Reviewer 1

This is a well written, interesting, short paper. Leaning into a variety of data collected 12 years apart, the authors offer a simple, clear message about active layer change. My comments are largely about accessibility in an attempt to help the authors reach the broadest audience that they can with their work, including moving up description and important of polygonal tundra a bit earlier, and more clearly coming back to the implications of their findings in the conclusions. While I have done some work in Alaska, my expertise in polygonal tundra is zero, and I struggled at the beginning of this paper with why this work was important as those details didn't come up until later, as you'll see in my minor comments below. My only scientific concern is the "artifacts" outlined in Figure 1 and 2 that aren't explained—if these data are artifacts, how are the authors sure the others are not? Overall, this is a nice paper worthy of publication.

Thank you, Dr. Singha, for your kind words and constructive comments and suggestions. We revised the manuscript to describe polygonal tundra and its importance earlier (in the Introduction), and to more clearly return to the potential implications of our findings in the Conclusions. We also clarified the mention of "artifacts" in Figures 1 and 2, explaining why these are classified as such, whereas other changes are considered real.

Minor comments:

Abstract: this might be obvious to the Cryosphere audience but isn't to me—why is the type of polygon important? This might be too much to build into an abstract, but it also makes the abstract a bit less compelling to me as I don't understand the 'why' yet.

We agree and have revised the abstract accordingly.

L21. Similarly, why only polygonal tundra landscapes? Other ones don't matter?

We improved the sentence.

L24-27. Again, this an accessibility thing—I imagine these sites might be familiar to many readers of Cryosphere, but it could be nice to have a map that labels them and perhaps also shows where these concerns (worldwide?) are most notable.

We added Figure S1 to show the location of Utqiagvik, Alaska, and revised the text to better explain the spatial context of study locations. Because the manuscript is limited to four figures, this additional map was placed in the Supplementary Material. To our knowledge, no existing data product delineates polygonal tundra extent at the Pan-Arctic scale. We therefore included regional-scale quantification of polygonal tundra in the revised manuscript.

L39. I still don't know why I care about polygon types, or even how many there are.

We revised the Introduction to present polygon types, extent, and importance earlier in the manuscript.

L41, 45/46. Again, map call out here would help. I don't entirely know where I am in Alaska.

See answer above.

L49-55. Aha! Here's what I was looking for. I'd move this into the introduction, personally, rather than study site, to help set the stage for people like me who don't work often enough in these settings to know these polygon types in and out.

We moved part of this information to the Introduction and added details on the importance of understanding changes in polygonal tundra.

L67. What does "the measurements" mean here? I don't know what measurements we're talking about yet. Also, HCP, LCP, FCP haven't been defined, although I can figure them out. That said, why have an acronym if you don't need it? If there's space, I'd just write these things out. I had trouble really even thinking about the results when I

got there because I had to remember what these acronyms were and then what that meant with respect to the results. It's fine in the figures, but would be nice to have written out in the captions, etc. just to make it easy on the reader.

We reduced the number of acronyms. We kept HCP, LCP and FCP in the captions only.

L74. What kind of geophysical data? Is a citation needed here, or are you talking about new data, in which case it should be in results? This section should more clearly outline the previous work that this paper's new data are being compared to. Is it all in Dafflon et al. (2016)? I can't quite follow and am not even sure what all data are being compared as things are a little disorganized.

We revised this section to more clearly outline the previous work to which this paper's new data are compared. The comparison dataset is presented in Dafflon et al. (2016).

L114. What is DSM? Digital surface model? I am not sure. Missing a period at the end of this sentence, too.

Digital surface model. We removed the acronym (except in the caption) and added a period.

L122-124. It's hard to know what to make of these values besides that they're comparable between times without some measure of noise. What's the error in the stacking/reciprocal/repeat measurements over these two data collection periods?

We added the error between normal and reciprocal measurements.

L130. Up to the authors, but this seems an unneeded acronym.

ALT acronym removed (except in captions).

L134. Specifically, differences in rank sums or medians, right? Statistical tests are all about differences, but different kinds of differences, so it helps to be specific.

The statistical significance of differences across polygon types was assessed using the Mann–Whitney U test, which compares the rank distributions of two groups. We improved the sentence in the manuscript. We also decided to provide more statistical metrics in the supplementary (Table S1), including the median paired change between 2011 and 2023 using the Wilcoxon signed-rank test (two-sided), and the Hodges–Lehmann median change with 95% CI.

Figures 1 and 2 captions. How do you know they are artifacts and that the other ones aren't? What do the colors even mean? I don't see a colorbar.

It is a binary map (though small-scale variability can give the appearance of color gradients). Field notes and visual observations suggest that the estimated pool expansions in the centers of low-centered polygons (LCPs) are artifacts of the machine-learning classification, likely caused by the complex, changing mixture of water, solids, moss, and sedge. We have added this explanation to the figure captions.

L165. Again, median ALT, yes? What's the actual p-value? 0.05 is a misused bright line, IMHO—it's more useful to see the actual p-value. There's a big difference between 0.01 and 0.044 in terms of my confidence.

