https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

A Framework for Automated Supraglacial Lake Detection and Depth Retrieval in ICESat-2 Photon Data Across the Greenland and Antarctic Ice Sheets

Philipp Sebastian Arndt[1] and Helen Amanda Fricker[1]

[1]Scripps Polar Center, University of California San Diego, 8885 Biological Grade, La Jolla, CA 92037, USA

Correspondence: Philipp Sebastian Arndt (parndt@ucsd.edu)

Abstract. Supraglacial lakes on the ice sheets have been linked to ice shelf collapse in Antarctica and accelerated flow of grounded ice in Greenland. However, it is difficult to quantify the impact of supraglacial lakes on ice dynamics accurately enough to predict their contribution to future mass loss and sea level rise. This is largely because ice-sheet-wide assessments of meltwater volumes rely on models that are poorly constrained due to a lack of accurate depth measurements. Various recent case studies have demonstrated that accurate supraglacial lake depths can be obtained from ICESat-2's ATL03 photon-level data product. ATL03 comprises hundreds of terabytes of unstructured point cloud data, which has made it challenging to use this bathymetric capability at scale. Here, we present two new algorithms – Flat Lake and Underlying Ice Detection (FLUID) and Surface Removal and Robust Fit (SuRFF) – which together provide a fully automated and scalable method for lake detection and depth determination from ATL03 data, and establish a framework for its large-scale implementation using distributed high-throughput computing. We report FLUID/SuRFF algorithm performance over two regions known to have significant surface melt – Central West Greenland and Amery Ice Shelf catchment in East Antarctica – during two melt seasons. FLUID/SuRFF reveals a total of 1249 lakes up to $25\,\mathrm{m}$ deep, with more water during higher melt years. In absence of ground truth data, manual annotation of test data suggests that our method reliably detects melt lakes whenever a bathymetric signal is discernible, and estimates water depths with a mean absolute error of $0.28\,\mathrm{m}$. These results imply that our proposed framework has the potential to generate a comprehensive data product of accurate meltwater depths across both ice sheets.

1 Introduction

Earth is warming and both of its ice sheets (Greenland and Antarctica) are losing mass to the ocean at increasing rates (Rignot et al., 2011; Smith et al., 2020), leading to sea level rise. There is growing evidence that some of this retreat is irreversible, thus committing coastal communities to embrace costly sea level rise adaptation strategies for decades and centuries to come (DeConto et al., 2021; Garbe et al., 2020; Gregory et al., 2020; Nordhaus, 2019). To address the resulting societal challenges, policy makers and coastal planners require accurate and actionable sea level rise projections (Moon et al., 2020). However, the projected contribution from the ice sheets is highly uncertain, to the point that the Sixth Assessment Report of the United Nations Intergovernmental Panel on Climate Change designated it as *"deep uncertainty"* (IPCC AR6; Fox-Kemper, 2021). Building confidence in projections of the ice sheets' contribution to future sea level rise requires better mechanistic understand-

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

ing of relevant mass balance processes for inclusion in ice sheet models (Golledge, 2020; Aschwanden et al., 2021). However, ice sheet wide details of many of these processes are poorly known because they have been under observed in both space and time.

Supraglacial melting on the ice sheets is one example of a process which has a potentially important contribution to future sea level rise projections, yet has been under observed so is poorly understood. In a warming climate, supraglacial lakes have the potential to trigger positive feedback loops and catastrophic collapse (Hill et al., 2024), yet the underlying mechanisms and associated likelihoods are vaguely defined due to a lack of high-quality observations (Arthur et al., 2020). In particular, models that attempt to capture the influence of supraglacial hydrology on ice sheet behavior require accurate estimates of volumes of pooled surface meltwater as input (Zwally et al., 2002; Parizek and Alley, 2004; Krawczynski et al., 2009; Robel and Banwell, 2019). However, there are few direct observations of supraglacial lake depths, which leads to errors in total water volume estimates. This introduces biases into model inputs for meltwater flow, impacting projections of future ice sheet evolution (Melling et al., 2024). To ensure that coupled hydrological-dynamical models accurately represent the underlying physics, it is important to find a method to acquire lake depths that are accurate, and also spatially and temporally and continuous.

Recently, NASA's Ice, Cloud and land Elevation Satellite (ICESat-2) laser altimeter has made it possible to directly obtain accurate water depth measurements due to its green light being able to penetrate water, which allows its sensor to register the elevation of photons that were reflected from both the lake surface and the lakebed (Fig. 1; e.g., Fair et al., 2020; Fricker et al., 2021; Xiao et al., 2023). However, ICESat-2's photon-level data product (ATL03) is limited to discrete, one-dimensional ground tracks that are coarsely spaced on the ground and have a relatively long repeat period of three months. While this means that ICESat-2 data alone cannot be used to monitor melt lake volumes, several case studies have shown that ICESat-2 depth measurements can be used to constrain parameters in models that estimate lake volumes from satellite imagery (Datta and Wouters, 2021; Leeuwen, 2023). This implies that large-scale extraction of supraglacial lake depths from ICESat-2 photon-level data in combination with concurrent optical satellite imagery would make it possible to use statistical learning methods to generate a well-constrained data-driven model for ice-sheet-wide lake volume estimation (Melling et al., 2024). Previous ICESat-2 studies have been limited to applying depth estimation methods to a handful of manually picked lakes or data granules, with no clear pathway to large-scale computational implementation across the ATL03 data catalog, which comprises hundreds of terabytes of unstructured point cloud data (Neumann et al., 2023b). To address this challenge, we have created a fully automated and scalable algorithm for lake detection and depth determination from ICESat-2 data. Here, we present this algorithm, demonstrate its performance in both Greenland and Antarctica, and provide a framework for its large-scale implementation using distributed high-throughput computing.

2 Background

2.1 How supraglacial lakes affect ice sheet mass loss

Supraglacial water has different roles in Greenland and Antarctic ice sheet mass loss processes, largely because of its different spatial extent on each ice sheet. Across most of the Greenland Ice Sheet, meltwater pools in supraglacial lakes that extend

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

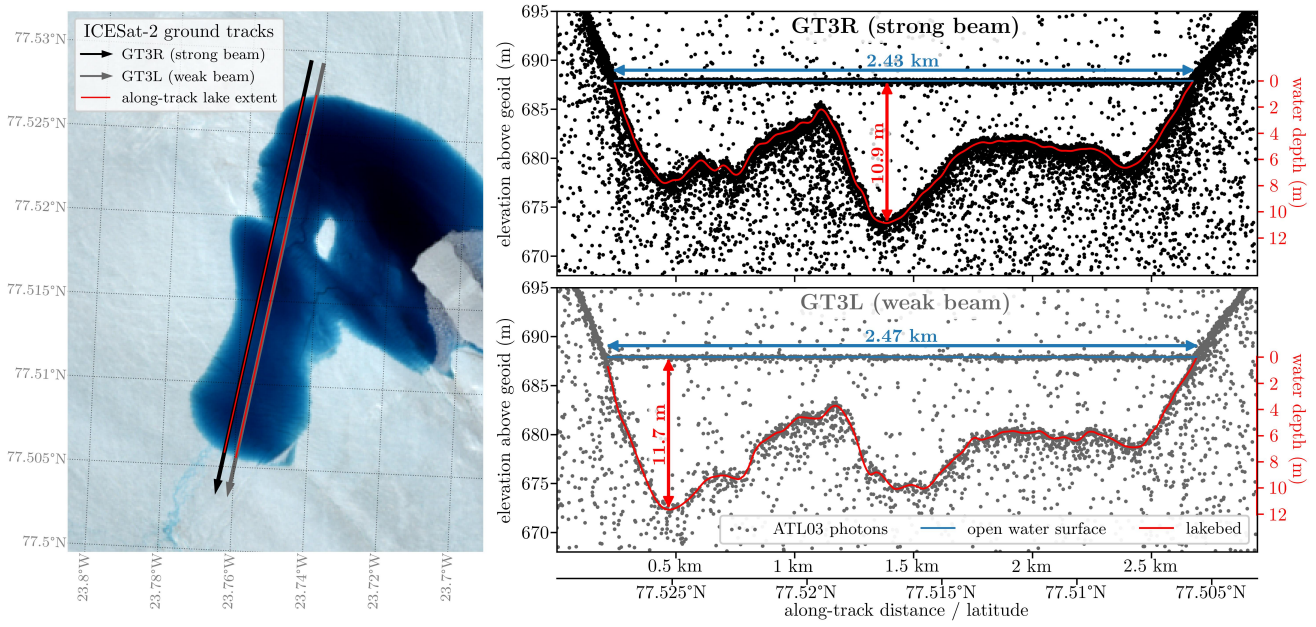

Figure 1. ICESat-2 ATL03 data over a supraglacial lake, showing a particularly strong bathymetric signal. (Data from ICESat-2 Track 406 on 20 July 2021; granule: ATL03_20210720053125_04061205_006_01.h5. The imagery in the left panel is a Sentinel-2 scene from the same day: S2B_MSIL2A_20210720T151809_N0301_R068_T27XVG_20210720T175839.)

from the ice margins up to about 2000 m on the plateau, and are forming further inland as temperatures increase (Leeson et al., 2015; Tedstone and Machguth, 2022). On the Antarctic Ice Sheet, pooling of surface meltwater in lakes is not as pervasive and is mostly observed at low elevations near the grounding lines of the ice shelves (Stokes et al., 2019), with large regional and interannual variability (Arthur et al., 2022). Pooling and storage of meltwater in supraglacial lakes can affect ice sheet mass loss directly or indirectly in four ways:

1. *Surface runoff:* Supraglacial lake drainage and transport of water off the ice sheet through surface streams or englacial pathways contributes to mass loss directly as surface runoff. This is already a significant component of the Greenland Ice Sheet surface mass balance (The IMBIE Team, 2020), but has also been observed on the Antarctic Ice Sheet (Bell et al., 2017; Warner et al., 2021; Trusel et al., 2022) and could become more significant in the future (Kingslake et al., 2017; Bell et al., 2017). Such runoff also results in surface elevation lowering, which further increases meltwater production by exposing the ice sheet surface to the higher temperatures found at lower elevations (Levermann and Winkelmann, 2016; Bell et al., 2018).

2. *Surface albedo lowering:* Supraglacial lakes lower the surface albedo, which can further accelerate melting and result in a temperature increase in the adjacent ice (Tedesco et al., 2012; Ryan et al., 2017; Stokes et al., 2019).

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

3. *Bedrock lubrication:* On grounded ice, rapid drainage of surface lakes by hydrofracture delivers pulses of meltwater to the base of the ice sheet, which lubricates the bedrock and causes acceleration of ice flow due to enhanced basal sliding. This is a well-studied phenomenon in Greenland (e.g., Das et al., 2008; Bartholomew et al., 2010; Tedesco et al., 2013; Maier et al., 2023) but has recently also been observed on the Antarctic Ice Sheet (Tuckett et al., 2019), where it could become an increasingly important mechanism as future warming will cause its hydrology to become more similar to Greenland's current ablation zone (Bell et al., 2018).

4. *Ice shelf collapse:* In Antarctica, the ponding and draining of supraglacial lakes can weaken and fracture the floating ice shelves (Munneke et al., 2014; Banwell and Macayeal, 2015; Banwell et al., 2019; Lai et al., 2020), which, in extreme cases, has been linked to their collapse by hydrofracture (MacAyeal et al., 2003; Scambos et al., 2004; Banwell et al., 2013). The resulting loss of buttressing back-stresses leads to accelerated discharge of upstream grounded ice into the ocean, which causes sea level rise (De Angelis and Skvarca, 2003; Scambos et al., 2004; Rignot et al., 2004; Rott et al., 2018). It has been hypothesized that these melt-driven hydrofracture processes could expose marine ice cliffs that are sufficiently tall and weakened to be prone to mechanical failure, which would trigger buoyancy-driven calving and could therefore lead to sustained, rapid ice sheet collapse, referred to as the Marine Ice Cliff Instability (MICI; Bassis and Walker, 2012; Pollard et al., 2015; DeConto and Pollard, 2016; Bassis et al., 2021, 2024).

Incorporating those processes through which surface meltwater ponding affects ice dynamics into ice sheet models can drastically increase projected future sea level rise (Martin et al., 2019; Edwards et al., 2021), yet they currently rely on poorly constrained parametrizations, making projections highly uncertain (Robel et al., 2019; Pattyn and Morlighem, 2020). This means that there is an urgent need to improve our understanding of the key underlying physical processes based on accurate observations (Hanna et al., 2024).

2.2 Observations of supraglacial lake depths

In-situ observations of melt lake depths are scarce (e.g., Tedesco and Steiner, 2011) due to the challenging logistics and planning required to collect such data. Supraglacial hydrological systems on ice sheets form seasonally in some of Earth's most remote and inaccessible locations, and they can rapidly evolve in complex and unpredictable patterns (Dirscherl et al., 2020; Gantayat et al., 2023), making survey planning difficult. Therefore, to obtain ice-sheet-wide estimates of meltwater lake depths for each melt season, it is necessary to rely on satellite remote sensing techniques (Moussavi et al., 2016; Melling et al., 2024). Besides ICESat-2's novel capability to directly observe water depths from photon refraction, various methods have been used to indirectly estimate lake depths from satellite data, which all have different advantages and disadvantages.

One such method is to apply a **radiative transfer equation** (RTE, Philpot, 1987, 1989) to estimate lake depth from optical imagery (Sneed and Hamilton, 2007; Moussavi et al., 2020; Leeson et al., 2020). This approach has been widely used since optical imagery provides continuous spatial coverage at short temporal intervals and because it is assumed that its physics-based principles hold everywhere, which makes it possible to apply it at scale. However, the RTE approach relies on poorly constrained choices of water attenuation coefficients and lakebed albedo, and makes simplifying assumptions such as: no

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

suspended particulate matter; a homogeneous lakebed albedo; no surface disturbances caused by wind; and the water column has vertically homogeneous optical properties (Brodskỳ et al., 2022). As a result, it has been found that the RTE approach can significantly over- or underestimate lake depths (Fricker et al., 2021; Melling et al., 2024).

Another approach to estimating lake depths is using **empirical models** derived from regression of in-situ depth measurements with optical imagery (Tedesco and Steiner, 2011; Legleiter et al., 2014; Pope et al., 2016). However, the regression coefficients of these models are limited to the spatial area of the original in-situ measurements, making them impractical for application on a larger, ice-sheet-wide scale.

A third approach is to use **digital elevation models** (DEMs) of a lake's bed topography that were acquired before it filled or after it drained, and then to determine its fill level from imagery (Moussavi et al., 2016; Yang et al., 2019b). While this has the advantage of being independent of the optical properties of the water column, currently available DEM acquisitions are sporadic and the method cannot account for changes in the lakebed topography between acquisitions, due to, for example, bottom ablation (Tedesco et al., 2012). Because this approach requires acquisitions from before a lake fills or after it drains, it is not suitable for perennial lakes that freeze over and are buried in winter without draining (Koenig et al., 2015; Schröder et al., 2020; Leppäranta et al., 2013), and it cannot be directly applied to lakes on floating ice shelves, where any filling and draining events result in a hydrostatic adjustment that bends the ice surface (Scambos et al., 2009; Warner et al., 2021).

3 Data and methods

We use individual photon data from the ICESat-2 instrument, provided in a data product known as Global Geolocated Photons (ATL03; Sect. 3.1). Our method to extract water depths consists of two algorithms that are run consecutively on ATL03 data: (1) Flat Lake and Underlying Ice Detection (FLUID, Sect. 3.2) automatically detects the locations of supraglacial lakes in the point cloud data; (2) Surface Removal and Robust Fit (SuRRF, Sect. 3.3) determines the along-track depth of each detected lake. To automatically compute results for large numbers of ATL03 data files over extensive geographical regions, we use the Open Science Grid's (OSG) Open Science Pool for distributed High-Throughput Computing (Sect. 3.4; OSG, 2006; Pordes et al., 2007). Figure 2 summarizes the various steps of this method in a flowchart, and we describe each of them in more detail below.

3.1 ICESat-2

ICESat-2 was launched in September 2018 and carries the Advanced Topographic Laser Altimeter System (ATLAS) instrument, a photon-counting green light ($532\,\mathrm{nm}$) laser altimeter operating at a frequency of $10\,\mathrm{kHz}$, which results in a $0.7\,\mathrm{m}$ along-track resolution (Markus et al., 2017). ATLAS divides the laser pulse it emits into six beams, forming three beam pairs, each of which consist of a weak and a strong beam. The footprint of each beam is about $11\,\mathrm{m}$ in diameter on the ground (Magruder et al., 2021a). The six resulting ground tracks (GTNXs) are referred to as GT1L, GT1R, GT2L, GT2R, GT3L and GT3R, where the number N refers to the beam pair from left to right in direction of flight, and the X refers to the left ("L") or right ("R") track within each pair (Neumann et al., 2022). The three track pairs are separated $3.3\,\mathrm{km}$ on the ground, and the two

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

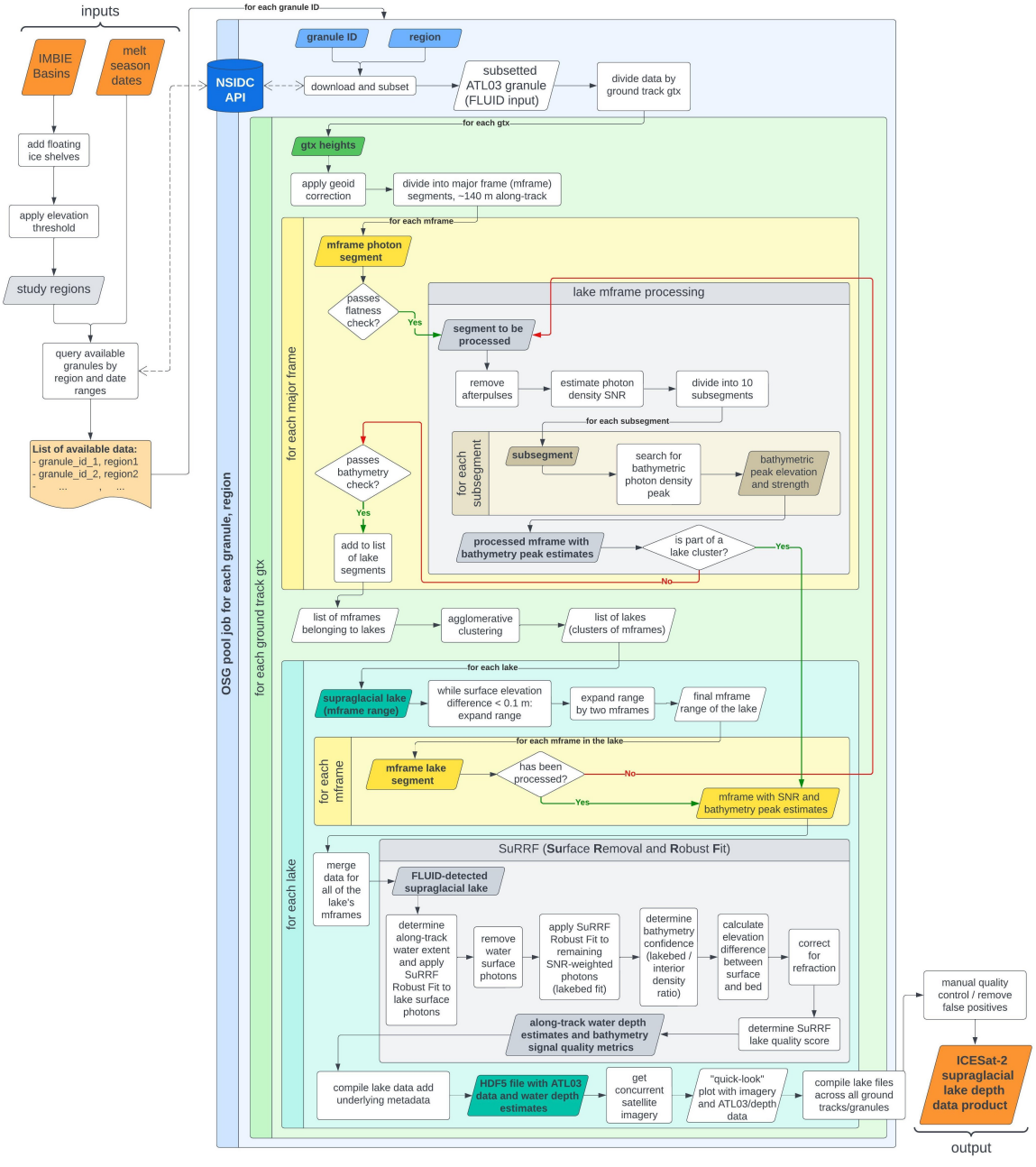

Figure 2. Flowchart of the FLUID/SuRRF framework for detecting and determining the depths of supraglacial melt lakes in ICESat-2 data for any melt season over any drainage basin of the Antarctic or Greenland ice sheets. All modules in the blue box can be parallelized for large amounts of input data granules, as a batch of compute jobs on a platform for distributed High-Throughput Computing, such as the OSG Open Science Pool.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

tracks within a pair are separated by 90 m each. The ICESat-2 spacecraft can fly in either forward or backward orientation, and flips between the two approximately twice a year (Neumann et al., 2019; Smith et al., 2019). In forward orientation, ATLAS's strong beams – numbered 1, 3 and 5 – are on the right side of each beam pair in the direction of flight and point to GTs 3R, 2R and 1R, respectively. Similarly, after a yaw flip to the backward orientation, the strong beams are on the left side of each beam pair and point to GTs 1L, 2L and 3L. This means that the data acquired along a particular GTNX can be associated with a strong or a weak beam, depending on the spacecraft orientation at the time of data acquisition. The ATLAS receiver uses photomultiplier tubes (PMTs) designed to detect individual reflected photons, with 16 independent timing channels for each strong beam and four for each weak beam (Yang et al., 2019a). The strong beams have four times more energy than the weak beams, resulting in a correspondingly higher count of laser photons returned per shot.

