the data of Optical Land Imager and further analyze it with a machine learning approach to classify the extent of clear ice. Random Forest classifier with minimum Root means square error (0.1 to 0.4) used through SNAP environment. Results indicate satisfactory spatial distribution of clear ice in higher elevations (> 5000 meters). 10 % area difference percentage exhibited in overall extent; however, 28 glaciers (area > 5 km^2) showed variation in the extent and confirmed the localized heterogeneity. Overall accuracy (82% to 83 %) and kappa coefficient values (0.64 to 0.65) confirm the role of individual bands of OLI. It is concluded that the glaciers in the sub-basin have an overall stable clear-ice extent except for variations in terminal ends. Meanwhile, machine learning has a significant role in the automatic extraction of clear ice when exploited with the OLI.

Keywords: Machine Learning; Random Forest; Mountain Glacier; Himalaya; Operational Land Imager: Climate Change: Hunza Sub-Basin.

1. Introduction

Mountainous glaciers have become one of the essential climate indicators and a sensitive interface to the changes in climate factors, such as temperature, precipitation, and snowfall, which are significantly related to the changes in glacier physical properties (Haireti et al., 2015). Information about glaciers becomes more pertinent to study for deriving local climate information in areas where climate stations and instrumentations are unavailable and inaccessible (Bolch et al., 2012). Besides the physical changes, mountainous glaciers are also a major contributor to providing fresh water to the downstream basins (Huss et al., 2017). Therefore, continuous monitoring of mountainous glaciers is needed to understand the dynamics of climate change and water security. Comprehensive information about the glacier attributes is available in glacier inventory, which has become the baseline data for climate change assessment (IPCC, 2013), hydrological modeling and prediction, river run-off estimation, sea-level rise, and glacier modeling (Kraaijenbrink et al., 2017). To date, there are many glacier inventories established, and sharing the large and regional-scale glacier data with public, such as the Glacier Area Mapping for Discharge from the (Missions and Catalogue, 2022) Asian Mountains

(GAMDAM) glacier inventory (Sakai, 2019), Chinese Glacier Inventory (Guo et al., 2015), Randolph Glacier Inventory (Pfeffer et al., 2014), Glacier inventory of Pamir and Karakoram (Mölg et al., 2018b) and Status of Glacier in the Hindu-Kush Karakoram Himalaya (HKH) inventory (Bajracharya and Shrestha, 2011; Minora et al., 2016).

The Hindu Kush Himalayas (HKH) is the primary mountain range in Central Asia, hosting the most concentrated collection of mountain glaciers in the world. These glaciers account for almost one-sixth of the global ice cover, making it the world's third Pole (Bajracharya and Shrestha, 2011; Dyhrenfurth, 1955). This region stretches over 60,000 km^2 and has three mighty mountainous ranges with dynamic climatic zones. Glaciers in the Karakoram and western Himalaya region mainly contribute fresh water for the Indus basin. These glaciers contribute not only to agriculture production, accounting for 90%, but also provide 13 Gigawatts of hydroelectricity to the downstream population (Cook et al., 2013). The Karakoram glaciers occupy a more significant part of the glaciated area in the Upper Indus Basin (UIB), which includes half of the glacier's population in the HKH (Khan et al., 2015). In UIB, the Hunza sub-basin is famous for its heterogeneous glacier characteristics but has possessed stable glacier extent since 1970 with a mass budget of -0.09±0.03 m.w.e.a-1 from 1970 to 2000 (Bolch et al., 2017). Despite being focused on research studies, the Hunza sub-basin is experiencing extreme topographic and climate conditions along with frequent glacier surging. Moreover, there is a lack of quantitative and qualitative knowledge of regional glaciology in terms of its topography, climate, and glacier variations (Baig et al., 2018). This glacier range experiences stability in glacier extents and exhibits the famous Karakoram Pamir Anomaly, which contrasts with the global climate change trends (Syed et al., 2018). Hence, knowledge about the current glacier extent and the nature of the changes in the Hunza basin is important to support downstream hydrology and water resource management. Therefore, there is a need for continuous monitoring and updating of the basin's glacier inventory.

Spaceborne remote sensing is an effective tool to monitor any inaccessible large mountainous glacier region, and multispectral remote sensing with medium resolution becomes the typical remote sensing variant in glacier studies (Paul et al., 2015). Optical remote sensing spanned over four decades of continuous glacier monitoring missions. It started with the launch of Earth Resource Tech (Krimmel and Meier, 1975), and in 2013, the multispectral Operational Land Imager (OLI) onboard in Landsat 8 mission with improved sensing capabilities is a significant addition (Syed et al., 2018). This improved capacity of Landsat 8 is equipped with additional bands compared to its predecessor through optical and thermal regions. Furthermore, it is superior in terms of all spatial and spectral resolutions, and the radiometric resolution of 16 bits is a significant improvement over traditional 8-bit Landsat data (Linghong et al., 2016).

The latest glacier inventories of UIB, including the Karakoram glaciers of the Hunza sub-basin, are largely developed based on optical satellite imageries (Khan et al., 2015). An essential aspect of these inventories is the delineation of clear ice and debris cover separately with different underlying methodologies. The mapping of clean ice is important to quantify the changes over time as a requirement by the World Glacier Monitoring System (ICIMOD, 2010). Therefore, most of the previous studies, i.e., Mölg et al. (2018a), Robson et al. (2015), and Guo et al. (2015) for mapping of the glaciers have delineated the clear ice surfaces separately.

The regional inventory reported by Mölg et al. (2018a) used a well-established semi-automatic band ratio method to classify the clean ice and snow part of the glacier surface by exploiting Landsat-7 spectroscopic capabilities and discussed the status of glaciers from 2001 to 2010 of Central Karakoram National Park. Bajracharya and Shrestha (2011) has extracted the clear ice features from Landsat-5 and Landsat-7 imageries from observation in the year 2005±3 through the Normalized Difference Snow Index (NDSI). The traditional NDSI method is susceptible to misclassification errors, and therefore, Bajracharya and Shrestha (2011) further

refined the results using the Normalized Difference Vegetation Index (NDVI), Land and Water Mask (LWM), hue for misclassified vegetation, water bodies, and shadows respectfully. Local inventory with the long status of glaciers from 2001 to 2010 in Central Karakoram was created by Mölg et al. (2018a), who followed the recommendations from a previous study by Paul et al. (2009) to use the band ratio approach for compiling the snow and ice information in the inventory. The latest version, 6.0 of the globally referenced Randolph Glacier Inventory (RGI), which is an unchanged version of 5.0 for the Karakoram region, was composed of contributions from regional inventories, and most of the clear ice information was extracted based on band ratio methods in the inventories (Arendt et al., 2017). Many studies have used the band ratio method for extracting the clear ice information from satellite imagery for the global, regional, and local inventories, and the same has been exploited in larger glacier extents of the Hunza basin (Khan et al., 2015). Threshold values of band ratios, which are determined empirically, are used to differentiate glacier ice and non-glaciated surfaces. However, this approach is unsuccessful when the ice needs to be separated from snow, water bodies, and debris-covered glacier terrain (Racoviteanu et al., 2009). It is also challenging to fix the threshold value because a minor variation in the value may increase the number of misclassified pixels (Paul et al., 2013). Moreover, the band ratio method also fails when glaciers need to be separated from urban areas (Lary et al., 2016). To complement the band ratio method and improve the accuracy and efficiency of clean ice extraction, the Machine Learning (ML) approach was introduced in previous studies during the last decade (Zhang et al., 2019). In addition to that, improved classification approaches in remote sensing have been developed recently, using machine learning algorithms like Random Forest (RF), Support Vector Machine, Artificial Neural Networks, etc. (Hussain and Khan, 2020). At the same time, ML is also good for training models adequately, even on data sets of high dimensions and with poor signal-to-noise ratios (Maxwell et al., 2018). RF as a machine learning algorithm is a promising classifier for mapping different earth features using remotely sensed data (Krishna et al., 2018), and many applications have been reported for advancements in mapping landcover changes and water using RF (Wessels et al., 2016; Mueller et al., 2016). The potential of RF to analyze Glacier Lakes Out Burst phenomena for spectrally variable target classes such as glacial lakes (Veh et al., 2018) makes it a viable choice of method to replace the single parametric approaches for spectrally variable mountainous glaciated surfaces. The application of machine learning to delineate clean ice glaciers using the recent Landsat OLI imageries is yet unexplored, and this study is motivated by the radiometric and spectral quality of Landsat 8 and the advantages of machine learning to map mountainous glaciers, particularly clean ice at higher accuracy for updating the HKH inventory (Syed et al., 2018).

This study aims (a) to provide an overview of the current extent of clear ice in the Hunza sub basin to inquire about the stability of the heterogeneous glacier in the basin, b) to explore the application of machine learning to apply all the combinations of band ratio to extract clear ice surface from the given optical data over the study area, and c) to assess the efficiency of the adopted methodology of machine learning by analyzing the accuracy and validation for extraction of clear ice in mountainous regions through optical data.

