



Detecting Himalayan glacial lake outburst floods from Landsat time series

Georg Veh^{a,*}, Oliver Korup^a, Sigrid Roessner^b, Ariane Walz^a

^a Institute of Earth and Environmental Sciences, University of Potsdam, D-14476 Potsdam, Germany

^b GFZ German Research Centre for Geosciences, Remote Sensing Section, Telegrafenberg, D-14473 Potsdam, Germany

ARTICLE INFO

Keywords:

Random Forest
Fuzzy classification
Land cover maps
Change detection
Change points
Lakes
Sediment tails
Hindu Kush Himalayas (HKH)

ABSTRACT

Several thousands of moraine-dammed and supraglacial lakes spread over the Hindu Kush Himalayan (HKH) region, and some have grown rapidly in past decades due to glacier retreat. The sudden emptying of these lakes releases large volumes of water and sediment in destructive glacial lake outburst floods (GLOFs), one of the most publicised natural hazards to the rapidly growing Himalayan population. Despite the growing number and size of glacial lakes, the frequency of documented GLOFs is remarkably constant. We explore this possible reporting bias and offer a new processing chain for establishing a more complete Himalayan GLOF inventory. We make use of the full seasonal archive of Landsat images between 1988 and 2016, and track automatically where GLOFs left shrinking water bodies, and tails of sediment at high elevations. We trained a Random Forest classifier to generate fuzzy land cover maps for 2491 images, achieving overall accuracies of 91%. We developed a likelihood-based change point technique to estimate the timing of GLOFs at the pixel scale. Our method objectively detected ten out of eleven documented GLOFs, and another ten lakes that gave rise to previously unreported GLOFs. We thus nearly doubled the existing GLOF record for a study area covering ~10% of the HKH region. Remaining challenges for automatically detecting GLOFs include image insufficiently accurate co-registration, misclassifications in the land cover maps and image noise from clouds, shadows or ice. Yet our processing chain is robust and has the potential for being applied on the greater HKH and mountain ranges elsewhere, opening the door for objectively expanding the knowledge base on GLOF activity over the past three decades.

1. Introduction

Melting glaciers in the Hindu Kush Himalayan (HKH) mountain ranges feed several thousand moraine-dammed and supraglacial lakes (Ives et al., 2010; Nie et al., 2017). Embedded in loose debris and surrounded by sources of falling debris and ice, many of these water bodies are prone to glacial lake outburst floods (GLOFs) (Clague and Evans, 2000). GLOFs can release and transport millions of cubic meters of water and sediment within few hours (Bajracharya et al., 2007; Cenderelli and Wohl, 2001; Wang et al., 2012). Quaternary outburst floods in the HKH have been shaping major valley trains for thousands of years (Korup and Tweed, 2007; O'Connor et al., 2013; Scherler et al., 2014). GLOFs have also killed several hundreds of people in the past decades and caused substantial damage to infrastructure, hydropower stations, livestock and farmland (Kattmann, 2003; Richardson and Reynolds, 2000; Yamada and Sharma, 1993). Data on loss and damage are crude, though Nepal and Bhutan may have suffered the highest socio-economic impacts by historic GLOFs worldwide (Carrivick and Tweed, 2016). In any case, GLOFs clearly rank among the most publicised glacial hazards in the Himalayas (Richardson and Reynolds, 2000).

Difficult access and high alpine conditions make detailed field-based

monitoring of lakes prone to outburst impractical; several studies thus resorted on measuring lake bathymetry, dam material, and the surrounding topography (Fujita et al., 2013; Wang et al., 2012; Worni et al., 2013). Moreover, data on historic GLOFs in the HKH are scarce and vague about outburst parameters. Local GLOF inventories often contradict each other, at least judging from data that we collected on 36 GLOFs from moraine-dammed lakes in the Himalayas since the 1950s (Ives et al., 2010; Komori et al., 2012; Liu et al., 2014; Wang et al., 2012; Table 1).

Current research aims at linking global climate warming to glacier melt, and the formation and changes of meltwater lakes, including the probability of catastrophic lake outburst (Harrison et al., 2017). Negative glacier mass balances (Brun et al., 2017) and increases in glacial lake number and area (Nie et al., 2017; Song et al., 2017; Zhang et al., 2015) have characterized many parts of the HKH over the past decades, and thawing permafrost in glacier dams and surrounding rock walls may further destabilise the glacial lake system (Haeberli et al., 2017). While all these observations are in line with a hypothesized increase in GLOF frequency, this remains difficult to test given commonly observed rates of up to one event per year, and only a few dozen reliably

* Corresponding author.

E-mail address: georg.veh@uni-potsdam.de (G. Veh).

Table 1

Documented GLOFs between 1988 and 2016. ID corresponds to labels in Fig. 1. We visually assessed whether drainage was complete (C) or partial (P).

ID	Lake	Country	E [°]	N [°]	Elevation [m a.s.l.]	Loss in lake area [m ²]	Date	Type of drainage	Source
1	Chorabari	India	79.06	30.75	3881	11,700	2013-06-17	C	Allen et al. (2016); Das et al. (2015)
2	Zanaco	TAR/China	85.37	28.66	4737	66,600	1995-06-06	C	Liu et al. (2014)
3	Zhangzangbo 2	Nepal	86.06	28.08	4501	10,800	2016-07-07	C	Cook et al. (2017); Gimbert et al. (2017)
4	Sabai Tsho	Nepal	86.84	27.74	4492	163,800	1998-09-03	P	Lamsal et al. (2015); Osti and Egashira (2009)
5	Lemthang Tsho	Bhutan	89.58	28.07	4273	53,100	2015-06-28	C	Gurung et al. (2017)
6	Chongbaxia Tsho	TAR/China	89.74	28.21	5028	227,700	Spring–Summer 2001	P	Komori et al. (2012)
7	Tshojo glacier	Bhutan	90.16	28.10	4273	81,900	2009-07-29	P	Yamanokuchi et al. (2011)
8	Luggye Tsho	Bhutan	90.28	28.09	4623	140,400	1994-10-07	P	Fujita et al. (2008); Watanbe and Rothacher (1996)
9	Gangri Tsho III	Bhutan	90.81	27.90	4826	26,100	Spring–Summer 1998	P	Komori et al. (2012)
10	Ranzeria Co	TAR/China	93.53	30.47	5051	246,600	2013-07-05	P	Sun et al. (2014)
11	Tsho Ga	TAR/China	94.00	30.83	4760	140,400	2009-04-29	P	Nie et al. (pers. comm., 2017)

documented events (Carrivick and Tweed, 2016; Harrison et al., 2017). This mismatch could reflect a censoring bias such that only extreme events and their impacts have been reported.

Clearly, a database of past events as complete as possible is essential for robust and reliable GLOF hazard assessment (Emmer et al., 2016). Time series from satellite imagery find widespread use for compiling multi-temporal glacial lake inventories, especially for rapidly expanding lakes that are thought to have an elevated outburst potential (Nie et al., 2017; Wang et al., 2015a, 2011). To our knowledge, no study has systematically explored the Landsat archive for retrospective GLOF detection in the HKH, although it offers a largely continuous, nearly 30-year time series with regional coverage every 16 days. For tracing past GLOFs, we build on the experience that lakes most often disappeared or shrank abruptly and exposed debris fans and sediment tails in river channels downstream. Only Komori et al. (2012) used these two indicators to visually scan satellite archives for unreported GLOFs in the Bhutan Himalayas. Since glacial lakes often re-fill or re-expand within few years after an outburst, previously used mapping intervals of five to ten years might be too coarse to detect GLOFs from lake inventories (Zhang et al., 2015). Dense cloud cover during the monsoon, lake freezing in winter, and mountain shadows are the main challenges for pursuing the glacial lake area over time. Multiple noise-free images per year may be desirable to detect reliably sudden lake changes, but remain rare in the Himalayan weather conditions. Expert-based manual mapping from multi-temporal medium to high resolution (< 30 m) imagery has so far offered high-quality lake inventories, but is resource-intensive and thus restricted to few selected glacial lakes (Shrestha et al., 2013; Wang et al., 2015a; Yao et al., 2012) or single catchments (Bolch et al., 2008; Che et al., 2014; Jain et al., 2012). Semi-automatic mapping using chains of decision rules along band and topographic indices allows for monitoring of glacial lakes over larger areas, but requires time-consuming post-processing (Gardelle et al., 2011; Li and Sheng, 2012; Song et al., 2016). Machine learning classifiers such as Random Forests (RF) have rapidly advanced the mapping of changing land cover and water bodies (Mueller et al., 2016; Rover et al., 2012; Tulbure et al., 2016), thereby accompanying a high potential for GLOF detection. Random Forests (Breiman, 2001) are ensemble classifiers that use bagging to grow and aggregate multiple independent decision trees from a bootstrap sample of predictor variables. The classifier can deal with non-monotonic and non-linear relationships between the predictors and response variables, and is robust against overfitting (Rodriguez-Galiano et al., 2012). Hence, RF are a powerful alternative to single, parametric classifiers (Waske and Braun, 2009), especially for spectrally variable target classes such as glacial lakes of differing depth and turbidity. Random Forests offer fuzzy or probabilistic class memberships, which offer richer information about the likelihood of change in land-cover time series (Foody and Boyd, 1999; Metternicht, 1999).

Change detection of water bodies with Landsat time series focused either on long-term trends of lake growth or shrinkage (Fraser et al., 2014; Nitze and Grosse, 2016) or on the estimation of flooding frequencies (Mueller et al., 2016; Tulbure et al., 2016). Automatically extracting distinct events of rapid lake decrease, as is the case for GLOFs, has rarely been of interest (Olthof et al., 2015). Change-point detection in Landsat time series is well-established for forest disturbance mapping, where pixels of vegetation indices are scanned for level shifts (Hermosilla et al., 2015; Kennedy et al., 2010) or structural breaks in fitted harmonic models (DeVries et al., 2015; Verbesselt et al., 2012). However, alternative techniques are required, as these approaches are difficult to apply to Himalayan glacial lakes where indices such as the Normalized Difference Water Index (NDWI; McFeeters, 1996) share similar spectral characteristics with clouds or shadows (Li and Sheng, 2012).

Our aim is to develop, validate and apply a technique to automatically detect past Himalayan GLOFs. We present a processing chain that traces losses in lake areas from nearly three decades of seasonal Landsat imagery building on (1) a Random-Forest based land cover classification and (2) a novel, likelihood based change-point algorithm to approximate the time stamp of GLOFs. We apply this processing chain to a spatial subset of the HKH, validate our method with documented GLOFs and present newly detected GLOFs. Our search includes sediment tails downstream of drained lakes, allowing us to trace the location, timing, and size of GLOFs, and thus contributing to a more complete GLOF inventory of the Himalayas.

2. Study area

Of all 36 documented GLOFs over the past seven decades, we could visually identify eleven GLOFs in Landsat images (Fig. 1). We obtained information on the date, location, and type of drainage for each GLOF, using the drained lake area as a key metric for comparing pre- and post-GLOF images (Table 1).

These GLOFs occurred in four different regions (Fig. 1) between the central-western Himalayas of northern India (a), the central Himalayas of Nepal and Bhutan (b and c), and the eastern Nyainqentanglha Mountains of China (d). The number of present-day moraine-dammed and supraglacial lakes in these areas is challenging to establish. Estimates for the whole HKH range from 2276 (Fujita et al., 2013) to > 8000 (Ives et al., 2010), depending on definition, mapping scale, and size of study area. In the central Himalayas, glacial lakes grew by 23% in size between 1990 and 2015. Lakes grow less rapidly in area in the western (5.0–5.4%) and eastern Himalayas (7.7–11.1%) (Nie et al., 2017), and mostly tied to glacier melt (Gardelle et al., 2013; Käb et al., 2012; Song et al., 2017; Wang et al., 2015b).

