

To the reviewers and editor,

We would like to thank the reviewers again for their valuable feedback, which has helped us make several meaningful improvements to both the dataset and the manuscript.

To make the data more transparent, accessible, and reusable, we added shapefiles delineating the areas of interest (AOIs) in which supraglacial lakes were detected, added a README.txt, and included additional TIFF files indicating the sensor contribution to each classification, so that users can clearly see how the results were derived. We also converted the TIFF files to cloud-optimized GeoTIFFs, ensuring easier access and full STAC compatibility for the wider community.

On the manuscript side, we incorporated additional literature references to better position our work within the current state of knowledge, re-phrased several sections for improved clarity, and added new figures in the supplementary material to make our results easier to follow and understand.

We believe these changes directly address the reviewers' comments and have substantially strengthened the quality, transparency, and usability of both the manuscript and the underlying dataset. Please find our answers to your comments in blue below. Line numbers refer to the final manuscript without track changes.

Kind regards,
Celia Baumhoer (in-behalf of all co-authors)

Reviewer 2

Overall Baumhoer et al. present a convincing and competent study of lake extent, which a neatly organised file directory. Using a modified U-Net they advance on some previous lake-detection studies by increasing the frequency of retrieval and the overall year-range. Their goal is straightforward, and they straightforwardly achieve it through well-explained methodology and nice illustrations. Their findings have implications for future lake-sensing studies, and for reassessment of which years had the greatest lake extent in Antarctica.

There are quite a few minor grammatical points and typos below (more may be hidden) but I would be happy to see this published follow these minor corrections.

Having gone through some of the data files I'm content that these show what they are meant to show, though I have of course not been able to quality check every folder. However, I am not really clear what the units in the annual mosaics represent. I looked at the George VI one and it went from 1 - 7, presumably years where it got to that extent? I don't know exactly what ESSD criteria are on this, but it would be good if the dataset was self standing, that is, could be made sense of without the attached paper. In that respect a short readme.txt with the most relevant information (such as units, projection info, etc.) could be a useful addition.

Thank you very much for your feedback. Please find our responses below.

Thank you for the idea of a README.txt. We added the file to the PolarLakes folder describing the basic dataset details:

README

=====

PolarLakes

This dataset provides bi-weekly supraglacial lake extents for 22 Antarctic ice shelves between 2015 to 2024.

For a detailed description of the dataset, please refer to:

Baumhoer, C. A., Koehler, J., Wouters, B., Lhermitte, S., Dietz, A. J., and Kuenzer, C.: PolarLakes: A Bi-Weekly Dataset of Supraglacial Lakes on Antarctic Ice Shelves from Multi-Sensor Satellite Observations (2015-2024), Earth System Science Data Discussions, 1-24, <https://doi.org/10.5194/essd-2025-569>, 2026.

Dataset description

This dataset contains bi-weekly and annual mosaics of supraglacial lake extent, derived per ice shelf ("shelfname"). All raster files are provided in coordinate reference system EPSG:3031 (Antarctic Polar Stereographic) with a pixel size of 10 m.

Folder and file structure

<shelfname>/

Top-level folder named after the ice shelf. Contains the three subfolders described below.

<shelfname>_biweekly_mosaics/

Binary rasters of bi-weekly maximum lake extent.

File naming: yyyy-mm-01_max_extent.tif / yyyy-mm-02_max_extent.tif

yyyy - year

mm - month

01 - first half of the month (days 1-15)

02 - second half of the month (day 16 to end of month)

Pixel values:

0 = no lake / not classified as ponded water

1 = maximum lake extent detected in this bi-weekly period

<shelfname>_biweekly_contribution_mosaics/

Same spatial and temporal structure as

<shelfname>_biweekly_mosaics, but pixel values indicate which satellite sensor(s) contributed to the classification, rather than a binary lake/no-lake mask.

File naming: yyyy-mm-01_max_extent.tif / yyyy-mm-02_max_extent.tif
(same convention as above)

Pixel values:

1 = classified using Sentinel-1 only

2 = classified using Sentinel-2 only

3 = classified using Sentinel-1 and Sentinel-2

<shelfname>_annual_mosaics/

Annual rasters summarizing the bi-weekly frequency of surface ponding during the melt season (November-March).