2. Study area and Datasets

1.1. Study Area

The study area is the Huzza sub basin of UIB, located in the upper North of HKH region and surrounded by the mountain range of Karakoram see Figure 1. The area extends between longitude of 74 to 76 degrees East and latitude of 35 to 37-degree North with the elevation ranges from 2000 to 8500 meters. The mean elevation of glaciers is between 4000 to 6000 m. This area has been listed in glacier inventories developed by Mölg et al. (2018b) and Bajracharya and Shrestha (2011). The former inventory has a total area of 13987 km^2, derived from the observations of Landsat-5 TM and Landsat-7 ETM+ between 1998 to 2002. It identified a total of 421 glaciers, each with an

area of above 02 km2. The latter is developed from imageries of Landsat-5 MSS and Landsat-7 ETM+ taken for the year 2005±3 and listed 453 glaciers in the total area of 12677 km2 for the clear surface glaciers.

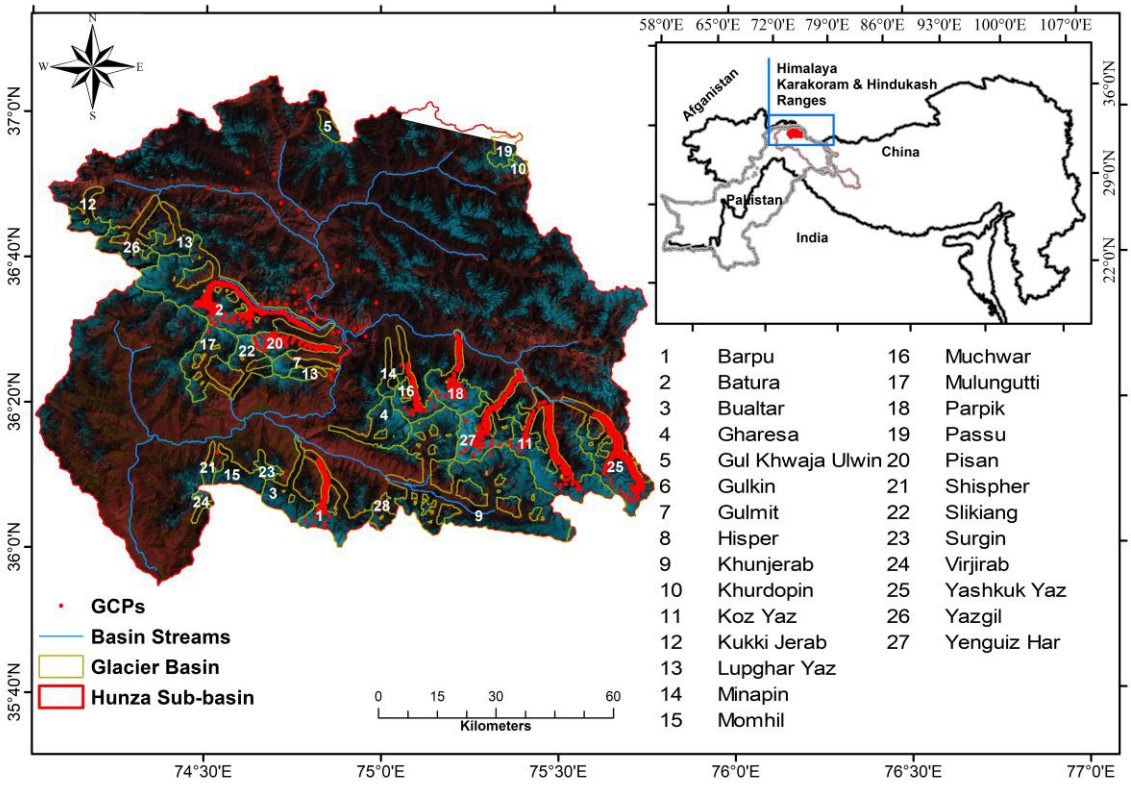

Fig. 1. Study area map showing geographic extent of HKH and Upper Indus Basin with sample glaciers exposed.

Climatically, the region is influenced by two sources of precipitation, i.e., monsoon and depression coming from the west, which rise to 600 mm at an altitude of 4400 and above (Immerzeel et al., 2015). This pattern of precipitation is responsible for summer and winter accumulation in the glaciated surfaces (Wiltshire, 2014) which variates the magnitude of clear ice surface in a specific glacier. Any change in glacier magnitude triggers glacier surging phenomena, that is why glaciers of the Karakoram range in the region are famous for the frequent surges, as reported in many of the studies, which complicate analysis to quantify the glacier statistics (Bolch et al., 2017).

2.2. Satellite data

The Landsat 8 series, equipped with the Operational Land Imager (OLI), was launched in February 2013. OLI is the continuation of archiving moderate resolution earth observation data, which has a significant improvement compared to the previous series of Landsat 5 & 7; push broom sensor type, long array of detectors, and additional bands of Cirrus and Coastal Aerosol are some of the sensor improvements provided by OLI (Markham et al., 2010). Glacier extractions have been performed using Landsat commissioned sensors for the last three decades, and the application of OLI sensor for glacier extraction gets more attraction due to adjusted bands, specifically the adjustment of the Near Infrared band to exclude at least 825 nm of moisture absorption (Wang et al., 2017).

2.3. Preparation of data using Google Earth Engine

Google Earth Engine (GEE), coupled with advances in data storage and computational

capabilities, is a paradigm shift in space-borne remote sensing data storage (Chastain et al., 2019). It is a cloud-based integrated platform designed to empower remote sensing scientists to utilize cloud computing resources (Gorelick et al., 2017) and provides user-friendly access through commonly used Google Mail credentials. The advent of GEE services makes it feasible to access, manipulate, and analyze tremendous volumes of open-access earth observation data. Availability of Top of Atmospheric (ToA) reflectance images of OLI for most of the sites across the globe are one of the tremendous contributions of GEE, and the ToA data is processed by extracting the calibration coefficient from image metadata by exploiting the approach discussed in (Chander et al., 2009). Moreover, the capability of GEE minimized the burden on the user's part to remove differences due to the atmospheric effects, which are needed to enhance the reliability of the data (Zhang et al., 2018).

Table 1: List of selected OLI Bands of Landsat-8

Band Number	Band Detail	Bandwidth (μm)	Spatial Resolution (m)
2	Blue	0.452-0.512	30
3	Green	0.533-0.590	30
4	Red	0.636-0.673	30
5	NIR	0.851-0.879	30
6	SWIR1	1.566-1.651	30
7	SWIR2	2.107-2.294	30

Therefore, in this study, selected spectral bands of OLI with ToA data were extracted for the required spatial and temporal extent through GEE (spatial extent of Hunza sub-basin and temporal extent of the ablation period for the year 2018 (July 15 to Oct 15)). The details of spectral characteristics of OLI bands are given in Table 1. All data extracted through GEE is sorted, filtered, and cloud-masked to the desired spatial and temporal extent. Table 2 lists the available Landsat images that are exploited in GEE from World Reference System 2 (WRS2) of the Hunza sub-basin over the ablation period of 2018.

Table 2: Available Images of Landsat 8 for the study area for spatial and temporal extent

WRS 2 (Path x Row)	**Date of Acquisition**	**Cloud Cover**
PATH 149 & ROW 35	19-JULY-2018	36.54%
	04-AUG-2018	02.57%
	20-AUG-2018	17.03%
	05-SEP-2018	55.96%
	21-SEP-2018	10.56%
	07-OCT-2018	30.65%
PATH 149 & ROW 35	19-JULY-2018	23.06%
	04-AUG-2018	10.01%
	20-AUG-2018	16.87%
	05-SEP-2018	74.94%
	21-SEP-2018	12.04%
	07-OCT-2018	30.00%

2.4. Glacier Inventories

The glacier boundaries are essential for glacier identification, and the interlinked international glacier inventories of RGI Version 5.0 provide the completeness of coverage for the glaciated region of the study area. However, the boundaries are not recommended for any change assessment and measurement of any rate of change in the glacier area (Arendt et al., 2017). Therefore, only for the completeness of the glacier area in the Hunza basin were RGI glacier boundaries covering the basin clipped from the data set. To account for the previous status of the glacier's clear ice surface, data on glacier outline for clear ice were downloaded from the ICIMOD portal (Bajracharya and Shrestha, 2011) and linked GIS files of the Glacier inventory of Pamir and Karakoram (Mölg et al., 2018b) for years 2002 to 2008 and 1998 to 2002 respectively.

2.5. Ancillary data

Shuttle Radar Topographic Mission (SRTM) data version 3.0 provides topographic information about slope, aspect, and elevation in 30-meter resolution to complement other optical imageries and vector data of glaciers (Alifu et al., 2015). Along with optical and vector data for the study area, SRTM DEM version 3.0, downloaded from USGS Earth Explorer, is used to extract the topography (slope, aspect, and elevation) of the study area. The SRTM DEM Version 3.0 data was downloaded from USGS Earth Explorer. Voids occurred in the earlier versions of SRTM DEM, which are filled later on version 3.0, using interpolation and hole-filling algorithms through ancillary. Now, the elevation model is provided with an overall accuracy of ±16 m with a 95% confidence level (Elkhrachy, 2018). Additional vector data for the HKH region and UIB boundary was downloaded from the regional database of ICIMOD was acquired from Shabeh ul (2016).

All datasets were defined with the coordinate system of the World Geodetic System (WGS 84) and projected with Universal Transverse Mercator (UTM) Zone 43 North. Furthermore, three topographic parameters (elevation, slope, and aspect) were generated along with the extraction of NDSI and eight different combinations of band ratios (Red/SWIR1, Red/SWIR2, Blue/SWIR1, Blue/SWIR2, Green/SWIR1, Green/SWIR2, NIR/SWIR1 and NIR/SWIR2) from the multispectral data of OLI. Ground Control points (GCPs) were collected over the clear ice surface by exposing clear ice using RGB composite Figure 3 of Landsat-8 with a combination of NIR, Red, and Green (Khan et al., 2015).

