Review for

Permafrost thaw follows spatially organized, nonlinear trajectories in ice-wedge terrain

This study investigates the differences in thaw trajectories within polygonal permafrost landscapes in a 50 km^2 study site in the Arctic Coastal Plain in Northern Alaska, over a time series covering four dates in 17 years. The authors find that the progression of degradation is not linear or uniform throughout the studied landscape and time frame, but that the trajectories are tied to the geomorphological setting of the immediate surroundings. As such, this study presents a novel and highly relevant addition to researching the spatial and temporal patterns and drivers of permafrost thaw. The overall quality of the paper is high, as the authors have clearly described their methods, results, and the conclusions drawn, making it possible to trace and reproduce most steps of their study design. Nevertheless, I suggest a few additions, corrections, and clarifications to the manuscript, before I would recommend acceptance for publication in The Cryosphere.

A. Major comments

1. What is the reason that you selected only those four years for your time series? Were there no other cloud-free summer images available in any of the other years?
2. I believe that it would be essential to include some meteorological observations into your study:
    i. Determining the degradation stage in your study is significantly dependent on the presence and share of water in troughs. You do mention the common practice of using peak summer period observations (lines 375-378), but how do you account for individual (out-of-the-ordinary) precipitation events that might have occurred shortly before a given acquisition date? I can imagine that especially class *Drying* is highly dependent on precipitation events right before image-acquisition dates.
    ii. You even mention the possibility of temporally nonlinear responses due to episodic climate extremes (line 350).
    iii. The study site description would generally benefit from some more details on the governing climate and permafrost properties (+ see comment #B52).
3. Do you have any LCP terrain in your study area? Does your approach translate to such study areas as well? Would be good to mention the scope of applicability of this study, so that it may be applied and adapted for significantly bigger parts of the Arctic.
4. Some thoughts on preserving the original resolution of the input image (lines 135-136): I understand not wanting to lower the resolution as that would lose some of the spatial detail that is present in the images. However: a higher resolution will allow you to pick up on smaller features, including smaller trough ponds. This could lead to the detection of ponds in higher-res images, that would be missed in lower-res images (especially given that the size of many trough ponds would only be a few pixels in size anyway and therefore close to the detection threshold). Looking at the percentage of Advanced Degradation areas (Fig 3) and comparing it to the spatial resolution of the corresponding year's satellite images, there is a clear positive correlation. I don't know how large this influence is, but it would be worthwhile to at least address in the Discussion.
5. Some thoughts on degradation stages (151-157, Fig 2):
    i. Did you have any cells with narrow and wet troughs?
    ii. And looking at the 'Advanced Degradation' stage according to Liljedahl et al. (2016), where we see 'fully'-connected and thus drained networks. Does this relate to

Wickland et al. (2020) and lines 170-176? How would you confidently distinguish these from the 'Initial Degradation' stage in your study?

iii. Lines 381-386: here you are even highlighting that in progressing degradation, surface water can drain, yet you consider such areas as *Drying*. It's not fully clear to me if this is part of a critical discussion of your results, but I think this would be a good place to elaborate more on the topic.

iv. How do you explain the cells that went from *Advanced Degradation* to *Undegraded* over the course of the observation period (see Fig. 3)?

