Summary

We thank both referees for their helpful comments and we have incorporated their feedback in the revisions. The manuscript has undergone a substantial revision and has incorporated a couple of new sections, as well as some changes to the methodology. These include:

1. Removing GLDAS-CLSM and GLDAS-Noah, as during our revision process, we came to realize that the discontinuities related to GLDAS-Noah and GLDAS-CLSM were related to differences in the driving meteorology between Version 2.0 and Version 2.1 of the products. These have been replaced with FLDAS, which has more a consistent forcing meteorology.

2. Remapping all products to the ERA-Interim resolution (0.75°) instead of the GLDAS resolution (1°).

3. The addition of a new section on soil temperature trends for the individual products (Section 4.4 – Multi-Annual Trends). We have also split our results on the variability of seasonal extremes into a section focused on the individual products (Section 4.5), and a section focused on the ensemble mean (Section 5.3).

4. An exploration of elevation impacts on soil temperature performance, described in the methods (Section 3.3) and Section 4.3 – Spatial Variability of the results.

5. Removing JRA55 from the ensemble mean soil temperature calculation, as its inclusion was found to dramatically increase the bias and RMSE of the ensemble mean.

Reviewer 1

P2, L43: Cao et al., 2022 presented improved ERA5-Land soil temperature in permafrost regions using an optimized multi-layer snow scheme. Please consider citing the reference.

We thank the referee for this helpful suggestion and have included Cao et al. (2022) as a reference pertaining to snow thermal insulation in ERA5-Land.

P9, L205: Please use "discontinuous permafrost" rather than "extensive discontinuous permafrost". See Zhang et al., 2000.

Corrected

P21, L353: In latex, $^\circ$C rather than $^\circ$~C

Corrected

P27, L445: Is it possible to show the change time of data assimilation for each product? This could be added to Table 1.

During our revision process, we came to realize that the discontinuities related to GLDAS-Noah and GLDAS-CLSM were related to differences in the driving meteorology between Version 2.0 and Version 2.1 of the products. In the case of CFSR, we diagnosed the changes in soil temperature variability to be mainly related to snow cover issues discussed in Section 4.4 – Multi-Annual trends. Thus the discussion about changes in upper-atmosphere assimilation are no longer relevant to the trends discussion, or to Section 6.2 in the discussion.

Reviewer 2

1. The manuscript is too strongly focused on establishing the ensemble mean product. To me, the novel science of the study is to rigorously document and compare the performance of the individual products.
2. In the main paper, I suggest presenting these six numbers for the best-performing product in each category, the most common (e.g. ERA-5) and the overall best-performing products that some users could actually be encouraged to use, and the ensemble mean product.
3. With these numbers, users will have a good basis to decide which product may be most applicable for their application.
4. L. 298: there is a critical difference between "most products" and "all products", so the authors need to add more information on this throughout the entire section. If only most products are worse, then at least one will be similarly good or better, so users can be directed to use this product.
5. For other metrics, single products seem to do better, Fig. 2 for example suggests that ERA-5 with respect to the bias is significantly better than the ensemble mean.

The authors have interpreted these five comments as relating to two main questions:

1. Why did the authors choose to focus the paper on the Ensemble Mean product, rather than a detailed characterization of the performance of individual soil temperature products from reanalysis and land data assimilation system (LDAS)? And

2. Why do the authors not recommend any one individual product as the best choice for estimating soil temperatures in the extratropical northern hemisphere, and the Arctic?

In the following paragraphs, we explain our answers to these two points in detail.

It was clear from previous studies and from our validation of soil temperatures over the extratropical northern hemisphere, that no single reanalysis product provided adequate performance over all regions and times of year (see Figure 2, skill scores). The noticeable decline in performance over the cold season, and over higher latitudes (Figure S1) is present in all products, as evidenced by the substantially lower skill scores relative to the warm season. This made it impractical to recommend a single product and led to the exploration of blending the suite of products together. Blending of multiple observation-based data products is becoming common practice in many subfields of climate science and has been demonstrated to reduce random errors and improve overall performance relative to any individual product (e.g. Mudryk et al. 2015; Dorigo et al., 2017; Gruber et al., 2019; Beck et al., 2019; Cao et al., 2019). This motivated the development of the ensemble mean product and, therefore, a detailed investigation into the performance of the ensemble mean product was required and became a major focus of the paper.

