

0.1 σ Estimation Method and Discussions

The general idea of the optimal σ estimation method is to back diffuse a depth series defined on an interval where the time span (i.e. the number of peaks and troughs expected in that section) is known. This allows us to use the diffusion length as a tuning parameter, to find the diffusion length estimate which generates the right number of peaks and troughs and fulfills the imposed constraints in the back diffused depth series. If more than one diffusion length meet these constraints, the largest diffusion length to still fulfill the constraints is sought after. This diffusion length is then assumed to be the optimal guess on the diffusion length in that interval, which allows for a temperature estimate, following the temperature dependence of the diffusion length, as described in Section ??.

The algorithm consists of two modules, where one module describes the numerical back diffusion, given an inputted depth series, core specification and specific σ_0 estimate (which is either manually inputted or estimated from the spectral analysis). The flowchart describing the processes carried out in this module can be seen in Figure 0.1. Many of the sections in this module are only necessary in the initialization of the algorithm as these parts do not change if the inputted diffusion length estimate is changed. In Figure 0.1 anything carried out above the *Frequency Filters* block to the left is only computed once, and the same with anything to the right of it, except the σ_0 estimate. The density and diffusion profile calculations, the spectral analysis and the Wiener filter construction is inherent to the depth series alone, and these analyses are carried out as previously described in this thesis.

The second module is responsible for the optimization. This module examines the parameter space containing the diffusion length estimates, and utilizes a constrained direct search method to find the optimal diffusion length estimate. This method is illustrated through a flowchart in Figure 0.2.

0.1.1 title

0.1.2 Constrained Peak Detection

The main constraint when back diffusing these depth series is the number of years between the two volcanic events detected in the ice. The number of especially winters between the two events are fixed to $N_P = 33$, as this number of winter is expected in the interval, even with a two month variation from the estimated event position. The counted number of summers can on the other

hand vary a bit.

For better peak detection, a number of other constraints have been implemented. Since the data is a proxy for a continuous physical process, the temperature, it is reasonable to set up constraints representing some of the logical expectations for this type of signals.

Firstly, restrictions may be demanded of the distance between peaks. The annual layer thickness, λ_A , may help with setting some limitations on the peak distances, as it is not likely for the average peak distance to be much smaller than λ_A . The peaks are expected to show the same annual cycle as the rest of the signal, representing summers.

Secondly, for the prominence of the peaks, i.e. the amplitude of the signal at a depth, it is assumed that individual peaks may not have a prominence of less than a certain percentage of the standard deviation of the signal at the given depth. This makes certain that smaller peaks or troughs, maybe representing a warm period in a winter or a cold period in summer, are not counted as annual peaks or troughs. This is one way to constrain peak prominence, but a more efficient and accurate way might be to consider the amplitude of the entire ice core signal. As it may be assumed that the amplitude and prominence at a given depth, will be somewhat smaller than the prominence of the peaks at a shallower depth, due to the general diffusion in the firn. By analyzing the amplitude attenuation of the ice core, the average attenuation at a given depth could be used as a restriction for the peak prominence. This is something that would have been implemented, if time had allowed it.

Thirdly, a constraint on how the general pattern of the trough and peak detection must look is imposed. As the temperature variations represent the change from summer to winter, it is assumed that the general pattern must be to detect a peak P , then a trough T , then a peak P , and so on, creating a pattern of $...PTPTPTP...$. Since the deposition time of volcanic material in the ice is assumed to be Gaussian, the pattern is not restricted to start with either a peak or a trough, as this may vary when drawing a location from the distribution. Thus, the number of peaks is set at $N_P = 33$ and the number of troughs is accepted with a variation, as long as the general pattern is intact.

Finally, the diffusion length estimate is kept at a positive value with an upper limit, that can be set manually, depending on the conditions of the site and the depth.

The optimal parameter choices for the constraints are presented in:

N_P	33
N_T	33 ± 1
Peak(trough) prominence	50 % of SD_{signal}
Peak(trough) distance	50 % of λ_A
$[\sigma_{\min}, \sigma_{\max}]$ [cm]	[0, 15]
Pattern	...PTPTPTPTP...

Table 0.1: The general constraints used in the method to optimize the diffusion length estimate.

0.1.3 Corrections on the Computed σ estimate

As mentioned in Section ??, the estimated diffusion length needs to be corrected for a number of different influences: the sampling diffusion, the ice diffusion at a given depth and the thinning function at a given drill site.

This section will describe how each of these influences have been taken care of in the work.

0.1.3.1 Ice Diffusion, σ_{ice}

0.1.3.2 Sampling Diffusion, σ_{dis}

0.1.3.3 Thinning Function $S(z)$

0.1.4 Constraints

This section will contain a walk through of the general differences between constrained and non-constrained optimization of the diffusion length. In Figure 0.4 an example of the results of the optimization with and without constraints can be seen.

0.2 Method Testing and Stability

Throughout this section a number of different tests of the algorithm will be presented. The tests are performed to examine the stability of the method,

the accuracy of the Laki and Tambora positions and how the choice of parameters(interpolation, spectral transform type, λ_A estimation, peak prominence and peak distance) changes the resulting diffusion length estimate.

