Dr. Frank McDermott

Associate Editor

*Geochimica et Cosmochimica Acta*

October 9, 2017

Dear Dr. McDermott,

Re: Manuscript reference No.GCA-D-17-00136

Please find attached a revised version of our manuscript “Molecular diffusion of stable water isotopes in polar firn as a proxy for past temperatures”, which we would like to re-submit for publication as an article in *Geochimica et Cosmochimica Acta*.

We thank you and the three reviewers for your comments and suggestions which has helped improving the manuscript. Specific comments to each reviewer is provided on the next pages. Text in bold is the reviewer’s comment. Besides language improvements, two major changes are implemented in the revised manuscript:

1. **A seasonal temperature signal** that propagates down through the firn is included in the diffusion and densification model (Sec. 2 and Appendix B in manuscript). The seasonal temperature variation affects the firn diffusion length nonlinearly due to the saturation vapor pressure exponential dependence on temperature. The firn temperature profile is obtained by numerical solution of the heat equation. In the previous manuscript, an isothermal firn column was assumed.
2. **Uncertainties in the diffusion, densification and ice flow models** are now included. Previously, the presented precisions were based on the estimation of the diffusion length from data. It was therefore assumed that the accumulation rate, surface density, close-off density, surface pressure, ice flow thinning and ice diffusion were known (they had a constant value for each ice core section). In the revised manuscript, each of these parameters have an uncertainty used in the temperature reconstructions (Table 4.1 in paper). This facilitates a better comparison between the single and differential diffusion techniques in case the single diffusion length methods are more sensitive to uncertainties in the diffusion model.

Although the model used in this study has changed significantly by implementing a seasonal signal and including uncertainties in the diffusion, densification and ice flow model, the main conclusion remains the same; the single diffusion length techniques perform better than the differential diffusion methods. The difference in precision between the techniques has however decreased.

As suggested by reviewer 2, the power spectra of the ice core data are now presented in Appendix F in the manuscript. In order to shorten the paper, we want to ask the opinions of the reviewers and the editor, if all (or some of) the appendices should be supplementary material instead? We have also updated a reference in the test with the fractionation factor parameterizations. We now refer to the fully published study by Lamb et al (2017) instead of the 2015 version. The used parameterization is identical so it does not change the results and conclusion.

As suggested by reviewer 1, we have included a figure that shows the impact of high temperatures on diffusion length (Fig. 2.2 in manuscript). This figure has already been published in Gkinis et al (2014), a paper which we authored. As the figure provides a good illustration on how both the accumulation rate and temperature affect the diffusion length, we ask the editor if he wants us to include the figure in this manuscript or refer to it.

We hope that the revisions in the manuscript and our accompanying responses will be sufficient to make our manuscript suitable for publication in *Geochimica et Cosmochimica Acta.*

We look forward to hearing from you.

Yours sincerely,

*Christian Holme*

Address: Juliane Maries Vej 30, 2100 Copenhagen. Denmark

Ph: + 45 21 67 43 13

E-mail: christian.holme@nbi.ku.dk

**Reviewer #1:**

**This manuscript examines a variety of methods for inferring temperature from diffusion of**

**water stable isotopes. The authors use both synthetic and ice core data to assess the reconstructions of temperature. They conclude that despite some theoretical advantages of using differential diffusion lengths, the single diffusion length methods are more robust.**

**This is a strong paper and should be published after minor revisions. The science is sound and well described.**

**This is a technical study and the conclusions are a little underwhelming, but I suppose the authors don't want to negatively bias methods while the field is still emerging. However, I would have liked at least a paragraph of forward looking discussion about how their testing of different methods should impact future measurements of ice cores.**

**[note from GCA Executive Editor: As the primary aim of GCA is to present studies of fundamental significance and broad relevance for understanding geochemical systems, mechanisms, and processes, this is a reasonable request and should be addressed specifically in the revised manuscript.]**

We thank reviewer 1 and the editor for their suggestions. We have added a section called “outlook with respect to ice core measurements” in the Discussion. Here we address how measurements of the δ17O signal is expected to be a common output from analyzed ice cores. As we showed with synthetic data, such a signal can also be used to reconstruct temperatures. Especially the differential diffusion length of δ17O and δD showed higher precision than that of δ18O and δD. Such measurements however, require that laboratories around the world have access to well calibrated standards. Calibration protocols for δ17O have been suggested (Schoenemann et al, 2013) although there is still a lack of δ17O values for the International Atomic Energy Agency standards VSMOW and SLAP. We also suggest that when future CFA operations are planned, consistency especially with respect to the depth scale registration should be taken into account as a parameter equally if not more important than the maximum achieved resolution or measurement precision.

**While I understand choosing climate conditions applicable to Central Greenland and the East Antarctic Plateau because there is ice core data available, I was disappointed that the authors did not take advantage of being able to create synthetic data from warmer and wetter climates. Many ice cores are being drilled closer to the margins of the ice sheet and the work would have greater applicability if it focused on where ice coring was going rather on where it's been.**

We understand the remark of the reviewer and therefore we have added a contour plot in the manuscript that includes diffusion length calculations for a variety of conditions (Fig. 2.2 in manuscript). As the synthetic data approach is a very simplistic (though still very valuable) we choose to restrict our synthetic data calculations for conditions similar to those of our existing real ice core data sets so we can provide the reader with meaningful comparisons.

**The manuscript could also use a quick edit from a native English speaker as some of the idioms and transitions are slightly off (e.g. "wavelengths in the order of 50cm" should be "wavelengths on the order of 50cm" or the use of "nowadays"). I've noted some of these at the beginning, but there are many I didn't note.**

Thank you for your comments and corrections. We have worked more with the language.

**A few other general points:**

**1) There was little if any discussion of the initial power spectrum at deposition. This became particularly apparent in the removal of the annual cycle, which struck me as odd the one signal that you know the frequency of just gets thrown away. Is there no way to use the annual signal in a constructive way?**

We have added a comment about the initial power spectrum at deposition (Sec. 3.1 lines 174-177).

With respect to the annual signal we can comment the following. With the exception of the GRIP dataset, most isotopic time series did not result in an observable annual spectral signal (in this study 5 out of 13 time series had an annual peak). Diffusion appears to kill the annual signal early in the firn for most of the sections we study here (we refer the reviewer to the Appendix F in the manuscript where time series and their power spectral estimates are presented). An assessment of diffusion, based on the attenuation of the the annual signal would require that the magnitude of the isotopic annual cycle is known and constant through time. This is by far not guarantied for any of the sites we studied as changes in the hydrological cycle have likely been frequent through time.

