The data obtained through various experimental measurements are easily compared with a time series, as they typically show some quantity measured all along the depth of an ice core. This depth is often, at short intervals, treated as a regular linear time series thus making it possible to use some of the known signal analysis methods. Of course, when considering the entirety of an ice core, the linearity disappears as thinning and compression makes the depth series non linear. But when considering short lengths of core it is possible to estimate a linearity, assuming conformity in this specific layer.

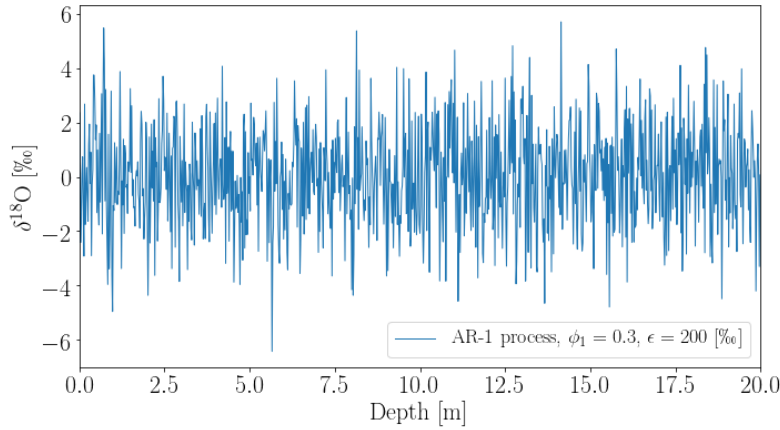## 0.1 Diffusion Illustrated through Synthetic Data

SIGNAL-SYNTHDATA: Write about synthetic data generation.

Figure 0.1

0.2 Back Diffusion

Due to diffusion in firn and ice, some of the water isotopic signal is lost. Some of this signal can be restored by investigating the diffusion process, and through filtering and deconvolution techniques(REFERENCES). For the data of this thesis two different restoration techniques are presented: a spectral method, determining the effect of mixing and diffusion as a spectral filter(REFERENCES), and a kernel restoration method much like the ones used to restore pixel resolution in images (REFERENCES).
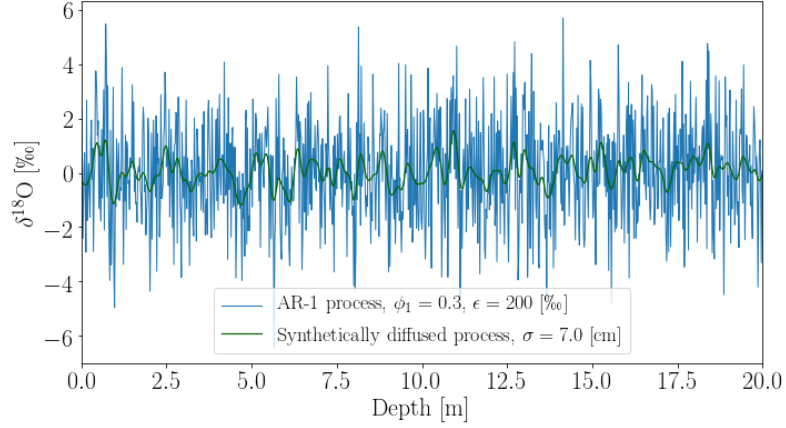
Figure 0.2

0.2.1 Spectral Analysis

0.2.1.1 Power Spectral Densities

A very useful tool for analyzing signals exhibiting oscillatory effects is analysis of the signals power spectrum. Instead of considering the signal in time, it is transformed to the spectral domain, where it is possible to obtain an estimate of both the signal and the underlying noise. This is crucial for enhancing the signal and filtering away noise. But to be able to examine these effects, first the data must be transformed. A range of different methods may be used to compute the frequency transform of the depth series, here I present the three I have been working with. Since the data are discrete and experimental, I will be presenting the discrete and applicable mathematical models.