The primary metric used in the evaluation is the skill score developed by Taylor et al. (2001), which is preferred over singular metrics (such as bias or RMSE), as it incorporates both RMSE as well as aspects of soil temperature variability (normalized standard deviation), similar to the Taylor Diagram. The ensemble mean product showed substantially higher skill than all other products over the cold season (Figure 2), and better performance over higher latitudes (Figure S1); particularly for deeper soil layers. Singular metrics such as bias can be misleading over the annual mean, owing to offsetting errors. For example, the small overall bias of ERA5 and ERA5-Land in Figure 2 are due to offsetting errors over the study region, where soils are too warm over much of the permafrost region, and too cold over more southern regions (See Figure S3); similar to the findings of Cao et al. (2020).

I am missing a clear overview (e.g. a Table) over the key numbers, which for me are: (a) the bias and (b) the RMSE for (1) the entire time period, (2) annual mean values (considering the entire time period), and (3) monthly mean values (considering the entire time period). In the main paper, I suggest presenting these six numbers for the best-performing product in each category, the most common (e.g. ERA-5) and the overall best-performing products that some users could actually be encouraged to use, and the ensemble mean product. Furthermore, for future reference, the authors should provide these numbers for all considered products in the supplement.

Thank you for this suggestion. We have included a table with the annual mean, cold season and warm season bias, RMSE and skill score for each product here (Tables R1 and R2), and in the supplement (Table S1 and S2) and make reference to this table in Section 4.1 (Extratropical Northern Hemisphere Mean) of the revised manuscript. We decided against showing monthly mean values because over such a short time period there would be too much variability between stations to make the averages meaningful. This is due to the very large range of latitude and continentality between stations and the associated differences in climate and the duration/extent of snow cover.

Table R1. Summary of the near surface mean bias, RMSE, and skill score for each product for the Extratropical Northern Hemisphere. Metrics are separated into an annual mean metric, and a metric for the cold and warm seasons. The best performing product is listed in bold for each metric, and season. The 95% confidence interval is also included. Multiple bold values may appear if the confidence interval of multiple products overlap for a particular metric and season.

Metric	Product	Annual	Cold Season	Warm Season
Bias				
	CFSR	-2.63 ± 0.13° C	-2.37 ± 0.19° C	-2.8 ± 0.19° C
	ERA-Interim	-2.83 ± 0.21° C	-4.92 ± 0.34° C	**-0.69 ± 0.34° C**
	ERA5	-0.73 ± 0.14° C	**-0.91 ± 0.24° C**	**-0.58 ± 0.24° C**
	ERA5-Land	**0.29 ± 0.14° C**	**0.89 ± 0.22° C**	**-0.35 ± 0.22° C**
	FLDAS	-2.62 ± 0.14° C	-3.42 ± 0.26° C	-2.18 ± 0.26° C
	JRA55	-2.15 ± 0.17° C	4.16 ± 0.31° C	-7.04 ± 0.31° C
	MERRA2	-1.55 ± 0.12° C	-1.88 ± 0.21° C	-1.33 ± 0.21° C
	Ensemble Mean	-1.68 ± 0.12° C	-2.1 ± 0.20° C	-1.32 ± 0.20° C
RMSE				
	CFSR	3.89 ± 0.14° C	3.74 ± 0.15° C	3.65 ± 0.15° C
	ERA-Interim	4.55 ± 0.26° C	5.88 ± 0.33° C	**2.17 ± 0.33° C**
	ERA5	**2.75 ± 0.15° C**	3.14 ± 0.18° C	**2.15 ± 0.18° C**
	ERA5-Land	**2.51 ± 0.17° C**	**2.67 ± 0.21° C**	**2.13 ± 0.21° C**
	FLDAS	4.01 ± 0.13° C	4.60 ± 0.17° C	3.12 ± 0.17° C
	JRA55	7.68 ± 0.17° C	5.87 ± 0.21° C	8.95 ± 0.21° C
	MERRA2	2.90 ± 0.13° C	3.30 ± 0.15° C	**2.33 ± 0.15° C**
	Ensemble Mean	2.85 ± 0.13° C	3.17 ± 0.14° C	**2.31 ± 0.14° C**
Pearson Correlation				
	CFSR	0.959 ± 0.003	0.655 ± 0.018	0.931 ± 0.018
	ERA-Interim	0.971 ± 0.002	0.731 ± 0.014	0.905 ± 0.014
	ERA5	0.970 ± 0.003	0.564 ± 0.021	0.950 ± 0.021
	ERA5-Land	0.970 ± 0.003	0.644 ± 0.015	0.954 ± 0.015
	FLDAS	0.970 ± 0.002	0.734 ± 0.014	0.939 ± 0.014
	JRA55	0.552 ± 0.008	0.570 ± 0.015	0.402 ± 0.015
	MERRA2	**0.978 ± 0.002**	0.674 ± 0.015	0.954 ± 0.015
	Ensemble Mean	**0.982 ± 0.002**	**0.793 ± 0.011**	**0.957 ± 0.011**
Skill Score				
	CFSR	0.946 ± 0.007	0.689 ± 0.019	0.902 ± 0.019
	ERA-Interim	0.896 ± 0.012	0.62 ± 0.022	0.955 ± 0.022
	ERA5	0.951 ± 0.008	0.685 ± 0.017	0.955 ± 0.017
	ERA5-Land	0.949 ± 0.009	0.601 ± 0.022	0.95 ± 0.022
	FLDAS	0.935 ± 0.008	0.691 ± 0.017	0.938 ± 0.017
	JRA55	0.536 ± 0.012	0.564 ± 0.017	0.647 ± 0.017
	MERRA2	0.954 ± 0.008	0.73 ± 0.016	**0.960 ± 0.016**
	Ensemble Mean	**0.958 ± 0.007**	**0.805 ± 0.014**	**0.961 ± 0.014**

