Review of "The Greenland Ice-Marginal Lake Inventory Series from 2016 to 2023" (essd-2025-18) by Penelope How and others

Summary
The authors present a time-varying inventory of ice-marginal lakes across Greenland that extends a past static 2017 inventory. The authors rely on this earlier work heavily for the present methods, but the methods are generally well described. However, the authors treatment of icebergs in the lake was unclear in the current version, which could have significant implications for the interpretation of lake surface temperature and area results described here. While I described several "major" issues below and numerous minor points, I consider my comments to reflect a suggestion of "minor revisions", as everything can be addressed by text better describing methods, clearly acknowledging potential issues/data artifacts, and a a few new analyses of the existing inventory (rather than anything requiring revision of the inventory itself).

Major comments
Sec 3.2: from this section and inspection of Figure 5, it doesn't seem like you are masking out icebergs from your water temperature estimates, is that right? If I'm wrong, please try to more clearly state this. If you aren't masking icebergs, how much of an effect could this have on your temperature estimates?

L243: Do you have any idea of the prevalence of false positives vs false negatives? If I understand this sentence correctly, you are taking data from a false positive rate (i.e., automated method says there's a lake, but there's not) and are then using this value as a uniform ± error. But do you know if the method misses lakes just as often as it makes them up? It seems like this could lead to substantially different errors on the positive vs. negative sides.

Also, it would be good to have an estimate of error in lake area and its change. I imagine the error in lake area is much smaller (in % terms) than the error in number that you discuss here. Presenting both of these errors could allow you to say, "while absolute number is somewhat uncertain, uncertainty in lake area change is small (if true), suggesting that the lake number error is primarily attributable to varying detection of small lakes". That would help bolster the utility of this dataset, which I imagine could be undermined by only reporting the ±36% error in lake number.

Figure 3: For bottom plots > the average lake area is far noisier than I'd expect given the many lakes being used to compute these averages. What do you attribute these large changes to? To me, the first thing that comes to mind is data processing artifacts like large lakes being split into multiple pieces in some years. Without a commentary on what underlies the high variability, it is hard to know how much to trust it as a physically meaningful value vs. data noise.

Figure 5: From looking at this image, it doesn't seem like you're doing any hole filling to remove icebergs floating in the lake? This seems like it could impose substantial variability in a lake's area following large calving events > can you comment on this somewhere?

Figure 6: Given the high noise level here, some measure of uncertainty/variation (perhaps interquartile range) would be helpful for ascertaining whether the plotted changed in the mean are due to real variation vs. noise. To me, this data seems like it would be better shown in box plot (or violin plot) form, so we could get some sense of the distribution of data, which is not easily grasped at present (aside from seeing that there is a lot of variation from lake to lake).

Minor comments

L39: The Shugar paper doing global ice-marginal lake mapping made an estimate for meltwater retention in lakes > it should probably be mentioned even if there are reasons why the estimate is imperfect (that can also be mentioned)

L73: Is this the GLO90 DEM, or what static DEM are you referring to?

Sec 2.2: I am not sure if this section is required by ESSD, but I personally didn't get much out of it (think most will be repeated later in more detail?) and think the whole thing could be deleted. Otherwise, the utility/value of this section should be made clearer.

Table 1: I suspect you mean 20 and 30% cloud cover limit across the whole image, not on some kind of pixel-wise basis > is that correct? It could be worth spelling this out

L129: It is not clear to me what you mean by "where positive classifications adhere to all thresholds" > please reword or clarify

Table 3: It would be nice to include the names of each band (e.g., "B2 = blue") in the caption for people less used to working with Sentinel data. I think maybe you are saying that it is considered water if it's less than that threshold value?

L168-178: Do you have a sense how much of this variation in lake number is due to processing artifacts (e.g., one lake classified as two in some years due to data issues) as opposed to physically meaningful variations? It seems like you could do some kind of intersection/spatial join to test this.

L195: It is interesting that many more lakes shrunk in size > this seems at adds with what is seen in many areas. Are you doing this analysis on a last yea r- first year basis? This would make your results really sensitive to noise in those years. Have you tried doing a linear fit to all lake area data at a site and making this threshold based off a rate of change rather than an absolute change? That

seems like it would be more resistant to noise. Regardless, do you have any physical interpretation of why shrinking lakes are more common? Maybe this comes later.

Sec 4.3 (related to Figure 6 comment): please discuss if icebergs are masked during temperature estimates. If not, it is unclear what these data mean. Also, have you analyzed lake temperature trends (or year-to-year temperature variability) on a lake-by-lake basis? This seems like it would add a lot to this section. As is, it is a little unclear what the lake surface temperature adds to the story here.

L229: I imagine this is described in more detail in How 2021, but it is unclear in the current manuscript how the different delineation methods are incorporated. How do you blend the datasets when they have inevitably somewhat differing shapes? In general, how do you choose whether just one or multiple methods are used to delineate a given lake?

Sec 6.1: I think this is all true, but much applies to a time static lake inventory, so it might be useful to better articulate what having the time variation adds here.