Each melt season contains at most 12 bi-weekly periods

(12 x ~2 weeks spans November 1 to March 31), so one bi-weekly detection of ponding equals a count of 1.

Two file types are provided per melt season:

`sum_<startyear>-11-01_<endyear>-03-31_<shelfname>.tif`
Per-pixel sum of bi-weekly ponding detections over the melt season indicated by the dates in the file name (November 1 of <startyear> through March 31 of <endyear>). Values range from 0 (never ponded) up to a maximum of 12.

`total_<startyear>-11-01_<endyear>-03-31_<shelfname>.tif`
Per-pixel bi-weekly frequency of ponding, i.e. the number of bi-weekly observation periods available/used for that pixel over the same melt season. Because the observation time frame differs between shelves (see dates in file name), the maximum possible value varies from shelf to shelf.

Medium points

The abstract and title make it seem as though the Ross and Ronne-Filchner ice shelves are included (well, in fact all Antarctic ice shelves). I don't think the title has to be changed, but it would be good for the abstract to include that it's rather a select subset of shelves.

We agree that this is miss leading. Unfortunately, Ross had to be excluded due to limited Sentinel-2 coverage further south and Ronne-Filchner did not experience a large amount of surface ponding. We added to the abstract L17: “To address this gap, we introduce the PolarLakes dataset, generated using a deep learning-based workflow that automates the detection and mapping of supraglacial lakes across 22 Antarctic ice shelves experiencing most frequent ponding”

Methods are presented very well, I appreciate the inclusion of the first methods paragraph.

Thank you, this is nice to read.

In the discussion there is some coverage of how full data is only available from (I think) 2017 onwards. I think a short sentence or two on how to/the importance of replicating the methods could be useful. One could see a need in the future to compare say 2035 melt formation to the heady days of 2020 and keeping methodology uniform would be useful here.

Sentinel-2 data is only available from the melt season 2016-2017. The Sentinel-1 and Sentinel-2 missions are currently continued by ESA and in very far future will probably come with a higher spatial resolution. This means the method can be replicated in the future by using the code we provide and the trained weights (see Dataset Availability section for access). Furthermore, we added a sentence to the conclusion L526: “The PolarLakes dataset provides new insights into supraglacial lake dynamics, facilitated by its longer temporal coverage and higher temporal resolution which can be extended into the future by the provided workflow and trained weights.”

Slush is omitted, when it's inclusion would of course be useful in interpreting lake extent and the slightly blurry line between the two. A bit more discussion about why this wasn't included would be nice (i.e. methodological impossibility, upcoming work, etc.).

We agree that including slush is an important next step. Within this one-year project it was not possible to include it but could be done in the future. We mention this now in the limitations section L461 “Although slush is recognized as an important contributor to surface ponding on

ice shelves (Dell et al., 2024), we were unable to include these regions in the PolarLakes dataset. This omission stems from several factors such as SAR data cannot yet confidently map slush, no labelled training data are currently available for such a segmentation task, and non-linear increase in labelling effort for a three-class segmentation task. We acknowledge that including slush, as well as hydrological channels, represents an important direction for future extensions of the dataset. “

Minor points

Table 1. 'This study' could be added as an additional line to aid comparison. Where the dataset DOI is a dash, a brief comment could be added to say how/if the data is publicly accessible.

Thanks, this is a good idea. We added a row for this study and mention where the Stokes dataset can be found.

Table 2. Please make the scores as % or 0-1 consistent throughout the text.

We consistently report percentages now.

Fig. 1 'identify errors in regions that were only ponded any four weeks' appears grammatically incorrect. Missing a word? Otherwise, this figure is really nice and I appreciate the file extensions at the end.

For clarity we re-phrased to L279: “We followed the stringent approach of Dell et al. (2024) to mask out errors introduced by clouds and cloud shadows by removing pixels classified as surface meltwater for two time steps or less (any two bi-weekly intervals) throughout the entire study period”

Fig. 3. As with the rest of the figures this is really nice, but the train and test sites are hard to spot. Suggest a different colour outline?

We changed the color outline for better visibility.

