Below we list point-by-point responses to all reviewer comments. We are grateful for all the comments by the reviewers, which has helped us clarify and improve this manuscript. Our replies are in *italics* and all changes to the text are in *blue*. This reply is very similar in contents to the one posted on the Clim Past Discussion page. We have, however, now included an analysis of the decadal to multi-decadal response to volcanic forcing to better accommodate the comment by Referee #2. In this text figure numbering, pages and lines numbers refer to the original manuscript. A version of the manuscript with tracked changes follows this reply.

Referee #1

The authors in this interesting study fuse measurements of isotopic composition of oxygen with simulated results from an isotope enabled global climate model to reconstruct seal level pressure and temperature changes over the 1241-1970 period. The model gives spatial maps of pressure and temperature for each year, in an ensemble way, which is raked according to minimization of the distance between the simulated and measured PCs of the oxygen isotope. Then the authors infer volcanic and solar influences in the reconstruction. The propoced method is novel and may return some useful information about external forcings over the last millennium. The manuscript is generally well written and figures are presented in a crystal way. In the following, I am listing some thoughts/comments which hopefully will increase clarity of the manuscript.

Thank you for the positive review, and the useful comments which has helped us to clarify our manuscript.

Major comments In my view, this study is more about evaluating the proposed methodology rather than detecting volcanic and solar signals. For example, the length of the text dedicated to volcanic effects in the reconstruction is a single paragraph only (1st paragraph page 7). Given the extended length of the detected response compared to other studies I would strongly encourage authors to provide more details. For example, to what extend results are influenced by the performance of ECHAM to eruptions. How the response looks if picked up actual model years not reconstructed? An additional validation method would be to examine the sensitivity of the results to the model choice. The PMIP3 simulations are ideal to this this purpose but I admit is a lot of additional effort.

While it is true that we spend quite some space in the manuscript on evaluating the method, we believe that this is necessary, since this is the first paper presenting this method. With respect to the response to the volcanic forcing: As written line 2-3, page 5 of the manuscript "The temporal succession of the reshuffled model fits do not resemble the order of the years in the original model [ed: run]...". This is also true for the timing of the volcanic and solar forcing used for the model simulation. This means that the response of the ECHAM5/MPI-OM model to volcanic forcing is not directly relevant for the reconstructed response to volcanic eruptions. We have clarified this in the Methods text. For reference, the response to volcanic forcing of the ECHAM5/MPI-OM model has been studies extensively (e.g. Zanchettin et al. 2012, Guðlaugsdóttir et al., accepted).
The reviewer suggests that our method is tested using the PMIP3 simulation. It would indeed be desirable with such a test, however this would require the PMIP3 model suite first to have isotope diagnostics implemented in the code. Swingedouw et al. (2017) (cited in the Introduction) reviewed the PMIP3 models' NAO response to major tropical eruption of the past millennium of the PMIP3 models and found a large scatter in the models' response, with little coherence with the reconstructed NAO response by Ortega et al. (2015). Our reconstructed response to volcanic forcing is fairly straight forward to describe, and we feel the result it self does not benefit further

from an extended description. However, following the reviewers suggestion we have added more details in the Discussion section concerning these questions.

Inserted, page 5, line 8: "Since the timing of the forcings used for the model simulation is not a factor influencing the reconstruction, it also means that the model performance for the response to forcings does not influence the reconstruction."

Inserted, page 8, line 6: "As shown in the comparison of the response to volcanic forcing between our reconstruction and NAOmc (Figure 3 c), the mean response look qualitatively very similar in the two reconstructions. However, as already mentioned, due to the preserved high frequency variability our reconstruction shows both a more immediate and persistent NAO response to volcanoes. This underlines the importance of producing climate reconstructions that do preserve high frequency variability, in particular if the reconstruction is used as baseline for model evaluation. In our analysis of the reconstructed response to volcanic eruptions we choose eruptions larger than or of similar magnitude as the 1991 Pinatubo eruption (< -6 Wm^2*). As discussed by Swingedouw et al. (2017) climate effects of smaller eruptions can be difficult to detect due to stochastic climate variability. We find that we can detect the an impact on reconstructed NAO from tropical eruptions selected in the range from -4 Wm*^{2}* to -8 Wm*^{2}* (Sigl et al., 2015), yielding a significant positive NAO one year after the eruptions, on average. This appears to be the limit of detection for our reconstruction, possibly owing both to the partly stochastic variability of the NAO and to noise in the reconstruction."*

Added value of reconstruction The added value of the reconstruction is not properly assessed. The reader is left wondering whether similar skills (e.g. correlations) could be obtained by the model simulation alone. The authors should at least mention the how much correlations between the modelled NAO and 20CR or other reconstructions are improved in the reconstruction. Likewise for the modelled PC2.

The variability of the NAO is to a large extent driven by internal variability, although a proportion is driven by external forcing and interaction with the ocean (Hurrell et al., 2003). Running an atmospheric model with SSTs prescribed from observations, as well as observed aerosols and stratospheric forcing yields a correlation to observed NAO of only about 0.25 (Compo et al., 2011). This is the reason for the many efforts to reconstruct the NAO, e.g. as reviewed by Pinto and Raible (2012), and part of the motivation of our study (page 2, line 24-30).
We compare with the two NAO reconstructions by Ortega et al. (2015) as well as the NAO reconstruction Luterbacher et al. (2004) in Table 1. We also compare the skill in relation to observed NAO (Jones et al., 1997) of our NAO reconstruction to that of the skill of Ortega et al. (2015), where our reconstruction shows slightly higher skill, and with much improved year-to-year variability (page 6 , line 27-35). We are no aware of any other published reconstruction of PC2 of SLP that we might compare with.
In the revised manuscript we now highlight also the comparison the NAO reconstruction by Luterbacher et al. (2004), which was not discussed in the text before.

Reformulated, page 6, line 21: "For an independent comparison we use reconstructed NAO and gridded of SLP (Luterbacher et al., 2001), and reconstructed temperature (Luterbacher et al., 2004) over Europe."

Inserted, page 6, line 24: "For the NAO reconstruction by Luterbacher et al. (2001) our reconstruction shows slightly lower correlation on interannnual time-scales compared to the correlation to NAOmc, but similar correlation on decadal to multi-decadal time-scales (Table 1)."

Use of solar forcing There is an extensive description of the TSI forcing used (is there

SSI variability?) in the model which differs from those recommended in the PMIP3. I guess the continuity of the Muscheler reconstruction throughout the last 500 or so years does not immediately make it superior over the other reconstructions. I am wondering however if the choice of TSI really matters, given that the reconstruction does not really benefit by the timing of TSI forcing (authors comment on that in p.5 l6). TSI and other forcings are just increasing the phase space of the surface patterns which may or may not be chosen eventually in the reconstruction. How different the reconstruction would be by feeding in model data from a constant forcing (control) simulation?

The TSI reconstruction is described in detail here since it, as mentioned by the reviewer, differs slightly from the PMIP3 reconstruction, but does however agree well with the updated PMIP4 forcing. Please note that the model run was performed before the PMIP4 forcing was available. This is a TSI forcing only, which ECHAM5 divides in 4 bands (1 visible+UV and 3 near infra red) (Roeckner et al., 2003). As the reviewer points out the exact forcing of the model is not important for the reconstruction. Although we find that the timing of the forcings driving the model simulation are not crucial for the reconstruction, and there is no preference to forced versus unforced model years, the combined effect of all the forcings increases the model variance, which produces a more diverse sampling space for the method to pick from. We have clarified the role of model forcings in the methods section (see reply to previous comment), and added a paragraph in the Discussion on the role of the model run.
However, for the analysis of the reconstructed longterm solar response we also use this TSI reconstruction, and here the timing and amplitude of the solar cycles are important. In this case the timing is crucial and our solar activity reconstruction is based on ^{14}C data that should not contain any dating uncertainties.

Inserted, page 8, line 6: "The model simulation used for our reconstructed translates the climate variability recorded in the Greenland ice cores to climate variability in the North Atlantic region. In the initial test of the isotope variability it is shown, that the spatio-temporal δ^{18}O variability of the ice cores is well represented by the model (Figure S6). This is a fundamental prerequisite which allows us to match the modeled δ^{18}O to the ice core δ^{18}O year-by-year. While the skill of the reconstruction is higher in the vicinity around Greenland, the reconstruction shows significant correlations to reanalysis data wide spread across the North Atlantic region. This skill depends on i) the integrative nature of the δ^{18}O as recorded in the ice cores, and represented by the modeled δ^{18}O ii) the modeled atmospheric teleconnection patterns in terms of temperature and circulation, and iii) how these patterns are connected to modeled δ^{18}O for Greenland. Clearly, the reconstruction is strongly dependent on the climate model when it comes to whether or not it is possible at all to use our method, and when it comes to the skill of reconstructed spatial patterns. The resolution of our model model simulation is relatively course and using a higher resolution simulation could improve the representation of several processes. For example, vapor transport to dry polar regions is often inhibited in models with courser resolution, resulting in too little precipitation in the interior of ice sheets and a positive bias in δ^{18}O (Masson-Delmotte et al., 2008, Sjolte et al., 2011). This is related to cloud parameterizations and course resolution models having difficulties in explicitly representing frontal zones in connection with synoptic weather systems. The orography in course resolution models is also more smooth, loosing orographical features such as the southern dome of the Greenland ice sheet, which also affects atmospheric circulation and small scale spatial variability. In our approach we match the modeled PCs of δ^{18}O, meaning that we are matching regional scale patterns in δ^{18}O, which partly addresses the problem of matching course model output to site specific proxy data. However, having a higher resolution model simulation could for example improve the spatio-temporal representation of Greenland δ^{18}O, allowing more than 3 PCs to be fitted, and generally giving a better representation of temperature, pressure and precipitation in the reconstruction. For reasons discussed above, it would be desirable using different GCMs to test for model dependencies of the reconstruction, as well as testing for added *value of ensemble reconstructions with several different GCMs. Doing these tests is presently limited by the availability of millennium length simulations using isotope enabled GCMs."*

Top-down mechanism is missing (related to the previous comment) I am pretty sure the ECHAM5 version does not properly resolves the so called "top-down" mechanism of SSI forcing on the stratosphere and subsequence changes in the troposphere and surface. To my understanding this is a key component of the proposed solar-NAO link. Given that weakens of the model it is not surprising that little evidence is found between NAO and solar forcing. So, the argument that solar forcing has little effect on NAO is not well justified. Likewise, in the discussion (p.8 last paragraph), the argument about increasing blocking comes hand in hand with the stratospheric response and the "top-down" mechanism, with is missing here. So it is likely that you may get right (to some extent) responses from wrong reasons.

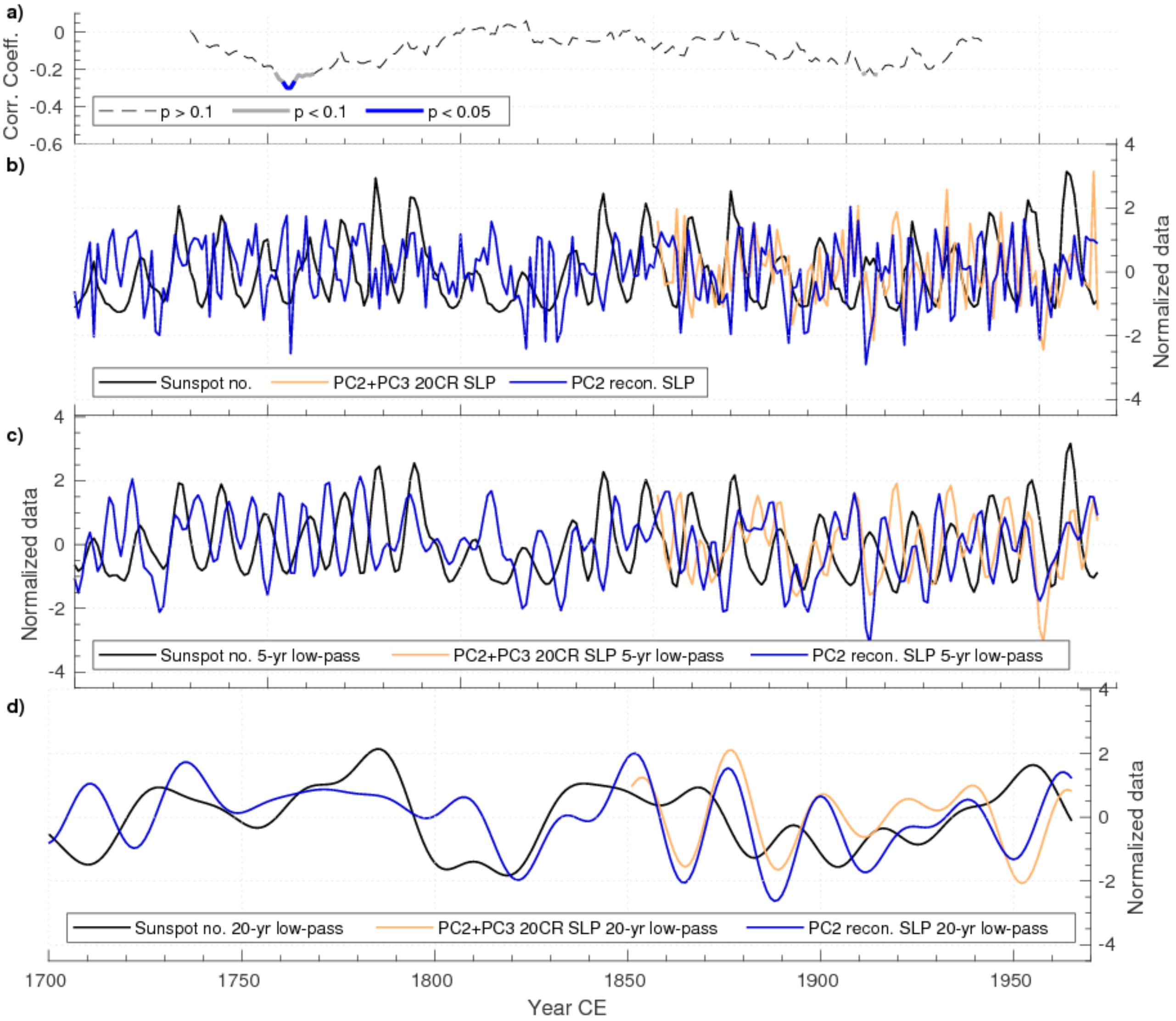
The reviewer is right in that the top-down hypothesis for solar influence of climate has been put forward in several studies, and the solar influence on the stratosphere is well-documented. However, as the model run is reshuffled (based on real-world isotope data), the reconstruction largely depends on the isotope-climate relation in the model and how this is matched to the ice core variability (see also replies to comments above). This means that whether or not the model has a realistic representation of stratosphere-troposphere coupling is less important, provided that the model results give a realistic ensemble of possible tropospheric circulation regimes. Given the limited representation of the stratosphere in the 19-layer version of the ECHAM5 model it is also unlikely that we can investigate a possible top-down mechanism from the reconstruction alone.

Solar responses and time lags The 11-yr solar cycle SLP anomalies in Figure 4 do not seem to support the results of Gray et al., 2013. It is the different period that change signals? Could you please calculate signals from 1850 as to be consistent with the HadSLP2 years? The argument that signals maximize with 5 years lag is strange. 20CR shows almost opposite signals in lag 5 compared to lag 0, an antisymmetry not obvious in the reconstruction (Figure 4). My guess is that the correlation of PC2 with the 11-yr happened to maximize at lag 5. Thinking linearly, I would expect a negative correlation at lead time +1 years or lag 0, but in table 2 I see just zero. Even at lag 5 the explained variance by the 11-yr solar forcing is tiny. Authors should comment on the amazing correlation between reconstructed PC2 and 20CR PC2+PC3 (figure 6c). Why is it so?

As we see it, the response seen in Gray et al. (2013) Figure 1, could be consistent with our results. With no lag, their results show negative correlations across the British Ilse and Scandinavia (although not significant), which is consistent to the high-pressure anomaly we see in the reconstructed response (anomalies for solar min-max, so the anomalies are reverse compared to the regression in Gray et al.). Gray et al. also see a stronger, and reversed, response at 5-year lag, which is significant. There are a number of differences between the study of Gray et al. (2013, 2016) and our study i) Gray et al. uses a regression model, while we pick low vs. high solar activity ii) the time period chosen by Gray et al. is 1870-2010, and iii) Gray et al. uses HadSLP2 data. The analysis of the solar response is sensitive to the method, the time period chosen, and the climate data set. Since our reconstruction only goes to 1970, we cannot perform a consistent comparison to Gray et al., and if we chose a short period, e.g. 1870-1970, instead of 1700-1970, the response is qualitatively similar but weaker and not significant in the reconstruction. The noise level in any reconstruction can be expected to be higher than in observations, and the solar influence in climate is difficult to identify even in observations. So we rely on the long-term average response from solar cycle to solar cycle to average out noise and get a mean response. Comparing the areas of

significant response in Figure 4 and Figure 5, we find the reconstructed response to the 11-yr cycle, with and without time lag to be consistent with the response seen in the 20th Century Reanalysis. The analysis shows strongest correlation with 5-year time lag between PC2 and the sunspot number, but correlations are significant with 3-7 year time lag, meaning that this 5-year lag should not be taken as an absolute number, but the average maximum lag correlation. We do actually see a weak negative correlation between PC2 and the sunspot number with no time lag, which is also evident when looking at the time series. We have now added a version of figure 6 with no time lag in the supplementary (see figure below), which illustrates the weaker and more inconsistent correlation. In particular when looking at the moving window correlation, which we added on request by Referee # 2.
We agree with the reviewers comment that the variance explained by solar forcing is small. However, we do not think that this makes our conclusions less important. It is known that large part of climate variability is driven by internal processes and are hence unforced. Thus, we cannot expect high correlation coefficients, although the correlations between reconstructed PC2 and solar forcing at decadal ($r = 29$, $p < 0.01$) and centennial time scales ($r = 0.6$, $p < 0.01$) are not negligible. It is important to identify and understand the forced variability in the climate system, as this allows for a better understanding of past climate, and a better skill in predicting future climate.
Page 6, line 4-10 we discuss the correlation between reconstructed PC2 and 20CR PC2+PC3. We have clarified this in the revision.

Inserted, page 6, line 9: "This indicates that the variability projected on PC2 and PC3 of the reanalysis data is partly summarized in PC2 of the reconstruction."



New figure in revised supplementary. Reconstructed PC2 of SLP plotted with the sunspot number (no time lag). ***a*** *Moving 61-point correlation between reconstructed PC2 of SLP and sunspot number.* ***b*** *Time series of sunspot number, PC2+PC3 of 20CR SLP (see text and Table 2) and PC2 of reconstructed SLP.* ***c*** *Same as* ***b****, except filtered with a 5-year low-pass filter.* ***d*** *Same as* ***b****, except filtered with a 20-year low-pass filter.*

Other It is a bit strange to present anomalies in solar minima minus maxima. Most of the studies I am aware of show max-min. Please describe the method inferring significance. It is mentioned in figures but should be clearly described in the manuscript. Same for the type of filtering applied on PC2s.

We show min-max because we are interested in the response to solar minimum context of the discussion of the role of solar forcing during the Little Ice Age. Hence, we think it is most convenient for the reader to see the response to solar minimum as also done in similar studies (Adolphi et al., 2014; Ineson et al., 2011; Martin-Puertas et al., 2012).
We have added a paragraph in the Methods section describing methods used for significance, and filtering in the manuscript.

Inserted page 5, line 24: “Statistical tests and filtering of data. We test the significance for anomalies of climate field variables with a two-tailed Student’s t-test. For low-pass and band-pass filtering of data series we use a Fast Fourier Transform approach in case the data is used for correlation analysis. In Figure 2 b we use a ‘loess’ filter to smooth the data for visualization of the multi-decadal variability. When calculating significance for correlations of filtered data we use the method by Ebisuzaki (1997) to take autocorrelation into account.

P.3 L.16 time lag of 4-6 years: This is solar maximum, no? (see my comment above)

This is referring to the time lag response with this analysis done on PC2, using the time series of PC2 and sunspot number, so there is no choice of solar maximum or minimum in the analysis. This has been clarified in the revision.