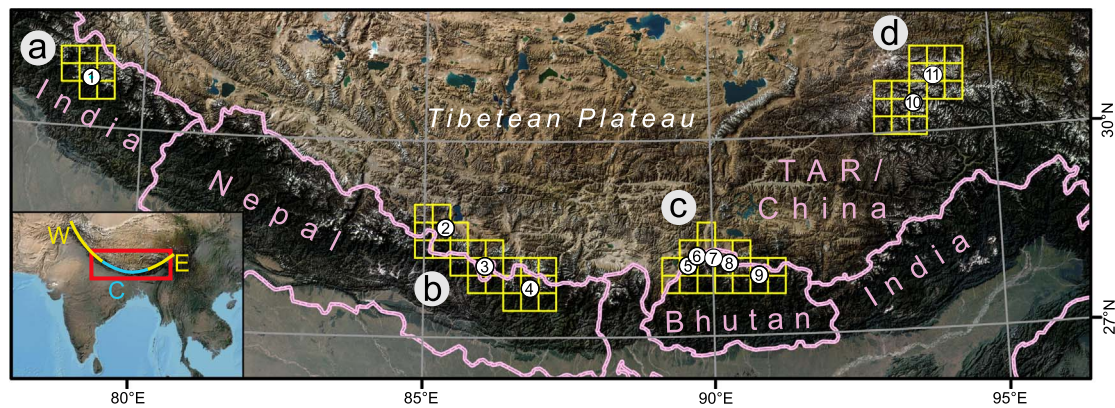


Fig. 1. Documented GLOFs (1–11, Table 1) between 1994 and 2016 in the HKH. Study regions a–d feature 30 km × 30 km tiles (yellow squares; see Fig. 3 for more details). Inset shows the geographic setting of the HKH, divided into a western (W), central (C) and eastern (E) segment. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3. Data and methods

3.1. Data

Our processing chain builds on six open-source data sets, including image and topographic data in raster format and two glacier inventories in vector format (Table 2).

The eleven reference GLOFs are covered by 19 Landsat scenes. We scanned the entire Landsat archive on the *EarthExplorer* web portal for TM, ETM+ and OLI images with < 60% cloud cover and time stamps between September and November to avoid excess cloud cover during summer monsoon or snow and ice cover in winter. We downloaded 2491 images that were radiometrically corrected to Top-of-Atmosphere (TOA) reflectance, and geometrically corrected to processing level L1T by the ESPA Landsat data processing platform.

The first images were acquired in late 1988 by TM, and coverage remained limited in the 1990s. The deployment of ETM+ in 1999 brought more coverage, though the failure of the Scan Line Corrector in 2003 caused image data gaps. The shutdown of TM in 2012 was compensated with the launch of OLI in 2013 (Fig. 2), so that the time series spans 29 years in total. The central HKH is most densely captured with up to 429 images per pixel (Fig. 3-A).

We used the *CFmask* product, a C implementation of the Function of Mask, originally designed for detecting clouds in Landsat images (Zhu and Woodcock, 2012). Zhu et al. (2015) extended *CFmask* to classify each Landsat scene into Cloud, Shadow, Ice and Snow, Water and Clear (i.e. without atmospheric disturbance), and to assign three categories of classification confidence. We also used the 30 m resolution digital elevation model (DEM) from the Shuttle Radar Topography Mission

(SRTM) after filling voids with data from the ALOS World 3D–30 m digital surface model, and smoothing the data with a 9 × 9 Gaussian filter. For data on glaciers we used the ICIMOD and the Randolph Glacier Inventory (RGI) (Bajracharya et al., 2011; Pfeffer et al., 2014). We filled the missing Chinese territory of the ICIMOD inventory with RGI data, so that our inventory contains > 50000 glaciers with a total area of ~9400 km².

3.2. The processing chain

To automatically detect GLOFs from Landsat time series, we used a two-step approach: first, we trained a Random Forest model to classify land cover in all tiles; second, we used a likelihood-based change-point algorithm to identify GLOF pixels with distinct changes from water to land (Fig. 4).

3.2.1. Data preprocessing

The study area covers six UTM zones so that we first re-projected the images to the centrally located UTM zone 45N with nearest neighbour resampling to preserve the original digital numbers of the image bands. We divided the study area into 30 km × 30 km tiles to increase data reading speed for parallel processing. We processed those eleven tiles containing documented outbursts, and their 57 neighbouring tiles, resulting in a total of 68 tiles (Fig. 1). Our analysis focused on lakes impounded by moraines or coalescing on debris-covered glaciers, so that we reduced the search range to valley floors < 3 km downstream of glaciers in our inventory. We excluded areas above the median elevation of each glacier, assuming that these steep and rugged areas do not host lakes.

Table 2

Data sets used in study.

Data set	Temporal coverage	Data format	Resolution [m]	Source
Landsat imagery	1988–2016	6 raster bands (TM and ETM+) 9 raster bands (OLI)	30	United States Geological Survey (USGS) https://earthexplorer.usgs.gov
<i>CFmask</i> products	1988–2016	Single-raster band	30	Earth Resources Observation And Science (EROS) Center Science Processing Architecture (ESPA) https://espa.cr.usgs.gov
SRTM DEM	2000	Single-raster band	30	USGS https://earthexplorer.usgs.gov
ALOS World 3D DEM	2006–2012	Single-raster band	30	Japan Aerospace Exploration Agency (JAXA) http://www.eorc.jaxa.jp/ALOS/en/aw3d30/data/index.htm
Randolph Glacier Inventory, V 5.0, Region 13–15	2006–2010	Multipart polygon shape file	–	Global Land Ice Measurements from Space (GLIMS), see Pfeffer et al. (2014)
ICIMOD Glacier Inventory	2005 ± 3 years	Multipart polygon shape file	–	International Centre for Integrated Mountain Development (ICIMOD), see Bajracharya et al. (2011)

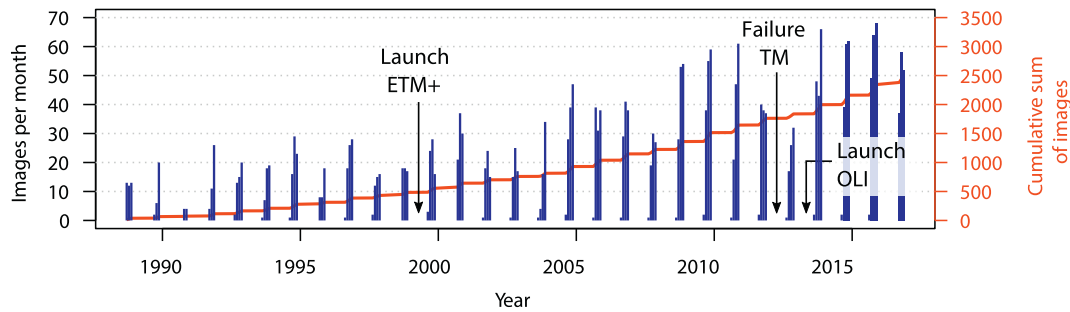


Fig. 2. Number and approximate time stamps of Landsat images per month for the study period (1988–2016).

3.2.2. Step 1: Random Forest classification

We defined six land cover classes that are potentially relevant for detecting GLOFs. *Water* and *Sediment* were the mandatory classes as we assumed that outbursts of glacial lakes expose sediments in the lake basin and downstream. Three additional classes (*Ice & Snow*, *Cloud* and *Shadow*) are sources of image noise in the alpine landscape. Finally, a *Land* class subsumed all other classes not contained in the other five classes. The eleven tiles containing reference GLOFs offered training data for our classification. We chose the first cloud- and ice-free observation after each GLOF and one randomly selected image before. The adjacent 57 tiles supplied test sites for detecting previously unreported GLOFs. We applied a two-fold stratification by land-cover class and confidence band from the *CFmask* product, and distributed 30 random points per stratum. We manually assigned each point to one of the six land cover classes by comparing the Landsat image with Google Earth images without knowing the classification from the *CFmask* product. Classes *Water*, *Cloud* and *Sediment* were undersampled due to misclassifications in the *CFmask* product; we solved with additional

random points using the lake inventory of Zhang et al. (2015), and a visually defined threshold of 2750 in the SWIR1 band for *Sediment* and *Cloud* samples. Sampling the *Sediment* class avoids confusion with the spectrally similar *Cloud* class. We then merged the *Sediment* and *Land* samples into a *Land & Sediment* class, resulting in a total of five land cover classes as response variable for the RF classifier.

Landsat TM and ETM + share the same spectral band width that differs from OLI. We therefore divided the sampled pixels into TM5/ETM + and OLI subsets. For TM/ETM + (OLI), we collected in total 6573 (5911) point samples, including 617 (633) *Water*, 1822 (2078) *Land & Sediment*, 1324 (1052) *Ice & Snow*, 1088 (945) *Cloud* and 1722 (1203) *Shadow* samples. We extracted the digital numbers from the predictor variables, involving all spectral bands, brightness temperature, and ten ancillary band and topographic indices (Table 3).

For fitting the two RF models (one for TM5/ETM +, one for OLI), we grew 1000 trees and randomly selected four predictors at each split. We assessed the performance of the classifiers via spatial cross-validation: we rearranged the samples into different tile combinations, always

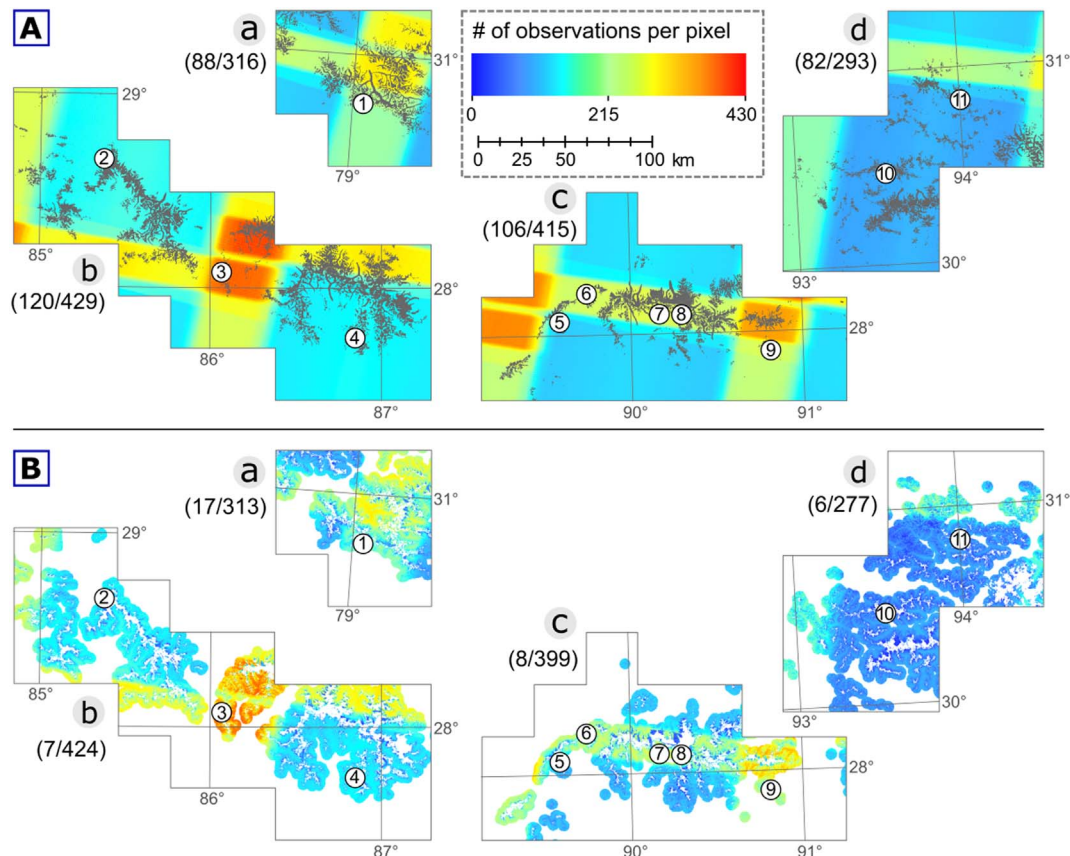


Fig. 3. A) Total number of Landsat observations per pixel for the four study regions (Fig. 1); glaciated areas are depicted in grey. B) Remaining observations for time-series analysis after pre-processing and noise removal (Step 2). Numbers in parentheses are minimum and maximum observations for each study region; numbers 1–11 refer to reference GLOFs in Table 1.