3. Methodology

The research framework for this study is illustrated in Figure 2; the three major steps are further explained in the subsequent sections.

3.1. GEE Pre-processing

Java scripted the Code Editor environment of GEE, which was used for the collection of Tier 1 calibrated TOA reflectance of selected OLI optical bands through Engine Snippet. It is important to note that OLI raw scenes in GEE contain digital numbers in radiance scale, and the conversion to ToA is a linear transformation that accounts for solar elevation and seasonally variable Earth-Sun distance. This transformation approach, as discussed in (Chander et al., 2009) and handled by ee.Algorithms.Landsat. TOA. In calculating the extent of clear ice in optical remote sensing data, cloud cover is always an issue for mountainous glacier monitoring (Ke et al., 2015). To avoid the impact of cloud cover for glacier extraction, the cloud Score algorithm was used for OLI data with a threshold value of 25% through ee .Algorithms.Landsat. SimpleCloudScoremethod.

3.2. Ground Control Points Selection and Percentile Analysis

Ground Control Points (GCPs) are placed on the glaciated surface of the study area, referred to in Figure 3 based on exposed clear ice in visual interpretation of False color composite (FCC) of OLI with bands combination of SWIR, NIR, band 3 for red, blue and green channels respectively. This particular band combination enables clear discrimination between snow, ice, and other

features (Kaab et al., 2014). GCPs needed to predict the threshold value for specific band ratios as training classes to classify clear ice using random forest analysis. Band statistics are used to estimate the minimum value using the percentile analysis by mapping the GCPs, and it shows the percentage of mapped pixels less than the calculated value. A total of 1233 well-distributed GCPS over exposed clear ice was used as a training set for the total population of pixels showing the clear ice.

Minimum threshold values for each Band ratios (BR) training classes were calculated using GCPs. The required training class was characterized with pixels exceeding a given threshold value, and was defined as the minimum percentile of GCPs through Sentinel Application Platform, i.e. 5th percentile of all GCPs over a specific BR (meaning 5% of GCPs placed over pixels of training class equal to or less than a threshold value).

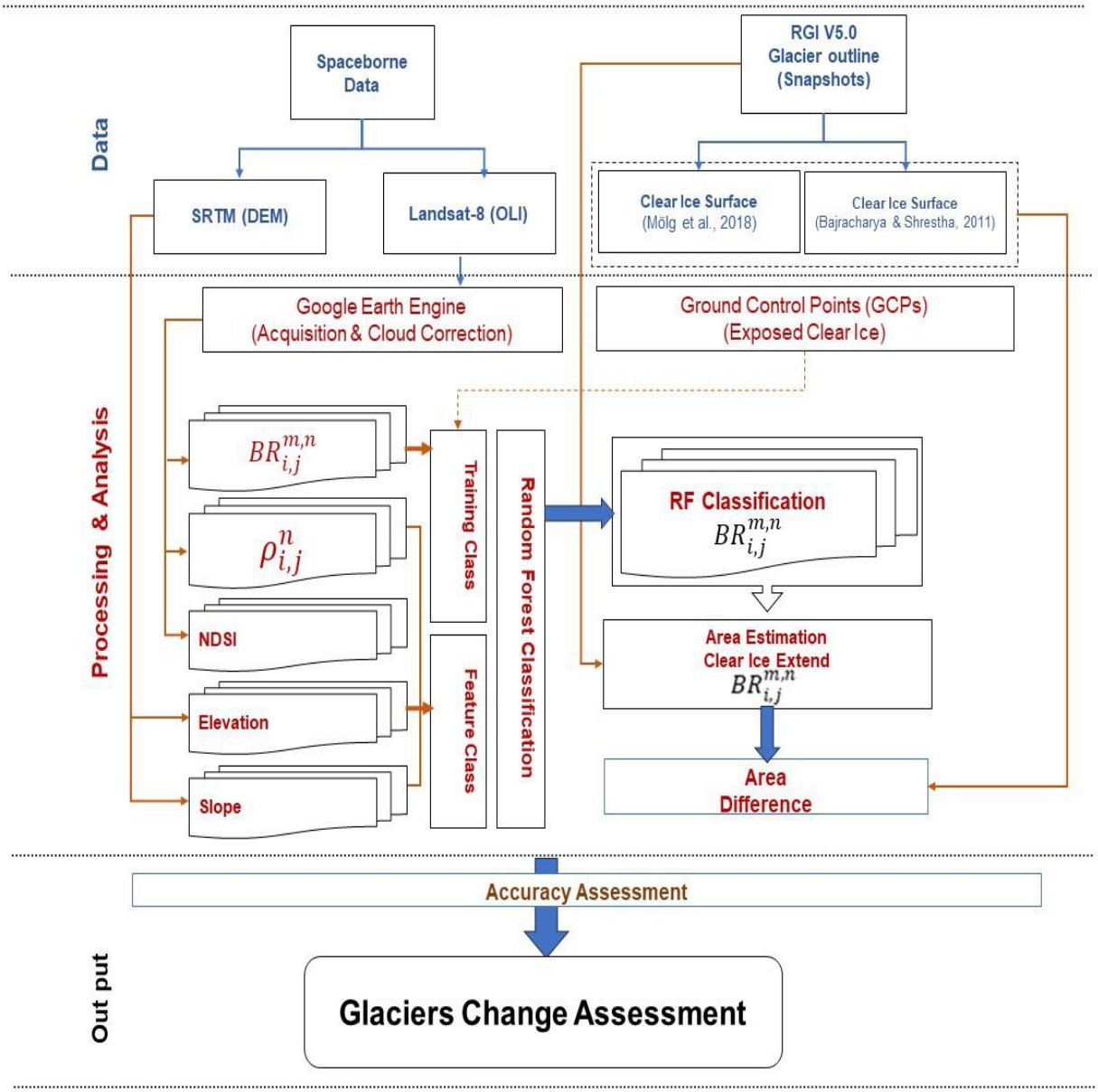

Fig. 2. Flow chart for extraction of clear ice from Hunza sub-basin

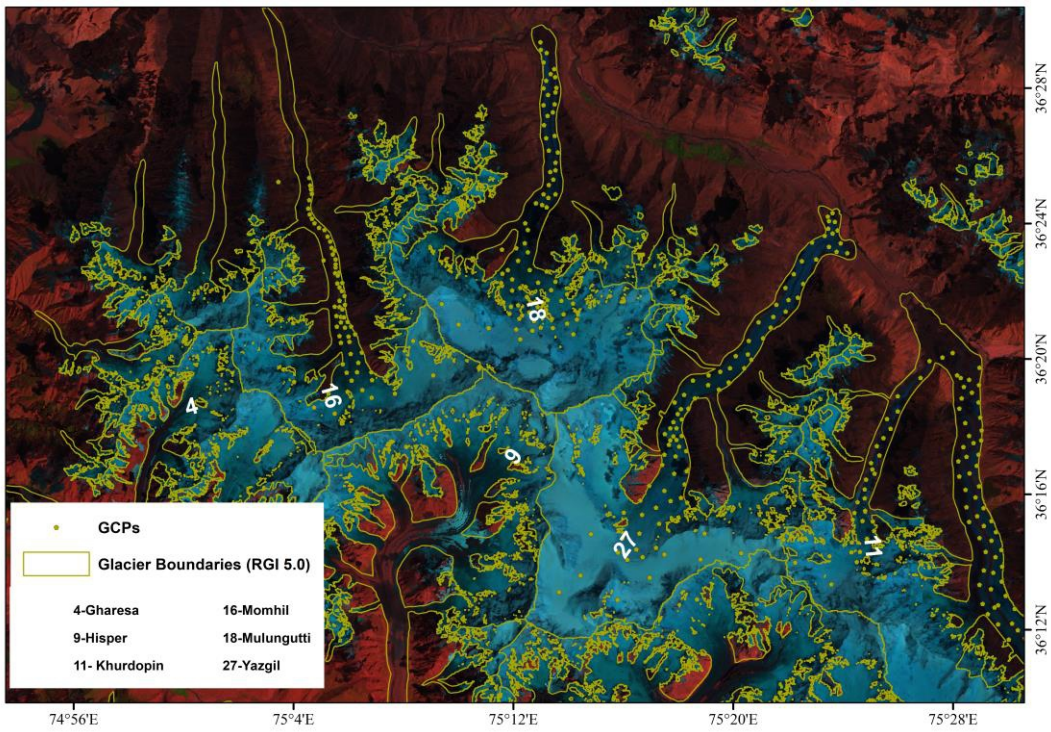

Fig. 3. Example of sample glaciers exposed with clear ice surface in RGB composite of OLI with band combination of SWIR, NIR & Green over ablation period of year 2018. Yellow dots are GCPs over clear ice used as training data.

