



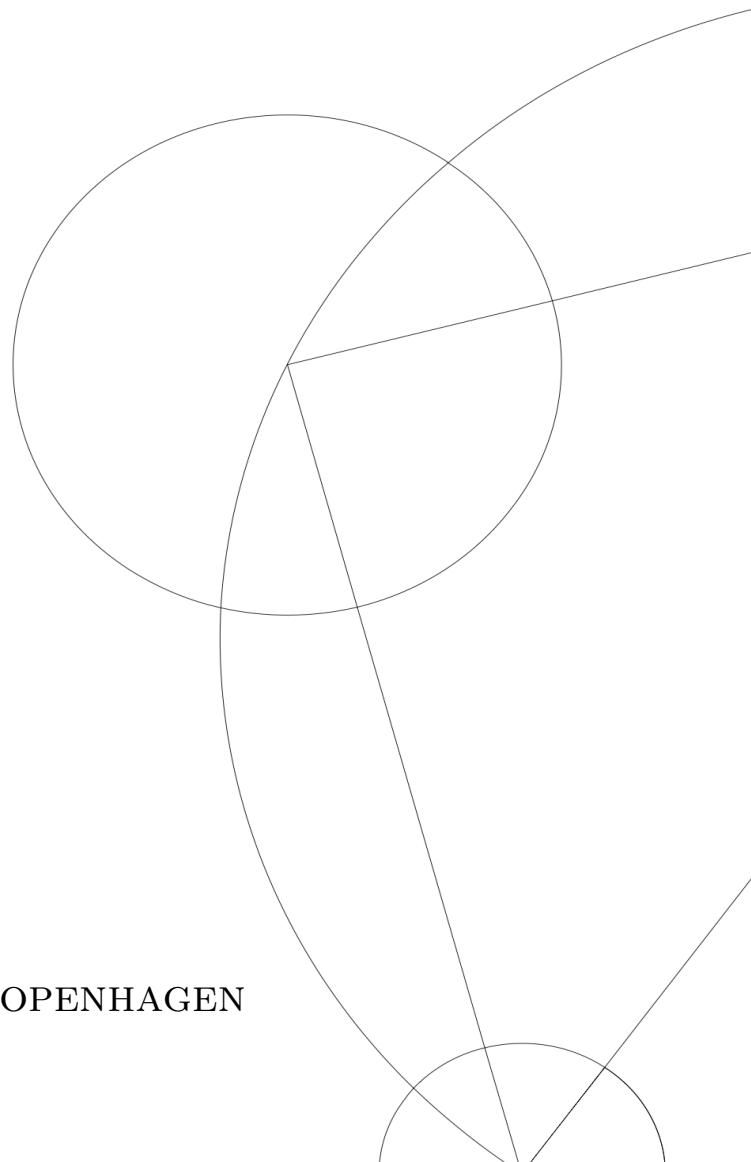
LAKI TO TAMBORA

Signal Restoration and Pattern Recognition in Ultra High Resolution Volcanic and Isotopic Signals

MASTER'S THESIS

Written by *Thea Quistgaard*
August 20th, 2021

Supervised by
Vasileios Gkinis and James Avery



UNIVERSITY OF COPENHAGEN



UNIVERSITY OF
COPENHAGEN

FACULTY: SCIENCE

INSTITUTE: Niels Bohr Institute

AUTHOR(S): Thea Quistgaard

EMAIL: Thea.Quistgaard@gmail.com

TITLE AND SUBTITLE: Laki to Tambora
- Signal Restoration and Pattern Recognition in Ultra
High Resolution Volcanic and Isotopic Signals

SUPERVISOR(S): Vasileios Gkinis and James Avery

HANDED IN: August 20th, 2021

DEFENDED: September, 2021

NAME _____

SIGNATURE _____

DATE _____

Darkness

I had a dream, which was not all a dream.
The bright sun was extinguish'd, and the stars
Did wander darkling in the eternal space,
Rayless, and pathless, and the icy earth
Swung blind and blackening in the moonless air;
Morn came and went—and came, and brought no day.

Lord Byron (1788–1824), July 1816



Figure 1: J. M. W. Turner: *The Eruption of the Soufriere Mountains in the Island of St Vincent, 30 April 1812*, 1815

Acknowledgments

Firstly, I would like to thank my supervisors Vasileios Gkinis and James Avery for guidance and never ending ideas and inputs. It has been an honour to work with so inspiring people, across of two different fields. Though I have not had the capacity to implement all ideas debated and juggled, I appreciated all the shared knowledge and curiosity, so apparent in both of you. In addition, I send all my thanks and appreciation to all the wonderful people situated at PICE. Through lockdowns and dark winters, through laughter and sun on the terrace, it has been a pleasure to be surrounded with that many sweet, engaged and caring people. A special thanks to Meg Harlan, Mikkel Rasmus Schmidt, Rebekka Frøystad and all the other students at PICE for making it feel like home. I would also like to send my thanks to my loving family, Hans, Bodil and Daniel, who has - through my entire life - been the wind in my back and the first ones to pop open a bottle of champagne. I could not have done this without you. A warmhearted thanks to Marcel Løvendahl Jørgensen, you will always be in my heart, thank you for all the support in the last many years. Furthermore, I send my appreciation to all my wonderful friends, Signe, Sofie, Josephine, Nadia, Simon, Mikkel, and many many more. So many people have set their mark on my studies at the Niels Bohr Institute, both personally and professionally, and it is a time, I will always treasure.

Abstract

A difficulty when studying data from ice cores, is the processes of densification and diffusion throughout the firn column. These processes attenuates some of the signal of the measured data, i.e. the water isotopic ratios or the chemical compositions of the ice. This work focuses on restoring the most likely signal in a given depth section of an ice core, through a method referred to as 'back-diffusion'. This method attempts to restore a signal by back diffusing it with an estimated diffusion length, σ . This diffusion length is affected by a number of different parameters, especially interesting by the temperature at deposition. So to give a qualified estimate on the temperature at a given time, it is of utmost importance to estimate the diffusion length as accurate and precise as possible. The restoration is achieved through a number of different computational methods, with a specific focus on a constrained optimization routine. This routine assumes a number of constraint over the depth series, especially the number of years expected in the section. This dating is made through Electrical Conductivity Measurements, where the two historically well documented volcanic eruptions of Laki (1783) and Tambora (1815) are visible. The method is tested on five different ice cores, from the Alphabet ice core series, all shallow cores drilled in the vicinity of the Crête ice core. The method has room for improvement, especially some of the simpler assumptions like the constraints and the optimization routine could benefit from further development. Moreover, the work carried out in this thesis leaves room for additional examination and development, but lays a good foundation to be used as a stepping stone in future research.

Contents

Darkness	i
Contents	iv
List of Figures	viii
List of Tables	x
1 Introduction	1
1.1 Ice Cores	1
1.2 A Rare Gem	2
1.3 Using the Rare Gem	4
1.4 Reading Guide	5
1.5 Software	6
2 Ice Theory	9
2.1 Water Isotopes	9
2.1.1 δ Notation	10
2.1.2 Relation to Temperature	11
2.2 Densification and Diffusion	11
2.2.1 Densification	11
2.2.1.1 HL in This Thesis	12
2.2.2 Diffusion	13
2.2.2.1 In Firn	14
2.2.2.2 In Solid Phase	16
2.2.3 Reconstruction of temperatures	17
2.3 CFM	19
2.3.1 Iso-CFM	20
2.3.1.1 Diffusion Length Profiles	21
2.4 Temperature Estimation	23
2.5 ECM and DEP	23

2.5.1	ECM	24
3	Data	27
3.1	Volcanic Horizons	27
3.1.1	Corrected Depth	29
3.1.2	Gaussian Distribution	29
3.2	Selection	30
3.2.1	AWI B-cores	31
3.2.2	Crete Area	31
3.2.2.1	Data Specifications	33
4	Signal Analysis & Comp. Meth.	37
4.1	Data as a Time Series	38
4.2	Spectral Analysis	39
4.2.1	PSD	39
4.2.1.1	Spectral Transforms	40
4.3	Back Diffusion	42
4.3.1	Signal Restoration	42
4.3.1.1	Spectral Filtering	45
4.3.1.2	Final Restoration	48
4.4	Peak Detection	49
4.5	Constraint: ALT from Spectral Analysis	51
4.5.1	Estimating λ_A	51
4.5.2	Change in l_{sec}	53
4.5.3	Change in l_{shift}	55
4.6	Splines and Interpolation	56
4.6.1	Interpolation	57
4.6.2	Interpolation in this Project	58
5	Method and Discussion	61
5.1	σ Estimation Method	62
5.1.1	Module 1: Initialization and Back Diffusion	63
5.1.2	Module 2	66
5.1.3	Constrained Peak Detection	68
5.2	Testing and Stability	70
5.2.1	Constraints or No Constraints	70
5.2.2	Effects of Interpolations	71
5.2.2.1	Interpolation 1	71
5.2.2.2	Interpolation 2	73
5.2.3	Spectral Transforms	75

5.2.4	LT locations	78
5.3	Upgrades	79
6	Results	85
6.1	Annual Layer Thickness	85
6.2	Diffusion Lengths	86
6.2.1	σ Constraints or No constraints	86
6.2.2	Diffusion Length V N Peaks	86
6.2.3	σ v. Spectral Transforms	87
6.2.4	σ Constant or Variable	88
6.2.5	Final σ Estimates	90
6.3	Temperature Estimates from Data	90
6.3.1	Iso-CFM Possibilities	91
7	Conclusion and Outlook	93
7.1	Conclusions	93
7.2	Outlook	95
Bibliography		97
APPENDICES		104
Appendices		105
Appendix II: AWI B-cores		106
Core Specifications		106
Appendix III: Alphabet Cores		110
Core Specifications		112
Appendix IV: Spectral Transforms		113
DFT & FFT		113
NUFT		115
DCT		116
NDCT		117
Appendix V: Splines and Interpolations		119
Existence and Uniqueness		119
Polynomial		120
Piecewise Polynomial		121
Appendix VII: Synthetic Diffusion		124
Appendix IIX: HL Empirical Model		125
Appendix IX: Iso-CFM		129
Diffusivity		129

CONTENTS

vii

Diffusion Length Profiles	131
Appendix X: DEP	134

List of Figures

1	<i>The Eruption of the Soufriere Mountains, Turner 1815</i>	i
1.1	Summers and Winters between Laki and Tambora	3
1.2	General idea flowchart	5
1.3	Map of Greenlandic ice core locations	5
2.2	Fractionation	10
2.1	10 m of Crête ice core with dating	10
2.3	Densification process	11
2.4	Five theoretical density profile examples	13
2.5	Site A density profile, data and theory	13
2.6	Diffusion	14
2.7	Rate terms in the diffusion equation	21
2.8	Analytical solutions of diffusion length, Alphabet cores	22
2.9	Five different analytical diffusion length profiles	23
2.10	Conductivity and $\delta^{18}\text{O}$ measurements from Site A	24
3.1	Summers and Winters between Laki and Tambora	28
3.2	ECM and $\delta^{18}\text{O}$ data between Laki and Tambora, Site A	28
3.3	Corrected depth positions of Laki and Tambora events	30
3.4	Corrected Tambora event, Site A	31
3.5	Corrected Laki event, Site A	31
3.6	ECM and $\delta^{18}\text{O}$ data between Laki and Tambora, Site B23	32
3.7	Map of Greenlandic ice core locations	32
3.8	Alphabet Cores Map	32
3.9	All Alphabet core $\delta^{18}\text{O}$ profiles in full	33
3.10	All Alphabet core $\delta^{18}\text{O}$ profiles, depths from Laki to Tambora	34
4.1	Flowchart of back diffusion method	43
4.2	FFT, DCT, NDCT, Site A	44
4.3	FFT, DCT, NDCT PSDs, Site A	44

4.4	Diffusion length profile, LT	45
4.5	FFT, DCT, NDCT PSDs with Fit, Site A	48
4.6	Wiener filter	48
4.7	Frequency filters example, Site A	49
4.8	10 m of Crête core with peaks and troughs	50
4.9	λ estimation example, Site B, in frequency domain	52
4.10	λ for full cores	53
4.11	λ depth profiles, Different l_{sec}	54
4.12	λ_{LT} vs. section length	55
4.13	λ depth profiles, different s_{sec}	56
4.14	$\lambda_{\text{LT}0}$ vs. s_{sec}	57
4.15	Even resampling of unevenly sampled data, Site A	58
4.16	Different resamplings of Site A	59
5.1	Diffusion length profile, LT	62
5.2	Flowchart of initialization method	64
5.3	Deconvolution with σ_{Theo} and σ_{Fit} , Site A	65
5.4	Deconvolution with σ_{Theo} and σ_{Fit} , Site B	65
5.5	Diffusion length profile, fit at LT	66
5.6	Flowchart of optimization method	67
5.7	σ vs. N peaks	68
5.8	Site B constrained vs. unconstrained	71
5.9	BD data, Site A, resamplings before deconvolution	72
5.10	σ vs. resampling size before deconvolution	74
5.11	BD data, Site A, resamplings after deconvolution	75
5.12	σ vs. resampling size after deconvolution	76
5.13	Spectral transform's effect on σ	76
5.14	Speed and σ estimates, spectral transforms	77
5.15	σ variations, varying Laki and Tambora	79
5.16	σ variations, varying only Laki	80
5.17	σ variations, varying only Tambora	80
5.18	2 month st. dev. variation of volcanic events locations	81
5.19	2 month variation of event locations, Site A	81
6.2	σ vs. N Peaks	88
6.1	σ vs. N Peaks	88
6.3	Deconvolution with σ_{constant} and $\sigma(z)$	88
6.4	σ , theoretical and optimal	90
6.5	Steady state temperature distributions	92

1	AWI B-cores density, theory and data.	108
2	AWI B-cores, theory and averaged data.	109
3	Alphabet cores, in full.	110
4	Alphabet cores, Laki to Tambora depth.	111
5	Monomials	121
6	Illustration of AR-1 process.	124
7	AR-1 process with diffusion.	124
8	AR-1 process in the spectral domain, raw and diffused.	125
9	Five different modelled density profiles.	128
10	Rate terms of the diffusion equation.	132
11	Modelled diffusion lengths of Alphabet cores.	134
12	Five different modelled diffusion length profiles.	135

List of Tables

3.1	Corrected Laki and Tambora positions	29
3.2	Core specifications for core drilled at Site A	35
5.1	Constraint parameters	70
5.2	Tested sample resolutions for resampling before back-diffusion . . .	73
6.1	Estimated λ at LT depths	85
6.2	Average sample size at LT	86
6.3	σ Estimates Given Unconstrained and Constrained Method	87
6.4	σ Estimates Given Spectral Transforms	89
6.5	Final σ Estimates	91
6.6	Steady state temperature estimates	91
7.1	Final σ Estimates	94
7.2	Steady State Temperature Estimates	95
1	Core specifications for AWI B-cores	106
2	Core specifications for core drilled at Site A.	112

Chapter 1

Introduction

The introduction to this work contains a motivation on why the research is of relevance, on what basis the general idea is based upon, and how the method is carried out. Along with this, it contains a brief walk through of the content of the entire thesis, and a short description of what software was developed, and where it is available.

Contents

1.1	Ice Cores	1
1.2	A Rare Gem	2
1.3	Using the Rare Gem	4
1.4	Reading Guide	5
1.5	Software	6

1.1 Ice Cores as a Window to the Past

The studies of ice cores have revealed much information and knowledge about the dynamics of the world's past climate, atmosphere and geology through measured proxies such as isotopic [13, W. Dansgaard, 1964], [28, S. Johnsen et al., 2001] and chemical compositions, conductivity [46, Moore et al., 1990], and dust measurements [34, F. Lambert et al., 2008]. By disclosing information about our past, the analysis of ice cores leads us to a greater understanding of the behaviors of the Earth system, which opens up for possibilities of modeling and predicting the future that lies ahead of us.

When analyzing ice cores it is most important that a relationship between depth and age is accurately established, as these timescales are of the essence when building empirical models and reconstructing paleoenvironments [8, E. Capron, 2013]. Dating of ice cores can be attempted through a variety of methods: visual inspection of annual layers in data, known volcanic events detected in the ice [67, B. M. Vinther et al., 2006] or radiocarbon dating [1, S. Aciego et al., 2011], to name a few.

