## 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 #1

This paper provides an evaluation of active layer thickness and permafrost extent as simulated by the NASA Catchmant Land Surface Model driven by MERRA-2. The model-generated dataset of permafrost conditions is evaluated again site data, global data and remotely (plane) sensed ALT. The comparison to the remotely sensed ALT is probably the most innovative part of the paper, but it also suffers from some drawbacks because the remotely sensed data conspicuously lack spatial variability. The paper is written clearly, the analysis is honest (not obviously trying to hide model shortcomings – but sometimes the assessment of the dataset quality seems a bit too optimistic), the figures are relevant and informative. The paper is a useful contribution, but some aspects detailed in the following could be improved.

We thank the reviewer for the careful reviewing. We understand the reviewer's concern about the small spatial variability with the remotely sensed ALT retrievals. The ALT retrievals were produced by the current algorithm developed by Chen et al., 2018. Somewhat larger spatial variability is presented in the original retrievals but is smoothed and reduced after aggregating to the scale of the Catchment Land Surface Model (CLSM) at 81 km$^2$, as also mentioned in the original manuscript. In addition, the radar penetration depth is not large enough to detect deeper thawed to frozen condition of the soil. All in all, while we eventually expect to further improve the retrieval algorithm, these are the results we have right now, and as discussed further below, their inclusion in this paper in their early form does, we feel, have value. We plan to tamp down our "optimistic" tone when discussing the ALT retrievals. We trust this first intercomparison of ALT among model results, remotely sensed retrievals and in-situ observations could provide useful insights to the research community. Please see specifically our response to R1C11, R1C12 and R2C7 below.

General:

- There are lots of global permafrost simulations that are driven by reanalysis-based meteorologies. What is really the added value of this one? The fact that is uses MERRA? In that case, could you say more about specific strengths and weaknesses of MERRA, please? More generally, simulations with other metorological forcing data, and comparison to other model-generated permafrost data sets (e;g. within the Permafrost Carbon Network) could be interesting.

R1C1: We thank the reviewer for encouraging us to explicitly highlight our contributions. Regarding the second, more general comment, a detailed quantitative comparison against existing permafrost data sets simulated by other land models, for example the land models participated in the Permafrost Carbon Network (PCN) (offline mode) and the Coupled Model Intercomparison Project phase 5 (CMIP5) (coupled mode), is beyond the scope of this paper, which is already a bit long.  (Note that aspects of such a general analysis have already been reported in literature (Peng et al., 2016;Wang et al., 2016b;Koven et al., 2013)). Inspired by the reviewer's comment, though, we plan to add a brief discussion to our manuscript that compares our dataset with others in terms of spatiotemporal resolution and simulated results (see our response in R3C6 for details). We will also summarize our dataset's particular strengths in our discussion section (will be the new section 5).

Specifically, we will add the discussion below.

a. General comparison between this work and existing model-generated permafrost data sets forced by other meteorological forcing data (both in uncoupled and coupled mode):

a.1 Regarding resolution (will be added into section 1)

"Most of these land models were run at coarse spatial resolution, e.g., ranging from 0.5° × 0.5° to 1.8° × 3.6° for LSMs participating in the Permafrost Carbon Network (PCN) (Wang et al., 2016a) and from 0.188° × 0.188° to 4.10° × 5° for the models participating in the Coupled Model Intercomparison Project phase 5 (CMIP5) (Koven et al., 2013; https://portal.enes.org/data/enes-model-data/cmip5/resolution). They thus lack some of the higher resolution information implicit in our data and are, as a result, perhaps less comparable to in-situ observations taken at the point scale. Other types of numerical models have been run at relatively higher resolution, but not globally; such simulation domains were limited to regional scales (e.g., 2 km × 2 km in Jafarov et al., 2012 covering Alaska;1 km × 1 km in Gisnas et al., 2013 covering Norway) as necessitated by the availability of ancillary data and the heavy computational burden. A unique aspect of our contribution to the existing permafrost datasets is thus global coverage at a moderately high spatial resolution."

