Overview

We thank the reviewers for their constructive comments and suggestions. The manuscript will be appropriately revised in response to the reviewers' comments (see the point-by-point expected responses below). As requested by the reviewers, we compared our modeling results with some existing permafrost data sets by calculating evaluation metrics that can be compared directly against matching results reported in the literature. In addition, we also conducted several new simulations that further assess the impact on ALT of the model soil layer configuration, the soil organic carbon content, and its vertical distribution.

In summary, the planned modifications to the text can be categorized as follows:

a) Novelty and added value:
   See R1C1 (i.e., Reviewer 1, Comment 1), R1C2, and R3C6

b) Comparison with other model-generated permafrost data sets:
   See R1C1, R1C2, R3C6, R3C29 and R3C30

c) Rephrasing "optimistic" discussion about ALT results:
   See R1C10, R1C11, R1C12, R3C2, R3C31

d) New sensitivity experiments and uncertainty discussion:
   See R1C3, R1C20, R2C8 and R2C12

e) Add specific evaluation metrics instead of using description words:
   See R3C6, R3C8, R3C24, R3C29 and R3C31

Throughout the discussion below, the text is colored as follows:

Black: Reviewer comment
Blue: Expected author response
Red: Expected text to be inserted into the revised manuscript

For reference, our response to comment "m" by reviewer "n" is labeled "R[n]C[m]".

Reviewer #2

Overall comments:

This paper used in-situ data and a remote sensing based ALT (active layer thickness) data to evaluate a model-based ALT dataset. Overall, I think it is a useful study. The analysis was done in a comprehensive way, and the results were acceptable. However, the analysis regarding the model uncertainty is somewhat general – considering we already have a good knowledge of the ability of global land models in NH permafrost simulation. I think, the study could benefit from more in-depth discussions on this aspect. The details were provided below.

We thank the reviewer for the careful reviewing. We absolutely agree with the reviewer about the importance of model uncertainty regarding ALT estimation. To further examine model uncertainty, we conducted new tests on sensitivity to organic carbon content and its vertical distribution. We also replaced the vegetation climatology at several Mongolian sites with satellite-based, time-variant LAI to investigate the impact of inter-annual variations in vegetation. Please also see our response to Reviewer #1 (R1C21) regarding the discussion about the soil layer configuration.

Major comments:

1. Page 4 Paragraph 3: I have questions regarding how the ALT was calculated. The paper indicates here it was calculated based on the simulated ice content. Does the model consider unfrozen water in frozen soils? If it does, please provide information on how the model calculates unfrozen water content. If not, this definition will be same as using a 0_C temperature threshold for thawed-to-frozen depth calculation. This information is especially important for the deep soils due to year-round low temperatures and coarse vertical resolution of the model at deeper depths.

R2C1: In our model, the ice fraction is unity if soil is fully frozen (i.e., $T < 0^{o}C$ and $f_{ice} = 1$), and it is zero if the soil is fully thawed (i.e., $T > 0^{o}C$ and $f_{ice} = 0$). The reviewer comment refers to the situation where the soil temperature is exactly at the freezing point and the soil is partially frozen (i.e., $T = 0^{o}C$ and $0 < f_{ice} < 1$). In the latter case, frozen and thawed soil and water coexist. This situation always occurs during freeze-to-thaw and thaw-to-freeze transitions. This is because in soil layers that are as thick as those used in the model, the phase transition does not occur instantly (relative to the model time step). More specifically, the model uses heat content as the prognostic variable, from which the temperature and ice fraction are diagnosed. Therefore, our calculation of ALT is not the same as simply using a 0ºC degree threshold to determine the thawed-to-frozen depth. Rather, we identify the deepest (fully or partially) thawed layer and then calculate the thawed-to-frozen depth based on the ice fraction within the layer. We will modify the paragraph about the ALT calculation as follows.

> "Precisely, the thawed-to-frozen depth is calculated as:

$$z_{\mathrm{bottom}}(l) - f_{\mathrm{ice}}(l, t) \times \Delta z(l), \qquad (1)$$

where layer l is the deepest layer that is fully or partially thawed, $z_{\mathrm{bottom}}(l)$ represents the depth at the bottom of layer l, $f_{ice}(l,t)$ is the fraction of ice in layer l at time t, where $f_{ice}(l, t) \in [0 \;\; 1]$, and $\Delta z(l)$ is the thickness of layer l. To identify layer l we use a 0ºC degree temperature threshold. Specifically, T > 0ºC degree indicates that a layer is fully thawed, T = 0ºC degree indicates that a layer is partially thawed, and T < 0ºC degree indicates that a layer is fully frozen. That is, layer l is the deepest layer that satisfies $T(l) \geq 0$ºC. Equation (1) then expresses that the thawed-to-frozen depth is equal to the bottom depth of the layer l but adjusted upward according to the ice fraction within the partially thawed layer l."

The above declaration seems contradictory to "The use of the 0_C degree threshold in CLSM for determining the thawed or frozen state of the soil may explain the model's underestimation of ALT." (Page 9, Paragraph 1). So I am confused what methods were actually used to determine the thawing depth/ALT. Please clarify.