6. Some more quantitative statistics (and their significance) for your spatial observations tying degradation stages to geomorphic landforms would be very helpful (this relates to e.g., Figure S9, Table S7, "pronounced spatial clustering" (line 408), …)
7. Lines 284-286: This observation is a bit difficult to make here, as you don't actually know that the degradation follows a linear trend (rapid degradation can trigger through disturbance events, e.g., lake drainage, fires, but also extreme temperature or precipitation events/years) and you have differently-spaced time steps (see also comment #B32).
8. Overall, you are describing 17 years of observations on degradation states, yet, you tie them to millenia-scale landscape changes. Does this imply that the terrain that was *Undegraded* in 2006 has been in that state since the early Holocene (e.g., lines 341-345, line 417)?
9. Could you discuss some more the relevance of your local study in the Conclusions? E.g., how do your observations translate to other sites; what would users need to consider when applying your approach elsewhere; who benefits from these spatially very detailed observations (rather than looking at the regional trend only); …?
10. There's quite a few AI sentences in here. I don't mind the use of LLMs for enhancing the understandability of scientific manuscripts in general, but sometimes the language gets too flowery and ends up making it harder to understand instead. Or it generates sentences that sound logical but make no semantic or argumentative sense once you look at them more closely and with expert knowledge. See for example: comments #B13, #B14, #B15, #B39, #B44, #B53, or sentences in lines 341-342, lines 404-405.

B. Minor suggestions to enhance understanding

1. I'd suggest adding details in the title that make it clear that you have looked at only a small study site, and over a limited time period: e.g.: *Recent permafrost thaw follows spatially organized, temporally nonlinear trajectories in (North Alaskan) eolian ice-wedge terrain*.
2. It would be helpful to readers if the Abstract contained more specific information, e.g.:
   i. Line 10: "Using a decadal time series **(2006–2023)**"
   ii. Line 11: "across a 50-km^2 **study site on the Arctic Coastal Plain of Northern Alaska**."
   iii. Some more quantitative/statistical results would be nice here as well (see also comment #A6).
3. Line 23 & 2x line 44: ice wedges *melt* as there is a phase change from solid to liquid.
4. Lines 33-35: The semantic reasoning combined with the choice of terminology makes these two sentences a bit weird: "networks of massive ground ice" → either "networks of massive ice wedges cover […]", "massive networks of ice wedges cover […]", or "millions of square km are characterized by high ground-ice content". The latter option would also build a better transition to the next sentence.

5. Line 36: "**These** mm-wide cracks" → The cracks as a semantic object have not been introduced yet, so **These** does not refer to anything yet. Maybe replace with "**Initiating as mm-wide cracks in cold and dry winter soils**, they are infilled [...]. As ice wedges grow **throughout many seasonal cycles**, they create [...]"
6. Lines 37-38: It's not the "polygonal networks" that span 5-30 m in diameter, but each of the polygons in the network.
7. Line 39: change "making the loss of ice-wedges through degradation" to "making the loss of **wedge ice** through **permafrost warming**"
8. Lines 44-45: missing citation
9. Line 57: what exactly is meant with "trough features"?
10. Lines 62-63: The second half of this sentence is a bit difficult to understand when reading the paper for the first time. "the overall magnitude of thaw expansion" is roughly the same as "the cumulative outcome of landscape change", so there is no real information in this sentence. I assume you want to use this to lead up to your more spatially and temporally detailed analysis, but even so, "the **overall magnitude** of thaw expansion" when using your analysis, will also be an aggregate number, and thus only give you the "the **cumulative outcome** of landscape change". I think with some other choices of words this can be clarified; but as is, I fear it is a bit confusing and not concrete enough for a first-time reader especially.
11. Lines 69-70: again, semantics: "Landscape heterogeneity" Sounds like it is referring to the morphological landscape classes (that will only be introduced later). But given the context of the next sentences, I assume you mean the heterogeneity of polygon network morphologies/classes?
12. Line 73: remove "elevated"
13. Lines 76-78: I would remove (or clarify) this sentence, as it a) seems to contradict itself/some of your study results and b) therefore makes more sense to consider after your study results have been shown, not in the Introduction. Alternatively, you could consider adding/changing to "[...] rather than **just** the annual-to-decadal **degradation dynamics that are commonly observed**."
14. Lines 82-83: Nitpicky, but: "The complex spatiotemporal trajectory of thaw" is the same as "the non-linear evolution of permafrost thaw", making the sentence's argument: "Characterizing A, is essential not only for A, but also for B". You wouldn't need to repeat the A part.
15. Line 90: I don't generally understand the difference between (i) and (ii) in this enumeration. Could you clarify? Does the enumeration or this three-tiered structuring show up somewhere again? Could be helpful to match these three points to the three subchapters of *3. Results*.
16. Chapter 2.1 – 2.3: The study site + the location and overlap of the satellite images + how these overlap and result in the 50 km^2 overlap site would immensely benefit from a study site map.
17. Chapter 2.2: In your methods chapter, you use the passive voice a lot. This makes it unclear to the reader if the work was conducted by you, or if it had been done before in another study by someone else. If you are describing work done by you, it is common to use the active voice, otherwise, if done outside this study, some citations would be helpful.
18. Lines 108-110: Change the order of the last two sentences.
19. Table 1: Could you change the date format to YYYY-MM-DD
20. Table 1: the percentage of training tiles/area from the whole share of the image would be a more helpful metric.
21. Line 116: "The model **was trained to segment** images into [...]"

