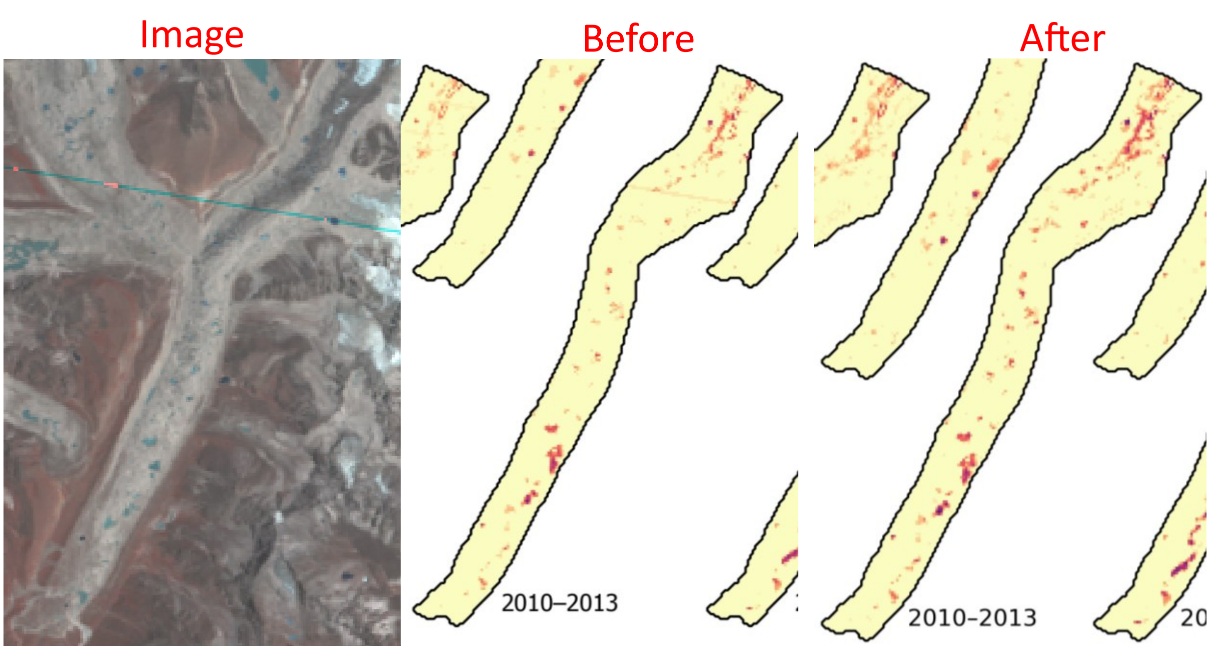
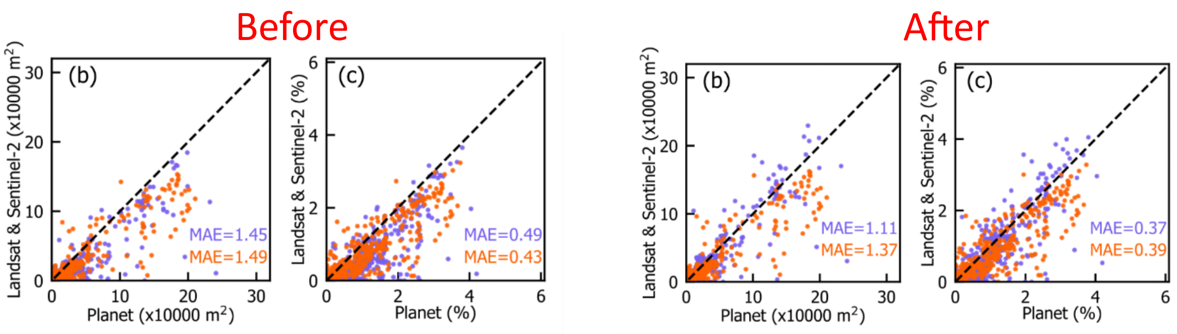
Dear Referee,

