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

Dear Editor,

We would like to thank you for providing significant comments that have helped clarify and make this paper a stronger submission. The reviewers' and your comments aided us in splitting our manuscript between the existing literature, the results that we build on which were previously published in Atchley et al, the results which are new here, and discussion about those results which lead to future work. As such we have made significant effort to reorganize the paper to clarify this, and hope this reorganization helps readers to properly place this work in context.

For ease of reference, the paper is now organized as:

1. Introduction
2. Methodology
    2.1. Model
    2.2. Previous calibration from Atchley et al (2015)
    2.3. Soil property uncertainty quantification
    2.4. Permafrost projections through 2100
    2.5. Permafrost metrics
        2.5.1. Active layer thickness (ALT)
        2.5.2. Annual thaw depth-duration ($\overline{\mathcal{D}}$)
        2.5.3. Annual mean liquid saturation ($\overline{S}_l$)
        2.5.4. Stefan number (S_T)
    2.6. Comparison to climate uncertainty
3. Results
    3.1. Ensemble of calibration-constrained soil parameter combinations
    3.2. Permafrost thaw projection uncertainty
    3.3. Comparison to climate model structural uncertainty
    3.4. Dependence of permafrost projections on soil parameters
4. Discussion and Conclusions

This is a more logical presentation of the material and we thank you for prompting this revision.

Additionally, based on the your suggestion, we have moved the 3 figures that were in the appendix to the supplement, and two additional figures that were in the main text to the supplement. Your suggestion has allowed us to focus our analysis and results and simplify our presentation to readers. We thank you for this suggestion. Figure 3 has been significantly modified to include an evaluation of the 2013 calibration from Atchley et al. (2015) with currently available 2014 data. Figure 5 has been significantly modified to include snow depths.

All responses and revisions are detailed below. We provide a tracked-changes pdf to aid in identifying these changes. Due to the large number of changes, the section numbers in the tracked-changes pdf are not correct (latexdiff was not able to keep up with changes). Please

refer to the revised manuscript for those. The tracked-changes pdf contains revisions from the original submission, therefore including revisions from both reviewer comments and the current revisions detailed below. To aid in identifying revisions, line numbers are provided below indicating the location in the tracked-changes pdf of revisions. These will be different than in the revised manuscript, but will allow easier identification of changes. Given the extent of the revisions, some revisions apparent in the tracked-changes pdf are not directly called out in the responses below.

We feel the manuscript is a much clearer description of our work, and thank the editor and other reviewers for their careful attention, constructive comments and critical analysis of our work. We realize that this is extremely time consuming.

Regards,
Authors

Reviewer comments (italic)
Our abbreviated responses to reviewers
Editor's comments (blue)
Our responses (green)

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.

We agree that our manuscript does not provide many details of the site, but instead focuses on uncertainty quantification relying on previous publications for site information. However, in order to provide the reader with a better understanding the site at Barrow, AK we added descriptions on lines 111-126, we also added information about the snow cover at the site here, lines 119-123. We again thank you for your suggestion and believe that this change has made for a more complete manuscript.

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.

We thank the editor for catching our misunderstanding of the reviewer comment. To clarify, in Atchley et al. (2015), a two-step calibration process was used. The first step involved a subsurface only calibration step where we set up the calibration model with top temperature boundary condition using measured 2cm deep borehole temperatures. This is the calibration used in the analysis of the current manuscript. Atchley et al. (2015) did perform an additional coupled surface/subsurface calibration of surface energy balance parameters, but the subsurface parameters were not varied. Our current analysis builds off of the subsurface calibration results. Now that we are aware of the intent of the reviewers comment we realize that there is confusion as to why we didn't compare actual 2013 data with the CESM projected data. We didn't see this as necessary because neither the actual forcing data for 2013 nor the CESM projection for 2013 are a result of our simulations, but rather inputs. We believe that a comparison of atmospheric variables is not applicable to the current manuscript, since the purpose of our manuscript is to evaluate the role of subsurface uncertainty in ALT projection and not the difference between climate model projections and actual data, which, of course, is the subject of a very large existing literature. However, now realizing this point may be lost in the reader we clarify this, the Methodology section now contains a section devoted to describing the previous calibration in Atchley et al and the model used there (Section 2.2, lines 173-204) while the model used for the projections is described in Section 2.1 and referred to in a subsection of the Methodology section dedicated to describing the projections and the model used there (Section 2.4). This clarifies that we are referring to two models, the calibration model and the projection model. In addition, we have expanded our discussion of the climate forcing data to include a comparison to long-term averages at the site.

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]

We thank both the editor and reviewer for addressing how uncertainty in snow depth can affect subsurface temperatures, especially during winter. This is important research that we believe needs to be addressed, and indeed has been addressed somewhat in literature already (Zhang, 2005; Hinkle and Hurd, 2006; Langer et al., 2013; Atchley et al., 2015). However, the focus of this manuscript is the influence of subsurface soil, property uncertainty in ALT as stated in the title, and at lines 2 and 92-94. Nevertheless, given the importance of snow to the subsurface thermal regime we now summarize the snow model in Section 2.1 (lines 161-165). Furthermore

the time series of the ensemble snow depth is also shown in figure 5. Again we thank the editior and reviewer here for pointing out needed research that we believe deserves its own work.

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.

We agree that a longer temperature record is desirable and that representing multiple landscape units is important, which is precisely why we calibrate the three dominant topographical features in polygonal tundra, the relatively dry rim, intermediate center, and consistently wet trough. While there is a lot of data available at Barrow, the high vertical resolution borehole temperature in adjacent rim, centers, and troughs, which we need is limited. We now have 2014 data available. In order to test whether or not 2014 data would significantly alter the calibration, we have collected and processed the 2014 data, performed forward simulations, and compared to simulated and measured temperatures. Our results indicate that the 2013 calibrated parameters produce 2014 temperatures that are as consistent with measured temperatures as for 2013.

This comment has led us to update figure 3 to include the 2014 evaluation of the 2013 calibration, greatly facilitating reader's ability to evaluate the utility of the calibrated parameter beyond the calibration period. This is discussed on lines 200-206 and 411-417. The approach of splitting the available data into a calibration period and an evaluation period is considered best practice for this type of parameter estimation study, and we are confident it significantly improves the paper. We thank the reviewer and editor for prompting this contribution to the paper.

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.

We agree with the reviewer that a joint inversion of temperature and water content would be better. Unfortunately, water content data was not collocated with the temperature boreholes at our study site. This is often the case that modelers must make the most out of the data that is available. Given the relative ease with which temperatures can be collected, we imagine that this will be a recurring scenario for many modelers for years to come. Additionally, it should be noted that temperatures are dependent on moisture content. As temperatures are available throughout soil column, temperature at depth is a direct response to the thermal conductivity of the mixture continua – thermal conductivity is strongly dependent on water/ice content, and so temperature gradients with depth incorporate significant soil moisture information. This is discussed now on lines 105-110 in greater detail. It should also be noted that in performing decoupled calibrations (separate calibrations on polygon center, rim, and trough profiles) in Atchley et al. (2015) that the mismatch during thaw and freeze-up were less pronounced. The mismatch increases during the coupled calibration, as the fit is a compromise to fitting temperatures in all three profiles while requiring that the soil parameters are the same. The coupled calibration provides soil properties that are generalized across profiles. The cost of this generalization is compromises in the fit to measured temperatures. The coupled and decoupled calibration approaches both have their merits. We chose the coupled calibration approach in order to obtain more generalized soil properties that perhaps do have greater mismatch during freeze and thaw periods, but do capture ALT in a generalized fashion across the polygon microtopography. We have added a discussion to lines 425-442, that discusses this in greater detail. We also state in the Conclusions and Discussions section that collocated temperature and water content measurements would better constrain soil properties lines 727-735.

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).

In this comment the reviewer addresses two concerns – 1. That the model does not describe the data, and 2. That there might be many other local minima, and by not finding them all we are underestimating uncertainty. We feel we have responded to the second concern appropriately in our initial response. However, we agree that we first concern was not appropriately addressed. We thank the editor for catching this omission and prompting us to provide a proper response to the reviewer. Given the significant complexities and nonlinearities in the integrated surface/subsurface thermal hydrology processes modeled, the general challenge of adequately accounting for subsurface heterogeneity in models, and that we are using a single set of model parameters in 1-D models at rim, center, and trough locations, we believe the fit between modeled and observed temperatures is remarkable good. Transient mismatch during freezing/thawing periods, although interesting scientifically and possibly pointing to neglected processes such as lateral flow, will not have an impact on our effect of interest, active layer depth. We've modified the manuscript to include this discussion.

While the reviewer is correct in pointing out that the biggest errors are in the most dynamic time periods, this is not surprising and does not mean that the model is not useful. No models are perfect; this one certainly isn't. However, as shown in Atchley et al, it is significantly better than several other, simpler models that were explored. We do plan to investigate these discrepancies more in the future using 2D and 3D models to ascertain the effect of lateral flow and thereby enabling the use of water table data.

Additionally, in performing decoupled calibrations (separate calibrations on polygon center, rim, and trough profiles) in Atchley et al. (2015), the mismatch during thaw and freeze-up were less pronounced. The mismatch increases during the coupled calibration, as the fit is a compromise to fitting temperatures in all three profiles while requiring that the soil parameters are the same. The coupled calibration provides soil properties that are generalized across a polygon. The cost of this generalization is compromises in the fit to measured temperatures. The coupled and decoupled calibration approaches both have their merits. We chose the coupled calibration approach in order to obtain more generalized soil properties that perhaps do have greater mismatch during freeze and thaw periods, but still capture ALT in a generalized fashion across the polygon microtopography.

We have added a discussion to section 3.1 that discusses this in greater detail (lines 425-442). We also discuss that collocated temperature and water content measurements would better constrain soil properties, but that these types of datasets were not available and will not always be available (lines 105-110 and 727-735).

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

As stated above, the paper has been reorganized into a clearer structure based on the comment. We agree that the paper is more easily digested in this form, and appreciate the editor's effort in helping make this a better manuscript.

- 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?

Again we thank the editor for suggesting a re-organization as this did indeed help eliminate un-necessary figures and draw attention to necessary illustrations. Because of the re-organization the temperature confidence band time series previously in the Appendix, the convergence analysis plot, and the permafrost metric correlation paired plot have been moved to supplemental information. All other plots are crucial to the results and would be referred to too often to be considered appropriate for supplemental information, or they provide a unique visual perspective on results.

- All figures need to be checked for correct format

We have closely evaluated all included figures to ensure that they are the in the correct format for the Journal.

- All figure captions should include the necessary information about the displayed data series

Captions have been reviewed for completeness and augmented in many cases. Please refer to captions in Figures 1, 2, 3, 5, 6, 7, 8, 9, 10, 11, and 12.

- Important information from the GMD paper (Atchley et al) needs to be included, if necessary for understanding the content of this paper

We agree that our manuscript needs to summarize content in Atchley et al. In order to provide the reader with a better understanding the site at Barrow, AK we added descriptions at line 111-126, we also added information about the snow cover at the site here, lines 119-124. We again thank you for your suggestion and believe that this change has made for a more complete manuscript. The snow model from Atchley et al is now summarized in Section 2.1 (lines161-165).

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?

A citation has been added here to clarify that we are still referring to the Atchley et al. paper

here and not the current manuscript. The calibration is described in Section 4 of Atchley et al. The restructuring of the paper also makes this more apparent since there is now a separate section (section 2.2) titled "Previous calibration from Atchley et al. (2015)". A clear reference to the additional surface/subsurface performed by Atchley et al. is now on lines 190-194.

-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?

We quantify the uncertainty due to a set of available climate models to provide some context for the soil property uncertainty that we observe. Reporting the magnitude of uncertainty without a reference point for comparison can be difficult to interpret. In our experience, most arctic researchers are familiar with the uncertainty associated with structural differences between climate models. Therefore, this comparison allows readers to gauge relative effect of soil property uncertainty with respect to another well-known source of uncertainty. Text has been added to lines 23-24, 101-102, 608-609. This is also discussed in section 2.6 (Comparison to climate 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.

The ranges are based on an extensive literature review and field data from the BEO. Since there are many conflicting reports on material properties, we use ranges that accommodate the majority of the available information. It should also be noted that the ranges in Table 1 do not represent the uncertainty used in the analysis. The ranges listed in Table 1 are used as parameter bounds in the calibration, where the end result is a set of calibrated parameters. The Null-Space Monte Carlo analysis then collects samples within the parameter bounds that are also consistent with the temperature data. These ranges will be a subset of the calibration ranges, as presented in histogram format in Figure 1. This is discussed on lines 373-377. The distributions indicated in Figure 1 should be taken as the uncertainty in material properties used in the projections. A description of the references and locations used to inform the parameter ranges are now provided on lines 365-372.

-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 Imnavait 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).

Again we thank the editor for the suggestion of placing our parameter values from literature in a global context as this provides a broader significance for this manuscript. Here we would like to point out that the parameter values were attained from sites that spanned both Alaska and Canada. Furthermore, as was suggested by the editor, informing our parameter range selection from Siberian sites also helps to increase the significance from this work, and while we didn't initially include Siberian sites we have worked to consider them in revisions of this work and our ongoing work. The locations of referred literature is now detailed on lines 365-372. Based off of this suggestion and the need to clarify calibration parameter space from the above comment, we now discuss this at lines 355-357.

-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.

This section is now part of the Results section. We thank the reviewer for prompting us to make this clarifying revision.

-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.

We admit that the manufacturer provides a resolution of 0.1C for their probes. However, Vladimir Romanovsky's lab has repeatedly calibrated these probes and measured higher accuracy – down to 0.02°C. We assume the discrepancy is conservatism on the part of the manufacturer. In accordance with the manufacturer and the editor's comment, we have changed the number to 0.1°C (lines 405-407).

-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?

We thank the editor for catching this error in the text. “trends that are not purely random” has been replaced with “trends during the freeze-up and thaw of the active layer” (lines 427-428). We agree that exploring the limits of uncertainty further through the effects of microtopography is important. Some of this type of analysis can be found in Atchley et al. We agree that this is important and incomplete work and plan to explore this using 2D and 3D models in the future. In particular, we believe that it will be necessary to include lateral flows to further refine the thermal hydrology representations. In order to keep the current research results concise, this type of analysis has not been performed here, instead we have produced parameters that generalize characteristics across the microtopography of the site in a coupled calibration. This is now discussed on lines 430-442. This is also discussed on lines 183-188.

-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. 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.

We provide a reference to Kane et al. (2001) (line 442), which describes non-conductive heat transfer mechanisms in frozen soil. It is speculative at this point that this is the cause of the mismatch, but supported by literature and worth mentioning.

-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).

This definition has been added to section 2.5.1 (line 233-234). We thank the editor for bringing this definition to our attention.

-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.

This sentence has been removed to avoid confusion.

-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.

The Active Layer itself, while important, is also a means to predicting carbon decomposition and feedbacks to the broader climate. As such, biogeochemists are extremely interested in understanding ALT dynamics, and discussions on how our ALT understanding impacts SOM decomposition is very relevant to properly capturing the implications of this work. This discussion helps biogeochemists to place the relevance of this work to their field, and broadens the audience of this work. We appreciate the editor's point that D does not account for biological activity below 0 C. A statement has been added to the definition of the metric along those lines (lines 253-255). This, along with the other stated limitation, will provide beogeochemists with the necessary caveats to understand what information D provides and how it can inform their work. Our hope is that providing such clearly defined hydrothermal metrics from our work, although currently limited in some respects, will lead to further integration and cross-communication within the arctic modeling community.

-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

Since we agree that this is not necessarily true in all cases, this sentence has been removed.

-L 326-331
Suggest omitting this section. Simply state that hydrology is coupled to biogeochemical fluxes.

This section has been modified as suggested in order to make a more concise paper (Section 2.5.3).

-L 314 Permafrost metric
The detailed description of the permafrost parameters and the rationale why using them should be in the into/method section (prior to results).

This section has been moved to the Methodology section (Section 2.5 and subsections).

-L 322 Annual thaw depth duration
How does this number consider the importance of earlier spring/summer thawing of AL?

Since the metric in question integrates the thaw depth over the entire year, it directly accounts for earlier thawing of the AL in a warming climate. Indeed, this is the point of including duration as well as averaged depth. Admittedly, we don't address changes in vegetation, which would be affected by the longer growing season. We include a discussion of this limitation in the revised manuscript. As is, the metric does provide a well-defined and objective metric that synthesizes a large amount of information regarding permafrost thaw, with clear limitations that

we describe. Nevertheless, this type of metric is ideal for synthesizing a large quantity of information and providing an integrated quantity characterizing general permafrost characteristics and how those characteristics change over time and their uncertainty.

