Supraglacial lake evolution on Tracy and Heilprin Glaciers in northwestern Greenland from 2014 to 2021

Yefan Wang and Shin Sugiyama

Institute of Low Temperature Science, Hokkaido University, Sapporo, Japan

Abstract: Supraglacial lakes form on the Greenland ice sheet during the melting season as meltwater accumulates in topographic depressions on the surface. They have the potential to increase ablation by reducing albedo and enhance ice flow when water drains to the bed. Despite recent progress in understanding the surface hydrology of the Greenland ice sheet, the seasonal and year-to-year variations of supraglacial lakes remain poorly understood. Here, we employed a machine learning method trained with random forest classifiers on Landsat 8 and Sentinel-2 imagery within Google Earth Engine for detecting supraglacial lakes. We applied this method to two large marine-terminating glaciers (Heilprin Glacier and Tracy Glacier) in northwestern Greenland to investigate the distribution, inter-annual and seasonal evolutions of supraglacial lakes between 2014 and 2021. Over the eight study years, cumulated maximum lake extent on Heilprin Glacier (22.84 km^2) was approximately three times larger than that of Tracy Glacier (7.60 km^2) despite the similarity in the glacier areas (654 km^2 and 540 km^2). The lake extent exhibited significant interannual variability controlled by summer PDD (sum of positive degree days), and the upper limit of lake formation coincided approximately with ELA (equilibrium line altitude). Greater lake areas were observed in 2016 and 2019, but obviously below-average areas in 2017–2018. The variabilities were pronounced in the region above 800 m a.s.l., particularly at Tracy Glacier. On the seasonal scale, lakes begin forming in early- to mid-June and disappear towards the end of August. The lake formation was triggered by the initiation of runoff in the regions below

800 m a.s.l., while it coincided with the onset of surface melt at the higher elevations due to the presence of ice/snow-buried lakes. Cascading rapid lake drainage events were observed in the mid-elevation areas (400−800 m a.s.l.) of Heilprin Glacier, which induced an abrupt decrease in the lake area at the beginning of July in the years of 2019 and 2020. Our study demonstrates the potential of the machine learning method to analyze supraglacial lakes in dense satellite images. The results imply inland expansion of supraglacial lakes under warming climate in Greenland and a possible interaction between draining lake water and glacier dynamics.

Key words: Supraglacial lake; Greenland ice sheet; Machine learning; Random forest; Google Earth Engine; Calving glacier

1 Introduction

Meltwater runoff from the Greenland ice sheet has substantially increased in the past two decades (Slater et al., 2021; Tedstone and Machguth, 2022). It was accompanied by an expansion of the visible runoff area in the ice sheet marginal region by 29% between the years 1985 and 2020 (Tedstone and Machguth, 2022). Part of the meltwater produced on the ice sheet surface accumulates in surface topographic depressions within the ablation zone, forming supraglacial lakes during the melt season (Echelmeyer et al., 1991). The formation of supraglacial lakes has a significant influence on the ice sheet mass balance and ice dynamics. Firstly, the lakes enhance the surface melting by 100–170% in comparison to the surrounding bare ice area, due to the lower albedo (Lüthje et al, 2006; Tedesco et al, 2012). Secondly, a rapid lake water drainage event may cause ice flow acceleration, due to elevated basal water pressure and reduced basal friction (Chudley et al., 2019; Das et al., 2008; Doyle et al., 2013; Stevens et al., 2015).

The formation and subsequent development of supraglacial lakes are controlled by several factors, including topography, climate, ice velocity, and ice thickness (Williamson et al., 2018b). Given that sufficient meltwater is produced, the location of supraglacial lake formation is primarily controlled by surface topography (Lüthje et al., 2006). Since surface topography is often affected by bed geometry, the position of lakes on the ice surface is largely controlled by the underlying bedrock topography (Lampkin and Vanderberg, 2011). The surface depressions provide potential locations for lake development, while lake ponding is controlled by water availability. According to studies conducted in southwestern Greenland (Johansson et al., 2013; Fitzpatrick et al., 2014), small lakes tend to cluster in near-terminus low-elevation areas in the form of water-filled crevasses, whereas relatively large lakes tend to form annually at the same topographic depressions at higher elevations (~1000–1500 m a.s.l.). In more inland regions (above ~1600 m), supraglacial lakes are rarely observed because meltwater is unavailable under low temperatures (Sundal et al., 2009; Johansson et al., 2013; Fitzpatrick et al., 2014; Yang et al., 2021). The formation of lakes shows seasonal progress from lower elevations to higher elevations as the temperature increases (Chu, 2014). Due to current climate warming in the Arctic, the appearance of supraglacial lakes is likely to migrate to higher elevations where surface slopes are gentle and ice speeds are slow (Leeson et al., 2015; Ignéczi et al., 2016). Therefore, investigating the lake evolution and drainage is crucial for assessing their potential impact on the mass balance and dynamics of the Greenland ice sheet.

To investigate supraglacial lakes in large spatial and temporal scales, remote sensing is an effective and commonly employed approach. For instance, MODIS (Moderate Resolution Imaging Spectroradiometer) images have been widely used in supraglacial lake studies (Box & Ski, 2007; Selmes et al., 2011; Sundal et al., 2009). They are available in a high temporal resolution (daily), but relatively low spatial resolution (250 or 500 m) of MODIS hinders the identification of small lakes (< 0.125 km^2). Smaller lakes can be studied with relatively high-resolution ASTER (15 m) and Landsat (30 m) images (Sneed and Hamilton, 2007; McMillan et al., 2007; Georgiou et al., 2009). However, the temporal resolutions of these images are insufficient for short-term lake variations represented by rapid drainage events (Williamson et al., 2017). Since 2014, the Sentinel-1 SAR data are available for the observation of supraglacial lakes in the polar regions (Miles et al., 2017; Dirscherl et al., 2021a; Li et al., 2021; Jiang et al., 2022; Shu et al., 2023). SAR images have an advantage in observation under cloud cover or during polar night, but the radar image suffers from the effects of speckle noise, rendering image processing relatively complex (Miles et al., 2017). After the launch of Landsat 8 and Sentinel-2, combination of optical images acquired by these two satellites enabled us to study supraglacial lakes in higher spatial and temporal resolutions (Williamson et al., 2018a; Dell et al., 2020; Moussavi et al., 2020).

For mapping supraglacial lakes on satellite imagery, the threshold-based classification is the most commonly adopted approach. This method identifies water pixels based on thresholds prescribed for multiple spectral bands or spectral indices. The Normalized Difference Water Index adapted for ice ($NDWI_{ice}$, Yang and Smith, 2013), which is based on an empirically selected red/blue reflectance threshold (typically > 0.2–0.5), is the most widely used spectral index for supraglacial lake identification (e.g. Fitzpatrick et al., 2014; Williamson et al., 2018a). However, the NDWI-based approach sometimes brings about misclassification as it is affected by cloud cover, cloud shadows, and other spectrally similar classes such as slush, blue ice, shaded rocks, and shaded snow. To avoid misclassifications, the multiple-threshold approach has been used in supraglacial lake research (e.g. Spergel et al., 2021; Tuckett et al., 2021; Arthur et al., 2022; Corr et al., 2022). For example, Moussavi et al. (2020) employed the Normalized Difference Snow Index (NDSI) based on green and shortwave infrared (SWIR) bands (Hall et al., 1995) and some other spectral indices for more robust lake classification. This approach led to successful mapping of supraglacial lakes along the entire Antarctic coastal margin with fewer misclassification errors (Moussavi et al., 2020). Despite its ability to distinguish lakes from other water-like features, the threshold-based method is prone to misclassification when applied to larger areas and longer temporal scales because the predefined thresholds may vary in different locations and seasons (Dell et al., 2021).

In recent years, machine learning has been applied for mapping supraglacial lakes. The machine learning approach utilizes a greater amount of spectral information compared to conventional approaches and identifies the most useful information for accurate classification (Dell et al., 2021). The methodology includes Cart, Support Vector Machine (SVM), Random Forest (RF), among which RF outperforms others (Halberstadt et al., 2020) and has been employed for mapping supraglacial lakes in Greenland (Yuan et al., 2020; Hu et al., 2021) and Antarctica (Dirscherl et al., 2020; Dell et al., 2021). The RF methodology reduces the misclassification of lakes (Dirscherl et al., 2020; Halberstadt et al., 2020), particularly because it distinguishes slush from shallow lakes (Dell et al., 2021). As a more advanced machine learning technology, deep learning techniques have also been introduced for mapping glacial lakes. Currently, the CNN (convolutional neural network) architectures are the most widely employed method in glacial lake study. With an advantage of accurate classification of ice, water, and rock surfaces, U-Net (Qayyum et al., 2020; Wu et al., 2020) and CoAtNet (Xu et al., 2023) have been used for mapping proglacial lakes on debris-covered glaciers in the High Mountain Asia. In the polar regions, Yuan et al. (2020) utilized CNN and two machine learning methods (RF and SVM) to extract supraglacial lakes in southwest Greenland based on the Landsat 8 imagery. Subsequently, the U-Net architecture was applied to the Sentinel-1 and Sentinel-2 imagery to study supraglacial lakes in Antarctica (Dirscherl et al., 2021a; Nie et al., 2023) and Greenland (Jiang et al., 2022; Lutz et al., 2023).

A number of lake studies have been reported in southwestern Greenland (e.g., Fitzpatrick et al., 2014; Williamson et al., 2018; Yang et al., 2021), where supraglacial lakes are widely distributed (Hu et al., 2021; Zhang et al., 2023). Lakes are also abundant in northern Greenland, but investigations are relatively sparse and limited in several large glaciers, i.e. Ryder Glacier (Otto et al., 2022), Petermann Glacier (Macdonald et al., 2018), Nioghalvfjerdsfjorden (79° N Glacier, Turton et al., 2021; Lutz et al., 2023), and Humboldt Glacier (Rawlins et al., 2023). Based on the analyses of Landsat or Sentinel-2 imagery, these studies demonstrated recent inland expansion of supraglacial lakes as projected by numerical modelling (Leeson et al., 2015). To better understand the mechanisms of such a trend, there is a need to increase temporal resolution and spatiotemporal coverage by applying an advanced machine learning technique on all available imagery. Moreover, no detailed lake study has been performed in northwestern Greenland, which has been the focus of a Japanese multi-disciplinary research project (Sugiyama et al., 2021). In this study, we present an automatic method for monitoring the evolution of supraglacial lakes on the Greenland ice sheet using the RF machine learning methodology. To achieve our relatively simple goal to distinguish water on a glacier surface, RF is expected to perform as accurately as deep learning (Yuan et al., 2020). The methodology is implemented in Google Earth Engine (GEE), which is a freely available cloud-computing platform provided by Google, allowing users to efficiently access and analyze satellite imagery and geospatial datasets on a planetary scale (Gorelick et al., 2017). By integrating the RF algorithm in GEE, the process model training, image classification, and validation were carried out directly on GEE. Thus, the complexity in the operation and local computational power were greatly reduced. We applied this automatic method for mapping supraglacial lakes on Tracy Glacier and Heilprin Glacier, two major marine-terminating glaciers in Inglefield Bredning, northwestern Greenland. By using Sentinel-2 and Landsat 8 imageries, a high-temporal resolution (sub-weekly) record was generated for lake formation and development on the glaciers. Based on records of multi-year supraglacial lake extent, we examined intra-annual and inter-annual evolution of supraglacial lakes in the study region. By comparing the results with glaciological and climatic datasets, we also investigate the potential factors controlling lake formation and distribution on Greenlandic marine-terminating glaciers.

2 Study site

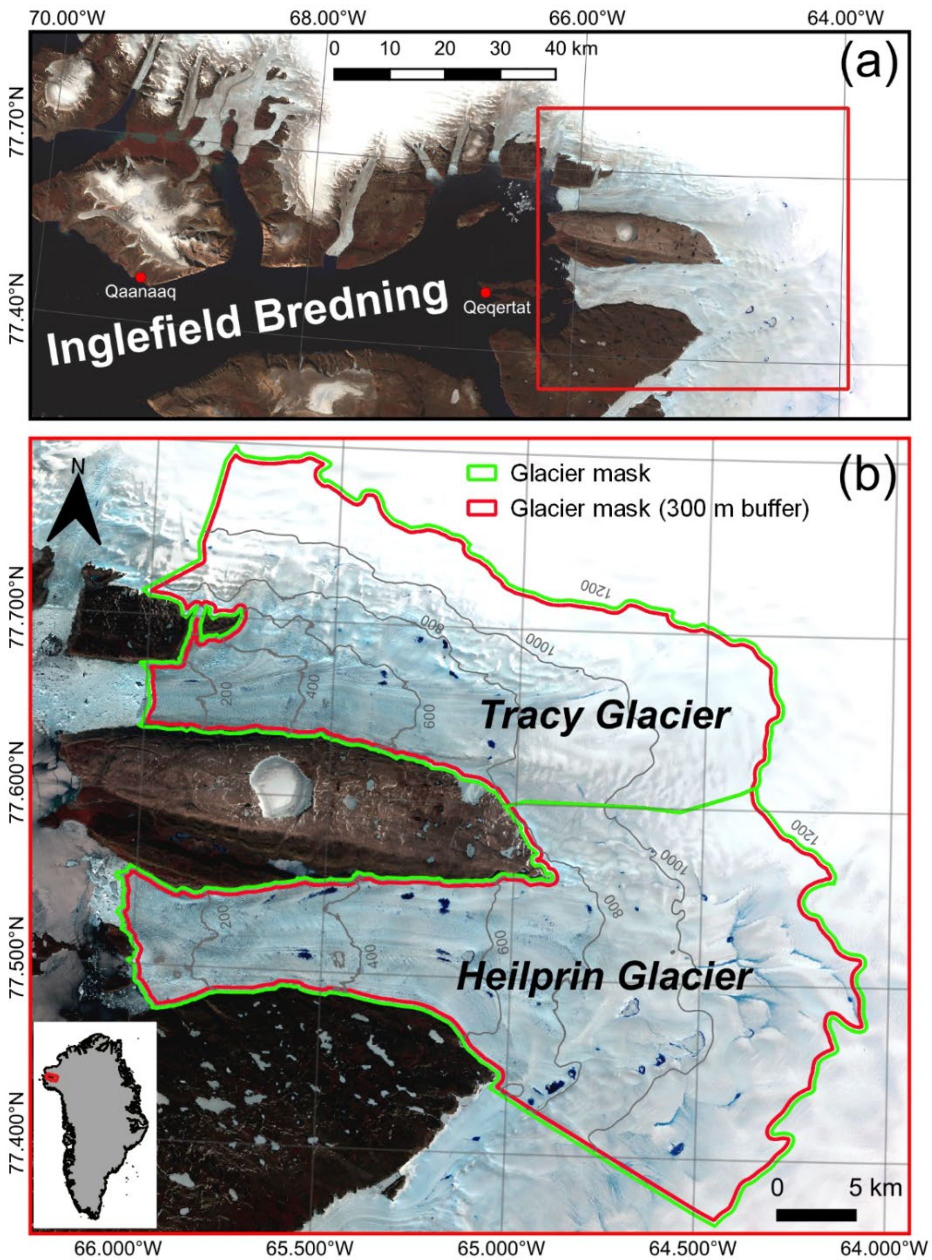

Figure 1. (a) Overview of the glaciers studied and the surrounding area. The background is a mosaicked Sentinel-2 image acquired during July–August 2020. The red box corresponds to the area shown in (b). (b) Details of Tracy and Heilprin Glaciers. The areas surrounded by the red and green boundaries are the glacier masks described in Section 3.1. The background is a Sentinel-2 image acquired on June 27, 2020.

Tracy and Heilprin Glaciers flow into the eastern corner of Inglefield Bredning, a ~80-km long and ~20-km wide glacial fjord in northwestern Greenland (Fig. 1). They are the two largest glaciers in the fjord system with the width of the glaciers near the front of 5 and 8 km in 2020. Although these two glaciers are only separated by a ~10 km-wide bedrock ridge, they exhibit remarkably different patterns in the changes in terminal position, ice speed, and surface elevation (Porter et al., 2014; Willis et al., 2018). From 1892 to 2009, Tracy Glacier retreated by more than 15 km. In contrast, Heilprin Glacier retreated by less than 4 km during the same period (Dawes and Van As, 2010). From 1999 to 2014, the total retreat distance of Tracy and Heilprin Glaciers was 5.7 and 2 km, respectively (Sakakibara and Sugiyama, 2018). The surface elevation change is also more pronounced on Tracy Glacier. where ice thinned at a rate of 3.91 m a^{-1} between 2001 and 2018 (cf. 0.51 m a^{-1} on Heilprin Glacier) (Wang et al., 2021). The difference in the thinning rate was most significant near the glacier fronts, where Tracy Glacier thinned by 9.9 m from 2016 to 2017, five times higher than that of Heilprin Glacier (1.8 m) during the same period (Willis et al., 2018). Meanwhile, the ice speeds of both glaciers accelerated during the years 2000–2014, but the acceleration of Tracy Glacier (51 m a^{-2}) was almost four times greater than that of Heilprin Glacier (13 m a^{-2}) (Sakakibara and Sugiyama, 2018). The significant difference in the behaviors of these two glaciers has been attributed to the deeper grounding line of Tracy Glacier, which makes the glacier front more exposed to deep warm water and therefore subject to stronger submarine melting (Porter et al., 2014; Wang et al., 2021; Willis et al., 2018).