a.2 Regarding model performance in simulated permafrost extent (particularly the deficiency in western Siberia) (will be added into section 4.4)

"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."

a.3 Regarding model performance in simulated ALT (will be added into section 4.4)

"The existing literature on simulated ALT fields (e.g., Dankers et al. (2011), Lawrence et al. (2012) and Guo et al. (2017)) reveals a general tendency for models to overestimate ALT climatology at the global scale. The CLSM-simulated ALT fields appear to be among the better simulation products."

b. Some improvements of MERRA-2 compared to MERRA (will be added into section 1):

"MERRA-2 has been found to be skillful in its simulation of near-surface atmospheric conditions (Reichle et al., 2017a;Reichle et al., 2017b;Bosilovich et al., 2015;Bosilovich et al., 2017) and to show improvements in the representation of cryospheric processes compared with its predecessor MERRA (Gelaro et al., 2017). In particular, MERRA-2 assimilates substantially more satellite observations and employs more physically reasonable hydrology representations for glaciated land surfaces compared to MERRA, and it also uses observation-based, seasonally-varying sea ice albedo as opposed to MERRA's fixed value of 0.6 (Gelaro et al., 2017). A recent study shows that MERRA-driven permafrost and ALT simulation results are inferior to those driven by other reanalysis-based forcing data sets, particularly those from the NOAA Climate Forecast System Reanalysis (CFSR) and European Centre for Medium-Range Weather Forecasts Re-Analysis Interim (ERA-I) (Guo et al., 2017). We note that our MERRA-2-driven permafrost simulation results, while potentially better than those we might have obtained with MERRA forcing, are still lacking (e.g., in western Siberia). The potential superiority of MERRA-2 forcing compared to MERRA forcing in the context of permafrost simulation remains unknown."

c. Summary of novelty and added value of this work (will be added into section 5):

"The permafrost dataset presented herein can be considered unique in terms of its daily temporal resolution combined with a relatively high spatial resolution at the global scale (i.e., 81 $km^2$). The dataset, which is derived from a state-of-the-art reanalysis, shows reasonable skill in capturing permafrost extent and in adequately estimating ALT climatology (aside from that at the Mongolian sites). With its resolution and available variables (ALT, subsurface temperature at different depths), it could prove valuable to many future permafrost analyses."

"This work also provides a first comparison between two highly complementary approaches to estimating permafrost: model simulation and remote sensing. The remote sensing approach is still relatively new, and many aspects still need to be worked out. It is important, though, to begin considering the modeling and remote sensing approaches side by side, as both should play important roles in permafrost quantification in the years to come. Indeed, once the science fully develops, joint use of modeling and remote sensing (e.g., through the application of downscaling

> methods) should allow the generation of more accurate permafrost products at even higher resolutions."

- Some words about potential uses of this dataset could be nice.

R1C2: We thank the reviewer for this suggestion. Relevant discussion about the potential uses of this dataset will be added to the end of this manuscript:

> "For example, these ALT estimates are highly relevant to the assessment of the regional water budget and can be helpful for monitoring groundwater changes at a wide variety of scales (Evans et al., 2018;Evans et al., 2015). In addition, these data can potentially contribute to ecological studies focused on the dynamics of microbial activity and soil respiration in cold regions, on vegetation migration/adaptation in response to climate change, and so on."

Specific points:

- Page 2, line 15: "simulations… with the land surface model ( Dankers et al., Guimberteau et al., Tao et al.)" -> these are different models. The sentence is misleading, and confusingly, you write "…and other numerical models" afterwards…

R1C3: We will revise the relevant sentences as follows.