As a matter of fact, especially for the lower accumulation sites as Dome C, Dome F and EDML we are not sure if it is even applicable to use the term annual cycle for the isotopic signal of the precipitation. At these sites the isotopic signal is more likely comprised of a few events per year some of which are also likely to be of an origin other than cyclonic (relocation of snow or clear sky precipitation). We see this not as a limitation but as an indication of how powerful the diffusion technique is due to its independence to such complications.

**2) The use of warm and humid to describe a site with an annual average temperature of 30C**

**and an annual average accumulation of 22cm per year is very confusing. Not even among ice core sites would this be considered warm and humid. Why not just use CG central Greenland and EAP East Antarctic Plateau since this is what you really mean. In addition, CD is not representative of much of Antarctica, just the East Antarctic Plateau.**

We agree that the terminology may seem confusing and we have changed the names CD to Case A and WH to Case B.

**3) The ability of the single diffusion length methods to reconstruct synthetic data to within 0.5C seems quite good. But as I think about it more, I wonder if this is really something to be excited about. Given that the amount of temperature change within the Holocene is about 0.5C, does this imply that water stable isotope cannot be used to infer anything about temperature in the Holocene? The synthetic tests don't really address the question of what amount of temperature change can be confidently identified. For instance, what percentage of the time can a synthetic data set created with a surface temperature of 31C be distinguished from a data set create with 30C?**

In the revised version of the manuscript we have performed a sensitivity study that incorporates uncertainties related to the firn diffusion, densification and ice flow models . The resulting precisions of the temperature estimations are found to be in the order of 1 C while the calculated accuracies are in the order of approximately 0.2 C

In our view it is not straightforward to talk about the validity of the method for temperature reconstructions during the Holocene epoch. Reviewer 1 suggests that the amount of temperature change is in order of 0.5 C. What does this actually mean? Is this a maximum peak to peak value of the temperature signal during the Holocene, a trend over the last 8-9,000 years and which place on the map does it refer to. We find the case of Holocene to be not a simple one, as ice core temperature reconstructions are to our view relatively poor for this climate epoch. Reconstructions based on other proxies for the case of Greenland for example, indicate sizable temperature differences during the Holocene. Either way, we consider an approx. 3 C cooling trend from 7-8,000 years BP until present to be a signal that several ice core studies seem to be agreeing to for Greenland (Dahl-Jensen et al, 1998 and Vinther et al, 2009).

**Specific comments:**

**Line 30 add "of the water isotope diffusion methods" so as not to imply that the water isotope diffusion methods are better than other methods, i.e. borehole temperature reconstructions**

**Line 36 change "invaluable". Maybe use unique or some other qualifier, but something with less subjectivity.**

**L43 delete "however"**

**L45 reword this sentence. Don't start with "However"; I'm not sure what "more physical principles" means, should it read "such as"?**

**L51 seminal seems strong for a paper read only by the very small community into water isotope diffusion**

**L57 " More specifically" does not seem like the correct transition**

**L59 " Nowadays" is a bit archaic and informal, no?**

We thanks the reviewer for the comments and all of it has been corrected.

**L62 A reference to the Univ. of Colorado system is noticeably absent here. Jones et al., 2017, JGR.**

A reference to Jones et al. (2017) AMT is now included. We prefer to use the reference from *Atmospheric Measurement Techniques* as it is the one that describes the CFA system utilized at the University of Colorado.

**L76 define what you mean by flavors. I think you can make this word choice work, but not without explaining how "flavor" applies to water isotope diffusion**

By flavors, we meant that each technique represents a different species of the diffusion-based temperature proxy. We have rephrased and restructured the sentence (Sec. 1, line 76)

**L83 Sentence beginning "For some cases" does not make sense to me. Also, the Holocene Climate Optimum is a bit nebulous, so you should define what you mean.**

We have defined it like Dahl-Jensen et al. (1997) as a warmer period spanning from 5-9 ka with temperatures up to 3C warmer than present day (Sec. 1, line 84).

**L85 don't call your own work "interesting". Let the reader decide that.**

**L 92-93 If you use "can" then "possibly" is redundant**

**L94 delete "that have dealt with the matter". This is implied**

**L97 I like this approach and feel you have struck a good balance with what information is included in the main, what's in the appendix, and what is referenced**

**L117 use different word than "sensor"**

The suggested revisions have been made.

**L137 I wonder why the authors only considered such cold sites. Given that much of the paper focuses on synthetic data, why not include coastal sites with warmer temperature and higher accumulation rates. This would make the work more broadly applicable, particularly since the vast majority of ice cores (not necessarily deep ice cores) are being drilled in locations warmer and wetter than either of the CG or EAP situations.**

As commented already we restricted our synthetic data tests to the kind of conditions that are representative of the real ice core data we had in our availability. We would like to point out that water isotopic data sets of high quality and resolution from below the firn-ice transition are not easily found even more if one is interested in dual measurements of δ18O and δD.

**Figure 2.1 I favor more detailed figure captions so that a reader can scan the manuscript and know what's going on in a figure without needing to find lots of details in the text. In this case, provide the accumulation rates and temperatures for WH and CD. I'd also like a sentence describing the main takeaway from the figure; something like: The increase in diffusion length with depth from vapor transport is partially offset by compaction of the firn which moves the ice closer together.**

**Figure 2.1 why is there a kink at ~20m depth. Does this have to due with the transition from Zone 1 to Zone 2 densification?**

Yes the kink around 20 m is due to the transition from zone 1 to zone 2 densification. We have now clarified that in the text (Sec. 2.1, line 126). We have also updated the caption of Fig. 2.1 and in general made more detailed captions. With respect to the extended captions we would like to point out that some additional text can be added though the manuscript is already quite lengthy as it is right now and we would like to avoid repetitions of the text in the captions. We hope this is understandable.

**Figure 2.1 make the legend lines bigger**

**L150 " borehole temperature profiles" is the more common description**

This has been clarified (Sec. 2.2, line 145).

**L150 This sentence confuses me. Do you just mean that the temperatures you use in your model are based on borehole temperature profiles? The sentence implies that you are obtaining the diffusivities from the high precision borehole temperature profiles. But really you just need a number to plug into the Ramseier equation.**

This is indeed what we are implying. We used the high precision borehole temperature profiles with depth when we solve the ice diffusion equation. The reason why we use the borehole temperature profile is that when approaching the deeper parts of the core the warmer ice temperatures enhance the effect of ice diffusion which then becomes an important and progressively dominating factor in the calculations. In order to perform the ice diffusion calculations for these deeper parts we certainly do not need a high precision (10-3 K) temperature measurement. Practically for the parts of the cores we are working with here a first order estimate of the ice temperature would be satisfactory as the ice is still very cold and solid ice diffusion extremely low. We have clarified that in Sec. 2.2.

