Dear Referee,

Thank you for taking the time to review our study and provide constructive feedback, suggestions which will improve the quality and understanding of the paper. Below we provide detailed responses to each of your comments, with our responses in blue.

This study provides insight into the evolution of ice-marginal lakes in the Alaska and NW Canada between 1984 and 2019 by employing supervised classification and semi-automated lake area delineation from Landsat images. The authors present novel findings and the text reads well. I find this study thematically suitable and of potential interest for the readers of The Cryosphere.

I have four more general (methodology-related) comments and a couple of specific ones.

The first general comment is related to inventory building. Using a semi-automated classification, you could possibly be missing some lakes identified as other features (false negatives); while you eliminated possible false positives by manual assignment of qualitative characteristics, this won't help you dealing with false negatives in a systematic way. Optimally, mapping outcomes of semi-automated classification would be checked against existing (e.g. sub-regional) inventory (you mentioned some in the intro), or manually prepared subset (e.g. manual mapping of 100 lakes to see the performance of semi-automated approach in terms of possible false negatives).

We agree that false negatives could be an issue within our dataset, though we minimize the likelihood by using more inclusive thresholds, then removing polygons manually. To add clarity in the manuscript, we will add a supplemental figure to illustrate what polygons look like after classification, after thresholding, and after manual cleaning to better explain the process. We also agree that checking outlines against an existing inventory would be useful, however, due to differences in methodology, criteria, and time period, comparison would introduce new uncertainties. Apart from the published study, we did compare our lakes to the dataset published by Field et al. (2021; https://doi.org/10.5194/tc-15-3255-2021), and found that 78 of the 85 manually delineated lakes were identified in our inventory. For the seven lakes not included, two are ice-dammed lakes (Summit Lake and Snow Lake) which are known to spend a significant amount of time drained or ice-filled, and therefore unidentifiable in our imagery. The other missed lakes are small, variable, marginal lakes which are likely missed due to differences in imagery dates (they are not visible in our mosaics). We will provide more explicit information and examples in the text about the types of lakes our inventory misses (i.e. drained, ice-filled). We decided not to include these types of lakes, even if we know their location, to avoid mixing methods and because although we know they exist, we can not provide an outline (since they don't appear in our Landsat mosaics) and therefore cannot assess change using our dataset. We aim to create a robust and repeatable dataset. Thus, our inventory presents a clear minimum in the number and location of lakes, with known false negatives which we will describe in the text.

The second comment is related to possibly missed fill-drain events (outbursts) typical for ice-dammed lakes (briefly mentioned on L104-107). This is actually quite important issue in my opinion (especially for formulating outburst hazard implications); I'm wondering whether any insight can be gained from histograms of pixel values used for mosaicking (for instance if two peaks of values in bare land and water spectra can indicate there was a lake outburst)? Please provide more discussion on this issue of possibly short-lived lakes (maybe a separate discussion section)

We agree that ice-dammed lakes that experience frequent drainage are very important for hazard implications, however, for this study, we focus on documenting decadal-scale patterns in lake area changes. Our dataset uses 5-year mosaics to detect a generalized outline for that period, which is unfortunately not sufficient to detect drainage events occurring on shorter timescales (often every 1-3 years). This is outside the scope of this study and will be addressed in a subsequent study. To further address this point, we will modify our introduction so that it is clear we are looking at decadal scale changes rather than rapid drainage events. As mentioned in our response above, we will also describe in greater detail why some lakes which are known to drain and fill are not included in our study.

The third one is related to dam type classification scheme. Your classes (Section 2.3.1) are defined in a clear, straightforward way. However, my experience from Peru is that I've been often facing cases where assignment to one of the classes was not at all straightforward in reality. For instance, I frequently observed lakes dammed by bedrock dam with discontinuous moraine cover (I ended up classifying these lakes as lakes as 'combined dams'). Sometimes, it was not possible to assign a dam to any of the classes, e.g. because of low quality / poor resolution of satellite imagery (and so I introduced 'not specified' dam category). Moreover, lake dam type can change in time (e.g. a shift from ice-dammed to bedrock-dammed is not rare). I'm also wondering whether you have observed any possibly landslide-dammed lakes in your inventory? Please comment on / discuss whether you've been facing similar issues when manually classifying lake dam types.