Ingrefield Bredning has been the focus of a Japanese research project on glacier and ocean changes, and their impact on marine ecosystems and human society (Sugiyama et al., 2021). There are two settlements in the fjord system. Qaanaaq is a village with ~650 inhabitants located at the northern flank of the fjord (Fig. 1a). Qeqertat is a smaller settlement with ~30 residents situated in the inner part of the fjord (Fig. 1a) near the fronts of our study glaciers (Flora et al., 2018). Inhabitants of the isolated fjord system primarily depend on subsistence hunting and fishing for their livelihoods (Straneo et al., 2022). Tracy and Heilprin Glaciers are the principal sources of freshwater discharge into the fjord system (Mankoff et al., 2020a, 2020b). Meltwater discharge from the glaciers plays a key role in generating nutrient-rich and highly productive aquatic ecosystems (Kanna et al., 2020, 2022), thereby supporting the local biological environment (Matsuno et al., 2020). Consequently, these two glaciers are important for the residents in the region.

3 Data and methods

3.1 Glacier masks

The analysis was carried out on Heilprin and Tracy Glaciers in the regions below 1200 m a.s.l. (Fig. 1b). Glacier boundaries reported by Mouginot and Rignot (2019) were revised by visual inspection of Sentinel-2 images taken in late August 2021 (green in Fig. 1b). Areas below 1200 m a.s.l. were delineated based on ArcticDEM Mosaic (Porter et al., 2018). To avoid the potential influence of seasonal snow cover, areas within 300 m of the boundaries were eliminated from the study area (red in Fig. 1b). The relatively conservative glacier masks with a 300 m buffer would not cause loss of information because lakes rarely form near glacier margins, based on the inspection of optical images.

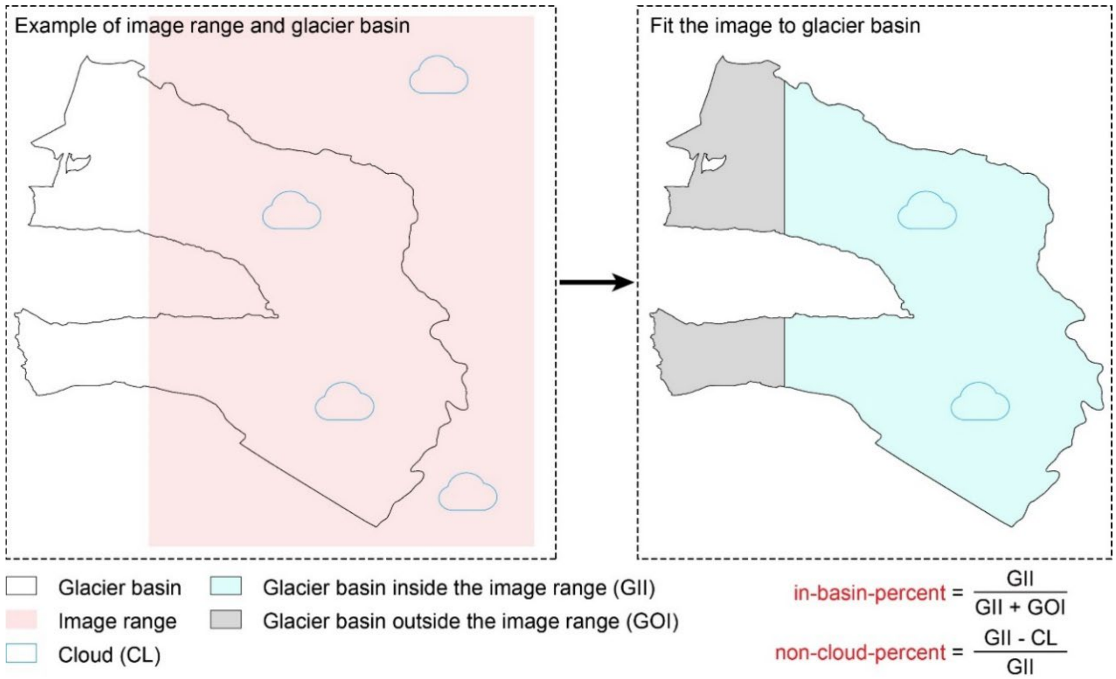

Figure 2. Schematic illustrations of the calculation of parameters for in-basin-percent and non-cloud-percent.

3.2 Satellite imagery acquisition

In this study, we used Landsat 8 images acquired from the product of Level-1 Tier 1 Top of Atmosphere (TOA) and Sentinel-2 images from the product of Level-1C (TOA). Both datasets are available for analysis through GEE. The TOA reflectance values are known to better represent surface conditions on the ice sheets because the pixel values are not influenced by differences in image acquisition conditions (Pope et al., 2016), and have been used previously in supraglacial lake studies (e.g. Moussavi et al., 2020; Williamson et al., 2018a). Although recently released Level-2 surface reflectance products may perform better, only a limited number of images are available for Landsat 8 and only available since May 2019 for Sentinel-2. We first selected images that cover the glacier masks from May to September in the periods 2014–2021 and 2016–2021 for Landsat 8 and Sentinel-2, respectively. Next, images with a sun elevation angle less than 20° were removed from the image collections because spectral difference between water surface and surrounding features is insignificant under low light conditions (Halberstadt et al., 2020; Moussavi et al., 2020; Dirscherl et al., 2020).

To eliminate images covering a limited portion of the glaciers, we introduced two parameters: “in-basin-percent” and “non-cloud-percent”. The in-basin-percent indicates how much of the image covers the glacier mask (Fig. 2), whereas the non-cloud-percent represents the portion of the image area unaffected by clouds. To calculate the non-cloud-percent, we obtained a cloud-likelihood score for every pixel in every image via the GEE built-in algorithm “simpleCloudScore” for Landsat 8 imagery. The “simpleCloudScore” algorithm determines a cloud-likelihood score in the range 0–100, using a combination of brightness, temperature, and NDSI. The score cannot be used as a robust cloud detector but is mainly intended to compare multiple views at the same point for relative cloud likelihood (https://developers.google.com/earth-engine/apidocs/ee-algorithms-landsat-simplecloudscore). In this study, we arbitrarily set a cloud score of 40 as the threshold, based on the visual interpretation of the Landsat images, meaning that pixels with a value higher than 40 are treated as clouds for Landsat 8 images. For Sentinel-2 images, we similarly obtained a cloud probability value (between 0 and 100) for every pixel in every image using the dataset of “Sentinel-2: Cloud Probability”, which is also available from GEE. This dataset is developed using the “Sentinel-2 Cloud Detector Library”. Details are available through Sentinel Hub (https://github.com/sentinel-hub/sentinel2-cloud-detector). As suggested by the data provider and our visual inspection, we set 65 as the probability threshold. By using these thresholds at pixel scale, we were able to estimate the cloud coverage within the glacier mask as “non-cloud-percent” (Fig. 2). Because our glacier mask was smaller than the area covered by a single Landsat 8 or Sentinel-2 image, knowing the cloud coverage within the glacier mask was helpful in selecting the images suitable for our study. By removing images with less than 90% of in-basin-percent and non-cloud-percent, we generated a Landsat 8 image collection with 166 images and a Sentinel-2 image collection with 317 images. The average values of in-basin-percent/non-cloud-percent were 99.4/98.4% for Landsat 8 and 99.2/98.7% for Sentinel-2, respectively.

3.3 Lake area delineation

To derive the lake extent from Landsat 8 and Sentinel-2 images, we employed a pixel-based RF supervised classification approach on GEE. An overview of the methodology is shown in Figure 3, and the processing steps are introduced in the following sections.

Table 1. Spectral bands and indices used as predictors to derive training data for Landsat 8 (L8) and Sentinel-2 (S2).

Predictors	**Formula**	**Satellites**	
		L8	S2
	Spectral bands		
blue		○	○
green		○	○
red		○	○
nir		○	○
swir1		○	
swir2		○	
vre (1–4)[1]			○
	Spectral indices		
$NDWI_{ice}$[2]	(blue − red) / (blue + red)	○	○
NDSI[3]	(green – swir1) / (green – swir1)	○	○
$AWEI_{sh}$[4]	blue + 2.5 × green – 1.5 × (nir + swir1) – 0.25 × swir2	○	○
$AWEI_{nsh}$[5]	4 × (green – swir1) – 0.25 × nir – 2.75 × swir2	○	○
TC_{wet}[6]	0.1509 × blue + 0.1973 × green + 0.3729 × red + 0.3406 × nir + 0.7112 × swir1 – 0.4572 × swir2	○	○
$SAVI_{mod}$[7]	(nir – red) / (nir + red + 1) × 2	○	○
NWI[8]	(blue – nir – swir1 –swir2) / (blue + nir + swir1 + swir2)	○	○
SI_{mod}[9]	(blue − nir) / (blue − nir)	○	○

Note: (1) Vegetation Red Edge bands of Sentinel-2 Imagery (band 5, 6, 7 and 8A). (2) Normalized Difference Water Index (Moussavi et al., 2020; Williamson et al., 2018a; Yang and Smith, 2013). (3) Normalized Difference Snow Index (Hall et al., 1995; Dozier, 1989). (4) Automated Water Extraction Index with the option of shadow (Feyisa et al., 2014). (5) Automated Water Extraction Index with the option of dark area removal (Feyisa et al., 2014). (6) Tasseled Cap for wetness (Crist and Cicone, 1984). (7) Modified Soil-Adjusted Vegetation Index (Huete, 1988). (8) New Water Index (Feng, 2012). (9) Modified Shadow Index (Li et al., 2016)

3.3.1 Training data generation

Training labels are required to support the supervised classification algorithms during the model training. To identify water pixels within the glacier mask, we created binary training labels as water and non-water. For each of the Landsat 8 and Sentinel-2 datasets, we selected training images (Table S1) from the image collections described in 3.2. In order to accurately identify seasonally changing water surface characteristics, the training images were selected from images covering the entire melt season (May–September). Furthermore, we deliberately chose typical images containing cloud and shade within the glacier masks, in order to enrich the training sample of non-water. For each of the training images, training regions were defined manually by drawing polygons in GEE. The regions were evenly distributed inside the glacier mask and contained information regarding water and non-water (including snow, ice, cloud, and shade).

Following the method described in Dirscherl et al. (2020), spectral bands and spectral indices were calculated and used as predictors for the classification process in discriminating between water and non-water. We employed 14 and 16 predictors for Landsat 8 and Sentinel-2 as listed in Table 1. Details of the spectral bands of the two satellites are listed in Tables S2 and S3. For the Sentinel-2 imagery, we harmonized the resolution of the spectral index calculation, by resampling data at 10-

m grid points with a bilinear resampling algorithm. All Landsat 8 images were resampled from a resolution of 30 m to 10 m to match the Sentinel-2 imagery for the time series generation (see Section 3.4 for details). Lastly, predictor values of the pixels within the pre-defined training regions were derived and fed into the classifier as training data. The relative importance of each predictor used for the RF classifier was determined using GEE.

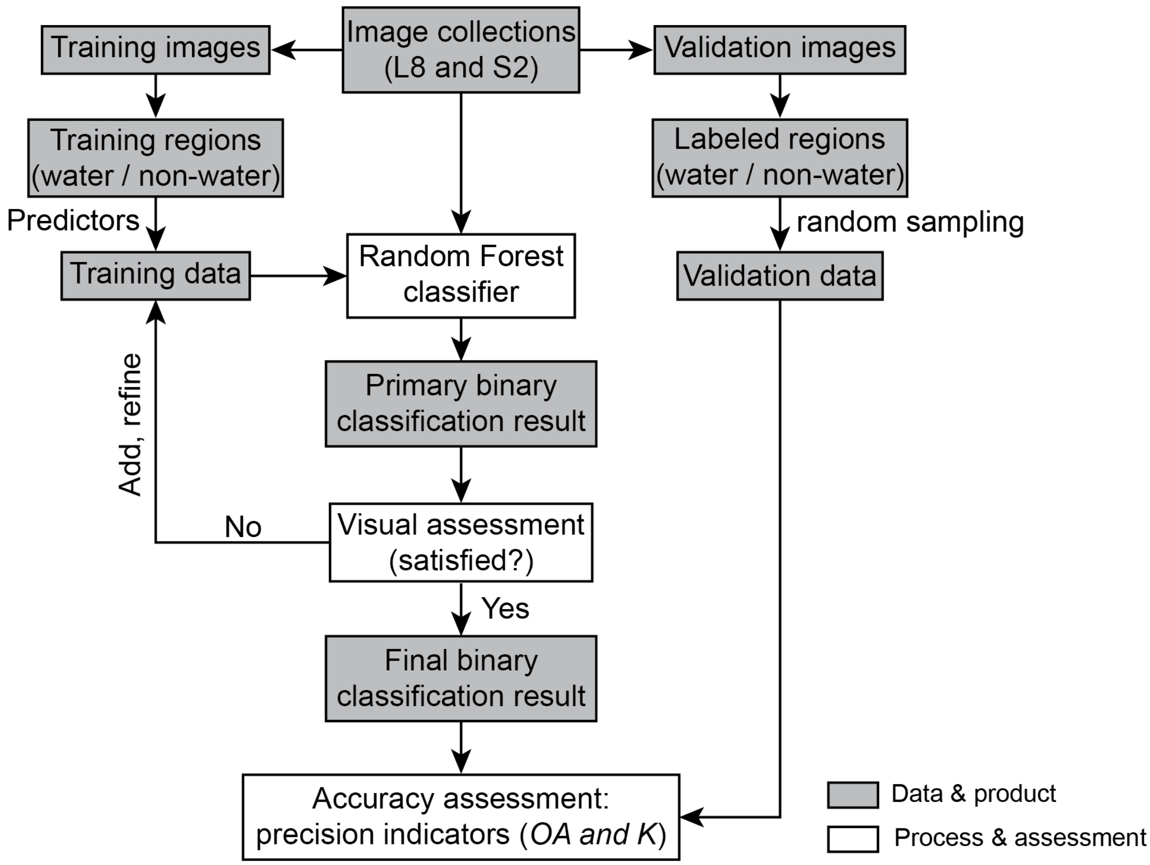

Figure 3. Flowchart illustrating lake area delineation by the RF method.

3.3.2 Image classification

We applied the RF classifier (Breiman, 2001) to classify each pixel of the selected satellite images. The RF algorithm consists of multiple uncorrelated random decision trees, which are bootstrapped and aggregated to classify a dataset, using the mode of predictions from all decision trees (Pal, 2005; Belgiu and Drăguţ, 2016). The RF method has been widely applied in satellite remote sensing, not only due to its robust performance, but also because of its computational efficiency and relatively easy implementation, compared to other machine learning methods (Belgiu and Drăguţ, 2016; Chan and Paelinckx, 2008; Denisko and Hoffman, 2018). Recent studies used this method for mapping supraglacial lakes on satellite images (Dell et al., 2021; Dirscherl et al., 2020, 2021b; Halberstadt et al., 2020).

The classification was carried out in GEE using an RF classifier trained with a built-in algorithm (ee.Classifier.smileRandomForest). The number of decision trees is an input parameter in the RF classifier. The overall accuracy of the classification improves as the number of decision trees increases (Breiman, 2001). However, as the number of trees increases, the computations become increasingly complex. As a result, there is an increased risk of exceeding computational limits in GEE and overfitting. We balanced the computational complexity and improved our level of accuracy, by setting the number of trees at 96 and 128 for Landsat 8 and Sentinel-2 imageries, respectively, after conducting multiple tests in GEE. Initial binary classification results were visually compared with the original optical image. For some obvious false classification areas (e.g. cloud shade, cloud, and water extent under thin cloud), we added or refined the training samples to improve the accuracy of the classification. This process was repeated until the classification became stable and a sufficiently high level of accuracy was achieved (Fig. 3).

3.3.3 Accuracy assessment

The performance of the classification was assessed using independent validation datasets (Fig. 3). Validation images were selected from the image collections described in Section 3.2, but not from those used as training images. We selected five pairs of Landsat 8 and Sentinel-2 images, which were taken within < 10 min of each other (Table S1). The images were visually inspected within the glacier masks to confirm that no obvious differences (e.g. fast cloud movement or rapid lake drainage event) existed between the image pairs.