0.2.1 No constraints versus constraints

0.2.2 Effects of Interpolations

0.2.2.1 Interpolation of Data Before Deconvolution

The first interpolation is needed, if the fast spectral transforms FFT or FCT are used, as one of the conditions of the algorithms is that the data are evenly spaced. At first, this was implemented in the analysis, but this had the risk of excluding some information that might lie in the unevenly sampled data. Later, the method was abandoned in favor of implementing a nonuniform spectral transform (NUFT or NDCT), which is slower than the FFT and FCT, but carries all information from the unevenly sampled signal into the spectral domain. Luckily, this nonuniform transform needs only be carried out once, as the inverse transform, i.e. resampling in time domain, can be done uniformly without loss of information and any future spectral transforms can then be performed through FFT or FCT.

Even though the first interpolation method was later abandoned, some analysis was carried out with it to examine the effect of the size of the resampled, interpolated data on the final diffusion length estimate. Examples of a resampled signal can be seen in Figure ?? and Figure ?. Figure ?? shows how sample resolution affects information from the signal. The higher sampling resolution, the more information is retained. But higher sampling resolution also means more data to be analyzed, which might slow down any analysis algorithms developed. This might create some headache if an entire ice core length of a couple thousand meters should be examined, but for this study only a few meters are of interest, and thus it should not create delays in the computation time.

To examine the effect of the resampling resolution on the final diffusion length estimate when conducting a spline interpolation before carrying out the back-diffusion, the full diffusion length analysis has been performed with 100 new interpolation resampling sizes in the range $[\Delta_{\min}; \Delta_{\max}]$. This gives an idea of the stability of the method considering both sample size of the raw data and resampling by interpolation. The minimum and maximum interpolation samplings are presented in Table 0.2 and an illustration of the test results can

be seen in Figure 0.5.

Figure 0.5 shows that if the sampling size of the interpolation is decreased, it becomes difficult for the algorithm to determine a diffusion length that fulfills the constraints. This is due to the spectral transforms and back diffusion method being sensitive to smaller variations that the spline interpolation introduces to the signal, as can be seen in the first panel in Figure 0.3. Furthermore, Figure 0.5 shows a less stable diffusion length estimate as the resampling size increases, and a general tendency to result in higher diffusion lengths as many features become washed out in the signal and needs more enhancement by cranking up the diffusion length estimate, see final panel in Figure 0.3. The interpolation before deconvolution is only necessary for running the method with the spectral transforms that are based on uniform sampling, i.e. the FFT and the DCT. For these two methods a choice of interpolation size can be made, but the general setting in the algorithm is to resample at the smallest sampling size found in the depth interval, Δ_{\min} .

0.2.2.2 Interpolation of Data After Deconvolution

The second interpolation is carried out after deconvoluting and back-diffusing the signal, but before detecting peaks. Splines are especially effective when trying to find features like peaks in data which underlying signal is continuous, smooth and differentiable, but the sampling is discrete and thus the data are discrete and non-smooth. The isotopic signal under examination here is assumed to be truly smooth and continuous throughout the core - unless any gaps are present. Thus the cubic spline interpolation is a good tool for estimating a higher resolution version of the final back-diffused data series to use for peak detection. This makes the detection of peaks and troughs more precise, as there might not be a discrete data point exactly at the top of a peak, but the spline interpolation then estimates where the most likely top of the peak must be, on the basis of the existing data. Examples of three different interpolation samplings are presented in Figure 0.6. The effect of resampling after deconvolution on the final diffusion length estimate is illustrated in Figure ??.

As the actual isotopic signal is continuous and that the discretization is introduced by different measurement samplings, it is assumed that the spline interpolation after deconvolution results in a more likely peak detection, when decreasing the numerical resampling size. Therefore, for the method, a resample size of $\Delta_{\min}/2$ is chosen for interpolation after back diffusion and

Site	Δ_{\min} [cm]	Δ_{\max} [cm]	Δ_{OG} [cm]
A	1.0	10.6	3.8-4.0
B	1.0	11.7	3.8-4.0
D	1.0	12.0	3.7-3.9
E	1.0	11.4	4.1-4.4
G	1.0	10.3	4.0-4.2

Table 0.2: Minimal and maximal new sample resolution used for testing interpolation before back-diffusion. Each test is run with 100 different new sample resolutions between Δ_{\min} and Δ_{\max} .

before peak detection. In Figure 0.7 The diffusion length estimate versus the resampling size after deconvolution and shows a convergence towards a fixed diffusion length as sampling size is decreased, and a much noisier diffusion length estimates as sampling size is increased. Furthermore, at certain larger sampling sizes the sought after pattern of peaks and troughs is not even reached. This resampling is carried out both for the FFT, DCT and NDCT spectral analysis methods, as the inverse NDCT can be resampled uniformly without any loss of information.

0.2.3 Spectral Transform's Effect on Diffusion Length

0.2.3.1 Visual Inspection

0.2.3.2 Speed Examination

0.2.4 Laki and Tambora as Gaussian Distributions

0.2.4.1 Varying both Laki and Tambora Position

0.2.4.2 Varying only Tambora Position

0.2.4.3 Varying only Laki Position

0.2.4.4 Variation Corresponding to 2 Months

0.3 Possible Algorithms Upgrades

This section will discuss ideas on how to improve on both the method and the algorithm

