JOURNAL OF GEOPHYSICAL RESEARCH, VOL. ???, XXXX, DOI:10.1029/,

Estimating diffusion lengths of water isotopes from ice core data measured by continuous flow analysis

Emma Kahle,[1] Christian Holme,[2], and everyone else

Corresponding author: Emma Kahle, Earth and Space Sciences Department, University of Washington, Seattle, WA, USA. (eckahle@uw.edu)

[1]Department of Earth and Space Sciences,

University of Washington, Seattle,

Washington, USA.

[2]CIC

Abstract.

We examine high-resolution water isotope data sets from continous flow analysis (CFA) of the West Antarctic Ice Sheet (WAIS) Divide (WDC) and South Pole (SPC) ice cores. Spectral analysis of water isotope data reveals damping of high-frequency variations associated with diffusive smoothing of the isotopic profile in the firn layer of an ice sheet. This diffusion of water isotope ratios in ice cores can provide information about past firn conditions. The data spectra of windowed sections of CFA data from WDC and SPC show different characteristics as compared with similar discretely-sampled data sets, affecting the estimation of high-frequency damping due to the firn diffusion process. The cause of this spectral difference is not known, but we suggest possible eplanations in the CFA measurement system and in natural firn proceses. Without knowing the origin of this spectral difference, we can still parametrize a method to estimate the extent of diffusion in the spectra. In this study we use a modified method to estimate extent of diffusion in order to efficiently and accurately produce diffusion estimates for full ice core CFA data sets.

1. Introduction

To understand past and future climate change, water isotope data from ice cores have long been used as climate proxies, based on the temperature-dependent distillation of water isotope ratios (e.g. $\delta^{18}O$) in the atmosphere (cite papers). This indirect method for obtaining temperature records through past glacial and interglacial cycles relies on empirical correlations that only approximate the physical processes involved. A more direct approach uses the signal of isotope diffusion preserved in the ice, overcoming the many issues related to conventional isotope ratios [*Johnsen*, 2000].

1.1. Stable Water Isotope Diffusion

Water isotope diffusion occurs primarily in the firn layer, snowfall in the upper tens of meters of an ice sheet that has yet to be fully compressed into ice. As firn is permeable, water can diffuse through in the vapor phase, damping the seasonal variations and high-frequency noise of the original isotope signal. As the diffusion process depends on temperature, a temperature record can be obtained by measuring the extent of diffusion that has occurred. This method is independent of conventional assumptions about isotope fractionation before deposition, and thus improves on conventional ice core temperature methods. Beyond information about past temperature, diffusion estimates also provide constraints on past firn conditions and total ice thinning. With these motivations, estimations of diffusion have been made for ice cores in both Greenland and Antarctica [*Johnsen*, 2000; *Simonsen et al.*, 2011; *Gkinis et al.*, 2014; *Jones*, 2016].

1.2. Continuous Flow Analysis

Most ice core water isotope data have been measured discretely by melting vertical sections of the core. The isotope ratio of each discrete sample is measured on mass spec or laser spec. Recently, ice core water isotope measurements are often measured on continuous flow analysis (CFA) systems. These systems continously feed the melting core directly into a laser spectrometer, yielding much higher resolution data than discrete measurements. Depending on the amount of effort committed, discrete analyses can produce results ranging from one-meter for entire ice cores to 2.5-cm resolution for smaller sections of ice cores. Meanwhile, CFA analyses can easily produce results at half-cm resolution for an entire ice core record. This higher resolution greatly improves the ability to analyze diffusion lengths, which depends on information in the high frequencies of the data spectrum.

A number of studies have estimated diffusion lengths from discretely sampled water isotopes ([*Johnsen*, 2000; *Simonsen et al.*, 2011; *Gkinis et al.*, 2014; *Van der Wel et al.*, 2015]). While spectra from discrete data match well spectra theoretically derived synthetic data [*Holme et al.*, 2017], spectra from CFA data have an additional characteristic in the mid-frequency range (5^-1 to 10^-1 cm^-1). Figure 5 demonstrates this difference with many discrete spectra and many CFA spectra plotted on top of one another (Christian's figures comparing spectra). The CFA data spectrum has a transition slope from the damped frequencies into the system noise. The highest resolution discretely sampled data sets do not show this transition in the spectrum, suggesting it is produced within the CFA system. Mixing and memory effects are known to take place within the CFA system and can explain this shape in the spectrum.

Since estimating diffusion length depends on the shape of the spectrum, the presence of this transition can affect diffusion results. In this paper we describe how to estimate diffusion lengths on CFA data without being affected by this transition. We start with the approaches developed in *Johnsen* [2000] and *Gkinis et al.* [2014] and propose an alteration to accommodate CFA water isotope data.

1.3. Water Isotope Data

We use data from the WAIS Divide and South Pole ice cores, both measured at the Institute of Arctic and Alpine Research at the University of Colorado. INSTAAR uses a CFA system to measure the water isotope ratios of $\delta^{18}O$, $\delta^{17}O$, and δD at a resolution of half-cm. For complete details on the INSTAAR CFA set up, see *Jones CFA* [2016].

2. Diffusion Theory

From the fundamental physics of advection and diffusion, the effects on the profile of water isotopes in the ice core can be described mathematically. The majority of water isotope diffusion occurs in the firn layer where interconnected air pathways allow water vapor to diffuse vertically through the firn column. After firn densification has sealed off the bubbles in the ice, vapor diffusion ceases and solid ice diffusion takes over, with a diffusivity orders of magnitude smaller than that of vapor diffusion.

The isotopic profile changes through the firn layer of an ice sheet due to the affects of diffusion across isotopic gradients and densification of the firn. As stated by *Johnsen* [1977] and subsequently used in several diffusion studies ([*Johnsen*, 2000; *Simonsen et al.*, 2011; *Gkinis et al.*, 2014; *Van der Wel et al.*, 2015; *Jones*, 2016; *Holme et al.*, 2017]), these changes to the isotopic profile can be described by Fick's second law, the basic

advection-diffusion equation:

$$\frac{\partial \delta}{\partial t} = D\frac{\partial^2 \delta}{\partial z^2} - \dot{\epsilon} z \frac{\partial \delta}{\partial z} \tag{1}$$

where δ is the isotope ratio, D is the diffusivity coefficient, z is the vertical coordinate assuming an origin fixed on a sinking layer of firn, and $\dot{\epsilon}$ is the vertical strain rate. The coordinate system is fixed to a sinking layer of firn, so the term $\dot{\epsilon} z$ can be thought of as the velocity in this advection-diffusion framework. Considering a layer of firn with $z = 0$ at the vertical midpoint, $\dot{\epsilon} z$ is the speed at which a point distance z from $z = 0$ approaches the origin. A solution for the isotopic profile at depth z in the firn column and at time t is given by:

$$\delta(z,t) = S(t)\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \delta(z,0) \exp\left(\frac{-(z-u)^2}{2\sigma^2}\right) du \tag{2}$$

where $S(t)$ is the total thinning the layer has experienced due to ice flow from $t = 0$ to $t = t'$:

$$S(t') = exp\left(\int_0^{t'} \dot{\epsilon}(t) dt\right) \tag{3}$$

For analytical and statistical derivations of the solution in Equation 2 see Appendices A and B.

The amount of smoothing applied to the signal in Equation 2 can be quantified as a diffusion length, or the average vertical distance traveled by a water molecule before it reaches the close-off depth. See Appendix C for a statistical derivation of diffusion length. *Johnsen* [1977] gives the solution to diffusion length, σ, as a differential equation:

$$\frac{d\sigma^2}{dt} - 2\dot{\epsilon}(t)\sigma^2 = 2D(t) \tag{4}$$

Following *Gkinis et al.* [2014], Equation 4 can be solved and written as a function of density as:

$$\sigma^2(\rho) = \frac{1}{\rho^2} \int_{\rho_0}^{\rho} 2\rho^2 \left(\frac{d\rho}{dt}\right)^{-1} D(\rho) d\rho \tag{5}$$

3. Estimating Diffusion Length from Data (methods)

With a mathematical understanding of water isotope diffusion, we can match diffusion lengths from ice core data to modeled diffusion lengths to learn about past firn conditions, including temperature, densification, and ice thinning. To make this comparison, we require an estimate of diffusion length from the ice core data. Equation 2 demonstrates that the ice core data, $\delta(z,t)$, is the convolution of the initial isotope signal $\delta(z,0)$ with a Gaussian filter of standard deviation σ, the diffusion length:

$$\mathcal{G} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-z^2}{2\sigma^2}\right) \tag{6}$$

Rather than solve this convolution to find the diffusion length σ in the Gaussian filter $\mathcal{G}$, we take the Fourier transform to simplify to multiplication:

$$\delta(z,t) = \delta(z,0) * \mathcal{G} \qquad \Rightarrow \qquad \hat{\delta}(z,t) = \hat{\delta}(z,0)\hat{\mathcal{G}} \tag{7}$$

where $*$ represents the convolution and $\hat{}$ represents the Fourier transform. Furthermore, the Fourier transform of a Gaussian remains a Gaussian:

$$\mathfrak{F}(\mathcal{G}) = \hat{\mathcal{G}} = \exp\left(\frac{-k^2\sigma^2}{2}\right) \tag{8}$$

With the assumption that the initial isotope signal $\delta(z,0)$ can be approximated as white noise [*Gkinis et al.*, 2014], we can simply fit a Gaussian curve in frequency space to the spectrum of some window of ice core data and solve for its standard deviation to find the diffusion length σ for that window of data. The standard deviation σ_{std} of the Gaussian

in frequency space is related to the diffusion length σ by:

$$\sigma = \frac{1}{2\pi\sqrt{2}\sigma_{std}} \tag{9}$$

Appendix D derives this conversion factor.

Fitting the data spectrum is complicated by the water isotope measurement process. Beyond ice thinning and natural diffusion in the firn and solid ice, the spectra of ice core data are further affected by the measurement process (refer back to Figure 5). *Jones* [2016] avoids the effect of this measurement noise by identifying the frequency at which measurement noises begins to affect the spectra and only including lower frequencies in the spectrum fit. Figure **??** shows examples of this fitting technique for data from the WAIS Divide ice core [*Jones*, 2016] and from the South Pole ice core. The Gaussian fit ignores the data spectrum beyond a chosen cut-off frequency where measurement noise is determined to dominate the signal. This technique relies on the fact that measurement noise dominates only at the highest frequencies and does not wipe out the damping signal left by natural diffusion.