R2C2: The 0ºC degree is used to determine the deepest thawed layer l, and then the ALT is calculated by $z_{\mathrm{bottom}}(l) - f_{\mathrm{ice}}(l, t) \times \Delta z(l)$ as explained above (R2C1). We will modify the text as follows:

"The use of the 0ºC degree temperature threshold in CLSM (along with the ice fraction) for determining the model's ALT is thus only an approximation. These and other simplifications contribute to the model's underestimation of ALT."

2. Page 5 Paragraph 3: The spin-up scheme is questionable, though the authors themselves acknowledged this. Why do the authors using the meteorology for the entire 36-year period for spin-up? If the design is to reduce the uncertainty introduced by a single-year surface meteorology, spin-up using the first few years during the period will be more acceptable.

R2C3: As the reviewer points out, we recognized this issue and discussed it in the original manuscript. No spin-up procedure is entirely problem-free. Using a shorter period for spin-up as suggested by the reviewer would exaggerate in the resulting initial conditions any anomalies that occur during the spin-up period. The ultimate solution would be to construct a realistic historical forcing dataset over hundreds of years with a dynamic geothermal flux applied to the bottom boundary of soil column (e.g., Sapriza-Azuri et al., 2018). However, this approach is hardly feasible and would still not assure absolutely correct initial conditions. Please also see our response to Reviewer #1 (R1C9).

3. I have questions regarding the vegetation effects on permafrost simulation in Northern Alaska (Page 11, Paragraph 2). Those 4 northern flights were dominated by "dwarf trees" as indicated by Fig. 2b (really?). Moreover, the changes in simulated maximum snow depth due to vegetation in those flights were much smaller comparing with the experiment homogenizing the forcing data (Fig. 6c). So I would expect the impact due to snow changes for the homogenizing vegetation experiment would be smaller comparing with the experiment homogenizing surface forcing, while this is not the case shown in Fig. 6a-b. Can the authors explain why?

R2C4: As we mentioned in the manuscript, the homogenization is applied cumulatively. Before we homogenized vegetation in this experiment, we already homogenized the forcing. The differences between HomF and HomF&Veg are then attributed to the changes in vegetation parameters (specifically LAI and vegetation height).

Fig.6a-b indeed demonstrates that the impact due to snow changes for the homogenizing vegetation experiment (differences between HomF and HomF&Veg) would be smaller compared to the experiment homogenizing surface forcing (differences between Baseline and HomF). Figure 7b also illustrates this in a quantitative way.

On the other hand, the Alaska North slope is dominated by tundra, while the vegetation map in CLSM indicates mostly dwarf trees or shrubs (Fig. 2b). Would this introduce uncertainties to the analysis on the contribution of different factors (i.e. forcing data, vegetation and soil) in Fig. 6?

R2C5: The vegetation class is only one of several model inputs. The land cover class used in the study is derived from the USGS Global Land Cover Characteristics Data Base Version 2.0 (GLCCv2). In addition to vegetation class, the model uses vegetation height, leaf-area index (LAI), greenness fraction and albedo, which are all obtained from other satellite-based sources that reflect realistic conditions for tundra. Put differently, while the modeled vegetation class may suggest the presence of dwarf trees, the typically low (satellite) LAI values in northern Alaska will instruct the model that the tree cover is extremely sparse in this region. Please refer to Table 1 in (Tao et al., 2017) to see all the data sources. We will add one sentence to Section 4.1 to clarify this.

> "Note that although the vegetation class (Figure 2b) suggests the presence of dwarf trees over the Alaska North Slope, the actual satellite-based LAI, vegetation height, greenness fraction and albedo will still instruct the model that the tree cover is extremely sparse in this region. The data sources for these vegetation-related boundary conditions can be found in Table 1 in Tao et al. (2017)."

Also, from Fig. 6, it does not seem to me that homogenizing soil parameters has much bigger impact on simulated ALT and surface soil temperature than homogenizing vegetation (at least for the northern flights), as the authors indicated in Fig. 7. Maybe I miss something here?

R2C6: We will add the following discussion to Section 4.2.

> "Also, over the northern transects (ATQ, BRW and DHO) the soil impact on ALT (difference between HomF&Veg&Soil in black and HomF&Veg in red) appears smaller than that of the vegetation (difference between HomF in green and HomF&Veg in red), as shown in Figure 6a. But the integrated impact along the transects as shown in Figure 7b indicates that the soil influence clearly outweighs the influence of vegetation, since the changes in vegetation parameters do not have much impact at several other transects, including HUS, KYK, COC, AMB, IVO

> and the first half of ATQ, where the vegetation conditions might be similar to those used for homogenizing. ”

4. Most of the value for AirMOSS radar retrievals, I think, is in its ability to characterize land surface heterogeneity. Simply averaging AirMOSS data to a much coarseresolution (i.e. 9km in this study) to compare with the land model simulations is not very insightful in terms of exploring the value of this dataset. I agree with the authors that the current AirMOSS retrievals seem having large uncertainties; most notably, its ALT retrievals were in a very narrower range. However, the inconsistency of the ALT spatial pattern at some of AirMOSS flights may come from the model itself. For example, at the DHO flightâ˘Aˇ Tthis is the flight with most in-situ sites available, the model ALT generally increases from the north to the south, while the in-situ data show large variability, and do not show a clear increasing trend from the north to the south (Fig. 5a). There are also a number of studies pointing out that ALT is extremely variable at local scale. Therefore, analysis using a dataset like AirMOSS in this aspect would be more valuable.