Reformulated, page 3, line 16: “We achieve the strongest correspondence between the solar forcing and reconstructed PC2 of SLP with a time lag of 5 years, indicating that an ocean-atmosphere feedback is in play.”

P.3 l.23 “” fully prescribed CO2”: what do you mean? How it differs in E1-COSMOS?

For the E1-COSMOS ensemble the model is configured with a carbon cycle module which drives part of the variability in the atmospheric CO2. We have specified this difference between the E1 ensemble and our run in the revised version of the manuscript and added Table S1 to the main text.

p.3 parag.2.1: What is the horizontal resolution of the model. How many layers in the stratosphere?

The atmospheric component of the model, ECHAM5, is a spectral model and we ran it in T31 resolution, corresponding to 3.75° x 3.75° (lon, lat), with 19 vertical layers. 5 of these layers are in the stratosphere. We have included this information in the revised version of the manuscript.

p.5. l.1: “fitting the PCs of. . .”: please make it clear what is the criterion.
The sum of PCs? Individual PCs?

This is referring to the time series of PCs of the ensemble members evaluated using Eq. 1. We have reformulated this part in the revision.

Reformulated, page 4, line 32 to page 5 line 1: “We define the sorted model output as ensemble members, such that the best fitting model year, for each year of the ice core data, belongs to ensemble member 1, the second best fitting model years belong to ensemble member 2, and so forth. Using a Chi-square goodness-of-fit test, with respect to the measure in Eq. 1, we evaluate the ensemble members against the PCs of Greenland ice core $\delta^{18}O$ and reject model fits with likelihood $p > 0.01$ of not fitting the ice core data. This leaves us with 39 time series of reshuffled model data fitted to the ice core data.”

p.5 l.16: to me it seems that only DYE-3 shows high correlation. In other high accumulation sites correlation is much lower.

Indeed, high accumulation and multiple cores from the same location help constrain the noise. This appears to be the deciding factors. We have reformulated this in the revision.

Referee #2

——————— General comments ———————

This paper is proposing a new methodology for reconstruction atmospheric variability modes over the last millennium. This technique is using a climate model simulation including the online computation of the variation in concentration of oxygen isotopes, which are then compared with ice cores records. This comparison is then allowing, through a statistical approach, to find the winter atmospheric circulation that fits the best with the observations. Using this approach, the authors propose a new reconstruction of the two first modes of the variability of the atmospheric circulation over the North Atlantic sector. Then, the authors compare their reconstruction with time series of volcanic eruptions and solar forcing and find a clear signature of the former on the first mode, and a smaller signature of the latter on the second mode.

This is a very interesting study, proposing an innovative approach to reconstruct past climatic modes of variability. The analysis led is thorough and the results found are very impressive. As compared to a former reconstruction of the NAO (first mode of atmospheric circulation variability in winter), the new one exhibits slightly better scores (not sure the difference is really significant) of validation and is also showing no loss of variance when coming back in time. The study is also providing a reconstruction of the second mode of variability, usually denominated as the East Atlantic Pattern (EAP). These are very useful reconstructions for the climate community, and the manuscript clearly deserves publication.

We thank the reviewer for the positive review and the many constructive comments and suggestions.

I have noticed quite a number of issues mainly related with the clarity of the explanations, and also, from time to time, I have noticed some too enthusiastic evaluation of their work by the author, while a scientific work requires objectivity and discussion of strengths but also of potential weaknesses and caveats of the methods, which are not sufficiently discussed at the moment.

In particular, I think it would be worth discussing:

1) The fact that the model used is quite coarse resolution, so that it may have large biases, potentially even in its representation of the atmospheric circulation and of its variability modes, which need to be better depicted.

We have added a paragraph in the discussion about the role of the climate model in the reconstruction.

Inserted, page 8, line 6: "The model simulation used for our reconstructed translates the climate variability recorded in the Greenland ice cores to climate variability in the North Atlantic region. In the initial test of the isotope variability it is shown, that the spatio-temporal $\delta^{18}O$ variability of the ice cores is well represented by the model (Figure S6). This is a fundamental prerequisite which allows us to match the modeled $\delta^{18}O$ to the ice core $\delta^{18}O$ year-by-year. While the skill of the *reconstruction is higher in the vicinity around Greenland, the reconstruction shows significant correlations to reanalysis data wide spread across the North Atlantic region. This skill depends on i) the integrative nature of the $\delta^{18}O$ as recorded in the ice cores, and represented by the modeled $\delta^{18}O$ ii) the modeled atmospheric teleconnection patterns in terms of temperature and circulation, and iii) how these patterns are connected to modeled $\delta^{18}O$ for Greenland. Clearly, the reconstruction is strongly dependent on the climate model when it comes to whether or not it is possible at all to use our method, and when it comes to the skill of reconstructed spatial patterns. The resolution of our model model simulation is relatively course and using a higher resolution simulation could improve the representation of several processes. For example, vapor transport to dry polar regions is often inhibited in models with courser resolution, resulting in too little precipitation in the interior of ice sheets and a positive bias in $\delta^{18}O$ (Masson-Delmotte et al., 2008, Sjolte et al., 2011). This is related to cloud parameterizations and course resolution models having difficulties in explicitly representing frontal zones in connection with synoptic weather systems. The orography in course resolution models is also more smooth, loosing orographical features such as the southern dome of the Greenland ice sheet, which also affects atmospheric circulation and small scale spatial variability. In our approach we match the modeled PCs of $\delta^{18}O$, meaning that we are matching regional scale patterns in $\delta^{18}O$, which partly addresses the problem of matching course model output to site specific proxy data. However, having a higher resolution model simulation could for example improve the spatio-temporal representation of Greenland $\delta^{18}O$, allowing more than 3 PCs to be fitted, and generally giving a better representation of temperature, pressure and precipitation in the reconstruction. For reasons discussed above, it would be desirable using different GCMs to test for model dependencies of the reconstruction, as well as testing for added value of ensemble reconstructions with several different GCMs. Doing these tests is presently limited by the availability of millennium length simulations using isotope enabled GCMs."*

2) The reconstruction is finally relying on very few data, only 8 cores from Greenland, while the Atlantic sector is huge. Validation score are relatively high (even though a correlation of 0.5 stands for less than 30% of the variance explained. . .), but we can easily envisage that more data will be helpful to improve the reconstruction. A discussion of this will be useful I think.

We have added a paragraph in the discussion about the selection of proxy records, number of records and quality of records (dating, resolution, ...).

Inserted, page 8, line 6 (after model discussion): "We selected the proxy records for this study based on the criterion of having seasonal resolution, small dating uncertainty, a long time span and a wide regional spread. In order to provide a quantitative link to the isotope enabled GCM we selected only isotope-based proxies. For the time being, this leaves us with the 8 Greenland ice cores used in this study. Other seasonal resolution ice cores from Greenland are available, but only covering a limited time span, and comparing to these cores shows that the reconstructed $\delta^{18}O$ also compares well to the isotopic variability at these sites (Figure S2). However, including more Greenland ice cores of similar quality would generally improve the signal to noise ratio of the reconstruction, and such records should be included if available for subsequent studies. Obtaining seasonal resolution in ice core data is mainly limited by the accumulation rate and seasonality of precipitation, which depends on the regional climate setting of the drill site (Vinther et al., 2010, Zheng et al., 2018). Including other archives than ice cores would give a more widespread regional coverage, potentially providing better constraints on circulation patterns and climate trends. Some oxygen isotope records from tree rings in Sweden (e.g. Edwards et al., 2017) and speleothems from the European alps (e.g. De Jong et al., 2013) covering the past millennium primarily reflect winter climate conditions. Both records in these examples have 5-year resolution, and the speleothem record has hiatuses, which reflect some of the challenges in using these proxy records. However, *there could be benefits of using a larger selection of data, despite the different temporal resolution (Steiger and Hakim, 2016)."*

3) The link with solar forcing is far from being straightforward, and the authors push their conclusions a bit too far I should say. Also, the variability is not necessarily related with external forcing, and large amount of variance can be purely stochastic (i.e. internal to the climate system). While lots of research is led on this topic, notably to explain the little ice age, this is simply not discussed at all here. What is the percentage of variance in our reconstruction that is not related to any external forcing? This is not easy to isolate, but a rough estimate would be interesting (e.g. Wang et al. 2017).

According to our statistical tests (taking into account autocorrelation on filtered data) the connection we find between solar forcing and the reconstruction is significant on a range of time scales (e.g. time lagged response in Table 2). We agree that variability in the climate system, and in particularly atmospheric variability, is partly governed by stochastic processes. However, we never claim that all of the variability is explained by the forcing, this is also not what the results show. The reviewer does not specify which studies we fail to discuss. We have tried to include additional relevant studies in the revision including Schurer et al. (2014) and Wang et al. (2017).
The study by Wang et al. (2017) is very different from our study as their reconstruction is represented by a single index of multi-decadal ocean variability, compared to our reconstruction, which is a climate field reconstruction of atmospheric variability. Their attributing of forced variability is done 30-yr low-pass filtered data for both forcings and reconstruction. Low-pass filtering of the volcanic forcing, comprised entirely of spikes of durations no longer than 3 years, alters the properties of the time series, creating an artificial forcing signal already before the eruptions happen due to the smoothing of the filter. This might be less of an issue when dealing with ocean indecies with slower variability. Given the abrupt atmospheric response to volcanic eruptions seen in our reconstruction we are reluctant to introduce this type of analysis. However to accommodate the comment by the reviewer we have estimated the decadal to multi-decadal impact on NAO using a time lag correlation analysis of band-pass filtered data. This gives a maximum correlation of 0.33 with time lag of 3-4 years, corresponding to 10% of the reconstructed decadal to multi-decadal NAO variability being explained by volcanic forcing.
In the case of solar forcing the correlation is up to 0.6 between PC2 of reconstructed SLP, corresponding to ~40% of explained variance at centennial time scales. As we are dealing with a reconstruction it is very uncertain how this explained variance translates to the real world, for example due to the explained variance of the reconstructed PCs having a different distribution compared to reanalysis data. We have added the time-lag analysis of decadal to multi-decadal impact of volcanic forcing to the results and also discuss the variance in the reconstruction explained by external forcing.

With a proper discussion of the following points, I think the manuscript will clearly deserve to be published in climate of the past.

——————— Specific points: ———————
- p.1, l.1: the two first sentences can be interpreted as almost contradictory. Can you please reformulate them to clarify what you have in mind here?

Corrected.

- P.1, l. 3: add "of" before "the effects"

Corrected.

- P. 1, l. 3: “A positive phase. . .”: This is not true. Some of the historical eruptions have been followed by negative phase of the NAO (e.g. Agung) and the link between NAO and volcanic eruptions over the historical era is very small and sensitive to the selected eruptions (cf. Swingedouw et al. 2017, cited in the ms.). Please clarify.

Reformulated. This is the average response to tropical eruptions.

- P. 1, l. 10: you should specify here that this is for winter season.

Corrected.

- P. 1, l. 16: “we observe a similar response” is not clear at all. Please reformulate what you mean here. The next sentence concerning “blocking frequency” is also quite unclear. Please avoid the word “likely” as it has a very specific meaning in IPCC report, which is not the one used here, since I have not seen any proper analysis of changes in blocking frequency in the manuscript.

Reformulated, page 1, line 16-18: “On centennial time scales we observe a similar response in circulation as for the 5-year time-lagged response, with a high-pressure anomaly across North America and south of Greenland. This anomalous pressure pattern could be due to an increase in blocking frequency, possibly linked to a weakening of the subpolar gyre.”

- P. 1, l. 19: You’ll need to define what you mean here by little ice age in terms of time period.

Corrected. As the reviewer probably knows, there are some variations in how LIA is defined. Following others (e.g. Moffa-Sanchez et al., 2014) we define it 1450-1850 CE.

- P. 1, l. 19: “a clear link” is quite subjective sentence. Can you be more specific (i.e. quantitative: the correlation is quite small so that the link is not that clear I would say).

Reformulated. We now mention the range of correlation between solar forcing and reconstructed PC2 on decadal ($r = 0.29$, $p < 0.01$ (5-year time lag)) to centennial timescales ($r = 0.6$, $p < 0.01$).

- P. 1, l. 20: no “s” to “pattern”

Corrected.

- P. 2, l. 1-2: It could be worth to specify here winter season, since the whole paper is focused on this season.

Corrected.

- P. 2, l. 6-7: “Past changes. . ..”: Variability in atmospheric circulation is usually believed to be mainly stochastic, with a weak imprint of external forcing. This is why seasonal or decadal prediction are so difficult: the atmosphere is dominated by chaos, with very small part of the variance that is predictable. . . This is in contradiction with the word “attributed” from your sentence. Please clarify. Also, please add some references to substantiate the claim from this sentence.

We have reformulated this sentence to clarify that we are discussing climate proxy studies that attribute variability in their records to external forcing, e.g. Wang et al. (2017).

- P. 2, l. 10: add coma after “millennium”

Corrected.

- P. 2, l. 14-15: “The NAO has also. . .”: This sentence is very confusing. Please rephrase by just saying that solar variations are slightly correlated with SLP over the historical era in a few areas.

Reformulated. Many studies claim a NAO response to solar forcing, and it is common to refer to a NAO-like pattern. This is what we are discussing here.

- P. 2, l. 23-24: what is also very important is the amplitude of the influence i.e. the variance explained by solar variability. It is possible this forcing has a slight influence, but how does it compare to the noise, i.e. the purely stochastic variation of the NAO. The signal-to-noise ratio is indeed key to evaluate here.

We agree with his comment by the reviewer, which is also why we emphasize the preserved high frequency variability in our reconstruction. We assume that the reviewer is referring to line 22-23.

Inserted, page 2, line 14: “The observed NAO does not show a consistent correlation to the solar forcing (Gray et al., 2013)”

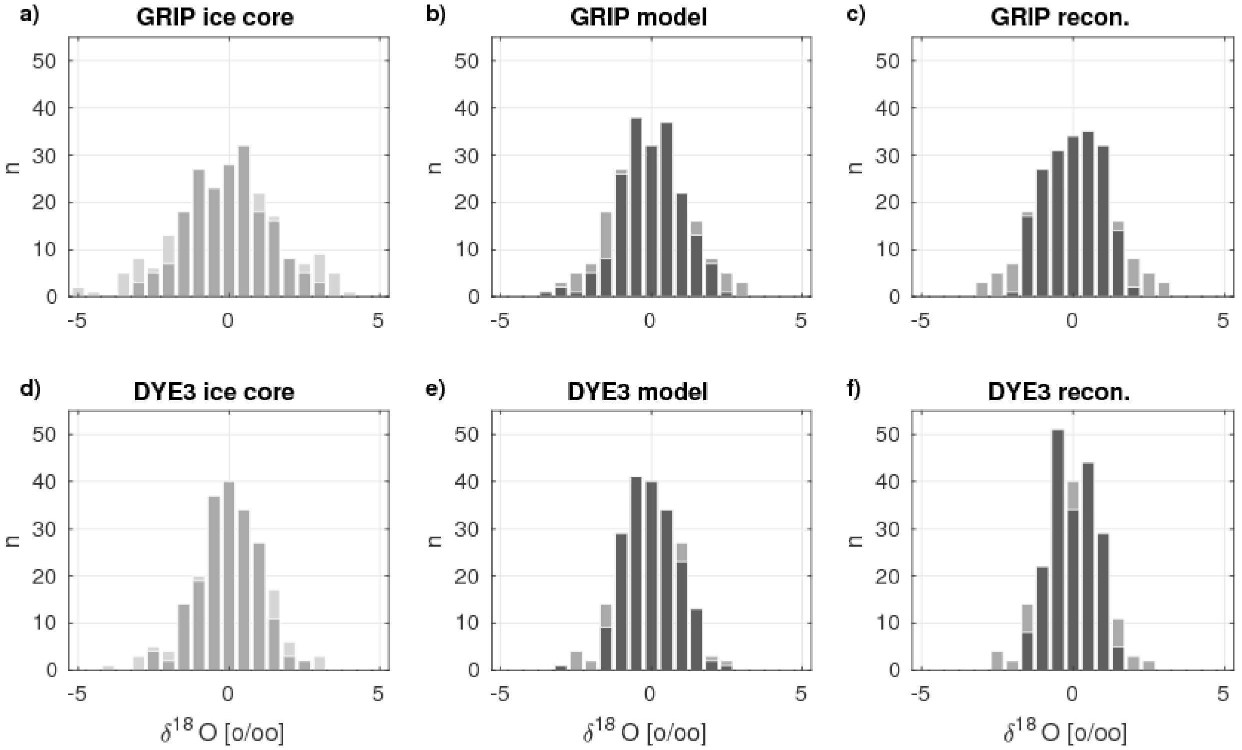
- P. 3, l. 1: replace “particularly” by “particular”.

Corrected.

- P. 3, l. 21-25: can you please provide a few more information on this model like its resolution? A brief description of its biases in the region analyzed will be also enlightening to evaluate the limit of this model for the exercise performed. How does the isotopic variations in the ice core sites compare with observations? How many grid points are covering Greenland?

The atmospheric component of the model, ECHAM5, is a spectral model and we ran it in T31 resolution, corresponding to 3.75° x 3.75° (lon, lat), with 19 vertical layers. 5 of these layers are in the stratosphere. 50 grid points are covering Greenland. We have added this information on model resolution, and additional information on model performance in the methods section. Also, supplementary Figure S1, Table S1 and Table S2 is moved to the manuscript to help clarify the methods section. We now also discuss the range of isotope variability in model and reconstruction compared to ice core. We use stacked ice core data from DYE3 and GRIP to take into account postdepositional noise in the ice core data. We will show histograms in the revised supplementary (see figure below). Even stacking 5-6 ice cores does probably not take into account all noise in the ice core data, which is part of the motivation of not using stacked data and also matching the PCA loadings and not matching directly to the ice core data of individual cores.

Inserted, page 3, line 21-24: “The performance of the atmospheric component of the model used in this study, ECHAM5-wiso, was evaluated for the Arctic region and Antarctica using different configurations of spatial resolution by Werner et al. (2011). For the configuration used in this study (T31) the model has a warm bias and is not depleted enough in δ^{18}O, however the climatological relation between δ^{18}O and temperature compares well to observations, despite the relatively course resolution.”



New figure in revised supplementary. Histograms of ice core $\delta^{18}O$ and modeled $\delta^{18}O_{pw}$ for winter covering the period 1778-1970. ***a*** *Histogram of stacked GRIP $\delta^{18}O$ (6 cores, gray), with $\delta^{18}O$ for a single core plotted in the background in light gray to illustrate reduction of variability from stacking.* ***b*** *modeled GRIP $\delta^{18}O$ (dark gray)with stacked ice core $\delta^{18}O$ plotted in the background for comparison.* ***c*** *same as* ***b****, but for reconstructed $\delta^{18}O$.* ***d-f****, same as* ***a-c****, but for DYE3. This figure both illustrates the inherent noise in ice core data, which can be reduced by using multiple cores, and shows that the model reconstruction covers most of the range of the ice core variability. Stacking even more ice core records (if available) could possible reduce the variability further. DYE3 has about twice the annual accumulation rate of GRIP, which also reduces the scatter.*

- P. 4, l. 11: this section 2.3 is key to the paper. Nevertheless, I find it a bit difficult to follow. They are very few references that support the method presented, so that it seems that this approach is entirely new. Is that correct? If so, I would like to have a longer description of it and a few examples to fully understand how the method works. Furthermore, I’m wondering how sensitive the results to the X2 measure are, which sounds a bit arbitrary and not very well justified.

Our method is indeed entirely new. We have extended the description of important parts of the method. The Chi-square probability is a standard measure for the goodness of fit of a given model to an observation. It is eventually based on squared distances between model and observation (like the RMSE), but allows us to infer the probability for each ensemble member of not matching the variability of the ice core PCs. Admittedly, the cutoff at $p > 0.01$ is somewhat arbitrary, but it provides us with a reasonable number of ensemble members, and the p-value rapidly ramps up to ~1 at ensemble member 60. Hence, other distance measures may provide a different number of ensemble members, but as long as they are based on quadratic distances, they would identify the same ranking of model-data matches.

- p. 5, l. 2: The authors depict “39 time series”. This is not crystal clear what the time is here and what the members are. Can you please be more specific to improve this description (a scheme could be useful as well for instance).