0.3.1 Peak Detection

0.3.2 Optimization Routine

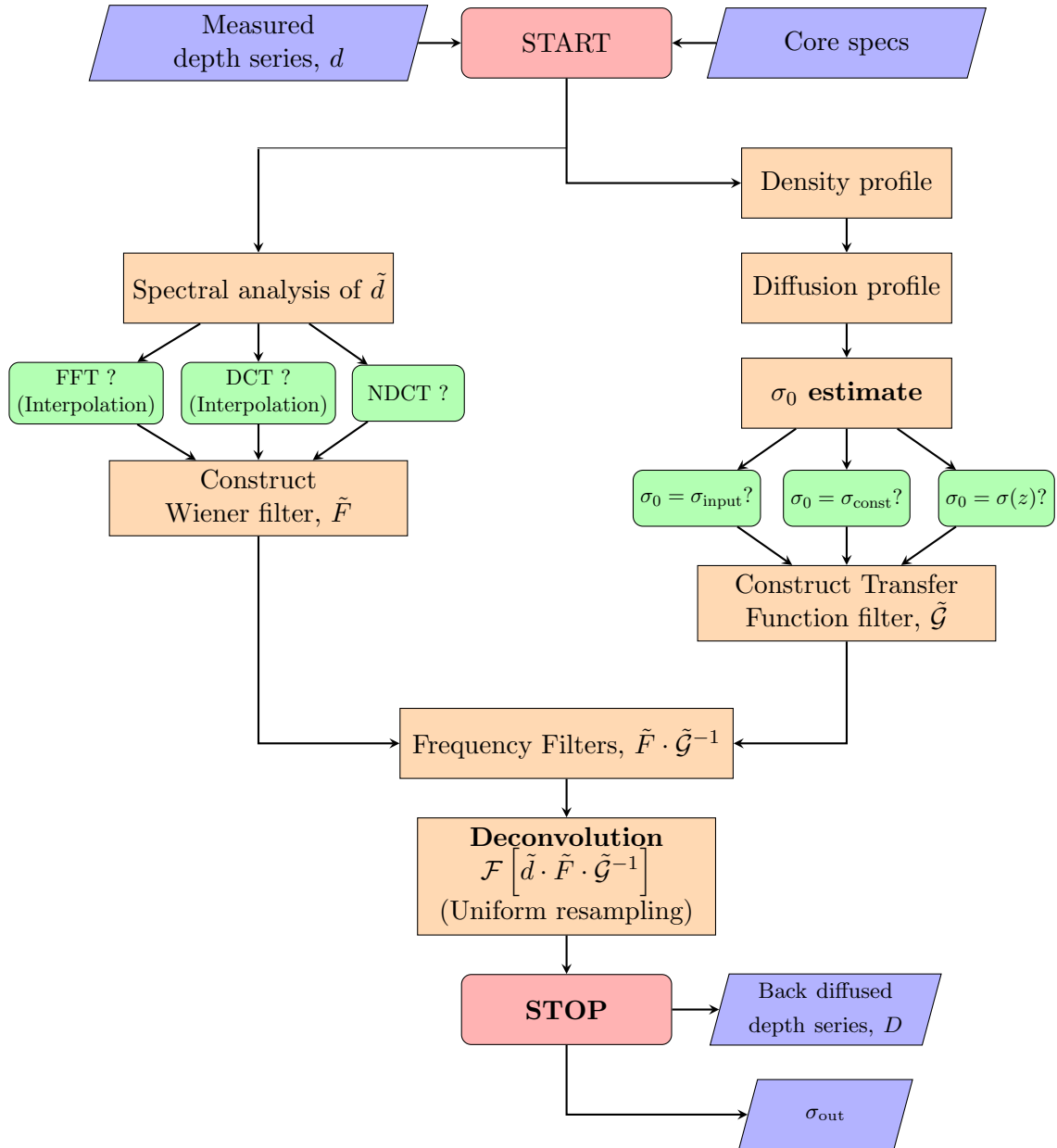


Figure 0.1: Flowchart of initialization method for back diffusion of a depth series given a diffusion length estimate.