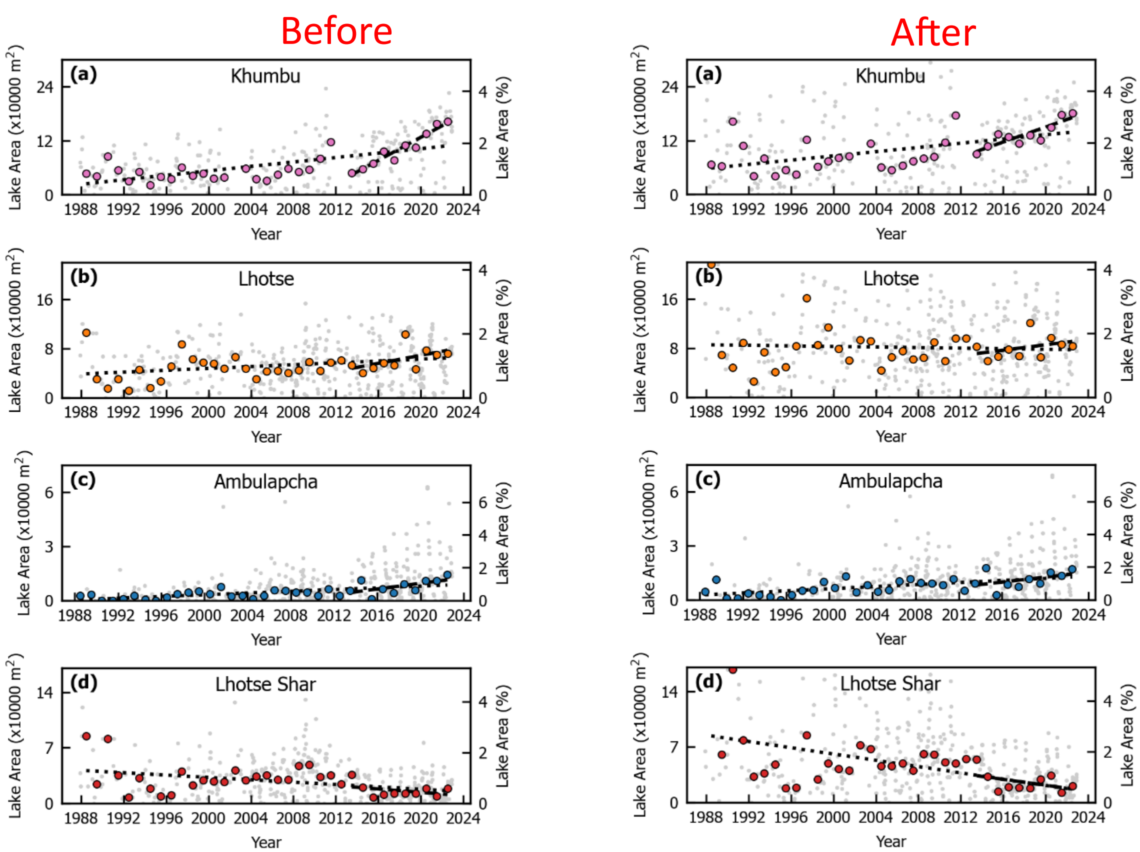
Thank you for taking the time to review our manuscript and to provide constructive feedback. We feel that the changes made in response will substantially improve the quality and value of the paper. We provide detailed responses to each of your comments below, with our responses in blue.

While formatting our responses, we found a small error in the code used to compute the optimized NDWI threshold for Landsat and Sentinel-2 imagery. The green-band shadow masking applied to Landsat and Sentinel-2 images (lines 164-165 in the initial submission) was mistakenly not applied to the training/validation dataset prior to computing the optimized NDWI threshold for Landsat and Sentinel-2 images. After fixing this omission, the optimized thresholds found are changed slightly (due to the low-brightness shadowed areas being commonly mis-classified as water without this shadow masking applied). These new NDWI thresholds are: 0.137 (Landsat 5 & 7), 0.188 (Landsat 8 & 9), and 0.250 (Sentinel-2), compared to the original values of 0.172, 0.226, and 0.260. Additionally, we introduced an additional filtering step to remove occasional linear artifacts in a handful of Landsat 5 images (see the example below). This filtering step removes pixels where the difference in surface reflectance between green and near-infrared is greater than 0.2. his threshold was chosen to remove the majority of these artifacts while leaving no discernible difference in true water identification.