-L 347 ff, modified Stefan number
I do not understand why this is beneficial?

The Stefan number quantifies the partitioning of energy into the component that causes changes in subsurface temperature (conduction) versus what is consumed in thawing (latent heat of phase change). Decreases in the Stefan number, as projected here, indicate that the component consumed in thawing dominates the ratio, leaving less energy available to increase the temperature of the soil. The metric is objective and provides an easily interpreted perspective on permafrost thaw in a warming climate. This metric will be of particular interest in identifying the relative importance of conduction vs latent heat processes as permafrost thaws. This is a basic indicator of energy partitioning in the subsurface that permafrost researchers should understand and consider. Text along these lines has been added to lines 16-18, 273-275, 712-714.

-Figure 6
Why only air temperature? Please add snow depth.

We thank the editor for the suggestion. Snow depth has been added to this plot allowing readers to evaluate its effects on thaw depth.

-Figure 7:
Why “interannual variability” when only days 285-291 are shown?

We apologize for the confusion here. The plot did state that it is the intra-annual uncertainty. Based on your comment, we have modified all references to intra-annual uncertainty to predictive or parametric uncertainty throughout the text. We hope that this is less confusing and feel that it is more appropriate. We also ensure that we are clear that inter-annual variability is due to yearly variations in single climate model (e.g., CESM). We thank the editor for prompting this clarifying revision.

Please give information on which subplot you are referring to.

We thank the editor for identifying this error. The reference to ‘subplot’ has been removed.

Harp Atchley Painter Coon Wilson Romanovsky Rowland
Manuscript prepared for The Cryosphere
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 10 November 2015

Effect of soil property uncertainties on permafrost thaw projections: A calibration-constrained analysis

Dylan R. Harp[1], Adam L. Atchley[1], Scott L. Painter[2], Ethan T. Coon[1], Cathy J. Wilson[1], Vladimir E.Romanovsky[3], and Joel C. Rowland[1]

[1]Earth and Environmental Sciences Division, Los Alamos National Laboratory, Los Alamos, NM, USA
[2]Climate Change Science Institute, Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA
[3]Geophysical Institute, University of Alaska Fairbanks, USA

Correspondence to: Dylan R. Harp (dharp@lanl.gov)

Abstract. The ~~effect~~ effects of soil property uncertainties on permafrost thaw projections are studied using a three-phase subsurface thermal hydrology model and calibration-constrained uncertainty analysis. The Null-Space Monte Carlo method is used to identify soil hydrothermal parameter combinations that are consistent with borehole temperature measurements at the study site, the Barrow Environmental Observatory. Each parameter combination is then used in a forward projection of permafrost conditions for the 21^{st} century (from calendar year 2006 to 2100) using atmospheric forcings from the Community Earth System Model (CESM) in the Representative Concentration Pathway (RCP) 8.5 greenhouse gas concentration trajectory. A 100-year projection allows for the evaluation of ~~intra-annual uncertainty~~ predictive uncertainty (due to soil ~~properties~~ property (parametric) uncertainty) and the inter-annual climate variability due to year to year differences in CESM climate forcings. After calibrating to measured borehole temperature data at this well-characterized site, soil property uncertainties are still significant and result in significant ~~intra-annual~~ predictive uncertainties in projected active layer thickness and annual thaw depth-duration even with a specified future climate. ~~Intra-annual uncertainties~~ Inter-annual climate variability in projected soil moisture content and Stefan number are small. A volume and time integrated Stefan number decreases significantly~~in~~ , indicating a shift in subsurface energy utilization in the future climate ~~, indicating that~~ (latent heat of phase change becomes more important than heat conduction~~in future climates~~). Out of 10 soil parameters, ALT, annual thaw depth-duration, and Stefan number are highly dependent on mineral soil porosity, while annual mean liquid saturation of the active layer is highly dependent on the mineral soil residual saturation and moderately dependent on peat residual saturation. By comparing the ensemble statistics to the spread of projected permafrost metrics using different climate mod-

els, we quantify the relative magnitude of soil property uncertainty to another source of permafrost uncertainty, structural climate model uncertainty. We show that the effect of calibration-constrained uncertainty in soil properties, although significant, is less than that produced by structural climate model uncertainty for this location.

~~This research was supported by the Next-Generation Ecosystem Experiments Arctic (NGEE-Arctic) project (DOE ERKP757) funded by the Office of Biological and Environmental Research in the US Department of Energy Office of Science and Los Alamos National Laboratory's Laboratory Directed Research and Development (LDRD) Arctic project (LDRD201200068DR). This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05- 00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.~~

1 Introduction

Increasing Arctic air and permafrost temperatures (Serreze et al., 2000; Jones and Moberg, 2003; Hinzman et al., 2002; Romanovsky et al., 2007), the resulting increase in the thickness of soil that thaws on an annual basis (Romanovsky and Osterkamp, 1995), and the potential for greenhouse gas release due to the ensuing decomposition of previously frozen organic carbon (Koven et al., 2011; Schaefer et al., 2011) provide motivation for developing robust numerical projections of the thermal hydrological trajectory of Arctic tundra in a warming climate. Projections of permafrost thaw and the associated potential for greenhouse gas release from the accelerated decomposition of previously frozen carbon are subject to several sources of uncertainty, including (but not limited to) structural uncertainties in the climate models; uncertainty about the model forcings/inputs in the future (scenario uncertainty in the typology of Walker et al. (2003)); parametric uncertainties in soil and surface properties that control the downward propagation of thaw fronts; and structural uncertainties in the surface and subsurface thermal hydrological models.

Previous efforts to characterize uncertainty in permafrost thaw projections have mostly focused on climate model structural uncertainties and climate ~~scenario~~ model uncertainties, presumably because of an implicit assumption that those two sources of uncertainty overwhelm the other sources. However, recent large-scale model comparisons suggest that a substantial portion of projected permafrost uncertainties is a result of structural model differences in land surface/subsurface schemes (Slater and Lawrence, 2013; Koven et al., 2013), particularly how subsurface thermal hydrologic processes are represented (Koven et al., 2013) rather than simply climate variation. Although those studies focused on structural uncertainty in surface and subsurface models and not on soil property uncer-

tainty, the reported sensitivity to the subsurface model suggests that uncertainty in soil properties may also contribute significantly to overall uncertainty in thaw projections.

The bulk hydrothermal properties of soil that control the active layer thickness (ALT, i.e. the depth of soil that thaws on an annual basis) (Neumann, 1860; Stefan, 1891; Romanovsky and Osterkamp, 1997; Peters-Lidard et al., 1998; Kurylyk et al., 2014) vary among sites and locally within a single site, in particular being sensitive to the local organic matter content and bulk porosity (Letts et al., 2000; Price et al., 2008; O'Donnell et al., 2009; Hinzman et al., 1991; Chadburn et al., 2015a). Langer et al. (2013) identify the soil composition uncertainties, particularly the soil ice/water content, to have the largest effect on ALT. Intermediate to large-scale thermal simulations of ALT are known to be sensitive to soil properties (Hinzman et al., 1998; Rawlins et al., 2013). Because of this sensitivity, large-scale Earth System Models (ESMs) were recently updated to include layers of moss and peat in order to better represent subsurface thermal conditions (Beringer et al., 2001; Lawrence and Slater, 2008; Wania et al., 2009; Subin et al., 2012; Ekici et al., 2014; Chadburn et al., 2015b). Despite the recognition of soil property uncertainty and heterogeneity as important contributors to uncertainties in permafrost conditions and extent, global and regional studies that address permafrost future conditions and extent typically apply broad soil texture classifications, such as those defined by Clapp and Hornberger (1978) and Cosby et al. (1984), to parameterize soil properties (Lawrence and Slater, 2008), usually without consideration of soil property uncertainty (Lawrence and Slater, 2005; Hinzman et al., 1998; Shiklomanov et al., 2007; Koven et al., 2013; Rinke et al., 2008).

Soil property uncertainty is different from many other sources of projection uncertainty (e.g. climate ~~scenario~~ model uncertainty) in that uncertainties in soil properties may be reduced by a combination of site characterization (Hinzman et al., 1998) and model calibration (Romanovsky and Osterkamp, 1997; Nicolsky et al., 2009; Jiang et al., 2012; Atchley et al., 2015). Initial steps in that direction have been taken. For example, Romanovsky and Osterkamp (1997) calibrate thermal soil properties using a purely conductive thermal model using measured temperatures at several sites and Nicolsky et al. (2009) perform a sensitivity analysis of a calibration (data assimilation) approach to identify its ability to recover thermal soil properties using a 1D thermal model and apply the calibration approach to several sites. Atchley et al. (2015) recently demonstrated an iterative approach for using site characterization data to simultaneously refine thermal hydrology model structure and estimate model parameters. Their approach was applied to the Barrow Environmental Observatory, but could be used at other sites to improve model structure and parameter assignments in the regional or global context.

Recognizing that permafrost projections are sensitive to subsurface model representations and that soil property uncertainties may be reduced through characterization and parameter estimation, a natural next step is to quantify how such activities will impact overall uncertainties in permafrost thaw projections in comparison to other sources of uncertainty. Here we address that question. Specifically, we consider how uncertainties in soil hydrothermal properties propagate to uncertainties in

numerical projections of permafrost thaw at a well-characterized site. We go beyond a traditional unconstrained uncertainty quantification and focus on the residual uncertainties that remain after soil parameters have been carefully calibrated to borehole temperature data. The intent of the current work is to develop initial insights into how effective site characterization activities might be at reducing uncertainties associated with soil parameters. We show that with future climate specified and with the advantage of calibration targets from a well-characterized site, significant uncertainties remain in projected ALT and other metrics important for carbon decomposition in the future climate. We evaluate both predictive uncertainty and inter-annual climate variability. We show that this residual uncertainty is significant, albeit less than that associated with uncertainties in future climate.

~~The methodology is described in Sect. 2. A brief description of our thermal hydrology process model is presented in Sect. 2.1. The generation of the ensemble of calibration-constrained parameter combinations~~ We focus on temperature data as they are one of the easiest and most common types of soil data to collect at field sites and are often used for early site characterization. While many sites may have other types of measurements available, such as water and ice content measurements, many of these are more difficult to obtain at regular temporal intervals for extended periods of time. The incorporation of other types of data, such as water and ice content measurements, would be expected to reduce soil property uncertainty, however this is not investigated here.

The arctic site in this investigation is the polygonal tundra within the Barrow Environmental Observatory (BEO) on the Seward Peninsula. In particular, we focus on NGEE-Arctic site "area C" which contains degraded permafrost characterized by ~50 cm deep troughs and shallow low centers. The polygonal tundra of the BEO is classified as a lowland, cold continuous permafrost system with a range of polygonal types and states, which includes intact low center polygons to degraded ice wedges and associated high center polygons. Much of the polygonal tundra contains an organic rich surface layer of peat overlaying a silty loam soil. Due to a low evaporative demand soils remain moist, despite relative low annual precipitation, of which the bulk falls in the summer months (Liljedahl et al., 2011). The snowpack over the microtopography at the site is redistributed to a relatively level surface by strong winds, resulting in the deepest snowpack over troughs. Snow depth measurements collected around the site on May 2, 2013 were between 20-40 cm for centers, 10-20 cm for rims, and 40-60 cm for troughs while the average snow density was 326 kg/m^3 (Atchley et al., 2015). While our investigation focuses on the polygonal tundra within the BEO, other arctic landscape types are also prevalent (hillslopes, lakes, pingos). The importance of soil properties and the dominant influence of particular soil properties may change in landscapes other than polygonal tundra.

The methodology is described in Sect. ~~3.1.Permafrost~~ 2, including: the model description (Sect. 2.1); a review of the calibration performed in Atchley et al. (2015) (Sect. 2.2); soil property uncertainty quantification approach (Sect. 2.3); permafrost projection approach (Sect. 2.4); description of permafrost thaw projection metrics ~~are described~~ (Sect. 2.5); and method of comparison to climate uncertainty

(Sect. 2.6). Results are presented in Sect. ~~2.5.The~~ 3, including: the ensemble of calibration-constrained parameter combinations (Sect. 3.1); predictive uncertainty and trends in permafrost thaw projections ~~are presented in~~ (Sect. 3.1~~. Sect. 3.1 presents the~~); comparison of soil property and climate model uncertainty ~~. A correlation analysis identifying the level of dependence~~ (Sect. 3.1); and correlation analysis between soil parameters and projection metrics ~~is presented in~~ (Sect. 3.1). Conclusions and discussion of the analysis are in Sect. 4.

2 Methodology

2.1 Model

We use the Arctic Terrestrial Simulator (ATS) to numerically solve the coupled groundwater flow, thermal, and surface energy balance equations. ATS is an integrated thermal hydrological code developed specifically for Arctic permafrost applications. It implements the modeling strategy outlined by Painter et al. (2013) using the multiphysics framework Arcos (Coon et al., 2015b) to manage model complexity in process rich simulations such as these. Various components of ATS have already been described elsewhere, therefore, only a brief summary is provided here.

In the subsurface, the ATS solves nonlinear conservation equations for water and energy, using a three-phase (air-water-ice), single-component representation (Karra et al., 2014), which is a simplification of a more general two-component (water and representative gas phase) model (Painter, 2011). A recently developed constitutive model (Painter and Karra, 2014) is used to partition water between ice and liquid phases in unsaturated or saturated conditions. The partitioning model relates unfrozen water content below the nominal freezing point to the unfrozen soil moisture characteristic curve, thus avoiding empirical freezing curves. The model has been successfully compared to a variety of laboratory experiments on freezing soils (Painter and Karra, 2014; Karra et al., 2014; Painter, 2011). The Material Component model defines thermal conductivities and is described in detail in Appendix A of Atchley et al. (2015). Surface boundary conditions use a "fill and spill approximation", where we allow up to 4 cm of water to pond on the surface; all additional ponded water may run off the domain. The surface and subsurface thermal hydrology systems are coupled using continuity of pressure, mass flux, temperature, and energy flux, in a thermal extension of the coupling strategy presented in (Coon et al., 2015a). Additionally, we use a surface energy balance (Hinzman et al., 1998; Ling and Zhang, 2004; Atchley et al., 2015) in which surface latent and sensible heat, incoming and outgoing radiation, and conducted heat terms, along with incoming precipitation and outgoing evaporation are tracked. Finally, a dynamic, snow model is incorporated for tracking snow aging and consolidation, with resulting effects on albedo and melt (Atchley et al., 2015). As described in Sect. 4.4 of Atchley et al. (2015), the snow model accounts for snow redistribution over the microtopography of the site and depth hoar formation. Additional details about the snow model are described in detail in Appendix B of Atchley et al. (2015). Not represented within this system are

carbon cycle and vegetation processes, including long-term changes of peat composition, variability in peat thickness, and evolving microtopography due to degradation of ice wedges.

The subsurface domain is represented by a 2 cm layer of moss, followed by a 10 cm layer of peat, and approximately 50 m mineral soil layer. The required climate forcings for the ATS models are precipitation of snow and rain, air temperature, wind speed, relative humidity, and incoming short and longwave radiation.

2.2 Previous calibration from Atchley et al. (2015)

The uncertainty quantification is performed around a previous calibration by Atchley et al. (2015). Atchley et al. (2015) used 1D column models representing the dominant microtopographical features (center, rim, and trough of polygonal ground) to calibrate hydro-thermal soil parameters using soil temperatures at the ~~Barrow Environmental Observatory (BEO)~~ BEO measured by the Next Generation Ecosystem Experiments Arctic (NGEE-Arctic) team during calendar year 2013. Initial conditions for the models were generated by completely freezing the fully saturated model from below and then allowing the initial conditions to emerge over a 10-year spin-up simulation using daily air temperatures averaged from 10 years of data as the top boundary condition. This process allowed a shallow vadose zone to develop consistent with field observations. The calibration considered temperatures measured at 9 depths from 10 to 150 cm.

The calibration was performed in a coupled fashion where each 'model run' of the calibration consisted of simulating center, rim, and trough column models with the same soil parameter values for peat and mineral soil. This coupled calibration identifies soil parameters that provide a generalized fit, compromising in a least squares sense to match the data from all three models. An implicit assumption of the coupled calibration is that the soil properties ~~are~~ may be adequately represented as independent of the microtopography. Atchley et al. (2015) first calibrated subsurface properties using 2 cm deep temperatures measured in 2013 as Dirichlet boundary conditions and temperatures measured at the considered depths as calibration targets. Then Atchley et al. (2015) performed an additional surface/subsurface calibration ~~was performed~~ to verify that the surface energy balance model is capable of producing surface temperatures consistent with measurements. The coupled surface/subsurface model allows the use of future climate ~~scenarios~~ models as model forcings to drive hydro-thermal permafrost projections.