ICESat-2 repeats its orbit every 91 days, after completing 1,387 distinct reference ground tracks (RGTs). Over land ice, ATLAS routinely points to these RGTs near its nadir to acquire repeat measurements (Magruder et al., 2021b). The satellite began targeting the planned RGTs in late March 2019, once the on-orbit pointing calibrations were finalized and updated in the onboard pointing control systems (Martino et al., 2019). Consequently, any observations during the 2018-19 Antarctic melt season do not align with the planned repeat tracks. ICESat-2 was in "safe-hold" from 26 June through 9 July 2019, which means that no data was collected during this 14-day period coinciding with the 2019 Greenland melt season.

Over shallow (< 30 m) and non-turbid water bodies, ATLAS's green light is able to pass through the water column, which means that signal photons can be reflected from both the flat open-water surface and the lakebed (Fair et al., 2020; Fricker et al., 2021). Most land ice applications use the ATL06 data product designed for glacier and ice sheet surfaces (Smith et al., 2019). However, the ATL06 algorithm provides only one surface and thus cannot be used to extract meltwater depths. Also, over melt lakes ATL06 inconsistently tracks either the water surface or the lakebed, so the results are ambiguous (Fricker et al., 2021). To overcome this limitation and track both surfaces, our technique relies on the elevations of individual photons, which are distributed in the Global Geolocated Photons (ATL03) data product (Neumann et al., 2023b). To keep the size of each individual ATL03 data file (or "granule") manageable, each RGT orbit of ICESat-2 data is divided into 14 granule regions (Neumann et al., 2023a). This means that each granule is limited to approximately 30 min of along-track data and rarely exceeds 10 GB in size. ATL03 reports geolocated photon attributes such as longitude, latitude, along-track distance and height for each individual photon detection event, thus providing an along-track point cloud of photon locations. Geophysical corrections (such as geoid height) are reported at a 20 m along-track segment rate, and parameters related to on-board data processing (such as telemetry window ranges) are reported at the 50 Hz (≈ 140 m along-track) "major frame" rate (Martino et al., 2022b). In an ATL03 point cloud, signal photons being reflected both from a lake's water surface and its lakebed results in characteristic double returns, which are used by FLUID to detect supraglacial lakes and by SuRRF to generate along-track lake depth estimates. While the strong beam data have a higher signal to noise ratio, we have designed our FLUID/SuRRF method to work well with both strong and weak beams whenever a bathymetric return from the lakebed is discernible.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

3.2 Supraglacial lake detection in ATL03: the FLUID algorithm

The Flat Lake and Underlying Ice Detection (FLUID) algorithm takes an ATL03 granule as input, searches for locations that contain potential supraglacial lakes with a bathymetric signal and then returns along-track segments of the data across those lakes. FLUID exploits two unique characteristics of supraglacial lakes in ATL03 data:

1. photons which are reflected back from an open water surface cluster around a flat line (Sect. 3.2.1, Fig. 3) and
2. a bathymetric return signal must present as a secondary peak in photon density below such a flat surface (Sect. 3.2.4, Fig. 6).

To search for supraglacial lakes in ATL03, FLUID divides the photon data into $140\,\mathrm{m}$ along-track segments aligning with ICESat-2's "major frames" and selects those that satisfy both of the above requirements. Then, adjacent major frames are iteratively clustered into larger along-track data segments that likely represent an entire supraglacial lake each (Sect. 3.2.5).

3.2.1 FLUID step 1: identification of flat water surfaces

This step uses the fact that the surface slope of a stationary body of open water is close to zero in geopotential coordinates (i.e. relative to the geoid), in contrast to the surrounding ice sheet or ice shelf which mostly have slopes greater than $0.01°$ (Shen et al., 2022; Fan et al., 2022). This simple property enables a computationally inexpensive calculation (a "flatness check") to be applied to geoid-corrected ATL03 height data to check for possible candidate lakes.

To perform this flatness check, we apply the ATL03-provided geoid correction to photon heights, and divide the data of each of ICESat-2's six ground tracks into approximately $140\,\mathrm{m}$ along-track segments aligning with ICESat-2's major frames. For each major frame, we bin photon elevations in $0.01\,\mathrm{m}$ intervals and smooth the resulting histogram using a gaussian filter with a standard deviation of $0.05\,\mathrm{m}$, then normalize it by dividing by its largest value. If the smoothed histogram has a single peak, we record the elevation of this peak h_{peak} as the surface elevation at which a flat surface reflector would be located. In the case of a melt lake with a bathymetric return signal, it is possible that the return from the lakebed is stronger than the surface return. Therefore, if the smoothed histogram has multiple peaks with prominence > 0.1, we choose h_{peak} from the two most prominent peaks, and set it to the elevation of the one located at a higher elevation.

The flatness check is based on ratios between photon densities d_i that we calculate for various elevation bands around h_{peak} (Appendix A). As illustrated in the lower panels of Fig. 3, d_0 is the photon density within an elevation band of $\pm w_{\mathrm{peak}} = 0.1\,\mathrm{m}$ around the photon density peak. If a major frame contains the flat surface of a lake, then most of the surface signal photons should be contained in this "lake surface elevation band", making d_0 significantly larger than the photon density in surrounding elevation bands (see Fig. 3, panel b). d_1 and d_2 are the photon densities within elevation bands of width $w_{\mathrm{buffer}} = 0.35\,\mathrm{m}$ just below and above the lake surface elevation band, respectively. Due to multiple scattering in the water column of a lake, we expect that over supraglacial lakes the photon density just below the surface (d_1) can take on larger values than the photon density just above the surface (d_2). d_3 is the photon density within the entire telemetry window except for the lake surface elevation band, and d_4 is the photon density between the top of the lake surface elevation band and the top of the telemetry

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

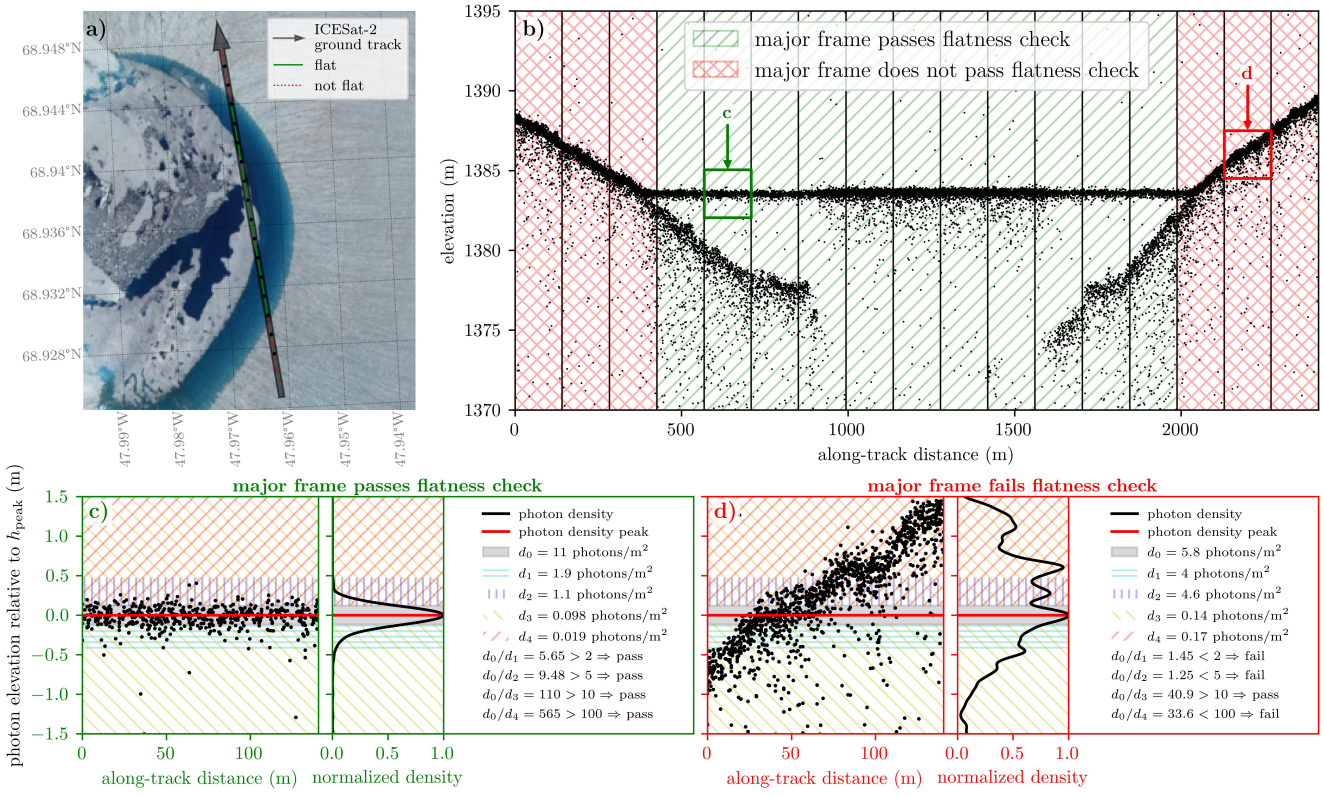

Figure 3. FLUID "flatness check" applied to every ATL03 major frame for identifying potential supraglacial lakes. a) Ground track of an along-track segment of ATL03 data over the Greenland Ice Sheet, with a partially ice-covered supraglacial lake; b) Corresponding along-track photon elevations, with major frame boundaries marked by vertical black lines and flatness check outcomes shown in hatching; c) and d) Photon density ratios for a passing and a failing segment, respectively. (Data from ICESat-2 Track 216 GT1L on 12 July 2019 and centered at 68.9370° N, 47.9657° W; granule: ATL03_20190712052659_02160403_006_02.h5, imagery: Sentinel-2 on 13 July 2019)

window. Over a lake, most of the telemetry window outside the lake surface elevation band contains only background noise photons, so we expect that the photon densities d_3 and d_4 need to be significantly smaller than the surface photon density (d_0) if the major frame contains a flat lake surface. Since d_3 can still contain photons below the lake surface due to multiple scattering and a bathymetric signal, we expect that over supraglacial lakes d_3 can take on larger values than d_4. Based on these
205 assumptions, and empirical observations of the relative photon densities in the above-described elevation bands in a variety of ATL03 granules over melt lakes and their surroundings, we defined the following thresholds on the density ratios that need to hold for a major frame to pass the flatness check: $d_0/d_1 \geq 2$, $d_0/d_2 \geq 5$, $d_0/d_3 \geq 10$ and $d_0/d_4 \geq 100$. Panels a and b of Fig. 3 illustrate the outcome of the flatness check for all the major frames within a short along-track segment of ATL03 data that contains a partially ice-covered supraglacial lake. Panels b and c show examples of major frames that pass the flatness check
210 and fail the flatness check, respectively, and illustrate the elevation bands that were used to calculate photon density ratios.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

The presence of a flat surface within a major frame is a necessary condition for detecting supraglacial bathymetry data, but it is not sufficient. There are many types of surfaces that would pass the flatness test but are not supraglacial lakes with a bathymetric signal. This includes areas of slush, frozen-over supraglacial lakes covered in ice and snow, any areas of sea ice, ocean water or ice-marginal lakes erroneously included in the ice mask used for subsetting data, and other short along-track sections over firn or glacial ice that happen to be extremely flat by chance. For example, a lake may have partial ice cover, which prevents ICESat-2 from obtaining a bathymetric return (Fig. 3). However, since the ice cover here appears to be thin and flat, the corresponding major frames still pass the flatness check despite the absence of any useful bathymetry data in those segments. This means that the flatness check presented in this chapter serves as a preliminary screening method, helping to efficiently narrow down the number of along-track segments that could potentially contain useful supraglacial bathymetry data. This process makes it computationally feasible to determine whether a bathymetric signal is actually present by performing more complex operations on only the data that remain after checking for a sufficiently flat surface. The following sections describe these methods, which are at first only applied to those major frames that passed the initial flatness check.

3.2.2 FLUID step 2: removal of afterpulses

The second step removes artefacts in the ATL03 photon data known as "afterpulses", which appear as additional lines below and parallel to the primary surface return, due to the specifics of the ATLAS sensor (Luthke, 2023; Lu et al., 2021; Martino et al., 2022a). Afterpulses only become noticeable when the sensor is nearly or fully saturated, which means they often appear in ATL03 data over supraglacial lakes because smooth open water surfaces can result in specular reflection. Figure 4 (panels d-h) shows an example of ATL03 data over a supraglacial lake in which these afterpulses are clearly visible below the flat water surface. There are three different mechanisms that can cause afterpulses:

1. *Dead-time* afterpulses appear in saturated pulses due to the ATLAS receiver channels only being able to register one photon event roughly every $3\,\mathrm{ns}$. If the return signal is strong enough that all receiver channels register photon events during a time span shorter than this "dead-time", ATLAS cannot register any photons until the receiver channels have recovered. This means that for saturated pulses, afterpulses can appear in intervals of about $3\,\mathrm{ns}$ of photon flight time, equivalent to $\sim 0.45\,\mathrm{m}$ of elevation (Lu et al., 2021).

2. *Internal reflection* afterpulses are found in ATL03 data around $2.36\,\mathrm{m}$, $4.27\,\mathrm{m}$ and $6.59\,\mathrm{m}$ below the surface return (Martino et al., 2022a). These are due to optical reflections internal to the ATLAS receiver.

3. *PMT ionization* afterpulses appear as a broad peak ~ 12–$40\,\mathrm{m}$ below the surface when pulses are strongly saturated and cause ionization of the Photomultiplier Tubes (PMTs), which triggers false photon detection events.

Since all of these afterpulses present as secondary peaks in photon density below the primary surface return, they can be mistaken for or obscure any real bathymetric signal returns. Therefore, they need to be removed before determining whether a bathymetric signal is present in the data. ATL03 provides the parameter `quality_ph` that is designed to allow users to filter out afterpulses. However, this parameter does not remove most dead-time afterpulses and naively removes all data more than

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

$2\,\mathrm{m}$ below the surface for saturated returns (Neumann et al., 2022). This means that using the ATL03-provided `quality_ph` flag is not appropriate when searching for sub-surface return signals in saturated pulses, as it would fail to remove dead-time afterpulses that could be misidentified as bathymetric signals and could remove actual bathymetric signals at depths greater than $2\,\mathrm{m}$ (Fig. 4 panel e). Therefore, we developed an improved afterpulse removal routine that is tailored to bathymetric applications.

Afterpulse removal: We first estimate the saturation level of each pulse, based on the sensor dead-time t_{dead} which is provided for each beam in the ATL03 data product. Let n_{ch} be the number of receiver channels, i.e. $n_{\mathrm{ch}} = 4$ for weak beams and $n_{\mathrm{ch}} = 16$ for strong beams. If the total number of photons in a pulse $n_{\mathrm{ph}} \geq n_{\mathrm{ch}}$ we can calculate the minimum vertical distance spanned by any n_{ch} photons, and denote it by Δh. We then estimate the sensor saturation ratio as $r_{\mathrm{sat}} = t_{\mathrm{dead}} c/(2\Delta h)$ if $n_{\mathrm{ph}} \geq n_{\mathrm{ch}}$ and zero otherwise. This means that for saturated pulses ($r_{\mathrm{sat}} \geq 1$), all receiver channels registered a photon within a time frame of $t_{\mathrm{dead}}/r_{\mathrm{sat}}$. For all saturated pulses, we calculate the elevation of the saturated return, h_{sat} as the mean elevation of the n_{ch} photons that span Δh.

To determine the typical locations of afterpulses relative to h_{sat} in saturated pulses, we compiled a data set of saturated pulses over melt lakes using an earlier version of FLUID (Arndt and Fricker, 2022), that did not include afterpulse removal. For each saturated pulse, we subtracted h_{sat} from the photon elevations and created a histogram of photon counts weighted by r_{sat} (Fig. 4, panels a-c). The strong peaks in this histogram correspond to the elevations at which afterpulses occur relative to the surface. We found that dead-time afterpulses occurred at four depths: $AP_1^{(\mathrm{dead})} = 0.55\,\mathrm{m}$, $AP_2^{(\mathrm{dead})} = 0.92\,\mathrm{m}$, $AP_3^{(\mathrm{dead})} = 1.50\,\mathrm{m}$ and $AP_4^{(\mathrm{dead})} = 1.85\,\mathrm{m}$. Only the first two of the internal reflection afterpulses were strong enough to significantly contaminate bathymetric data in saturated or near-saturated pulses, at $AP_1^{(\mathrm{ir})} = 2.46\,\mathrm{m}$ and $AP_2^{(\mathrm{ir})} = 4.25\,\mathrm{m}$ below the surface. While the third internal reflection afterpulse is also visible at $AP_3^{(\mathrm{ir})} = 6.52\,\mathrm{m}$, it appears that this afterpulse is not typically strong enough to be confused for a bathymetric return signal. The broad peak associated with PMT ionization around $AP^{(\mathrm{ion})} \approx 29 \pm 15\,\mathrm{m}$ only became noticeable at the typical length scales of melt lakes when $r_{\mathrm{sat}} > 3.5$. For such strongly saturated pulses, we simply discarded any photons $> 12\,\mathrm{m}$ below the surface.

