Responses to Reviewer #2

This manuscript looks to understand sources of uncertainty in monthly seasonal snow cover trends, particularly that from selection of dataset, time series length, choice of reference period, and dataset adjustment. A motivating factor is to have snow cover trends be better characterized for climate assessments such as the Arctic Report Card and IPCC reports. The results demonstrate how dataset choice impacts snow cover trends, and I think a study like this is important, particularly to have more robust results in climate assessments. However, some edits are needed before this manuscript is ready to be accepted.

Thanks for your careful review, and apologies for the editorial errors. We have cleaned these up and responded to all comments as outlined below, and made a number of other minor changes to improve the paper which are noted in the track changes version of the revised manuscript. All line references refer to lines in the version with changes tracked.

Line 29: You mention that unlike sea ice, snow disappears in the Arctic every summer. You should clarify that snow permanent snow/ice remains in glaciers. Further, how are glacier areas handled in this study? I did not see that mentioned.

We have clarified the wording on line 29 to state:
"Unlike the proportion of sea ice cover that presently survives the summer melt season (at least for the immediate future), terrestrial snow (outside of glaciers and sporadic snow patches) disappears completely from the Arctic land surface every summer."

In Section 2.1, we added new text to describe how permanent snow and ice areas were treated: "Grid cells with perpetual snow cover were retained in the analysis. While this would have an appreciable effect on snow mass analysis, there is a minimal effect on snow extent at the pan-Arctic scale." (lines 116-118)

Line 45: Does the IPCC assessment break down the confidence level by month or season? If so, are all months of changes to Arctic snow cover extent 'high confidence' or is there variability throughout the year?

At the time of initial drafting we were focused on the IPCC SROCC, but we have updated the text to reflect statements in the Sixth Assessment Report (Working Group 1), focusing on why trends at the hemisphere scale have been assessed with 'very high confidence' but Arctic trends in SROCC were assessed only with 'high confidence'.
Revised text:
"Both the 5th (Vaughn et al., 2015) and 6th (Gulev et al., 2021) IPCC Assessment Reports assigned 'very high confidence' to observed spring snow extent reductions at the hemispheric scale (IPCC calibrated uncertainty language is described in Mastrandrea et al., 2010). The SROCC SPM assessment, however, statement assigned only 'high confidence' to observed changes to Arctic snow cover extent (IPCC, 2019). Inconsistent autumn trends due to the choice of snow dataset (e.g. Brown and Derksen, 2013) and large inter-product differences in spring snow extent magnitude and trends (e.g. Brown et al., 2010) precluded the attribution of 'very high confidence'." (lines 45-51).

New citations:
Vaughan, D.G., J.C. Comiso, I. Allison, J. Carrasco, G. Kaser, R. Kwok, P. Mote, T. Murray, F. Paul, J. Ren, E. Rignot, O. Solomina, K. Steffen and T. Zhang, 2013: Observations: Cryosphere. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth

Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Gulev, S.K., P.W. Thorne, J. Ahn, F.J. Dentener, C.M. Domingues, S. Gerland, D. Gong, D.S. Kaufman, H.C. Nnamchi, J. Quaas, J.A. Rivera, S. Sathyendranath, S.L. Smith, B. Trewin, K. von Schuckmann, and R.S. Vose, 2021: Changing State of the Climate System. In Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 287–422, doi:10.1017/9781009157896.004.

Line 67: Please add references. Many studies/reviews have given examples of why/how snow is so difficult to measure, model, remotely sense.

Good point – we added references to this statement, which now reads: "Snow on the ground is a deceptively challenging variable to measure (Brown et al., 2021), model (Rutter et al., 2009), and remotely sense (Frei and Lee, 2010)." (lines 70-71)

Brown, R., Smith, C., Derksen, C., and Mudryk, L.: Canadian in situ snow cover trends 1955-2017 including an assessment of the impact of automation, Atmos.-Ocean, 59, doi: 10.1080/07055900.2021.1911781, 2021.

Frei, A., and Lee, S: A comparison of optical-band based snow extent products during spring over North America, Remote Sens. Environ. 114, 1940–1948, doi: 10.1016/j.rse.2010.03.015, 2010.

Rutter, N, and 50 others: Evaluation of forest snow processes models (SnowMIP2), J. Geophys. Res., 114, D06111, doi:10.1029/2008JD011063, 2009.

Lines 95-99: It's unclear to me here which (both?) version of the NOAA-CDR you are using, whether the 180 km or the 24 km one.