~~In order to make projections of hydro-thermal permafrost conditions, we use the surface/subsurface model of . We use the Community Earth System Model (CESM) driven by the Representative Concentration Pathway 8.5 (RCP8.5) greenhouse gas concentration trajectory from year 2006 to 2100 as atmospheric forcings for the surface energy balance of the model. In this way, we hold the climate scenario constant to isolate the effect of soil property uncertainty. RCP8.5 corresponds to a business as usual warming scenario with 8.5 Wm^{-2} forcing by 2100.~~ The calibration data period is limited to calendar year 2013 since at the time of calibration, this was the only full year

of high-resolution borehole temperatures available at the site (Atchley et al., 2015). Subsequently, year 2014 data has become available. To verify that the calibration has extracted the hydrothermal properties of the system independent of the climatic conditions during the calibration, we evaluated the ability of the calibrated parameters to produce forward simulations that are consistent with 2014 data. This evaluation is presented in the results section.

2.3 Soil property uncertainty quantification

We generated an ensemble of 1,153 calibration-constrained parameter combinations by the Null-Space Monte Carlo (NSMC) method (Doherty, 2004). The NSMC approach samples from insensitive regions of the parameter space (i.e. the null space) determined by an eigenanalysis of parameter sensitivities calculated at the calibration point. Based on analysis of ensemble forward simulations of the calibration year (2013) and a convergence analysis of the ~~95th~~ 95% confidence band of simulated temperatures, we consider all parameter combinations in the ensemble calibrated and equally consistent with measured temperatures.

2.4 Permafrost projections through 2100

In order to make projections of hydro-thermal permafrost conditions, we use the surface/subsurface model described in Sect. 2.1. We use the Community Earth System Model (CESM) (Gent et al., 2011) driven by the Representative Concentration Pathway 8.5 (RCP8.5) greenhouse gas concentration trajectory (Moss et al., 2008) from year 2006 to 2100 as atmospheric forcings for the surface energy balance of the model. The CESM output was adjusted to fit current climate at the BEO. In this way, we hold the climate scenario constant to isolate the effect of soil property uncertainty. RCP8.5 corresponds to a business as usual warming scenario with 8.5 Wm^{-2} forcing by 2100.

The projection simulations took on the order of several hours (~2-4 hours) on a Linux cluster with 3.2 GHz processors. We used the Model Analysis ToolKit (MATK) Python module (http://matk.lanl.gov) to facilitate the concurrent execution of the ensemble of ATS models on Los Alamos National Laboratory high performance computing clusters.

2.5 Permafrost metrics

Predictive uncertainty of projections is determined by comparison of permafrost metrics at year 2006 and for the last decade of the projections (2091 through 2100). The metrics include (1) ALT, (2) annual thaw depth-duration ($\overline{\mathcal{D}}$), (3) annual mean liquid saturation ($\overline{S}_l$), and (4) a modified Stefan number (S_T) and are described ~~in detail in Sect. 2.5.~~ below.

2.5.1 Active layer thickness (ALT)

In general, ALT is defined as "The layer of ground subject to annual thawing and freezing in areas underlain by permafrost" (http://www.uspermafrost.org/glossary.php). Permafrost has also been defined

as the region of the subsurface that remains at or below 0°C for two or more years. 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 in accord with the general definition and similar to Koven et al. (2011). Given the discrete nature of our mesh, and the nonlinear nature of vertical soil temperature profiles near 0°C, we determine ALT as the bottom of the deepest thawed mesh cell (temperature above 0°C) for the year.

2.5.2 Annual thaw depth-duration ($\overline{\mathcal{D}}$)

ALT controls the amount of organic carbon experiencing thaw and thus microbially induced decomposition during a year. Because ALT is defined as the maximum thaw depth, it does not include information on duration of thaw. To quantify increasing duration of thaw in future climate (i.e., the effects of earlier thaw and later freeze-up) as well as increasing depth, a new metric is introduced here: the mean annual thaw depth $\overline{\mathcal{D}}$, defined as

$$\overline{\mathcal{D}} = \frac{1}{365} \int \int H(T(z,t)) dz dt \tag{1}$$

where H is the heavyside function (1 if $T(z,t)$ is above 0°C, 0 otherwise), z is depth in meters, and t is time in days. The fraction on the right side of Eq. (1) normalizes the metric by the 365 days in a year. We express $\overline{\mathcal{D}}$ with units of $\mathrm{m^3 m^{-2}}$ to indicate that this metric defines the volume of thawed soil per unit area. $\overline{\mathcal{D}}$ is a rough proxy for the potential for soil organic matter decomposition. It merges the amount of unfrozen soil and duration that soil is above freezing temperature for a given year. Therefore, the metric does not account for biological activity that occurs below 0°C, which is generally considered to be greatly reduced (Mikan et al., 2002; Davidson and Janssens, 2006), but has been observed in permafrost soils (Sachs et al., 2008). It is noted that, while the annual amount of decomposition is likely correlated with $\overline{\mathcal{D}}$, the two quantities are not directly proportional because soil temperature and moisture will also change and affect the decomposition rates in future climates. Nevertheless, uncertainty in $\overline{\mathcal{D}}$ is of interest as it is indicative of uncertainty in future decomposition rates.

2.5.3 Annual mean liquid saturation ($\overline{S_l}$)

The annual mean liquid saturation $\overline{S_l}$ is defined as

$$\overline{S_l} = \frac{\int \int H(T(z,t)) S_l(z,t) dz dt}{\int \int H(T(z,t)) dz dt} \tag{2}$$

where $S_l(z,t)$ is the liquid saturation as a function of depth and time. $\overline{S_l}$ quantifies the spatially and temporally averaged liquid saturation in the unfrozen soil for a given year. Note that the denominator

in Eq. (2) is the annual thaw depth-duration metric $\overline{\mathcal{D}}$ from above, except without dividing by 365. Liquid saturation within the active layer is of interest because of its control on decomposition rates, coupling hydrology to biogeochemical fluxes.

2.5.4 Stefan number (S_T)

We propose an extension of the Stefan number from the form in Kurylyk et al. (2014) to one that incorporates intra-annual temporal changes and stratified soil properties. The Stefan number is the ratio of subsurface sensible to latent heat. In the current context, this refers to the amount of subsurface heat exchange that results in a change in temperature versus the amount that is consumed in the isothermal conversion of ice to liquid water. The Stefan number provides information about the form of subsurface energy utilization in permafrost regions and is fundamental to a basic understanding of permafrost thaw mechanisms.

In its most basic form, the Stefan number is defined as

$$S_T = \frac{c_b \Delta T}{L_f}. \tag{3}$$

where c_b is the bulk specific heat of the material and L_f is the latent heat of fusion of water (334,000 J kg^{-1}). Kurylyk et al. (2014) define the Stefan number for the permafrost problem as

$$S_T = \frac{c_b \rho_b (T_s - T_f)}{S_{wf} \rho_w \phi L_f} \tag{4}$$

where ρ_b is the density of the thawed zone, T_s is the surface temperature, T_f is the temperature of freezing or thawing soil (taken as 0°C), S_{wf} is the liquid saturation in the thawed zone that was frozen, and ρ_w is the density of liquid water. Kurylyk et al. (2014) use this definition to evaluate the thermal regime of analytical solutions of soil thaw. We expand this definition here to include the increased detail available in our numerical simulations as

$$S_T = \frac{\int\int c_b(z)\rho_b(z)\, H\left(\frac{dT}{dt}\right) \frac{dT}{dt} dz dt}{\rho_{ice} L_f \int\int H\left(-\frac{dS_{ice}}{dt}\right)\left(-\frac{dS_{ice}}{dt}\right) \phi(z) dz dt} \tag{5}$$

where S_{ice} is ice saturation. The integrations are performed over the entire year (i.e. from Jan. 1 through Dec. 31). Equation 5 expands on Eq. (4) to allow the consideration of details of transient heating and cooling throughout the year and stratified hydrothermal soil properties within the soil profile.

2.6 Comparison to climate uncertainty

To provide a reference point for the effect and magnitude of soil property uncertainty, we also perform ATS projections forcing the energy balance model with atmospheric projections from CESM, INM-CM4 (INM) (Volodin et al., 2010), BCC-CSM1-1 (BCC) (Ji, 1995), MIROC (Watanabe et al.,

2010), CanESM2 (CAN) (Verseghy, 1991), and HadGEM2-CC (HAD) (Jones et al., 2011; Bellouin et al., 2011; Collins et al., 2011) climate models based on RCP8.5 using the calibrated (fixed) soil parameters from Atchley et al. (2015). Using the calibrated soil parameters in these simulations isolates the effect of structural climate uncertainty. We compare permafrost projection uncertainty due to the NSMC ensemble of soil parameters (hydrothermal soil property uncertainty) and to the variability between climate models (structural climate uncertainty).

The soil property uncertainty in this analysis is parametric and can be considered more aleatoric/probabilistic in nature. 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.

3 ~~Model~~Results

~~We use the ATS computer code to simulate surface/subsurface thermal hydrology processes. ATS is an integrated thermal hydrological code developed specifically for Arctic permafrost applications. It implements the modeling strategy outlined by using the multiphysics framework Arcos to manage model complexity in process rich simulations such as these. Various components of ATS have already been described elsewhere, therefore, only a brief summary is provided here.~~

~~In the subsurface, the ATS solves nonlinear conservation equations for water and energy, using a three-phase (air-water-ice), single-component representation , which is a simplification of a more general two-component (water and representative gas phase) model . A recently developed constitutive model is used to partition water between ice and liquid phases in unsaturated or saturated conditions. The partitioning model relates unfrozen water content below the nominal freezing point to the unfrozen soil moisture characteristic curve, thus avoiding empirical freezing curves. The model has been successfully compared to a variety of laboratory experiments on freezing soils . Surface boundary conditions use a "fill and spill approximation", where we allow up to 4 cm of water to pond on the surface; all additional ponded water may run off the domain. The surface and subsurface thermal hydrology systems are coupled using continuity of pressure, mass flux, temperature, and energy flux, in a thermal extension of the coupling strategy presented in . Additionally, we use a surface energy balance in which surface latent and sensible heat, incoming and outgoing radiation, and conducted heat terms, along with incoming precipitation and outgoing evaporation are tracked. Finally, a dynamic, single-layer snow model is incorporated for tracking snow aging and consolidation, with resulting effects on albedo and melt . Not represented within this system are carbon cycle and vegetation processes, including long-term changes of peat composition, variability in peat thickness, and evolving microtopography due to degradation of ice wedges.~~

~~The subsurface domain is represented by a 2 cm layer of moss, followed by a 10 cm layer of peat, and approximately 50 m mineral soil layer. The required climate forcings for the ATS models are precipitation of snow and rain, air temperature, wind speed, relative humidity, and incoming short and longwave radiation.~~

4 ~~Creation of ensemble of soil parameter combinations~~

3.1 Ensemble of calibration-constrained soil parameter combinations

In order to determine the effect that calibration-constrained soil property uncertainty can have on long term projections of permafrost conditions, we performed an uncertainty quantification around the calibrated soil parameters of Atchley et al. (2015). The strategy involved identifying a representative set of parameter combinations that all produce simulated temperatures that are consistent with observed temperatures. We use Null-Space Monte Carlo (NSMC) (Tonkin and Doherty, 2009), a form of calibration-constrained Monte Carlo, to accomplish this goal. NSMC was selected based on its sampling economy given the computational burden of the simulations involved.

A subset of the 16 soil parameters from the calibration of Atchley et al. (2015) are included here and presented in Table 1. The top pressures of the center and trough profiles from the calibration ~~(parameters *toppresctr* and *topprestrg* in)~~ are not included here as these are internally calculated in the surface/subsurface ATS model. The van Genuchten water retention parameters ~~(α_{vgpeat}, α_{vgmin}, m_{vgpeat}, m_{vgmin} in)~~ are not included either as they were found to significantly exceed their physical boundaries during NSMC sampling. This is an indication that these are highly insensitive parameters and do not significantly effect simulated temperatures. This may be explained by the fact that these parameters control the shape of the water retention curve, but that this influences thermal properties of the soil only for a limited time near freeze-up or thaw.

This leaves the 10 soil parameters listed in Table 1. The parameters $\Theta_{r,peat}$ and $\Theta_{r,min}$ are van Genuchten soil moisture characteristic residual saturations (Van Genuchten, 1980). K_{peat} and K_{min} are material thermal conductivities for peat organic matter and mineral grains within the soil layers. ~~These are not bulk thermal conducitivities for the soil layers, but are used in their calculation.~~ Bulk thermal conductivities are a function of material thermal conductivities and are sensitive to ice, liquid and gas saturation, which is calculated within ATS as described in Appendix A of Atchley et al. (2015). $A_{peat,fr}$, $A_{peat,un}$, $A_{peat,fr}$, and $A_{peat,un}$ are ~~emperical~~ empirical exponents describing the dependence of frozen (fr) and unfrozen (un) Kersten numbers (i.e. ratios of partially to fully saturated thermal conductivities) to ice and liquid saturation states, respectively (Painter, 2011). Bulk thermal conductivities for peat and mineral soil layers are calculated within ATS using the Material Component model defined by Atchley et al. (2015) with the parameters listed in Table 1. 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

Table 1. NSMC parameter minimum and maximum bounds, units, and descriptions

Parameter	Min	Max	Units	Description
ϕ_{peat}	0.7	0.93	–	Peat porosity
ϕ_{min}	0.19	0.76	–	Mineral porosity
$\Theta_{r,peat}$	0.04	0.4	m^3m^{-3}	Peat residual liquid saturation
$\Theta_{r,min}$	0.05	0.25	m^3m^{-3}	Mineral residual liquid saturation
K_{peat}	0.05	0.38	$Wm^{-1}K^{-1}$	Peat thermal conductivity
K_{min}	0.2	4.0	$Wm^{-1}K^{-1}$	Mineral thermal conductivity
$A_{peat,fr}$	0.1	3.0	–	Frozen peat thermal conductivity shape parameter
$A_{peat,un}$	0.1	1.5	–	Unfrozen peat thermal conductivity shape parameter
$A_{min,fr}$	0.1	3.0	–	Frozen mineral thermal conductivity shape parameter
$A_{min,un}$	0.1	1.5	–	Unfrozen mineral thermal conductivity shape parameter

through literature review and field observations from the BEO (Imnavait Creek and Kuparuk River, Alaska (Hinzman et al., 1991, 1998); large-scale pan-arctic modeling efforts (Beringer et al., 2001; Lawrence and Slater, 2008); Capricorn Fen, Northern Quebec (Letts et al., 2000); Gailbraith Lake, Northern Alaska (Overduin et al., 2006); Bonanza Creek, Delta Junction, and Washington Creek, Interior Alaska (O'Donnell et al., 2009); Siksik Creek, Northwest Territories (Quinton et al., 2000); Franklin Bluffs, West Dock, Imnavait Creek, Northern Alaska (Nicolsky et al., 2009); Fort Simpson, Scotty Creek, Northwest Territories and Wolf Creek, Yukon Territory (Zhang et al., 2010); Samoytov Island, Lena River delta, Siberia (Chadburn et al., 2015b)).

To a lesser degree, other parameters were also found to exceed their physical boundaries during NSMC sampling. Therefore, we used the intersection of the null space and parameter boundaries as our criterion to accept samples. The generation of 20,000 samples within the null space resulted in 1,153 samples within the parameter boundaries. Samples outside of the parameter boundaries were discarded.

Figure 1 presents histograms while Fig. 2 presents paired plots of the NSMC ensemble soil parameters. In the matrix of plots in Fig. 2, parameter histograms are plotted along the diagonal (also presented in greater detail in Fig. 1), paired scatterplots in the lower triangle, and Pearson correlation coefficients are presented in the upper triangle. In Fig. 1, it is apparent that K_{peat} followed by $A_{peat,un}$ are the most constrained parameter by the NSMC analysis. The rest of the parameters span significant portions of their range. This indicates that ~~their~~ there are many combinations of parameters that result in calibrated temperatures. Many of the histograms are seen to butt up against

their boundaries, indicating that these are parameters where the extent of the null space exceeds their range.

