Response to reviewers' comments on manuscript TC-2022-78: High-resolution imaging of supraglacial hydrological features on the Greenland Ice Sheet with NASA's Airborne Topographic Mapper (ATM) instrument suite

Editor: Dr. Louise Sandberg Sørensen

We use the following color and font coding scheme in our response:

Referee's comments

Response: authors' response to comments.

Dear Dr. Sørensen,

Our detailed response to the two anonymous referees' comments (RC1 and RC2) are below. We would like to thank both referees for their constructive and helpful feedback and also for completing their reviews in a timely manner during challenging times.

Best regards,

Michael Studinger

Response to the Referee 1 comments (RC1) manuscript TC-2022-78: High-resolution imaging of supraglacial hydrological features on the Greenland Ice Sheet with NASA's Airborne Topographic Mapper (ATM) instrument suite

We thank the referee for the positive general comments and helpful technical corrections. We have revised our manuscript accordingly.

Technical Corrections

Line 22: insert comma between 'months' and 'seasonal'.

Response: Done.

Line 34: insert comma between 'scale' and 'the network'.

Response: Done.

Line 46: insert comma after 'cryoconite'.

Response: Done.

Line 47: insert comma after 'crevasses'.

Response: Done.

Lines 48-50: suggest rewrite of the sentence beginning 'Yang and Smith'.

Response: We have rephrased this sentence to make it clearer:

Yang and Smith (2013) emphasize the need to gain a better understanding of supraglacial streams, which is limited due to supraglacial streams' relatively narrow width (~1-30 m) making detection difficult using moderate resolution satellite imagery.

Line 50: suggest rewrite to 'However, the recent increase in availability of…'

Response: Done.

Line 51: insert comma after 'streams'.

Response: Done.

Line 52: insert comma after 'wide'.

Response: Done.

Lines 65-67: suggest rewrite of the sentence beginning 'Airborne'.

Response: Done.

Line 74: insert comma after 'strategy'.

Response: Done.

Line 76: insert comma after 'Spring campaigns'. Should 'summer' be capitalised as 'Spring' is, for consistency? Should 'flow' be 'flown'.

Response: Done. We have also capitalized "Winter" in the same paragraph for consistency.

Lines 80-83: suggest rewrite of the sentence beginning 'We analyze'.

Response: We have split this sentence into two sentences to make it easier to read.

Line 93: should 'angle' be 'angles'?

Response: Yes. Done.

Line 114 & 140: Section headings are confusing. Are these sections both 'Methods'? If so, suggest re-label as 3.1 and 3.2. Methods section could then have an introduction as other sections do.

Response: We have modified the headings for Sections 3 and 4 to better reflect the contents of both sections and to make clear that both are descriptions of methods. Section 3 describes an image-based analysis and classification method, and Section 4 is using the lidar waveforms for water depth estimates. We feel this difference justifies having them in separate sections. The new section headings are:

3 Image-based methods: automatic identification of supraglacial hydrological features

4 Lidar waveform-based methods: estimating lake and stream bathymetry

Figure 3: '[ ]' is confusing following '$NDWI_{ice}$'.

Response: $NDWI_{ice}$ is a dimensionless quantity which is typically indicated by empty brackets such as "[ ]". We refer how to handle dimensionless quantities to the Copernicus copy editor if the manuscript gets accepted for publication.

Line 134: replace 'expense' with 'expensive'.

Response: Done.

Line 139: suggest remove 'the purpose of'.

Response: Done.

Line 157: suggest replace 'Over' with 'For' and 'in estimating' with 'to estimate'.

Response: Done;

Line 158: insert comma after 'cases'.

Response: Done.

Lines 161-163: suggest rewrite of sentencing beginning 'For some'.

Response: We have split this sentence into two sentences to make it easier to read.

Line 178: suggest insert 'the' before 'ground test'.

Response: Done.

Lines 257-259: suggest rewrite of the sentence beginning 'The gaps…'.

Response: We have deleted "and", which was misplaced and made this sentence confusing.

Line 266: suggest replace 'reaching maximum' with 'reach the maximum'.

Response: Done.

Line 297: suggest add 'the' before 'recording'.

Response: Done.

Line 300: insert comma after 'lake', before 'there is a'.

Response: Done.

Line 309: replace 'shows' with 'show'.

Response: Done.

Line 312: suggest replace 'part' with 'parts'.

Response: Done.

Line 316: suggest remove 'back'.

Response: Done.

Line 317: suggest insert 'those' between 'as' and 'found'.

Response: Done.

Line 338: suggest insert 'as' between 'acts' and 'a specular'.

