

Mapping Retrogressive Thaw Slumps Using Single-Pass TanDEM-X Observations

Philipp Bernhard[✉], *Graduate Student Member, IEEE*, Simon Zwieback[✉], Silvan Leinss[✉], *Member, IEEE*, and Irena Hajnsek, *Fellow, IEEE*

Abstract—Vast areas of the Arctic host ice-rich permafrost, which is becoming increasingly vulnerable to terrain-altering thermokarst in a warming climate. Among the most rapid and dramatic changes are retrogressive thaw slumps. These slumps evolve by a retreat of the slump headwall during the summer months, making them detectable by comparing digital elevation models over time using the volumetric change as an indicator. Here, we present and assess a method to detect and monitor thaw slumps using time series of elevation models applied on two contrasting study areas in Northern Canada. Our two-step method is tailored to single-pass InSAR observations from the TanDEM-X satellite pair, which have been acquired since 2011. For each acquisition, we derive a digital elevation model and uncertainty estimates. In the first step, we difference digital elevation models and detect the significant elevation changes using a blob-detection algorithm. In the second step, we classify the detections into those due to thaw slumps and other causes using a simple thresholding method (accuracy: 78%), a random forest classifier (87%), and a support vector machine (86%). When our method is applied to other areas, the classifiers should be trained with data from part of the study area or with data obtained from similar areas in terms of topography, vegetation, and thaw slump characteristics to achieve the best performance. The obtained locations of thaw slumps can be used as a starting point to extract important slump properties, such as the headwall height and the volumetric change, which are currently not available on regional scales.

Index Terms—Abrupt thaw, banks island, digital elevation model (DEM) differencing, DEM generation, interferometry, Mackenzie river delta, permafrost, remote sensing, retrogressive thaw slumps (RTSs), single-pass radar interferometry, synthetic aperture radar (SAR), thermokarst.

I. INTRODUCTION

ABOUT one-quarter of the landmass in the Northern Hemisphere are underlain by permafrost, which is becoming increasingly vulnerable to rapid thaw in a warming climate [1], [2]. Rapid permafrost degradation has major impacts on the local hydrology and ecosystems, and it can also reinforce climate

change by mobilizing organic carbon, leading to the emission of the greenhouse gases CO₂ and CH₄. One important land surface characteristic arising from rapid thaw are retrogressive thaw slumps (RTS). RTSs are characterized by a steep headwall, which can reach several tens of meters in height (see Fig. 1) [3]. During the summer, the ice in the headwall melts and sediments get transported downslope, leading to a continuous retreat of the headwall. In the context of recent warming, an increase in the rates and size of RTSs in permafrost regions of Northern Canada and Alaska has been found [3]–[7]. Nevertheless, on the pan-Arctic scale, the prevalence and rates of thaw slumping remain poorly constrained, and so does their contribution to climate change. This is mainly due to the remote landscape and the severe climate conditions in the Arctic, making remote sensing techniques highly important for studying RTSs in these areas.

In recent years, remote sensing techniques have been explored to map and investigate RTSs on large scales ranging from aerial photographs [8], laser scanning (LiDAR) [9], [10], and satellites in the optical and microwave domain [10]–[14]. The most detailed and highest resolution can be achieved using aerial photographs and LiDAR measurements from airplanes, but at the disadvantage of high costs and no pan-Arctic coverage. A large-scale coverage of vast lands is only possible by remote sensing satellites. In this regard, optical satellites can detect active RTSs by identifying disturbed soil due to the movement of the headwall and the transport of sediments downslope. [15], [16]. Due to the need for cloud-free observation as well as solar irradiation, the amount of usable data is limited. Additionally, the implementation of RTS detection methods is especially difficult in the vegetation sparse Arctic tundra regions.

Another option for measuring changes over time are synthetic aperture radar (SAR) systems, which have great potential for a pan-Arctic monitoring of permafrost degradation. Since these systems use an active radar, there is no need for solar irradiation, making it possible to observe the Arctic landscape year around and are furthermore independent of cloud cover. To measure elevation changes over time, differential as well as topographic approaches have been used. In the differential approach, SAR observations are taken from the same position at different times (repeat-pass InSAR) and measure ground movement on a millimeter- to centimeter-scale, e.g., thaw subsidence by permafrost thaw [17], [18] or glacier movement [19] under the assumption that surface properties do not change significantly between acquisitions. Since RTSs can change strongly over

Manuscript received December 19, 2019; revised March 12, 2020 and May 4, 2020; accepted May 26, 2020. Date of publication June 8, 2020; date of current version June 25, 2020. (Corresponding author: Philipp Bernhard.)

Philipp Bernhard and Silvan Leinss are with the Institute of Environmental Engineering, ETH Zurich, 8093 Zurich, Switzerland (e-mail: bernhard@ifu.baug.ethz.ch; leinss@ifu.baug.ethz.ch).

Simon Zwieback is with the Geophysical Institute, University of Alaska Fairbanks, Fairbanks, AK 99775 USA (e-mail: szwieback@alaska.edu).

Irena Hajnsek is with the Institute of Environmental Engineering, ETH Zurich, 8093 Zurich, Switzerland, and also with the German Aerospace Center (DLR) e.V. Microwaves and Radar Institute, 82234 Wessling, Germany (e-mail: irena.hajnsek@dlr.de).

Digital Object Identifier 10.1109/JSTARS.2020.3000648



Fig. 1. Aerial photograph of an RTS close to Inuvik (photo by Simon Zwieback taken in summer 2018) with headwall and scar zone.

timescales of months, investigating the headwall movement over a multiyear time frame is not feasible using repeat-pass interferometry due to the effect of temporal decorrelation on the signal [11].

The topographic InSAR approach from single-pass observations can provide large-scale maps of thaw slump activity and volumetric changes. Here, two observations are taken with a spatial separation (baseline) between the satellites. The phase difference between the two observations can then be used to generate digital elevation models (DEM). In the single-pass bistatic mode, two satellites fly close to each other and take observations at the same time where one antenna emits a radar signal that gets then observed by both. This has the advantage that the temporal decorrelation as well as the atmospheric phase contribution are greatly reduced. By comparing DEMs that are generated from a series of single-pass observations, temporal elevation changes can be measured. DEMs from single-pass InSAR are a unique but largely untapped data source for monitoring thaw slumps. Till now, the single-pass method has been mainly used for the quantification of volumetric changes after volcanic eruptions [20] and the estimation of glacier mass balances [21].

Pan-Arctic repeated single-pass InSAR data have been acquired by the TanDEM-X pair since 2011. TanDEM-X is a high-resolution single-pass interferometry satellite mission that was launched by the German Aerospace Center (DLR), with the purpose of generating a high-resolution global DEM [22]. The satellite pair started observations in 2011 and are still operational until today. From 2011 to 2018, the global land areas are observed at least three times. Both satellites carry an active synthetic radar (SAR) operating in the X-band at a wavelength of 3.11 cm. A planimetric resolution after averaging of about 10–12 m and, depending on the distance between the satellites, vertical height resolutions of the order of about 2 m can be achieved. This accuracy is better than a typical RTS headwall height of 5–10 m and can provide accurate estimates of the thaw slump topography.

The TanDEM-X archive provides unparalleled opportunities for mapping thaw slumps and their environmental impacts on large scales. By using DEM time series, RTS drivers and controls

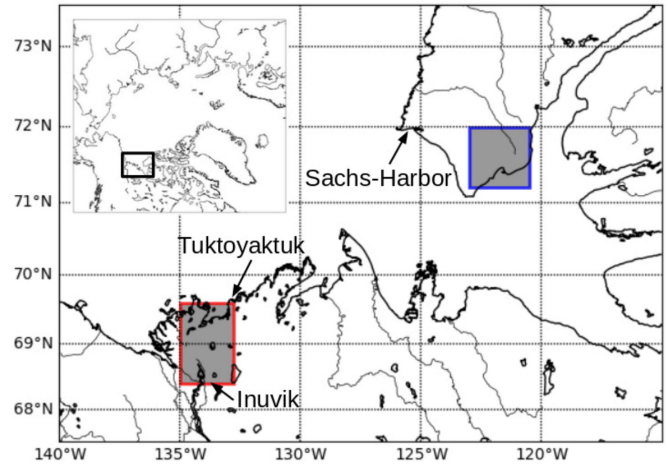


Fig. 2. Outline of the study area in the Northwest Territories in Northern Canada. The red and blue rectangles correspond to the MRD and the Banks study sites, respectively.

for their development and growth can be investigated. The possibility of measuring the induced volumetric change makes it potentially possible to estimate the amount of mobilized and displaced materials, including organic carbon, nutrients, and sediment. Since the TanDEM-X observations cover the whole pan-Arctic landscape, a method to extract these information has the potential to greatly improve our knowledge of RTSs activity on large scales. The detection and classification of elevation changes in the Arctic is an important step toward this goal.

Here, we develop and assess an RTS mapping procedure using data obtained by the TanDEM-X satellites. For each observation, we generate a DEM and estimate the uncertainties. In a first step, we difference the generated DEMs to track elevation changes over time. The induced volumetric changes of RTSs with small headwall heights and low retreat rates are of the order of the systematic errors. We implement methods to deal with errors due to low coherences and waterbodies and apply a detection algorithm on the generated elevation change maps. Nevertheless, several detections are not related to active RTSs. In a second step, we use statistical and machine learning classification methods to remove detections that are not caused by active RTSs. To validate our method, we use reference classifications from high-resolution optical data and apply our method on two contrasting study areas with different topographies, vegetation and RTS characteristics. To explore the applicability on large scales, we quantify the spatial transferability by training the classifiers on one area and evaluate it on the other. We additionally investigate the use of DEMs generated at different times of the year and assess the data availability and accuracy needed for large-scale mapping.

II. STUDY AREAS

For our study, we choose two contrasting areas in Northern Canada (see Fig. 2) that show differences in climatic conditions, vegetation, as well as topographic characteristics. The first area is limited by the Tuktoyaktuk Coastlands to the north, the start of the tree-covered region to the south (Inuvik), the Husky Lakes

to the East, and the Mackenzie River Delta in the west (in the following named as “MRD”). It spans about 8500 km². The second study area is located on the southeastern part of Banks Island with a size of about 7000 km² (in the following named as “Banks”). In the south and west, the Banks study area is limited by the Arctic Ocean. The soil in both study areas contains large amounts of ground ice, making it vulnerable to the development of RTSs [23], [24]. The RTSs in the two study areas differ in their morphology and size. In recent years, increases in RTS abundance and activity have been reported from both areas [4], [7], [12], [24].

The MRD study site is located in the low Arctic and the climatic conditions change from forested vegetation in the south toward a shrub tundra in the north. Due to the ice-rich ground, thermokarst features, such as thermokarst lakes, degraded high-centered polygons and active-layer detachment slides are common in the area. Likewise, RTSs are prevalent and occur mostly along lake shores. The largest thaw slumps can reach heights of up to 15 m and affect areas of several hectares in size. However, most slumps are much smaller with headwall heights of < 10 m and affected areas of < 2 hectare [25].

The Banks study site is located in the high Arctic and is only sparsely vegetated. The maximum elevation of 350 m is reached toward the center of the island. In the summer month, the water from snow melt and precipitation flows from the more mountainous area toward the sea in trenches and valleys. In winter, winds redistribute the snow leading to large snow accumulations. Due to the glacial and periglacial history, the ground has a high ice content leading to a widespread development of thermokarst features. The typical headwall heights of RTSs ranges from 2–15 m [23], [26].

III. DEM GENERATION

At the core of our processing chain are TanDEM-X bistatic single-pass observations, delivered by the German Aerospace Center (DLR). To obtain DEMs with reasonable vertical accuracy, we exclude observations with height of ambiguities (HoA) larger than 80 m. The incidence angles range from 36°–44°. All observations were taken in an ascending orbit with look direction to the right in single-polarization mode (VV or HH). In total, our dataset contains 149 bistatic single-pol TanDEM-X observations (MRD: 87, Banks: 62) acquired between 2011 and 2017. These observations cover our study areas with about four to five observations of which at least one observation is taken during winter before 2013 and one during winter 2016/2017 (see Fig. 3). We generate DEMs for each observation with a final planimetric resolution of 10–12 m and vertical accuracies of about 2 m in areas with high coherences. To improve our DEM generation process, we additionally use the TanDEM-X 12-m DEM product as reference DEM. The interferometric processing was done using the *Gamma Remote Sensing* software [27]. All further processing steps are implemented using the programming language *Python*. An overview of the processing chain can be seen in Fig. 4.