R2C7: We agree with the value of AirMOSS radar retrievals in terms of (theoretically) being able to represent the spatial variability of ALT. Note, however, that Figure 4 also compares the radar retrievals at the site scale with in-situ observations and demonstrates that the radar retrievals exhibit too little variability also at their native resolution.

The differences in the spatial patterns of the AirMOSS ALT retrievals and the simulated ALT suggest that neither radar remote sensing nor modeling is perfect. As we mentioned in the manuscript, the radar sensing depth (about 60cm) strongly constrains the retrieval accuracy. We expanded on the analysis by adding a new Table 3, which provides several evaluation metrics for ALT restricted to less than 60cm. The table suggests that the radar retrievals are in better agreement with in-situ observations especially at model scale when only using sites that have ALT less than or equal to 60cm. We will add several sentences to Section 4.1:

> “Excluding the sites with in-situ ALT measurements that exceed 60 cm, the overall mean bias for the AirMOSS retrievals at the model scale (site scale) drops to -0.01 m (0.02 m), and the correlation coefficient at the model scale (site scale) increases to 0.64 (0.20). In contrast, the CLSM simulation results show a bias of 0.01 m and a zero correlation coefficient at the same sites.”

Please also see our response in R1C12.

5. It would be more interesting if the authors could provide more insightful analysis regarding the model ALT uncertainties or the correlation analysis, including:

a) Why does the model show much stronger correlation with maximum SWE in portions of NH permafrost region than with air temperature?

R2C8a: Good question. We will add a sentence into Section 4.3 of the revised manuscript.

> "One possible explanation is that the warming impact of the current climate has not yet had an impact on subsurface heat transfer over these areas because the insulation provided by the snow pack prevents such an influence."

b) It would be helpful if the authors could give more explanations regarding why the model fails in western Siberia? Those areas also include continuous permafrost, so I do not think it is too challenging for global models to capture the permafrost distribution there.

R2C8b: As indicated in Figure 1b, all four types of permafrost (i.e., continuous, discontinuous, sporadic and isolated) are present in the western Siberia. The literature suggests that other global models also missed this portion of permafrost, including CLM3 and CCSM3, although the updated versions of these models (i.e., CLM4 and CCSM4) demonstrated improved performance (Lawrence et al., 2012). Similarly, Guo et al. (2017) also reported underestimations in permafrost extent in western Siberia simulated by CLM4.5 when driven with three different reanalysis forcings (CFSR, ERA-I and MERRA) and improved performance when using forcing data from a different reanalysis. We will add some discussion to compare our simulation with other existing works:

> "Note that some other global models, such as CLM3 and CCSM3 as reported in Lawrence et al. (2012), also missed this area of permafrost and that updated versions of these models (i.e., CLM4 and CCSM4) showed improved performance in this regard (Lawrence et al., 2012). Guo et al. (2017) reported underestimated permafrost extent simulated in western Siberia by CLM4.5 driven by three different reanalysis forcings (i.e., CFSR, ERA-I and MERRA), and they showed an improved simulation of permafrost extent in this area when using another reanalysis forcing, the CRUNCEP (Climatic Research Unit - NCEP) (Guo and Wang, 2017). Guimberteau et al. (2018) found similar improvements stemming from the use of CRUNCEP forcing. We leave for further study whether the MERRA-2 forcing data is responsible for the western Siberia deficiency seen in our own results."

c) The ALT trends shown in Fig. 13a seem not very consistent with the trends of temperature indices shown in this figure. Have the authors explore the changes in snow cover duration? A longer snow free season generally leads to warmer soil temperature and thus deep ALT esp. in the southern area.

R2C8c: We did examine the trend in snow cover duration (see Figure R4). While some areas show a trend in snow cover duration, this trend does not seem correlated with the trend in ALT.

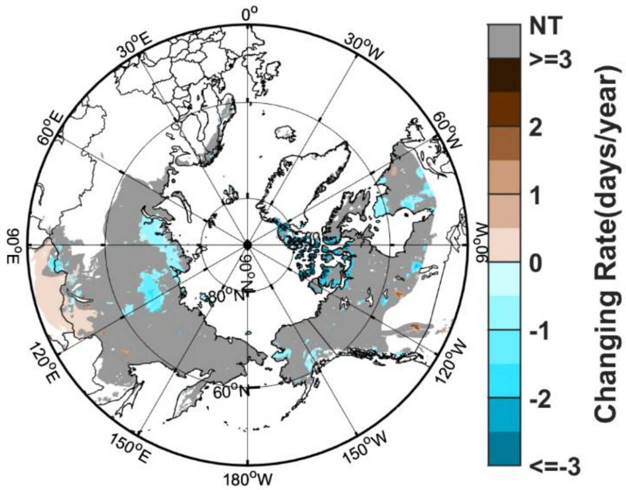

Figure R4: Spatial distribution of trend in snow persistence days when daily mean snow depth > 25cm.