A difficulty, that arises when dating ice cores, is the effects of diffusion through the firn column. Both gas and water molecules can diffuse through the firn which presents a number of obstacles for the continued dating. Firstly, the diffusion of gases in firn, through air pockets connected to the atmosphere, makes the age of the gas in the ice younger than the age of the firn at the same depth. Secondly, the diffusion of solid state molecules present in the firn, like water molecules, erases some of the signal, when measuring different properties of the ice. This erasure is commonly described through the average diffusion length of a molecule at a given depth, σ . This work focuses in particular on the densification and diffusion processes affecting water isotopic ratios in the firn.

The diffusion length σ is affected by a variety of parameters: the depth, the annual average accumulation, the ice flow and - especially interesting - the temperature [26, C. Holme et al., 2018]. By understanding which parameters influence the behavior of σ , it might be possible to use this signal erasure obstacle to gain more knowledge about paleoconditions: if it is possible to empirically estimate a diffusion length at a given depth, it may be achievable to reconstruct the temperature for this time interval.

The goal of this thesis is to establish a method for estimating the diffusion length for a given depth section. This is achieved through different analyses, and through a hidden gem in the ice cores.

1.2 A Rare Gem of Knowledge

In June 1783 the Icelandic volcano Laki erupted explosively [60, J. Steingrímsson, 1998]. The eruption led to an eight-month long emission of volcanic aerosols into the European airspace, bringing climatic and sociological disruptions with it [62, R. Stone, 2004], [64, T. Thordaldson and S. Self, 2004]. Although a cataclysmic catastrophe for much of Europe, with extreme weather, famine

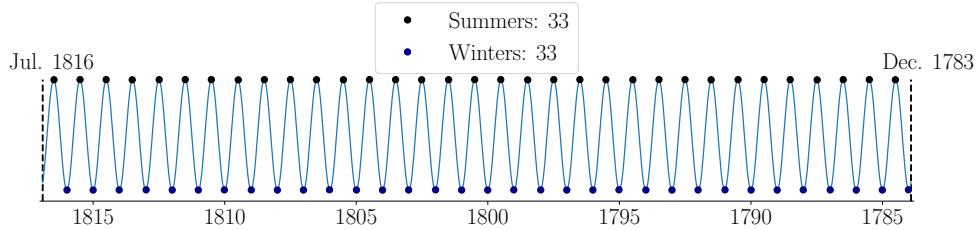


Figure 1.1: Visualization of pattern of summers and winters in the time span between the Laki and Tambora volcanic depositions in Greenland.

and higher death rates, the violence of the impact on the life of Europeans, led to this eruption being very well documented and recorded across most of Europe.

Later, in April of 1815 on the Southern Hemisphere, the eruption of the Indonesian volcano Tambora resulted in a series of events just as, or even more fatal than, the Laki eruption [49, C. Oppenheimer, 2003]. Tens of or even hundreds of thousands died either during the eruption or in direct consequences hereof, from starvation or epidemic diseases. Furthermore, the eruption, following a number of decades with heavy volcanic activity [11, J. Cole-Dai et al., 2009], left its mark on the climate of the entire Earth system, disrupting global temperatures. In Europe, the following year of 1816 became known as *The Year Without a Summer* [49], and the apocalyptic climate affected not only the crops and human necessities, but also the artistic and sociological environment of Europe. Writers like Lord Byron (*Darkness* 1816, 6) and Mary W. Shelley (*Frankenstein* 1815-1818) became inspired by the cold and dark weather [39, A. Marshall, 2020], and painter J. M. W. Turner was clearly influenced by the change of colour of the world, to a more yellow, brown and gloomy ambiance of the 1816 European summer, for example in the painting *The Eruption of the Soufriere Mountains in the Island of St Vincent, 30 April 1812*, 1815, see Figure 1 [73, C. Zerefos, 2007].

Both volcanic events were not only landmarks in European history, but quite literally also left their marks on the Greenlandic ice sheet, by deposition of volcanic material through precipitation. These layers of snow with extra high content of volcanic material shows to be a rare gem of knowledge, hidden in the ice. When measuring ice cores this can be utilized along with a different property of the data signals available through ice core analysis, namely the seasonality of a given parameter.

Some measured signals in the ice cores contain annual cycles. For example, the water isotopic ratios in the firn are sensitive to temperature [31, J. Jouzel, 1997], leading to a clear summer-to-winter cycle. This makes it relatively easy to date shallow ice cores as the cycles can be counted, but as diffusion takes place in the firn column, some of this signal is washed away. Luckily, another method can be utilized to date the ice: detection of known volcanic events through electrical conductivity measurements. This reveals a quite unique gem of knowledge: by knowing the time of a certain volcanic event, either through historical observations or through previous ice core synchronization, and matching this with the depth of the detected event in the ice, it is possible to set some very certain dates on the timescale of the ice.

An example of this type of event dating, which is used in this work, is by examining the volcanic eruptions of Laki and Tambora. Both eruptions are, as described, very well historically documented and are visible and detectable in a great number of ice cores [?, H. Clausen, 1988], [35, C. Langway, 1988]. The deposition in Greenlandic ice cores has been estimated to be in December 1783 for Laki and in July 1816 for Tambora, yielding 33 summers and winters between the two events, see Figure 1.1 [11, J. Cole-Dai et al., 2009], [69, L. Wei, 2008]. This does not only make it possible to generally date and synchronize different ice cores, but it also allows for in depth analysis of the diffusion and densification processes in the ice.

1.3 Utilizing the Rare Gem

By considering an isotopic depth series situated between two volcanic events, it is possible to back diffuse this series over the known time span in years - or even months - using the diffusion length as a tuning parameter. This is an optimal way to empirically estimate the diffusion length of a given depth interval which makes it possible to obtain a temperature reconstruction of this interval, as σ is temperature sensitive [26].

The goal is thus to reconstruct the lost signal by a back diffusion scheme, tuning σ of the diffusion process, until the known actual number of winters/summers between the events can be counted as peaks and valleys in the depth signal. Then the estimated optimal diffusion length can be used to make a temperature estimate of the given interval. The back diffusion method is built on both empirical models and signal analysis of the measured data. A simplified flowchart of the general idea is illustrated in Figure 1.2.

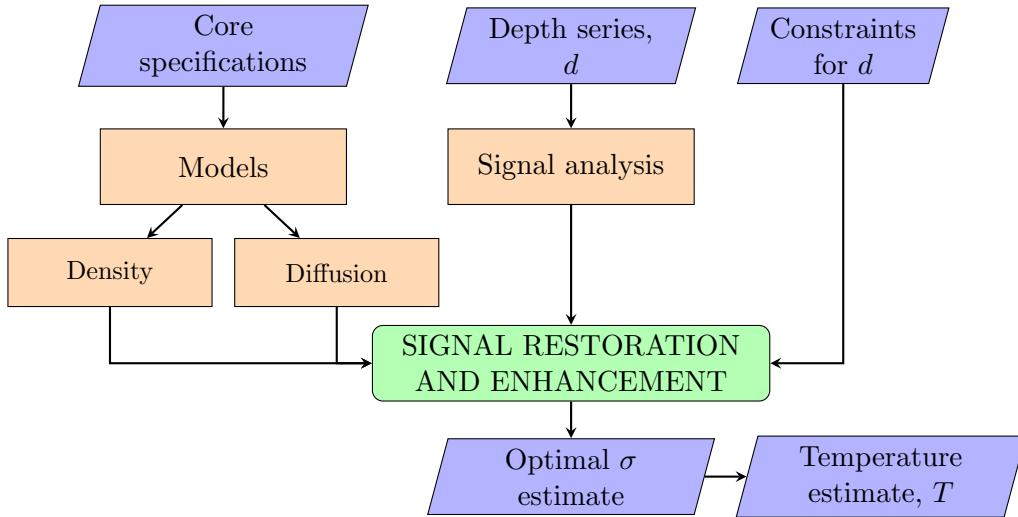


Figure 1.2: Illustration of the general idea, restoring signal by tuning the diffusion length until the expected number of years is detectable in the depth interval.

The data under consideration in this thesis is mostly shallow ice cores, namely the Alphabet cores drilled in the vicinity of the 405 m Crête ice core [?, H. Clausen and C. Hammer, 1988].

The spatial locations of Crête, the alphabet cores and three other major ice cores, the NEEM [20], NGRIP [48] and EGRIP [52] cores, can be seen in Figure 1.3.

1.4 Reading Guide

In this thesis, Chapter 2 contains an introduction to the theory of diffusion of water isotopes in ice cores along with theoretical and empirical methods for modeling densification and diffusion profiles. Following, still in Chapter 2, is a brief examination of different experimental methods for detection of deposited volcanic material and which methods have been used for the data under inspection. The chosen data are then presented in Chapter 3, along with an argumentation of why they were selected. Then a thorough presentation of the data and signal analysis along with important computational methods are presented in Chapter 4. These different tools are then combined in the method description in Chapter 5, depicting a walk-through and testing of the

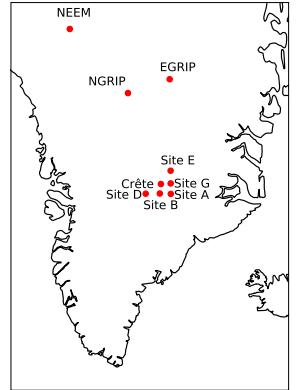


Figure 1.3: Location of Alphabet cores along with some major ice cores, NEEM, EGRIP and NGRIP.

final algorithm developed for estimating the diffusion length given the specific number of years. The final method is tested, and further discussed, developed and fine tuned, still in Chapter 5, and results from the final iteration of the method are presented along with a statistical analysis of variations in the final estimates in Chapter 6. On the basis of these results, finally, a basic temperature reconstruction of the examined depth intervals for the ice cores is presented. Finally, a walk through of the most important conclusions and an outlook to future research is given in Chapter 7.

Chapter 2 contains already existing material, as it is a walk through of the theory of ice cores developed throughout the last century. Chapter 3 contains, as mentioned, a description of data, which is per se not new material, as the ice cores in focus were drilled and analyzed in the 1970's and 1980's, but some small corrections were made to the estimates of the locations of the volcanic events. Chapter 4 contains a presentation of the existing computational methods utilized, but also a walk through of how the choices of methods and parameters affect the final method. In Chapter 5 the newly developed back diffusion method is presented along with a discussion of various subjects in the thesis.

1.5 Software

All computational analysis carried out in this work is implemented through Python v3.8.10. All code is available on GitHub repository by T. Quistgaard, Master's Thesis, (2021), GitHub repository: https://github.com/TheaQG/AWI_Bcores_Analysis. The main modules implemented are:

- Herron-Langway densification model, Section 2.2.1.1, in file `HL_AnalyticThea_class.py`.
- Diffusion length profile model, Section 2.3, in files `DiffusionProfiles_calculations.py` and `Diffusivity.py`.
- Interpolation methods, Sections 4.6 and 5.2.2, in file `Interpolation_Class.py`.
- Signal attenuation and annual layer thickness estimation, Section , in file `SignalAttenuation.py`.
- Spectral transforms and analysis, along with general deconvolution/back diffusion method, Sections 4.2, 4.3 and 5.1.1, in file `Decon.py`.

- The final optimization module, along with the constrained peak detection method, Sections 5.1.2 and 5.1.3, in file `BackDiffuse_LT.py`.
- Final temperature estimates, Chapter 6, in files `sigmaSolver.py` and `TemperatureEstimates.py`.

The modules are all described in depth in the corresponding sections referred to, and the connections between modules are illustrated in Figures 5.2 and 5.6. Along with the presented files are a number of files containing code for testing the different modules and for generating the final results.

Chapter 2

The Theory of Ice Cores

The ice core theory on which the methods and analyses of this thesis is built will be presented in this chapter. This includes the basis of water isotopic ratios, densification and diffusion in ice, temperature estimation from diffusion lengths, and electrical conductivity measurements.

Contents

2.1	Water Isotopes	9
2.1.1	δ Notation	10
2.1.2	Relation to Temperature	11
2.2	Densification and Diffusion	11
2.2.1	Densification	11
2.2.2	Diffusion	13
2.2.3	Reconstruction of temperatures	17
2.3	CFM	19
2.3.1	Iso-CFM	20
2.4	Temperature Estimation	23
2.5	ECM and DEP	23
2.5.1	ECM	24

2.1 Water Isotopes

A corner stone in ice core analysis, which helps lay the basis for paleo climate research, is the measurements of isotopic composition of the water which

makes up the ice or of the encapsulated air in bubbles throughout the ice [19, V. Gkinis et al., 2011]. Water isotopes are sensitive to temperature changes and can thus be used as a proxy for paleo temperature and climate along with being used as dating parameters, since the annual cycles often are detectable in water isotope data, as is clearly the case with the water isotopic ratios depicted in Figure 2.1.

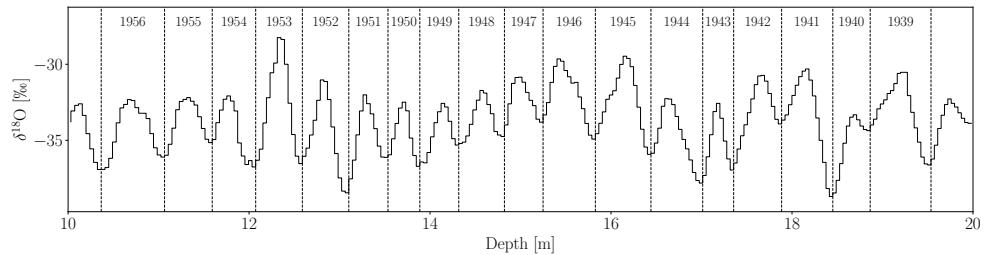


Figure 2.1: Ten meters of the top of Cret\acute{e} ice core, with identification and dating of 19 annual layers.

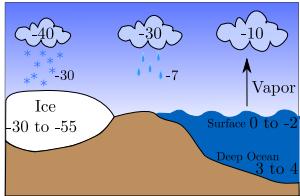


Figure 2.2: Fractionation through evaporation, transportation and precipitation along with typical water isotopic ratios in [\%].

2.1.1 δ Notation

Water isotopic ratios, i.e. the ratio of the minority isotope, H_2^{18}O or H_2^{17}O ($^2\text{H}_2\text{O}$), compared to the majority isotope, H_2^{16}O ($^1\text{H}_2\text{O}$), are used to report the quantities of isotopes in a sample relative to the ratio of a given reference water sample. This is commonly expressed in the δ -notation as:

$$\delta^i = \frac{iR_{sample}}{iR_{reference}} - 1 \quad (2.1)$$

where $^{18}R = \frac{n_{^{18}\text{O}}}{n_{^{16}\text{O}}}$, $^2R = \frac{n_{^2\text{H}}}{n_{^1\text{H}}}$ and $^{17}R = \frac{n_{^{17}\text{O}}}{n_{^{16}\text{O}}}$. Here n is the abundance of the given isotope.

Besides the isotopic quantities $\delta^{17}\text{O}$, $\delta^{18}\text{O}$ and $\delta^2\text{H} = \delta\text{D}$, both deuterium excess and $\Delta^{17}\text{O}$, known as ^{17}O excess, can be of interest [33, K. Lamb et al., 2017]. Deuterium excess is often used as a measure of the kinetic fractionation processes, taking place in the water vapor formation of polar precipitation, to which it is especially sensitive, giving an indicator of the conditions during precipitation formation, and thus giving a pointer to the source of the water vapor.

The naturally occurring oxygen isotopes are approximately divided into 99.759% ^{16}O , 0.037% ^{17}O and 0.204 % ^{18}O [3], and historically the focus

has been on the water isotopic ratio of $\delta^{18}\text{O}$, due to the higher abundance making it easier to measure than for example $\delta^{17}\text{O}$. One of the challenges of modern glaciology is to develop a stable, precise and accurate measuring technique for $\delta^{17}\text{O}$. This work examines the data from ice cores drilled and measured in the past century, and only $\delta^{18}\text{O}$ has been measured for these cores.