For each bistatic single-pass TanDEM-X observation, we started with the single-look complex (CoSSC) product delivered

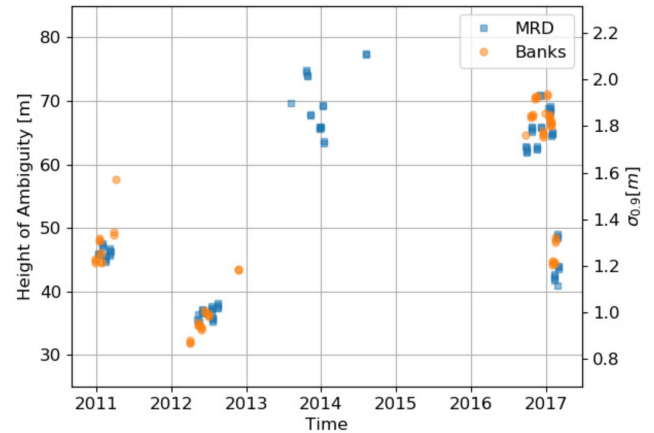


Fig. 3. HoA and the expected standard error corresponding to an assumed coherence of 0.9. The lowest HoAs are around 30–40 m. During winter of 2010/2011, the spring and summer of 2012, and twice during the winter 2016/2017, both study areas have a total coverage. Parts of the MRD study site was also observed, in 2013 and 2014.

by DLR and verified the coregistration using a cross-correlation algorithm. We compute the interferograms after applying range spectral filtering to allow for the incorporation of large baseline interferograms. To reduce speckle noise, we apply a multilooking window of size 4 pixels \times 4 pixels, reducing the original resolution from 2–3 m to 8–12 m. The complex coherence was computed based on the flattened interferogram where the phase contribution due to a flat topography was removed. Additionally, we compute a simulated interferogram as well as a layover and shadow mask based on the orbit information and the reference DEM. To minimize errors due to phase unwrapping, we first subtract the observed interferogram from the simulated one before applying the minimum cost flow algorithm [28]. The unwrapped phases are converted to elevation heights in meters. We use the orbit information for the orthorectification and apply the layover and shadow mask to remove areas that are affected by layover and shadow effects. The obtained DEMs contain height offsets and tilts, as well as horizontal misalignments on the order of a few meters. We do an initial correction of these two errors using the reference DEM to obtain an estimate of the absolute vertical offset by computing the mean difference. We fit a second degree polynomial to the spacial difference to remove tilts.

The DEMs were split into tiles to facilitate the comparison of DEMs acquired from different orbits. The tiles have an extent of about 13 \times 13 km with a buffer of 1 km to avoid edge effects.

IV. ACCOUNTING FOR RANDOM AND SYSTEMATIC ERRORS

When comparing DEMs generated from single-pass interferometric observations, the elevation differences contain several error sources that have to be considered. These can arise due to errors in the DEM generation, such as an inaccurate DEM registration and remaining large-scale trends, a random error in each DEM, related to phase inaccuracies in the individual observations as well as systematic elevation biases related to water, snow, or vegetation. These effects can lead to apparent

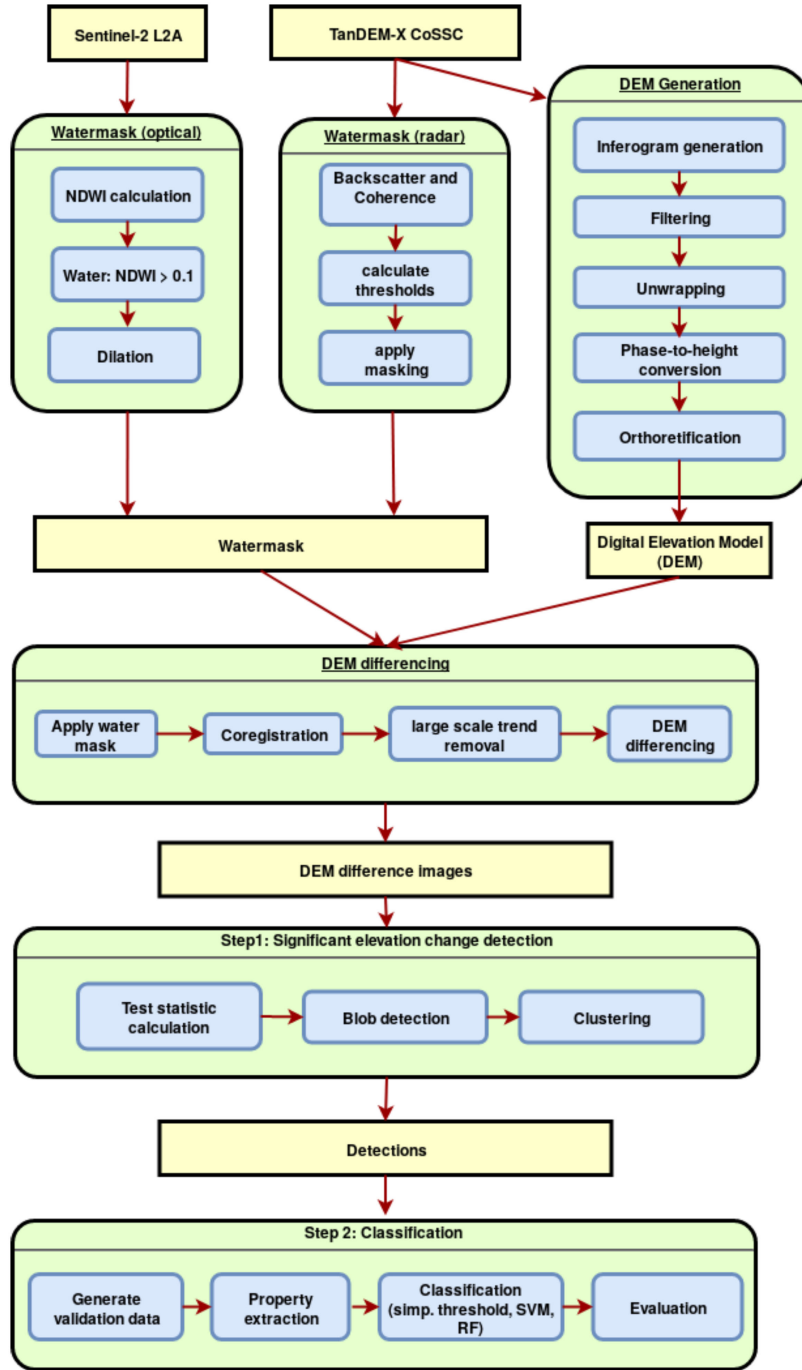


Fig. 4. Overview of our processing chain. We use the TanDEM-X CoSSC data as input for the DEM generation process and the waterbody mask generation. Additionally, the Sentinel-2 Level 2 A (L2A) product is used for the waterbody mask generation in areas where there is no TanDEM-X summer observation available. After applying DEM differencing, we detect, in a first step, significant elevation changes. This includes the computation of a test statistic, a blob detection, and a clustering algorithm. In a second step, the detections are classified to find elevation losses that are caused by RTSs. Here, three classifiers, namely a simple threshold method, a support vector machine (SVM), and a random forest (RF) classifier, are used.

elevation changes in single-pass interferometric observations of several meters.

A. High-Precision DEM Registration

Our generated DEMs can contain weak large-scale trends, an absolute height offset, and misalignments. We tried to account

for these errors during the DEM generation process, but in areas experiencing large random error, such as over water bodies, small trends, offsets, and misalignments remain. We thus employ a second DEM registration step where we again try to remove these errors.

We correct for misalignments in a first step using the Nuth and Kääb method, which exploits the shift-induced relation

between the elevation differences and terrain aspect [29]. To achieve a good result, the random errors need to be smaller than the shift signal, and prior to the correction, we apply a waterbody and coherence mask (see Section IV-B and IV-D) to remove areas with large errors. We obtain good registration results if topographic features, such as deep trenches and large and steep hills, are present but the registration becomes error prone for areas with only little topography. For these areas, mostly in the MRD study area, we additionally orthorectify the backscatter intensities and remove the misalignment using a cross-correlation algorithm. This improves the registration in these areas. After correcting for the obtained shift, we again apply a large-scale trend removal and correct for a vertical offset. These steps are repeated until no misalignments and shifts can be detected.

B. Random Height Errors

The heights are affected by random errors that are due to noise in the interferometric phases. The coherence γ is a measure of the quality of the interferometric phase between the two observations and takes values between 0 (low phase quality) and 1 (high phase quality). For the single-pass bistatic TanDEM-X observations, interferometric coherence loss due to temporal decorrelation, ambiguities, baseline decorrelation, and relative shifts of the Doppler spectra are highly suppressed ($\gamma > 0.95$) and the main contribution to the decorrelation is due to volume decorrelation and low backscatter intensities (low signal-to-noise ratio) [22], [30].

The coherence estimate can be translated into an estimate on the elevation error using the Cramer–Rao bound

$$\sigma_h^\gamma = \frac{\sqrt{1 - \gamma^2}}{\gamma\sqrt{2L}} \frac{\text{HoA}}{2\pi} \quad (1)$$

where HoA is the height of ambiguity, defined as the height change in one interferometric phase cycle and L is the number of looks to reduce speckle noise.

Since we are normalizing our height measurements in the generation of the test statistic image, we also include measurements with low coherence. Due to the large errors in the elevation measurements at very low coherence values, which become problematic for phase unwrapping and for the Cramer–Rao bound equation (1) to be applicable, we mask for each observation the pixels with an estimated coherence of less than 0.3 [31].

C. Systematic Errors

When comparing DEMs generated at different times, water, snow, and vegetation can generate apparent elevation gains or losses of several meters. Water bodies induce large errors for several reasons. During the ice-free summer months, low backscatter intensities from the smooth water induce low coherences and a large random phase error. Additionally, the nonzero along-track baseline decorrelation can lead to spurious movement-induced phase contributions [25]. During winter, floating ice causes a strong reflectance due to the large differences between the dielectric properties of ice and water. This leads to high

values in the radar backscatter as well as the coherence values of the measured elevation at the ice–water interface. On the other side, if the ice freezes completely to the ground (bedfast ice), the dielectric contrast drops leading to low backscatter and coherence values [32]. Additionally, waterbodies can induce real changes in elevation by changes in the water level and the formation or drainage of lakes.

In terms of snow, the winter temperatures in the Arctic fall to average monthly temperature values of below -20°C resulting in a dry snow pack and radar waves can propagate through without being strongly affected, measuring the elevation at the ground. On the contrary during the time of snow melt, the water content in the snow pack increases and interacts with the radar waves by reflection and absorption. This leads to an apparent change in elevation as well as to a decrease in the coherence and thus in the accuracy of the height measurements [33].

In regard to vegetation, shrub heights along lake shores or along trenches and streams can reach several meters in height in the more southerly Arctic tundra regions. When the landscape is frozen, the microwaves interact little with the vegetation. In contrary, in the summer months, leaf growth and an increased water content in the vegetation can induce an apparent elevation change as well as large errors due to volume decorrelation [25].

Errors in DEMs generated from observations when the landscape is frozen can be dealt with by masking low coherence measurements and the utilization of a waterbody mask. In DEMs generated from observations taken when the landscape is not frozen, errors are more extensive and the best strategy is to remove these observations from the analysis. Due to the climatic conditions, in our study areas, we can in general assume a frozen landscape between October and May and therefore exclude observation taken in the months from April to September for the general RTS detection procedure. To investigate how strong the influence of these error sources are, we additionally apply our processing scheme on a dataset including the April to September observations.

D. Waterbody Mask

To remove errors due to waterbodies, we generate a waterbody mask. If a radar observation is available during the summer months (July and August), we apply a threshold-based masking, using the measured backscatter intensities as well as the measured coherence. This has the advantage of having a high resolution and a good coregistration to the DEMs. Due to the short time frame of an ice- and snow-free landscape, TanDEM-X observations for the generation of a water mask are only available for parts of the MRD study area. For the remaining parts, we use Sentinel-2 multispectral observations for the waterbody mask generation.

To generate the waterbody mask based on TanDEM-X observations, we determine our backscatter threshold value using the histogram of the measured backscatter intensities. For areas with a high abundance of lakes, we can identify two peaks, one coming from land surfaces (high) and the other coming from water surfaces (low). We set the minimum value between the two distributions as our threshold value [34]. If no peaks can

be identified, we use a threshold value of 0.4 of the average intensity. The coherence threshold was set to 0.3 [35].