Accordingly, we plan to modify the relevant sentence as follows:

> "It is possible that in such cases, the computed trends are strongly affected by snowpack variability, even though maximum SWE itself does not tend to show a significant trend in these areas (not shown), neither the snow cover duration (not shown). "

d) Page 18, Paragraph 3: I do not quite agree with the authors' explanation why the model fails in the Mongolian sites. For those sites, the model simulated ALT is generally less than 1.4 m, therefore, the coarse resolution at deeper soils in the model set-up (i.e. layer 5: 1.4-3m) should not be a major contributing factor there. Much drier conditions, sparse vegetation, and perhaps uncertainties in soil texture data (eps. in deep soils), I think, are more likely contributing more to the model uncertainties.

R2C8d: The reviewer is correct about the soil configuration. We conducted new tests using different soil configurations at several Mongolian sites. Please see the results and discussion in our response to the Reviewer #1 (R1C21).

The reviewer also raised a very good point regarding the influence on ALT of soil wetness, vegetation and uncertainties in soil texture for deep soils. We first examined realistic satellite-based LAI data at from Moderate Resolution Imaging Spectroradiometer (MODIS) MCD15A2H product and the Advanced very-high-resolution radiometer (AVHRR) AVH15C1 product (see Table R3). Figure R5 then shows the time series of the MODIS and AVHRR LAI, along with the LAI climatology used in the model at one CALM Mongolian site (M11). A post-processing procedure that included quality screening and gap filling was applied to the two satellite LAI products. The CLSM LAI climatology is used for the years that MODIS data is not available (1980 to 2002).

Table R3 – Information of satellite-based LAI products from MODIS and AVHRR.

Sensor	Dataset	Product	Resolution	Temporal Granularity	Temporal Extent
MODIS	MCD15A2H	Leaf Area Index and Fractional Photosynthetically Active Radiation	500 m	8-day Composites	July 2002 - Present
AVHRR	AVH15C1	Leaf Area Index and Fraction of Absorbed Photosynthetically Active Radiation	0.05deg	Daily	June 1981 - Present

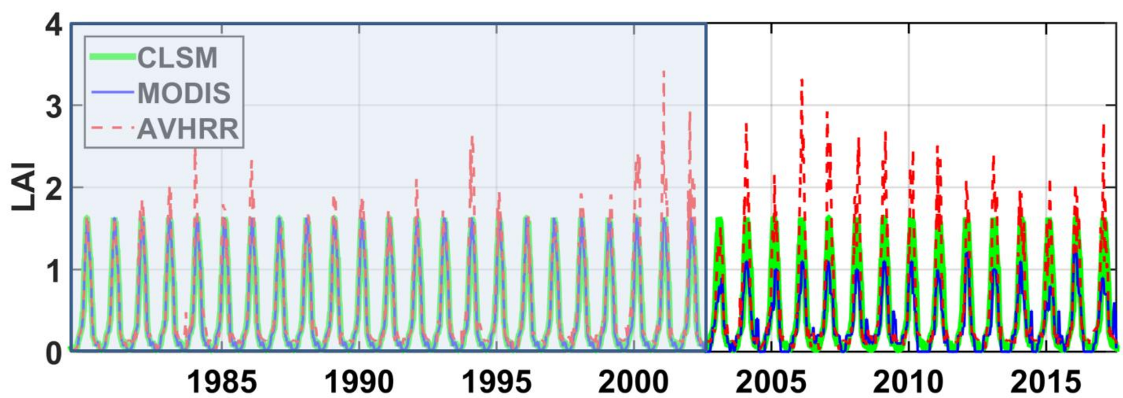

Figure R5: The LAI time series at one Mongolian site (M11). Green line represent the original LAI climatology used in the CLSM. Blue and red dash line represent the realistic (time-varying) LAI data from MODIS and AVHRR.

As illustrated by Figure R5, MODIS shows smaller LAI than AVHRR over the valid period after 2002. The LAI climatology used in the model is inbetween of the two products. The differences between the LAI climatology used in the model and realistic LAI products would cause differences in energy and water partitioning at the land surface via impacting surface albedo. We conducted a simple test to examine the impact of using the more realistic, inter-annually varying vegetation inputs on the winter surface albedo and thus the snow accumulation process, which in turn would impact ALT estimates. Specifically, we replaced the LAI climatology with the satellite-based, inter-annually varying LAI products in the model, but turned off the impacts in summer, i.e., not affecting the snow-free albedo. The simulation results show only minimal differences in the estimated ALT. That is, the winter surface albedo when using realistic satellite LAI products does not differ very much from that using the original LAI climatology. However, we speculate that large differences in summer could have significant impact on ALT estimation. We leave further investigation for future work. We will add one sentence in the manuscript to bring up this issue.