Fig. 4 Do you have a higher res version of this? It becomes quite blurry quite quickly. There is an errant grey line in the bottom right and the x-axis on b is painful to read. Otherwise really nice, though suggest decimal degrees rather than minutes and seconds for panel a. Usually two y-axis suggest different values are displayed, but both of these go 0-100 suggesting directly comparable? Unclear but if so only one necessary and the comparability info should go in the caption.

We will provide vector image files for publication with higher resolution than the compressed Word version. We changed to decimal degrees and now mention in the caption that also the median cloud cover is given in percent. We would like to keep the two axis as one percentage axis alone wouldn't distinguish between the valued measured (contribution to classification and cloud cover).

Fig. 5. I don't think the 1-1, 1-2, 1-3 notation has been previously introduced ? Also (%) technically needed as x-axis units.

We provide the unit for the F1-Score and added an explanation for the numbers (relates to the specific test set).

Fig. 6 may be blind, but can't see any pink on this figure

We improved the color coding as this issue was also raised by another reviewer. Indeed, in these examples no single S1 contribution is shown. We now mention this in the figure caption, to make it more clear.

Fig. 7 date axes are always hard to get right. There is scope here to simplify things by removing years from the lower x and just having months (as the years are above).

Thanks for raising this point. We would like to keep the year in the x-axis as we use the hydrological year here and it is not as intuitive when stating the month only to which year it belongs.

Fig. 8. I would say this is much more legible than 7, but could be made even more so by removing the year from the labels in the lower x axis, as this is included in the shading at the top, though note that your years here are not correctly aligned.

As mentioned above we report here the hydrological year, hence they are aligned. For clarity, we mention this in the figure caption.

Fig. 9. Is the bi-weekly recurrence in days? Does a bi-weekly recurrence of 15 or above make sense in this context?

Each count refers to one two-week period. We clarify this in the figure caption now by stating L431 “Biweekly supraglacial lake recurrence during the austral summer, where each recurrence count represents lake occurrence over a two-week interval.”

l1. Does EGU have any guidance on capitalised titles? I was under the impression full capitalisation was American.

I couldn't find any guideline on this but I am sure that during the final proof reading stage ESSD staff will correct if necessary so it complies to the journal standard.

l17. The pedant in me notes that if there are no bi-weekly records, there could still be weekly records (which would be better). Suggest stating what the current best is and then saying you do better.

Thank you for pointing out the vague phrasing. We now say L16 “To date, only monthly pan-Antarctic dataset exists that capture supraglacial lake extents until 2021. To address this gap, we introduce the PolarLakes dataset, generated using a deep learning-based workflow that automates the bi-weekly detection and mapping of supraglacial lakes across 22 Antarctic ice shelves experiencing most frequent ponding”.

l30. Good to specify very briefly what these variations are if possible (space, time, etc.).

We point this out more clearly now L31: “The record shows substantial year-to-year variability in ponding and lake formation, with peak extents occurring in different years across the Antarctic Peninsula and East Antarctica. Continent-wide, the maximum extent occurred in 2020 during the observed period.”

l31. To *this* end?

Corrected.

l35. References should go year and then alphabetical, at present they are just placed at random.

Corrected.

l38. And at all in winter??

All studies focus on the summer months. During polar night in winter only SAR data could spot lakes. But lakes will be over frozen and could then only be detected by longer wavelength SAR data (L-Band, P-Band). In return, these wavelengths could also detect firn aquifers and a separation would be necessary here.

l43. Scope here to talk about general ice shelf fracturing in addition to full ice shelf collapse. See for example Banwell et al. (2024, <https://doi.org/10.1017/jog.2024.31>) for more localised and Banwell et al. (2013, <https://doi.org/10.1002/2013GL057694>) for chain reaction drainage.

Thank you for adding this additional perspective. We extended this section to L45: “This can cause meltwater-induced ice shelf fracture (Banwell et al., 2013, 2024) and may ultimately result in the sudden collapse of ice shelves (Lai et al., 2020; Leeson et al., 2020).”

l46. This has been quantified for Greenland, but it isn't actually so important overall (Ryan et al., 2025, <https://doi.org/10.1038/s41467-025-62503-5>).