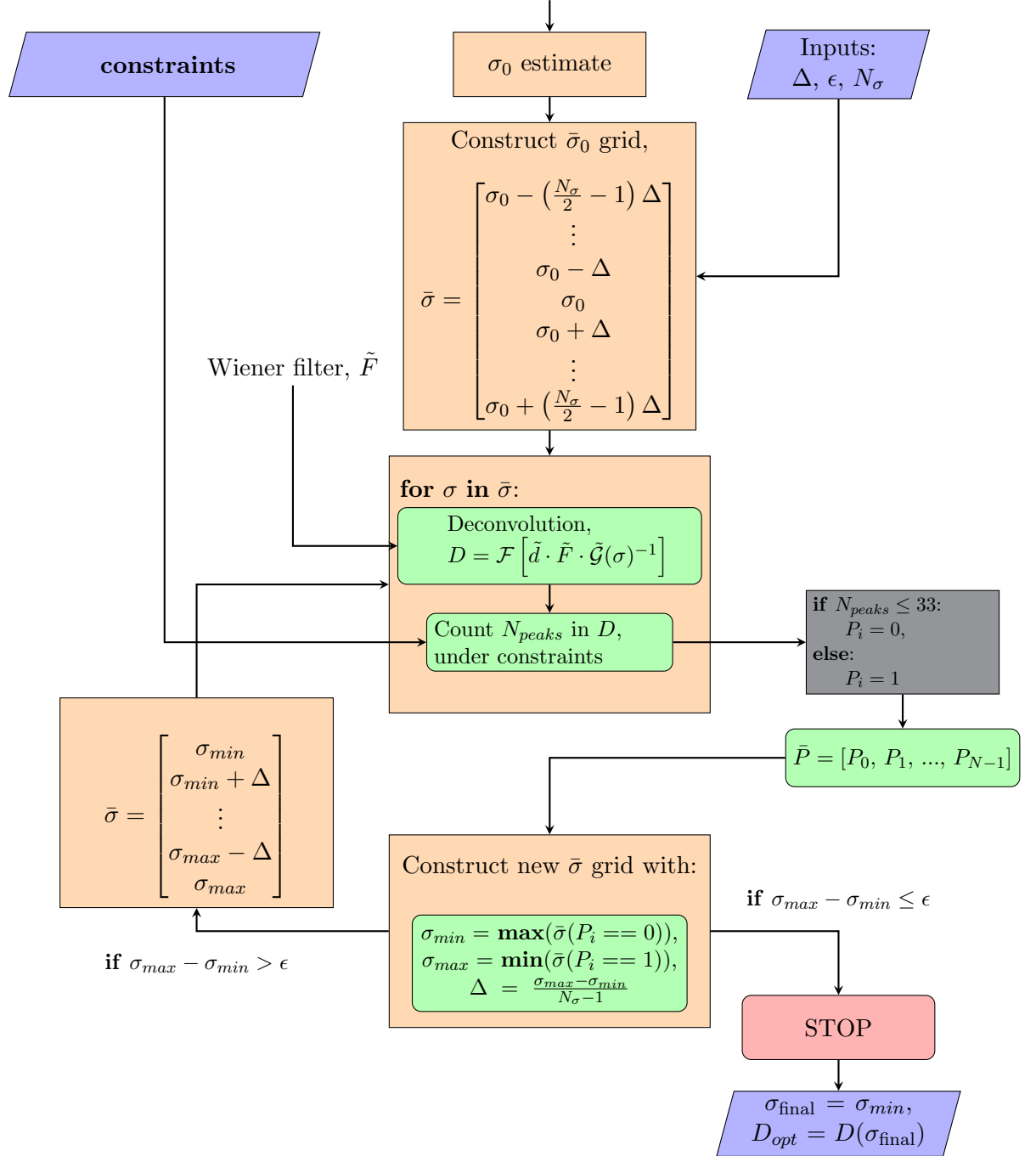


Figure 0.2: For Method (optimization)



Figure 0.3: Site A, illustration of the effect of five different cubic spline resamplings before deconvolution. Original sample sizes lie in the interval [3.80; 4.00] [cm]. Resampling at smaller sample sizes show a tendency to restore some signal frequencies that are not necessarily inherent in the original signal. The resample should thus not be chosen too small as this would introduce some false signal into the results.

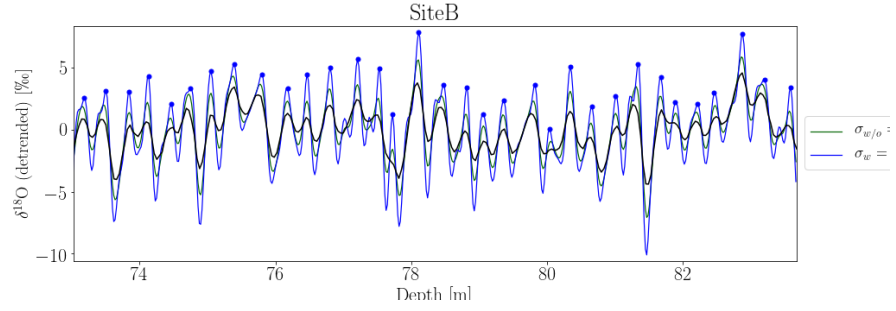


Figure 0.4: Example of how the imposed constraints effect the final diffusion length estimate for the Laki to Tamborad depth section of the core drilled at Site B. The black line shows the data, the green the back diffused data using a method with less constraints, and the blue shows the back diffused data when using the imposed constraint. The blue dots represents the peaks counted in the constrained method.