With the locations of the main afterpulses known, we can use a simple empirical method to remove afterpulses from ATL03 data for all major frames that passed the flatness check and for which at least 100 photons are attributed to saturated pulses. For each major frame, we follow the above weighted histogramming procedure and examine the heights of the seven most prominent peaks; if any of these peaks align with the relative elevations of the known afterpulses, we consider it evidence for likely afterpulsing and remove any photons that belong to saturated pulses in that elevation band (Fig. 4, panel h).

While more sophisticated approaches for afterpulse removal are certainly possible, we found that in practice our purely empirical approach strikes a good balance between effectively removing enough afterpulse photons to prevent bathymetric surface fitting methods from considering afterpulses a signal, while also retaining enough photons to prevent removal of actual signal returns whenever it is possible to discern the two.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

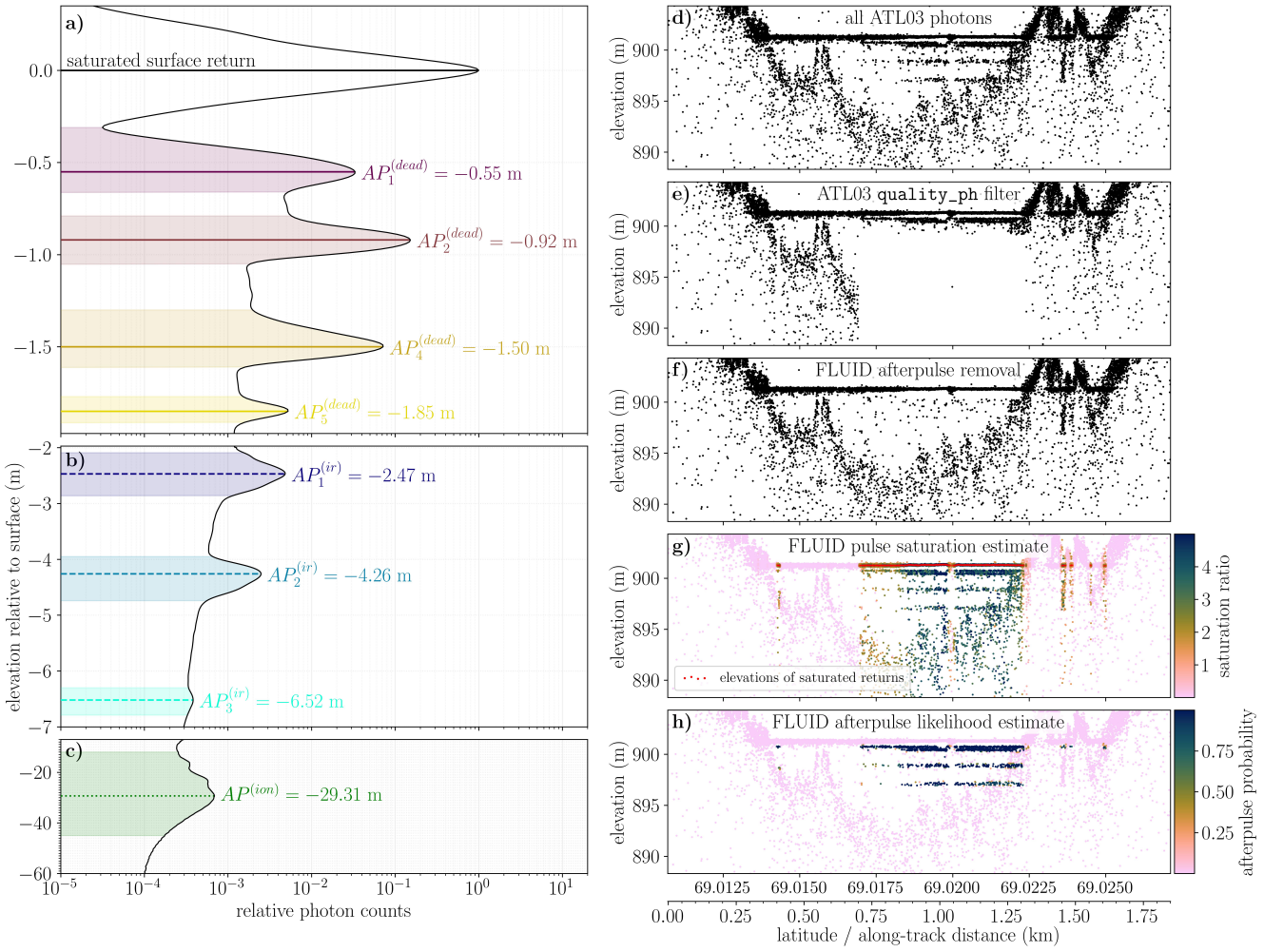

Figure 4. FLUID afterpulse removal. (a) to (c): Histogram of photon elevations in saturated pulses over melt lakes, relative to the elevation at which saturation occurred. The secondary peaks that appear below the saturated surface return are afterpulses that are caused by (a) dead-time of the ATLAS sensor; (b) internal reflections in the instrument; and (c) PMT ionization. Note that (a), (b) and (c) have different vertical scales. d) to h): Implementation of FLUID's afterpulse removal for a short along-track segment of ATL03 data that contains a supraglacial lake with sections of highly saturated (specular) pulses. Known locations of ATLAS afterpulses (a) to (c) are used to remove likely afterpulse photons. (Data from ICESat-2 Track 1222 GT2L on 17 June 2019 and centered at 69.0189° N, 49.0444° W; granule: ATL03_20190617064249_12220303_006_02.h5)

3.2.3 FLUID step 3: photon signal confidence estimation

Once flat "candidate" segments have been identified and afterpulses have been removed, the next step is to assign a signal confidence score to the remaining photons. ATL03 contains many noise photons from various sources, such as solar background noise, atmospheric backscatter, and multiple scattering in translucent media (Neumann et al., 2019; Yang et al., 2023). Release

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

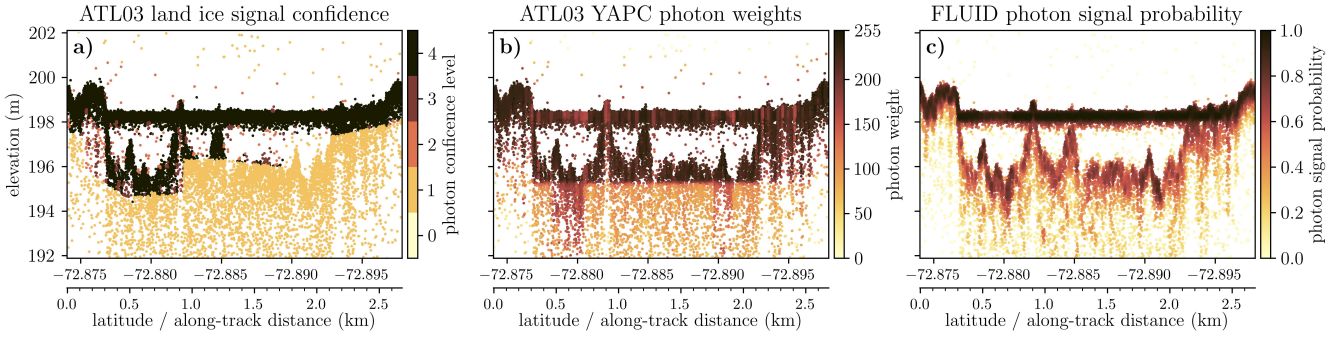

Figure 5. A comparison between existing photons signal confidence estimates and our method for melt lakes in FLUID. (Data from ICESat-2 Track 81 GT2L on 2 January 2019 and centered at 72.8859° S, 67.3082° E; granule: ATL03_20190102184312_00810210_006_02.h5)

006 of the ATL03 data product provides two measures that can help discriminate between signal and noise photons. Over the ice sheets, the `signal_conf_ph` parameter gives an estimate of how likely it is that a photon is part of the land ice surface signal, based on slant histogramming (Neumann et al., 2019). This parameter, however, does not consider the possibility of two distinct reflective surfaces that are both signal, and therefore often labels lakebed return photons over supraglacial lakes as noise (Fig. 5, panel a). Since release 006, ATL03 also includes the `weight_ph` parameter, which provides a local metric for relative photon density based on the Yet Another Photon Classifier method (YAPC; Neumann et al., 2022; Sutterley and Gibbons., 2021). For each target photon, the YAPC weight calculation is based on a rectangular window $\pm 3\,\mathrm{m}$ in elevation around the photon location. This can result in sharp photon weight discontinuities $3\,\mathrm{m}$ above and below highly reflective flat surfaces, which are inconsistent with relative local photon density (Fig. 5, panel b). Due to these drawbacks of the ATL03-provided parameters, we developed a new density-based method for photon signal confidence estimation that is more accurate over melt lakes (Fig. 5, panel c). This method is based on the inverse euclidean distances between a photon and its k-nearest neighbors within a search radius that depends on the background noise rate (Appendix B).

In FLUID, we implement this photon signal probability estimation using a KD-tree approach for querying nearest neighbors of photons, applied to individual major frames. To calculate photon densities within a major frame, we consider additional photons within a sufficiently wide buffer in along-track distance to avoid penalizing photons that are near the major frame margins by not taking into account all their nearest neighbors. Panel c) of Fig. 5 shows the resulting density-based photon signal probabilities for a supraglacial melt lake on Amery Ice Shelf.

3.2.4 FLUID step 4: secondary bathymetry peak detection

To determine which major frames amongst the ones that pass the flatness check are likely to provide useful bathymetry data, FLUID checks for secondary peaks in photon density below the flat surface return. To do so, we divide each major frame into 10 along-track sub-segments of equal length (i.e. about $14\,\mathrm{m}$ per sub-segment). For each sub-segment we use the FLUID photon-level signal probabilities to calculate photon signal confidence as an empirical smooth function of elevation. To determine

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

whether a potential bathymetric signal is present below the lake surface, we determine the elevation of the most prominent below-surface peak in this function for every sub-segment. Based on the along-track locations, elevations and prominences of all peaks that were identified in a given major frame, we define four quality heuristics q_i for different components that we found to affect the overall quality of the bathymetric return (Appendix C). The q_i all take on values between zero and one, with higher values implying a "better" bathymetric signal. We designed the expressions for the quality heuristics such that q_1 penalizes major frames with smaller numbers of detected subsurface peaks, q_2 penalizes major frames with less prominent peaks, q_3 penalizes major frames with a very large overall spread of peak elevations, and q_4 penalizes major frames with peak elevations that do not align along a smooth surface. We then define the overall bathymetric quality summary q_s of a major frame as the product of the four q_i. We consider the secondary bathymetric peak in photon density strong and coherent enough to pass the bathymetric signal check for any major frames with $q_s \geq 0.1$.

We illustrate this procedure for an along-track segment over Central West Greenland that contains a supraglacial lake with a bathymetric return signal that varies in strength along the ground track (Fig. 6). In this example, most of the major frames that belong to the supraglacial lake pass the bathymetric signal check, with bathymetric photon density peaks smoothly tracing the apparent lakebed. However, two of the major frames within the lake do not have a strong-enough signal of photons reflected from the lakebed to be passing the bathymetric signal check during this step, in which FLUID initially considers each major frame in isolation.

3.2.5 FLUID step 5: along-track aggregation of lake segments

Given the collection of major frames that individually pass the bathymetric signal check along a ground track, FLUID aggregates major frames into clusters that likely represent a transect of an entire supraglacial lake each. To achieve this, we use an agglomerative clustering scheme based on two simple assumptions: (i) the water surface elevation within a single lake should be nearly constant along the ground track, and (ii) the along-track extent of supraglacial lakes does not exceed a few kilometers. At the start of the clustering process, each major frame that passed the initial bathymetric return check is considered a singleton cluster with a water surface elevation h_{surf} equal to the single major frame's surface photon density peak h_{peak}, and major frame start and end IDs $m_{\text{start}} = m_{\text{end}}$ that are both equal to the single major frame's ID. This means that a cluster can be expressed as

$$\mathcal{C}^{(i)} = \left\{ h_{\text{surf}}^{(i)}, m_{\text{start}}^{(i)}, m_{\text{end}}^{(i)} \right\} \tag{1}$$

where the index $i \in 1, 2, \ldots, n_{\text{clusters}}$ is assigned to the i^{th} cluster when sorting all n_{clusters} clusters by their respective values of m_{start}. Since major frame IDs are numbers that strictly increase with along-track distance, this means that $m_{\text{end}}^{(i)} < m_{\text{start}}^{(i+1)}$ for all clusters. Now, clusters that are adjacent to each other in along-track coordinates are compared in a pairwise fashion. For all uneven numbers $i < n_{\text{clusters}} - 1$, if

$$\left| h_{\text{surf}}^{(i)} - h_{\text{surf}}^{(i+1)} \right| \leq \Delta h_{\text{max}} = 0.1\,\text{m} \tag{2}$$

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

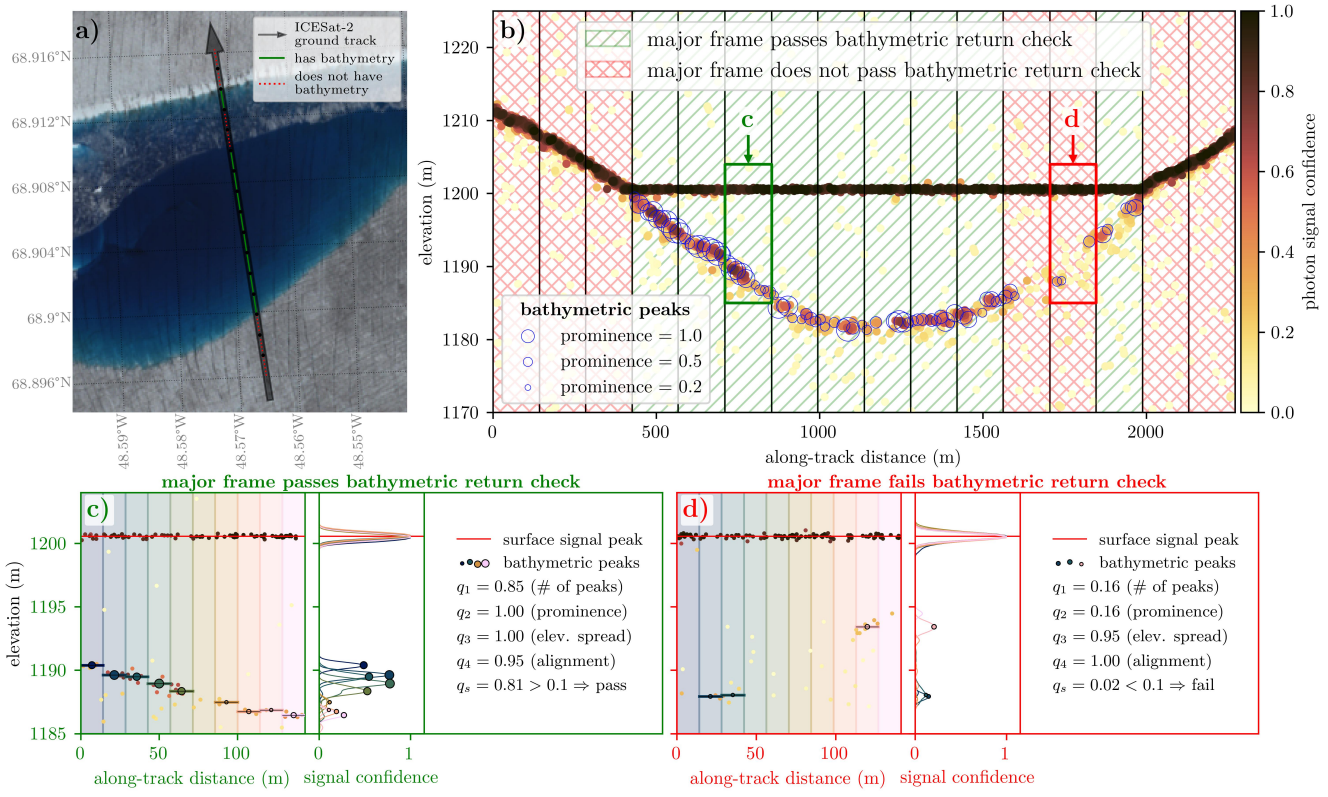

Figure 6. FLUID's bathymetric signal check run on every ATL03 major frame that has passed the flatness check. a) Ground track for an along-track segment of ATL03 data over the GrIS, which contains a supraglacial lake that has a thin partial ice cover near its northern shore. b) Corresponding along-track photon elevations, and locations of detected bathymetric photon density peaks. The vertical black lines are the major frame boundaries, and hatching indicates whether major frames passed the bathymetric signal check or not. c) and d) Bathymetric peak-finding procedure from photon density and associated values of the bathymetric return quality heuristics defined in the text, for a passing and a failing major frame, respectively. (Data from ICESat-2 Track 277 GT3R on 16 July 2019 and centered at 68.9062° N, 48.5689° W; granule: ATL03_20190716051841_02770403_006_02.h5, imagery: Sentinel-2 on 16 July 2019)

and

$$h^{(i)}_{\text{surf}} - h^{(i+1)}_{\text{surf}} \leq \Delta m_{\text{max}} = 10. \tag{3}$$

clusters $\mathcal{C}^{(i)}$ and $\mathcal{C}^{(i+1)}$ are merged into a new cluster

$$\mathcal{C}^{(i')} = \left\{ \left(h^{(i)}_{\text{surf}} + h^{(i+1)}_{\text{surf}} \right) / 2, m^{(i)}_{\text{start}}, m^{(i+1)}_{\text{end}} \right\}. \tag{4}$$

If this does not result in any two clusters being merged, then the same pairwise comparison is carried out for all even numbers $i < n_{\text{clusters}} - 1$. After an iteration of merging clusters, the indices of the remaining n'_{clusters} clusters are re-set to $1, 2, \ldots, n'_{\text{clusters}}$, and the same procedure is repeated until no more clusters can be merged based on the conditions above.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

The resulting final clustering is now considered the set of supraglacial lakes that have been found on each ground track. Note that for simplicity we here consider an "ICESat-2 lake" to be just an along-track segment of a single ICESat-2 ground track which contains data over a melt lake. If multiple ICESat-2 ground tracks contain data over the same body of surface water, the distinct ground track segments are still considered different "ICESat-2 lakes" for the purpose of this algorithm. For example, the two ATL03 profiles acquired by the two neighboring ground tracks of the center beam pair shown in Fig 1 would be considered two distinct "ICESat-2 lakes" here despite ICESat-2 having acquired their underlying data during the same overpass over the same body of water.

Since every lake that was detected this way is characterized by an along-track range of major frame IDs $[m_{\text{start}}, m_{\text{end}}]$, FLUID extends these ranges outwards to make sure that no data were missed near the edges of any lake. To do so, each lake's range is extended to include any major frames for which $|h_{\text{peak}} - h_{\text{surf}}| < 0.2\,\text{m}$ as long as such major frames exist within three major frames of the lake's range. At the end of this process we add another four major frames as a buffer, two to each side of the lake. Since this expansion of the along-track ranges of the lakes can create lakes that overlap, the set of buffered lakes is corrected by separating partially overlapping lakes at the midpoint of their along-track overlap and removing any lakes that fully overlap with another lake. We apply FLUID steps 2,3 and 4 (afterpulse removal, signal confidence estimation and bathymetric return check) to all major frames in the buffer (i.e. now included in the along-track range of a lake but had not initially passed the flatness check). The resulting final set of lakes across all ground tracks in the input granule is the output of FLUID.

