The Cryosphere Discussions

The Cryosphere Discuss., referee comment RC2
https://doi.org/10.5194/tc-2022-5-RC2, 2022


# Comment on tc-2022-5

Anonymous Referee #2

---

Referee comment on "Validation of pan-Arctic soil temperatures in modern reanalysis and data assimilation systems" by Tyler C. Herrington et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2022-5-RC2, 2022

---

## General comments

The authors perform a comprehensive validation of current reanalysis and LDAS model products using a set of in situ soil temperature measurements across the northern hemisphere. Validation of soil temperatures in large-scale reanalysis data products is an important task given their frequent use in global ecological and hydrology models and permafrost simulations, for example. It is crucial that the underlying biases are clearly acknowledged by the users of such products. A novelty of this study is that it validates multiple reanalysis soil temperature products at the same time and covers the entire pan-Arctic area.

The authors find that most reanalysis products are considerably cold-biased and that the biases and inter-model variability are larger during the cold season compared to the warm season. The authors also compile an ensemble from the assessed reanalysis data and show how it interestingly overperforms most/all individual models. The authors list potential future applications of the ensemble mean product, but I would wish to see a bit more discussion on its current usability, given that the recorded biases remain quite high and display some regional patterns. The underlying reasons for these are addressed in the manuscript but not how the biases would affect, e.g., permafrost simulations where a bias or RMSE of above 2° C can have notable implications.

The article is very well written and figures are great. At places the text is hard ro follow (especially Section 4.3, see detailed comments below) owing to the multiple simultaneous comparisons: near surface vs. at depth soil temperatures, cold season vs. warm season, permafrost vs. no to little permafrost, North America vs. Eurasia, and DJF vs. JJA. I suggest the authors to make sure all sections are clearly defined.

I recommend the publication of this manuscript after the authors have considered my

minor suggestions and comments.

**Specific comments**

L104: The authors suggest that their study is *"To the authors' knowledge, this one of the first studies to compile a comprehensive set of in situ soil temperature measurements across the Eurasian and North American Arctic, from multiple diverse sparse networks".* While it may be true that this is true for the "one of the first" part, it should be noted that the compilation is not totally novel, given that similar in situ temperature datasets have been compiled not only by Cao et al. (2020, in the references) but also, e.g., by Karjalainen et al. (2019) and Ran et al. (2022) who used mostly the same data sources, albeit computing temperatures averages for a much larger depth (several meters deep in permafrost but also in non-permafrost soils). Moreover, Lembrechts et al. (2020) have published a global soil temperature compilation of soil and near-surface temperatures. I suggest the authors to consider if their statement needs some elaboration, e.g., does the compiled dataset differ from previous datasets in some ways.

The authors recognize the notably different sampling size for North America but retain from explaining why no more data were collected, apart from mentioning the overall data scarcity in northern Canada, to correct the imbalance between North America and Eurasia. Based on the previous data compilations (see above), there should be suitable measurement time series available from North America. Notwithstanding, the authors satisfactorily show how the sampling imbalance does not affect the fundamental conclusions (LL378-391).

LL140-141: *"Panel B of Figure 1 shows the spatial standard deviation of monthly surface soil temperatures for grid cells with more than two stations included."* However, in Figure 1b, grid cells with two stations are also shown. Also, I remain unsure whether there are any grid cells with >1 stations in Eurasia?

L236: Reference should be to Fig. S1, right?

LL 236-238: *"The mean bias and RMSE are typically 1â□¦ C to 3â□¦ C smaller over North America, relative to the permafrost zone in Eurasia (see Figure S3); however with fewer grid cells over North America, the uncertainty is also larger - as evidenced by the larger error bars."* I cannot see the mentioned difference between North America and Eurasia in Figure S3 (biases in ERA5-Land) or in the associated bar graph (ERA5-Land) in Figure S2. However, all products considered the said difference between regions is visible. Consider checking and editing the text so that it corresponds to the results shown in Figure S3.

L239: What correlations, the ones between measurements and reanalysis temperatures? A slight elaboration would help the reader to see that what are compared in the sentence.

LL240-241: I also struggled with this sentence. What is the opposite situation here? It is hard to follow the comparisons between permafrost and little to no permafrost, as well as near-surface and at depth temperatures at the same time, especially since the results are not shown.

LL243-246: Are these results related to the permafrost binning? It's fine if they are not, but overall Section 4.3 is at times hard to follow because it deals with both permafrost binning and regional comparisons.

L405: Instead of the cold season standard deviations, should you not refer here to cold stations/observations? That is, figure 6 does not distinguish between warm and cold season.

L261: The ensemble mean product is not properly addressed until deep into the results (validation) section. I suggest presenting the ensemble mean product and its calculation procedure already in the early stages (possibly inside section 2.1.).

L303: I find "coastal regions" not the ideal term here because the regions with the highest variability are far more than that. In winter, greatest variation associates with the coldest regions, yet not exclusively either. Could the variation here be related to snow cover duration / snow thickness as mentioned elsewhere in the text?

**Technical corrections**

L61: Please, open the abbreviation GLDAS-CLSM already here.

LL80-83: Check grammar of the sentence. Maybe delete the word "that" at line 81?

L191: Figure 2 does not have panels C and D.

Figure 3: This is a nice figure with lots of information in it. The letters in “Correlation coefficient” are clumped together and could be corrected.

Figure 4: Stratification of the values in histograms is not explained. Please add it to the caption.

Figure 5: Y-axis is a bit messy. Consider adjusting the interval at which temperatures are denoted.

Figure 8: DJF missing from Panel A label.

L286: NH à northern hemisphere

L290: Why are ensemble mean at depth temperatures not shown? Could be part of the supplement. Figure 9 also shows at depth results, so it would be interesting to see how the models reconstruct frozen ground in JJA, although it is acknowledged that this is not explicitly representative of permafrost.

L366: Please put Gruber et al. 2018 inside parentheses.

L369-370: *"Moreover, the impact of snow cover on soil temperature is generally more pronounced over permafrost regions (regions of seasonal frost).*" Is something missing here? Should it be "compared to regions of seasonal frost" or what is the idea?

LL418-419: Could you elaborate, what does it mean *"is being explored"?*

L428: Please provide a url for the ensemble mean dataset on the ADC.

L583: Database title and url missing.

**References**

Karjalainen et al. 2019 Circumpolar permafrost maps and geohazard indices for near-future infrastructure risk assessments. Scientific Data https://doi.org/10.1038/sdata.2019.37

Lembrechts et al. 2020 SoilTemp: A global database of near-surface temperature. Global Change Biology https://doi.org/10.1111/gcb.15123.

Ran et al. 2022 New high-resolution estimates of the permafrost thermal state and hydrothermal conditions over the Northern Hemisphere. Earth System Science Data https://doi.org/10.5194/essd-14-865-2022