We manually inspected the image pairs and labeled the water areas with polygon within GEE, treating the labeled area as the ground truth. The labeled water area was applicable to the Landsat 8 and Sentinel-2 image pairs since the lake conditions are nearly the same. For every validation scene, we randomly sampled 1500 pixels within the labeled water area, as water pixels. We also sampled 1500 pixels randomly from off-water areas within 300 m of the lake margin, as non-water pixels. These non-water pixels are important for the validation because misclassification tends to occur in areas surrounding water (Dirscherl et al., 2020). To ensure that the sample pixels represented various environmental objects, we further selected 1500 non-water pixels randomly from cloud, shade, and slush areas. The selected water and non-water pixels were taken as a manually labeled binary result (hereafter, referred to as actual result). As an example, manually labeled lakes and lakes detected by the RF-based method on 6 August 2021 are compared in Fig. S1.

The accuracy assessment was performed by comparing the actual result with the prediction by the RF classifier. It is shown as a confusion matrix (Table 2, Stehman, 1997). Following the methods employed by Dirscherl et al. (2020) and Teluguntla et al. (2018), we calculated the user accuracy (*UA*) and producer accuracy (*PA*), which were used to evaluate the performance of the RF classifier for each class. Additionally, overall accuracy (*OA*) and Cohen's Kappa (*K*) were computed to assess the overall performance of the RF classifier (Dirscherl et al., 2020; Landis and Koch, 1977; Teluguntla et al., 2018).

Table 2. A framework of the confusion matrix used for the accuracy assessment.

	Actual	*UA*

		water	non-water	
Predict	water	True positive (*TP*)	False negative (*FN*)	*TP* / (*TP* + *FN*)
	non-water	False positive (*FP*)	True negative (*TN*)	*TN* / (*FP* + *TN*)
	PA	*TP* / (*TP* + *FP*)	*TN* / (*FN* + *TN*)	

The precision indicators (*OA* and *K*) were calculated as follows (*EA*: expected accuracy, a parameter used for calculating *K*):

$$OA = \frac{TP + TN}{TP + FN + FP + TN} \quad (1)$$

$$K = \frac{OA - EA}{1 - EA} \quad (2)$$

$$EA = \frac{(TN + FP) \times (TN + FN) + (FN + TP) \times (FP + TP)}{(TN + FN + FP + TP) \times (TN + FN + FP + TP)} \quad (3)$$

3.4 Time series generation

We obtained two binary data sets from Landsat 8 and Sentinel-2 imageries at a 10 m resolution through the lake area delineation. Each individual supraglacial lake was identified by grouping water pixels connected by a common corner (Dell et al., 2020). The MATLAB function "bwconncomp" was used for this purpose. Small clusters (≤ 18 pixels in total) and linear features (≤ 3 pixels wide) were removed because these features typically represent slush or supraglacial streams rather than lakes (Moussavi et al., 2020; Pope et al., 2016; Williamson et al., 2018a). Benefited from the overlap of satellite tracks in the polar region, a combination of Sentinel-2 and Landsat 8 imageries enables us to monitor the supraglacial lakes at a sub-week temporal scale. Moreover, dual-satellite monitoring has been used in recent studies to track the supraglacial lake evolution in Greenland (Williamson et al., 2018a) and Antarctica (Dell et al., 2020; Moussavi et al., 2020). Here, we followed the method of Williamson et al. (2018a) to generate a time series of the lake area between 2014 and 2021 using Landsat 8 and Sentinel-2 imagery. To test the correlation between Landsat 8 and Sentinel-2 lake areas, we compared the areas of 1249 lakes obtained from 21 pairs of Landsat 8 and Sentinel-2 images acquired within < 20 minutes of each other. The result showed a significant correlation between the two sets of lake areas ($r^2 = 0.996$, $p = 0.000$, RMSE = 0.010 km^2) (Fig. S2). We then fused the Landsat 8 and Sentinel-2 lake masks between 2014 and 2021 into one time series. When both Landsat 8 and Sentinel-2 lake masks were available on the same day, the Sentinel-2 lake mask was preferentially used because of its higher resolution. In the case of two Sentinel-2 lake masks being available on the same day, we employed images with a higher non-cloud-percent.

3.5 Climate data

The formation and development of supraglacial lakes were compared with air temperature, surface snowmelt, runoff, and surface mass balance (SMB). Air temperature data at Qaanaaq station, situated approximately 75 km west of the termini of Tracy and Heilprin Glaciers (77.48°N, 69.38°W; 16 m a.s.l., Fig. 1a), is available from NOAA's National Centers for Environment Information (https://www.ncei.noaa.gov/). Daily mean temperature from June to August was utilized to compute the summer positive degree day sum (PDD). Snowmelt, runoff and SMB modelled by the Regional Atmospheric Climate Model version 2.3p2 (RACMO2.3p2) are provided as a downscaled dataset with a resolution of 1 km (Noël et al., 2016; 2019). The RACMO2.3p2 product was utilized to compute the three-day moving average surface snowmelt and runoff rates for three elevation bands (0–400, 400–800, 800–1200 m a.s.l.). The initial dates of snowmelt and runoff (IDM and IDR) were defined in each elevation band, based on the first observation of snowmelt and runoff greater than 4 mm w.e. d^{-1} for five consecutive days in each year. The initial date of lake formation (IDL) was defined in each elevation band and for the whole glacier area, based on the first observation of lake area greater than 0.2 km^2 for three consecutive observations in each year.

3.6 Ice speed and topography

We extracted the ice speed of the glaciers from the NASA MEaSUREs ITS_LIVE time-averaged (1985–2018) velocity mosaic dataset (Gardner et al., 2019). Surface elevation and topography were analyzed using the ArcticDEM Mosaic dataset (Porter et al., 2018). Additionally, surface depressions (>0.05 km^2) were identified by applying the MATLAB package TopoToolbox (Schwanghart and Scherler, 2014) on the ArcticDEM.

4 Results

4.1 Evaluation of the classification method

The RF classifier provides a quantitative measure of each predictor's contribution to the classification result. This is useful in evaluating the importance of each predictor. The results showed that all predictors (spectral bands and spectral indices) contributed toward the classification of water and non-water features in both satellite imagery datasets (Fig. S3). For the Landsat 8 imagery, $NDWI_{ice}$ made the most significant contribution, with a relative importance of over 10% (Fig. S3a). Spectral bands blue, swir2, swir1, and spectral indices $SAVI_{mod}$, NDSI, also contributed significantly (> 7%) to the classification (Fig. S3a). For the Sentinel-2 imagery, the most important contributors were NDSI and $NDWI_{ice}$ with a relative importance of 8.1% (Fig. S3b). Spectral bands vre4 (band 8A), blue, and vre2 (band 6) showed a relative importance of over 6% (Fig. S3b). The remaining predictors showed similar importance, ranging between 6.0–6.8% for Landsat 8 (Fig. S3a) and 5.3–6.0% for Sentinel-2 imagery (Fig. S3b).

The results of the performance assessment showed a similarly high level of overall accuracy, namely 98.5% and 98.6% for Landsat 8 and Sentinel-2, respectively (Table S4). Another indicator, Kappa coefficient, also showed similarly high values for Landsat 8 (0.97) and Sentinel-2 (0.97). As for producer accuracy, the performance was slightly lower for water class than non-water class, both on Landsat 8 (97.3% vs 99.0%) and Sentinel-2 imageries (96.73% vs 99.46%). However, user accuracy was similar for water and non-water classes on Landsat 8 (98.1% vs 98.7 %) and Sentinel-2 imageries (98.9% vs 98.4%).

Despite the relatively strict criterion to eliminate the influence of clouds, a small number of the selected images were extensively covered by thin clouds. This was due to the inability of the “Cloud Probability” of Sentinel-2 imagery and “simpleCloudScore” for Landsat 8 imagery to identify thin clouds. In such cases, thin clouds sometimes hindered the discrimination of water as examples shown in Figure S4. Nevertheless, thin clouds were found mostly over lake-free regions. Moreover, the could-induced underestimation of a lake area can be realized by analyzing the continuous time series of the data. Thus, the effect of thin clouds on our study was insignificant.

4.2 Spatial extent and distribution of supraglacial lakes

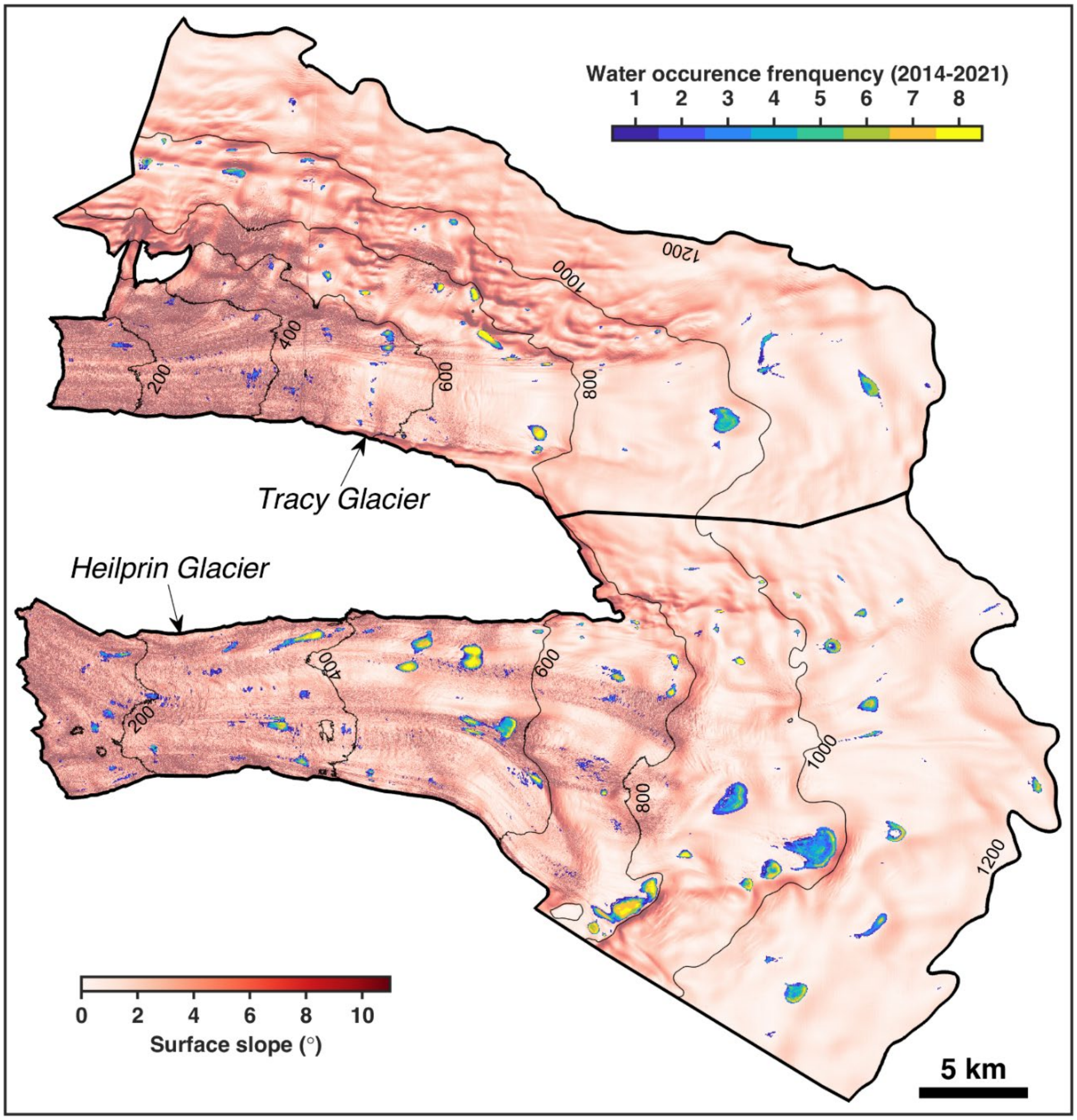

Figure 4. Spatial distribution of supraglacial lakes on Tracy and Heilprin Glacier. The blue-yellow color scale shows the recurrence frequency of water pixels between 2014 and 2021. The background red color scale shows the surface slope. The thick black line shows the glacier boundary before the introduction of a 300-m marginal buffer described in 3.1. Elevation contours (thin grey lines) derived from ArcticDEM Mosaic (Porter et al., 2018).

To understand the maximum lake extent over the eight melt seasons, Figures 4 and 5a show the lake extent cumulated over the period 2014–2021. Supraglacial lakes were found in every 200 m elevation band on both Heilprin and Tracy Glaciers (Fig. 4). Although the study areas of the two glaciers are similar, the lake extent on Tracy Glacier is only one-third of that on Heilprin Glacier (Table 3). The lake area increased with elevation, reaching the largest area in the elevation band 800–1000 m a.s.l. both at Heilprin (5.94 km^2) and Tracy Glaciers (2.28 km^2) (Fig. 5a). Despite the glacier area increases at higher elevation (Fig. 5b), lake area dropped in the elevation band 1000–1200 m a.s.l. at Heilprin (3.66 km^2) and Tracy Glaciers (1.37 km^2) (Fig. 5a).

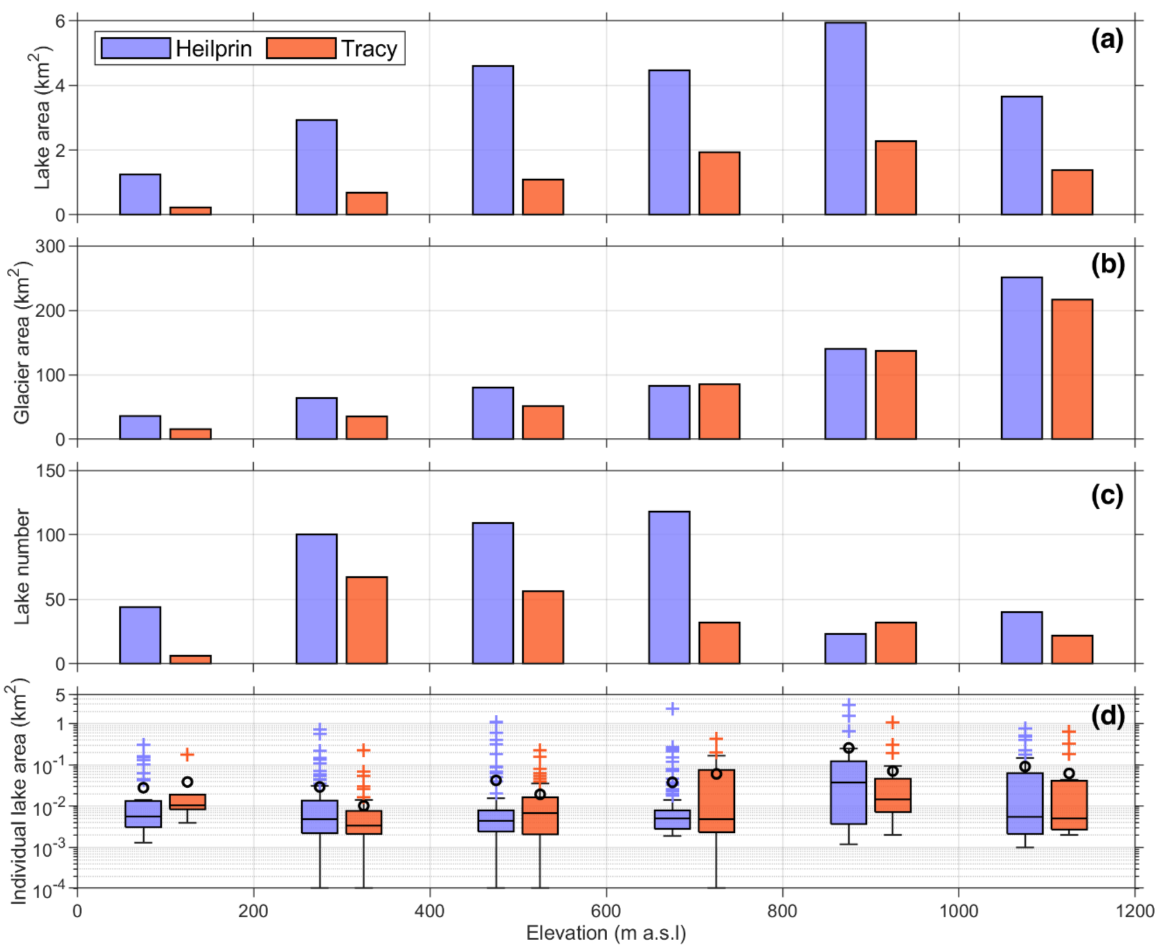

Figure 5. (a) Lake area, (b) glacier area, (c) lake number, and (d) individual lake areas for Heilprin and Tracy Glaciers in each 200 m elevation band. The lake data are cumulated over the period of

2014–2021. The glacier areas in (b) are calculated based on the glacier boundary and elevation contours in Figure 4. In (d), the black circle indicate the mean individual lake area in each 200 m elevation band for each glacier.