All analyses and figures which used Landsat and Sentinel-2 data have been updated to reflect these changes. Items which have changed are: Figures 4, 6, 10, 11, and 12 (based on figure numbering in the original manuscript submission), Supplementary Figures S12-S19, and Tables 1, S3, and S4. The figures below (Figure 6 c-d and Figure 11 a-d) provide an example of the changes which have occurred.





This study brings together a range of satellite image sources to characterise the dynamics of surface lakes (or ponds) on debris-covered glaciers in the Everest-region of Nepal. The analysis is neatly divided between looking at the seasonal dynamics of these ponds using daily Planet imagery, and the long-term (decadal) patterns captured in the Sentinel and Landsat archives. The result is a considerable dataset that builds on previous work on the same glaciers and adds an element of detail with the availability of the finer-resolution imagery. I support its publication, but I do think more could be made of the analysis/data presentation so that future studies, which look to build on this one in coming years, can readily use it as a baseline for comparison.
In particular, the authors should consider:

1. Adding summary statistics to the main text (or, if you prefer, in a table). What are the mean, max and min lake areas for example, per glacier per year, as well as mean, max and min number of lakes? Which glacier hosts the most lakes/greatest area? Is that the same every year? How many lakes are ephemeral vs permanent? How much of the overall lake area do the ephemeral lakes account for (and therefore how important are they, relatively speaking?). These sorts of stats help the reader to interpret the patterns you talk about in the text in general terms, as well as providing concrete values for future studies to use as comparison.

We appreciate this suggestion, which was similarly brought up by the second reviewer. We agree that these are important findings to highlight in the context of this project. We have elaborated on these concepts by including a table in the main text that provides information of the number of lakes, lake area, and number of permanent lakes for each glaciers using the PlanetScope product. Additionally, supplementary tables provide this information broken down by individual years. We have expanded the Results section to highlight some of these findings, and have included further discussion of them at various points in the Discussion section.

2. Providing more information on the life-cycle of these smaller lakes that appear to be responsible for the seasonal patterns you show. For example, and since you have already gone to the trouble of correcting for surface displacement, can you elaborate on how frequently lakes appear and then drain, how long they last (more or less than a single season?), how often they coalesce, and whether it is the same ones that reappear each time, or new ones that emerge? Kneib et al., 2021 do a nice job of this for ice cliffs as an example. This will tell us more about the processes that are driving the surface changes on these glaciers and add significant value to your manuscript.

Thank you for this suggestion. Along with the glacier-specific lake number and area which we discussed in the comment above, we have added statistics which investigate the permanence and recurrence frequency of lakes using the PlanetScope dataset. This line of investigation has provided important context to highlight and contextualize the wide range of behavior which these dynamic features can show.

3. Whether the inclusion of the UAV data is necessary – I'm not sure at present it adds anything to the key story – if anything it detracts from it. Consider re-packaging it as a ground validation dataset for the Planet imagery (see below)?

We appreciate your comments and suggestions regarding the use of UAV-derived datasets here. We agree that using the UAV-derived imagery as a ground truth for the coarser-resolution imagery would be a valuable approach. Unfortunately, the Planet-derived dataset (2017-2022) does not overlap with the field observations (May 2023), and extending the timeline of the Planet-derived dataset for this purpose is not feasible due to the quota limits and effort which would be needed.