> "Another issue affecting our ALT comparisons is the climatological representation of vegetation parameters such as LAI used in CLSM. Additional investigation (not

shown) revealed large differences between the LAI climatology used in CLSM and more realistic, time-varying LAI products at several Mongolian sites."

Without any further information about the soil parameters in deep soils, we could not conduct further tests. Here we provided the results on the sensitivity of soil organic carbon and the vertical distribution in Figure 6. The figure reveals that a deeper ALT results from reducing the SOC content and from using a very different vertical distribution profile that arbitrarily concentrates less carbon in the top soil. Indeed, changing the vertical distribution profile for SOC content plays an almost equivalent role to changing the SOC content.

This further confirms the reviewer's comment regarding the importance of vertical variation of soil properties. Thus, we will add one sentence here to bring up the issue about vertical variation in soil parameters.

"Besides the vertical distribution of soil carbon, the vertical variation in other soil hydrological properties (e.g. soil texture, porosity, hydraulic conductivity, etc.) should also play a significant role since they all affect soil thermal conductivity and heat capacity."

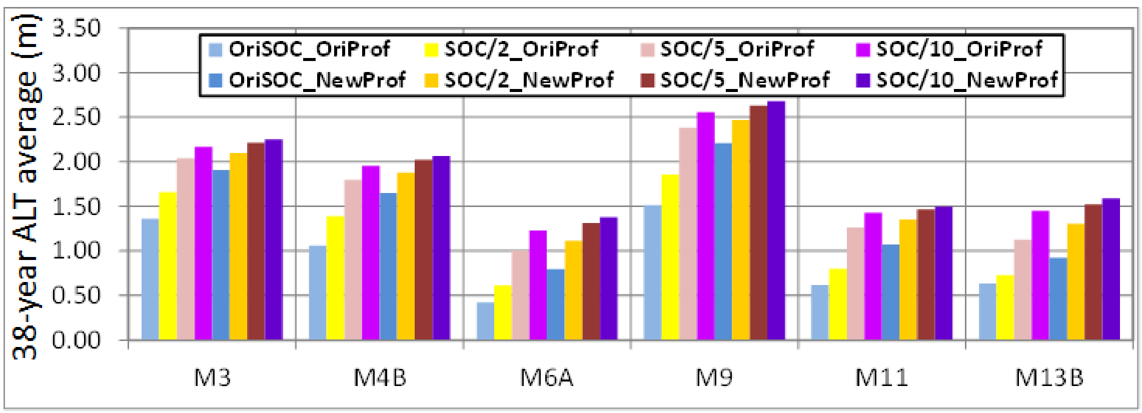

Figure R6: Simulation results at six Mongolian sites with different soil carbon contents and vertical carbon distributions. "OriSOC_OriProf" – Original soil organic carbon (SOC) content vertically distributed with the original profile as used in baseline simulation. "SOC/N_OriProf" – Reduced soil organic carbon content (by dividing the original SOC content by N) vertically distributed using the original profile. "OriSOC_NewProf" – Original SOC content vertically distributed with a new profile which arbitrarily concentrates less carbon in top soils. "SOC/N_NewProf" – Reduced SOC content (by dividing the original SOC content by N) vertically distributed using the new profile.

Minor comments:

1. Fig. 4: The text in Section 2.3 indicates most of the comparisons against AirMOSS data will be at the 4 flights in the Alaska north slope. So I wonder why COC flight was included in this figure esp. considering the AirMOSS retrievals were not included?

R2C9: We only have AirMOSS retrievals for IVO, ATQ, BRW and DHO, not for COC. However, our model provides results here, thus we included COC to add two additional measurements to compare with model results (Figure 4b).

2. Fig. 4 a-b: were the sites arranged according to the latitudinal changes?

R2C10: We arranged the sites aligning with the flight direction. We will also add this into the caption of Figure 4.

> "The sites are arranged aligning with the flight direction."

3. Page 9, Line 21: Does MERRA-2 not provide air temperature at 2-m surface height?

R2C11: MERRA-2 does provide output of hourly 2-m air temperature. However, the land model within the MERRA-2 system is forced with the air temperature in the lowest (atmospheric) model layer (TLML), and the 2-m temperature is simply diagnosed from TLML and the surface temperature. For consistency, the off-line (land-only) model simulations presented here were likewise driven with TLML from MERRA-2. In any case, the sentence in question was not necessary and only caused confusion, so we will delete it.

4. Section 3.2 and Figure 6: Part of the IVO flight lies in the Brooks mountain range with very low SOC content (Fig. 2c). It may be not a good representative of the average conditions in this area, at least considering SOC variability.

R2C12: The point at which we extracted the soil parameters does have a sort of intermediate SOC (greenish color in Fig.2c) which is also shown in Figure 5b. As mentioned in the original manuscript, however, we actually used an arbitrary intermediate SOC value which is 40 kg/m^2.

