Reply to Referee #2. Author comments in blue italics.

Anonymous Referee #2

Received and published: 11 January 2020

The manuscript by Sjolte et al. investigates a new innovation in the rapidly developing field of paleoclimate data assimilation. Specifically, they investigate the potential of reconstructing seasonal fields using subannually resolved ice-core (and to a lesser extent, tree ring) data from the North Atlantic. The manuscript is well-written, well-illustrated and generally well-organized, and the results are interesting, and suited to Climate of the Past. I do however, have a few concerns and and suggested additions to the manuscript that I'd like to see addressed.

- *We thank the reviewer for the positive comments and interest in out manuscript, as well as the detailed comments which helped us greatly improve the manuscript.*

Major Issues

In general, in my opinion, the primary weakness of the manuscript is that the exploration of the reanalyses is rather limited. For example, in the authors subdivide the ice cores into a group of 8 that extends from 1241-1970 and a larger group that is shorter (1777-1970). However the reconstructions are only analyzed in the context of instrumental data. No results from prior to 1850 are shown in the manuscript or supplement, except for figure S3, which is specifically focused on the tree ring sites. In evaluating this technical approach, it is important for readers to be able to see how the longer term variability compares to other reconstructions from the region, and to consider and discuss how the seasonal assimilations affect long-term variability, and the potential climatic implications of that. Given that this approach creates a field reconstruction; these results could be compared to regional temperature reconstructions, NAO reconstructions, and more, and give the readers a better sense of how this approach compares with previous efforts.

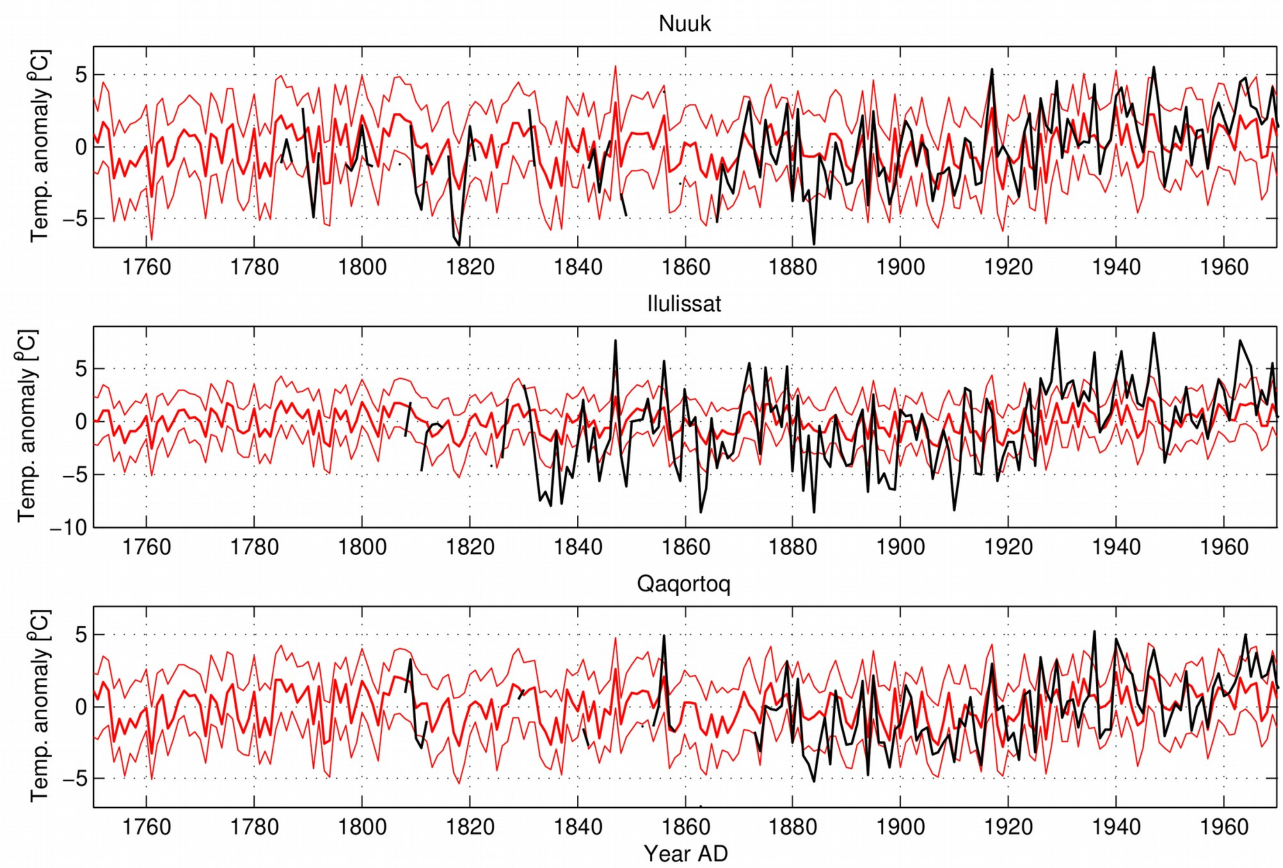
- *We agree that an in-depth comparison to other reconstructions of the longer term variability would be very interesting. For the DJF NAO reconstruction covering 1241-1970 this is already done by Sjolte et al. (2018). The purpose of this paper is to test detailed aspects of seasonal variability and which factors affect the skill of seasonal reconstructions. These tests are only possible with observation-based data where we have full control on the time scale and seasonality. Furthermore, a full-scale comparison to the long term variability of other reconstructions is a whole study in itself, and this manuscript is already crowded by many figures and results. As also pointed out by Referee #1 the scope of this study should be better defined. We will do so in the revised manuscript.*

At present, the evaluation of the results is restricted to spatial comparison of the first three PCs with instrumental data, temporal comparison of the same thing. I was glad to see SSTs averaged and compared to instrumental data, but feel like the comparison was ultimately very limited.