I'm not entirely sure how best to incorporate this comment. We now also provide a perspective on Greenland L50: “This self-reinforcing melt-albedo feedback enhances solar energy absorption and accelerates surface melt and has been studied in Greenland (Ryan et al., 2025) and Antarctica (Dell et al., 2024).”

l60. I would try to draw more comparisons to Greenland literature, which is more advanced in terms of lake formation. Some citations would be... TODO

l62. The above comment applies to this paragraph, too.

Thank you for this comment. We agree that drawing more comparisons to the Greenland literature, where research on lake formation is more advanced, would strengthen this section. It appears this comment may not have been fully finished at the time of submission (the citation suggestions were left as "TODO"). We would be happy to incorporate any specific references the reviewer had in mind. In the meantime, we have added a few references to the Greenland literature in the introduction at L49.

“Additionally, supraglacial lakes influence the surface energy balance through lowering surface albedo by replacing high reflective snow and ice by water. This self-reinforcing melt-albedo feedback enhances solar energy absorption and accelerates surface melt and has been studied in Greenland (Ryan et al., 2025) and Antarctica (Dell et al., 2024). Supraglacial lake drainage potentially can accelerate grounded ice flow on the AP (Leibrock et al., 2025; Scambos et al., 2004; Tuckett et al., 2019) but it still remains unclear whether similar drainage affects ice dynamics in West and East Antarctica by reaching the bed and altering basal conditions (Arthur et al., 2020). But there is recent evidence from field observations that indicate acceleration of grounded ice driven by surface derived meltwater also in East Antarctica (Sugiyama et al., 2026). In contrast, Sundal et al. (2011) found that in Greenland, once subglacial drainage becomes efficient under sustained high meltwater input, ice flow velocity can be reduced rather than enhanced.”

l65. The Scambos reference would be better suited in the paragraph above.

We shifted the sentence with the Scambos reference in the upper paragraph where it indeed fits better.

l70. I would just like to use the references on this line to emphasise that citation order should be sorted out throughout this paper

Thank you for raising this issue again. We went through the entire manuscript and corrected the citation order where the automated order by the citation software was wrong.

l71. Ration

Corrected to band ratios.

l74. TOA has not been introduced, and is only used once. I suggest just putting top of atmosphere.

Corrected.

l87. Would 'AnarcticLakes' not be a bit more descriptive?

We would like to keep the name PolarLakes as an extension of the Algorithm and dataset generation to Greenland would be a possible extension of the dataset.

l92. How, exactly? I think this is a broadly useful dataset, but I'm generally not a fan of papers saying things can help improve ice sheet modelling unless there is actually a clear and direct link.

Thank you for raising this point which was also mentioned by another reviewer. We explain this further now by saying L107: “The PolarLakes dataset provides high resolution information on recent supraglacial lake dynamics giving insights into slow lake drainage, refreezing and lake expansion. The biweekly resolution of this dataset offers new opportunities for evaluating surface mass balance models against observed melt-pond timing and extent, for constraining firn air content, and for assessing ice-shelf vulnerability to hydrofracture when combined with additional earth observation datasets such as velocity and crevasse mapping.”

l92. Could be rounded out to say, i.e. 'when combined with observations/models of ice shelf collapse and ice sheet velocity, PolarLakes can provide critical insights into...!'

See comment above.

l107. A bit of further information regarding how 'frequent surface melt' is obtained would be appreciated.

We added the reference to Section 2.5.1 where we explain the shelf selection in detail. Extended to in L131: “The PolarLakes dataset covers 22 ice shelves experiencing frequent surface melt (see **Fehler! Verweisquelle konnte nicht gefunden werden.** for ice shelves included in the dataset and Section **Fehler! Verweisquelle konnte nicht gefunden werden.** for shelf selection).”

l121. Good to specify which package the Apply Orbit File function comes from. Ditto with how thermal noise is removed and with the speckle filtering.

As also raised by another reviewer, we shifted the information of using ESA SNAP software to the beginning of the paragraph. I think now it should be clear where the function is coming from.

l127. Ok I see this is introduced here, good to put this earlier.

See comment above.

l131. AOI = area of interest?? Only used once so suggest not using the acronym. Citation shouldn't be fully parenthesised.

Written out now.

l133. If we turn digital numbers into an acronym then nothing is safe from this treatment. Also only used once.