Based on the maximum lake extent cumulated over the eight melt seasons (Fig. 4), we observed 434 and 215 supraglacial lakes on Heilprin and Tracy Glaciers, respectively (Table 3). On Heilprin Glacier, the number of lakes increased with altitude, reaching the maximum (118) in the elevation band 600–800 m, whereas it sharply decreased above 800 m (Fig. 5c). On Tracy Glacier, the maximum number of supraglacial lakes (67) occurred in the elevation band 200–400 m a.s.l. and gradually decreased upglacier (Fig. 5c). Mean lake areas were 0.053 km^2 and 0.035 km^2 for Heilprin and Tracy Glacier, respectively (Fig. 5c). Only 2% of the lakes were larger than 0.5 km^2. They were mainly distributed in areas higher than 400 m a.s.l. (Fig. 5d). The largest lakes on Heilprin (2.81 km^2) and Tracy Glaciers (1.09 km^2) were both located at an elevation of 950 m a.s.l (Fig. 5d).

Table 3. Statistics for topography and supraglacial lakes on Heilprin and Tracy Glaciers.

	Heilprin Glacier	Tracy Glacier
Glacier area below 1200 m a.s.l. (km^2)	654.15	540.62
Average surface slope (°)	2.5	2.8
Maximum lake area (km^2)	22.84	7.60
Total lake number	434	215
Mean lake area (km^2)	0.05	0.04
Largest lake area (km^2)	2.81	1.09
Fraction of water recurred least (1–2 times)	44.1 %	49.9 %
Fraction of water recurred most (7–8 times)	13.9 %	9.3 %

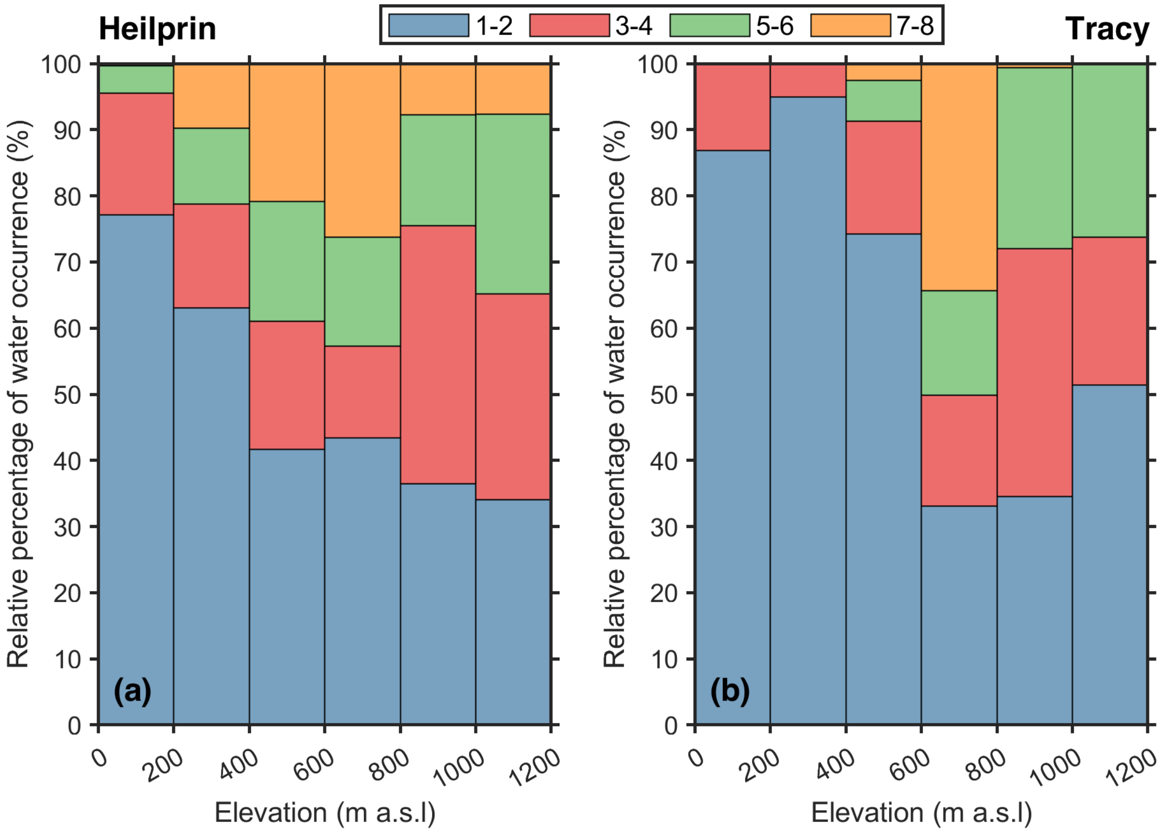

Figure 6. Fractions of water occurrence frequencies over the studied eight years (2014–2021) in each 200 m elevation band for (a) Heilprin and (b) Tracy Glaciers.

Table 3 and Figure 6 present the occurrence of lake formation during the eight years of the study. Over 50% of the water pixels occurred more than twice on both of the glaciers (Table 3). Furthermore, 13.9% and 9.3% of water pixels on Heilprin and Tracy Glaciers were observed almost every year (7−8 times) (Table 3). Water pixels were observed more randomly near the glacier terminus, whilst appearing repeatedly at the same locations upstream (Fig. 6). On Heilprin Glacier, over 70% of the water pixels in the elevation band 0–200 m a.s.l. appeared only once or twice (Fig. 6a). In the upper reaches, water formed at the same location more frequently, i.e. more than 7% of the water pixels recurred 7–8 times in the area above 200 m a.s.l. On Tracy Glacier, > 70% of the water pixels appeared only once or twice in the region below 600 m a.s.l., whereas > 50% of the water pixels appeared more than twice in the upper region (Fig. 6b). Water pixels on Tracy Glacier recurred most frequently in the middle of the glacier. Specifically, 34% of water pixels were observed at least 7–8 times in the elevation band 600–800 m a.s.l., whereas such pixels represented only 3% at other elevations (Fig. 6b).

4.3 Seasonal lake evolutions

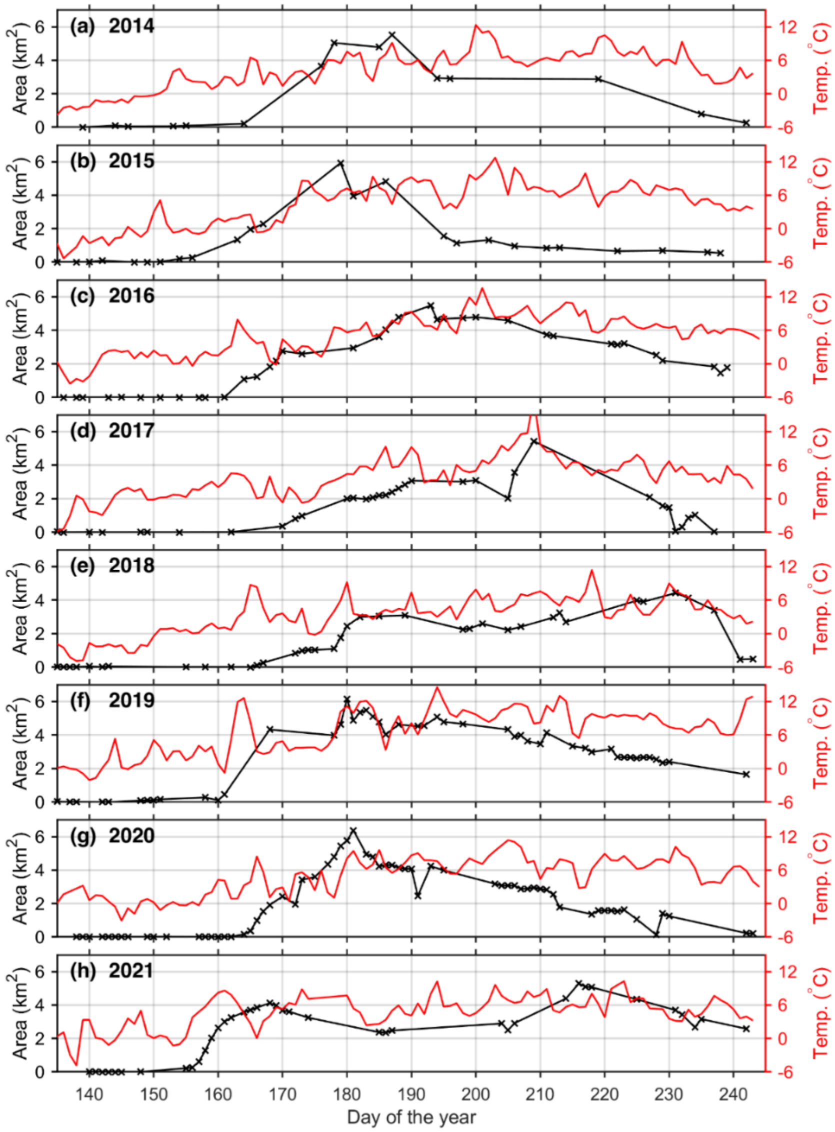

Figure 7. Time series of the total lake area on Heilprin and Tracy Glaciers (black) and daily mean air temperature at Qaanaaq station (red during the melting seasons from 2014 to 2021).

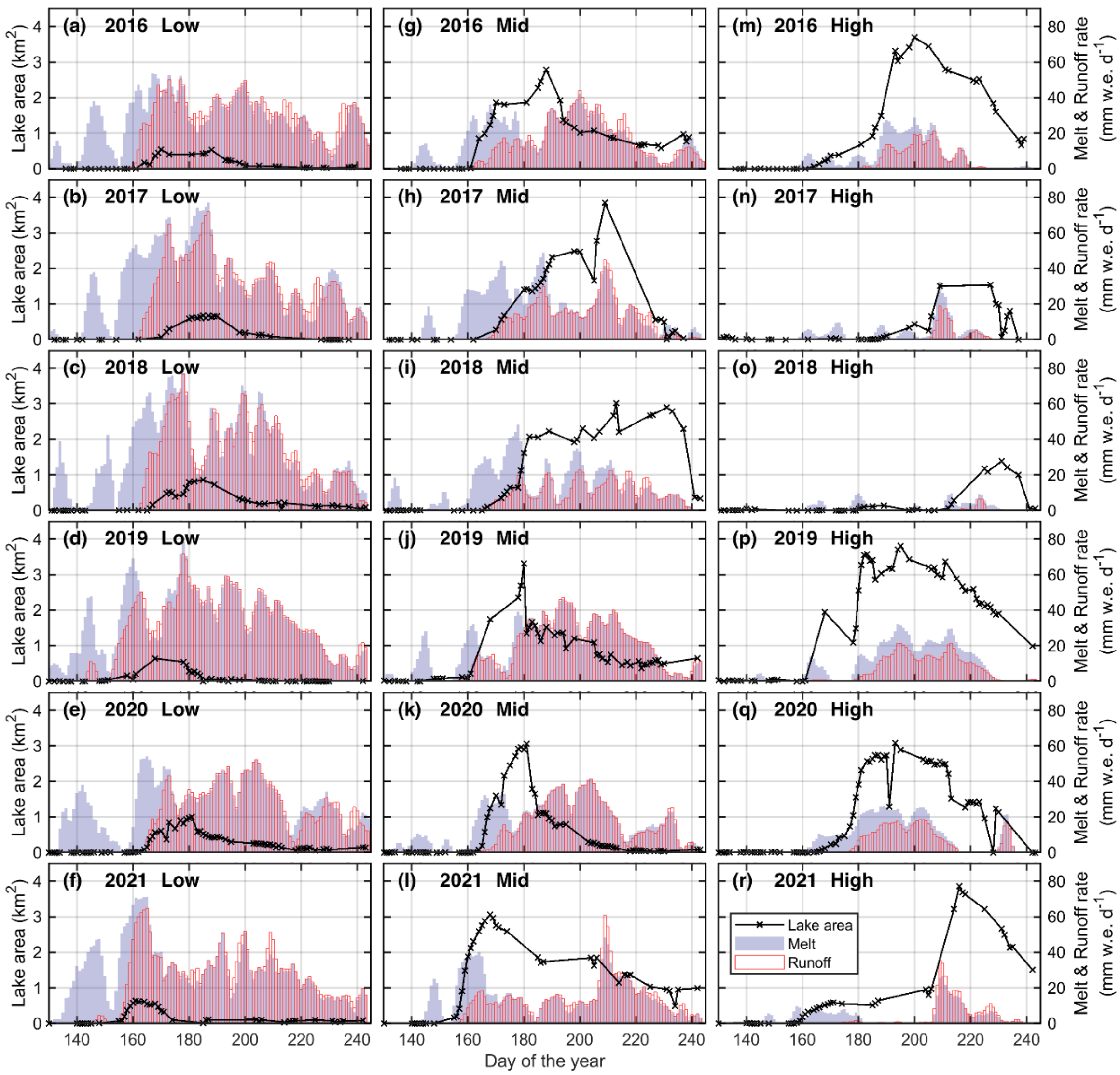

Figure 8. Time series of total lake areas (×) and 3-day moving average RACMO melt and runoff rates (bar graph) in low (0–400 m a.s.l.), mid (400–800 m a.s.l.), and high (800–1200 m a.s.l.) elevation areas on Tracy and Heilprin Glaciers between 2016 and 2021.

Over the eight melting seasons, the IDL was usually observed in early- to mid-June (days 155–170) (Fig. 7). Nevertheless, the pattern of subsequent lake growth was significantly different each year. In the years 2014–2016 and 2019–2020, the lake area increased rapidly, peaking in late June to early July (days 180–190) (Figs. 7a–c, f–h). In 2017 and 2018, the seasonal increase in lake area was relatively slow, with the peak occurring in late July or even August (Figs. 7d–e). In 2021, although lake formation started early and increased substantially by day 170, the growth of the lake slowed down. The maximum area was not reached until early August (day 216) (Fig. 7h). After reaching its maximum, the lake area gradually decreased toward the end of the melt season. Most of the lakes disappeared by the end of August (~day 240), except for a few large lakes in the high-elevation areas on Heilprin Glacier, which lasted until the last available image.

Lakes usually started to form with the first sharp rise in temperature (>5 °C increase within five days) and lake areas generally fluctuate in accordance with variations in air temperature (Fig. 7). However, in the years 2015, 2016, 2019, and 2020, the correlation between air temperature and lake area was observed only until the lake area peaked. (Figs. 7b–c, f–g). Following the peak, the lake area progressively decreased, regardless of temperature fluctuations. In contrast, variations in the lake area in the years 2017, 2018, and 2021, did correspond with changes in air temperature, right until the end of the melt season (Figs. 7d–e, h). The disappearance of the lakes was due to either surface refreezing, snow cover or drainage into the englacial or subglacial environment. Surface freezing and snow cover occurred typically in areas higher than 800 m a.s.l. in the late melt season or after heavy snowfall. Lakes were gradually covered by floes growing from the lake margin or by a frozen water surface expanding from the lake center (e.g. Fig. S5). Lake drainages were generally observed prior to August in low and mid-elevation areas. Most of the drainage events took place over the course of several weeks, but some lakes drained rapidly through moulins or crevasses within one or two days (e.g. Fig. S7).

Time series of lake area showed significantly different seasonal patterns in low (0–400 m a.s.l.), middle (400–800 m a.s.l.) and high (800–1200 m a.s.l.) elevation bands (Fig. 8). At low elevation, the onset of lake formation was the earliest, but the maximum lake area was limited to < 1 km² throughout all six melt seasons. Lakes in this region were relatively short-lived, i.e. most of the lakes disappeared before day 200 (Figs. 8a–f). In the mid-elevation areas, lakes began forming at a rate similar to that in the low-elevation areas, but the lake area increased at a faster rate and reached a greater maximum area (Figs. 8g–i). In 2019 and 2020, the lake area declined rapidly after its peak, losing more than half its area in two days (Figs. 8j, k). Inspection of the images showed a cluster of drainage events during these periods. Lake formation in high-elevation areas began later than in the lower reaches (Figs. 8m–r). Specifically, in 2017 and 2018, lake formation began in late July, >30 days later than the onset in the lower reaches. The maximum lake area in these years was limited to ~1.5 km^2, which was less than 40% of those observed at middle elevation (Figs. 8n, o). In the years 2016, 2019, and 2020, the lake areas in the high-elevation area showed a similar seasonal variation to those in low- and mid-elevation areas. In these years, the maximum areas in the high-elevation area exceeded those in the lower areas (Figs. 8m, p, q).