3.3. Band Ratios and Snow index

Multispectral data during the ablation period (July 15–Oct 15, 2018) with a minimum cloud of 25% acquired from the GEE environment, used to develop 08 BRs; the visible and NIR bands are used with SWIR bands to extract the ratio reflectance over the clear ice surface. This combination of bands has been used extensively in previous studies due to their good performance in discriminating clear ice surfaces from the rest of the earth's features (Paul et al., 2016). The Normalized Difference Snow Index (NDSI) was calculated by exploiting universally accepted visible and SWIR bands (Hall and Riggs, 2011) using raster calculation based on the equations (1 & 2). However, NDSI is less favorable than the simple band ratio.

$$BR_{i,j}^{m,n} = \frac{\rho_{i,j}^{m}}{\rho_{i,j}^{n}} \qquad (1)$$

where $BR_{i,j}^{m,n}$ is Simple Band Ratio of m^{th} and n^{th} band for of OLI for the pixel location *(i,j)*, $\rho_{i,j}^{m}$ and $\rho_{i,j}^{n}$ is ToA reflectance of respective bands.

$$NDSI_{i,j} = \frac{\rho_{VIS} - \rho_{SWIR}}{\rho_{VIS} + \rho_{SWIR}} \qquad (2)$$

where $NDSI_{i,j}$ is Normalize Differential Snow Index of pixel location *(i,j)* for OLI, ρ_{VIS} ToA reflectance of visible band and ρ_{SWIR} is ToA of Short-Wave Infrared band.

3.4. Topography Extraction

The Digital Elevation Model is the most common dataset to extract the glacier topography. Among several elevation models, SRTM DEM has a significantly high performance in extracting correct elevations in different altitude regimes (Tobias et al., 2005). Slope information extracted from DEM is usually exploited for the observation of clear ice of glacier surfaces (Linghong et al., 2016). Slope and elevation information extracted through spatial analysis from commonly used SRTM DEM with an absolute vertical accuracy of ±16 m and a horizontal positional accuracy

of about ±20m (Rabus et al., 2003).

3.5. Training and Feature Classes

Training classes were created through the utilization of threshold values, taking advantage of ground control points (GCPs). Meanwhile, the RF classifier was employed, with a training class and feature class serving as inputs. The resultant output of RF predicted classification consisted of two classes (clear ice and non-glaciated surface). In the final stage of this step, the area percent difference was computed by comparing the extracted area from the data of historical inventories and the predicted area calculated through RF analysis. The last step in the framework is an illustration of the current extent for clear ice and area difference percentage for change analysis. This gives the understanding of glacier stability; however, accuracy assessment, as an intermediate step, is carried out prior to providing the final output of the study.

3.6. Selection of Training and Feature Classes

BRs as training class of RF classifier require minimum threshold values. These minimum threshold values were determined through statistics of percentile analysis (Aczel and Sounderpandian, 2017) of 1233 GCPs placed over a clear ice surface. In the context of the RF classifier, feature classes also refer to the sets of input variables. Therefore, seven spectral bands of OLI, spectral indices (NDSIs), rest of BRs, and topographic parameter (slope, elevation, and aspect) are used as feature classes for RF to learn and predict the clear ice surface accordingly. BRs and NDSIs were calculated using expressions (1) and (2), respectively based on OLI data and the topographic parameters extracted from SRTM.

3.7. Random Forest Classification

The study focused on extracting clear ice surfaces of the mountainous glaciated region using RF classifier. RF, as a machine learning algorithm generates binary trees for the training classifier and thereafter votes for the popular class through aggregation (Breiman, 2001). Training samples and features class are the main ingredients of RF, while the decision tree is constructed from the input (Ali et al., 2012). Zhang et al. (2019) provided a general framework of RF classifier. In this study, RF classifier has been applied using supervised classification in the SNAP tool environment (Onojeghuo et al., 2018). The classifier was trained on raster data with a training sample set limited to the default value of 5000. An iteration of 10 decision trees was used for the minimum error of classification. In this process, the minimum value for raster training class was fixed based on the extracted values obtained through percentile analysis. In the final stage of RF construction, selected features, as described above were used along with training samples. The output of the process provided the classes whereas root means square error (RMSE) parameter was used as an evaluator to measure the training class (Guan et al., 2012a).

3.8. Accuracy Assessment

The accuracy of results derived with remotely sensed data is usually calculated in terms of the producer's, user's accuracy, and overall accuracy, and these parameters are examined through the confusion matrix (Shao and Wu, 2008). Hence, results for clear ice delineation extracted using RF classifier in the study were evaluated using these parameters, and the Kappa coefficient was also used as a parameter of accuracy assessment, calculated as in Equation (3) (Congalton, 2001).

$$\text{Kappa} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} * x_{i+})}{N^2 - \sum_{i=1}^{r}(x_{i+} * x_{i+})} \quad (3)$$

where r is the number of rows in the error matrix, x_{ii} is the number of observations in row i and column i, x_{i+} and x_{i+} are the marginal totals of row i and column i, respectively, and N is the total number of observations.

4. Results

4.1. RF Classification Results

At the first stage of RF classification, stability analysis was carried out for every training class as shown in Figure 4. For every individual training class, four aspects of feature classes: spectral features (07), spectral indices

(02), all other BRs (07), and topography (slope & elevation) captured as selected feature. The stability of RF is demonstrated by Root Mean Square Error (RMSE), and minimum stable RMSE appears after 10 trees for each training class. Whereas the range varies from 0.4 to 0.2, and the minimum value of 0.29 is noted for the $BR^{4,6}$. The feature ranking is also displayed in Figure 4, which demonstrates the importance of a feature for the training class (Guan et al., 2012b). According to ranking profile, all the selected features have different behavior for different training classes; however, the ranking of NDSI seems to be consistent for all the training classes which induce the importance of snow indices for the application of RF and confirms that the clear ice extraction primarily relies on NDSI (Bajracharya and Shrestha, 2011).

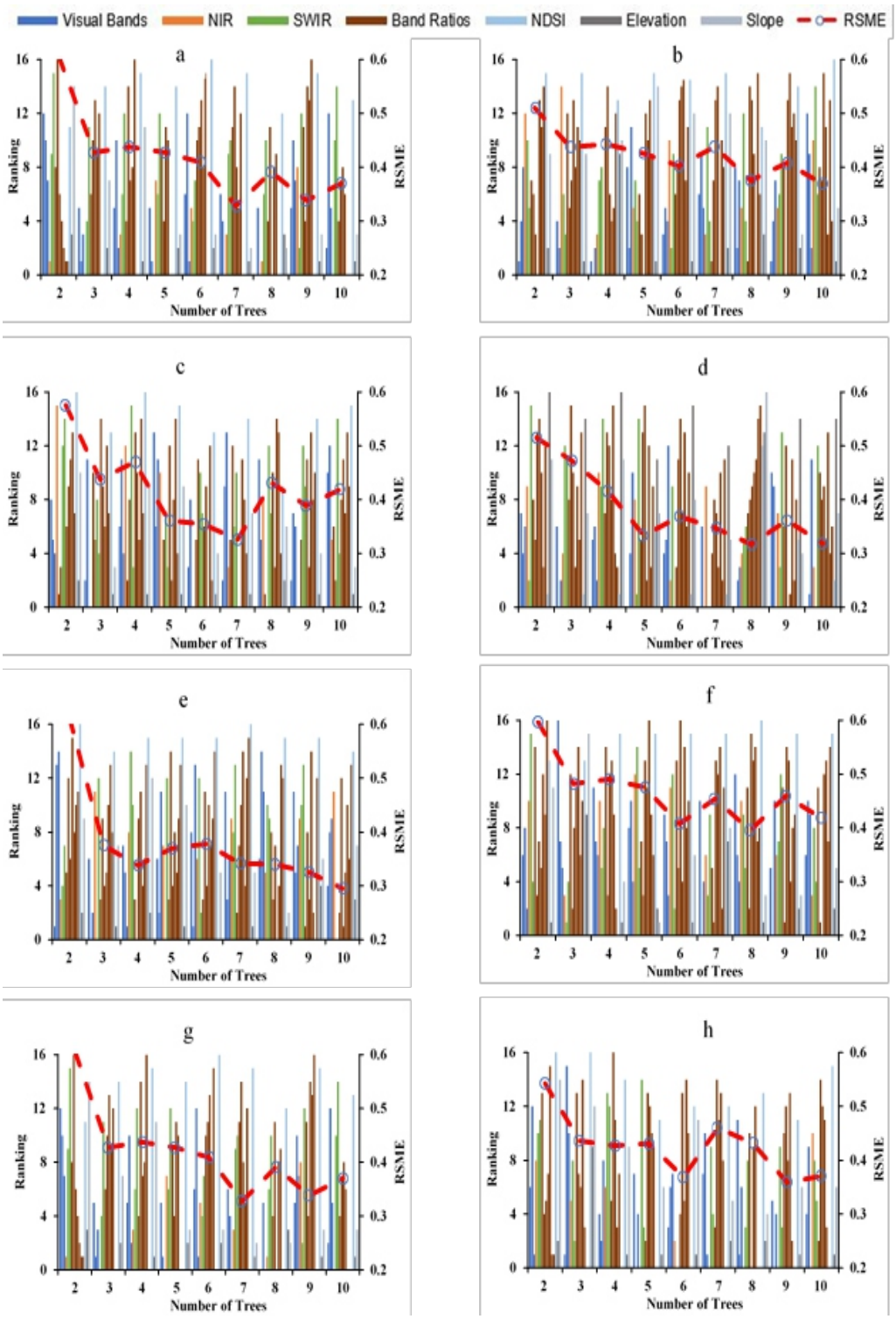

Fig. 4. Stability analysis of RMSE for RF classification and ranking of feature classes for $BR_{i.j}^{m,n}$ as training class where BR is band ration of m and n band of OLI for the pixel location i,j (a) $BR_{i.j}^{2,6}$ (b) $BR_{i.j}^{3,6}$ (c) $BR_{i.j}^{4,6}$ (d) $BR_{i.j}^{2,7}$ (e) $BR_{i.j}^{3,7}$ (f) $BR_{i.j}^{4,7}$ (g) $BR_{i.j}^{5,6}$ (h) $BR_{i.j}^{5,7}$ Figure 5a shows spatial distribution of clear ice over the glaciated region of the Hunza basin for all the training classes, continuous pattern of snow and ice visible in the higher elevations of the study area which shows an agreement with Hewitt (2005). However, the terminal end of clear ice has variation as compared to the previous inventories for some of the glaciers, as shown in Figure 5b, and it is due to the variation in the snow coverage as reported by Minora et al. (2016). Still there are some misclassified areas for each of RF results which needed to be removed manually. Figure 5c shows an area of clear ice predicted through RF using a different combination of BR and area difference percentage from the previous inventory for the year 2000 ± 2 years, which varies from 2% to 10% and indicates the stability in glaciers of the basin (Bolch et al., 2017).

