Reviewer 3

I thank the reviewer for their constructive comments and especially for (i) the suggestion of the key references that I was unaware of and (ii) the suggestion of focusing on the quality of the 20CRv3 data, which I would like to meet 'half-way' in a revised paper (see final comment).

GENERAL REMARKS

In the manuscript "Spatial and temporal changes in autumn Eurasian snow cover and its relationship with the Arctic Oscillation" the author uses long-term reanalysis data to subdivide the correlations between autumn snow cover and winter AO sign regionally. This topic is of great importance and fits the scope of the journal.

That said, I am really torn about the quality of this manuscript. I really want to love it, but I find myself wondering about the goal and the impact.

In general, I agree with the previously published reviews that the results are not substantially new and it is difficult for me to see the key point of the exercise.

However, I can see many positive and useful parts in this paper. I think it would be very helpful for the community to have an updated evaluation of SD in 20CRv3, versus the older products and stations, kind of like an update of the cited Wegmann et al 2017 study with the nice bias removal procedure shown here. And then that can be looked at regionally and you can discuss about why and how 20CRv3 might be improved or not, you could go into more detail for other seasons etc. I think this would be of great value for this community. So basically everything until 4.2 I like.

With 4.2 the science becomes unfocused and diffuse. In the end you have a huge array of mediocre findings in a very complicated form. All of this regionalisation could be, in my opinion, be performed by EOF analysis of the whole field, similar to what Han & Sun did in 2019 (see the link below). You mention and show that there are decadal changes in correlation, but offer no process that might explain that. You remain too superficial for novelty in my eyes. Your KEY result in my eyes, after filtering a lot of text and data for meaning, is the west east anti-correlation pattern between autumn SC and AO. Unfortunately, Han & Sun and Wegmann et al. looked into this already to some degree and you do not put your results in the context of those studies at any point. This is an important oversight.

So what is the goal of this manuscript? Is it to find the most impactful regions for seasonal prediction? Is it to find an explanation for the decadal variability? Is it to evaluate snow in 20CRv3? Is it to do something different, maybe in terms of method, season and region that previous studies did? The goal and thus the evaluation if you reached this goal stays elusive. One major difference to previous

studies is the use of 20CRv3, but again, you did not really show the reader how different that product is from 20CRv2c and 20CRv2.

So my recommendation is that you sharpen your manuscript (agreeing with the two other Reviews) and pinpoint towards your goal. I personally would recommend the goal to be SD evaluation rather than looking at dynamics, simply because arguing for novelty in the dynamics part seems difficult from my side of things.

A few notes on the overall structure:

I think the discussion and conclusion sections are misused.

The discussion section in this manuscript is actually just another result section.

I disagree with this statement as I think the discussion can be a place to expand on the key results. However, following the updated methodology and the new references the discussion will be rewritten with an emphasis on comparing previous work, especially the SC indices of Han & Sun (2018) and the subsequent paper by Wegmann et al. (2020), as updated using the 20CRv3 data over the longer time period (see Table HS below).

Table HS. Proportion of decades that the Han & Sun (2018) SC indices derived from 20CRv3 have positive and negative correlations with the 20CRv3 AO, and the proportion that are statistically significant based on a one-tailed test ($p < 0.10$): *** is $p < 0.01$, ** is $p < 0.05$, * is $p < 0.10$. The period covered is 1831-40 to 2006-2015.

Snow Cover Index	Positive (%)	Negative (%)	Pos. (significant) (%)	Neg. (significant) (%)
Oct_SNOWI	62.36	37.64	12.94	11.76
Nov_SNOWI	31.18	68.82	5.88	32.94***
Nov_SNOWI (west)	75.29	24.71	32.94***	4.12*
Nov_SNOWI (east)	34.71	65.29	7.65	18.72

Key points from this table:

1. The proportion of decades with a significant Oct_SNOWI-AO relationship is not statistically significant at all.
2. The proportion of decades with a significant negative Nov_SNOWI-AO relationship is statistically significantly high ($p < 0.01$) but not significantly low for a positive relationship.
3. Dividing the November dipole into its two components, it appears that only the western area is statistically related to the AO, which matches the new findings of this work.
4. The SAI_NOV_SW in the new analysis (cf. new Fig. 8 and Table_new below) provides a slightly better predictor for the winter AO than the Han & Sun Nov_SNOWI in the 20CRv3 data in that (i) the proportion of of significant decadal SC-AO correlations is statistically significantly high for one sign of the relationship and significantly low for the other, indicating that the relationship is more temporally stable and (ii) the frequency of time when there is a significant relationship is higher.

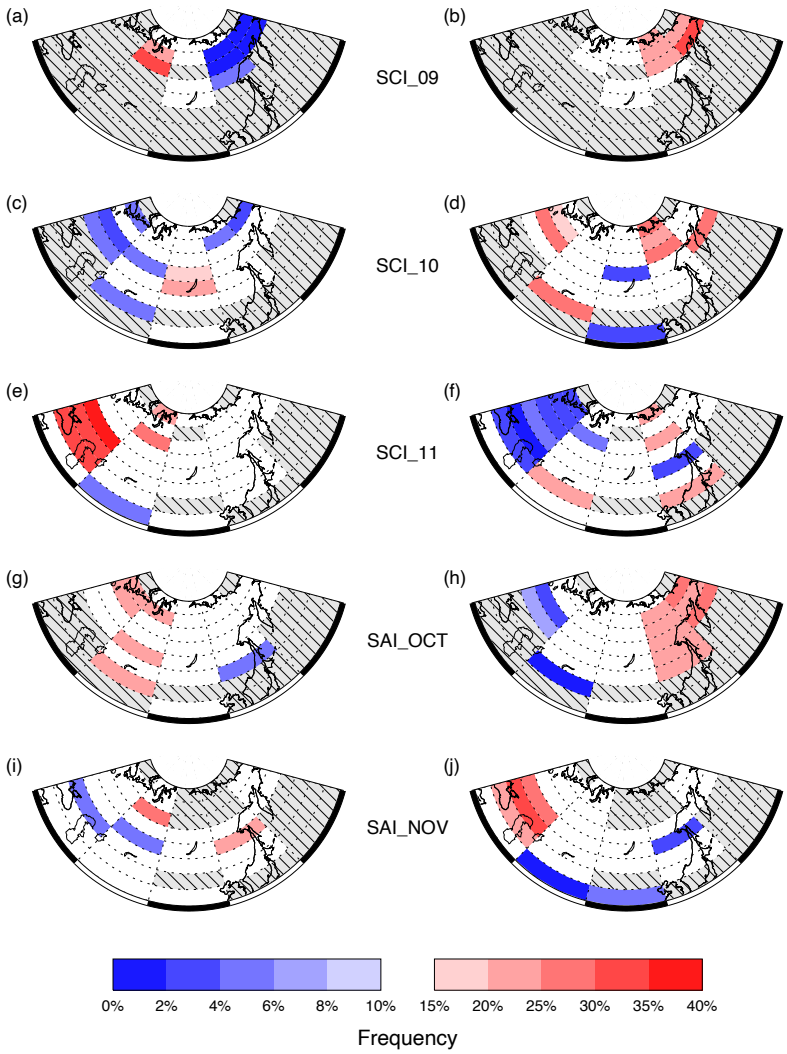

Figure 8. The frequency of decades with statistically significant correlations with the AO ($p < 0.10$) from the 20CRv3 data (1831-1840 to 2005-2014) for the five SC indices. Frequencies for positive correlations are shown on the left in red and negative correlations on the right in blue. Hatched subregions are where no value was calculated because data availability was less than 50%.

Table_new. Proportion of decades SC indices derived from 20CRv3 have positive and negative correlations with the 20CRv3 AO, and the proportion that are statistically significant based on a one-tailed test ($p < 0.10$): *** is $p < 0.01$, ** is $p < 0.05$, * is $p < 0.10$. The period covered is 1831-40 to 2006-2015.