Table R2. Summary of the near surface mean bias, RMSE, and skill score for each product for the Extratropical Northern Hemisphere. Metrics are separated into an annual mean metric, and a metric for the cold and warm seasons. The best performing product is listed in bold for each metric, and season. The 95% confidence interval is also included. Multiple bold values may appear if the confidence interval of multiple products overlap for a particular metric and season.

Metric	Product	Annual	Cold Season	Warm Season
Bias				
	CFSR	-3.12 ± 0.16° C	-1.63 ± 0.19° C	-4.48 ± 0.19° C
	ERA-Interim	-2.71 ± 0.22° C	-2.26 ± 0.29° C	-2.99 ± 0.29° C
	ERA5	-0.54 ± 0.16° C	**0.27 ± 0.20° C**	**-1.34 ± 0.20° C**
	ERA5-Land	**0.10 ± 0.17° C**	1.33 ± 0.21° C	**-1.10 ± 0.21° C**
	FLDAS	-3.09 ± 0.19° C	-2.72 ± 0.26° C	-3.71 ± 0.26° C
	JRA55	-1.82 ± 0.19° C	**0.49 ± 0.27° C**	-3.67 ± 0.27° C
	MERRA2	-2.08 ± 0.15° C	-0.70 ± 0.19° C	-3.30 ± 0.19° C
	Ensemble Mean	-1.91 ± 0.15° C	-0.95 ± 0.18° C	-2.82 ± 0.18° C
RMSE				
	CFSR	4.54 ± 0.16° C	2.98 ± 0.16° C	5.28 ± 0.16° C
	ERA-Interim	4.05 ± 0.22° C	3.42 ± 0.27° C	4.13 ± 0.27° C
	ERA5	**2.73 ± 0.15° C**	**2.10 ± 0.17° C**	**2.98 ± 0.17° C**
	ERA5-Land	**2.76 ± 0.17° C**	**2.28 ± 0.20° C**	**2.82 ± 0.20° C**
	FLDAS	4.68 ± 0.15° C	3.99 ± 0.18° C	4.83 ± 0.18° C
	JRA55	4.24 ± 0.16° C	3.26 ± 0.14° C	4.86 ± 0.14° C
	MERRA2	3.44 ± 0.16° C	**2.17 ± 0.15° C**	4.00 ± 0.15° C
	Ensemble Mean	3.19 ± 0.14° C	**2.20 ± 0.13° C**	3.69 ± 0.13° C
Pearson Correlation				
	CFSR	0.817 ± 0.012	0.77 ± 0.018	0.749 ± 0.018
	ERA-Interim	0.909 ± 0.007	0.857 ± 0.012	0.847 ± 0.012
	ERA5	0.906 ± 0.01	0.849 ± 0.014	0.871 ± 0.014
	ERA5-Land	0.901 ± 0.01	0.86 ± 0.011	**0.881 ± 0.011**
	FLDAS	0.857 ± 0.013	0.84 ± 0.014	0.833 ± 0.014
	JRA55	0.781 ± 0.009	0.865 ± 0.011	0.753 ± 0.011
	MERRA2	0.91 ± 0.008	0.856 ± 0.013	0.869 ± 0.013
	Ensemble Mean	**0.922 ± 0.007**	**0.892 ± 0.011**	**0.871 ± 0.011**
Skill Score				
	CFSR	0.792 ± 0.019	0.768 ± 0.019	0.768 ± 0.019
	ERA-Interim	0.852 ± 0.019	0.751 ± 0.019	0.794 ± 0.019
	ERA5	**0.872 ± 0.017**	0.825 ± 0.018	0.843 ± 0.018
	ERA5-Land	0.839 ± 0.02	0.781 ± 0.022	**0.864 ± 0.022**
	FLDAS	0.759 ± 0.022	0.697 ± 0.02	0.782 ± 0.02
	JRA55	0.794 ± 0.016	0.685 ± 0.018	0.762 ± 0.018
	MERRA2	0.81 ± 0.021	0.814 ± 0.02	0.803 ± 0.02
	Ensemble Mean	0.866 ± 0.017	**0.853 ± 0.015**	0.841 ± 0.015