Response: Done.

Appendix:

Figure A2: Y-axis labels should contain a unit within '[ ]'.

Response: Normalized signal amplitude is a dimensionless quantity which is typically indicated by empty brackets such as "[ ]". We refer how to handle dimensionless quantities to the Copernicus copy editor if the manuscript gets accepted for publication.

Line 380: In Figure A2 caption the sentence 'For T6 28084 waveforms were stacked and for T7 48703' should be rewritten to be more clear.

Response: Done.

Line 423: suggest insert comma after 'In pressurized aircraft'.

Response: Done.

Line 441: The line 'https://doi.org/10.5281/zenodo.6248436 (Studinger et al., 2022).' appears out of place, with multiple blank lines above it.

Response: This is how the Copernicus MS Word template formats the text. We trust that Copernicus will properly format the manuscript during the typesetting process if the manuscript gets accepted for publication and therefore don't see a need to change this for the manuscript version in review.

Line 457: suggest insert comma after 'altimetry'.

Response: Done.

Line 480: Should 'Sensitivity analysis' be presented as a heading like 'Range bias', 'Gaussian fit' etc.? If so, it should be formatted as sections above and the colon removed.

Response: Agreed. Done.

Response to the Referee 2 comments (RC2) manuscript TC-2022-78: High-resolution imaging of supraglacial hydrological features on the Greenland Ice Sheet with NASA's Airborne Topographic Mapper (ATM) instrument suite

We thank the referee for the positive comments and the very helpful specific suggestions. We have revised our manuscript accordingly.

Major comments:

[1] If I have understood the paper correctly, the primary contribution is methods for retrieving the bathymetry of supraglacial lakes and streams at high resolution from airborne lidar measurements. The introduction provides a strong review of existing methods for estimating lake and stream depth and their limitations. However, the motivation for this study would be strengthened by a little bit more discussion of the science that would be enabled by these more accurate or higher resolution bathymetric measurements. Would this improve our understanding of total water budget and volume of melt impounded in the surface network? Energy balance controls on channel morphology and evolution? Evidence for past prior hydrofracture events in flooded closed basins? A brief discussion of why bathymetry in particular is worth improving would be useful.

> *Response:* The motivation for this study comes from the limitations of existing methods outlined in the introduction. We believe that our statement (lines 351 – 354 and abstract) provides sufficient justification for this study: "The accuracy and resolution from airborne measurements compared to spaceborne sensors provides critical complementary information that can support calibration and validation of spaceborne methods and provide information necessary for high-resolution process studies of the supraglacial hydrological system on the GrIS that currently cannot be achieved from spaceborne observations alone."
>
> We cannot predict how our colleagues will use the method described in the manuscript or what scientific questions they choose to answer, but if the editor determines such statements would be beneficial, we will add a sentence according to the referee's comment. We feel that speculating what our colleagues may do beyond our statement in the manuscript does not belong in a scientific paper. The limitations of existing methods discussed in the introduction are a valid justification in our view to submit this manuscript for consideration to be published in a scientific journal.

[2] The study makes some general statements about "methods to study supraglacial hydrological features" (see for example lines 69-70). The contribution really seems to be to studying the bathymetry of these features and I would encourage the authors to be precise about this throughout the paper.

> *Response:* We have modified the sentence to "…to develop methods that allow the study ***of the bathymetry*** of supraglacial hydrological features…" We have checked the manuscript but did not

find any other general statements regarding the study of hydrological features that would require clarification.

[3] I would encourage the authors to move Figure B3 to the main text early in the paper (perhaps as a new Figure 3), as it would provide a very effective roadmap for the methods and might clarify the overall structure of the paper and the approach for the reader early on.

Response: The reference to Fig. B3 at the end of Section 4 in the main text was unfortunate and may have triggered the referee's comment. We have moved the reference to the flow diagram to the beginning of Section 3 so that readers who are interested in a graphic overview of the processing flow can look at the flow diagram while others can continue reading. We are leaning towards leaving the flow diagram in the Appendix since the manuscript is already heavy on figures (9 in the main section with several multi-panel figures). During manuscript preparation, we have discussed if a flow diagram would even be necessary and therefore believe the Appendix is a more suitable place for this figure than the main section. If the editor determines this figure should be moved to the main section of the manuscript we will do so.

[4] Is it possible to characterize the uncertainty in the final bathymetric estimates based on the uncertainty in the various calibration and correction steps, the SNR, etc? What is the typical error or bias introduced by picking the peaks of the Gaussian fit, rather than the direct waveform peaks? In general, the paper would be strengthened by a dedicated section discussing possible sources of uncertainty in the bathymetric estimates and how they might be identified, mitigated, or quantified. Figure B2 may partially address this question, but it did not get much discussion or explanation in the text.