2.1.2 Water Isotopes in the Earth System

Due to H_2^{18}O being slightly heavier than H_2^{16}O , the isotopic signal is depleted of the heavier isotopes through transportation of the water vapor before it is deposited at the ice sheets. A clear annual signal in the $\delta^{18}\text{O}$ measurements can be observed due to the seasonal temperature changes from summer to winter at deposition time. This is the main isotopic feature under consideration for this thesis. The greatest issue with this type of signal is that there is a loss of information over time as the snow becomes more compact and diffusion processes wash out some of the signal throughout the ice depth. Thus to be able to examine and analyze isotopic signals more accurate and precisely, the processes of densification and diffusion need to be well understood.

2.2 Densification and Diffusion

Deposited snow will over time be compressed and at last be compacted completely to glacial ice, see Figure 2.3. Any ice state between snow and glacial ice is referred to as *firn*. Throughout the firn column the important processes of densification and diffusion take place. Both processes need to be well understood and examined when analyzing ice core data, as diffusion and densification play a large role in thinning of annual layers due to compression of snow to ice and in washing out the measured signals through diffusion in the firn. Specifically in this thesis the diffusion processes are carefully examined, as the diffusion length, σ can be used as a proxy for paleotemperatures [26], [32].

2.2.1 Densification

Densification is the process of compression of snow to ice. It affects the annual layer thickness in the data as snow will be compacted to a smaller volume under pressure from the firn column above until it reaches a solid ice state with an, almost, constant density.

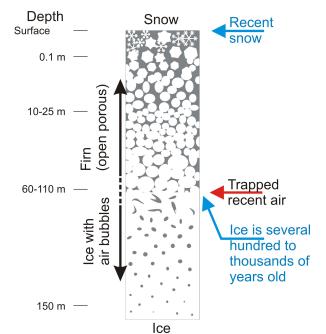


Figure 2.3: Illustration of the densification process of a firn column, from snow deposition to glacial ice. Image source [6].

Commonly three stages of densification are described in the firn column [25, M. Herron and C. Langway, 1980]. The first stage is between the initial precipitated snow density and the 'critical density' at $0.55 \frac{\text{Mg}}{\text{m}^3}$, the second stage is between critical density and the close-off density at $0.82 - 0.84 \frac{\text{Mg}}{\text{m}^3}$, and the third stage is from close-off and all the way through the glacial ice. The close-off depth refers to the depth, where all air pores in the firn column become closed-off from the atmosphere at the surface, and separate bubbles start to form in the ice.

At the first stage the densification is mostly due to grain settling and packing and the densification rate is very rapid. At the second stage, the snow is close to isolating air bubbles. At the third stage, the dominating densification taking place is by the compression of air bubbles [25]. For these three stages it has early been of interest to develop a model to determine the depth-density profile, which is dependent on snow accumulation rate and temperature. The focus was on developing an empirical model for the first and second stages of densification, as they are the most dramatic sections of the firn column considering densification and diffusion.

A number of different densification models have been developed [5, J. Barnola, 1991], [22, C. Goujon, 2003], but for this thesis work, only one model is used, namely the Herron-Langway empirical model, first presented by Michael M. Herron and Chester C. Langway Jr in [25, Herron and Langway, 1980]. The basic idea of the model will be presented in the following. For a more thorough mathematical presentation, the reader is referred to Appendix 7 or the original paper.

2.2.1.1 Herron-Langway Model in This Thesis

The Herron-Langway model (from now on *HL model*), is an attempt to describe the densification process through a firn column, given some initial conditions, mainly the temperature and annual accumulation rate at the site. The model is generally derived from the suggestion that during the process of densification, the proportional change in airspace is linearly related to the change in stress due to the overburdened snow:

$$\frac{d\rho}{\rho_{\text{ice}} - \rho} = \text{const.} \rho dh. \quad (2.2)$$

with ρ being the density of the firn, ρ_{ice} the density of glacial ice and h the depth.

This implies, by integration, a linear relationship between $\left[\frac{\rho}{\rho_{ice} - \rho} \right]$ and h . Through empirical analysis of depth-density measurements from a number of different ice cores, and the basic acceptance of Equation 2.2, the rate equations $\left(\frac{d\rho}{dt} \right)_{zone\ 1}$ and $\left(\frac{d\rho}{dt} \right)_{zone\ 2}$ can be derived, see Appendix 7. From these rate equations, and given temperature, T , accumulation rate, A , and initial snow density, ρ_0 , it is then possible to estimate the density and age of the snow at any given depth.

In this thesis, the HL model is used to generate an estimate of the density profile for each ice core site under examination. The HL model has been incorporated in the computational works of this thesis as a class module, which computes a diffusion length profile, given initial parameters. An example of five different density profiles computed with this module, given present day conditions at five different ice core drilling sites as initial conditions can be seen in Figure 2.4. These profiles are then used to give a first diffusion length profile estimate for each core, which is used in the further analysis.

For the purpose of this thesis, an extra module has been added to the HL model class, which makes it possible to incorporate any measured data in the density profile estimation, so as to best approach the actual profile at the drilling site. This module is implemented by allowing an intake of two arrays consisting of depth and density measurements, which are then used to fit the model, using a least square optimization from `scipy`, `scipy.optimize.leastsq`. An example of a density profile estimation, both using the measured data and using only the model, can be seen in Figure 2.5.

2.2.2 Diffusion

In general diffusion refers to the mixing of material, e.g. molecules, by net movement from high to low concentration due to somewhat random stochastic processes, see Figure 2.6. When considering ice cores, the diffusion taking place is the displacement of the molecules that make up the ice and firn, for example water molecules. The diffusion processes are subject to varying conditions and are caused by different effects at different depths. In the first few meters of an ice core, the main diffusion is due to mixing and settling of snow particles due to surface conditions and gravity, and later, the total diffusion is more subject isotopic gradients in the ice [29, S. Johnsen et al., 2000].

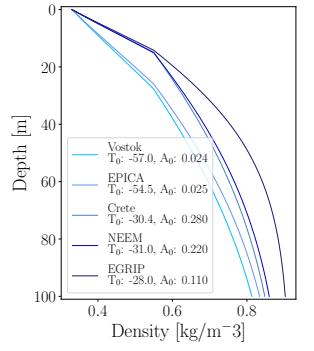


Figure 2.4: Density profile examples given five different initial conditions representing present day conditions at the five different ice core locations. Temperature, T_0 , is in $^{\circ}\text{C}$ and accumulation, A_0 , is in meter of water equivalent per year.

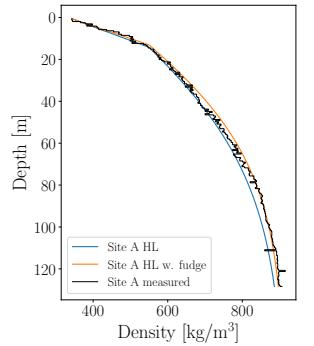


Figure 2.5: Density profiles from ice core Site A near Crête. Both purely modelled profile and profile fitted to the inputted depth-density measurements are presented.

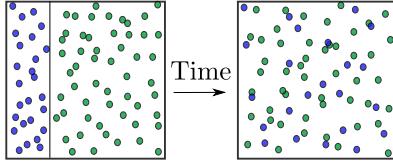


Figure 2.6: Example of diffusion over time, where two materials are mixed.

The diffusion processes are different for the porous firn and for the later glacial ice, described in the different stages above and below close-off depth.

2.2.2.1 In Firn

The diffusion processes in firn can be used to describe the attenuation of a given signal, e.g. a water isotopic signal, due to vapor phase diffusion in the porous firn column. This vapor phase process takes place in the air pockets of the material from time of deposition to pore close-off [29].

To develop accurate knowledge of paleo climate and temperatures it is of great importance to understand this process, as a reconstruction of the part of the signal lost will reveal finer details in the signal and thus a more detailed knowledge of past times.

Diffusion is modeled through Fick's 2nd law [15, A. Fick, 1851], which describes the change in concentration of a substance with time, due to diffusion:

$$\frac{\partial \phi}{\partial t} = D_{\text{dif}}(t) \frac{\partial^2 \phi}{\partial z^2} - \dot{\epsilon}_z(t) z \frac{\partial \phi}{\partial z} \quad (2.3)$$

where ϕ is the concentration of the substance, t is time, D_{dif} is the diffusion coefficient, z is the position, and $\dot{\epsilon}_z$ is the strain rate. Focusing on water isotopic ratios, the water isotopic signal can be assumed as a concentration, $\phi \rightarrow \delta$, resulting in a diffusion equation as:

$$\frac{\partial \delta}{\partial t} = D_{\text{dif}}(t) \frac{\partial^2 \delta}{\partial z^2} - \dot{\epsilon}_z(t) z \frac{\partial \delta}{\partial z} \quad (2.4)$$

Through depth and time driven attenuation due to diffusion there is a loss of information which can be used to gain information of the underlying processes. Since the diffusion constant and the vertical strain rate $\dot{\epsilon}_z(t)$ in Fick's 2nd law are dependent on temperature and accumulation on site, the

information loss process can infer the temperature of firn and accumulation on site.

The attenuated, directly measured, isotopic signal, $\delta(z)$, can be described as the convolution between the initial isotopic signal, $\delta'(z)$, and a Gaussian filter, $\mathcal{G}(z)$. The Gaussian function describes the stochastic diffusion process. The signal is furthermore multiplied by the thinning function, $S(z)$, which describes the total thinning of a given layer at depth z due to the vertical strain from the above firn column.:

$$\delta(z) = S(z)[\delta'(z) * \mathcal{G}(z)] \quad (2.5)$$

where

$$S(z) = e^{\int_0^z \dot{\epsilon}_z(z') dz'} \quad (2.6)$$

and

$$\mathcal{G}(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2\sigma^2}} \quad (2.7)$$

In the Gaussian filter, the variance σ is referred to commonly as the diffusion length: the distance a water molecule is displaced along the z-axis. This quantity is directly related to both $D_{\text{dif}}(t)$ and $\dot{\epsilon}_z(t)$ (the strain rate being approximately proportional to the densification rate in the column). Thus an accurate estimate of the diffusion length is crucial for describing the diffusion process. The change of diffusion length over time is given as

$$\frac{d\sigma^2}{dt} - 2\dot{\epsilon}_z(t)\sigma^2 = 2D_{\text{dif}}(t) \quad (2.8)$$

by [27, Johnsen, 1977], which also states that in the case of firn and assuming a site with little ice flow, the vertical strain rate can be approximated with a simple strain rate, only dependent on the density and its time evolution

$$\dot{\epsilon}_z(t) \approx -\frac{d\rho}{dt} \frac{1}{\rho}, \quad (2.9)$$

where ρ is the density and $\frac{d\rho}{dt}$ is the densification rate. With this approximation, the solution to Eq. 2.4 describing evolution of the diffusion length in the firn column, can be found, defined only through density and densification rates, as:

$$\sigma^2(\rho) = \frac{1}{\rho^2} \int_{\rho_0}^{\rho} 2\rho'^2 \left(\frac{d\rho'}{dt} \right)^{-1} D_{\text{dif}}(\rho') d\rho'. \quad (2.10)$$

Certain densities and corresponding depths are of special interest as they indicate a specific stage of the firn and ice column [25]. At top and bottom, we find the two extremum densities of settled snow, $\rho_{\text{snow}} = 330 \frac{\text{kg}}{\text{m}^3}$, and ice, $\rho_{\text{ice}} = 917 \frac{\text{kg}}{\text{m}^3}$. In between these two there are two more densities of importance: the critical density, $\rho_{\text{Cr}} = 550 \frac{\text{kg}}{\text{m}^3}$, describing the transition between the two firn stages (see Section 2.2.1), and the pore close off density, $\rho_{\text{co}} = 330 \frac{\text{kg}}{\text{m}^3}$, describing the density at which air pockets in firn will seal off from each other to form single bubbles. As mentioned, from the close-off density, further densification will be due to compression of these closed off air bubbles until the density reaches ρ_{ice} . If we assume that the diffusion constant, $D(\rho)$, and the densification rate, $\frac{d\rho}{dt}$ are known, then it is possible to give an estimate of the diffusion length profile by integrating from top, at density ρ_0 , to pore close-off depth, ρ_{co} .

2.2.2.2 In Solid Phase

When firn reaches the solid state of ρ_{ice} below close-off depth, the isotope diffusion is driven not as much by densification any more, but by isotopic gradients within the ice crystal lattice structure. This diffusion process is much slower than the diffusion in vapor phase taking place in firn, and thus does not contribute as much to the information loss and attenuation of the signal. For solid ice, at $\rho \geq \rho_{\text{ice}}$, the diffusion constant is only dependent on temperature, and can be described through an Arrhenius type equation as [54, Ramseier, 1967], [30, Johnsen et al., 2000]:

$$D_{\text{ice}} = 9.2 \cdot 10^{-4} e^{-\frac{7186}{T}} \left[\frac{\text{m}^2}{\text{s}} \right] \quad (2.11)$$

The diffusion length in solid state ice is then given from the diffusion constant in ice and the thinning function as:

$$\sigma_{\text{ice}}^2(t) = S(t)^2 \int_0^t 2D_{\text{ice}}(t')S(t')^{-2} dt' \quad (2.12)$$

From Equations (2.10) and (2.12) it is possible to model a diffusion length profile given surface conditions from a ice core drill site. This profile along with the depth-density profile will be the theoretical grounds for much of the empirical work carried out in this thesis.

2.2.3 Reconstruction of temperatures

Reconstruction of paleotemperatures can be attempted through a number of various techniques [31], [26], [16]. For this work, the focus is on restoring a signal by single isotopic back-diffusion.

By the deconvolution theorem [41], convolution in time domain is equal to multiplication in the frequency domain. According to Equation (2.5), the transfer function describing the diffusion in the frequency domain, will be the Fourier transform of the Gaussian filter:

$$\mathcal{F}[\mathcal{G}(z)] = \hat{\mathcal{G}}(f) = e^{-\frac{k^2\sigma^2}{2}}, \quad k = 2\pi f = \frac{2\pi}{\Delta_z} \quad (2.13)$$

where Δ_z is the discrete sampling size. This filter, of course depending on Δ_z , keeps larger wavelength frequencies (> 50 cm) unaltered but attenuates short wavelengths (< 20 cm) heavily, which is exactly the effect of diffusion on the isotopic signal. In the Gaussian filter σ represents the diffusion length.

An estimate of the diffusion length σ can be made from the power spectral density (PSD) of an isotopic time series. In the frequency domain a PSD composed of an initial signal, a filter function and a white noise term can be described as:

$$P_s = P_0(k)e^{-k^2\sigma^2} + |\hat{\eta}(k)|^2, \quad f \in [0, f_{Nq}] \quad (2.14)$$

where the diffused and noise-affected signal, P_s , is equal to the original signal, $P_0(k)$, times a filter, $e^{-k^2\sigma^2}$ (our previously inspected Gaussian filter), plus a noise term, $|\hat{\eta}(k)|^2$, over a frequency space ranging from zero to the Nyquist frequency, f_{Nq} [58, C. Shannon, 1949]. The Nyquist frequency is dependent on the sampling resolution by $f_{Nq} = \frac{1}{2\Delta_z}$.

The noise term, often categorized as white noise, but red noise is also seen in isotopic signals [26], is given as

$$|\hat{\eta}(k)|^2 = \frac{\sigma_n^2 \Delta_z}{|1 - a_1 e^{ik\Delta_z}|^2} \quad (2.15)$$

Equation 2.15 describes an autoregressive process of the first order (white noise), with a_1 being an AR-1 coefficient. An AR-n process describes the evolution of a stochastic time series \mathbb{X} where the next time step, X_t is dependent on the last n points, $\{X_{t-n}, \dots, X_{t-1}\}$, and is mathematically defined

as:

$$X_t^{(n)} = C + \sum_{i=1}^n \phi_i X_{t-i} + \epsilon_t \quad (2.16)$$

where C is a constant, $\bar{\phi} = \{\phi_1, \dots, \phi_n\}$ are the model parameters and ϵ_t is the noise added to the given time step. The AR-1 process thus describes a series where each new point is only dependent on the last point before:

$$X_t^{(1)} = C + \phi_1 X_{t-1} + \epsilon_t. \quad (2.17)$$

and the power spectral density of the AR-1 process is, corresponding to Eq. 2.15:

$$S^{(1)}(f) = \frac{\sigma_z^2}{|1 - \phi_1 e^{-2\phi_1 f}|^2} \quad (2.18)$$

The spectral estimate of the time series, \mathbb{P}_s , can be computed via a number of different numerical schemes, in this work the focus is on Fourier and Cosine transforms.