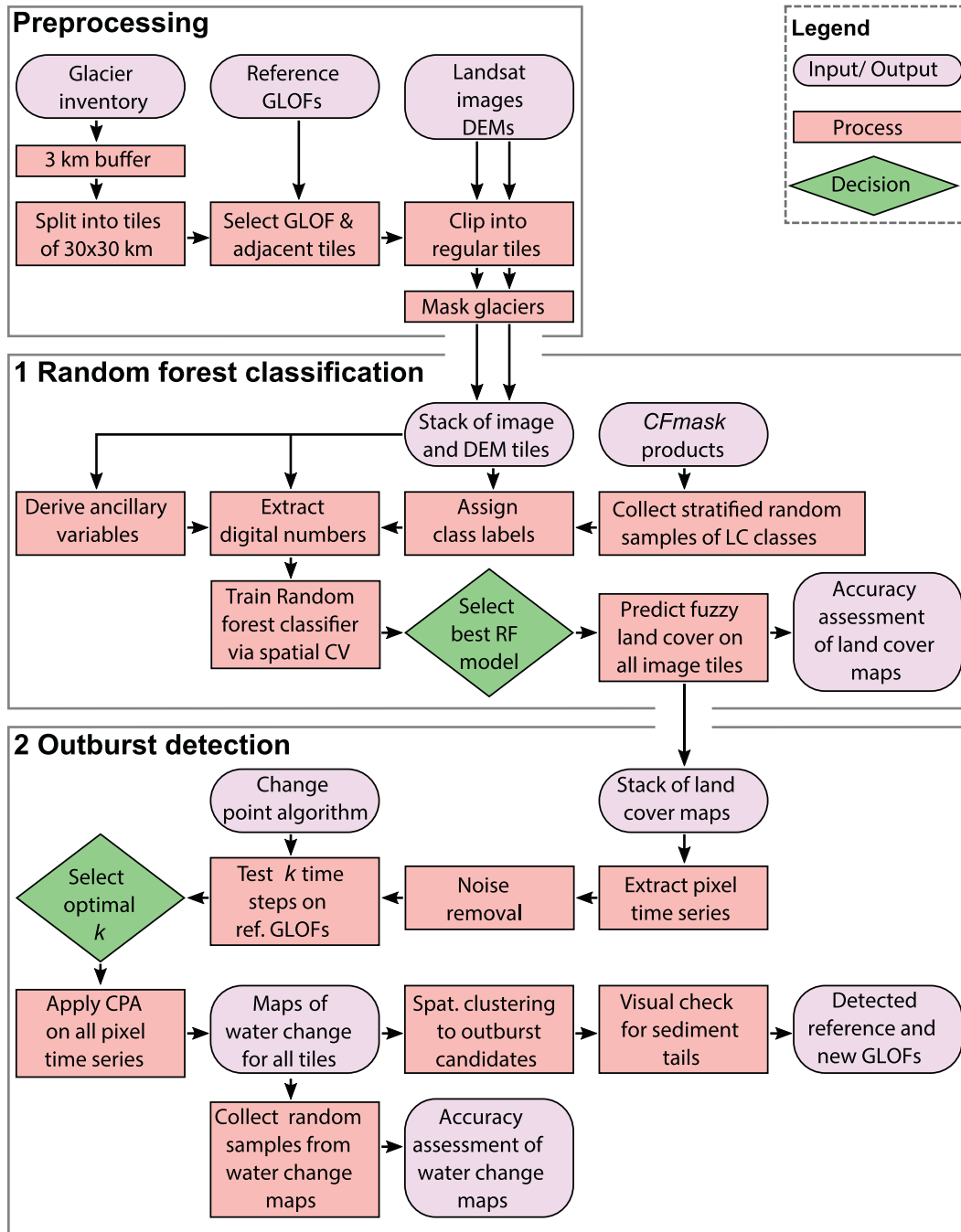


Fig. 4. Flowchart of processing chain for automated GLOF detection. Abbreviations: CPA – change point algorithm (see Section 3.2.3); RF – Random Forest; LC – land cover; CV – cross-validation.

using seven tiles as training and four tiles as test data sets. We re-sampled the training and test data sets to 200 and 100 random instances per class, respectively. In each cross-validation run, the RF models were fitted to the 200 training observations and subsequently used for prediction on the 100 hold-out samples in the test data. We evaluated the classifier performance for each class using a log loss function (1) that penalizes divergences between predictions and manually assigned classes:

$$\text{logloss} = -\frac{1}{N} \sum_{i=1}^N y_i \log p_i + (1 - y_i) \log(1 - p_i) \quad (1)$$

where N is the number of pixels, y_i is an indicator function of whether pixel i belongs to the target class ($y_i = 1$; $y_i = 0$ otherwise), and p_i is the

proportional membership of pixel i over the predicted class. We added a ten and five-fold cost of misclassifying *Water* and *Land & Sediment* to emphasise their importance for GLOF detection.

Labelling mixed pixels with hard classes potentially introduces an interpreter bias during class assignment whose influence can be decreased by cross-validation. We repeated cross-validation 100 times and chose the run with the lowest median log loss as the optimal sample combination. We trained the final RF models with all samples from these tiles and evaluated the accuracy on all samples from the remaining test tiles. For the best two classifiers, we converted the fuzzy class memberships into a hard classification and estimated the error matrices. We report the Overall Accuracy (OA) as well as Producer's Accuracy (PA) and User's Accuracy (UA), and used these two models to predict land cover on all tiles.

Table 3
Ancillary variables for Random Forest training.

Variable	Reference
Normalized Difference Cloud Index (NDCI)	Martinuzzi et al. (2007)
Modified Soil Adjusted Vegetation Index 2 (MSAVI ₂)	Qi et al. (1994)
Normalized Difference Vegetation Index (NDVI)	Rouse Jr et al. (1974)
Normalized Difference Water Index (NDWI)	McFeeters (1996)
Normalized Difference Snow Index (NDSI)	Riggs et al. (1994)
Hillshade	Conrad et al. (2015)
DEM	USGS (EarthExplorer)
Slope	Zevenbergen and Thorne (1987)
Topographic Position Index (TPI)	Weiss (2001)
Aspect	Zevenbergen and Thorne (1987)

3.2.3. Step 2: outburst detection from pixel time series

Our key indicator for GLOFs is the pixel-based change from *Water* to *Land & Sediment*. Each pixel from the stack of land cover maps has a time series of membership probabilities for one of five land-cover classes (Fig. 5-A). We defined class-membership probabilities of *Cloud*, *Shadow* or *Ice & Snow* > 0.5 as noise and excluded them from further analysis (Fig. 5-B), resulting in significant spatial differences of valid observations (Fig. 3-B). Filtered pixel time series with > 10% of their original length could potentially include a change event. Yet simply flagging each time step at which the class membership changes from *Water* to *Land & Sediment* would overestimate the number of change points, since we cannot assume correct classifications of water and land throughout the time series. Indeed, most classification errors form irregular spikes in the time series (Fig. 5-B).

We developed a change-point algorithm (CPA) to remove outliers and to detect major transitions from water to land for each pixel. We approximate the likelihood of change $p(C)$ by multiplying the likelihoods of belonging to the *Water* and *Land & Sediment* class before and

after each time step i , respectively:

$$p(C) = \prod_k^{i-1} P(W) \prod_i^k P(LS) \quad (2)$$

where $P(W)$ is the predicted probability of belonging to class *Water* for k time steps before i , and $P(LS)$ is the probability of belonging to the *Land & Sediment* class for k time steps after i . The likelihood of change $p(C)$ at time step i must exceed a specified threshold $T(C)$ to qualify as a change point. We set the threshold as

$$T(C) = 0.5^{2k} \quad (3)$$

where 0.5 represents the random chance of belonging to a certain class and $2k$ is the CPA bandwidth. We tested different sizes of k time steps ranging from 2 to 10 to identify optimal values for detecting GLOFs. We then generated maps of water change for each tile where all pixels exceeding $T(C)$ were labelled with the date of the change point.

We followed the recommendations of Olofsson et al. (2013) to compute stratified error matrices that consider the proportion of land change area. We applied a stratified random sampling approach by collecting 20 samples for each of the *Water to Land* and *No Change* class from the eleven water change maps, resulting in a total of 440 samples. The *No Change* class was subsampled from pixels where the change-point algorithm revealed no transition. We visually identified the correct class on all images and labelled the pixels with the class value. After estimating the error matrix from raw sample counts, we weighted the matrix cells with the total area computed for the two classes to estimate the accuracy of area change.

We used the density-based clustering technique DBSCAN (Ester et al., 1996) to spatially aggregate pixels within a three-year period as change objects. DBSCAN joins spatially separated change pixels and removes spurious change pixels in regions with low pixel density. We set a minimum of four pixels and a maximum of 150 m to form a cluster of abruptly disappeared lake pixels. For all outburst candidates, we

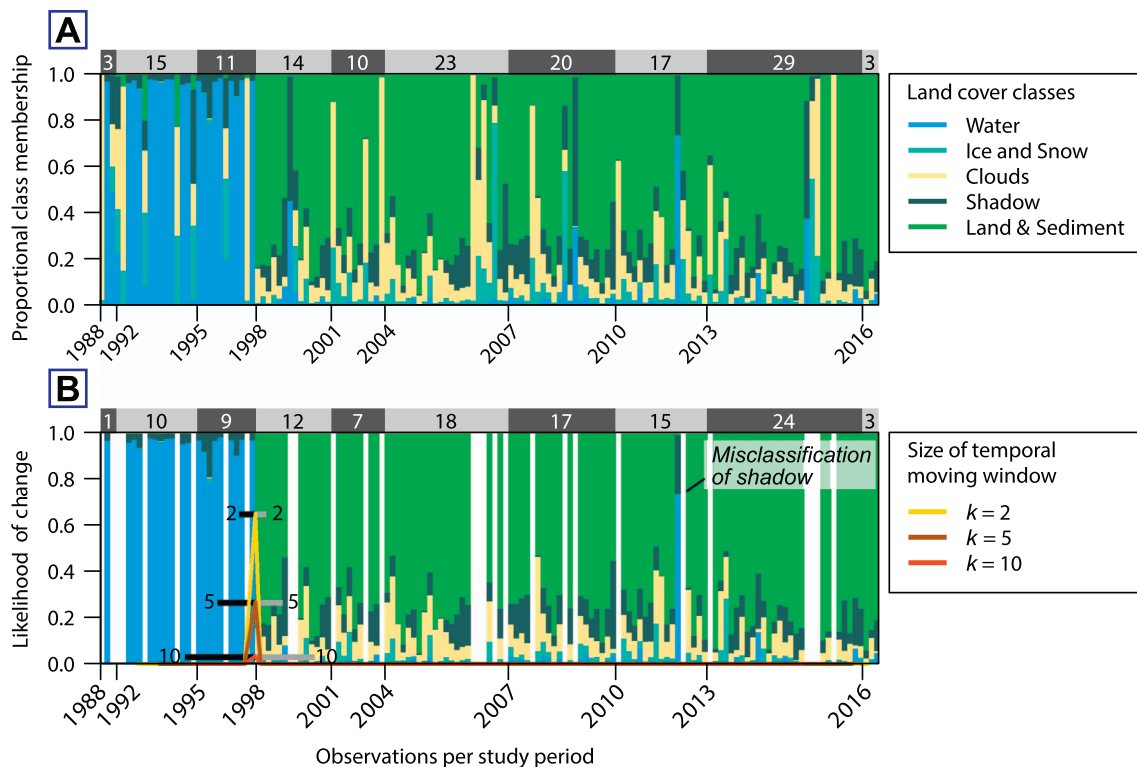


Fig. 5. Example of pixel time series with a change from water to land and sediment for Sabai Tsho GLOF, October 1998. A) Raw time series of estimated membership probability for the five land-cover classes. Note the differing number of observations in three-year intervals at the bar on top. B) Filtered pixel time series results of a change-point algorithm with different values of k (see text).