In this study we follow the fitting approach of *Gkinis et al.* [2014], which uses a least-squares technique to fit the summation of multiple functions to the spectra of ice core data. The first function is a Gaussian, representing the damping of firn diffusion, and the second function is an autoregressive noise signal of order 1 (AR-1). Adding together these two functions avoids having to choose a cut-off frequency for the fit. Since it does not require any subjective choices, this multi-function technique can be fully automated and produce results in much less time and with much less effort than the cut-off technique.

The multi-function technique has been used effectively to estimate diffusion lengths on many discretely sampled data sets [*Gkinis et al.*, 2014; *Holme et al.*, 2017]. It has been used

on short sections of CFA data sets, but in some cases required extra boundary conditions in the parameterizations. However, when applying this technique to newer CFA data sets with a higher signal to noise ratio, the transition zone in the CFA spectra affects the ability to effectively fit a Gaussian to the data. Figure 6 illustrates this issue with examples of CFA spectra from WAIS Divide and from South Pole. The presence of the transition zone force both the Gaussian and noise functions to attempt to accommodate that part of the spectrum. The Gaussian function is pulled to a greater standard deviation, corresponding to a smaller diffusion length, and the noise level is raised slightly at lower frequencies.

4. Improving the fit

In order to best fit the spectra of CFA data, we have examined two different approaches. The first approach removes the extra transition zone by masking it with white noise. The second approach parameterizes the total PSD by including an extra function in the fitting procedure. Both techniques are introduced and tested on ice core data from WDC and SPC.

4.1. Adding white noise

The main difference between the WDC/SPC CFA data and the discrete data is the resolution and precision. This difference migrates into the PSD by letting the CFA data having a lower white noise baseline for the high frequencies compared to that of the discrete data (refer to a figure). Previous studies of water isotope diffusion have not had the same precision and resolution as the measurements of WDC and SPC. Despite of lower precision, diffusion-based temperature reconstructions have still proved possible [*Gkinis et al.*, 2014; *Holme et al.*, 2017]. Hence, whether the transition zone is some climate or

measurement induced process, one solution is to mask the transition zone in the CFA data. This is achieved by adding Gaussian distributed noise in the time domain which increases the white noise level in the frequency domain. This masks the red noise transition and the power spectrum is now similar to that of older CFA measurements. The effect of this method can be seen in Fig. 1.

We have constructed a sensitivity test in order to quantify the most optimal choice of noise level. The test consist of a 500 year moving window. For each window section, a total of 100 diffusion lengths are estimated by adding white noise with an increasing baseline to the data. As a result, a diffusion length estimate is available for each noise level as a function of age/depth. By adding too much noise to the data we risk masking the climate signal. Hence, the optimal noise level is reached when the diffusion length estimates stop changing with increasing noise - the gradient approaches 0. We have plotted the gradient in diffusion length with respect to added noise in Fig 2. For WDC, adding a mean noise level of 0.4 ‰ for δ^{18}O and of 3.0 ‰ for δD seem to be the most optimal solution.

4.2. Parameterizing a Multi-Function Fit or Parametrerizing the red noise transition

The second approach concerns a parameterization of the total spectrum. In order to best fit the spectra of CFA data, we try three different parameterizations of a multi-function fit of all frequencies. In addition to the Gaussian that represents firn diffusion and the autoregressive noise function, we try adding a third function to the total fit to take up any slack in the fit and allow the main Gaussian and the noise level to relax into better fits for the low and high frequencies, respectively. First, we try adding a second autoregressive noise function, second, we try adding a second Gaussian curve, and third

we try adding the PSD of a folded normal distribution (FND):

$$P_{FND} = P_{0_{FND}} \cdot e^{-(k\cdot\sigma_{FND})^2} \cdot |\,[1 - \Phi(-ik\sigma_{FND})]\,|^2, \tag{10}$$

where $\Phi(\sigma_{FND}) = 1/2 \cdot \mathrm{erfc}(-\sigma_{FND}/2)$. Here P_0 and σ_{FND} are the two parameters that are being estimated from the fitting. The justification of fitting a folded normal distribution is that it assimilates memory induced by the CFA system. Our goal here is not to explain physically what this part of the spectrum represents, but simply to find the optimal parameterization to fit the data spectrum. The optimal parameterization will allow the firn Gaussian to better fit the lower-frequency part of the spectrum, and thus to get a better estimate of diffusion length compared to Figure 6.

Figure 7 shows the parameterization that sums a Gaussian curve and two autoregressive noise functions to create a total fit for the data spectrum. With the inclusion of this second Gaussian in both WDC and SPC, the total fit is visually improved, as compared to the single-Gaussian fit in Figure 6. Figure 8 shows the parameterization that sums two Gaussian curves and one autoregressive noise function. Similarly, there is a visual improvement in the total fit as compared to the single-Gaussian fit. Finally, Figure **??** shows the parameterization that sums one Gaussian, one autoregressive noise function, and one cumulative Gaussian function.

5. Results - Goodness of fits

The different fitting procedures are evaluated by calculating the adjusted goodness of fit ($\bar{R}^2$) between the spectra and parameterizations. The $\bar{R}^2$ is a goodness of fit estimation (R^2) that takes into account the number of fitted parameters (p):

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1}, \tag{11}$$

where n is the sample size. The $\bar{R}^2$ values enables a comparison between the parameterizations that use six fitting parameters with the parameterization that uses five parameters (the addition of white noise). The $\bar{R}^2$ values are plotted as a function of age in Figure 10. It is evident that all the parameterizations provide good fits to the data, but the best fits are obtained with the extra Gaussian or the FND. For simplicity and computational efficiency, the Gaussian curve should be preferred over the FND.

6. Validation of parameterization

The first way of examining the significance of the results is by comparing the diffusion length estimates from the noise adding method with the estimates from utilizing the parameterization method. In Figure 1, the diffusion length estimates are plotted with respect to age. It is evident that both methods reconstruct similar diffusion lengths, but the noise adding method seems to be more stable in the glacial period. The higher stability is attributed to the fact that the measured water isotope signal seems to contain several noisy sections of half a meter or so, which will affect the spectra. By adding white noise to the data, the noisy sections are masked, making the fit more stable.

Another way of validating the results is by comparing with the diffusion lengths estimated with the cut-off technique presented in *Jones* [2016]. While the cut-off technique requires the choice of cut-off frequency and is less efficient, it avoids the effect of the transition zone because that part of the spectrum is not included in the fit. Since these results are not affected by the transition zone in the spectrum, we can use them as a benchmark with which to validate the effectiveness of the double-Gaussian fit. This comparison is made in Figures 3 and 4. It is evident that the estimated diffusion lengths from this paper are very similar to the presented results in *Jones* [2016]. There are a few peaks that stand

out such as an age of around 12,000 years. These peaks were not found in *Jones* [2016] likely due to their lower resolution (not a moving window).

7. Discussion

Figure 9 and Table **??** demonstrate that for the CFA data sets for WAIS Divide and South Pole, the double-Gaussian parameterization of the multi-function fitting technique works better than the single-Gaussian formulation. The cut-off fitting technique provides validation of the fits, but is overall less efficient than the multi-function technique and requires a partially subjective choice of cut-off value.

While including this second Gaussian in the multi-function technique improves fit, it remains to be explained how to physically interpret the transition region of the spectra. Possible origins of this transition include noise generated in the CFA measurement system, noise generated naturally in the firn, or some combination of both. In this section we explore each of these possibilities and how to gain insight into the origin of this noise.

7.1. Possible Noise Origin: CFA Measurement System

There are many possible sources of noise throughout the CFA measurement system that could contribute to the data spectrum bewteen periods of 5 to 10 cm. Mixing and memory effects are known to occur throughout the system as sample water travels to the instrument through tubing and different reservoirs. *Jones CFA* [2016] ran standards of ice through the CFA system and reported system diffusion lengths of 0.7 cm and 0.8 cm for $\delta^{18}O$ and δD, respectively. However, system diffusion is unlikely to be the cause of the transition noise because, mathematically, it is expected to increase the diffusion in the main Gaussian, rather than add this red noise in the 5 to 10 cm range.

A second source of this red noise in the CFA system is the Picarro instrument. (Cite Gkinis thesis) showed that the Picarro measurement can be affected by variations in water concentration in the instrument cavity. The INSTAAR CFA system has been carefully calibrated and set up to ensure water concentrations remain at a level that does not affect the measurement. However, it is possible that small fluctuations in water concentration could still add noise even if they do not change the isotope value significantly. This explanation is a possibility – they've tested for it, but it could be more complex.

Errors in depth registration within the CFA system could produce noise in this frequency range. We tested this possibility by adding random noise to a depth series and inspecting the resulting frequency spectra. This test showed that this source of noise does/does not remain a possibility.

Figure 12 demonstrates the variability through the core of the diffusion length associated with the second Gaussian. The CFA system remains constant through the analysis of the entire core, and thus any effects solely due to the system would be expected to remain constant as well. This reasoning suggests that, according to Figure 12, there is a significant contribution from climate variation to the transition noise. However, inspection of individual spectra shows that there is an implicit dependence of the diffusion length of the second Gaussian on that of the main Gaussian. Figure 13 shows that with a greater natural diffusion length, the transition zone is shifted toward lower frequencies and thus the diffusion length of the second Gaussian is greater as well. Conversely, with a smaller natural diffusion length, the transition zone is shifted toward higher frequenices and the second Gaussian diffusion length is also smaller. Therefore the transition noise may be affected directly by climate influence on the lower frequencies in the spectrum.

To better isolate CFA system effects, a good approach is to make double measurements on ice from SPC. This could include measuring the same meters of ice discretely at half-cm resolution as well as continuously on the CFA at the saem resolution. This would give us two data sets of the same resolution to compare the effects of the CFA directly to discretely sampled data.

7.2. Possible Noise Origin: Natural Firn Diffusion Processes

Outline Christo's grain lifetime idea and show that it could theoretically explain this part of the spectrum.