To give a data based estimate on the diffusion length a fit to these estimated spectral data, P_s , is found through for example a least square optimization, from which the parameters P_0 , σ , a_1 , σ_η^2 can be estimated and the diffusion length σ can be calculated by least-square minimization of the misfit between \mathbb{P}_s and P_s , see Section 4.3.

This estimated diffusion length needs to be corrected: the obtained $\hat{\sigma}$ is affected by two further diffusion processes, taking place respectively in the ice and in the experimental sampling:

- **Sampling diffusion:** This diffusion is due to the sampling method. Sampling at a certain discrete resolution - be it discrete sections or resolution in CFA system due to step or impulse response - gives an additional diffusion length of

$$\sigma_{dis} = \frac{2\Delta_z^2}{\pi^2} \ln\left(\frac{\pi}{2}\right) \quad (2.19)$$

- **Ice diffusion** When below the close-off depth, a correction for the ice diffusion must also be made.

So to obtain the actual diffusion length estimate from the raw data, both the sampling and the ice diffusion need to be subtracted from $\hat{\sigma}^2$, and a scaling factor due to thinning from the strain must be introduced:

$$\sigma_{\text{firn}}^2 = \frac{1}{S(z)^2} \hat{\sigma}_{\text{firn}}^2 = \frac{\hat{\sigma}^2 - \sigma_{\text{dis}}^2 - \sigma_{\text{ice}}^2}{S(z)^2} \quad (2.20)$$

From Equation (2.20) it is clear that the accuracy of the diffusion length estimate σ_{firn}^2 is dependent on the correction terms σ_{dis}^2 , σ_{ice}^2 and the thinning function $S(z)$. The correction term related to the discrete sampling method used in isotope analysis, σ_{dis}^2 is generally a well managed parameter, as the discretization of the measurements is known.

At relatively shallow depths, as the ones under examination in this thesis, the term σ_{ice}^2 is relatively small compared to the total diffusion length estimate $\hat{\sigma}^2$ and can be either neglected or easily accounted for with simple assumptions on ice flow and borehole temperature. For this thesis, the ice diffusion is neglected at the depths which are operated on.

The final correction parameter, the thinning function $S(z)^2$, has a strong influence on the final diffusion length estimate, and errors from the ice flow modelling will be propagated to the diffusion length estimate, and finally to the temperature estimate.

2.3 The Community Firn Model

The Community Firn Model (CFM) is an open-source modular Python framework for firn-modelling. It was first developed by [61, Stevens et al, 2020] and later adapted to a different version with a focus on water isotopic diffusion by [17, Gkinis et al, 2019] under the name Iso-CFM, the latter version used in this thesis.

The original CFM is modular, meaning that, firstly, it is easy to choose which physical processes should be included in the modelling, and, secondly, it allows for the user to develop new modules that can easily be integrated in the framework, which is what [17] utilized in their work on the Iso-CFM. The main focus of the CFM is on modelling the evolution of firn density and temperature using a Lagrangian (firn parcel-following) grid, where each parcel (model volume) represents a layer of firn with uniform properties. In this work

it is carried out by first assuming that the accumulation rate at the site under consideration is constant and that the firn-density profile is in a steady state (Sorge's law). The densification rate is computed at each step in the model, using any of the given previously published firn-densification models, provided in the modules, and the firn density is updated in each time step. Following, the firn temperature evolution is computed through a coupled heat-diffusion model, by the use of a finite-volume method, [51, Patankar, 1980]. The model then proceeds by adding a new layer on top of the simulated firn column, described by the provided input parameters (temperature, accumulation rate, density), and the bottom volume of the grid is removed.

For this work, the CFM is only utilized with steady state conditions (i.e. constant accumulation rate and temperature), but for future works many different dynamic assumption could be made for more accurate depiction of densification and diffusion processes along with more precise temperature estimates.

The user starts by specifying input parameters(firn-densification physics, time-step size, surface boundary conditions and more) in a .json file. A model run is then started by a "spin-up" which determines a steady state model used as the initial condition for the main model run. Thus, for the initial model, a steady state analytic firn-densification Herron-Langway model [25, Herron and Langway, 1980] is used to calculate depth-density and depth-age profiles, using the forced steady state parameters temperature, T_0 and accumulation, \dot{b}_0 . Then stepping forward one time step, it uses the specified densification model to evolve the firn in time. An in depth description of the entire CFM can be found in [61, Stevens et al.].

2.3.1 Iso-CFM

The Iso-CFM building on the Community Firn Model, is a tool for estimating firn diffusion rates of water isotopes, $\delta^{18}\text{O}$, $\delta^{17}\text{O}$ and δD , developed by [?, Gkinis et al., 2021] for use in [?]. It requires two main inputs, temperature and accumulation rate, and no prior knowledge of the isotopic signal is required. The model provides computation of the diffusion lengths for the mentioned water isotopes. As previously discussed, the diffusion lengths are a metric for the smoothing a signal has undergone and can, along with deconvolution techniques also discussed in this thesis, be used to reconstruct some of the signal that has been otherwise attenuated.

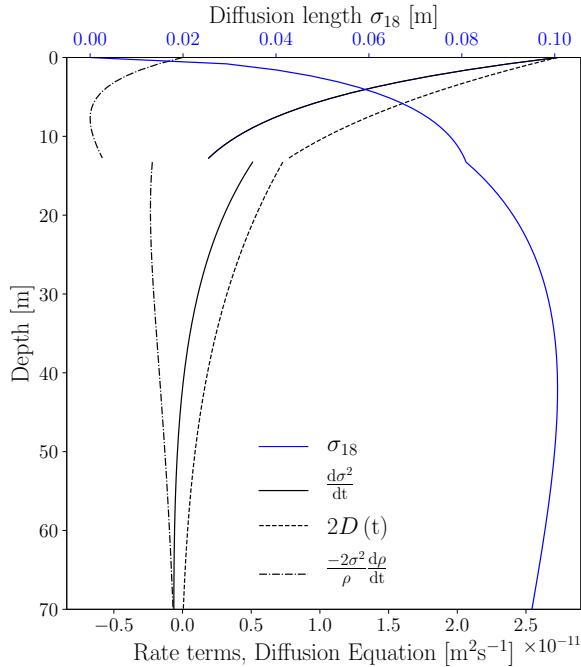


Figure 2.7: Contribution of the diffusion(dashed) and densification(dot-dashed) terms from Eq. 52 to the final analytical diffusion length solution (blue).

2.3.1.1 Diffusion Length Profiles

The iso-CFM computes a numerical solution for σ using a time-stepping scheme, as is the case for the original CFM, to estimate the most likely diffusion length profile at a given site. From each time step the CFM computes $\frac{d\rho}{dt}$ and T , and the iso-CFM uses these results to calculate the quantity $\frac{d\sigma^2}{dt}$:

$$\frac{d\sigma^2}{dt} = 2 \left(D(t) - \frac{\sigma^2}{\rho} \frac{d\rho}{dt} \right) \quad (2.21)$$

Equation (52) shows that the diffusion length signal throughout the ice is a result of two processes, opposing each other: the always positive diffusivity term $D(t)$, and the densification process contributing negatively to the change over time, $-\frac{\sigma^2}{\rho} \frac{d\rho}{dt}$. After a certain depth, the densification term comes to dominate and thus the entire equation becomes negative and the value of the diffusion length is decreasing, see Figure 10.

To simplify the work of this thesis, the numerical module of the CFM

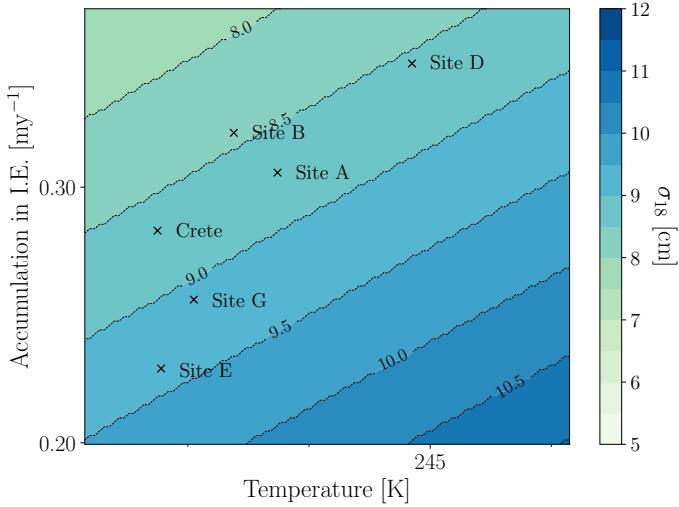


Figure 2.8: Crete and surrounding Alphabet cores, as their analytical solutions place them according to observed temperature and accumulation rate.

and the iso-CFM has not been implemented in the final computations, and the diffusion length profiles referred to in the rest of the project are calculated through an analytical method, using equations derived from Equation (52) analytically. A short walk-through of the derivations are presented in Appendix 7 as they are described in [18, Gkinis et al., 2021]. Five examples of diffusion length profiles given different conditions are presented in Figure 12, with the same conditions as used in 9.

The analytical equations derived in Appendix 7 have been used for creating a contour plot of the analytical solutions for σ_{18} at the close-off density, ρ_{co} . This can be seen in Figure 11. The plot shows six different ice cores drilled in the proximity of the Crête ice core drill site and their diffusion length analytical solutions.

These analytical equations are used to compute diffusion lengths to compare with the optimal diffusion length estimates computed from the raw data. One could advantageously explore the iso-CFM further to numerically compute the diffusion lengths with different temperature and accumulation forcing to recreate a diffusion length profile corresponding to the most likely at a given drill site. Since the iso-CFM do consist of many different modules all with different possibilities for parameterisation, it is outside the scope of

this project to develop more advanced iso-CFM diffusion length estimates for the examined cores, and only the previously described, simple diffusion length profile estimate will be used in the further work. In-depth methodology and results from the iso-CFM can be found in [18, Gkinis et al., 2021].

2.4 Temperature Estimation

Through the theoretically based analytical (or numerical) estimates of the diffusion length, σ_{model} , and the diffusion length estimate from the isotopic signal of a depth section of an ice core, σ_{firn} , a temperature estimate of can be given. This estimate made by numerically solving Equation (2.10) with T as the unknown variable:

$$\left(\frac{\rho_{\text{co}}}{\rho_{\text{ice}}}\right)^2 \sigma_{\text{model}}^2(\rho = \rho_{\text{co}}, T(z), A(z)) = \sigma_{\text{firn}}^2. \quad (2.22)$$

For the temperature estimates made in this project a secant numerical method[53] was used, through the Python SciPy package `scipy.optimize.newton`. Only the function it self is provided and no derivative is given, so the `scipy.optimize.newton` uses the secant method to find a zero of the function passed. If a derivative was given, the packages would use a Newton-Raphson[53] scheme to find the zero of the function.

2.5 Dating of Ice Cores: ECM and DEP

Isotopic composition analysis has already been presented as a tool for analyzing ice cores to examine past temperatures, climate and atmospheric composition. Another method is through electrical conductivity measurements (ECM) [24, C. Hammer, 1980] which is much more sensitive to violent volcanic eruptions and deposition of volcanic material. This sensitivity makes it possible to use historically dated eruptions visible in the ice cores as volcanic horizons, and thus making dating of the ice core more precise and absolute. In Figure 2.10 two aligned depth series can be seen: the water isotope profile and the conductivity profile. The conductivity profile shows clear peaks, which are a result of the increased conductivity due to volcanic material deposited along with the snow at the ice core drill site.

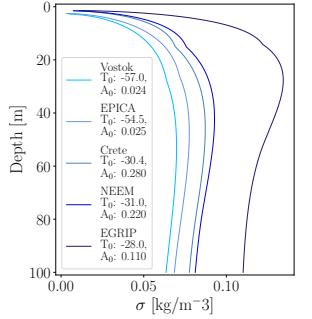


Figure 2.9: Analytically calculated diffusion length profile examples given five different initial conditions representing present day conditions at the five different ice core locations. Temperature, T_0 , is in $^{\circ}\text{C}$ and accumulation, A_0 , is in meter of water equivalent per year.

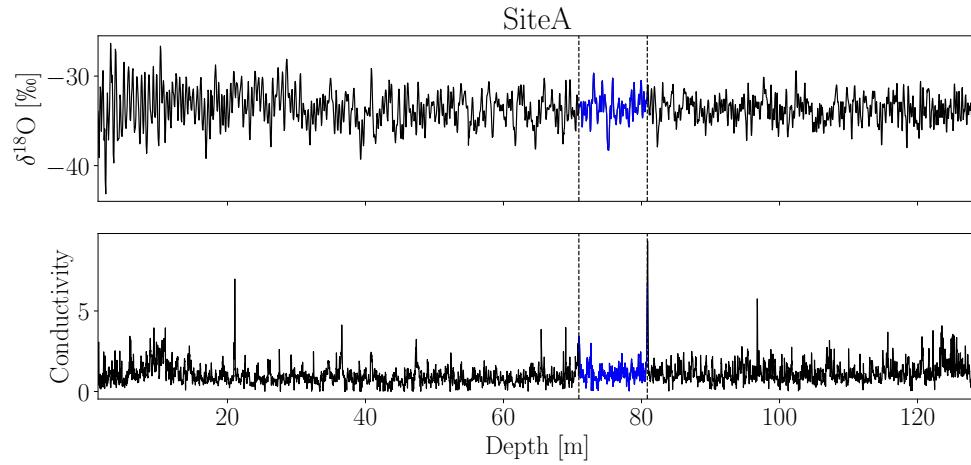


Figure 2.10: Water isotope measurements aligned with conductivity measurements from the ice core drilled at Site A.

2.5.1 Electrical Conductivity Measurements

The conductivity of ice arises from the current emerging due to the build-up of space charges in the ice structure. This conductivity can be analyzed by measuring the electrical current (DC) - induced by the electric potential and the acid balance - between two electrodes which are moved along the ice cores length. This current is related to the acid impurity concentration (pH), in the form of H_3O^+ concentration, of the ice core. Higher levels of acid impurity concentration are due to volcanic eruptions. Large amounts of volcanic gases, i.e. SO_2 , in the atmosphere oxidize and combine with water to form acid, i.e. sulphuric acid, which is then washed out of the air due to precipitation. Thus it is made possible to recognize volcanic horizons in ice cores, and - if the spatial location of the eruption is known - from the amount of acid, the magnitude of the eruption can also be estimated.

The measured current can then be transformed into acidity by a calibration curve relating the current, in μA , to the acidity, in $\mu\text{equivalents H}_3\text{O}^+$ per kilogram. To find the calibration parameters, the current and the acidity must be measured - the current through the above mentioned method, and the acidity through pH measurements of melted ice core samples. The pH measurements must further be corrected for any CO_2 induced H^+ ions [63, K. Taylor et al., 1993]. The relation between acidity $[\text{H}^+]$ (corrected for CO_2 induced H^+) and current I can be expressed in two ways:

- $[H^+] = (0.017 I^2 + 1.2) \mu\text{equiv. H}^+/\text{kg}$
without a 50% correction for CO_2 surplus.
- $[H^+] = (0.045 I^{1.73}) \mu\text{equiv. H}^+/\text{kg}$
with a 50% correction for CO_2 surplus.

The salt concentration in the ice can be estimated from measurements of the specific conductivity κ of the melted samples. The salt contribution hereto can be expressed as:

$$\kappa_s = \kappa - \kappa(\text{H}^+) - \kappa(X^-) - \kappa(\text{HCO}_3^-) \quad (2.23)$$

where the three later terms correspond to the contributions from H^+ (through pH measurements) and its anions¹, HCO_3^- and any other anions X^- . The anion concentration will be equal to the cation concentration, which in this case is only H^+ concentration. Disregarding low acidity samples, the concentration of HCO_3^- is negligible and thus $\text{concentration}(X^-) \approx \text{concentration}(\text{H}^+)$. The current is thus heavily influenced on/determined by the H^+ concentration, and thus it is approximated that the salt concentration has no influence on the current readings, which is fortunate, since the ECM method only responds to acidity, and not to salt and ammonia concentrations. This is one of the method's limitations, which the later dielectric profiling method has taken into account [45, J. Moore and S. Fujita, 1993], [44], see Appendix 7.