> "Simulations and/or predictions with a variety of land surface models (LSMs) have been used to quantify large-scale permafrost patterns (i.e., distributions and thermal states) and their interactions with a warming climate. LSMs utilized for this include, for example, the Joint UK Land Environment Simulator (JULES, Dankers et al., 2011), the ORganizing Carbon and Hydrology in Dynamic EcosystEms (ORCHIDEE) - aMeliorated Interactions between Carbon and Temperature (ORCHIDEE-MICT, Guimberteau et al., 2018), the Catchment Land Surface Model (CLSM, Tao et al., 2017), and the Community Land Model (Lawrence and Slater, 2005;Alexeev et al., 2007;Nicolsky et al., 2007a;Yi et al., 2007;Lawrence and Slater, 2008;Lawrence et al., 2008;Lawrence et al., 2012;Koven et al., 2013;Chadburn et al., 2017;Guo and Wang, 2017)."

- Page 2, line 20-24: Strictly speaking, the fact that 2017 set records doesn't mean that permafrost conditions will change. 2017 is only one year. It's the long-term trends that matter (2017 is of course consistent with that trend)

R1C4: We agree with the Reviewer regarding this point. The message we had tried to convey is that the warming trend of our climate seems to have increased in recent years and that given the associated exacerbation in permafrost thawing, monitoring permafrost in a timely manner is critical. We will modify the relevant sentences as follows. Please also see our response to next comment (R1C5).

> "In addition, given the apparent climate warming seen in recent years (exemplified by the fact that the average Arctic air temperature in 2017 (ending in September) was the second warmest on record since 1900 [Arctic Report Card; http://www.arctic.noaa.gov/Report-Card/Report-Card-2017] and that 2017 was the warmest year on record for global ocean temperatures (Cheng and Zhu, 2018), important reductions in permafrost might be occurring as well."

- Page 2, line 23: "Some aspects of the current global permafrost thermal states are ... still unknown": can you elaborate on that, please?

R1C5: We will clarify this by expanding this sentence as shown below:

> "However, current global permafrost thermal states (i.e., permafrost temperature, ice content and degradation rates across much of Northern latitudes) are arguably still unknown. Monitoring permafrost degradation in a timely manner is particularly critical for ecosystem management and for various policy decisions."

- Page 3, line 10: "extensive challenges" sounds bizarre to my non-native speaker's ears

R1C6: We will replace the end of the sentence with "combined with the many challenges of physical process modelling."

- Page 3, line 16: Could you say a few words specifically about high-latitude performance? Advantages, drawbacks // other reanalyses?

R1C7: We will add some relevant discussion about the performance of MERRA-2 in high-latitude regions:

> "In particular, MERRA-2 assimilates substantially more satellite observations and employs more physically reasonable hydrology representations for glaciated land surfaces compared to MERRA, and it also uses observation-based, seasonally-varying sea ice albedo as opposed to MERRA's fixed value of 0.6 (Gelaro et al., 2017)."

- Page 3, line 26: Chen et al. is a paper in review. Can you reassure the reviewer that these retrievals are independent from the data produced here? One or two sentences would be nice anyway even if Chen et al. 2018 will be available to the reader soon.

R1C8: The AirMOSS radar retrievals of ALT we used here were produced by the algorithm described and analyzed in detail by Chen et al. 2018. The retrievals here are indeed identical to those produced by Chen et al. 2018, though here we examine them from a different perspective. We will add some text to distinguish the scope of this study from that of Chen et al. 2018 and to emphasize the data independence:

> "Although based on the same set of ALT retrievals, Chen et al. (2018) mainly focus on the development and improvement of the retrieval algorithm, whereas this study emphasizes using the remotely sensed ALT estimates to characterize and assess the

> spatial variability of the modelling results. The ALT retrievals and the modelling results are fully independent."

- Page 5, line 13: because you later speak about the spinup in the trend analysis, it might be interesting to say a few words about this here. The looping through the 36 years cannot given the same soil temperatures as you would normally have if you have realistic spinup data.

R1C9: We thank the reviewer for the suggestion. We will add two sentences here.

> "One caveat about this looping procedure, by the way, is worth mentioning. Because it makes use of the warmer conditions of the last few decades, it might produce a warmer initial condition than a set of initial conditions produced with realistic historical forcing over hundreds of years (e.g., Sapriza-Azuri et al., 2018). This will affect our trend analysis, as discussed in section 4.5."