Another possibility is that the red noise transition originates naturally in the firn rather than in the CFA system during the measurement. This effect could be explained by considering a more complex model of water isotope diffusion. In the *Johnsen* [2000] model, all water molecules are treated identically, experiencing the same amount of time in the vapor phase relative to that spent in the solid ice phase. In reality, different molecules experience some range of time in the vapor phase relative to the solid phase. Some molecules may be trapped inside ice grains throughout their advection down the firn column, while others may remain in the vapor phase and on the surface of grains. Previous estimates (Johnsen and Whillans and Grootes?) claim that the lifetime of grains is short enough that no grains exist long enough to trap molecules and significantly affect the bulk diffusion length. However, if the lifetime of grains is longer than their original estimates, this process could prevent all molecules from spending the same amount of relative time in the vapor phase, and thus could affect the diffusion length.

The *Johnsen* [2000] model has been shown to work well in capturing a bulk diffusion length by fitting a Gaussian curve to the data spectrum, but the added complexity of

allowing a range of individual diffusion lengths could contribute to the red noise transition. Mathematically, the *Johnsen* [2000] model matches a diffusion length and corresponding Gaussian curve to the bulk water isotopes in the firn in a given window of data. A model that allows individual molecules to experience a range of relative time spent in the vapor phase corresponds to a collection of Gaussian curves. Each Gaussian curve is weighted by the number of molecules that experienced that amount of diffusion and the collection of weighted Gaussian curves sum to a non-Gaussian filter. In theory, this collection is infinite as it represents an infinite range of possible amounts of time spent in the vapor phase. In practice, this inifite collection can be approximated by a finite collection. As shown above in the multi-function parameterization, perhaps as few as two Gaussian curves could sufficiently represent the infinite summation.

8. Application Example: South Pole Glacial-Interglacial Temperature

The data set from the South Pole ice core provides new information about climate in central Antarctica. From previous ice cores in the climatically distinct East Antarctica and West Antarctica, we have temperature estimates based on conventional assumptions about Rayleigh distillation. However, model estimates do not match with temperature estimates. The South Pole spatially bridges the gap between those climatically distinct regions and diffusion-based temperature estimates will provide new insight into this model-data discrepancy.

We use the method described above to calculate a first estimate of the glacial-interglacial temperature change at the South Pole. A complete water isotope data set is not yet available from the South Pole ice core, but we can use two 20-meter windows of data, from the Holocene and Late Glacial respectively, to make a diffusion-based estimate of

the temperature change at this site. Using the adding-noise method, we calculate a change of about XX degrees, and using the multi-function parameterization method, we calculate a change of about 8 degrees C. These methods agree well with one another, and provide the first insight the glacial-interglacial change at the South Pole. With the complete water isotope data set in the future, we will be able to carefully calculate a temperature history, rather than only a temperature difference, to learn more about the climate in central Antarctica and how well it does or does not agree with model simulations.

9. Conclusions

In this study we examined the diffusion of water isotope data from the WAIS Divide and South Pole ice cores measured on the CFA system at INSTAAR. We observed that spectra from these CFA data, unlike spectra from comparable discetely sampled data, have a unique transition zone in the mid-frequencies. We found that the most effective way to estimate diffusion lengths on these CFA data sets was with a double-Gaussian parameterization of the multi-function fitting technique. This method is efficient in terms of both time and effort required, and also effectively fits the data spectrum with the transition zone.

The origin of the noise in the transition zone is unknown. It could be caused by effects in the CFA measurement system, by effects in natural diffusion in the firn, or by some combination of both. Further work is required to isolate these issues.

Appendix A: Analytical Derivation of Diffused Isotope Profile

Here is a complete derivation of the desired solution that I found in the slides of an MIT course "Mathematical Methods for Engineers" from 2006, which

I found at http://ocw.mit.edu/courses/mathematics/18-086-mathematical-methods-for-engineers-ii-spring-2006/readings/am54.pdf. I've written up the steps as outlined by the MIT course, but added in many more detailed steps that are omitted on the course slides.

This derivation starts with the heat equation, with diffusivity D, and therefore is still not quite the solution I seek, which would also include an advection term. Equation (47) starts off with the basic heat equation.

$$\frac{\partial u}{\partial t} = D\frac{\partial^2 u}{\partial x^2} \tag{A1}$$

We can solve this PDE using the common technique of separation of variables. We assume that the solution $u(x,t)$ can be written as the product of two functions, one that depends only on t and one that depends only on x:

$$u(x,t) = G(t)E(x) \tag{A2}$$

We can take derivatives of $u(x,t)$ this solution and rewrite equation (47) as

$$G'E = GE'' \tag{A3}$$

$$\frac{G'(t)}{G(t)} = \frac{E''(x)}{E(x)} \tag{A4}$$

Since these ratios are equal, they must be equal to some constant, and we can find a family of solutions that will satisfy equation (50):

$$E(x) = De^{\imath kx} \quad E''(x) = -Dk^2 e^{\imath kx} \tag{A5}$$

$$G(t) = e^{-k^2 t} \quad G'(t) = -k^2 e^{-k^2 t}$$

check solutions:

$$\frac{G'(t)}{G(t)} = -k^2$$

$$\frac{E''(x)}{E(x)} = -k^2$$

Now we can write our solution as

$$u(x,t) = De^{\imath kx}e^{-k^2t} \tag{A6}$$

Now account for all solutions from all linear combinations by integrating over all k:

$$u(x,t) = \frac{1}{2\pi D}\int_{-\infty}^{\infty}\hat{u}_0(k)e^{\imath kx}e^{-k^2t}\mathrm{d}k \tag{A7}$$

Here $\hat{u}_0(k)$ is included to satisfy the initial conditions $u(x,0)$ at $t = 0$.

check:

$$u(x,0) = \frac{1}{2\pi D}\int_{-\infty}^{\infty}\hat{u}_0(k)e^{\imath kx}\mathrm{d}k \tag{A8}$$

Using the definition of the inverse Fourier transform:

$$x(t) = \frac{1}{2\pi D}\int_{-\infty}^{\infty}\hat{x}(\omega)e^{\imath\omega t}\mathrm{d}\omega \tag{A9}$$

Comparing the above, we find the right hand side is exactly equal to our arbitrary initial condition:

$$u(x,0) = u(x,0) \tag{A10}$$

So we have successfully derived the general solution as shown in equation (53).

To derive the fundamental solution for a specified initial condition, let us first consider an initial condition that is a delta function, or a point source: $u(x,0) = \delta(x)$ at $t = 0$. This is a nice initial condition because its Fourier transform is $\hat{u}_0(k) = 1$.

Plugging this initial condition into equation (53):

$$u(x,t) = \frac{1}{2\pi D}\int_{-\infty}^{\infty}e^{\imath kx}e^{-k^2t}\mathrm{d}k \tag{A11}$$

Next take the partial derivative with respect to x of both sides:

$$\frac{\partial u}{\partial x} = \frac{1}{2\pi D}\int_{-\infty}^{\infty} \imath k e^{\imath kx} e^{-k^2 t} \mathrm{d}k \tag{A12}$$

Now rearrange to separate for integration by parts:

$$\frac{\partial u}{\partial x} = \frac{1}{2\pi D}\int_{-\infty}^{\infty} \left(\imath e^{\imath kx}\right)\left(k e^{-k^2 t}\right) \mathrm{d}k \tag{A13}$$

where the first grouping will be u and the second grouping will be dv. To compute the integration by parts we have:

$$\begin{aligned} u &= \imath e^{\imath kx} & v &= -\frac{1}{2t} e^{-k^2 t} \\ \tfrac{\mathrm{d}u}{\mathrm{d}k} &= -x e^{\imath kx} & \frac{\mathrm{d}v}{\mathrm{d}k} &= k e^{-k^2 t} \end{aligned}$$

From the equation for integration by parts,

$$\int_{-\infty}^{\infty} u\mathrm{d}v = (uv)_{-\infty}^{\infty} - \int_{-\infty}^{\infty} v\mathrm{d}u \tag{A14}$$

we get:

$$\frac{\partial u}{\partial x} = \frac{1}{2\pi D}\left[\left(-\imath e^{\imath kx}\frac{1}{2t}e^{-k^2}\right)_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{1}{2t} e^{-k^2 t} x e^{\imath kx} \mathrm{d}k\right] \tag{A15}$$

The term on the left inside the square brackets goes to zero when evaluated from $-\infty$ to ∞, and we are left with:

$$\frac{\partial u}{\partial x} = -\frac{1}{4\pi t D}\int_{-\infty}^{\infty} e^{-k^2 t} x e^{\imath kx} \mathrm{d}k \tag{A16}$$

Now compare equation (58) to the expression for $u(x,t)$ in equation (54) to see the simple result

$$\frac{\partial u}{\partial x} = -\frac{xu}{2t} \tag{A17}$$

This is now a linear ODE that is solved by:

$$u = Ce^{-x^2/4Dt} \tag{A18}$$

We can solve for c based on conservation of the initial condition through time:

$$\begin{aligned}\int_{-\infty}^{\infty} u(x,t)\mathrm{d}x &= \int_{-\infty}^{\infty} u(x,0)\mathrm{d}x \\ &= \int_{-\infty}^{\infty} \delta(x)\mathrm{d}x \\ &= 1\end{aligned}$$

Apply this conservation criterion to solve for c:

$$\int_{-\infty}^{\infty} ce^{-x^2/4tD}\mathrm{d}x = 1 \tag{A19}$$

$$c = \frac{1}{\int_{-\infty}^{\infty} e^{-x^2/4tD}\mathrm{d}x} \tag{A20}$$

This is the integral of a Gaussian, which is known:

$$\int_{-\infty}^{\infty} e^{-ax^2}\mathrm{d}x = \sqrt{\frac{\pi}{a}} \tag{A21}$$

Here $a = \frac{1}{4tD}$, and thus we can solve for c as

$$c = \frac{1}{\sqrt{4tD\pi}} \tag{A22}$$

So now our solution to equation (60) can be written

$$u(x,t) = \frac{1}{\sqrt{4\pi tD}} e^{-x^2/4tD} \tag{A23}$$

This is the fundamental solution from a single point source (recall that our initial condition used in this solution was a single delta function).