Lake formation started at approximately the same time as the initiation of runoff in low and mid-elevation areas (Fig. 8a–l). In the high-elevation areas, however, lakes began to form approximately at melt onset rather than runoff (Fig. 8m–r). Although the melt and runoff rates were greater at lower elevation, they did not result in a larger lake area in the low-elevation area (Figs. 8a–f). While the lake area at low elevation was relatively insensitive to melt/runoff, a short spike in melt or runoff coincided with a significant increase in lake area in the high-elevation area (Figs. 8m–r).

4.4 Interannual lake evolutions

The maximum lake area showed slightly different interannual variations at Heilprin and Tracy Glaciers (Fig. 9). On Heilprin Glacier, the lake area varied between 7.45 km^2 in 2015 and 12.41 km^2 in 2019 (Fig. 9a). The year-to-year variation was more pronounced on Tracy Glacier, where the lake area in 2017 (1.65 km^2) and 2018 (1.75 km^2) was only 40% of the lake extent in 2019 (4.05 km^2) (Fig. 9b).

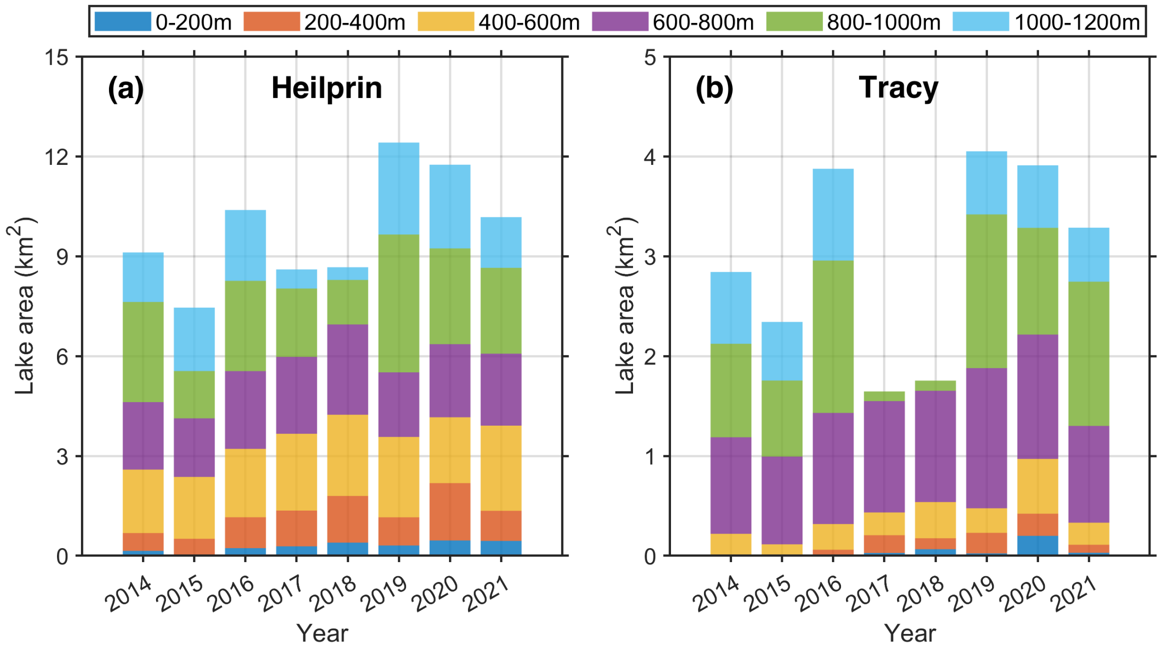

Figure 9. Maximum lake areas observed each year within individual 200 m elevation bands for (a) Heilprin and (b) Tracy Glaciers.

The amplitude of the interannual lake area variations was greater in areas of higher elevation. On Heilprin Glacier, the lake area below 800 m a.s.l. ranged from −27% (2015) to +23% (2018) of the mean over the study period, whereas that in the upper reaches ranged from −59% (2018) to +65% (2019) (Fig. 10a). On Tracy Glacier, the lake area below 800 m a.s.l. varied from −35% (2015) to +45% (2020), while a more significant variation ranging from −93% (2017 and 2018) to +70% (2016) was observed above 800 m a.s.l. (Fig. 10b). Supraglacial lakes were most widely distributed above 800 m a.s.l. (Figs. 4 and 5a), with a mean coverage accounting for 42% and 43% of the maximum lake areas on Heilprin and Tracy Glaciers, respectively. However, lake extents above 800 m were anomalously small in 2017 and 2018 (Fig. 9), accounting for only ~30% (Heilprin) and 6% (Tracy) of the maximum lake areas. In these two years, even no lakes were observed in the 1000–1200 m a.s.l. elevation band of Tracy Glacier (Figs. 9b and S6d–e). Relatively high summer

PDD (561 and 710 °C day) in 2016 and 2019 caused lake extents above 800 m a.s.l. to be at their highest during the study period. (Fig. 10). In contrast, relatively low summer PDD in 2017 and 2018 (439 and 395 °C day) resulted in the smallest lake extent (Fig. 10). Therefore, year-to-year variations in the maximum lake area were predominantly affected by lake formation above 800 m a.s.l., which was controlled by summer PDD.

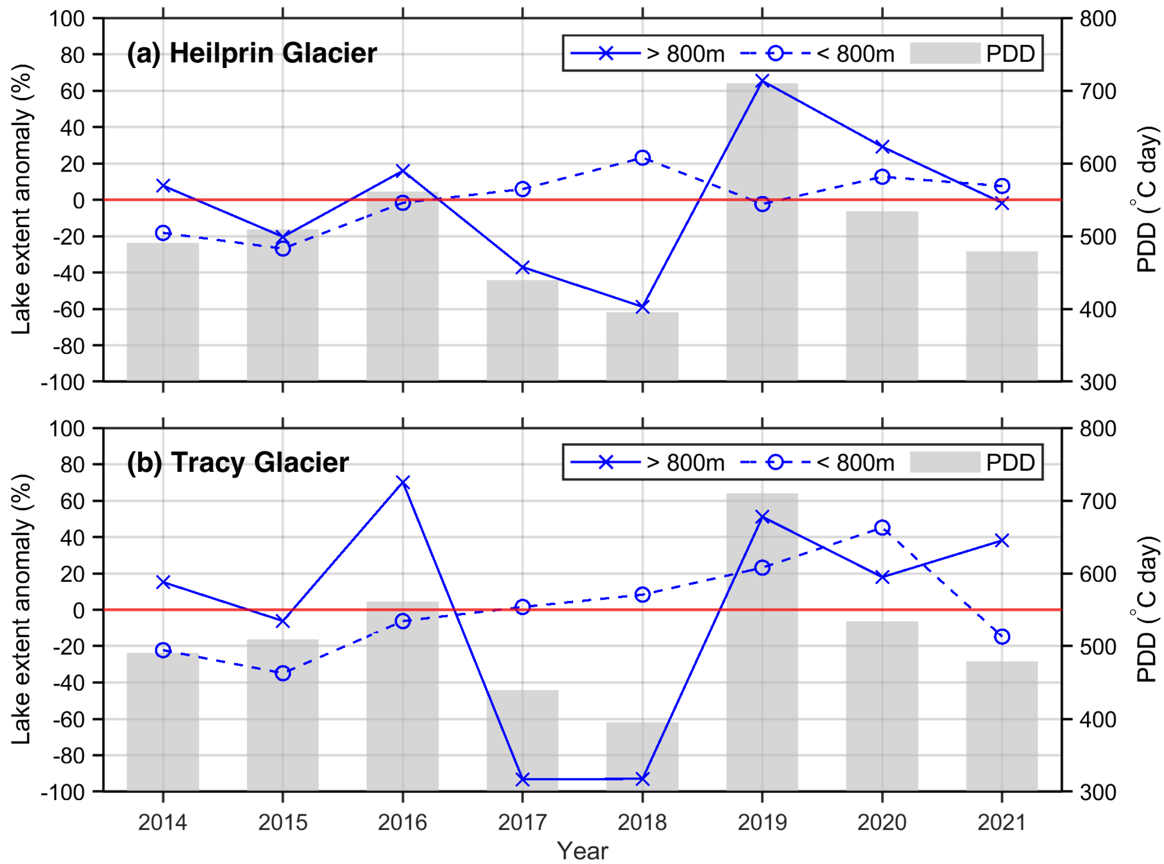

Figure 10. Lake extent anomalies above (×) and below 800 m a.s.l. (○) for (a) Heilprin and (b) Tracy Glaciers. The bar graph denotes the positive degree day sum (PDD) between June–August at Qaanaaq station each year.

5 Discussion

5.1 Spatial distribution of supraglacial lakes

On the Greenland ice sheet, supraglacial lakes are distributed with a significant heterogeneity. The greatest lake area is reported in southwestern and northeastern Greenland, whereas it is smaller in southeast Greenland (Selmes et al., 2011; Hu et al., 2021; Zhang et al., 2023). These observations are consistent with a relatively small extent of bare ice in southeastern Greenland (Ryan et al., 2019). Individual lakes in southwestern, northwestern, and northeastern Greenland are characterized by relatively large area (0.5–20 km^2), whereas lakes in central western, northern, and southeastern Greenland are smaller (< 10 km^2) (Hu et al., 2021). Supraglacial lakes form primarily in surface depressions. Because supraglacial lakes form in surface depressions generated under the influence of subglacial topographic undulations, lakes tend to form annually at the same location (Echelmeyer et al., 1991; Sergienko, 2013). In our study glaciers, we find no obvious evidence of supraglacial lake migration with ice flow except for some water-filled crevasses near the glacier fronts. This is in contrast to lakes on floating glacier tongues, where lakes migrate downglacier with ice flow (e.g. Petermann Glacier, Macdonald et al., 2018; 79° N Glacier, Turton et al., 2021; Ryder Glacier, Otto et al., 2022). The spatial distribution of lakes on Tracy and Heilprin Glaciers aligns with those of surface depressions, particularly in higher elevation areas (Fig. 11a). Throughout the study period, large lakes have formed repeatedly within these depressions (Fig. 4). These findings are in line with a previous report by Ignéczi et al. (2016), which suggested that surface depressions provide favorable locations for the development of supraglacial lakes. Therefore, we assume that surface depressions derived from DEMs have the potential to predict future lake distributions in high-elevation areas. According to the results of our study, lake formation at surface depressions is more sporadic in the high-elevation areas as compared to the mid-elevation areas. However, as climate warming continues, greater amounts of water are available and bare ice is more exposed in these regions. Therefore, a larger proportion of surface depressions will be filled with water in the high-elevation areas and the upper limit of lake growth is expected to migrate inland (Howat et al., 2013; Leeson et al., 2015; Ignéczi et al., 2016).

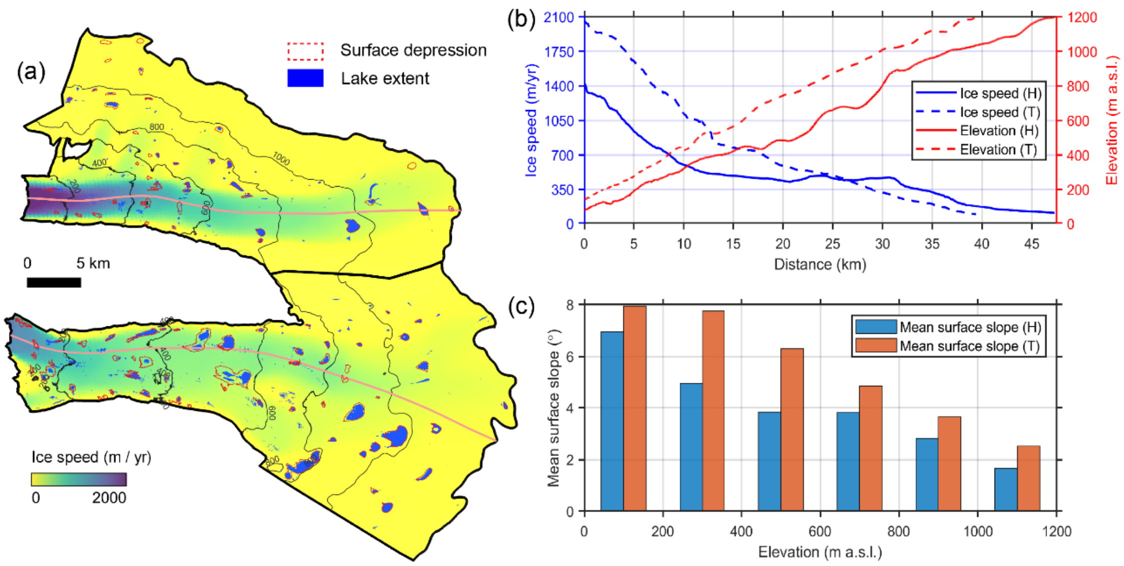

Figure 11. (a) Maximum lake extent observed from 2014 to 2021 (blue patches) and surface depressions (> 0.05 km^2) (red dashed line polygons) on Heilprin and Tracy Glaciers. The color scale shows ice speed distribution from ITS_LIVE time-averaged (1985–2018) velocity mosaic dataset. The two magenta curves are flowlines used in (b). (b) Ice speed and elevation along the flowlines of Heilprin (solid) and Tracy Glaciers (dashed) shown in (a). (c) Mean surface slope in each 200 m elevation band for Heilprin (blue) and Tracy Glaciers (red).

The largest lake extent was observed in the elevation band 800–1000 m a.s.l. for both Heilprin and Tracy Glaciers (Fig. 5a). Additionally, the largest individual lake occurred in this elevation band (Fig. 5d). Ice speed in this elevation band was slower than the lower elevation bands both in Heilprin and Tracy Glaciers (Figs. 11a and b). Moreover, the mean surface slope in this elevation band was 2.8° and 3.6° in Heilprin and Tracy Glaciers, respectively, indicating the regions are flatter than the lower elevation bands (Figs. 11b and c). Most likely, the slow movement of ice and gentle slope of the surface facilitated the growth of supraglacial lakes in this elevation band. Although ice velocity and surface slope further decrease in the 1000–1200 m a.s.l. elevation band and glacier area continues to increase, the lake extent decreases in the highest elevation band (Figs. 5a and b). We assume that the growth of supraglacial lakes in this elevation range was limited by the insufficient availability of meltwater, as well as meltwater absorption by snow layers (Koenig et al., 2015; Ignéczi et al., 2016). On Qaanaaq ice cap, situated 75 km west of the glaciers studied, the equilibrium line altitude (ELA) ranged from 862–1001 m a.s.l. between 2012 and 2016 (Sugiyama et al., 2014; Tsutaki et al., 2017). Therefore, it is likely that meltwater generated above 800 m a.s.l. was absorbed by snow during the early melt season. In summary, ice surface topography provides potential locations for supraglacial lake formation, while meltwater availability controls actual lake formation (Chu, 2014; Arthur et al., 2020).

The results in Section 4.2 reveal that lakes in the low-elevation area are smaller and more randomly distributed than those at higher elevation (Fig. 6). Crevasses are densely distributed near the front of calving glaciers (Fig. S9) because of tensile stresses generated by a stretching flow regime (Benn et al., 2007). Since crevasses are typically narrow, deep, and have a distinctive elongated shape (Colgan et al., 2016), it is difficult for large supraglacial lakes to form near the glacier front. Lakes tend to grow larger in relatively flat and crevasse-free areas (Turton et al., 2021; Arthur et al., 2020; Sundal et al., 2009). Although meltwater-filled crevasses were observed in the early melt season, the relatively fast ice flow of the studied glaciers (>1000 m a^{-1} near the terminus) facilitated lake drainage due to crevasse opening (Colgan et al., 2016). Therefore, formation of new lakes and rapid drainage in crevasses were frequently observed near the terminus, resulting in random distribution of lakes in the region. In the more inland areas, on the other hand, relatively flat surfaces and lower ice speed facilitate the growth of lakes as long as meltwater is available on the glacier surface (Sundal et al., 2009; Hu et al., 2022; Jiang et al., 2022; Zhang et al., 2023).

A higher number of lakes and larger lake areas were observed on Heilprin Glacier as compared to Tracy Glacier (Figs. 5a and c). Ice speed of Tracy Glacier is approximately 1.5 times greater than that of Heilprin Glacier (Sakakibara and Sugiyama, 2018; 2020), and ice near the front is thinning at a rate more than five times faster than Heilprin Glacier (Willis et al., 2018; Wang et al., 2021). Consequently, the glacier surface topography of Tracy Glacier was more prone to change. As a result of rapidly changing ice conditions, lakes downstream from Tracy Glacier are smaller and distributed more randomly (Figs. 6 and 8a). Moreover, our observations indicate a more extensive distribution of surface depressions and flatter surface slopes on Heilprin Glacier (Figs. 11). This, in turn, facilitated the formation of larger lakes. We, therefore, attribute the more abundant lake formation on Heilprin Glacier to a combination of glacier dynamics and topographic factors.