**L166 " in" should be "on"**

**L178 this notation comes out of nowhere**

**L232 add comma after f\_lamda**

Thank you. This has been corrected.

**L232 is there really nothing useful to be gained from the annual signal? This raises the question of why the power spectral density is the way it is.**

See previous explanation.

**L270 is the ice diffusion the same for all isotopologues? What is the evidence for this?**

Yes. The self-diffusion coefficient has been found to be the same for monocrystals (Ramseier, 1967). The various parameterizations of diffusivity available in the literature, result in very small spread in the calculated solid ice diffusion lengths. In Gkinis et al. 2014 (supp. material) such a comparison is performed. Moreover, the ice diffusion is of negligible influence in this study.

**L284 delete the comma and change "can be" to "is"**

**L285 define dxs. This is the first usage**

This has been clarified.

**Reviewer #2:**

**Review of "Molecular diffusion of stable water isotopes in polar firn as a proxy for past**

**temperatures" by Holme et al. Holme et al. compare several methods of reconstructing paleotemperatures from water isotope diffusion lengths. Based on synthetic and real ice core data, they argue that single diffusion length estimates are more reliable than differential methods, both in terms of their precision and their accuracy. The comparison of these different methods is a good and timely contribution to the literature, and should be performed. However, due to the strong reliance on synthetic data, it is unclear to me whether these conclusions will actually hold for real world studies. I think additional work is needed to provide realistic estimates of the accuracy of the method when applied to real world problems.**

**I personally have not worked on the problem of water isotope diffusion, but I've been following the field out of interest. In principle the method is very original and powerful, but I've always had the feeling that several key uncertainties are being ignored, which means that the cited uncertainties are much too optimistic. The 0.5 degree accuracy claimed for the single profile method seems unrealistic to me. The method shows up to 34 degrees offset for modern sites (Fig. 5.2), which seems like a much more realistic and meaningful estimate of the accuracy to me.**

**My main concern is that there is likely a lot of missing physics in the forward diffusion model, which is not accounted for. This is particularly problematic when using synthetic data, as I will argue below.**

**Some missing physics includes:**

**1) The work assumes that the initial deposited isotope record has a more or less white noise spectrum. This assumption is made implicitly in the work of Holme et al., and never really addressed. This assumption is baked into Eq. 4.2. We know for certain that the annual cycle has a big imprint, which is briefly addressed (but not included in the synthetic data, as far as I can tell). Besides that, Greenland ice cores are strongly affected by the NAO, and the WAIS core by ENSO/PDO variability as well as the SAM. We know that all these modes of variability have considerable power in the sub annual bands. Also, snow is deposited in finite events (up to**

**several cm of precip each), which leads to considerable autocorrelation. How does this influence the accuracy of the methods?**

**There have been studies of isotopes in surface snow and vapor. Can the assumption of whiteness be proven or disproven?**

We would like to thank reviewer 2 for his insightful comments and suggestions on our work. As a general comment we would like to say that some of the concerns of reviewer 2, although very reasonable may have been caused by the absence of any visualised examples of isotopic data and/or fitted power spectral estimates. In the revised version of the manuscript we have included plots of both the isotopic time series as well as their adjacent power spectral estimates and the model fits. We believe that this information will give a better insight into our work and make it easier to understand why we believe some of the potential issues reviewer 2 is talking about are of lesser importance.

Indeed one of the assumptions of the model used here is that the distribution of the power spectral density of the isotopic signal at t = 0 is flat and resembled by white noise. This may in fact be slightly different than the term reviewer 2 is using ie *initial deposited isotope record*. The term *record* may be confused with a time series that spans several meters in the firn and this is not what we assume here.

We would kindly disagree with reviewer 2 that *we know* ***for certain*** *that the annual cycle has a big imprint.* Especially for the low accumulation Antarctic sites included in our study such a statement would most likely not be true. In both the isotopic time series as well as their adjacent power spectral estimates we see no indications of a spectral signature of the annual signal. It is beyond the scope of this work to examine why this is the case but a combination of the following could be true. i) the annual signal is heavily attenuated already in the first few meters due the extremely low accumulation, ii) post depositional effects as for example snow relocation result in a very low signal to noise ratio for these time series iii) there may not be a periodic annual signal in precipitation to start with due to the low number of precipitation events that comprise the total signal and the fact that part of the precipitated snow is not of cyclonic system origin. For the data sections of this study where the accumulation rate is higher (typical for Greenland sites and WAIS -D) only a few cases indicated some presence of a spectral signature of the annual cycle for which we were able to correct as described in the manuscript. It is important to note that the influence of the annual spectral peak is localised around the frequencies corresponding to the annual layer thickness and does not affect the overall shape of the spectrum. For the rest of the series the annual signal was completely diminished and therefore not apparent in the spectra. We believe that including figures of the power spectral densities of the ice core sections under consideration will clear out some of the doubts of the reviewers regarding the spectral estimates and their fits.

In our view the case of other climate modes such as NAO, ENSO/PDO and SAM is even more straightforward as no observable spectral features are observed in any of our case studies. This does not necessarily mean that there is no imprint in those bands to start with, but our analysis does not indicate this and these signals are either too weak to noticeably affect the fits of the assumed model (i.e. diffusion plus noise) or they cannot be resolved at all because their power lies lower than the measurement noise (we now address this in Sec. 5.2.1, line 457-462).

Any possible departure from the here assumed model due to strong spectral imprints of climatic signals would be manifested as an increased misfit between the model in use and the power spectral estimates. Here, we examine the effect of climatic signals on the diffusion estimates of the NGRIP δ18O core 1 (this test has not been included in the manuscript in order to not further lengthen the paper ). This was done by evaluating how well we fitted the power spectral densities of the data. If the performance changed with time, it could be a result of changing spectral signatures that affects the diffusion length. As we are concerned with multiannual variability, we only evaluated the goodness of fit (GoF) for frequencies lower than the annual signal (Fig. A.2 in Appendix A of the cover letter). The GoF with respect to age can be seen in Fig. A.1. Here it is evident that the goodness of fit is rather constant and does not change with depth/time. We can therefore not see the effect of multiannual variability in the power spectrum of δ18O. We would of course not be able to analyze the effect if it had a constant power over the entire ice core record. However, this does not seem likely.

With respect to the autocorrelation function we would actually expect quite the opposite. The lower the number of the discrete events that comprise the “annual signal” the faster the terms of the autocorrelation function would approach zero. To some extent we believe we see this effect when comparing the Antarctic (low accumulation and typically fewer precipitation events) with Greenlandic (higher accumulation, more precipitation events) sites. The former seem to resemble the diffusion plus white noise model in the best way while the latter seem to show a more red-noise behavior.