Thank you for pointing this out, as you have a unique appreciation for the difficulty in determining dam type. It is true that some dams are quite difficult to determine, and we classified them using our best judgement. We will add an additional paragraph within our discussion discussing the challenges in determining dam type, and emphasizing that our classification is our best interpretation, though it is limited by possible mixed dam types and poor imagery resolution. We will revise the explanation of Figure 2 as we wish to display typical lake behavior of each dam type, not necessarily provide an example of how we determined the dam type.

We do document a shift in dam type for some lakes, particularly from ice-dammed to bedrock-dammed (Figure R1.1). As each lake from each time step has its own polygon and dam type classification, this information is available within the dataset. However, for change over time, only one dam type can be classified for each individual lake, and therefore we use the most recent dam type for analysis. We do not observe any landslide-dammed lakes within our inventory. They likely do exist in Alaska, but our inventory is limited to lakes within 1 km of the RGI.

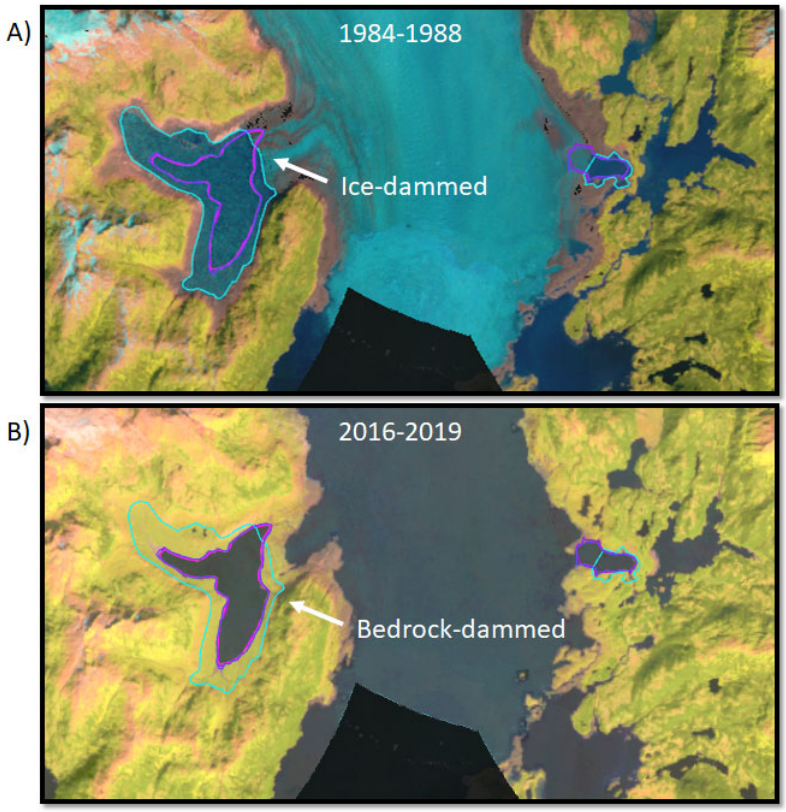

Figure R1.1. Example in shift from ice-dammed lake in 1984–1988 (A) to bedrock-dammed lake in 2016–2019 (B) for Terentiev Lake, which was dammed by the Columbia Glacier.

It is not fully clear how disappeared lakes (and there are many in fact) are considered and treated in statistics of total lake area change (e.g. Tab. 2), see also my specific comments; please provide more methodological details on that

Disappeared lakes are considered to have an area of zero km^2 after they disappear. We will include more explicit details in the revised manuscript about which lakes are included in which statistics (all inventoried lakes vs lakes only present in the most recent time period).

- - -

Specific comments:

L33: GLOF may also result from dam overtopping; dam breach is a sub-type (one of mechanisms) of dam failure in my understanding

Thank you-- we will reword to say “when a lake dam suddenly fails or is overtopped”.

L56: can supraglacial lake also be located on debris-free glacier?

We only identified supraglacial lakes which are located within debris on debris-covered glaciers. We will be more explicit about this classification.

L85: you may consider confronting results of these previous Alaska-focusing studies in separate discussion section

We will add a discussion section focusing on Alaska-specific studies, both comparison to region-wide studies, as well as studies of a few specific lakes, as was also recommended by the second reviewer.