Table 4

Error matrix for mapped (columns) and predicted (rows) land cover samples, as well as Producer's (PA) and User's (UA) accuracy for all land cover classes as discretisation of fuzzy to hard Landsat TM/ETM+ classification. The accuracies were highest for the *Water* class (blue) and lowest for the cloud class (red) due to confusion with the *Land & Sediment* class.

		Reference							
		Water	Land & Sediment	Ice & Snow	Cloud	Shadow	Total	PA	UA
Prediction	Water	201	2	8	0	5	216	0.931	0.901
	Land & Sediment	0	560	16	40	19	635	0.882	0.927
	Ice & Snow	1	6	585	10	27	629	0.930	0.927
	Cloud	5	29	10	211	6	261	0.808	0.808
	Shadow	16	7	12	0	681	716	0.951	0.923
	Total	223	604	631	261	738	2457		

visually checked the corresponding images for sediment tails downstream of disappeared lake areas. For newly detected GLOFs we gathered the information as in Table 1 and grouped the events into two categories based on our degree of confidence.

4. Results

4.1. Accuracy of land cover maps

The best models from 258 spatial cross-validation runs had a median weighted log loss of 0.1098 for Landsat TM/ETM+, and 0.0868 for Landsat OLI. According to this loss function, the two RF models would assign a class-membership probability of 0.896 (TM/ETM+) and 0.917 (OLI) for a particular observation. The estimated error matrices (Tables 4 and 5) from hard land cover classes indicate a high overall accuracy of 91% (TM/ETM+) and 91% (OLI). The *Water* class is characterized by a comparable (good) UA and PA for TM/ETM+ (0.931/0.901) and OLI (0.929/0.945), confirming a high separability of water from all other land cover classes. Both classifiers showed their lowest performance for the *Cloud* class, mainly because of confusion with the *Land & Sediment* class. While PA and UA for *Ice & Snow* were similar, the results for *Land & Sediment* and *Shadow* diverged between the two classifiers. The number of samples per class in the training and test tiles differed between OLI and TM/ETM+, causing different sample sizes in the error matrices (Tables 4 and 5).

4.2. Optimal time steps for change-point detection

For the eleven reference GLOFs, the change point algorithm showed decreasing mean success rates with increasing number of consecutive observations after noise removal (Fig. 6-A). Success rates were highest for $k = 2$ (85%), then stepped towards a plateau for $k \in [3,4,5]$ (76%–73%) and dropped below 50% for $k \geq 8$. The rate of additionally detected change pixels (Fig. 6-B) fell more steeply, showing a distinct kink at the transition to $k = 3$. We seek to maximize the rate of correctly detected change pixels, while minimizing the rate of additionally detected change pixels. We propose an optimum number of $k = 3$ time steps for change detection owing to the high rate of detected GLOF pixels and, at the same time, to the moderate rate of additionally

detected pixels in the corresponding tiles.

4.3. Accuracy of water change maps

Considering $k = 3$ time steps, our algorithm mapped 0.18% of the total area in the eleven GLOF tiles as a change from *Water to Land* (Table 6). The error matrix of the raw sample counts from the strata *Water to Land* shows that 70 out of 220 change pixels did not change. Area-weighted sample fractions had an overall accuracy of 99% (Table 7). The target class *Water to Land* class had a UA of 67% and a PA of 12%, while the *No Change* class had UA and PA of > 99%.

4.4. Detection of GLOFs

The change point algorithm classified 38,009 change pixels in 68 image tiles. After spatial aggregation to candidate objects with DBSCAN, we investigated the images for sediment tails downstream, successfully identifying ten of the eleven reference GLOFs. For nine of the eleven reference GLOFs, our algorithm correctly detected > 78% of the pixels that we had manually mapped as change pixels (Fig. 6-A). The 2009 GLOF from Tshojo Glacier, Bhutan, was detected with a success of 30%, whereas the 2013 outburst from Lake Chorabari, India, remained undetected. The lake changes predicted by our detection algorithm agree well with independent manual mapping (Fig. 7) to less than two pixels in width.

The change-point algorithm identified ten previously unreported GLOFs (Table 8), with nine of them in the central Himalayas and one in the Nyainqentanglha Mountains (Fig. 8). Seven of these newly detected lake sources had distinct sediment tails downstream (Fig. 10-A and B), which were more difficult to identify for three medium-confidence locations (Fig. 10-C). We could identify outbursts from lake change maps involving as few as eight pixels (7200 m², Table 8-H). We also detected a lake upstream of Luggye Tsho (Bhutan, Table 8-G) that had two outbursts with a major event in 1991, followed by refilling and a second minor event in 2010.

We find that our method underestimates lake area changes by 9.4% on average; most of the lakes showed partial drainage (Table 8). Drained lakes areas of the newly detected GLOFs are below 120000 m², corresponding to half of the maximum size of previously documented GLOFs (Fig. 9).

Table 5

Error matrix from mapped (columns) and predicted (rows) land cover samples, as well as Producer's (PA) and User's (UA) accuracy for all land cover classes as discretisation of fuzzy to hard Landsat OLI classification.

		Reference							
		Water	Land & Sediment	Ice & Snow	Cloud	Shadow	Total	PA	UA
Prediction	Water	171	3	2	0	8	184	0.929	0.945
	Land & Sediment	0	691	6	6	7	710	0.973	0.857
	Ice & Snow	1	9	292	4	9	315	0.927	0.945
	Cloud	3	74	7	395	1	480	0.823	0.975
	Shadow	6	29	2	0	218	255	0.855	0.897
	Total	181	803	307	405	243	1943		

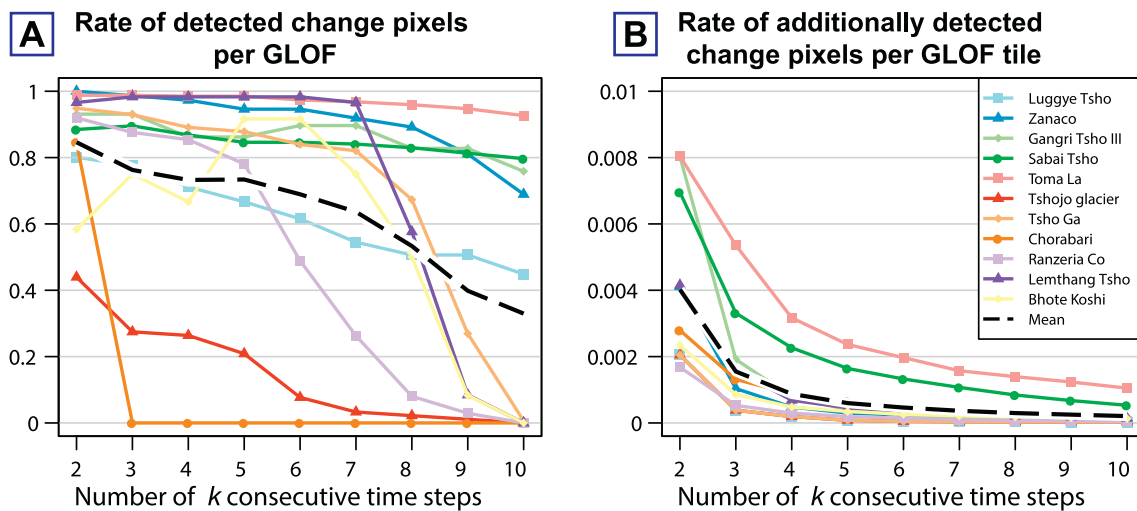


Fig. 6. True positive rate expressed as the rate of detected change pixels by window size for all reference GLOFs; B) Rate of additionally detected change pixels within the GLOF tile expressed as the number of additionally detected pixels detected pixels divided by the number of all valid pixels per GLOF tile.

Table 6

Error matrix for a total of 440 stratified samples randomly collected from the land cover change maps of the eleven GLOF tiles. *Water to Land* is the change between the classes indicating a GLOF; *No Change* includes stable land cover over of *Water* and *Land*, as well as all other changes in classes *Cloud*, *Shadow* and *Ice & Snow*. Map Area refers to the sum of all pixels of the change classes multiplied by a pixel size of 900 m². W_i is the proportion of the area mapped as change class i , divided by the total area.

	Reference				Map area [km ²]	W_i
		Water to Land	No Change	Total		
Prediction	Water to Land	150	70	220	9.532	0.002
	No Change	2	218	220	5246.078	0.998
	Total	152	288	440	5255.61	1

Table 7

Estimated error matrix based on raw sample counts from Table 6. Cell entries of the change classes *Water to Land* and *No Change* represent the estimated proportion of area calculated as the product of the raw sample fraction and class weights W_i given in Table 6.

	Reference				UA	PA	Overall
		Water to land	No change	Total			
Prediction	Water to Land	0.0012	0.0006	0.0018	0.6667	0.1165	0.9903
	No Change	0.0091	0.9891	0.9982	0.9909	0.9994	
	Total	0.0103	0.9897	1			

5. Discussion

5.1. Data quality

The success of our processing chain for automatically detecting past GLOFs in the HKH rests on a number of assumptions and simplifications. Time-series analysis requires co-registered Landsat imagery that we assumed to be sufficiently honoured by consistently using L1T imagery. Precise pixel alignment is most relevant for partial drainage where the decrease in lake area is only several pixels wide (Tables 1 and 8). The resampling of overlapping Landsat scenes to the master projection (UTM Zone 45N) may also shift pixels by up to one pixel width, but this error does not negatively affect our change detection. We rarely observed such artefact offsets of lake shorelines that the change point

algorithm interpreted as change events from water to land, causing an overestimation of change events. However, nearest neighbour is the only interpolation method for image resampling that preserves the original digital numbers (Parker et al., 1983). Other techniques such as cubic convolution or spline interpolation would blur the distinct transition between land and lake (Parker et al., 1983), compromising our ability to detect minor GLOFs.

5.2. Quality of fuzzy land cover maps

We chose a RF classifier because it is flexible, easy to use, robust against overfitting, while offering soft class membership estimates that express some of the uncertainties in the data. Moreover, RF performed robustly on comparable or smaller training data sets than ours (Rodríguez-Galiano et al., 2012). We avoided simpler and more transparent (parametric) classifiers, as these often perform worse in land-cover classification (Waske and Braun, 2009; Xu et al., 2014). Alternatives such as Support Vector Machines and Artificial Neural Networks have no or weakly justified probabilistic output, while algorithm parametrization becomes less intuitive (Lippitt et al., 2008). Future work might compare the performance of different classifiers in our workflow, although such benchmarking depends heavily on the choice of data (Belgiu and Drăguț, 2016). The error matrices (Tables 4 and 5) indicate that confusing *Clouds* with *Land* is the most prominent classification error. Even a penalization could not fully avoid misclassifications. Foga et al. (2017) suggested masking clouds from images with *CFmask*, but our tests with *CFmask* merely confirmed the known issue of confusing clouds with bright sediment surfaces. A possible solution might involve first predicting land cover on training images, followed by updating the classifier with samples from misclassified pixels along the lines of reinforcement learning. Collecting training data is the most labour-intensive part in our processing chain, though it remains doubtful whether adding thousands of new training data would substantially increase our classification accuracy (Zhu et al., 2016). Spatial cross-validation with bootstrapping improves the performance of the RF classifier in diverse geographical settings, and is promising for transferring our approach to the whole HKH region.