For the optical waterbody mask, we select observations without cloud cover at the end of summer to avoid misclassification due to clouds and a frozen snow-covered landscape. We use the Normalized Difference Water Index with a threshold value of 0.1 to separate between water and nonwater [36], [37]. The optical waterbody masks have a resolution of 10×10 m, which is on the same order as the radar observations. Furthermore, the horizontal alignment of the TanDEM-X reference DEM and the Sentinel-2 observations are less than 20 m, such that no additional coregistration step is necessary [38]–[40]. To minimize errors due to horizontal misalignments and emergent vegetation in the littoral zone, we apply a morphological filter (dilation) of two pixels on the optical water masks to slightly increase the water area. If several Sentinel-2 observations are available, we generate a waterbody mask for each observation and combine them using an “OR” operation to obtain the final mask.

V. DETECTION AND CLASSIFICATION

We compute elevation changes by differencing the DEMs and apply a waterbody mask to remove water-induced artifacts. In a first step, we normalize the data using the estimated coherences and detect significant elevation changes using a blob detection method followed by a clustering algorithm to merge detections in the same locations over time. In a second step, we classify the detections using a simple threshold method and two machine learning classifiers, namely RF and SVM, to separate the detections into induced by active RTSs and other causes.

A. Step 1: Detection of Significant Elevation Changes

To detect significant elevation changes, we apply DEM differencing on our data. Since the largest changes occur between DEMs taken further apart in time, we compute all possible difference maps between DEMs generated by observations separated by more than two years.

The error on the measured elevation depends on several properties and can be estimated by σ_h^γ . We normalize the elevation changes using the Gaussian standard error

$$d^{\text{norm}} = \frac{h_1 - h_2}{\sqrt{\sigma_1(\hat{\gamma}_1, \text{HoA}_1)^2 + \sigma_2(\hat{\gamma}_2, \text{HoA}_2)^2}} \quad (2)$$

where d^{norm} is the normalized elevation difference and $h_{1,2}$ the measured elevation of the first and the second DEM, respectively. A pixel by pixel thresholding approach yields many wrong detections since headwall heights can be only of the order of a few meters and thus only slightly above the expected standard error. Since typical thaw slumps are larger than a resolution cell, we additionally use the spatial size of the slumps by computing a test score T over a moving window (Ω_k)

$$T_k = \frac{\sum_{i \in \Omega_k} d_i^{\text{norm}}}{n(\Omega_k)} \quad (3)$$

where n indicates computing the number of valid elements. We compute the test scores for square windows of size 3, 5, 8, and 15 pixels. On each of the test score images, we run a blob

TABLE I
PARAMETER SETS FOR THE BLOB DETECTION ALGORITHM

	σ_{\min}	σ_{\max}	σ_{num}	threshold [10^{-3}]
low	5, 6, 7, 9	35	50	4.5, 3.2, 2.5, 1.3
mid	6, 7, 8, 10	35	50	3.5, 2.5, 2., 1.
high	7, 8, 9, 11	35	50	3., 2., 1.5, 0.7

The multiple entries correspond to different window sizes of 3, 5, 8, and 15 pixels.

detection algorithm based on the determinant of the Hessian using a Gaussian kernel [41]. Here, four parameters have to be set: σ_{\min} indicating the minimum standard deviation of the Gaussian kernel, σ_{\max} for the maximum standard deviation of the Gaussian kernel, and σ_{num} for the number of intermediate values as well as a threshold based on typical RTS sizes (see Table I). The two most important parameters are σ_{\min} and the threshold, since these determine the smallest blobs in size and intensity that are detected. We choose three different settings (low, mid, high) to investigate the dependence of the parameters in relation to the number of RTS detections.

The exact location of the blob-detected RTSs can vary over time. To group blob detections corresponding to the same active RTS, we apply a hierarchical clustering scheme with a maximum cluster size of 40 pixel, corresponding to about 250 m [42]. Only very large slumps exceed this limit, which would lead to two or more cluster locations corresponding to the same RTS. The individual clusters are counted as one detection. We average the positions of the individual location of the blob detections inside a cluster to obtain the final detection location used in the further analysis. An example of the detection locations, elevation loss, and the computed test score for part of the MRD study area is shown in Fig. 5.

1) *Accuracy Assessment (Step 1)*: Two important quantities that can be used to evaluate the detection method are the positive predictive value (PPV), indicating the number of detected RTSs in relation to the total number of detection, as well as the false negative rate corresponding to the number of active RTSs that were not detected.

Knowledge of the proportion of RTS and non-RTS detections is important since a high number of the latter needs a more accurate second classification step. We thus compare the detection results for the different blob detection parameters based on the PPV. To estimate the PPV, we use high-resolution optical imagery time series (Sentinel-2 and Planet RapidEye) to determine if a detection is caused by active RTSs. Sentinel-2 data are available starting in 2015 with a 10-m resolution in the optical and infrared bands. The RapidEye satellites provides data over all our study time period with a 5-m resolution in the optical and infrared bands [43]. The criteria for classifying a detection as an active RTS are the exposure of bare soils, a movement over time, their location related to the possibility of sediment removal as well as the presence of a headwall [3]. This allocation into RTS and non-RTS detections is also used as class labels in the second classification step.

To get an estimate on the false-negative rate, we use reference data from the study [7] for the Banks study area. The Google Earth Engine Landsat time-laps dataset was used to identify

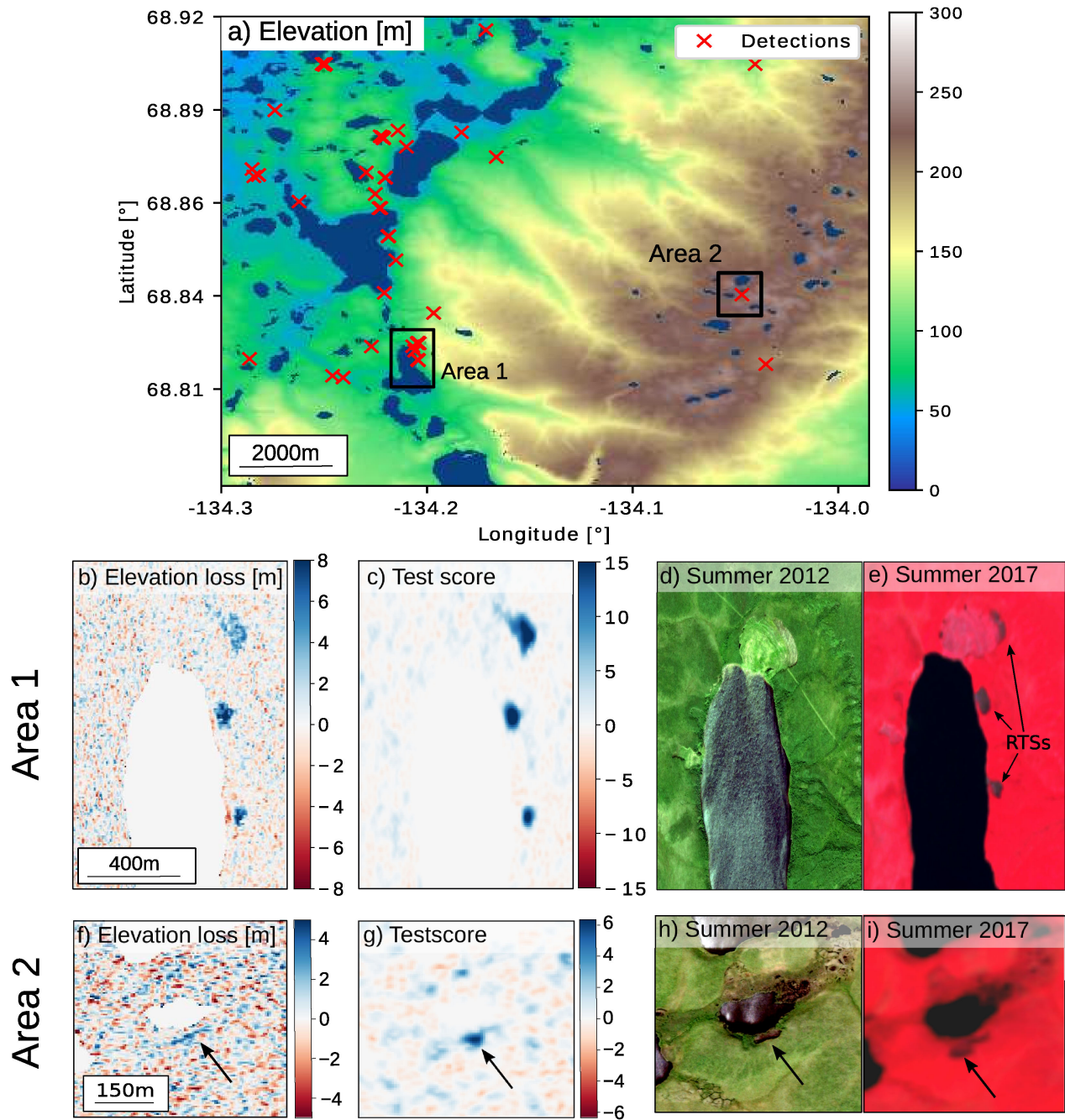


Fig. 5. Examples of Step 1: Detection of significant elevation changes. (a) Locations of detections for part of the MRD study region with the mid-parameter setting. (b) Elevation change between observations taken on February 15, 2011 and October 25, 2016. Three blue spots showing elevation losses are visible to the east of the masked lake. (c) Calculated test-statistic values. The three areas corresponding to strong elevation losses show high values in the test score statistic. (d) World-View Google Earth image from summer 2012. One RTS is visible at the top of the lake. (e) Sentinel-2 false color image taken on June 30, 2017. Compared to the World-View Image of 2012, three RTSs at the location show elevation losses. (f) Elevation change between observations taken on February 15, 2011 and October 25, 2016. The arrow indicates the location of an RTS visible in the Google Earth World-View image of 2017. (g) Calculated test-statistic values. The slumps along the lake is much better distinguishable and was detected using the mid parameter setting, but missed in the low setting. (h) World-View Google Earth image from summer 2012. (i) Sentinel-2 false color image taken on June 30, 2017. Compared to the World-View Image of 2012, a small increase in the area without vegetation is visible.

active slumps by their change over time (1985–2015). We include all RTSs that show signs of movement after the year 2011 (“Lewkowicz sample”). It is to note that due to the use of optical data, not every year is guaranteed to contain usable observations and that the cutoff between active and nonactive RTSs is governed by the relatively coarse resolution. The exact

year of stabilization and activation of RTSs in the Lewkowicz sample can thus be erroneous. Nevertheless, for each of the 982 RTSs in the Lewkowicz sample, we look for a counterpart in our detections. Some RTSs moved several hundred meters since their initiation or merged over time and the location given in the Lewkowicz sample and our detection location was separated

by a considerable amount. We again use optical data to follow the slump movement over time and assign the detections accordingly.

B. Step 2: Classification

To map active RTSs, we classify the previously detected candidates with significant elevation changes using three different classifiers. Our first choice for a classifier is a simple threshold method based on the total volumetric change. In recent years, the use of machine learning algorithms to classify remote sensing data showed an increase in the classification accuracies compared to previous methods [44]. Since our problem can be formulated as a supervised two-class classification problem with the additional limitation that the amount of trainings data is small, we choose two machine-learning classifiers, namely an RF and an SVM, that are generally among the best-performing classification algorithms in these type of classification problems [45]–[48]. For the implementation of the machine learning algorithms, we use the *Python scikit-learn* module [49]. The class labels (“RTS,” “non-RTS”) of the individual detections are determined by the accuracy assessment performed in Step 1.