When considering a signal, it may be of interest to investigate how the energy of said signal is distributed with frequency. The total power is defined as:

$$\text{Total Power} = \int_{-\infty}^{\infty} |X(\tau)|^2 \, d\tau. \tag{1}$$

Using Parseval's theorem (REFERENCE) (assuming that the signal has a finite total energy), the power of the signal can alternatively be written as

$$\int_{-\infty}^{\infty} |X(\tau)|^2 \, d\tau = \int_{-\infty}^{\infty} |\tilde{X}(\tau)|^2 \, df \tag{2}$$
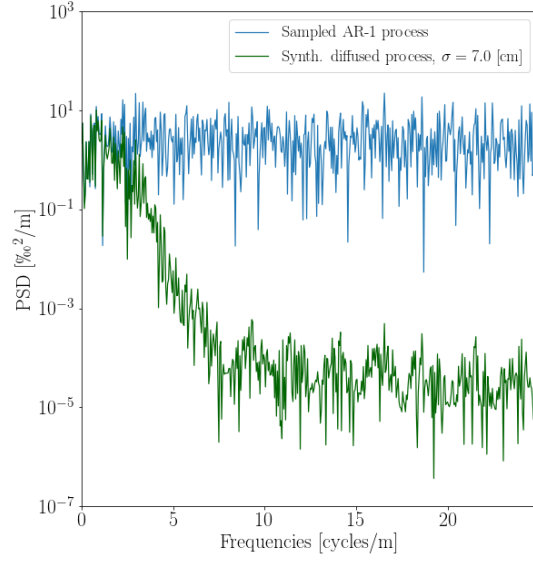
Figure 0.3

where $\tilde{X}(f)$ is the spectral (Fourier) transform of the signal, from time to frequency domain, defined as:

$$\tilde{X}(f) = \int_{-\infty}^{\infty} X(\tau)e^{2\pi i f \tau}\, d\tau \tag{3}$$

and the inverse spectral (Fourier) transform, from frequency to time domain, defined as:

$$X(t) = \int_{-\infty}^{\infty} \tilde{X}(f)e^{-2\pi i f \tau}\, df. \tag{4}$$

Both $X(t)$ and $\tilde{X}(f)$ represent the same function, just in different variable domains. Often, the angular frequency ω is used instead, with the relation between ω and f being $\omega \equiv 2\pi f$, giving the Fourier and inverse Fourier transforms as:

$$\begin{aligned}
\tilde{X}(\omega) &= \int_{-\infty}^{\infty} X(t)e^{i\omega\tau}\, d\tau \\
X(\tau) &= \int_{-\infty}^{\infty} \tilde{X}(\omega)e^{-i\omega\tau}\, d\omega
\end{aligned} \tag{5}$$

From Equation 2 we can interpret the integrand on the right hand side $|\tilde{X}(f)|^2$ as a density function, describing the energy per unit frequency. This is a property which is able to reveal much information about the considered signal, and it is useful to define this as the (one-sided) Power Spectral Density:

$$P_X(f) \equiv |\tilde{X}(f)|^2 + |\tilde{X}(-f)|^2 \qquad 0 \leq f < \infty \qquad (6)$$

This entity ensures that the total power is found just by integrating over $P_X(f)$ from 0 to ∞. When the function is purely real, the PSD reduces to $P_X(f) = 2|\tilde{X}(f)|^2$.

In the above the transform used to define the PSD was presented as the Fourier transform. When working with discrete data, as is very common when analyzing real world data, there are a number of different ways of estimating the PSD. In the following three different methods will be presented, all used in this thesis.

SIGNAL-
BACKDIFF: RE-
THINK THIS
PART. DO NOT
USE TIME ON ALL
THE CALCULA-
TIONS. WRITE
THE GENERAL
IDEAS OF THE
METHODS AND
STATE HOW TO
CALCULATE/-
COMPUTE. SMALL
CODE SNIP TO
GIVE GENERAL
IDEA.

RETHINK THIS PART. DO NOT USE TIME ON ALL THE CALCULATIONS. WRITE THE GENERAL IDEAS OF THE METHODS AND STATE HOW TO CALCULATE/COMPUTE. SMALL CODE SNIP TO GIVE GENERAL IDEA.

