

Harp et al. Effect of soil property uncertainties on permafrost thaw projections: a calibration-constrained analysis

*Reviewer comments ( italic)*

Your Responses to reviewer

My comments (blue)

Questions and comments / reviewer #1

*2. It would be better to include a site information section for Barrow. It can explain the site conditions in particular climate, snow distribution and vegetation cover as well as soil characteristics for the observational location.*

Reviewer #1 had commented on the lack of Barrow site characteristic information. Some additional information is given in your additional text (lines 91-100), however, it still lacks spatial and temporal snow cover variability, as well as snow cover physical information. Thus, please summarize the necessary information in this paper. Furthermore, please consider adding the time series of snow cover (as suggested in comment 4). Time series are available from Barrow.

Though there is information about the snow cover in your GMD paper, this paper needs to include this information to be understood.

*3. As I understand, the CESM outputs are used to drive the surface/subsurface model for calibration period (2013). Why not using the observed climate or at least showing the difference between observed and modeled atmospheric variables?*

The observed climate was used for the calibration in Atchley et al., 2015. The CESM outputs are used to drive the projections for which no data are available. We have modified the abstract to clearly state that "measured" borehole temperatures were used (line 11 of attachment)...

I do not understand why the addition of „measured“ borehole temperatures answers the point of the reviewer (what other borehole data could be used for calibration?)

Also, your response “The observed climate was used for the calibration in Atchley et al., 2015” – does not respond to the question of differences between observed or modeled atmospheric parameters.

As the reviewer suggests, please include a statement on the differences (observed/modeled) as text and/or figure in the supplementary material/appendix.

*4. What about the snow depth time series comparison? That would give important information on changes and timing of saturation as well as other metrics.*

Please also comment on snow physical processes/parametrization used in your model (see also comment 1 above). Not knowing your snow physical properties can induce large uncertainties in permafrost temperatures (see also Langer et al. 2013)- please clarify your methods used and uncertainties introduced.

[Again, your GMD paper results need to be summarized here]

*5. Why did you choose to calibrate for a single year of observational data? Wouldn't it be more useful to include as much observation as possible to constrain the parameters? Are there no available observations from other years?*

Yes, calibrating for multiple years would be ideal. However the subsurface data needed to calibrate the model was not available prior to September of 2012, and the calibration was done during 2014 prior to that year's data becoming available. The only complete year of data was for calendar year 2013. A sentence has been added to the Methodology section to explain this to readers (lines 115-116 of attachment). We thank the reviewer for pointing out that this was not clearly stated previously.

I agree with the reviewer that one single year of observational data is not enough (in addition to the fact that very similar temperature data are used from one landscape unit). Thus I encourage you to include several years.

There should be ample of temperature data available at Barrow.

Questions and comments / reviewer #2

*1. It is obvious from the high parameter uncertainty (and not surprising for a soil physicist), that temperature data alone is not sufficient to get a well confined parameter set. As freezing and thawing of porous media is a tightly coupled process where heat and water transport interact, there is obviously information missing about the total water content of the material. Additionally, the information content in the calibration data is quite low as can be seen in figure A-1 to A-3. The temperature is constant for long periods of time as a consequence of the zero-curtain effect or isolation by snow.*

*I am pretty sure that an in-depth survey (e.g. with virtual data) would show that temperature measurements at fewer locations combined with measurements of water and ice content would give a parameter set with much less uncertainty.*

*Thus the availability of only temperature data should be mentioned as one of the main reasons for the uncertain predictions.*

The manuscript quantifies the uncertainty in the case where only temperature measurements are available, a common scenario given the relative ease with which temperature measurements can be obtained compared to many other types of data. The soil property uncertainty would be expected to decrease if other types of data were incorporated, such as ice and water content. To ensure that this point is clear to the reader, a paragraph has been added to the introduction (lines 91-96) and the existing discussion has been augmented in the discussion and conclusions section (line 552).

This comment is not addressed satisfactorily in your response, as well as in the additional text.

*2. Even with a total of 16 calibrated parameters the model is obviously not at all capable of describing the data. The authors refer to the fraction of temperature measurements which are in the 95 percent confidence band.. I would expect that a thorough analysis of the response surface of the objective function should show a number of local minima. However, due to the high computational effort, the authors concentrated in this paper on investigation of the uncertainty around a single calibration point, which might result in an underestimation*

*of the uncertainty.*

However, in our inspection of the uncertainty produced by NSMC around the single calibration point, we discovered that parameter combinations spanned the majority of the parameter space (refer to Figure 2). Investigation of demarcation between null space and calibration space described on lines 291-299 indicated that the inclusion of parameter combinations outside the selected null space resulted in larger simulated temperature ranges than warranted. We therefore concluded that applying NSMC to a single calibration point does not underestimate the soil property uncertainty in our case, even though this will not necessarily be true in other cases. We have added a paragraph on lines 246-250 to clarify this to the reader.

The critical point of the reviewer was that the model cannot reproduce the temperatures accurately, especially during freeze thaw cycles as well as for the summer thawed period where the model predicts warmer/thawed temperatures (especially visible in figures of the Appendix A2).