4.2. Glaciers Change Assessment

Final results were obtained by mapping the spatial distribution of predicted results with clear ice distribution presented for years from 1998-2002 by Mölg et al. (2018a) and it is observed that during last two decades i.e. from 2000 to 2018 the overall clear ice glacial extent has not faced a significant decline as presented in Figure 5c and area difference percentage varies (i.e. 0-10%) which is in contrast to the results provided by Hewitt (2005). However, the outcome of change assessment for 27 renowned glaciers having (area > 5.0 km2 have slightly different scenarios as compared to overall ice surface decline. Figure 6 shows the change assessment for the glaciers regarding area predicted using different BRs. It is interestingly noted that the glacier with small areas has observed a positive area difference percentage as compared to the glacier with a large extent of clear ice. This prediction supports one of the findings of the study by Hewitt (2005). Apart from the findings for the stability of overall glaciers, the terminal end of clear ice has significant variations, if compared with the previously developed inventories of years 1998 to 2000 and 2003 to 2008 as presented in Figure 7 which highly supports the vulnerability of glacier extents in the lower elevations in the basin (Khan et al., 2018).

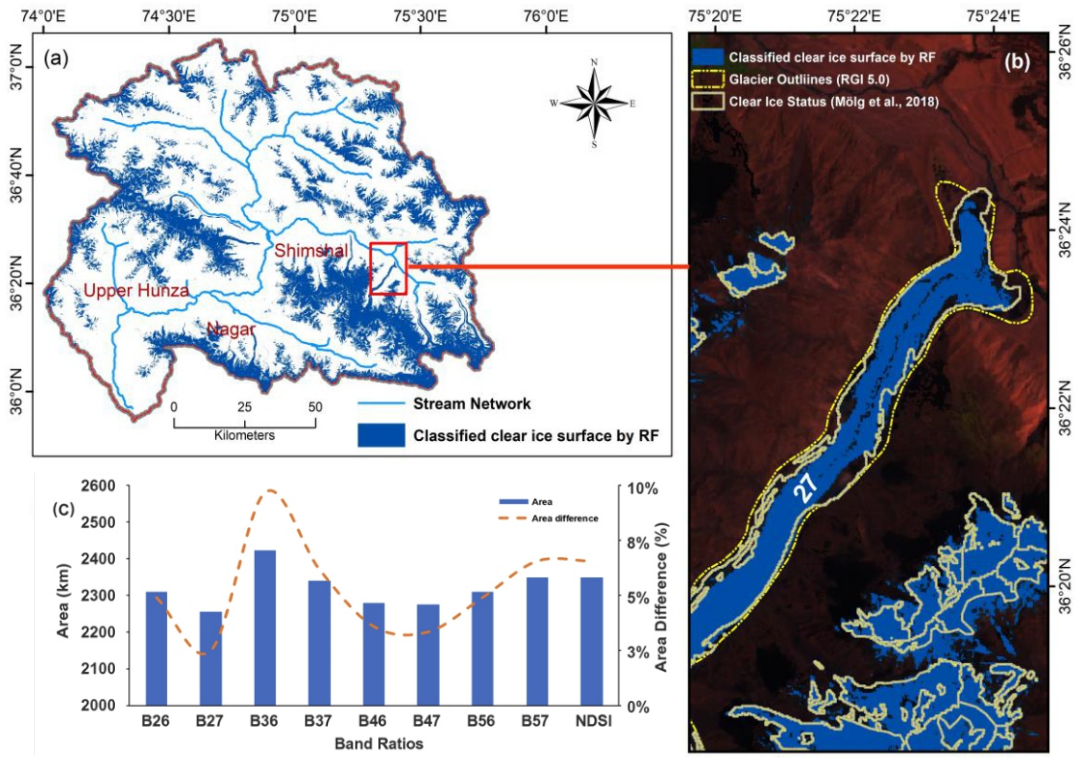

Fig. 5. Primary results of RF Classifier (a) Spatial distribution of clear ice (b) Example of variation in Terminal End of clear ice (c) Area difference percentage of BR and inventory of -2000+

Fig. 6. Corresponding Area Difference Percentage of predicted clear ice surface with area predicted for years 1998 to 2002 by Moleg (2018).

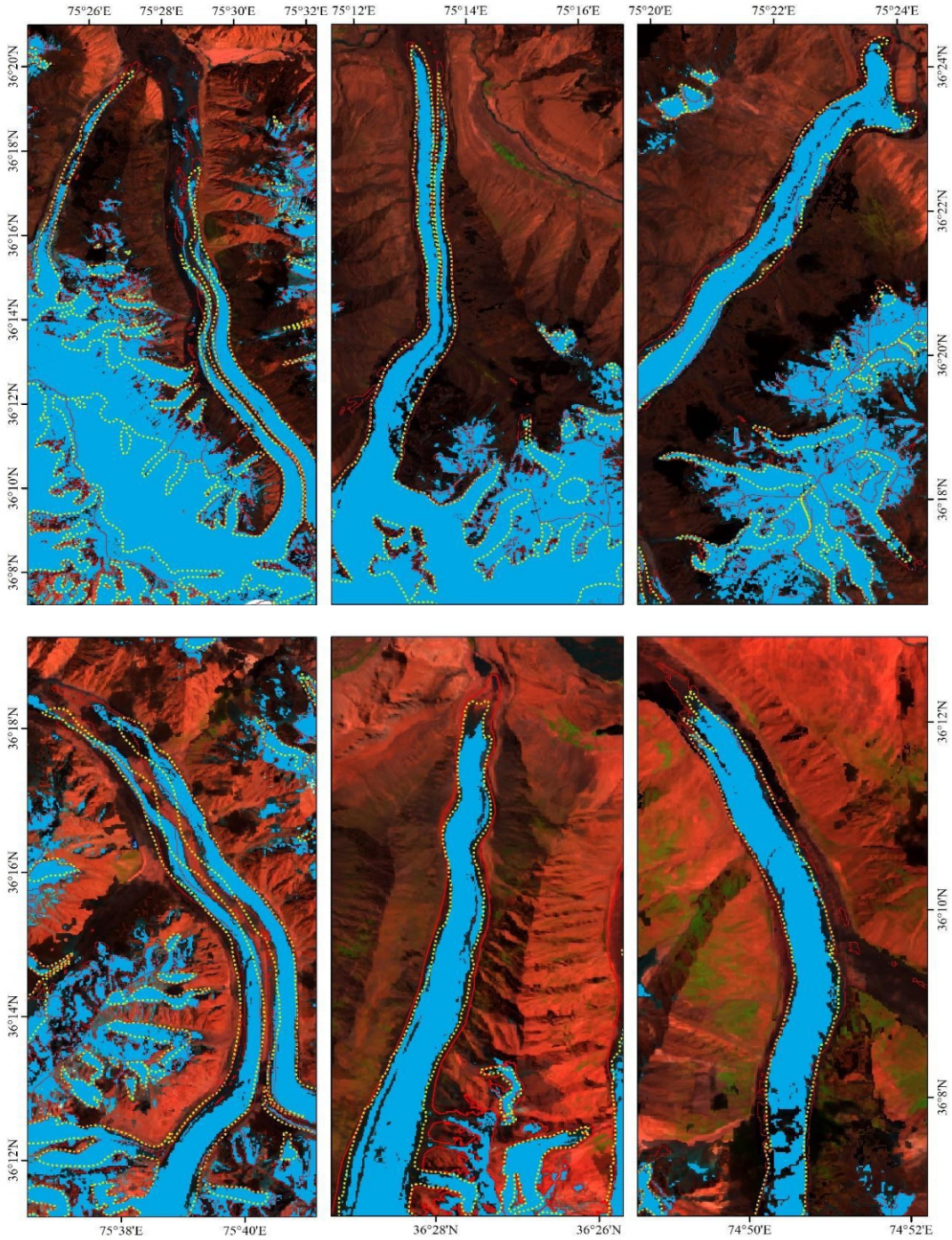

Fig. 7. Example of some glaciers with variation in clear ice terminus. (a) Khurdopin Glacier (b) Mulungutti Glacier (c) Yazgil Glacier (d) Virjerab Glacier (e) Pasu Glacier (f) Barpu Glacier

4.3. Accuracy Assessment

Accuracy of each BR to extract clear ice by exploiting RF classifier was estimated using the 402 selected reference samples. The sample points are then characterized in confusion matrix as given in Table 3 to summarize the accuracy of the results. From the list of the BR in Overall accuracy for all the bands lies between 82% and 83% with Kappa values ranging from 0.64 to 0.65 when used as training class in RF classifier. While the user's accuracy for all BRs ranged from 81%to 83% is less than that of producer's accuracy ranging from 83% to 87%. $BR^{5,6}$ exhibits maximum producer's accuracy of 87% in contrast to the accuracy of 83 % for, $BR^{3,6}, BR^{4,6} and\ BR^{5,6}$.