We also conducted two additional simulations using a very large and a very small SOC value everywhere. The results are shown in Figure R7 below. The "IntermC" used the same SOC that was used for homogenization (i.e., 40 kg/m^2). "LowC" and "HighC" used the lowest (10kg/m^2) and highest (120kg/m^2) SOC values found along the transects as shown in Figure 5b and also in Figure R7b. Figure R7a reveals that the model sensitivity to soil carbon is much larger for lower SOC than for higher SOC, and easily gets saturated for high SOC (i.e., larger than ~100 kg/m^2). However, all of this depends on the vertical soil carbon distribution profile used. Please also see R2C8d.

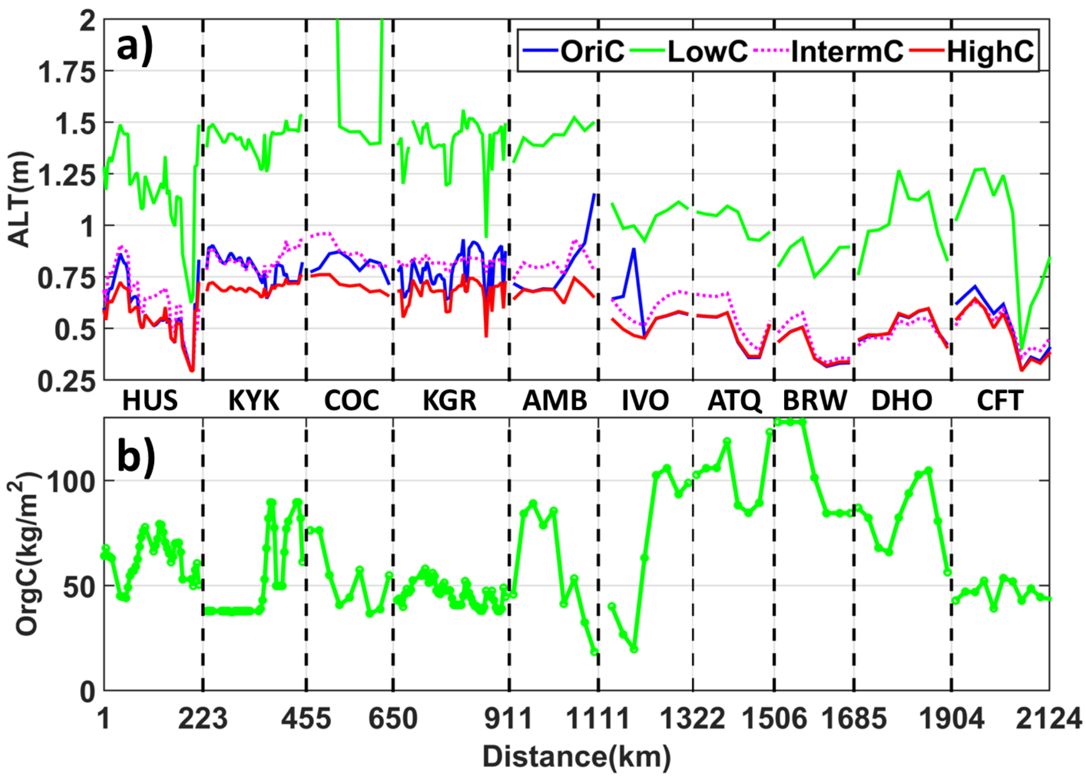

Figure R7: a) similar to the Figure 5a in the original manuscript but showing model sensitivity to organic carbon content along the AirMOSS flight transect. b) the organic carbon content along the transect.

We will add a sentence into the Section 3.2.

> "Our investigation reveals that the model sensitivity to soil carbon content is much larger for lower SOC than for higher SOC, and easily gets saturated for high SOC (i.e., larger than ~100 kg/m^2) (not shown). Thus, we trust 40 kg/m^2 is an appropriate value representing an intermediate SOC condition."

5. Section 3.2: Would it be easier to follow if the description regarding the idealized experiments was included in the Methods section?

R2C13: We thank the reviewer for the helpful suggestion. We will introduce the idealize experiments in Section 2 by adding a new Methods section (section 2.5).

6. Fig. 7 is not very informative. I suggest summarizing the results in a table.

R2C14: We feel that Fig.7 is informative, which is also supported by Reviewer #1 (R1C15). The figure displays the actual values and therefore also serves as a table. We therefore opt to keep the figure in its current form.

Reference

Alexeev, V. A., Nicolsky, D. J., Romanovsky, V. E., and Lawrence, D. M.: An evaluation of deep soil configurations in the CLM3 for improved representation of permafrost, Geophys Res Lett, 34, 10.1029/2007gl029536, 2007.

Bosilovich, M. G., Akella, S., Coy, L., Cullather, R., Draper, C., Gelaro, R., Kovach, R., Liu, Q., Molod, A., Norris, P., Wargan, K., Chao, W., Reichle, R., Takacs, L., Vikhliaev, Y., Bloom, S., Collow, A., Firth, S., Labow, G., Partyka, G., Pawson, S., Reale, O., Schubert, S. D., and Suarez, M.: MERRA-2: Initial Evaluation of the Climate, NASA Technical Report Series on Global Modeling and Data Assimilation, NASA/TM-2015-104606, Vol. 43, National Aeronautics and Space Administration, Goddard Space Flight Center, Greenbelt, Maryland, USA, 2015.