5.2 Seasonal evolution of supraglacial lakes

Lake formation began in low-elevation areas in early June (days 155–165, Figs. 7–8) and distribution expanded to higher elevation as the melting season progressed. This observation is consistent with those recorded in northern Greenland on Petermann Glacier (81°N in northwestern Greenland, Macdonald et al., 2018), 79°N Glacier (79°N in northeastern Greenland, Turton et al., 2021; Sundal et al., 2009), and Ryder Glacier (81°N in northwestern Greenland, Sundal et al., 2009; Otto et al., 2022). In lower latitude, on Russell Glacier (67°N in western Greenland, Fitzpatrick et al., 2014; Sundal et al., 2009) and in the Paakitsoq region (69°N in western Greenland, McMillan et al., 2007), lakes began to form as early as May. Generally, air temperature and winter snowpack govern surface melt and runoff, which, in turn, control the initiation of lake formation (Arthur et al., 2020; Fitzpatrick et al., 2014; Macdonald et al., 2018; Tuckett et al., 2021). Our study demonstrates a close relationship between the occurrence of runoff and the appearance of supraglacial lakes in the lower reaches of the glaciers (Fig. 8). A significant correlation was found between IDL and IDR in low (0–400 m a.s.l.) and middle (400–800 m a.s.l.) elevation areas (Fig. 12a). Typically, there is a 20 to 30-day delay between the onset of surface melt and runoff in these elevation ranges (Fig. 8). During this period, meltwater runoff is inhibited by percolation into layers of snow and firn (Harper et al., 2012; Chu, 2014; Tedstone and Machguth, 2022). Supraglacial lakes form when snow and firn layers are saturated with meltwater, resulting in the onset of runoff. In contrast to the general agreement between the onset of supraglacial lake formation and runoff, lake formation at high elevation was significantly earlier than that of runoff, occurring only slightly after the onset of surface melting (Figs. 8 and 14b). This observation can be explained by the existence of supraglacial lakes under the snow layer. In low- and mid-elevation areas, lakes typically fully drain or freeze before the end of the melting season. At high elevation, some large and deep lakes stay unfrozen over the winter under a layer of snow or thin ice (Koenig et al., 2015; Dunmire et al., 2021, Figs. S5b–c). In the following melting season, the lakes are exposed as soon as the snow or ice cover melts, even before the onset of runoff. While some lakes appear immediately after the onset of surface melt, following the above-mentioned process, a substantial increase in lake area occurs only after the onset of runoff. This suggests that most of the lakes in high-elevation areas follow the formation pattern observed at middle or low elevation.

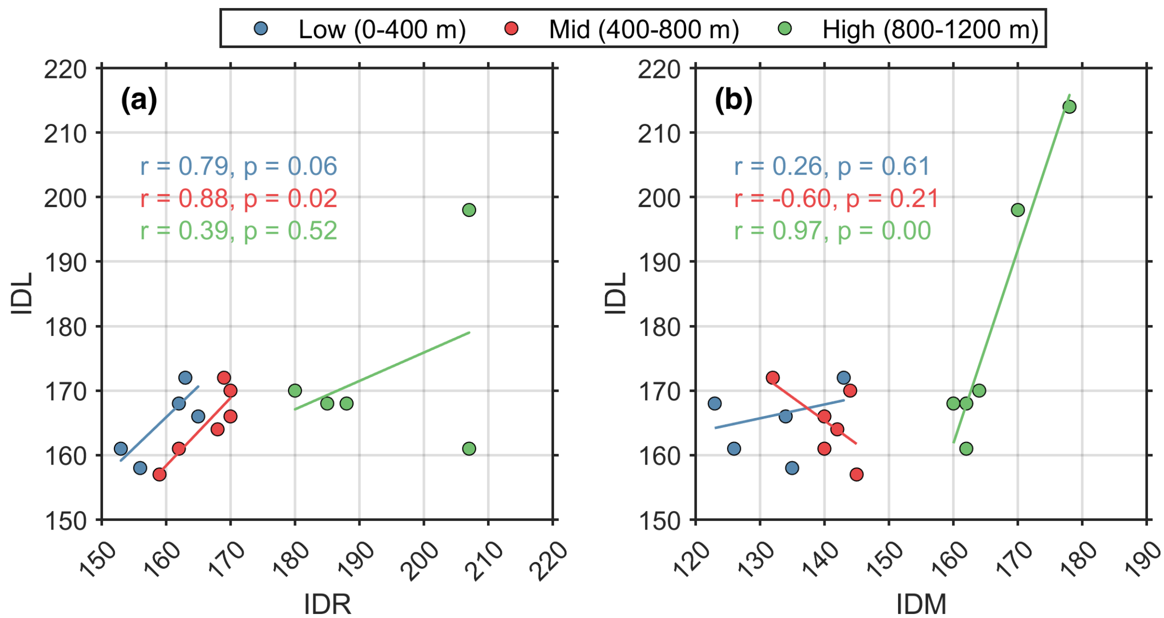

Figure 12. Correlation between the initial date of lake formation (IDL) and those of (a) runoff (IDR) and (b) surface melt (IDM). Each marker shows the observation in the period 2016 to 2021. The marker colors indicate the elevation bands. Data from 2018 is excluded, since runoff observations did not satisfy the conditions used to define the start day as described in Section 3.5.

Seasonal variations in lake area showed significantly different patterns in the three elevation ranges. Lakes in the low-elevation area formed early in the melt season, but the lake area remained small (Figs. 8a–f). This is because most of the lakes formed in crevasses and drained frequently. In the mid-elevation area, the lake area increased rapidly immediately after the onset of the formation. It subsequently dropped in July and August when air temperature, melt and runoff rates were still high. For example, more than 50% of the lake area was lost within a week or less in the years 2019 and 2020 (Figs. 8j and k). These rapid changes were induced by one or more drainage events which were commonly observed in the Greenland ice sheet (e.g. Selmes et al., 2011; Williamson et al., 2018b). In the high-elevation area, the lake area increased under the condition of relatively low melt and runoff rates (Figs. 8m–r). The large lake area typically lasted until the end of the melt season, suggesting that only a few lakes drain into the subglacial drainage system.

Rapid lake drainage occurs when surface water penetrates the ice, establishing a surface-to-bed connection via hydraulic fractures (Das et al., 2008; Stevens et al., 2015; Chudley et al., 2019). During our observations, the most pronounced rapid lake drainage occurred at 474 and 436 m a.s.l. on Heiprin Glacier on 29–30 June 2019 and 2–4 July 2020 (Fig. S8). The supraglacial lakes lost nearly all of their area in one or two days, as seen from satellite images, but the actual drainage time may have been shorter as observed previously (Bartholomew et al., 2012; Chudley et al., 2019). Several lakes located within a distance of 10 km simultaneously drained during the events (Fig. S8), suggesting that drainage of a lake triggered cascaded drainage events through short-term stress/strain perturbations transmitted over a distance (Christoffersen et al., 2018; Fitzpatrick et al., 2014). Alternatively, the simultaneous lake drainage events might have been triggered by seasonal speed up of the glacier. Flow acceleration near the glacier front enhances a stretching flow regime, which facilitates crevasse formation and lake drainage. In either case, complex interactions are expected between lake drainage and glacier dynamics (e.g., Chudley et al., 2019; Fitzpatrick et al., 2014; Poinar and Andrews, 2021). To understand the triggering mechanism and to forecast the timing of the lake drainage, further investigation is required on the lakes (depth, volume, morphology), glacier dynamics (ice velocity, strain), and meteorological conditions (Williamson et al., 2018b). Furthermore, meltwater input to the glacier bed was suggested as the trigger of the summer speedup observed on Tracy and Heilprin Glaciers (Sakakibara and Sugiyama, 2019). Thus, further investigation on the location and timing of the lake drainage sheds light on the mechanism of seasonal glacier speed variations in Greenland.

The Landsat-8 and Sentinel-2 data sets enabled us to investigate the seasonal evolution of supraglacial lakes with a sub-weekly resolution for most of the study period. However, cloud cover resulted in data gaps longer than ten days (e.g., days 209–227 in 2017, days 214–225 in 2018, days 187–204 in 2021, Fig. 7). Supraglacial lakes might have expanded or even drained after a precipitation event during the cloud cover (Andrews et al., 2018). Furthermore, optical sensors are not able to capture lake conditions during the polar night when water storage and lake drainage are still possible (Benedek and Willis, 2021). The use of Sentinel-1 SAR imagery may fill the gaps in optical imagery and help us more continuous monitoring of supraglacial lake evolutions.

5.3 Annual evolution of supraglacial lakes

Annual variations in the maximum lake area were consistent with those observed above 800 m a.s.l. (Figs. 9–10). This is because the lake area was predominantly controlled by large lakes formed in the high-elevation area. The development of supraglacial lakes at higher elevation is largely dependent on year-to-year variations in climatic conditions. For example, lake areas above 800 m a.s.l were notably smaller in 2017 and 2018 (Fig. 9). In these two years, the PDDs during June–August were 440 and 395 °C day, which were substantially lower than the mean over the other years (548 °C day, Fig. 10). Conversely, PDD exceeded 700 °C day in 2019, the year with the greatest lake extent above 800 m a.s.l. The annual lake development above 800 m a.s.l. of the two glaciers showed significant correlations with the summer PDD (Fig. 13). In contrast to the large variations in the upper reaches, lake extent below 800 m a.s.l. showed only small annual variations (< 25% in Heilprin and < 50% in Tracy Glaciers, Fig. 10). Meanwhile, no significant correlation was obtained between the lake area and PDD in the regions (Fig. 13).

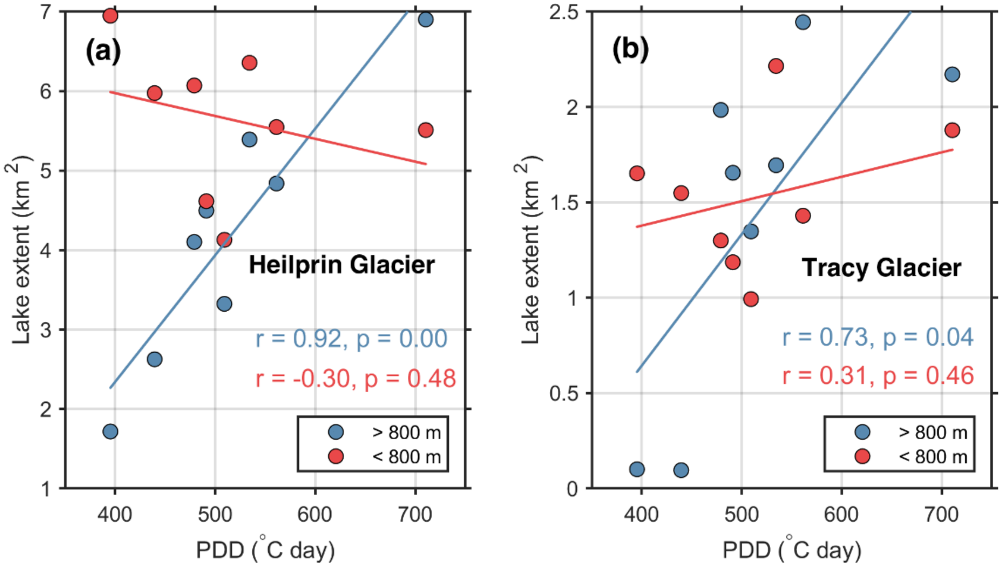

Figure 13. Correlation between summer PDD and lake extent above (blue) and below (red) 800 m a.s.l. in (a) Helprin Glacier and (b)Tracy Glacier.

This result is consistent with previous studies of northern Greenland, which reported the formation of lakes in low-elevation areas, even in relatively cold years (Macdonald et al., 2018; Turton et al., 2021). The clear contrast above and below 800 m a.s.l. is attributed to glacier surface conditions. Based on the SMB computed by RACMO, ELA of the studied glaciers in 2018 and 2019 were 600–800 and 1000–1200 m a.s.l., respectively (Fig. 14). The elevations are similar to the upper limit of surface water availability in northwestern Greenland in 2018 (897 m) and 2019 (1245 m) (Zhang et al., 2023). The ELAs in 2018 and 2019 are also consistent with the upper limits of supraglacial lakes observed in other years (2016–2017 and 2020–2021, Figs. S7), implying that lake formation was facilitated by bare ice exposed on the glacier surface below ELA. Turton et al. (2021) reported a similar observation in northeastern Greenland, where the development of supraglacial lakes was controlled by SMB and the upper limit of lake formation coincided approximately with ELA. Total glacier area over the entire Greenland ice sheet was significantly smaller in 2018, the year characterized by colder summer than in 2019 (Zhang et al., 2023) and 2016–2017 (Hu et al., 2021). The elevation limit of surface meltwater availability was about 140 m lower in 2018 (1407 m a.s.l. in 2018, 1545 m a.s.l. in 2019) (Zhang et al., 2023).

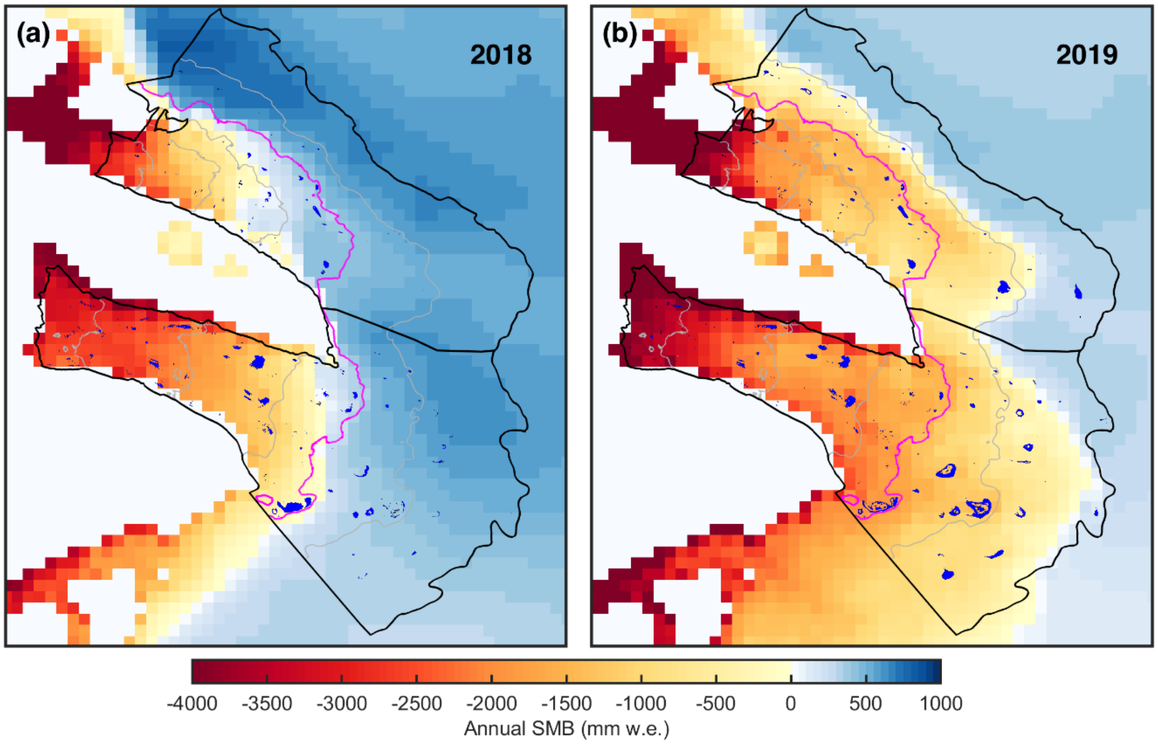

Figure 14. The annual surface mass balance in the study area computed by RACMO (Noël et al., 2019) for the periods (a) from September 2017 to August 2018, and (b) from September 2018 to August 2019. The blue patches represent the maximum lake extent each year. The black lines are glacier boundaries and the gray lines are elevation contours at every 200 m, with 800 m a.s.l. highlighted in magenta.

6. Conclusion

In this study, we applied a machine learning method for mapping supraglacial lakes using satellite imagery. Two medium-resolution optical satellite datasets (Sentinel-2 and Landsat 8) were processed within GEE to generate a high spatial and temporal resolution record of lake area on Heilprin and Tracy Glaciers in northwestern Greenland over the period 2014–2021. We trained RF classifiers by using 14 and 16 predictors selected from spectral bands and indices of Landsat 8 and Sentinel-2 imageries, respectively. The glacier area was classified into two classes (water and non-water) with an overall accuracy greater than 98% for both satellite datasets. Because a good agreement was obtained in the lake areas between the two datasets ($r^2 = 0.996$, $p = 0.000$, RMSE $= 0.010$ km^2), the datasets were merged into a series that covers the two glaciers over eight melt seasons.