Now let us consider the possibility that our initial condition delta function is instead located a different point $x = s$, such that

$$u(x,0) = \delta(x-s) \quad \text{at} \quad t = 0 \tag{A24}$$

Then the argument of the exponential in our solution shifts by s, i.e.

$$e^{-x^2/4tD} \quad \text{becomes} \quad e^{-(x-s)^2/4tD} \tag{A25}$$

Because of linearity, any initial condition $u(x,0)$ can be written as the combination of point sources:

$$u(x,0) = \int_{all\ S} \delta(x-s)u(s,0)\mathrm{d}s \tag{A26}$$

The solution for an initial condition extending over all points in x can be written as an integral of the responses to $\delta(x-s)$:

$$u(x,t) = \frac{1}{\sqrt{4\pi tD}} \int_{-\infty}^{\infty} u(s,0)e^{-(x-s)^2/4tD}\mathrm{d}s \tag{A27}$$

And finally we have the solution we have been looking for. The one exception is that this solution does not include the $S(t)$ term that accounts for the thinning of the ice because this solution was derived from the pure diffusion equation, without accounting for any advection.

Appendix B: Statistical Derivation of Diffused Isotope Profile

Lasaga derives the same solution through a statistical framework, starting with the idea of a discrete random walk. A particle starts at $z = 0$ and at each time step can move a distance L either to the right with probability p or to the left with probability $q = 1 - p$. We would like to know what the probability is that after N steps the particle is at some position $z = mL$, where $-N \leq m \leq N$.

Let us define the number of steps the particle takes to the right, n_R, and the number of steps the particle takes to the left, n_L. With these definitions we can write:

$$N = n_R + n_L \qquad\qquad m = n_R - n_L \tag{B1}$$

and thus:

$$n_R = \frac{1}{2}(N+m) \qquad\qquad n_L = \frac{1}{2}(N-m) \tag{B2}$$

We can write the probability, $P_N(m)$ of arriving at position $z = mL$ after N steps as the product of the probability of taking a particular sequence of steps to that position, $p^{n_R}q^{n_L}$, times the number of different sequences of steps that will end at that position (because multiple sequences of steps will end at the same z position, i.e. RRRL, RRLR, RLRR, and LRRR all end at $m = 2$):

$$P_N(m) = p^{n_R}q^{n_L}\frac{N!}{n_R!n_L!} \tag{B3}$$

Plugging in equations (5) and (6),

$$P_N(m) = \frac{N!}{[\frac{1}{2}(N+m)]![\frac{1}{2}(N-m)]!}p^{n_R}q^{n_L} \tag{B4}$$

We can now make a couple of assumptions to simplify equation (8). First, we can assume that $p = q = \frac{1}{2}$, that the particle is equally likely to step to the left or right. Second, we can assume that the particle has taken many steps ($N \gg 1$) and that the number of steps taken is much greater than the distance from the starting point ($N \gg m$). With these assumptions, we can use Stirling's Approximation:

$$\ln n! = n\ln n - n + \frac{1}{2}\ln n + \frac{1}{2}\ln 2\pi \tag{B5}$$

Rewrite equation (8) with this approximation and set $p = q = \frac{1}{2}$ to get:

$$\begin{aligned}\ln P_N(m) \;=\; & N\ln N - \frac{1}{2}(N+m)\ln[\frac{1}{2}(N+m)] \\ & -\frac{1}{2}(N-m)\ln[\frac{1}{2}(N-m)] + \frac{1}{2}\ln N - \frac{1}{2}\ln[\frac{1}{2}(N+m)] \\ & -\frac{1}{2}\ln[\frac{1}{2}(N-m)] - \frac{1}{2}\ln(2\pi) + N\ln\frac{1}{2}\end{aligned} \tag{B6}$$

Simplify with the following:

$$\ln(N+m) \;=\; \ln\left[N\left(1+\frac{m}{N}\right)\right] \tag{B7}$$

$$= \ln N + \ln\left(1 + \frac{m}{N}\right)$$

For small x,

$$\ln(1+x) \sim x - \frac{1}{2}x^2 \tag{B8}$$

and thus,

$$\ln\left(1 + \frac{m}{N}\right) = \frac{m}{N} - \frac{1}{2}\frac{m^2}{N^2} \tag{B9}$$

So, equation (11) becomes:

$$\ln(N+m) = \ln N + \frac{m}{N} - \frac{1}{2}\frac{m^2}{N^2} \tag{B10}$$

and, similarly,

$$\ln(N-m) = \ln N - \frac{m}{N} - \frac{1}{2}\frac{m^2}{N^2} \tag{B11}$$

Plugging equations (13) and (14) into equation (10) and simplifying, we get:

$$\ln P_N(m) = -\frac{1}{2}\ln N + \frac{1}{2}\ln 2 - \frac{1}{2}\ln \pi - \frac{m^2}{2N} + \frac{m^2}{2N^2} \tag{B12}$$

Neglecting the term of order $\frac{1}{N^2}$ and combining the remaining terms leaves us with:

$$P_N(m) = \left(\frac{1}{\pi N}\right)^{1/2} e^{-m^2/2N} \tag{B13}$$

Equation (16) shows the Gaussian curve expected from diffusion from a point source, based only on the assumptions of a random walk and of many time steps.

This equation describes only the discrete probability of the particle taking steps of size L. If we want an expression for the continuous probability $W(z,t)$ that the particle will be at any point z at time t, we must generalize m to z and N to t.

Define the relations:

$$z = mL \qquad N = vt \tag{B14}$$

where v is the frequency of steps.

Because the particle can only reach an even numbered position at an even time step or an odd numbered position at an odd time step (i.e. N and m are both even or both odd), $P_N(m)$ is the discrete probability of the particle being anywhere between $z = mL$ and $z = (m+2)L$. Thus the discrete probability can be related to the continuous probability as:

$$P_N(m) = W(z,t)2L \tag{B15}$$

Using this relation and plugging in known expressions:

$$\begin{aligned} W(z,t) &= \frac{P_N(m)}{2L} \\ &= \frac{1}{2L}\left(\frac{2}{\pi N}\right)^{1/2} e^{-(m^2/2N)} \\ &= \frac{1}{2L}\left(\frac{2}{\pi vt}\right)^{1/2} e^{-[(m^2/L^2)/2vt]} \\ &= \frac{1}{(2\pi L^2 vt)^{1/2}} e^{-(z^2/2vL^2t)} \end{aligned} \tag{B16}$$

And with the definition $D \equiv \frac{1}{2}vL^2$,

$$W(z,t) = \frac{1}{2\sqrt{\pi Dt}} e^{-(z^2/4Dt)} \tag{B17}$$

We now have an expression for the continuous probability of particles diffusing to a particular point at a particular time that originate from a point source. However, we would like to know what this solution will look like for a continuous source with spatial extent. Because the diffusion equation is linear, it obeys the law of superposition, and

we can generalize our point source solution by treating a continuous source as the sum of individual source slabs of arbitrarily small width dz.

If our point of interest is located at z' and one of the slabs making up the continuous source is located at z, we can write the contribution of this single slab to the concentration of particles at point z' as

$$c(z',t) = \frac{c_0}{2\sqrt{\pi D t}} \exp\left(-\frac{(z'-z)^2}{4Dt}\right) \mathrm{d}z \tag{B18}$$

where c_0 is the initial concentration at point z.

We can sum the contributions from all the slabs making up the continuous source by integrating the above equation over the z values of the entire source. For a source of infinite extent we get:

$$c(z',t) = \int_{-\infty}^{\infty} \frac{c_0}{2\sqrt{\pi D t}} \exp\left(-\frac{(z'-z)^2}{4Dt}\right) \mathrm{d}z \tag{B19}$$

Defining the diffusion length $\sigma \equiv \sqrt{2Dt}$, pulling the constants outside the integral, and replacing c with δ, z' with z, and z with u, we see that we have arrived at the solution presented by Johnsen in equation (2):

$$\delta(z,t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \delta(z,0) \exp\left(\frac{-(z-u)^2}{2\sigma^2}\right) \mathrm{d}u \tag{B20}$$

The only difference between the above and equation (2) is the inclusion of the thinning function $S(t)$, which is included as a correction for the ice thinning, which has not yet been taken into account by this statistical derivation. Otherwise, following Lasaga's statistical approach, we have now derived the same diffusion solution as presented in Johnsen.

Appendix C: Statistical Derivation of Diffusion Length

The diffusion length, σ, which is the standard deviation of the Gaussian convolved with the original signal, can also be derived in a statistical manner. Lasaga starts by introducing

the probability $W(Z,\tau)$ that an atom at position z at time t will be at position $z+Z$ at time $t+\tau$ through the effect of diffusion. If the concentration of profile at time t is known everywhere, then the concentration after diffusion acting over τ amount of time can be written as:

$$c(z,t+\tau) = \sum_{all\ \ Z} c(z-Z,t)W(Z,\tau) \tag{C1}$$

This equation accounts for particles arriving at location z from all other locations $z-Z$. Expanding the concentration terms as a Taylor series,

$$c(z,t+\tau) = c(z,t) + \tau\frac{\partial c}{\partial t} + \dots \tag{C2}$$

and

$$c(z-Z,t) = c(z,t) - Z\frac{\partial c}{\partial z} + \frac{Z^2}{2}\frac{\partial^2 c}{\partial z^2} + \dots \tag{C3}$$

Plugging these two concentration expressions into equation (68),

$$\begin{aligned} c(z,t) + \tau\frac{\partial c}{\partial t} + \dots = c(z,t)\sum_{all\ \ Z} W(Z,\tau) - \frac{\partial c}{\partial z}\sum_{all\ \ Z} ZW(Z,\tau) \\ + \frac{1}{2}\frac{\partial^2 c}{\partial z^2}\sum_{all\ \ Z} Z^2 W(Z,\tau) + \dots \end{aligned} \tag{C4}$$

The particles must exist somewhere in space at time τ, so by the definition of probability:

$$\sum_{all\ \ Z} W(Z,\tau) = 1 \tag{C5}$$

We can also use the definition of averages to write

$$\sum_{all\ \ Z} ZW(Z,\tau) = \langle Z \rangle \tag{C6}$$

and

$$\sum_{all\ \ Z} Z^2 W(Z,\tau) = \langle Z^2 \rangle \tag{C7}$$

Here $\langle \quad \rangle$ represents a statistical average. We can now write the basic diffusion equation (1) as:

$$\frac{\partial c}{\partial t} = \frac{\langle Z^2 \rangle}{2\tau} \frac{\partial^2 c}{\partial z^2} - \frac{\langle Z \rangle}{\tau} \frac{\partial c}{\partial z} \tag{C8}$$

Comparing equations (1) and (73), we can write new statistical expressions for the diffusion coefficient and the velocity:

$$v = \frac{\langle Z \rangle}{\tau} \tag{C9}$$

$$D = \frac{\langle Z^2 \rangle}{2\tau} \tag{C10}$$

Equation (75) can be rearranged as

$$\langle Z^2 \rangle^{\frac{1}{2}} = \sqrt{2D\tau} \tag{C11}$$

$$= \sigma$$

This equation is equivalent to the diffusion length σ, which is used to describe the "extent" of diffusion. Equation (76) gives some insight into the statistical meaning of the diffusion length as the root mean square of the vertical position.