Snow Cover Index	Positive (%)	Negative (%)	Pos. (significant) (%)	Neg. (significant) (%)
SCI_09_NE	26.47	73.53	1.76**	27.64***
SCI_11_SW	76.47	23.53	38.34***	4.12*

And the conclusion is the discussion. However, some key context is missing here.

The conclusion is missing.

Again, I disagree with this statement: while the second half of the conclusion will be changed to reflect the new results, I see no problem with the first sections that briefly outline the methodology used and summarise the principal results regarding temporal changes in the different SC indices.

I was constantly confused by the direction of the correlation. For me more snow to negative AO should be a negative correlation.

As mentioned in the text (line 160) it is only the SAI indices where the sign of the correlation is reversed, and this follows from the work of Cohen & Jones (2011) and Peings et al. (2013). It was originally done in Cohen & Jones (2011) so that an SAI time series could be easily compared to a winter AO time series: there was a strong negative correlation between the rate of snow advance and the subsequent winter AO in the period encompassed by the earlier work (cf. their Fig. 1). The SCI indices are such that more snow leading to a negative AO is indeed a negative correlation. I have kept the same definition as the previous literature but would be happy to change the sign of the SAI indices if the editor thinks it appropriate.

SPECIFIC REMARKS

L12: You point out the novelty of going back to 1836, but here you only mention the last 50 years. I am missing some coherence between the two statements.

Depending on what other changes will be made to the Abstract, I will change the start of the sentence to, “The greatest changes in the snow advance indices (SAI) occur over the past 50 years, in which there has been a slowing …”

L34: "There is often" is too imprecise. When does this dipole happen in the season, when was it strong and why and what is the direction of the gradient? Is it usually a lot of snow in the west and no snow in the east? Does the sign flip interannually? Decadally?

With the addition of the new references, this sentence can be expanded. “There is often an east-west dipole in SC anomalies across Eurasia during November with one pole in eastern Europe and the opposite pole in southeast Siberia and northern Mongolia (e.g., Gastineau et al., 2017; Han and Sun, 2018; Zhang et al., 2023). Typically, there is a marked longitudinal SC gradient with the strength of this varying at multidecadal frequencies, perhaps associated with North Atlantic SST variability (Wegmann et al., 2020).

L45: Upward trend of what? Do you mean positive trend?

I have changed the sentence to, “… the upward trend in CDR-derived SC from 1992 …”

L47: When you say "in reality", what data product do you refer to, if we assume that the CDR is "not reality"?

I have changed the sentence to, “In reality, observations indicate that autumn SC …”

L57: First you mention that the Arctic air is "warmer", then you mention that the cyclones bring "Cold" Arctic air. This is confusing.

I have rewritten the sentence as, “In southern Siberia (50-60°N) these cyclones transport cold air into the region from the north or east and induce negative temperature anomalies that allow SC to persist (Bednorz and Wibig, 2017).

L78: I would summerize Arctic amplification and Climate Change to anthropogenic global warming.

I disagree, as they are not solely a response to anthropogenic activity: e.g., Zhou et al. (2024) state that, “natural variability has substantially modulated the degree of Arctic amplification.”

Zhou, W., Leung, L. R., and Lu, J.: Steady threefold Arctic amplification of externally forced warming masked by natural variability, Nat. Geosci., 17, 508–515, https://doi.org/10.1038/s41561-024-01441-1, 2024.

L108: In this context, Wegmann et al 2021 investigated an interesting middle ground, where free running seasonal forecast models were induced with different snow regimes (https://wcd.copernicus.org/articles/2/1245/2021/). They found a small but measurable impact onto the NAO sign.

I will add a new sentence to the revised manuscript, “In the ‘middle ground’, Wegmann et al. (2021) compared free-running models with different initial SC regimes and demonstrated a small but measurable impact on the winter AO.”