Table 3: Confusion Matrix Results for the classification of clear ice (Class-1) and non-glaciated surface (Class-2) using Band Ratio as raster training dataset for OLI Images to extract clear ice surface in Hunza sub basin

Training Class (Band Ratio)	Class	1	2	Total	Commission	Omission	Accuracy			Kappa
							Producer	User	Overall	
$BR_{ij}^{2,6}$	1	170	39	209	19%	15%	85%	81%	83%	0.65
	2	31	162	193	15 %	19%	81%	84%		
$BR_{ij}^{2,7}$	1	169	39	208	19%	14%	86%	81%	82%	0.65
	2	32	162	194	15%	20%	80%	84%		
$BR_{ij}^{3,6}$	1	173	41	214	19%	14%	86%	81%	83%	0.66
	2	28	160	188	15%	20%	80%	85%		
$BR_{ij}^{3,7}$	1	166	36	202	18%	17%	83%	82%	82%	0.65
	2	35	165	200	14%	18%	82%	83%		
$BR_{ij}^{4,6}$	1	167	37	204	18%	17%	83%	82%	82%	0.65
	2	34	164	198	14%	18%	82%	83%		
$BR_{ij}^{4,7}$	1	167	37	204	18%	17%	83%	83%	82%	0.65
	2	34	164	198	14%	18%	82%	82%		
BR_{ij}^{56}	1	174	42	216	19%	13%	87%	81%	83%	0.66
	2	27	159	186	15%	21%	79%	85%		
$BR_{ij}^{5,7}$	1	172	42	214	20%	14%	86%	80%	82%	0.65
	2	29	159	188	15%	21%	79%	85%		

5. Discussion

The stability of mountain glaciers is expressed as an indicator of climate change and its varying effects on the global climate (Haeberli et al., 2007). The general trend indicates a retreat in glaciers on a global scale (Zemp et al., 2015), whereas many of studies highlight heterogenous trend in glaciers of Karakoram range (Gardelle et al., 2012). This study builds on existing research to quantify the clear ice extent using ML approach by leveraging the RF classification algorithm. The study aims to overcome the limitations of traditional methods and provide more reliable glacier stability statistics. However, the challenge of required number of samples to train the model affects the accuracy of the model. The results contribute to the ongoing discussion about the unique behavior of Karakoram glaciers and provide an updated inventory reflecting their condition in 2018.

5.1. Ice Extend Stability of Glaciers

OLI data serves as the foundation dataset for the study, which focuses on extracting the clear ice extent of Karakoram glaciers using an advanced machine learning (ML) classification technique. The OLI Data has also been utilized by several other glacier studies. In contrast to the current study, which uses chosen glaciers over 2000 meters above sea level, Rastner et al. (2017) used OLI (Level 1T) data in the Russian Artic region, which is characterized by small valley glaciers in an elevation range of 1300–1600 meters above sea level. Li et al. (2023) conducted a study on the glaciers of Karakoram using images from SPOT, which has a 60 km x 60 km scene coverage area constraint. When compared to OLI images, such a data source is not favorable across a large area extent (Franklin et al., 2011). The findings in this study shed light on the stability of these high glaciers of Karakoram during the previous ten years and help to measure their current clear ice extent. Although it is widely acknowledged that glaciers are receding globally, pertinent research suggests that the glaciers in the Karakoram range have exhibited a degree of relative stability (Li et al., 2023). According to this study's findings, which are consistent with earlier research of Minora et al. (2016), the area difference percentages for small and large glaciers are positive and insignificant, respectively, ranging from 0 to 10% (over a ten-year period).

5.2. Accuracy of Random Forest

The study utilizes the Random Forest (RF) classification algorithm to address the limitations of the simple band ratio method commonly used for ice surface extraction in glaciated areas (Zhang et al., 2019). A unique approach in this study is training the classifier on raster data of the desired band ratio with the threshold value. This contrasts with the more common practice of using vector data for classifier training by Khan et al. (2020), Talukdar et al. (2020) and others. By employing raster data, the study avoids the labor-intensive task of labeling training data. The accuracy assessment is explained with the reference of most popular metrics i.e. overall accuracy and kappa coefficient however, rest of the accuracy parameters are also recorded. Whereas most studies like, Thanh Noi and Kappas (2017) and Lu et al. (2020) stick with only few parameters for explanation of accuracy assessment. Further, the algorithm's learning ability in this study capping the maximum overall accuracy at 83% is generally acceptable to validate the results (Ismail and Jusoff, 2008). The achieved accuracy in this study is lower than that of some other glacier-related studies (Zhang et al., 2019) which is because the limitations of the study. Foremost limitation is setting the training sample size to a default value of 5000 pixels which presented a significant impact. Furthermore, machine learning is the sensitivity to seasonal and atmospheric variations, such as cloud cover and snow whereas, present study processed 12 scenes (table 2) and found only 5 scenes with less 30 percent cloud cover/ Despite these limitation, the study significantly provides an updated inventory of Karakoram glaciers with a timestamp of 2018, contributing valuable data to the existing referenced glacier inventories established by Bajracharya and Shrestha (2011), Mölg et al. (2018a) and Minora et al. (2016).

6. Conclusion

This study has explored the extraction of clear ice extent in most heterogenous glaciated sub-basin of Hunza in UIB of HKH to discuss the stability of glaciers in the basin. The improved mountain glacier monitoring ability of Landsat-8 with its 12-bit quantized multispectral data has been utilized, and the dataset has been pre-processed through GEE. RF classifier as a machine learning technique for automatic extraction of clear ice has been exploited by considering BRs and topographic information from Landsat-8 OLI and STRM DEM data, respectively. The results demonstrated that the extraction of clear ice has an acceptable overall accuracy and kappa coefficient, whereas the area difference percentage between predicted results and previous glacier inventories for clear ice surface is presented for different combinations of BRs. The results indicate that glacier ice is the basin has not significantly changed overall for the last 20 years, differences are, however, observed in 28 renowned glaciers (area > 5.0 km2), confirming higher variation in some of the specific glaciers in the area. The most interesting outcome of the results is that the change in clear ice extent is inversely proportional to the area of the glaciers, which means that the smaller glacier has much more variations as compared to the bigger glaciers. However, the change in the position of the terminal end of clear ice in most of the glaciers verifies the decline of clear ice in lower elevations, which is a common phenomenon in the basin. Furthermore, variations in the predicted area through different BRs also confirm the role of selected bands for band ratios, and in this connection, visible (band2) and SWIR wavelengths show the stability of glaciers with minimized variation, i.e., 0.3 % of the area difference percentage. The accuracy of classification results lies in between 82% and 83% for BRs with a kappa value of 0.5 to 0.6, showing slightly different behaviors of BRs to classify clear ice extent. However, the lower user accuracy (80%) in comparison of its producer accuracy (86%) justifies the least contribution of the NIR bands for the classification.

Authors' Contribution

Syed Najam ul Hassan: Writing of original draft–Data Collection, Analysis Visualization, Validation, Resources, Project administration, Methodology. Mohd Nadzri Md. Reba: Writing–review & editing, Supervision, Software, Resources, Project

administration, Methodology, Investigation. Aftab Ahmed Khan: Writing-review & editing, ML software Management, Validation of ML Models.

References

Aczel, A. D., & Sounderpandian, J. (2017). Introduction and Descriptive Statistics. In *Complete business statistics*. Boston, MA: Irwin/McGraw Hill.

Ali, J., Khan, R., Ahmad, N., & Maqsood, I. (2012). Random forests and decision trees. *International Journal of Computer Science Issues (IJCSI)*, *9*(5), 272.

Alifu, H., Tateishi, R., & Johnson, B. (2015). A new band ratio technique for mapping debris-covered glaciers using Landsat imagery and a digital elevation model. *International Journal of Remote Sensing*. *36*(8), 2063–2075. https://doi.org/10.10808/2150704X.2015.1034886

Arendt, A., Bliss, A., Bolch, T., Cogley, J. G., Gardner, A., Hagen, J.-O., Hock, R., Huss, M., Kaser, G., & Kienholz, C. (2017). *Randolph Glacier inventory–A dataset of Global glacier outlines: Version 6.0: Technical report, Global land ice measurements from space*.

Baig, S. U., Khan, H., & Din, A. (2018). Spatio-temporal analysis of glacial ice area distribution of Hunza River Basin, Karakoram region of Pakistan. *Hydrolog ical Processes*, *32*(10), 1491-1501. https://doi.org/10.1002/hyp.11508

Baig, S. U., Tahir, A. A., Din, A., & Khan, H. (2018). Hypsometric properties of mountain landscape of Hunza River Basin of the Karakoram Himalaya. *Journal of Mountain Sciences*. *15*, 1881–1891. https://doi.org/https://doi.org/10.1007/s11629-018-4849-x

Bajracharya, S. R., & Shrestha, B. R. (2011). *The status of glaciers in the Hindu Kush-Himalayan region*. International Centre for Integrated Mountain Development (ICIMOD).