Thank you for pointing out this apparent disconnect between the area-volume relationship and seasonal evolution of lake areas in our submission. As we outline below, we feel that it is a valuable contribution to our collective understanding of the hydrology of these debris-covered glacier systems, but needs to better justified and connected than we have done previously. Making observations of supraglacial lake area is a relatively straightforward undertaking from remote sensing observations, and is thus the primary tool used for studies such as this one. However, understanding the volume of water stored in these features if often the end goal, as the volume storage (and change) is more important from the perspective of understanding hydrologic fluxes and potential hazards. The use of area-volume scaling relationships is typically used to convert from observed area to estimated volume of lakes (Watson et al., 2018; Cook and Quincey, 2015), but these relationships are far from perfect and can result in large errors. Further, the area-volume relationship for supraglacial lakes/ponds of the size seen on the glaciers in this study (<20000 m^2) is based almost entirely on the dataset from Watson et al. (2018) of bathymetry measurements of 24 lakes on Khumbu Glacier. Very few other measurements of SGL bathymetry/geometry exist in this region, in part because the effort required to gain these measurements is so large. The observations which we provide in this study of the area-volume relationship of two dynamic SGLs on Ambulapcha Glacier are an important addition to this sparse dataset, particularly the finding that the geometry of these lakes fit within the range of existing observations.

Given this context, we feel that this data is a valuable component of this manuscript. We have elected to move Figure 8 (illustrating the area-volume relationship of these lakes) to the supplementary material as we feel that the other results and discussion presented in the manuscript are more important for readers to focus on.

4. Reducing the number of figures overall (including in Supplementary) and condensing the text where possible (particularly methods)

Thank you for this helpful suggestion, which was similarly brought up by the second reviewer (regarding limiting the number of figures and text in the main text). We agree that the original Methods section was overly comprehensive and included information not explicitly needed for interpretation of the findings. In order to focus on the most important steps we have elected to move portions of the Methods section into the Supplementary Materials. Specifically, we have moved the PlanetScope "Data Access and Cleanup" as well as "Filtering and smoothing" sections into the supplement.

Additionally, we have reduced the numbers of figures in the main text by moving Figures 2, 6, and 9 into the supplementary material. We elected to move Figures 2 and 6 because they are the most methodology-focused, and provided the least value for interpretation of our findings compared to other figures. As discussed in our response to your comment on inclusion of UAV-data, we moved Figure 9 (which investigated the area-volume scaling of the Ambulapcha lakes) to the supplement as well.

While we acknowledge that there are many figures provided in the supplementary material, we feel that these are all important to include in order to be transparent and thorough in our presentation of our findings. The majority of supplementary figures are to provide glacier-specific recreations of Figures 8 and 11 (from the original figure numbering). Including Fig. S3-S10 in necessary to highlight how the seasonal pattern of lake expansion during the winter, which we identified from the PlanetScope-derived SGL dataset (one of the main findings of this study), is repeated on almost all glaciers in all years, and is not just found on Lhotse Glacier. The inclusion of the Landsat-derived decadal changes in SGLs on each glacier (Fig. S13-S19) is important because this a relatively data-sparse region, and it could provide important context to future researchers investigating these glaciers either from remote sensing observations or planning field work in this region.

5. Being more explicit about the % errors on your lake areas, rather than presenting it as a proportion of the debris-covered area.

Thank you for this suggestion. We agree that it is important to be explicit in the error estimations, and that presenting errors as both a percentage of the total lake area and the total debris-covered area is important. In order to present our findings in a similar manner to previous studies we have amended the text throughout to include these numbers, including recalculating error estimates for Table S3 (comparing PlanetScope-derived SGL area to Landsat and Sentinel-2 derived areas for coincident imagery).

More minor comments:

Line 11: I interpret 'annual' variation to mean from one year to the next, but I think you're meaning from one season to the next here?
Thank you for this clarifying comment. This was our intention and we have revised it accordingly.

Line 20: present -> presents
Done.

Line 50: can you add a sentence here to underline the importance of understanding these seasonal cycles that your paper characterises?
Thank you for this suggestion. We have elaborated on this concept in the revised manuscript.