In addition, I would like to see a similar quantification for the multiannual trends. With these numbers, users will have a good basis to decide which product may be most applicable for their application. Sect. 4.2/4.3: Please add another section on multi-annual trends, as for the ensemble mean.

The authors thank the reviewer for this suggestion, and have added a new section (Section 4.4 – Multi-Annual Trends) outlining the decadal soil temperature trends for each product. We have included a new figure (Figure 6 in the manuscript) showing various aspects of the annual mean soil temperature trends for each product, alongside station estimates where available, with three supplementary figures displaying maps of the annual mean, DJF, and JJA soil temperature trends (Figure S5 – S7).

Most products generally show annual mean warming (positive soil temperature trends) over North America, with a pocket of regional cooling over Western North America. Over Eurasia, most products, show warming over the annual mean, though CFSR, and the European reanalyses show a region of cooling, particularly over higher latitudes. Similar to skill score, and RMSE, products show greater disagreement over higher latitudes, and during winter.

Fig. 2: It says "bias\RMSE" in the figure, and bias in the caption. I guess this is the bias, not the RMSE?

Yes, this was a typo, and should have read "bias". Figure 2 has been revised to include both bias and RMSE.

Fig. 3: please add (e.g. a histogram with the) number of data points per bin; please add to the caption for which time period (I guess monthly?) the individual values are obtained (same for the following figures 4 and 5).

Yes, the standard deviation is based on the monthly soil temperatures within each station soil temperature bin. We have added this to captions of Figure 3 and the new Figure 4 (standard deviation figure). Figures 3 and 4 now also include a histogram with the number of data points per bin.

Fig. 4: the quality of this figure is very poor, both graphically and content-wise. I cannot really see what to get out of the figure, other than a blob of blue and red dots giving the max/min-range. The authors could consider binning the data as in the previous figures and presenting whisker plots with mean, standard deviation (10/90percentiles) and min/max. At least for the means, it seems to me that this information would be pretty much the same as presented in the previous Fig. 3. The authors should consider if this figure can be moved to the supplement (after drastically improving its quality as suggested above).

We agree that Figure 4 was hard to read and did not really add much new information. This figure has been removed in the revised manuscript.

Fig. 6: to make it easier for the reader, please add a sentence to the caption where a “1:1 match” would be located, i.e. explain the star in the figure better than just “Reference”. I also suggest renaming “standard deviation” to “standard deviation of data set”.

A product would line up on the 1:1 line if the timeseries at all stations matched perfectly, and have added a sentence to the caption to this effect. We’ve also renamed “Reference” to “Station” and “standard deviation” to the “standard deviation of the dataset”.

(Note this is now Figure 5 in the revised manuscript)

Fig. 8: Please provide more information in the caption, what does “1 model” mean, how is this calculated?

“1 model” refers to the average RMSE and Pearson Correlation of the reanalysis and LDAS individual products themselves (not including the ensemble mean).