0.2.1.2 Spectral Transforms

- DFT/FFT

- NUDFT

- DCT

- NDCT

- (MEM)

0.2.2 Spectral Filtering

0.2.2.1 Wiener Filtering

Through spectral analysis it is possible to treat the noise of the signal consistently. The goal is to create spectral filters which enhances the signal while
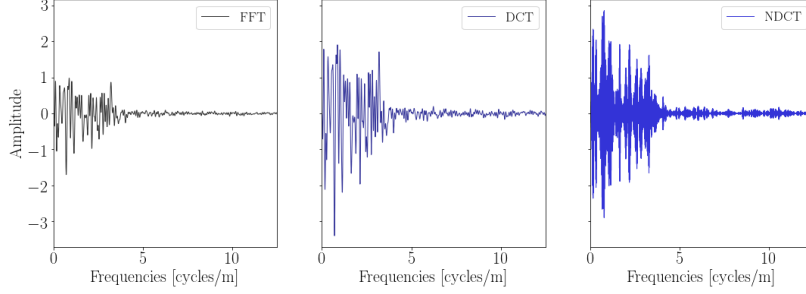
Figure 0.4: Examples of three different spectral transforms, FFT, DCT, NDCT, performed on the depth series between Tambora and Laki eruptions from Site A.
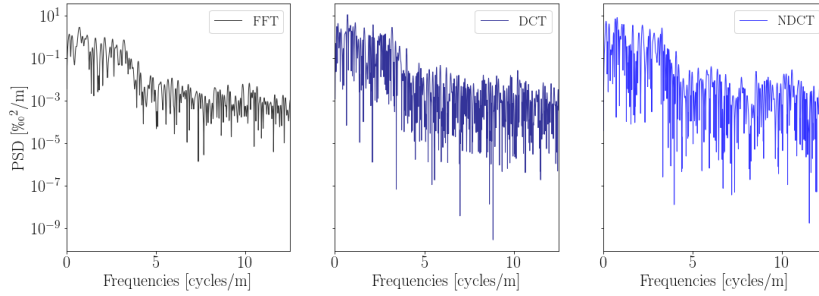

Figure 0.5: Examples of power spectral densities related to the three different spectral transforms, FFT, DCT, NDCT, seen in Figure 0.4.

minimizing the effect of the noise, thus increasing the signal-to-noise ratio (SNR).

Theoretically, without any diffusion, the change in isotopic concentration would be described through a step function, going from one constant concentration to another. This step function can be described by the Heaviside function:

$$D(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq \end{cases} \tag{7}$$

In reality, a number of different mixing processes change this step function, and the measured signal will be a smooth curve, $s(t)$, which corresponds to

the convolution of $S(t)$ with the mixing response function $M(\tau)$

$$d(t) = \int_{-\infty}^{\infty} D(\tau) \cdot M(t - \tau) d\tau \tag{8}$$

0.2.3   Signal Restoration by Optimal Diffusion Length

0.2.3.1   Kernel Estimation

As is well known, in the spectral domain, convolution is multiplication and the mixing is described as the multiplication between the Fourier transform of S and M:

$$\tilde{d} = \tilde{D} \cdot \tilde{M} \tag{9}$$

By differentiation with respect to time, the mixing filter M is unaffected, and differentiation of the measured system response, the Heaviside function, S' is a delta function, which Fourier transformed is unity, leading to:

$$\tilde{d}' = \tilde{D}' \cdot \tilde{M} = \tilde{M} \tag{10}$$

The mixing filter can thus be determined by measuring the system response to a step function, differentiating performing Fourier transform of the result d'.

After determination of the mixing filter $\tilde{M}$, the unmixed signal D can be estimated in theory by inverse Fourier transform of

$$\tilde{D} = \tilde{d} \cdot \tilde{M}^{-1} \tag{11}$$

During the mixing, cycles with short wavelengths are heavily washed out, and through the restoration in Eq. 11, the amplitudes corresponding to these wavelengths are heavily amplified by the filter. This method though has a drawback, which is that when the measurements contain noise, the restored signal will be dominated by high-frequency noise, greatly amplified by the mixing filter. Thus it is a problem of retaining as much (short wavelength) signal as possible while simultaneously attempting to amplify the high-frequency noise as little as possible. This optimal trade-off can be found by creating an optimum filter for the considered measured isotopic signal:

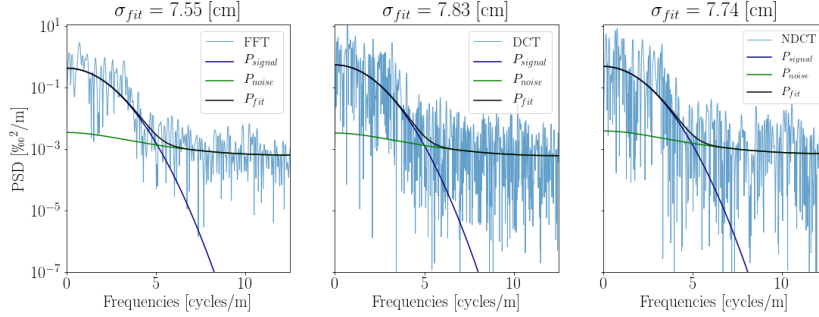$$\delta_M(z) = \delta_m(z) + \eta(z) \tag{12}$$

Figure 0.6: Noise, signal and total fit to PSD, illustrating the construction of the Wiener Filter, see Sec. 0.4.

This optimal (Wiener) filter $\tilde{F}$, defined for each wave number $k = 2\pi\omega$, is presented as the ratio between pure signal and pure signal plus noise described in Power Spectral Densities as:

$$\tilde{F}(k) = \frac{|\tilde{\delta_m}(\omega)|^2}{|\tilde{\delta_m}(\omega)|^2 + |\tilde{\eta}(\omega)|^2} \tag{13}$$

In this work, the power spectral densities of the signal and the noise, respectively, are determined through analysis of the power spectral density of the combined signal/noise PSD.

The PSD of the noise free measured signal, $|\tilde{\delta_m}(\omega)|^2$, is assumed describe as

$$|\tilde{\delta_m}(\omega)|^2 = P_0 e^{-k^2 \sigma_{\text{tot}}^2} \tag{14}$$

where σ_{tot}^2 describes the total estimated diffusion length of the mixing.

The noise is assumed to be red noise, described by an autoregressive process of first order, AR1:

$$|\tilde{\eta}(\omega)|^2 = \frac{\sigma_\eta^2 \Delta z}{|1 + a_1 \exp(-2\pi i \omega \Delta z)|^2} \tag{15}$$

where σ_η^2 is the variance of the red noise, a_1 is the AR1 coefficient and Δz is the resolution of the time/depth data. It is then possible to estimate the parameters P_0, σ_{tot}^2, σ_η^2 and a_1 by curve fitting, separately, the two expressions in Eq. 14 and 15 to the data. The estimated parameters are varied to find the optimal guess to use for the filter.

0.3 Estimation of Annual Layer Thickness

0.4 Enhanced Resolution and Restoration of Signal

0.4.1 Interpolation of Data

For the purpose of this thesis, interpolation of data needs to be fast, efficient and result in a function as smooth as possible. The last criterion is due to the knowledge of the nature of the data. The measurements are not continuous but should indeed in theory be so. Thus a good choice for interpolation of the data examined in this thesis would be the cubic spline interpolation. An instance of a such interpolation can be seen in Figure **??**.

Cubic spline interpolation has been used in two instances during this analysis, both times through the `Python SciPy` package `scipy.interpolate.CubicSpline`. Firstly, to assure equally spaced data points, so as to be able to perform a useful frequency analysis through spectral transformation, see Section, **??**. Secondly cubic spline interpolation was used to improve on peak detection in the final back diffused data. The final data have a rather low resolution, leading to an initial guess of peak positioning that might be shifted due to the discretization. Through cubic spline interpolation it is possible to construct a smooth estimate of a signal of higher resolution, leading to a peak positioning estimate that might be less shifted, see Figure **??**.

> SIGNAL-INTERP: REFERENCE!!

0.4.2 Detrending and Standardising

> SIGNAL-STANDARD: Think about if this is necessary. Maybe work into recursivity and constraints in Method?

0.4.3 Cycle Length Estimation of Detrended Signal

> SIGNAL-CYCLE: Think about if this is necessary. Maybe work into recursivity and constraints in Method?