- Page 8, line 15 and following: The assessment might be a bit too optimistic here: Basically one sees that the ALT is between 0.2 and 1m both in obs and simulations, not much more. Is there a significant correlation at all?

R1C10: We will modify this sentence as shown below.

> "Figure 4b, c demonstrates that in some ways, the CLSM-simulated results roughly agree, to first order, with the in-situ observations. The overall mean bias of simulated ALT relative to the in-situ measurements is -0.05 m. Nevertheless, the scatter (blue) in Fig. 4c is large, and the corresponding correlation coefficient is quite weak (0.27)."

- Page 8, line 21 and following: The AirMOSS ALT retrievals. Bascically the retrievals are the same everywhere! Around 0.45 m. No variability. Are they actually of any use?

R1C11: As we mentioned in the original manuscript, relatively larger variability with the ALT retrievals is seen at its native resolution (Figure 3a), but this larger variability was smoothed out through aggregation to the model scale at 81 km$^2$, as we expected. In addition, as also mentioned in the manuscript, these retrievals cannot exceed the P-band radar sensing depth of about 60cm, and thus for the shallow permafrost here, the averaged ALT retrievals appears to be around 0.45 m everywhere.

We emphasize that this is a first attempt to compare remote sensing ALT data with modeling results. An expected future direction is to take advantage of the detailed heterogeneity information in the remote sensing data to downscale model results directly or to improve modeling skill indirectly. We will add several sentences about the potential use of these ALT retrievals into the discussion section (see our response to R1C1, part c, above). Please also see our response to the next comment (R1C12).

- Page 8, line 32 and following ("Excluding: : :"): Yes, OK, but then there is still no correlation. Values are just around 0.45 m and the mean ALT of the remaining sites just happens to be around that value.

R1C12: With further analysis, we find that the correlation coefficient between the ALT retrievals and the in-situ observations, while small at the site scale, is larger than that for modeled ALT at the model scale, both for all sites and for sites with measured ALT below 60cm. We will discuss these findings and include the new Table 3 in the revised manuscript:

> "For the AirMOSS retrievals, the overall ALT bias is -0.11 m at the site scale and -0.12 m at the model scale. While the correlation coefficient with the in-situ observations is only 0.05 at the site scale, it is 0.61 at the model scale. "
>
> "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."

*Table R1– Evaluation metrics for model-simulated ALT and AirMOSS retrievals for 2015. (New Table 3)*

| Metric | All sites | | | Sites with ALT measurements < 60 cm | | |
|---|---|---|---|---|---|---|
| | CLSM-simulated ALT (model scale) | AirMOSS ALT retrievals (model scale) | AirMOSS ALT retrievals (site scale) | CLSM-Simulated ALT (model scale) | AirMOSS ALT retrievals (model scale) | AirMOSS ALT retrievals (site scale) |
| RMSE (m) | 0.17 | 0.17 | 0.21 | 0.12 | 0.06 | 0.08 |
| Bias (m) | -0.05 | -0.12 | -0.11 | 0.01 | -0.01 | 0.02 |

| | | | | | | |
|---|---|---|---|---|---|---|
| R | 0.27 | 0.61 | 0.05 | -0.00 | 0.64 | 0.20 |

- Page 9, line 9-10: “Further investigation: : :”: You could nevertheless elaborate a little bit on this. Are there any common characteristics of sites with thick active layer (dry soil, highly conducive soil, southward sloping, etc.) that the model doesn’t get?

R1C13: Here we specifically meant some investigation on the zero-curtain problem with the model. Nevertheless, we will amend the text to read:

> “The use of the $0^{o}$C degree threshold in CLSM for determining the thawed or frozen state of the soil may explain in part the model’s underestimation of ALT, as may the lack of an explicit treatment of local aspect, errors in assigned model parameters, and so on.”

(Note that we will move this paragraph to the discussion section, as suggested by the Reviewer #3.)