(Note: This figure has been removed in the revised manuscript)

Fig. 9: The axis description of Panel B is almost in the Panel A figure; the color axis in Panel B is not specified. This is poor manuscript preparation by the authors.

We have corrected this, and the figure now includes a colorbar in Panel B.

(Note that this figure is now Figure 12 in the revised manuscript)

Sect. 6.1: This is a key question: “why are the reanalysis products so bad and how can they be improved”. Despite going in much detail with the validation, it appears that the manuscript cannot add much new insight, pretty much the entire section is about other published studies.

We thank the reviewer for this question. We have used the revision process to reassert the principal contributions of our study, which are articulated in the revised manuscript and summarized here. This study is the first to validate soil temperatures from all major modern reanalysis systems across the extratropical northern hemisphere. This represents a multi-faceted, complex challenge, requiring assembly of a wide variety of reference datasets from sparse observing networks with large differences in data quality and availability. Previous studies generally limited their analysis to a restricted geographical area, or to one product, allowing for more detailed explorations of specific phenomena, and the drivers of bias in these regions. We also present a comprehensive quantification of product performance across the seasonal cycle, comparing mid-latitude regions and permafrost regions separately. This study is also the first to investigate the value of an ensemble mean (blended) soil temperature product.

Given the extent of the challenge involved in achieving this first set of objectives, a detailed investigation of the primary drivers of bias in each product is beyond the scope of this study. However, our research paves the way for these detailed process studies in future by identifying the sign, magnitude, timing and location of biases.

Sect. 6.2: How does this impact the calculated trends?

Discontinuities in reanalysis products typically affect the upper atmosphere (e.g. Hersbach et al., 2020; Shuangguan et al., 2019), though discontinuities have been noted in the deep soil moisture and deep soil temperature values of MERRA2 (Bosilovich et al., 2015), and in high latitude precipitation (Reichle et al., 2017). Most modern reanalysis products employ overlapping spin-up periods to reduce such issues (Hersbach et al., 2020), and during our analysis we were unable to detect any obvious discontinuities (i.e., those emerging above the magnitude of internal variability) in any soil temperature product at the grid cell, regional or hemispheric scale.

We believe that the changes in the variability noted in CFSR in Section 6.2 in the previous version of the manuscript were due in large part to the anomalous soil temperatures in 2009 and 2010 (caused by issues with CFSR snow depth values during these years); discussed in Section 4.4 – Multi-Annual Trends.

Sect. 6.3: This is a potentially important issue which should be explored some more in the methods section. The authors should at least analyze the altitude of the observation sites with respect to the average altitude at typical scales of reanalysis grid cells, and possibly exclude observations e.g. from "mountain sites" much higher than the average altitude. For those, the comparison is meaningless, as they should (on average) be significantly colder than the reanalysis products, even if these were perfect.

We thank the reviewer for this suggestion. We investigated the impacts of elevation by separating stations based on their elevation (as obtained from the Copernicus 90m DEM), and grouped stations into three elevation zones (0m – 500m, 500m – 1000m and >1000m). Owing to the small sample size of higher elevation stations (Table R3), we combined the mid and higher elevation stations together in Table R4. While the referee is correct that the RMSE in elevation is substantially larger at higher elevation stations, and that most products underestimate the elevation (Table R3), the mean performance for soil temperature is not substantially different in low- or higher-grid cells with an elevation at or above 500m (Figure R1).

We have added a subsection in the Methods section (Section 3.3) explaining how we assessed the impact of elevation on product performance and a brief discussion of any minor differences in performance over the elevation bands has been added in Section 4.3 – Spatial Variability.

Table R3. Number of grid cells in each elevation range for the near surface and at depth.

Elevation Range	**Near Surface Grid Cells**	**Depth Grid Cells**
Below 500m	310	275
500m – 1000m	105	87
1000m +	15	15

Table R4. Average elevation RMSE (in metres), along with the 95% confidence interval, for each product as a function of station elevation. Biases are calculated for low elevation stations (below 500m), and for stations above 500m.

