

The Cryosphere Discuss., referee comment RC1
<https://doi.org/10.5194/tc-2021-281-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on tc-2021-281

Chris Derksen (Referee)

Referee comment on "Temporal stability of long-term satellite and reanalysis products to monitor snow cover trends" by Ruben Urraca and Nadine Gobron, The Cryosphere Discuss., <https://doi.org/10.5194/tc-2021-281-RC1>, 2021

This study uses a reference dataset of point snow depth measurements to assess the performance and stability of snow extent and snow cover duration from reanalysis and satellite-derived products. This is important to quantify because changes to the quality and quantity of satellite data and the data sources assimilated into reanalysis can introduce spurious trends and temporal discontinuities into multi-decadal time series. The analysis is focused on ERA5 and the NOAA snow chart climate data record (NOAA-CDR), which are two widely used datasets that provide snow information back to the 1960s. Overall, I found the analysis to be comprehensive in scope, sound in the overall approach, and clearly explained.

I have a number of both major and minor comments, mostly in an effort to further clarify the methods and tighten the messaging. This was a really enjoyable paper to review, thanks to the authors for their efforts.

Major Comments

Lines 61-68: Some additional context/examples could be provided in this paragraph. First: "The transition between different sensors (e.g., JAXA GHRM5) or increasing the number of satellite sources used (e.g., IMS, NOAA CDR)..." It may not be clear to some readers that the IMS product is actually manually derived by analysts from multiple sources of satellite imagery (as opposed to an objective retrieval like the JAXA product). This is noted later on line 88, but this could be mentioned in this introductory paragraph. Second: The ESA GlobSnow and Snow CCI products are derived from the passive microwave satellite record, which is composed of SMMR + SSM/I + SSMIS data, which is another example of how discontinuities can be introduced through changing instruments during the satellite era. (Incidentally, we have found there are differences in the validation statistics for Snow CCI SWE performance related to the different passive microwave sensors. This work is under review, but it would be interesting to also include the Snow CCI dataset in the analysis you present in this work.)

Figure 1: It's unfortunate no data from Canada were used in this study (particularly in the context of the trend analysis in Figure 9, which gives the impress of negative trends in the Eurasian sector and no trends over Canada, which is not the case). There is an updated snow depth dataset for Canada described here: Brown, R., C. Smith, C. Derksen, and L. Mudryk. 2021. Canadian in situ snow cover trends 1955-2017 including an assessment of the impact of automation. Atmosphere-Ocean. DOI: 10.1080/07055900.2021.1911781. For future reference, the Canadian Historical Daily Snow Depth Database should soon be available here (or contact the authors of the above paper): <https://catalogue.ec.gc.ca/geonetwork/srv/eng/catalog.search#/metadata/63dca4bb-a29a-43b0-828b-7eccb03de456>

Section 2.2: How did you ensure that the snow depth observations retained for analysis were not assimilated into ERA5? This issue must be addressed specifically in the text to ensure independence between the reanalysis and validation datasets.

Section 2.4: Very interesting comparison with the analysis of Hori et al (2017). I'm not fully clear on how the SCF surrounding the station was determined: "In this study, we used the SCF in the surroundings of the station measured at RIHMI stations to analyze the correlation between SD at the station and the surrounding SCF (Fig. 2)." Was IMS data used to determine SCF? What distance was used around the station (IMS pixels contained by coarser ERA5/ERA5-Land pixels as described in Section 2.3)?

Section 2.4.1: It is noted that "Stability was evaluated by analyzing how the annual bias in both SD and SCD changed temporally." and that stability was analyzed separately for the RIHMI and GHCN networks. But how were step changes statistically determined (the vertical lines in Figures 4 and 5)? Line 198: Why was the interval of four years selected to compare the bias difference before and after a step discontinuity? Was there any testing performed to confirm that this was some sort of ideal number?

Section 2.5: What is the justification for including the stations which failed the spatial representative test in the trend analysis?

Section 3.1: I appreciate the effort taken to quantify the spatial representativeness of the point measurements. This is a long standing problem in the validation of gridded snow products at variable resolutions, which is usually acknowledged but not addressed analytically. So these results are very interesting...

Line 291: "This suggests that the H-TESEL land model used in both ERA5 and ERA5-Land tends to systematically overestimate SD, most likely due to an excessive snowfall, when no data is assimilated (ERA5 before 1979, ERA5 above 1500 m, ERA5-Land)." I find the messaging in this sentence to be confusing. If the overestimation is related to H-TESEL, this implies that uncertainty in snow parameterizations in the model lead to overestimation of snow depth, but then the problem of excessive snowfall is mentioned.

Does this not imply that precipitation bias is the source of the positive SD bias as opposed to the land model?

Section 3.2.2: The bias trend in the NOAA CDR in fall is an important finding, and corroborates previous work which found similar issues with this product in this season. This is important because numerous studies continue to cite a positive trend in October snow extent over Eurasia, despite increasing multiple lines of evidence (this study provides a new line of evidence) which outline inhomogeneity in the NOAA CDR. I found lines 335-340 to be somewhat confusing, and suggest this text be edited for clarity. The study of Mudryk et al (2017) could also be considered, which showed (1) the NOAA CDR trends in October and November are non-physical and not consistent with other datasets, and (2) NOAA CDR trends in spring are stronger than other datasets. (Mudryk, L., P. Kushner, C. Derksen, and C. Thackeray. 2017. Snow cover response to temperature in observational and climate model ensembles. *Geophysical Research Letters*. 44, doi:10.1002/2016GL071789.)

