

This is a thorough presentation of a great dataset that I'm sure will be very beneficial to the community. I have two 'major'/general comments and many minor comments.

Major comments:

Throughout the manuscript, I struggled to understand the contribution that Sentinel-1 provides. Since you choose to include it, I assume it is useful. Figure 4b indicates that it contributes significantly a lot of time and Table 2 and Figure 5 indicate good performance as measured by an F1 score. However, there are many instances in the text where you state or imply that the Sentinel-1 mapping is not useful or reliable. None of the map figures indicate much of Sentinel-1 contribution; where it does contribute, it doesn't seem to agree much with Sentinel-2 (Figure 4b; though acknowledging the images are often on different days within the two-week period). I think this could be resolved by simply providing a figure showing clearly what is mapped as a pond in Sentinel-1 and how it compares with Sentinel-2, ideally also with comparison to some kind of ground-truth/test data.

There are now several datasets of surface meltwater presence in various forms, resolutions and time periods. I appreciated the discussion of the differences and similarities between this dataset and some of those, but I think the manuscript would be significantly strengthened by some quantitative comparison when/where suitable overlaps exist. I realise the definition of melt/ponded water differs in some datasets.

Finally, can you add detail regarding your plans/funding to maintain and update this dataset, unless the intention is for it to remain 'static'.

Dataset:

- 1) There seems to be an errant total_<start_year>_<end_year>_<shelf>.tif in each of the annual mosaic directories, which seems identical to the total_YYYY-mm-dd_YYYY-mm-dd_<shelf>.tif file
- 2) I think the data would be more readily adopted by other users if they were provided also/instead as NetCDF files, so that temporal analysis doesn't require loading in many different files and parsing many filenames to retrieve dates.

Minor comments:

Line 39: Please cite Tuckett et al. (2025) here as well. <https://www.nature.com/articles/s41558-025-02363-5>

Line 48: Although I agree that meltwater drainage to the bed of grounded ice on the AP probably does lead to transient acceleration of grounded ice, I think the evidence for it is still ambiguous, so I recommend rephrasing this sentence accordingly.

Line 49: There are now some field observations indicating acceleration of grounded ice (though only 1 km upstream of the grounding line) driven by surface derived meltwater:
<https://www.nature.com/articles/s41467-026-72724-x>.

Line 55: "supraglacial lake formation is expected to expand further inland and across higher elevations, raising concerns over the long-term stability of Antarctic ice shelves" I'm not sure the second part follows logically from the first part? In any case, I think this too much of a simplification

of the current state of understanding. Please consider adding more detail and nuance to this based on e.g. van Wessem et al (2023): <https://www.nature.com/articles/s41558-022-01577-1>

Line 57: perhaps also add interannual variations and response to changes in the firn/snowpack ability to store/refreeze meltwater? As per Tuckett et al. (2025) and perhaps Dunmire et al. (2024): <https://www.nature.com/articles/s43247-024-01255-4>

Table 1: please add Tuckett et al. (2025) here and to the associated text.

Dataset name – “PolarLakes”: are there plans to expand this dataset beyond Antarctic ice shelves? If not, then perhaps something like IceShelfLakes would be a more suitable name? That might come up more in Google searches too. Up to you of course.

Line 91: “recent supraglacial lake dynamics giving insights on critical hydrological processes”. Can you be more specific here? “dynamics” and “hydrological processes” are quite vague and when it comes to supraglacial lakes readers will likely think of hydrofracture and rapid lake drainage, but since that process typically occurs over hours rather than weeks, I suspect you’re referring to lake growth and slow drainage/refreeze here?

Line 92: I think the dataset will also be very useful to those interested in ice sheet surface mass balance and firn processes.

Line 108: What do you mean by “accessible”? Aren’t they all free online?

Line 121: Consider naming the software/tool used first, as it’s not clear until the end of the paragraph where one can find this function.

Line 121: I suspect it makes no difference, but the type/precision of orbit data used should probably be stated.

Lines 121-125: please provide a few more details on the thermal noise removal, radiometric calibration and speckle filtering algorithms (or citations for each). If the radiometric calibration just used the image annotation data, then there’s no need to provide details on that.

Line 122: Should “naught” be “nought”?