**2) There is considerable uncertainty in the densification model and tortuosity parameterization. There is a large literature from gas studies that looks into tortuosity (in models, measurements, and firn pumping experiments). These suggest that tortuosity is highly site specific, as well as scale dependent (for example, cm scale tortuosity is different from mscale).**

**This will skew the reconstructions (it's not clear in which direction…). This bias will not show up in synthetic data tests, because both the forward and inverse model assume the same firn properties.**

It is correct that several studies suggest that the tortuosity is site specific. The parameterization that we use in this study is also site specific through the densification model (surface density, accumulation rate and temperature). We acknowledge that the tortuosity parameterization is simplified, but that does not imply that it is not useful. A new study by Schaller et al (presented at EGU, 2017) shows that the trapping of bubbles in a single layer is solely determined by its total porosity and thereby independent of depth. Their results agree well with the Schwander et al (1988) model that we use. Even though there is an uncertainty in such parameterizations, it might be a fine approximation when used in integrated estimates. In the revised manuscript, we test the influence of such uncertainties by performing sensitivity tests where we vary some of the parameters associated with the densification model. This is done in tests with both synthetic data and ice core data.

**3) Does the seasonal temperature cycle matter? This drives both intense grain metamorphism (depth hoar) and influences vapor pressure and diffusivity (exponentially!). Density layering in general is probably important, because the diffusion length is not a single value, but presumably varies within each annual layer. This means that the final PSD is not Gaussian, but a superposition of Gaussians with different sigma values.**

The seasonal cycle has a small influence, and it was previously excluded in order to have a simple model. In the revised manuscript, we have included a seasonal temperature signal that propagates down through the firn (Appendix B in manuscript). This firn temperature profile is used in the diffusion model resulting in a difference in the diffusion length value of approx. 0.1 cm at maximum. This difference translates to a temperature difference of 0.5 C for the CD and WH scenario (now case A and B) respectively.

**4) Recently, it was shown that at WAIS Divide the diffusion length changes through time in a very unpredictable way (doi: 10.1002/2016JF003938). Also, field tests suggest large differences between modeled and observed diffusion (doi:10.5194/tc910892015). All these observations argue for missing physics. I do not pretend to know what such physics would be; I just think that these observations should caution us that the physics is not as simple as Johnsen et al. (2000) would have us believe.**

We are aware of the WAIS divide results published recently in Jones et al. Even though it is beyond the scope of this review exchange to discuss other studies in detail we would like to comment the following.

1. The WAIS divide study is based on data obtained with a CFA system for the span of the whole core. It is to our knowledge the first reconstruction of this kind (based solely on CFA data for the full span of the ice core). With this in mind, we would like to point out that such melting systems introduce additional mixing the magnitude of which is (a) not always easy to characterise and (b) can possibly be variable during the span of a measurement campaign which for the case of WAIS-D was about 3-4 years. A more detailed comment on isotopic analysis strategies and how they impact diffusion studies has been added to the manuscript under the subsection “Outlook with respect to ice core measurements”. Jones et al unfortunately have not presented a temperature curve based on their diffusion length reconstruction. Nevertheless we think that it may be preliminary to claim that the past temperature curve from WAIS divide should be expected to look like any other ice core based temperature curve from Antarctica. In that sense, we would wish to reject the term *unpredictable* for anything that comes out of WAIS. We need to keep in mind that isotope diffusion thermometry gives integrated firn column temperatures and that WAIS divide is a very dynamic site from the ice flow point of view with very likely strong elevation changes over the last 25,000 years. As a result the diffusive properties of every point in the WAIS divide core are likely to represent a firn column of a different spatial origin. This opens the possibility that a diffusion based temperature reconstruction could in principle look very different.
2. We would agree that in general field and laboratory experiments have often the potential to provide a better insight into physical processes due to the intended better control of some of the parameters of the experiment. However we feel the need to point out that in order to simulate processes that normally would take much longer time in nature, such experiments lead to shortcuts that sometimes can potentially have a negative impact in the quality of the experiment results, in fact adding more complications than simplifying the process. Having ourselves experience from working in a laboratory based study (van der Wel, 2011) similar to van der Wel et al. (2015) we know from first hand that generating artificial snow and building an artificial snow pack can be extremely complicated even in the controlled environment of a laboratory and thus discrepancies between the experiment and the natural process can arise. For the case of van der Wel et al. (2015) we would also like to point out that the isotopic spike used in the experiment was extremely enriched isotopically (>1000‰ in δD) in order to achieve diffusive rates fast enough to be able to measure within the time frames of a PhD project. Field measurements in porous media (see Hassanizadeh and Leijnse 1995) indicate that in cases of very high gradients there can be observed a departure from the traditional Fickian diffusion and the diffusive mass flux becomes non-linearly dependent with respect to the concentration gradients. The diffusivities measured in these cases are 50-70% lower than those expected assuming Fickian transport. Additionally one of the assumptions made for the derivation of the diffusivity parameterization in Johnsen et al 2000 is that vapor space and solid matrix are at isotopic equilibrium at all times. We would be very cautious assuming such a statement is true when isotopic gradients of this order are at play. Last but not least, if our approach and calculation were as much off as the study of van der Wel 2015 indicates, our temperature reconstructions would also be far off for all the ice core sites studied. On the contrary, the temperatures we estimate, are for all the ice core sites and for most of the diffusion techniques used in our study, fairly close to the modern temperature of each site and certainly very far from the 40-60% offset that van der Wel suggest in their work.

**The use of synthetic data is interesting in principle, because it allows testing in a controlled environment. But it can also be very misleading, because it assumes that the forward model (i.e., how the diffused isotope records are created) is exactly known which I think is a fallacy, see above. Perhaps I misunderstood, but it seems that in the synthetic tests the exact same firn properties (density, tortuosity) are used in the forward model, as in the "inverse" model (with the inverse model I mean the calculations used to go from sigma to temperature). I think this artificially improves the accuracy, and favors the single diffusion length methods (perhaps explaining why they look so good in the synthetic comparison). The differential diffusion methods should be less sensitive to these errors, because both isotopes diffuse through the same firn (tortuosity, density), canceling out some of these uncertainties. What I would advise the authors to do is to build some of this missing physics into their forward model that generates the synthetic data series. For example, they could use random number generators to control annual accumulation rates, temperatures, precip events, firn density layering and densification rates, tortuosity parameterizations, fractionation factors, etc. The isotope input spectrum should have an annual cycle, as well as a "redness" that is randomly selected. These random perturbations are not known to the part of the model that interprets the synthetic data. Then when the authors use the synthetic data to reconstruct T, there are unknown "biases" in their data, which is a much more realistic test that currently performed.**

Reviewer 2 is absolutely right that forward and “inverse” test pretty much do the same thing. We need to point out that we do not see the initial synthetic data tests as anything more than a first order basic test. One of the reasons we included it in the study was because our experience in fitting high resolution water isotope spectras has taught us that a large bias in the method can originate from bad data points in the time series (particularly outliers). Therefore we felt it was beneficial to test how well our fitting routines would work in the case of just white measurement noise of various magnitudes. The uncertainties we obtained from these simplified tests are absolutely the highest boundary of precision we can achieve only based on the limitations of our isotopic analysis equipment. We were satisfied to see that these uncertainties were very low but we agree that those numbers can be misleading as they completely disregard the level of understanding we have with respect to firn densification and ice flow uncertainties.