Though DEP measurements might have been more sensitive to salt and ammonia concentrations and thus would have given a different signal, the data available for conductivity for the ice cores under consideration in this thesis are made through the ECM method. In Figure 2.10 the raw ECM signal can be seen, aligned with the isotopic signal. Marked in blue is the estimated depth between the two eruptions Laki and Tambora.

Through electrical conductivity measurements it is possible to observe the very clear effects of some volcanic events in ice cores. Particles from the eruption are quickly transported from the source, since the atmospheric airflow will scatter the particles all over the atmosphere at a relatively high speed. Thus the dust (particle) and ECM signals pick up the volcanic signal faster than for example the isotopic signals. The isotopic signal reacts much slower, as it must be subjected to a change in global - and then following local -

²Anions are molecules losing a number of electrons to become negatively charged.
Cations are molecules that gain a number of electrons to become positively charged.

temperature, which might first show after a number of years. Thus ECM, DEP and dust measurements are good records to use for dating ice cores. Some eruptions are only great enough to show in ice cores located close to the volcanic source, while others are of a magnitude impacting the entire globe [35], [69], [11], thus showing in almost all ice core records. These volcanic horizons are specifically good for synchronizing records, which is essential for developing knowledge about the geographically varying climate, temperatures and hemispherical dependency of the past.

Chapter 3

Isotopic Data: Laki to Tambora as Seen in Five Ice Cores.

The data under examination in this thesis, namely the Alphabet ice cores near the Crête drill site, were chosen because of the relatively high resolution and because of the existence of both water isotopic data and electrical conductivity measurements. The data and the selection of data is presented in this chapter, along with some corrections applied to the original conclusions from [10, Clausen & Hammer, 1988].

Contents

3.1	Volcanic Horizons	27
3.1.1	Corrected Depth	29
3.1.2	Gaussian Distribution	29
3.2	Selection	30
3.2.1	AWI B-cores	31
3.2.2	Crete Area	31

3.1 Determining Time of Volcanic Material Deposition

For this thesis, two volcanic horizons have been in focus, namely the eruption of the Icelandic volcano Laki in 1783 and the Indonesian volcano Tambora

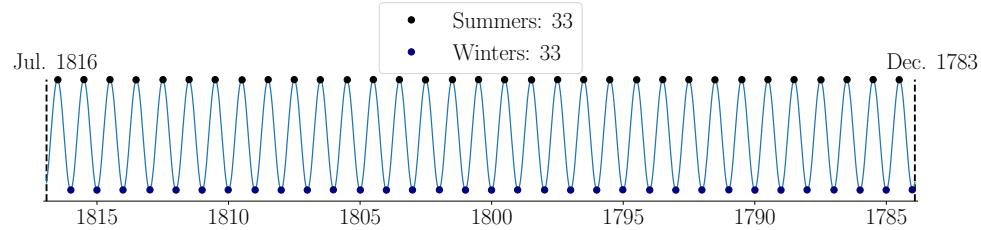


Figure 3.1: Illustration of expected summers and winters in the time span between the Laki and Tambora volcanic depositions in Greenland.

in 1815. Due to delay in atmospheric transport ([69, Wei et al., 2008], [11, Cole-Dai et al, 2009]), the volcanic material from Tambora was first deposited around the summertime of 1816. This reveals a time span between volcanic material deposition from the two eruptions of 33 summers (peaks) and 33 winters (troughs). The essence of this project is to restore as much of the diffused signal as possible, while obeying the constraint, that the isotopic signal must correspond to this time span, that is, the signal must exhibit 33 peaks and 33 troughs.

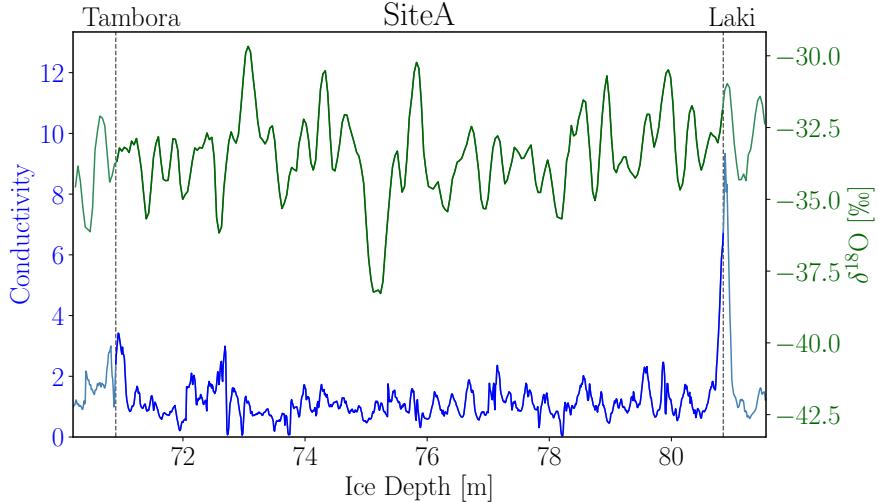


Figure 3.2: Conductivity measurements aligned with water isotope measurements at the depth corresponding to the time between eruptions at Laki and Tambora.

In Figure 3.2, the isotopic and the conductivity signals have been aligned from the assumed positions of the Laki and Tambora depositions.

The Laki deposition signal is especially well-defined, as it originates from a volcano in the vicinity of the Greenlandic ice core. The Tambora signal on the other hand is more smudged and has a wider signal with a less well-defined peak. This is due to the large hemispheric distance between Greenland and Indonesia, and the time it thus takes for the volcanic material to be distributed through the atmosphere and finally deposited at Greenland.

Specifically considering the constraint involving the 33 troughs, there might be room for some error, as it is clear that if the deposition happened just a bit earlier or later than expected, there could be one or two troughs more or less.

3.1.1 Corrected Depth Estimate

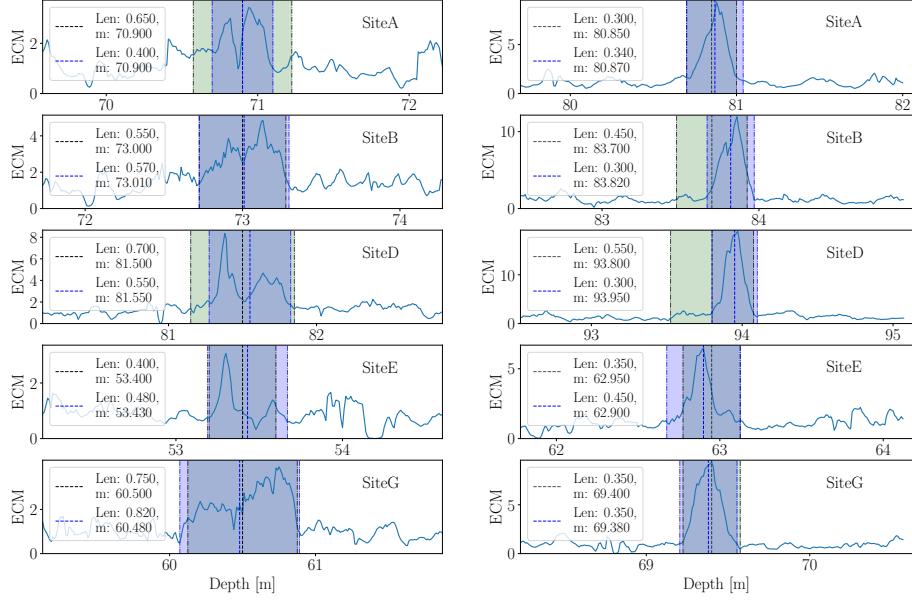
[10, Clausen & Hammer, 1988] presented estimates of the Laki and Tambora positions from the ECM measurements previously presented in this thesis. These positions have in this thesis been revisited and corrected to a more fitting central value, which can be seen in Figure 3.3 and Table 3.1. In the table the subscript 'CH' refers to values as found in [10, Clausen & Hammer, 1988] and subscript 'TQ' refers to new estimates as found in this thesis. m corresponds to what is estimated to be middle of the event and s corresponds to the estimated width of the event.

Core	Tambora				Laki			
	m_{CH} [m]	m_{TQ} [m]	s_{CH} [m]	s_{TQ} [m]	m_{CH} [m]	m_{TQ} [m]	s_{CH} [m]	s_{TQ} [m]
Site A	70.90	70.90	0.65	0.40	80.85	80.87	0.30	0.34
Site B	73.00	73.01	0.55	0.57	83.70	83.82	0.45	0.30
Site D	81.50	81.55	0.55	0.70	93.80	93.95	0.55	0.30
Site E	53.40	53.43	0.40	0.48	62.95	62.90	0.35	0.45
Site G	60.50	60.48	0.75	0.82	69.40	69.38	0.35	0.35

Table 3.1: Original and corrected middle, m , and width, s , values for Laki and Tambora deposition events.

3.1.2 Gaussian Distribution

From the corrected middle and width estimates for the volcanic events observed in ECM data, it is possible to investigate what happens to the further analysis if the event location is moved from a fixed point. This gives a possibility of estimating the locations as Gaussian distributions, with a mean equal



(a) Corrected depth positions and duration for Tambora deposition event.

(b) Corrected depth positions and duration for Laki deposition event.

Figure 3.3: Corrected depth positions durations for the volcanic signal of Laki and Tambora in the ECM data. Green shades correspond to previously estimated depths, from [10, Clausen & Hammer, 1988], and blue shades correspond to newly estimated positions and durations.

to the middle value and a standard deviation of $1/4$ (Tambora) or $1/5$ (Laki) of the width, and thus makes it possible to draw the Laki and Tambora estimated locations from these distributions. This enables further analysis of the developed method, as the stability can be tested by varying the positions of the volcanic events within these Gaussian distributions. To be able to use the developed method, it might be necessary to make the constraints of 33 peaks and 33 troughs soft but the stability of these constraints is also something that can be investigated through this method. Figures 3.5 and 3.4 show examples of Gaussian distributions generated from investigation of midpoint and width of the volcanic events.

3.2 Selection of Isotopic Data

The method developed in this project is very general and can hopefully be used for information reconstruction and diffusion length estimation in a great

number of different ice cores. But to develop a general algorithm one must first test it on specific data sets. The focus was chosen to be mainly on a number of shallow ice cores, the Alphabet cores near the Greenlandic ice core Crête [10], and especially on the core drilled at Site A, see Figures 3.7 and 3.8 for location of the different cores examined.

3.2.1 AWI B-cores: Core B23

Before choosing to focus mainly on the Alphabet cores, some time was spent on examining a number of cores of length between 100-175 m drilled during the North Greenland Transverse (NGT) between 1993 and 1995 in northern Greenland, from now on referred to as the AWI (Alfred-Wegener-Institut) B-cores, [70, Weissbach et al. 2016]. These were primarily chosen due to their great spatial coverage of an area of roughly 10 % of the Greenland ice sheet. This could have proven very useful for using the method developed here to estimate a spatial-temporal map of the covered area in the period between the eruptions of Laki and Tambora. Unfortunately the data from the AWI B-cores were not of high enough quality to meet the requirements of the following data analysis. Of the twelve AWI B cores available, only seven had corresponding electrical conductivity measurements with recognizable Laki and Tambora signals. Out of these seven only three were of adequate quality and resolution to subsequently be analyzed, see Appendix 7. The $\delta^{18}\text{O}$ and electrical conductivity profiles of one of the three high-quality cores from the NGT can be seen in Figure 3.6.

3.2.2 Crete and Surrounding Alphabet Cores: Site A

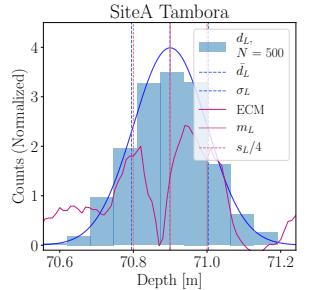


Figure 3.4: Example of Gaussian distribution of the volcanic event from Tambora, generated from observations of the ECM data, Site A. μ_T for the distribution is set to be equal to the middle point, s_T , and the standard deviation, σ_T^2 is set to be $s_T/4$.

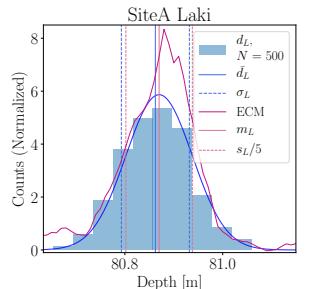


Figure 3.5: Example of Gaussian distribution of the volcanic event from Laki, generated from observations of the ECM data, Site A. μ_L for the distribution is set to be equal to the middle point, s_L , and the standard deviation, σ_L^2 is set to be $s_L/5$.

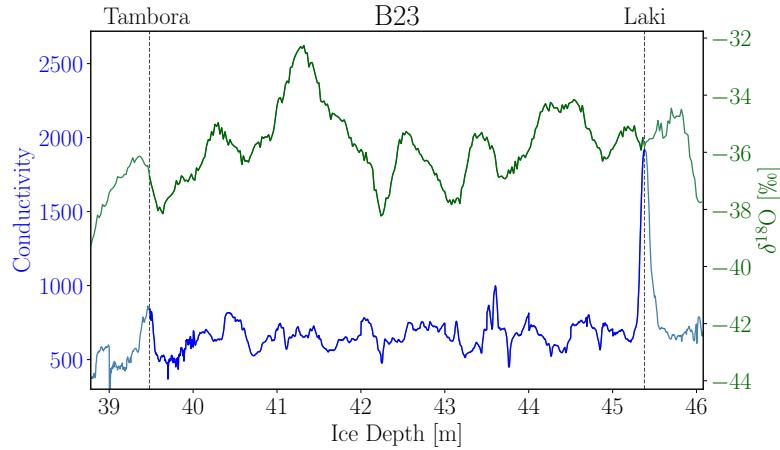


Figure 3.6: $\delta^{18}\text{O}$ and conductivity profile of the AWI B-core B23. The dashed lines represent the suggested locations of the Laki and Tambora eruptions as matched in [70, Weissbach et al. 2016]

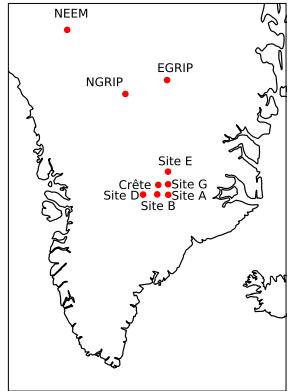


Figure 3.7: Location of Alphabet cores along with some major ice cores, NEEM, EGRIP and NGRIP.

The cores drilled in 1984-85 around the Crête core consist of the 400 m Crête core obtained in 1974 [9, H. Clausen et al., 1988] and eight shallow cores of varying length, between 25 m and 130 m, drilled in the Crête vicinity with a spatial coverage of 150×150 km, [?, Clausen, Gundestrup, Johnsen 1988]. See Figures 3.7 and 3.8 for spatial positioning of the ice cores.

Three cores were not of use for this project, Site C and Site F due to their shallow maximal depth, and the Crête core, as the ECM data were missing. The remaining five cores make up the cores in focus of this project. They are all well-documented, [10, Clausen & Hammer, 1988], [?, Clausen, Gundestrup, Johnsen 1988], and of high resolution making them ideal for the data and signal analysis used in the scope of this thesis. The Alphabet cores under consideration are presented in their full length in Figure 3, and at the

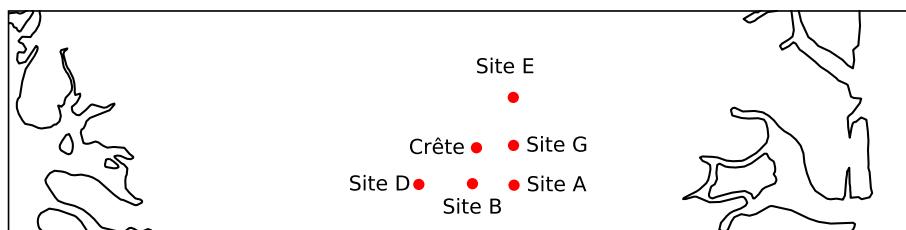


Figure 3.8: Map of spatial locations of the Alphabet cores analyzed in this thesis.