Applying NSMC to multiple calibration locations is often suggested (Tonkin and Doherty, 2009). In the calibration performed by Atchley et al. (2015), multiple local minima were identified. However, based on the broad range of parameter combinations with limited correlations and the fact that most parameters span most of their range, we conclude that the NSMC analysis from this single calibration point sufficiently captures the soil property uncertainty.

The correlations imposed by the NSMC sampling are evident by inspecting the Pearson correlation coefficients and scatterplots in Fig. 2. The strong correlations that are present are a result of a balancing act between parameters to achieve a least squares fit to measured temperatures. For example, the relatively strong negative correlation between K_{peat} and K_{min} (correlation of -0.81) is due to the fact that deeper temperatures in the soil profiles are controlled by the effective thermal conductivity. Therefore, there are numerous (negatively correlated) combinations of K_{peat} and K_{min} that produce similar effective thermal conductivities resulting in good matches to measured temperatures. Many other correlated parameter pairs are also apparent, most with significantly lower correlations. There are also many uncorrelated parameter pairs (e.g. ϕ_{peat} and K_{peat}) indicating a complete lack of interaction between the parameter pairs. The following analysis of permafrost projection uncertainty is conditional on the NSMC correlations presented here, and any conclusions take these correlations into account. References to Fig. 2 are made in the following sections explaining some of the impacts of these correlations.

The range in RMSE values is from around 0.55 to 0.65°C. The accuracy of the temperature probes are ±~~0.02~~0.1°C. Therefore, the percentage of the RMSE that may be attributable to measurement imprecision is around ~~2-3~~15-18%.

Figure 3 presents the evaluation of the calibration against 2014 data and the 95% confidence band of temperatures for the NSMC ensemble. ~~Figure ?? presents the convergence analysis for the NSMC ensemble based on the confidence band inclusion ratio (i.e.the ratio of measured temperatures within the 95th~~The evaluation is presented as time series of temperatures where the fit between 2013 measured and calibrated temperatures can be compared to the fit between 2014 measured and simulated temperatures. By inspection of the plots, it is apparent that the match during the evaluation period is similar to the match during the calibration period (1st, 3rd, and 5th plots for the center, rim, and trough, respectively). This provides an initial indication that the calibration has extracted the hydrothermal relationships from the system and can be applied to years with different climate conditions than the calibration period.

The other plots in Figure 3 contain the corresponding 95% ~~confidence band~~ confidence bands for 2013 temperatures. We performed a convergence analysis of the ensemble ~~simulated temperatures). The relatively stable confidence band inclusion ratio after~~ by calculating the ratio of measurements included in the 95% confidence band as the number of ensemble members increased. We found that

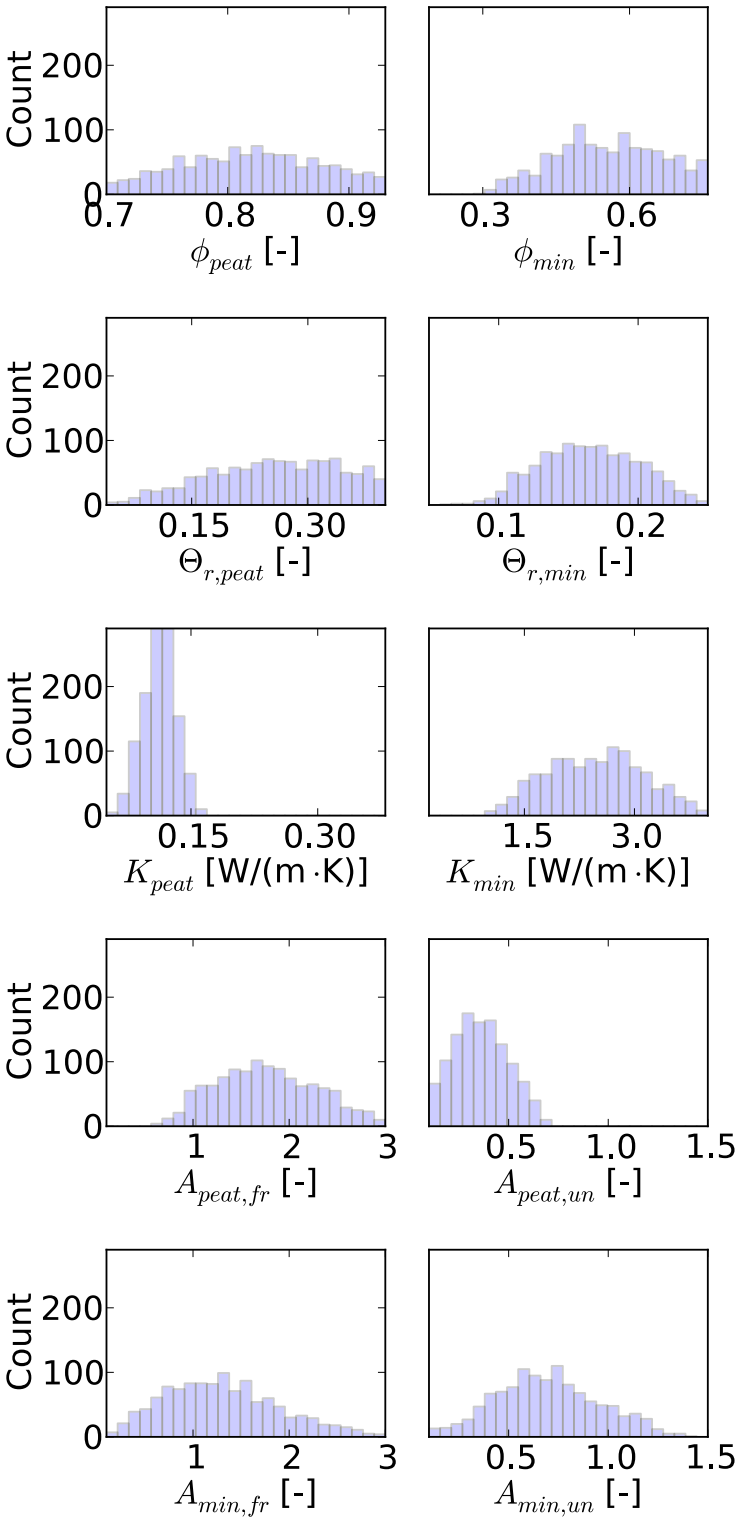


Figure 1. Histograms of calibration-constrained hydrothermal soil parameter combinations obtained by NSMC sampling

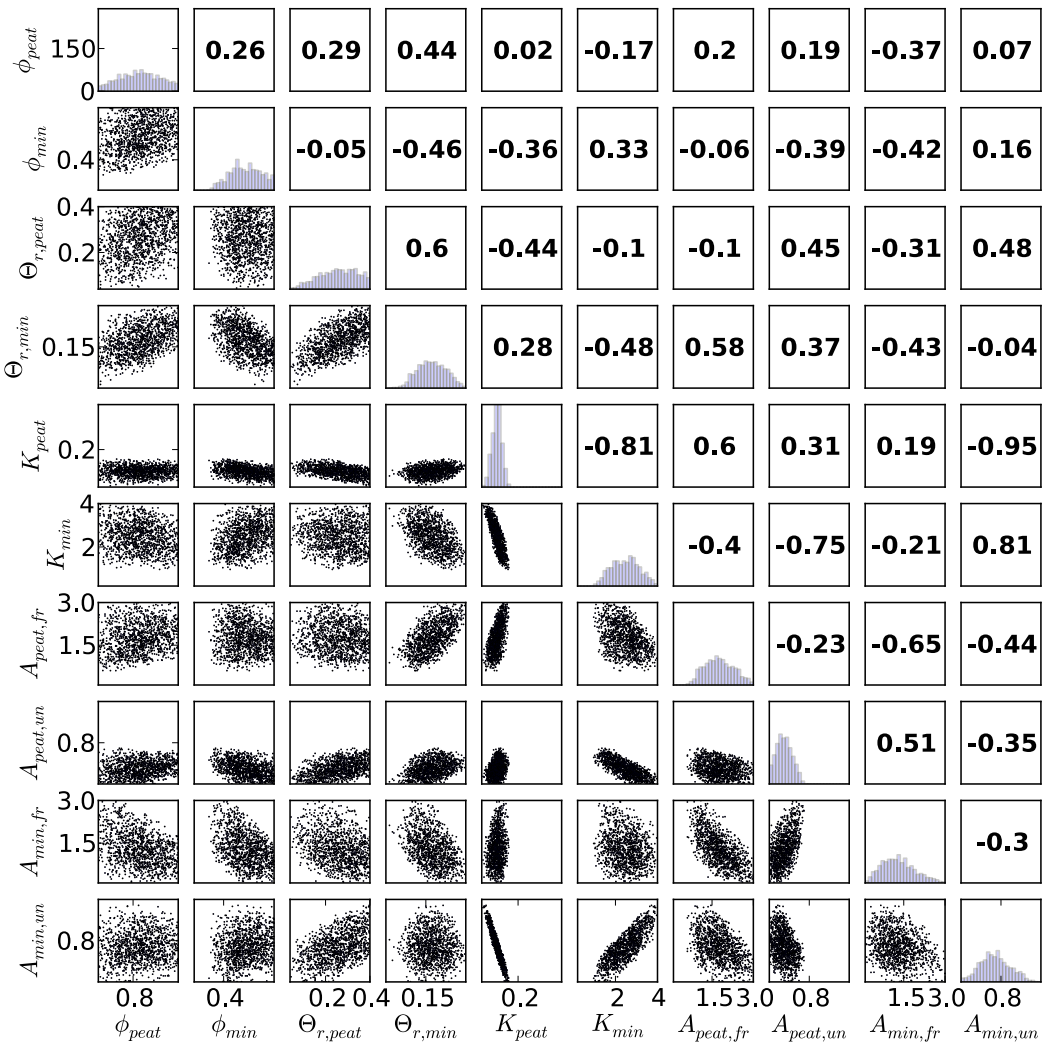


Figure 2. Matrix of paired plots of calibration-constrained hydrothermal soil parameter combinations obtained by NSMC sampling. Parameter histograms are plotted along the diagonal, paired scatterplots in the lower triangle (2D projections of the null space), and Pearson (linear) correlation coefficients in the upper triangle. The histogram counts for all histograms are indicated along the ordinate axis of the upper left plot.

the ratio stabilized after the ensemble reached more than around 800 ~~ensemble members~~members. This indicates that the ensemble has converged and that more samples are not necessary. A plot of the convergence analysis is provided in the Supplement to this article, Fig. S1.

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 ~~discrepencies~~ discrepancies are due to difficulties in capturing trends ~~that are not purely random~~during the freeze-up and thaw of the active layer. The low values are primarily due to the 95% confidence band missing measured values at deep measurements apparent in Figs. ~~??, ??, and ?? in Sect. ??~~. S2, S3, and S4 in the Supplement to this article. A lack of overlap is apparent during thawing (around day of year 150) and freeze-up (around day of year 320), and is particularly evident in the rim profile in Fig. 3. These discrepancies are reduced in the decoupled calibrations (calibrations on individual profiles) (Atchley et al., 2015). We choose to use the coupled calibration parameters in order to obtain soil property values that provide a generalized characterization of the soil properties across the microtopography at the site. The expense of such a generalized characterization is a compromised fit across profiles. The discrepancies are also less pronounced in the center profile, which is the model used for the forward projections. Many physical processes may be leading to this result that become more pronounced in the coupled calibrations as parameter values are given less freedom to mask missing physical processes. For one, the exposed sides of the rim and subsequent lateral heat flow are not explicitly modeled and may at least partially explain the ~~discrepency~~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 (Kane et al., 2001).

~~NSMC conventionally involves a recalibration step, where a few Levenberg-Marquardt iterations are applied to each NSMC sample, often using existing sensitivities from the calibration point. Based on the RMSE values of the ensemble and the percentages of measured temperatures within the 95confidence band, we consider all the unmodified NSMC samples to be calibrated and do not apply this step. These observations also led to the assumption that all NSMC samples are equally consistent with measured temperatures as opposed to using a weighting scheme.~~

An initial ensemble created using Latin Hypercube Sampling with 1,000 samples postprocessed to include parameter combinations with RMSE's below various thresholds indicated that to achieve a convergent ensemble using Latin Hypercube Sampling would be ~~computational~~ computationally prohibitive. An additional NSMC analysis was performed with a more restrictive null space (only 2 eigenvectors out of 10 included in the null space). This ensemble did not require postprocessing based on RMSE, since all the RMSE values were deemed sufficiently small. This analysis resulted in over-correlated parameters. We therefore chose a loosely constrained NSMC (5 out of 10 eigenvectors included in the null-space) excluding samples with RMSE greater than 0.65°C. We considered other RMSE cutoffs, but selected 0.65°C based on achieving a confidence band inclusion ratio and ensuring that simulated temperatures for 2013 were as consistent near the active layer base as possi-

ble across the ensemble. ALT in 2013 was around 40 cm (refer to Figs. ~~??, ??, and ??).~~ S2, S3, and S4 in the Supplement).

NSMC conventionally involves a recalibration step, where a few Levenberg-Marquardt iterations are applied to each NSMC sample, often using existing sensitivities from the calibration point. Re-calibration of the ensemble members was not performed to avoid reducing the simulated temperature uncertainty (lowering the RMSE values) beyond what we deem warranted given the uncertainties involved in measurements and model structure and to avoid the introduction of bias in the ensemble. Based on the RMSE values of the ensemble (< 0.65°C) and the percentages of measured temperatures within the 95% confidence band, we consider all the unmodified NSMC samples to be calibrated and do not apply this step. These observations also led to the assumption that all NSMC samples are equally consistent with measured temperatures as opposed to using a weighting scheme.

~~Calibration-constrained ensemble convergence analysis based on the ratio of measured temperatures within the 95confidence band for ensemble simulated temperatures.~~

~~The projection simulations took on the order of several hours (~2-4 hours)on a Linux cluster with 3.2 GHz processors. We used the Model Analysis ToolKit (MATK) Python module (http://matk.lanl.gov) to facilitate the concurrent execution of the ensemble of ATS models on high performance computing clusters.~~

4 ~~Permafrost metrics~~

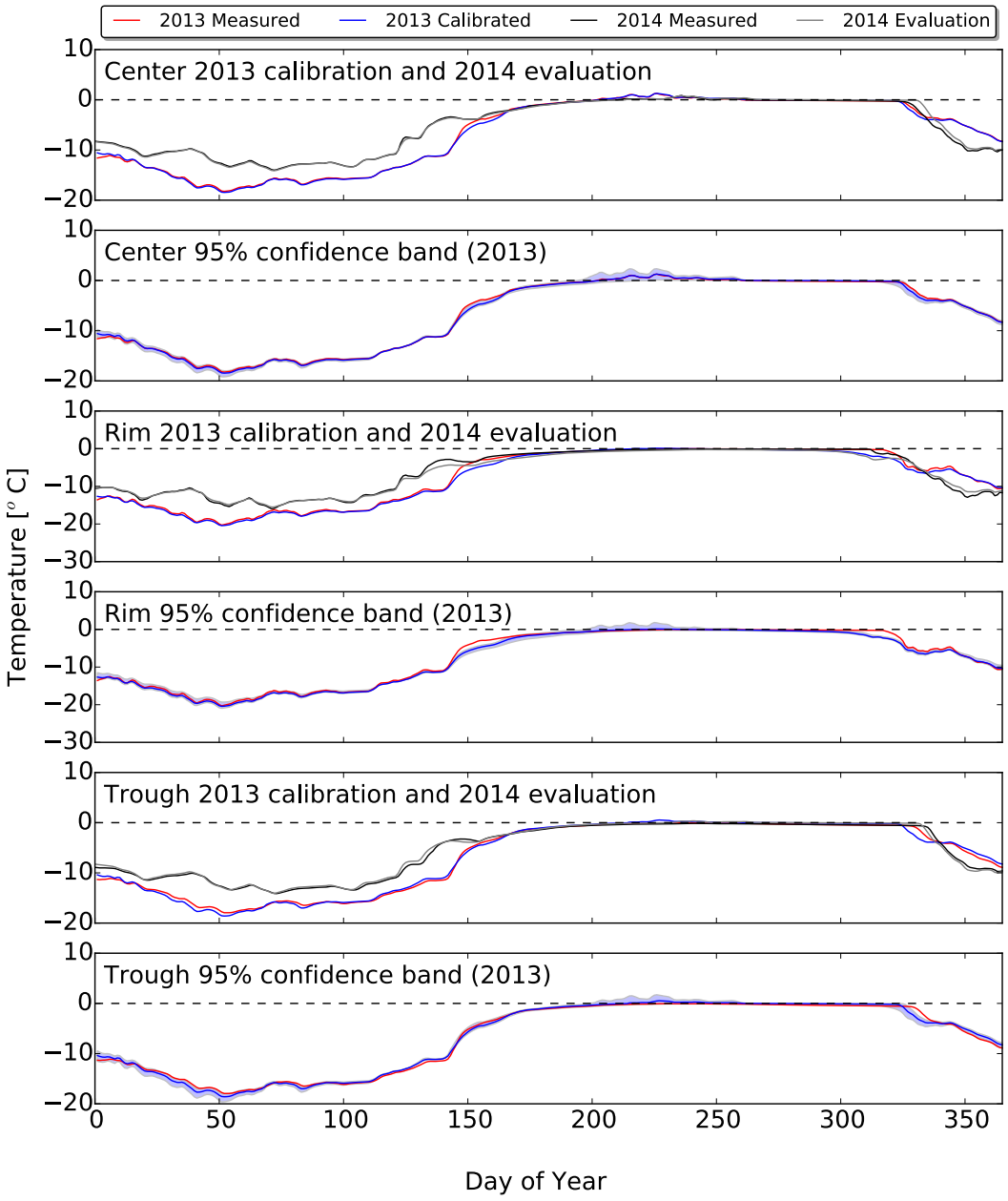
3.1 ~~Active layer thickness (ALT)~~

~~Permafrost is traditionally defined as the region of the subsurface that remains at or below 0°C for two or more years. 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 . Given the discrete nature of our mesh, and the nonlinear nature of vertical soil temperature profiles near 0°C, we determine ALT as the bottom of the deepest thawed mesh cell (temperature above 0°C) for the year.~~