1) *Feature Extraction*: We calculate several features at the detection location based on the DEM and the elevation change over time. To characterize the elevation change over time, we use the volume change between the first and last valid observations (V_{tot}) and the integrated volumetric change (V_{int}). The volumetric change between an observation m and an observation n in a window Ω , centered at the detection location, can be computed by

$$V_{m,n} = \sum_{i \in \Omega} (h_m^i - h_n^i) \quad (4)$$

where h_m^i and h_n^i indicate the elevations in pixel i measured by observation m and observation n , respectively. The total volumetric change V_{tot} in a window Ω is then defined as

$$V_{\text{tot}} = V_{1,n_{\text{obs}}} \quad (5)$$

where the available DEMs are ordered by observation date from 1 (earliest observation) to n_{obs} (latest observation). We computed V_{int} by subtracting all available observations from the earliest one and obtain the final value by integrating over time

$$V_{\text{int}} = \sum_{n=2}^{n_{\text{obs}}} \frac{V_{1,n-1} + V_{1,n}}{2} \cdot \Delta t_{n,n-1} \quad (6)$$

where $\Delta t_{n,n-1}$ is the number of seconds between observations n and $n-1$. The features based on the DEM values are the circular variance (ϕ_{var}), the mean slope (m_{mean}), and the maximum slope (m_{max}). The circular variance is defined as

$$\phi_{\text{var}} = 1 - \sqrt{\left(\sum_{i \in \Omega} \cos(\phi_i)\right)^2 + \left(\sum_{i \in \Omega} \sin(\phi_i)\right)^2} \quad (7)$$

where ϕ_i denotes the aspect in radians and Ω denotes an area around the location where ϕ_{var} is computed. ϕ_{var} is a measure of how variable the orientation of the slope vectors are with values close to 0 if all slope vectors point in the same direction and close to 1 for random orientations. The motivation to select this

TABLE II
OVERVIEW OF GENERATED FEATURES

Property	based on	sensitive to
V_{tot}	DEM change	volumetric change
V_{int}	DEM change	volumetric change
ϕ_{var}	DEM	trenches, flat areas, slopes
m_{max}	DEM	flat areas, slopes
m_{mean}	DEM	flat areas, slopes

properties is based on how active RTSs appear and evolve as well as to find typical sources for detections that are not caused by active RTSs. Since RTSs occur along lake and sea shores as well as at steep slopes due to the need of sediment removal they typically show a low circular variance and moderate to high values in the slope. On the contrary, wrong detections in flat and polygon covered areas show high values in the circular variance and low slope values. The circular variance is additionally sensitive to small trenches and depressions. An overview of the calculated features can be seen in Table II.

For the feature computation, the window size Ω has to be set. It should be big enough, such that detections due to large errors are distinguishable from small RTSs induced ones. If the window is too big, the elevation changes of small RTS movements do not show up in a significant way. To balance these two factors, we used a window size Ω of 5 pixels \times 5 pixels, corresponding to a square with length of approximately 20 m for the feature computation. To compute the features ϕ_{var} , m_{mean} , and m_{max} , we used the reference DEM as input for the computation. To avoid numerical difficulties in the training of the machine learning methods, we normalize our features to values around 0 ± 1 . ϕ_{var} , m_{mean} , and m_{max} can easily be expanded to $[-1, 1]$. V_{int} and V_{tot} have a more dynamic range and we use the mean and standard deviation of the training data to normalize them.

2) Classifiers:

a) *Thresholding on V_{tot} (Simple Threshold)*: In a first step, we analyze the use of a simple threshold on the total volumetric change. To determine the threshold value, we use the receiver operator characteristic (ROC) curve that is commonly used to evaluate binary classifiers [50]. The ROC curve is generated by plotting the sensitivity against $(1 - \text{Specificity})$ for different threshold settings. To determine the best threshold, we use the Youden index J defined as

$$\begin{aligned} J &= \max_t (\text{Sensitivity}(t) + \text{Specificity}(t) - 1) \\ &= \max_t (\text{TPR}(t) - \text{FPR}(t)) \end{aligned} \quad (8)$$

where t indicates the threshold, TPR the true positive rate, and FPR the false positive rate [51], [52].

b) *Random Forest (RF)*: Our first choice as a machine learning classifier is an RF classifier based on the generation and combination of ordinary binary classification trees [53]–[56]. In the initialization, hyperparameters, such as the number of decision trees (N_{tress}), maximum tree size (d_{max}), and a splitting criterion, need to be set. For setting the values of N_{tress} and d_{max} , we use a grid search on the training data. For the splitting criterion, we use the Gini index. We apply bootstrapping on the input data to reduce the variance and improve generalization [57].

TABLE III
RESULTS OF HYPERPARAMETER ESTIMATION USING
A FIVEFOLD CROSS VALIDATION

Trainings-Set	Classifier	Hyperparameters
Total	SVM	C: 1, γ : 1
	RF	N_{trees} : 1000, d_{max} : 5
	simp. thres.	t : 2897 m ³
Banks	SVM	C: 100, γ : 0.1
	RF	N_{trees} : 100, d_{max} : 7
	simp. thres.	t : 2705 m ³
MRD	SVM	C: 10, γ : 0.1
	RF	N_{trees} : 100, d_{max} : 10
	simp. thres.	t : 3479 m ³

The advantage of this classifier is the ability to understand the classification behavior by visualizing some of the decision trees as well as the possibility of computing a feature importance ranking.

c) *Support Vector Machine (SVM)*: As a second machine learning classifier, we use an SVM [58]–[60]. Here, each detection is characterized by n features and represented as a vector in a n -dimensional space. The classification is achieved by separating the points by an $n - 1$ dimensional hyperplane. The criterion for finding this hyperplane is to maximize the signed distance of the data points in each class to the hyperplane. Nonlinear decision boundaries can also be represented by an SVM, provided a suitable kernel function is used. For the kernel, we use a radial basis function with a free kernel parameter γ , which is a typical choice for SVMs. Additionally, a penalty parameter C has to be defined indicating the weight of the error term in the classification.

3) *Training and Hyperparameter Estimation*: We split our data into a training set containing 70% of the sample. The remaining 30% are used for evaluation. For all classifiers, several hyperparameters need to be set. We use a grid search approach with a fivefold cross validation to evaluate different options. The RF classifier was initialized with different number of trees N_{trees} : [3, 5, 10, 20, 100, 1000] and tree size d_{max} : [3, 5, 7, 10, 20, 100]. For the SVM, we use different options for the penalty parameter C : [0.1, 1, 10, 100, 1000, 10 000] and for the radial-basis kernel, the gamma parameter γ : [0.0001, 0.001, 0.01, 0.1, 1]. We also use a fivefold cross validation to determine the optimal threshold for the simple threshold method by averaging the results. The result of the hyperparameter estimation can be seen in Table III.

The RF and SVM classifiers are then trained on the whole training dataset. We evaluate the accuracy of the methods using the evaluation set. To investigate differences in the two study areas, we train and evaluate the classifiers on the two study areas separately.

Additionally, to infer the spatial transferability of the classifiers, we generate classifiers that are trained in one area and evaluate it on the other.

4) *Accuracy Assessment (Step 2)*: After the training of the classifiers on the training set, we assess the accuracy of the classifiers using the evaluation set. To assess the accuracy, the overall accuracy (OA), defined as the number of correctly classified detections divided by the total number of detections, the

TABLE IV
RESULT OF STEP 1 (CHANGE DETECTION) FOR THE TWO STUDY AREAS WITH
DIFFERENT BLOB DETECTION PARAMETER SETTINGS

Study Area	params	Total	N_{RTS}	PPV
Total	low	897	605	68%
	mid	1302	954	73%
	high	3837	-	-
Banks	low	777	551	71%
	mid	1075	813	76%
	high	3283	-	-
MRD	low	120	104	87%
	mid	227	141	62%
	high	554	186	34%

Due to the high number of detections for the high parameter setting on Banks, a manual assessment of the positive predictive value (PPV) was not feasible.

average accuracy (AA), defined as the average of the accuracies of the two classes, and the Cohen's kappa coefficient (κ), which measures the degree of classification agreement, are used [61].

VI. RESULTS

We first present the results of the detection prior to classification, and then the combined two-step approach. For the latter, we contrast the performance of the simple thresholding, RF, and SVM approaches and investigate the transferability across the study regions. We additionally investigate the influence of using DEMs, which are generated during the spring and summer, on the number of RTS and non-RTS detections.

A. Change Detection

In the first step, we investigate the dependence of the parameter settings low, mid, and high of the blob detection algorithm. Here, we focus on two objectives: First, we want to detect as many significant elevation changes due to active RTSs as possible; but second, limit the number of non-RTS detections since these have to be weed out in the following classification step.

The result of the change detection can be seen in Table IV. The total dataset is defined as the combination of the data from both study areas. With the low parameter setting, we found in total 897 detections. Of these detections, 605 could be related to active RTSs, corresponding to a PPV of 68%. For the midparameter setting, the number of detections increased to 1302 with an improvement of the PPV to 73%. Decreasing the threshold values further (high parameter setting), the number of detections increased significantly to 3837. On the Banks study site, it was not possible to manually classify the detections due to their high number. An overview of the location of the detections for the midparameter setting of the two study sites can be seen in Fig. 6. On MRD, non-RTS detections are typically found in vegetated trenches along streams and rivers, in polygon covered areas as well as along vegetated lake shores (for example, see Fig. 9). On the Banks study, these non-RTS detections are often found in deep valleys and along hillslopes and ridges.

In the MRD study area, anthropogenic activities, such as gravel pits, for road construction triggered significant height changes. We found 49 detections in the midparameter setting that are related to anthropogenic activities. Considering the whole

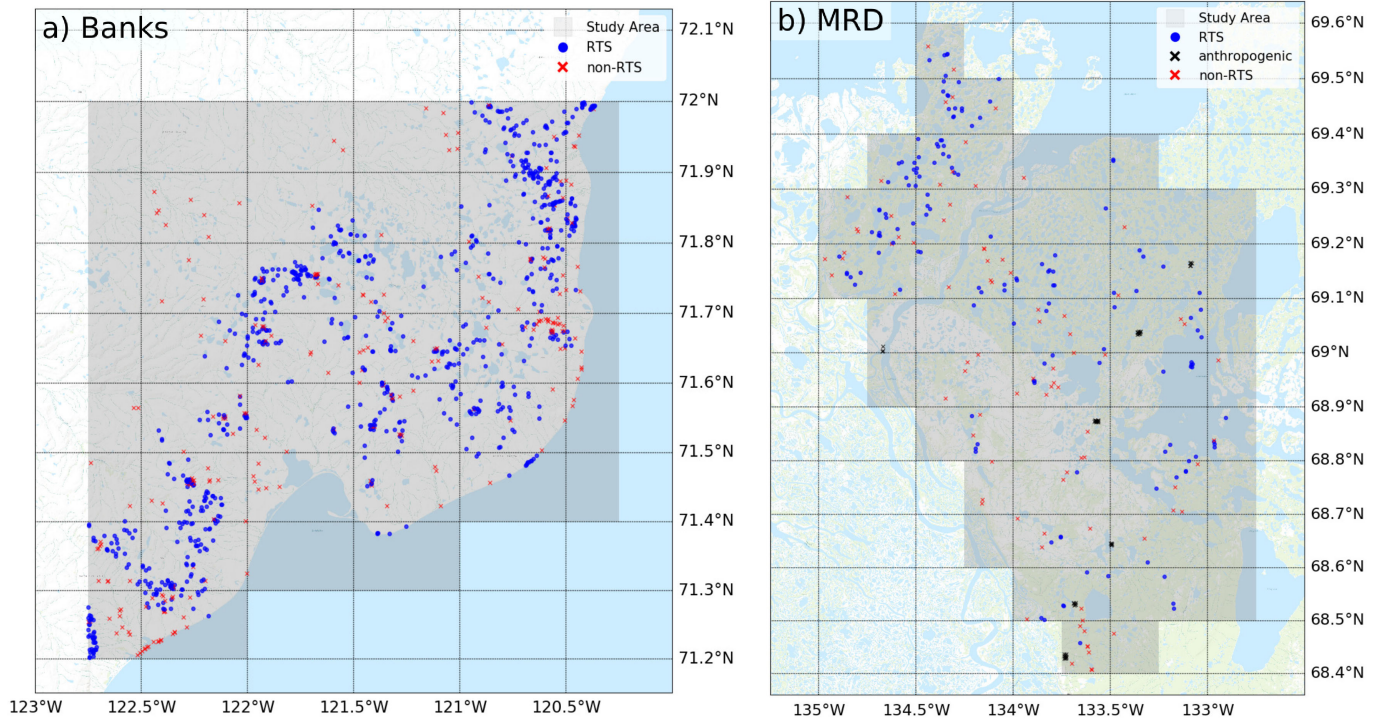


Fig. 6. Detection result with the midparameter setting (Step 1). (a) Banks study region and (b) MRD study region. On MRD, several detection could be attributed to anthropogenic causes, mainly due to gravel pits for road construction.

pan-Arctic landscape, such detections are very unusual and not a concern for large-scale applications and we removed these detections for the computation of the detection rates and the further analysis.