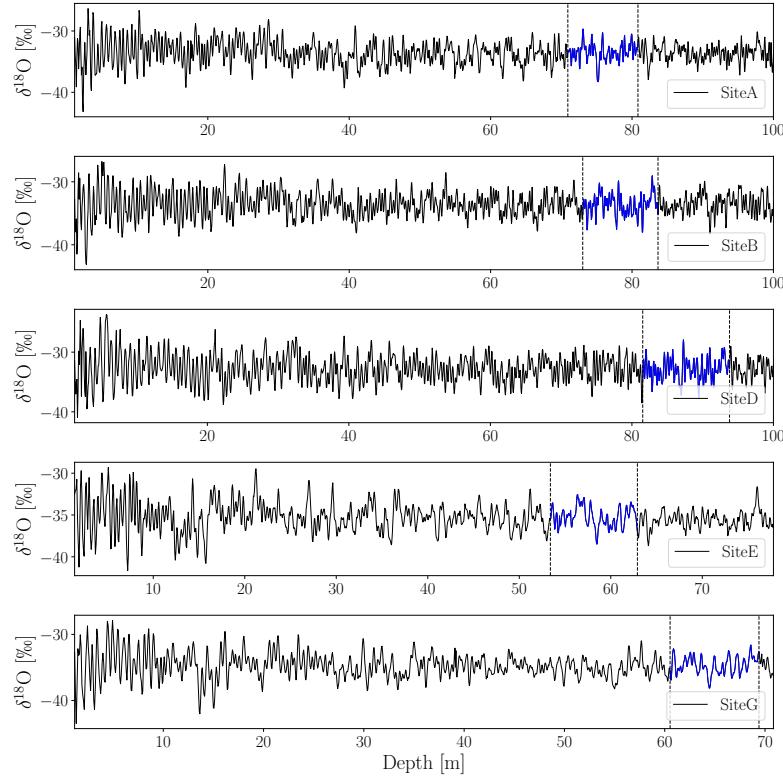


Figure 3.9: All Alphabet cores examined in this thesis in their full length. Blue sections correspond to depth of Laki to Tambora

Laki to Tambora depth sections in Figure 4.

3.2.2.1 Data Specifications

The general specifications of all five cores, i.e. drill site information and volcanic event locations, are presented in Table 3.2 [10]. d describes depth of event, A describes accumulation rate, T describes temperature, ρ describes density at a given depth and s describes the width of a given event. Subscripts L and T stand for volcanic events Laki and Tambora, respectively, subscript 0 describes initial surface condition, and superscripts CH and TQ represent the original and corrected values for depth and width of events.

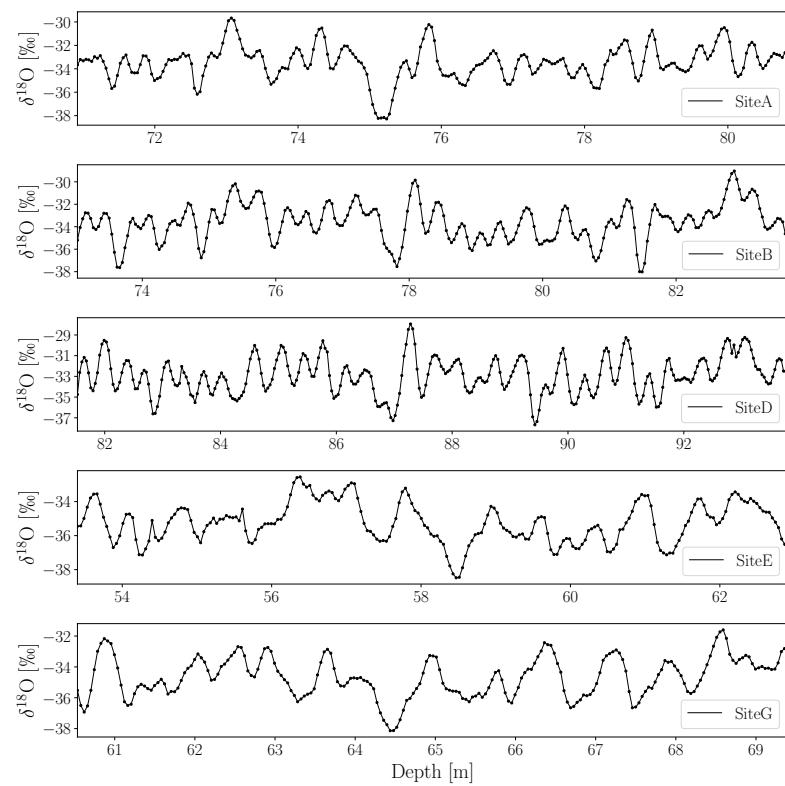


Figure 3.10: All Alphabet cores used in thesis at the depth corresponding to Laki to Tambora.

		Site A	Site B	Site D	Site E	Site G
d_L^{CH}	[m]	80.85	83.7	93.8	62.95	69.4
d_T^{CH}	[m]	70.9	73.0	81.5	53.4	60.5
A_0 W.E.	[m]	0.307	0.327	0.365	0.225	0.251
A_0 I.E.	[m]	0.282	0.300	0.335	0.206	0.230
T_0	[°C]	-29.41	-29.77	-28.3	-30.37	-30.1
ρ_0	[kg m ⁻³]	343.0	355.0	350.0	325.0	-
z_0		0.55	0.55	0.825	0.675	-
s_L^{CH}	[cm]	30.0	45.0	55.0	35.0	35.0
s_T^{CH}	[cm]	65.0	55.0	70.0	40.0	75.0
ρ_L	[kg m ⁻³]	836.0	841.0	857.0	786.0	807.0
ρ_T	[kg m ⁻³]	812.0	816.0	839.0	749.0	778.0
d_L^{TQ}	[m]	80.87	83.82	93.95	62.9	69.38
d_T^{TQ}	[m]	70.90	73.01	81.55	53.43	60.48
s_L^{TQ}	[cm]	34.0	30.0	30.0	45.0	35.0
s_T^{TQ}	[cm]	40.0	57.0	55.0	48.0	82.0

Table 3.2: Core specifications for core drilled at Site A. d describes depth of event, A describes accumulation rate, T describes temperature, ρ describes density at given depth and s describes the width of a given event. Subscripts L and T stand for volcanic events Laki and Tambora, respectively, subscript 0 describes initial surface condition, and superscripts CH and TQ represent the original and corrected values for depth and width of events.

Chapter 4

Signal Analysis and Computational Methods

This work is based on the main idea that a diffused depth series can be attempted reconstructed through various analysis of especially the frequency spectrum of the signal and through theoretical and empirical knowledge of diffusion and densification processes in snow, firn and ice. To utilize this idea, the specific analysis methods used must be well understood. This chapter focuses on the signal analyses and computational methods used in the further analysis, and discusses some of the challenges encountered.

Contents

4.1	Data as a Time Series	38
4.2	Spectral Analysis	39
4.2.1	PSD	39
4.3	Back Diffusion	42
4.3.1	Signal Restoration	42
4.4	Peak Detection	49
4.5	Constraint: ALT from Spectral Analysis	51
4.5.1	Estimating λ_A	51
4.5.2	Change in l_{sec}	53
4.5.3	Change in l_{shift}	55
4.6	Splines and Interpolation	56
4.6.1	Interpolation	57
4.6.2	Interpolation in this Project	58

This chapter only focuses on the hands on methods used and does not go into depth with the next part of the work, which uses the back diffusion method along with a constrained optimization to estimate an optimal diffusion length, σ , for a given depth section. The optimization method and algorithms are presented in the next chapter, Chapter 5. This chapter focuses on spectral transform and analysis of a depth series, Section 4.2, frequency filtering and deconvolution as means to back diffusion, Section 4.3, peak detection, Section 4.4, annual layer thickness as constraint for peak detection, Section 4.5, and spline interpolation, Section 4.6.

4.1 Isotopic Data as a Time Series

The presented data obtained through various experimental measurements, electrical conductivity measurements and water isotopic 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.

This section will contain a detailed presentation of the signal analysis and computational methods used to obtain the restored and enhanced signal that is sought after. In Figure 4.1 the general back diffusion method for signal restoration and enhancement is presented, with the specific modules described in this section highlighted in bold text and with an YellowGreen colour.

The method presented in Figure 4.1 is a subpart of the final algorithm that will be presented in the subsequent chapters. Figure 4.1 describes the flow of the general method, through back diffusing a measured depth series given empirically modelled σ . Thus it describes a single process, which is implemented in the optimization algorithm later on, where the diffusion length is the unknown parameter that is being optimized.

In this chapter spectral analysis and frequency filtering are presented as the main signal analysis methods used. The description of spectral analysis also contains a walk through of how back diffusion is carried out through

spectral analysis. As computational tools this chapter presents spectral transforms, interpolation and constrained peak detection, all methods used as instruments to improve accuracy and precision of the analysis.

4.2 Spectral Analysis of Time Series

A very useful tool for analyzing signals exhibiting oscillatory effects is analysis of the signals power spectrum. Instead of considering the signal in time (or depth), 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 is presented the three that has been implemented and utilized throughout this thesis. Since the data are discrete and experimental samples, it will be the discrete and applicable mathematical models that are presented.

4.2.1 Power Spectral Densities

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:

$$\text{Total Power} = \int_{-\infty}^{\infty} |X(\tau)|^2 d\tau. \quad (4.1)$$

Using Parseval's theorem [50, M. Parseval, 1806] (assuming that the signal has a finite total energy), the power of the signal can alternatively be written as

$$\int_{-\infty}^{\infty} |X(\tau)|^2 d\tau = \int_{-\infty}^{\infty} |\tilde{X}(f)|^2 df \quad (4.2)$$

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

$$\tilde{X}(f) = \int_{-\infty}^{\infty} X(\tau) e^{2\pi i f \tau} d\tau \quad (4.3)$$

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

$$X(\tau) = \int_{-\infty}^{\infty} \tilde{X}(f) e^{-2\pi i f \tau} df. \quad (4.4)$$

Both $X(\tau)$ 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:

$$\begin{aligned}\tilde{X}(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} X(t) e^{i\omega\tau} d\tau \\ X(\tau) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{X}(\omega) e^{-i\omega\tau} d\omega\end{aligned}\tag{4.5}$$

From Equation 4.2 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:

$$P_X(f) \equiv |\tilde{X}(f)|^2 + |\tilde{X}(-f)|^2 \quad 0 \leq f < \infty\tag{4.6}$$

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 different ways of estimating the PSD. In the following a number of different transforms will be presented briefly, all used in this thesis. For a more in depth description and discussion of the individual transforms, see Appendix 7.

4.2.1.1 Spectral Transforms

- **DFT/FFT** In the above section the general description of the continuous Fourier Transform and its inverse were presented. When considering discrete functions, as is generally the case with measured data, the Fourier transform of the measurements will also be discrete and represented over a finite interval $[a; b]$ for functions with a period of $L = b - a$, described continuously as $C_N \int_a^b X(\tau) e^{i\frac{2\pi}{L}\omega\tau} d\tau$ with C_N being the appropriate normalization. The discrete version of the Fourier transform is referred to as the Discrete Fourier Transform (DFT). It transforms the discrete signal into a sum of separate components contributing at different frequencies.

The DFT and its inverse for a data series consisting of N discrete points are defined as:

$$\tilde{X}_n \equiv \sum_{k=0}^{N-1} X_k e^{2\pi i k \frac{n}{N}} \quad (4.7)$$

$$X_n \equiv \frac{1}{N} \sum_{n=0}^{N-1} \tilde{X}_n e^{-2\pi i k \frac{n}{N}} \quad (4.8)$$

The DFT is dependent on the sampling interval, Δ_z , which limits the bandwidth to frequencies smaller in magnitude than the so-called Nyquist critical frequency, $f_{NQ} \equiv \frac{1}{2\Delta_z}$. That is $\tilde{X}(f) = 0$ for $|f| \geq f_{NQ}$. Thus when considering a signal with frequencies both inside and outside this Nyquist interval, the spectral information outside of the interval will be falsely interpreted as being inside the interval - this is called aliasing, and gives rise to an increased power at the daughter frequency of the aliased frequencies.

Computation of the DFT can be very slow, but a number of different algorithms have been developed for fast and efficient computation of the DFT, and the one considered in this thesis is referred to as the Fast Fourier Transform (FFT). The FFT implemented in this work is the `scipy.fft`.

- **DCT** The full Fourier transform is designed to process complex-valued signals, always producing a complex-valued spectrum, as the cosine and sine functions each contain individual information of the spectrum and constitute a complete set of basis functions. But a purely real-valued signal has a symmetric Fourier spectrum, meaning that it is only necessary to compute half the number of spectral coefficients, without losing any signal information. In this work, the data analyzed is purely real-valued. This can be utilized by only using the cosines as basis functions, describing the purely real part of the signal - which is exactly what is needed here. The FCT is commonly computed through the FFT algorithm, and in this thesis the `scipy.fft.dct` is used.
- **NDCT** Both the FFT and the FCT work under the assumptions that data is equispaced. This is not always the case when considering real world data, and when the data is nonuniform, the DCT is described as

$$\tilde{X}_k = 2 \sum_{n=0}^{N-1} X_n \cos \left(2\pi f_k \left(p_n + \frac{1}{2N} \right) \right), \quad 0 \leq k < M-1 \quad (4.9)$$

with, in the most general case, nonuniformly spaced signal, p_0, \dots, p_{N-1} , data and frequency data, f_0, \dots, f_{M-1} . The inverse of NDCT, the INDCT, is computed as:

$$X_k = \frac{\tilde{X}_0}{\sqrt{N}} + \sqrt{\frac{2}{N}} \sum_{n=1}^{N-1} \tilde{X}_n \cos \left(\left(p_n + \frac{1}{2N} \right) 2\pi f_k \right), \quad 0 \leq k < N - 1 \quad (4.10)$$

The data under consideration in this project is rarely exactly equispaced and previously, the spectral analysis made on isotopic depth series has assumed that the sampling size differences were of an order that could be ignored, and assumed that the samplings were uniform. When working with large data sets this is understandable, as it can slow down the analysis if the FFT or FCT could not be used. This can lead to a loss of information, as in nonuniformly sampled data, some features may be erased by assuming uniformity.

For this project though, the data sets are not of great numbers and the differences in and effects of using FFT, FCT(mostly just referred to as DCT) and NDCT has been examined.

4.3 Back Diffusion Through Spectral Analysis

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. For the data of this thesis a spectral method, determining the effect of mixing and diffusion as a spectral filter [29, S.Johnsen et al., 2000], [55, S. Rasmussen et al., 2005] is used as a restoration technique.

4.3.1 Signal Restoration by Optimal Diffusion Length

When considering a water isotopic depth series as the ones under examination in this thesis, it is possible to restore some of the signal lost to diffusion. This signal restoration technique will forwardly be referred to as 'back-diffusion', as it approximates the inverse of the diffusion process.