We can write equation (76) in its integrated form as:

$$\sigma^2 = \int_0^t 2D(\tau) \mathrm{d}\tau \tag{C12}$$

Taking the derivative $\frac{d}{dt}$ of both sides:

$$\frac{d\sigma^2}{dt} = 2D(t) \tag{C13}$$

And we can now compare to equation (6) from Gkinis et al 2014, based directly off of Johnsen 1977:

$$\frac{d\sigma^2}{dt} - 2\dot{\epsilon}\sigma^2 = 2D(t) \tag{C14}$$

Comparing equations (78) and (79), we have derived the same expression for sigma with the advection term neglected.

Appendix D: Derivation of Sigma Conversion Factor

This section derives the conversion factor between the standard deviation of the Gaussian in frequency space and the sigma diffusion length in the depth domain.

We start with the Gaussian (Green's Function) that describes diffusion:

$$G = \sqrt{\frac{1}{2\pi\sigma^2}} \exp\left(\frac{-x^2}{2\sigma^2}\right) \tag{D1}$$

This Gaussian is convolved with the original isotope signal in depth/time space to yield the diffused signal. Alternatively, this Gaussian can be Fourier transformed into frequency space to easily multiply the fft of the signal to, equivalently, yield the diffused signal.

The definition of transforming a Gaussian to frequency space is simple:

$$g(x) = \sqrt{\frac{\pi}{a}} \exp\left(\frac{-\pi^2 x^2}{a}\right) \Rightarrow \hat{g}(f) = \exp\left(-af^2\right) \tag{D2}$$

From Eq (1) above, we can solve for a by

$$\sqrt{\frac{1}{2\pi\sigma^2}} = \sqrt{\frac{\pi}{a}} \qquad\qquad a = 2\pi^2\sigma^2 \tag{D3}$$

This expression for a yields the Gaussian we began with in (1). Now we can find the Fourier transform of (1) using the definition in (2).

$$\hat{G}(f) = \exp\left(-2\pi^2\sigma^2 f^2\right) \tag{D4}$$

We can use this to find, for some frequency, the amplitude A_σ of the diffused signal from the amplitude A_0 of the original signal as follows:

$$A_\sigma = A_0 \exp\left(-2\pi^2\sigma^2 f^2\right) \tag{D5}$$

We now have $\hat{G}(f)$ but we want $\hat{G}(k)$, a Gaussian in terms of k, where k is the wavenumber and $k = 2\pi f$, or $f = \frac{k}{2\pi}$. Plugging in this definition of k, we get:

$$\hat{G}(k) = \exp\left(-2\pi^2\sigma^2\frac{k^2}{2^2\pi^2}\right) \qquad \hat{G}(k) = \exp\left(\frac{-k^2\sigma^2}{2}\right) \tag{D6}$$

And again in terms of amplitude of a specific frequency,

$$A_\sigma = A_0 \exp\left(\frac{-k^2\sigma^2}{2}\right) \tag{D7}$$

This result agrees with Eqs (14) and (15) in Gkinis et al 2014.

As I've shown, the Gaussian filter yields the resulting amplitude of the diffused signal. In calculating diffusion length, we are fitting the power spectral density of the signal, and not the amplitude. Thus, we need to convert (6) and (8) to reflect the filter applied to the PSD rather than to the amplitude. Given that the power of a signal is defined as the square of the signal, we have

$$P_\sigma = P_0 \exp\left(-4\pi^2\sigma^2 f^2\right) \tag{D8}$$

and

$$P_\sigma = P_0 \exp\left(-k^2\sigma^2\right) \tag{D9}$$

where P_0 is the power spectral density of the original signal and P_σ is the power spectral density of the diffused signal.

Up to this point, σ has remained consistent and still refers to the standard deviation of the Gaussian in the depth before the Fourier transform, which is the diffusion length. When fitting a Gaussian to the isotope data, we do so in the frequency domain. Thus our fit gives us standard deviation of the Gaussian in frequency space, σ_{std}, which is not the same as σ diffusion length, the standard deviation of the Gaussian in the depth domain.

We need to know how to convert from σ_{std} back to σ, which tells us the standard deviation of (1), and thus diffusion length.

To find the standard deviation in frequency space, we must convert Eq (11) to be in the standard Gaussian form $G(f) = \exp\left(\frac{-f^2}{2\sigma_{std}^2}\right)$.

$$\exp\left(-k^2\sigma^2\right) = \exp\left(\frac{-f^2}{2\sigma_{std}^2}\right) \tag{D10}$$

$$-k^2\sigma^2 = \frac{-f^2}{2\sigma_{std}^2} \tag{D11}$$

$$-2^2\pi^2 f^2\sigma^2 = \frac{f^2}{2\sigma_{std}^2} \tag{D12}$$

$$\sigma^2 = \frac{1}{2^2\pi^2 2\sigma_{std}^2} \tag{D13}$$

$$\sigma = \frac{1}{2\pi\sqrt{2}\sigma_{std}} \tag{D14}$$

This is the conversion between the standard deviation σ_{std} of the Gaussian fit in the frequency domain, and the diffusion length σ, the standard deviation of the Gaussian in the depth domain.

Acknowledgments. This work was partially supported by a grant from the Spanish Ministry of Science and Technology.

References

Holme, C., Gkinis, V., and B. M. Vinther (2017), Molecular diffusion of stable water isotopes in polar firn as a proxy for past temperatures, *Journal*, *number*(), from-to.

Johnsen, S. (1977), *Integral Equation Methods in Scattering Theory*, John Wiley, New York.

Hsiao, G. C., E. P. Stephan, and W. L. Wendland (1991), On the Dirichlet problem in elasticity for a domain exterior to an arc, *J. Comput. Appl. Math.*, *34*(1), 1–19.

Lu, P., and M. Ando (2012), Difference of scattering geometrical optics components and line integrals of currents in modified edge representation, *Radio Sci.*, *47*, RS3007, doi: 10.1029/2011RS004899.

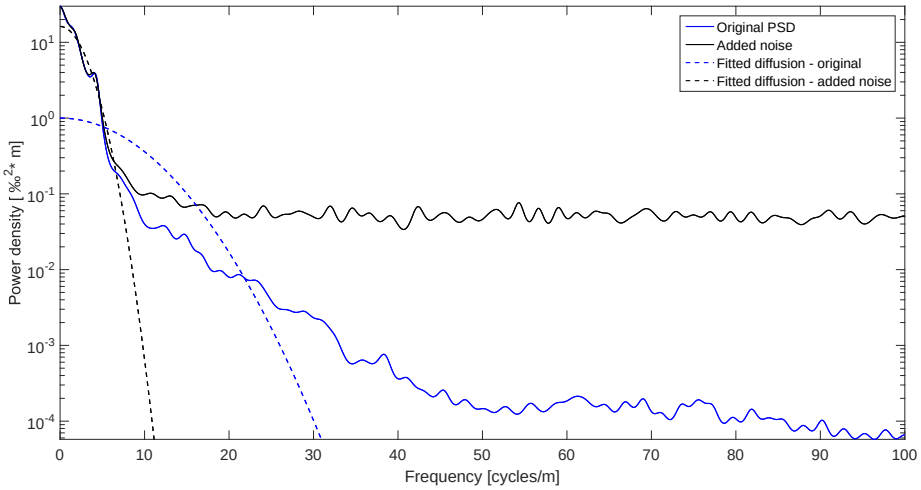

Figure 1. PSD of δD for a WDC section (solid curves). Blue curve represents the measured signal and the black curve represents the modified signal. The dashed lines represent the fitted Gaussian function which corresponds to the firn diffusion.

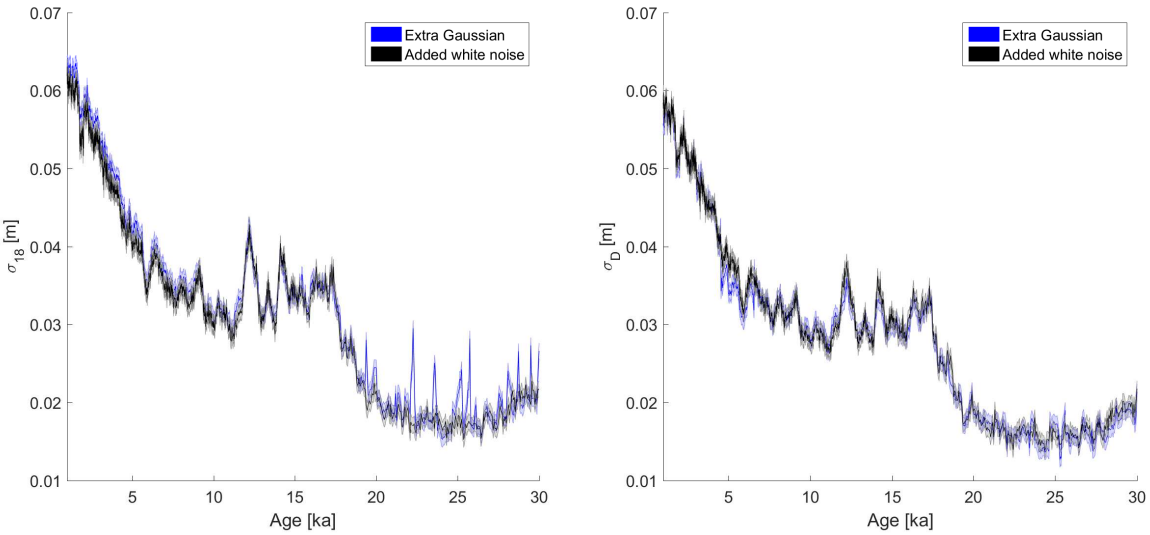

Figure 1. WDC diffusion lengths of δ^{18}O (left) and δD (right). The method that has an extra Gaussian is represented by the blue curve and the addition of white noise is represented by the black curve.