- Page 11, line 1-3 – meteorological forcing dominant control: Of course. Could anyone seriously expect something different?

R1C14: Yes. We will add “as expected” into this sentence.

- Figure 7: Good that this is quantified in such a way here. Much more synthetic &interesting than figure 6.

R1C15: Thank you. We agree.

- Page 13, line 12: Correlation might increase if time steps with snow on the ground but air temp > 0_C are not counted in Tcum - it’s the soil surface temperature that counts, not the air temperature.

R1C16: The reviewer makes a reasonable suggestion, and we thus recomputed the correlation map using this modified rule. We found that the map did not differ very much from the original one. We will keep the original figure as is.

- Page 14, line 5 : Problems in mountain areas: Snow forcing might be severely in error in these regions

R1C17: Yes. We will add one sentence here.

> “In addition, MERRA-2 snow forcing might be severely erroneous in these regions.”

- Page 14, line 12: “The reasons: : :” – I have probably missed the information: How deep is the model soil column?

R1C18: As mentioned in section 2.1, the depth ranges of the six soil layers are 0-0.1m, 0.1-0.3m, 0.3-0.7m, 0.7-1.4m, 1.4-3m, and 3-13m, respectively. We checked the CALM sites over western Siberia and found that the ALT observations there are basically below 2 m. Therefore, the depth of the model's soil column is not an issue, and we will delete this speculation from the sentence. The new sentence below will instead be added.

> "The reasons for this particular deficiency are unclear; perhaps the initial thermal conditions over western Siberia were too warm, or perhaps MERRA-2 overestimates current air temperatures in this region."

- Page 16, line 15, Mongolian ALT trends: How can you have a 25 cm/year trend over 17 years? That would mean that ALT increases by over 4 m over that period. That's quite improbable. These data are very suspicious.

R1C19: We share the reviewer's concerns about the quality of these data. Below in Table R2 we provide the time series of the actual ALT measurements at the three Mongolian sites M1, M1 A and M3 (https://www2.gwu.edu/~calm/data/north.html). Because M1 and M1 A are within the same CLSM modeling grid cell, we used the average of their two time series. Time series of observed and simulated ALT at the two grid cells containing these three sites are plotted in Figure R1. The observed and simulated ALT trends at the two grid cells correspond to the two dots showing the extraordinarily large observed trends in Figure 14a in the original manuscript. Note the simulated ALT trends were calculated using ALT estimates only in years when observed ALTs are available, as also mentioned in the manuscript.

*Table R2 – Observed ALT (cm) at three Mongolian sites.*

| Site Code | 1996 | 1997 | 1998 | 1999 | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| M 1 | 345 | 350 | 355 | 345 | 350 | 340 | 355 | - | 375 | 365 | 380 | 380 | 370 | 350 |
| M1 A | 390 | 390 | - | 430 | 485 | 400 | 450 | 600 | 770 | 710 | 820 | 820 | - | - |
| M3 | - | 610 | 620 | 600 | 610 | 660 | 720 | 760 | 770 | 750 | 760 | 770 | 800 | - |