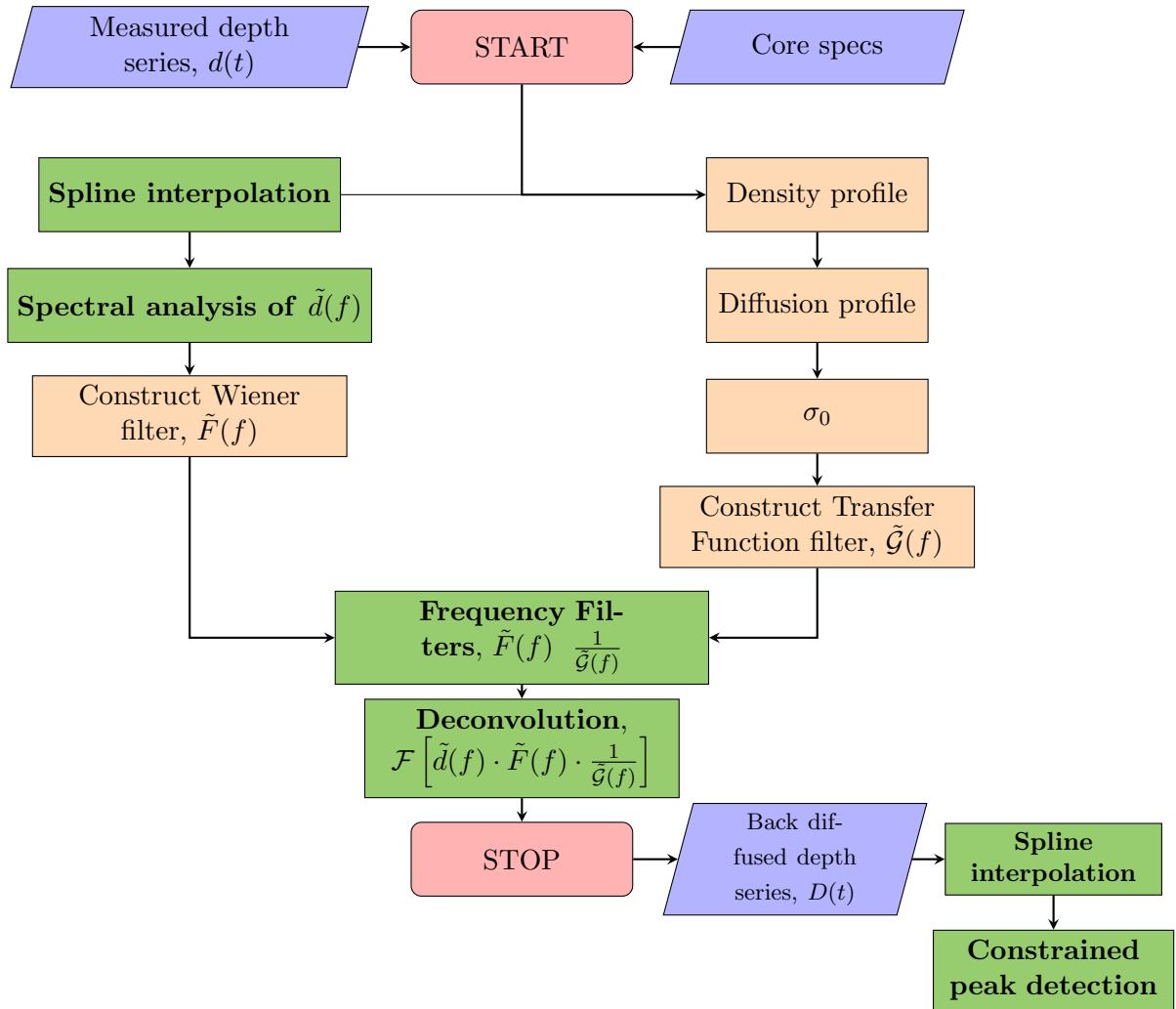


Figure 4.1: Flowchart of back diffusion method and peak detection, highlighted in green are the computational methods examined in this chapter: spectral transform and analysis of a depth series, Section 4.2, frequency filtering and deconvolution as means to back diffusion, Section 4.3, peak detection, Section 4.4, annual layer thickness as constraint for peak detection, Section 4.5, and spline interpolation, Section 4.6.

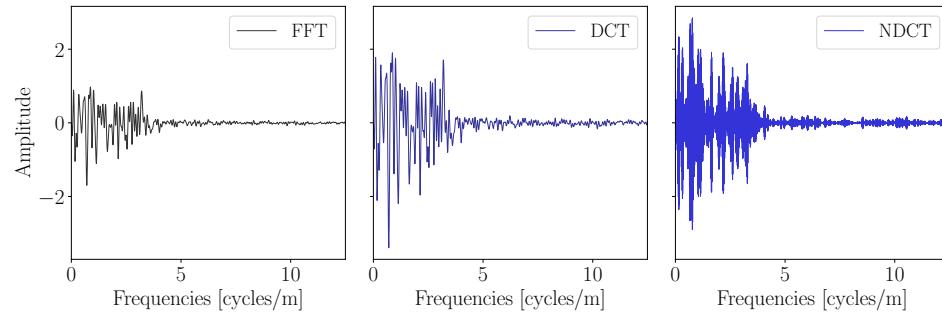


Figure 4.2: Examples of three different spectral transforms, FFT, DCT, NDCT, performed on the depth series between Tambora and Laki eruptions from Site A.

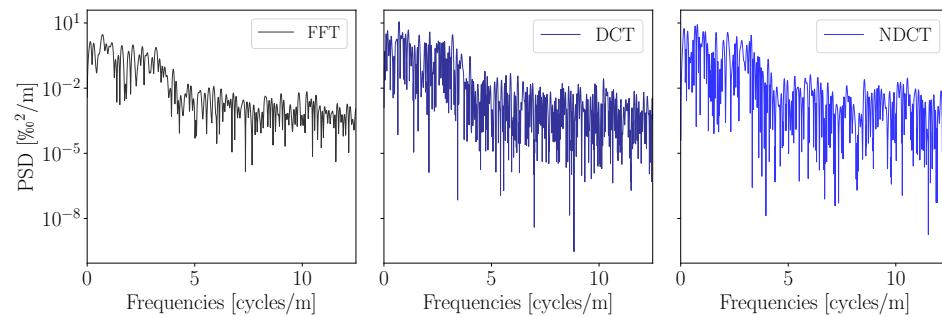


Figure 4.3: Examples of power spectral densities related to the three different spectral transforms, FFT, DCT, NDCT, seen in Figure 4.2.

The spectral filtering technique assumes the measured signal to be composed of a noise free signal and a noise signal. During the signal restoration process, a first estimate of the diffusion length σ of the depth series can be made, since the PSD of the noise free measured signal is described through $\text{PSD}_{\text{signal}} = P_0 e^{-k^2 \sigma_{\text{tot}}^2}$, where σ_{tot} describes the most likely mean diffusion length of the total mixing in the system. When estimating through spectral analysis, the diffusion length estimate will be a constant, even though the diffusion length is varying with depth. It is still a good estimate for the entire depth series, as the depth sections under consideration here show a relative small change in theoretically expected diffusion length, see Figure 4.4 showing the modelled diffusion length profile for Site A with a variation of 0.2 cm.

4.3.1.1 Spectral Filtering

When examining a real-world signal with focus on the frequency domain one will quickly run into the problem of noise. Different signals are prone to different types of noise, and must thus be treated in a fitting manner. Some spectra are prone to random low frequency noise and some to high, while others again might have a well-defined noise spectrum inherently. Understanding the nature of the noise is crucial for further signal analysis, as the signal-to-noise ratio (SNR) needs to be sufficiently high to be able to accurately separate signal and noise from each other. By understanding the noise, and modelling and estimating it, it can become possible to generate a filter which might minimize the noise and enhance the signal. This next section will go into detail with how to construct a filter fitting to the data at hand, and builds upon the theory presented in the previous chapter, along with [29] and [56]. This filtering will also be able to give an estimate of the diffusion length at a given depth section, due to the inherent nature of the signal and noise.

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 minimizing the effect of the noise, thus increasing the SNR.

Estimating the Unmixed Signal 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:

$$D(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases} \quad (4.11)$$

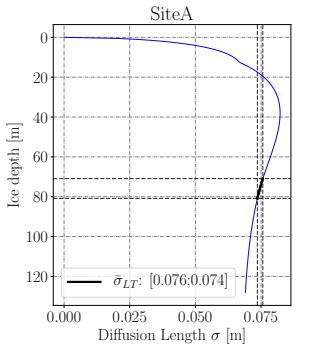


Figure 4.4: Diffusion length profile with section between Laki and Tambora highlighted.

In reality, a number of different mixing processes change this step function, and the measured signal will be a smooth curve, $d(t)$, which corresponds to the convolution of $D(t)$ with the mixing response function $M(\tau)$

$$d(t) = \int_{-\infty}^{\infty} D(\tau) \cdot M(t - \tau) d\tau \quad (4.12)$$

where t describes the time or depth with which the signal $D(t)$ varies with, and $M(t - \tau)$ in the integral describes the effect of the mixing response function at all different time or depth differences.

As is well known by the convolution theorem, in the spectral domain, convolution is multiplication and the mixing is described as the element wise multiplication between the Fourier transform of D and M :

$$\tilde{d}(f) = \tilde{D}(f) \odot \tilde{M}(f) \quad (4.13)$$

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

$$\tilde{d}'(f) = \tilde{D}'(f) \odot \tilde{M}(f) = \tilde{M}(f) \quad (4.14)$$

The mixing filter can thus be determined by measuring the system response to a step function, differentiating, and performing Fourier transform of the result $\tilde{d}'(f)$.

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

$$\tilde{D}(f) = \tilde{d}(f) \odot \frac{1}{\tilde{M}(f)} \quad (4.15)$$

Construction of Wiener Filter During the mixing, cycles with short wavelengths are heavily washed out, and through the restoration in Eq. 4.15, the amplitudes corresponding to these wavelengths are heavily amplified by the filter, since the restoration filter is chosen to reconstruct the wavelengths otherwise washed out by diffusion. 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 $\delta_M(z)$, consisting of the noise free signal $\delta_m(z)$ and the noise $\eta(z)$:

$$\delta_M(z) = \delta_m(z) + \eta(z) \quad (4.16)$$

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:

$$\tilde{F}(k) = \frac{|\tilde{\delta}_m(\omega)|^2}{|\tilde{\delta}_m(\omega)|^2 + |\tilde{\eta}(\omega)|^2} \quad (4.17)$$

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 described as

$$|\tilde{\delta}_m(\omega)|^2 = P_0 e^{-k^2 \sigma_{\text{tot}}^2} \quad (4.18)$$

where σ_{tot}^2 describes the total estimated diffusion length of the mixing. In ice core isotopic data, the noise in a signal is often due to random relocation of the snow deposited. This would lead to an assumption of the noise in the data under consideration here being described by white noise. But based on the conclusions of [26, C. Holme et al., 2018], stating that high accumulation Greenlandic ice cores, as the ones under consideration here exhibits red noise, the noise is assumed to be red noise, and is described by an autoregressive process of first order, AR1:

$$|\tilde{\eta}(\omega)|^2 = \frac{\sigma_\eta^2 \Delta_z}{|1 + a_1 \exp(-2\pi i \omega \Delta_z)|^2} \quad (4.19)$$

where σ_η^2 is the variance of the red noise, a_1 is the AR1 coefficient and Δ_z is the resolution of the time/depth data.

Since the power spectrum of the full signal $\delta_M(z)$ is a sum of the noise free signal and the noise, and the noise and signal are assumed described through Eq. 4.18 and 4.19, it is possible to estimate the parameters P_0 , σ_{tot}^2 , σ_η^2 and a_1 by curve fitting to the power spectrum. The estimated parameters are varied to find the optimal guess to use for the filter. An example of the constructed Wiener filter can be seen in Figure 4.6 on both a linear and a double logarithmic scale.

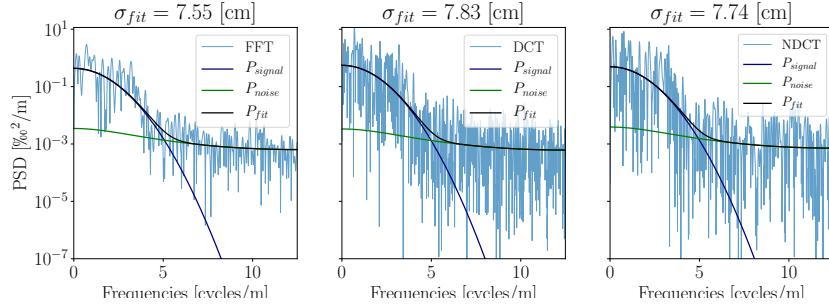
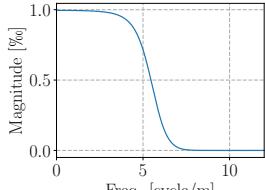
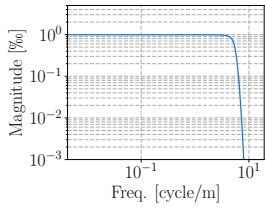


Figure 4.5: Noise, signal and total fit to PSD, illustrating the construction of the Wiener Filter.



(a)



(b)

Figure 4.6: (a) Wiener filter on linear scale. (b) Wiener filter on double logarithmic scale.

4.3.1.2 Final Restoration

After finding the best fit (noise and signal) to the spectral data, it is possible to construct an optimal restoration filter, $\tilde{R}(f)$, which contains two separate filters. One is the Wiener filter, $\tilde{F}(f)$, which is described in the above section and the other is a Gaussian filter constructed to amplify certain frequencies. This filter, the transfer function of the system, is constructed to specifically amplify the frequencies heavily attenuated by the diffusion process and is described as

$$\mathcal{G}(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-z^2}{2\sigma^2}}, \quad (4.20)$$

in the time (depth) domain and, since the Fourier transform of a Gaussian is still a Gaussian, in the frequency domain it is described as:

$$\tilde{\mathcal{G}}(f) = \mathcal{F}[\mathcal{G}(z)] = e^{\frac{-(2\pi f)^2\sigma^2}{2}}. \quad (4.21)$$

Finally, the constructed frequency restoration filter, in the frequency domain, is a product of the Wiener filter, from Equation (4.17), and the transfer function, from Equation (4.21):

$$\tilde{R}(f) = \tilde{F}(f) \frac{1}{\tilde{\mathcal{G}}(f)} \quad (4.22)$$

which results in a rewrite of equation 4.15 to:

$$\tilde{D}(f) = \tilde{d}(f) \tilde{F}(f) \frac{1}{\tilde{\mathcal{G}}(f)} \quad (4.23)$$

An example of the final frequency restoration filter, as it changes with diffusion length estimate inputted in the transfer function, $\tilde{G}(f)$, can be seen in Figure 4.7.

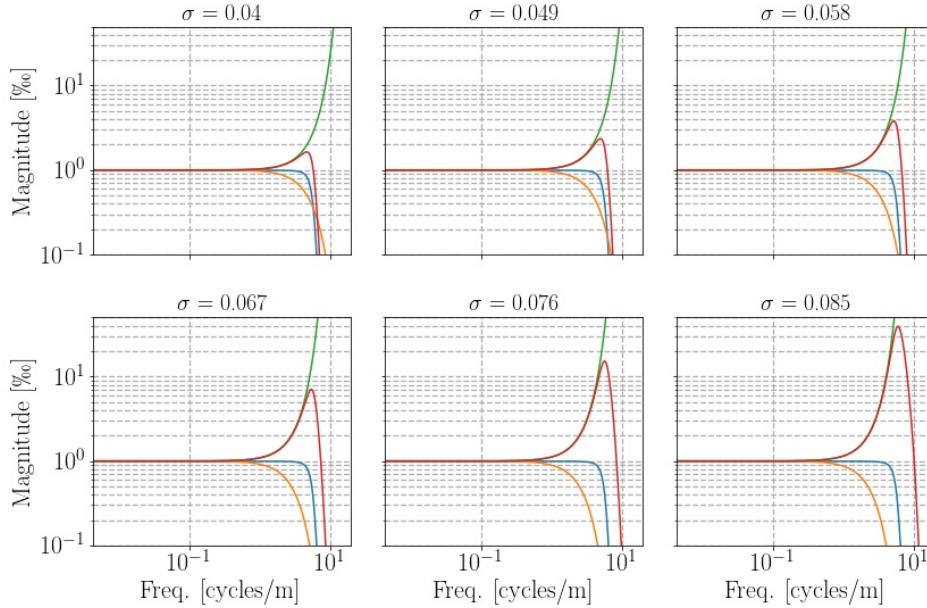


Figure 4.7: Frequency filter examples ranging from diffusion length 0.04 m to 0.085 m.

4.4 Peak Detection

After attempting to reconstruct the diffused signal as best as possible, the next question is, how to determine the estimated number of years in a given back diffused depth section.

Knowing that water isotopic data are a proxy for temperature, the most obvious way to determine annual layers in the signals is by detecting peaks and troughs. During colder periods, e.g. winter, the air masses arriving at the ice core sites have formed more precipitation before reaching the sites, and the vapor that results in this final precipitation is then more depleted of heavy isotopes, resulting in lower isotopic values, troughs in Figure 2.1. The precipitation falling during warmer conditions, e.g. summer, is correspond-

ingly less depleted of the heavy isotopes, and results in higher isotopic values, peaks in Figure 4.8.