The maximum extent of the lake in Heilprin Glacier (22.84 km^2) was three times greater than that of Tracy Glacier (7.60 km^2). The greater lake area in Heilprin Glacier can be attributed to slower ice flow and more frequent surface depression on a relatively flat surface. The formation of supraglacial lakes in regions below 800 m a.s.l. was triggered by the initiation of runoff, while it coincided with the onset of surface melt in higher elevations due to the presence of ice/snow-buried lakes. At low elevations (0–400 m a.s.l.), relatively small supraglacial lakes formed at random locations with little influence of annual climatic variations. These lakes are usually drained by crevasse opening before the middle of the melting season. At mid-elevations (400–800 m a.s.l.), larger supraglacial lakes developed annually at the same locations within surface depressions. Cascading lake drainage events sometimes occurred in this area. At high elevations (800–1200 m a.s.l.), the mean individual lake areas were the largest. The lake development was correlated with the summer PDD, and the upper limit of the lake formation coincided approximately with ELA.

Most of the lakes in this area remained without drainage and frozen lakes were buried by snow at the end of the melting season.

Our study demonstrated that our RF-based machine learning method is an effective method for mapping supraglacial lakes on medium-resolution optical satellite datasets. This method has the potential to monitor surface water features over a greater spatial scale. The study results suggest an inland expansion of supraglacial lakes under a warming climate in Greenland and highlight the possible interaction between draining lake water and glacier dynamics.

Data availability. Landsat 8 imagery is freely provided by USGS (United States Geological Survey). Sentinel-2 imagery is freely provided by ESA (European Space Agency). ArcticDEM is freely provided by the University of Minnesota Polar Geospatial Center. The data mentioned above in this research is directed accessed through the Data Catalog of Google Earth Engine (https://developers.google.com/earth-engine/datasets/). Air temperature data at Qaanaaq station is acquired through NOAA's National Centers for Environment Information (https://www.ncei.noaa.gov/). ITS_LIVE ice speed data is freely provided by NASA and available at https://its-live.jpl.nasa.gov/. The RACMO2.3p2 product was provided by Dr. Brice Noël.

Code availability. The Google Earth Engine script used for random forest supervised classification is available at https://code.earthengine.google.com/1a1cffbb31f58c8c71cc5e7014f344ab.

Acknowledgements. We thank Dr. Brice Noël for providing the RACMO2.3p2 product. Thanks are also to the Scientific Editor Menghua Wang for handling the paper and two anonymous referees for constructive review comments. This research was supported by Arctic Challenge for Sustainability research project II (ArCS II: JPMXD1420318865), ArCS II International Research Exchange Program, and Japanese Government Scholarship, funded by the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT).

References

Andrews, L.C., Hoffman, M.J., Neumann, T.A., Catania, G.A., Lüthi, M.P., Hawley, R.L., Schild, K.M., Ryser, C., & Morriss, B.F. (2018). Seasonal evolution of the subglacial hydrologic system modified by supraglacial lake drainage in western Greenland. *Journal of Geophysical Research: Earth Surface,* https://doi.org/10.1029/2017JF004585

Arthur, J.F., Stokes, C.R., Jamieson, S.S.R., Carr, J.R., & Leeson, A.A. (2020). Distribution and seasonal evolution of supraglacial lakes on Shackleton Ice Shelf, East Antarctica. *The Cryosphere, 14*, 4103-4120, https://doi.org/10.5194/tc-14-4103-2020

Arthur, J.F., Stokes, C.R., Jamieson, S.S.R., Rachel Carr, J., Leeson, A.A., & Verjans, V. (2022). Large interannual variability in supraglacial lakes around East Antarctica. *Nature Communications, 13*, 1711, https://doi.org/10.1038/s41467-022-29385-3

Bartholomew, I., Nienow, P., Sole, A., Mair, D., Cowton, T., & King, M.A. (2012). Short-term variability in Greenland Ice Sheet motion forced by time-varying meltwater drainage: Implications for the relationship between subglacial drainage system behavior and ice velocity. *Journal of Geophysical Research: Earth Surface, 117*, https://doi.org/10.1029/2011JF002220

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing, 114*, 24-31, https://doi.org/10.1016/j.isprsjprs.2016.01.011

Benedek, C.L., & Willis, I.C. (2021). Winter drainage of surface lakes on the Greenland Ice Sheet from Sentinel-1 SAR imagery. *The Cryosphere,* 15, 1587-1606, https://doi.org/10.5194/tc-15-1587-2021

Benn, D.I., Warren, C.R., & Mottram, R.H. (2007). Calving processes and the dynamics of calving glaciers. *Earth-Science Reviews, 82*, 143-179, https://doi.org/10.1016/j.earscirev.2007.02.002

Box, J.E., & Ski, K. (2007). Remote sounding of Greenland supraglacial melt lakes: implications for subglacial hydraulics. *Journal of Glaciology,* 53, 257-265, https://doi.org/10.3189/172756507782202883

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

Chan, J.C.-W., & Paelinckx, D. (2008). Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sensing of Environment, 112*, 2999-3011, https://doi.org/10.1016/j.rse.2008.02.011

Christoffersen, P., Bougamont, M., Hubbard, A., Doyle, S.H., Grigsby, S., & Pettersson, R. (2018). Cascading lake drainage on the Greenland Ice Sheet triggered by tensile shock and fracture. *Nature Communications, 9*, 1064, https://doi.org/10.1038/s41467-018-03420-8

Chu, V.W. (2014). Greenland ice sheet hydrology:A review. *Progress in Physical Geography: Earth and Environment, 38*, 19-54, https://doi.org/10.1177/0309133313507075

Chudley, T.R., Christoffersen, P., Doyle, S.H., Bougamont, M., Schoonman, C.M., Hubbard, B., & James, M.R. (2019). Supraglacial lake drainage at a fast-flowing Greenlandic outlet glacier. *Proceedings of the National Academy of Sciences, 116*, 25468-25477, https://doi.org/10.1073/pnas.1913685116

Colgan, W., Rajaram, H., Abdalati, W., McCutchan, C., Mottram, R., Moussavi, M.S., & Grigsby, S. (2016). Glacier crevasses: Observations, models, and mass balance implications. *Reviews of Geophysics, 54*, 119-161, https://doi.org/10.1002/2015RG000504

Corr, D., Leeson, A., McMillan, M., Zhang, C., & Barnes, T. (2022). An inventory of supraglacial lakes and channels across the West Antarctic Ice Sheet. *Earth Syst. Sci. Data,* 14, 209-228, https://doi.org/10.5194/essd-14-209-2022

Crist, E.P., & Cicone, R.C. (1984). A Physically-Based Transformation of Thematic Mapper Data--The TM Tasseled Cap. *IEEE Transactions on Geoscience and Remote Sensing, GE-22*, 256-263, https://doi.org/10.1109/TGRS.1984.350619

Das, S.B., Joughin, I., Behn, M.D., Howat, I.M., King, M.A., Lizarralde, D., & Bhatia, M.P. (2008). Fracture Propagation to the Base of the Greenland Ice Sheet During Supraglacial Lake Drainage. *Science, 320*, 778-781, https://doi.org/10.1126/science.1153360

Dawes, P.R., & van As, D. (2010). An advancing glacier in a recessive ice regime: Berlingske Bræ, North-West Greenland. *Geol. Surv. Den. Greenl. Bull, 20*, 79-82, https://doi.org/10.34194/geusb.v20.4986

Dell, R., Arnold, N., Willis, I., Banwell, A., Williamson, A., Pritchard, H., & Orr, A. (2020). Lateral meltwater transfer across an Antarctic ice shelf. *The Cryosphere, 14*, 2313-2330, https://doi.org/10.5194/tc-14-2313-2020

Dell, R.L., Banwell, A.F., Willis, I.C., Arnold, N.S., Halberstadt, A.R.W., Chudley, T.R., & Pritchard, H.D. (2021). Supervised classification of slush and ponded water on Antarctic ice shelves using Landsat 8 imagery. *Journal of Glaciology*, 1-14, https://doi.org/10.1017/jog.2021.114

Denisko, D., & Hoffman, M.M. (2018). Classification and interaction in random forests. *Proceedings of the National Academy of Sciences, 115*, 1690-1692, https://doi.org/10.1073/pnas.1800256115

Dirscherl, M., Dietz, A.J., Kneisel, C., & Kuenzer, C. (2020). Automated Mapping of Antarctic Supraglacial Lakes Using a Machine Learning Approach. *Remote Sensing, 12*, 1203, https://doi.org/10.3390/rs12071203

Dirscherl, M., Dietz, A.J., Kneisel, C., & Kuenzer, C. (2021a). A Novel Method for Automated Supraglacial Lake Mapping in Antarctica Using Sentinel-1 SAR Imagery and Deep Learning. *Remote Sensing,* 13, 197, https://doi.org/10.3390/rs13020197

Dirscherl, M.C., Dietz, A.J., & Kuenzer, C. (2021b). Seasonal evolution of Antarctic supraglacial lakes in 2015–2021 and links to environmental controls. *The Cryosphere, 15*, 5205-5226, https://doi.org/10.5194/tc-15-5205-2021

Doyle, S.H., Hubbard, A.L., Dow, C.F., Jones, G.A., Fitzpatrick, A., Gusmeroli, A., Kulessa, B., Lindback, K., Pettersson, R., & Box, J.E. (2013). Ice tectonic deformation during the rapid in situ drainage of a supraglacial lake on the Greenland Ice Sheet. *The Cryosphere, 7*, 129-140, https://doi.org/10.5194/tc-7-129-2013

Dozier, J. (1989). Spectral signature of alpine snow cover from the landsat thematic mapper. *Remote Sensing of Environment, 28*, 9-22, https://doi.org/10.1016/0034-4257(89)90101-6

Dunmire, D., Banwell, A.F., Wever, N., Lenaerts, J.T.M., & Datta, R.T. (2021). Contrasting regional variability of buried meltwater extent over 2 years across the Greenland Ice Sheet. *The Cryosphere, 15*, 2983-3005, https://doi.org/10.5194/tc-15-2983-2021

Echelmeyer, K., Clarke, T.S., & Harrison, W.D. (1991). Surficial glaciology of Jakobshavns Isbræ, West Greenland: Part I. Surface morphology. *Journal of Glaciology, 37*, 368-382, https://doi.org/10.3189/S0022143000005803

Feng, D. (2012). A New Method for Fast Information Extraction of Water Bodies Using Remotely Sensed Data. *Remote Sensing Technology and Application, 24*, 167-171,

Feyisa, G.L., Meilby, H., Fensholt, R., & Proud, S.R. (2014). Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sensing of Environment, 140*, 23-35, https://doi.org/10.1016/j.rse.2013.08.029

Fitzpatrick, A.A.W., Hubbard, A.L., Box, J.E., Quincey, D.J., van As, D., Mikkelsen, A.P.B., Doyle, S.H., Dow, C.F., Hasholt, B., & Jones, G.A. (2014). A decade (2002–2012) of supraglacial lake volume estimates across Russell Glacier, West Greenland. *The Cryosphere, 8*, 107-121, https://doi.org/10.5194/tc-8-107-2014

Flora, J., Johansen, K.L., Grønnow, B., Andersen, A.O., & Mosbech, A. (2018). Present and past dynamics of Inughuit resource spaces. *Ambio,* 47, 244-264, https://doi.org/10.1007/s13280-018-1039-6

[dataset] Gardner, A. S., M. A. Fahnestock, and T. A. Scambos, 2019 [update to time of data download]: ITS_LIVE Regional Glacier and Ice Sheet Surface Velocities: Version 1. Data archived at National Snow and Ice Data Center; https://doi:10.5067/6II6VW8LLWJ7.

Georgiou, S., Shepherd, A., McMillan, M., & Nienow, P. (2009). Seasonal evolution of supraglacial lake volume from ASTER imagery. *Annals of Glaciology,* 50, 95-100, https://doi.org/10.3189/172756409789624328

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment, 202*, 18-27, https://doi.org/10.1016/j.rse.2017.06.031

Halberstadt, A.R.W., Gleason, C.J., Moussavi, M.S., Pope, A., Trusel, L.D., & DeConto, R.M. (2020). Antarctic Supraglacial Lake Identification Using Landsat-8 Image Classification. *Remote Sensing, 12*, 1327, https://doi.org/10.3390/rs12081327

Hall, D.K., Riggs, G.A., & Salomonson, V.V. (1995). Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. *Remote Sensing of Environment, 54*, 127-140, https://doi.org/10.1016/0034-4257(95)00137-P

Harper, J., Humphrey, N., Pfeffer, W.T., Brown, J., & Fettweis, X. (2012). Greenland ice-sheet contribution to sea-level rise buffered by meltwater storage in firn. *Nature, 491*, 240-243, https://doi.org/10.1038/nature11566

Howat, I., De la Pena, S., Van Angelen, J., Lenaerts, J., & Van den Broeke, M.J.T.C. (2013). Brief Communication" Expansion of meltwater lakes on the Greenland ice sheet". *The Cryosphere, 7*, 201-204, https://doi.org/10.5194/tc-7-201-2013

Hu, J., Huang, H., Chi, Z., Cheng, X., Wei, Z., Chen, P., Xu, X., Qi, S., Xu, Y., & Zheng, Y. (2022). Distribution and Evolution of Supraglacial Lakes in Greenland during the 2016–2018 Melt Seasons. *Remote Sensing, 14*, 55, https://doi.org/10.3390/rs14010055

Huete, A.R. (1988). A soil-adjusted vegetation index (SAVI). *Remote Sensing of Environment, 25*, 295-309, https://doi.org/10.1016/0034-4257(88)90106-X

Ignéczi, Á., Sole, A.J., Livingstone, S.J., Leeson, A.A., Fettweis, X., Selmes, N., Gourmelen, N., & Briggs, K. (2016). Northeast sector of the Greenland Ice Sheet to undergo the greatest inland expansion of supraglacial lakes during the 21st century. *Geophysical Research Letters, 43*, 9729-9738, https://doi.org/10.1002/2016GL070338

Jiang, D., Li, X., Zhang, K., Marinsek, S., Hong, W., & Wu, Y. (2022). Automatic Supraglacial Lake Extraction in Greenland Using Sentinel-1 SAR Images and Attention-Based U-Net. *Remote Sensing,* 14, 4998, https://doi.org/10.3390/rs14194998

Johansson, A.M., Jansson, P., & Brown, I.A. (2013). Spatial and temporal variations in lakes on the Greenland Ice Sheeta. *Journal of hydrology, 476*, 314-320, https://doi.org/10.1016/j.jhydrol.2012.10.045

Kanna, N., Sugiyama, S., Ando, T., Wang, Y., Sakuragi, Y., Hazumi, T., Matsuno, K., Yamaguchi, A., Nishioka, J., & Yamashita, Y. (2022). Meltwater Discharge From Marine-Terminating

Glaciers Drives Biogeochemical Conditions in a Greenlandic Fjord. *Global Biogeochemical Cycles,* 36, e2022GB007411, https://doi.org/10.1029/2022GB007411

Kanna, N., Sugiyama, S., Fukamachi, Y., Nomura, D., & Nishioka, J. (2020). Iron Supply by Subglacial Discharge Into a Fjord Near the Front of a Marine-Terminating Glacier in Northwestern Greenland. *Global Biogeochemical Cycles,* 34, e2020GB006567, https://doi.org/10.1029/2020GB006567

Koenig, L.S., Lampkin, D.J., Montgomery, L.N., Hamilton, S.L., Turrin, J.B., Joseph, C.A., Moutsafa, S.E., Panzer, B., Casey, K.A., Paden, J.D., Leuschen, C., & Gogineni, P. (2015). Wintertime storage of water in buried supraglacial lakes across the Greenland Ice Sheet. *The Cryosphere, 9*, 1333-1342, https://doi.org/10.5194/tc-9-1333-2015

Lampkin, D.J., & VanderBerg, J. (2011). A preliminary investigation of the influence of basal and surface topography on supraglacial lake distribution near Jakobshavn Isbrae, western Greenland. *Hydrological Processes, 25*, 3347-3355, https://doi.org/10.1002/hyp.8170

Landis, J.R., & Koch, G.G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics, 33*, 159-174, https://doi.org/10.2307/2529310

Leeson, A.A., Shepherd, A., Briggs, K., Howat, I., Fettweis, X., Morlighem, M., & Rignot, E. (2015). Supraglacial lakes on the Greenland ice sheet advance inland under warming climate. *Nature Climate Change, 5*, 51-55, https://doi.org/10.1038/nclimate2463

Li, H., Xu, L., Shen, H., & Zhang, L. (2016). A general variational framework considering cast shadows for the topographic correction of remote sensing imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 117*, 161-171, https://doi.org/10.1016/j.isprsjprs.2016.03.021

Li, W., Lhermitte, S., & López-Dekker, P. (2021). The potential of synthetic aperture radar interferometry for assessing meltwater lake dynamics on Antarctic ice shelves. *The Cryosphere,* 15, 5309-5322, https://doi.org/10.5194/tc-15-5309-2021