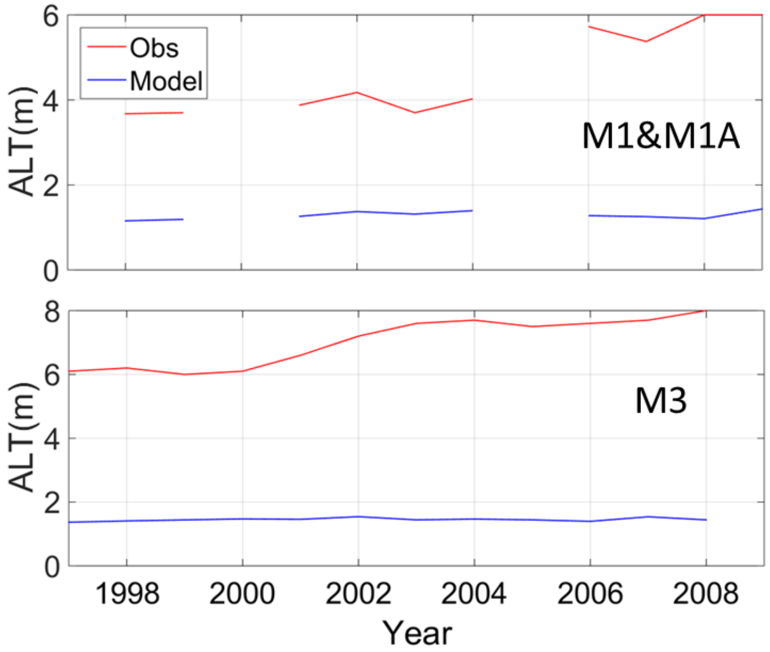

*Figure R1: Time series of observed and simulated ALT at Mongolian sites collocated with two simulation grid cells, i.e., M1&M1A (upper) and M3 (bottom). The calculated ALT trends from observations and simulation are 24.38 cm/yr and 1.13 cm/yr, respectively, for the grid cell containing M1&M1A and are 19.69 cm/yr and 0.51 cm/yr, respectively, for the grid cell containing M3.*

We attempted to contact the PI responsible for these data (Dr. Natsagdorj Sharkhuu from the Institute of Geography, Mongolian Academy of Sciences); however, the email address (provided here https://www2.gwu.edu/~calm/data/webforms/mg_f.html) is apparently obsolete, and the email delivery failed. We were thus unable to investigate further the data quality.

In any case, as indicated by the reviewer, this issue calls at the very least for a specific caveat about these data, which we will add:

> "A particular caveat is required regarding the data from the Mongolian sites, given the unusual observed trends there. Attempts to contact the data providers to attain more detailed information for data evaluation were unsuccessful, and accordingly our confidence in these particular data is limited."

- Page 18, line 20 : Â´n ..addition of soil layers Â˙z : Would that be so difficult? Tests with more levels would really be interesting, but if they are really difficult, I refrain from asking for such test to be carried out.

R1C20: In response to this comment we have conducted such tests. Our general conclusions are consistent with other studies in terms of how soil configuration affects permafrost modeling (e.g., Alexeev et al., 2007;Lawrence et al., 2008;Sapriza-Azuri et al., 2018;Nicolsky et al., 2007b;Dankers et al., 2011).