On Banks, we could additionally estimate the number of RTSs that were not detected. For the mid-parameter setting, we found that 719 of the 982 RTSs in the Lewkowicz sample had a matching detection in our sample (false negative rate: 26%). For the low parameter setting, the false negative rate increased to 52% (471/982 detected). Considering that typical RTSs on Banks have headwall heights of about 5–10 m with maximum retreat rates of 6–8 m per year, the calculated non-RTS detection rate of 26% for the midparameter setting indicates a good performance of our detection method [24], [26], [62], [63]. On MRD, no current RTS inventory is available, but past studies found typical headwall heights of about 3 m and mean retreat rates of about 2 m per year [4], [6], [25]. Since these values are just at the border of our detection limit, the 141 RTSs that could be detected show the potential of our method to be applied on areas with relatively small RTSs.

We used the midparameter setting for the further analysis since it provides the most balanced result between the detected RTS and the non-RTS detections. The following classification step is intended to weed out the non-RTS detections.

B. Classification

In the second step, we classify the detections that were found by the blob detection algorithm in Step 1 to improve the accuracy of our method. We investigated the performance of three

different classification approaches, which differ in their ability to distinguish between the detected significant elevation changes due to active RTSs and other causes (see Fig. 8). The SVM and RF machine learning approaches archived higher accuracies, but they were also less transferable across regions [see Fig. 10(a)].

1) *Thresholding on V_{tot} (Simple Thresholding)*: As a first classifier, we investigated the possibility of setting a threshold on the total volumetric change V_{tot} . The distribution of V_{tot} of RTS and non-RTS detection can be seen in Fig. 7(a). The distribution of the total volumetric change of RTS detection is shifted toward higher values than the non-RTS detection. We investigate the use of setting a threshold by calculating the true positive rate (TPR), false positive rate (FPR), and number of remaining detections (N_{tot}) after removing detection below the different thresholds values. We removed detections with $V_{\text{tot}}^i < V_{\text{tot}}^{\text{thres}}$ where $V_{\text{tot}}^{\text{thres}}$ indicates the threshold and V_{tot}^i indicates the value of the i th detection. The obtained result for different thresholds values can be seen in Fig. 7(b). In general, increasing the thresholds leads to a decrease in the number of classified thaw slumps. A strong increase and decrease in the number of true positives and false positives, respectively, is visible between 0 and about 3000 m³ indicating that mostly false positives are removed. To determine an optimal threshold, we used the ROC curve and compute the maximum value of the Youden J index [see Fig. 7(c)]. By using a fivefold cross validation, we found an average threshold value of $t = 2897$ m³ for the total dataset, $t = 3479$ m³ for MRD, and $t = 2705$ m³ for Banks. Applying these thresholds on the respective test sets yields overall accuracies of 0.78 for the total dataset and 0.83 for MRD and 0.75 for Banks (see Table V).

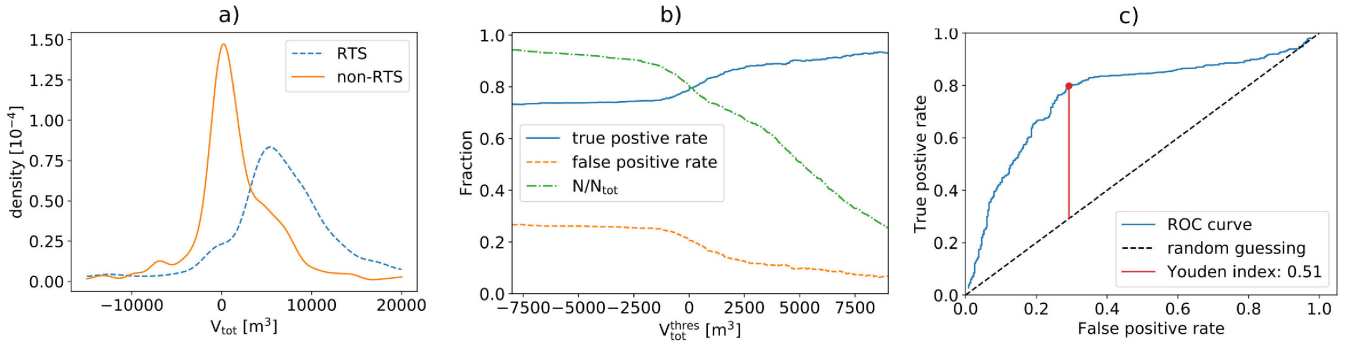


Fig. 7. Simple thresholding method for the total dataset. (a) Kernel density plot of the feature V_{tot} . The V_{tot} values corresponding to non-RTS detections show an earlier peak in the distribution than for the RTS detections. (b) False-positive rate, true positive rate, and normalized number of detections remaining after setting a threshold on the total volumetric change ($V_{\text{tot}}^{\text{thres}}$). (c) Plot of the ROC curve for the training dataset for the simple threshold method. The best threshold value was found at a Youden index value of $J = 0.51$ corresponding to 2897 m^3 .

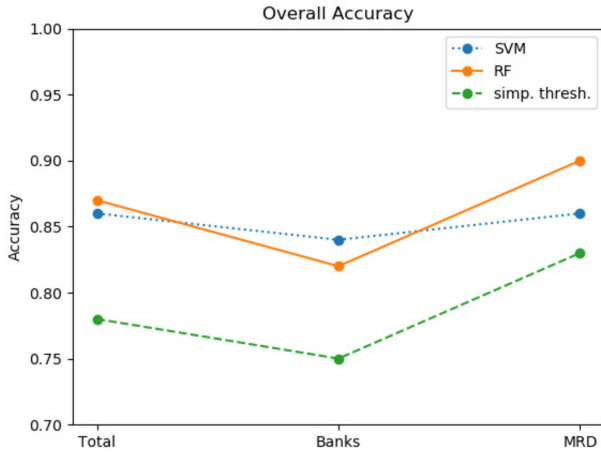


Fig. 8. Overall classification accuracies of the three classifiers for the total dataset (total), Banks, and MRD. The two machine learning classifiers (RF and SVM) perform better than the simple thresholding method. The classifier performance is better on MRD than on Banks.

TABLE V
ACCURACY ASSESSMENT OF SIMPLE THRESHOLDING, SVM, AND RF CLASSIFIERS WITH CALCULATED OA, AA, AND COHEN'S KAPPA COEFFICIENT (κ)

Dataset	Classifier	OA	AA	κ
Total	SVM	0.86	0.84	0.56
	RF	0.87	0.84	0.54
	simp. thres	0.78	0.64	0.47
MRD	SVM	0.86	0.87	0.65
	RF	0.90	0.88	0.82
	simp. thres	0.83	0.74	0.78
Banks	SVM	0.84	0.80	0.43
	RF	0.82	0.78	0.48
	simp. thres	0.75	0.68	0.44

2) *RF and SVM*: The results of the RF and SVM classifier can be seen in Table V. Both machine-learning classifiers show higher classification accuracies than the simple threshold method. The RF classifiers perform slightly better (OA = 0.87) than the SVM classifier (OA = 0.86) on the total dataset. When using the data of the two study areas separately, the MRD area shows a higher classification accuracy for both classifiers. On

TABLE VI
ACCURACY ASSESSMENT OF SVM, RF CLASSIFIERS, AND SIMPLE THRESHOLD METHOD FOR DIFFERENT TRAINING AND EVALUATION SETS WITH CALCULATED OA, AA, AND COHEN'S KAPPA COEFFICIENT (κ)

Training	Evaluation	Classifier	OA	AA	κ
MRD	Banks	SVM	0.73	0.63	0.27
		RF	0.61	0.58	0.12
		simp. thres	0.75	0.74	0.40
Banks	MRD	SVM	0.80	0.78	0.55
		RF	0.77	0.76	0.48
		simp. thres	0.82	0.82	0.62

Banks, the SVM performs slightly better than the RF classifier with OA of 0.84. The AA values are generally close to the overall accuracies indicating no bias in terms of a better classification of RTSs or non-RTSs. The calculated κ values are in the range of 0.42–0.82, indicating a moderate to good performance of the classifiers. An example area showing the result of the RF classifier can be seen in Fig. 9.

3) *Spatial Transferability*: To investigate the spatial transferability we applied the classifiers that are trained on one study site and used it to evaluate it on the other [see Fig. 10(a) and Table VI]. For the simple threshold method, the performance stays about the same, since the optimal threshold values for the two study sites are only minimally different. On the contrary, the accuracies of the SVM and RF classifiers drop below the simple threshold method with the RF showing a stronger reduction than the SVM.

For the RF classifier, the feature importance values for the different features can be calculated [see Fig. 10(b)]. On MRD, the most important feature is the total volumetric change followed by the features based on the DEM (ϕ_{var} , m_{max} , m_{mean}). On Banks, the total volumetric change as well as the integrated volumetric change shows a higher importance than the features based on the DEM.

On MRD, wrong classifications are mainly found in vegetated trenches and in areas showing signs of polygons (see Fig. 9). In contrary, RTSs occur mainly along lake shores. At the Banks study, area larger topographic features are present in the landscape, leading to non-RTS detections. Since RTSs

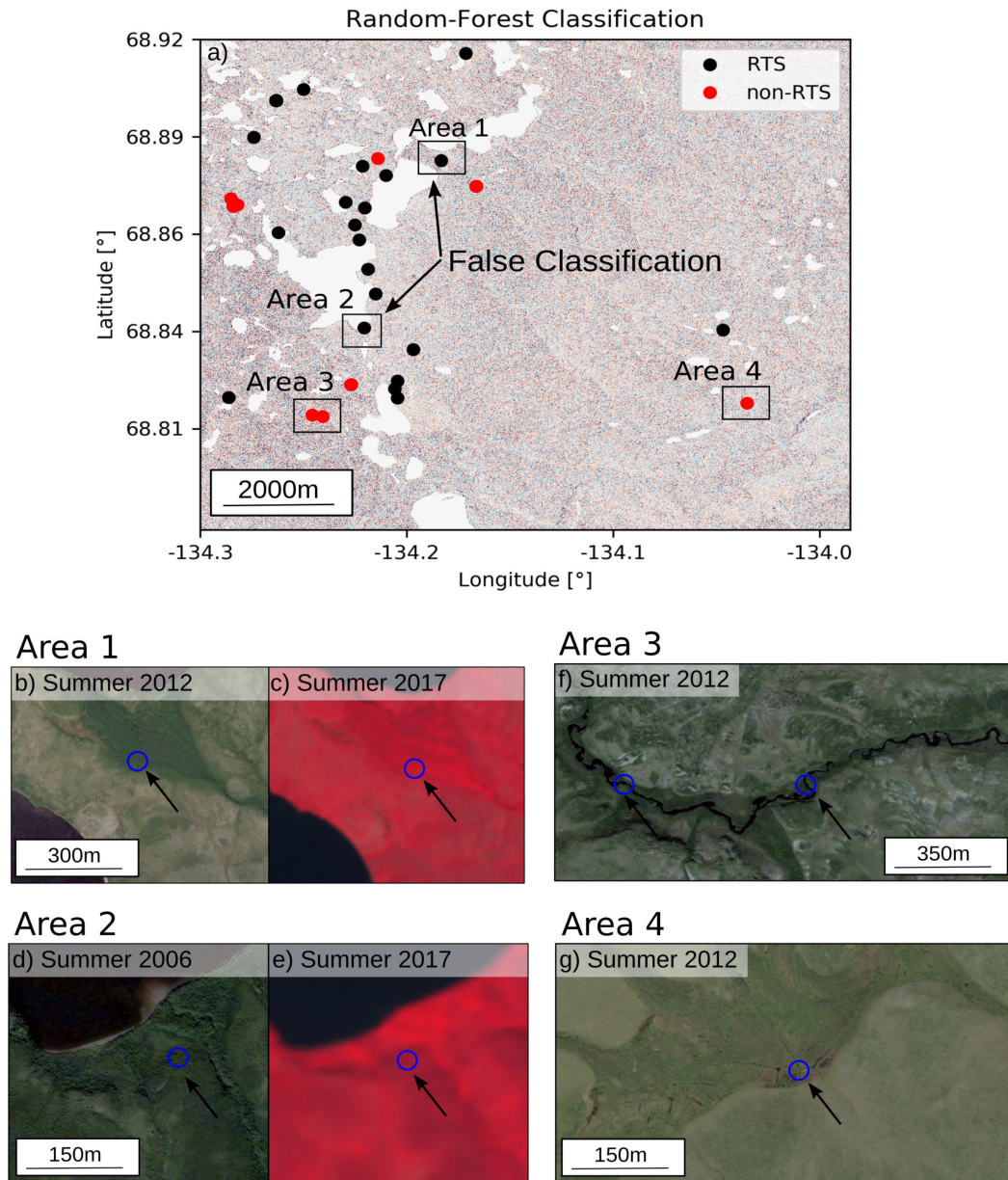


Fig. 9. Parts of the MRD area showing the result of the RF classification. Area 1 and Area 2 show two examples of typical false classifications. Area 3 and Area 4 show typical false detections that are classified correctly by the RF classifier. (a) Locations of detections and the generated classification in RTSs and non-RTSs by the RF classifier. The underlying image shows the elevation loss between February 15, 2011 and October 25, 2016. (b) Google Earth WorldView image of Area 1 taken in summer 2012. It shows a strongly vegetated area on the top, transitioning to a more sparsely vegetated area in the lower part. An old stabilized RTS is visible in the lower left part. The false classification as RTS (blue circle) is located at the transition between the strongly and sparsely vegetated area. The area around the detection shows a uniform slope gradient toward the NE. (c) Sentinel-2 false color image of Area 1. The uniform red color at the detection location indicates no bare soil. When compared to the WorldView Image (b) of 2012 no thaw slump is visible at the detection location (blue circle). (d) Google Earth WorldView image of Area 2 taken in summer 2006. The false classification as RTS is located at a vegetated area at a lake shore. When compared to the Sentinel-2 false color image of 2017 (e) no RTS activity exposing bare soil is visible. The location of the false classification is similar to locations where typical RTSs are found. (f) Google Earth WorldView Image of 2012 with two false detections (blue circle) along a small stream that are correctly classified as non-RTS by the RF classifier. (g) WorldView image taken in summer 2012 of Area 4. The false detection is located at a polygon covered flat area. The RF classifier correctly classified the detection as non-RTS.

occur mainly along hillslopes, the features based on the DEM properties cannot distinguish between RTSs and other detections. The dependence on the input features is also visible in the drop-off of the accuracies when the classifiers were trained on the MRD study area and evaluated on Banks. The DEM features are assigned high importance in the training but are not useful when applied on the detections found on Banks. Since the simple threshold method only uses the total volumetric change, it shows

a relatively constant performance, independent of what area was used for training and evaluation.