Section 3.3: "Both ERA5 and ERA5-land use a threshold (5 cm) larger than the one applied to the stations (2.5 cm)..." In reading Section 2.4, I was wondering about the impact of these different thresholds on the validation analysis. I understand the decision to apply 2.5 cm to the snow depth measurements because this is supported by Figure 2 and is consistent with Hori et al (2017). But calculating bias with slightly different thresholds to convert SD to SCD seems problematic. Can you report on any sensitivity analysis which determines how the bias calculations are related to the choice of threshold as applied to the snow depth measurements?

Section 3.5: I suggest moving this into Section 4, because it is largely discussion material and does not present new analysis.

Conclusions: The key result with respect to ERA5 is clearly stated on line 460: "In the reanalysis, data assimilation creates a trade-off between accuracy and stability." For applications like NWP, the instantaneous best estimate is the highest priority, but this of course does not ensure the temporal consistency required for climate monitoring. The key result for the NOAA CDR is communicated less clearly: "Overall, most of the trends/discontinuities observed are larger than the actual snow trends and the GCOS stability requirements, making these products inappropriate for climate applications without correction, particularly ERA5." I suggest re-phrasing this to provide an assessment more clearly focused on the NOAA CDR. This study provides a new line of evidence that autumn trends are very problematic in this dataset, but there are seasons and regions in which the product is suitable for climate analysis (e.g. MAM as shown in Figure 10b).

Section 4 could also highlight that studies continue to claim there is a positive trend in autumn snow extent based solely on the NOAA CDR (<https://doi.org/10.1126/science.abi9167>) and do not acknowledge the literature which has identified problems with this dataset, so your study once again points out that this dataset is problematic in the autumn.

Minor Comments

Line 18: change 'snow cover decrease is aggravated' to 'snow cover decrease is coincident to decreasing snow depth...'

Line 30: not clear what is meant by 'global circulation'.

Paragraph 1 of the Introduction: The Stocker et al (2013) reference for snow trends and snow-albedo feedback is a little out of date. Updated SAF estimates are in the IPCC SROCC Chapter 3, and the Thackeray et al (2019) paper provides a fairly current review. (Meredith, M., M. Sommerkorn, S. Cassotta, C. Derksen, A. Ekaykin, A. Hollowed, G. Kofinas, A. Mackintosh, J. Melbourne-Thomas, M.M.C. Muelbert, G. Ottersen, H. Pritchard, and E.A.G. Schuur, 2019: Polar Regions. In: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)].)

Line 33: suggest changing to '...such as the Arctic and high elevations.'

Line 33: "Notably, only 11 long-term stations are available in the Southern Hemisphere (SH)." Very interesting! Is there a reference for this statement?

Line 46: Is there a reference for the S-NPP VIIRS dataset, as is provided for the others in this list?

Line 50: This is a very minor point, but the most recent citation for the GlobSnow dataset (v3) is: Luojus, K., J. Pulliainen, M. Takala, J. Lemmetyinen, C. Mortimer, C. Derksen, L. Mudryk, M. Moisander, P. Venäläinen, M. Hiltunen, J. Ikonen, T. Smolander, J. Cohen, M. Salminen, K. Veijola, and J. Norberg. 2021. GlobSnow v3.0 Northern Hemisphere snow water equivalent dataset. Scientific Data. doi: 10.1038/s41597-021-00939-2.

Line 87: "Since 2004, ERA5 also assimilates the IMS product but only over altitudes below 1500 m." Could add a reference to the Orsolini et al (2017) paper here.

Line 101: "...but snow observations are not directly assimilated." This is a small point but make clear that both the in situ snow depth and the IMS data are not assimilated into

ERA5-land.

Line 114: Some older citations could be added to provide readers with more background on the CDR and IMS: Robinson, D., K. Dewey, and R. Heim. 1993. Global snow cover monitoring: an update. *Bulletin of the American Meteorological Society*. 74(9): 1689-1696. Helfrich, S., D. McNamara, B. Ramsay, T. Baldwin, and T. Kasheta. 2007. Enhancements to, and forthcoming developments in the Interactive Multisensor Snow and Ice Mapping System (IMS). *Hydrological Processes*. 21: 1576-1586.

Line 120: remove 'around'

Line 141: typo 'sires'

Line 223: "...stations are located either on peaks (Fig. 3b) or in the valley..." This wording is quite specific. Perhaps just emphasize that elevation gradients around the stations create uncertainty?

Line 267: I would not refer to the NOAA CDR as having a "retrieval algorithm" since it is analyst-derived. How about "The positive bias is explained by changes in the analysis approach to produce the snow charts, which since 1999..."

Line 270: can you provide a reference to the NOAA CDR product manual?

Line 273: "...but a positive trend is observed since 1990 in fall and winter." Add a reference to Figure A2 here.

Line 285: The study of Mortimer et al (2020) focuses on the ERA5 discontinuity in 2004, not 1980. (please double check the other citations)

Line 324: Instead of Derksen, 2014, could cite Brown and Derksen (2013) here.

Line 327: change 'algorithm' to 'analysts'

Line 392: "In regions such as Europe, spring SCD reductions add up to the decreasing SD,

increasing, even more, the annual SCD trends.” Awkward wording. I think the point is that in Europe both SD and SCD are decreasing, with the trend towards shallow snow depth amplifying the shorter SCD. In Russia, the snow cover season is shortening, despite positive SD trends in some areas, which means the spring melt signal driven by warming temperatures overrides any increase in snow accumulation during the winter.

Line 453: I very much appreciate the comment that while multi-product ensembles are preferred for historical trend analysis, it is still important to quantify the performance of individual products over time.