3.3 Supraglacial lake depth determination: the SuRRF algorithm

To estimate the along-track depth of each detected lake, we developed the Surface Removal and Robust Fit (SuRRF) algorithm, updated from Fricker et al. (2021). The central idea of SuRRF is to use a robust fitting procedure (Sect. 3.3.1) to first fit a smooth line to all photons returned from the lake surface and the surrounding topography (Sect. 3.3.2), then remove all photons that are part of the water surface and fit another smooth line to the remaining photons to determine the location of the lakebed (Sect. 3.3.3). The along-track water depth estimate is the elevation difference between the fits to the water surface and the lakebed, corrected for the refractive index for the speed of light in water (Sect. 3.3.5). In Fig. 7, we illustrate the main steps of SuRRF using an example of an ATL03 along track segment that was determined to be a supraglacial lake by FLUID. We summarize the main steps of SuRRF in the following sections.

3.3.1 SuRRF Robust Fit

In SuRRF steps 1 and 2 (Sect. 3.3.2 and 3.3.3), we use a tailored robust nonparametric regression method to locate the lake surface and lakebed. The SuRRF Robust Fit is based on Locally Weighted Regression and Smoothing Scatterplots (LOWESS, Cleveland, 1979), which is applied to the data iteratively, while removing outliers during each iteration to converge to an along-track fit that smoothly tracks the elevation of highest photon density. For each evaluation location x_{fit} in along-track coordinates, we fit a locally weighted n^{th}-degree polynomial regression to the photons that are at a distance of at most x_{max} in along-track coordinates. The value of x_{max} is the minimum distance from x_{fit} within which there are n_{ph} photons i with a nonzero FLUID-derived signal confidence p_i, or a minimum along-track fitting window length of x_{min}, whichever is larger.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

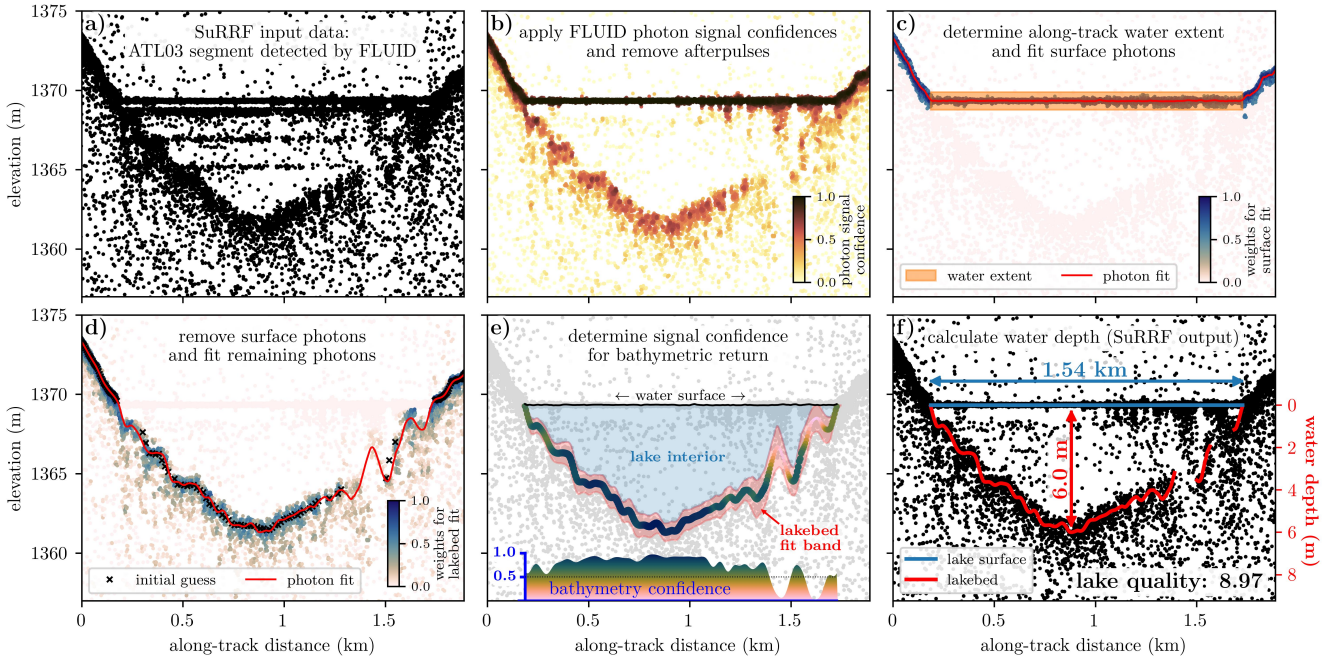

Figure 7. SuRRF algorithm for determining supraglacial lake depth from an along-track segment of ATL03 data that was detected by FLUID. (Data from ICESat-2 Track 277 GT2L on 13 July 2020 and centered at 68.1923° N, 48.5134° W; granule: ATL03_20200713115804_02770803_006_01.h5

To achieve smooth local weighting, the along-track weight of a photon i at location x_i is calculated using the tri-cube weight function $w_i^{(\mathrm{x})} = \left(1 - |(x_i - x_{\mathrm{fit}})/x_{\mathrm{max}}|^3\right)^3$. If the method is provided with an initial guess for the first iteration, we calculate the residuals e_i as the difference between each photon's elevation and the linear interpolation of the initial guess elevations to each photon's along-track location. In this case, we consider only photons whose absolute residuals are at most h_{max}, and calculate their residual-based weights as $w_i^{(\mathrm{h})} = \left(1 - (|e_i|/h_{\mathrm{max}})^3\right)^3$. In absence of an initial guess, we set $w_i^{(\mathrm{h})} = 1$ for all photons in the first iteration. The photon weights that are used for the regression are $w_i = p_i w_i^{(\mathrm{x})} w_i^{(\mathrm{h})}$. We evaluate the resulting regression model at each fit location x_{fit} to obtain an along-track estimate of the fit to the photon heights. For each consecutive iteration, we calculate the residuals e_i as the difference between each photon's elevation and the elevation of the linearly interpolated elevation fit of the previous iteration. Let σ be the standard deviation of residuals, weighted by the previous iteration's weights. To achieve a robust fit, we now consider only photons whose absolute residuals are at most n_{std} standard deviations, and calculate the residual-based weights $w_i^{(\mathrm{h})}$ as defined above, using $h_{\mathrm{max}} = n_{\mathrm{std}}\sigma$. We run this nonparametric weighted regression for a number of n_{iter} iterations to obtain a smooth fit that tracks the along-track elevation of highest photon density.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

3.3.2 SuRRF step 1: fitting the lake surface

To fit an elevation profile to just the surface of a lake detected with FLUID, we calculate the along-track extent of photons that belong to the flat lake surface, remove all photons below the lake surface, and then apply the SuRRF Robust Fit to the remaining photons (Fig. 7, panel c). We determine the extent of the open water surface by calculating the photon density within an elevation band of $\pm 0.225\,\mathrm{m}$ around the lake's surface elevation at $1\,\mathrm{m}$ along-track resolution, as well as the corresponding photon densities within the remaining telemetry window and within $2\,\mathrm{m}$ above the surface elevation band. We smooth all photon densities using an along-track gaussian window with a $15\,\mathrm{m}$ standard deviation. We then consider a location along the ground track to contain a water surface if the photon density within the surface elevation band is at least ten times as large as the other photon densities for any continuous along-track section of at least $100\,\mathrm{m}$ in length. Within the resulting estimate of along-track lake extent, we remove all photons that are more than $0.4\,\mathrm{m}$ below the lake's surface elevation from the surface fit by setting their signal confidence $p_i = 0$.

To obtain an along-track fit to the lake's surface and its surrounding topography, we apply the SuRRF Robust Fit (Sect. 3.3.1) to all remaining photons with a signal confidence > 0.5 at evenly spaced locations x_{fit} along the ground track, at $5\,\mathrm{m}$ resolution. In this step, we use a linear regression ($n = 1$) and run it for $n_{\mathrm{iter}} = 10$ iterations, with no initial guess. We here choose $x_{\mathrm{min}} = 20\,\mathrm{m}$ and let n_{ph} decrease linearly from $n_{\mathrm{ph}}^{(\mathrm{start})} = 300$ in the first iteration to $n_{\mathrm{ph}}^{(\mathrm{end})} = 100$ in the last iteration. Similarly, we choose $n_{\mathrm{std}}^{(\mathrm{start})} = 10$ and $n_{\mathrm{std}}^{(\mathrm{end})} = 4$. We illustrate this surface-fitting procedure by showing an example of a supraglacial lake in ATL03 with photons color-coded by their surface fit weights p_i, along with the detected along-track water surface extent and the final smooth photon fit to the lake and surrounding topography (Fig. 7, panel c).

3.3.3 SuRRF step 2: fitting the lakebed

To fit an elevation profile to just the lakebed, we remove all photons that belong to the lake's water surface, and then again apply the SuRRF Robust Fit to the remaining photons (Fig. 7, panel d). In this case, we remove all photons that fall within the along-track water surface extent that was determined in the previous step, and are located at an elevation of $0.35\,\mathrm{m}$ below the lake's surface or higher. To provide the SuRRF Robust Fit with an initial guess, we combine the locations of the bathymetric peaks found by FLUID that have a peak prominence value of at least 0.5 and fall within the lake's along-track surface extent with those locations of the smooth surface fit from Sect. 3.3.2 that fall outside of the lake's along-track surface extent. We then smooth the values of the initial guess using a running mean with a window of five data points to decrease the influence of any potential outliers. To decrease the influence of near-surface photons from multiple scattering, we further reduce the signal confidence values p_i of any photons that are between a lower bound of one meter above the initial guess and an upper bound of the lake surface elevation by multiplying them by a factor that linearly decreases from one to zero between the lower and the upper bound. To fit the lakebed, we apply the SuRRF Robust Fit (Sect. 3.3.1) to the same evaluation locations x_{fit} that were used to fit the lake surface. We here use a third-degree polynomial regression ($n = 3$) and SuRRF Robust Fit parameters $n_{\mathrm{iter}} = 20$, $x_{\mathrm{min}} = 100\,\mathrm{m}$, $n_{\mathrm{std}}^{(\mathrm{start})} = 10$, $n_{\mathrm{std}}^{(\mathrm{end})} = 3$, and $h_{\mathrm{max}} = 10\,\mathrm{m}$ in the first iteration. In this step we choose different values for $\left(n_{\mathrm{ph}}^{(\mathrm{start})}, n_{\mathrm{ph}}^{(\mathrm{end})}\right)$ depending on ATLAS beam strength: $(200, 100)$ for strong beam data and $(100, 50)$ for weak beam data.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

We illustrate this lakebed-fitting procedure by showing photons color-coded by their lakebed fit weights p_i, along with the initial guess and the final smooth photon fit to the lakebed (Fig. 7, panel d).

3.3.4 SuRRF step 3: bathymetry signal confidence estimation

To estimate the signal confidence of the fit to the lakebed, we calculate the photon density ratio between the lower half of the interior of the lake and the elevation band within $\pm n_{\text{std}}^{(\text{end})}\sigma$ of the last iteration of the lakebed fit for each fit location $x_{\text{fit}} \pm 5\,\text{m}$ along the ground track. Here, we consider the interior of the lake to be the elevation range between the top of the elevation band of the lakebed fit and the surface elevation of the lake. For any along-track points for which there are no lakebed photons or for which the elevation band of the lakebed fit includes the lake surface, we set the ratio to 1. We then set the bathymetry confidence to one minus the density ratio, clip it to the range $[0, 1]$, set it equal to one wherever the lakebed fit is at a higher elevation than the lake surface elevation (i.e. wherever the estimated water depth is zero), and smooth it using an along track gaussian filter with a standard deviation of $10\,\text{m}$. Wherever the elevation range of the interior of the lake is less than the width of the elevation band of the lakebed fit, we further decrease the confidence by multiplying it by the ratio between the two elevation ranges. We illustrate this bathymetry signal confidence estimation procedure by visualizing both the elevation band of the final lakebed fit and the interior of the lake, and showing the resulting along-track confidence estimates for the bathymetric return (Fig. 7, panel e).

3.3.5 SuRRF step 4: water depth calculation and lake quality

To determine the along-track water depth, we take the difference between the lake's surface elevation and the fit to the lakebed, and divide it by the refractive index for the speed of $532\,\text{nm}$ light in $0\,°\text{C}$ freshwater (≈ 1.336; Mobley, 1995). For any locations along the lake segment where the final lakebed fit (Sect. 3.3.3) returns a higher elevation than the surface elevation of the lake, we record a water depth of zero meters (i.e. no water is present). We do not correct water depths for the effect of lakebed return geolocation errors caused by refraction, since ICESat-2 is nadir-pointing to its reference ground tracks over land ice, making the water depth correction due to the angle of refraction negligibly small (≈ 0.003 of the total water depth for the slightly off-nadir-pointing outer beam pairs, which is about $9\,\text{cm}$ for a water depth of $30\,\text{m}$; Parrish et al., 2019). Final along-track water depths can be selected by applying a threshold to the bathymetry signal confidence (Sect. 3.3.4). Here, we select a confidence threshold of 0.5. For the example shown, this results in a maximum lake depth of $6.0\,\text{m}$ and gaps in along-track depth data in locations where no bathymetric return is evident (Fig. 7, panel f).

To give a lake-by-lake indication of data quality, we provide an estimate for a summarized quality measure for each lake. Let $h_x^{(\text{surf})}$ and $h_x^{(\text{bed})}$ be the surface and bed fits at along-track measurement location x. For all x where $\Delta h_x = h_x^{(\text{surf})} - h_x^{(\text{bed})} > 0$, we calculate a histogram of photon counts within a $5\,\text{m}$ along-track window for 300 elevation bins that are evenly spaced between $h_x^{(\text{bed})} - \Delta h_x$ and $h_x^{(\text{surf})} + \Delta h_x$. We then normalize the associated bin elevations $\tilde{h}$ such that $h_x^{(\text{bed})}$ corresponds to $\tilde{h} = 0$ and $h_x^{(\text{surf})}$ corresponds to $\tilde{h} = 1$, and take the per-bin sum across all x. We smooth the resulting elevation-normalized histogram using a gaussian filter with a standard deviation of three bins, and calculate the quality ratio r_q as the ratio between the value at $\tilde{h} = 0$ and the mean of the first quartile of the lowest values in $0 > \tilde{h} > 1$.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

The "quality ratio" can be considered an along-track average estimate for the photon density ratio between the lakebed and the lowest-photon-density part of the interior of the lake. We classify lakes with $r_q \leq 2$ as "zero quality" lakes. Similarly, we classify lakes with $r_q > 2$ as "high quality" lakes, for which we report the lake quality score as $r_q - 2$. This means that a lake is assigned a non-zero quality by SuRRF if the along-track averaged strength of the return signal from the lakebed is at least twice as large as the along-track averaged background noise rate within the interior of the lake. While "zero quality" lakes might still include a clear bathymetric return along a small part of their associated along-track segments, a quality score of zero is meant to indicate that there may be significant issues with data quality. In the example shown, this results in SuRRF classifying the given lake as high quality, with a score of 9.0 (Fig. 7, panel f). We show more examples of FLUID/SuRRF output lakes and their associated quality scores in a range from 0.2 to 115.5 in Figs. 9 and 10 (Sect. 4.1).

3.4 Computational implementation of FLUID/SuRRF

To facilitate large-scale use of FLUID/SuRRF, we implemented the algorithms as a Python routine that can be run on any ATL03 data granule, and developed a framework that allows for estimating all ICESat-2 lake depths for a given region of interest and time span. Given a polygonal region of interest (e.g.; a particular glacier, ice shelf, drainage basin, or other study region) and time span (e.g.; a typical melt season), we use the National Snow and Ice Data Center's Data Access and Service API (NSIDC API; NSIDC, 2021) to query for a list of all available ATL03 data granules that satisfy the spatio-temporal search parameters. To obtain all desired ICESat-2 lake depths, we can subset these ATL03 granules to the region of interest, run FLUID/SuRRF on each subsetted granule individually, and collect all melt lake depth output data. This allows for parallel processing of data granules.

To apply FLUID/SuRRF to all identified granules in an efficient, cost-effective and reproducible manner, we use the Open Science Grid (OSG) Open Science Pool for distributed High-Throughput Computing (dHTC; OSG, 2006; Pordes et al., 2007) Since batches of OSG compute jobs run on heterogeneous hardware, we run all jobs in a Singularity container (Kurtzer et al., 2017) that we designed for use with FLUID/SuRRF. We run one OSG compute job per ATL03 granule, where each job receives as input the producer ID of the granule and a shapefile of the corresponding region of interest. Each job makes a request to the NSIDC API to subset the specified granule to the given shapefile, and downloads the subsetted granule. The job then runs FLUID/SuRRF on the downloaded granule for each of ICESat-2's six ground tracks, and sends back individual HDF5 files of output water depths for each lake that was detected by FLUID.

Each output file reports water depth estimates at a $5\,\mathrm{m}$ along-track resolution with associated values for longitude, latitude, along-track distance, bathymetry signal confidence and elevations of the lakebed and surface fit to the photon data. We also include lake properties such as: surface elevation, SuRRF quality score, and various metadata such as the granule name, beam, time of data acquisition, center longitude/latitude. For reference, we add the underlying ATL03 photon heights and locations with FLUID estimates of photon signal probability, saturation level and afterpulse probability, as well as calculated FLUID parameters at the major frame rate. In addition to each lake's data file, we also create an associated "quick look" plot of the photon data with surface and lakebed fits and the ground track shown over the closest available cloud-free Landsat 8/9 or Sentinel-2 imagery (e.g.; Fig. 9, a-j and 10, a-j). The availability of these for all returned lakes makes it possible to add a final

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

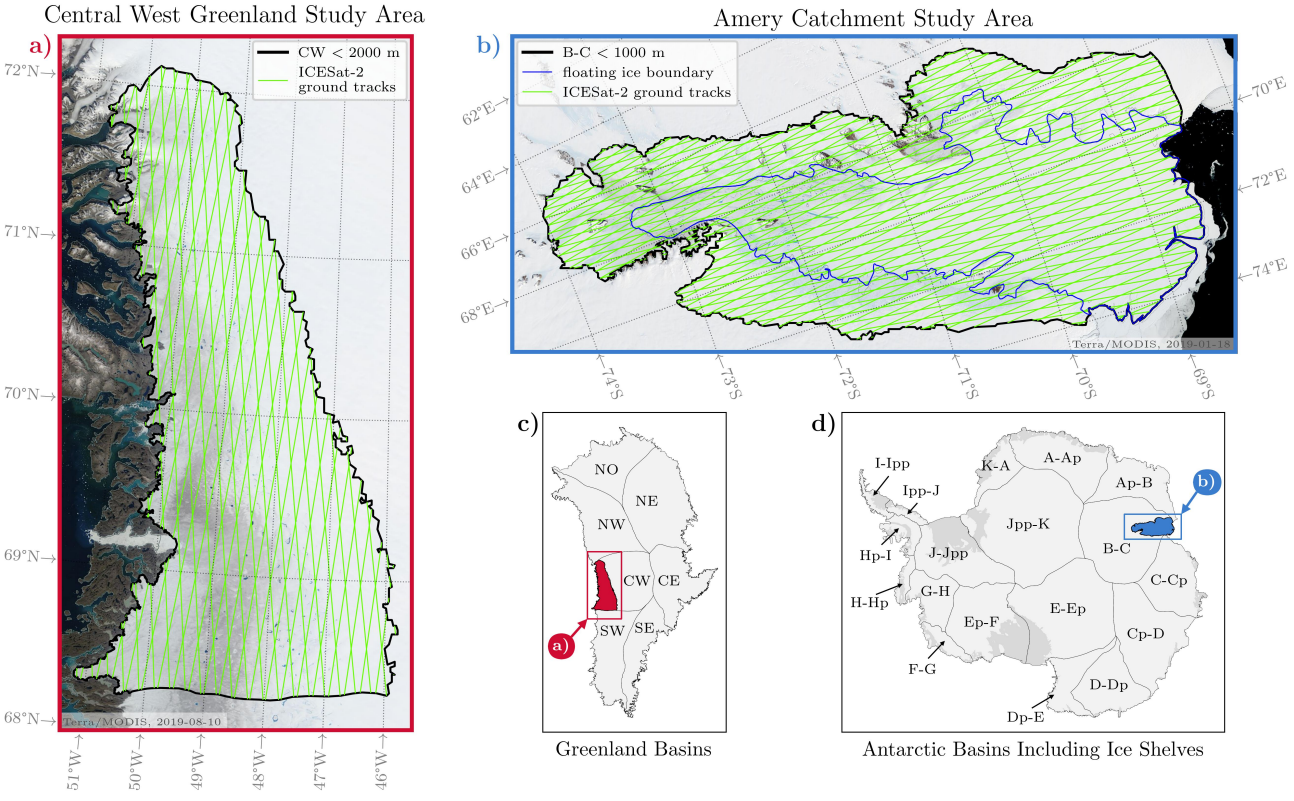

Figure 8. Study regions for testing the FLUID/SuRFF framework. a) and b) Maps of the two regions chosen for this study. The black outlines show the boundaries of the regions which were obtained by thresholding the corresponding ice sheet drainage basins (c and d) by the elevations shown in the legends. The green lines show ICESat-2 reference ground track coverage.

manual quality control step to our method, based on visual inspection of the plots. We use this to remove clear false positives from the output data.