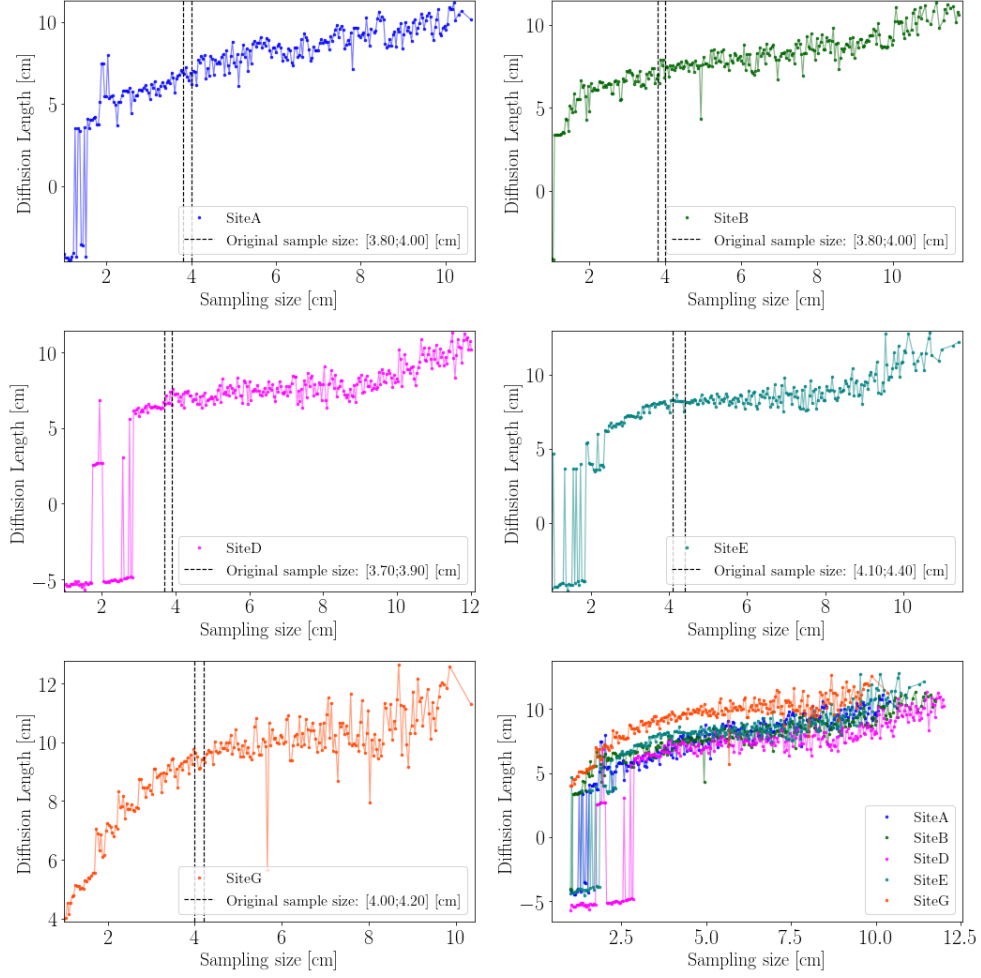


Figure 0.5: Diffusion length estimates versus resamples through cubic spline interpolation before deconvolution for Alphabet cores from sites A, B, D, E and G.

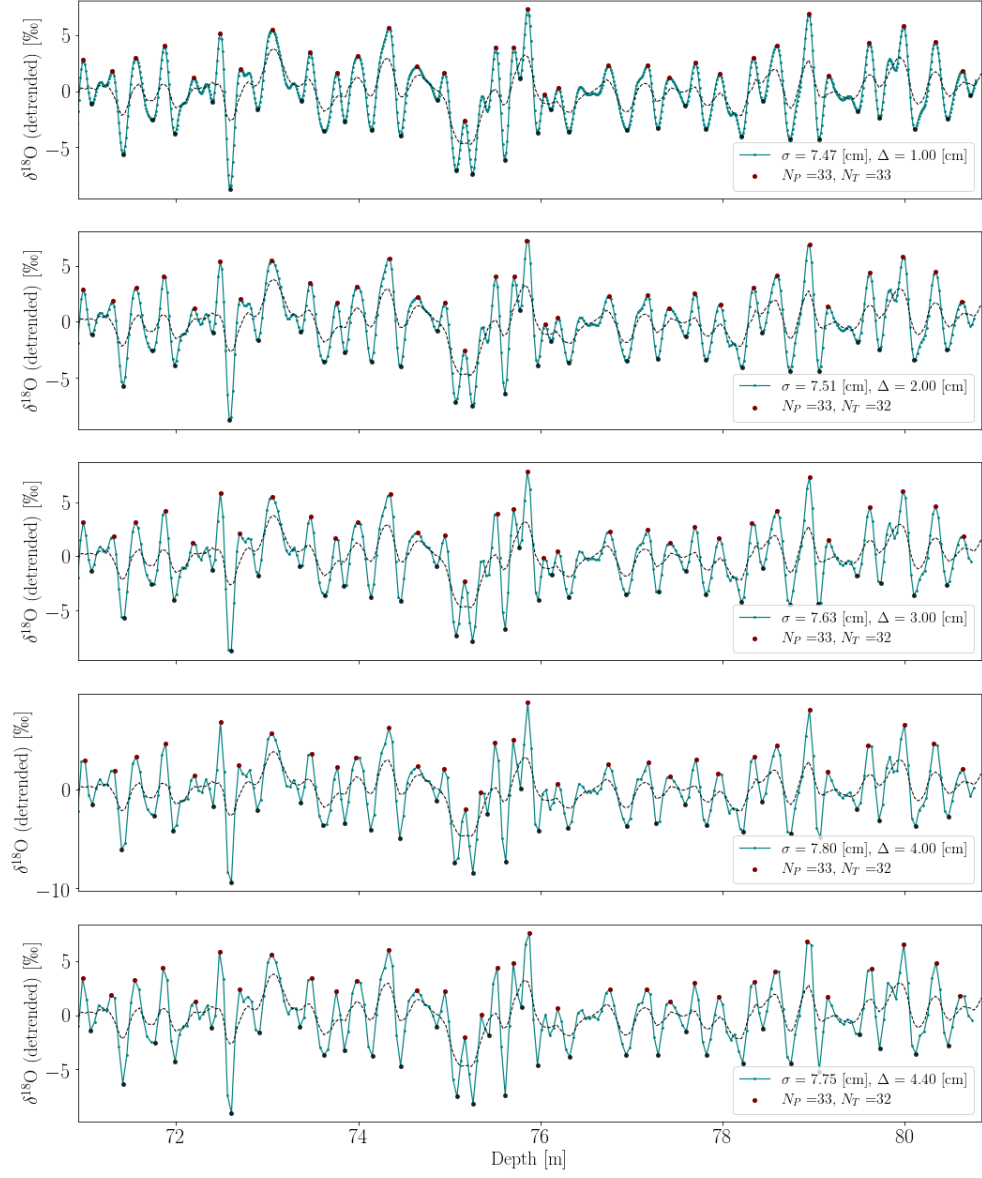


Figure 0.6: Site A, effect of cubic spline interpolation after signal has been deconvoluted. The interpolation is introduced to make peak detection more stable.