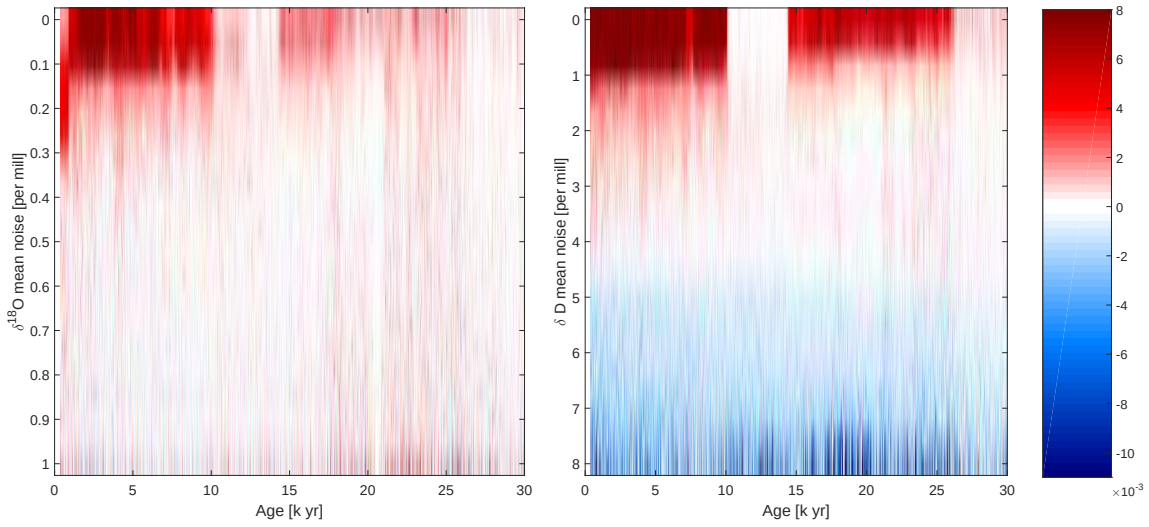

Figure 2. Gradient in estimated diffusion lengths with respect to noise level and age.

Table 1. Placeholder table

Treatments	Response 1	Response 2
Treatment 1	0.0003262	0.562
Treatment 2	0.0015681	0.910
Treatment 3	0.0009271	0.296

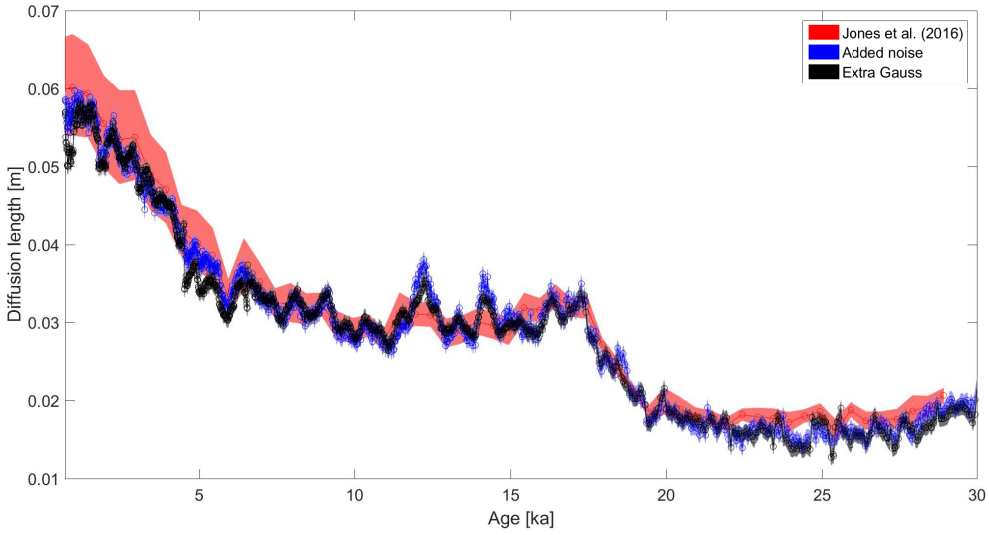

Figure 3. Estimated WDC diffusion lengths of δD compared with *Jones* [2016]

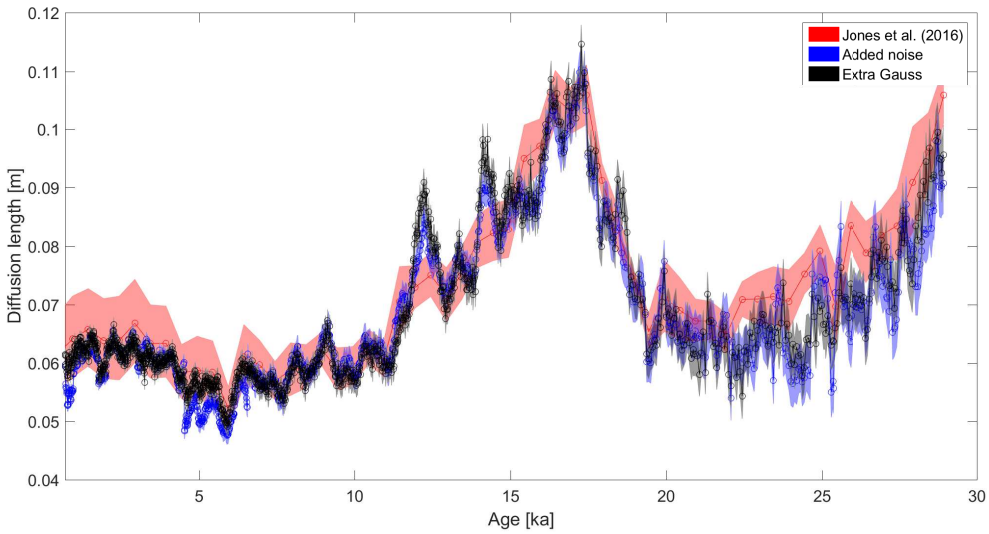

Figure 4. Thinning corrected WDC diffusion lengths compared of δD with *Jones* [2016]

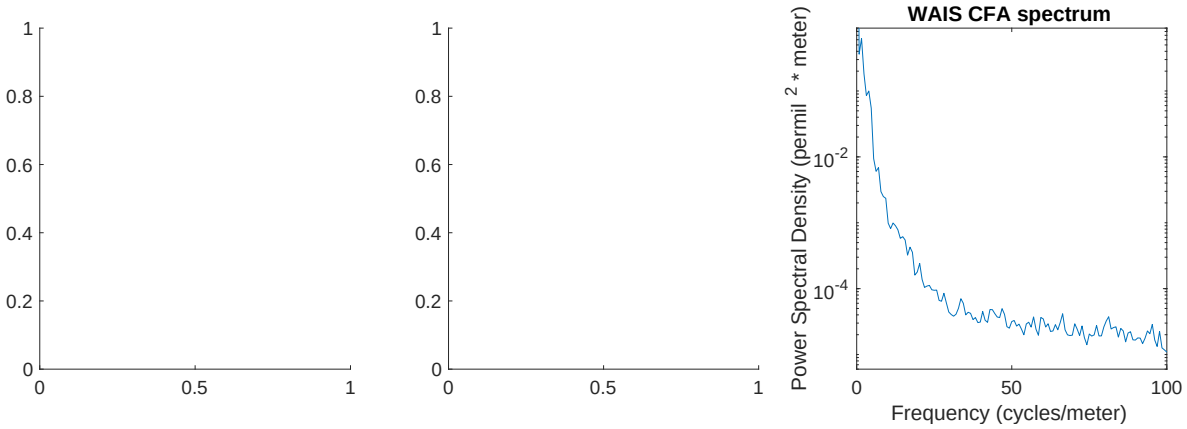

Figure 5. Replace with better figures from Christian comparing many discrete spectra and CFA spectra to illustrate the effect of the transition zone