Response: We respond to the several questions/concerns raised by RC2 in [4] separately:

- **Is it possible to characterize the uncertainty in the final bathymetric estimates based on the uncertainty in the various calibration and correction steps, the SNR, etc?**

The lack of an uncertainty characterization was an oversight, and we agree with the referee that this issue needs to be discussed in the manuscript. We have added a paragraph at the end of section 4 that discussed the uncertainty in our water depth estimates:

"In general, for waveforms with a lake surface and lake bottom return pulse, the uncertainty of the water depth estimate is primarily a function of the uncertainty of the range estimate of the Gaussian tracker used. We use the data from the T6 and T7 ground tests for this campaign (available at https://doi.org/10.5281/zenodo.6248436) to estimate the uncertainty of the Gaussian range estimate to a target with a known distance. The standard deviation of the difference in air between the Gaussian range estimate and the true range to the target is ± 0.7 cm when averaged over the entire range of weak pulses just above the detection threshold through heavily saturated pulses. The total uncertainty of the relative range estimate between the surface and the bottom return is the square root of the sum of the squares of the uncertainties of the two ranges estimates being subtracted and is ± 1 cm. For water depth estimates that only have a lake bottom return, additional uncertainty is added the directly scales with the height of the capillary waves on the lake surface. "

Strictly speaking, we don't know the temperature and density profiles of the water column along the path of the laser beam. The refractive index of water at 532 nm is generally considered to be 1.333 (e.g., George M. Hale and Marvin R. Querry, Optical Constants of Water in the 200-nm to 200-µm Wavelength Region, Appl. Opt. 12, 555-563, 1973). We found no evidence in the literature that the uncertainty of this estimate would be big enough to impact our range estimate in a significant way and therefore consider the uncertainty in propagation velocity of the laser light in water as negligible for the water depth estimate.

- **"What is the typical error or bias introduced by picking the peaks of the Gaussian fit, rather than the direct waveform peaks?"**

Using the maxima from digitized transmit and receive waveforms limits the time-of-flight resolution to ± 1 sample of the digitization interval. The uncertainty of this approach depends on the location of the maximum amplitude between the two neighboring samples around the maximum, which is not known. If waveform saturation occurs this uncertainly can reach several samples. For the ATM system, sampled at 4 GHz, ± 1 sample is equal to 0.25 ns. Geodetic-quality laser altimeters, such as the ATM system, require much higher time-of-flight resolution than ± 1 sample for precise range determination. To achieve sub-sample time and range resolution, centroid-based and Gaussian-fit estimates are commonly used in laser altimetry and are considered superior compared to simplistic maximum-based time-of-flight estimates. We have not estimated what the range improvement of a Gaussian tracker is compared to a maximum-based estimate since the method is widely used in laser altimetry and digital signal processing rather than maximum-based estimates.

- **"In general, the paper would be strengthened by a dedicated section discussing possible sources of uncertainty in the bathymetric estimates and how they might be identified, mitigated, or quantified. Figure B2 may partially address this question, but it did not get much discussion or explanation in the text."**

We have approximately tripled the length of the discussion of the sensitivity analysis in Appendix B and feel this level of detail should stay in the Appendix and only be for readers who are interested in that level of detail. We have also modified the text regarding the sensitivity analysis in Section 4 to clarify our simulations. We have also expanded on the uncertainty characterization in the main section (see above).

[5] The authors argue that errors in the NDWI classification are not problematic, because the lidar analysis will look for surface and bottom returns and so even areas that might be misclassified as water ultimately will not be assigned a depth estimate because there will only be a surface return in those areas. However, the authors later go on to discuss how they estimate the location of the surface return when only a bottom return is visible in the lidar waveform. This raises the question of how they classify as single return as being a bottom return rather than a surface return and whether that classification relies on the assumption that the NDWI classification is accurate. Similarly, I wonder if there would be cases where the lidar waveform might have multiple peaks caused by something other than penetration through water (could complicated crevasse networks or ice damage cause similar returns?). Overall, the paper might benefit from some discussion of when and where (what types of physical environments) the authors would expect the assumptions inherent in their workflow to breakdown. Similarly, it would

be good to mention whether this workflow is applicable to every generation of the ATM data currently published, or if different seasons would require any tuning of the methods.

