

Reply to reviewer comment #2

In this study the authors use high-resolution ice-core records of aerosols, water isotopes, and layer thickness from Greenland to examine phasing of different aspects of the climate system during Dansgaard-Oeschger events. They use objective techniques to tease out small leads and lags between the noisy records, showing that changes in calcium aerosols and layer thickness lead changes in sodium aerosols and water isotope ratios. They conclude that these lags suggest that changes in sea ice extent did not occur before other changes in the climate system, namely atmospheric circulation. The manuscript is very well written, the analysis is careful, and the discussion is well-argued. The results are quite impressive and should be of wide interest to the community. I have a couple questions and concerns that I hope the authors will address and then several minor questions.

Thank you!

(somewhat) Major questions: The authors have made a compelling observation through their analysis of leads and lags between the proxy records. They attribute these differences in timing to aspects of the climate system that have different influence on the proxies. My first two questions have to do with whether these leads and lags could arise from other factors. I suspect that these concerns are not too important to the conclusions of this study.

1. How does the signal-to-noise ratio of the record influence fitting the transition model and the identification of starting, mid, and end points? Here I'm thinking of the SNR as quantified by the size of the transitions compared to the variance within the stadials and interstadials. I realize the fitting procedure takes into account the interannual noise and its autocorrelation. But if you have an idealized known ramp function with different levels of background noise, will the model find the same starting, mid, and endpoints?

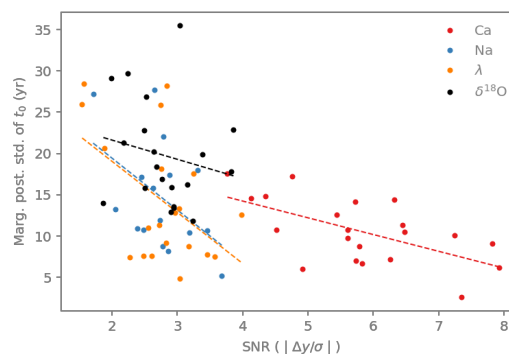
One could imagine that an increased background noise could lead to the identification of time-shifted transition points depending on the fitting technique. One then risks conflating difference in the timing of signals between records with difference in one's ability to detect signals between records. I cannot tell from the description of the transition model alone how much of an issue this is to this analysis. Doing some simple tests with a couple different ramp-fitting and significant change detection techniques (though ones less sophisticated than the technique used by the authors), I find that different levels of interannual noise can influence the timing of a fitted transition, though not in all circumstances.

My particular worry here is that the substantially higher noise (interannual variability) in the Na records could lead to the identification of a delayed onset or shorter transitions compared to the Ca and other records. My worry is heightened slightly in that the relative timing seems to correspond to the level of background noise (at least visually): the $\delta^{18}\text{O}$ and Na timing are most similar among the proxies and also both appear to have much lower SNR (more noise) compared to the Ca and layer thickness records.

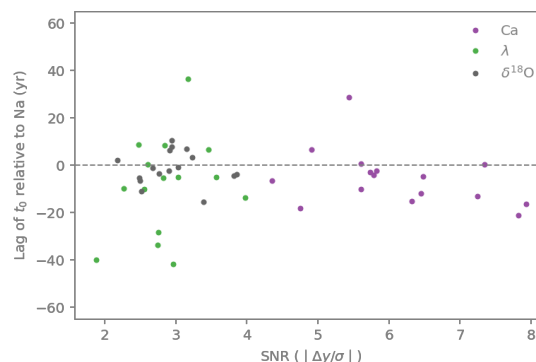
Do the mean lags depend on the amplitude of the DO event or the length of the transition between stadial and interstadial? This could be the case if the ramp-fitting depends on the amplitude of background noise. I realize this could be hard to determine since one needs to look at many events at once to see the mean lags. But a scatter plot of lags vs. event magnitude or fitted ramp duration could be informative.

I suspect that this concern is entirely accounted for by the very careful analysis of the marginal posterior densities of the onsets, midpoints, and endpoints for each proxy and the comparison between proxies, that the authors have already performed. It would however be helpful to see the influence, or the demonstration of the absence of influence, of the SNR on the fitting procedure given that the conclusions rest on the difference in timing with respect to Na in particular. I'd find it very useful to see this demonstrated on artificial ramp signals (of varying duration) where the true timings are known explicitly, with varying SNR, and especially with the SNRs relevant to d18O, lambda, Ca, and Na. It seems crucial to know that different SNR alone can not account for the difference in timing identified by the fitting procedure.