5.3. Challenges of change-point detection

Calculating likelihoods in the change point algorithm assumes identically and independently distributed time steps. However, pixel time series involving multiple observations per year are temporally correlated and not equidistant due to noise. Other studies on change-

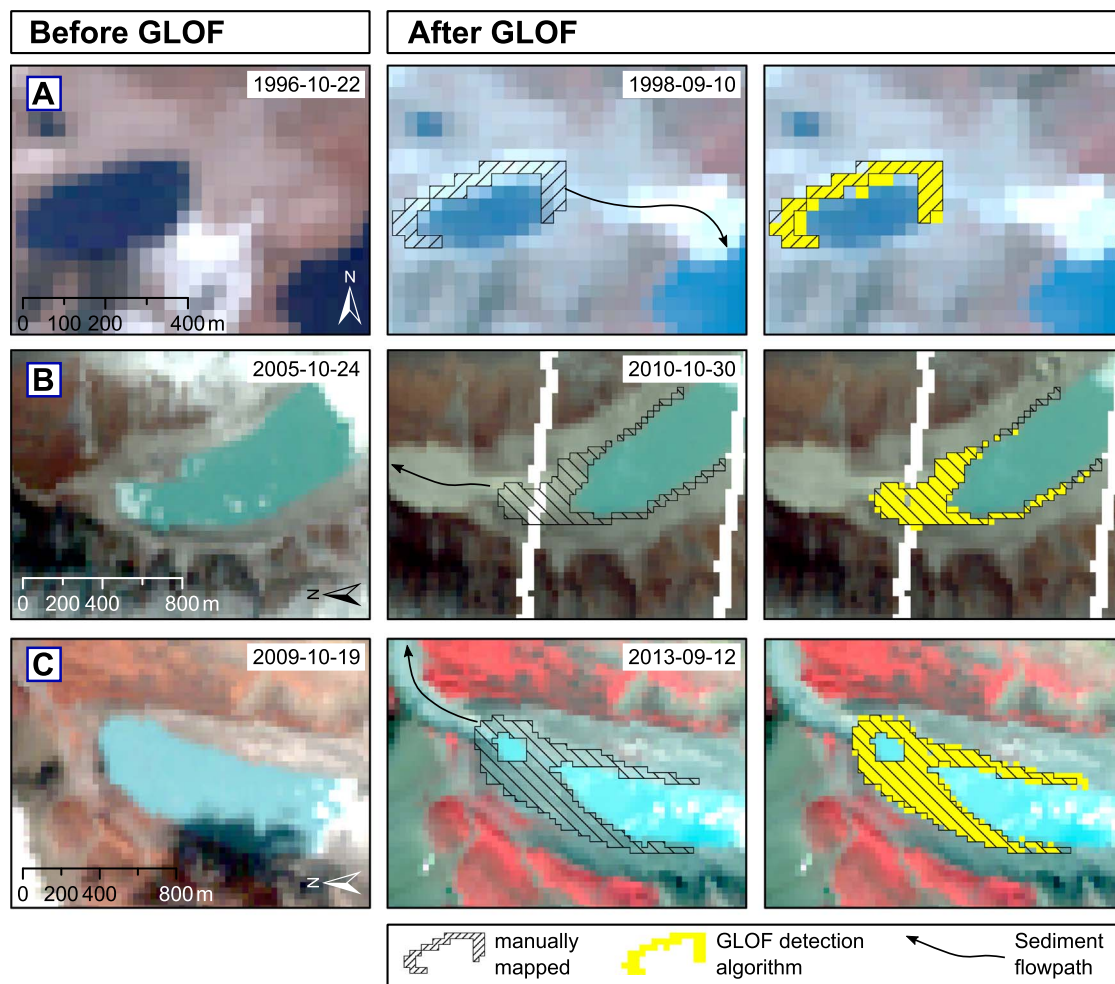


Fig. 7. Glacial lake outbursts from A) Gangri Tsho III in 1998, B) Tsho Ga in 2009 and C) Ranzeria Co in 2013, showing the last noise-free image before (left panels) and the first noise-free image after the event (middle and right panels). The appearance of sediment tails downstream is among our key criteria for detecting GLOFs. The post-GLOF image from Tsho Ga shows artefacts from the failure of the scan-line corrector of Landsat ETM+.

point detection visually selected only one suitable image per year (Pflugmacher et al., 2012) or generated annual to five-year image composites (Griffiths et al., 2014; Schroeder et al., 2011), likely concealing GLOFs within certain seasons (as for Sabai Tsho in September 1998 or Luggye Tsho in October 1994). Data availability for the reference GLOFs suggests that four to seven years without any clear view of the lakes are common (Tsho Ga and Ranzeria Co), and that high

overpass rates (Gangri Tsho III) do not necessarily increase the chance of acquiring suitable images (Fig. 11).

The highly varying data availability in the HKH region led us to develop a change-point algorithm that differs from previously proposed methods; for example, for boreal environments comparable data gaps are less frequent (Hermosilla et al., 2016). Choosing $k = 3$ time steps of consecutive observations for computing the likelihoods of *Water* and

Table 8

Previously unreported GLOFs from 1988 to 2016. ID corresponds to labels in Fig. 8. Elevations were extracted from SRTM data. Type of drainage is either complete (C) or partial (P). Column sums are given by Σ , and column means by μ .

ID	Country	Long [°]	Lat [°]	Elevation [m a.s.l.]	Degree of confidence	Area change manually mapped [m ²]	Area change predicted by CPA [m ²]	Relative difference manual vs. predicted [%]	Maximum age	Minimum age	Difference between noise free images	Type of drainage
A	China	85.479	28.663	5196	High	48,600	36,900	−24.1	1995-11-01	1996-10-02	336	P
B	China	86.447	27.929	5192	Medium	109,800	72,900	−33.6	1992-09-22	1992-11-01	40	P
C	Nepal	86.709	27.95	4730	Medium to high	107,100	107,100	0	1993-11-20	1994-08-19	272	C
D	Nepal	86.776	27.955	5015	High	13,500	10,800	−20.0	1998-11-02	1999-08-01	272	C
E	Nepal	86.844	27.793	5309	Medium	44,100	41,400	−6.1	2006-10-07	2008-09-26	720	P
F	China	90.226	28.28	5301	High	115,200	136,800	+18.75	2007-11-21	2008-09-20	304	P
G	Bhutan (1st)	90.325	28.088	4706	High	57,600	52,200	−9.4	1990-11-14	1991-09-30	320	P
	(2nd)			4706	High	23,400	20,700	−11.5	2009-11-18	2010-10-04	320	P
H	Bhutan	90.419	27.904	5142	High	11,700	7200	−38.5	1997-11-01	1998-11-04	368	P
J	China	94.324	30.685	4117	High	18,000	21,600	+20.0	2002-10-24	2003-10-11	352	P
						$\Sigma = 441,900$	$\Sigma = 400,500$	$\mu = -9.4$				$\mu = 330$

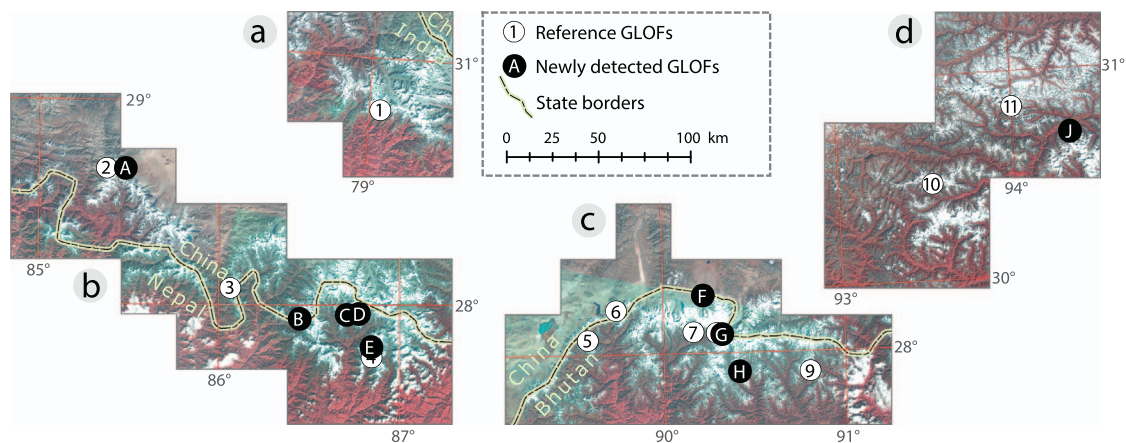


Fig. 8. All reference GLOFs (numbers 1–11) and newly detected GLOFs (letters A–J).

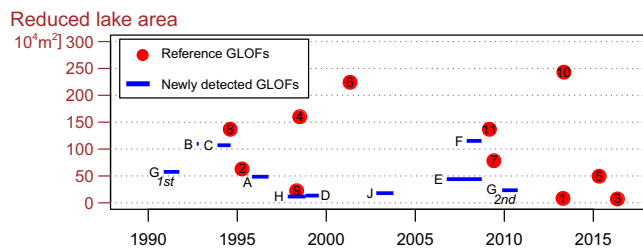


Fig. 9. Comparison of drained lakes areas linked to previously documented (reference; Table 1) and newly detected GLOFs in a training area covering $\sim 10\%$ of the HKH. Blue bars give intervals between noise-free images for newly detected GLOFs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Land is one of the few empirical steps in our processing chain, and may need further testing for larger areas. Our processing chain substantially decreases the influence of user-dependent decisions compared to other change-point algorithms that depend on a wide range of empirical parameters (Jamali et al., 2015; Kennedy et al., 2010). With the current parameter setting, we seem to overestimate the number of change pixels (Table 6), causing low PA of changed area (Table 7). Increasing k may decrease time-consuming visual validation, but also increase the risk of overlooking GLOFs. By contrast, any automated detection becomes difficult with less than three consecutive observations. The 2013 GLOF from Lake Chorabari, India, remained undetected because no noise-free image was available between lake filling and drainage (Fig. 12).

Future work may involve a more flexible handling of time steps such that k becomes more (less) sensitive in regions where the number of observations in the pixel time series is low (high). The Chorabari time series also demonstrates how glacial lakes change their size without producing outburst floods. Desiccation, infilling, or subsurface drainage are processes that shrink or remove proglacial lakes in the HKH. Similarly, ponds on debris-covered glaciers can grow and shrink annually (Miles et al., 2017; Watson et al., 2016), though these and other, monsoonal, changes are beyond the scope of this study, given the seasonal preference of our satellite imagery.

Finally, determining whether a lake has drained catastrophically remains a time-consuming interactive step at the end of our processing chain. In most cases, Landsat imagery was our only data source for identifying tell-tale sediment tails. The resolution of Landsat data is often insufficient to reveal details about sediment dynamics in the outburst paths. In this context, the newly available Sentinel-2 data improve the situation for detecting GLOF events, offering a higher spatial resolution (up to 10 m) and repeat rate of acquisition. Combining multi-sensor and multi-resolution datasets have become

popular to map high-mountain glacial outlines and flow velocities (Holzer et al., 2015; Kääb et al., 2015; Shukla et al., 2010), but may also unlock more regional detail of glacial lake dynamics and resolve more, probably smaller events. Other drainage mechanisms such as vertical drops in lake level or the emptying of subglacial water pockets elude our analysis of optical imagery, but may be resolved with laser altimetry or radar techniques (Smith et al., 2009; Wingham et al., 2006). In any case, the ten newly detected GLOFs redefine the lower envelope of their potential activity in the HKH region.