As a result, we have modified the test with synthetic data so uncertainties related to the densification process, diffusion and ice flow model are incorporated. We have also incorporated a seasonal signal in the diffusion model, but we still generate synthetic data based on an AR-1 data series without an annual peak. We don’t think forcing an annual peak in the time series will make the synthetic data a better representative of water stable isotope time series. In most cases, there is not an annual peak present in the power spectral densities of the ice core data (See Appendix F in the manuscript).

**The authors use an AR(1) process currently, but it is clear that the model does not introduce any meaningful level of redness (See Figure C.2). Where do the autocorrelation parameters come from?**

In the previous Fig. C.2, it might have looked like the AR-1 data didn’t include any meaningful level of redness. However, that was a result of the sampling process masking the redness. We have included a subfigure (Fig. D.2 in the manuscript) to show the initial power spectrum before sampling – here the redness of the signal is visible. Furthermore, the spectra in Appendix F shows that there is not necessarily any redness left after diffusion, post depositional noise and sampling (Dome C for instance).

The selected autocorrelation parameter of 0.3 is based on experience from power spectra of ice cores. Moreover, Gkinis et al. (2014) showed that the autocorrelation parameter did not significantly influence the diffusion length.

**My gut feeling is that the single diffusion methods are much more sensitive to such "unknown physics", because they will cancel out to some degree in the differential methods. How does the inverse model deal with accumulation rates? Are these assumed to be known exactly? Here some noise should be added too, since these are never known precisely. The same is true for the thinning function. There are different types of uncertainties: unknown firn physics, fractionation factors, accumulation rates, ice thinning, etc. In general, it would be interesting to know how sensitive these methods are for each of the sources of uncertainty. That could help decide which method is most suitable under which condition.**

While it is true that some of the densification parameters ‘cancel out’ when using the differential diffusion methods, the influence of the accumulation/thinning function is basically the same. One also needs to keep in mind that being a differential signal Δσ2 typically presents a lower signal to noise ratio and in general its estimation from ice core data is more tedious than that of the single diffusion length. With respect to the influence of the densification and ice thinning uncertainties we refer the reviewer to our previous answers related to the additional sensitivity tests we have included in the revised version of the manuscript.

**More detailed comments:**

**Line 44: assuming a linear relationship is of course not the same as using the spatial slope. The temporal slope is also a linear one.**

We have rephrased it to modern day linear relationship.

**Line 47: Schwander 1988 is not the correct reference here. Thermal fractionation wasn't discovered until a decade later.**

Thank you for noticing that. That reference has been deleted.

**Line 51: While Johnsen 2000 is undoubtedly a good paper, there were many others that made just as seminal contributions, like Whillans and Grootes 1985, and Johnsen 1977.**

We have updated the text with these references.

**Line 85: HCO is poorly defined, give the age!**

We have defined it and provided an age (see Sec. 1, line 84)

**Eq 2.1: Technically D is a function of depth also, particularly in the presence of layering. Also, this equation is an approximation, because delta is not as mass conserving quantity.**

D is also a function of depth but as the equation already is being evaluated within a considered layer, we think it’s implied when we express it as dependent on time. We have now stated that it is an approximation that assumes the concentration is similar to the water isotope ratio signal (e.g. δ18O).

**Equation 2.2: is this an exact solution, or an approximation?**

This is an exact solution.

**Line 117 and elsewhere: This is confusing. How can sigma be used to reconstruct both T and A? It has to be either one. You cannot reconstruct two independent parameters from a single number.**

We agree with reviewer 2 that inferring both temperature and accumulation from one value of diffusion length would constitute an underconstrained problem. This is more than clear. We would like however to point out that the main influence of the accumulation is via the linear impact of the ice flow thinning function. Assuming that a temperature profile is inferred based on diffusion length estimates this linear dependence will work as a “lever” for the whole profile thus giving some combined information on both the temperature and the ice flow thinning. This is not the same as saying that one diffusion length point can result in simultaneous estimates of both A and T but rather that the information extracted by a **full temperature profile estimation** combined with existing estimates of temperature based on other thermometry techniques (this could be nitrogen isotopes fractionation or borehole thermometry) and used as control or tuning points can provide some very valuable information for the ice flow thinning function. We have made a modification in the text so our claims on the matter are better explained (Sec. 3.1, line 214-222).

**Line 126-127. This is conceptually sloppy. Both drho/dt and D do not care about accumulation rate, of course. drho/dt cares about T and overburden pressure (not A), and D cares about temperature and firn connectivity.**

Our statement was based on the parameterizations that we use in the study. We have changed it in Sec. 2.1, line 119-120 so the description is conceptually correct.

**Line 128: again, I don't think you could do both T and A, unless you have some other source of info.**

See previous comment.

**Line 129: what are the typical measurement precisions for these various isotopes?**

They are 0.07, 0.50 and 0.05 permil for respectively δ18O, δD and δ17O. We define that in Appendix D, line 705 that describes the generation of the synthetic data.

**Line 139: above the closeoff depth ice diffusion also happens of course, within the grains.**

We agree with the reviewer that ice diffusion happens above the close-off depth. We just wrote that below the close-off depth, diffusion occurs in solid ice driven by the isotopic gradients within the lattice of the ice crystals. As the ice diffusivity is 4 orders of magnitude smaller in the firn, it is negligible.

**Figure 2.1: Why Fig 2.1? please change to Fig. 1**

The figure number matches the section number. From the Geochimica et Cosmochimica Acta *author information pack*, it says that we should refer to equations, tables and figures in such a way for appendices. We therefore assumed that it should be same for the rest of the manuscript. It can be changed if required.

**Eq. 4.2: Shouldn't P\_0 be a function of (k) also? It seems to be a constant here.**

Yes, but we assume that P\_0 is constant for each section due to strong depositional noise. We have clarified this in Sec. 3.1 line 174-177.

