The Cryosphere Discussions

The Cryosphere Discuss., referee comment RC2
https://doi.org/10.5194/tc-2022-78-RC2, 2022

Comment on tc-2022-78

Anonymous Referee #2

Referee comment on "High-resolution imaging of supraglacial hydrological features on the Greenland Ice Sheet with NASA's Airborne Topographic Mapper (ATM) instrument suite" by Michael Studinger et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2022-78-RC2, 2022

This paper describes as workflow for combining high-resolution optical aerial imagery and airborne lidar measurements to identify supraglacial water features and estimate their depth. It then shows examples of the obtained results over three different water features with different characteristics and discusses the capabilities and limitations of the method in each of these environments. In particular, the authors argue that airborne data can provide a high-resolution data at large spatial scales that fill a gap between spaceborne and field-based measurements.

This work is technically correct to the best of my knowledge and provides a useful for studying supraglacial bathymetry from the existing OIB ATM data record. The authors have published their codes, which makes the method easily reproducible. However, the manuscript lacks a robust quantification of uncertainty in the bathymetric measurements and could benefit from some discussion of where these methods might break down.

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 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.

[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.

[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.

[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.

[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.

Minor Comments:

Line 85 – 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.

Line 119 – 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?

Line 210 – 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.

Line 355 – 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?

Line 480 – 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?

Line 480 – 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?