Bolch, T., Kulkarni, A., Kaab, A., Huggel, C., Paul, F., Cogley, J. G., Frey, H., Kargel, J. S., Fujita, K., Scheel, M., Bajracharya, S., & Stoffel, M. (2012). The State and Fate of Himalayan Glaciers. *Science*, *336*(6079), 310-314.

Bolch, T., Pieczonka, T., Mukherjee, K., & Shea, J. (2017). Brief communication: Glaciers in the Hunza catchment (Karakoram) have been nearly in balance since the 1970s. *Cryosphere*, *11*(1). https://doi.org/10.5194/tc-11-531-2017

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5-32. https://doi.org/10.1023/A:1010933404324

Chander, G., Markham, B. L., & Helder, D. L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, *113*(5), 893-903.https://doi.org/10.1016/J.RSE.2009.01.007

Chastain, R., Housman, I., Goldstein, J., & Finco, M. (2019). Empirical cross sensor comparison of Sentinel-2A and 2B MSI, Landsat-8 OLI, and Landsat-7 ETM+ top of atmosphere spectral characteristics over the conterminous United States. *Remote Sensing of Environment*, *221*, 274-285. https://doi.org/10.1016/J.RSE.2018.11.012

Congalton, R. G. (2001). Accuracy assessment and validation of remotely sensed and other spatial information. I*nternational Journal of Wildland Fire*, *10*(4), 321–328.

Cook, E. R., Palmer, J. G., Ahmed, M., Woodhouse, C. A., Fenwick, P., Zafar, M. U., Wahab, M., & Khan, N. (2013). Five centuries of Upper Indus River flow from tree rings. *Journal of Hydrology*, *486*, 365-375. https://doi.org/10.1016/J.JHYDROL.2013.02.004

Dorothy K Hall, & George A Riggs. (2011). Normalized-Difference Snow Index (NDSI). In *Encyclopedia of Snow, Ice and Glaciers*. https://doi.org/10.1007/978-90-481-2642-2

Dyhrenfurth, G. O. (1955). *To the Third Pole: The History of the High Himalaya*. W. Laurie. https://books.google.ie/books?id=NcO1AAAAIAAJ

Elkhrachy, I. (2018). Vertical accuracy assessment for SRTM and ASTER Digital Elevation Models: A case study of Najran

city, Saudi Arabia. *Ain Shams Engineering Joournal*, *9*(4), 1807-1817. https://doi.org/10.1016/J.ASEJ.2017.01.007

Franklin, S. E., He, Y., Pape, A., Guo, X., & McDermid, G. J. (2011). Landsat-comparable land cover maps using ASTER and SPOT images: a case study for large-area mapping programmes. *International Journal of Remote Sensing*, *32*(8), 2185-2205. https://doi.org/10.1080/01431161003674642

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. https://doi.org/10.1038/ngeo1450

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. https://doi.org/10.1016/J.RSE.2017.06.031

Guan, H., Yu, J., Li, J., & Luo, L. (2012). Random forests-based feature selection for land-use classification using lidar data and ortho-imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIX-B7, 203-208. doi: 10.5194/isprsarchives-XXXIX-B7-203-2012

Guo, W., Liu, S., Xu, J., Wu, L., Shangguan, D., Yao, X., Wei, J., Bao, W., Yu, P., Liu, Q., & Jiang, Z. (2015). The second Chinese glacier inventory: Data, methods and results. *Journal of Glaciology*, *61*(226), 357-372. https://doi.org/10.3189/2015JoG14J209.

Haeberli, W., Hoelzle, M., Paul, F., & Zemp, M. (2007). Integrated monitoring of mountain glaciers as key indicators of global climate change: the European Alps. *Annals of Glaciology*, *46*, 150-160. https://doi.org.10.3189/172756407782871512

Haireti, A., Ryutaro, T., & Brian, J. (2015). A new band ratio technique for mapping debris-covered glaciers using Landsat imagery and a digital elevation model. *International Journal of Remote Sensing*, *36*(8), 2063-2075. https://doi.org/10.1080/2150704X.2015.1034886

Hewitt, K. (2005). The Karakoram Anomaly? Glacier Expansion and the "Elevation Effect", Karakoram Himalaya. *Mountain Research and Development*, *25*(4), 332-340. https://doi.org/10.1659/0276-4741(2005)025[0332:TKAGEA]2.0.CO;2,

Hussain, D., & Khan, A. A. (2020). Machine learning techniques for monthly river flow forecasting of Hunza River, Pakistan. *Earth Science Informatics*, *13*, 939–949. https://doi.org/10.1007/s12145-020-00450-z

ICIMOD. (2010). *Glacier mapping and monitoring tools and Techniques*. Himal Doc. https://lib.icimod.org/record/26904

Immerzeel, W. W., Wanders, N., Lutz, A. F., Shea, J. M., & Bierkens, M. F. P. (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. https://doi.org/10.5194/HESS-19-4673-2015

IPCC. (2013). Climate Change 2013 The Physical Science Basis: *Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*.

Ismail, M. H., & Jusoff, K. (2008). Satellite data classification accuracy assessment based from reference dataset. I*nternational Journal of Geological and Environmental Engineering*, *2*(3), 23–29.

Kaab, A., Bolch, T., Casey, K., Heid, T., Kargel, J. S., Leonard, G. J., Paul, F., & Raup, B. H. (2014). Glacier mapping and monitoring using multispectral data. In *Global Land Ice Measurements From Space* (pp. 75–112). Springer.

Ke, L., Ding, X., & Song, C. (2015). Heterogeneous changes of glaciers over the western Kunlun Mountains based on ICESat and Landsat-8 derived glacier inventory. *Remote Sensing of Enviroment*, *168*, 13-23. https://doi.org/10.1016/j.rse.2015.06.019

Khan, A. A., Jamil, A., Hussain, D., Taj, M.,

Jabeen, G., & Malik, M. K. (2020). Machine-Learning Algorithms for Mapping Debris-Covered Glaciers: The Hunza Basin Case Study. *IEEE Access*, *8*, 12725-12734. https://doi.org/10.1109/ACCESS.2020.2965768

Khan, A., Naz, B. S., & Bowling, L. C. (2015). Separating snow, clean and debris covered ice in the Upper Indus Basin, Hindukush-Karakoram-Himalayas, using Landsat images between 1998 and 2002. *Journal of Hydrology*, *521*, 46-64. https://doi.org/10.1016/j.jhydrol.2014.11.048

Kraaijenbrink, P. D. A., Bierkens, M. F. P., Lutz, A. F., & Immerzeel, W. W. (2017). Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers. *Nature*, *549*(7671), 257-260. https://doi.org/10.1038/nature23878

Krimmel, R, M., & Meier, M, F. (1975). Glacier Applications of Erts Images. J*ournal of Glaciology*, *15*(73), 391–402. https://doi.org/10.3189/S00221430000345 0X

Krishna, G., Sahoo, R. N., Pradhan, S., Ahmad, T., & Sahoo, P. M. (2018). Hyperspectral satellite data analysis for pure pixels extraction and evaluation of advanced classifier algorithms for LULC classification. *Earth Science Informatics*, *11*(2), 159-170. https://doi.org/10.1007/s12145-017-0324-4

Lary, D. J., Alavi, A. H., Gandomi, A. H., & Walker, A. L. (2016). Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, *7*(1), 3–10. https://doi.org/10.1016/j.gsf.2015.07.003

Li, J., Sun, M., Yao, X., Duan, H., Zhang, C., Wang, S., Niu, S., & Yan, X. (2023). A Review of Karakoram Glacier Anomalies in High Mountains Asia. *Water*, *15*(18), 3215. https://doi.org/10.3390/w15183215

Linghong, K., Ding, Xi., Zjang, L., Hu, J., Shum, C. K., & Lu, Z. (2016). Compiling a new glacier inventory for southeastern Qinghai–Tibet Plateau from Landsat and PALSAR data. *Journal of Glaciology*, *62*(233), 579-592. https://doi.org/10.1017/jog.2016.58

Lu, Y., Zhang, Z., & Huang, D. (2020). Glacier mapping based on random forest algorithm: A case study over the eastern Pamir. *Water (Switzerland)*, *12*(11), 1–25. https://doi.org/10.3390/w12113231

Markham, B. L., Dabney, P. W., Murphy-Morris, J. E., Pedelty, J. A., Knight, E. J., Kvaran, G., & Barsi, J. A. (2010). The landsat data continuity mission operational land imager (OLI) radiometric calibration. *2010 IEEE International Geoscience and Remote Sensing Symposium*, 2283–2286. https://doi.org/10.1109/IGARSS.2010.5652789

Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, *39*(9), 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

Minora, U., Bocchiola, D., D'Agata, C., Maragno, D., Mayer, C., Lambrecht, A., Vuillermoz, E., Senese, A., Compostella, C., Smiraglia, C., & Diolaiuti, G. A. (2016). Glacier area stability in the Central Karakoram National Park (Pakistan) in 2001–2010: The "Karakoram Anomaly" in the spotlight. Progress in Physical Geography. *Earth and Environment*, *40*(5), 629-660. https://doi.org/10.1177/0309133316643926

Missions, S., & Catalogue, S. M. (2022). *Landsat 9*. Retrieved April.

