Response to Editor Prof. Chris Derksen

Thanks very much for your thorough revisions to the manuscript. All three of the Reviewers have provided a second round of comments. Please revise the manuscript with due attention to their comments, after which the final editorial decision will be taken. Reviewer 2 has a number of concerns still to be addressed, so I will give this Reviewer the opportunity to provide a final round of comments. Thanks very much for ongoing efforts to revise the manuscript. Best regards, Chris Derksen

R: Thank you very much for your time and effort in handling our manuscript entitled "***Which global reanalysis dataset represents better in snow cover on the Tibetan Plateau?***" (MS No.: ***egusphere-2024-82***). We appreciate your granting the opportunity to address the reviewers' comments and improve our manuscript.

We are grateful for the valuable feedback and guidance provided by the three reviewers. We agree with their suggestions and have revised our manuscript accordingly. Now the revisions have been completed, and we believe the revised manuscript effectively addresses all the reviewers' concerns. Our point-by-point responses are attached in the files, summarizing the specific modifications made to the manuscript. Additionally, we have re-ordered the figures and adjusted the reference format to meet the requirements of The Cryosphere. We have also carefully checked the manuscript for any typographical errors and ensured that there were no changes to the co-authors and their affiliations. The datasets and equations used in the manuscript remain unchanged.

We hope that the revised manuscript has been improved satisfactorily and it can be accepted for publication in The Cryosphere. We welcome further review and guidance from you and the reviewers to address any remaining issues and ensure the quality and completeness of the manuscript. Thank you again for your consideration and guidance. We look forward to receiving your feedback soon.

Anonymous Referee #1

Which global reanalysis dataset represents better in snow cover on the Tibetan Plateau?

The authors put a lot of effort in the revision of their manuscript by addressing all raised issues.

R: Thank you very much for recognizing our revisions to the manuscript. Your constructive suggestions have been greatly inspiring and have significantly helped us enhance the scientific rigor of our results. We have corrected the data resolution processing method and re-evaluated all results throughout the manuscript. Additionally, we have restructured the manuscript to segregate the dataset evaluation and bias source analysis into distinct sections within the results, ensuring a clearer logical flow. The following are our point-by-point responses to the comments, starting with "R:".

Major points:

• **L261:** Good that you brought everything to a common grid. However, bilinear interpolation is not suited to upscale data over multiple orders of magnitude. Especially for HMASR at 500m resolution. The recommended procedure is to simply aggregate the data to approximately a similar resolution (e.g, from a 0.1 to 0.5 grid by averaging 5 by 5 pixels), and then using the bilinear interpolation just to match the final grids (origin, extent, and/or slight resolution mismatch). (you can see this effect clearly in Fig1a, where SPIRES is much too noisy)

R: We previously overlooked that bilinear interpolation is unsuitable for magnifying data across multiple orders of magnitude. Therefore, we used the grid averaging approach to directly process high-resolution HMASR, ERA5L, SPIReS, and TPMFD data into a 0.5°×0.5° grid. Additionally, we used the grid averaging method to process ERA5 and GLDAS data, which could be converted from a 0.25° to a 0.5° resolution. Only MERRA2, JRA55, and CRAL were processed using bilinear interpolation. After reprocessing all the data, we re-evaluated and updated the results throughout the manuscript.

• Capturing sign and significance of trends over 17 years is not the best assessment of temporal agreement. It would be great to see some time series plots of the variation of SCF with water years by dataset, potentially also by basin. This could be easily summarized with values of bias and correlations, to give you some more information on temporal agreement between datasets.

R: We have added the annual SCF time series for SPIReS and the eight reanalysis datasets, as shown in Fig. 4. Additionally, we analyzed the temporal agreement between the reanalysis datasets and SPIReS from the perspective of annual time series correlation. This information for each basin is provided in Supplementary Fig. S5. We provide specific description for evaluating the temporal performance of the reanalysis datasets in Section 3.1.1.

Minor points:

• **L314:** parentheses three (two) is weird. Please spell out.

R: We revised the unclear statement: "For HMASR, CRAL, and CFSR, which do not include their parameterization among these three methods, we derived three additional SCF datasets. MERRA2, ERA5, ERA5L, JRA55, and GLDAS derive another two SCF datasets."

• Figure 5 does not match with previous ones: MERRA2 seems to have little bias, while in the previous plots it showed a high negative SCF bias.