C. DEM Errors in Spring and Summer

Spring and summer acquisitions are less suitable for thaw slump mapping due to seasonal vegetation and late-lying snow packs that persistent long into the summer. On MRD several

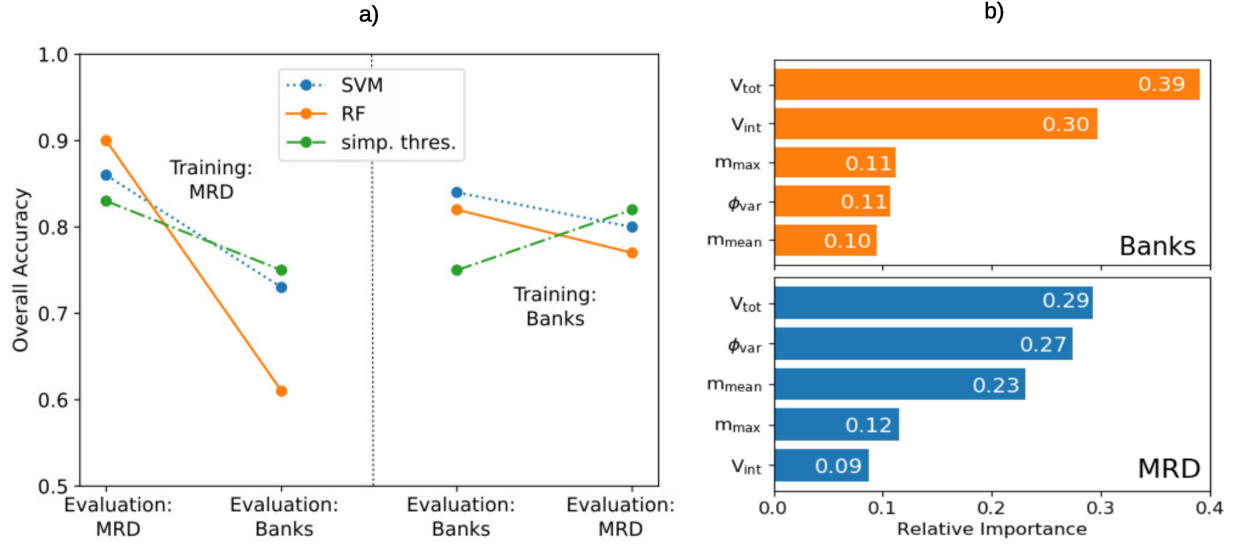


Fig. 10. (a) Overall classification accuracies of the classifier when training and evaluation were done on different study areas. The left panel shows the result when the classifiers were trained on the MRD data and applied on the two study areas. All classifiers show a reduction in the accuracies with the RF classifiers showing the strongest drop from 90% to 61% accuracy. The right panel shows the result when the classifiers were trained on the Banks data and applied on the two study areas. The reduction in the accuracies is not as strong as when training was done on the MRD data, with even an increase in the OA for the simple threshold method. (b) Feature importance values generated by the RF classifiers for the two study sites. On both study sites, the most important feature is the total volumetric change. The features based on the DEM show a higher importance on the MRD study site in contrary to the Banks study site.

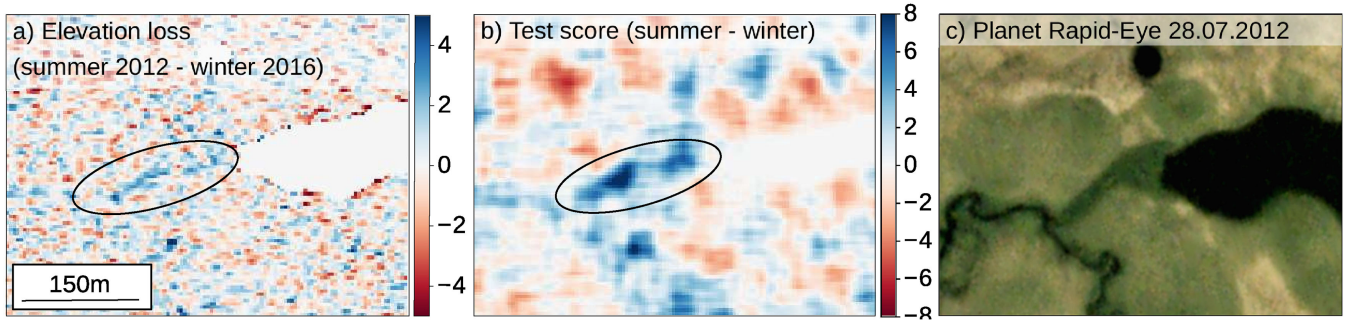


Fig. 11. Example of influence of vegetation on elevation measurements on MRD. (a) and (b) Measured elevation change and calculated test statistic between DEMs generated from observation taken on July 16, 2012 during the summer and on October 25, 2016 when the landscape was frozen. (c) Planet observation taken on July 28, 2012 of the same area, indicating that area, which shows signs of an elevation change, is covered by vegetation.

TanDEM-X observation were taken during the summer of 2012. When using the generated DEMs for the detection of significant elevation changes we found many detection caused by vegetation. An example of the influence of vegetation on MRD can be seen in Fig. 11. The area showing significant elevation changes in the test score image is covered by vegetation and water is potentially flowing below the canopy, inducing wrong elevation measurements. On Banks, several TanDEM-X observation were taken in spring 2012, leading to many false detections. An example area can be seen in Fig. 12. Wet late-lying snow packs can be identified as the reason for many apparent elevation changes, indicated by low backscatter intensities in valleys and trenches. Snow packs potentially also accumulate and persist in RTSs. This would lead to false classifications and to false estimations of the volumetric changes.

Consequently, the investigation of using DEMs generated from observations that were taken during a time when the landscape is not frozen (April to September) showed increase in detections for both study areas. The increase in detections was

stronger for the mid and high, than for low parameter setting. For the low parameter setting, the number of detections for MRD increased from 120 to 207 detections with a decrease in the PPV from 87% to 82%. Of the 87 new detections, 27 are additional RTSs. For the mid parameter setting, the change was more pronounced with an increase of 227 to 545 and a decrease in the PPV from 62% to 47% with 67 new RTS detections. On Banks using the low parameter setting, the number of detection increased from 777 to 1402 with a decrease in the PPV from 71% to 48% with 124 additional RTSs. For the midparameter setting, the number of detections increased significantly to 4791 and a manual assessment of the detections (accuracy assessment Step 1) in RTS and non-RTS detection was not possible.

VII. DISCUSSION

We present the first mapping algorithm for RTSs based on the volumetric changes induced by thaw slump activity. This is in contrast to prior studies where the 2-D information from optical

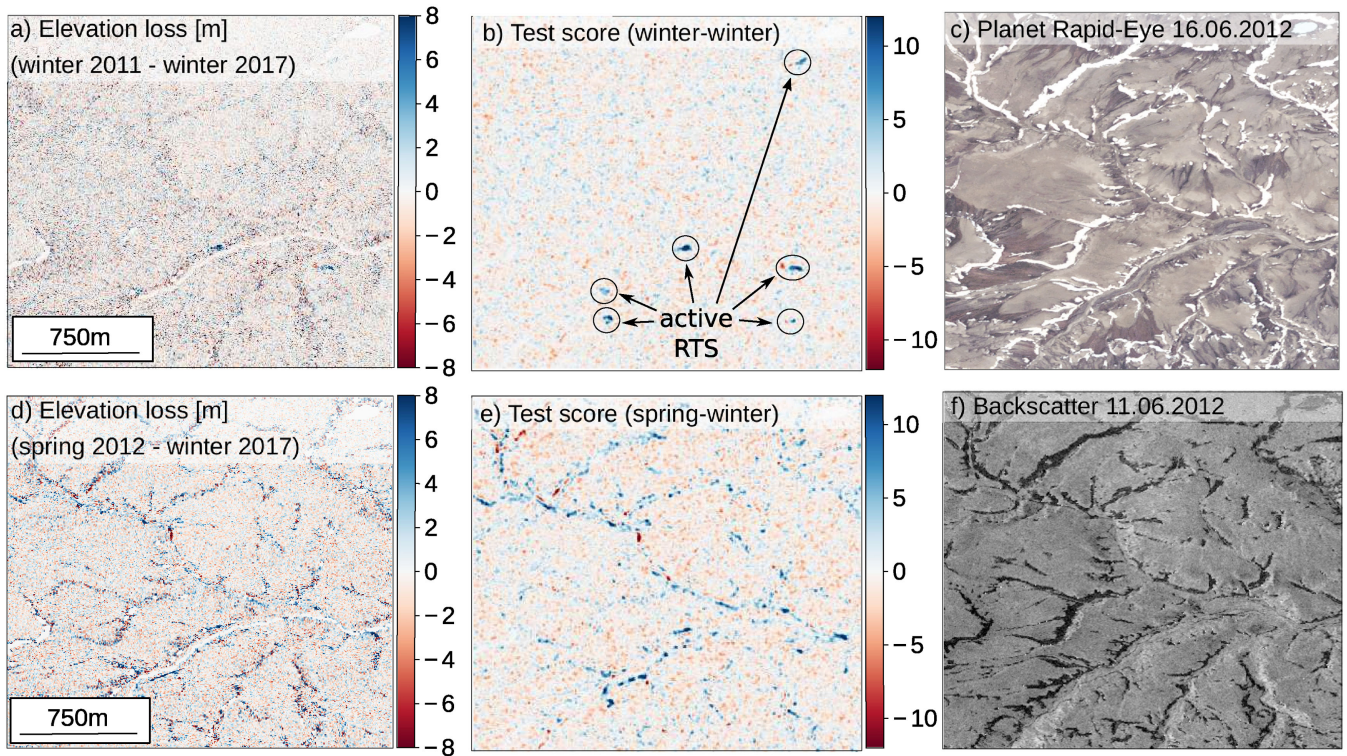


Fig. 12. Influence of snow on elevation measurements on Banks. (a) and (b) Elevation change and computed test score, respectively, for DEMs generated between January 22, 2011 and January 19, 2017. Several significant height changes are visible corresponding to active RTSs. The circles show the locations of active RTSs that were detected. (d) and (e) Elevation change and computed test score, respectively, for DEMs generated between June 11, 2012 and January 19, 2017. The measurements show many disturbances especially at locations in trenches. (c) True-color Planet RapidEye observation taken on June 16, 2012. Several late-lying snow packs at similar location than the elevation changes in the TanDEM-X difference image are visible. (f) Influence of wet snow is also visible in the backscatter intensity image for the TanDEM-X observation taken on June 11, 2012. The darker areas corresponding to lower backscatter values and are in similar areas then the snow areas in the optical Planet true-color image.

satellites based on time-series analysis on Landsat data [15] and deep learning methods on Planet data [16] have been used. The growth of the scare zone and the accompanied vegetation change is used as an RTS indicator. Our two-step approach is tailored to TanDEM-X single-pass SAR observations obtained at different times to generate DEMs and estimate the volumetric change by differentiation. The use of the volumetric change as an RTS indicator is a more direct indication of RTS activity and furthermore independent on vegetation coverage, which is especially important for the vegetation sparse Arctic tundra regions.