We do completely agree with this concern and thank you for bringing this up. You are correct in that the signal to noise ratio as defined by the amplitude of the ramp divided by residual standard deviation varies widely between the different parameters and that Ca has by far the highest SNR of the records. In our analysis this is reflected in the posterior standard deviations of the timing estimates for the individual parameters: The higher the SNR, i.e. the clearer the transition, the lower the posterior standard deviation. To illustrate this, the Figure below shows the marginal posterior standard deviations of the start point of the transitions as functions of the SNR for the NGRIP record.



As we propagate these uncertainties to the lead/lag calculations and the combined estimates we take this into account. This is also clear when plotting the leads/lags relative to Na as a function of the SNR in the respective parameter which does not show any relationship with the SNR:



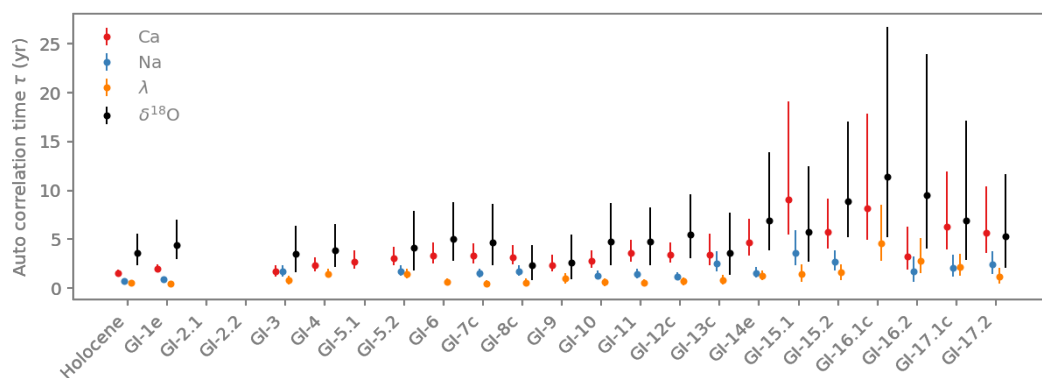
We have added this discussion and the two Figures above to the Appendix of the manuscript.

2. Can the authors rule out the influence of water isotope diffusion on the difference in timing of the d18O signals and those of Ca and layer thickness? Based on analysis of NGRIP (Gkinis et al 2014) and a similar site in Antarctica (Jones et al 2017, 2018), I'd guess diffusion lengths are on the order of 5-10cm through this interval and so are not insignificant compared to the annual layer thickness. Such diffusion lengths can have meaningful influence on the inter-annual and even decadal variability in the water isotope record (Jones et al 2018). I imagine that correcting the records used here for the potential influence of diffusion, if that would even be sensible, is far beyond the scope of this study. However, it seems entirely reasonable to estimate the influence (if any) of the smoothing implied by diffusion on the timing identified by the transition model fitting procedure. If you take identical idealized ramps, and smooth one with a time-scale reflective of water isotope diffusion lengths, will the fitting procedure identify the same start, mid, and end points? I suspect these effects, if any, are small, though we are only talking about lags of a few years.

Yes, the smoothing by the isotope diffusion would influence the shape of the transition. It is however difficult to say what the exact effect on the estimated onsets would be as it also increases the SNR by reducing the high-frequency variability that is captured by the noise term.

That being said, in the case of DO onsets, where the d18O signal rapidly increases, any diffusion/smoothing of that signal would yield to a slight shift towards earlier onsets, i.e. would reduce the lag between Ca and d18O.

Nevertheless, this is an interesting and important thing to check. To do so, we examined the auto-correlation time that is estimated by the fitting algorithm. Any diffusion of the signal would lead higher auto-correlation times estimated for d18O than for all the other parameters, even though the AR(1) process is not an ideal description for the autocorrelation introduced through gaussian smoothing.