R: We found that MERRA2 shows an underestimation both spatially and seasonally, which is consistent with the results throughout the manuscript. Spatially, MERRA2 underestimations SCF across the entire TP, with values in the Indus basin reaching as low as -0.5 (Fig. 2a). Once considering the TP as a whole, the mean SCF climatology of MERRA2 is 0.05, which is 0.09 lower than the 0.14 observed in SPIReS (Fig. 1b). Seasonally, treating TP as a whole explains why the MERRA2 SCF bias is around 0.1 in Fig. 5. Additionally, the SCF bias of 0.1 is significant compared to the SPIReS climate average SCF of 0.14.

• Figure 9 is a bad choice for visualization, since it's extremely hard to compare opposite bars. Consider a table or simple bar/point chart. Also bias as metric would be interesting, to see if there is a relationship with snow height. Did you assess this by basin? Maybe different parametrizations might work better in different areas (shallow or variable snowpack as in the ITP vs more deep snowpack as in Amu and Indus, etc.)

R: We have simplified the complex Fig. 9 into a bar chart. Additionally, in Supplementary Fig. S7, we provide the spatial distribution of the SCF climatological bias calculated offline using different parameterization methods for all reanalysis datasets. This metric visually illustrates the improvements in SCF due to parameterization. Previously, we overlooked the impact of parameterization on SCF across different basins. Further investigation revealed that the influence of parameterization between basins is minimal. We present these findings in Supplementary Fig. S8 and briefly discuss them in Section 3.2.2 of the manuscript.

• **L701:** "from the internet" is vague, consider providing a doi or url for each.

R: We revised the vague statements in the data availability statement and provided the URLs for all reanalysis datasets.

Anonymous Referee #2

Which global reanalysis dataset represents better in snow cover on the Tibetan Plateau?

Major revision

Please refer to the edited manuscript (egusphere-2024-82-referee-report.pdf) for detailed editorial suggestions.

R: Thank you very much for your detailed editorial suggestion. These clear and detailed recommendations have helped me understand your suggestions better, allowing me to revise accordingly. The following are our point-by-point responses to your comments, starting with "R:".

I thank the authors for carefully responding to my comments and those of the other reviewers. As stated in my initial review, comparison of SCF products over the TP is an important endeavour. This manuscript attempts to provide such an evaluation and includes some interesting analysis of the impact of SCF parameterization (going from SD/SWE to SCF). However, even with the additions made to the manuscript I continue to struggle to follow the logic and get lost in the weeds of the results section. The authors struggle to interpret and present their results in a succinct and logical manner. It's not clear to me if this issue stems from language or from a lack of understanding of the statistical tests used.

R: In the revised manuscript, we removed numerous complicated expressions and aimed to analyze the conclusions with clearer logic and more concise language.

The text could be significantly reduced. I appreciate it is challenging to summarize the performance of several different reanalysis products over various spatial and temporal (annual and seasonal) domains. Nevertheless, it is important to synthesize the results to digestible amount of text and figures.

R: We have condensed the introduction, results, and conclusions sections of the

manuscript, reducing verbose text. In the results section, we merged discussions of SCF climatological and seasonal spatial variations into section 3.1.1. Additionally, the meteorological forcing effects on SCF climatology and seasonal biases are now analyzed in a single paragraph within section 3.2.1. The analysis of parameterization throughout the manuscript has been consolidated into section 3.2.2. We used clear language and figures to present the results in an easily understandable manner.

The authors removed many of their broad conclusions that were not supported by evidence which is appreciated. However, it was still difficult to review the quality of some of the interpretations and analysis due to the paper structure. I found myself having to examine the figures, go back to the text, and try to reconcile the two and figure out what the authors are trying to say. This was also the case in the first round of revisions, but I was hopeful that by the second round it would be easier. For this reason, I did not provide too many comments on the content and validity of the conclusions.

R: As suggested, our results section is now divided into three parts: evaluation, attribution, and optimization.

In the evaluation section (3.1), we focus on assessing the performance of the reanalysis dataset in spatial (3.1.1) and temporal (3.1.2) aspects. In the temporal performance evaluation, we added a description of the time series correlation, as shown in Fig.4.

In the attribution section (3.2), we specifically describe the impacts of meteorological factors (3.2.1) and parameterization (3.2.2) on the SCF biases of the reanalysis dataset. The impact of data assimilation has been separately discussed in the discussion section.