A. Two-Step Algorithm for Mapping Thaw Slumps With TanDEM-X

In a first step, we detected significant elevation change in the DEM difference images. The computation of a test statistic, additionally using the spatial size of the thaw slumps, followed by blob detection algorithm and a clustering step, successfully extracted spatially extensive physical height changes from isolated artifacts due to noise. The detection step needs to strike a balance between the number of RTS detections in relation to non-RTS detections. This is a particular challenge for small thaw slumps since two counter acting objectives have to be considered. On the one hand, we want to detect RTSs with properties as close

as possible to the detection limit but decreasing the thresholds and thus the amount of detected RTSs leads on the other hand to an increase in the number of non-RTS detections. Since in the following classification step RTS and non-RTS detection are separated, we selected the midparameter setting as the optimal threshold, in which also slumps with retreat rates and headwall heights just at the detection limit are detected but nevertheless very small or barely active RTSs are missed. The PPV value of 73% of the total dataset using the midparameter setting indicates a good performance of our algorithm in narrowing down the sites of potential slump activity. Nevertheless, 27% of the detections could not be related to RTSs and a second classification step is needed to separate RTS and non-RTS detections.

To weed out non-RTS detection, we investigated the use of three classifiers and investigated their transferability between the study areas. All classifiers showed a good performance (OA: 75%–90%) in separating between RTS and non-RTS detections. The performance of the machine learning classifiers (SVM, RF) was about 5%–10% better than for the simple threshold method. The classifiers performed better on the MRD study area than on Banks Island. This difference can be explained using the feature important values generated by the RF classifier [see Fig. 10(b)]. On MRD, the most important feature is the total volumetric change followed by the features based on the DEM. On the contrary on the Banks data, the

DEM features show relatively low importance values. This hints toward a difference in the topographic locations of RTS and non-RTS detections in the two study sites. On MRD, wrong detections are mainly found in vegetated trenches and in areas showing signs of polygons, whereas RTSs occur mainly along lake shores. Since the features ϕ_{var} and m_{mean} are features detecting these differences, they show a high importance for distinguishing between RTS and non-RTS detections. In contrast to the MRD study site, Banks Island has much larger topographic features, generating many wrong detections in similar places as typical RTSs. Consequently, the features based on the DEM were not helpful in distinguishing between RTS and non-RTS detections.

The assessment of our algorithm requires high-quality training data, which are difficult to obtain. Our validation data were subject to three limitations. First, the volumetric changes that characterize slump activity are not directly visible in the ≈ 10 m resolution optical satellite images we used for validation. In particular, slowly stabilizing slumps are difficult to distinguish from stabilized slumps in optical data, and recent small slumps are difficult to detect. These errors may have falsely increased the number of non-RTS detections (moderately active slump not included in the reference data) as well as the PPV thus the number of detected RTSs (stable slumps wrongly characterized as active). Second, for the estimation of the false negative rate, the reference data from Banks Island did not include the final two years of our study period, such that potentially more active slumps have been missed. Third, the sample size of the training data was limited due to the considerable effort required for manual mapping. While the amount was sufficient for the low to moderate complexity machine learning classification algorithms we employed, the training data requirements of more powerful deep learning methods will be more difficult to meet in practice. Training data requirements are also an important consideration for extending our algorithm to larger scales.

B. Opportunities and Limitations for Large-Scale Mapping

Our two-step method can potentially be applied on a pan-Arctic scale, thanks to the repeated blanket coverage by TanDEM-X. However, the nonuniform temporal coverage and available viewing geometries have to be considered. These properties are related to the global TanDEM-X acquisition strategy [64].

In terms of the temporal coverage, we generally have at least two observations in the years 2011/2012 and two in 2016/2017 covering the whole pan-Arctic areas with HoA's of below 80 m. Nevertheless, one total coverage of the MRD study area was obtained in summer 2012 and for the Banks study area in spring 2012. When we included DEMs generated from these observations during which the landscape was not frozen, the number of detections increased significantly. The additional detections are related to vegetation covered areas and wet late-lying snow packs (see Figs. 11 and 12). The induced elevation changes are on the order of typical RTS movements. DEMs generated during this time frame can also lead to false volume change estimates, for example, from snow accumulation at the foot of the slump

headwall. Due to these problems, we found that the best strategy is to remove these observations from the analysis. The need for observations that are taken when the landscape is frozen reduces the number of usable observations. This implies that if one or two observations are taken in the spring and summer time, no or only one observation pair is well suited for the RTS detection and property extraction leading to gaps in the pan-Arctic coverage. We expect an improved data availability in the future due to the current second global TanDEM-X acquisition plan with an additionally coverage of the global permafrost region in the winter with HoA's of 35–45 m [65].

Our study area is observed in the same viewing geometry (right looking, ascending). Hill slopes and trenches are influenced by the radar viewing angle, increasing or decreasing the backscatter values and thus the coherence as well as the spatial resolution. Additionally, the right looking, ascending viewing geometry leads to a potential bias in terms of the orientation of the detected RTSs related to layover and shadow. For example, if the headwall is oriented away from the sensor, the headwall and part of the slump floor cannot be observed by the radar waves. Due to the complex interplay between these factors and the limited amount of validation data regarding location, headwall height, and movement of RTSs on large scale, we could not quantify the sensitivity of our method regarding possible biases. The absence of descending observations is related to the acquisition strategy for generating the global TanDEM-X DEM in which terrain with relatively flat topography was observed with lower variability in different observation modes than for more difficult mountainous terrain. In areas with relatively flat topography, the effect of layover and shadow on the detection accuracy should be minor, but when using our detections to investigate RTS drivers and controls as well as for the computation of the induced volumetric changes the potential error due to the viewing geometry needs to be investigated in detail. Similarly, the application to more mountainous terrain needs to be investigated separately and it is likely that additional methods dealing with layover, shadow, and the DEM coregistration need to be implemented.

When extending our RTS detection procedure to different study sites, the topography as well as vegetation characteristics need to be similar to our study areas. Knowledge about the expected location of RTSs as well as non-RTS detections can improve the classification accuracy. For example, since typical non-RTS detections on MRD are located in trenches and flat polygon covered areas, the two machine learning algorithms showed a higher classification accuracy than the simple threshold method. But when applying the classifier trained on the MRD study on the Banks data, the classification accuracy dropped significantly [see Fig. 10(a)]. To achieve the best performance, the use of machine learning classifiers trained on part of the study area or on a similar area regarding typical RTSs and landscape characteristics (e.g., topography and vegetation) should provide the best result. Nevertheless, using a simple threshold method on the volumetric change with a threshold at about 3000 m^3 can significantly improve the result without the need for trainings data. Additional datasets based on optical data, such as the ArcticDEM or optical satellites (Sentinel-2, RapidEye), have

the potential to further improve the detection rates as well as provide additional information on the detected RTSs.

VIII. CONCLUSION

We presented and assessed a method to detect active RTSs, using for the first time the volumetric change as an RTS indicator by applying DEM differencing. Our suggested approach is applicable on flat and medium mountainous terrain and provides an important step toward an RTS detection and monitoring method on the pan-Arctic scale.

In our methods first step, we isolated significant height changes using a statistical multiscale approach that is intended to discard spurious changes induced by measurement noise. In total, 1302 significant height changes were detected but reference data showed that 27% are due to processes other than thaw slumps. In a second step, we classified those significant height changes into RTS and non-RTS using thresholding and two machine learning algorithms. The machine learning classifiers exploited the characteristic properties of RTSs and common non-RTS detections, such as their size and their topographic position. We found that these differed markedly between the two study regions. For example, non-RTS detections in the MRD study area were often found in flat areas in contrary to the Banks study area where many non-RTS detections were located close to large topographic features. Accordingly, the transferability of the algorithms was limited: a simple threshold-based algorithm outperformed the SVM and the RF classifiers when transferred to the other study region. Provided that training data from a similar study region are available, our algorithm is expected to perform with accuracies of 85%–90% using the SVM and RF classifiers. Since the TanDEM-X observations cover the whole pan-Arctic landscape, the availability of an RTS detection method makes it possible to generate large-scale inventories of RTSs. Such inventories have the potential to be used as a starting point to measure RTS-induced volumetric changes and estimate the amount of mobilized materials including organic carbon, nutrients, and sediments.

REFERENCES

- [1] E. A. G. Schuur *et al.*, "Climate change and the permafrost carbon feedback," *Nature*, vol. 520, no. 7546, 2015, pp. 171–179.
- [2] G. Grosse *et al.*, "Vulnerability of high-latitude soil organic carbon in North America to disturbance," *J. Geophys. Res., Biogeosci.*, vol. 116, no. G4, 2011, Art. no. G00K06.
- [3] H. Lantuit and W. H. Pollard, "Fifty years of coastal erosion and retrogressive thaw slump activity on Herschel Island, southern Beaufort Sea, Yukon Territory, Canada," *Geomorphology*, vol. 95, no. 1/2, pp. 84–102, 2008.
- [4] T. C. Lantz and S. V. Kokelj, "Increasing rates of retrogressive thaw slump activity in the Mackenzie Delta region, N.W.T., Canada," *Geophys. Res. Lett.*, vol. 35, no. 6, 2008, Art. no. L06502.
- [5] M. N. Gooseff, A. Balser, W. B. Bowden, and J. B. Jones, "Effects of hillslope thermokarst in Northern Alaska," *Eos*, vol. 90, no. 4, pp. 29–30, 2009.
- [6] S. V. Kokelj, T. C. Lantz, J. Kanigan, S. L. Smith, and R. Coutts, "Origin and polycyclic behaviour of tundra thaw slumps, Mackenzie Delta region, Northwest Territories, Canada," *Permafrost Periglacial Processes*, vol. 20, no. 2, pp. 173–184, 2009.
- [7] A. G. Lewkowicz and R. G. Way, "Extremes of summer climate trigger thousands of thermokarst landslides in a High Arctic environment," *Nature Commun.*, vol. 10, no. 1, 2019, Art. no. 1329.
- [8] D. Lacelle, J. Bjornson, and B. Lauriol, "Climatic and geomorphic factors affecting contemporary (1950–2004) activity of retrogressive thaw slumps on the Aklavik plateau, Richardson mountains, NWT, Canada," *Permafrost Periglacial Processes*, vol. 21, no. 1, pp. 1–15, 2010.
- [9] T. B. Barnhart and B. T. Crosby, "Comparing two methods of surface change detection on an evolving thermokarst using high-temporal-frequency terrestrial laser scanning, Selawik River, Alaska," *Remote Sens.*, vol. 5, no. 6, pp. 2813–2837, 2013.
- [10] J. L. Ramage, A. M. Irrgang, U. Herzschuh, A. Morgenstern, N. Couture, and H. Lantuit, "Terrain controls on the occurrence of coastal retrogressive thaw slumps along the Yukon Coast, Canada," *J. Geophys. Res., Earth Surface*, vol. 122, pp. 1619–1634, 2017.
- [11] N. Short, B. Brisco, N. Couture, W. Pollard, K. Murnaghan, and P. Budkewitsch, "A comparison of TerraSAR-X, RADARSAT-2, and ALOS-PALSAR interferometry for monitoring permafrost environments, case study from Herschel Island, Canada," *Remote Sens. Environ.*, vol. 115, no. 12, pp. 3491–3506, 2011.
- [12] A. C. A. Rudy, S. F. Lamoureux, P. Treitz, and A. Collingwood, "Identifying permafrost slope disturbance using multi-temporal optical satellite images and change detection techniques," *Cold Regions Sci. Technol.*, vol. 88, pp. 37–49, 2013.
- [13] A. Brooker, R. H. Fraser, I. Olthof, S. V. Kokelj, and D. Lacelle, "Mapping the activity and evolution of retrogressive thaw slumps by tasseled cap trend analysis of a Landsat satellite image stack," *Permafrost Periglacial Processes*, vol. 25, no. 4, pp. 243–256, 2014.
- [14] D. Lacelle, A. Brooker, R. H. Fraser, and S. V. Kokelj, "Distribution and growth of thaw slumps in the Richardson Mountains-Peel Plateau region, northwestern Canada," *Geomorphology*, vol. 235, pp. 40–51, 2015.
- [15] J. Nitze, I. Grosse, G. Jones, B. M. Romanovsk, and V. E. Boike, "Remote sensing quantifies widespread abundance of permafrost region disturbances across the Arctic and Subarctic," *Nature Commun.*, vol. 9, no. 1, 2018, Art. no. 5423.
- [16] L. Huang, J. Luo, Z. Lin, F. Niu, and L. Liu, "Using deep learning to map retrogressive thaw slumps in the Beiluhe region (Tibetan Plateau) from CubeSat images," *Remote Sens. Environ.*, vol. 237, 2020, Art. no. 111534.
- [17] T. Strozzi, U. Wegmuller, C. Werner, and A. Kos, "TerraSAR-X interferometry for surface deformation monitoring on periglacial area," in *Proc. Int. Geosci. Remote Sens. Symp.*, 2012, pp. 5214–5217.
- [18] S. Antonova *et al.*, "Thaw subsidence of a Yedoma landscape in Northern Siberia, measured *in situ* and estimated from TerraSAR-X Interferometry," *Remote Sens.*, vol. 10, no. 4, 2018, Art. no. 494.
- [19] T. Strozzi, F. Paul, A. Wiesmann, T. Schellenberger, and A. Kääh, "Circum-Arctic changes in the flow of glaciers and ice caps from satellite SAR data between the 1990s and 2017," *Remote Sens.*, vol. 9, no. 9, 2017, Art. no. 947.
- [20] J. Kubanek, M. Westerhaus, A. Schenk, N. Aisyah, K. S. Brotospito, and B. Heck, "Volumetric change quantification of the 2010 Merapi eruption using TanDEM-X InSAR," *Remote Sens. Environ.*, vol. 164, pp. 16–25, 2015.
- [21] H. Rott, D. Floricioiu, J. Wuite, S. Scheiblauer, T. Nagler, and M. Kern, "Mass changes of outlet glaciers along the Nordenskjöld Coast, northern Antarctic Peninsula, based on TanDEM-X satellite measurements," *Geophys. Res. Lett.*, vol. 41, pp. 8123–8129, 2014.
- [22] G. Krieger *et al.*, "TanDEM-X: A satellite formation for high-resolution SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 11, pp. 3317–3341, Nov. 2007.
- [23] H. M. French, "The banks island tundra," in *Landscapes and Landforms of Western Canada*. Cham, Switzerland: Springer, 2017.
- [24] R. A. Segal, T. C. Lantz, and S. V. Kokelj, "Acceleration of thaw slump activity in glaciated landscapes of the Western Canadian Arctic," *Environ. Res. Lett.*, vol. 11, no. 3, 2016, Art. no. 034025.
- [25] S. Zwieback, S. Kokelj, F. Günther, J. Boike, G. Grosse, and I. Hajnsek, "Sub-seasonal thaw slump mass wasting is not consistently energy limited at the landscape scale," *Cryosphere*, vol. 12, no. 2, pp. 549–564, 2018.
- [26] A. C. Rudy, S. F. Lamoureux, S. V. Kokelj, I. R. Smith, and J. H. England, "Accelerating thermokarst transforms ice-cored terrain triggering a downstream cascade to the ocean," *Geophys. Res. Lett.*, vol. 44, no. 21, pp. 11080–11087, 2017.
- [27] C. Werner, U. Wegmuller, T. Strozzi, and A. Wiesmann, "GAMMA SAR and interferometric processing software," Eur. Space Agency (Special Publication), vol. 461, pp. 211–219, 2000.
- [28] M. Costantini, "A novel phase unwrapping method based on network programming," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 3, pp. 813–821, May 1998.