Line 52: I'm not sure if it is 'in-situ' or 'in situ' but be consistent
Thank you for pointing this out. We have changed all instances to the correct "in situ".

Line 66 (Section 3): this is a very long section! Can you be a bit more concise and focus the text on key points?
We have shortened the methods section by moving some information in to the supplementary material, leaving only the most important aspects of the methodology in the main text.

Line 68-69: remove sentence starting 'Each individual source…' – it's superfluous
We have changed this sentence to read: "Each source allows…"

Line 75: Unmanned -> Uncrewed
Done

Line 75-76: if that's what your UAV data helped with, I'm not sure it is evident in this paper? Consider revising?
Thank you for this suggestion. We have rephrased this sentence to better reflect the use of in situ observations to provide context of smaller-scale surface features and local topography of the glaciers.

Line 114: maybe spell out the NDWI using mathematical notation?
Done.

Line 116-117: does this methodological step not fall down for larger lakes given all pixels within that 150 m buffer will have similar NDWI values? Can you clarify?
Thank you for this clarifying question. In our testing, none of the lakes in our study area are big enough to see this effect (Figure 5 highlights the largest lakes we encounter). This length was tuned empirically to be as small as possible while not seeing this effect on the large lakes.

However, if this method was used on larger lakes (for example, Imja Tsho) a larger buffer size would be necessary.

Line 190 (Section 3.3): I'm not convinced that the inclusion of these field data adds value to the manuscript. How about using them instead to assess the uncertainty in using your PlanetScope imagery as the 'truth' for the other coarser resolution datasets? At 3 m spatial resolution there is still some ambiguity as to exactly where the lake margins lie – and with your drone imagery you can put a figure on that, which will be of use to anyone using Planet as a validation dataset going forward?
Thank you for this suggestion regarding the best use of our field observations. We agree that using the UAV-derived imagery as a ground truth for the coarser-resolution imagery would be a valuable approach. Unfortunately, the Planet-derived dataset (2017-2022) does not overlap with the field observations (May 2023), and extending the timeline of the Planet-derived dataset for this purpose is not feasible due to the quota limits and effort which would be needed. Please see our response to the major comments (above) for more discussion on the inclusion of these materials.

Line 200: can you be sure that with an error estimate of +/- 1 pixel these manual delineations are more robust than the semi-automatic results? Normally one would use a finer-resolution dataset than that being evaluated to produce these values…? See point above.
We believe that the manual delineations of SGLs in Planet imagery are more robust than any individual lake product from the semi-automated approach. If we were to manually delineate the lakes on every Planet image we have, then we would have higher confidence in the results. But the time required to do so is of course prohibitive. We consider our validation dataset more as a guideline of how well the semi-automated results match up with that "best possible" dataset of all manually-derived products, and a confirmation that the results and seasonal variations that we find from the semi-automated results are in fact real.

Line 213: can you also state the % error on the lake areas? It's useful to know within the context of the whole debris-covered area, but probably more important is what it means for the data you present in the plots (and ideally, these would have error bars on too)
Thank you for this suggestion. We have now included the % error relative to total lake area. In the text (25%, 33%, and 13% of the total lake-covered area on each glacier, dropping to 7.6%, 19.1%, and 8.3% during the September-February months).

Line 228-230: same point as immediately above
This stated error is calculated with respect to the total lake area. We have updated the text to make this clear.

Line 255-256: this same point has been made three times in quick succession. Maybe remove this sentence?
Thank you for pointing this out. We have edited this section to make it less redundant.

Line 278-285: this feels a bit odd in the context of the satellite-based observations.
Please see our response to the major comments point #3 above.

Line 309-317: Not discussing your results in comparison to Watson et al., 2016, who used similarly fine-resolution imagery along with Landsat to look at seasonal patterns, seems like a bit of an obvious omission here?
Thank you for pointing out the unintentional omission here. We have added a paragraph here to elaborate on direct comparisons between our PlanetScope-derived results with their high-resolution imagery.