3.1 ~~Annual thaw depth-duration ($\overline{D}$)~~

~~ALT controls the amount of organic carbon experiencing thaw and thus microbially induced decomposition during a year. Because ALT is defined as the maximum thaw depth, it does not include information on duration of thaw. To quantify increasing duration of thaw in future climate as well as increasing depth, a new metric is introduced here: the mean annual thaw depth $\overline{D}$, defined as~~

$$= \frac{1}{365} \int \int H(T(z,t)) dz dt$$



~~Time-series of temperature at 40 cm depths for the polygonal center, rim and trough profiles. Measured values are shown in red, calibrated in blue, and the 95confidence band is the shaded blue region.~~

Figure 3. Time-series of temperature at 40 cm depths plotted as a function of the day of the year for the polygonal center, rim and trough profiles. Alternating plots include measured values from the BEO for 2013 (red line) and 2014 (grey line) and simulated temperatures from the 2013 calibration (blue line) and 2014 evaluation (black line). Every other plot contains the 2013 95% confidence band for the NSMC ensemble as a shaded light blue region.

where H is the heavyside function (1 if $T(z,t)$ is above 0°C, 0 otherwise), z is depth in meters, and t is time in days. The fraction on the right side of Eq. (1) normalizes the metric by the 365 days in a year. We express $\overline{\mathcal{D}}$ with units of m^3m^{-2} to indicate that this metric defines the volume of thawed soil per unit area. Of course, 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. $\overline{\mathcal{D}}$ is a rough proxy for the potential for soil organic matter decomposition. It merges the amount of unfrozen soil and duration that soil is above freezing for a given year. It is noted that, while the annual amount of decomposition is likely correlated with $\overline{\mathcal{D}}$, the two quantities are not directly proportional because soil temperature and moisture will also change and affect the decomposition rates in future climates. In addition, the soil organic matter content in soils generally decreases with depth, which is not accounted for in the $\overline{\mathcal{D}}$ metric. Nevertheless, uncertainty in $\overline{\mathcal{D}}$ is of interest as it is an important control on uncertainty in future decomposition rates.

3.1 Annual mean liquid saturation ($\overline{S_l}$)

The annual mean liquid saturation $\overline{S_l}$ is defined as

$$\iota = \frac{\int\int H(T(z,t))S_l(z,t)dzdt}{\int\int H(T(z,t))dzdt}.$$

$\overline{S_l}$ quantifies the spatially and temporally averaged liquid saturation in the unfrozen soil for a given year. Note that the denominator in Eq. (2) is the annual thaw depth-duration metric $\overline{\mathcal{D}}$ from above, except without dividing by 365. While frozen soil (i.e.soil below 0°C) in our models contain a residual liquid saturation, this is not included in $\overline{S_l}$ (refer to Eq. (2)). Liquid saturation within the active layer is of interest because of its control on decomposition rates. In particular, decomposition may be slower in dry conditions, and oxygen limitations in saturated or nearly saturated conditions may cause methane production to be favored over CO_2 production. Therefore, $\overline{S_l}$ provides an indication of the potential rate of decomposition as well as an indication of the chemical form of the resulting greenhouse gas produced in the active layer.

3.1 Stefan number (S_T)

We propose an extension of the Stefan number from the form in to one that incorporates intra-annual temporal changes and stratified soil properties. The Stefan number is the ratio of subsurface sensible to latent heat. In the current context, this refers to the amount of subsurface heat exchange that results in a change in temperature versus the amount that is consumed in the isothermal conversion of ice to liquid water. In its most basic form, the Stefan number is defined as

$$S_T = \frac{c_b \Delta T}{L_f}.$$

~~where c_b is the bulk specific heat of the material and L_f is the latent heat of fusion of water (334,000 J kg^{-1}). define the Stefan number for the permafrost problem as~~

$$S_T = \frac{c_b \rho_b (T_s - T_f)}{S_{wf} \rho_w \phi L_f}$$

~~where ρ_b is the density of the thawed zone, T_s is the surface temperature, T_f is the temperature of freezing or thawing soil (taken as 0°C), S_{wf} is the liquid saturation in the thawed zone that was frozen, and ρ_w is the density of liquid water. use this definition to evaluate the thermal regime of analytical solutions of soil thaw. We expand this definition here to include the increased detail available in our numerical simulations as~~

$$S_T = \frac{\int\int c_b(z)\rho_b(z)\, H\left(\frac{dT}{dt}\right)\frac{dT}{dt}dzdt}{\rho_{ice} L_f \int\int H\left(-\frac{dS_{ice}}{dt}\right)\left(-\frac{dS_{ice}}{dt}\right)\phi(z)dzdt}$$

~~where S_{ice} is ice saturation. The integrations are performed over the entire year (i.e. from Jan. 1 through Dec. 31). Equation 5 expands on Eq. (4) to allow the consideration of details of transient heating and cooling throughout the year and stratified hydrothermal soil properties within the soil profile.~~

4 ~~Permafrost thaw projection uncertainty~~

3.1 Permafrost thaw projection uncertainty

Figure 4 present boxplots of permafrost metrics for the first year (2006) and the last decade (2091-2100) of the projections. Individual boxplots for each year present the ~~intra-annual predictive uncertainty~~ predictive uncertainty (due to parametric soil property uncertainty), while comparisons between boxplots for each metric indicate the inter-annual climate variability of the projections for the specified climate ~~scenario~~model. We present the ~~first year~~ year 2006 as an indication of the ~~intra-annual uncertaintyat the beginning of the projections~~initial parametric uncertainty.

Boxplots of ALT are shown in Fig. 4a. The median ALT increased from approximately 30 cm in 2006 to nearly 0.9 m by the end of the century. The ~~intra-annual~~ predictive uncertainty in ALT also increases significantly from the beginning to later years of the projections. The ~~intra-annual~~ inter-annual variability of ALT projections is dependent on climate, as warmer years (e.g. 2094) have greater ALT and larger uncertainty than cooler years. This is apparent in Fig. 5 where the ensemble thaw depth statistics (median and 95% confidence band)~~and~~ , CESM8.5 air temperature, and ensemble snow depth statistics (95% confidence band) times series are plotted together for comparison.

Boxplots of annual thaw depth-duration ($\overline{\mathcal{D}}$) are presented in Fig. 4b. The ~~intra-annual~~ predictive uncertainty in $\overline{\mathcal{D}}$ during the last decade of the projections is significantly greater than for the first year (2006). As expected, the inter-annual trends in $\overline{\mathcal{D}}$ and ALT are similar. Also, the uncertainty of $\overline{\mathcal{D}}$ is relatively larger during warmer years than cooler years, similar to ALT.

Boxplots of the annual mean liquid saturation ($\overline{S}_l$) are presented in Fig. 4c. The ~~intra-annual~~ predictive uncertainty in $\overline{S}_l$ actually decreases slightly from the first year to the last decade. Also, in general, the last decade is slightly wetter than 2006, but only marginally so. Therefore, this hydrothermal analysis does not indicate that the partitioning of carbon decomposition between CO_2 and CH_4 will change significantly as permafrost thaws. However, other factors affecting carbon decomposition not considered here could affect the partitioning of carbon decomposition end products.

Boxplots of the Stefan number (S_T) are presented in Fig. 4d. In 2006 the soil profiles for the majority of the ensemble are latent heat dominated. However, some Stefan numbers are greater than 1, with values ranging from around 0.3 to 1.4 (from around 3 times the latent heat as sensible heat to 1.4 times the sensible heat as latent heat). However, by the last decade, nearly all Stefan numbers are 0.2 or less (at least 5 times as much, and up to 20 times as much latent heat as sensible heat). This indicates a fundamental change in the way that the active layer processes energy between the beginning and later years of the projections. The thermal regime of the active layer becomes significantly more dominated by latent heat during the projections. The amount of energy that is utilized in creating a temperature gradient in the soil profile becomes proportionately smaller compared to the amount of energy consumed in the isothermal melting of ice. This is at least partially due to the approximately 3 times increase in the quantity of ice that is melted during later years of the projections. Perhaps the most significant result of this change is the temperature regime of the underlying permafrost in decreased seasonal temperature variations and their depth of penetration. ~~Intra-annual~~ Predictive uncertainty appears to decrease from 2006 compared to the last decade, but this is likely due to the Stefan number approaching its lower limit.

To further illustrate ~~intra-annual~~ predictive uncertainty of the ALT projections, temperature profiles at the time of ALT for year 2100 are presented in Fig. 6. Summary statistics (median and 5th and 95th percentiles) for 2006 are presented for reference. The discrete surface temperatures categorized by day of year (colors) reflect the fact that the surface temperature is highly dependent on the climate/air temperature for a given year, which is the same for all projections. The increase in median ALT from around 30 cm to around 0.9 m from 2006 to 2100 is also apparent in this figure. The difference in the temperature regime within the profile is apparent in these figures as well by the ~~curve~~ curvature near the surface in most of the profiles in 2100 compared to 2006. This indicates that as the climate warms and the day of year when ALT occurs becomes later in the year (day of year ALT occurs in 2006 projections is from 246 to 260), the surface temperature at that time will be cooler. This increase in lag time from the surface temperature to the active layer base is a result of the thermal wave traveling a greater distance to reach the permafrost. This may also be due to relative changes in the temperature gradient within the active layer and the permafrost as the ALT increases leading to delayed freeze from below.

Figure 7 shows similar plots to Fig. 6, but in this case, statistical measures of the ensemble are plotted. Statistical representation of the temperature profiles in Fig. 6 are plotted in Fig. 7a, along with

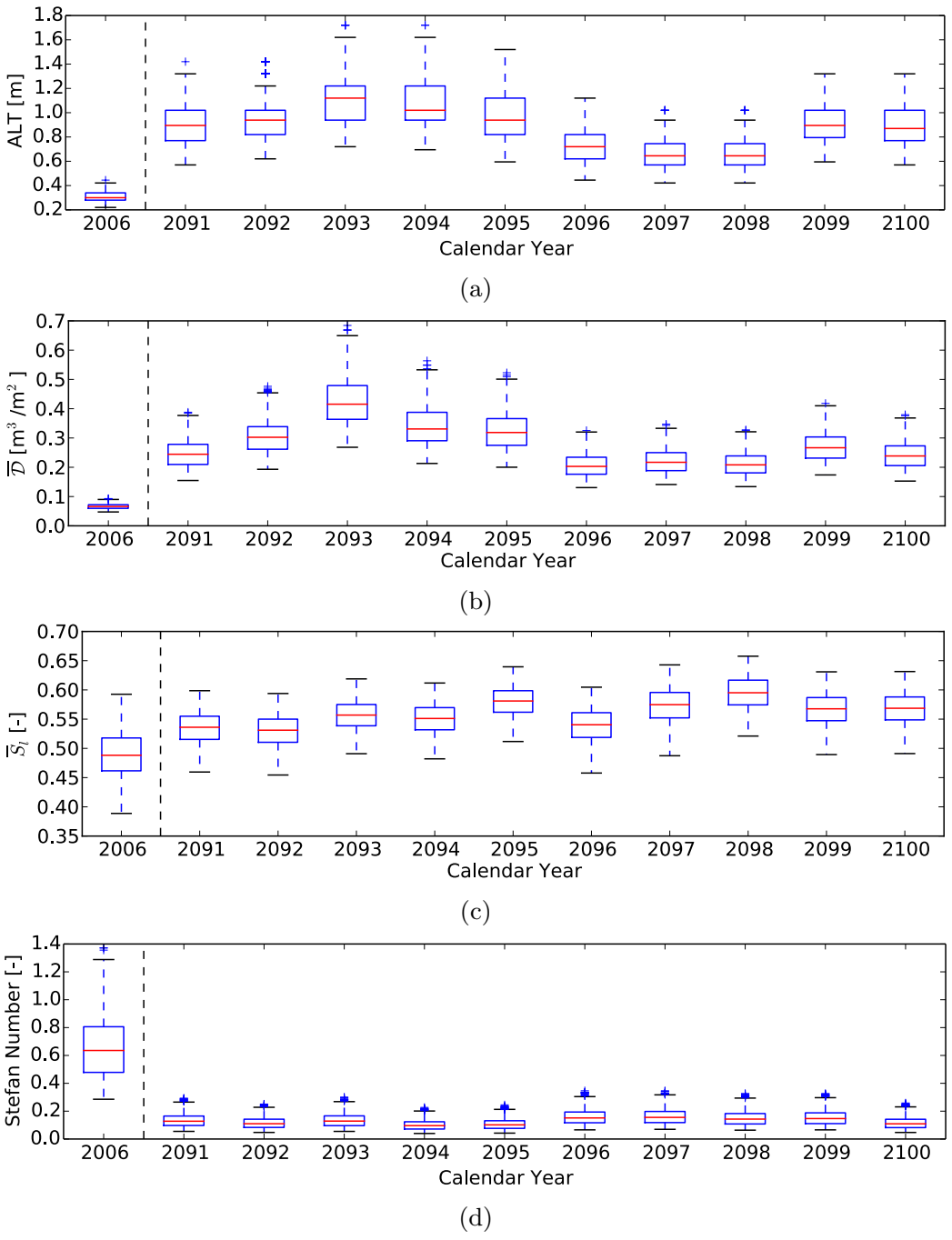


Figure 4. Boxplots of projected metrics including (a) ALT, (b) annual thaw depth-duration, (c) annual mean liquid saturation, and (d) Stefan number for year 2006 and from 2091 to 2100. The bottom and top of the boxes are the first and third quartiles, the red lines are medians, the whisker lengths are 1.5 times the interquartile range (50%), and the plus symbols are outliers.