- [29] C. Nuth and Kääb, "Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change," *Cryosphere*, vol. 5, no. 1, pp. 271–290, 2011.
- [30] P. Rizzoli, M. Martone, H. Rott, and A. Moreira, "Characterization of snow facies on the Greenland ice sheet observed by TanDEM-X interferometric SAR data," *Remote Sens.*, vol. 9, 2017, Art. no. 315.
- [31] A. Ferretti, A. Monti-Guarnieri, C. Prati, and F. Rocca, "InSAR principles: Guidelines for SAR interferometry processing and interpretation (TM-19, Feb. 2007)," ESA Publications, Noordwijk, The Netherlands, 2007.
- [32] C. Elachi, M. L. Bryan, and W. F. Weeks, "Imaging radar observations of frozen Arctic lakes," *Remote Sens. Environ.*, vol. 5, pp. 169–175, 1976.
- [33] T. Nagler and H. Rott, "Retrieval of wet snow by means of multitemporal SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 21, pp. 754–765, Mar. 2000.
- [34] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognit.*, vol. 19, no. 1, pp. 41–47, 1986.
- [35] A. Wendleder, B. Wessel, A. Roth, M. Breunig, K. Martin, and S. Wagenbrenner, "TanDEM-X water indication mask: Generation and first evaluation results," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 1, pp. 171–179, Feb. 2013.
- [36] S. K. McFeeters, "The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features," *Int. J. Remote Sens.*, vol. 17, no. 7, pp. 1425–1432, 1996.
- [37] G. Kaplan and U. Avdan, "Object-based water body extraction model using Sentinel-2 satellite imagery," *Eur. J. Remote Sens.*, vol. 50, no. 1, pp. 137–143, 2017.
- [38] B. Wessel, TanDEM-X Ground Segment—DEM Products Specification Document EOC, DLR, Oberpfaffenhofen, Germany, Public Document TD-GS-PS-0021, issue 3.1 ed., EOC, DLR, Oberpfaffenhofen, Germany, 2016. [Online]. Available: <https://tandemx-science.dlr.de/>
- [39] S. Baillarin *et al.*, "Sentinel-2 level 1 products and image processing performances," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2012, pp. 7003–7006.
- [40] M. Pandzic, D. Mihajlovic, J. Pandzic, and N. Pfeifer, "Assessment of the geometric quality of Sentinel-2 data," *ISPRS—Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XLI-B1, pp. 489–494, 2016.
- [41] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-Up Robust Features (SURF)," *Comput. Vision Image Understanding*, vol. 3, no. 110, pp. 346–359, 2008.
- [42] S. C. Johnson, "Hierarchical clustering schemes," *Psychometrika*, vol. 32, pp. 241–254, 1967.
- [43] Planet Team, "Planet application program interface: In space for life on earth," San Francisco, CA, USA, 2017.
- [44] A. E. Maxwell, T. A. Warner, and F. Fang, "Implementation of machine-learning classification in remote sensing: An applied review," *Int. J. Remote Sens.*, vol. 39, no. 9, pp. 2784–2817, 2018.
- [45] M. Pal and P. M. Mather, "An assessment of the effectiveness of decision tree methods for land cover classification," *Remote Sens. Environ.*, vol. 86, no. 4, pp. 554–565, 2003.
- [46] G. M. Foody and A. Mathur, "A relative evaluation of multiclass image classification by support vector machine," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1335–1343, Jun. 2004.
- [47] M. Pal and G. M. Foody, "Evaluation of SVM, RVM and SMLR for accurate image classification with limited ground data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 5, pp. 1344–1355, Oct. 2012.
- [48] R. L. Lawrence and C. J. Moran, "The AmericaView classification methods accuracy comparison project: A rigorous approach for model selection," *Remote Sens. Environ.*, vol. 170, pp. 115–120, 2015.
- [49] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
- [50] B. J. McNeil and J. A. Hanley, "Statistical approaches to the analysis of receiver operating characteristic (ROC) curves," *Med. Decis. Making*, vol. 4, no. 2, pp. 137–150, 1984.
- [51] W. J. Youden, "Index for rating diagnostic tests," *Cancer*, vol. 3, pp. 32–35, 1950.
- [52] M. Greiner, D. Pfeiffer, and R. D. Smith, "Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests," *Preventive Veterinary Med.*, vol. 45, no. 1, pp. 23–41, 2000.
- [53] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [54] M. Pal, "Random forest classifier for remote sensing classification," *Int. J. Remote Sens.*, vol. 26, no. 1, pp. 217–222, 2005.
- [55] D. R. Cutler *et al.*, "Random forests for classification in ecology," *Ecology*, vol. 88, no. 11, pp. 2783–2792, 2007.
- [56] M. Belgiu and L. Dragut, "Random forest in remote sensing: A review of applications and future directions," *ISPRS J. Photogrammetry Remote Sens.*, vol. 114, pp. 24–31, 2016.
- [57] Y. Grandvalet, "Bagging equalizes influence," *Mach. Learn.*, vol. 55, no. 3, pp. 251–270, Jun. 2004.
- [58] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [59] M. Pal and P. M. Mather, "Support vector machines for classification in remote sensing," *Int. J. Remote Sens.*, vol. 26, no. 5, pp. 1007–1011, 2005.
- [60] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 3, pp. 247–259, 2011.
- [61] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Meas.*, vol. 20, no. 1, pp. 37–46, 1960.
- [62] H. M. French, "Active thermokarst processes, eastern banks island, Western Canadian Arctic," *Can. J. Earth Sci.*, vol. 11, no. 6, pp. 785–794, 1974.
- [63] A. G. Lewkowicz, "Headwall retreat of ground-ice slumps, Banks Island, Northwest Territories," *Can. J. Earth Sci.*, vol. 24, no. 6, pp. 1077–1085, 1987.
- [64] P. Rizzoli *et al.*, "Generation and performance assessment of the global TanDEM-X digital elevation model," *ISPRS J. Photogrammetry Remote Sens.*, vol. 132, pp. 119–139, 2017.
- [65] M. Bachmann, D. Borla Tridon, M. Martone, F. Sica, S. Buckreuss, and M. Zink, "How to update a global DEM - acquisition concepts for TanDEM-X and tandem-L," in *Proc. Eur. Conf. Synthetic Aperture Radar*, 2018, pp. 1–5.



Philipp Bernhard (Graduate Student Member, IEEE) received the M.Sc. degree in physics, in 2016 from ETH Zürich, Zürich, Switzerland, where he worked on galaxy simulations and the evolution of super massive black holes. He is currently working toward the Ph.D. degree in physics with ETH Zürich.

His main research interest are on permafrost thaw in the Arctic with a special emphasis on retrogressive thaw slumps. For this, he works on the generation of microwave digital elevation models, SAR data processing, and on the fusion of different remote sensing products for large-scale applications.



Simon Zwieback received the M.Sc. degree in geodesy and geophysics from the Vienna University of Technology, Vienna, Austria, and the Ph.D. degree in environmental engineering from the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland, in 2016.

He is currently a Research Associate with the University of Alaska Fairbanks, Fairbanks, AK, USA. His main research interests are in the hydrology and geomorphology of periglacial environments and in the development of forward and inverse models in microwave remote sensing. He is also interested in the uncertainty characterization of remote sensing products.



Silvan Leinss (Member, IEEE) received the Diploma degree in physics, in 2008 from the University of Constance, Konstanz, Germany, where he worked with femtosecond terahertz spectroscopy and light scattering in colloidal suspensions of random media, and the Ph.D. in radar remote sensing, in 2015 from ETH Zürich, Zürich, Switzerland, with focus on interferometric and polarimetric analysis of snow.

From 2009 to 2010, he was with light scattering statistics of Bose–Einstein condensates at ETH Zürich. Since then he worked as a remote sensing PostDoc and Scientist on snow and glacier dynamics at ETH Zürich.



Irena Hajnsek (Fellow, IEEE) received the Diploma degree (honors) from the Free University of Berlin, Berlin, Germany, in 1996, and the Dr. degree (honors) from the Friedrich Schiller University of Jena, Jena, Germany, in 2001.

Since November 2009, she has been a Professor of Earth Observation with the Swiss Federal Institute of Technology (ETH), Zürich Institute of Environmental Engineering, Zürich, Switzerland, and, at the same time, the head of the Polarimetric SAR Interferometry research group with the German Aerospace Center

Microwaves and Radar Institute, Wessling, Germany. Since 2010, she has been the Science Coordinator of the German satellite mission TanDEM-X and proposed satellite mission Tandem-L. Her main research interests include electromagnetic propagation and scattering theory, radar polarimetry, SAR and interferometric SAR data processing techniques, and environmental parameter modeling and estimation.

Dr. Hajnsek was the Technical Program Co-chair of the IEEE IGARSS 2012 in Munich, Germany, and IEEE IGARSS 2019 in Yokohama, Japan. Since 2013, she has been a member of the IEEE GRSS AdCom, and since 2016, she has been the Vice-President of the IEEE GRSS Technical Committees.