22. Line 119: “Loss of trough water can occur during stabilization of degradation”. Do you mean drainage from connecting troughs and trough networks? In the classic model, stabilization occurs partly *because* trough water has drained. Some rephrasing might be helpful to clarify the relationship.
23. Line 122: Could you add the year of the orthomosaics and DEMs you used here?
24. Line 122: “in accurately identify**ing** trough […]”
25. Lines 127-128: This sounds a lot like overfitting. Which, in the context of this study, is absolutely acceptable and actually beneficial for the results. However, it might be helpful to mention, so that others are aware of this in case they are interested in using your trained model on other images.
26. Line 140: Could you add here a sentence on if the model was evaluated on the whole/merged maps or on the individual tiles?
27. Figure 1 is a very nice figure!
28. Lines 146-147: I’d be very interested in a table containing the results of these metrics! Doesn’t necessarily need to go into the paper, but maybe in the Supplement? Or in the accompanying dataset repository. These metrics could be quite interesting for e.g., modelers as well.
29. Line 178: “for each **30 m by 30 m** spatial aggregation cell”
30. Line 180: Up until here, it still has not been made clear what the different pathway/trajectory categories are, or if there even are distinct categories. I think it would be beneficial to mention this setup a lot earlier on, for example in lines 89-91. Throughout the Methods chapter, I have found myself wondering a few times, where the whole setup is leading to.
31. Lines 181-194: Fig. 4E is such a great Figure, and it would be helpful to already have it available here. Could you either add it here in larger, or at least refer to it in this paragraph? And add the little ice-wedge icons next the degradation stage segmented map in Fig. 2.
32. Lines 196-198: Technically you are identifying the first year that you detect the change, not necessarily the year that the change occurred.
33. Line 203: What sensor and what year was the satellite image from? Can any of the landforms have changed over the observation period (e.g., lake drainage/expansion)?
34. Lines 206-213: This paragraph comes a little out of nowhere and is quite technical all of a sudden. Could you add an introductory sentence that explains why these steps were done?
35. Lines 225-227: Could you add the numbers/percentages here?
36. Line 229: “substantial portion of the landscape”: again, a number would be good.
37. Generally, I’m a fan of Sankey diagrams, and I initially thought it’s super cool here as well, but after looking into a bit more detail, I might conclude that it’s actually not ideal in this case; so here are some thoughts on Figure 3:
    i. The distances between the dates in the diagram are not accurate to the time that has passed between them in reality (not technically wrong, but separating them according to the time that has passed would give the reader an significant visual help in interpreting the changes a little better in the correct temporal frame).
    ii. You write (lines 239-240): “Sankey diagram illustrates […], revealing multiple, non-linear pathways of landscape change.” However, since it is not possible to trace the trajectory of individual cells throughout the four dates, but only the changing shares, it doesn’t actually reveal the specific “pathways”, only the overall share of degradation statuses within the landscape over time. Together with Table S6 it is a lot more informative

though, so I'd suggest a) changing the wording here a little bit, and b) referring to S6 in the caption here as well.

iii. Could you change the date format to YYYY-MM-DD?