#### Editorial comments

- The structure of paper should follow the order intro, methods, results, discussion, conclusion (as outlined in ..) with clear headers. See also TC guidelines: [http://www.the-cryosphere.net/for\\_authors/manuscript\\_preparation.html](http://www.the-cryosphere.net/for_authors/manuscript_preparation.html)
- The manuscript includes (too) many figures. Please differentiate the figures into the relevant sections (method/results/conclusion/appendix)- which ones are essential results and which ones can be moved into the appendix or supplementary information?
- All figures need to be checked for correct format
- All figure captions should include the necessary information about the displayed data series
- Important information from the GMD paper (Atchley et al) needs to be included, if necessary for understanding the content of this paper

#### Major comments

-L 128-128

Then an additional surface/subsurface calibration was performed to verify that the surface energy balance model is capable of producing surface temperatures consistent with measurements.

Where is this shown?

-L 158-162

The climate model uncertainty is epistemic in nature due to a lack of knowledge regarding modeling of atmospheric phenomena. These distinctions do limit comparisons that can be drawn between these two uncertainties. However, the comparison is relevant for our purposes to provide a frame of reference for soil property uncertainty to one of the other current, primary sources of permafrost thaw uncertainty.

I do not understand the rationale here- why look at different climate models when looking at soil property uncertainty?

-L 203-204

A subset of the 16 soil parameters from the calibration of Atchley et al. (2015) are included here and presented in Table 1.

How big is this range (I expect a small range)? Is this reasonable? There are lots of data available from Barrow, not only from polygons.

-L 220-225

The minimum and maximum parameter boundaries are modified from the calibration for the NSMC sampling (the parameter ranges are reduced in most cases) to physical limits identified through literature review and field observations from the BEO (Hinzman et al., 1991, 1998; Lawrence and Slater, 2008; Letts et al., 2000; Beringer et al., 2001; Overduin et al., 2006; O'Donnell et al., 2009; Quinton et al., 2000; Nicolsky et al., 2009; Zhang et al., 2010).

Please clarify “from the BEO”(Hinzman et al. from Innvait Creek, Overduin et al. from Gailbraith lake,..). These literature citations are from various sites in Alaska, but do not cover a wider literature review (for example Siberian sites).

-L 231 ff

Figure 1 presents histograms while Fig. 2 presents paired plots of the NSMC ensemble soil pa

Are these now results? Not clear.

-L 253-254

The range in RMSE values is from around 0.55 to 0.65\_C. The accuracy of the temperature probes are 0.02\_C.

The accuracy of your temperature is at best 0.1°C. Please correct and report the corrected percentage of the RMSE.

-L 261-263

The measured temperatures are within the 95% confidence band 79% of the time for the center, 59% for the rim, 46% for the trough, and 61% overall. The primary causes of these discrepancies are due to difficulties in capturing trends that are not purely random.

Why the differences? What is meant with “ trends that are not purely random”.

It looks that especially the phase change in spring is often not well reproduced.

You are using temperature data from center, rim, trough, thus these sites should differ in their (unfrozen) volumetric water contents because of their microtopography. Can you explore the limits of uncertainty further?

-L 266-271

Many physical processes may be leading to this result. For one, the exposed sides of the rim and subsequent lateral heat flow are not explicitly modeled and may at least partially explain the discrepancy. During the thaw, a lack of advective transport of heat by liquid water through the pore space created by sublimation during the winter (not included in the model) may result in warmer measured temperatures.

Please support this statement either through other citations or results.

-L 296-298

The ALT defined that way would be the minimum of the maximum annual thaw depth over each two year moving window. We use a less arbitrary definition for the ALT here as the annual maximum thaw depth, similar to Koven et al. (2011).

The definition of active layer is “The layer of ground subject to annual thawing and freezing in areas underlain by permafrost (<http://www.uspermafrost.org/glossary.php>).

-L310-312

..., this can be reduced to simply meters, however, it must be recognized that the metric is averaged over the entire year including while the soil column is completely frozen.

Please correct sentence.

-L 312

D is a rough proxy for the potential for soil organic matter decomposition.

Freeze curve times are excluded from this, but activity also possible below °C (within freezing and thawing curves when soil is not completely frozen).

Why is discussion on soil organic decomposition included here? It is not part of the model results in this paper. I suggest removing this discussion, including discussion on decomposition, and speculations on future soil moisture/temperature.

-L 316

In addition, the soil organic matter content in soils generally decreases with depth, which is not accounted for in the D metric.

This is not a correct assumption for permafrost soils, see for example Schirrmeister et al. (2011).

Schirrmeister, L., G. Grosse, S. Wetterich, P. P. Overduin, J. Strauss, E. A. G. Schuur, and H.-W. Hubberten (2011), Fossil organic matter characteristics in permafrost deposits of the northeast Siberian Arctic, *J. Geophys. Res.*, 116, G00M02, doi:10.1029/2011JG001647

-L 326-331

Suggest omitting this section. Simply state that hydrology is coupled to biogeochemical fluxes.

-L 314 Permafrost metric

The detailed description of the permafrost parameters and the rationale why using them should be in the intro/method section (prior to results).

-L 322 Annual thaw depth duration

How does this number consider the importance of earlier spring/summer thawing of AL?

-L 347 ff, modified Stefan number

I do not understand why this is beneficial?

-Figure 6

Why only air temperature? Please add snow depth.

-Figure 7:

Why “interannual variability” when only days 285-291 are shown?

Please give information on which subplot you are referring to.