Product	**Avg Elevation RMSE (0m - 500m)**	**Avg Elevation RMSE (500m+)**
CFSR	147.73 ± 24.33	589.31 ± 61.38
ERA5	140.64 ± 21.91	583.10 ± 61.63
ERA5-Land	155.55 ± 28.76	587.35 ± 61.62
ERA-Interim	163.49 ± 17.93	580.54 ± 59.14
FLDAS	76.43 ± 14.49	156.74 ± 42.53
JRA55	144.00 ± 16.73	579.53 ± 59.17
MERRA2	76.05 ± 10.78	143.97 ± 32.69

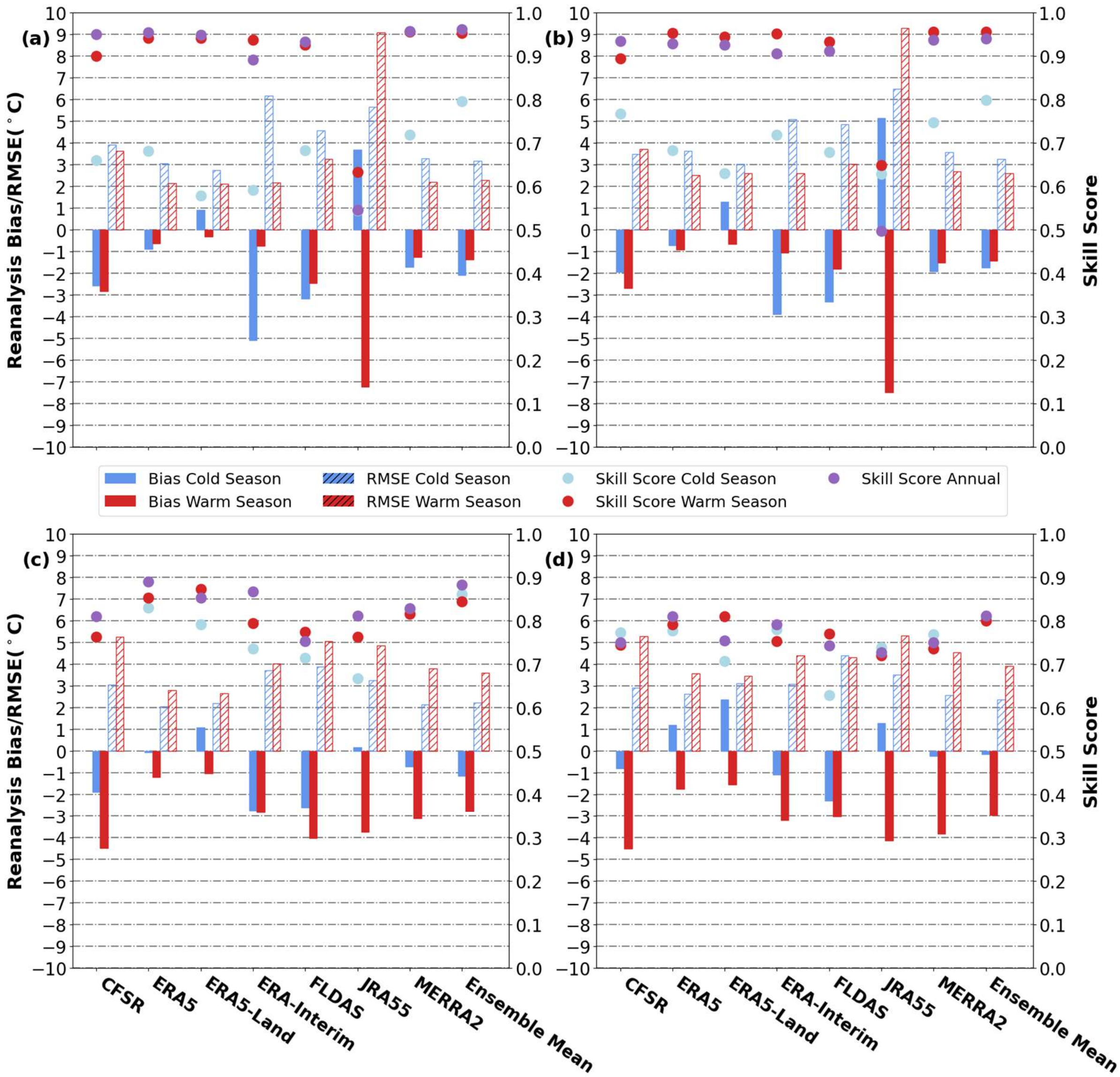

Figure R1. Bias, RMSE and skill score values for grid cells containing stations with an elevation between 0m and 500m (panels A and C) and for elevations 500m+ (panels B and D). The near surface is shown in panels A and B, while the depth is shown in panels C and D.