This should be 39 ensemble members, based on the best match of PC1, PC2 and PC3. We have clarified description of this section.

Reformulated, page 4, line 32 to page 5 line 1: "We define the sorted model output as ensemble members, such that the best fitting model year, for each year of the ice core data, belongs to ensemble member 1, the second best fitting model years belong to ensemble member 2, and so forth. Using a Chi-square goodness-of-fit test, with respect to the measure in Eq. 1, we evaluate the ensemble members against the PCs of Greenland ice core $\delta^{18}O$ *and reject model fits with likelihood* $p > 0.01$ *of not fitting the ice core data. This leaves us with 39 time series of reshuffled model data fitted to the ice core data."*

- P. 5, l. 2: "reshuffled model": a model is a tool with which you can perform simulations, providing climate variables. Thus, I'm not sure that "reshuffled model" is a proper terminology. I think, you are speaking here of the "reshuffled output from the simulation". This remark is true throughout the manuscript where model is used inconsistently.

The term we are using here is "reshuffled model fits" which could more accurately be "model output reshuffled to fit the ice core data". We have reformulated this sentence, and gone through the manuscript to correct for inconsistent terminology.

- P. 5, l. 29: "Note. . .": This is not a very expression. Normally all what is written in the manuscript is worth to be noted.

Corrected.

- P. 5, l. 30: "is associated": it is not clear what is substantiating this claim. Please clarify.

Reformulated. "... can be associated".

-P. 5, l. 30: "Figure S3": I find the pattern of EOF2 very different between model simulation and reanalysis. This should be said somewhere.

The difference in EOF patterns between model and 20CR is discussed page 6, line 4-10. We have clarified this discussion in the revision.

Inserted, page 6, line 9: "This indicates that the variability projected on PC2 and PC3 of the reanalysis data is partly summarized in PC2 of the reconstruction."

- P. 5, l. 33: add a coma after "reconstructions"

Corrected.

- P. 6, l. 10: what about the biases of the model? I believe this can also explain your differences between model simulations and reanalysis! Climate models are not providing perfect representation of reality, the opposite is true.

Here the expression "intrinsic model variability" includes model biases. We now discuss the role of the model run in the discussion. See also reply to major comment 1.

- P. 6, l. 14 add "in" after "consists"

We have added "of" after "consists".

- P. 7, l. 16: "Scandinavian blocking-type pattern". You should support this by a reference. Also, I would have argued that a Scandinavian blocking is referring to a change in frequency and is usually an anticyclone that remains blocked over the Scandinavia, following the weather regime approach (e.g. Vautard et al. 1990, Ortega et al. 2014 in line with ice core analysis and weather signature). Can you please clarify this sentence?

We have reformulated this sentence and added references (Ortega et al. 2014, Rimbu et al. 2017). Under the assumption that increased blocking frequency also results in higher average pressure in this region, we associate this pattern with the Scandinavian blocking pattern.

- P. 7, l. 22: "it should be noted". Avoid this type of subjective comment.

Corrected.

- P. 7, l. 30: "looks slightly different": this is a very subjective judgments, and I would rather say that they have hardly anything in common. Can you please provide a more objective metric of their similarity (spatial correlation for instance)?

As stated in line 28, we are comparing the 5-year time lagged response to the long-term response (Figure 4 b versus 4 c). We now refer to these figures for clarity.

- P. 7, l. 31: "wave structure". OK, but the signs are almost opposite in Fig. 4 a and c. . . So, this is not very convincing as a similar pattern!

See reply to comment above. The temperature response (Figure 4 e, and 4 f) is also very similar, and we now include this point in the comparison of the long-term and short-term response.

- P. 8, l. 3: maybe state that this secondary pattern is usually denominated as the EAP. Also, it is better to avoid "likely", except when the meaning is in line with IPCC precise definition.

We have reformulated this, also referring to the EAP pattern at an earlier point in the manuscript, and also add the range of correlations between PC2 and solar activity.

- P. 8, l. 14: this positive feedback is very weak, and hardly significant in the observations. . . (which is why the NAO is so difficult to predict at the seasonal scale, and the recent improvements seem to be more related with tropical teleconnection and frontal dynamics around the Gulf Stream region. . .).

For very strong volcanic eruptions, estimated to be well beyond the magnitude of eruptions during the instrumental era, forcing the NAO to a positive state for 2-3 years, the reinforcement feedback could play a bigger role. We have softened up this point in the revised discussion.

Reformulated, page 8, line 12-13: "... can best be explained ..." changed to "... could be explained ..."

- P. 8, l. 30: this line is very speculative and not supported by any references. The influence of the North Atlantic SST on atmospheric circulation is mainly not significant in winter season in the observation, or at least largely debated, so this hand-made explanation sounds a bit speculative I should say. I would at least replace "likely" l. 28 by "possible" given its high level of speculation.

Corrected.

- P. 9, l. 1: "Although the authors. . ." Indeed, they do not. . . They rather propose that the LIA would have been intrinsic and related to (unforced) rapid changes of the SPG. Furthermore, detection-attribution analysis (e.g. Schurer et al. 2014) found almost no signature of solar forcing on NH temperature, implying that LIA was hardy forced by solar variations. You should discuss this as well I would say. How do you reconcile this with your interpretations?

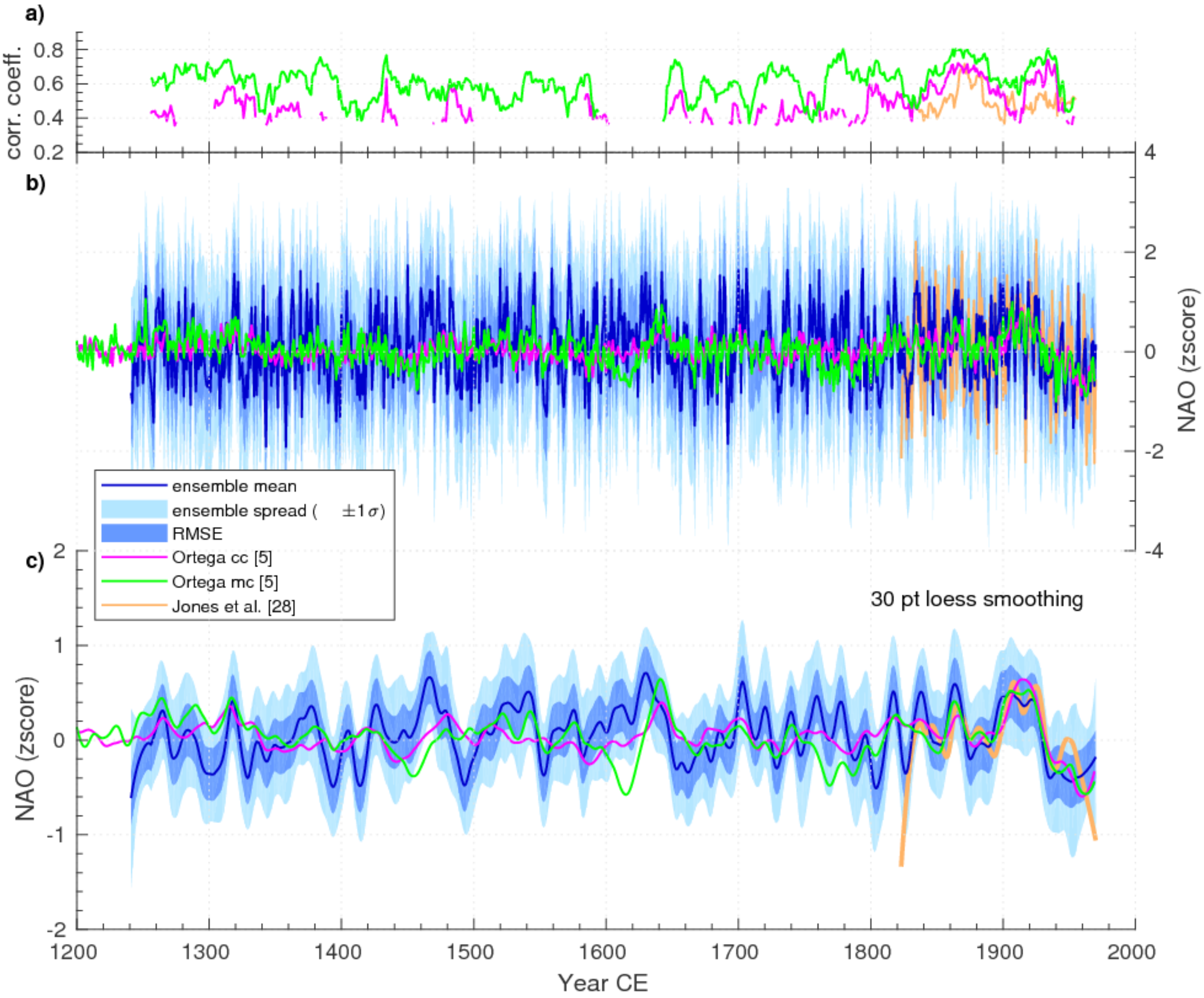
The study by Moreno-Chamarro et al., shows the connection between the SPG and the anomalies in SLP. The weakening of the SPG might happen unforced in their model (the test of solar forcing is not specified), but we see good indications that the changes in atmospheric circulation (PC2) is coinciding with changes in solar activity on a range of time scales. The model study by Schurer et al. (2014) was, as mentioned by the reviewer, done on NH mean temperature. If we would average the North Atlantic response to solar forcing in reconstructed temperature we would likely also see no significant response as some areas get warmer, while others get colder. We have addressed these points in the revised discussion.

Inserted, page 9, line 3: "One explanation could be that low solar activity is the preconditioning factor in reality, causing the response to solar forcing seen in our reconstruction, while the climate response to solar forcing might not be fully captured by the MPI-ESM (Mitchell et al., 2015) used by Moreno-Chamarro et al. (2017)."

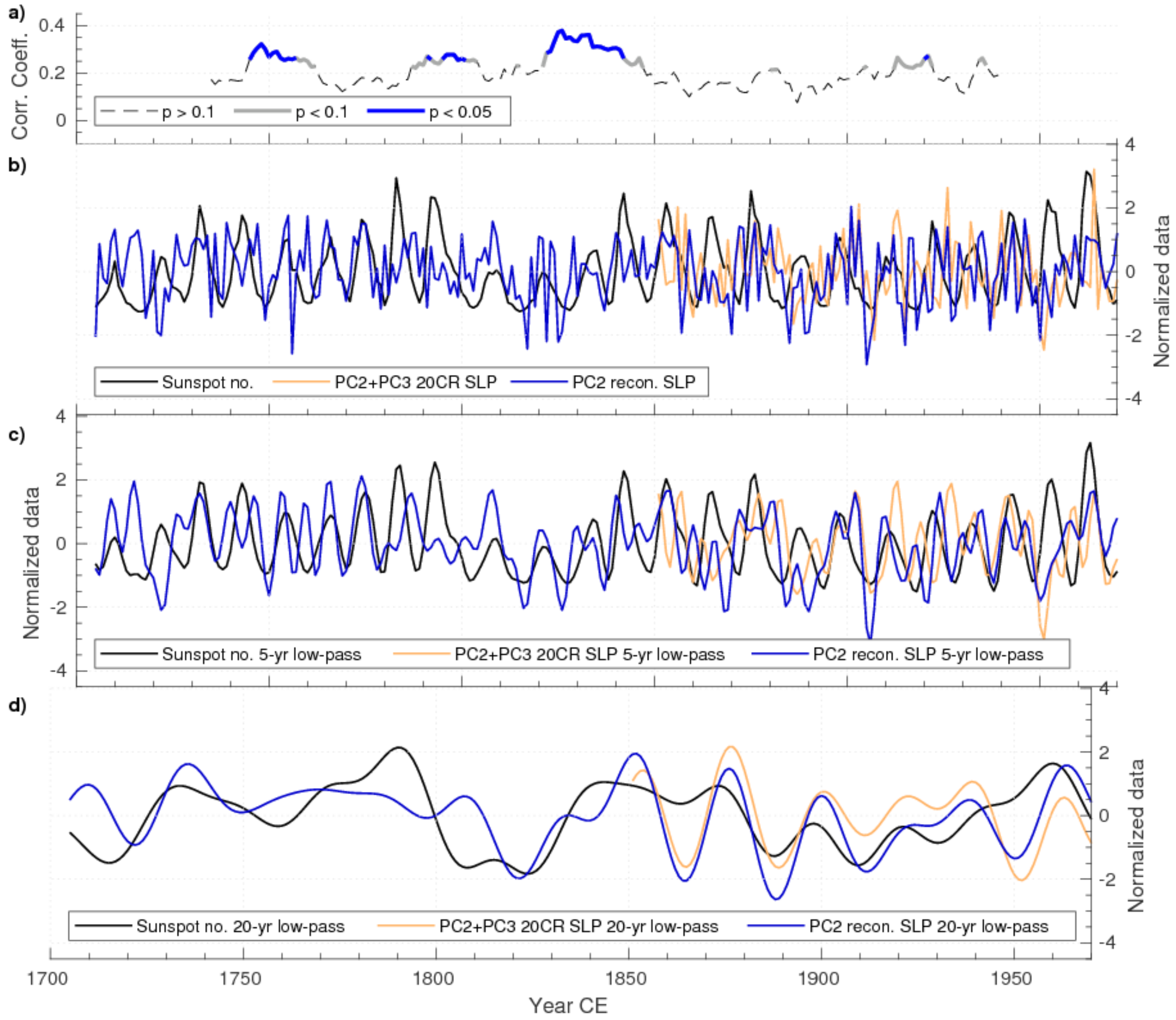
Inserted, page 9, line 4: "The complexity is also reflected in a non-uniform temperature response to solar forcing, with both regional warming and cooling. This also means that part of this signal will be smoothed out if such analysis is carried out on hemispherical mean temperature (e.g. Schurer et al., 2014)."

- Figure 1: what about adding a sliding window correlation here? It could also be useful for Figure 2 and Figure 6. It will help to evaluate when the time series are well-correlated by another mean than just by eye.

Following the suggestion of the reviewer we have added moving window correlations to the Figure 2 and 6, and the information from these new plots are included in the text (see plots below). In Figure 2 we used a 31-year window to show the variable strength in correlation, while we chose a 61-year window in Figure 6 to get reasonably clear correlations (~60 years cover ~5 solar cycles).



Revised Figure 2. Comparison of instrumental and proxy-based NAO reconstructions. ***a*** *Moving 31-point correlation between reconstructed NAO from this study and NAOcc (magenta), NAOmc (green) (Ortega et al., 2015) and observed NAO (yellow) (Jones et al., 1997). Only significant correlations are plotted ($p < 0.05$).* ***b*** *Ensemble mean reconstructed NAO (PC1 of reconstructed SLP (Hurrell et al., 2013)) with error estimated by ensemble spread and RMSE, compared to observed NAO (Jones et al., 1997) and NAO reconstructions by Ortega et al. (2015). The amplitude of all time series are scaled to fit the decadal variability of the observed NAO.* ***c*** *Same as* ***b****, except filtered with a 30 point 'loess' filter.*



Revised Figure 6. Reconstructed PC2 of SLP plotted with the 5-year lagged sunspot number. ***a*** *Moving 61-point correlation between reconstructed PC2 of SLP and sunspot number shifted for a 5-year time lag.* ***b*** *Time series of sunspot number shifted for a 5-year time lag, PC2+PC3 of 20CR SLP (see text and Table 2) and PC2 of reconstructed SLP.* ***c*** *Same as* ***b****, except filtered with a 5-year low-pass filter.* ***d*** *Same as* ***b****, except filtered with a 20-year low-pass filter.*

References

de Jong, R., C. Kamenik, M. Grosjean, Cold-season temperatures in the European Alps during the past millennium: variability, seasonality and recent trends, Quat. Sci. Rev., 82 (2013), pp. 1-12

Edwards, T. W. et al. Seasonal variability in Northern Hemisphere atmospheric circulation during the Medieval Climate Anomaly and the Little Ice Age. Quat. Sci. Rev. ***165****, 102–110 (2017).*

Masson-Delmotte, V., et al. (2008), A review of Antarctic surface snow isotopic composition: Observations, atmospheric circulation, and isotopic modeling, J. Clim., 21(13), 3359–3387, doi:10.1175/2007JCLI2139.1.

Mitchell, D. M., Misios, S. , Gray, L. J., Tourpali, K. , Matthes, K. , Hood, L. , Schmidt, H. , Chiodo, G. , Thiéblemont, R. , Rozanov, E. , Shindell, D. and Krivolutsky, A. (2015), Solar signals in CMIP-

5 simulations: the stratospheric pathway. Q.J.R. Meteorol. Soc., 141: 2390-2403. doi:[*10.1002/qj.2530*](https://doi.org/10.1002/qj.2530)

Ortega, P., Swingedouw, D., Masson-Delmotte, V., Risi, C., Vinther, B., Yiou, P., Vautard, R., Yoshimura, K., 2014. Characterizing atmospheric circulation signals in. Greenland ice cores: insights from a weather regime approach. Clim. Dyn. 43 (9–10), 2585–2605.

Rimbu, N., Lohmann, G., Werner, M., & Ionita, M. (2017). Links between central Greenland stable isotopes, blocking and extreme climate variability over Europe at decadal to multidecadal time scales. Climate Dynamics, 49, 649–663.

Roeckner, E., Bäuml, G., Bonaventura, L., Brokopf, R., Esch, M., G., Hagemann, S., Kirchner, I., Kornblueh, L., Manzini, E., Rhodin, A., Schlese, U., Schulzweida, U., and Tompkins, A.: The atmospheric general circulation model ECHAM5. Part I: Model description., Tech. Rep. Rep. 349, 127 pp., Max Planck Institute for Meteorology, available from MPI for Meteorology, Bundesstr. 53, 20146 Hamburg, Germany, 2003.

Schurer, A. P., Tett, S. F. B. & Hegerl, G. C. Small influence of solar variability on climate over the past millennium. Nature Geosci. 7, 104–108 (2014).

Sjolte, J., G. Hoffmann, S. J. Johnsen, B. M. Vinther, V. Masson-Delmotte, and C. Sturm (2011), Modeling the water isotopes in Greenland precipitation 1959–2001 with the meso-scale model REMO-iso, J. Geophys. Res., 116, D18105, doi:10.1029/2010JD015287.

Steiger, N. and Hakim, G.: Multi-timescale data assimilation for atmosphere–ocean state estimates, Clim. Past, 12, 1375-1388, https://doi.org/10.5194/cp-12-1375-2016, 2016.

Zheng, M., Sjolte, J., Adolphi, F., Vinther, B. M., Steen-Larsen, H. C., Popp, T. J., and Muscheler, R.: Climate information preserved in seasonal water isotope at NEEM: relationships with temperature, circulation and sea ice, Clim. Past Discuss., https://doi.org/10.5194/cp-2018-8, in review, 2018.

Wang, J., Yang, B., Ljungqvist, F. C., Luterbacher, J., Osborn, T. J., Briffa, K. R., and Zorita, E.: Internal and external forcing of multidecadal Atlantic climate variability over the past 1,200 years, Nat. Geosci., 10, 512–517, 2017.

Werner, M., P. M. Langebroek, T. Carlsen, M. Herold, and G. Lohmann (2011), Stable water isotopes in the ECHAM5 general circulation model: Toward high-resolution isotope modeling on a global scale, J. Geophys. Res., 116, D15109, doi:10.1029/2011JD015681.

Solar and volcanic forcing of North Atlantic climate inferred from a process-based reconstruction

Jesper Sjolte[1], Christophe Sturm[2], Florian Adolphi[1,3], Bo M. Vinther[4], Martin Werner[5], Gerrit Lohmann[5], and Raimund Muscheler[1]

[1]Department of Geology – Quaternary Science, Lund University, Sölvegatan 12, 223 62, Lund, Sweden.
[2]Department of Geological Sciences, Stockholm University, SE-106 91 Stockholm, Sweden.
[3]Climate and Environmental Physics & Oeschger Centre for Climate Change Research, Physics Institute, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland
[4]Centre for Ice and Climate, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen Oe, Denmark.
[5]Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Sciences, Bussestr. 24, 27515 Bremerhaven, Germany.