Line 131: AOI hasn’t been defined yet

Line 135/136: I struggled to understand this. What does “cloud masking algorithm within the image acquisition strip of the scene” mean? Is a scene comprised of multiple strips?

Line 143: please specify the original resolution of the SW infrared bands.

Line 150: “the dataset” – training or validation or both?

Line 155: perhaps refer the reader to Section 3.1 for more details

Line 179: As someone who has never used a UNet, I have no idea what “the bottleneck” means.

Line 185: Should that be “factor of 10” rather than 0.1? i.e. the learning rate gets smaller, not bigger.

Line 193: Can you show some of these prediction probability tiles next to the corresponding image, and perhaps a graph showing the deviation from the ‘ground truth’ for a sample of probability levels? This would give the reader another measure of uncertainty in the mapping.

Line 210: should the “observed in January 2017” be cited?

Line 215: can you specify what is “a sufficient amount”?

Line 247: “covered extracted” – is that a typo?

Line 251: in the Sentinel-1 imagery, how do you know what is water and what is not water? From Figure 1 and other Sentinel-1 imagery I’ve seen, ponded water is clearly not always dark. I can see you used Sentinel-1 maximum extent maps for support, but presumably there is a time difference between those maps and the Sentinel-1 images, during which time cold snaps or snowfall could change the lake extent.

Section 3.2: I struggled to understand the rationale behind this confidence flag. It seems to suggest that you don’t assign much confidence to the Sentinel-1 derived lake extents. That’s fair enough, but it doesn’t seem to tally with the results shown in Table 2 and Figure 5, which indicate that Sentinel-1 performs well.

Line 315: should that be “the annual **maximum** extent”?

Line 335: should that be “intra-annual”?

Line 337: typo? “much stronger”

Line 350: I think the “research” you refer to here should be cited (presumably one of the references from the line above?)

Line 375 and elsewhere: Re lakes forming “in depressions”, consider adding “rather than crevasses”, so pedants like me don’t cleverly point out that all lakes must form in depressions whether they have smooth edges or not.

Line 384-385: I can’t see a comparison between ponded area/locations and meltwater production here, so I’m not sure you can support this claim. An alternative explanation is that there was the same amount of melt but a lower capacity in the snowpack to accommodate the meltwater (perhaps because of reduced snowfall that year or because of year-on-year depletion of firn air content).

Line 388: but isn’t that the opposite to what you have found?

Line 394-395: I couldn’t follow this sentence. What is significantly smaller than the fused product? And how can the fused product increase the amount of detections from Sentinel-1, which is what this sentence seems to imply?

Figure 1: is the reader supposed to be able to read the legend in the U-Net Classification panels? I tried to zoom in but still couldn’t at max zoom. Perhaps worth removing that or making a bigger pop-out of those panels? As it is, I don’t think that panel provides any information.

Figure 1: the label “errors in regions that were only ponded any four weeks” is a little hard to parse – the main text description on lines 240-242 is much clearer (though note the typo – “be-weekly”)

Figure 4b: on line 225-226, you state that only a small amount of lake extents are detected with Sentinel-1, but Figure 4 seems to suggest that Sentinel-1 often contributes significantly and sometimes wholly to the lake detection. Though all of the map images provided do indicate a low contribution from Sentinel-1. Please can you clarify this in the text?

Figure 4b: there are quite a few occasions in this timeseries where there is a mix of contributions from Sentinel-1 and Sentinel-2, but with very little or no overlap. From the provided images (e.g. Figure 4a,

Figure 6) I couldn't identify where these differences in mapping occur. Please can you provide an example? This could potentially go in the supplementary information.

Figure 7: This is a nice figure but it isn't really used in the text. There is a passing reference in Section 5. Also, perhaps it's some kind of optical illusion but are there some dots (e.g. 2022-12) with a blue colour intermediate between "medium confidence" and "high confidence"?

Figure 10 and others like it: I find it very hard/impossible to distinguish between the flat colours used for "overlap" and "sentinel-2" and actual meltwater in the satellite image that has not been mapped. So when I look at Figure 10, I can't really see where the differences in mapping are. Are there other ways you could present this information? Perhaps just show the classifications in a separate panel to the satellite image? Or manually map the ponded water and graph/tabulate the differences to that?

Table 2: Can you comment specifically on the reasons for the much lower F1 scores are Larsen C and Hull for Sentinel-2?