6. Conclusions

Our newly developed processing chain allows a robust retrospective detection of glacial lake outburst floods (GLOFs) from Landsat time series in the Hindu Kush Himalayan (HKH). Fuzzy land-cover classification with Random Forests and a likelihood-based change point algorithm promise high accuracies in detecting independently documented GLOFs. We successfully detected ten of eleven reported GLOFs, as well as ten previously unreported ones. Though we analysed only 10% of the HKH, we were able to expand the current GLOF inventory by 91% for the past three decades. Ubiquitous cloud, ice and shadows in the HKH lead to data gaps that prohibit a comprehensive tracking of glacial lakes at regular intervals. We used shrinking lakes and exposed sediment tails as key indicators for GLOF detection that we time-stamped with a change-point algorithm that computes the likelihood of change from three consecutive observations of water to land. Thus exploring all valid observations in the stack of land cover maps is a novel approach for detecting changes in pixel time series and holds promise for mountainous environments with patchy Landsat coverage. Data gaps and the limits of detection caused by 30 m resolution Landsat imagery call for multi-sensor and multi-resolution approaches including higher resolution imagery such as Sentinel-2 for future regional studies on GLOF detection. We robustly configured our processing chain using techniques of spatial-cross validation with bootstrapping for classifier training such that our methods are transferable to the whole HKH and mountain ranges elsewhere. The detection of ten previously unreported GLOFs confirms our initial hypothesis that the existing GLOF inventories underestimate significantly the number of GLOFs in the Himalayas. We thus set the basis towards a more complete GLOF inventory for the HKH, and towards objectively and systematically filling the gaps in the hitherto censored chronology of past GLOFs.

Acknowledgements

This research was funded by Deutsche Forschungsgemeinschaft (DFG) within the graduate research training group NatRiskChange (GRK 2043/1) at the University of Potsdam (<http://www.>

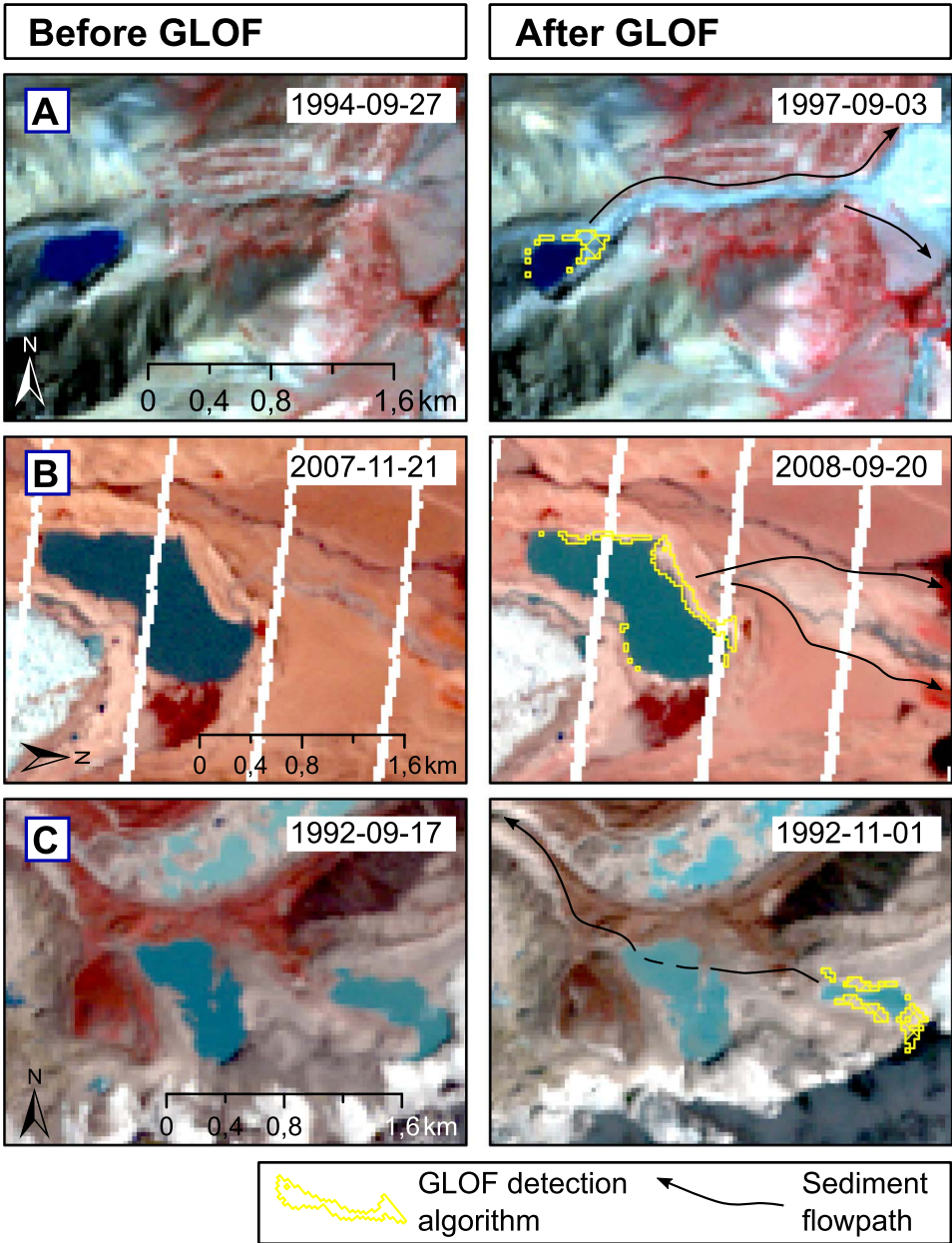


Fig. 10. Examples of newly detected GLOFs with high confidence (A and B) and medium confidence (C). Note the bright reflectance of sediments from the partially drained lakes in panels A and B (see Table 8-A and F). Panel C shows a distinct decrease of lake area, but uncertain outburst path. The change in colour of the lake downstream implies a change in sediment concentration as an indicator for the GLOF event.

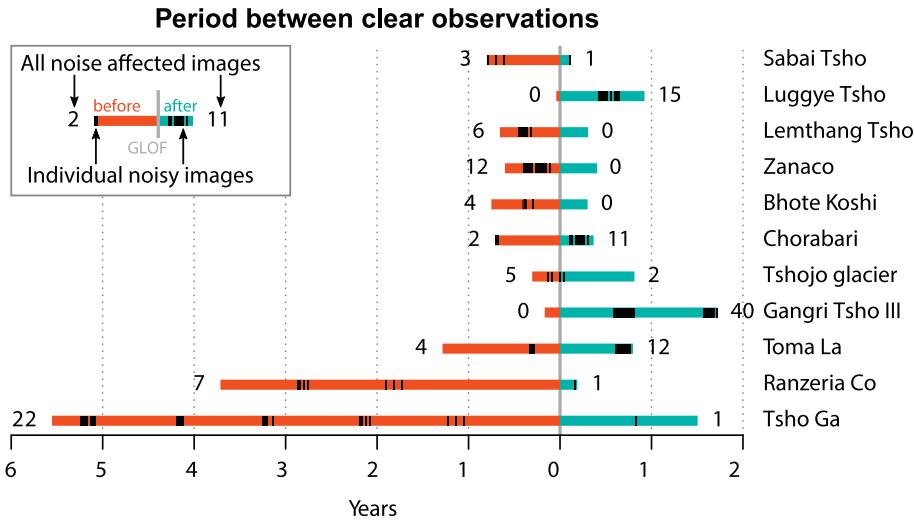


Fig. 11. Period in years between clear images before (red) and after (cyan) the reference GLOFs. Numbers at the end of the bars indicate the number of unsuitable images due to noise induced by clouds, shadows or lake freezing. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

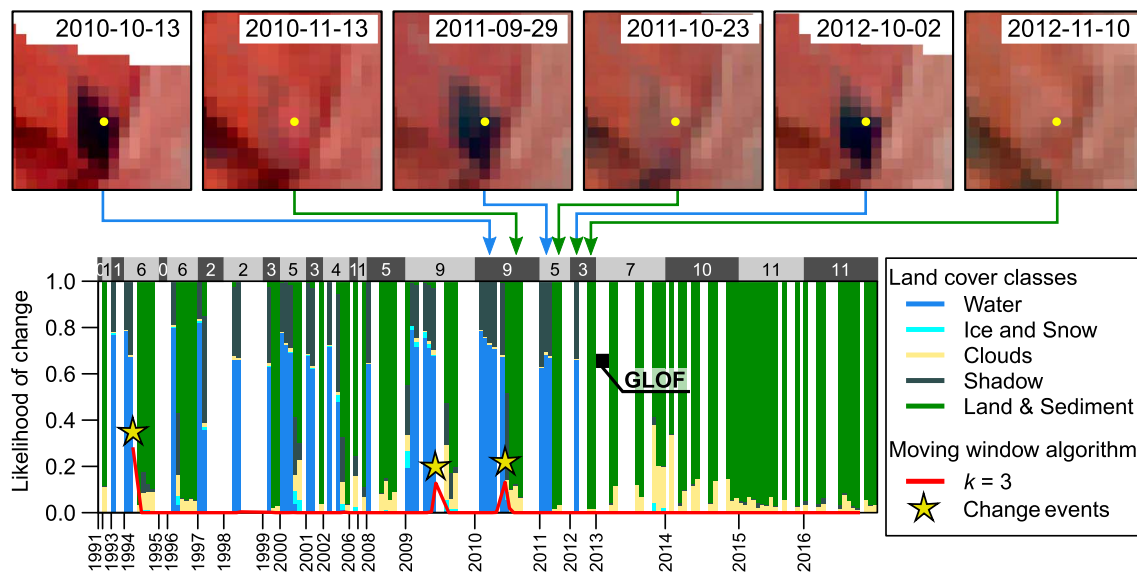


Fig. 12. Noise-filtered pixel time series from Lake Chorabari (Table 1-1), extracted from yellow location pixel in top-row panels. The algorithm correctly detected several changes from water to land (stars), but missed the GLOF because of noisy data in that year. A higher number of noise-free observations per year (see grey-shaded bar) allows better change detection. The lake had appeared and disappeared several times before the GLOF. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

natriskchange.de). SRTM and Landsat data were available from the U.S. Geological Survey. Landsat imagery was preprocessed via the Earth Resources Observation and Science (EROS) Center Science Processing Architecture (ESPA). Our computations were entirely based on the statistical software R (<http://www.r-project.org/>). We thank Adam Emmer and two anonymous referees for their helpful comments and Berry Boessenkool for computational support.