3.5 Study regions and time span

To evaluate the performance of our method, specifically whether it is able to capture spatial and temporal variability while reliably extracting supraglacial lake depths at scale, we focus on one drainage basin on each of the ice sheets and compare a high-melt with a low-melt season for each. For both the Greenland and Antarctic ice sheets, we define our study regions using the Ice sheet Mass Balance Inter-comparison Exercise (IMBIE) drainage basins (Fig. 8, Mouginot et al., 2017; Mouginot and Rignot, 2019). Since we do not expect significant surface meltwater pooling beyond a certain elevation, we apply elevation thresholds to the drainage basins prior to running FLUID/SuRRF.

In Greenland, we focus on the Central West (CW) drainage basin, and compare ICESat-2 lake depths between the exceptionally warm 2019 melt season (Tedesco and Fettweis, 2020) and the 2020 melt season, which experienced comparatively

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

little surface melt and runoff (Druckenmiller et al., 2021). During these two summers, Central West Greenland experienced a particularly stark contrast in observed surface runoff elevation limits, with surface runoff extending to significantly higher elevations in 2019 than in 2020 (Tedstone and Machguth, 2022). For Central West Greenland, we use an elevation threshold of $2000\,\mathrm{m}$ based on Zhang et al. (2023), who reported a mean elevation limit of surface water of $1609\,\mathrm{m}$ above sea level in this region during the anomalously warm 2019 melt season. We apply this threshold based on the ArcticDEM digital elevation model (Morin et al., 2016).

In Antarctica, we focus on the B-C drainage basin, including Amery Ice Shelf and its surrounding grounded ice catchment, which on average experiences more meltwater pooling than any other Antarctic ice shelf. We compare the 2018-19 and 2020-21 melt seasons, which exhibit positive and negative anomalies in terms of open-water melt extent, respectively (Tuckett et al., 2022). For the Amery catchment, we use an elevation threshold of $1000\,\mathrm{m}$, based on Tuckett et al. (2021), who reported $> 95\,\%$ of lakes at elevations below $500\,\mathrm{m}$ in the 2004–20 time period, and only a handful of small lakes above $1000\,\mathrm{m}$ even during high-melt summers. We apply this threshold based on the Reference Elevation Model of Antarctica (REMA) digital elevation model (Howat et al., 2019).

Our two study areas cover latitudes from $68.2°$ N to $72.1°$ N in Greenland and latitudes from $68.4°$ S to $74.0°$ S in Antarctica, meaning that ICESat-2 track spacing is very similar over the two regions. The total area of the Central West Greenland study region is about $650{,}000\,\mathrm{km}^2$ with a coverage of 50 distinct ICESat-2 reference ground tracks, and the area of the Amery catchment study region is about 3.5 million km^2 with a coverage of 74 distinct ICESat-2 reference ground tracks.

For Greenland, we consider the annual melt season to be the 5-month period between the first day of May and the last day of September of a given year. Similarly, for Antarctica, we define the melt season to be the 5-month period between the first day of November and the last day of March of the following year. Based on these spatiotemporal parameters, we processed a total of 447 ATL03 granules with a total size of $1.15\,\mathrm{TB}$, amounting to a total along-track distance about $760{,}000\ \mathrm{km}$ and comprising nine billion individual photon locations.

4 Results and discussion

Using FLUID, we identified a total of 1249 supraglacial lakes over our two study areas in the available ATL03 data during the four melt seasons we considered (Table 1). We found that FLUID reliably detects potential supraglacial lakes, with the number of detected lakes varying with the strength of the melt season and their locations aligning well with imagery-derived melt extents (Sect. 4.1). Along-track lake depths determined by SuRRF agree well with manually annotated data, with deeper lakes in Central West Greenland than in the Amery catchment (Sect. 4.2). Our method is effective for detection and depth determination of supraglacial lakes over the ice sheets; however, it is not designed for ICESat-2 bathymetry over other targets, for which different methods have been developed (Sect. 4.3). Applying our method at an ice-sheet-wide scale and combining the results with satellite imagery would make it possible to develop data-driven models for accurate estimation of the volume of pooled surface meltwater across the ice sheets at high resolution and spatial coverage (Sect. 4.4).

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

Table 1. Summary statistics for the ICESat-2 lakes extracted by FLUID/SuRRF for our regions and melt seasons of interest.

	Amery catchment (B-C)		Central West Greenland (CW)	
surface melt season	2018-19	2020-21	2019	2020
melt amount	high	low	high	low
total lakes	721	28	325	175
high-quality lakes	165	5	196	109
percent high quality	23 %	18 %	60 %	62 %
median lake depth	1.85 m	1.48 m	2.77 m	3.43 m
maximum lake depth	10.4 m	17.3 m	25.8 m	15.1 m

4.1 FLUID lake detection and accuracy

4.1.1 FLUID lake detection

Out of the 1249 supraglacial lakes that we detected in total, 500 were located in Central West Greenland and 749 in the Amery catchment. The number of lakes that we detected using FLUID varied with strength of the melt season (Figs. 9 and 10). Over Central West Greenland, we identified 325 lakes during the high-melt 2019 boreal summer versus only 175 during the low-melt 2020 boreal summer. Over the Amery catchment, we identified 721 lakes during the high-melt 2018-19 austral summer versus only 28 during the low-melt 2020-21 austral summer.

Across all the data that we analyzed, FLUID and SuRRF determined on average 0.12 % of total distance along the ICESat-2 ground tracks to be meltwater surfaces. During the high-melt summers this number was 0.22 % and 0.21 % for Central West Greenland and the Amery catchment, respectively. The corresponding numbers for low-melt summers were 0.088 % and 0.0057 %. SuRRF assigned a non-zero quality score to 475 of these lakes, indicating that they likely contain high-quality bathymetric measurements. The fraction of non-zero quality lakes was significantly higher for Central West Greenland (61 %) than for the Amery catchment (22.7 %). This is likely because the locations of supraglacial lake basins on Greenland's grounded ice are primarily controlled by bedrock topography (Lampkin and VanderBerg, 2011), which allows for well-defined lake basins to develop in the same locations every year. In contrast, lake basins on Antarctica's floating ice shelves are more difficult to distinguish from their much flatter surrounding topography, on which they usually form sporadically and in different locations each year (Arthur et al., 2022).

4.1.2 Accuracy of FLUID lake detection

To validate ICESat-2-derived supraglacial lake extents, we calculated the maximum surface meltwater extent for each of the melt seasons independently using Landsat 8 imagery, based on the methods detailed in Tuckett et al. (2021) (blue regions in Fig. 9 and 10). We found a high correspondence between these estimates: For Central West Greenland, the ground tracks of

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

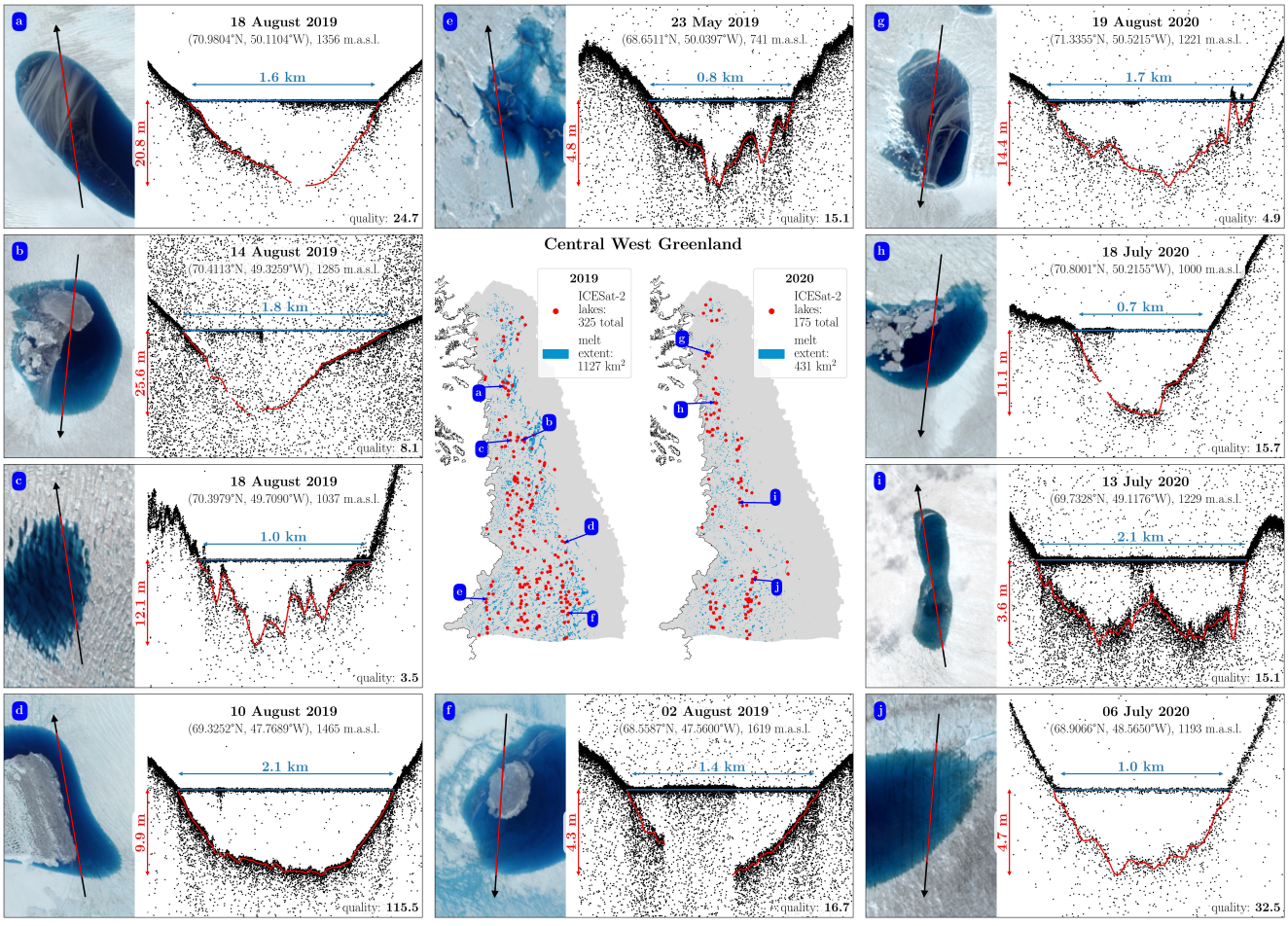

Figure 9. Center maps: FLUID/SURFF algorithm testing in Greenland. Locations of melt lakes detected in ATL03 data for the Greenland Ice Sheet's Central West drainage basin for melt seasons 2019 and 2020, mapped over the corresponding seasons' maximum meltwater extent from Landsat 8. Panels a-j: Examples showing the underlying ATL03 photon clouds and water depths calculated by SuRRF, for some of the lakes shown on the maps. Numbers in lower right of panels are SuRFF lake quality scores.

97.0 % of high-quality ICESat-2 lakes coincided with the Landsat-8-derived maximum seasonal melt extent. The corresponding percentage for the Amery catchment was 95.8 %.

Since there are no ground truth data of water depths available for any ICESat-2 data over supraglacial lakes, it is necessary to evaluate the performance of our method by manually examining the data in representative granules. We report results for ICESat-2 track 81, GT2L on 2 January 2019 over the Amery catchment, which was also studied in Fricker et al. (2021) and Xiao et al. (2023). To evaluate whether FLUID detects all supraglacial lakes that have bathymetric data in this ATL03 ground track, we manually inspected the data for any evidence of meltwater and determined whether any such along-track contained a clearly discernible return from a lakebed. Out of 25 along-track segments with meltwater, we judged that only eight were

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

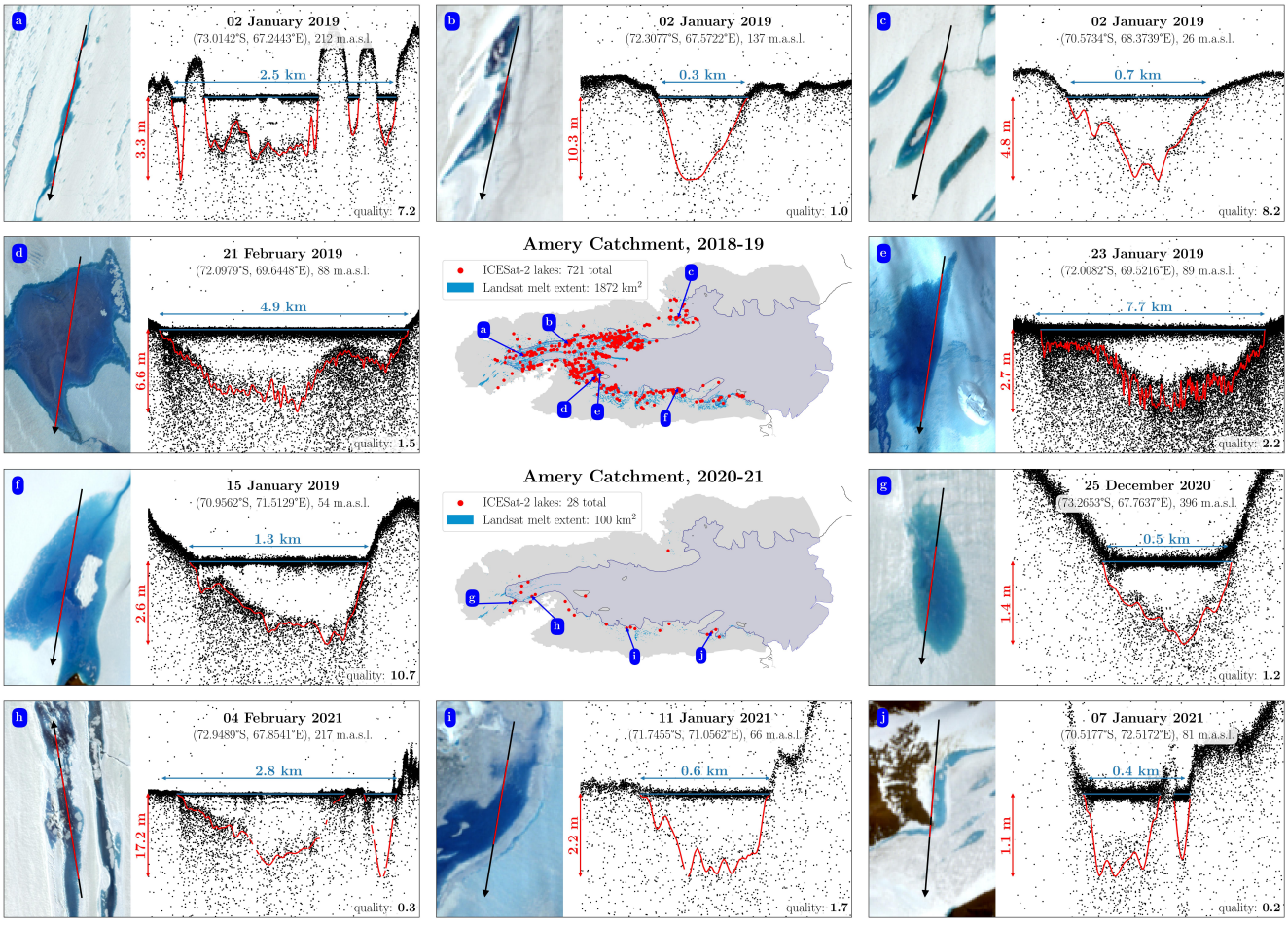

Figure 10. Center maps: FLUID/SuRFF algorithm testing in Antarctica. Locations of melt lakes that FLUID detected in ATL03 data for the Antarctic Ice Sheet's B-C drainage basin for melt seasons 2019-20 and 2020-21, mapped over the corresponding seasons' maximum meltwater extent from Landsat 8. Panels a-j: Examples showing the underlying ATL03 photon clouds and water depths calculated by SuRRF, for some of the lakes shown on the maps. Numbers in lower right of panels are SuRFF lake quality scores.

supraglacial lakes with a clear bathymetric signal. FLUID found 16 potential melt lakes within the same ground track, which all showed evidence of meltwater, but only the eight that we had also manually picked were classified as high-quality lakes.

Out of all lakes that FLUID detected, we deemed 308 to be false positives, many of which showed evidence of surface water but no clearly distinguishable bathymetric return. However, we found that lakes classified by SuRRF as "high quality" contain few false positives, with only 15 such lakes that we manually removed from the data. The majority (11) of these were due to our study region erroneously extending past the calving front of the Amery Ice Shelf in Antarctica or marine-terminating outlet glaciers in Greenland, or due to ice-marginal lakes being included in the study region. Of the remaining four false positives, three were likely due to random noise and one showed a supraglacial lake but the data seemed to be affected by a "Did Not

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

Finish Major Frame" data transfer error in the ATLAS Photon Counting Electronics (Magruder et al., 2024). While testing our algorithm, we found that false positives may also arise when ICESat-2's footprint includes two surfaces at different elevations for a substantial along-track distance, which is possible when the ground track passes over a flat calving front or crevasses with a bottom return at an acute angle. Where the study region extends past an ice shelf calving front, FLUID/SuRRF can also falsely classify the bathymetric return of a submerged "bench" as a supraglacial lake (Buck, 2024).