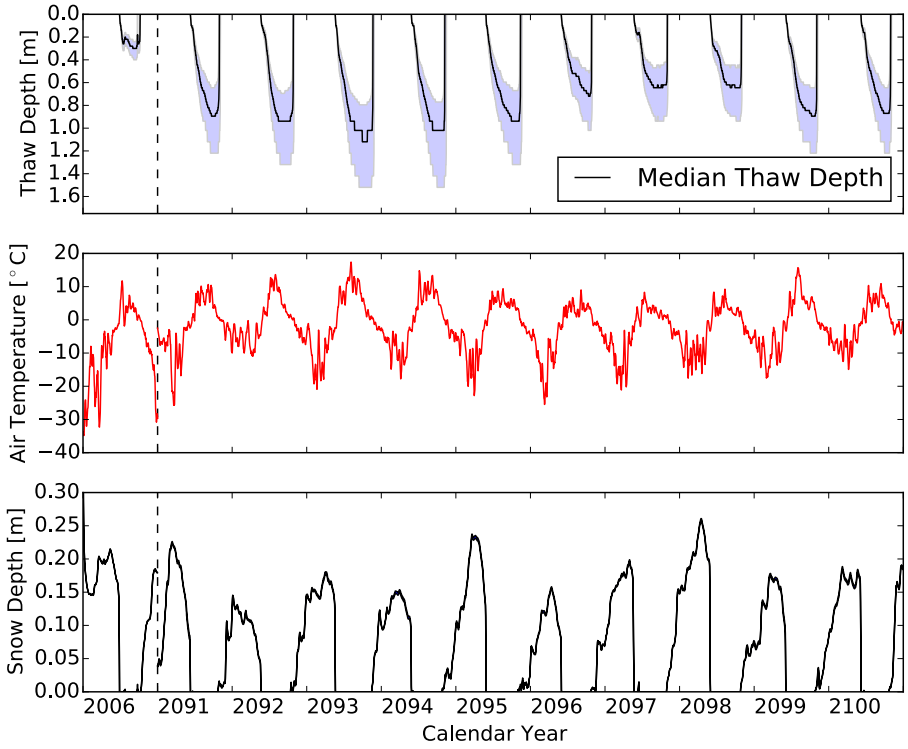


Figure 5. Thaw depth~~and~~ , air temperature, and snow depth time series for years 2006 and 2091 through 2100. The black line in the top plot is the median thaw depth of the ensemble and the blue shaded region is the 95% thaw depth confidence band for the ensemble. The black region in the bottom plot is the 95% snow depth confidence band for the ensemble.

bulk thermal conductivity (Fig. 7b) and ice (Fig. 7c), liquid (Fig. 7d), and gas (Fig. 7e) saturation profiles when ALT occurs in 2006 and 2100. The variation in thermal conductivity and saturation states further illustrates the ~~intra-annual projection~~ predictive uncertainty due solely to soil properties. Substantial shifts in ~~intra-annual~~ predictive uncertainty are also apparent from 2006 to 2100. In Fig. 7a, it is apparent that the thermal conductivity in the soil profile decreases from 2006 to 2100 due to the loss of the more thermally conductive ice from the profile, thereby inhibiting the propagation of the thermal wave. The deepening of the permafrost table is apparent in Fig. 7c as a deepening of the ice saturated region. Note that liquid saturations for mineral soil remain at its residual values below 0°C and that residual liquid saturations ($\Theta_{r,peat}$ and $\Theta_{r,min}$) are variable parameters within the uncertainty quantification (refer to Table 1). As a result, the ice saturation within the permafrost region is variable within the ensemble. In Figs. 7d and 7e, it is apparent that the liquid and gas satura-

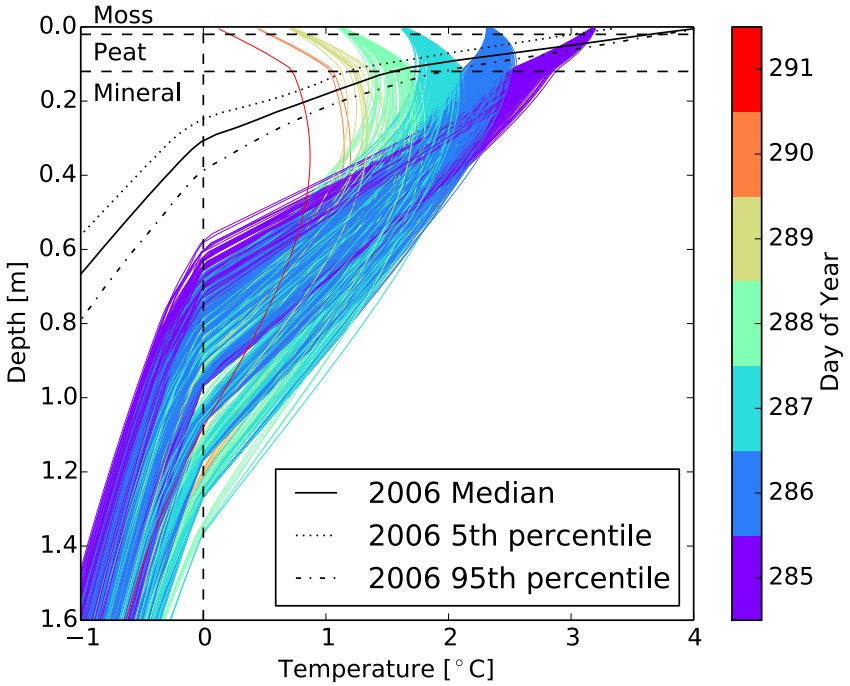


Figure 6. ~~Intra-annual~~ Predictive uncertainty due to soil properties for depth profiles of temperature for the ensemble when ALT occurs for calendar year 2100. The 2006 median and 5th and 95th percentiles are ~~presented in subplot~~ plotted for reference. Day of year when ALT occurs for 2006 is from 246 to ~~260.~~260 (not indicated in the plot).

tions both increase as ice is converted to liquid and void space becomes available with the deepening of the permafrost table.

4 ~~Comparison to climate model structural uncertainty~~

3.1 Comparison to climate model structural uncertainty

In this section, we provide a frame of reference to the effect of soil property uncertainty on permafrost thaw projections by comparison to the uncertainty currently present in climate models. Without such a comparison, the relative contribution of soil property uncertainty would be difficult to gauge. Figure 8 presents histograms of projection metrics collected from each ensemble sample for years 2091 through 2100 (a total of 11,530 values, i.e. 1,153 samples $\times$ 10 years). This combines the ~~intra-annual~~ predictive uncertainty for the last decade of the projections. The 95% confidence band of the calibration-constrained ensemble for each metric is indicated by dashed vertical lines in each plot. Below the histograms are the values obtained using atmospheric forcing data from CESM, INM, BCC, MIROC, CAN, and HAD climate models to drive the ATS models with the calibrated

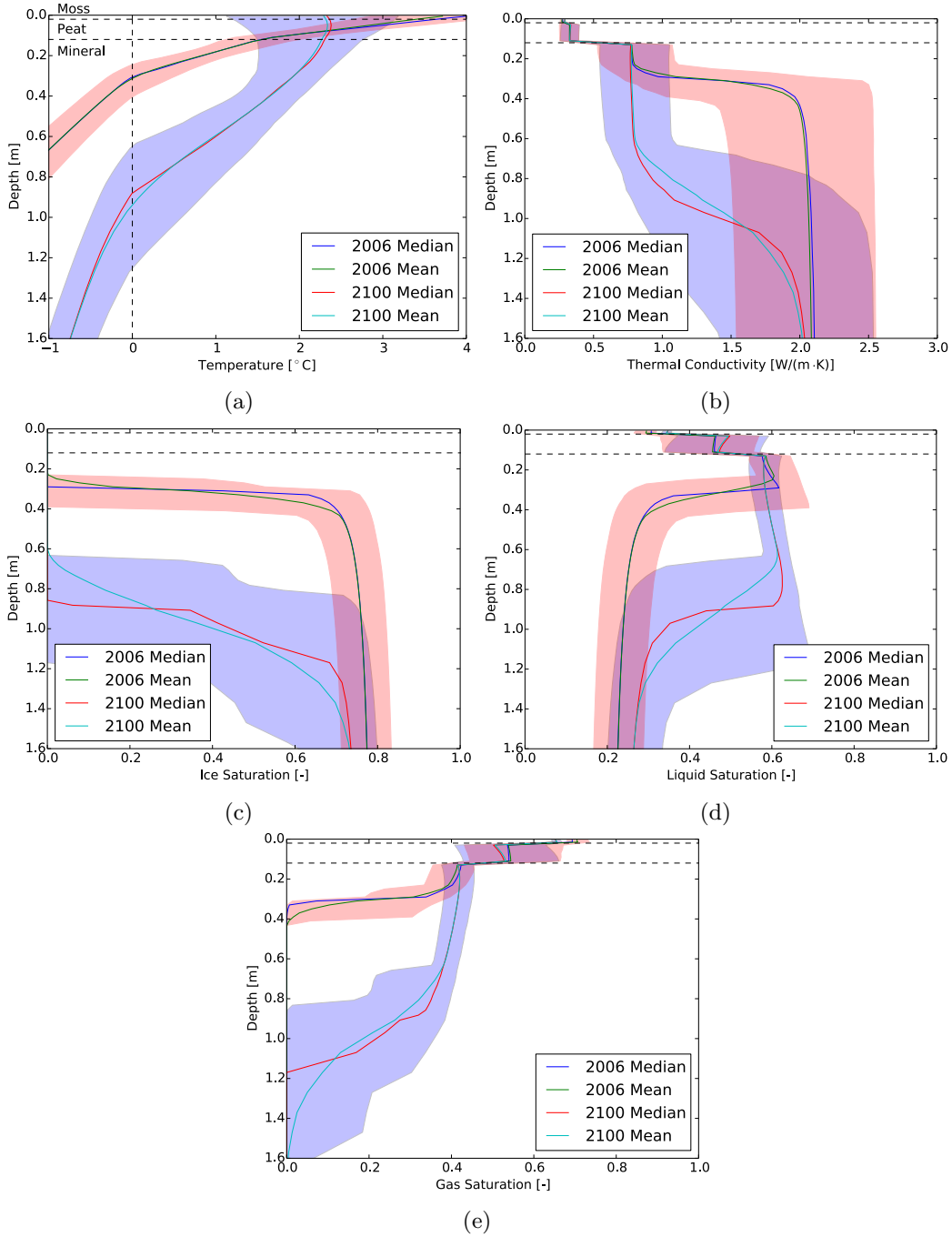


Figure 7. ~~Intra-annual predictive~~ Predictive uncertainty due to soil property uncertainty for depth profiles of ensemble statistical quantities when ALT occurs for calendar years 2006 and 2100. The shaded regions are the 95% confidence intervals for 2006 (red) and 2100 (blue).

(fixed) soil parameters for the same years, 10 values each. BCC has only 9 values as we could only obtain its ~~data~~ output through year 2099. These values provide a sampling of current climate model structural uncertainty due to varying assumptions and numerical representations of atmospheric phenomena.

Note that the CESM values lie within the support of the calibration-constrained ensemble histograms in all cases. This is expected since the calibration-constrained ensemble is forced using the CESM ~~scenario~~model. Similarly, the supports of calibration-constrained ensemble histograms for other climate ~~scenarios~~ models would be expected to encompass the calibrated soil parameter values (circles in Fig. 8) as well. This indicates that different climate ~~scenarios~~ models will result in different magnitudes of projection uncertainty due to soil property uncertainty. For example, if the calibration-constrained ensemble was simulated using MIROC, the magnitude of the projection uncertainty of $\overline{\mathcal{D}}$ (Fig. 8b) could be as much as 4-5 times larger than for CESM. This indicates the interactive effect that soil property and structural climate model uncertainties have on projection uncertainty and that these forms of uncertainty are not easily decoupled.

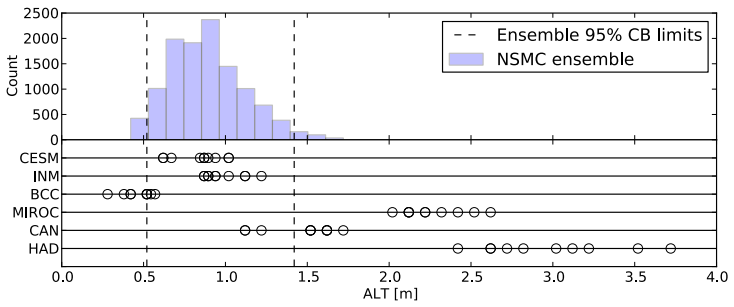
These plots present both the magnitude of projection uncertainty due to ~~only~~ soil property uncertainty based on CESM atmospheric projections (histograms) and to ~~only~~ structural climate model uncertainty (circles). By comparing the ensemble 95% confidence bands for the metrics to the range of values across the climate models, it is apparent that structural climate model uncertainty has a greater impact on projection uncertainty than soil property ~~uncertianty~~uncertainty. The ratios of the ensemble 95% confidence band width and the range between the minimum and maximum values for climate models are 26% for ALT, 9% for $\overline{\mathcal{D}}$, 45% for $\overline{S}_l$, and 80% for S_T. As explained above, if a different climate model had been used for the ensemble calculations, these percentages would be different.

~~4 Dependence of permafrost projections on soil parameters~~

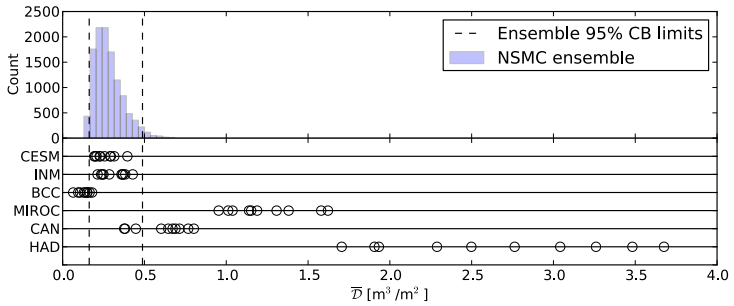
3.1 Dependence of permafrost projections on soil parameters

~~Figure ?? presents paired plots of calibration-constrained projections for year 2100. The diagonals are projection histograms, the lower triangle contains paired scatterplots, and the upper triangle contains the Pearson correlation coefficients between matrix pairs. The samples are discrete in ALT due to the mesh discretization. The mesh cell thickness increases with depth, and the active layer is determined as the depth to the bottom of the deepest unfrozen cell (i.e.with a temperature above 0°C).~~

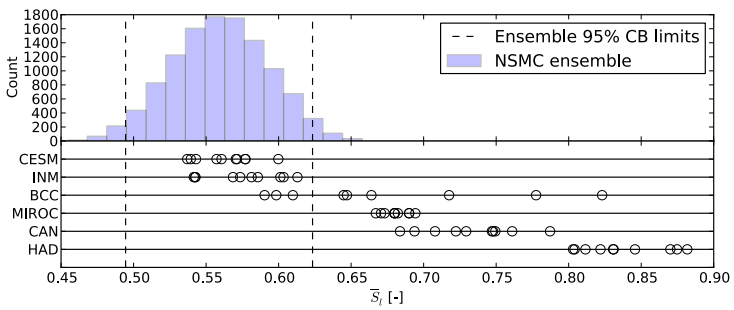
~~From this figure, it is apparent that all the metrics are positively correlated.~~ Based on a correlation analysis, all the permafrost metrics are positively correlated, with lower correlations between annual mean liquid saturation and the other metrics. A paired plot of the permafrost metrics is provided in the Supplement to this article for additional detail (Fig. S5). The correlation between ALT and $\overline{\mathcal{D}}$ is



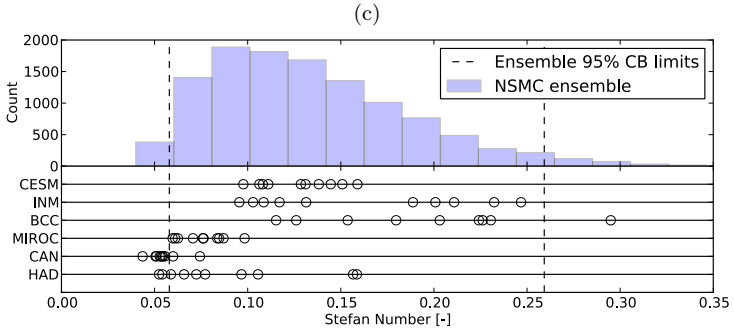
(a)



(b)



(c)



(d)

Figure 8. Comparison of (a) ALT, (b) annual thaw depth-duration, (c) annual mean liquid saturation, and (d) Stefan number projection uncertainty due to soil property uncertainty (histograms) and structural climate model uncertainty (circles). Histograms include calibration-constrained ensemble values for years 2091 to 2100 (11,530 values) based on the CESM8.5 climate ~~scenario~~model. Open circles below the histograms are values for the various climate ~~scenarios~~ models for the same years using the calibrated soil parameters (10 values each, except for BCC which has 9). ~~Ensemble~~ NSMC ensemble 95% confidence band (CB) limits are indicated as vertical dashed lines.

expected given the definition of $\overline{\mathcal{D}}$ as a metric defining the quantity and duration of unfrozen soil. The correlation of $\overline{S}_l$ to ALT is a result of the deeper portions of the thicker ALT scenarios having slightly increased levels of saturation, which is apparent in the liquid saturation statistical profiles in Fig. 7d for year 2100. The correlation between $\overline{\mathcal{D}}$ and $\overline{S}_l$ can be explained by a similar argument. Increased levels of saturation lead to higher bulk thermal ~~conductivy~~ conductivity of the mineral soil layer, resulting in thicker ALT and larger $\overline{\mathcal{D}}$ due to increased energy flux. Correlations between S_T and the other projection metrics indicate that as ALT increases, resulting in increased annual thaw depth-duration $\overline{\mathcal{D}}$ and annual mean liquid saturation $\overline{S}_l$, the system becomes increasingly latent heat dominated. This is due to the fact that more energy is required to thaw greater depths of frozen soil ~~each year~~in later years.

