Answers to editor: TC-2021-281

Temporal stability of long-term satellite and reanalysis products to monitor snow cover trends

Ruben Urraca and Nadine Gobron

PUBLIC EDITOR COMMENTS

Dear authors,

Thank you very much for the thorough revisions in response to the two reviewers' comments. Some of those comments were major ones, and I appreciate the new analyses you carried out in response to those major issues. I appreciate in particular the inclusion of all new Canadian stations, which increased the number of stations analysed from 470 to 527; and the efforts to obtain and remove from your validation the list of stations that are assimilated in ERA. I think indeed this was an excellent comment by reviewer 1 and commend you for the effort made to address it. I wonder if this list could be made public as an appendix to the paper – upon clarification with the ERA team. I also think that the sensitivity analysis included in response to the comment on how to identify discontinuity has strengthened the paper.

Given that the paper underwent major revisions, I am sending it back to the reviewers. I feel that some more explanations (and possibly analysis) are needed to address reviewer's 1 comment on why the authors decided to include the stations which failed the spatial representative test in the trend analysis, as those trends identified will be very local. I will first leave this to the reviewer to evaluate.

Once again, thanks for the thorough revision.

All my best,

Answer: Regarding the concern of how low spatially representative stations affect the trend analysis:

Spatial representativeness is critical when comparing satellite pixels to in-situ measurements (point-to-pixel comparison), as it introduces a spatial mismatch between satellite values and in-situ values. Therefore, stations where this mismatch error is significant (due to a high spatial variability and/or a coarse resolution of satellite pixel) need to be discarded or corrected using an up-scaling procedure.

However, stations with low spatial representativeness can still be useful for other applications such as trend analysis. This has been discussed by previous studies such as *Schwartz et al 2017.* A good example are stations on the coast (Norwegian coast, Canadian coast or Eastern Russia). All of them are excluded for the point-to-pixel comparison because the satellite pixel includes the sea. However, the trends observed are coastal locations are still valid from a climate perspective.

Following the editor comment, we have compared the decadal trends obtained in the group of low spatially representative stations against those obtained in the group of spatially representative stations (Figure 1). We do not observe any inconsistency between both groups.

We have also clarified in the Methods sections the limitations of using in-situ data for evaluating global trends:

"Note that the density of stations was too low for a complete analysis of NH snow cover trends. Even in regions with good coverage, the heterogeneous density of the stations as well as their different spatial representativeness also prevent the calculation of spatial representative trends. However, our main goal

was to estimate the trend magnitude to evaluate the significance of the artificial trends and discontinuities introduced by each product."

References

M. Schwarz, D. Folini, M. Z. Hakuba, M. Wild. Spatial Representativeness of Surface-Measured Variations of Downward Solar Radiation. JGR Atmospheres 2017.https://doi.org/10.1002/2017JD027261