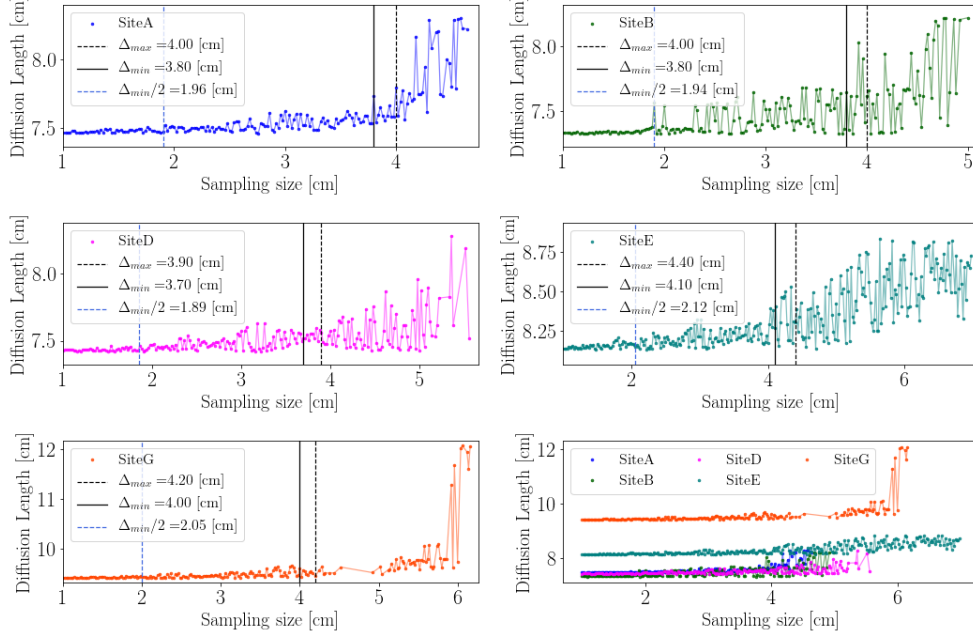


Figure 0.7: Final diffusion length estimate, given new resample by cubic spline interpolation after deconvolution for Alphabet cores from sites A, B, D, E and G.. The original sample size interval is illustrated as black vertical lines.

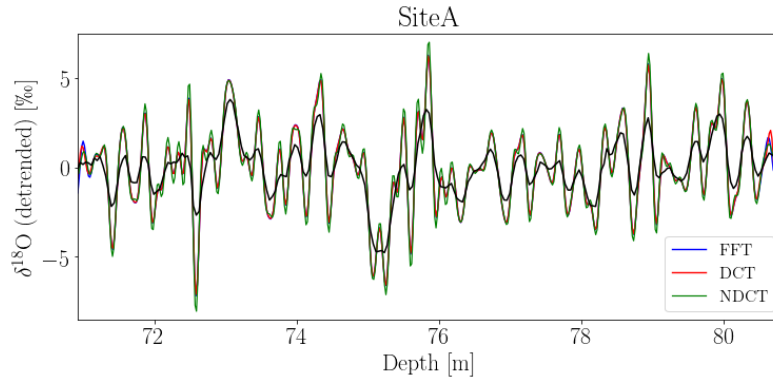


Figure 0.8: A visual example of the differences in final back diffused data when using different spectral transforms.

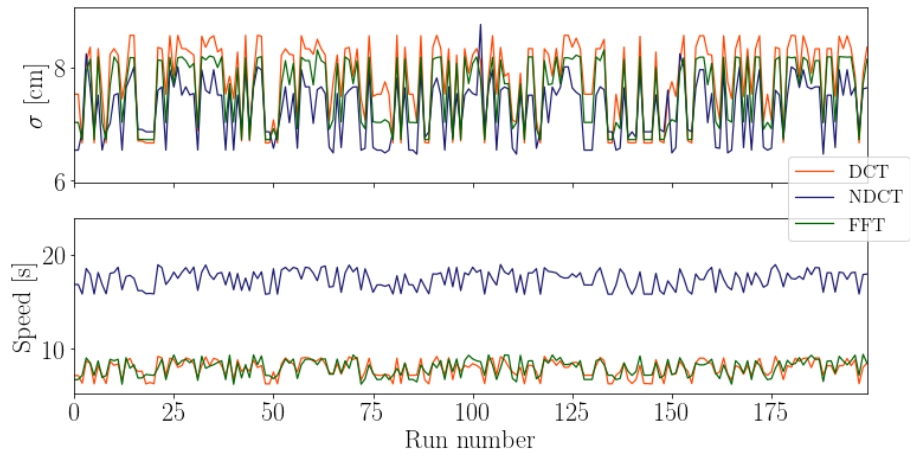


Figure 0.9: Diffusion length estimate along with speed of algorithm given the three different spectral transforms examined in this work. The volcanic event depths have been drawn from a Gaussian distribution with a width corresponding to 1 month around the estimated mean depth of the event.