In the optimization section (3.3), we analyze the capability of combining the reanalysis dataset to optimize SCF.

The authors added descriptions of the methods which are appreciated. However, it would help readability if the methods-related text in the results section was moved to and integrated with this new methods section. While it can be helpful to remind the

readers of the objectives and the methods, I found there was a bit too much explanation in the results section and the text loses focus. Further, there are still elements of the methods that aren't clearly laid out. I have provided a suggested template in the detailed comments.

R: We have moved the methods-related text in the results section to the method description section and integrated it with the original method. Additionally, we revised the method description for the SCF accuracy assessment using the provided method description template. In the attribution analysis, we have simplified extensive discussions on topics such as snowfall and temperature mechanisms to "Overestimated snowfall contributes to heightened snow accumulation, while underestimated temperatures can impede the snowmelt process, leading to an overestimation of snow cover (Liu et al., 2022).". Furthermore, the discussion on the impact of data assimilation on SCF bias has been moved to section 4.1.

I feel there is good work in this paper, and it should be published at some point. Unfortunately, in its current form I do not feel it would be a useful contribution to the community. I really want to see this manuscript succeed and hope the authors are able to re-work the text.

R: We have extensively revised the manuscript based on your suggestions. We hope that the revised manuscript can fit the publication standards.

I have provided detailed editorial suggestions directly on the revised manuscript. I hope the authors can build on some of these suggestions to produce a revised manuscript that is able to communicate this interesting work to the broader community.

R: We have carefully read your detailed editorial suggestions and made corresponding adjustments and improvements to the manuscript.

Specific comments

-Please refer to the edited manuscript (egusphere-2024-82-referee-report.pdf) for further suggestions. The absence of detailed suggestions doesn't mean revisions are not recommended. Not all sections received the same level of attention.

R: Thank you once again for your detailed editorial suggestions on our manuscript. We have adopted most of your suggestions, as reflected in the revised version of the manuscript. Here, we will explain the sections where we have different viewpoints from your suggestions:

Comment on line 80 ("but" instead of "and"): Shortcomings in the studies by Orsolini et al. (2019) and Li et al. (2022) include incomplete evaluation of the reanalysis datasets and dimensions assessed. We consider these two aspects to be parallel rather than contrasting relationships. Moreover, taking into account your comment on line 85, we have revised this sentence to "Their studies considered the SCF accuracy of a limited number of reanalysis datasets and lacked multidimensional evaluation that considers aspects such as regional variations and annual trends, as well as an in-depth analysis of the impact of parameterization on SCF bias."

Comment on line 285 (to determine?): The original statement was ambiguous due to language issues. It has been revised to: "The Mann-Kendall (MK; Kendall, 1975; Mann, 1945) test was used to assess the significance of annual trends. Since the sign (+ or –) may impact the robustness of the trend analysis results, we employed the Consistency Index (CI; Zhang et al., 2021) to compare the agreement in SCF annual trend signs between the reanalysis datasets and SPIReS.".

Comment on line 436 (how so? Do you mean that JRA55 has similar snowfall and temperature biases to E5/EL/CF but differing SCF estimates so differences in SCF between JRA and these other 3 products unlikely to be due to forcing data. One possible explanation could be differences in paramaterization. This is explored in section X. (be clear that paramaterization is only one of many possible explanations for the differences)): We have removed the discussion about JRA55 snowfall and temperature bias from the revised manuscript and placed the

parameterization discussion in Section 3.2.2. The original statement indicated that in JRA55, the magnitude of snowfall and temperature bias is similar to ERA5. However, the spatial distribution of significant correlation between snowfall bias and SCF bias, as well as temperature bias and SCF bias, differed from ERA5 in few areas. This suggests that factors other than snowfall and temperature, such as parameterization, may impact SCF bias in JRA55. However, the effects of snowfall and temperature are not excluded.

-Suggest trying to separate pure results from discussion and interpretation to help the flow of the manuscript.

R: We have separated results from discussion and interpretation.

-Just say correlation rather than writing R.

R: We have replaced "R" with "correlation" throughout the manuscript.

Methods

-Move line 337 end to line 329 to methods.

R: We revised this paragraph according to the PDF editorial suggestions.