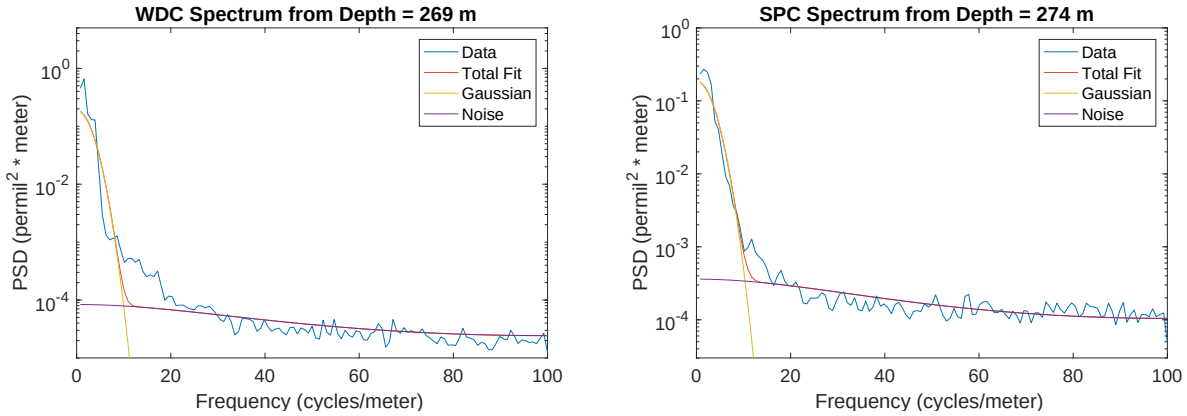

Figure 6. WDC and SPC spectra fits with single-Gaussian multi-function technique

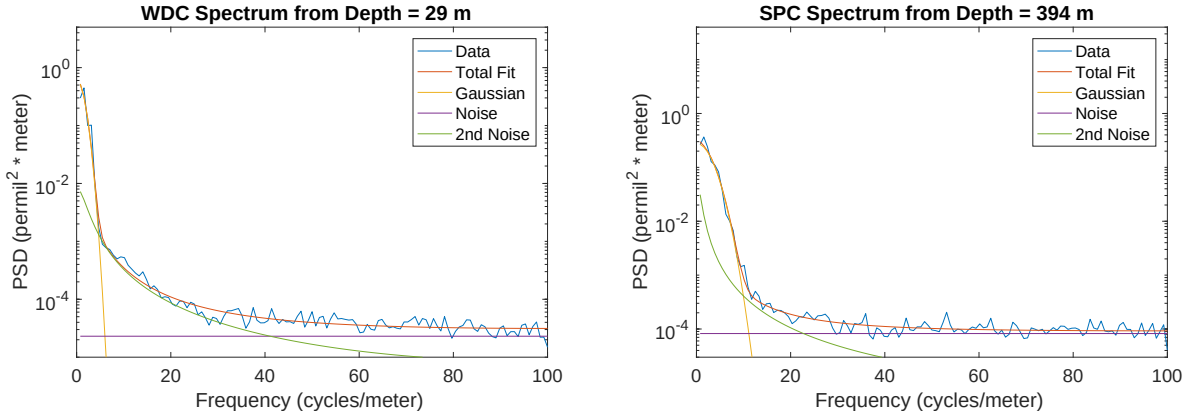

Figure 7. WDC and SPC spectra fits with a single Gaussian and two autoregressive noise functions

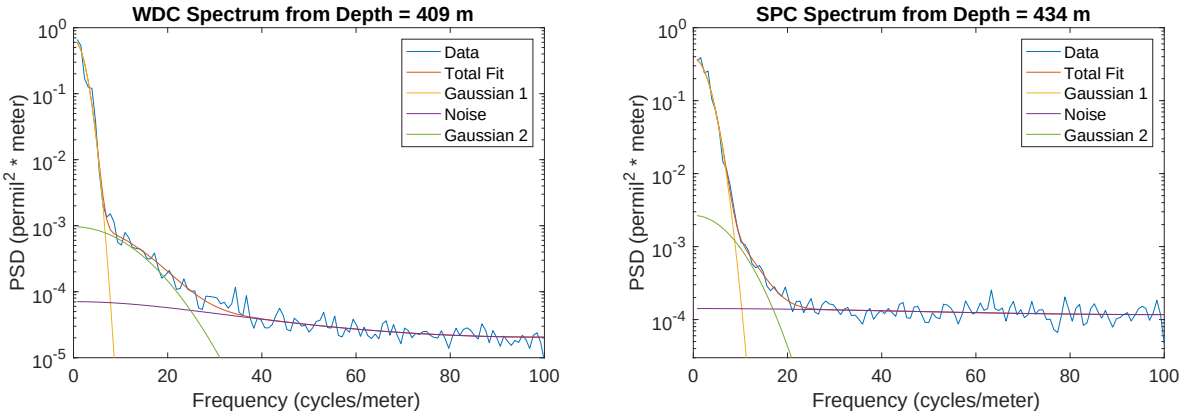

Figure 8. WDC and SPC spectra fits with double-Gaussian multi-function technique

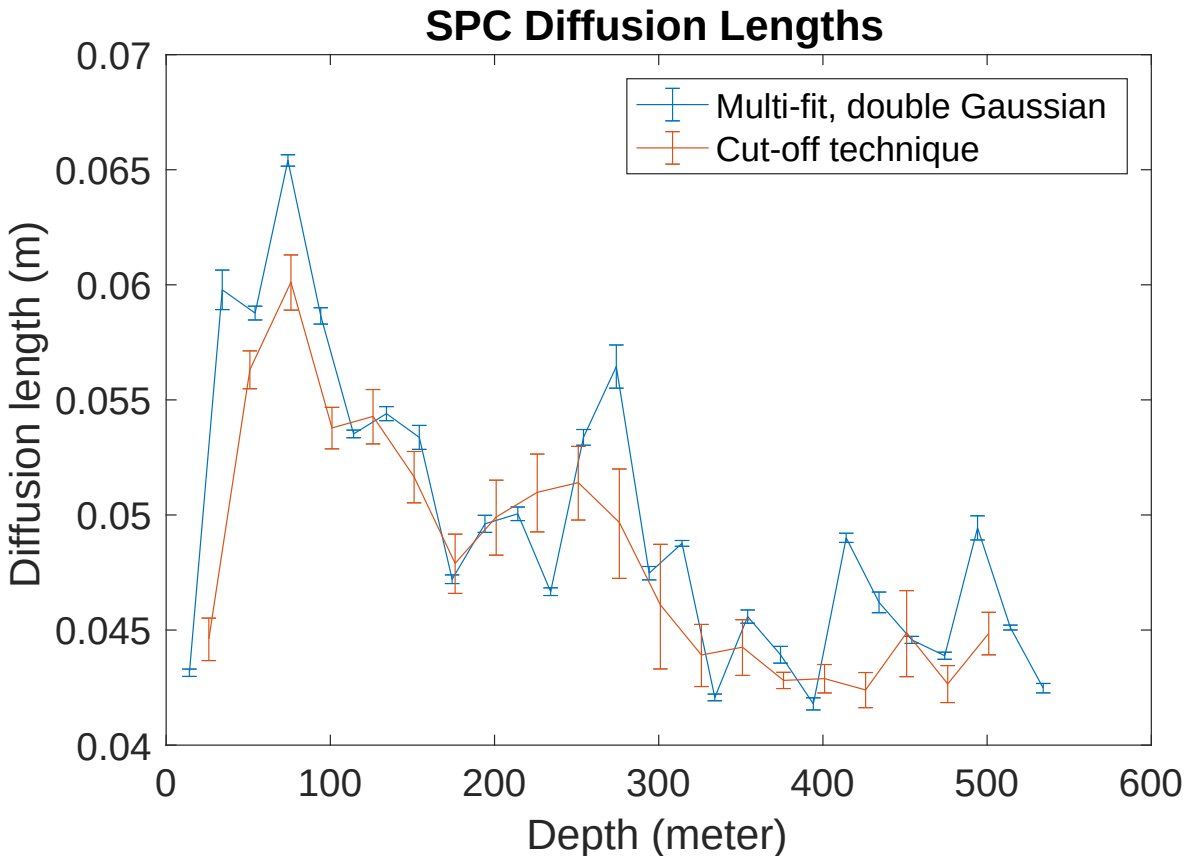

Figure 9. Comparison of diffusion length σ as calculated by cut-off technique and multi-function technique (single and double Gaussian) for WDC and SPC

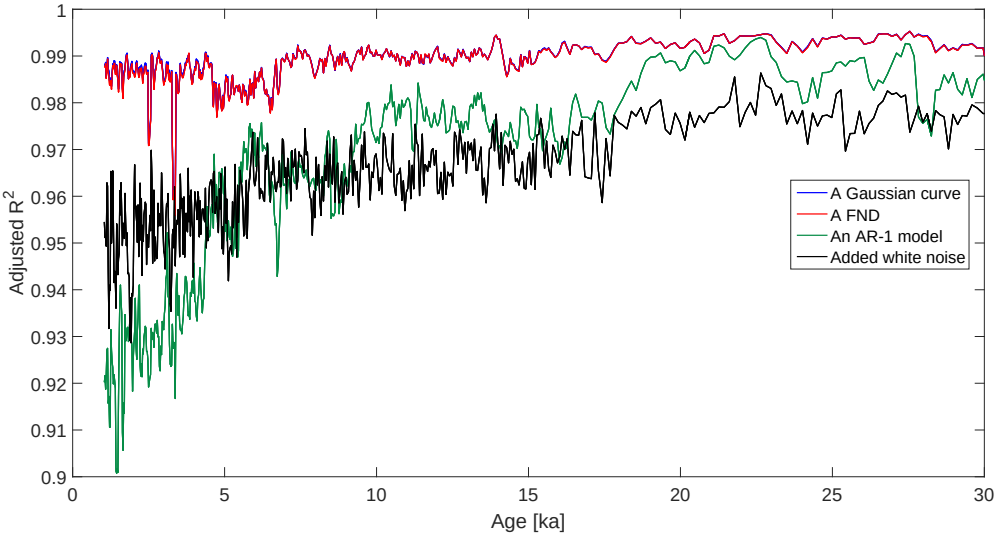

Figure 10. The adjusted goodness of fit calculations through age.

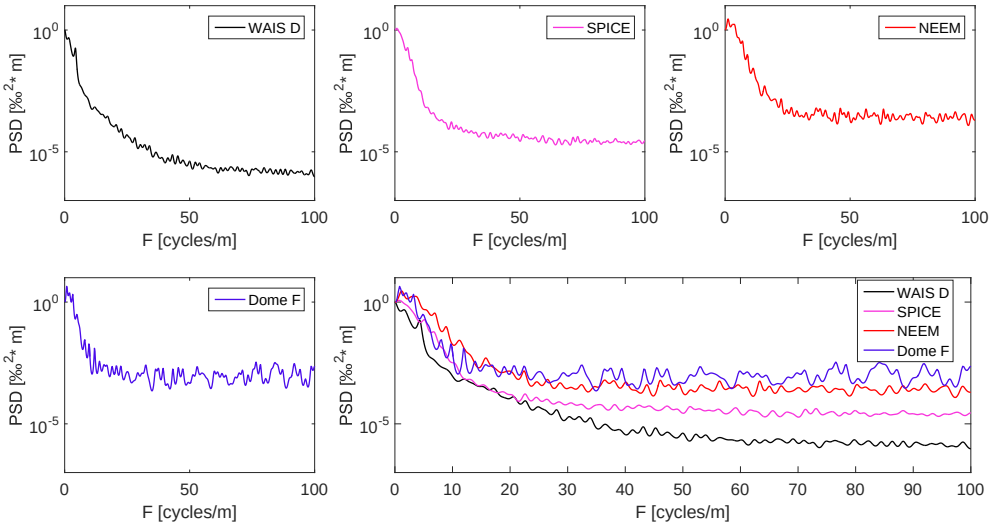

Figure 11. Normalized power spectral densities of δD records measured with different CFA systems.