L125: Building up on Peings and Douville, Wegmann et al 2020 looked at the link between Eurasian snow and NAO and found that the snow impact is highest during Arctic warm periods, when the Albedo impact of additional snow is strong (https://esd.copernicus.org/articles/11/509/2020/).

I will add an additional sentence at the end of the paragraph, “In contrast, Wegmann et al. (2020) determined that a negative AO signal following a strong November west-to-east SC gradient was valid throughout the last 150 years. Moreover, the strongest correlations occurred during periods of high November snow variability coincident with strong Arctic warming at the beginning and end of the 20th Century.” This is something that I aim to analyse further in the new Discussion utilising the longer 20CRv3 data.

L290: I think it is worth mentioning in this manuscript, that the grid size of ERA5 is substantially smaller than than of 20CRv2c, so we would expect better agreement between a station and ERA5 vs a station and 20CRv2c.

The sentence will be rewritten as, “SD in earlier versions of 20CR is known to be biased high (Wegmann et al., 2017) and, in addition, the coarser spatial resolution of 20CR means we might expect poorer agreement than in ERA5.”

Also you assume here that the errors of 20CRv2c are carried over to 20CRv3. Until now in the manuscript, you did not show the reader that that is indeed the case. Maybe that would be a worthwhile information for the community to have.

The pair of figures below, which are a direct comparison with snow depth from observations, suggest that the high bias in 20CRv2c is indeed carried through to 20CRv3. I suggest that I make this point and reference these figures in the supplementary material.

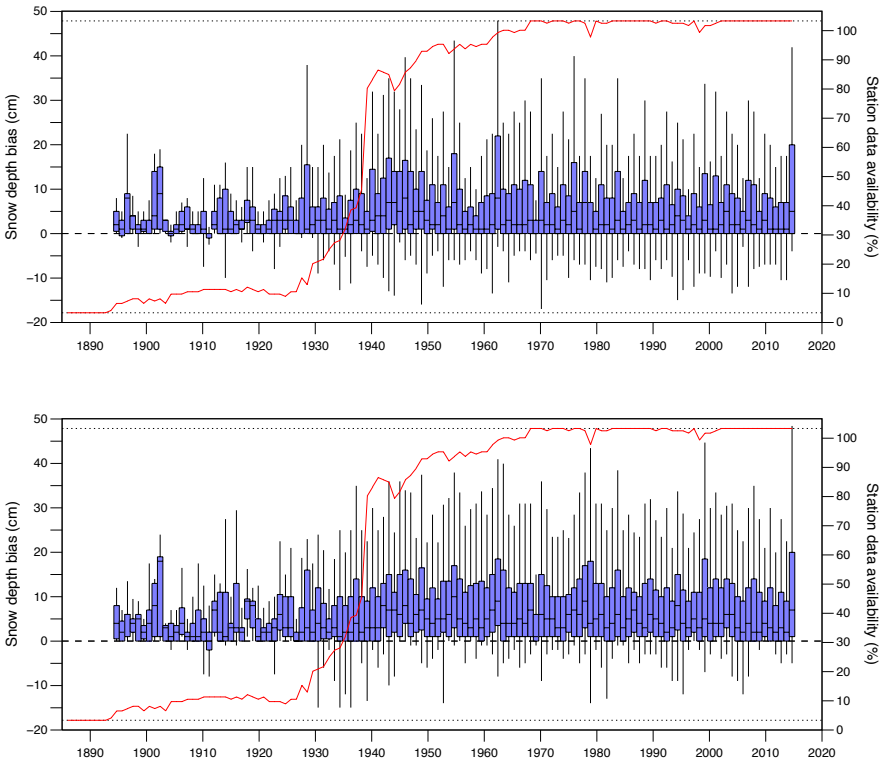

Figure. Bias of the median October-November snow depth in 20CRv2c (top) and 20CRv3 (bottom) from 125 stations

L694 to end: I think your manuscript would really benefit from discussing the work of Han & Sun 2019 (https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018JD028443), Wegmann et al 2020 and Wegmann et al 2021 regarding this west-east difference in snow impact. In there you will also find some initial explanations for what could be the reason for this behavior. In there you will also find some answers to your last sentence in this manuscript.