L94: A separate figure (workflow) depicting individual steps of the procedure, input and output data would be beneficial for readers

Thank you for this suggestion – we will add a workflow figure to clearly show what steps we used.

L112: can 'supraglacial debris' located on a glacier be distinguished from 'just debris' located elsewhere based on the spectral profile?

We did note spectral differences between supraglacial debris and 'just debris' (Figure R1.2), likely due to supraglacial debris generally being wetter and colder than off-glacier debris. Supraglacial debris often appeared as false positives after initial classification, as it had a spectral profile with more similarities to water, particularly in the NIR and SWIR bands (likely due to surficial melt associated with debris cover, dependent on debris thickness).

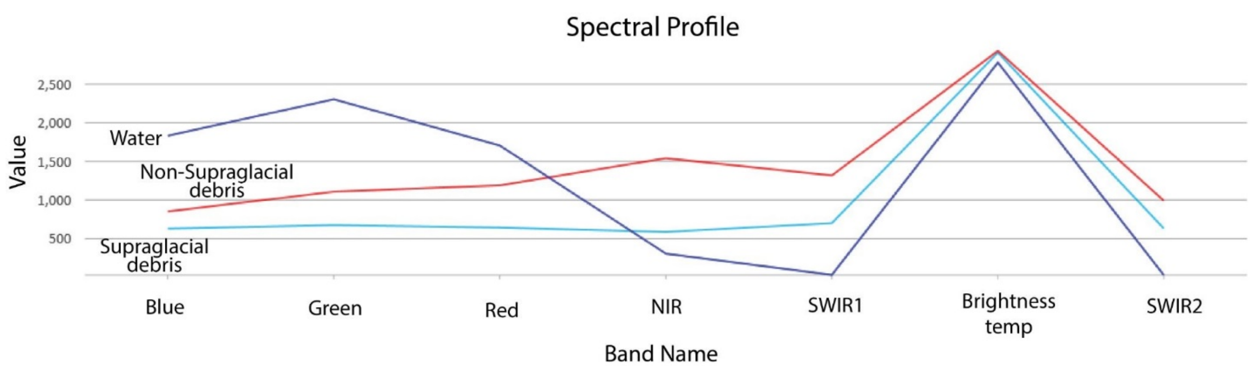

Figure R1.2. Examples of spectral profiles for water (dark blue), non-supraglacial debris (red), and supraglacial debris (light blue).

L117: pixel is areal unit (doesn't need to be squared)

This has been fixed – we were trying to describe a 7.5 by 7.5 pixel area, but we will just say an ~ 55 pixel area instead -- thank you.

L119: maybe you could specify date of images RGI is based on in your study area

The source imagery for the RGI in Region 01 is mostly from 2004-2010 (Kienholz et al., 2015). This information will be added to the description.

Fig. 2: These examples with false-color images are not very illustrative in terms of distinguishing different lake dam types; (e-f) instead of (d-f)

Thank you – here we aim to provide examples of typical lake behavior for each dam type, rather than examples of how we determined the dam type. We will fix our text to reflect this, and will correct the caption to (e-f).

Fig. 4: part a: you show drainages of lakes for individual periods (e.g. 11 drainages of moraine-dammed lakes in 1984-1988) – does it mean that you actually have insights into the within-period lake dynamics (and it is not blurred by mosaicking as described in methodology)? Please clarify

The example of 11 lakes in 1984-1988 is supposed to represent 11 lakes that are present in this time period which drain in a subsequent time period. We understand how this can be confusing and will add text to clarify that this does not represent within-period dynamics, but rather changes between periods.

L208: please unify Number of lakes (e.g. Fig. 4a) and frequency (e.g. Fig 5a) or explain the difference

These two represent the same thing (the number of lakes within each time period or bin), and the axes will be changed to reflect that.

L214: how did you actually deal with possibly merging lakes? Have you observed any such a case? Please comment

Lakes which merge or split at some point in time are given the same Lake ID, so that we do not get a false signal of a lake disappearing when the lakes merge or appearing when the lakes split. This happens most frequently for supraglacial lakes.