References

- Allen, S.K., Rastner, P., Arora, M., Huggel, C., Stoffel, M., 2016. Lake outburst and debris flow disaster at Kedarnath, June 2013: hydrometeorological triggering and topographic predisposition. *Landslides* 13, 1479–1491. <http://dx.doi.org/10.1007/s10346-015-0584-3>.
- Bajracharya, B., Shrestha, A.B., Rajbhandari, L., 2007. Glacial lake outburst floods in the Sagarmatha region: hazard assessment using GIS and hydrodynamic modeling. *Mt. Res. Dev.* 27, 336–344. <http://dx.doi.org/10.1659/mrd.0783>.
- Bajracharya, S.R., Maharjan, S.B., Shrestha, F., 2011. Clean Ice and Debris Covered Glaciers of HKH Region.
- Belgiu, M., Drăguț, L., 2016. Random forest in remote sensing: a review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* 114, 24–31. <http://dx.doi.org/10.1016/j.isprsjprs.2016.01.011>.
- Bolch, T., Buchroithner, M.F., Peters, J., Baessler, M., Bajracharya, S., 2008. Identification of glacier motion and potentially dangerous glacial lakes in the Mt. Everest region/ Nepal using spaceborne imagery. *Nat. Hazards Earth Syst. Sci.* 8, 1329–1340.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45, 5–32.
- Brun, F., Berthier, E., Wagnon, P., Käb, A., Treichler, D., 2017. A spatially resolved estimate of High Mountain Asia glacier mass balances from 2000 to 2016. *Nat. Geosci.* 10, 668–673. <http://dx.doi.org/10.1038/ngeo2999>.
- Carrivick, J.L., Tweed, F.S., 2016. A global assessment of the societal impacts of glacier outburst floods. *Glob. Planet. Chang.* 144, 1–16. <http://dx.doi.org/10.1016/j.gloplacha.2016.07.001>.
- Cenderelli, D.A., Wohl, E.E., 2001. Peak discharge estimates of glacial-lake outburst floods and “normal” climatic floods in the Mount Everest region, Nepal. *Geomorphology* 40, 57–90.
- Che, T., Xiao, L., Liou, Y.-A., 2014. Changes in glaciers and Glacial Lakes and the identification of dangerous Glacial Lakes in the Pumqu River basin, Xizang (Tibet). *Adv. Meteorol.* 2014, 1–8. <http://dx.doi.org/10.1155/2014/903709>.
- Clague, J.J., Evans, S.G., 2000. A review of catastrophic drainage of moraine-dammed lakes in British Columbia. *Quat. Sci. Rev.* 19, 1763–1783.
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.* 8, 1991–2007. <http://dx.doi.org/10.5194/gmd-8-1991-2015>.
- Cook, K., Andermann, C., Gimbert, F., Hovius, N., Adhikari, B., 2017. Impacts of the 2016 outburst flood on the Bhote Koshi River valley, central Nepal. *Geophys. Res. Abstr.* 19.
- Das, S., Kar, N.S., Bandyopadhyay, S., 2015. Glacial lake outburst flood at Kedarnath, Indian Himalaya: a study using digital elevation models and satellite images. *Nat. Hazards* 77, 769–786. <http://dx.doi.org/10.1007/s11069-015-1629-6>.
- DeVries, B., Verbesselt, J., Kooistra, L., Herold, M., 2015. Robust monitoring of small-scale forest disturbances in a tropical montane forest using Landsat time series. *Remote Sens. Environ.* 161, 107–121. <http://dx.doi.org/10.1016/j.rse.2015.02.012>.
- Emmer, A., Vilimek, V., Huggel, C., Klimeš, J., Schaub, Y., 2016. Limits and challenges to compiling and developing a database of glacial lake outburst floods. *Landslides*. <http://dx.doi.org/10.1007/s10346-016-0686-6>.
- Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. In: *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*. Presented at the 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96), pp. 226–231.
- Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dilley, R.D., Beckmann, T., Schmidt, G.L., Dwyer, J.L., Joseph Hughes, M., Laue, B., 2017. Cloud detection algorithm comparison and validation for operational Landsat data products. *Remote Sens. Environ.* 194, 379–390. <http://dx.doi.org/10.1016/j.rse.2017.03.026>.
- Footy, G.M., Boyd, D.S., 1999. Detection of partial land cover change associated with the migration of inter-class transitional zones. *Int. J. Remote Sens.* 20, 2723–2740. <http://dx.doi.org/10.1080/014311699211769>.
- Fraser, R., Olthof, I., Kokelj, S., Lantz, T., Lacelle, D., Brooker, A., Wolfe, S., Schwarz, S., 2014. Detecting landscape changes in high latitude environments using Landsat trend analysis: 1. Visualization. *Remote Sens.* 6, 11533–11557. <http://dx.doi.org/10.3390/rs6111533>.
- Fujita, K., Suzuki, R., Nuimura, T., Sakai, A., 2008. Performance of ASTER and SRTM DEMs, and their potential for assessing glacial lakes in the Lunana region, Bhutan Himalaya. *J. Glaciol.* 54, 220–228.
- Fujita, K., Sakai, A., Takenaka, S., Nuimura, T., Surazakov, A.B., Sawagaki, T., Yamanokuchi, T., 2013. Potential flood volume of Himalayan glacial lakes. *Nat. Hazards Earth Syst. Sci.* 13, 1827–1839. <http://dx.doi.org/10.5194/nhess-13-1827-2013>.
- Gardelle, J., Arnaud, Y., Berthier, E., 2011. Contrasted evolution of glacial lakes along the Hindu Kush Himalaya mountain range between 1990 and 2009. *Glob. Planet. Chang.* 75, 47–55. <http://dx.doi.org/10.1016/j.gloplacha.2010.10.003>.
- Gardelle, J., Berthier, E., Arnaud, Y., Käb, A., 2013. Region-wide glacier mass balances over the Pamir-Karakoram-Himalaya during 1999–2011. *The Cryosphere* 7, 1263–1286. <http://dx.doi.org/10.5194/tc-7-1263-2013>.
- Gimbert, F., Cook, K., Andermann, C., Hovius, N., Turowski, J., 2017. Using seismic arrays to quantify the physics of a glacial outburst flood and its legacy on upland river dynamics. *Geophys. Res. Abstr.* 19.
- Griffiths, P., Kuemmerle, T., Baumann, M., Radeloff, V.C., Abrudan, I.V., Lieskovsky, J., Munteanu, C., Ostapowicz, K., Hostert, P., 2014. Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites. *Remote Sens. Environ.* 151, 72–88. <http://dx.doi.org/10.1016/j.rse.2013.04.022>.
- Gurung, D.R., Khanal, N.R., Bajracharya, S.R., Tsering, K., Joshi, S., Tshering, P., Chhetri, L.K., Lotay, Y., Penjor, T., 2017. Lemthang Tsho glacial Lake outburst flood (GLOF) in Bhutan: cause and impact. *Geoenviron. Disaster* 4. <http://dx.doi.org/10.1186/s40677-017-0080-2>.
- Haeblerli, W., Schaub, Y., Huggel, C., 2017. Increasing risks related to landslides from degrading permafrost into new lakes in de-glaciating mountain ranges. *Geomorphology* 293, 405–417. <http://dx.doi.org/10.1016/j.geomorph.2016.02.009>.
- Harrison, S., Kargel, J.S., Huggel, C., Reynolds, J., Shugar, D.H., Betts, R.A., Emmer, A.,