**Line 171: I am confused by the use of AR(1) processes. In Eq. 4.2, the noise term represents the instrumental noise floor, correct? Shouldn't this be white noise? That's what's assumed elsewhere on Line 561. Also, this AR(1) is distinct from the AR(1) used in appendix C to estimate the input time series, correct?**

Ideally, the noise in Eq. 4.2 (now Eq. 3.2) represents the white measurement noise. But real ice core measurements sometimes have a more red noise behavior (Appendix F in manuscript). That is why we fit the entire noise of the data. We have clarified that in Sec. 3.1 line 176.

We generate time series using an AR-1 model with predefined variance and autoregressive parameter. After diffusion and sampling we add white measurement noise.

**Line 174/Fig 4.2: Could you show us some more real data PSDs, instead of synthetic ones? Ideally at least one PSD per core site (can be an appendix). Of course the fit to the synthetic data is good, because the synthetic data has the assumption of being Gaussian baked into it. I'm curious to see how well your method fits the actual databased spectrum at various sites. Also, do the spectra at Greenland and Antarctic sites look very different?**

We strongly agree with this suggestion. We believe that many of the questions and/or doubts of the reviewers with respect to the goodness of the fits, the presence and/or influence of the annual signal or other multiannual signals can be better addressed by also providing an insight into the isotopic data and the estimated power spectral densities. We have included all the fitted spectra used in this study in Appendix F in the manuscript. Here you can see that the spectral-based techniques perform well on real ice core data.

The spectra from Greenland and Antarctica do not look very different. The only visible difference is that the noise level looks almost completely white for the Antarctic ice cores.

**Line 179: This is a little unclear. What is the difference between regular P and fancy P? Is regular P the model, and fancy P the observations?**

Regular P was defined as the model and fancy P as the observations. We have clarified that further in order to avoid misunderstandings (Sec. 3.1 line 184).

**Line 179: It appears that the background noise level is used as a fitting parameter. Isn't the noise level known from the instrument used?. Do the fitted noise levels agree with the known analytical precision of the instrument?**

The background noise is used as a fitting parameter because some ice core data have a more red noise behavior (Appendix F in manuscript) than that of white measurement noise. If we fit a white noise line at the highest frequencies it agrees well with the analytical precision of the instrument. This has also been shown by Gkinis (2011).

**Line 196: I assume integrating Eq. 2.7 includes all the parameterizations used for tortuosity, density, thinning, etc? Please be more clear about this.**

That is correct and we defined that earlier in Sec. 2, line 131-138 (now Sec. 2.1, line 115-123).

**Line 205-213: is the single diffusion method more sensitive to this than the differential one?**

No. The same correction is used in both cases. The diffusion length ratio is the only technique independent of densification processes and ice flow.

**Line 218: the annual peak is clearest when there is little diffusion, this is obvious. But the enhanced power in the annual band (in the input spectrum) should be preserved under all circumstances, correct? It may just be smeared out at high diffusion sites, but the power is there. It seems to me that at all sites the annual band should be disregarded from the spectral fitting. Please confirm how this is done. Also, is the annual band avoided in all analyses done here?**

It is not quite clear what reviewer 2 means with the term *input spectrum*, however we would partly agree that no matter how strong the diffusive process is, the remaining power of the annual band should be there stored in the ice. This of course would be only true if there is an enhanced signal in the interval of frequencies corresponding to an annual signal. Even though we cannot exclude this possibility completely we have some serious doubts that any significantannual signal exists for the Antarctic sites (see our previous comments on the frequency of precipitation events and post depositional noise). Even if we assume that such an annual signal exists, a diffusion length estimate for the site of Dome C for present conditions is approx. 6 cm ice eq. (very similar for Dome F). The diffusion transfer function with this diffusion length would attenuate wavelengths of 3 cm (roughly the accumulation at Dome C and Dome F) by 10 orders of magnitude. It is technically true that this very low signal is still there in the ice but neither our measurement nor the spectral estimation techniques in use can get down to this sensitivities.

The case of Greenlandic sites with higher accumulation is of course different and as we have shown both because an isotopic signal is much more likely to exist in firn (that is pre diffusion) and because the higher accumulation rates result in a lower attenuation of the annual signal. A Gaussian transfer function with a diffusion length equal to 8 cm typical for present day conditions at sites like NorthGRIP and GRIP would leave 5 % of the power of the signal.

We have only disregarded the annual band from the spectral fitting if it had a spectral signature. The annual band was removed in 5 out of 13 data sets. This is clarified in Sec. 4.2, line 360-361 and the power spectra can be seen in Appendix F in the manuscript.

In the remaining cases, we don’t exclude the annual bands as inspection of the spectra shows no evidence or indication of a signal close to the annual layer thickness. Fitting the diffusion plus noise signal results in a good fit of the estimated power spectral densities.

**Why was there no annual peak added to the synthetic series?**

The purpose of the synthetic data test was to investigate the theoretical spectral model described in Eq. 4.2 in the manuscript (now Eq. 3.2). Furthermore, as we have explained in our answer above, the majority of the real ice core data sets we have worked with shows no sign of an annual signal in the spectral- (and actually neither in the depth-) domain. Lastly, addition of an annual signal requires assumptions about the initial magnitude of these cycles, how this magnitude changes with time and what its seasonality behavior looks like. We feel that one of the main advantages of the technique is that it should in principle not need any knowledge on the initial pre-diffusion isotopic signal of the precipitation and in combination with the picture we get from real ice core data we decided that addition of an annual signal was unnecessary. We have included a seasonal temperature variation in the firn diffusion model to account for the effect of the saturation vapor pressure non-linear dependence on pressure

**Equation 4.7: What value of df\_lamba do you use?**

We chose a weight function of w(f\_lambda – 0.5 < f < f\_lambda +3 ) = 0 in all cases. We decided to filter out more of the higher frequencies. We have clarified that in the manuscript (Sec. 4.2, line 360-361).

**Figure 4.2: please use real data here, synthetic ones are not as interesting.**

We agree with reviewer 2 that presenting the time series and power spectra of the real ice core data sections is important and gives a honest view on how well the spectral model performs. We have provided real ice core data in Appendix F. We prefer using the synthetic data as it facilitates a better comparison between the different techniques and power spectra and it also offers a clear indication of how an idealized scenario of a diffusion dominated power spectral density looks like. In that way readers can also judge for themselves on the quality of the reconstructions and spectral fits on the real ice core data spectras

**Figure 4.5: are these real or synthetic data? I guess you know my preference by now :)**

Fig. 4.5 (now Fig. 3.5) presents synthetic data.

**Section 4.3: since this method performs so poorly, perhaps it can be left out to shorten the MS?**

We would prefer not to remove this technique from the study, as this is the first time the method is utilized on ice core data.