- *Please note that we also compare to the station-based NAO (Figure 7c and 7d), as well as temperature data from Greenland and Iceland (Table 4). Again, the motivation for using only observation-based data is that this is the only data where there is no uncertainly with respect to sub-annual temporal resolution. See also reply above.*

The other major weakness of the manuscript, that I believe should be able to address, was the representation of uncertainty. The methodological approach to uncertainty quantification; an ensemble based approach, is reasonable. I was disappointed however that the results were not presented in the manuscript. Every figure in the manuscript, except for the first two, could, and should, have uncertainty ranges (like 95

- *Correlation maps (e.g. Figure 3 and 4) are done for the ensemble mean field. The correlation between individual ensemble members is relatively low, and the weather/climate signal only really emerges in the ensemble mean. Also, including this spread in a map would make the figure very hard to read, and other important features are lost.*
- *For the maps of the EOF patters (e.g Figure 5) the 2nd and 3rd EOFs also only emerge clearly for the ensemble mean data. For this reason, including the ensemble spread for the time series of the PCs (Figure 6) is not really feasible. And would also make the already busy figures hard to read.*
- *We do discuss the uncertainty in Section 5, but we do see the point of the Referee and we will expand this discussion to include the points we mention above, and also include a figure to exemplify the ensemble spread of temperature (see figure below including ensemble mean and spread (red) compared to DJF temperature for Greenland coastal stations (black)).*



Additional issues/notes

I'm a little confused about how the analog matching is working, based on figure 1. Specifically, are any adjustments made to the model-output before calculating the EOFs of δ 18 O? If not, I'm confused about how there is such fine spatial structure in the model δ 18 O, given that it has 3.5 degree grid cells. In fact, I think it would be helpful to see the outlines of the gridcells on the lower half of Figure 1. Maybe there's enough resolution there, but I found it confusing. I'm also pretty surprised about how comparable the modeled and observed δ 18 O EOFS are, they're nearly identical. I'm not particularly familiar with this region and proxy, but model-proxy EOF comparisons this similar are exceptionally rare, unless one was forced/derived from the other, and I'd be interested to learn more about this.

- *There are no adjustments made to the model output before comparing the EOFs in Figure 1. This is the whole point – that we can match the modeled patters to ice core data without tuning/calibration.*
- *The model grid is relatively course, corresponding to ~400 km at the Equator. However, since the grid gets denser at higher latitudes, there are quite a few grid points covering Greenland. To illustrate this, we follow the Referee's advise and add the grid to Figure 1 d-f.*
- *The match of the model to the patterns is indeed very good, and without this the method would not work. The good match means that the average variability of the modeled d18O is realistic on the regional scale. In general ECHAM5-wiso has been found to score very high in model-data comparisons using isotope enabled models (e.g. Steen-Larsen et al. (2017).*

Here's a suggestion that might be beyond the scope of this manuscript, but that I think is interesting: have you considered trying to assimilate different proxies for different seasons, but for the same assimilation? It would be really interesting to see what an annual reconstruction looks like where tree rings were assimilated for summer, while ice cores were simultaneously assimilated for winter – i.e., do the analog matching differently for each season but find the years that match both optimally.

- *This is certainly an interesting question, while it is beyond the scope of this study. It would require quite a lot of testing and also put strong constraints on the seasonal variability of the model and which model analogues that can be chosen. As indicated by Figure 2, there is limited co-variability between the seasons, however there is potentially some additional information on the climate variability to be gained from this approach. In the study by Tardif et al. (*https://doi.org/10.5194/cp-15-1251-2019*) they use seasonal proxies to reconstruct the annual variability, however there is no analysis of seasonal reconstructions in that study, so it is hard to know to which extent they are successful.*

Minor issues:
Line 7. "Reconstructs" should be "Reconstructions"

- *Corrected*

32: 18-O should have the standard superscript formatting

- *Corrected*

85: "extend" should be "extent"

- *Corrected*

328: "depended" should be "dependent"

- *Corrected*

385: "particularly" should be "particular"

- *Corrected*

Figure 7: Add some additional labels to the panels to help differentiate. It took me awhile to figure out why c and d were separated.

- *Correceted*

References.
Steen-Larsen, H. C., C. Risi, M. Werner, K. Yoshimura, and V. Masson-Delmotte (2017), Evaluating the skills of isotope-enabled general circulation models against in situ atmospheric water vapor isotope observations, J. Geophys. Res. Atmos., 122, 246–263, doi:10.1002/2016JD025443.