Bosilovich, M. G., Robertson, F. R., Takacs, L., Molod, A., and Mocko, D.: Atmospheric Water Balance and Variability in the MERRA-2 Reanalysis, Journal of Climate, 30, 1177-1196, 2017.

Brown, J., Ferrians, O., Heginbottom, J. A., and Melnikov, E.: Circum-Arctic Map of Permafrost and Ground-Ice Conditions, Version 2. [Permafrost Extent]. NSIDC: National Snow and Ice Data Center, Boulder, Colorado USA., 2002.

Chadburn, S. E., Burke, E. J., Cox, P. M., Friedlingstein, P., Hugelius, G., and Westermann, S.: An observation-based constraint on permafrost loss as a function of global warming, Nat Clim Change, 7, 340-+, 2017.

Chen, R. H., Tabatabaeenejad, A., and Moghaddam, M.: A time-series active layer thickness retrieval algorithm using P-and L-band SAR observations, Geoscience and Remote Sensing Symposium (IGARSS), 2016 IEEE International, 2016, 3672-3675,

Chen, R. H., Tabatabaeenejad, A., and Moghaddam, M.: Retrieval of permafrost active layer properties using time-series P-band radar observations, IEEE Transactions on Geoscience and Remote Sensing. (In Review), 2018.

Cheng, L., and Zhu, J.: 2017 was the warmest year on record for the global ocean, Adv Atmos Sci, 35, 261-263, 2018.

Dankers, R., Burke, E. J., and Price, J.: Simulation of permafrost and seasonal thaw depth in the JULES land surface scheme, Cryosphere, 5, 773-790, 2011.

Evans, S. G., Ge, S., and Liang, S.: Analysis of groundwater flow in mountainous, headwater catchments with permafrost, Water Resources Research, 51, 9564-9576, 10.1002/2015wr017732, 2015.

Evans, S. G., Ge, S., Voss, C. I., and Molotch, N. P.: The Role of Frozen Soil in Groundwater Discharge Predictions for Warming Alpine Watersheds, Water Resources Research, 54, 1599-1615, 10.1002/2017wr022098, 2018.

Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G. K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), Journal of Climate, 30, 5419-5454, 2017.

Gisnas, K., Etzelmuller, B., Farbrot, H., Schuler, T. V., and Westermann, S.: CryoGRID 1.0: Permafrost Distribution in Norway estimated by a Spatial Numerical Model, Permafrost Periglac, 24, 2-19, 2013.

Guimberteau, M., Zhu, D., Maignan, F., Huang, Y., Yue, C., Dantec-Nedelec, S., Ottle, C., Jornet-Puig, A., Bastos, A., Laurent, P., Goll, D., Bowring, S., Chang, J. F., Guenet, B., Tifafi, M., Peng, S. S., Krinner, G., Ducharne, A., Wang, F. X., Wang, T., Wang, X. H., Wang, Y. L., Yin, Z., Lauerwald, R.,