Response: We respond to the several questions/concerns raised by RC2 in [5] separately:

- **"However, the authors later go on to discuss how they estimate the location of the surface return when only a bottom return is visible in the lidar waveform. This raises the question of how they classify as single return as being a bottom return rather than a surface return and whether that classification relies on the assumption that the NDWI classification is accurate."**

We have modified the text and added wording in Section 3 that expands on the reliability and limitations of identifying single returns as either bottom or surface returns. As stated in the manuscript we are aware that no classification algorithm is perfect and that false detections based on an $NDWI_{ice}$ threshold alone are present in our data (lines 131 to 135) and therefore do not rely in our analysis on all classifications being accurate.

We have modified the text and in Section 3 and added the following wording: "Therefore, false detections based on $NDWI_{ice}$ alone can be eliminated during the lidar analysis, however, they require the presence of a well-defined lake surface from surrounding lidar footprints and cannot be performed on an individual footprint basis."

- **"Similarly, I wonder if there would be cases where the lidar waveform might have multiple peaks caused by something other than penetration through water (could complicated crevasse networks or ice damage cause similar returns?). Overall, the paper might benefit from some discussion of when and where (what types of physical environments) the authors would expect the assumptions inherent in their workflow to breakdown."**

The method presented in our paper is a general method that requires detailed knowledge of the lidar instruments used to produce the data sets that most users don't have. Our intention was to provide interested users with a foundation to develop their own methods that need to be tailored for the specific interests they have in studying the bathymetry of supraglacial hydrological features. There are certainly cases that will cause complex multi-peak waveforms that might be mistaken for lake surface and bottom returns, but as stated above including a well-defined water surface from neighboring lidar footprints and additional lidar parameters, such as increased pulse widths from returns from steep crevasse walls, can be used to tailor the individual user's requirement for scientific analysis. The methods presented here is a research and analysis project. Developing a turn-key ready methods that covers all possible cases is beyond the scope of this paper.

We feel we have already addressed the above-mentioned issue in our response to other concerns from the reviewer.

- **Similarly, it would be good to mention whether this workflow is applicable to every generation of the ATM data currently published, or if different seasons would require any tuning of the methods.**

The described workflow is obviously only applicable for campaigns that have waveform data products available at NSIDC as well as the Version 2 CAMBOT L1B product.

We have added a sentence at the end of Section 2.2 that clarifies that: "The footprint-based surface classification described in the next section requires the CAMBOT L1B, Version 2 imagery. The geolocated and orthorectified images from the Digital Mapping System (DMS), that are available before the CAMBOT Version 2 imagery was added, do not have the required spatial accuracy over ice sheets for reliable footprint-based surface classification."

We have also added a paragraph to Section 3 that describes the possible need for adjusting the $NDWI_{ice}$ threshold for different campaigns. See also our response to "Minor comment (line 119)" below.

<u>Minor Comments:</u>

<u>Line 85</u> – I suppose it does not matter that much, but it is not clear why the fall campaign needs to be discussed since none of the data contributed to the methods development or analysis presented in this paper.

Response: As mentioned in the introduction the temporal and spatial evolution of the supraglacial hydrological system is a topic of scientific interest. We assume most readers would ask the question: why have the authors not analyzed the Fall 2019 campaign for their paper? The answer is we have and in the spirit of scientific transparency we feel it is beneficial to report a negative result. Mentioning this negative finding also partially answers comment [5] from the referee: "Similarly, it would be good to mention whether this workflow is applicable to every generation of the ATM data currently published, or if different seasons would require any tuning of the methods." It shows that there are limitations to which campaigns this method can be applied that are unrelated to data availability.

<u>Line 119</u> – How reliable/consistent is the 0.05 NDWI threshold? How radiometrically stable are the CAMBOT images? Would this threshold require tuning for data from different seasons or flight dates?

Response: We have added a sentence to Section 3 to clarify the $NDWI_{ice}$ threshold selection:

"CAMBOT images are not radiometrically calibrated. CAMBOT is a passive instrument that uses sunlight as the source of illumination. During flight, the camera operator adjusts shutter speed, aperture, and the camera's sensitivity to light (ISO number) to minimize motion blur and optimize exposure over the dynamic range of the camera sensor. The $NDWI_{ice}$ threshold used in this manuscript is an empirical parameter like the 10% threshold of pixels within a CAMBOT frame (line 117 in the original manuscript) exceeding the 0.05 $NDWI_{ice}$ threshold. Both parameters were determined to be suitable for the Spring 2019 campaign and are not absolute thresholds but appear to be stable estimates for the duration of the campaign. It is likely that different campaigns with different light conditions will require adjustment of these thresholds."