Good point. We agree we oversimplified it in the text. We have revised the text to be as precise as the caption.

L241. Spell out CALM on first use.

Done.

Conclusions: While a good summary of the paper, I think the authors miss an opportunity both in their results and conclusions to make really explicit what the implications of their results are, especially around issues around carbon cycling as mentioned in the abstract. At a time when some people seem to think climate change isn't a thing or doesn't matter, a quick highlight, say for a U.S. congressperson's staffer who might see this, would be a worthy add.

We added a sentence to be more explicit on possible implications.

Reviewer 2:

The study "Decadal changes in topography, surface water and subsurface structure across an Arctic coastal tundra site" by Bachman et al. presents an extensive field data set that characterizes the evolution of arctic polygonal tundra upon warming over a decade. Such comprehensive data evaluations are rare and the study is well-written and presented. I recommend publication in TC after some revisions.

We thank the reviewer for the kind words and for the constructive comments and suggestions.

Major comments:

1. Please provide an analysis of the meteorological conditions in the two years, focusing especially on summer precipitation and potentially air temperatures, which are at least some kind of proxy for evapotranspiration. The main question is whether the increase in surface water is indeed due to permafrost degradation and ground ice melt, or whether 2023 was just a wet year. I don't have doubts in the overall analysis and the interpretation of the findings, but this is important context for the findings.

We added Figure S3 to compare air temperature and rainfall in 2011, 2013, and 2023, and discuss it in the manuscript. We confirm that we believe the main reason for the increase in surface water is localized subsidence at multiple sites. We have clarified this point in the text.

2. The study investigates three polygon types, which at least to some degree are also development stages of polygons in a warming climate. However, the study treats them as static over the ten investigated years and also only provides single values for the subsidence, ALT, etc., which suggests (or at least leaves the possibility open) that the polygons are subsiding homogenously in space and time. In reality, the subsidence and ALT changes should be different for the different parts of the polygons, depending on their development stage. While I would not require additional analysis, the authors should mention and discuss this issue wherever appropriate throughout the entire paper. The authors could for example just provide the elevation difference maps in Fig. 1 and 2, and then discuss whether the subsidence patterns are controlled by the polygon structures or not.

We agree that in ideal conditions we should separate each polygon and each trough, rim, and center of each polygon to assess where changes occur. Yet, this is difficult, primarily because some of our datasets cover only the main transect and others are point-scale measurements with limited spatial resolution (incl., active layer thickness). We added a sentence linking our classification of changes by polygon type with the general distribution of changes along the transect (i.e. trough vs. more diffuse changes). In addition to the DSM difference along the transect (Figure 1i), we included a DSM difference map in the Supplementary Material (Fig. S4) (see response to minor comment). Finally, we note that our classification into HCP, FCP, and LCP is based on their likely different developmental stages. We treated them as static types in this study because we did not observe any polygon transitioning from one type to another during our measurement period.

Minor comments:

L. 24: provide a very short definition of " degrading" ice wedges in brackets, e.g. negative net ice mass balance or similar.

Done.

L. 25: Is "thermokarst pool" the same as "thermokarst lakes and ponds"?

We use "thermokarst pool" to refer to small thermokarst ponds. Some of these features may be troughs as well. To improve consistency, we have reduced our use of "pond."

L. 46/47: use either metric system or the one with feet and miles. I think TC might even require metric?

We improved consistency by using metrics only.

L. 48: reword this phrase: "geomorphology" cannot define "microtopography". What you mean is more the "formation mechanism/processes which define the microtopography".

We improved the sentence accordingly.

L. 62: use stronger formulation than "indicate"

We replaced it with demonstrate.

L. 73: define whether this is by total volume, relative to the volume of the drained sample, of by weight.

It is volumetric. We clarified.

L. 91: provide some more information on how the accuracy was determined.

We added details.

Fig. 1: Please also provide the difference image between the two shots for the entire stripe.

We added the DSM difference in the Supplementary Material (Fig. S4). The DSM difference along the center line (along the ERT) is already shown in Figure 1i. Because we are limited in the number of figures allowed in a short communication, we included the additional map as supplementary material rather than adding a figure or expanding the size of the main figures.

Fig. 1i: explain how this is determined. Is this the average over the stripe, or the position at a single line, e.g. where the ALT was measured.

This is where ALT was measured. We have clarified this in the manuscript.

Fig 1j: I think "permafrost thaw" is misleading, this is rather the "change in permafrost table".

We avoided using "permafrost table" because we believe readers may think about permafrost table depth, which depending on the reference point (ground surface or fixed elevation) may not account for subsidence. We have now replaced permafrost thaw with permafrost table elevation at several locations across the paper. And we modified the caption of figure 1 to clarify that by permafrost thaw (in the figure) we refer to changes in permafrost table elevation, inferred from changes in ALT and ground elevation.

Fig. 3e: is this really 0.5-1.5m, i.e. it overlaps with the 0.5-1.0m of the previous figure? If not, please correct!

That is correct. We used 0.5–1.5 m to evaluate the average over a thicker depth interval.

L. 216: thaw depth is more common

Done.