~~Matrix of paired plots of calibration-constrained ensemble projections for year 2100. Parameter histograms are plotted along the diagonal, paired scatterplots in the lower triangle, and Pearson correlation coefficients in the upper triangle. The range of counts for all histograms are as indicated along the ordinate axis of the upper left plot.~~

Figures 9, 10, 11, and 12 explore correlations between the calibration-constrained parameters and projected metrics. These figures ~~plot~~ contain scatterplots between hydro-thermal soil parameters and projection metrics for year 2100. The discrete nature of the samples with respect to ALT mentioned above due to the mesh discretization is also apparent in Fig. 9. Pearson correlation coefficients for each soil parameter/projection metric pair are presented on each scatterplot. The points are colored by $\overline{\mathcal{D}}$ in Fig. 9 and by ALT in Figs. 10, 11, and 12 to ~~further~~ illustrate the correlations between metrics ~~already presented in~~ (see also Fig. ~~??~~S5 in the Supplement). Peat parameters are presented along the left column and mineral soil parameters along the right column of each figure.

Some strong correlations are apparent in Figs. 9, 10, 11, and 12 with coefficients greater ~~that~~ than 0.9. Many of these correlations confirm our qualitative understanding of the model. It is apparent that in many cases projection metrics have stronger dependencies on the mineral soil porosity (ϕ_{min}) and residual saturation ($\Theta_{r,min}$) parameters compared to the corresponding peat parameters (ϕ_{peat} and $\Theta_{r,peat}$). Dependence on the other parameters is less predictable. For example, decreasing mineral soil porosity (ϕ_{min}) increases the bulk thermal conductivity of the mineral soil due to the relatively large thermal conductivity of the mineral soil grains, leading to larger ALT (top right plot in Fig. 9).

We determine linear dependency coefficients of projection metrics to calibration-constrained parameters using ordinary least squares. We limit the analysis to soil parameter/projection metrics ~~exibiting~~ exhibiting moderate to strong correlation ($|\rho| > 0.7$). Table 2 presents the intercept and slope coefficients from the analysis, along with their 95% confidence intervals. All coefficients in Table 2 are significant at the 1% level. The coefficient of determination (R^2) is presented indicating the portion of the variance explained by the regression for each case. Note that since we use ordinary least squares including an intercept, the R^2 is simply the square of the correlation coefficients

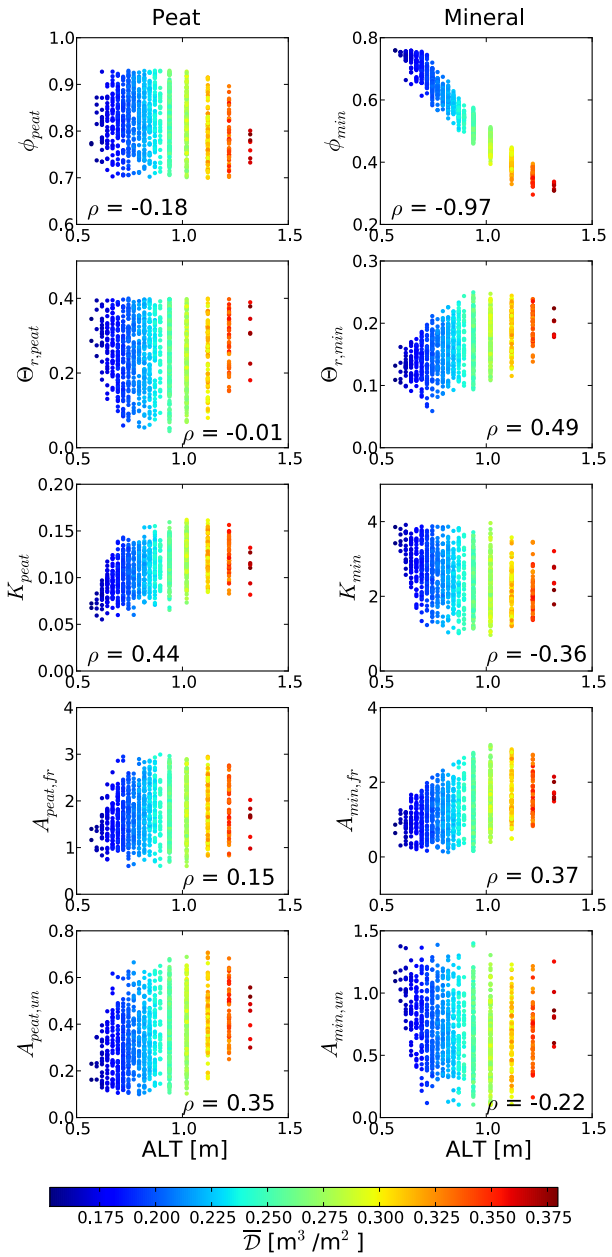


Figure 9. Scatterplots between calibration-constrained parameters and projected ALT for year 2100. Soil parameters associated with peat are on the left and with mineral soil on the right (refer to column headings). Colors represent annual thaw depth-duration. The associated Pearson correlation coefficient ρ is indicated in each plot. The discrete nature of the ALT is due to the computational mesh discretization.

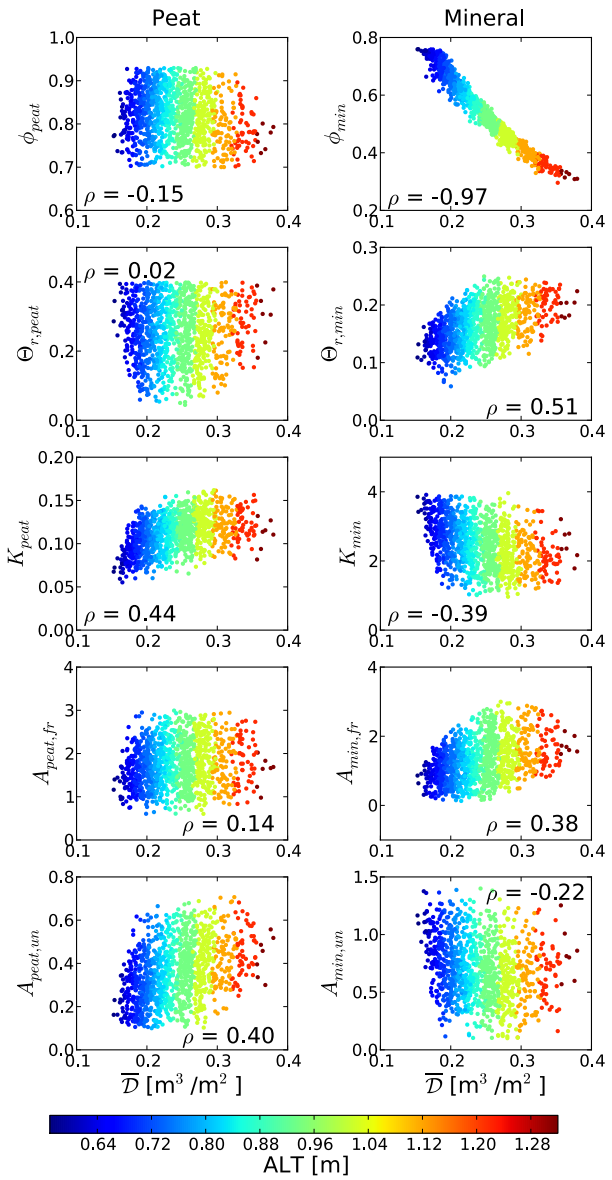


Figure 10. Scatterplots between calibration-constrained parameters and projected annual thaw depth-duration. Soil parameters associated with peat are on the left and with mineral soil on the right (refer to column headings). Colors represent ALT. The associated Pearson correlation coefficient ρ is indicated in each plot.

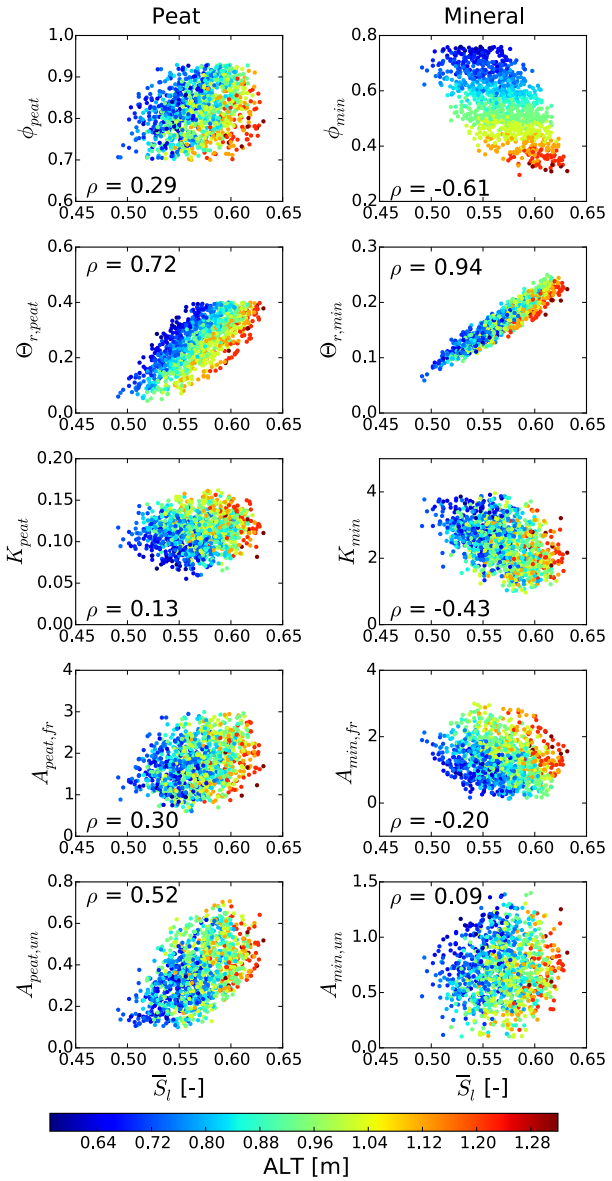


Figure 11. Scatterplots between calibration-constrained parameters and projected annual mean saturation. Soil parameters associated with peat are on the left and with mineral soil on the right (refer to column headings). Colors represent ALT. The associated Pearson correlation coefficient ρ is indicated in each plot.

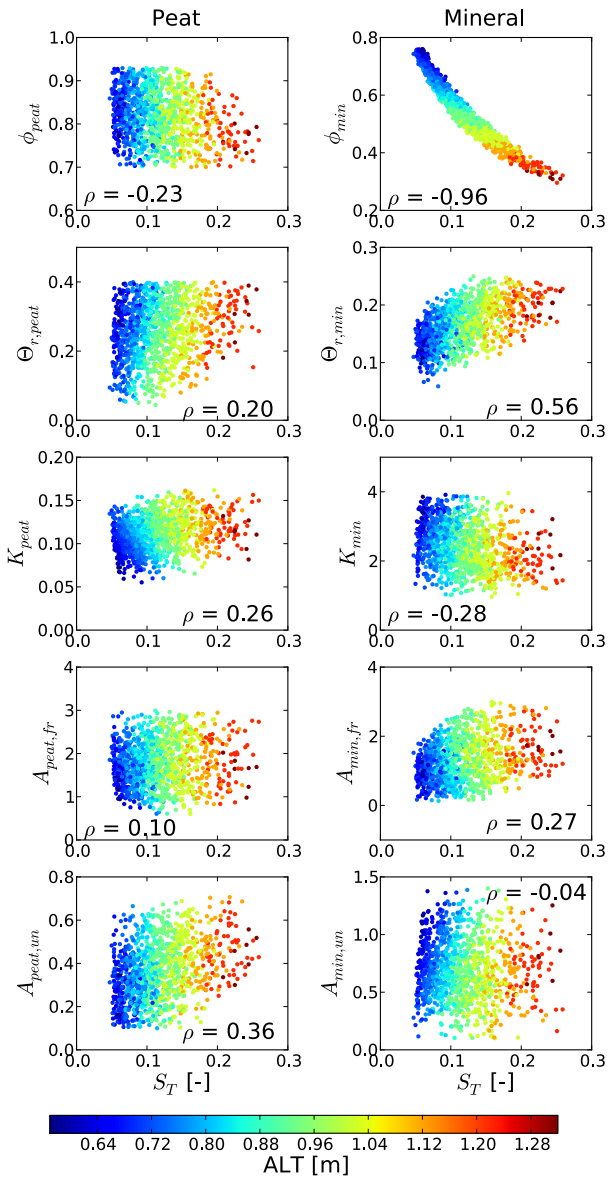


Figure 12. Scatterplots between calibration-constrained parameters and projected Stefan number. Soil parameters associated with peat are on the left and with mineral soil on the right (refer to column headings). Colors represent ALT. The associated Pearson correlation coefficient ρ is indicated in each plot.

Table 2. Linear regression intercept and slope coefficients for permafrost metrics as a function of calibration-constrained parameters

Metric	Parameter	Intercept	95% Conf. Int.	**Slope**	95% Conf. Int.	R^2
ALT	ϕ_{min}	1.66	1.65 – 1.67	**-1.39**	-1.41 – -1.37	0.95
$\overline{\mathcal{D}}$	ϕ_{min}	0.465	0.462 – 0.468	**-0.402**	-0.408 – -0.397	0.95
$\overline{S}_l$	$\Theta_{r,peat}$	0.510	0.506 – 0.513	**0.227**	0.215 – 0.240	0.52
	$\Theta_{r,min}$	0.452	0.450 – 0.455	**0.702**	0.687 – 0.717	0.87
S_T	ϕ_{min}	0.327	0.323 – 0.331	**-0.381**	-0.387 – -0.374	0.92

(ρ) presented in Figs. 9, 10, 11, and 12. Calibration-constrained parameters not included in Table 2 resulted in regressions with R^2 less than 0.5.

The slope coefficients are emphasized in bold in the table since these describe the first-order dependence of projection metrics on the calibration-constrained parameters. The slope coefficients describe the change in ALT given a unit change in the calibration-constrained parameter. For example, if ϕ_{min} increases by 0.1, we would estimate that ALT will decrease by around 0.14 m. These coefficients can be useful in gaging the impact of soil parameter changes on projection metrics.

4 Discussion and Conclusions

In summary, we extended previous calibration and model refinement work (Atchley et al., 2015) to quantify post-calibration uncertainty in soil properties and the impact of ~~that~~ uncertainty on projections of permafrost thaw. Using a model with parameters calibrated against data from the BEO, driving the NSMC ensemble of models using the CESM climate model in the RCP8.5 scenario, and comparing against a set of other climate models in the RCP8.5 ~~scenarion~~scenario, the following conclusions can be made:

- The median ALT and annual thaw depth-duration ($\overline{\mathcal{D}}$) of the calibration-constrained ensemble increase by around a factor of 3 by the end of the century.
- The effect of soil property uncertainty based on CESM atmospheric forcings is approximately 26% of the uncertainty caused by climate model structural uncertainty for ALT, 9% for $\overline{\mathcal{D}}$, 45% for $\overline{S}_l$, and 80% for Stefan number.
- ~~Intra-annual~~ Predictive uncertainty of ALT and $\overline{\mathcal{D}}$ due to soil property uncertainty increase significantly from the first year to the last decade of the projections
- ~~Intra-annual~~ Predictive uncertainty of soil moisture content due to soil property uncertainty is not significantly changed by the end of the century.

- ~~Intra-annual~~ Predictive uncertainty of the Stefan number due to soil property uncertainty decreases, but this is at least partially due to this metric approaching its lower boundary in the last decade.

- The ~~active layer~~ manner in which the active layer processes incoming energy changes significantly. The active layer moves to an increasingly latent heat dominated system due to larger quantities of frozen ground thawed each year.

- ALT, $\overline{\mathcal{D}}$, and S_T are highly dependent on ϕ_{min}, while $\overline{S}_l$ is highly dependent on $\Theta_{r,min}$ and moderately dependent on $\Theta_{r,peat}$.

Efforts to quantify the relative roles of ~~subsurface versus climate and scenario~~ soil property versus climate model uncertainty have only recently begun. We found that the effect of soil property uncertainties can be reduced to levels lower than the uncertainty generated by uncertainties in climate model structure through a process of calibration to field observations, model structural refinement (Atchley et al., 2015), and calibration-constrained uncertainty analysis. However, we had the advantage of high-resolution data from an unusually well-characterized site, which suggests that the residual uncertainty identified here using temperature data only is close to a practical limit.