Correspondence: Jesper Sjolte (jesper.sjolte@geol.lu.se)

Abstract. ~~External forcings are known to impact atmospheric circulation . However,~~ The effect of external forcings on atmospheric circulation is debated. Due to the short observational period the analysis of the role of external forcings~~based on observational data is hampered due to the short observational period, and~~ , making it difficult to assess the sensitivity of atmospheric circulation to external forcings as well as persistence the effects~~are debated. A positive phase of the~~ . In observations the average response to tropical volcanic eruptions is a positive North Atlantic Oscillation (NAO) ~~has been observed the followingwinter after tropical volcanic eruptions~~during the following. However, past major tropical eruptions exceeding the magnitude of eruptions during the instrumental era could have more lasting effects. Decadal NAO variability has been suggested to follow the 11-year solar cycle, and linkages has been made between grand solar minima and negative NAO. However, the solar link to NAO found by modeling studies is not unequivocally supported by reconstructions, and is not consistently present in observations for the 20th century. Here we present a reconstruction of atmospheric winter circulation for the North Atlantic region covering the period 1241-1970 CE. Based on seasonally resolved Greenland ice core records and a 1200-year long simulation with an isotope enabled climate model, we reconstruct sea level pressure and temperature by matching the spatio-temporal variability of the modeled isotopic composition to that of the ice cores. This method allows us to capture the primary and secondary modes of atmospheric circulation in the North Atlantic region, while, contrary to previous reconstructions, preserving the amplitude of observed year-to-year atmospheric variability. Our results show 5 winters of positive NAO on average following major tropical volcanic eruptions, which is more persistent than previously suggested. In response to decadal minima of solar activity we find a high-pressure anomaly over Northern Europe, while a reinforced opposite response in pressure emerges with a 5-year time lag. On ~~longer~~ centennial time scales we observe a similar response in circulation as for the 5-year time-lagged response~~. This is likely~~ , with a high-pressure anomaly across North America and south of Greenland. This anomalous pressure pattern could be due to an increase in blocking frequency~~and an associated~~ , possibly linked to a weakening of the subpolar gyre. The long-term response of temperature to solar minima shows cooling across Greenland, Iceland and Western Europe, resembling the cooling pattern during the Little Ice Age (1450-1850 CE). While our results show

~~a clear link~~ significant correlation between solar forcing and the secondary circulation ~~patterns~~pattern on decadal ($r = 0.29$, $p < 0.01$) and centennial timescales ($r = 0.6$, $p < 0.01$), we find no consistent relationship between solar forcing and NAO. We conclude that solar and volcanic forcing impacts different modes of our reconstructed atmospheric circulation, which can aid to separate the regional effects of forcings and understand the underlying mechanisms.

1 Introduction

Climate variability in the North Atlantic region can, to a large extent, be explained by different modes of atmospheric circulation. This is particularly true for the variability during winter. The dominant mode is the NAO, which is a measure of the strength and position of the westerly winds across the North Atlantic. Secondary modes of circulation include Atlantic Ridge and Trough -types of variability, characterized by mid-Atlantic and Scandinavian blockings, respectively (Hurrell et al., 2013). These different modes for example determine the severity of European winters (Hurrell et al., 2013). ~~Past changes in climate and~~ In paleoclimate studies climate variability and changes atmospheric circulation are often attributed to external forcing related to solar variability or volcanic eruptions (e.g., Ortega et al., 2015; Wang et al., 2017)). Using reanalysis of weather observations (1871-2008) it has been shown that a positive phase of the NAO occurs in the winters following major tropical volcanic eruptions, while climate models generally fail to reproduce this dynamical response to the forcing (Driscoll et al., 2012; Zambri and Robock, 2016; Swingedouw et al., 2017). For the past millennium, Ortega et al. (2015) found a positive NAO response in the second winter following the 11 largest tropical eruptions in their reconstructed NAO. However, more persistent climate effects of volcanic eruptions were found by Sigl et al. (2015) who inferred cooler European summer temperatures lasting up to 10 years following major tropical eruptions during the past 2,500 years, raising the question if a more persistent impact on winter circulation also can be expected. The ~~NAO has also been suggested to be influenced by solar forcing as~~ observed NAO does not show a consistent correlation to the solar forcing (Gray et al., 2013). However, anomalies in sea level pressure during 1950-2010 exhibit a NAO-like pattern correlated to the 11-year solar cycle (Ineson et al., 2011; Adolphi et al., 2014). It has been hypothesized that solar-induced anomalies in the stratosphere can propagate to the troposphere (Haigh and Blackburn, 2006; Kodera and Kuroda, 2005) possibly synchronizing NAO variability to the 11-year solar cycle (Thieblemont et al., 2015) with a maximum response lagging 2-4 years due to ocean memory effects (Scaife et al., 2013; Gray et al., 2013). Several paleoclimate studies have indicated solar influences on climate in the North Atlantic region on centennial time-scales (Bond et al., 2001; Sejrup et al., 2010; Moffa-Sanchez et al., 2014; Adolphi et al., 2014; Jiang et al., 2015), and it has been suggested that the climate conditions during the Little Ice Age was linked to negative NAO forced by low solar activity (Shindell et al., 2001; Swingedouw et al., 2011). However the nature of the solar influence in terms of mechanisms and dynamical response is debated.

Our understanding of past climate dynamics and the impact of external forcings relies heavily on analysis of past changes. Gridded data sets of climate variables based on meteorological observations have been developed for these purposes (Compo et al., 2011). Such reanalysis data sets are constrained in quality and coverage back in time due to sparse instrumental data, and we rely on climate reconstructions based on proxy data to go beyond the instrumental era. Previously, reconstructions of climate indices (such as e.g. the NAO) have been site-based, with the reconstruction essentially done by extrapolating observed empirical relationships back in time. This approach results in a wide spread of reconstructions of past atmospheric circulation modes (Pinto and Raible, 2012). Recently, efforts have been done to develop reanalysis-type climate reconstructions based on climate proxy data. However, those reconstructions are focused on annual mean data, not taking into account the migration of circulation patterns from summer to winter (Hurrell et al., 2013), and have low skill for atmospheric circulation over the North Atlantic (Hakim et al., 2016; Steiger et al., 2017).

Stable water isotope ratios in ice cores carry quantitative information about past climate (Johnsen and Vinther, 2007). For Greenland ice cores seasonal isotope variability is attainable. In ~~particularly~~particular, the winter isotope signal has been shown to be highly correlated to atmospheric circulation and temperature (Vinther et al., 2003, 2010). However, the strength of the relation between the main patterns of variability of the ice core isotope signal and the NAO has been suggested to vary in strength (Vinther et al., 2003), indicating that a simple (regression based) relation between the ice core records and NAO bears large uncertainties for reconstruction purposes.

Here we present a climate reconstruction for the North Atlantic region for winter covering 1241-1970 CE, and analyze the impact of solar and volcanic forcing. For our reconstruction we combine a simulation using a coupled atmosphere-ocean model with stable isotope diagnostics embedded in the hydrological cycle with eight seasonally resolved isotope records from Greenland ice cores. We do not calibrate the reconstructed meteorological variables to observations, as we solely rely on matching the modeled isotopic composition to the ice core data. Testing the reconstruction against reanalysis data and observations, the reconstruction has good skill not only for the NAO, but also for secondary circulation modes. We find the average response to major tropical volcanic eruptions to be a positive NAO for the five consecutive winters after eruptions, which is more persistent than previous studies have shown. However, we find no persistent relationship between solar forcing and the NAO. On the other hand, we find a strong impact of solar forcing on the secondary modes of circulation represented by the second principal component (PC2) of reconstructed sea level pressure (SLP). We achieve the strongest correspondence between the solar forcing and ~~atmospheric circulation with~~ reconstructed PC2 of SLP with a time lag of ~~4-6~~ 5 years, indicating that an ~~ocean-atmosphere~~ atmosphere-ocean feedback is in play. Taking this time lag into account we find a consistent relationship between PC2 of reconstructed SLP and solar forcing on decadal to centennial time scales.

2 Data and Methods

2.1 Model simulation

We use the isotope enabled version of the atmosphere-ocean model ECHAM5/~~MPIOM~~ MPI-OM (Werner et al., 2016) to simulate the period 800-2000 CE forced by greenhouse gases, volcanic aerosols, total solar irradiance, land-use and orbital forcing (see Table ~~S1). Except~~ 1). For our study ECHAM5 is run with a T31 spectral resolution (3.75^o x 3.75^o) with 19 vertical layers, 5 of which are in the stratosphere. Our simulation uses a similar set up as the E1 COSMOS ensemble by Jungclaus et al. (2010), except for an updated solar forcing (see section below) and fully prescribed CO_2~~the forcings are the same as for~~, where the E1 simulations incorporates a carbon cycle module. We apply the identical physical general ciculation models (GCMs) (ECHAM5/MPI-OM) as the E1 ~~COSMOS ensembleby Jungclaus et al. (2010), generally yielding~~ ensemble, as well as similar forcings, and our simulation generally yields a very similar climate ~~due to the identical physical GCM models (ECHAM5/MPIOM) applied~~as the E1 ensemble. The performance of the atmospheric component of the model used in this study, ECHAM5-wiso, was evaluated for the Arctic region and Antarctica using different configurations of spatial resolution by Werner et al. (2011). For the configuration used in this study (T31) the model has a warm bias and is not depleted enough in δ^{18}O, however the climatological relation between δ^{18}O and temperature compares well to observations, despite the relatively course resolution. Greenland is represented by 50 grid points in the simulation.

2.2 Solar forcing

The solar forcing record employed in this study is based on the solar modulation record inferred from the combined neutron monitor and tree-ring ^{14}C data (Muscheler et al., 2016, 2007). In contrast to Schmidt et al. (2011, 2012) it covers the last 2000 years and consistently uses the solar modulation record for the complete period, while in Schmidt et al. (2011, 2012) the ^{14}C-based record is combined with sunspot-based data for the period after 1850 CE. Therefore, our approach employs an internally self-consistent forcing record for the last 2000 years, and agrees well with the latest recommended solar forcing reconstruction for the past millennium (Jungclaus et al., 2017). The 11-year solar cycle is based on the neutron monitor and ^{14}C data for the last 500 years where the underlying data has a sufficiently high temporal resolution. For the period before 1500 CE an artificial 10.5-year cycle was added and the phasing was adjusted to allow for a smooth connection to the subsequent start of the data-based solar cycle. The amplitude of the solar cycle modulation was inferred from the data-based part of the record, i.e. by employing the relationship between the solar cycle amplitude and the longer-term solar modulation levels (11-year averages) during last 500 years. This solar modulation record was scaled to the total solar irradiance (TSI) record by Schmidt et al. (2011, 2012) including longer-term trends in TSI (see “MEA (back)”, Figure 8 in Schmidt et al. (2011)). This was done by linearly transforming the solar modulation to a TSI record in order to reproduce the long-term changes (11-year average) in TSI from the Maunder minimum to the most recent 50 years i.e. leading to a similar range of long-term TSI changes as suggested by Schmidt et al. (2011).

2.3 Climate reconstruction

We use winter seasonal means (Nov-Apr) for 8 Greenland ice cores (Vinther et al., 2010) for the period 1241-1970 CE (Table ~~S2~~2). All ice cores are synchronized via volcanic reference horizons, and the dating uncertainty is estimated to one year for the oldest parts of the records used here (Vinther et al., 2006). Under the assumption that δ^{18}O in precipitation is a result of a number of processes mainly determined by atmospheric variability, we treat each year of the model run (see above) as a sample in a sampling space relating δ^{18}O in precipitation with circulation, temperature, etc. By extracting precipitation weighted winter seasonal means (Nov-Apr) of δ^{18}O from the model at the 8 ice cores sites we can find the model years best matching the isotope pattern of each winter in the ice core data. In order not to over-fit noise in the ice core data (post depositional processes etc. (White et al., 1997)) we first perform a principal component analysis of the ice core δ^{18}O and the model δ^{18}O from grid cells covering the investigated ice core sites. We retain the first three principal components (PCs) explaining a total of 60% and 97% of variability in the ice core and model δ^{18}O, respectively. The loadings of the 3 PCs for the model data match the loading of the ice core data well (Figure ~~S1~~1). Notice that this step is done without performing any selection of the model data, meaning that the modeled spatio-temporal variability of the δ^{18}O in precipitation corresponds well to that of the ice core data. As we only explain part of the variability of the δ^{18}O data using the first 3 PCs, we use an ensemble approach to take into account that the matching of a given ice core δ^{18}O pattern will ~~results~~ result in a suite of well-matching model years. To match each year in the ice core data, the model data is evaluated using a χ^2-measure between the 3 PCs of Greenland ice core δ^{18}O and the 3 PCs of the modeled δ^{18}O:

$$\chi^2_{Match} = \frac{1}{3}\sum_{k=1}^{3}(PC(k)_{model} - PC(k)_{icecore})^2 \tag{1}$$

~~,~~ where $PC(k)_{model}$ and $PC(k)_{icecore}$ are the values from a given year of the normalized time series of model and ice core δ^{18}O PCs, respectively. Each model year is evaluated against each ~~ice core year~~ year of the ice core data and then sorted in ascending order of the quality of the fit. This creates 1201 (number of model years, see above) fits of model ~~data for each ice core year, and in turn~~ output for each year of the ice core data, i.e., 1201 resampled (with replacement) ~~model fits~~ and sorted model years for the entire length of the ice core data (1241-1970 CE). We ~~view each of the fits along the length~~ define the sorted model output as ensemble members, such that the best fitting model year, for each year of the ice core data~~as ensemble members~~, belongs to ensemble member 1, the second best fitting model years belong to ensemble member 2, and so forth. Using a Chi-square goodness-of-fit test, with respect to the measure in Eq. 1, we evaluate the ensemble members against the ~~ice core data~~ PCs of Greenland ice core δ^{18}O and reject model fits with likelihood $p > 0.01$ of not fitting the ~~PCs of Greenland δ^{18}O~~ice core data. This leaves us with 39 time series of ~~model fits~~reshuffled model data fitted to the ice core data. The temporal succession of the reshuffled model ~~fits do~~ output does not resemble the order of the years in the original model run, and there are no systematic preferences of the method to pick certain time periods of the model run to match certain time periods of the reconstruction. For example if we exclude the years 1851-2000 from the model run and perform the reconstruction using the remaining years, the reconstruction is almost identical to the reconstruction using all model years. This also means that the timing of the forcing used for the model run has no relation to the timing or impact of specific forcings in the reconstruction, i.e., the forcings of the model run effectively serve to produce enough range in order to match the simulation to the isotope variability of the ice cores. Since the timing of the forcings used for the model simulation is not a factor influencing the reconstruction, the model performance for the response to forcings does not influence the reconstruction. We treat the 39 model fits as an ensemble solution fitting the model output to the ice cores δ^{18}O, and we calculate the ensemble mean reconstructed δ^{18}O and the standard deviation showing the ensemble spread. For the target climate variables (SLP, T2m) we extract the DJF ensemble mean corresponding to the reconstructed δ^{18}O constituting the reconstruction of these variables. Note that using this approach the method is optimized to fit modeled δ^{18}O to ice core δ^{18}O records and, in contrast to presently existing reconstructions, it is not calibrated to match observations of any of the target variables of the reconstruction (i.e. SLP or T2m).

As a first test we correlate the reconstructed δ^{18}O to winter means of ice core data from 20 cores covering the last 200 years of the reconstruction, i.e. including ice core data not part of the reconstruction. The correlation shows a spread between 0.4 and

0.9 with the highest correlations found for the high accumulation sites, ~~supporting~~ and where multiple ice cores for the same site are used for the reconstructeion. This supports the idea that the skill of the model fit to the seasonal ice core data is largely limited by the signal-to-noise ratio of the ice core data (Figure ~~S2~~S1). High accumulation sites are generally less sensitive to wind scoring and post depositional diffusion. Despite using the variability of the PCs to fit the model output to the ice core data, the reconstructed range of δ^{18}O matches the ice core data well (Figure S2). Furthermore we test the fit of the reshuffled model δ^{18}O to the ice core data of the 8 sites to investigate if any time periods stand out, as well as the mean fit of the method across the whole period of the reconstruction. We find no trends in the performance in terms of fitting the modeled PCs to the ice core data PCs. As the method performs equally well during any period as during the instrumental period (post 1850) in terms of matching the ice core isotope variability, it is likely that our reconstruction of SLP and T2m is equally valid for any period as it is for the instrumental period. This conclusion can be drawn as the reconstruction is not calibrated to either SLP or T2m, and is only constrained by the ice core isotope variability.

2.4 Statistical tests and filtering of data

We test the significance for anomalies of climate field variables with a two-tailed Student's t-test. For low-pass and band-pass filtering of data series we use a Fast Fourier Transform approach if data is used for correlation analysis. In Figure 2 b we use a 'loess' filter to smooth the data for visualization of the multi-decadal variability. When calculating significance for correlations of filtered data we use the method by Ebisuzaki (1997) to take autocorrelation into account.

3 Results

3.1 Evaluation of climate reconstruction

Comparing to the 20th Century Reanalysis (Compo et al., 2011) (20CR) our reconstruction shows skill for SLP and T2m in the North Atlantic region (Figure ~~1~~ 2 a, b), the main mode of atmospheric circulation (NAO) as well as secondary circulation modes (Figure ~~1~~ 2 c, d and Table ~~1). Note that our reconstruction is completely independent of the reanalysis data set~~3). This is a completely independent test of the reconstruction as the reconstruction has not be calibrated to reanalysis data or observations. For T2m the pattern of significant correlations with the 20CR data ~~is~~ can be associated with the main circulation modes (Figure S3 and S4), albeit with decreasing skill with the distance to the ice cores. We interpret the high skill near Greenland as being due to the direct physical connection between the local temperature in Greenland and the temperature along the path of the vapor, and the isotopic signal in Greenland ice cores. Contrary to previous millennial scale reconstructions, our reconstructed NAO shows similar strength of year-to-year variability as the observed NAO indicating that the reconstruction preserves the known characteristic variability of the NAO (Figure ~~1~~ 2 d, e, Figure ~~2~~ 3 a). In addition to capturing the NAO, our reconstruction has skill in representing Atlantic ridge/trough-type variability as projected on PC2 of SLP over the North Atlantic. The PC2 pattern is also referred to as the East Atlantic pattern (Wallace and Gutzler, 1981). The correlation of the reconstructed SLP PC2 and PC2 of the 20CR SLP is 0.24 ($p < 0.01$) and increases to 0.53 ($p < 0.01$) on decadal time scales (Table ~~1~~3). However, comparing the SLP patterns of PC2 in our reconstruction and reanalysis (Figure S3) implies that the variability captured by in PC2 of the reconstruction is likely split between PC2 and PC3 in the reanalysis. Indeed, the reconstructed PC2 of SLP is also correlated to PC3 of the reanalysis SLP (corr. = 0.19, $p < 0.01$, Table ~~1~~3). Correlating the reconstructed PC2 of SLP against the sum of the reanalysis PC2 and PC3 shows increased correlations between the reconstruction and reanalysis, in particularly on decadal and bi-decadal time scales (Table ~~1).~~ 3). This indicates that the variability projected on PC2 and PC3 of the reanalysis data is partly summarized in PC2 of the reconstruction. The difference in distribution of the variability on the PCs in the reconstruction and reanalysis is likely due to i) the intrinsic variability of the model use for the reconstruction and ii) the reconstruction only captures the North Atlantic variability as recorded in the isotopic composition of the Greenland ice core data.