Joetzjer, E., Qiu, C. J., Kim, H., and Ciais, P.: ORCHIDEE-MICT (v8.4.1), a land surface model for the high latitudes: model description and validation, Geoscientific Model Development, 11, 121-163, 2018.
Guo, D. L., and Wang, H. J.: Simulated Historical (1901-2010) Changes in the Permafrost Extent and Active Layer Thickness in the Northern Hemisphere, Journal of Geophysical Research-Atmospheres, 122, 12285-12295, 10.1002/2017jd027691, 2017.
Guo, D. L., Wang, H. J., and Wang, A. H.: Sensitivity of Historical Simulation of the Permafrost to Different Atmospheric Forcing Data Sets from 1979 to 2009, Journal of Geophysical Research-Atmospheres, 122, 12269-12284, 2017.
Hugelius, G., Bockheim, J. G., Camill, P., Elberling, B., Grosse, G., Harden, J. W., Johnson, K., Jorgenson, T., Koven, C. D., Kuhry, P., Michaelson, G., Mishra, U., Palmtag, J., Ping, C. L., O'Donnell, J., Schirrmeister, L., Schuur, E. A. G., Sheng, Y., Smith, L. C., Strauss, J., and Yu, Z.: A new data set for estimating organic carbon storage to 3m depth in soils of the northern circumpolar permafrost region, Earth Syst Sci Data, 5, 393-402, 10.5194/essd-5-393-2013, 2013a.
Hugelius, G., Tarnocai, C., Broll, G., Canadell, J. G., Kuhry, P., and Swanson, D. K.: The Northern Circumpolar Soil Carbon Database: spatially distributed datasets of soil coverage and soil carbon storage in the northern permafrost regions, Earth Syst Sci Data, 5, 3-13, 10.5194/essd-5-3-2013, 2013b.
Jafarov, E. E., Marchenko, S. S., and Romanovsky, V. E.: Numerical modeling of permafrost dynamics in Alaska using a high spatial resolution dataset, Cryosphere, 6, 613-624, 10.5194/tc-6-613-2012, 2012.
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, 10.1175/Jcli-D-12-00228.1, 2013.
Lawrence, D. M., and Slater, A. G.: A projection of severe near‐surface permafrost degradation during the 21st century, Geophys Res Lett, 32, 2005.
Lawrence, D. M., and Slater, A. G.: Incorporating organic soil into a global climate model, Clim Dynam, 30, 145-160, 2008.
Lawrence, D. M., Slater, A. G., Romanovsky, V. E., and Nicolsky, D. J.: Sensitivity of a model projection of near-surface permafrost degradation to soil column depth and representation of soil organic matter, Journal of Geophysical Research-Earth Surface, 113, 10.1029/2007jf000883, 2008.
Lawrence, D. M., Slater, A. G., and Swenson, S. C.: Simulation of Present-Day and Future Permafrost and Seasonally Frozen Ground Conditions in CCSM4, Journal of Climate, 25, 2207-2225, 10.1175/Jcli-D-11-00334.1, 2012.
Nicolsky, D., Romanovsky, V., Alexeev, V., and Lawrence, D.: Improved modeling of permafrost dynamics in a GCM land‐surface scheme, Geophys Res Lett, 34, 2007a.
Nicolsky, D. J., Romanovsky, V. E., Alexeev, V. A., and Lawrence, D. M.: Improved modeling of permafrost dynamics in a GCM land-surface scheme, Geophys Res Lett, 34, 2007b.
Peng, S., Ciais, P., Krinner, G., Wang, T., Gouttevin, I., McGuire, A. D., Lawrence, D., Burke, E., Chen, X., Decharme, B., Koven, C., MacDougall, A., Rinke, A., Saito, K., Zhang, W., Alkama, R., Bohn, T. J., Delire, C., Hajima, T., Ji, D., Lettenmaier, D. P., Miller, P. A., Moore, J. C., Smith, B., and Sueyoshi, T.: Simulated high-latitude soil thermal dynamics during the past 4 decades, Cryosphere, 10, 179-192, 10.5194/tc-10-179-2016, 2016.
Reichle, R. H., Draper, C. S., Liu, Q., Girotto, M., Mahanama, S. P. P., Koster, R. D., and De Lannoy, G. J. M.: Assessment of MERRA-2 land surface hydrology estimates, Journal of Climate, 0, null, 10.1175/jcli-d-16-0720.1, 2017a.

Reichle, R. H., Liu, Q., Koster, R. D., Draper, C. S., Mahanama, S. P. P., and Partyka, G. S.: Land Surface Precipitation in MERRA-2, Journal of Climate, 30, 1643-1664, 10.1175/jcli-d-16-0570.1, 2017b.
Sapriza-Azuri, G., Gamazo, P., Razavi, S., and Wheater, H. S.: On the appropriate definition of soil profile configuration and initial conditions for land surface-hydrology models in cold regions, Hydrol Earth Syst Sc, 22, 3295-3309, 10.5194/hess-22-3295-2018, 2018.
Tao, J., Reichle, R. H., Koster, R. D., Forman, B. A., and Xue, Y.: Evaluation and Enhancement of Permafrost Modeling With the NASA Catchment Land Surface Model, Journal of Advances in Modeling Earth Systems, 9, 2771-2795, 10.1002/2017MS001019, 2017.
Wang, W., Rinke, A., Moore, J. C., Cui, X., Ji, D., Li, Q., Zhang, N., Wang, C., Zhang, S., Lawrence, D. M., McGuire, A. D., Zhang, W., Delire, C., Koven, C., Saito, K., MacDougall, A., Burke, E., and Decharme, B.: Diagnostic and model dependent uncertainty of simulated Tibetan permafrost area, Cryosphere, 10, 287-306, 10.5194/tc-10-287-2016, 2016a.
Wang, W. L., Rinke, A., Moore, J. C., Ji, D. Y., Cui, X. F., Peng, S. S., Lawrence, D. M., McGuire, A. D., Burke, E. J., Chen, X. D., Decharme, B., Koven, C., MacDougall, A., Saito, K., Zhang, W. X., Alkama, R., Bohn, T. J., Ciais, P., Delire, C., Gouttevin, I., Hajima, T., Krinner, G., Lettenmaier, D. P., Miller, P. A., Smith, B., Sueyoshi, T., and Sherstiukov, A. B.: Evaluation of air-soil temperature relationships simulated by land surface models during winter across the permafrost region, Cryosphere, 10, 1721-1737, 10.5194/tc-10-1721-2016, 2016b.
Yi, S. H., Woo, M. K., and Arain, M. A.: Impacts of peat and vegetation on permafrost degradation under climate warming, Geophys Res Lett, 34, 2007.
Yi, Y., Kimball, J. S., Chen, R. H., Moghaddam, M., Reichle, R. H., Mishra, U., Zona, D., and Oechel, W. C.: Characterizing permafrost active layer dynamics and sensitivity to landscape spatial heterogeneity in Alaska, The Cryosphere, 12, 145, 2018.