This Figure shows these estimates for all parameters in NGRIP. In comparison to Ca, d18O exhibits similar auto-correlation lengths throughout the record. Indeed all the records show an increase in auto-correlation length with increasing depth with is a direct consequence of the decrease in the resolution in terms of years due to the layer thinning.

We have added the following to the discussion (P12, L29 ff)

“Due to the small timing differences between the records it is worth noting that water isotope records from polar ice cores are subject to smoothing by diffusion. For the NGRIP isotope record the diffusion length at the end of the last glacial is on the order of five to ten centimeters (Gkinis et al., 2014), influencing the high frequency variability. For the analysis here the diffusion means that the rapid increase of the

$\delta^{18}\text{O}$ signal at the D–O onsets would be slightly shifted towards earlier times, leading to lower apparent lags between the aerosol records and the water isotope record. Thus, the inferred lead of Ca over $\text{d}18\text{O}$ can be regarded as a conservative estimate.”

Gkinis, V. et al. (2014), Water isotope diffusion rates from the NorthGRIP ice core for the last 16,000 years – Glaciological and paleoclimatic implications, Earth and Planetary Science Letters, 405132-141, doi:10.1016/j.epsl.2014.08.022

3. The most interesting conclusion of this study to my mind is the authors statement that "at face value, this sequence of events suggests that the collapse of North Atlantic sea-ice cover is not the initial trigger for the DO events..." Because of its potential wide interest, this statement deserves some scrutiny. It rests on the authors use of Na as a "qualitative indicator of sea ice cover" in the North Atlantic.

The discussion on Page 5, lines 3-28, highlights the debate over in the interpretation of Na very well. The authors interpretation is laid out on Page 5 lines 20-25. Markle et al 2018 find that most of the millennial variability in Antarctic sea salts can be explained simply by the changes in moisture rainout that are required to explain the water isotope record (these changes also explain most of the changes in Antarctic Ca variability and its relationship to water isotopes in both Antarctica and Greenland (c.f. their Figure 4)). This suggests comparatively small if any changes in sea salt source latitude or strength are needed to explain those Na records (though changes in those things are still possible of course). Is there evidence that this explanation for the sea salts is insufficient in Greenland? Are the observed changes in sea salt for example much larger than what one would expect from temperature dependent rainout alone? It was unclear to me from this discussion that the sea salt source strength (or even mean source latitude) should have a clear relationship to the sea ice edge.

Generally, for Greenland the same mechanism as described by Markle et al. (2018) explains most of the millennial-scale variability as shown in Schüpbach et al. (2018). However, applying the model of Markle et al. (2018) to the DO events would completely preclude any leads or lags between $\text{d}18\text{O}$ and the aerosols as well between the different aerosols. On shorter timescales dynamical changes in the (co-) transport of aerosols and moisture seem to play a relatively larger role in determining the concentrations observed in the ice cores. Here, this results in the lack of co-evolution between the investigated proxies and can explain the lack of coherence between the records in Markle et al. 2018 at the sub-millennial timescales.

Schüpbach, S. et al. (2018), Greenland records of aerosol source and atmospheric lifetime changes from the Eemian to the Holocene, Nature Communications, 9, 1-10, doi:10.1038/s41467-018-03924-3.

If the main way sea ice influences Greenland Na is through its relationship to the variables driving the rainout effect, then a change in sea ice extent doesn't necessarily mean one should have a coincident change in Greenland Na, particularly at interannual timescales.

For example, one can imagine a scenario in which sea ice extent begins to retreat at exactly the same time as the changes observed in Ca initiate. Coincident increases in temperature and moisture removal would cause a decrease in the amount of Na (and Ca) reaching Greenland (as described by the authors). However, if that change in sea ice extent caused an increase in sea salt source production or a northward migration of mean source latitude (both debatable

but reasonable, particularly the latter) this could temporally compensate for the increased removal. This combination of influences could lead to an apparent timing difference in the final Na signal observed in Greenland compared to the actual timing of sea ice changes (an example of the superposition of competing source and rainout factors on polar aerosols is given in the Supplement of Markle et al 2018 c.f. Figures S9). A somewhat similar scenario may be likely for the water isotopes, as changing sea ice extent could drive moisture source effects that could temporarily compensate for the decreased depletion driven by simultaneous changes over the ice sheet (these would be of the correct sign to compensate, though it would be hard to assess the potential size of this effect on the isotopes without analyzing deuterium excess records from the same cores). Quantitatively disentangling these influences may be well outside the scope of this study. But this does at least suggest a limitation to using Na as a qualitative indicator of sea ice, and should suggest some commensurate caution in the conclusions drawn based on that interpretation.