The reconstructed NAO shows strong multi-decadal variability, while no major trends are found on centennial time scales, as opposed to the reconstruction by Trouet et al. (2009), and in agreement with the NAO reconstructions by Ortega et al. (2015) (Figure ~~2~~3). The NAO reconstructions by Ortega et al. (2015) consists of two multi-proxy NAO reconstructions. Of these reconstructions one is calibration-constrained (NAOcc) and the other model-constrained (NAOmc) (Figure ~~2~~ 3 b), where NAOmc only uses proxies from sites that were estimated from model simulations to have a stable relation to the NAO. It should be noted that both the reconstruction by Ortega et al. (2015) and this study use some of the same Greenland ice core data (Crete, DYE-3, GRIP), which obviously could lead to a correspondence in variability. We find the NAO in our SLP reconstruction has best correspondence with NAOmc for interannual variability, while for multi-decadal time scales prior to the instrumental period, there is little coherency between our reconstruction and both NAOmc and NAOcc (Table ~~1~~3). This lack of coherence of multidecadal variability is similar to the aforementioned divergence between previous NAO reconstructions. For an independent comparison we used ~~gridded reconstructions of SLP (Luterbacher et al., 2001) and~~ reconstructed NAO and gridded reconstructed SLP over Europe (Luterbacher et al., 2001), as well as gridded reconstructed temperature (Luterbacher et al., 2004) over Europe. However, we restricted the comparison to 1659-1970 due to the methodological differences in the reconstructions for Europe prior and post 1659, and the use of Greenland ice core data for the European temperature reconstruction covering 1500-1658. For the NAO reconstruction by Luterbacher et al. (2001) our reconstruction shows slightly lower correlation on interannnual time-scales compared to the correlation to NAOmc, but similar correlation on decadal to multi-decadal time-scales (Table 3). The comparison between our reconstruction and these reconstructions for Europe shows similar pattern and correlation levels for SLP as with the 20CR data, and moderate, but significant, correlations for temperature (Figure S5).

Our reconstructed NAO shows higher correlation (corr. = 0.52, $p < 0.01$) to the observed NAO (Jones et al., 1997) (DJF, 1824-1970) than NAOmc and NAOcc, which have correlations to the observed NAO of 0.46 and 0.47, respectively. Even more important, the skill of our reconstructed NAO is achieved without calibrating to the observed NAO. In summary, we think that our reconstruction is the most suitable for analyzing the influence of volcanic eruptions and solar activity on circulation, because i) our reconstruction not only has good skill for the NAO, but also for the secondary modes of circulation, which, as we will show later, is crucial for investigating the impact of solar forcing ii) high frequency variability is preserved, making it possible better to detect rapid shifts in circulation after volcanic eruptions, and finally iii) our reconstruction is not calibrated to observed SLP or T2m, making it free of biases that could arise from tuning a reconstruction to observations during the instrumental era.

3.2 Response of atmospheric circulation to external forcing

In this section we investigate the reconstructed response in SLP and temperature to major tropical volcanic eruptions and solar variability. The mean post eruption SLP and T2m anomalies in response to 12 major tropical eruptions show the characteristics of a positive NAO (Figure ~~3~~ 4 a, b). On average, we find a significant positive NAO response during the five consecutive winters following the eruptions (Figure ~~3~~ 4 c). Performing the same analysis on NAOmc yields a similar response, while not reaching as high significance levels and persistence as our reconstruction, likely due to the attenuated year-to-year variability of NAOmc (Figure ~~1 c & 3~~ 2 c & 4 c, Figure S6). Due to the short time span between some of the volcanic eruptions it is not possible to consistently analyze any longer term response than five years for single eruptions, since it would limit the number of eruptions and, hence, the robustness of the statistical analysis. ~~However, together with the suggested 5-10 years of post eruption summer cooling in Europe (Sigl et al., 2015), our result points to~~ We analyzed the decadal to multi-decadal response to volcanic forcing by calculating the correlation between reconstructed NAO and volcanic forcing from tropical eruptions (Toohey and Sigl, 2017) after filtering both data series with a ~~persistence in the response to major tropical volcanic eruptionswe have yet to observe in the instrumental record~~band-pass filter. The correlation analysis of filtered data estimates the combined effect of the eruptions during the reconstructed time frame. Due to the very abrupt nature of the volcanic forcing, heavy filtering can introduce artificial forcing prior to the onset of the actual forcing. We find that smoothing the volcanic forcing data using a low-pass filter with no less than 1/10 cycle per year only has negligible effects on the analysis. Performing a time-lag correlation analysis on the band-pass filtered data (1/10 to 1/100 cycles per year) we obtain significant correlations ($p < 0.01$) with the reconstructed NAO lagging the volcanic forcing 1 to 6 years (Figure S7). This corresponds well to the results of the analysis of the mean NAO response for the 12 major tropical eruptions. Maximum correlation is reached at time lags of 3 to 4 years with a correlation of 0.33. This shows the cumulative effect of several volcanic eruptions which can cause trends in the NAO on longer timescales than single eruptions.

For the analysis of solar influences on circulation we analyzed the average response to the 11-year solar cycle and the multi-decadal to centennial solar variability. We calculated the difference between reconstructed SLP and T2m for years of low and high solar activity using the annual sunspot number (Clette and Lefèvre, 2016) (1700-1970) and a ^{14}C-based solar reconstruction (1241-1970) (see Sect. 2.2) for the short term and long term cycles, respectively. The response to the 11-year solar cycle (solar low minus high) is a ~~Scandinavian blocking-type pattern~~ high pressure over Scandinavia corresponding well to the pattern found for reanalysis data (Figure ~~4~~ 5 a, d and Figure ~~5~~ 6 a, c). This anomalous high pressure could be due to increased frequency in Scandinavian blockings, which has been shown to impact Greenland δ^{18}O (Ortega et al., 2014; Rimbu et al., 2017). Investigating the time lagged response to the 11-year solar cycle we find the strongest response in reconstructed SLP and T2m when lagging the solar forcing with 5 years, which also matches the time lagged pattern found in reanalysis data (Figure ~~4~~ 5 b, e and Figure ~~5~~ 6 b, d). This pattern projects on PC2 of the reconstructed SLP, also with the strongest correlation between forcing and response when lagging the sunspot data with 5 years. Taking this time lag into account yields a consistent relationship between solar forcing and PC2 of reconstructed SLP on decadal to multi-decaldal time scales (Figure ~~6, Table 2). It should be noted that the~~ 7, Table 4). The 5-year lagged response in circulation is not simply due to the response to the solar maximum approximately half a cycle later in the 11-year solar cycle, but a reinforced response. This is most clearly seen in the stronger correlations for the time lagged response, both for the original data and the filtered data (Table ~~2~~4, Figure 7 and Figure S8). The relation between PC2 of reconstructed SLP and solar forcing persists also for centennial variability, which is seen by comparing the circulation response to the ^{14}C-based solar reconstruction (Table ~~2, Figure S7~~4, Figure S9). The pattern of the response in SLP to the long-term solar minima is an Atlantic ridge-type pattern (anomalous high south of Greenland), which also projects on PC2 of SLP, with an associated cooling pattern for the western North Atlantic (Figure ~~4~~ 5 c, f). Compared to the 5-year lagged response to the 11-year cycle this pattern has the strongest response in SLP south of Greenland, with a similar, but more widespread cooling in the eastern North Atlantic. Even though the SLP response looks slightly different for short and long-term solar forcing variations, the main feature: a wave structure over the North Atlantic and Scandinavia, is consistent (Figure 5 b, c). This similarity in the 5-year lagged and long-term response can also be seen in the patterns of the temperature anomalies (Figure 5 e, f). The temperature response to the long-term solar minima is a cooling across Greenland, Iceland and western Europe during solar minima (Figure ~~4~~ 5 f). This cooling pattern corresponds well to the suggested cooling during the Little Ice Age in proxy records from Greenland (Stuiver et al., 1997), Iceland (Moffa-Sanchez et al., 2014) and Europe (Luterbacher et al., 2004). A NAO-type response to long-term solar forcing would give opposing temperature responses in Greenland and Europe, which is not the case. We find no consistent relation between our reconstructed NAO and solar forcing. Instead we would like to stress the importance of the connection between solar activity and the secondary circulation patterns, which ~~likely captures~~ possibly shows the main response to solar forcing on decadal to centennial time scales, with correlations of 0.29 ($p < 0.01$) and 0.6 ($p < 0.01$), respectively.

4 Discussion

The model simulation used for our reconstructed translates the climate variability recorded in the Greenland ice cores to climate variability in the North Atlantic region. In the initial test of the isotope variability it is shown, that the spatio-temporal δ^{18}O variability of the ice cores is well represented by the model (Figure 1). This is a fundamental prerequisite which allows us to match the modeled δ^{18}O to the ice core δ^{18}O year-by-year. While the skill of the reconstruction is higher in the vicinity around Greenland, the reconstruction shows significant correlations to reanalysis data wide spread across the North Atlantic region. This skill depends on i) the integrative nature of the δ^{18}O as recorded in the ice cores, and represented by the modeled δ^{18}O ii) the modeled atmospheric teleconnection patterns in terms of temperature and circulation, and iii) how these patterns are connected to modeled δ^{18}O for Greenland. Clearly, the reconstruction is strongly dependent on the climate model when it comes to whether or not it is possible at all to use our method, and when it comes to the skill of reconstructed spatial patterns. The resolution of our model model simulation is relatively course and using a higher resolution simulation could improve the representation of several processes. For example, vapor transport to dry polar regions is often inhibited in models with courser resolution, resulting in too little precipitation in the interior of ice sheets and a positive bias in δ^{18}O

(Masson-Delmotte et al., 2008; Sjolte et al., 2011). This is related to cloud parameterizations and course resolution models having difficulties in explicitly representing frontal zones in connection with synoptic weather systems. The orography in course resolution models is also more smooth, loosing orographical features such as the southern dome of the Greenland ice sheet, which also affects atmospheric circulation and small scale spatial variability. In our approach we match the modeled PCs of δ^{18}O, meaning that we are matching regional scale patterns in δ^{18}O, which partly addresses the problem of matching course model output to site specific proxy data. However, having a higher resolution model simulation could for example improve the spatio-temporal representation of Greenland δ^{18}O, allowing more than 3 PCs to be fitted, and generally giving a better representation of temperature, pressure and precipitation in the reconstruction. For reasons discussed above, it would be desirable using different GCMs to test for model dependencies of the reconstruction, as well as testing for added value of ensemble reconstructions with several different GCMs. Doing these tests is presently limited by the availability of millennium length simulations using isotope enabled GCMs.

We selected the proxy records for this study based on the criterion of having seasonal resolution, small dating uncertainty, a long time span and a wide regional spread. In order to provide a quantitative link to the isotope enabled GCM we selected only isotope-based proxies. For the time being, this leaves us with the 8 Greenland ice cores used in this study. Other seasonal resolution ice cores from Greenland are available, but only covering a limited time span, and comparing to these cores shows that the reconstructed δ^{18}O also compares well to the isotopic variability at these sites (Figure S2). However, including more Greenland ice cores of similar quality would generally improve the signal to noise ratio of the reconstruction, and such records should be included if available for subsequent studies. Obtaining seasonal resolution in ice core data is mainly limited by the accumulation rate and seasonality of precipitation, which depends on the regional climate setting of the drill site (Vinther et al., 2010; Zheng et al., 2018). Including other archives than ice cores would give a more widespread regional coverage, potentially providing better constraints on circulation patterns and climate trends. Some oxygen isotope records from tree rings in Sweden (e.g., Edwards et al., 2017) and speleothems from the European alps (e.g., de Jong et al., 2013)) covering the past millennium primarily reflect winter climate conditions. Both records in these examples have 5-year resolution, and the speleothem record has hiatuses, which reflect some of the challenges in using these proxy records. However, there could be benefits of using a larger selection of data, despite the different temporal resolution (Steiger and Hakim, 2016).

The comparison between the response to volcanic forcing between our reconstruction and NAOmc (Figure 3 c) shows that the mean response look qualitatively very similar in the two reconstructions. However, as already mentioned, due to the preserved high frequency variability our reconstruction shows both a more immediate and persistent NAO response to volcanoes. This underlines the importance of producing climate reconstructions that do preserve high frequency variability, in particular if the reconstruction is used as baseline for model evaluation. In our analysis of the reconstructed response to volcanic eruptions we choose eruptions larger than or of similar magnitude as the 1991 Pinatubo eruption ($<$ -6 Wm^{-2}). As discussed by Swingedouw et al. (2017) climate effects of smaller eruptions can be difficult to detect due to stochastic climate variability. We find that we can detect the an impact on reconstructed NAO from tropical eruptions selected in the range from -4 Wm^{-2} to -8 Wm^{-2} (Sigl et al., 2015), yielding a significant positive NAO one year after the eruptions, on average. This appears to be the limit of detection for our reconstruction, possibly owing both to the partly stochastic variability of the NAO and to noise in the reconstruction.

Model studies of the volcanic response to major tropical eruptions during the past millennium show a large spread in the modeled NAO/AO response, with either no consistent response or 1-2 years of significant response (Swingedouw et al., 2017; Zambri et al., 2017). In contrast to this we find a clear tendency for positive NAO for the five consecutive winters following the year of eruption as an average response to the 12 largest tropical eruptions during 1241-1970 CE. An immediate strengthening of the polar vortex following eruptions is in agreement with the observed response of atmospheric circulation (Driscoll et al., 2012), which then translates to a positive NAO as the stratospheric anomaly propagates to the troposphere. The presence of volcanic aerosols gradually tails off of during the first 2-3 years (Crowley and Unterman, 2013) and a more sustained positive NAO for up to five years ~~can best~~ could be explained via a positive ocean feedback through a tri-pole SST response to the strongly anomalous positive NAO (Cayan, 1992). Ongoing efforts to improve the simulation of volcanic forcing and response could help close the gap between models and observations as well as reconstructions (Zanchettin et al., 2016).

It has been suggested that the observed increase in blocking frequency over the North Atlantic in response to solar minima (Woollings et al., 2010) is coupled to a weakening of the polar night jet in response to a weaker stratospheric equator-pole temperature gradient. This mechanism could be in play on both decadal and centennial time scales. A recent study investigated the response in circulation to solar activity using a regression-based analysis between sunspot data and gridded observed SLP and SST data (Gray et al., 2016). The authors analyzed the time lagged response to solar forcing, and found that the solar response could be explained via two mechanisms (Gray et al., 2016). One involving the aforementioned stratosphere-troposphere coupling acting on time lags of 0-2 years, and one for time lags of 3-4 years involving ocean temperature anomalies being stored beneath the mixed layer and reinforced from the previous winter (Gray et al., 2016). The reinforcement of SST anomalies from year-to-year has also been shown in a simulation of the response to the solar 11-year cycle (Andrews et al., 2015). Such a mechanism could be the cause of the time lag we see in the reconstructed response to solar forcing, although we get the maximum response at 5-year time lag, compared to the 3-4 year time lag found in observations (Gray et al., 2016). However, this difference could be due to differences in the methodologies of the analysis of the response to solar forcing, and that the aforementioned study (Gray et al., 2016) is focusing on the NAO-like response seen in their analysis. A ~~likely~~ possible mechanism based on our findings is as follows. An initial increase in atmospheric blockings weakens the subpolar gyre (Moffa-Sanchez et al., 2014) (SPG), thereby decreasing the heat transport to the north-western North Atlantic giving favorable conditions for mid-Atlantic blocking. This pattern is reinforced year-by-year and the main atmospheric response shifts to the node of PC2 of SLP south of Greenland under sustained forcing conditions on longer time scales (Figure ~~4~~ 5 c). A recent model study (Moreno-Chamarro et al., 2017) suggested that the cooling during the Little Ice Age was connected to a weakening in the SPG, sustained by way of atmosphere-ocean feedbacks. Although the authors do not related this to solar forcing, but with preconditioned initial model variability, the anomalies in SLP and temperature associated with the weakening of the SPG are very similar to the reconstructed pattern of the response to long-term solar forcing. One explanation could be that low solar activity is the preconditioning factor in reality, causing the response to solar forcing seen in our reconstruction, while the climate response to solar forcing might not be fully captured by the MPI-ESM (Mitchell et al., 2015) used by Moreno-Chamarro et al. (2017).

Our study of the reconstructed North Atlantic winter circulation shows a complex response to solar forcing which is, in contrast to a prevalent hypothesis, not directly linked to the NAO. The complexity is also reflected in a non-uniform temperature response to solar forcing, with both regional warming and cooling. This also means that part of this signal will be smoothed out if such analysis is carried out on hemispherical mean temperature (e.g., Schurer et al., 2015). In our study we do not exclude that there could be an influence of the solar 11-year cycle on NAO. However, unlike for PC2 of reconstructed SLP, we find no consistent relationship between reconstructed NAO and solar forcing across multiple time scales. Furthermore, the results suggest that sustained longer-term solar forcing leads to a shift in the atmospheric circulation response compared to the response to the short-term forcing, likely due to feedback processes involving the ocean integrating the long term effects of anomalous atmospheric circulation.

In the analysis our reconstruction we find that solar and volcanic forcing impacts different modes of the atmospheric circulation during winter, which can aid to separate the regional effects of forcings and understand the underlying mechanisms. The reconstructed response to forcings can also serve as a baseline for climate model evaluation. Although atmospheric variability to a large extent is a stochastic process, the variability in our reconstruction also shows an over all significant impact of forcings. The squared correlation coefficient can provide an estimate for explained variance of the external forcings. Using this approach, tropical volcanic forcing accounts for about 10% of the decadal to multi-decadal variability of the reconstructed NAO, while solar forcing accounts for about 40% of the variability of PC2 of reconstructed SLP on centennial timescales.

Author contributions. J.S. developed the method, performed the analysis, conducted the model simulation and wrote the first version of the manuscript. C.S. initiated the study and contributed to setting up the model simulation. F.A. contributed to the method development and analysis. B.V. provided seasonal ice core data. M.W. provided technical support for model simulation and access to climate model. G.L. provided insights on model setup. R.M. contributed with solar activity reconstruction and provided insight on solar forcing of climate. All authors discussed and edited the manuscript.

Competing interests. The authors declare no competing interests.

Acknowledgements. This work was supported by the Swedish Research Council (grant DNR2011-5418 & DNR2013-8421 to RM), the Crafoord foundation and the strategic research program of ModEling the Regional and Global Earth system (MERGE) hosted by the Faculty of Science at Lund University. The model simulation was carried out at the AWI Computer and Data Center, Bremerhaven.

References

Adolphi, F., Muscheler, R., Svensson, A., Aldahan, A., Possnert, G., Beer, J., Sjolte, J., Bjorck, S., Matthes, K., and Thieblemont, R.: Persistent link between solar activity and Greenland climate during the Last Glacial Maximum, NATURE GEOSCIENCE, 7, 662–666, https://doi.org/10.1038/NGEO2225, 2014.

Andrews, M. B., Knight, J. R., and Gray, L. J.: A simulated lagged response of the North Atlantic Oscillation to the solar cycle over the period 1960–2009, Environmental Research Letters, 10, 054 022, http://stacks.iop.org/1748-9326/10/i=5/a=054022, 2015.

Bond, G., Kromer, B., Beer, J., Muscheler, R., Evans, M., Showers, W., Hoffmann, S., Lotti-Bond, R., Hajdas, I., and Bonani, G.: Persistent solar influence on north Atlantic climate during the Holocene, SCIENCE, 294, 2130–2136, https://doi.org/10.1126/science.1065680, 2001.

Bretagnon, P. and Francou, G.: PLANETARY THEORIES IN RECTANGULAR AND SPHERICAL VARIABLES - VSOP-87 SOLUTIONS, ASTRONOMY & ASTROPHYSICS, 202, 309–315, 1988.

Cayan, D. R.: LATENT AND SENSIBLE HEAT-FLUX ANOMALIES OVER THE NORTHERN OCEANS - DRIVING THE SEA-SURFACE TEMPERATURE, JOURNAL OF PHYSICAL OCEANOGRAPHY, 22, 859–881, https://doi.org/10.1175/1520-0485(1992)022<0859:LASHFA>2.0.CO;2, 1992.

Clette, F. and Lefèvre, L.: The New Sunspot Number: Assembling All Corrections, Solar Physics, 291, 2629–2651, https://doi.org/10.1007/s11207-016-1014-y, https://doi.org/10.1007/s11207-016-1014-y, 2016.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, Jr., B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Broennimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, O., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The Twentieth Century Reanalysis Project, QUARTERLY JOURNAL OF THE ROYAL METEOROLOGICAL SOCIETY, 137, 1–28, https://doi.org/10.1002/qj.776, 2011.

Crowley, T. J. and Unterman, M. B.: Technical details concerning development of a 1200 yr proxy index for global volcanism, EARTH SYSTEM SCIENCE DATA, 5, 187–197, https://doi.org/10.5194/essd-5-187-2013, 2013.