**Eq. 4.12: Define alpha**

This has been defined now.

**Line 321: Here I was confused. Is the noise white or AR(1), as claimed earlier?**

The time series is generated from an AR-1 model with predefined variance and autoregressive parameter. The added noise is white as that represents the measurement uncertainty. When we fit the noise in a power spectrum, we use the power spectral density of an AR-1 noise model. So technically we are using the same tool (an AR(1) process) but strictly the AR(1) used to generate synthetic data series is not the same as the AR(1) we use to fit the noise signal of the power spectral density.

**Line 324: As mentioned above, I think each time you generate synthetic data you should also perturb the firn physics, because this is an unknown in any real world problem.**

Done – see previous comments.

**Figure 5.1: Do you have any idea where the clear 17Delta^2 bias comes from?**

The bias is not evident in the new simulations. The bias in the previous manuscript was a result of a small offset in the estimate of the sigma\_17 value. It propagated into the differential diffusion technique. That is the reason the bias was not present when using the linear fit technique.

**Lines 368-371: I couldn't agree more! This makes all the methods look too good, and as I argued above, particularly the single diffusion case.**

**Table 5.2: The selection of depth ranges seems somewhat arbitrary. Why were these depths chosen?**

High resolution dual water isotopic data of high quality are not that easy to find in fact. For some sites (Dome C, Dome F) the section presented is basically all that exists at this high resolution. We focused on sections of 10-20 m, all from the Holocene and preferably from late Holocene in order for the boundary conditions of our reconstructions (temperature, accumulation, ice flow and firn characteristics) to be relatively close to present conditions

**Line 389: 0.5deg C seems too good to be true. Many uncertainties were not really considered, so this number is not all that meaningful. The way of things is that this number will be cited by others in future work without all the caveats needed to understand that this value holds only under very narrow conditions of synthetic data in a perfect world where all physics is understood. I think this should be revised upward by including all the relevant uncertainties, in order for this number to be a meaningful reflection on the accuracy of the method in real world applications.**

See previous comments.

**Table 5.1 It's not precisely clear how you go from sigma to T. What assumptions were made regarding the accumulation rates, for example? Were they precisely known?**

See previous comments.

**Line 417: Why not just compare the reconstructed temperature with the modern day temperature? I understand there can be differences, but comparing to present day is the only meaningful way to test the accuracy of the method. The synthetic data are artificially unbiased because you assume you understand all the physics. Comparing the reconstructed temp from real isotope data to observed temps is the only meaningful metric to my mind. The abstract claims that the single isotope methods are the most accurate, but this doesn't always seem true when I look at Fig. 5.2, 5.3 and 5.4.**

We have rephrased the section of the manuscript that refers to the method being unbiased and we mostly talk about accuracy and precision (Sec. 6.2). As we have explained earlier in our answers the sole purpose of the synthetic data tests is to provide a first order check that our spectral estimate and temperature reconstruction techniques operate in the simplest case possible. We still believe this is a test of some value as in our view there can easily be pitfalls in the spectral estimation, spectral fitting and temperature calculation. It also serves the purpose of conceptuallyintroducing the reader to the method with a more hands-on example. We have indeed extended the tests to include a fair amount of uncertainty in some key parameters of the Herron and Langway model.

We agree with the comment on the modern day temperature comparison. In fact we do perform such a comparison in the manuscript even though we are not very explicit in the text about it due to mainly two reasons. Strictly speaking, a true comparison between modern day temperature and a diffusion based reconstruction is not possible. This is because diffusion records an integrated firn temperature from surface to close-off and not an instantaneous temperature at the surface. Moreover a spectral estimate on the firn section of the core while diffusion is still under way cannot give a representative value for the diffusion length of a data section. Second, we simply did not have access to high res, high quality data just below the firn-ice transition. Only for some of the cores was it possible to obtain late Holocene sections. Nevertheless for the case of Dome C and Dome F it would be a rather reasonable statement to say that the estimated early Holocene temperature should not differ much from the present or late Holocene conditions.

**Line 485: I don't believe you have proved that the method is unbiased, because in the synthetic world you know all the physics and parameterizations, which is not true in the real world.**

We have changed the wording. See previous comment/answer.

**Line 502: Again, I don't think accuracy has been proven given that you didn't compare the temperatures reconstructed from the real data to real site temperatures.**

See previous comments.

**Eq. A7: Note that there have been multiple different parameterizations. This is a very old study. A more recent result to use would be doi:10.5194/tc83192014, or doi:10.3189/172756402781816582.**

We have acknowledged the two other studies at Eq. A7: doi:10.5194/tc-8-319-2014 and doi:10.3189/172756402781816582

**Line 556: what are these parameters for the AR(1) input based on? Actual data? The input spectrum of the synthetic data is very white, and does not have an annual cycle.**

The input parameters have been chosen empirically by examining power spectra similar to the warm and cold conditions. The input spectrum is red but it looks white due to the sampling process masking the redness. We have included an extra figure showing the raw power spectrum before sampling (Fig. D.2).

**Reviewer #3:**

**This manuscript is well written and convincing for the material it chooses to cover.**

**My main reaction, however, is that it leaves the problem only half analyzed and this leads to the final conclusion being a logical fallacy.**

**The Abstract and the Conclusion sections finish (lines 2930, 501502) by claiming "this study has demonstrated that methods based on the single isotope diffusion length result in the most accurate and precise estimates of past temperatures." But the authors have chosen not to analyze the factors that can make the single isotope methods problematic, in particular uncertainties in (1) firn densification and density-depth profiles as a function of climate, and (2) uncertainties in how vapor diffusivity varies as a function of density. These must be addressed!**

**The potential value of the multiple isotope analyses is to find methods independent of these uncertainties. The present analysis does a good job of revealing the inherent weakness of the multiple isotope methods, in turn.**

**This is an original and useful contribution. But ignoring the other side of the problem means that no conclusion can be reached about the relative or absolute merits of the different methods.**

We thanks reviewer 3 for the insightful comments. In the new manuscript, uncertainties related to the firn diffusion model, densification and ice flow are considered and included. The uncertainties are included in order to facilitate a better comparison between the single and differential techniques (in case the single diffusion methods are more sensitive to such uncertainties). We have also included a seasonal temperature signal that propagates through the firn column. The seasonal temperature variation affects the firn diffusion length nonlinearly due to the saturation vapor pressure exponential dependence on temperature. The firn temperature profile is obtained by numerical solution of the heat equation.