-Some suggested text for the methods section to more clearly lay out the study:

L261: For each 0.5x0.5 grid cell we calculate the climatological SCF over the full period and seasonally for each of the reanalysis datasets and SPIReS (e.g. Figure 1a). Absolute bias and correlation (Pearson correlation coefficient) were calculated from these values at both the basin and regional (TP) scales. [please correct me if I am wrong here and you used the time series of annual SCF] Spatial maps and summary values of the climatological SCF and its bias with respect to the SPIRES dataset are presented. Additionally, Taylor Diagrams are used to provide additional information regarding the RMSE and standard deviation ratio (STDR). The climatological SCF values for each grid cell were used as input to calculate the Taylor Diagram's component metrics -

correlation, RMSE, and Standard Deviation Ratio (STDR). [cleaned text from line 271 277 goes here ending with (closer to 1).]. The component metrics (correlation, RMSE, STDR) were summarized by the Taylor Skill Score (SS) as follows:

SS = ... [the SS equation goes here] can probably omit the rest of the text on line 280-282]

Similar to the absolute bias, Taylor Diagrams and SS were calculated for the full TP and each of the nine sub-basins for the full period and seasonally.

R: Thank you very much for providing the method description text. It clearly outlines our original methods and adds many details, which will greatly help us organize our result logical framework. We have revised all method descriptions using this template. For more details, please refer to the revised manuscript, section 2.3.1.

Methods Section 2.3.2 – suggest splitting into 3 separate short sections. This structure could help re-frame the results.

- Analysis of SCF bias sources
- Analysis of SCF parameterization
- Generation of ensemble datasets

R: We have divided the methods section into four subsections: 2.3.1 Evaluation of SCF accuracy for reanalysis datasets, 2.3.2 Analysis of SCF bias sources for meteorological forcings, 2.3.3 Analysis of SCF bias sources for parameterization method, and 2.3.4 Generation of combined optimal datasets.

Results

- Condense and streamline

R: We have extensively streamlined the manuscript.

- Separate results and discussion related to SCF parameterization from the text concerning temperature and precipitation forcing. For example, L436-437 you mention that met forcing can only explain some of the biases and then you go into a discussion

of parameterization. Instead, I suggest you mention that SCF parameterization could be a compounding factor and that it will be discussed in Section #.

R: As previously described, we reconstructed the structure of the manuscript. The effects of meteorological forcing and parameterization are described in two subsections.

- Lines 446-460 could be moved to discussion and/or the parameterization parts moved to a results piece related to parameterization thresholds.

R: As mentioned above, we have moved the content related to parameterization to section 3.2.2 in the results, and the content related to data assimilation to section 4.1 in the discussion.

- Consider moving seasonal SCF state and performance to follow immediately the annual/full period results before discussing the bias attribution. Bias attribution for seasonal case can follow or be integrated with the full period results.

R: We have organized the assessment of climatological and seasonal spatial variability in section 3.1.1. Subsequently, we moved the analysis of all SCF bias sources to section 3.2. In the bias source analysis, we combined the annual and seasonal dimensions.

- There are still difficulties in explaining and interpreting the results and metrics and some confusion between the time and space dimensions.

R: We have adequately addressed the results in the spatial and temporal dimensions of the manuscript.

Discussion

I did not provide specific or detailed comments on the discussion or conclusions as there is already enough to be re-worked in the first three sections.

R: We moved the discussion on data assimilation to this section. Additionally, we have streamlined and refined the text.

Conclusions

The authors expanded the conclusions at the request of one of the reviewers. While a good attempt, the new text needs to be cleaned and carefully edited (condensed). I did not provide detailed comments on the conclusions but hope the authors can build on suggestions provided elsewhere in the text.

R: Based on the restructured logic of the results, we have adjusted the structure of the conclusion. Additionally, we removed much of the repetitive text in the conclusion to express our points more concisely.

References

Kendall, M. G.: Rank Correlation Methods, J. Inst. Actuar., 75, 140–141, https://doi.org/10.1017/S0020268100013019, 1975.

Li, Q., Yang, T., and Li, L.: Evaluation of snow depth and snow cover represented by multiple datasets over the Tianshan Mountains: Remote sensing, reanalysis, and simulation, Int. J. Climatol., 42, 4223–4239, https://doi.org/10.1002/joc.7459, 2022.

Liu, Y., Fang, Y., Li, D., and Margulis, S. A.: How well do global snow products characterize snow storage in High Mountain Asia?, Geophys. Res. Lett., 49, https://doi.org/10.1029/2022GL100082, 2022.

Mann, H. B.: Nonparametric tests against trend, Econometrica, 13, 245–259, https://doi.org/10.2307/1907187, 1945.