Crowley, T. J., Zielinski, G., Vinther, B., Udisti, R., Kreutz, K., Cole-Dai, J., and Castellano, J.: Volcanism and the Little Ice Age, PAGES Newsletter, 16, 22–23, 2008.

de Jong, R., Kamenik, C., and Grosjean, M.: Cold-season temperatures in the European Alps during the past millennium: variability, seasonality and recent trends, Quaternary Science Reviews, 82, 1 – 12, https://doi.org/https://doi.org/10.1016/j.quascirev.2013.10.007, http://www.sciencedirect.com/science/article/pii/S0277379113003958, 2013.

Driscoll, S., Bozzo, A., Gray, L. J., Robock, A., and Stenchikov, G.: Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions, JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES, 117, https://doi.org/10.1029/2012JD017607, 2012.

Ebisuzaki, W.: A method to estimate the statistical significance of a correlation when the data are serially correlated, JOURNAL OF CLIMATE, 10, 2147–2153, https://doi.org/10.1175/1520-0442(1997)010<2147:AMTETS>2.0.CO;2, 1997.

Edwards, T. W., Hammarlund, D., Newton, B. W., Sjolte, J., Linderson, H., Sturm, C., Amour, N. A. S., Bailey, J. N.-L., and Nilsson, A. L.: Seasonal variability in Northern Hemisphere atmospheric circulation during the Medieval Climate Anomaly and the Little Ice Age, Quaternary Science Reviews, 165, 102 – 110, https://doi.org/https://doi.org/10.1016/j.quascirev.2017.04.018, http://www.sciencedirect.com/science/article/pii/S0277379117303475, 2017.

Gray, L. J., Woollings, T. J., Andrews, M., and Knight, J.: Eleven-year solar cycle signal in the NAO and Atlantic/European blocking, Quarterly Journal of the Royal Meteorological Society, 142, 1890–1903, https://doi.org/10.1002/qj.2782, http://dx.doi.org/10.1002/qj.2782, 2016.

Gray, L. J., Scaife, A. A., Mitchell, D. M., Osprey, S., Ineson, S., Hardiman, S., Butchart, N., Knight, J., Sutton, R., and Kodera, K.: A lagged response to the 11 year solar cycle in observed winter Atlantic/European weather patterns, JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES, 118, 13 405–13 420, https://doi.org/10.1002/2013JD020062, 2013.

Haigh, J. D. and Blackburn, M.: Solar influences on dynamical coupling between the stratosphere and troposphere, SPACE SCIENCE REVIEWS, 125, 331–344, https://doi.org/10.1007/s11214-006-9067-0, 2006.

Hakim, G. J., Emile-Geay, J., Steig, E. J., Noone, D., Anderson, D. M., Tardif, R., Steiger, N., and Perkins, W. A.: The lastmillennium climate reanalysis project: Framework and first results, JOURNAL OF GEOPHYSICAL RESEARCH-ATMOSPHERES, 121, 6745–6764, https://doi.org/10.1002/2016JD024751, 2016.

Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M.: An Overview of the North Atlantic Oscillation, pp. 1–35, American Geophysical Union, https://doi.org/10.1029/134GM01, http://dx.doi.org/10.1029/134GM01, 2013.

Ineson, S., Scaife, A. A., Knight, J. R., Manners, J. C., Dunstone, N. J., Gray, L. J., and Haigh, J. D.: Solar forcing of winter climate variability in the Northern Hemisphere, NATURE GEOSCIENCE, 4, 753–757, https://doi.org/10.1038/NGEO1282, 2011.

Jiang, H., Muscheler, R., Bjorck, S., Seidenkrantz, M.-S., Olsen, J., Sha, L., Sjolte, J., Eiriksson, J., Ran, L., Knudsen, K.-L., and Knudsen, M. F.: Solar forcing of Holocene summer sea-surface temperatures in the northern North Atlantic, GEOLOGY, 43, 203–206, https://doi.org/10.1130/G36377.1, 2015.

Johnsen, S. and Vinther, B.: {ICE} {CORE} {RECORDS} | Greenland Stable Isotopes, in: Encyclopedia of Quaternary Science, edited by Elias, S. A., pp. 1250 – 1258, Elsevier, Oxford, https://doi.org/http://dx.doi.org/10.1016/B0-44-452747-8/00345-8, http://www.sciencedirect.com/science/article/pii/B0444527478003458, 2007.

Jones, P., Jonsson, T., and Wheeler, D.: Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland, INTERNATIONAL JOURNAL OF CLIMATOLOGY, 17, 1433–1450, https://doi.org/10.1002/(SICI)1097-0088(19971115)17:13<1433::AID-JOC203>3.0.CO;2-P, 1997.

Jungclaus, J. H., Lorenz, S. J., Timmreck, C., Reick, C. H., Brovkin, V., Six, K., Segschneider, J., Giorgetta, M. A., Crowley, T. J., Pongratz, J., Krivova, N. A., Vieira, L. E., Solanki, S. K., Klocke, D., Botzet, M., Esch, M., Gayler, V., Haak, H., Raddatz, T. J., Roeckner, E., Schnur, R., Widmann, H., Claussen, M., Stevens, B., and Marotzke, J.: Climate and carbon-cycle variability over the last millennium, Climate of the Past, 6, 723–737, https://doi.org/10.5194/cp-6-723-2010, http://www.clim-past.net/6/723/2010/, 2010.

Jungclaus, J. H., Bard, E., Baroni, M., Braconnot, P., Cao, J., Chini, L. P., Egorova, T., Evans, M., González-Rouco, J. F., Goosse, H., Hurtt, G. C., Joos, F., Kaplan, J. O., Khodri, M., Klein Goldewijk, K., Krivova, N., LeGrande, A. N., Lorenz, S. J., Luterbacher, J., Man, W., Maycock, A. C., Meinshausen, M., Moberg, A., Muscheler, R., Nehrbass-Ahles, C., Otto-Bliesner, B. I., Phipps, S. J., Pongratz, J., Rozanov, E., Schmidt, G. A., Schmidt, H., Schmutz, W., Schurer, A., Shapiro, A. I., Sigl, M., Smerdon, J. E., Solanki, S. K., Timmreck, C., Toohey, M., Usoskin, I. G., Wagner, S., Wu, C.-J., Yeo, K. L., Zanchettin, D., Zhang, Q., and Zorita, E.: The PMIP4 contribution to CMIP6 – Part 3: The last millennium, scientific objective, and experimental design for the PMIP4 *past1000* simulations, Geoscientific Model Development, 10, 4005–4033, https://doi.org/10.5194/gmd-10-4005-2017, https://www.geosci-model-dev.net/10/4005/2017/, 2017.

Kodera, K. and Kuroda, Y.: A possible mechanism of solar modulation of the spatial structure of the North Atlantic Oscillation, Journal of Geophysical Research: Atmospheres, 110, n/a–n/a, https://doi.org/10.1029/2004JD005258, http://dx.doi.org/10.1029/2004JD005258, d02111, 2005.

Lefohn, A., Husar, J., and Husar, R.: Estimating historical anthropogenic global sulfur emission patterns for the period 1850-1990, ATMOSPHERIC ENVIRONMENT, 33, 3435–3444, https://doi.org/10.1016/S1352-2310(99)00112-0, 214th National Meeting of the American-Chemical-Society, LAS VEGAS, NEVADA, SEP 07-11, 1997, 1999.

Luterbacher, J., Xoplaki, E., Dietrich, D., Jones, P. D., Davies, T. D., Portis, D., Gonzalez-Rouco, J. F., von Storch, H., Gyalistras, D., Casty, C., and Wanner, H.: Extending North Atlantic oscillation reconstructions back to 1500, Atmospheric Science Letters, 2, 114–124, https://doi.org/10.1006/asle.2002.0047, http://dx.doi.org/10.1006/asle.2002.0047, 2001.

Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., and Wanner, H.: European seasonal and annual temperature variability, trends, and extremes since 1500, SCIENCE, 303, 1499–1503, https://doi.org/10.1126/science.1093877, 2004.

MacFarling Meure, C., Etheridge, D., Trudinger, C., Steele, P., Langenfelds, R., van Ommen, T., Smith, A., and Elkins, J.: Law Dome CO2, CH4 and N2O ice core records extended to 2000 years BP, GEOPHYSICAL RESEARCH LETTERS, 33, https://doi.org/10.1029/2006GL026152, 2006.

Marland, G., Boden, T. A., and Andres, R. J.: Global, regional and national emissions, in: Trends: a compendium of data on global change, Carbon Dioxide Information Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, TN, 2003.

Masson-Delmotte, V., Hou, S., Ekaykin, A., Jouzel, J., Aristarain, A., Bernardo, R. T., Bromwich, D., Cattani, O., Delmotte, M., Falourd, S., Frezzotti, M., Gallée, H., Genoni, L., Isaksson, E., Landais, A., Helsen, M. M., Hoffmann, G., Lopez, J., Morgan, V., Motoyama, H., Noone, D., Oerter, H., Petit, J. R., Royer, A., Uemura, R., Schmidt, G. A., Schlosser, E., Simões, J. C., Steig, E. J., Stenni, B., Stievenard, M., van den Broeke, M. R., van de Wal, R. S. W., van de Berg, W. J., Vimeux, F., and White, J. W. C.: A Review of Antarctic Surface Snow Isotopic Composition: Observations, Atmospheric Circulation, and Isotopic Modeling, Journal of Climate, 21, 3359–3387, https://doi.org/10.1175/2007JCLI2139.1, https://doi.org/10.1175/2007JCLI2139.1, 2008.

Mitchell, D. M., Misios, S., Gray, L. J., Tourpali, K., Matthes, K., Hood, L., Schmidt, H., Chiodo, G., Thiéblemont, R., Rozanov, E., Shindell, D., and Krivolutsky, A.: Solar signals in CMIP-5 simulations: the stratospheric pathway, Quarterly Journal of the Royal Meteorological Society, 141, 2390–2403, https://doi.org/10.1002/qj.2530, https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.2530, 2015.

Moffa-Sanchez, P., Born, A., Hall, I. R., Thornalley, D. J. R., and Barker, S.: Solar forcing of North Atlantic surface temperature and salinity over the past millennium, NATURE GEOSCIENCE, 7, 275–278, https://doi.org/10.1038/ngeo2094, 2014.

Moreno-Chamarro, E., Zanchettin, D., Lohmann, K., and Jungclaus, J. H.: An abrupt weakening of the subpolar gyre as trigger of Little Ice Age-type episodes, Climate Dynamics, 48, 727–744, https://doi.org/10.1007/s00382-016-3106-7, https://doi.org/10.1007/s00382-016-3106-7, 2017.

Muscheler, R., Joos, F., Beer, J., Müller, S. A., Vonmoos, M., and Snowball, I.: Solar activity during the last 1000 yr inferred from radionuclide records, Quaternary Science Reviews, 26, 82 – 97, https://doi.org/http://dx.doi.org/10.1016/j.quascirev.2006.07.012, http://www.sciencedirect.com/science/article/pii/S0277379106002460, 2007.

Muscheler, R., Adolphi, F., Herbst, K., and Nilsson, A.: The Revised Sunspot Record in Comparison to Cosmogenic Radionuclide-Based Solar Activity Reconstructions, Solar Physics, 291, 3025–3043, https://doi.org/10.1007/s11207-016-0969-z, http://dx.doi.org/10.1007/s11207-016-0969-z, 2016.

Ortega, P., Swingedouw, D., Masson-Delmotte, V., Risi, C., Vinther, B., Yiou, P., Vautard, R., and Yoshimura, K.: Characterizing atmospheric circulation signals in Greenland ice cores: insights from a weather regime approach, Climate Dynamics, 43, 2585–2605, https://doi.org/10.1007/s00382-014-2074-z, https://doi.org/10.1007/s00382-014-2074-z, 2014.

Ortega, P., Lehner, F., Swingedouw, D., Masson-Delmotte, V., Raible, C. C., Casado, M., and Yiou, P.: A model-tested North Atlantic Oscillation reconstruction for the past millennium, NATURE, 523, 71+, https://doi.org/10.1038/nature14518, 2015.

Paul, J., Fortuin, F., and Kelder, H.: An ozone climatology based on ozonesonde and satellite measurements, Journal of Geophysical Research: Atmospheres, 103, 31 709–31 734, https://doi.org/10.1029/1998JD200008, http://dx.doi.org/10.1029/1998JD200008, 1998.

Pinto, J. G. and Raible, C. C.: Past and recent changes in the North Atlantic oscillation, WILEY INTERDISCIPLINARY REVIEWS-CLIMATE CHANGE, 3, 79–90, https://doi.org/10.1002/wcc.150, 2012.

Pongratz, J., Reick, C., Raddatz, T., and Claussen, M.: A reconstruction of global agricultural areas and land cover for the last millennium, GLOBAL BIOGEOCHEMICAL CYCLES, 22, https://doi.org/10.1029/2007GB003153, 2008.

Rimbu, N., Lohmann, G., Werner, M., and Ionita, M.: Links between central Greenland stable isotopes, blocking and extreme climate variability over Europe at decadal to multidecadal time scales, Climate Dynamics, 49, 649–663, https://doi.org/10.1007/s00382-016-3365-3, https://doi.org/10.1007/s00382-016-3365-3, 2017.

Scaife, A. A., Ineson, S., Knight, J. R., Gray, L., Kodera, K., and Smith, D. M.: A mechanism for lagged North Atlantic climate response to solar variability, Geophysical Research Letters, 40, 434–439, https://doi.org/10.1002/grl.50099, http://dx.doi.org/10.1002/grl.50099, 2013.

Schmidt, G. A., Jungclaus, J. H., Ammann, C. M., Bard, E., Braconnot, P., Crowley, T. J., Delaygue, G., Joos, F., Krivova, N. A., Muscheler, R., Otto-Bliesner, B. L., Pongratz, J., Shindell, D. T., Solanki, S. K., Steinhilber, F., and Vieira, L. E. A.: Climate forcing reconstructions for use in PMIP simulations of the last millennium (v1.0), Geoscientific Model Development, 4, 33–45, https://doi.org/10.5194/gmd-4-33-2011, http://www.geosci-model-dev.net/4/33/2011/, 2011.

Schmidt, G. A., Jungclaus, J. H., Ammann, C. M., Bard, E., Braconnot, P., Crowley, T. J., Delaygue, G., Joos, F., Krivova, N. A., Muscheler, R., Otto-Bliesner, B. L., Pongratz, J., Shindell, D. T., Solanki, S. K., Steinhilber, F., and Vieira, L. E. A.: Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1.1), Geoscientific Model Development, 5, 185–191, https://doi.org/10.5194/gmd-5-185-2012, http://www.geosci-model-dev.net/5/185/2012/, 2012.

Schurer, A. P., Tett, S. F. B., and Hegerl, G. C.: Small influence of solar variability on climate over the past millennium (vol 7, pg 104, 2014), NATURE GEOSCIENCE, 8, 159, https://doi.org/10.1038/NGEO2342, 2015.

Sejrup, H. P., Lehman, S. J., Haflidason, H., Noone, D., Muscheler, R., Berstad, I. M., and Andrews, J. T.: Response of Norwegian Sea temperature to solar forcing since 1000 A.D., Journal of Geophysical Research: Oceans, 115, n/a–n/a, https://doi.org/10.1029/2010JC006264, http://dx.doi.org/10.1029/2010JC006264, c12034, 2010.

Shindell, D. T., Schmidt, G. A., Mann, M. E., Rind, D., and Waple, A.: Solar Forcing of Regional Climate Change During the Maunder Minimum, Science, 294, 2149–2152, https://doi.org/10.1126/science.1064363, http://science.sciencemag.org/content/294/5549/2149, 2001.

Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., Buentgen, U., Caffee, M., Chellman, N., Dahl-Jensen, D., Fischer, H., Kipfstuhl, S., Kostick, C., Maselli, O. J., Mekhaldi, F., Mulvaney, R., Muscheler, R., Pasteris, D. R., Pilcher, J. R., Salzer, M., Schuepbach, S., Steffensen, J. P., Vinther, B. M., and Woodruff, T. E.: Timing and climate forcing of volcanic eruptions for the past 2,500 years, NATURE, 523, 543+, https://doi.org/10.1038/nature14565, 2015.

Sjolte, J., Hoffmann, G., Johnsen, S. J., Vinther, B. M., Masson-Delmotte, V., and Sturm, C.: Modeling the water isotopes in Greenland precipitation 1959–2001 with the meso-scale model REMO-iso, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2010JD015287, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2010JD015287, 2011.

Steiger, N. and Hakim, G.: Multi-timescale data assimilation for atmosphere–ocean state estimates, Climate of the Past, 12, 1375–1388, https://doi.org/10.5194/cp-12-1375-2016, https://www.clim-past.net/12/1375/2016/, 2016.

Steiger, N. J., Steig, E. J., Dee, S. G., Roe, G. H., and Hakim, G. J.: Climate reconstruction using data assimilation of water isotope ratios from ice cores, Journal of Geophysical Research: Atmospheres, 122, 1545–1568, https://doi.org/10.1002/2016JD026011, http://dx.doi.org/10.1002/2016JD026011, 2016JD026011, 2017.

Stuiver, M., Braziunas, T., Grootes, P., and Zielinski, G.: Is There Evidence for Solar Forcing of Climate in the GISP2 Oxygen Isotope Record?, Quaternary Research, 48, 259 – 266, https://doi.org/http://dx.doi.org/10.1006/qres.1997.1931, http://www.sciencedirect.com/science/article/pii/S0033589497919311, 1997.

Swingedouw, D., Terray, L., Cassou, C., Voldoire, A., Salas-Mélia, D., and Servonnat, J.: Natural forcing of climate during the last millennium: fingerprint of solar variability, Climate Dynamics, 36, 1349–1364, https://doi.org/10.1007/s00382-010-0803-5, https://doi.org/10.1007/s00382-010-0803-5, 2011.

Swingedouw, D., Mignot, J., Ortega, P., Khodri, M., Menegoz, M., Cassou, C., and Hanquiez, V.: Impact of explosive volcanic eruptions on the main climate variability modes, GLOBAL AND PLANETARY CHANGE, 150, 24–45, https://doi.org/10.1016/j.gloplacha.2017.01.006, 2017.

Tanre, D., Geleyn, J.-F., and Slingo, J. M.: Aerosols and Their Climatic Effects, chap. First results of the introduction of an advanced aerosol-radiation interaction in the ECMWF low resolution global model, p. 133–177, Deepak Publishing, Hampton, VA, 1984.

Thieblemont, R., Matthes, K., Omrani, N.-E., Kodera, K., and Hansen, F.: Solar forcing synchronizes decadal North Atlantic climate variability, NATURE COMMUNICATIONS, 6, https://doi.org/10.1038/ncomms9268, 2015.

Toohey, M. and Sigl, M.: Volcanic stratospheric sulfur injections and aerosol optical depth from 500 BCE to 1900 CE, Earth System Science Data, 9, 809–831, https://doi.org/10.5194/essd-9-809-2017, https://www.earth-syst-sci-data.net/9/809/2017/, 2017.

Trouet, V., Esper, J., Graham, N. E., Baker, A., Scourse, J. D., and Frank, D. C.: Persistent Positive North Atlantic Oscillation Mode Dominated the Medieval Climate Anomaly, SCIENCE, 324, 78–80, https://doi.org/10.1126/science.1166349, 2009.

Vinther, B., Jones, P., Briffa, K., Clausen, H., Andersen, K., Dahl-Jensen, D., and Johnsen, S.: Climatic signals in multiple highly resolved stable isotope records from Greenland, Quaternary Science Reviews, 29, 522 – 538, https://doi.org/http://dx.doi.org/10.1016/j.quascirev.2009.11.002, http://www.sciencedirect.com/science/article/pii/S0277379109003655, 2010.

Vinther, B. M., Johnsen, S. J., Andersen, K. K., Clausen, H. B., and Hansen, A. W.: NAO signal recorded in the stable isotopes of Greenland ice cores, Geophysical Research Letters, 30, n/a–n/a, https://doi.org/10.1029/2002GL016193, http://dx.doi.org/10.1029/2002GL016193, 1387, 2003.