**Another disappointment of the submitted paper is that the authors do not take advantage**

**of the ice core data to quantify the accuracy of the thermometry methods. The reason,**

**apparently, is that temperature varies (with depth and/or time) in the firn column, whereas**

**Eq. 2.7 is for a steady state (although this is never stated explicitly. I might be misinterpreting this.) Why don't the authors use Eq. 2.5 itself and time dependent calculations of temperature and density to overcome this limitation??**

Even though we have implemented a seasonal cycle in the revised manuscript, we still don’t have a temperature forcing history in the firn diffusion model. In the presented analysis, we estimate a diffusion length value for a given section, and that diffusion length estimate represents the final diffusion length at pore close-off. We can therefore not use Eq. 2.5 to make time-dependent calculations as we only have the final value for a single ice core section (and no temperature forcing history). As a result, we are only able to reconstruct an average temperature for the past firn column (this has been further specified in Sec. 5.2.2, line 493-500).

Since we don’t know the exact temperature at a given depth/time, we can’t quantify the accuracy using ice core data. We are only able to evaluate the deviation from the present day annual mean temperature.

**The terrain covered by this analysis is already a narrow one for GCA. The authors can**

**probably have this manuscript accepted almost "as is" at a good journal focusing on**

**ice core or glacial studies. It would be much more useful to the readership, and to the**

**authors' scientific reputations, if GCA returns this manuscript to them with instructions to**

**expand the analysis to enable a useful and cogent conclusion. Then it will make a fine**

**GCA paper.**

Thank you. We have made a great effort in expanding the analysis.

**A few minor points:**

**1. the authors neglect to cite the Whillans and Grootes JGR 1985 analysis, which is**

**the source of most of our understanding of isotopic diffusivity**

We now cite the paper in the revised manuscript (Sec. 1, line 49).

**2. looking at Figure 5.2, is it just a coincidence that the Delta sigma^2 I method**

**(the green circles) does a good job of matching the modern temperatures? A better**

**job than any of the other methods?**

The Delta sigma^2 I method reconstructs matching temperatures in six out of nine cases. It is not clear if it is a coincidence or correct. The method does not always result in temperatures similar to the Delta sigma^2 II technique which had a slightly better performance with the synthetic data. We therefore advise to utilize both techniques in future ice core research. We have added this to the discussion (Sec. 5.2.2, lines 528-537). Please be reminded that the ice core sections we use come from depths that are not necessarily representative of the present site temperature (even though we can safely assume that it should not be much different than that). As a result comparing the deviations between diffusion based temperatures and present day site temperatures for the number of sites we consider is technically not a proper comparison.

**3. there is a recent paper by T.R. Jones et al. in JGR Earth Surface (2017) that looks at isotope diffusion in the WAIS Divide core. It is a case in which the relationship between diffusion length and climate does not seem to work as expected.**

The WAIS divide study is based on data obtained with a CFA system for the span of the whole core. It is to our knowledge the first reconstruction of this kind (based solely on CFA data for the full span of the ice core). With this in mind, we would like to point out that such melting systems introduce additional mixing the magnitude of which is (a) not always easy to characterise and (b) can possibly be variable during the span of a measurement campaign which for the case of WAIS-D was about 3-4 years. A more detailed comment on isotopic analysis strategies and how they impact diffusion studies has been added to the manuscript under the subsection “Outlook with respect to ice core measurements”. Jones et al unfortunately have not presented a temperature curve based on their diffusion length reconstruction. Nevertheless, we think that it may be preliminary to expect that the past temperature curve from WAIS divide should look like any other ice core based temperature curve from Antarctica. We need to keep in mind that isotope diffusion thermometry gives integrated firn column temperatures and that WAIS divide is a very dynamic site from the ice flow point of view with very likely strong elevation changes over the last 25,000 years. As a result the diffusive properties of every point in the WAIS divide core are likely to represent a firn column of a different spatial origin. This opens the possibility that a diffusion based temperature reconstruction could in principle look very different.

**Appendix A**

****

****

**References**

Dahl-Jensen, D., Mosegaard, K., Gundestrup, N., Clow, G. D., Johnsen, S. J., Hansen, A. W., Balling, N. (1998). Past temperatures directly from the Greenland ice sheet. *Science*. Vol: 282, pp: 268-271.

Gkinis, V. (2011). High resolution water isotope data from ice cores.PhD thesis, University of Copenhagen.

Gkinis, V., Simonsen, S. B., Buchardt, S. L., White, J. W. C., Vinther, B. M. (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.* Vol: 405, pp: 132-141.

Jones, T. R.; White, J. W. C.; Steig, E. J.; Vaughn, B. H.; Morris, V.; Gkinis, V.; Markle, B. R. & Schoenemann, S. W. (2017), 'Improved methodologies for continuous-flow analysis of stable water isotopes in ice cores', *Atmos. Meas. Tech.* **10**(2), 617--632.

Majid Hassanizadeh, S. & Leijnse, A. (1995), 'A non-linear theory of high-concentration-gradient dispersion in porous media', *Advances in Water Resources* **18**(4), 203--215.

Lamb, K. D., Clouser, B. W., Bolot, M., Sarkozy, L., Ebert, V., Saathoff, H., Möhler, O., Moyer, E. J. (2017). Laboratory measurements of HDO/H2O isotopic fractionation during ice deposition in simulated cirrus clouds. *PNAS*, Vol: 113, no: 22, pp: 5612-5617.

Schaller, C. F., Freitag, J., Sowers, T., Vinther, B., Weinhart, A., Eisen, O. (2017, unpublished). In-situ observation of bubble trapping in polar firn. G*eophysical Research Abstracts Vol. 19, EGU2017-14721-1, 2017 EGU General Assembly 2017*

Schwander, J., Stauffer, B., Sigg, A. (1988). Air mixing in firn and the age of the air at pore close-off. *Annals of Glaciology*. Vol: 10, pp. 141-145.

Van der Wel, L. G.; Gkinis, V.; Pohjola, V. A. & Meijer, H. A. J. (2011), 'Snow isotope diffusion rates measured in a laboratory experiment', *Journal of Glaciology* **57**(201), 30--38.

van der Wel, L. G., Been, H. A., van de Wal, R. S. W., Smeets, C. J. P. P., Meijer, H. A. J. (2015). Constraints of the δ2H diffusion rate in firn from field measurements at Summit, Greenland. *The Cryosphere*. Vol: 9, pp: 1089-1103.

Vinther, B. M. , Buchardt, S. L. , Clausen, H. B., Dahl-Jensen, D. , Johnsen, S. J. , Fisher, D. A. , Koerner, R. M., Raynaud, D., Lipenkov, V. , Andersen, K. K. , Blunier, T., Rasmussen, S. O., Steffensen, J. P., Svensson, A. M. (2009). Holocene thinning of the Greenland ice sheet. *Nature*. Vol 461, pp: 385-388.