<u>Line 210</u> – The flight lines in southern Greenland largely overly areas that are either above the typical visible surface runoff limit or where firn aquifers have been identified (see Miege, et al (2016) "Spatial extent and temporal variability of Greenland firn aquifers detected by ground and airborne radars"), so

it does not seem particularly surprising that there would not be surface lake detections in these areas, particularly in May.

Response: It seems the first part of the first sentence may have triggered the referee's comment. We have modified the first sentence by deleting "Despite the earlier increase in air temperatures in southern Greenland, compared to central west Greenland and presumed longer exposure to surface melt conditions at the transition from Winter to Spring,…", which hopefully addresses the referee's concern.

It is unclear to us what exactly the referee's concern is since the comment basically reiterates our statements from the manuscript (lines 209 – 2018 in initial submission): "Despite the earlier increase in air temperatures in southern Greenland, compared to central west Greenland and presumed longer exposure to surface melt conditions at the transition from Winter to Spring, we found no compelling evidence for liquid water along our flight lines in southern Greenland in the Spring 2019 data. A possible explanation for this is that the ice margin and location of our flight lines are at higher elevations in southern Greenland than they are in central west Greenland and therefore likely exposed to cooler temperatures (Fig. 2). This observation is consistent with the lakes farthest inland being observed in the Sermeq Kujalleq (Jakobshavn Isbræ) area near Ilulissat (Fig. 6) where the ice surface within the Sermeq Kujalleq drainage basin is at lower elevations (Fig. 2) and likely has experienced warmer air temperatures and therefore more melt days than the higher elevations farther south."

<u>Line 355</u> – Do you have a sense for the theoretical penetration depth that should be possible with the instrument, or is that too dependent on variable environmental parameters?

Response: The answer to this question would require exploring a complex set of parameters that can contribute to the theoretical maximum penetration depth that is beyond the scope of this manuscript. Environmental parameters such as the slope of the lake bottom within a laser footprint, reflectance of the lake bottom (cryoconite vs clean ice, see Fig. 1), and reflectance of the water surface (capillary waves vs flat) can all contribute to the maximum water depth that can be resolved. While environmental parameters play a role, instrument settings likely have an even bigger impact. As discussed in the paper the primary goal of the surveys was to determine ice surface elevation and change. During flight, the strength of the neutral density filters is adjusted to avoid detector saturation over bright ice surface targets. If bathymetry measurements were the main purpose, the neutral density filters would be adjusted to maximize penetration while risking saturation over snow and ice targets. Similarly, the trigger threshold for starting waveform recording could be lowered to capture weaker lake bottom returns that would only contribute to undesirable noise for ice surface elevation measurements. Given that the main purpose of all Greenland missions was surface elevation change, we do not have such data.

The last sentence of the manuscript is an honest reflection of our limited knowledge: "…the maximum water depth measured with the ATM optimized for snow and ice elevation measurement was around 7 meters. However, this could also reflect the maximum water depth encountered early in the melt season rather than an instrument limitation."

<u>Line 480</u> – Presumably it is also a function of the relative amplitude difference between the two peaks? Given the shape of a Gaussian pulse, you would need greater separation between two pulses with very

different amplitudes to separate the peaks, compared to two pulses of approximately equal amplitude. Can this be quantified and how much of a concern might it be for bathymetry measurements where there seems to be significant variation in the relative amplitude of the surface and bottom returns?

Response: We have modified the sentence to "…is **PRIMARILY** a function of ~~signal-to-noise ratio and~~ pulse separation…". The referee brings up an interesting aspect that we have considered in our work but have not discussed specifically in the manuscript because we didn't think it to be relevant. The simulations discussed in Fig. B2 demonstrate that pulse separation is the main factor, and our 30 cm minimum depth threshold is a conservative cutoff. Fig. 5c shows that relative differences in pulse amplitude are likely not an issue above the 30 cm depth threshold we use in the manuscript. We have added a sentence to Section 4 to emphasize the ability to resolve surface and bottom returns in overlapping pulses even if the signal amplitudes are very different.

<u>Line 480</u> – An in-text discussion of Figure B2 and how those calculations were carried out would be helpful. I do not fully understand where the 250,000 non-linear regressions come from. Is this from fitting stochastic realizations of the waveform with different noise levels?

Response: We agree with the referee's comment that a more detailed discussion of how the results presented in Fig. B2b were obtained is beneficial. We have clarified the description in the main section of the manuscript (Section 4) and have tripled the length of the description in Appendix B. See also our responses to the comment above and to Major Comment [4] regarding the uncertainty of our depth estimates and the sensitivity analysis.