The quantitative results shown here are specific to the site, available data, RCP trajectory assumption, and climate model. Nevertheless, the approach presented here is anticipated to be useful for understanding the impact that additional data collection might have on reducing uncertainty associated with other high-latitude permafrost sites. Potential directions for future work include the investigation on the impact that longer data streams and other types of observation might have on reducing uncertainties. In particular, the calibration against borehole temperature data was uninformative of certain water retention properties of the soils (van Genuchten α and m parameters). Therefore, co-located measurements of soil moisture would be useful to help constrain those parameters, and may reduce the uncertainty associated with the other soil properties as well. Moreover, given the known spatial variability in soil properties across the pan-Arctic (Hinzman et al., 1998; Rawlins et al., 2013), calibration-constrained soil property uncertainty across larger spatial scales warrants further investigations.

5 ~~Supplemental information~~

~~Figures ??, ??, and ?? present the 95th confidence band for NSMC ensemble temperatures during the calibration year for all depths. These figures present the complete data set from which Figure 3 was drawn, which presents the 40 cm depth values only (near the ALT in 2013) .~~

~~Time-series of temperature at specific depths for the polygonal center. Measured values from the field are shown as a red line, the mean of the NSMC sample as a blue line, and the 95confidence band is the shaded light blue region.~~

~~Time-series of temperature at specific depths for the polygonal rim. Measured values from the field are shown as a red line, the mean of the NSMC sample as a blue line, and the 95confidence band is the shaded light blue region.~~

~~Time-series of temperature at specific depths for the polygonal trough. Measured values from the field are shown as a red line, the mean of the NSMC sample as a blue line, and the 95confidence band is the shaded light blue region.~~

Acknowledgements. This research was supported by the Next-Generation Ecosystem Experiments Arctic (NGEE-Arctic) project (DOE ERKP757) funded by the Office of Biological and Environmental Research in the US Department of Energy Office of Science and Los Alamos National Laboratory's Laboratory Directed Research and Development (LDRD) Arctic project (LDRD201200068DR).

References

Atchley, A. L., Painter, S. L., Harp, D. R., Coon, E. T., Wilson, C. J., Liljedahl, A. K., and Romanovsky, V. E.: Using Field Observations to Inform Thermal Hydrology Models of Permafrost Dynamics with ATS (v0.83), Geoscientific Model Development Discussions, 8, 3235–3292, doi:10.5194/gmdd-8-3235-2015, 2015.

Bellouin, N., Collins, W., Culverwell, I., Halloran, P., Hardiman, S., Hinton, T., Jones, C., McDonald, R., McLaren, A., O'Connor, F., et al.: The HadGEM2 family of met office unified model climate configurations, Geoscientific Model Development Discussions, 4, 765–841, 2011.

Beringer, J., Lynch, A. H., Chapin III, F. S., Mack, M., and Bonan, G. B.: The representation of arctic soils in the land surface model: the importance of mosses, Journal of Climate, 14, 3324–3335, 2001.

Chadburn, S., Burke, E., Essery, R., Boike, J., Langer, M., Heikenfeld, M., Cox, P., and Friedlingstein, P.: Impact of model developments on present and future simulations of permafrost in a global land-surface model, The Cryosphere Discussions, 9, 1965–2012, 2015a.

Chadburn, S., Burke, E., Essery, R., Boike, J., Langer, M., Heikenfeld, M., Cox, P., and Friedlingstein, P.: An improved representation of physical permafrost dynamics in the JULES land surface model, Geoscientific Model Development Discussions, 8, 715–759, 2015b.

Clapp, R. B. and Hornberger, G. M.: Empirical equations for some soil hydraulic properties, Water resources research, 14, 601–604, 1978.

Collins, W., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Halloran, P., Hinton, T., Hughes, J., Jones, C., Joshi, M., Liddicoat, S., et al.: Development and evaluation of an Earth-system model–HadGEM2, Geoscientific Model Development Discussions, 4, 997–1062, 2011.

Coon, E., Moulton, J., Berndt, M., Manzini, G., Garimella, R., Lipnikov, K., and Painter, S.: Coupled Surface and Subsurface Hydrologic Flow using Mimetic Finite Differences, in review, Advances in Water Resources, 2015a.

Coon, E. T., Moulton, J. D., and Painter, S. L.: Managing complexity in land surface and near-surface process models, Environmental Modelling and Software, under review, 2015b.

Cosby, B., Hornberger, G., Clapp, R., and Ginn, T.: A statistical exploration of the relationships of soil moisture characteristics to the physical properties of soils, Water Resources Research, 20, 682–690, 1984.

Davidson, E. A. and Janssens, I. A.: Temperature sensitivity of soil carbon decomposition and feedbacks to climate change, Nature, 440, 165–173, 2006.

Doherty, J.: Model-Independent Parameter Estimation, User Manual:, 2004.

Ekici, A., Beer, C., Hagemann, S., and Hauck, C.: Simulating high-latitude permafrost regions by the JSBACH terrestrial ecosystem model, Geoscientific Model Development, 7, 631–647, 2014.

Gent, P. R., Danabasoglu, G., Donner, L. J., Holland, M. M., Hunke, E. C., Jayne, S. R., Lawrence, D. M., Neale, R. B., Rasch, P. J., Vertenstein, M., et al.: The community climate system model version 4, Journal of Climate, 24, 4973–4991, 2011.

Hinzman, L., Kane, D., Gieck, R., and Everett, K.: Hydrologic and thermal properties of the active layer in the Alaskan Arctic, Cold Regions Science and Technology, 19, 95–110, 1991.

Hinzman, L. D., Goering, D. J., and Kane, D. L.: A distributed thermal model for calculating soil temperature profiles and depth of thaw in permafrost regions, Journal of Geophysical Research: Atmospheres (1984–2012), 103, 28 975–28 991, 1998.

Hinzman, L. D., Bettez, N., Chapin, F., Dyurgerov, M., Fastie, C., Griffith, D., Hope, A., Huntington, H., Jensen, A., Kane, D., et al.: Evidence and implications of recent climate change in terrestrial regions of the Arctic, in: AGU Fall Meeting Abstracts, vol. 1, p. 0010, 2002.

Ji, J.: A climate-vegetation interaction model: Simulating physical and biological processes at the surface, Journal of Biogeography, pp. 445–451, 1995.

Jiang, Y., Zhuang, Q., and O'Donnell, J. A.: Modeling thermal dynamics of active layer soils and near-surface permafrost using a fully coupled water and heat transport model, Journal of Geophysical Research: Atmospheres (1984–2012), 117, 2012.

Jones, C., Hughes, J., Bellouin, N., Hardiman, S., Jones, G., Knight, J., Liddicoat, S., O'Connor, F., Andres, R. J., Bell, C., et al.: The HadGEM2-ES implementation of CMIP5 centennial simulations, Geoscientific Model Development, 4, 543–570, 2011.

Jones, P. D. and Moberg, A.: Hemispheric and large-scale surface air temperature variations: An extensive revision and an update to 2001, Journal of Climate, 16, 206–223, 2003.

Kane, D. L., Hinkel, K. M., Goering, D. J., Hinzman, L. D., and Outcalt, S. I.: Non-conductive heat transfer associated with frozen soils, Global and Planetary Change, 29, 275–292, 2001.

Karra, S., Painter, S., and Lichtner, P.: Three-phase numerical model for subsurface hydrology in permafrost-affected regions, The Cryosphere Discussions, 8, 149–185, 2014.

Koven, C. D., Ringeval, B., Friedlingstein, P., Ciais, P., Cadule, P., Khvorostyanov, D., Krinner, G., and Tarnocai, C.: Permafrost carbon-climate feedbacks accelerate global warming, Proceedings of the National Academy of Sciences, 108, 14 769–14 774, 2011.

Koven, C. D., Riley, W. J., and Stern, A.: Analysis of permafrost thermal dynamics and response to climate change in the CMIP5 Earth System Models, Journal of Climate, 26, 1877–1900, 2013.

Kurylyk, B. L., McKenzie, J. M., MacQuarrie, K. T. B., and Voss, C. I.: Analytical solutions for benchmarking cold regions subsurface water flow and energy transport models: One-dimensional soil thaw with conduction and advection, Advances in Water Resources, 70, 172–184, doi:10.1016/j.advwatres.2014.05.005, http://www.sciencedirect.com/science/article/pii/S0309170814000992, 2014.

Langer, M., Westermann, S., Heikenfeld, M., Dorn, W., and Boike, J.: Satellite-based modeling of permafrost temperatures in a tundra lowland landscape, Remote Sensing of Environment, 135, 12–24, 2013.

Lawrence, D. M. and Slater, A. G.: A projection of severe near-surface permafrost degradation during the 21st century, Geophysical Research Letters, 32, 2005.

Lawrence, D. M. and Slater, A. G.: Incorporating organic soil into a global climate model, Climate Dynamics, 30, 145–160, 2008.

Letts, M. G., Roulet, N. T., Comer, N. T., Skarupa, M. R., and Verseghy, D. L.: Parametrization of peatland hydraulic properties for the Canadian Land Surface Scheme, Atmosphere-Ocean, 38, 141–160, 2000.

Liljedahl, A., Hinzman, L., Harazono, Y., Zona, D., Tweedie, C., Hollister, R. D., Engstrom, R., and Oechel, W.: Nonlinear controls on evapotranspiration in arctic coastal wetlands, Biogeosciences, 8, 3375–3389, 2011.

Ling, F. and Zhang, T.: A numerical model for surface energy balance and thermal regime of the active layer and permafrost containing unfrozen water, Cold Regions Science and Technology, 38, 1–15, 2004.

Mikan, C. J., Schimel, J. P., and Doyle, A. P.: Temperature controls of microbial respiration in arctic tundra soils above and below freezing, Soil Biology and Biochemistry, 34, 1785–1795, 2002.

Moss, R. H., Babiker, M., Brinkman, S., Calvo, E., Carter, T., Edmonds, J. A., Elgizouli, I., Emori, S., Lin, E., Hibbard, K., et al.: Towards new scenarios for analysis of emissions, climate change, impacts, and response strategies, intergovernmental Panel on Climate Change (IPCC), Geneva (Switzerland), 2008.

Neumann, F.: Lectures given in the 1860's CF Riemann-Weber, Die partiellen Differential-gleichungen Physik, p. 121, 1860.

Nicolsky, D., Romanovsky, V., and Panteleev, G.: Estimation of soil thermal properties using in-situ temperature measurements in the active layer and permafrost, Cold Regions Science and Technology, 55, 120–129, 2009.

O'Donnell, J. A., Romanovsky, V. E., Harden, J. W., and McGuire, A. D.: The effect of moisture content on the thermal conductivity of moss and organic soil horizons from black spruce ecosystems in interior Alaska, Soil Science, 174, 646–651, 2009.

Overduin, P. P., Kane, D., and van Loon, W.: Measuring thermal conductivity in freezing and thawing soil using the soil temperature response to heating, Cold regions science and technology, 45, 8–22, 2006.

Painter, S. and Karra, S.: Constitutive model for unfrozen water content in subfreezing unsaturated soils, Vadose Zone Journal, 13, 2014.

Painter, S., Moulton, J., and Wilson, C.: Modeling challenges for predicting hydrologic response to degrading permafrost, Hydrogeology Journal, pp. 1–4, 2013.

Painter, S. L.: Three-phase numerical model of water migration in partially frozen geological media: model formulation, validation, and applications, Computational Geosciences, 15, 69–85, 2011.

Peters-Lidard, C., Blackburn, E., Liang, X., and Wood, E.: The effect of soil thermal conductivity parameterization on surface energy fluxes and temperatures, Journal of the Atmospheric Sciences, 55, 1209–1224, 1998.

Price, J. S., Whittington, P. N., Elrick, D. E., Strack, M., Brunet, N., and Faux, E.: A method to determine unsaturated hydraulic conductivity in living and undecomposed moss, Soil Science Society of America Journal, 72, 487–491, 2008.

Quinton, W., Gray, D., and Marsh, P.: Subsurface drainage from hummock-covered hillslopes in the Arctic tundra, Journal of Hydrology, 237, 113–125, 2000.

Rawlins, M., Nicolsky, D., McDonald, K., and Romanovsky, V.: Simulating soil freeze/thaw dynamics with an improved pan-Arctic water balance model, Journal of Advances in Modeling Earth Systems, 5, 659–675, 2013.

Rinke, A., Kuhry, P., and Dethloff, K.: Importance of a soil organic layer for Arctic climate: A sensitivity study with an Arctic RCM, Geophysical Research Letters, 35, 2008.

Romanovsky, V. and Osterkamp, T.: Interannual variations of the thermal regime of the active layer and near-surface permafrost in northern Alaska, Permafrost and Periglacial Processes, 6, 313–335, 1995.

Romanovsky, V. and Osterkamp, T.: Thawing of the active layer on the coastal plain of the Alaskan Arctic, Permafrost and Periglacial processes, 8, 1–22, 1997.

Romanovsky, V. E., Sazonova, T. S., Balobaev, V. T., Shender, N. I., and Sergueev, D. O.: Past and recent changes in air and permafrost temperatures in eastern Siberia, Global and Planetary Change, 56, 399–413, doi:10.1016/j.gloplacha.2006.07.022, http://www.sciencedirect.com/science/article/pii/S0921818106001974, 2007.

Sachs, T., Wille, C., Boike, J., and Kutzbach, L.: Environmental controls on ecosystem-scale CH4 emission from polygonal tundra in the Lena River Delta, Siberia, Journal of Geophysical Research: Biogeosciences (2005–2012), 113, 2008.

Schaefer, K., Zhang, T., Bruhwiler, L., and Barrett, A. P.: Amount and timing of permafrost carbon release in response to climate warming, Tellus B, 63, 165–180, 2011.

Serreze, M., Walsh, J., Chapin Iii, F., Osterkamp, T., Dyurgerov, M., Romanovsky, V., Oechel, W., Morison, J., Zhang, T., and Barry, R.: Observational evidence of recent change in the northern high-latitude environment, Climatic Change, 46, 159–207, 2000.

Shiklomanov, N. I., Anisimov, O. A., Zhang, T., Marchenko, S., Nelson, F. E., and Oelke, C.: Comparison of model-produced active layer fields: Results for northern Alaska, Journal of Geophysical Research: Earth Surface (2003–2012), 112, 2007.

Slater, A. G. and Lawrence, D. M.: Diagnosing present and future permafrost from climate models, Journal of Climate, 26, 5608–5623, 2013.

Stefan, J.: Über die Theorie der Eisbildung, insbesondere über die Eisbildung im Polarmeere, Annalen der Physik, 278, 269–286, 1891.

Subin, Z., Koven, C., Riley, W., Torn, M., Lawrence, D., and Swenson, S.: Effects of soil moisture on the responses of soil temperatures to climate change in cold regions, in: AGU Fall Meeting Abstracts, vol. 1, p. 0418, 2012.

Tonkin, M. and Doherty, J.: Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques, Water Resources Research, 45, 2009.

Van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, Soil science society of America journal, 44, 892–898, 1980.

Verseghy, D. L.: CLASS—A Canadian land surface scheme for GCMs. I. Soil model, International Journal of Climatology, 11, 111–133, 1991.

Volodin, E., Dianskii, N., and Gusev, A.: Simulating present-day climate with the INMCM4. 0 coupled model of the atmospheric and oceanic general circulations, Izvestiya, Atmospheric and Oceanic Physics, 46, 414–431, 2010.

Walker, W. E., Harremoes, P., Rotmans, J., Van Der Sluijs, J. P., Van Asselt, M. B. A., Janssen, P., and Mayer von Krauss, M. P.: Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support, Inegrated Assessment, 4, 5–7, 2003.

Wania, R., Ross, I., and Prentice, I.: Integrating peatlands and permafrost into a dynamic global vegetation model: 1. Evaluation and sensitivity of physical land surface processes, Global Biogeochemical Cycles, 23, 2009.

Watanabe, M., Suzuki, T., O'ishi, R., Komuro, Y., Watanabe, S., Emori, S., Takemura, T., Chikira, M., Ogura, T., Sekiguchi, M., et al.: Improved climate simulation by MIROC5: mean states, variability, and climate sensitivity, Journal of Climate, 23, 6312–6335, 2010.

Zhang, Y., Carey, S. K., Quinton, W. L., Janowicz, J. R., Pomeroy, J. W., and Flerchinger, G. N.: Comparison of algorithms and parameterisations for infiltration into organic-covered permafrost soils, Hydrology and earth system sciences, 14, 729–750, 2010.