Orsolini, Y., Wegmann, M., Dutra, E., Liu, B., Balsamo, G., Yang, K., De Rosnay, P., Zhu, C., Wang, W., Senan, R., and Arduini, G.: Evaluation of snow depth and snow cover over the Tibetan Plateau in global reanalyses using in situ and satellite remote sensing observations, The Cryosphere, 13, 2221–2239, https://doi.org/10.5194/tc-13-2221-2019, 2019.

Zhang, H., Zhang, F., Che, T., Yan, W., and Ye, M.: Investigating the ability of multiple reanalysis datasets to simulate snow depth variability over mainland China from 1981 to 2018, J. Clim., 34, 9957–9972, https://doi.org/10.1175/JCLI-D-20-0804.1, 2021.

Anonymous Referee #3

Which global reanalysis dataset represents better in snow cover on the Tibetan Plateau?

The manuscript is much improved in structure and writting. The main concerns I have raised earlier have been adequately addressed. I would suggest a minor revision before it can be published.

R: We greatly appreciate your positive feedback and constructive suggestions. Previously, we overlooked that bilinear interpolation is unsuitable for magnifying data across multiple orders of magnitude. We have revised the method for processing data resolution and reprocessed all results in the manuscript. Additionally, we have restructured the entire manuscript to divide the results section into three parts: dataset evaluation, bias source analysis, and dataset optimization. We hope that the revised manuscript can fit the scope of the journal standard. The following are our point-by-point responses to the comments, which start with "R:".

Specific comments

1) P3, L46: I donot see the necessity of the clarification that ground-based measurements achieve the best accuracy.

R: We removed the statement from the introduction suggesting that ground-based measurements provide the most accurate data. Instead, we revised this section: "Comprehensive ground-based measurements face challenges due to the complex terrain and harsh weather conditions on the TP (Yang et al., 2019), leading to issues of spatial representativeness.".

2) P3, L55: please give the typical temporal coverage of RS datasets

R: We have provided the typical temporal coverage of remote sensing datasets in L58 of change-tracked manuscript.

3) Table 1: Centre --> (Sponsoring) agencies

R: Changed as suggested.

4) P10, L237: what do you mean “surface elevation”?

R: In determining snowfall types, we followed the method described by Ding et al. (2014), where “surface elevation” refers to local altitude. We replaced “surface elevation” with “altitude” to enhance clarity and make the methodology clearer.

5) P11, L264: Do you really need to give the total grid number here?

R: We removed the description of the total grid number from the manuscript.

6) P12, L283: Remove the first sentence.

R: Removed as suggested.

7) P15, L363: Orsolini et al. (2019) used the snow depth from CMA as ground truth, which are mainly in areas with very limited snow. This is a possible reason that MERRA-2 was found outperform other reanalysis although MERRA2 seems to generally underestimate the snow depth over TP.

R: Orsolini et al. (2019) evaluated the ability of reanalysis datasets to simulate SD and SCF in the TP. They noted that it is JRA55, rather than MERRA2, that assimilates SD data. Therefore, the contradictory results in MERRA2 between our study and Orsolini et al. (2019) might arise from different evaluation metrics being employed. Orsolini et al. (2019) emphasized the high correlation of MERRA2 but overlooked the poorer STDR resulting from the overall underestimation of SCF, thus indicating poorer spatial performance.

References

Ding, B., Yang, K., Qin, J., Wang, L., Chen, Y., He, X., 2014. The dependence of precipitation types on surface elevation and meteorological conditions and its parameterization. J.

Hydrol. 513, 154–163. https://doi.org/10.1016/j.jhydrol.2014.03.038

Orsolini, Y., Wegmann, M., Dutra, E., Liu, B., Balsamo, G., Yang, K., De Rosnay, P., Zhu, C., Wang, W., Senan, R., Arduini, G., 2019. Evaluation of snow depth and snow cover over the Tibetan Plateau in global reanalyses using in situ and satellite remote sensing observations. The Cryosphere 13, 2221–2239. https://doi.org/10.5194/tc-13-2221-2019

Yang, M., Wang, X., Pang, G., Wan, G., Liu, Z., 2019. The Tibetan Plateau cryosphere: Observations and model simulations for current status and recent changes. Earth Sci. Rev. 190, 353–369. https://doi.org/10.1016/j.earscirev.2018.12.018