Vinther, B. M., Clausen, H. B., Johnsen, S. J., Rasmussen, S. O., Andersen, K. K., Buchardt, S. L., Dahl-Jensen, D., Seierstad, I. K., Siggaard-Andersen, M.-L., Steffensen, J. P., Svensson, A., Olsen, J., and Heinemeier, J.: A synchronized dating of three Greenland ice cores throughout the Holocene, Journal of Geophysical Research: Atmospheres, 111, n/a–n/a, https://doi.org/10.1029/2005JD006921, http://dx.doi.org/10.1029/2005JD006921, d13102, 2006.

Wallace, J. M. and Gutzler, D. S.: Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter, Monthly Weather Review, 109, 784–812, https://doi.org/10.1175/1520-0493(1981)109<0784:TITGHF>2.0.CO;2, https://doi.org/10.1175/1520-0493(1981)109<0784:TITGHF>2.0.CO;2, 1981.

Wang, J., Yang, B., Ljungqvist, F. C., Luterbacher, J., Osborn, T. J., Briffa, K. R., and Zorita, E.: Internal and external forcing of multidecadal Atlantic climate variability over the past 1,200 years, NATURE GEOSCIENCE, 10, 512+, https://doi.org/10.1038/NGEO2962, 2017.

Werner, M., Langebroek, P. M., Carlsen, T., Herold, M., and Lohmann, G.: Stable water isotopes in the ECHAM5 general circulation model: Toward high-resolution isotope modeling on a global scale, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2011JD015681, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JD015681, 2011.

Werner, M., Haese, B., Xu, X., Zhang, X., Butzin, M., and Lohmann, G.: Glacial–interglacial changes in $H_2{}^{18}O$, HDO and deuterium excess – results from the fully coupled ECHAM5/MPI-OM Earth system model, Geoscientific Model Development, 9, 647–670, https://doi.org/10.5194/gmd-9-647-2016, http://www.geosci-model-dev.net/9/647/2016/, 2016.

White, J. W. C., Barlow, L. K., Fisher, D., Grootes, P., Jouzel, J., Johnsen, S. J., Stuiver, M., and Clausen, H.: The climate signal in the stable isotopes of snow from Summit, Greenland: Results of comparisons with modern climate observations, Journal of Geophysical Research: Oceans, 102, 26 425–26 439, https://doi.org/10.1029/97JC00162, http://dx.doi.org/10.1029/97JC00162, 1997.

Woollings, T., Lockwood, M., Masato, G., Bell, C., and Gray, L.: Enhanced signature of solar variability in Eurasian winter climate, Geophysical Research Letters, 37, n/a–n/a, https://doi.org/10.1029/2010GL044601, http://dx.doi.org/10.1029/2010GL044601, l20805, 2010.

Zambri, B. and Robock, A.: Winter warming and summer monsoon reduction after volcanic eruptions in Coupled Model Intercomparison Project 5 (CMIP5) simulations, Geophysical Research Letters, 43, 10,920–10,928, https://doi.org/10.1002/2016GL070460, http://dx.doi.org/10.1002/2016GL070460, 2016GL070460, 2016.

Zambri, B., LeGrande, A. N., Robock, A., and Slawinska, J.: Northern Hemisphere winter warming and summer monsoon reduction after volcanic eruptions over the last millennium, Journal of Geophysical Research: Atmospheres, 122, 7971–7989, https://doi.org/10.1002/2017JD026728, http://dx.doi.org/10.1002/2017JD026728, 2017JD026728, 2017.

Zanchettin, D., Khodri, M., Timmreck, C., Toohey, M., Schmidt, A., Gerber, E. P., Hegerl, G., Robock, A., Pausata, F. S. R., Ball, W. T., Bauer, S. E., Bekki, S., Dhomse, S. S., LeGrande, A. N., Mann, G. W., Marshall, L., Mills, M., Marchand, M., Niemeier, U., Poulain, V., Rozanov, E., Rubino, A., Stenke, A., Tsigaridis, K., and Tummon, F.: The Model Intercomparison Project on the climatic response to Volcanic forcing (VolMIP): experimental design and forcing input data for CMIP6, Geoscientific Model Development, 9, 2701–2719, https://doi.org/10.5194/gmd-9-2701-2016, https://www.geosci-model-dev.net/9/2701/2016/, 2016.

Zheng, M., Sjolte, J., Adolphi, F., Vinther, B. M., Steen-Larsen, H. C., Popp, T. J., and Muscheler, R.: Climate information preserved in seasonal water isotope at NEEM: relationships with temperature, circulation and sea ice, Climate of the Past Discussions, 2018, 1–20, https://doi.org/10.5194/cp-2018-8, https://www.clim-past-discuss.net/cp-2018-8/, 2018.

Table 1. List of forcings for the ECHAM5/MPI-OM model simulation.

Greenhouse gases (CO_2, CH_4, N_2O)	MacFarling Meure et al. (2006)
Greenhouse gases (historical, anthropogenic)	Marland et al. (2003)
Ozone	Climatology of Paul et al. (1998)
Volcanic aerosols	Crowley et al. (2008)
Aerosol forcing	Background from Tanre et al. (1984) and post 1850 variations by Lefohn et al. (1999)
Total solar irradiance	Based on Muscheler et al. (2016, 2007) (see also Methods)
Land-use	Pongratz et al. (2008) with vegetation from Jungclaus et al. (2010) E1 ensemble member mil0010
Orbital forcing	Variation Seculaires des Orbites Planetaires (VSOP) analytical solution by Bretagnon and Francou (1988)

Table 2. Site details of ice core records used for the reconstruction (Vinther et al., 2010).(*) BC indicates the at the core was drilled to bedrock.

Drill site	Lat. (^{o}N)	Long. (^{o}W)	Elevation (m a.s.l.)	Acc. Rate (m ice/yr)	Time span
Crete	71.12	37.32	3172	0.289	551-1974
DYE-3 71	65.18	43.83	2480	0.56	1239-1971
DYE-3 79	65.18	43.83	2480	0.56	BC*-1979
GRIP 89-1	72.58	37.64	3238	0.23	918-1989
GRIP 89-3	72.58	37.64	3238	0.23	BC*-1989
GRIP 93	72.58	37.64	3238	0.23	1062-1993
Milcent	70.30	44.50	2410	0.54	1173-1973
Renland	71.27	26.73	2350	0.50	BC*-1988

Table 3. Correlations of reconstructed NAO and PC2 of reconstructed SLP, and observed NAO, 20CR PC2 and PC3 of SLP, as well as NAO reconstructions by Ortega et al. (2015) and Luterbacher et al. (2004). All correlations are for detrended data, and p-values are calculated with the random-phase test by Ebisuzaki (1997) to take into account auto-correlation.

	Annual (DJF)	10-year low-pass	20-year low-pass
Corr(NAO$_{recon}$, HadCRU) 1824-1970	0.52 ($p<0.01$)	0.68 ($p<0.01$)	0.70 ($p<0.01$)
Corr(NAO$_{recon}$, NAO$_{20CR}$) 1851-1970	0.44 ($p<0.01$)	0.44 ($p<0.01$)	0.46 ($p<0.01$)
Corr(SLP$_{PC2-recon}$, SLP$_{PC2-20CR}$) 1851-1970	0.24 ($p<0.01$)	0.53 ($p<0.01$)	0.58 ($p<0.01$)
Corr(SLP$_{PC2-recon}$, SLP$_{PC3-20CR}$) 1851-1970	0.19 ($p<0.01$)	0.57 ($p<0.01$)	0.66 ($p<0.01$)
Corr(SLP$_{PC2-recon}$, SLP$_{PC2+PC3-20CR}$) 1851-1970	0.30 ($p<0.01$)	0.67 ($p<0.01$)	0.84 ($p<0.01$)
Corr(NAO$_{recon}$, Ortega et al. MC) 1241-1969	0.49 ($p<0.01$)	0.43 ($p<0.01$)	0.37 ($p<0.05$)
Corr(NAO$_{recon}$, Ortega et al. MC) 1241-1820	0.47 ($p<0.01$)	0.36 ($p<0.05$)	0.15 ($p<0.2$)
Corr(NAO$_{recon}$, Ortega et al. CC) 1241-1969	0.36 ($p<0.01$)	0.35 ($p<0.01$)	0.40 ($p<0.01$)
Corr(NAO$_{recon}$, Ortega et al. CC) 1241-1820	0.29 ($p<0.01$)	0.21 ($p<0.05$)	0.12 ($p<0.2$)
Corr(NAO$_{recon}$, Luterbacher) 1659-1970	0.34 ($p<0.01$)	0.39 ($p<0.01$)	0.40 ($p<0.01$)

Table 4. Correlation between solar forcing and PC2 of SLP, with and without time lag. The first column indicate which data is used, and if any filtering is done to the data. Second column is correlation coefficients between solar forcing and PC2 of reconstructed SLP, with solar forcing either being represented by sunspot number or ^{14}C data. The third column is correlation coefficients between solar forcing and PC2 of reconstructed SLP, with solar forcing represented by sunspot number shifted for a lag of 5 years. All correlations are for detrended data, and p-values are calculated with the random-phase test by Ebisuzaki (1997) to take into account auto-correlation.

	No time lag	5-year time lag
Corr(Recon. PC2 SLP, sunspots) 1700-1970	-0.06 ($p>0.1$)	0.20 ($p<0.01$)
5-year low-pass filtered data: Corr(Recon. PC2 SLP, sunspots) 1700-1970	-0.07 ($p>0.1$)	0.29 ($p<0.01$)
20-year low-pass filtered data: Corr(Recon. PC2 SLP, sunspots) 1700-1970	0.30 ($p<0.05$)	0.53 ($p<0.01$)
20-500 year band-pass filtered data: Corr(Recon. PC2 SLP, solar activity (^{14}C)) 1241-1970	0.30 ($p<0.01$)	-
60-500 year band-pass filtered data: Corr(Recon. PC2 SLP, solar activity (^{14}C)) 1241-1970	0.60 ($p<0.01$)	-

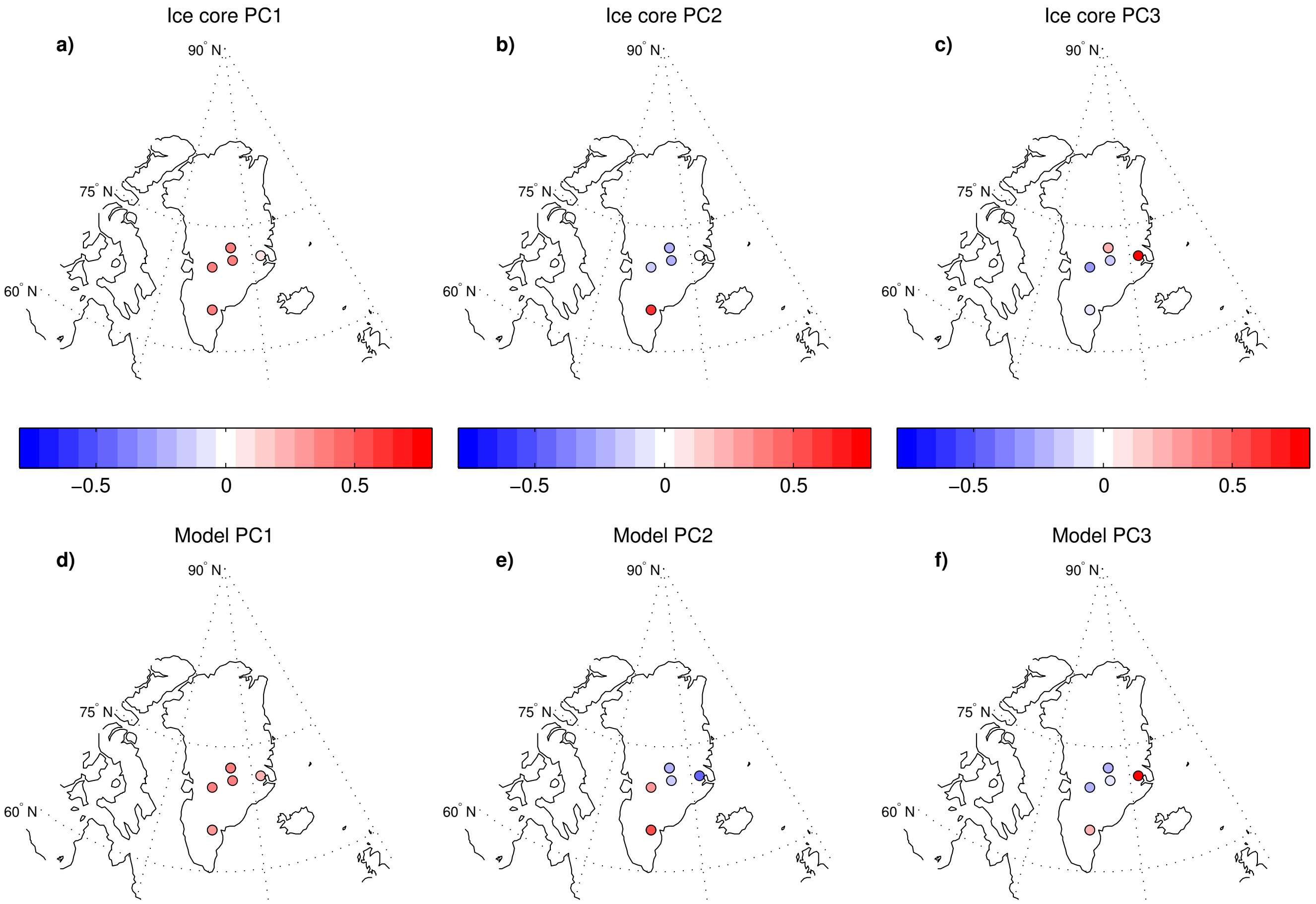


Figure 1. Spatial patterns of the three main modes of variability in the δ^{18}O ice core records used in this study and the modeled modes of variability at these sites. The pattern of the loadings on ice core δ^{18}O PCs are shown in (**a, b, c**) and modeled loadings on δ^{18}O PCs at the ice core sites are shown in (**d, e, f**).

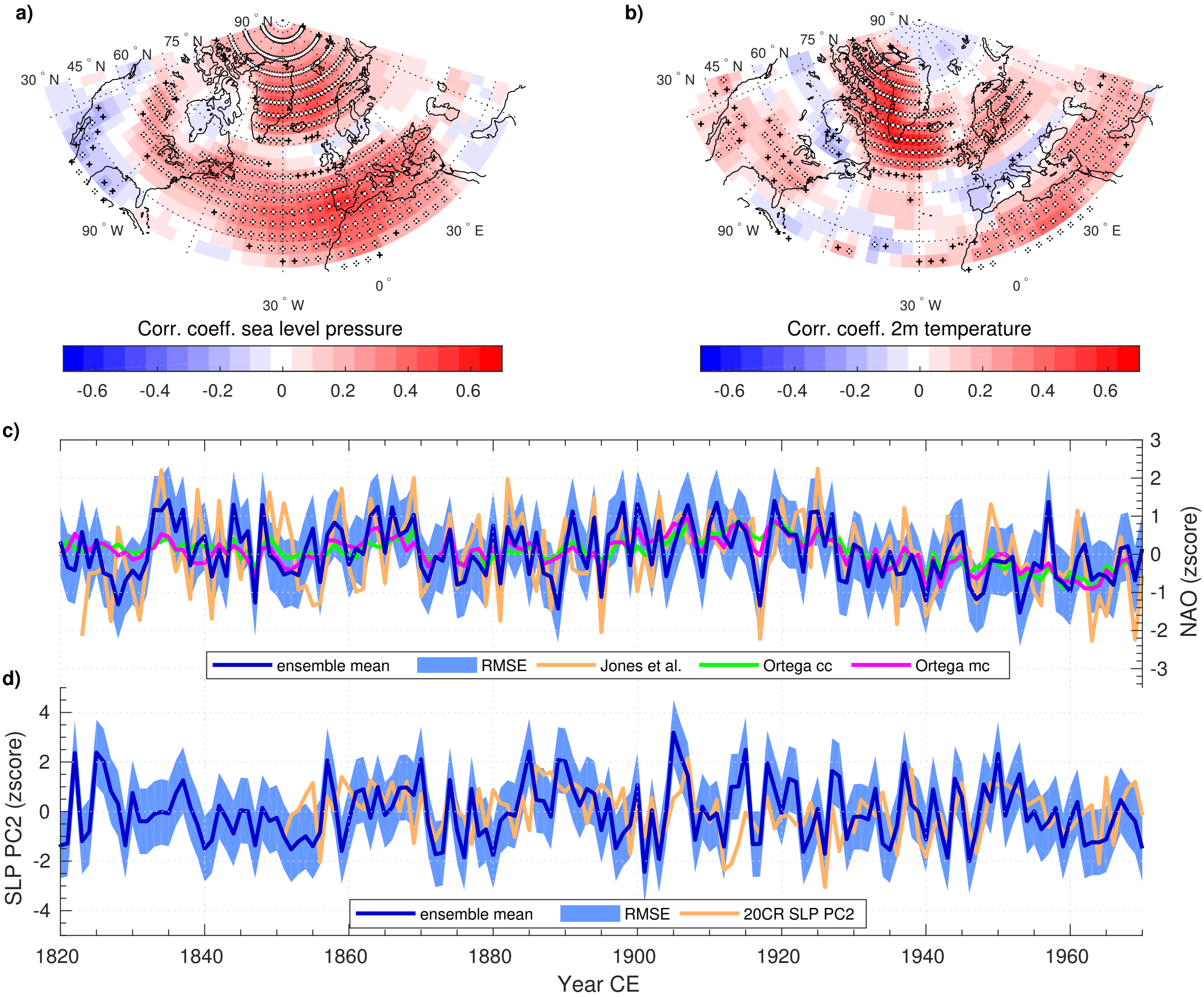


Figure 2. Evaluation of the winter circulation reconstruction. Grid point correlation between reconstructed DJF SLP (**a**) and T2m (**b**), and reanalysis data (Compo et al., 2011) (1851-1970) interpolated to the model grid (lat. x lon. ∼~~3.85~~3.75^o x ~~3.85~~3.75^o). The white stippling indicates significance $p < 0.05$, and black stippling indicates significance $p < 0.1$. **c** Ensemble mean reconstructed DJF NAO (PC1 of reconstructed DJF SLP (Hurrell et al., 2013)) with Root Mean Square Error (RMSE) compared to observed DJF NAO (Jones et al., 1997), NAOcc and NAOmc by Ortega et al. (2015). (**d**) Ensemble mean PC2 of reconstructed DJF SLP with RMSE compared to PC2 of reanalysis DJF SLP (Compo et al., 2011).