Figure 11: needs a colour scale to show frequency values
Thank you for pointing this out, we have included a colorbar in the revised version.

Line 348: can you explain (in the text) why there would be increased solar radiation absorption during winter months?
The use of "increased" was meant to refer to increased energy absorption in the areas of thin debris cover relative to clean-ice surfaces due to decreased albedo, however this was not clear in the original phrasing. We have edited this sentence by removing the word 'increased' here, and further clarifying the explanation in the following sentence.

Line 371: I agree the topographic characteristics of Lhotse are unique within this suite of glaciers, but is the predominant expansion of ponds at Khumbu not also a couple of km from the glacier terminus?
Thank you for pointing this out. Yes, the distance-from-terminus at which the major lake expansion is happening is similar on both Khumbu and Lhotse (and also similar to where SGL expansion initiated on Imja Glacier in the mid-1900s). The down-glacier characteristics (the steepness and lack of a terminal moraine) are the more unique aspect. We have edited this paragraph to put more emphasis on this.

Line 372: Figure S17 takes me to Lhotse Shar, not Lhotse
Thank you for pointing this out. We have fixed the citations and ordering of supplementary materials.

Line 399 (Section 6.5): I'm not sure this is packaged up in a fair way – the other papers/inventories you refer to here didn't attempt to delineate small ponds on the glacier surfaces as you have here – they all set a minimum pond area for detection, and were largely

focussing on what may otherwise be termed as a lake (i.e. much larger than a pond) because of their much broader spatial coverage. It's a bit like comparing apples with oranges in my mind.
Thank you for this suggestion. We have expanded this section of the discussion to provide a more thorough explanation of the previous studies which we are comparing our results to, as well as explaining more clearly that the minimum size threshold used in these studies is the reason for the differences. We feel that it is important to provide the comparisons here in order to highlight the limitations of our current knowledge of the regional-scale spatial and temporal variability in supraglacial hydrology of debris-covered glacial systems.

Figure S12: is it worth pointing our somewhere that the negative trend you identify on Lhotse Shar and Imja is at least partly a consequence of the glacier area shrinking (and Imja Tsho expanding) over the period of observation?
We do not incorporate time-varying glacier extents into our analysis. The products we present here are derived using a single outline corresponding to the modern ~2022 extent. which excludes Imja Tsho. The long-term negative trends for Lhotse Shar and Imja show the lake area for this 'patch' that would have been further removed from the terminus and proglacial lake in the 1980s and 90s. We discuss the implications of this, and the possible feedbacks between proglacial lake expansion and supraglacial hydrology, in the discussion (Section 6.3. Long Term Trends).

Figures S13-S19: these all need a legend to give meaning to the shades of red.
Thank you for pointing this out. We have updated the figures to include the legend.

Figures S13-S19: There are some suspicious areas in the uppermost part of your debris-cover boundary that look to be misclassification, rather than lakes. They are particularly apparent in the Lhotse Nup (2002-2005), Ama Dablam (1998-2001) and Ambulapcha (1998-2001) figures – could they be areas of wet snow? Do they translate through to your data presentation in the main text? Or do you believe them to be genuine…?
These areas that you have pointed out are most likely misclassified shadows, wet snow, or shadowed snow areas that were not masked out during our processing steps. While they are not genuine, these over-estimations of lake area have minimal effect on the long-term trends which we observe because they occur infrequently.

References

Kneib, M., Miles, E. S., Buri, P., Molnar, P., McCarthy, M., Fugger, S., & Pellicciotti, F. (2021). Interannual dynamics of ice cliff populations on debris-covered glaciers from remote sensing observations and stochastic modeling. Journal of Geophysical Research: Earth Surface, 126(10), e2021JF006179. https://doi.org/10.1029/2021JF006179

Watson, C. S., Quincey, D. J., Carrivick, J. L., & Smith, M.W. (2016). The dynamics of supraglacial ponds in the Everest region, central Himalaya. Global and Planetary Change, 142, 14–27. https://doi.org/10.1016/j.gloplacha.2016.04.008