Lutz, K., Bahrami, Z., & Braun, M. (2023). Supraglacial Lake Evolution over Northeast Greenland Using Deep Learning Methods. *Remote Sensing,* 15, 4360, https://doi.org/10.3390/rs15174360

Lüthje, M., Pedersen, L., Reeh, N., & Greuell, W. (2006). Modelling the evolution of supraglacial lakes on the West Greenland ice-sheet margin. *Journal of Glaciology, 52*, 608-618, https://doi.org/10.3189/172756506781828386

Macdonald, G.J., Banwell, A.F., & MacAYEAL, D.R. (2018). Seasonal evolution of supraglacial lakes on a floating ice tongue, Petermann Glacier, Greenland. *Annals of Glaciology*, 1-10, https://doi.org/10.1017/aog.2018.9

Mankoff, K.D., Noël, B., Fettweis, X., Ahlstrøm, A.P., Colgan, W., Kondo, K., Langley, K., Sugiyama, S., van As, D., & Fausto, R.S. (2020a). Greenland liquid water discharge from 1958 through 2019. *Earth Syst. Sci. Data,* 12, 2811-2841, https://doi.org/10.5194/essd-12-2811-2020

Mankoff, K.D., Solgaard, A., Colgan, W., Ahlstrøm, A.P., Khan, S.A., & Fausto, R.S. (2020b). Greenland Ice Sheet solid ice discharge from 1986 through March 2020. *Earth Syst. Sci. Data,* 12, 1367-1383, https://doi.org/10.5194/essd-12-1367-2020

Matsuno, K., Kanna, N., Sugiyama, S., Yamaguchi, A., & Yang, E.J. (2020). Impacts of meltwater discharge from marine-terminating glaciers on the protist community in Inglefield Bredning, northwestern Greenland. *Marine Ecology Progress Series,* 642, 55-65, https://doi.org/10.3354/meps13324

McMillan, M., Nienow, P., Shepherd, A., Benham, T., & Sole, A. (2007). Seasonal evolution of supra-glacial lakes on the Greenland Ice Sheet. *Earth and Planetary Science Letters, 262*, 484-492, https://doi.org/10.1016/j.epsl.2007.08.002

Miles, K.E., Willis, I.C., Benedek, C.L., Williamson, A.G., & Tedesco, M. (2017). Toward Monitoring Surface and Subsurface Lakes on the Greenland Ice Sheet Using Sentinel-1 SAR and Landsat-8 OLI Imagery. *Frontiers in Earth Science,* 5, https://doi.org/10.3389/feart.2017.00058

[dataset] Mouginot, J., & Rignot, E. (2019). Glacier catchments/basins for the Greenland Ice Sheet. *UC Irvine Dash: Irvine, CA, USA*, https://doi.org/10.7280/D1WT11

Moussavi, M., Pope, A., Halberstadt, A.R.W., Trusel, L.D., Cioffi, L., & Abdalati, W. (2020). Antarctic Supraglacial Lake Detection Using Landsat 8 and Sentinel-2 Imagery: Towards Continental Generation of Lake Volumes. *Remote Sensing, 12*, 134, https://doi.org/10.3390/rs12010134

Niu, L., Tang, X., Yang, S., Zhang, Y., Zheng, L., & Wang, L. (2023). Detection of Antarctic Surface Meltwater Using Sentinel-2 Remote Sensing Images via U-Net With Attention Blocks: A Case Study Over the Amery Ice Shelf. *IEEE Transactions on Geoscience and Remote Sensing,* 61, 1-13, https://doi.org/10.1109/TGRS.2023.3275076

Noël, B., Berg, W.J.v.d., Lhermitte, S., & Broeke, M.R.v.d. (2019). Rapid ablation zone expansion amplifies north Greenland mass loss. *Science Advances, 5*, eaaw0123, https://doi.org/10.1126/sciadv.aaw0123

Noël, B., van de Berg, W.J., Machguth, H., Lhermitte, S., Howat, I., Fettweis, X., & van den Broeke, M.R. (2016). A daily, 1 km resolution data set of downscaled Greenland ice sheet surface mass balance (1958–2015). *The Cryosphere, 10*, 2361-2377, https://doi.org/10.5194/tc-10-2361-2016

Otto, J., Holmes, F.A., & Kirchner, N. (2022). Supraglacial lake expansion, intensified lake drainage frequency, and first observation of coupled lake drainage, during 1985–2020 at Ryder Glacier, Northern Greenland. *Frontiers in Earth Science, 10*, https://doi.org/10.3389/feart.2022.978137

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing, 26*, 217-222, https://doi.org/10.1080/01431160412331269698

Poinar, K., & Andrews, L.C. (2021). Challenges in predicting Greenland supraglacial lake drainages at the regional scale. *The Cryosphere, 15*, 1455-1483, https://doi.org/10.5194/tc-15-1455-2021

Pope, A., Scambos, T.A., Moussavi, M., Tedesco, M., Willis, M., Shean, D., & Grigsby, S. (2016). Estimating supraglacial lake depth in West Greenland using Landsat 8 and comparison with other multispectral methods. *The Cryosphere, 10*, 15-27, https://doi.org/10.5194/tc-10-15-2016

Porter, D.F., Tinto, K.J., Boghosian, A., Cochran, J.R., Bell, R.E., Manizade, S.S., & Sonntag, J.G. (2014). Bathymetric control of tidewater glacier mass loss in northwest Greenland. *Earth and Planetary Science Letters, 401*, 40-46, https://doi.org/10.1016/j.epsl.2014.05.058

[dataset] Porter, Claire, et al., 2018, "ArcticDEM, Version 3", , Harvard Dataverse, V1, https://doi.org/10.7910/DVN/OHHUKH

Qayyum, N., Ghuffar, S., Ahmad, H.M., Yousaf, A., & Shahid, I. (2020). Glacial Lakes Mapping Using Multi Satellite PlanetScope Imagery and Deep Learning. *ISPRS International Journal of Geo-Information,* 9, 560, https://doi.org/10.3390/ijgi9100560

Rawlins, L.D., Rippin, D.M., Sole, A.J., Livingstone, S.J., & Yang, K. (2023). Seasonal evolution of the supraglacial drainage network at Humboldt Glacier, North Greenland, between 2016 and 2020. *The Cryosphere Discuss.,* 2023, 1-32, https://doi.org/10.5194/tc-2023-23

Ryan, J.C., Smith, L.C., van As, D., Cooley, S.W., Cooper, M.G., Pitcher, L.H., & Hubbard, A. (2019). Greenland Ice Sheet surface melt amplified by snowline migration and bare ice exposure. *Science Advances,* 5, eaav3738, https://doi.org/10.1126/sciadv.aav3738

Sakakibara, D., & Sugiyama, S. (2018). Ice front and flow speed variations of marine-terminating outlet glaciers along the coast of Prudhoe Land, northwestern Greenland. *Journal of Glaciology, 64*, 300-310, https://doi.org/10.1017/jog.2018.20

Sakakibara, D., & Sugiyama, S. (2020). Seasonal ice-speed variations in 10 marine-terminating outlet glaciers along the coast of Prudhoe Land, northwestern Greenland. *Journal of Glaciology, 66*, 25-34, https://doi.org/10.1017/jog.2019.81

Schwanghart, W., & Scherler, D. (2014). Short Communication: TopoToolbox 2 – MATLAB-based software for topographic analysis and modeling in Earth surface sciences. *Earth Surf. Dynam., 2*, 1-7, https://doi.org/10.5194/esurf-2-1-2014

Selmes, N., Murray, T., & James, T. (2011). Fast draining lakes on the Greenland Ice Sheet. *Geophysical Research Letters, 38*, https://doi.org/10.1029/2011GL047872

Sergienko, O.V. (2013). Glaciological twins: basally controlled subglacial and supraglacial lakes. *Journal of Glaciology, 59*, 3-8, https://doi.org/10.3189/2013JoG12J040

Shu, Q., Killick, R., Leeson, A., Nemeth, C., Fettweis, X., Hogg, A., & Leslie, D. (2023). Characterising the ice sheet surface in Northeast Greenland using Sentinel-1 SAR data. *Journal of Glaciology,* 1-12, https://doi.org/10.1017/jog.2023.64

Slater, T., Shepherd, A., McMillan, M., Leeson, A., Gilbert, L., Muir, A., Munneke, P.K., Noël, B., Fettweis, X., van den Broeke, M., & Briggs, K. (2021). Increased variability in Greenland Ice Sheet runoff from satellite observations. *Nature Communications, 12*, 6069, https://doi.org/10.1038/s41467-021-26229-4

Sneed, W.A., & Hamilton, G.S. (2007). Evolution of melt pond volume on the surface of the Greenland Ice Sheet. *Geophysical Research Letters,* 34, https://doi.org/10.1029/2006GL028697

Spergel, J.J., Kingslake, J., Creyts, T., van Wessem, M., & Fricker, H.A. (2021). Surface meltwater drainage and ponding on Amery Ice Shelf, East Antarctica, 1973–2019. *Journal of Glaciology, 67*, 985-998, https://doi.org/10.1017/jog.2021.46

Stehman, S.V. (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment, 62*, 77-89, https://doi.org/10.1016/S0034-4257(97)00083-7

Stevens, L.A., Behn, M.D., McGuire, J.J., Das, S.B., Joughin, I., Herring, T., Shean, D.E., & King, M.A. (2015). Greenland supraglacial lake drainages triggered by hydrologically induced basal slip. *Nature, 522*, 73, https://doi.org/10.1038/nature14480

Straneo, F., Slater, D.A., Bouchard, C., Cape, M.R., Carey, M., Ciannelli, L., Holte, J., Matrai, P., Laidre, K., & Little, C.J.O. (2022). An interdisciplinary perspective on Greenland's changing coastal margins. *Oceanography,* 35, 106-117, https://doi.org/10.5670/oceanog.2022.128

Sugiyama, S., Sakakibara, D., Matsuno, S., Yamaguchi, S., Matoba, S., & Aoki, T. (2014). Initial field observations on Qaanaaq ice cap, northwestern Greenland. *Annals of Glaciology, 55*, 25-33, https://dor.org/10.3189/2014AoG66A102

Sugiyama, S., Kanna, N., Sakakibara, D., Ando, T., Asaji, I., Kondo, K., Wang, Y., Fujishi, Y., Fukumoto, S., Podolskiy, E., Fukamachi, Y., Takahashi, M., Matoba, S., Iizuka, Y., Greve, R., Furuya, M., Tateyama, K., Watanabe, T., Yamasaki, S., Yamaguchi, A., Nishizawa, B., Matsuno, K., Nomura, D., Sakuragi, Y., Matsumura, Y., Ohashi, Y., Aoki, T., Niwano, M., Hayashi, N., Minowa, M., Jouvet, G., van Dongen, E., Bauder, A., Funk, M., Bjørk, A.A., & Oshima, T. (2021). Rapidly changing glaciers, ocean and coastal environments, and their impact on human society in the Qaanaaq region, northwestern Greenland. *Polar Science,* 27, 100632, https://doi.org/10.1016/j.polar.2020.100632

Sundal, A.V., Shepherd, A., Nienow, P., Hanna, E., Palmer, S., & Huybrechts, P. (2009). Evolution of supra-glacial lakes across the Greenland Ice Sheet. *Remote Sensing of Environment, 113*, 2164-2171, https://doi.org/10.1016/j.rse.2009.05.018

Tedesco, M., Lüthje, M., Steffen, K., Steiner, N., Fettweis, X., Willis, I., Bayou, N., & Banwell, A. (2012). Measurement and modeling of ablation of the bottom of supraglacial lakes in western Greenland. *Geophysical Research Letters, 39*, https://doi.org/10.1029/2011GL049882

Tedstone, A.J., & Machguth, H. (2022). Increasing surface runoff from Greenland's firn areas. *Nature Climate Change, 12*, 672-676, https://doi.org/10.1038/s41558-022-01371-z

Teluguntla, P., Thenkabail, P.S., Oliphant, A., Xiong, J., Gumma, M.K., Congalton, R.G., Yadav, K., & Huete, A. (2018). A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing, 144*, 325-340, https://doi.org/10.1016/j.isprsjprs.2018.07.017

Tsutaki, S., Sugiyama, S., Sakakibara, D., Aoki, T., & Niwano, M. (2017). Surface mass balance, ice velocity and near-surface ice temperature on Qaanaaq Ice Cap, northwestern Greenland, from 2012 to 2016. *Annals of Glaciology, 58*, 181-192, https://doi.org/10.1017/aog.2017.7

Tuckett, P.A., Ely, J.C., Sole, A.J., Lea, J.M., Livingstone, S.J., Jones, J.M., & van Wessem, J.M. (2021). Automated mapping of the seasonal evolution of surface meltwater and its links to climate on the Amery Ice Shelf, Antarctica. *The Cryosphere, 15*, 5785-5804, https://doi.org/10.5194/tc-15-5785-2021

Turton, J.V., Hochreuther, P., Reimann, N., & Blau, M.T. (2021). The distribution and evolution of supraglacial lakes on 79° N Glacier (north-eastern Greenland) and interannual climatic controls. *The Cryosphere, 15*, 3877-3896, https://doi.org/10.5194/tc-15-3877-2021

Wang, Y., Sugiyama, S., & Bjørk, A.A. (2021). Surface Elevation Change of Glaciers Along the Coast of Prudhoe Land, Northwestern Greenland From 1985 to 2018. *Journal of Geophysical Research: Earth Surface, 126*, e2020JF006038, https://doi.org/10.1029/2020JF006038

Williamson, A.G., Arnold, N.S., Banwell, A.F., & Willis, I.C. (2017). A Fully Automated Supraglacial lake area and volume Tracking ("FAST") algorithm: Development and application using MODIS imagery of West Greenland. *Remote Sensing of Environment,* 196, 113-133, https://doi.org/10.1016/j.rse.2017.04.032

Williamson, A.G., Banwell, A.F., Willis, I.C., & Arnold, N.S. (2018a). Dual-satellite (Sentinel-2 and Landsat 8) remote sensing of supraglacial lakes in Greenland. *Cryosphere, 12*, https://doi.org/10.5194/tc-12-3045-2018

Williamson, A.G., Willis, I.C., Arnold, N.S., & Banwell, A.F. (2018b). Controls on rapid supraglacial lake drainage in West Greenland: an Exploratory Data Analysis approach. *Journal of Glaciology, 64*, 208-226, https://doi.org/10.1017/jog.2018.8

Willis, J.K., Carroll, D., Fenty, I., Kohli, G., Khazendar, A., Rutherford, M., Trenholm, N., & Morlighem, M. (2018). Ocean-ice interactions in Inglefield Gulf: Early results from NASA's Oceans Melting Greenland mission. *Oceanography, 31*, 100-108, https://doi.org/10.5670/oceanog.2018.211

Wu, R., Liu, G., Zhang, R., Wang, X., Li, Y., Zhang, B., Cai, J., & Xiang, W. (2020). A Deep Learning Method for Mapping Glacial Lakes from the Combined Use of Synthetic-Aperture Radar and Optical Satellite Images. *Remote Sensing,* 12, 4020, https://doi.org/10.3390/rs12244020

Xu, J., Feng, M., Sui, Y., Yan, D., Zhang, K., & Shi, K. (2023). Identifying Alpine Lakes in the Eastern Himalayas Using Deep Learning. *Water,* 15, 229, https://doi.org/10.3390/w15020229

Yang, K., & Smith, L.C. (2013). Supraglacial Streams on the Greenland Ice Sheet Delineated From Combined Spectral–Shape Information in High-Resolution Satellite Imagery. *IEEE Geoscience and Remote Sensing Letters, 10*, 801-805, https://doi.org/10.1109/LGRS.2012.2224316

Yang, K., Smith, L.C., Cooper, M.G., Pitcher, L.H., van As, D., Lu, Y., Lu, X., & Li, M. (2021). Seasonal evolution of supraglacial lakes and rivers on the southwest Greenland Ice Sheet. *Journal of Glaciology,* 67, 592-602, https://doi.org/10.1017/jog.2021.10

Yuan, J., Chi, Z., Cheng, X., Zhang, T., Li, T., & Chen, Z.J.W. (2020). Automatic extraction of supraglacial lakes in southwest greenland during the 2014–2018 melt seasons based on convolutional neural network. *Water, 12*, 891, https://doi.org/10.3390/w12030891

Zhang, W., Yang, K., Smith, L.C., Wang, Y., van As, D., Noël, B., Lu, Y., & Liu, J. (2023). Pan-Greenland mapping of supraglacial rivers, lakes, and water-filled crevasses in a cool summer (2018) and a warm summer (2019). *Remote Sensing of Environment,* 297, 113781, https://doi.org/10.1016/j.rse.2023.113781