38. Some thoughts on Figure 4:
    i. Very, very, nice figure (the organization, the graphs, the colors, …)! (nevertheless, some thoughts ☺)
    ii. There's not too much to identify in panels A and C. A zoom in would be a nice add-on (but not essential). However, will you publish the data of the map in A? This will be very interesting for other users too, I assume.
    iii. Panel F: Can you explain again what the categories "No Trough", and especially, what "No Trajectory" are? Is "No Trough" the masked areas of lakes (given the white color), or is this an extra class? How is "No Trajectory" different from the aggregate class "Persistent Trajectories"?
    iv. Panel F: Overall, I think switching the bars around would be a lot more informative/interesting to readers. So instead of showing *Distribution of Landforms within each Trajectory*, I would instead show *Distribution of Trajectories within each Landform*. If we consider to potentially extrapolate this study, the underlying landforms in other areas are information that is simpler to get. So given that information, it would then be possible to estimate the amount/share of degradation pathways one could expect in those other areas, using your shares of trajectories within landforms. In your text (e.g., lines 253-261, lines 263-264, or Fig. 6), you also report and argue on this 'reversed' number, so that underlines the argument to show this reversed stacked bar plot instead actually.
    v. You could also consider making a combination of panels F and G; i.e., not having the 100% stacked bar plot, but integrating the share into the absolute bar plot of G. Of course, you wouldn't see much of the distribution in e.g., class *Undegraded*, but then that class doesn't contribute to the overall share anyway.
39. Line 255: From Fig 4F, it looks like *Drying* and *Oscillating* also occur at similarly high percentages within fluvial channels?
40. Lines 267-268: Consider removing the sentence "These younger geomorphic surfaces generally contain less developed ice-wedge networks and higher surface water cover." It is a bit out of context here and out of place in the Results section. Or explain the relationship to *No Troughs* and *Undegraded* class in more words.
41. Line 273: "rather than remaining confined to persistent hot spots" → was IWD expected to remain confined?
42. Lines 274-275: Can you use the same terminology here as you've used before? That makes it easier to relate to the specific classes you're talking about here.
43. Lines 288-295: It looks like these observations can be taken from numbers in Table S7? Could you link to this? And what year does Table S7 represent?
44. Lines 303-304: What does "without erasing geomorphic structure" mean?
45. Lines 311-327: This content would better fit to the Conclusions or Abstract, rather than the Discussion.
46. Fig. 6: How do the thaw trajectories at the bottom of your figure relate to your classes in Fig. 4F? Could you reuse your ice-wedge icons or make another visual link?

47. Line 365: “a heterogeneous mosaic of hydrologic, ecological, and biogeochemical states […]” → I’d remove *ecological* and *biogeochemical*, since you are not investigating this in your paper. Or change to: “a heterogeneous mosaic of hydrologic, **and implied through this:** ecological, and biogeochemical states […]”.
48. Lines 392-393: “[…] would be expected to exhibit slower rates of change **today** than terrain that began degrading more recently.”
49. Lines 401-402: Based on what was this assumption made?
50. Line 412: remove “consistently”.
51. Lines 421-443: The majority of this paragraph should go into Chapter 2.1 Study site.
52. Lines 443-445: “Our results suggest that […] contemporary warming drives [thaw trajectory] progression through time” → you did not look at any temperature data though (see also comment #A2).
53. Lines 465-469: You did not investigate “topography”, “sedimentary properties”, or “climate” in your study, so I would suggest removing this from the concluding sentence. Same for “climate forcing”, “local ecohydrological feedbacks” (line 478), and “contemporary local feedbacks” (line 481)
54. Will you publish the segmented maps only for the overlap site or for the entire satellite images (the 2900 km^2)? I think the latter would be an incredibly valuable resource for the community!