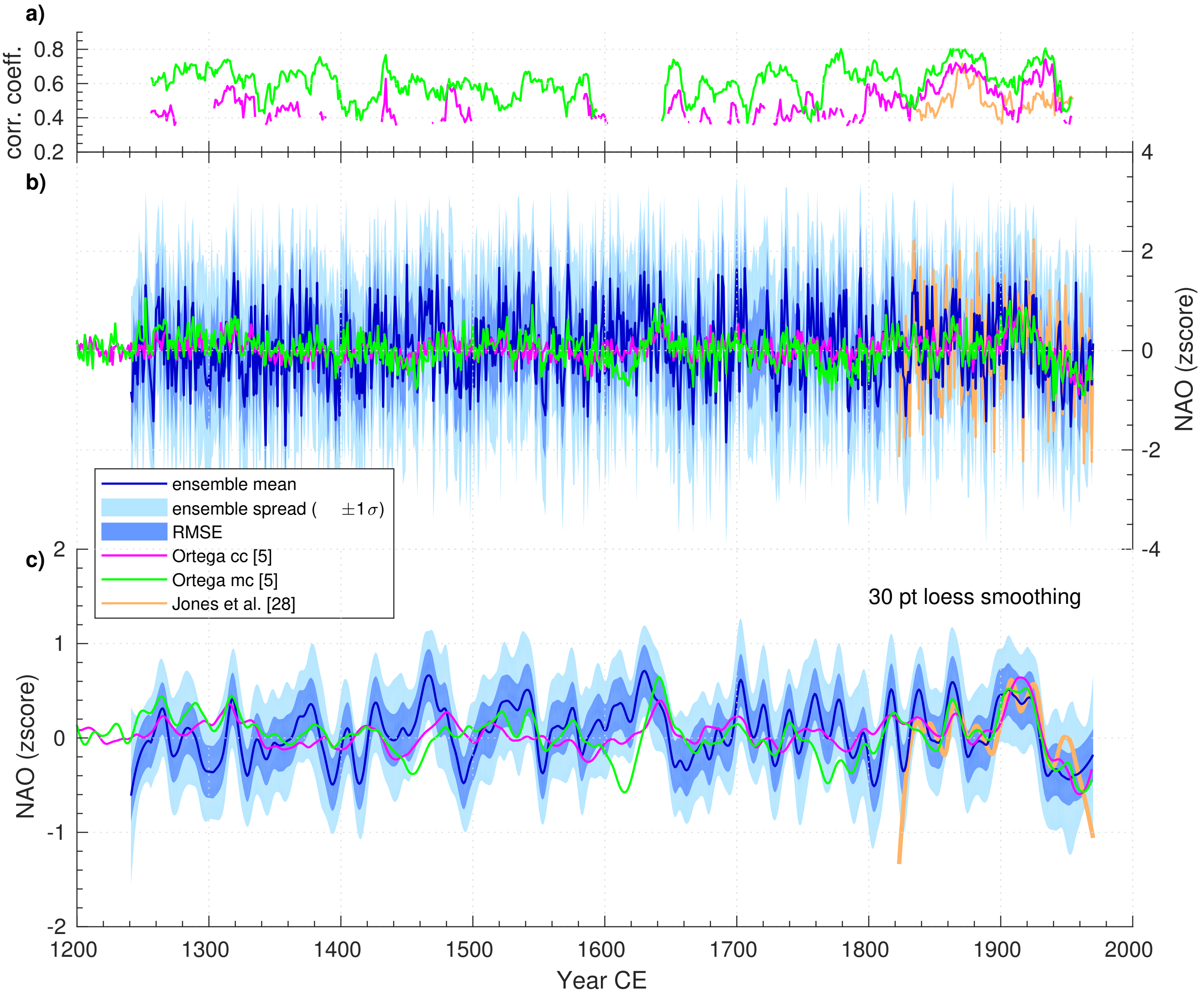


Figure 3. Comparison of instrumental NAO and proxy-based NAO reconstructions. **a** Moving 31-point correlation between reconstructed NAO from this study and NAOcc (magenta), NAOmc (green) (Ortega et al., 2015) and observed NAO (yellow) (Jones et al., 1997). Only significant correlations are plotted ($p < 0.05$). **a** Ensemble mean reconstructed NAO (PC1 of reconstructed SLP (Hurrell et al., 2013)) with error estimated by ensemble spread and RMSE, compared to observed NAO (Jones et al., 1997) and NAO reconstructions by Ortega et al. (2015). The amplitude of all time series are scaled to fit the decadal variability of the observed NAO. **~~b~~c** Same as **~~a~~b**, except filtered with a 30 point 'loess' filter.

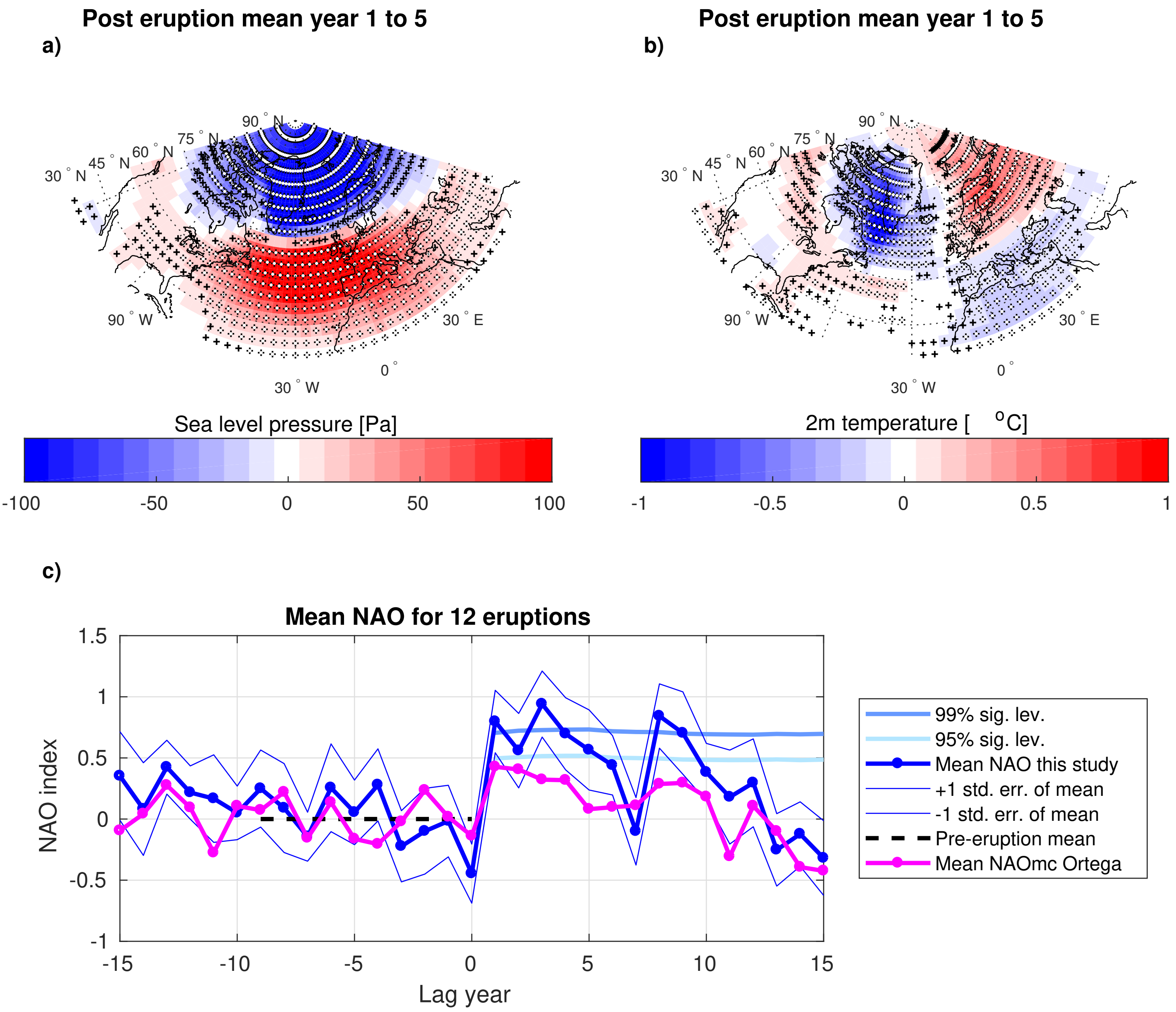


Figure 4. Superimposed epoch analysis of the mean response in atmospheric circulation to the 12 largest tropical volcanic eruptions (Sigl et al., 2015) (Table ~~S3~~S1). The response in SLP and T2m is normalized to the mean fields of the 10 years preceding the eruption. **a** Mean DJF SLP anomalies [Pa] for the first five post eruption years. **b** Mean DJF T2m anomalies [^{o}C] for the first five post eruption years. The white stippling indicates significant anomalies $p < 0.01$, and black stippling indicates significant anomalies $p < 0.05$ (two-tailed Student's *t*-test). **c** Mean response in reconstructed NAO (blue) with the time series normalized to the mean NAO of the 10 years preceding the eruption. For comparison the same analysis is carried out for the NAOmc reconstruction (magenta) by Ortega et al. (2015). The significance levels in **c** are estimated from 100,000 random samples of 12 years drawn from the reconstructed NAO. See Figure S6 for significance levels for NAOmc.

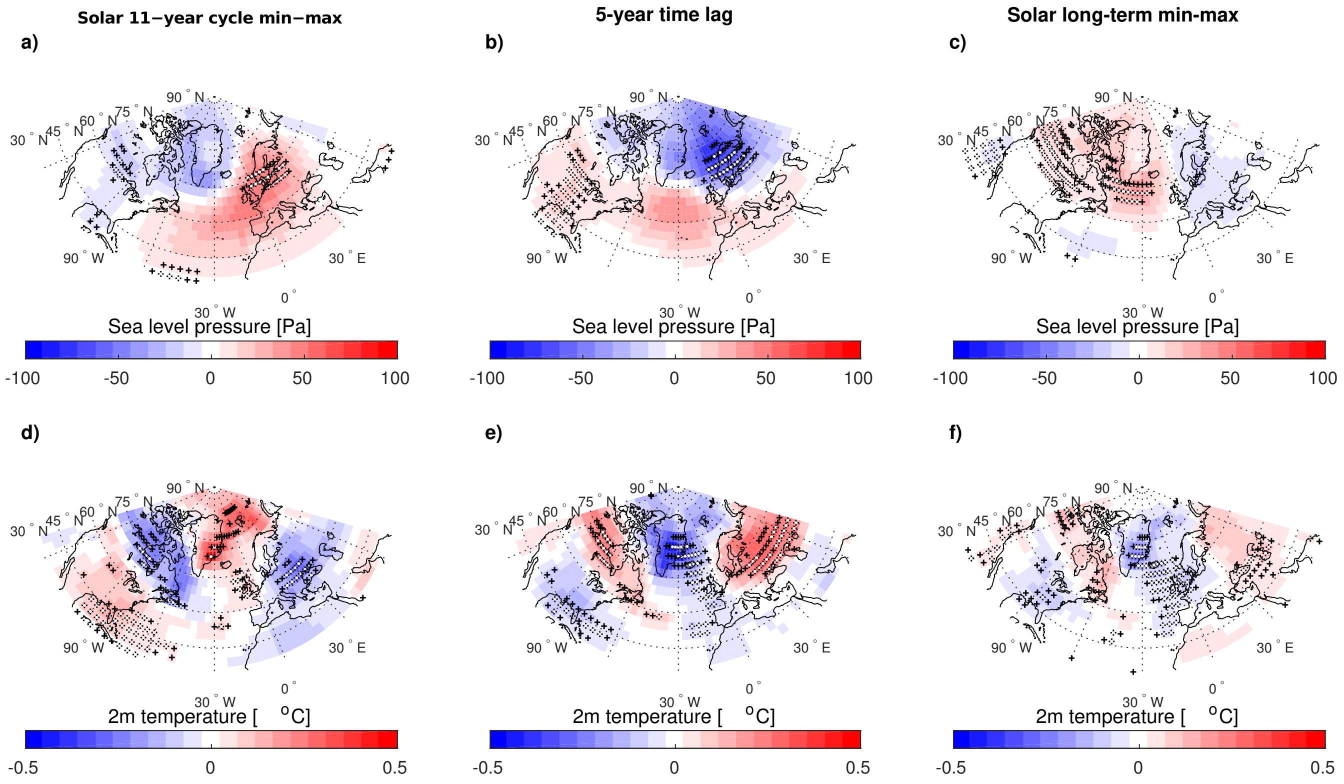


Figure 5. Reconstructed atmospheric response to solar forcing. **a** DJF SLP anomalies [Pa] in response to the 11-year solar cycle (solar min. minus solar max. defined in Figure ~~S8~~S10). **b** DJF SLP anomalies [Pa] in response to the 5-year lagged 11-year solar cycle (solar min. minus solar max.). **c** DJF SLP anomalies [Pa] in response to the long-term solar forcing (solar min. minus solar max. defined in Figure ~~S9~~S11). **d, e, f** corresponding figures to **a, b, c**, but for T2m [^{o}C]. The white stippling indicates significant anomalies $p < 0.05$, and black stippling indicates significant anomalies $p < 0.1$ (two-tailed Student's *t*-test).

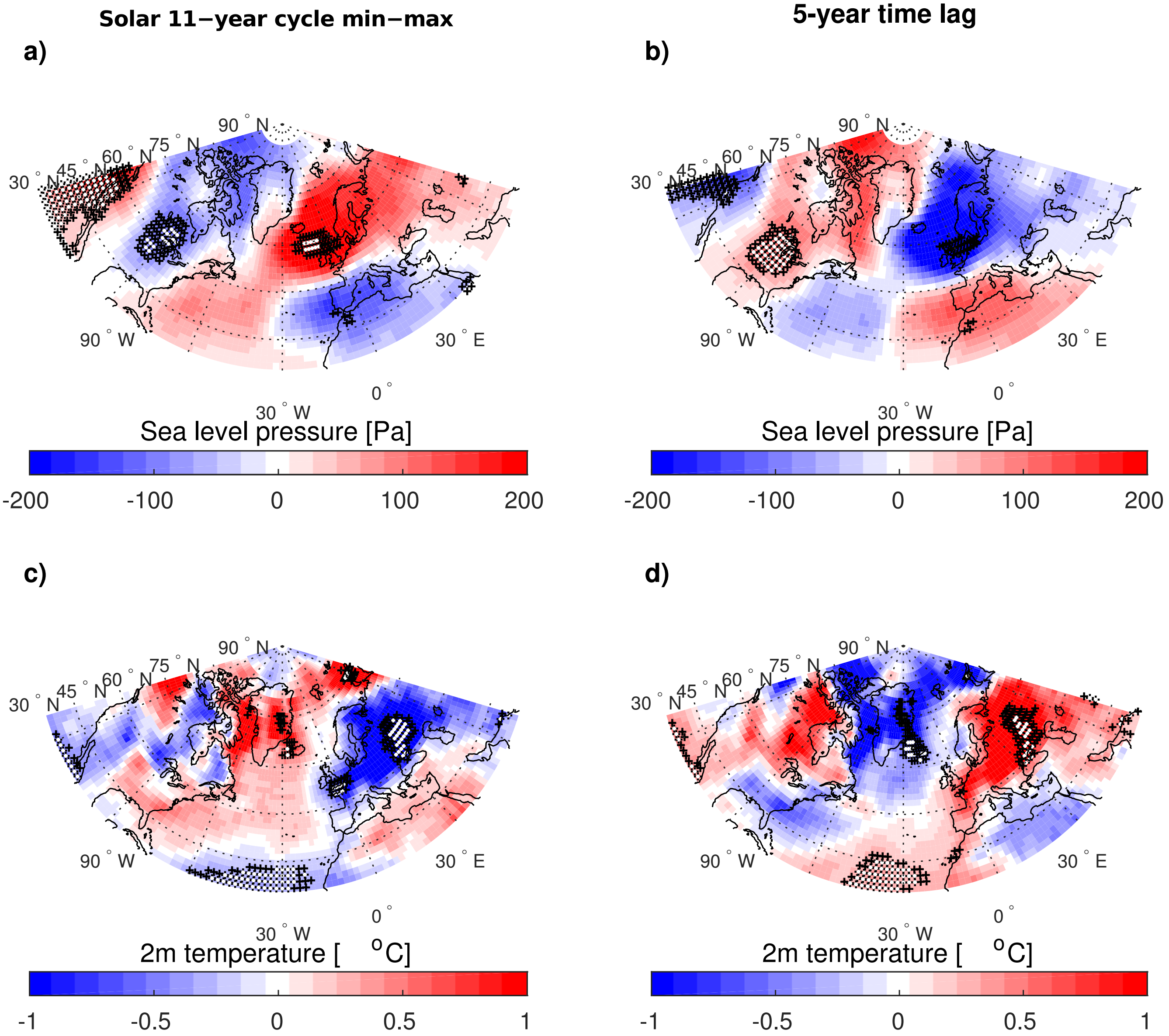


Figure 6. 20CR (1948-2010) atmospheric response to solar forcing. **a** DJF SLP anomalies [Pa] in response to the 11-year solar cycle (solar min. minus solar max. defined in Figure ~~S10~~S12). **b** DJF SLP anomalies [Pa] in response to the 5-year lagged 11-year solar cycle (solar min. minus solar max.). **c, d** corresponding figures to **a, b**, but for T2m [^{o}C]. The white stippling indicates significant anomalies $p < 0.05$, and black stippling indicates significant anomalies $p < 0.1$ (two-tailed Student's *t*-test). The time interval for this analysis in limited to 1948-2010 due to limitation of the data quality prior to this, although similar results can be achieved for the period 1851-2010.

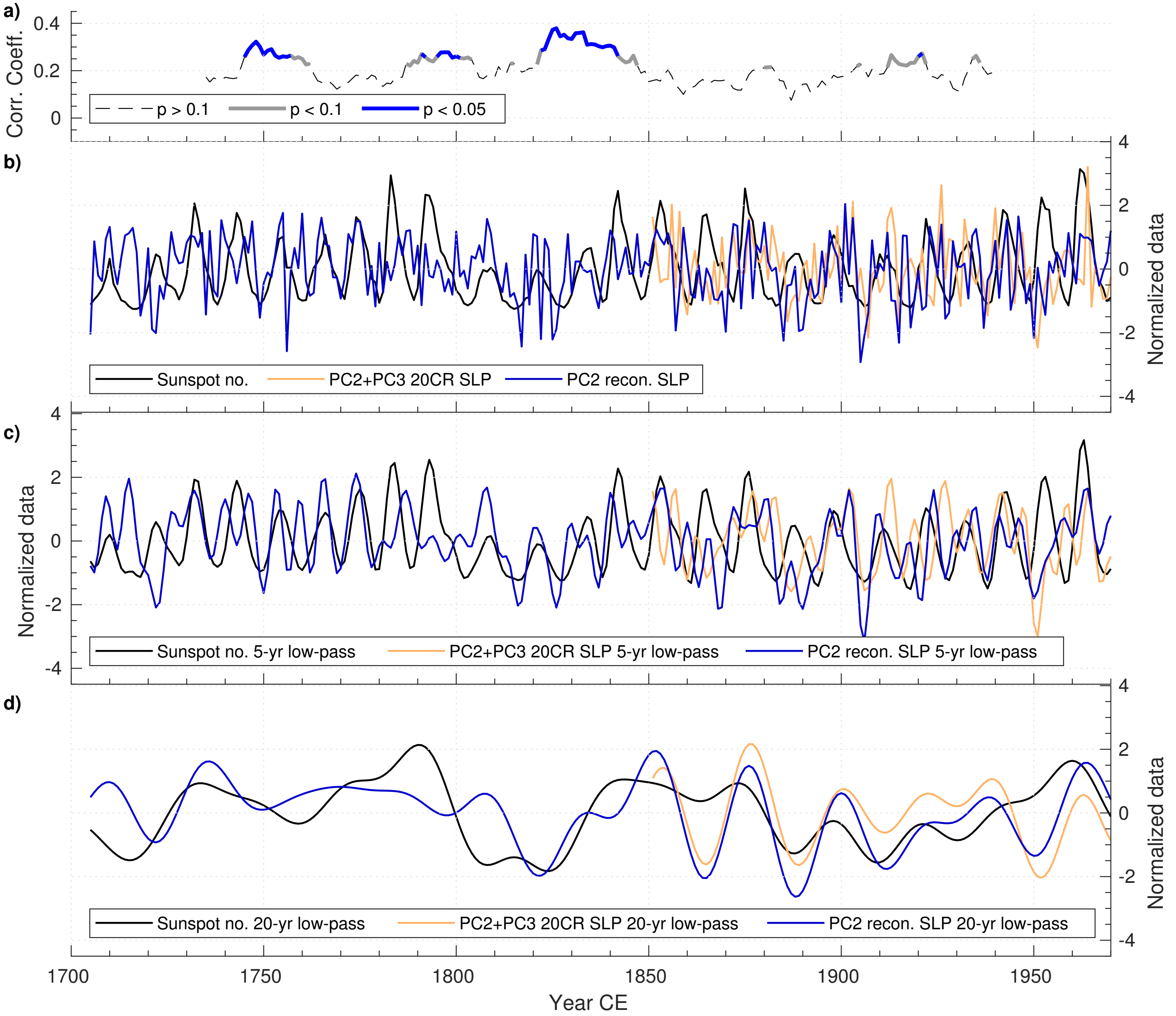


Figure 7. Reconstructed PC2 of SLP plotted with the 5-year lagged sunspot number. **a** Moving 61-point correlation between reconstructed PC2 of SLP and sunspot number shifted for a 5-year time lag. **b** Time series of sunspot number shifted for a 5-year time lag, PC2+PC3 of 20CR SLP (see text and Table ~~2~~4) and PC2 of reconstructed SLP. **~~b~~c** Same as **~~a~~b**, except filtered with a 5-year low-pass filter. **~~c~~d** Same as **~~a~~b**, except filtered with a 20-year low-pass filter.