Dear Anonymous Referee 2,

Thank you very much for your review and your constructive and detailed comments. The entire text of your comments (RC2) is shown below, followed by our responses (bold text). We hope that you will find the following explanations and proposed improvement to our manuscript a suitable response to the issues you have raised.

Throughout our responses, we refer to line numbers in the original manuscript and state revisions we have made to the manuscript. If we are invited by the editor to submit a revision, it will contain these changes.

Kind regards,

Nick Brown and Stephan Gruber

RC2 Major points to address

RC2: Some of the proposed metrics are not completely clear to me. I have noted specific questions below.

Specific questions are addressed below

RC2: Several sections are overly lengthy and would benefit from more concise writing, particularly the Introduction. Throughout the text, there are some redundancies which should be addressed

In addition to specific recommendations below, the introduction has been shortened by removing redundant material and tightening sentence structure.

RC2: The manuscript relies mainly on simulated time series (ERA) forcing a heat flow model (GlobSim) to derive and present the new metrics and then includes a case study that does not add much new information. Consider removing Section 5 or integrating it into the Discussion.

We believe that including real data as a part of a case study helps ground the research in reality and acts as a 'ground truth' of the simulation-driven theoretical development of the metrics. It also helps illustrate some of the differences in trend rates caused by measurement depth.

We have shortened section 5 to tighten the focus and have added a map of locations in the supplementary material as Figure S37.

RC2 Detailed comments

RC2: p. 5, chapt. 2.4: I do not fully understand the discussion about the talik. Please rephrase for clarity.

This paragraph has been rephrased as follows:

In cold permafrost, our estimate of the ALT coincides with the depth to the permafrost table, or top of permafrost (TOP). However, if a supra-permafrost talik develops, this is no longer true. In that case, TOP continues to deepen independently of ALT causing the two metrics to differ by an amount equal to the talik thickness. Here, we neglect any differences between the thermal (cryotic) and physical (freeze-thaw) definitions of talik or active layer boundaries.

RC2: p. 6: There are many equations — check whether all are necessary. Some measures are commonly known and do not need to be written out.

Four of the more obvious equations have been removed from sections 2.3–2.8.

RC2: p. 7, chapt. 2.8: I do not follow this paragraph; please clarify or rewrite.

Section 2.8 has been rewritten for clarity.

RC2: p. 10, last paragraph: Delete this paragraph, it is not necessary.

Paragraph has been deleted

RC2: p. 11, l. 275: Delete this paragraph. You do not need to continually preview what comes next.

Paragraph has been deleted

RC2: Figure 1: This is a nice figure. Consider indicating that the first six boxes represent “cold” sites and the next boxes “warm” sites. Alternatively, place cold sites in the left column and warm sites in the right column to facilitate comparison.

The image has been reordered so that the left column represents cold sites, and right column represents warm sites

RC2: p. 13: You have a subsection 4.3.1 but no 4.3.2. Either remove the subheading for 4.3.1 or add the missing subsection.

Subheading has been changed to 4.4

RC2: p. 4.5: This paragraph contains some redundancy — please tighten it.

Assuming this means subsection 4.5, the text has been reworked to reduce redundancy

RC2: Figure 3: I do not understand this figure; please improve the caption and figure clarity or simplify the presentation.

Figure has been simplified by removing coloured CDFs corresponding to different terrain types (these have been moved to supplementary materials). The retained CDF has added reference lines for 50th, 90th and 95th percentiles discussed in text. The caption has been reworded

RC2: p. 4.9: I do not follow this section completely. Why is this point important? Also, Fig. 4 is not readable — improve resolution and labeling.

This section aims to provides quantification for the reliability of metrics based on their signal-to-noise ratio (when used as indicators of changes to borehole heat content) and consistency. More specifically, we want to evaluate how much to *trust* that the direction and existence of change in borehole heat is predicted by change in a metric.

We recognize that calling this concept 'predictiveness' may contribute to some confusion and have replaced this with 'reliability' and SNR as defined above.

The introduction of this subsection has been significantly rewritten to clarify this point, and the remainder of the text has been updated for consistency.

The resolution of figure 4 and its labels has been increased to 600 DPI. Axis labels have been added to all bottom axes. Axis ticks and labels have been added to all left axes. The leftmost figure panel has been removed for clarity.

RC2: p. 4.10: Consider deleting this summary section; it seems unnecessary.

While this section does repeat some material from earlier parts of section 4, we argue that it serves a purpose by integrating the qualitative (sec 4.1-4.7) and quantitative (sec 4.8-4.9) analyses of the metrics to provide a final interpretation before moving on to the case study and discussion.

RC2: Chapter 5: See comment above about possibly removing this section. If you keep it, provide a map showing site locations — many readers will not know where "Schilthorn" and the other sites are located.

A map (Figure 5 in revised manuscript) has been added to show site locations and some text has been removed to shorten and clarify the section.

RC2: l. 571–572: The purpose of these lines is unclear, especially the sentence in line 572. Please reword or delete.

The lines have been deleted

Our results highlight the challenges of comparing trends between boreholes or regions. ~~In our simulations, the behaviour of each metric generally follows a similar trajectory. For MAGT, this is: rapid warming, reduced warming, no warming, and finally rapid warming if thaw progresses to the observation depth.~~ Most importantly, we do not change the magnitude of the warming…

RC2: Chapt. 6.3: The comment about bedrock sites echoes the original objective of the PACE boreholes, if I recall correctly — consider mentioning that here.

A sentence has been added pointing towards PACE as an exemplar of bedrock sites:

"The PACE project provides another example where boreholes in bedrock on mountain summits or plateaus were used for permafrost and climatic monitoring (Etzelmüller et al., 2020; Harris et al., 2009)."

RC2: Chapt. 6.5: The discussion seems to indicate talik development. If that is the intended message, state it clearly; if not, consider removing this paragraph.

Section heading changed to "Talik formation and accelerated permafrost degradation" removed. Two paragraphs of the section have been split out into a separate 'implications for monitoring' section for clarity.

RC2: Conclusions: Provide clear, concise conclusions, preferably as bullet points, rather than a long paragraph with references and summaries.

References and summaries have been removed. The conclusions have been reworked into bullet points, and some material has been moved into the discussion.

RC2: The manuscript addresses an important topic and presents useful new metrics. After clarifying the points above, tightening the text, and improving figure clarity, the manuscript looks ok for me.