Minor Comments

L. 22: Please revise this sentence, this not really correct. The first statement with the 800GtC is for permafrost soils only (i.e. only the permafrost, not active layer and permafrost-free areas in the permafrost region), while the second statement refers to the entire soil carbon pool in the permafrost region. The carbon pool in the atmosphere is around 850 GtC, so only the entire soil carbon pool is significantly larger. In addition, the numbers from Tarnocai et al. (2009) are somewhat outdated, better to use the estimates from Hugelius et al. (2014).

We thank the reviewer for highlighting this. The sentence has been modified to read “Roughly 1400 to 1600 gigatonnes of carbon (GtC) is estimated to be stored in soils in permafrost affected regions of the Northern Hemisphere (Hugelius et al., 2014).”

L. 175: please add abbreviation (SS)

Corrected.

Sect. 4.2 Change title to “Seasonal cycle”. “Temporal variability” would also include multi-annual temperature developments which is not analyzed here.

This section has been renamed to "Seasonal Cycle".

L. 304: close bracket.
Corrected.

L. 329/332/333: ??

These were referring to Figure 10. The missing figure numbers here have been corrected.

L. 482: make this a proper sentence.

This sentence has been split into two separate sentences as follows: "However, it is also apparent that the uncertainties arising from variations in the number of grid cells included in a station average are substantially smaller than the spread between reanalysis products. During the cold season, the uncertainty in soil temperatures associated with the spread between reanalysis products is often two to three times larger than the uncertainty arising from fluctuations in station availability."

L. 491: I have a hard time finding this in the results section, please specify what exact metric (monthly averages) this refers, and provide a reference to the section or figure where this is presented. For other metrics, single products seem to do better, Fig. 2 for example suggests that ERA-5 with respect to the bias is significantly better than the ensemble mean.

The RMSE of the new ensemble mean is approximately 0.5°C better than the next best product during the cold season over permafrost regions across both depths (Figure S1).

The low overall bias of ERA5 and ERA5-Land arises due to a cancellation of errors over the study region, with generally warm biases over areas underlain by permafrost, and predominantly cold biases over more southern regions – a finding similar to Cao et al. (2020). When the RMSE of these products are considered, it becomes apparent that they still show substantial errors, particularly over the cold season. Conversely at higher latitudes, ERA5 and ERA5-Land show substantial degradations to performance, where the ensemble mean outperforms them (Figure S1).

References

Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., Huffman, G. J., Adler, R. F., and Wood, E. F.: Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS, Hydrol. Earth Syst. Sci., 23, 207–224, https://doi.org/10.5194/hess-23-207-2019, 2019.

Bosilovich, M., Akella, S., Coy, L., Cullather, R., Draper, C., Gelaro, R., Kovach, R., Liu, Q., Molod, A., Norris, P. M., Wargan, K., Chao, W., Reichle, R., Takacs, L., Todling, R., Vikhliaev, Y., Bloom, S., Collow, A., Partyka, G. S., Firth, S., Labow, G., Pawson, S., Reale, O., Schubert, S. D., and Suarez, M.: MERRA-2: Initial Evaluation of the Climate, NASA, Maryland, USA, 2015.

Cao, B., Quan, X., Brown, N., Stewart-Jones, E., and Gruber, S.: GlobSim (v1.0): deriving meteorological time series for point locations from multiple global reanalyses, Geosci. Model Dev., 12, 4661–4679, https://doi.org/10.5194/gmd-12-4661-2019, 2019.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions, Remote Sensing of Environment, 203, 185–215, https://doi.org/10.1016/j.rse.2017.07.001, 2017.

Gruber, S., Brown, N., Stewart-Jones, E., Karunaratne, K., Riddick, J., Peart, C., Subedi, R., and Kokelj, S. V.: Permafrost Ground Temperature Report: Ground temperature and site characterisation data from the Canadian Shield tundra near Lac de Gras, Northwest Territories, Canada, Northwest Territories Geological Survey, Northwest Territories, Canada, 2019.

Shangguan, M., Wang, W., and Jin, S.: Variability of temperature and ozone in the upper troposphere and lower stratosphere from multi-satellite observations and reanalysis data, Atmos. Chem. Phys., 19, 6659–6679, https://doi.org/10.5194/acp-19-6659-2019, 2019.