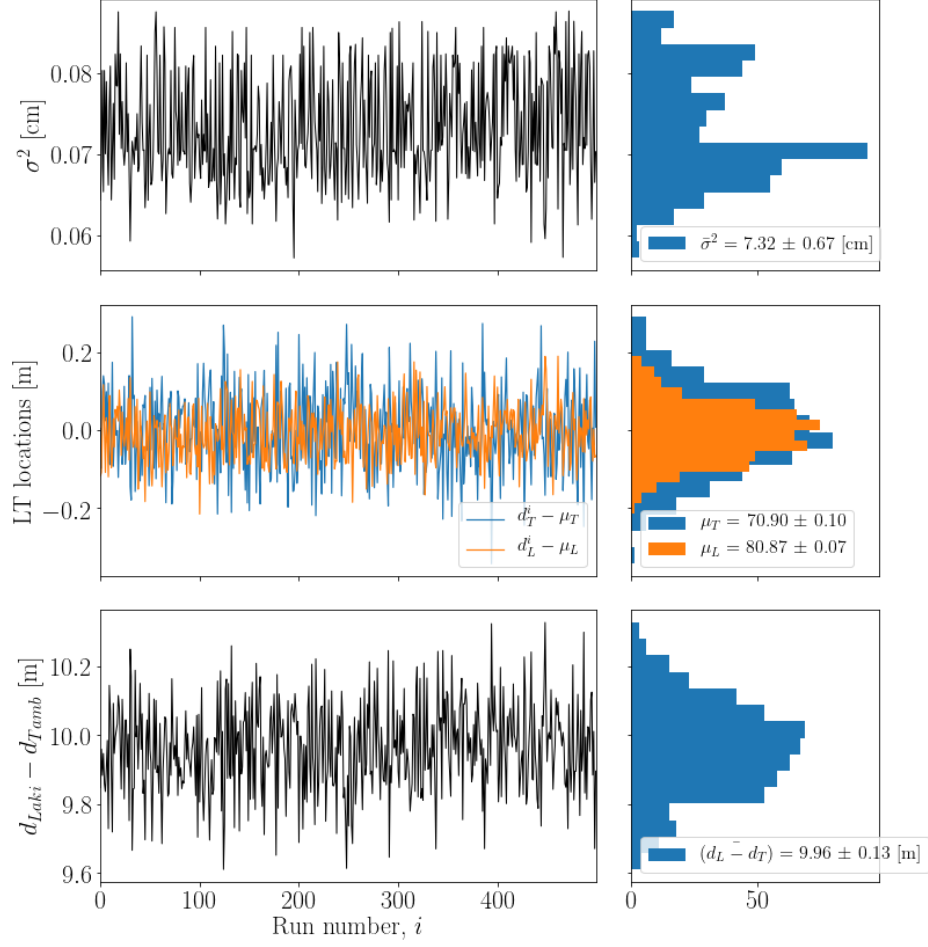


Figure 0.10: Diffusion length estimates when varying the depth locations of the volcanic events. The locations of both Laki and Tambora events have been drawn from Gaussian distributions as the ones presented in Section ??.

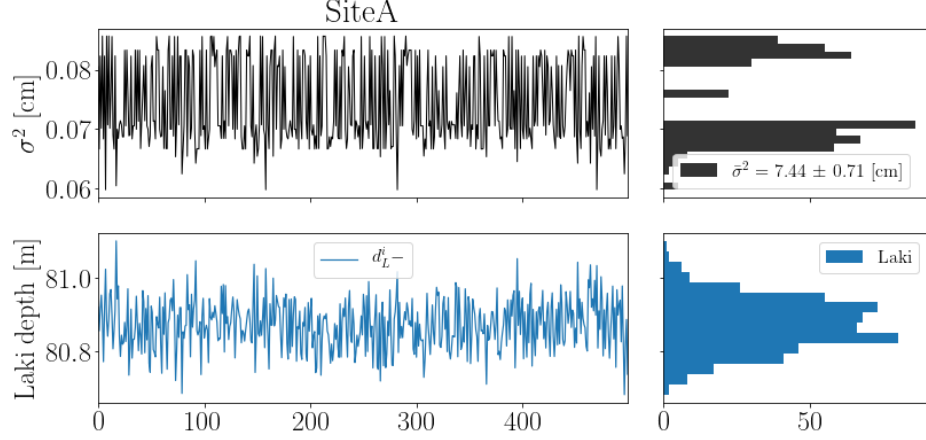


Figure 0.11: Diffusion length estimates for Site A when varying only the Laki volcanic event. The locations are drawn from Gaussian distributions as the ones presented in Section ???. The depth section is kept at a constant length, corresponding to the mean distance value, $\bar{d}_{\text{Laki}} - \bar{d}_{\text{Tambora}}$.