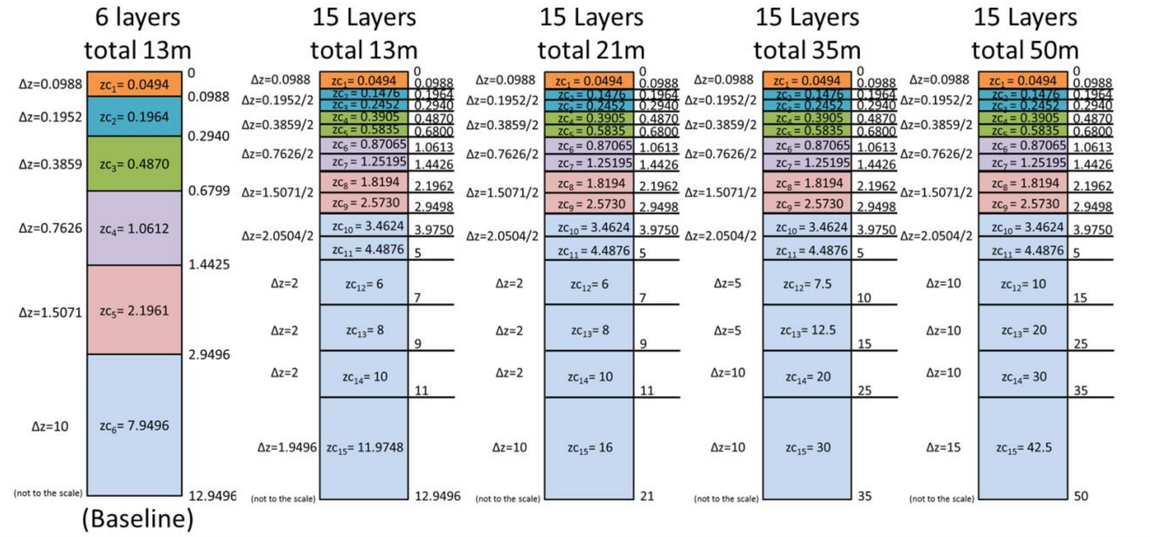

*Figure R2: Soil configurations we have newly tested.*

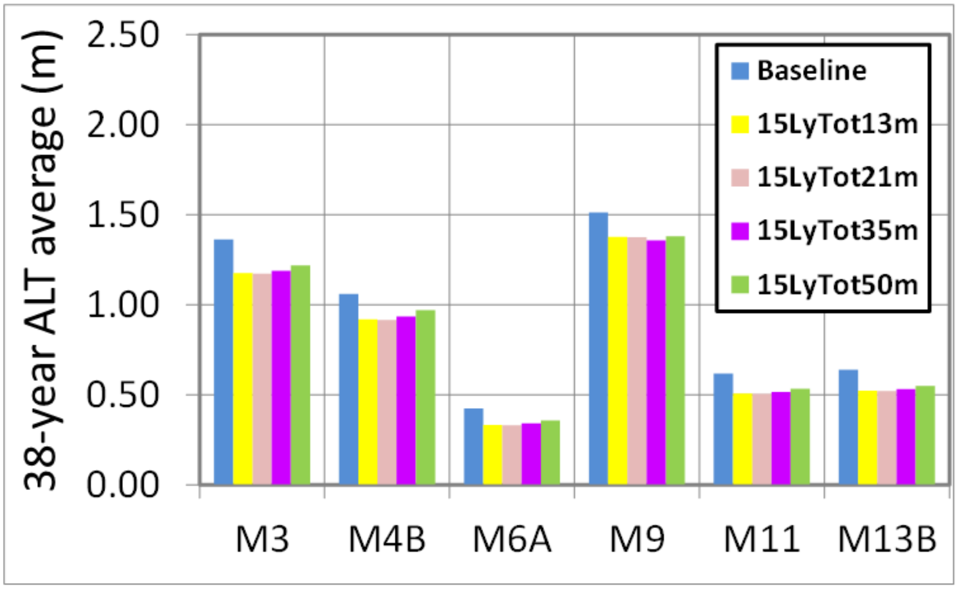

*Figure R3: Simulated ALT results at six Mongolian sites with 5 different soil configurations. (Note the baseline soil configuration contains 6 soil layers and a total soil depth about 13 m.)*

Specifically, we tested four new soil configurations with 15 soil layers and different soil depths (ranging from 13 m to 50 m). Figure R3 reveals that increasing the number of soil layers decreases ALT climatology at the tested sites, which is consistent with previous studies (e.g., Alexeev et al., 2007;Lawrence et al., 2008;Dankers et al., 2011). The figure also demonstrates that variations in total soil depth have only a small impact on the simulation of ALT, as reported in the previous studies. However, the zero flux we employ at our lower boundary for all the simulations might influence heat transfer in deep soils and thus might decrease the impact of using deeper soils.

Based on these tests, we will include some appropriate discussion in the text, though without adding a figure:

> "Local test simulations (not shown) with alternative model configurations indicate that increasing the number of soil layers may act to decrease somewhat the simulated ALT, suggesting that our values may be a little overestimated; however, based on results from a new study by Sapriza-Azuri et al.(2018), our use of a no-heat-flux condition at the bottom boundary rather than a dynamic geothermal flux may lead to underestimates of ALT. Such uncertainties should naturally be kept in mind when interpreting our results. Our supplemental simulations also suggest that increasing the total modeled soil depth has only a small impact on simulated ALT."

We will also explicitly mention our lower boundary condition in section 2.1:

> "A no-heat-flux condition is employed at the bottom of the model's soil column."

# 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.