Mölg, N., Bolch, T., Rastner, P., Strozzi, T., & Paul, F. (2018a). A consistent glacier inventory for Karakoram and Pamir derived from Landsat data: Distribution of debris cover and mapping challenges. *Earth System Science Data*, *10*(4), 1807–1827. https://doi.org/10.5194/essd-10-1807-2018

Mölg, N., Bolch, T., Rastner, P., Strozzi, T., & Paul, F. (2018b). G*lacier inventory of Pamir and Karakoram, link to GIS files*. PANGAEA. https://doi.org/10.5194/essd-2018-35

Mueller, N., Lewis, A., Roberts, D., Ring, S., Melrose, R., Sixsmith, J., Lymburner, L., McIntyre, A., Tan, P., Curnow, S., & Ip, A. (2016). Water observations from space: Mapping surface water from 25 years of Landsat imagery across Australia. *Remote*

Sensing of Environment, *174*, 341–352. https://doi.org/10.1016/J.RSE.2015.11.003

Onojeghuo, A. O., Blackburn, G. A., Wang, Q., Atkinson, P. M., Kindred, D., & Miao, Y. (2018). Mapping paddy rice fields by applying machine learning algorithms to multi-temporal Sentinel-1A and Landsat data. *International Journal of Remote Sensing*, *39*(4) , 1042-1067. https://doi.org/10.1080/01431161.2017.1395969

Paul, F., Barrand, N. E., Baumann, S., Berthier, E., Bolch, T., Casey, K., Frey, H., Joshi, S. P., Konovalov, V., Le Bris, R., Mölg, N., Nosenko, G., Nuth, C., Pope, A., Racoviteanu, A., Rastner, P., Raup, B., Scharrer, K., Steffen, S., & Winsvold, S. (2013). On the accuracy of glacier outlines derived from remote-sensing data. A*nnals oof Glaciology*, *54*(63), 171-182 https://doi.org/10.3189/2013AoG63A296

Paul, F., Barry, R. G., Cogley, J. G., Frey, H., Haeberli, W., Ohmura, A., Ommanney, C. S. L., Raup, B., Rivera, A., & Zemp, M. (2009). Recommendations for the compilation of glacier inventory data from digital sources. *Annals of Glaciology*, *50*(53), 119-126. doi:10.3189/172756410790595778

Paul, F., Bolch, T., Kääb, A., Nagler, T., Nuth, C., Scharrer, K., Shepherd, A., Strozzi, T., Ticconi, F., Bhambri, R., Berthier, E., Bevan, S., Gourmelen, N., Heid, T., Jeong, S., Kunz, M., Lauknes, T. R., Luckman, A., Merryman Boncori, J. P. & Van Niel, T. (2015). The glaciers climate change initiative: Methods for creating glacier area, elevation change and velocity products. *Remote Sensing of Environment*, *62*, 408-426. https://doi.org/10.1016/j.rse.2013.07.043

Paul, F., Winsvold, S. H., Kääb, A., Nagler, T., & Schwaizer, G. (2016). Glacier remote sensing using Sentinel-2. part II: Mapping glacier extents and surface facies, and comparison to Landsat 8. *Remote Sensing*, *8*(7), 575. https://doi.org/10.3390/rs8070575

Pfeffer, W. T., Arendt, A. A., Bliss, A., Bolch, T., Cogley, J. G., Gardner, A. S., Hagen, J.-O. O., Hock, R., Kaser, G., Kienholz, C., Miles, E. S., Moholdt, G., Mölg, N., Paul, F., Radic, V., Rastner, P., Raup, B. H., Rich, J., Sharp, M. J., & Sharp, M. J. (2014). The Randolph Glacier Inventory: a globally complete inventory of glaciers. *Journal of Glociology*, *60*(221), 537-552 . https://doi.org/10.3189/2014JoG13J176

Rabus, B., Eineder, M., Roth, A., & Bamler, R. (2003). The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. *ISPRS Journal of Photogrammetry and Remote Sensing*, *57*(4), 241–262.

Racoviteanu, A. E., Paul, F., Raup, B., Khalsa, S. J. S., & Armstrong, R. (2009). 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. https://doi.org/10.3189/172756410790595804

Rastner, P., Strozzi, T., & Paul, F. (2017). Fusion of multi-source satellite data and DEMs to create a new glacier inventory for Novaya Zemlya. *Remote Sensing*, *9*(11), 1122. https://doi.org/10.3390/rs9111122

Robson, B. A., Nuth, C., Dahl, S. O., Hölbling, D., Strozzi, T., & Nielsen, P. R. (2015). Automated classification of debris-covered glaciers combining optical, SAR and topographic data in an object-based environment. *Remote Sensing of Environment* , *170* , 372-387. https://doi.org/10.1016/j.rse.2015.10.001

Sakai, A. (2019). Brief communication: Updated GAMDAM glacier inventory over high-mountain Asia. *The Cryosphere*, *13*(7), 2043–2049. https://doi.org/10.5194/tc-13-2043-2019

Shabeh ul, H. (2016). Future Water Availability from Hindukush-Karakoram-Himalaya upper Indus Basin under Conflicting Climate Change Scenarios. *Climate*, *4*(3), 40. https://doi.org/10.3390/cli4030040

Shao, G., & Wu, J. (2008). On the accuracy of landscape pattern analysis using remote sensing data. *Landscape Ecology*, *23*(5), 505–511. https://doi.org/10.1007/s10980-008-9215-x

Syed, N. ul H., Md Reba, M. N., Hussain, D., & Ahmed, A. (2018). Spectral quality assessment of Landsat 8 and Sentinel 2

bands for glacier identification in Upper Indus Basin. *39th Asian Conference on Remote Sensing (ACRS)*.

Talukdar, S., Singha, P., Mahato, S., Shahfahad, Pal, S., Liou, Y. A., & Rahman, A. (2020). Land-use land-cover classification by machine learning classifiers for satellite observations-A review. In *Remote Sensing*, *12*(7), 1135. Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/rs12071135

Thanh Noi, P., & Kappas, M. (2017). Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. *Sensors (Basel, Switzerland)*, *18*(1), 18. https://doi.org/10.3390/s18010018

Tobias, B., Ulrich, K., & Jeffrey, O. (2005). Using ASTER and SRTM DEMs for studying geomorphology and glaciation in high mountain areas - Zurich Open Repository and Archive. In Marinko, O. (Ed.), *New Strategies for European Remote Sensing*. Millpress.

Veh, G., Korup, O., Walz, A., & Roessner, S. (2018). Detecting Himalayan glacial lake outburst floods from Landsat time series. *Remote Sensing of Environment*, *207*, 84-97. . https://doi.org/10.1016/j.rse.2017.12.025

Wang, H., Yang, R., LI, X., & CAO, S. (2017). Glacier parameter extraction using Landsat 8 images in the eastern Karakorum. IOP Conference Series: *Earth and Environmental Science*, *57*(1), 012004. httos://doi.org/10.1088/1755-1315/57/1/012004

Wessels, K., van den Bergh, F., Roy, D., Salmon, B., Steenkamp, K., MacAlister, B., Swanepoel, D., & Jewitt, D. (2016). Rapid Land Cover Map Updates Using Change Detection and Robust Random Forest Classifiers. *Remote Sensing*, *8*(11), 888. https://doi.org/10.3390/rs8110888

Wiltshire, A. J. (2014). Climate change implications for the glaciers of the Hindu Kush, Karakoram and Himalayan region. *The Cryosphere*, *8*(3), 941-958. https://doi.org/10.5194/tc-8-941-2014

Zemp, M., Frey, H., Gärtner-Roer, I., Nussbaumer, S. U., Hoelzle, M., Paul, F., Haeberli, W., Denzinger, F., Ahlstrøm, A. P., Anderson, B., Bajracharya, S., Baroni, C., Braun, L. N., Cáceres, B. E., Casassa, G., Cobos, G., Dávila, L. R., Delgado Granados, H., Demuth, M. N., & Vincent, C. (2015). Historically unprecedented global glacier decline in the early 21st century. *Journal of Glaciology*, *61*(228), 745–762. https://doi.org/DOI: 10.3189/2015JoG15J017

Zhang, H. K., Roy, D. P., Yan, L., Li, Z., Huang, H., Vermote, E., Skakun, S., & Roger, J.-C. (2018). Characterization of Sentinel-2A and Landsat-8 top of atmosphere, surface, and nadir BRDF adjusted reflectance and NDVI differences. *Remote Sensing of Environment, 215*, 482-494. https://doi.org/10.1016/J.RSE.2018.04.031

Zhang, J., Jia, L., Menenti, M., & Hu, G. (2019). Glacier Facies Mapping Using a Machine-Learning Algorithm: The Parlung Zangbo Basin Case Study. *Remote Sensing*, *11*(4), 452. https://doi.org/10.3390/rs11040452

Zhang, M., Wang, X., Shi, C., & Yan, D. (2019). Automated glacier extraction index by optimization of Red/SWIR and NIR /SWIR ratio index for glacier mapping using landsat imagery. *Water (Switzerland)*, *11*(6), 1223. https://doi.org/10.3390/w11061223