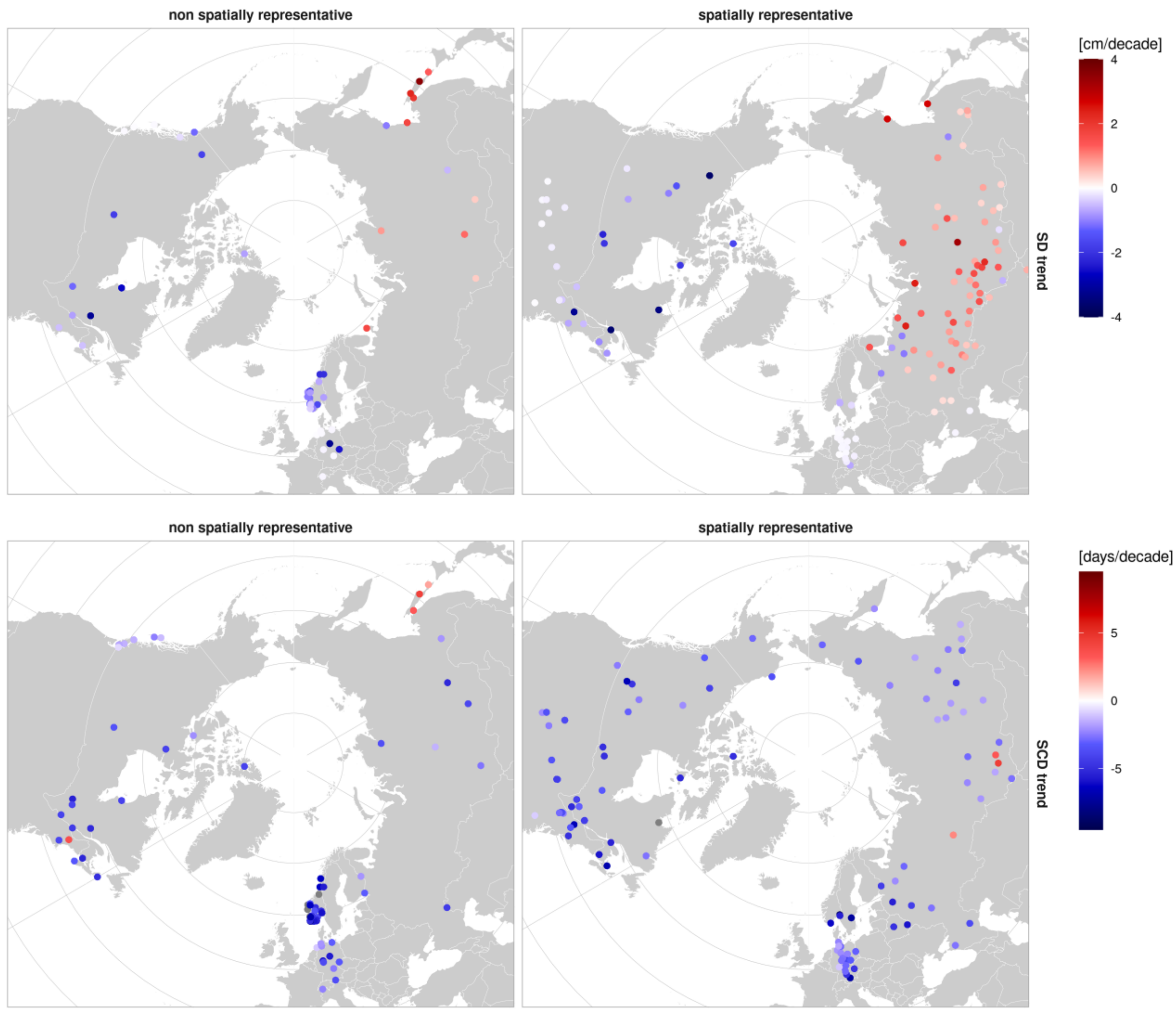

Fig 1 Decadal trends in snow depth (SD) and snow cover duration (SCD) from 1955 to 2015 based on in-situ measurements, splitting between spatially representative and non-spatially representative stations.

NON-PUBLIC EDITOR COMMENTS

Dear authors,

Thanks again for the thorough revisions, which have improved the paper considerably. Just a minor note here: I would have appreciated in few responses a short summary of how the results have changed as a result of the new analysis carried out, e.g. does inclusion of the new canadian stations or removal of almost half of the validation stations change your major results or conclusions?

Answer:

The study is composed of three main sections: (1) stability analysis of gridded products, (2) accuracy analysis of gridded products, and (3) snow trend estimation based on in-situ data. The changes mentioned by the editor affected the different sections as follows:

- Addition of Canadian stations.
    1. Stability analysis: The new Canadian stations corroborated that the ERA5 step discontinuity observed between 1977-1980 exists at global scale. Besides, the analysis of the stations assimilated by ERA5 over Canada corroborated that this discontinuity was most likely due to starting assimilating in-situ snow observations during those years (1977-1980).
    2. Accuracy analysis: Most Canadian stations were low spatially representative, so their influence in this section was negligible.
    3. Trend estimation: This is the section where the addition of Canadian data was most valuable, as now we have a better coverage of NH regions with seasonal snow. Negative trends in both snow depth and snow cover duration were obtained in most of the Canadian stations analyzed.
- Identification of stations assimilated by ERA5.
    1. Stability analysis: We used the list of stations assimilated by ERA5 to identify the cause of the step discontinuities observed in the temporal evolution of ERA5 bias (Figure 6). We could confirm that both the global discontinuity in 1977-1980 and the Asian discontinuity in 1991-1992 were caused by the addition of new snow in-situ measurements to the ERA5 model in those regions during those years.
    2. Accuracy analysis: The number of stations available was significantly reduced in order to guarantee the independence of our validation set. As expected, ERA5 performance metrics (bias, RMSE) slightly worsen after removing the stations assimilated by ERA5.
    3. Trend estimation: this part of the study was not affected, as decadal trends were calculated directly from the in-situ measurements.

Could you double check with the ERA team whether the list of stations assimilated can be made public in an appendix? I think it would make a great service to the community.

Answer: We have not been able to confirm with C3S/ECMWF whether the exact list of stations assimilated (and the exact year of inclusion) can be included in the appendix. They suggested that the full list of stations assimilated (not only snow but other variables) may be included in ERA5 documentation page in upcoming releases.

If this situation changes during the review process of the paper, we will add the corresponding table.

I would also appreciate some more explanations as to why you chose to evaluate only the ERA and ERA5 snow products in the introduction, compared to other available re-analysis. The Japanese reanalysis for instance cover almost the same period as ERA (from 1953 onwards). I feel it would lend strength to your manuscript.

Answer: The goal of the study is to analyze the different stability issues faced by satellite products, global reanalysis, and land reanalysis to monitor long term snow trends. Therefore, we selected one product of each type (the ones with the longest coverage and typically also the ones most used): NOAA CDR, ERA5 and ERA5-Land. We have summarized other long-term products available in the introduction, either reanalysis ones (JRA-55, MERRA-2) or satellite ones (JAXA GHRM5, ESA Globsnow, ESA CCI SWE).

At this point, adding a new gridded product would require a huge amount of work. Our study is based on daily snow depth/cover values, but reanalysis typically provides the data as monthly or hourly averages. In the case of JRA-55, we would need to download and process 3-hourly snow depth data for almost 70 years, which will require a lot of computational time.

Otherwise I look forward to the reviewers' second evaluation and hope the paper can be published soon. It contains some very important results that will be very useful to the community.

All my best,

Francesca