Even in the absence of competing influences, uncertainty in the linearity of the sodium-as-sea-ice-extent interpretation poses challenges. Does a change in the sea ice edge (or extent) of a given size have the same impact on Na in Greenland if the sea ice edge is at 55 degrees North (just for example) versus if the edge is at 65 degrees North? If relationship between changes in Greenland Na and the sea ice absolute position is (sufficiently) nonlinear then the changes in sea ice at the start of DO events may not be as detectable in Greenland Na as changes later in the event. This could lead to apparent lags of the Na signal to Ca, even if the change in sea ice initiates at the same time as the change in whatever drives Ca. I'm not sure I think this is particularly likely, but it is certainly plausible. Further, the impact of sea ice on climate isn't linear. Again a given size change in the sea ice edge when the edge is at 55 degrees North (for example) doesn't have the same impact on the surface radiation budget nor the buoyancy forcing of the overturning circulation as when the edge is at 65 degrees North. The relative timing of sodium changes in Greenland don't necessarily rule out sea ice changes as a potential "trigger" of DO events, if climatically meaningful sea ice changes can happen without influencing Greenland Na.

To be sure, there are innumerable, potentially ad hoc, explanations for the data, and it is not the authors responsibility to come up with and then weigh the merits of all such explanations. However the assertion that the lag in sodium seen in Greenland necessitates a lag in changes in sea ice extent with respect to shifts in atmospheric circulation, and that this in turn rules out sea ice changes as the trigger for DO events, is somewhat bold. That this "provides an essential benchmark for climate models" is bolder still. Both of these statements to my mind somewhat outpace the robustness in the interpretation of Na as a qualitative indicator of sea ice extent. I think some slight tempering of the conclusions is merited here and perhaps some discussion of the limitations to the interpretation, or much more compelling evidence is needed that a change in sea ice extent must be seen in the Greenland Na records regardless of other processes. To be clear, I do think the difference in timing of the aerosol species identified here is a compelling target for modeling. And I do think the climatic interpretation offered by the authors is plausible, and one to be taken seriously. It is just not clear to me that these data alone place that strong of constraints on the timing of sea ice changes.

We agree with the possible complicating factors you pointed out. However, it is difficult to imagine that a "climatically meaningful" sea-ice change can happen without influencing neither Na nor d18O in Greenland ice through the changing source signals, en-route rainout or changes in seasonality or a combination of these factors: Taking the transient DO simulation performed by Vettoretti and Peltier (2015,

2018) as an example, the change in the sea-ice coverage in the North Atlantic results in an immediate increase in moisture cycling (evaporation and precipitation) over the open ocean as well as an increase in temperature in Greenland. Nevertheless, we agree that barring any direct simulations using both isotope and aerosol models over DO-events, the direct test of our hypothesis will be very difficult. To account for this limitation, we have rephrased the relevant statements in the conclusions to be more careful (P13, L11 ff)

“The progression of environmental changes revealed in the Greenland aerosol records provides a good target for climate models that aim to transiently simulate DO events, preferentially explicitly modeling both water isotope and aerosol transport.”

Vettoretti, G., and Peltier, W. R. (2015), Interhemispheric air temperature phase relationships in the nonlinear Dansgaard-Oeschger oscillation, Geophysical Research Letters, 42,1180-1189, doi:10.1002/2014GL06289

Vettoretti, G., and Peltier, W. R. (2018), Fast Physics and Slow Physics in the Nonlinear Dansgaard--Oeschger Relaxation Oscillation, Journal of Climate, 31, 3423-3449, doi:10.1175/JCLI-D-17-0559.1

Minor points: Page 1, line 8: The authors say "from one of the cores".... Which one?

We have now clarified the sentence in the abstract.