I guess this comment refers to the “Sen2Cor” processor for the atmosphere correction. This is the name of the software I can't change. We put it into quotation marks to show this is a common name ([Sen2Cor v2.12 – STEP](#))

l140. Furthermore, and doesn't make sense

Sorry for the incorrect sentence. We removed “and”.

l150. Good to briefly say what classifies as pronounced surface melt. Also good to know what % of overall data is used as training data.

Pronounced surface melt mostly refers to the maximum extents observed during January. We clarify this now also in the text L177: “For Sentinel-1, the training and validation dataset includes image tiles covering periods of pronounced surface melt occurring mostly during January between 2016 and 2020, focusing on the austral summer months (1 December to 1 March).”

Only a tiny fraction of data was used for training. As we mention, we cover 13 sites which are parts of full Sentinel-1 scenes and for the dataset we processed 20,250 S1 scenes. With machine learning approaches which generalize well this upscaling from little training data is possible.

l166. A brief and parenthetical summary of these rules could be included following this sentence.

We extended the description to 193: “The labeling follows the supraglacial lake labelling ruleset published by Baumhoer and Koehler (2025) to ensure a consistent labeling strategy (mapping consolidated lakes also with thin ice cover, excluding slush, narrow streams and islands).”

l185. suggest 10^x for these numbers for improved legibility.

Corrected.

l193. Generally style guides suggest putting the % directly after a number (as with deg C). Unsure which one esd follows.

We handled this inconsistently and now put the % directly after the number.

l200. The Ross ice shelf is very large, could it have made sense to include a subset of it that did fulfil your requirements?

The biggest issue with the Ross Ice Shelf is that lakes occur along the Transantarctic Mountains. But where the lakes usually occur the Sentinel-2 coverage does not reach far enough South. Therefore, we excluded Ross.

l203. Including this information was one of my previous comments. Good to foreshadow this section by saying (see Section 2.5.1 for site selection...).

Thank you, we added this.

l210. Great, this covers a few of my earlier comments, but as above, good to foreshadow this section, or even move it all up into methods.

We refer to this selection now in the method section.

l241. be-weekly

Corrected.

l259. %s

Corrected as mentioned above.

l260. This indicates that,

Corrected.

267. The pedant within me dislikes that here scores are given 0-1 and above in %. Please choose wither or.

We made this consistent by always using percent.

l310. Nice

Thank you.

l337. more strongly

Corrected.

l347. Not clear what 'less significant' means in this context. If you're finding 2020 as a larger melt year than 2017 using your methodology that could be something worth mentioning in the abstract.

Sorry for the incorrect wording. Changed to “less pronounced”.

l356. ,,

Corrected.

l357. and as such these are not included

Corrected.

l360. Is there a reason you didn't include slush as well as lakes? May have helped with some uncertainty statements.

We agree that including slush would have been very valuable. As we already explained to Reviewer 1, unfortunately this was not possible due to time constraints of the project: “We agree that including slush is an important next step. Within this one-year project it was not possible to include it but could be done in the future. We mention this now in the limitations section” and added the following to the manuscript L461 “Although slush is recognized as an important contributor to surface ponding on ice shelves (Dell et al., 2024), we were unable to include these regions in the PolarLakes dataset. This omission stems from several factors such as SAR data cannot yet confidently map slush, no labelled training data are currently available for such a segmentation task, and non-linear increase in labelling effort for a three-class segmentation task. We acknowledge that including slush, as well as hydrological channels, represents an important direction for future extensions of the dataset. ”

l391. This is a good and thorough limitation section, but foreshadowing could be peppered throughout the manuscript. For example, the issues with consistent time series data is worth mentioning earlier when the data sets are introduced.

We mention the data availability and time series issue now right in the beginning at the end of the introduction L104: “...(please note that, depending on the region, Sentinel-2 data are available from either 2016 or 2017 onward making the time series only consistent after this point in time).”

l406 As in Dell et al.... ? and how (rather than just giving the citation)

We elaborate further on this point now L471: “To address this problem, we limited our analysis to regions that were detected more than two times throughout the entire time series as done in other studies such as Dell et al. (2024). From the total recurrence raster over the entire time period since 2015 all pixels ponded only once or twice on a bi-weekly interval were removed.”