L224: 130 disappearing lakes from 791 total lakes is quite high number; if these were GLOFs, you observe 16.4 GLOFs per 100 lakes, which is extremely high ratio

The majority (103/130, 79%) of the lakes which disappeared are supraglacial lakes, many of which likely did not produce a GLOF. We will add a section in the discussion addressing the high number of disappearing likes (mainly due to the high variability of supraglacial lakes). Excluding supraglacial lakes, the new ratio would be 3.4 GLOFs (or rather drainage events, as we don't know how each lake drained) per 100 lakes.

L234-255: I suggest to start with % of lakes which actually experienced change and describe them in more detail in this chapter; taking into account lakes which did not experience areal change is confusing (and resulting in median change of 0.00 km^2 which is not very useful insight in my opinion)

Thank you for this suggestion. We will focus the results mainly on lakes which experienced change, and be explicit about which part of the dataset we are describing in each section.

Tab. 2: isn't this statistics biased when disappeared lakes are not considered – I mean, If you would consider 791 lakes instead of 661 lakes in this table, the overall pattern of lake area would be different I guess (count disappeared lakes as lake area decrease); further, I suggest to mention also min and max values, so the reader can get an idea about the range of observed values (median is ok, but I'm also interested in extremes); please consider re-designing this table

Thank you for this suggestion. As 103 of the 130 lakes which drained are supraglacial, the main change in statistics would be for the supraglacial lakes. Min and max do seem like valuable numbers to include and will be added to the table.

L257-258: you mentioned that most of the lakes did not experienced detectable change – how can then median change on subregional level when considering all lakes be 0.04-0.06 km^2 (I would expect 0.00 as well km^2)?

Thank you for catching this discrepancy. These are the median changes for lakes with detectable change. As lakes which experienced change are dominated by moraine-dammed lakes, these have the largest influence on the subregional area change. This will be clarified in the text.

Tab. 3: an interesting indicator could be lake area per deglaciated area

This is a very interesting suggestion. However, we are not aware of any existing dataset documenting deglaciated area within each subregion and producing such a dataset would be outside the scope of this study.

Fig. 7: please consider plotting relative cumulative lake area against relative cumulative lake count (that could provide clear insights what % of the largest lakes (count) represent what % of total area); analogically to Lorenz curve

Thank you for this suggestion. The second reviewer also suggested separating out the large lakes (>10 km^2) for a subanalysis to more clearly demonstrate how large lakes contribute the largest change in area. The plot that you suggest would be beneficial, although we think it would be easier to implement as a replacement to Figure 5. We feel that adding this type of plot for Figure 7 would be difficult since we are showing change over time, and have limited space on the map.

L281: some part of discussion are more results (e.g. section 4.3)

We will relocate some of the temporal trends discussion section to the results.

Fig. 8: please specify how many lakes are plotted in this figure

We will add an "n=" for each of the subplots.

L315: 'loss of an ice for'?

This reads "loss of an ice dam for", trying to explain that ice dam loss was attributed to ice thinning for 82% and 62% of land-terminating and lake-terminated glaciers, respectively.

Fig. 10: captions on x axis are confusing (if this is a change rate between two periods, both should be included, e.g. (1984-1988 to 1997-2001); (1997-2001 to 2007-2011); and (2007-2011 to 2016-2019), or similar)

We will make these changes to the x axis in Fig. 10 for clarity, as we are displaying the change rate between two periods.

L395-400: I think that important control of possible transferability of observed evolutionary patterns is topographical (relief) similarity (shape of a hypsometric curve of a mountain range); please consider taking this aspect into discussion

Thank you for this interesting suggestion. Transferability to other mountain ranges is an important aspect of the discussion, but as this would require the addition of new analyses and methods, we feel it is outside the scope of our current study.

L414: please replace 'basins' by 'parts of the study area'

This has been changed, thank you.

L459: second?

Yes, second, thank you for catching this mistake.

- - -

To sum up, I'm convinced this is valuable contribution to our understanding to dynamics of lake evolution in deglaciating mountain landscapes of Alaska and NW Canada. This study is undoubtedly worthy publishing as soon as some revisions are made. I suggest moderate to major revisions (especially methodological issues should be clarified, see my general comments).

Kind regards

Adam Emmer (Uni Graz, Austria)

Thank you again for your thorough and insightful review. We appreciate the time and effort you've invested to help improve our manuscript.