Page 1 line 17: I think that the clause "In the course of the last glacial period" should be moved after the 2nd comma in this line (after "warming episodes,"). This is a possibly silly language thing but: the ice core records reveal that the events were in the last glacial period. It wasn't during the last glacial period that the ice core records revealed these events (that was in the 1980s!).

Done.

Page 1 line 20: This is a language choice the authors should feel free to ignore but "..going along with an almost doubling..." sounds strange to my ear, though I can't tell why. "...coincident with a near-doubling..." sounds better to me.

Adjusted accordingly.

Page 2 line 3: I'd move the "Also" at the start of the sentence to after "Northern Hemisphere".

Done.

Page 2 Lines 20-35: This is a very nice description of several related though distinct ideas. Thanks!

Thank you!

Page 3 lines 15-20: Are the Ca and Na records corrected for the sea-salt vs non sea-salt components of each (e.g. using average Na/Ca mass ratios in average crust and sea salt, as is

common in Antarctic records)? I ask because having looked into this once briefly, it seemed like the corrections often used in Antarctica lead to nonsensical results in Greenland, which was disturbing, and I wasn't sure why.

No, the records we use here are not "corrected" for the sea-salt and dust contributions. We chose not to do this for two reasons: First and foremost, the Ca/Na ratio in the dust in Greenland is quite different from the crustal average that is typically employed for Antarctic records. The reason for this is the large Ca content of Central Asian Dust. This is also true for Antarctica and has previously been extensively investigated by Bigler et al (2006).

Bigler, M. et al. (2006), Aerosol deposited in East Antarctica over the last glacial cycle: Detailed apportionment of continental and sea-salt contributions, Journal of Geophysical Research, 11110.1029/2005JD006469.

Page 6, Line 9: I think there is a missing "of" between "interpretation" and "phase".

Corrected, thanks.

Page 6, lines 20-21: This is interesting! Nice.

Thanks.

Page 6, lines 27: "on" should be "one" I think.

Corrected, thanks.

Page 6, lines 29: If you don't make the assumption that the DO-events have the same imprint in both cores, how would that effect any of your conclusions?

Making this assumption would allow to combine the estimates from the two cores into one. We chose not to do this to be able to gauge the robustness of our estimates, i.e. whether they agree between the cores.

Page 10, lines 25-30: Might be worth citing Fudge et al 2016 here, they show a strong divergence of accumulation rate and d18O in ice cores at millennial time scales.

Thank you for pointing out this reference. We added it to the manuscript with the qualification that it deals with Antarctica and not Greenland.

Page 11, lines 3-5: This is nice, convincing analysis.

Thank you.

Page 11, lines 11: I think this should be "...a coinciding..." Not "...an coinciding..."

Corrected, thanks.

Page 11, lines 21-22: How big are the additional changes in source need compared to the total variability across a DO event?

To explain the complete amplitude of the stadial/interstadial changes in the Ca concentration in the ice, a source strength change of a factor around 4 is needed. We have added this to the manuscript (P12, L9) and refer the Reviewer to Schüpbach et al. (2018) for further details:

Schüpbach, S. et al. (2018), Greenland records of aerosol source and atmospheric lifetime changes from the Eemian to the Holocene, Nature Communications, 91-10, doi:10.1038/s41467-018-03924-3.

Page 11, lines 24-35: This is really nice discussion.

Thank you.

Refs cited: Markle, Bradley R., et al. "Concomitant variability in high-latitude aerosols, water isotopes and the hydrologic cycle." *Nature Geoscience* (2018): 1.

Jones, Tyler R., et al. "Southern Hemisphere climate variability forced by Northern Hemisphere ice-sheet topography." *Nature* 554.7692 (2018): 351.

Jones, T. R., et al. "Water isotope diffusion in the WAIS Divide ice core during the Holocene and last glacial." *Journal of Geophysical Research: Earth Surface* 122.1 (2017): 290-309.

Fudge, T. J., et al. "Variable relationship between accumulation and temperature in West Antarctica for the past 31,000 years." *Geophysical Research Letters* 43.8 (2016): 3795-3803.

Gkinis, Vasileios, et al. "Water isotope diffusion rates from the NorthGRIP ice core for the last 16,000 years—Glaciological and paleoclimatic implications." *Earth and Planetary Science Letters* 405 (2014): 132-141.