- Glasser, N., Haritashya, U.K., Klimeš, J., Reinhardt, L., Schaub, Y., Willyshire, A., Regmi, D., Vilímek, V., 2017. Climate change and the global pattern of moraine-dammed glacial lake outburst floods. *Cryosphere Discuss.* 1–28. <http://dx.doi.org/10.5194/tc-2017-203>.
- Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., 2015. An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. *Remote Sens. Environ.* 158, 220–234. <http://dx.doi.org/10.1016/j.rse.2014.11.005>.
- Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Campbell, L.B., 2016. Mass data processing of time series Landsat imagery: pixels to data products for forest monitoring. *Int. J. Digital Earth* 1–20. <http://dx.doi.org/10.1080/17538947.2016.1187673>.
- Holzer, N., Vijay, S., Yao, T., Xu, B., Buchroithner, M., Bolch, T., 2015. Four decades of glacier variations at Muztagh Ata (eastern Pamir): a multi-sensor study including Hexagon KH-9 and Pleiades data. *Cryosphere* 9, 2071–2088. <http://dx.doi.org/10.5194/tc-9-2071-2015>.
- Ives, J.D., Shrestha, R., Mool, P., et al., 2010. Formation of Glacial Lakes in the Hindu Kush-Himalayas and GLOF Risk Assessment. ICIMOD Kathmandu.
- Jain, S.K., Lohani, A.K., Singh, R.D., Chaudhary, A., Thakural, L.N., 2012. Glacial lakes and glacial lake outburst flood in a Himalayan basin using remote sensing and GIS. *Nat. Hazards* 62, 887–899. <http://dx.doi.org/10.1007/s11069-012-0120-x>.
- Jamali, S., Jönsson, P., Eklundh, L., Årdö, J., Seaquist, J., 2015. Detecting changes in vegetation trends using time series segmentation. *Remote Sens. Environ.* 156, 182–195. <http://dx.doi.org/10.1016/j.rse.2014.09.010>.
- Kääb, A., Berthier, E., Nuth, C., Gardelle, J., Arnaud, Y., 2012. Contrasting patterns of early twenty-first-century glacier mass change in the Himalayas. *Nature* 488 (7412), 495–498.
- Kääb, A., Treichler, D., Nuth, C., Berthier, E., 2015. Brief Communication: Contending estimates of 2003–2008 glacier mass balance over the Pamir-Karakoram-Himalaya. *Cryosphere* 9, 557–564. <http://dx.doi.org/10.5194/tc-9-557-2015>.
- Kattelmann, R., 2003. Glacial lake outburst floods in the Nepal Himalaya: a manageable hazard? *Nat. Hazards* 28, 145–154.
- Kennedy, R.E., Yang, Z., Cohen, W.B., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms. *Remote Sens. Environ.* 114, 2897–2910. <http://dx.doi.org/10.1016/j.rse.2010.07.008>.
- Komori, J., Koike, T., Yamanokuchi, T., Tshering, P., 2012. Glacial lake outburst events in the Bhutan Himalayas. *Glob. Environ. Res.* 16, 59–70.
- Korup, O., Tweed, F., 2007. Ice, moraine, and landslide dams in mountainous terrain. *Quat. Sci. Rev.* 26, 3406–3422. <http://dx.doi.org/10.1016/j.quascirev.2007.10.012>.
- Lamsal, D., Sawagaki, T., Watanabe, T., Byers, A.C., McKinney, D.C., 2015. An assessment of conditions before and after the 1998 Tam Pokhari outburst in the Nepal Himalaya and an evaluation of the future outburst hazard: Photogrammetric analysis of Tam Pokhari before and after its outburst. *Hydrol. Process.* <http://dx.doi.org/10.1002/hyp.10636>.
- Li, J., Sheng, Y., 2012. An automated scheme for glacial lake dynamics mapping using Landsat imagery and digital elevation models: a case study in the Himalayas. *Int. J. Remote Sens.* 33, 5194–5213. <http://dx.doi.org/10.1080/01431161.2012.657370>.
- Lippitt, C.D., Rogan, J., Li, Z., Eastman, J.R., Jones, T.G., 2008. Mapping selective logging in mixed deciduous forest. *Photogramm. Eng. Remote Sens.* 74, 1201–1211.
- Liu, J.-J., Cheng, Z.-L., Su, P.-C., 2014. The relationship between air temperature fluctuation and glacial lake outburst floods in Tibet, China. *Quat. Int.* 321, 78–87. <http://dx.doi.org/10.1016/j.quaint.2013.11.023>.
- Martinez, S., Gould, W.A., González, O.M.R., 2007. Creating Cloud-free Landsat ETM + Data Sets in Tropical Landscapes: Cloud and Cloud-shadow Removal.
- McFeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* 17, 1425–1432. <http://dx.doi.org/10.1080/01431169608948714>.
- Metternicht, G., 1999. Change detection assessment using fuzzy sets and remotely sensed data: an application of topographic map revision. *ISPRS J. Photogramm. Remote Sens.* 54, 221–233.
- Miles, E.S., Willis, I.C., Arnold, N.S., Steiner, J., Pellicciotti, F., 2017. Spatial, seasonal and interannual variability of supraglacial ponds in the Langtang Valley of Nepal, 1999–2013. *J. Glaciol.* 63, 88–105. <http://dx.doi.org/10.1017/jog.2016.120>.
- 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 Sens. Environ.* 174, 341–352. <http://dx.doi.org/10.1016/j.rse.2015.11.003>.
- Nie, Y., Sheng, Y., Liu, Q., Liu, L., Liu, S., Zhang, Y., Song, C., 2017. A regional-scale assessment of Himalayan glacial lake changes using satellite observations from 1990 to 2015. *Remote Sens. Environ.* 189, 1–13. <http://dx.doi.org/10.1016/j.rse.2016.11.008>.
- Nitze, I., Grosse, G., 2016. Detection of landscape dynamics in the Arctic Lena Delta with temporally dense Landsat time-series stacks. *Remote Sens. Environ.* 181, 27–41. <http://dx.doi.org/10.1016/j.rse.2016.03.038>.
- O'Connor, J.E., Clague, J.J., Walder, J.S., Manville, V., Beebe, R.A., 2013. Outburst floods. In: *Treatise on Geomorphology*. Elsevier, pp. 475–510.
- Olofsson, P., Foody, G.M., Stehman, S.V., Woodcock, C.E., 2013. Making better use of accuracy data in land change studies: estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* 129, 122–131. <http://dx.doi.org/10.1016/j.rse.2012.10.031>.
- Olthoff, I., Fraser, R.H., Schmitt, C., 2015. Landsat-based mapping of thermokarst lake dynamics on the Tuktoyaktuk Coastal Plain, Northwest Territories, Canada since 1985. *Remote Sens. Environ.* 168, 194–204. <http://dx.doi.org/10.1016/j.rse.2015.07.001>.
- Osti, R., Egashira, S., 2009. Hydrodynamic characteristics of the Tam Pokhari glacial lake outburst flood in the Mt. Everest region, Nepal. *Hydrol. Process.* 23, 2943–2955. <http://dx.doi.org/10.1002/hyp.7405>.
- Parker, J.A., Kenyon, R.V., Troxel, D.E., 1983. Comparison of interpolating methods for image resampling. *IEEE Trans. Med. Imaging* 2, 31–39.
- Pfeffer, W.T., Arendt, A.A., Bliss, A., Bolch, T., Cogley, J.G., Gardner, A.S., Hagen, J.-O., Hock, R., Kaser, G., Kienholz, C., Miles, E.S., Moholdt, G., Mölg, N., Paul, F., Radić, V., Rastner, P., Raup, B.H., Rich, J., Sharp, M.J., 2014. The Randolph Glacier Inventory: a globally complete inventory of glaciers. *J. Glaciol.* 60, 537–552. <http://dx.doi.org/10.3189/2014JG13J176>.
- Plugmacher, D., Cohen, W.B., Kennedy, R.E., 2012. Using Landsat-derived disturbance history (1972–2010) to predict current forest structure. *Remote Sens. Environ.* 122, 146–165. <http://dx.doi.org/10.1016/j.rse.2011.09.025>.
- Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. *Remote Sens. Environ.* 48, 119–126.
- Richardson, S.D., Reynolds, J.M., 2000. An overview of glacial hazards in the Himalayas. *Quat. Int.* 65, 31–47.
- Riggs, G.A., Hall, D.K., Salomonson, 1994. A snow index for the Landsat Thematic Mapper and Moderate Resolution Imaging Spectroradiometer. In: Presented at the Geoscience and Remote Sensing Symposium, 1994. IGARSS '94. Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation. vol. 4. pp. 1942–1944. <http://dx.doi.org/10.1109/IGARSS.1994.399618>.
- Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS J. Photogramm. Remote Sens.* 67, 93–104. <http://dx.doi.org/10.1016/j.isprsjprs.2011.11.002>.
- Rouse Jr., J., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring Vegetation Systems in the Great Plains With ERTS.
- Rover, J., Ji, L., Wylie, B.K., Tieszen, L.L., 2012. Establishing water body areal extent trends in interior Alaska from multi-temporal Landsat data. *Remote Sens. Lett.* 3, 595–604. <http://dx.doi.org/10.1080/01431161.2011.643507>.
- Scherler, D., Munack, H., Mey, J., Eugster, P., Wittmann, H., Codilean, A.T., Kubik, P., Strecker, M.R., 2014. Ice dams, outburst floods, and glacial incision at the western margin of the Tibetan Plateau: a > 100 ky chronology from the Shyok Valley, Karakoram. *Geol. Soc. Am. Bull.* 126, 738–758.
- Schroeder, T.A., Wulder, M.A., Healey, S.P., Moisen, G.G., 2011. Mapping wildfire and clearcut harvest disturbances in boreal forests with Landsat time series data. *Remote Sens. Environ.* 115, 1421–1433. <http://dx.doi.org/10.1016/j.rse.2011.01.022>.
- Shrestha, B.B., Nakagawa, H., Kawaike, K., Baba, Y., Zhang, H., 2013. Glacial hazards in the Rolwaling valley of Nepal and numerical approach to predict potential outburst flood from glacial lake. *Landslides* 10, 299–313. <http://dx.doi.org/10.1007/s10346-012-0327-7>.
- Shukla, A., Arora, M.K., Gupta, R.P., 2010. Synergistic approach for mapping debris-covered glaciers using optical thermal remote sensing data with inputs from geomorphometric parameters. *Remote Sens. Environ.* 114, 1378–1387. <http://dx.doi.org/10.1016/j.rse.2010.01.015>.
- Smith, B.E., Fricker, H.A., Joughin, I.R., Tulaczyk, S., 2009. An inventory of active subglacial lakes in Antarctica detected by ICESat (2003–2008). *J. Glaciol.* 55, 573–595.
- Song, C., Sheng, Y., Ke, L., Nie, Y., Wang, J., 2016. Glacial lake evolution in the south-eastern Tibetan Plateau and the cause of rapid expansion of proglacial lakes linked to glacial-hydrogeomorphic processes. *J. Hydrol.* 540, 504–514. <http://dx.doi.org/10.1016/j.jhydrol.2016.06.054>.
- Song, C., Sheng, Y., Wang, J., Ke, L., Madson, A., Nie, Y., 2017. Heterogeneous glacial lake changes and links of lake expansions to the rapid thinning of adjacent glacier termini in the Himalayas. *Geomorphology* 280, 30–38. <http://dx.doi.org/10.1016/j.geomorph.2016.12.002>.
- Sun, M., Liu, S., Yao, X., Li, J., 2014. The cause and potential hazard of glacial lake outburst flood occurred on July 5, 2013 in Jiali County, Tibet. *J. Glaciol. Geocryol.* 36, 158–165.
- Tulbure, M.G., Broich, M., Stehman, S.V., Kommareddy, A., 2016. Surface water extent dynamics from three decades of seasonally continuous Landsat time series at sub-continental scale in a semi-arid region. *Remote Sens. Environ.* 178, 142–157. <http://dx.doi.org/10.1016/j.rse.2016.02.034>.
- Verbesselt, J., Zeileis, A., Herold, M., 2012. Near real-time disturbance detection using satellite image time series. *Remote Sens. Environ.* 123, 98–108. <http://dx.doi.org/10.1016/j.rse.2012.02.022>.
- Wang, W., Yao, T., Gao, Y., Yang, X., Kattel, D.B., 2011. A first-order method to identify potentially dangerous glacial lakes in a region of the southeastern Tibetan Plateau. *Mt. Res. Dev.* 31, 122–130. <http://dx.doi.org/10.1659/MRD-JOURNAL-D-10-00059.1>.
- Wang, X., Liu, S., Ding, Y., Guo, W., Jiang, Z., Lin, J., Han, Y., 2012. An approach for estimating the breach probabilities of moraine-dammed lakes in the Chinese Himalayas using remote-sensing data. *Nat. Hazards Earth Syst. Sci.* 12, 3109–3122. <http://dx.doi.org/10.5194/nhess-12-3109-2012>.
- Wang, W., Gao, Y., Iribarren Anaconda, P., Lei, Y., Xiang, Y., Zhang, G., Li, S., Lu, A., 2015a. Integrated hazard assessment of Cirenmaco glacial lake in Zhangzangbo valley, Central Himalayas. *Geomorphology*. <http://dx.doi.org/10.1016/j.geomorph.2015.08.013>.
- Wang, W., Xiang, Y., Gao, Y., Lu, A., Yao, T., 2015b. Rapid expansion of glacial lakes caused by climate and glacier retreat in the Central Himalayas. *Hydrol. Process.* 29, 859–874. <http://dx.doi.org/10.1002/hyp.10199>.
- Waske, B., Braun, M., 2009. Classifier ensembles for land cover mapping using multi-temporal SAR imagery. *ISPRS J. Photogramm. Remote Sens.* 64, 450–457. <http://dx.doi.org/10.1016/j.isprsjprs.2009.01.003>.
- Watanbe, T., Rothacher, D., 1996. The 1994 Lugge Tsho glacial lake outburst flood, Bhutan Himalaya. *Mt. Res. Dev.* 16, 77. <http://dx.doi.org/10.2307/3673897>.
- Watson, C.S., Quincey, D.J., Carrivick, J.L., Smith, M.W., 2016. The dynamics of

- supraglacial ponds in the Everest region, central Himalaya. *Glob. Planet. Chang.* 142, 14–27. <http://dx.doi.org/10.1016/j.gloplacha.2016.04.008>.
- Weiss, A., 2001. Topographic position and landforms analysis. Presented at the Poster presentation. In: ESRI User Conference, San Diego, CA.
- Wingham, D.J., Siegert, M.J., Shepherd, A., Muir, A.S., 2006. Rapid discharge connects Antarctic subglacial lakes. *Nature* 440, 1033–1036. <http://dx.doi.org/10.1038/nature04660>.
- Worni, R., Huggel, C., Stoffel, M., 2013. Glacial lakes in the Indian Himalayas — from an area-wide glacial lake inventory to on-site and modeling based risk assessment of critical glacial lakes. *Sci. Total Environ.* 468–469, S71–S84. <http://dx.doi.org/10.1016/j.scitotenv.2012.11.043>.
- Xu, L., Li, J., Brenning, A., 2014. A comparative study of different classification techniques for marine oil spill identification using RADARSAT-1 imagery. *Remote Sens. Environ.* 141, 14–23. <http://dx.doi.org/10.1016/j.rse.2013.10.012>.
- Yamada, T., Sharma, C.K., 1993. Glacier lakes and outburst floods in the Nepal Himalaya. *IAHS Publ.* 218, 319–330.
- Yamanokuchi, T., Tadono, T., Komori, J., Kawamoto, S., Tomiyama, N., 2011. Temporal monitoring of supraglacial lakes on Tshojo Glacier at Bhutan. In: IEEE International Geoscience and Remote Sensing Symposium, Vancouver, Canada, Paper.
- Yao, X., Liu, S., Sun, M., Wei, J., Guo, W., 2012. Volume calculation and analysis of the changes in moraine-dammed lakes in the north Himalaya: a case study of Longbasaba lake. *J. Glaciol.* 58, 753–760. <http://dx.doi.org/10.3189/2012JoG11J048>.
- Zevenbergen, L.W., Thorne, C.R., 1987. Quantitative analysis of land surface topography. *Earth Surf. Process. Landf.* 12, 47–56.
- Zhang, G., Yao, T., Xie, H., Wang, W., Yang, W., 2015. An inventory of glacial lakes in the third pole region and their changes in response to global warming. *Glob. Planet. Chang.* 131, 148–157. <http://dx.doi.org/10.1016/j.gloplacha.2015.05.013>.
- Zhu, Z., Woodcock, C.E., 2012. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sens. Environ.* 118, 83–94. <http://dx.doi.org/10.1016/j.rse.2011.10.028>.
- Zhu, Z., Wang, S., Woodcock, C.E., 2015. Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and sentinel 2 images. *Remote Sens. Environ.* 159, 269–277. <http://dx.doi.org/10.1016/j.rse.2014.12.014>.
- Zhu, Z., Gallant, A.L., Woodcock, C.E., Pengra, B., Olofsson, P., Loveland, T.R., Jin, S., Dahal, D., Yang, L., Auch, R.F., 2016. Optimizing selection of training and auxiliary data for operational land cover classification for the LCMAP initiative. *ISPRS J. Photogramm. Remote Sens.* 122, 206–221. <http://dx.doi.org/10.1016/j.isprsjprs.2016.11.004>.