4.2 SuRFF depth retrieval and accuracy

4.2.1 SuRFF: meltwater depths

The meltwater depths determined by SuRRF suggest that supraglacial lakes in Central West Greenland are generally deeper than those in the Amery catchment: The median of along-track maximum water depths of lakes in Greenland was $3.09\,\mathrm{m}$ versus only $1.84\,\mathrm{m}$ in Antarctica. Median lake depth in the Amery catchment was greater during the high-melt season ($1.85\,\mathrm{m}$ than it was during the low-melt season ($1.48\,\mathrm{m}$). In contrast, median lake depth in Central West Greenland was less during the high-melt season ($2.77\,\mathrm{m}$) than it was during the low melt season ($3.43\,\mathrm{m}$) (Fig. 11, a). We hypothesize that this opposite behavior is due to different relationships between lake elevations and depths in these regions. A linear regression of lake depth on surface elevation in our data suggests that lake depth decreases with elevation in Central West Greenland ($p = 0.12$), while it increases with elevation in the Amery catchment ($p = 0.093$) (Fig. 11, b). In Greenland, where supraglacial lake basin locations are controlled by bedrock topography, lakes in a low-melt season are more likely to form at lower elevations, where repeated filling, bottom ablation and draining during most prior melt seasons has resulted in well-defined, deep lake basins. In a high-melt season, surface melt extends further upwards above the ablation zone, where lake basins are less well-defined, forming shallower lakes and slush areas. On an Antarctic ice shelf, lakes in a low-melt season are more likely to form at lower elevations, where the ice is floating and the surface topography is very flat, resulting in mostly shallow lakes. In a high-melt season, melt extends further upward of the grounding line where lakes form in deeper basins that are controlled by the bedrock topography (similar to lakes in Greenland), thus resulting in deeper lakes.

4.2.2 Accuracy of SuRFF depth retrievals

For the 16 potential melt lakes that FLUID identified in ICESat-2 track 81, GT2L on Jan 2nd, 2019 over the Amery catchment (Sect. 4.1.2 Fricker et al., 2021), SuRRF assigned a nonzero lake quality score only to the eight data segments which we had manually determined to be supraglacial lakes with a clear bathymetric signal. For four of the lakes identified in this granule, Fricker et al. (2021) established manually annotated baseline depth estimates, which were used to assess the performance of automated algorithms in the absence of any ground truth data. In addition to these four lakes, Melling et al. (2024) used a similar method to establish manual estimates for five more lakes in Southwest Greenland. When run on the corresponding ATL03 granules, FLUID automatically detects all these lakes, and the SuRRF depth estimates track the general shape of the manually outlined lakebeds well (Fig. 12). This is also demonstrated by an average Pearson's correlation coefficient of $R = 0.991$ for the nine lakes examined here, with averages of $R = 0.987$ for the lakes on Amery catchment and $R = 0.994$ for

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

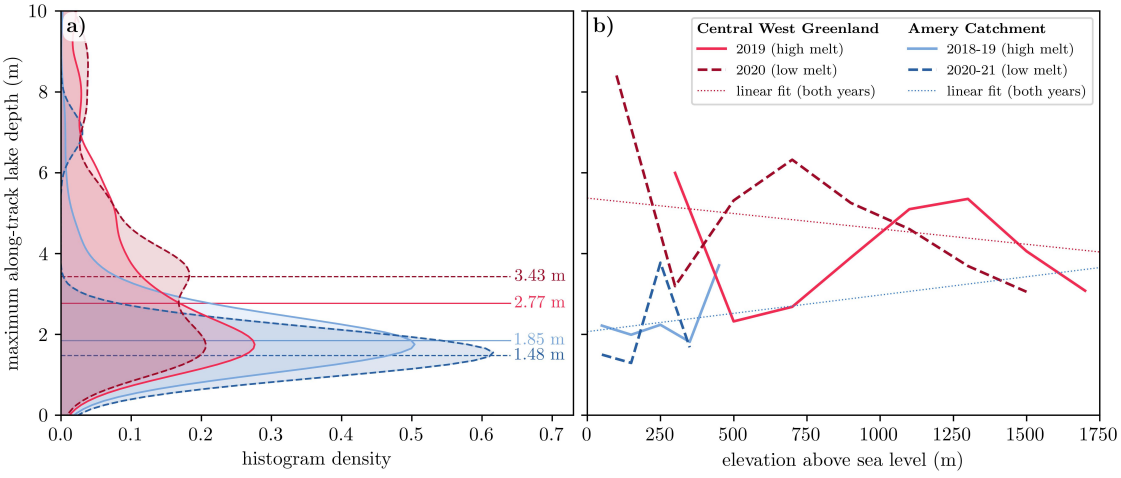

Figure 11. SuRRF lake depth statistics for two study regions in high and low melt years. a): Density of maximum lake depth distributions. The numbers shown are median lake depths during the corresponding melt season and study region. b): Elevation-binned means of maximum lake depths for each season and study region. The dotted lines are the linear fit to all data for the corresponding study region.

the lakes in Southwest Greenland. SuRRF depth estimates show an average bias of $0.27\,\mathrm{m}$ deeper than the manually picked values, with a mean absolute error (MAE) of $0.28\,\mathrm{m}$. For the lakes on Amery Ice Shelf, the average bias is $0.33\,\mathrm{m}$ and the MAE is $0.33\,\mathrm{m}$. For the lakes in Southwest Greenland, the average bias is $0.24\,\mathrm{m}$ and the MAE is $0.25\,\mathrm{m}$.

Based on the nine lakes that were used to assess the accuracy of depth estimates, most algorithms (including SuRRF) seem to estimate slightly greater depths than the manually picked values (Xiao et al., 2023). When the bathymetric returns in the ATL03 point cloud are "fuzzy" the difference between the manual baseline and SuRRF water depth estimates tends to become larger. The majority of algorithms for lake depth retrieval from ATL03 operate on the assumption that the elevation with the locally highest photon density corresponds to the elevation of the ice-water interface at the lakebed, yet most altimetry experts placed the secondary return signal higher up in the point cloud, where the photon density first significantly increases below the water surface (Fricker et al., 2021). We agree that the true location of the lakebed is likely at a higher elevation than the along-track elevation of highest photon density, since multiple scattering increases the photon density below the elevation of the lakebed. However, the elevation at which the photon density first significantly increases represents the highest point within ICESat-2's $11\,\mathrm{m}$ footprint and furthermore tracks the upper bound of ATLAS's single-photon time-of-flight uncertainty of $800\,\mathrm{ps}$ (Markus et al., 2017), which corresponds to about $12\,\mathrm{cm}$ in photon height uncertainty. We therefore believe that the true water depth falls somewhere in between our (deeper) SuRRF estimates the (more shallow) manual baseline estimates from Fricker et al. (2021) and Melling et al. (2024). This disparity in depth estimates demonstrates the need for in-situ meltwater depth data that can be used to reliably validate the accuracy of ICESat-2 estimates.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

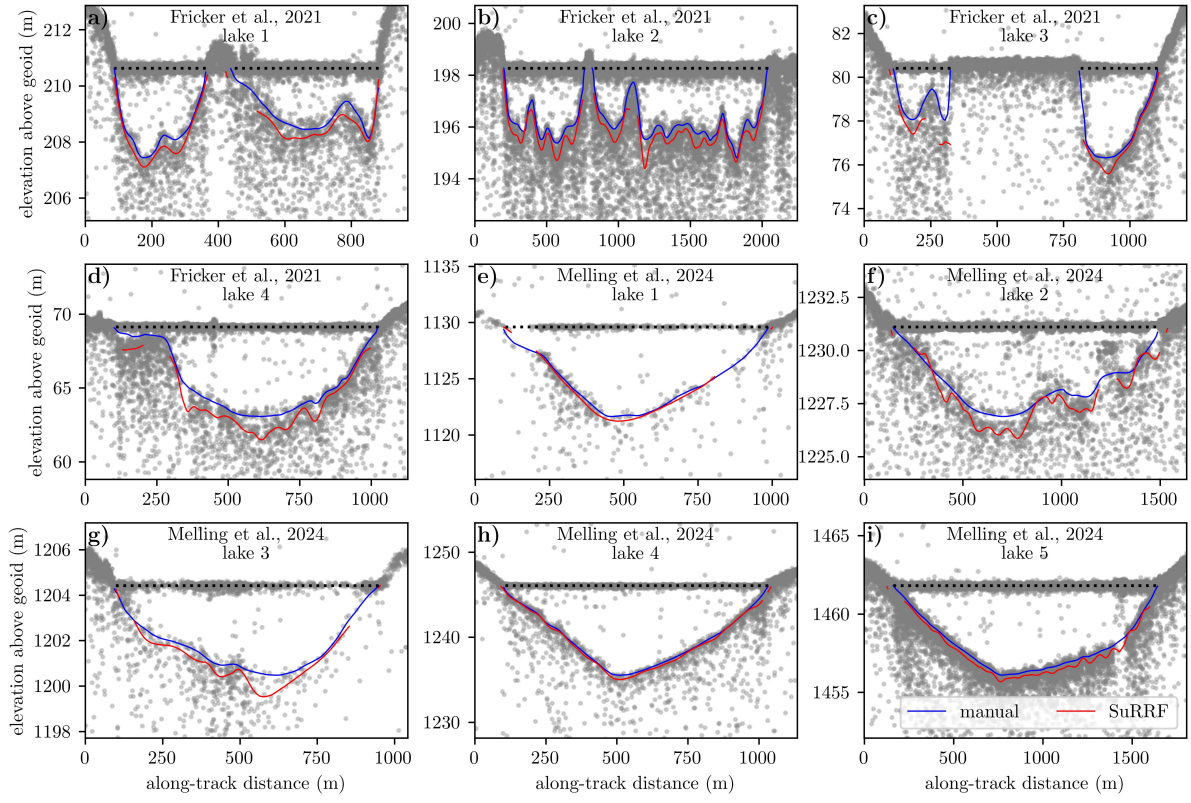

Figure 12. Comparison between SuRRF water depth estimates and the manually annotated ICESat-2 depths. a)-d): melt lakes on the Amery Ice Shelf with manual annotations from Fricker et al. (2021). e)-i): melt lakes in Southwest Greenland with manual annotations from Melling et al. (2024). SuRRF depth estimates are shown only where the estimated bathymetric signal confidence exceeds 0.5. To align along-track depths with the ATL03 photons (gray dots), water depth estimates were multiplied by the refractive index for the speed of light in water and subtracted from the surface elevation of each lake (black dotted lines).

4.3 ICESat-2 bathymetry over other targets

Beyond estimating the depth of supraglacial lakes, ICESat-2's bathymetric capabilities have been used for various other applications such as satellite-derived nearshore ocean bathymetry (e.g.; Parrish et al., 2019; Ma et al., 2020; Thomas et al., 2021), estimating water depths of inland waters (e.g.; Li et al., 2019; Xu et al., 2020; Jasinski et al., 2023) and tracking the evolution of melt pond depths on sea ice (e.g.; Farrell et al., 2020; Tilling et al., 2020; Herzfeld et al., 2023; Buckley et al., 2023). For nearshore and inland bathymetry applications, the locations of the desired bathymetry estimates are usually known *a priori* and the bottom topography is often considered constant, so bathymetric measurements can be accumulated over time and compared to non-concurrent validation data. In contrast, the ephemeral nature of supraglacial lakes on the ice sheets and melt ponds on sea ice makes it necessary to automatically detect their locations directly from ATL03 and makes it more difficult to reliably validate depth estimates. Our FLUID algorithm addresses this issue for supraglacial lakes on the ice sheets: it automatically

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

detects the locations of lakes by relying on the fact that the water surface reflection over open water presents as a flat line in geoid-corrected ICESat-2 elevation data, while the surrounding topography on the ice sheets is almost always sloped. This makes our algorithm efficient over the ice sheets by discarding most non-lake photon segments in the first "flatness check" step. However, it makes our algorithm less suited for detecting lakes on sea ice, where most segments over the open ocean and thin sea ice would likely pass the flatness threshold.

4.4 Future studies

Our ICESat-2-derived water depths are the first comprehensive dataset of supraglacial lake depths that were directly measured from a satellite; however, they are too sparsely spaced in space and time to allow for the calculation of lake volumes or for the continuous tracking of meltwater throughout the progression of a melt season. However, the large volume and wide variety of accurate water depth data that our method provides implies that the results presented here could be used to better constrain parameters in existing methods which estimate meltwater volumes from high-resolution, spatially continuous satellite imagery. Furthermore, our method could be used to extract pan-ice-sheet meltwater depths and combine them with concurrent satellite imagery, thus providing a training data set that would enable the development of data-driven models of the relationship between meltwater depth and satellite imagery reflectances, based on statistical learning methods.

5 Summary

Supraglacial lakes form seasonally around most of the margins of the Greenland and Antarctic ice sheets. In a warming climate, these lakes have the potential to significantly impact the future stability of both ice sheets through processes that are not yet understood well-enough to be included into models. To confidently project future sea-level rise, better satellite observations of surface meltwater are needed to enable science that produces a better mechanistic understanding of how ice dynamics are impacted by the pooling of surface meltwater in lakes. Until recently, any available methods used to estimate supraglacial lake depth from satellite data have been required to make strong assumptions and to use poorly constrained parameters, making it difficult to accurately assess the distribution of meltwater volumes across ice surfaces. Multiple case studies have successfully demonstrated that supraglacial lake depths can now be directly measured from photon refraction in ICESat-2's laser altimetry data. However, ICESat-2 data had not previously been used at scale for this purpose because its photon-level product comprises hundreds of terabytes of unstructured point cloud data along spatially discrete ground tracks, which makes it difficult to integrate the data with spatially continuous data in existing workflows.

To address this challenge, we have presented the fully automated, two-step FLUID/SuRRF algorithm for the detection and depth determination of supraglacial lakes on the ice sheets in ICESat-2 photon data, and proposed a computational framework that allows for its large-scale implementation across any desired ice sheet drainage basins and melt seasons. To test our method, we applied it to all available ICESat-2 data over two drainage basins, one on the Antarctic Ice Sheet and one on the Greenland Ice Sheet, for a high-melt and a low-melt summer each. We have demonstrated the following for our method:

1. It reliably detects supraglacial lakes based on the flatness of their surface and the presence of a lakebed return.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

2. There is a potential for false positives, but their impact can be effectively mitigated by filtering for the strength of the bathymetric signal.

3. Water depth estimates are accurate based on manual validation of the data, however, there is an urgent need for in-situ data for definitive ground truthing.

4. It can be applied at scale by leveraging distributed high-throughput computing.

5. The resulting data effectively capture spatial and temporal variability in meltwater extent and depth.

Our framework can be used to generate a comprehensive data product of supraglacial lake depths for Greenland and Antarctica since the launch of ICESat-2, which would enable the development of data-driven models of meltwater volumes in satellite imagery.

Code and data availability. The FLUID/SuRRF source code is freely available at https://doi.org/10.5281/zenodo.10905941 (Arndt and Fricker, 2024a). To execute this code, users need to create a free NASA Earthdata login for ICESat-2 data access. To implement the method on the Open Science Pool, users need to register for OSG Connect, which is freely available to researchers contributing to a US-based project at an academic, government, or non-profit organization. The source code contains a singularity container in which this version of FLUID/SuRRF can be executed. More information is included in the README file. The supraglacial lake meltwater depth estimates and associated "quick-look" plots for all 1249 lakes identified in this study are available at https://zenodo.org/doi/10.5281/zenodo.10901737 (Arndt and Fricker, 2024b). All data and code that is needed to reproduce the figures in this study are available at https://doi.org/10.5281/zenodo.10901826 (Arndt and Fricker, 2024c). ICESat-2 ATL03 data is available at NSIDC (Neumann et al., 2023b). Sentinel-2 and Landsat imagery were accessed via Google Earth Engine (Gorelick et al., 2017). Drainage basins for Greenland are available at DRYAD (Mouginot and Rignot, 2019). Drainage basins for Antarctica are available at NSIDC (Mouginot et al., 2017). ArcticDEM mosaics for elevation thresholding are available at https://www.pgc.umn.edu/data/arcticdem/ (Porter et al., 2023) and REMA DEM mosaics are available at https://www.pgc.umn.edu/data/rema/ (Howat et al., 2022). ICESat-2 ground track KML files are available at https://icesat-2.gsfc.nasa.gov/science/specs.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

Appendix A: Calculation of photon density ratios in FLUID flatness check

For a given major frame, let $l_{\text{mframe}} \approx 140\,\text{m}$ be the along-track length of the major frame, and let P be the set of photons with h_i denoting the geoid-corrected height of photon $i \in P$. We calculate the five photon densities used in the flatness check as:

$$d_0 = \frac{\sum_{i\in P}\left[|h_i - h_{\text{peak}}| \le w_{\text{peak}}\right]}{2w_{\text{peak}} l_{\text{mframe}}} \tag{A1}$$

$$d_1 = \frac{\sum_{i\in P}\left[0 < h_i - h_{\text{peak}} - w_{\text{peak}} \le w_{\text{buffer}}\right]}{w_{\text{buffer}} l_{\text{mframe}}} \tag{A2}$$

$$d_2 = \frac{\sum_{i\in P}\left[-w_{\text{buffer}} \le h_i - h_{\text{peak}} + w_{\text{peak}} < 0\right]}{w_{\text{buffer}} l_{\text{mframe}}} \tag{A3}$$

$$d_3 = \frac{\sum_{i\in P}\left[|h_i - h_{\text{peak}}| > w_{\text{peak}}\right]}{(h_{\text{max}} - h_{\text{min}} - 2w_{\text{peak}})\, l_{\text{mframe}}} \tag{A4}$$

$$d_4 = \frac{\sum_{i\in P}\left[h_i - h_{\text{peak}} > w_{\text{peak}}\right]}{(h_{\text{max}} - h_{\text{peak}} - w_{\text{peak}})\, l_{\text{mframe}}} \tag{A5}$$

where $w_{\text{peak}} = 0.1\,\text{m}$ is the width of the density peak, $w_{\text{buffer}} = 0.35\,\text{m}$ is the width of the buffer around the peak, and h_{min} and h_{max} are the bottom and the top of the telemetry window, respectively. Here $[\cdot]$ are Iverson brackets, which evaluate to 1 if the condition inside the brackets is true, and to 0 otherwise.

Appendix B: Calculation of photon signal confidences in FLUID

Let P be the set of photons within an along-track segment (e.g., a major frame), and for each photon $i \in P$, let x_i and h_i be the along-track distance and elevation, respectively. To give appropriate relative weights to the two spatial dimensions when calculating photon densities, we need to adjust along-track distance by an aspect ratio parameter r_a. We found that for typical supraglacial lakes, a value of $r_a = 30$ works well. Denote by $\tilde{x} = x/r_a$ the aspect-ratio-adjusted along-track distance. We now want to express a photon's signal confidence as the average of the inverse euclidean distances between the photon and up to k_{max} nearest neighbors within a search radius of r_s, where distances are normalized and clamped to this search radius. Let K_i be the set of photons including i and its k_{max} nearest neighbors, i.e. $K_i \subseteq P$ s.t. $|K_i| = k_{\text{max}} + 1$ and

$$d_{i,j} \ge \max_{n\in K_i} d_{i,n} \quad \forall j \in P \setminus K_i \tag{B1}$$

where

$$d_{i,j} = \sqrt{(\tilde{x_j} - \tilde{x_i})^2 + (h_j - h_i)^2} \tag{B2}$$

is the aspect-ratio-adjusted euclidean distance between photons i and j. Then, the signal probability of photon i is estimated as

$$p_i = \frac{1}{k_{\text{max}}}\left(k_{\text{max}} - \sum_{j\in K_i}\left(\frac{\min(d_{i,j}, r_s)}{r_s}\right) + 1\right). \tag{B3}$$

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

Note that this implies $p_i \in [0,1]$. In particular, $p_i = 0$ if photon i does not have any neighbors within a radius of r_s, and p_i would be equal to 1 only for a "perfect photon", whose location coincides exactly with that of k_{max} other photons. To effectively discriminate between signal and noise, we need to choose k_{max} and r_s in a way such that typical background noise photons are assigned a small target value, which we set to $p_{\text{noise}} = 0.05$.