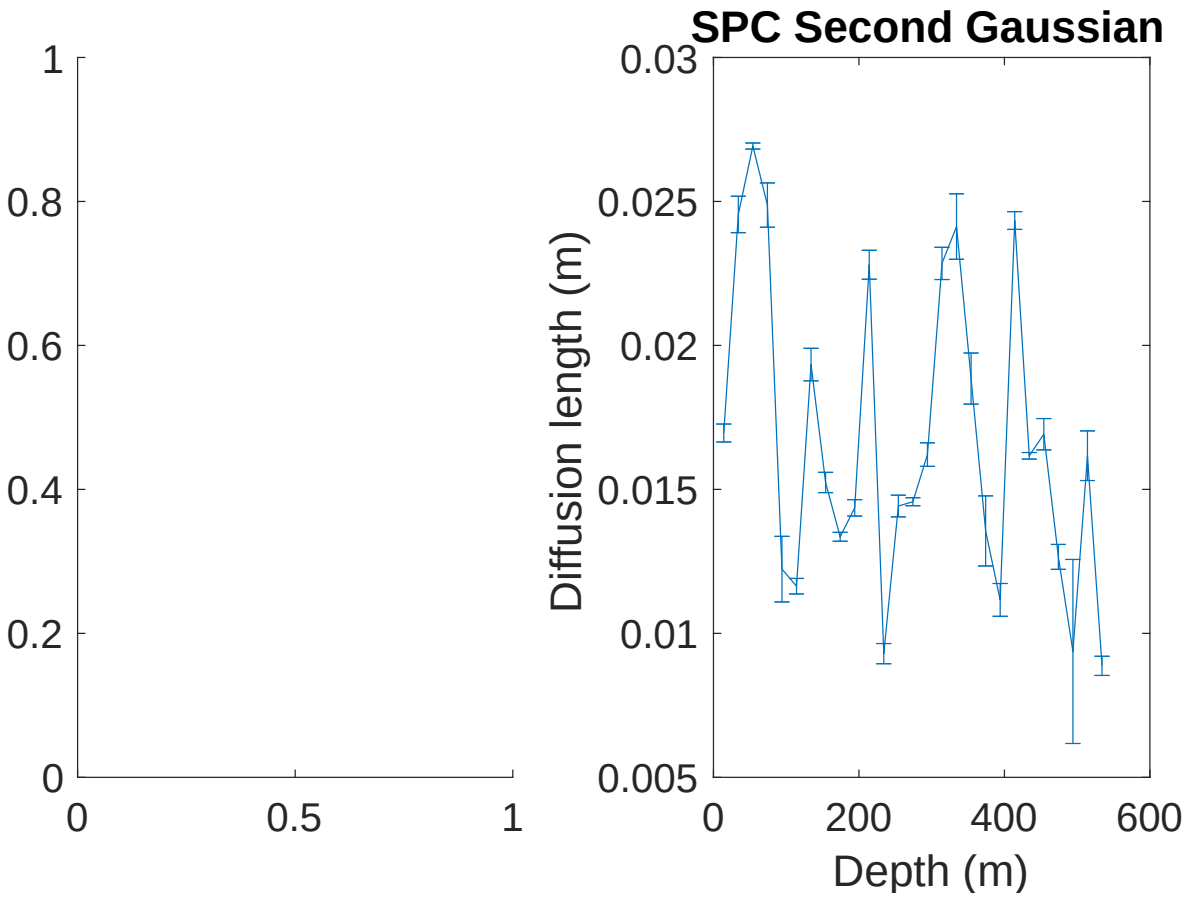

Figure 12. Plot the σ from the second Gaussian for WDC and SPC

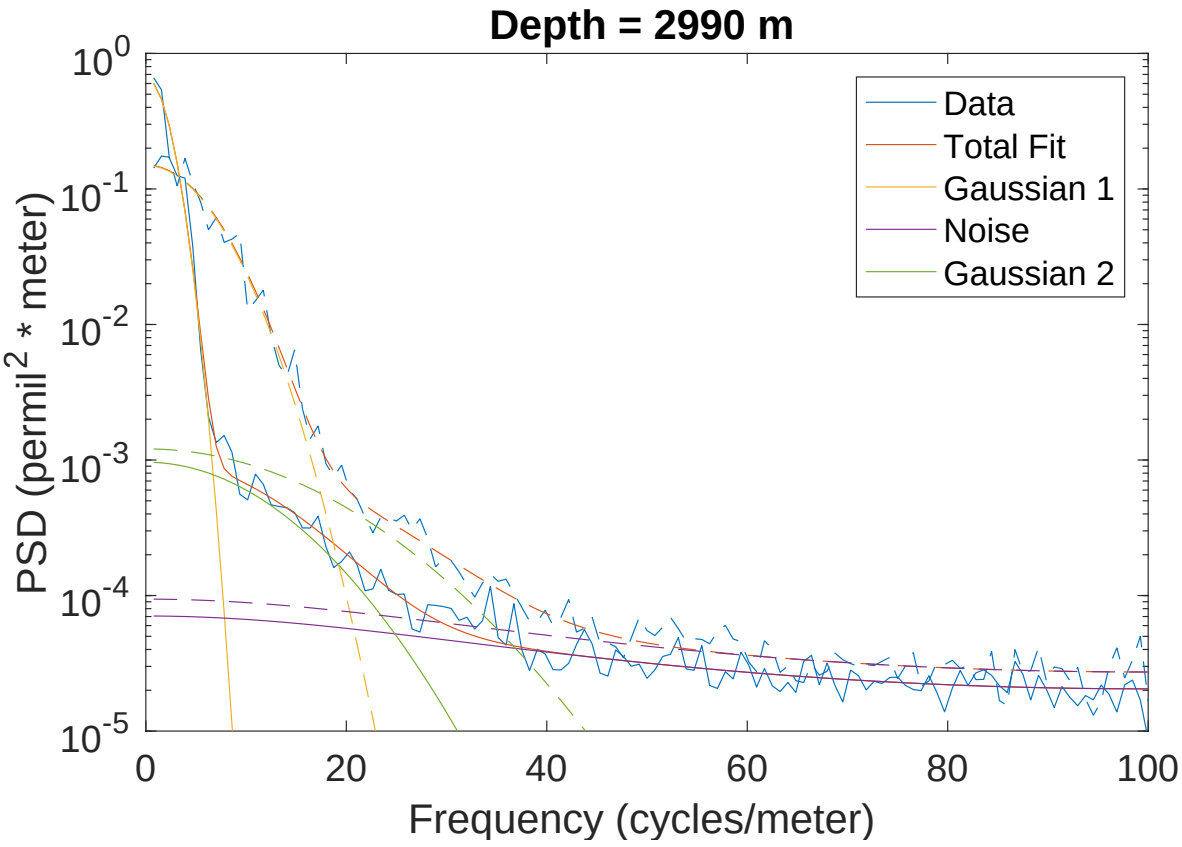

Figure 13. Showing how sigma2 depends on sigma1

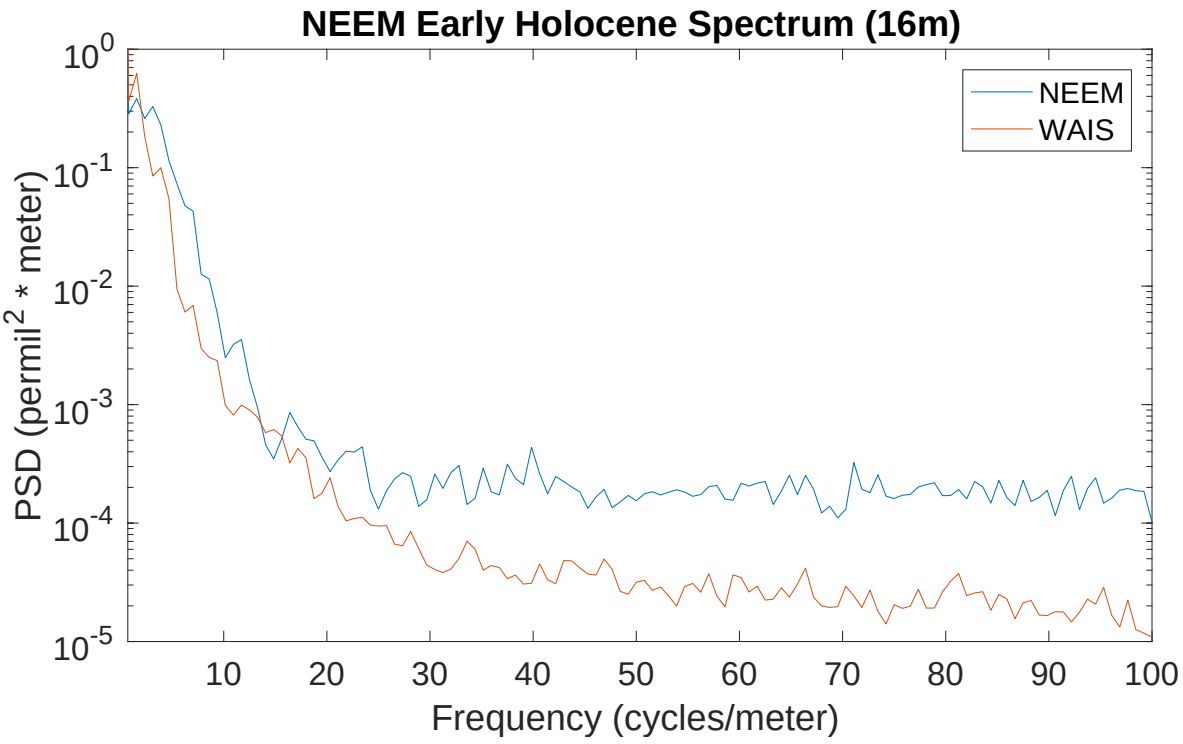

Figure 14. Spectra from NEEM 1382-1398 meters, CFA data measured on a different system