We use both versions as summarized in Table 1. The "NOAA-CDR" is used in the family of 'older' products. The "Rutgers 24 km" data is in the family of 'newer' products. We have better distinguished the names used to refer to them in the text and clarified the wording in lines 100-109:

"(5) the historical snow chart products (NOAA-CDR and Rutgers 24 km). The NOAA-CDR is a 190 km resolution binary snow cover product (Estilow et al., 2015) based on historical snow charts. These snow charts were originally produced by analysts who relied primarily on visible satellite imagery. Starting in 1998 their production switched to the 24 km Interactive Mapping System (IMS). The original NOAA-CDR was produced by digitizing the historical charts at 190 km resolution and merging them with the IMS charts coarsened to the same 190 km grid spacing. Recently, the pre-IMS snow charts from late 1980 onwards have been re-digitized at 24 km resolution to match that of the IMS charts. This process enabled production of a revised snow cover product with improved grid spacing (Robinson and Estilow, 2021). We denote the recent product as Rutgers 24 km and use it for our primary analysis over the 1981-2020 period. We denote the original product as NOAA-CDR and only use it when necessary to extent the time series to the 1967-1980 period (Figures 1-3) or when comparing older and updated products (Figures 7-8)."

Line 90: I believe the GlobSnow products (not sure about CCI) mask out alpine areas. Is that correct? Did you mask out alpine areas from the other datasets, too? If not, how much land is included in MERRA2, for example, that is excluded from the GlobSnow products? Further, it could be helpful to have an additional figure that maps what land areas are included in the analysis, assuming alpine and/or glacier areas are excluded.

We have added a description of mountain masking in the GlobSnow and Snow CCI datasets starting on line 99:
"Both the GlobSnow and Snow CCI datasets mask out complex terrain so snow mass was replaced in the mountain grid cells using the mean value from datasets 1 – 3 above (consistent with Mudryk et al., 2020)."

Line 103: Did you test the sensitivity of the selected 5 mm SWE threshold for determining snow extent? If so, please include a few statements on that. If not, was there a reason you selected 5 mm?

Background on the calculation of this threshold is provided in Krinner et al., 2018. We have clarified this in the text (line 113) and added the citation.
"Gridded daily snow cover is calculated for datasets 1-4 by applying a 5 mm SWE threshold to determine snow extent (background on the derivation of this threshold is provided in Krinner et al. 2018)."

Krinner, G., and 55 others: ESM-SnowMIP: Assessing snow models and quantifying snow-related climate feedbacks, Geophys, Model Dev. 11, 5027–5049, DOI: 10.5194/gmd-11-5027-2018, 2019.

Table 1: Do any of the Snow Cover Extent products have a SWE/depth limit for when they consider a grid cell to be snow covered? If so, please include a mention and reference to that.

The two snow extent datasets are the NOAA snow charts and JAXA JASMES. None of these datasets considers snow depth, rather they map snow extent directly.

Figure 3: Any thoughts for why the JASMES product has much more negative October SCE change than the other products?

The strong negative October trend in the JASMES datasets is consistent with Hori et al., 2017. This dataset has the highest October snow extent during the 1980s, with declines starting in the 1990s. These changes prior to the 21st century explain why the trend is already negative for the initial trend period ending in spanning 1980-2000 shown in Figure 3c. We think a full explanation is too much detail for this section, but we added a citation to Hori et al. (2017) to line 173.

Line 233: Dataset Adjustments is missing its subsession number/heading

The section number was added to make this "3.4 Dataset Adjustments" (line 249).

Line 274: Figure 5 makes it seem like June's average impact of reference year is below 3%, maybe only 2%. Am I reading the figure incorrectly?

You are correct, 2% more accurately reflects what is shown in Figure 5 and so this change was made (line 292).

Line 274: -1 should be an exponent here

This typo was corrected (line 292).

Line 276: For some datasets, there is an effect in the fall months. Consider saying that there is no net effect on average for all the products.

We have revised this text to reflect that changes are evident in both June and September: "The most recent reference periods considered (1991-2010; 2001-2017; see Table 2) have less snow extent than the reference periods which include 1981-2000. When expressed as a % change relative to these different baselines, the rates appear weaker stronger. The magnitude of this effect is small, but does reach 2% decade^{-1} in June and September." (lines 289-292)

Figure 5: The y axis indicates the opposite that is mentioned in the caption. It's labeled that positive values mean a weaker change and negative values mean a stronger change. Can you clarify which one it is?

As noted in our response to Reviewer 1, the caption to Figure 5 was correct but the descriptive text on the y-axis of was reversed. We have corrected Figure 5.

Line 364: Could you include a figure or more text to describe what could be missed by only considering monthly trends instead of daily trends?

We have added extra text to explain this more clearly (lines 380-388): "The analysis of data averaged to calendar months is common (e.g. the widely cited change in September sea ice extent), but of course this creates arbitrary and non-physical temporal boundaries to the data. The snow products assessed in this study are all available at a daily time step. The most dynamic period of Arctic snow extent change spans late May into early June (snow line retreat), and late October into early November (snow line advance). Because these snow extent changes are rapid, temporal precision is lost through the use of monthly averaged data. For example, when melt occurs during the last week of May and first week of June, the monthly averages will still be largely influenced by weeks with stable snow cover. At the very least, analysis should shift to weekly averaged data which would better capture year to year variability in the timing of snow extent changes."

Line 384: Are the Brown data products available to download?

The Brown products will be available for download from an open data archive soon, but at this point interested users need to contact us. We have now made this clear in the Data Availability text (line 414).