Assuming that a major frame contains only noise photons and a flat surface with a signal photon spread of $w_{\text{signal}} = 0.3\,\text{m}$, we can estimate the aspect-ratio-adjusted mean area per background photon (i.e. the inverse photon density) as

$$a = \frac{(h_{\text{max}} - h_{\text{min}} - 2w_{\text{signal}})\, l_{\text{mframe}}}{r_a \sum_{i \in P} [|h_i - h_{\text{peak}}| > w_{\text{signal}}]}. \tag{B4}$$

If background noise photons follow a random uniform distribution, this means that to find an expected k neighbors around a typical background photon, one would have to search within a radius of $r(k) = \sqrt{a(k+1)/\pi}$ around that photon. Using the fact that the average inverse distance of a point in a circle with radius R to it's origin is $R/3$ (Stone, 1991), we can set

$$r_s = \sqrt{3p_{\text{noise}}}\, r(k_{\text{max}}) = \sqrt{\frac{3ap_{\text{noise}}(k_{\text{max}}+1)}{\pi}} \tag{B5}$$

to ensure that $p_i \leq p_{\text{noise}}$ for typical background noise photons. We found that setting $k_{\text{max}} = 15$ strikes a good balance between being sensitive enough to small-scale signals, and being too sensitive to small amounts of noise photons clustering together by random chance.

Appendix C: Calculation of quality heuristics in FLUID bathymetry check

To detect any potential bathymetric returns in a major frame, we divide it into $n_{\text{seg}} = 10$ along-track sub-segments of equal length. For each sub-segment, we calculate the median of the FLUID photon-level signal probability p_i for photons i within $0.1\,\text{m}$ elevation bins, assigning a value of zero whenever there are no photons within any bin. This results in an empirical function that relates signal probability to elevation h, which we here denote by $p(h)$. In addition, we calculate the analogous function of photon density $d(h)$ as a simple histogram for $0.01\,\text{m}$ elevation bins. Let $d'(h) = d(h)\ \forall h : |h - h_{\text{peak}}| > w_{\text{signal}}$ and zero otherwise. Denote by $\cdot|_l$ a smoothing operator using a gaussian window with standard deviation of l. Now calculate signal confidence as a function of h as

$$c(h) = p(h)|_l \min\left(\frac{d(h)|_l}{\max_h d'(h)|_l}, 1\right), \tag{C1}$$

where $p(h)$ is linearly interpolated to match the domain of $d(h)$. Note that this implies $p(h) \in [0,1]\ \forall h$ in the telemetry window. To determine whether a potential bathymetric signal peak is present, we find peaks in $c(h)$. If there are at least two peaks with prominence $\rho \geq 0.1$, and the elevation of the peak closest to h_{peak} is found within an elevation difference of w_{signal}, then out of the peaks whose elevation is below $h_{\text{peak}} - w_{\text{signal}}$ the peak with the highest prominence is considered a potential bathymetric peak. Denote by n_{peaks} the number of potential bathymetry peaks found in the major frame, and let h_i be the elevation and ρ_i the prominence of peak $i \in [1, 2, \ldots, n_{\text{peaks}}]$. We only consider the major frame to be potentially part of a lake if $n_{\text{peaks}} \geq 3$. In

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

this case, we define the following heuristics for different components that affect the quality of the bathymetric return within the major frame.

$$q_1 = f^{3/2} \tag{C2}$$

$$q_2 = \min\left(\frac{\sum_i \rho_i}{n_{\text{peaks}}} 2(\min(\{2f, 1\}) - 1) + 1, 1\right) \tag{C3}$$

$$q_3 = \min\left(\frac{1}{\log_5(\max(\Delta h, 1.1))}, 1\right) \tag{C4}$$

$$q_4 = \frac{1}{1 + \frac{d_h}{\max(\Delta h, 0.5 n_{\text{seg}})}} \tag{C5}$$

where $f = n_{\text{peaks}}/n_{\text{seg}}$ is the fraction of sub-segments for which a potential bathymetric peak was detected, $\Delta h = \max_i h_i - \min_i h_i$ is the elevation spread of the detected peaks, and

$$d_h = -\sum_{i=2}^{n_{\text{seg}}-1} \frac{|h_i - h_{i-1}| + |h_{i+1} - h_i|}{2} \max\left(\text{sgn}(h_i - h_{i-1})\,\text{sgn}(h_{i+1} - h_i), 0\right). \tag{C6}$$

Note that the quality heuristics $q_i \in [0, 1] \forall i$, with higher values implying a "better" bathymetric signal. The expressions used here were derived by trial-and-error and designed such that q_1 penalizes major frames with smaller numbers of detected bathymetry peaks, q_2 penalizes major frames with less prominent peaks, q_3 penalizes major frames with a very large overall spread of peak elevations, and q_4 penalizes major frames with peak elevations that do not align along a smooth surface. Based on these quality heuristics for the major frame, we calculate a quality summary $q_s = \prod_{i=1}^{4} q_i$ for each major frame.

Author contributions. **Philipp S. Arndt:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Visualization, Writing – Original Draft, Funding acquisition **Helen A. Fricker:** Supervision, Writing – Review & Editing, Conceptualization, Project administration, Resources, Funding acquisition

Competing interests. The authors declare no competing interests.

Acknowledgements. We would like to thank the ICESat-2 Project Science Office and Science Team, particularly the Land Ice group for helpful input regarding ICESat-2 data and its peculiarities, the OSG Research Computing Facilitation Team for support with implementation of our method on the Open Science Pool, and the National Snow and Ice Data Center for help with data access. The authors were supported through the following grants: Future Investigators in NASA Earth and Space Science and Technology Award #80NSSC20K1666 and NASA ICESat-2 Science Team Awards #80NSSC20K0977 and #80NSSC23K0934. This research was done using services provided by the OSG Consortium (Pordes et al., 2007; Sfiligoi et al., 2009; OSG, 2006, 2015), which is supported by the National Science Foundation awards #2030508 and #1836650. The figures in this publication were produced using Scientific Color Maps (Crameri et al., 2020) where applicable. The writing in a handful of short sections of this manuscript were guided by the AI tools scite.ai and chatGPT.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

References

Arndt, P. and Fricker, H. A.: Towards Automated Generation of Ice-Sheet Wide Supraglacial Meltwater Depth Measurements from ICESat-2 Data, Using High-Throughput Computing, in: AGU Fall Meeting Abstracts, vol. 2022, pp. C35C–0895, 2022.

Arndt, P. and Fricker, H. A.: Source Code for: A Framework for Automated Supraglacial Lake Detection and Depth Retrieval in ICESat-2 Photon Data Across the Greenland and Antarctic Ice Sheets, https://doi.org/10.5281/zenodo.10905941, 2024a.

Arndt, P. and Fricker, H. A.: Data For: A Framework for Automated Supraglacial Lake Detection and Depth Retrieval in ICESat-2 Photon Data Across the Greenland and Antarctic Ice Sheets [Data set], https://doi.org/10.5281/zenodo.10901737, 2024b.

Arndt, P. and Fricker, H. A.: Data and Code for Figure Reproduction: A Framework for Automated Supraglacial Lake Detection and Depth Retrieval in ICESat-2 Photon Data Across the Greenland and Antarctic Ice Sheets, https://doi.org/10.5281/zenodo.10901826, 2024c.

Arthur, J. F., Stokes, C., Jamieson, S. S., Carr, J. R., and Leeson, A. A.: Recent understanding of Antarctic supraglacial lakes using satellite remote sensing, Progress in Physical Geography: Earth and Environment, 44, 837–869, 2020.

Arthur, J. F., Stokes, C. R., Jamieson, S. S., Rachel Carr, J., Leeson, A. A., and Verjans, V.: Large interannual variability in supraglacial lakes around East Antarctica, Nature communications, 13, 1711, 2022.

Aschwanden, A., Bartholomaus, T. C., Brinkerhoff, D. J., and Truffer, M.: Brief communication: A roadmap towards credible projections of ice sheet contribution to sea level, The Cryosphere, 15, 5705–5715, 2021.

Banwell, A. F. and Macayeal, D. R.: Ice-shelf fracture due to viscoelastic flexure stress induced by fill/drain cycles of supraglacial lakes, Antarctic Science, 27, 587–597, 2015.

Banwell, A. F., MacAyeal, D. R., and Sergienko, O. V.: Breakup of the Larsen B Ice Shelf triggered by chain reaction drainage of supraglacial lakes, Geophysical Research Letters, 40, 5872–5876, 2013.

Banwell, A. F., Willis, I. C., Macdonald, G. J., Goodsell, B., and MacAyeal, D. R.: Direct measurements of ice-shelf flexure caused by surface meltwater ponding and drainage, Nature communications, 10, 1–10, 2019.

Bartholomew, I., Nienow, P., Mair, D., Hubbard, A., King, M. A., and Sole, A.: Seasonal evolution of subglacial drainage and acceleration in a Greenland outlet glacier, Nature Geoscience, 3, 408–411, 2010.

Bassis, J., Berg, B., Crawford, A., and Benn, D.: Transition to marine ice cliff instability controlled by ice thickness gradients and velocity, Science, 372, 1342–1344, 2021.

Bassis, J. N. and Walker, C. C.: Upper and lower limits on the stability of calving glaciers from the yield strength envelope of ice, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 468, 913–931, 2012.

Bassis, J. N., Crawford, A., Kachuck, S. B., Benn, D. I., Walker, C., Millstein, J., Duddu, R., Åström, J., Fricker, H., and Luckman, A.: Stability of Ice Shelves and Ice Cliffs in a Changing Climate, Annual Review of Earth and Planetary Sciences, 52, 2024.

Bell, R. E., Chu, W., Kingslake, J., Das, I., Tedesco, M., Tinto, K. J., Zappa, C. J., Frezzotti, M., Boghosian, A., and Lee, W. S.: Antarctic ice shelf potentially stabilized by export of meltwater in surface river, Nature, 544, 344–348, 2017.

Bell, R. E., Banwell, A. F., Trusel, L. D., and Kingslake, J.: Antarctic surface hydrology and impacts on ice-sheet mass balance, Nature Climate Change, 8, 1044–1052, 2018.

Brodskỳ, L., Vilímek, V., Šobr, M., and Kroczek, T.: The Effect of Suspended Particulate Matter on the Supraglacial Lake Depth Retrieval from Optical Data, Remote Sensing, 14, 5988, 2022.

Buck, W. R.: The effect of ice rheology on shelf edge bending, EGUsphere, 2024, 1–17, 2024.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

Buckley, E. M., Farrell, S. L., Herzfeld, U. C., Webster, M. A., Trantow, T., Baney, O. N., Duncan, K. A., Han, H., and Lawson, M.: Observing the evolution of summer melt on multiyear sea ice with ICESat-2 and Sentinel-2, The Cryosphere, 17, 3695–3719, 2023.

Cleveland, W. S.: Robust locally weighted regression and smoothing scatterplots, Journal of the American statistical association, 74, 829–836, 1979.

Crameri, F., Shephard, G. E., and Heron, P. J.: The misuse of colour in science communication, Nature communications, 11, 5444, 2020.

Das, S. B., Joughin, I., Behn, M. D., Howat, I. M., King, M., Lizarralde, D., and Bhatia, M. P.: 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, 2008.

Datta, R. T. and Wouters, B.: Supraglacial lake bathymetry automatically derived from ICESat-2 constraining lake depth estimates from multi-source satellite imagery, The Cryosphere, 15, 5115–5132, 2021.

De Angelis, H. and Skvarca, P.: Glacier surge after ice shelf collapse, Science, 299, 1560–1562, 2003.

DeConto, R. M. and Pollard, D.: Contribution of Antarctica to past and future sea-level rise, Nature, 531, 591–597, 2016.

DeConto, R. M., Pollard, D., Alley, R. B., Velicogna, I., Gasson, E., Gomez, N., Sadai, S., Condron, A., Gilford, D. M., Ashe, E. L., et al.: The Paris Climate Agreement and future sea-level rise from Antarctica, Nature, 593, 83–89, 2021.

Dirscherl, M., Dietz, A. J., Kneisel, C., and Kuenzer, C.: Automated mapping of Antarctic supraglacial lakes using a machine learning approach, Remote Sensing, 12, 1203, 2020.

Druckenmiller, M. L., Moon, T. A., Thoman, R. L., Ballinger, T. J., Berner, L. T., Bernhard, G. H., Bhatt, U. S., Bjerke, J. W., Box, J. E., Brown, R., et al.: BAMS State of the Climate in 2020: The Arctic, Bulletin of the American Meteorological Society, 102, S263–S316, 2021.

Edwards, T. L., Nowicki, S., Marzeion, B., Hock, R., Goelzer, H., Seroussi, H., Jourdain, N. C., Slater, D. A., Turner, F. E., Smith, C. J., et al.: Projected land ice contributions to twenty-first-century sea level rise, Nature, 593, 74–82, 2021.

Fair, Z., Flanner, M., Brunt, K. M., Fricker, H. A., and Gardner, A.: Using ICESat-2 and Operation IceBridge altimetry for supraglacial lake depth retrievals, The Cryosphere, 14, 4253–4263, 2020.

Fan, Y., Ke, C.-Q., and Shen, X.: A new Greenland digital elevation model derived from ICESat-2 during 2018–2019, Earth System Science Data, 14, 781–794, 2022.

Farrell, S., Duncan, K., Buckley, E., Richter-Menge, J., and Li, R.: Mapping sea ice surface topography in high fidelity with ICESat-2, Geophysical Research Letters, 47, e2020GL090 708, 2020.

Fox-Kemper, B.: Ocean, cryosphere and sea level change, in: AGU fall meeting abstracts, vol. 2021, pp. U13B–09, 2021.

Fricker, H. A., Arndt, P., Brunt, K. M., Datta, R. T., Fair, Z., Jasinski, M. F., Kingslake, J., Magruder, L. A., Moussavi, M., Pope, A., et al.: ICESat-2 Meltwater Depth Estimates: Application to Surface Melt on Amery Ice Shelf, East Antarctica, Geophysical Research Letters, 48, e2020GL090 550, 2021.

Gantayat, P., Banwell, A. F., Leeson, A. A., Lea, J. M., Petersen, D., Gourmelen, N., and Fettweis, X.: A new model for supraglacial hydrology evolution and drainage for the Greenland Ice Sheet (SHED v1. 0), Geoscientific Model Development Discussions, 2023, 1–30, 2023.

Garbe, J., Albrecht, T., Levermann, A., Donges, J. F., and Winkelmann, R.: The hysteresis of the Antarctic ice sheet, Nature, 585, 538–544, 2020.

Golledge, N. R.: Long-term projections of sea-level rise from ice sheets, WIREs Climate Change, 11, https://doi.org/10.1002/wcc.634, 2020.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R.: Google Earth Engine: Planetary-scale geospatial analysis for everyone, Remote sensing of Environment, 202, 18–27, 2017.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

Gregory, J. M., George, S. E., and Smith, R. S.: Large and irreversible future decline of the Greenland ice sheet, The Cryosphere, 14, 4299–4322, 2020.

Hanna, E., Topál, D., Box, J. E., Buzzard, S., Christie, F. D., Hvidberg, C., Morlighem, M., De Santis, L., Silvano, A., Colleoni, F., et al.: Short-and long-term variability of the Antarctic and Greenland ice sheets, Nature Reviews Earth & Environment, pp. 1–18, 2024.

Herzfeld, U. C., Trantow, T., Han, H., Buckley, E., Farrell, S. L., and Lawson, M.: Automated Detection and Depth Determination of Melt Ponds on Sea Ice in ICESat-2 ATLAS Data—The Density-Dimension Algorithm for Bifurcating Sea-Ice Reflectors (DDA-bifurcate-seaice), IEEE Transactions on Geoscience and Remote Sensing, 2023.

Hill, E., Gudmundsson, G. H., and Chandler, D.: Regime-shifts in ice-shelf melt could trigger irreversible ice loss from the Antarctic Ice Sheet, Tech. rep., Copernicus Meetings, 2024.

Howat, I., Porter, C., Noh, M.-J., Husby, E., Khuvis, S., Danish, E., Tomko, K., Gardiner, J., Negrete, A., Yadav, B., Klassen, J., Kelleher, C., Cloutier, M., Bakker, J., Enos, J., Arnold, G., Bauer, G., and Morin, P.: The Reference Elevation Model of Antarctica - Mosaics, Version 2, https://doi.org/10.7910/DVN/EBW8UC, 2022.

Howat, I. M., Porter, C., Smith, B. E., Noh, M.-J., and Morin, P.: The reference elevation model of Antarctica, The Cryosphere, 13, 665–674, 2019.

Jasinski, M., SSAI, J. S., SSAI, D. H., Robbins, J., SSAI, J. N., Morison, J., Jones, B., Ondrusek, M., STAR, N., Pavelsky, T., et al.: Algorithm Theoretical Basis Document (ATBD) for Along Track Inland Surface Water Data ATL13 Release 006, https://nsidc.org/sites/default/files/documents/technical-reference/icesat2_atl03_atbd_v006.pdf, 2023.

Kingslake, J., Ely, J. C., Das, I., and Bell, R. E.: Widespread movement of meltwater onto and across Antarctic ice shelves, Nature, 544, 349–352, 2017.

Koenig, L. S., Lampkin, D., Montgomery, L., Hamilton, S., Turrin, J., Joseph, C., Moutsafa, S., Panzer, B., Casey, K., Paden, J. D., et al.: Wintertime storage of water in buried supraglacial lakes across the Greenland Ice Sheet, The Cryosphere, 9, 1333–1342, 2015.

Krawczynski, M., Behn, M., Das, S., and Joughin, I.: Constraints on the lake volume required for hydro-fracture through ice sheets, Geophysical Research Letters, 36, 2009.

Kurtzer, G. M., Sochat, V., and Bauer, M. W.: Singularity: Scientific containers for mobility of compute, PloS one, 12, e0177 459, 2017.

Lai, C.-Y., Kingslake, J., Wearing, M. G., Chen, P.-H. C., Gentine, P., Li, H., Spergel, J. J., and van Wessem, J. M.: Vulnerability of Antarctica's ice shelves to meltwater-driven fracture, Nature, 584, 574–578, 2020.

Lampkin, D. and VanderBerg, J.: 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, 2011.

Leeson, A., Shepherd, A., Briggs, K., Howat, I., Fettweis, X., Morlighem, M., and Rignot, E.: Supraglacial lakes on the Greenland ice sheet advance inland under warming climate, Nature Climate Change, 5, 51–55, 2015.

Leeson, A., Forster, E., Rice, A., Gourmelen, N., and Van Wessem, J.: Evolution of supraglacial lakes on the Larsen B ice shelf in the decades before it collapsed, Geophysical Research Letters, 47, e2019GL085 591, 2020.

Leeuwen, G. v.: The automated retrieval of supraglacial lake depth and extent from ICESat-2 photon clouds leveraging DBSCAN clustering, Master's thesis, Universiteit Utrecht, 2023.

Legleiter, C., Tedesco, M., Smith, L., Behar, A., and Overstreet, B.: Mapping the bathymetry of supraglacial lakes and streams on the Greenland ice sheet using field measurements and high-resolution satellite images, The Cryosphere, 8, 215–228, 2014.

Leppäranta, M., Järvinen, O., and Mattila, O.-P.: Structure and life cycle of supraglacial lakes in Dronning Maud Land, Antarctic Science, 25, 457–467, 2013.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

Levermann, A. and Winkelmann, R.: A simple equation for the melt elevation feedback of ice sheets, The Cryosphere, 10, 1799–1807, 2016.