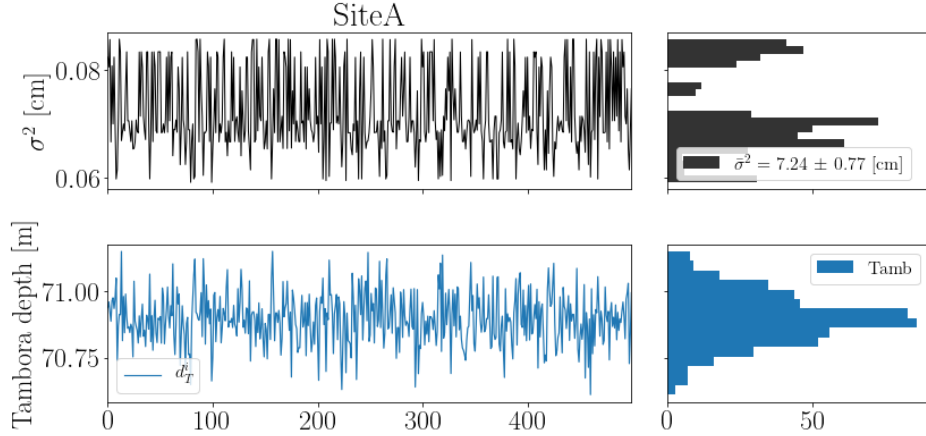


Figure 0.12: Diffusion length estimates for Site A when varying only the Tambora volcanic event. The locations are drawn from Gaussian distributions as the ones presented in Section ???. The depth section is kept at a constant length, corresponding to the mean distance value, $\bar{d}_{\text{Laki}} - \bar{d}_{\text{Tambora}}$.

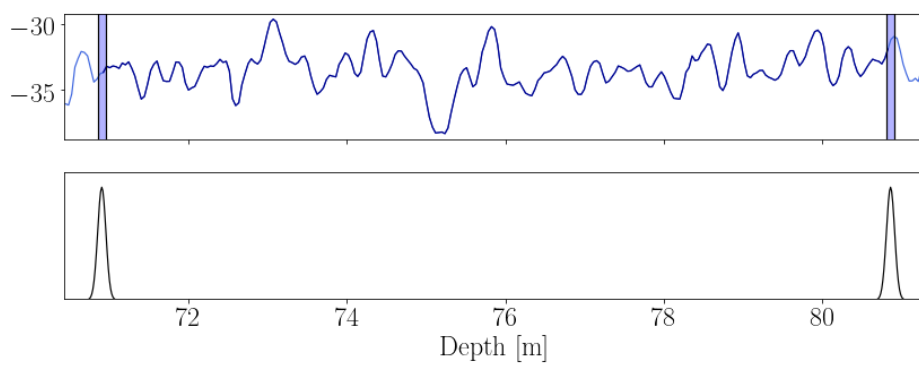


Figure 0.13: Illustration of the method implemented to manage the uncertainty of the exact depth location of the volcanic events. The method establishes a Gaussian distribution with a mean of the estimated middle of the volcanic event and a standard deviation of what corresponds to two months.

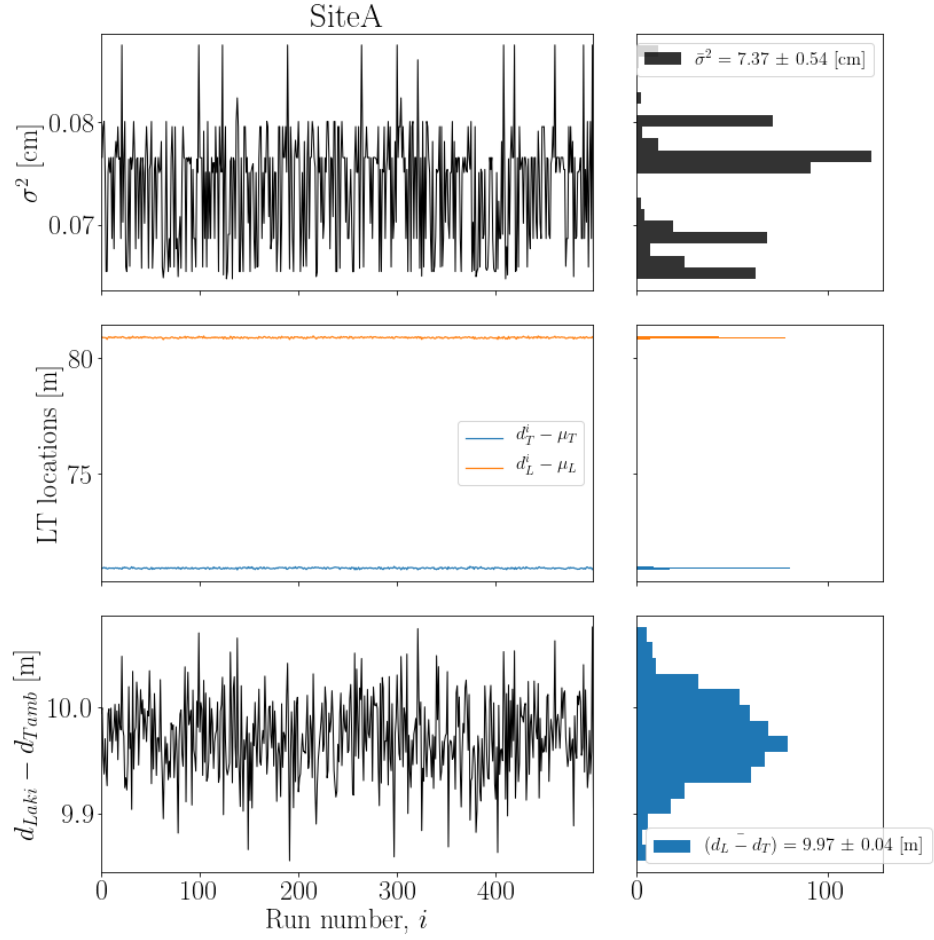


Figure 0.14: 500 runs with locations of Laki and Tambora events drawn from a Gaussian distribution with a standard deviation of two months. Using NDCT as spectral transform.