I agree completely with this and it’s clearly a significant oversight of mine to not include these highly relevant references: I have updated the scope of my journal alerts! Moreover, the previous work of

Han & Sun (2018) and Wegmann et al. (2020) negate some of the novel aspects of the work that I thought I was undertaking. Bearing this in mind, and the comments of Reviewer 3, I propose changing the emphasis of the paper to include an examination of how 20CRv3 compares against the previous version of 20CR (and possibly other datasets) when calculating changes in the Eurasian SC indices and their relationship with the AO. The latter component will include an analysis of the Han & Sun SC indices with the new dataset (see above) and a comparison of their temporal variability with the work of Wegmann et al. (2020). For example, an equivalent plot to their Fig. 7b using 20CRv3 is shown below (Figure W). Comparing the two plots indicates a broad similarity, with statistically significant negative SC-AO relationships in the 1920s and the present. Differences include some short periods of positive correlations in the 20CRv3 plot that are not present in 20CRv2c. Moreover, being able to extend the time series back further using 20CRv3 reveals that periods of positive correlation were actually dominant during the mid 19th Century, thus indicating that the negative Han & Sun (2018) November dipole-AO relationship is less temporally invariant than previously thought.

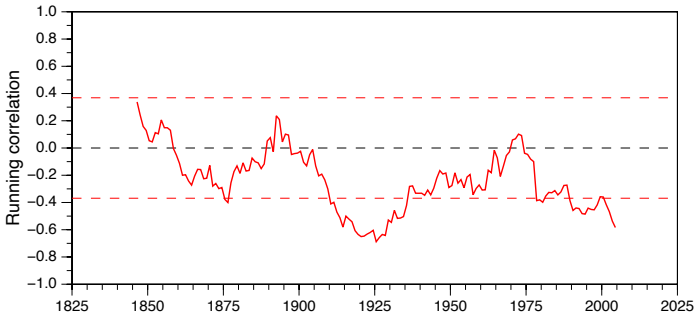

Figure W. Running 21-year correlations between the Han & Sun (2018) November dipole index and winter AO. The dashed horizontal lines of the same colour indicate significance levels at $p < 0.10$ for a one-tailed test.

There is also the 'new' result regarding the statistically significant negative SCI_09-AO relationship in northeast Eurasia (cf. Table_new above), which appears robust, both in terms of having a significantly high frequency of decades with a significant negative SC-AO correlation and a significantly low frequency of decades with a significant positive SC-AO correlation, and which demonstrates some fairly regular decadal variability (see Figure New below) that can be examined further in comparison to potential external forcing factors.

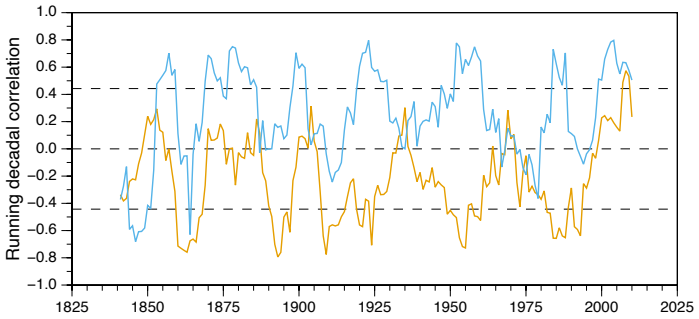

Figure New. Running decadal correlations between the SCI_09_NE and SCI_11_SW and winter AO. The dashed horizontal lines of the same colour indicate significance levels at $p < 0.10$ for a one-tailed test.

A new title for the paper might be:

An examination of changes in autumn Eurasian snow cover indices and the temporal stability of their relationships with the Arctic Oscillation using adjusted 20th Century Reanalysis version 3 snow depth data.