Li, Y., Gao, H., Jasinski, M. F., Zhang, S., and Stoll, J. D.: Deriving high-resolution reservoir bathymetry from ICESat-2 prototype photon-counting lidar and landsat imagery, IEEE Transactions on Geoscience and Remote Sensing, 57, 7883–7893, 2019.

Lu, X., Hu, Y., Yang, Y., Vaughan, M., Palm, S., Trepte, C., Omar, A., Lucker, P., and Baize, R.: Enabling value added scientific applications of ICESat-2 data with effective removal of afterpulses, Earth and Space Science, 8, e2021EA001 729, 2021.

Luthke, S.: ATL03 Version 3 Known Issues, NASA Goddard Space Flight Center, 2023.

Ma, Y., Xu, N., Liu, Z., Yang, B., Yang, F., Wang, X. H., and Li, S.: Satellite-derived bathymetry using the ICESat-2 lidar and Sentinel-2 imagery datasets, Remote Sensing of Environment, 250, 112 047, 2020.

MacAyeal, D. R., Scambos, T. A., Hulbe, C. L., and Fahnestock, M. A.: Catastrophic ice-shelf break-up by an ice-shelf-fragment-capsize mechanism, Journal of Glaciology, 49, 22–36, 2003.

Magruder, L., Brunt, K., Neumann, T., Klotz, B., and Alonzo, M.: Passive ground-based optical techniques for monitoring the on-orbit ICESat-2 altimeter geolocation and footprint diameter, Earth and Space Science, 8, e2020EA001 414, 2021a.

Magruder, L., Neumann, T., and Kurtz, N.: ICESat-2 Early Mission Synopsis and Observatory Performance, Earth and Space Science, 8, e2020EA001 555, 2021b.

Magruder, L. A., Reese, A. R., Gibbons, A., Dietrich, J. T., and Neumann, T. A.: ICESat-2 Onboard Flight Receiver Algorithms: On-orbit Parameter Updates the Impact on Science Driven Observations, Authorea Preprints, 2024.

Maier, N., Andersen, J. K., Mouginot, J., Gimbert, F., and Gagliardini, O.: Wintertime Supraglacial Lake Drainage Cascade Triggers Large-Scale Ice Flow Response in Greenland, Geophysical Research Letters, 50, e2022GL102 251, 2023.

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., et al.: The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): science requirements, concept, and implementation, Remote sensing of environment, 190, 260–273, 2017.

Martin, D. F., Cornford, S. L., and Payne, A. J.: Millennial-scale vulnerability of the Antarctic ice sheet to regional ice shelf collapse, Geophysical Research Letters, 46, 1467–1475, 2019.

Martino, A., Field, C. T., and Ramos-Izquierdo, L.: ICESat-2/ATLAS instrument linear system impulse response, Authorea Preprints, 2022a.

Martino, A. J., Neumann, T. A., Kurtz, N. T., and McLennan, D.: ICESat-2 mission overview and early performance, in: Sensors, systems, and next-generation satellites XXIII, vol. 11151, pp. 68–77, SPIE, 2019.

Martino, A. J., Bock, M. R., Gosmeyer, C., Field, C., Neumann, T. A., Hancock, D. W., Jones, R., Dabney, P., Webb, C., Lee, J., and Pingel, A.: Ice, Cloud, and Land Elevation Satellite (ICESat-2) Project Algorithm Theoretical Basis Document (ATBD) for ATL02 (Level-1B) Data Product Processing, Version 6, ICESat-2 Project, https://doi.org/10.5067/7LI3JNHLHB6X, 2022b.

Melling, L., Leeson, A., McMillan, M., Maddalena, J., Bowling, J., Glen, E., Sandberg Sørensen, L., Winstrup, M., and Lørup Arildsen, R.: Evaluation of satellite methods for estimating supraglacial lake depth in southwest Greenland, The Cryosphere, 18, 543–558, 2024.

Mobley, C. D.: Handbook of optics: The optical properties of water, Handbook of optics, 1, 43, 1995.

Moon, T., Scambos, T., Abdalati, W., Ahlstrøm, A. P., Bindschadler, R., Gambill, J., Heimbach, P., Hock, R., Langley, K., Miller, I., et al.: Ending a sea of confusion: Insights and opportunities in sea-level change communication, Environment: Science and Policy for Sustainable Development, 62, 4–15, 2020.

Morin, P., Porter, C., Cloutier, M., Howat, I., Noh, M.-J., Willis, M., Bates, B., Willamson, C., and Peterman, K.: ArcticDEM; a publically available, high resolution elevation model of the Arctic, in: Egu general assembly conference abstracts, pp. EPSC2016–8396, 2016.

Mouginot, J. and Rignot, E.: Glacier catchments/basins for the Greenland Ice Sheet, https://doi.org/10.7280/D1WT11, 2019.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

Mouginot, J., Scheuchl, B., and Rignot., E.: MEaSUREs Antarctic Boundaries for IPY 2007-2009 from Satellite Radar, Version 2, https://doi.org/10.5067/AXE4121732AD, 2017.

Moussavi, M., Pope, A., Halberstadt, A. R. W., Trusel, L. D., Cioffi, L., and Abdalati, W.: Antarctic Supraglacial Lake Detection Using Landsat 8 and Sentinel-2 Imagery: Towards Continental Generation of Lake Volumes, Remote Sensing, 12, 134, 2020.

Moussavi, M. S., Abdalati, W., Pope, A., Scambos, T., Tedesco, M., MacFerrin, M., and Grigsby, S.: Derivation and validation of supraglacial lake volumes on the Greenland Ice Sheet from high-resolution satellite imagery, Remote sensing of environment, 183, 294–303, 2016.

Munneke, P. K., Ligtenberg, S. R., Van Den Broeke, M. R., and Vaughan, D. G.: Firn air depletion as a precursor of Antarctic ice-shelf collapse, Journal of Glaciology, 60, 205–214, 2014.

Neumann, T. A., Martino, A. J., Markus, T., Bae, S., Bock, M. R., Brenner, A. C., Brunt, K. M., Cavanaugh, J., Fernandes, S. T., Hancock, D. W., et al.: The Ice, Cloud, and Land Elevation Satellite–2 Mission: A global geolocated photon product derived from the advanced topographic laser altimeter system, Remote sensing of environment, 233, 111 325, 2019.

Neumann, T. A., Brenner, A., Hancock, D., Robbins, J., Gibbons, A., Lee, J., Harbeck, K., Saba, J., Luthcke, S., and Rebold, T.: Ice, Cloud, and Land Elevation Satellite (ICESat-2) Project Algorithm Theoretical Basis Document (ATBD) for Global Geolocated Photons ATL03, Version 6, ICESat-2 Project, 2022.

Neumann, T. A., Brenner, A., D. Hancock, J. R., Gibbons, A., Lee, J., Harbeck, K., Saba, J., Luthcke, S. B., and Rebold., T.: User guide for ATLAS/ICESat-2 L2A Global Geolocated Photon Data, Version 6, https://nsidc.org/sites/default/files/documents/user-guide/atl03-v006-userguide.pdf, https://doi.org/10.5067/ATLAS/ATL03.006, accessed: 2024-04-16, 2023a.

Neumann, T. A., Brenner, A., D. Hancock, J. R., Gibbons, A., Lee, J., Harbeck, K., Saba, J., Luthcke, S. B., and Rebold., T.: ATLAS/ICESat-2 L2A Global Geolocated Photon Data, Version 6, https://doi.org/10.5067/ATLAS/ATL03.006, 2023b.

Nordhaus, W.: Economics of the disintegration of the Greenland ice sheet, Proceedings of the National Academy of Sciences, 116, 12 261–12 269, 2019.

NSIDC: Programmatic Data Access Guide, NASA National Snow and Ice Data Center Distributed Active Archive Center Data Access and Service API, https://nsidc.org/data/user-resources/help-center/programmatic-data-access-guide, accessed: 2024-03-24, Last updated: December 2021, 2021.

OSG: OSPool: Serving Open Science throughput computing, https://doi.org/10.21231/906P-4D78, 2006.

OSG: Open Science Data Federation: Providing data access and transfer services for Open Science, https://doi.org/10.21231/0KVZ-VE57, 2015.

Parizek, B. R. and Alley, R. B.: Implications of increased Greenland surface melt under global-warming scenarios: ice-sheet simulations, Quaternary Science Reviews, 23, 1013–1027, 2004.

Parrish, C. E., Magruder, L. A., Neuenschwander, A. L., Forfinski-Sarkozi, N., Alonzo, M., and Jasinski, M.: Validation of ICESat-2 ATLAS Bathymetry and Analysis of ATLAS's Bathymetric Mapping Performance, Remote Sensing, 11, 1634, 2019.

Pattyn, F. and Morlighem, M.: The uncertain future of the Antarctic Ice Sheet, Science, 367, 1331–1335, 2020.

Philpot, W. D.: Radiative transfer in stratified waters: a single-scattering approximation for irradiance, Applied Optics, 26, 4123–4132, 1987.

Philpot, W. D.: Bathymetric mapping with passive multispectral imagery, Applied optics, 28, 1569–1578, 1989.

Pollard, D., DeConto, R. M., and Alley, R. B.: Potential Antarctic Ice Sheet retreat driven by hydrofracturing and ice cliff failure, Earth and Planetary Science Letters, 412, 112–121, 2015.

Pope, A., Scambos, T. A., Moussavi, M., Tedesco, M., Willis, M., Shean, D., and Grigsby, S.: Estimating supraglacial lake depth in West Greenland using Landsat 8 and comparison with other multispectral methods, The Cryosphere, 10, 15, 2016.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

Pordes, R., Petravick, D., Kramer, B., Olson, D., Livny, M., Roy, A., Avery, P., Blackburn, K., Wenaus, T., Wurthwein, F., et al.: The open science grid, in: Journal of Physics: Conference Series, vol. 78, 2007.

Porter, C., Howat, I., Noh, M.-J., Husby, E., Khuvis, S., Danish, E., Tomko, K., Gardiner, J., Negrete, A., Yadav, B., Klassen, J., Kelleher, C., Cloutier, M., Bakker, J., Enos, J., Arnold, G., Bauer, G., and Morin, P.: ArcticDEM - Mosaics, Version 4.1, https://doi.org/10.7910/DVN/3VDC4W, 2023.

Rignot, E., Casassa, G., Gogineni, P., Krabill, W., Rivera, A., and Thomas, R.: Accelerated ice discharge from the Antarctic Peninsula following the collapse of Larsen B ice shelf, Geophysical research letters, 31, 2004.

Rignot, E., Velicogna, I., van den Broeke, M. R., Monaghan, A., and Lenaerts, J. T.: Acceleration of the contribution of the Greenland and Antarctic ice sheets to sea level rise, Geophysical Research Letters, 38, 2011.

Robel, A. A. and Banwell, A. F.: A speed limit on ice shelf collapse through hydrofracture, Geophysical Research Letters, 46, 12 092–12 100, 2019.

Robel, A. A., Seroussi, H., and Roe, G. H.: Marine ice sheet instability amplifies and skews uncertainty in projections of future sea-level rise, Proceedings of the National Academy of Sciences, 116, 14 887–14 892, 2019.

Rott, H., Abdel Jaber, W., Wuite, J., Scheiblauer, S., Floricioiu, D., Van Wessem, J. M., Nagler, T., Miranda, N., and Van Den Broeke, M. R.: Changing pattern of ice flow and mass balance for glaciers discharging into the Larsen A and B embayments, Antarctic Peninsula, 2011 to 2016, The Cryosphere, 12, 1273–1291, 2018.

Ryan, J. C., Hubbard, A., Box, J. E., Brough, S., Cameron, K., Cook, J. M., Cooper, M., Doyle, S. H., Edwards, A., Holt, T., et al.: Derivation of high spatial resolution albedo from UAV digital imagery: application over the Greenland Ice Sheet, Frontiers in Earth Science, 5, 40, 2017.

Scambos, T., Fricker, H. A., Liu, C.-C., Bohlander, J., Fastook, J., Sargent, A., Massom, R., and Wu, A.-M.: Ice shelf disintegration by plate bending and hydro-fracture: Satellite observations and model results of the 2008 Wilkins ice shelf break-ups, Earth and Planetary Science Letters, 280, 51–60, 2009.

Scambos, T. A., Bohlander, J., Shuman, C. u., and Skvarca, P.: Glacier acceleration and thinning after ice shelf collapse in the Larsen B embayment, Antarctica, Geophysical Research Letters, 31, 2004.

Schröder, L., Neckel, N., Zindler, R., and Humbert, A.: Perennial supraglacial lakes in Northeast Greenland observed by polarimetric SAR, Remote Sensing, 12, 2798, 2020.

Sfiligoi, I., Bradley, D. C., Holzman, B., Mhashilkar, P., Padhi, S., and Wurthwein, F.: The pilot way to grid resources using glideinWMS, in: 2009 WRI World congress on computer science and information engineering, vol. 2, pp. 428–432, IEEE, 2009.

Shen, X., Ke, C.-Q., Fan, Y., and Drolma, L.: A new digital elevation model (DEM) dataset of the entire Antarctic continent derived from ICESat-2, Earth System Science Data, 14, 3075–3089, 2022.

Smith, B., Fricker, H. A., Holschuh, N., Gardner, A. S., Adusumilli, S., Brunt, K. M., Csatho, B., Harbeck, K., Huth, A., Neumann, T., et al.: Land ice height-retrieval algorithm for NASA's ICESat-2 photon-counting laser altimeter, Remote Sensing of Environment, 233, 111 352, 2019.

Smith, B., Fricker, H. A., Gardner, A., Medley, B., Nilsson, J., Paolo, F. S., Holschuh, N., Adusumilli, S., Brunt, K. M., Csathó, B., Harbeck, K., Markus, T., Neumann, T., Siegfried, M. R., and Zwally, H. J.: Pervasive ice sheet mass loss reflects competing ocean and atmosphere processes, Science, 368, 1239–1242, https://doi.org/10.1126/science.aaz5845, 2020.

Sneed, W. and Hamilton, G. S.: Evolution of melt pond volume on the surface of the Greenland Ice Sheet, Geophysical research letters, 34, 2007.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

Stokes, C. R., Sanderson, J. E., Miles, B. W., Jamieson, S. S., and Leeson, A. A.: Widespread distribution of supraglacial lakes around the margin of the East Antarctic Ice Sheet, Scientific reports, 9, 1–14, 2019.

Stone, R. E.: Some average distance results, Transportation Science, 25, 83–90, 1991.

Sutterley, T. C. and Gibbons., A.: pyYAPC: Python interpretation of the NASA Goddard Space Flight Center YAPC ("Yet Another Photon Classifier") algorithm, https://doi.org/10.5281/zenodo.6717591, 2021.

Tedesco, M. and Fettweis, X.: Unprecedented atmospheric conditions (1948–2019) drive the 2019 exceptional melting season over the Greenland ice sheet, The Cryosphere, 14, 1209–1223, 2020.

Tedesco, M. and Steiner, N.: In-situ multispectral and bathymetric measurements over a supraglacial lake in western Greenland using a remotely controlled watercraft, The Cryosphere, 5, 445–452, 2011.

Tedesco, M., Lüthje, M., Steffen, K., Steiner, N., Fettweis, X., Willis, I., Bayou, N., and Banwell, A.: Measurement and modeling of ablation of the bottom of supraglacial lakes in western Greenland, Geophysical Research Letters, 39, 2012.

Tedesco, M., Willis, I. C., Hoffman, M. J., Banwell, A. F., Alexander, P., and Arnold, N. S.: Ice dynamic response to two modes of surface lake drainage on the Greenland ice sheet, Environmental Research Letters, 8, 034 007, 2013.

Tedstone, A. J. and Machguth, H.: Increasing surface runoff from Greenland's firn areas, Nature Climate Change, 12, 672–676, 2022.

The IMBIE Team: Mass balance of the Greenland Ice Sheet from 1992 to 2018, Nature, 579, 233–239, 2020.

Thomas, N., Pertiwi, A. P., Traganos, D., Lagomasino, D., Poursanidis, D., Moreno, S., and Fatoyinbo, L.: Space-Borne Cloud-Native Satellite-Derived Bathymetry (SDB) Models Using ICESat-2 And Sentinel-2, Geophysical Research Letters, 48, e2020GL092 170, 2021.

Tilling, R., Kurtz, N., Bagnardi, M., Petty, A., and Kwok, R.: Detection of melt ponds on Arctic summer sea ice from ICESat-2, Geophysical Research Letters, 47, e2020GL090 644, 2020.

Trusel, L. D., Pan, Z., and Moussavi, M.: Repeated tidally induced hydrofracture of a supraglacial lake at the Amery Ice Shelf grounding zone, Geophysical Research Letters, p. e2021GL095661, 2022.

Tuckett, P., Ely, J., Sole, A., Livingstone, S., Jones, J., Lea, J., and Gilbert, E.: Continent-scale mapping reveals a rise in East Antarctic surface meltwater, Research Square, 2022.

Tuckett, P. A., Ely, J. C., Sole, A. J., Livingstone, S. J., Davison, B. J., Melchior van Wessem, J., and Howard, J.: Rapid accelerations of Antarctic Peninsula outlet glaciers driven by surface melt, Nature Communications, 10, 4311, 2019.

Tuckett, P. A., Ely, J. C., Sole, A. J., Lea, J. M., Livingstone, S. J., Jones, J. M., and van Wessem, J. M.: 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, 2021.

Warner, R. C., Fricker, H. A., Adusumilli, S., Arndt, P., Kingslake, J., and Spergel, J. J.: Rapid formation of an ice doline on Amery Ice Shelf, East Antarctica, Geophysical research letters, 48, e2020GL091 095, 2021.

Xiao, W., Hui, F., Cheng, X., and Liang, Q.: An automated algorithm to retrieve the location and depth of supraglacial lakes from ICESat-2 ATL03 data, Remote Sensing of Environment, 298, 113 730, 2023.

Xu, N., Ma, Y., Zhou, H., Zhang, W., Zhang, Z., and Wang, X. H.: A method to derive bathymetry for dynamic water bodies using ICESat-2 and GSWD data sets, IEEE Geoscience and Remote Sensing Letters, 19, 1–5, 2020.

Yang, G., Martino, A. J., Lu, W., Cavanaugh, J., Bock, M., and Krainak, M. A.: IceSat-2 ATLAS photon-counting receiver: initial on-orbit performance, in: Advanced photon counting techniques XIII, vol. 10978, pp. 48–55, SPIE, 2019a.

Yang, J., Zheng, H., Ma, Y., Zhao, P., Zhou, H., Li, S., and Wang, X. H.: Background noise model of spaceborne photon-counting lidars over oceans and aerosol optical depth retrieval from ICESat-2 noise data, Remote Sensing of Environment, 299, 113 858, 2023.

https://doi.org/10.5194/egusphere-2024-1156
Preprint. Discussion started: 15 May 2024

Yang, K., Smith, L. C., Fettweis, X., Gleason, C. J., Lu, Y., and Li, M.: Surface meltwater runoff on the Greenland ice sheet estimated from
1020 remotely sensed supraglacial lake infilling rate, Remote Sensing of Environment, 234, 111 459, 2019b.

Zhang, W., Yang, K., Smith, L. C., Wang, Y., van As, D., Noël, B., Lu, Y., and Liu, J.: 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, 113 781, 2023.

Zwally, H. J., Abdalati, W., Herring, T., Larson, K., Saba, J., and Steffen, K.: Surface melt-induced acceleration of Greenland ice-sheet flow, Science, 297, 218–222, 2002.