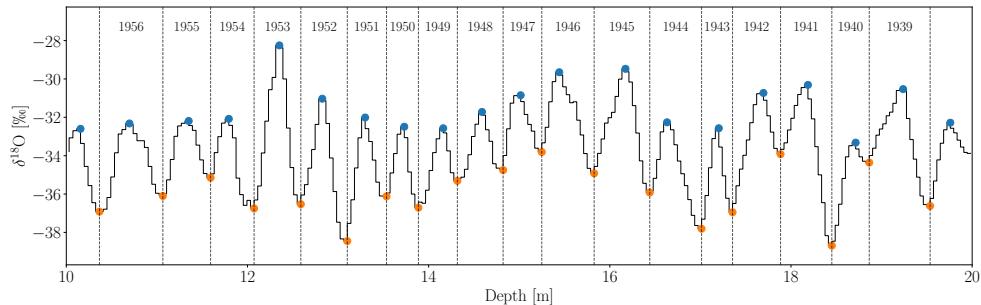


Figure 4.8: Ten meters of the top of Cret\acute{e} ice core, with identification and dating of 19 annual layers, with peaks(blue) corresponding to summers and troughs(orange) corresponding to winters.

Peak detection and layer counting has previously been carried out by visual inspection of the ice core depth signals, but as computers and algorithms have become more integrated in data analysis, it is now more common to use different automatic methods [71, J. Wheatley, 2015], [72, M. Winstrup, 2012], [56, S. Rasmussen et al., 2014]. Developing and implementing layer counting and peak detection algorithms can be done in a number of different ways, but for this project, at first a very simple method has been initially implemented and later the method has been improved and optimized through a number of different constraints. One could also use different pattern recognition techniques to achieve even more qualified seasonal cycle detection, and later some of these methods will be presented.

The most na\"{i}ve approach, and the one first implemented in this project, to peak detection is to simply find local maxima by comparing neighbouring values. When examining point d_i , the point is deemed a local maxima, if $d_{i\pm 1} < d_i$. Local minima, troughs, can be found in exactly the same manner by finding minima as $d_{i\pm 1} > d_i$. A very simple constraint for this method is to keep a required minimal distance between peaks, so that two peaks cannot be detected within a point distance of Δd_{\min} . For example at a depth of 12 m in Figure 4.8 two troughs can be seen, but only one is chosen, as they are within the threshold distance to each other, which here is set to $\Delta d_{\min} = 7$ points. Here, the lowest of the two troughs is chosen. The threshold distance can be chosen in different ways, for this short section it has been chosen through

visual inspection, but more generally it can be chosen by examining some of the intrinsic properties of the signal like the estimated annual layer thickness, λ_A , as explained in the following Section 4.5.

4.5 Constraint: Annual Layer Thickness Estimates

For a more reliable peak detection, it is of interest to impose some constraints representing the known nature of the signal, i.e. that it is continuous, that it should exhibit cyclical behavior or that the signal must decrease in amplitude with depth.

A feature, important to impose as constraint, that is also possible to examine through spectral analysis, is the annual layer thickness, λ_A , i.e. the distance between peaks or troughs in the annual isotopic signal [67, B. Vinther et al., 2006]. Due to densification processes and ice flow, λ_A is expected to decrease with depth, as the layers are thinning with depth. The annual layer thickness will be of use for the further analysis, as λ_A can be used as a constraint for the distance between peaks and troughs when counting years in the depth signal. Thus the goal is to estimate an annual layer thickness at the depth section between the Laki and Tambora events, which can then be used to set up a distance constraint. This section will present the method used for estimating the annual layer thickness piecewise throughout an entire ice core, and a discussion of the different parameters that can be varied when estimating λ_A . Especially the length of the depth section used in the spectral analysis, l_{sec} , and the length of the shift between sections, l_{shift} , are considered. Specifically the effect on the annual layer thickness estimate at the depth between Laki and Tambora by changes in l_{sec} and l_{shift} are investigated to make a qualified decision on what the optimal values for these parameters are.

4.5.1 Estimating λ_A Through Spectral Analysis

The transfer function of the diffusion process, Equation (4.21), can be rewritten as a function of the wavelength of the isotopic signal, λ :

$$\tilde{\mathcal{G}}(f) = e^{\frac{-2\pi^2\sigma^2}{\lambda(f)^2}} \quad (4.24)$$

The wavelength of the signal corresponds roughly to the annual layer thickness, λ_A , in the signal. Regarding the isotopic signal as a wave, it is clear

from both theoretical knowledge of the processes and visual inspection of the data, that the diffusion and densification have an effect on the magnitude and the wavelength/frequency respectively. The diffusion attenuates some of the waves magnitude, and the densification shortens the wavelength of the signal. As presented in the previous sections of this chapter, the general effects of the densification and diffusion processes can be examined by spectral analyzing the entire core. This is done by dividing the signal into a number of sections of an (almost) fixed length, l_{sec} and shifting this section down through the core with a shift of length l_{shift} . The length is not exactly fixed, as the sample sizes vary through the core, which results in small variations in section and shift lengths.

For a such depth section, the PSD of the signal is computed through the use of DCT, NDCT, FFT and MEM (Maximum Entropy Method, [66, T. Ulrych and T. Bishop, 1975]). All four spectral transforms are utilized to be able to give an uncertainty estimate on the annual layer thickness. Since the frequency of the seasonal cycle is assumed to be the most prominent in the spectrum, thus λ_A is detected as the most prominent peak, see Figure 4.9. Though as diffusion affects the signal more and more with depth, the annual signal becomes slightly more difficult to single out in the spectrum, and other features might result in more prominent peaks at lower frequencies. This has been dealt with in this thesis by simply assuming that the frequency in section l_{sec}^i must be larger than the frequency in l_{sec}^{i-1} , due to knowing that the densification and thinning in the firn column must result in a continuous decrease of annual layer thickness. The analysis is then carried out sequentially going through the depth segments

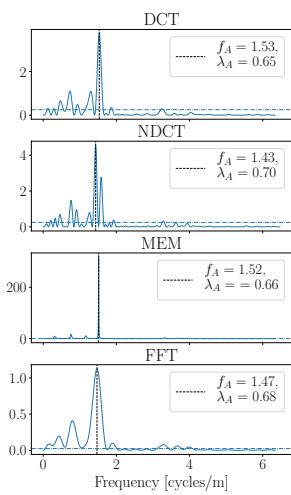


Figure 4.9: Example of annual layer thickness estimation for section of 5 meters at depth [0;5] m, Site B.

Figure 4.10 shows the λ_A estimates of the entire cores drilled at site A, B, D and G. The section length is 5 m and the shift is 4 m, to avoid missing any frequencies in the transitions between sections. Along with the estimated λ_A s, an estimate of the average $\bar{\lambda}_A$ value at the depth corresponding to between Laki and Tambora events has also been shown.

The final λ_{LT} used as a constraint on peak distance in the final algorithm is determined as an average of the estimated λ_A falling inside the Laki to Tambora depth interval.

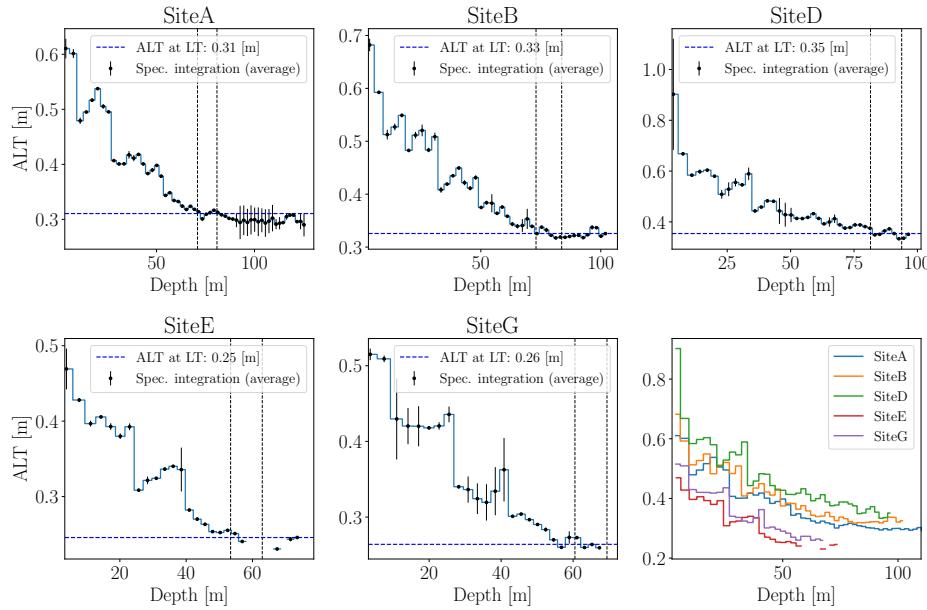


Figure 4.10: Annual layer thickness estimates, λ , calculated through spectral integration of 5 m sections of the entire core lengths. The average is based on using both FFT, DCT and NDCT spectral estimations. Blank spaces, akin to the sections below 60 m depth at Site E shows problematic areas, where the spectral estimation have had trouble with determining λ . The black vertical lines show the depth sections between Laki and Tambora events.

4.5.2 Change in section length

When performing spectral analysis on a time series, the length of the section under examination is crucial for the outcome of the analysis. If the section is too short to contain the information hidden in the signal, then it will be difficult to detect patterns and frequencies of interest. If the section is too long though, noise might dominate the spectrum, or longer term cycles that are not of interest for this analysis might become visible and dominant. Furthermore, for a nonlinear signal, as the isotopic depth series are, the linearity does not hold for very long sections, as the cycles are compressed and diffused. Thus it is of interest to examine the effect of the section length on the annual layer thickness estimates. The λ_A estimates were performed for sections lengths of 1.0 to 25.0 m. Above this upper limit, there were too much noise in the spectrum to determine a λ_A .

In Figure 4.11 different λ_A depth profiles can be seen, given five differ-

ent section lengths. All estimates are made with a shift of 1 m. Considering the shortest section length of $l_{sec} = 1.00$ m, it is clear that the estimation method picks up some noise resulting in a highly varying λ_A estimate throughout the cores. but, generally, the other four section lengths seem to pick up the same annual cycles, except for at Site E and G, where the two largest section lengths of 15.55 m and 20.39 m give some very long annual layer thicknesses at some depths.

Considering Figure 4.12 the λ_A estimates of the Laki to Tambora depth sectionsshow low stability at short section lengths and increasing accuracy and precision as the section length increases, until at some point, especially clear in Site E, but also visible in Site G and Site A, the accuracy start to drop as the variance of the λ_{LT} estimate increases.

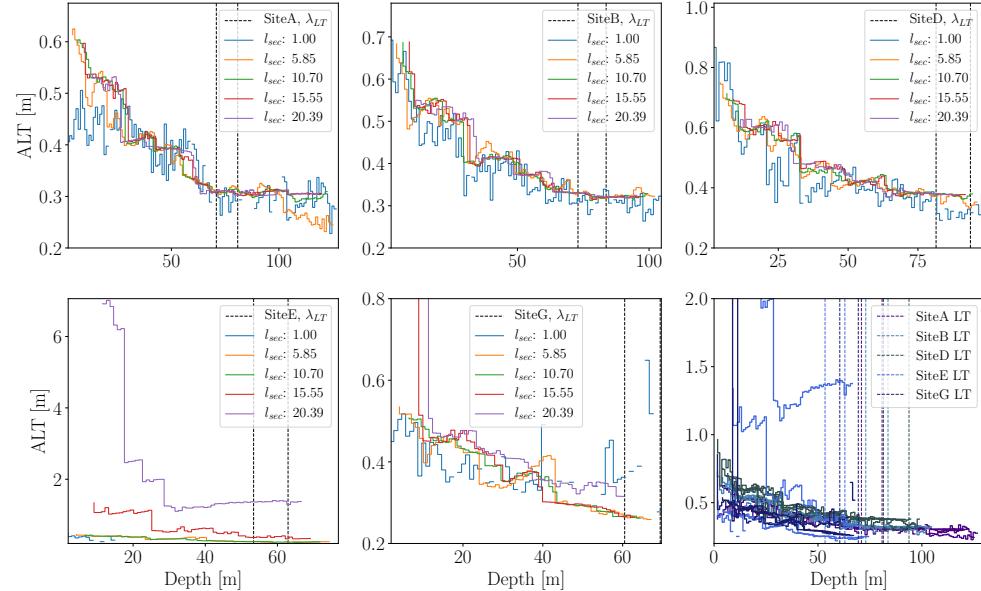


Figure 4.11: Examples of the λ depth profile given different section lengths. All profiles are computed with a shift of 1 m.

Based on the stability analysis of the section lengths, the general section length to estimate λ_A in this work was set to $\lambda_A = 7$ m, but with possibility to vary.

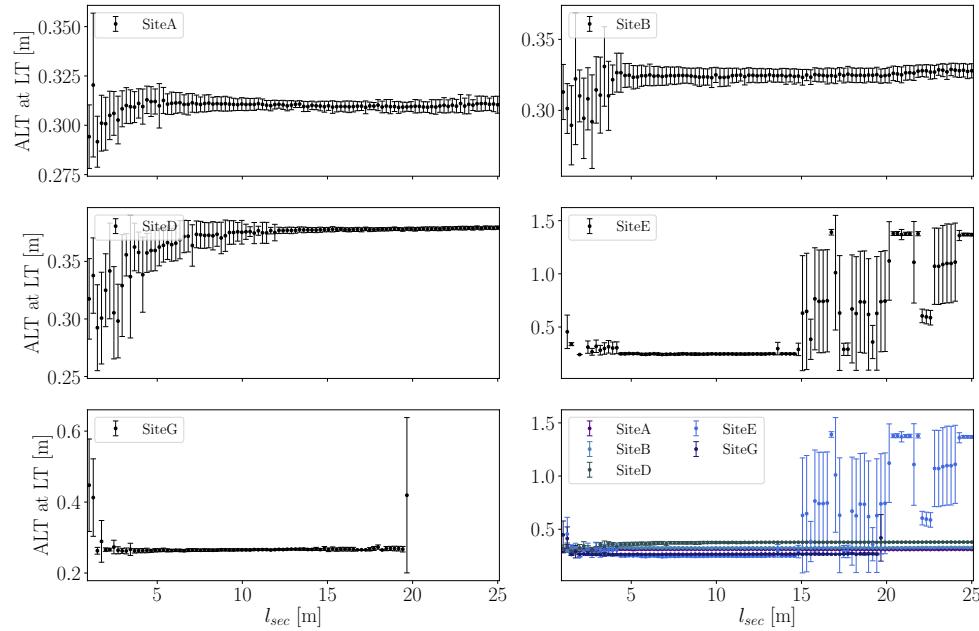


Figure 4.12: The annual layer thickness estimates in the section between Laki and Tambora as computed with different section lengths, l_{sec} , used in the spectral analysis. The section is shifted 1 m and then computed again. The λ_{LT} is then calculated as a mean of the λ estimates falling into the Laki to Tambora depth section.

4.5.3 Change in shift length

The change in shift length is assumed to be more an issue of how detailed the λ_A depth profile needs to be versus how computationally fast it is needed to be. If the shift is small, there will be a need to perform the spectral analysis many times before the entire depth is covered, but many small varying features may become visible throughout the core. All shift lengths are examined with a section length of $l_{sec} = 7$ m.

For this project though, a rough estimate of the λ_{LT} at a given depth is needed, and the accuracy and details is not as important. Looking at Figure 4.13 it is obvious that for most cores there is a clear exponential attenuation effect on the layer thickness. This effect is visible for all different shift lengths, which might point to that, for this work at least, a very small shift is not necessary.

Interestingly, though, is that especially for Site E and Site G, there

are some depths where the λ_A estimates seem to be increasing instead of decreasing. Luckily for this work, these depth are not of interest, and the specific depth between the Laki and Tambora events seem to be stable at almost any shift length. Only site E presents some troubles where λ_A has presented itself difficult to estimate exactly at the Laki to Tambora depth. For this specific core it could therefore be necessary to decrease the shift length to be able to get more estimates in the Laki to Tambora area.

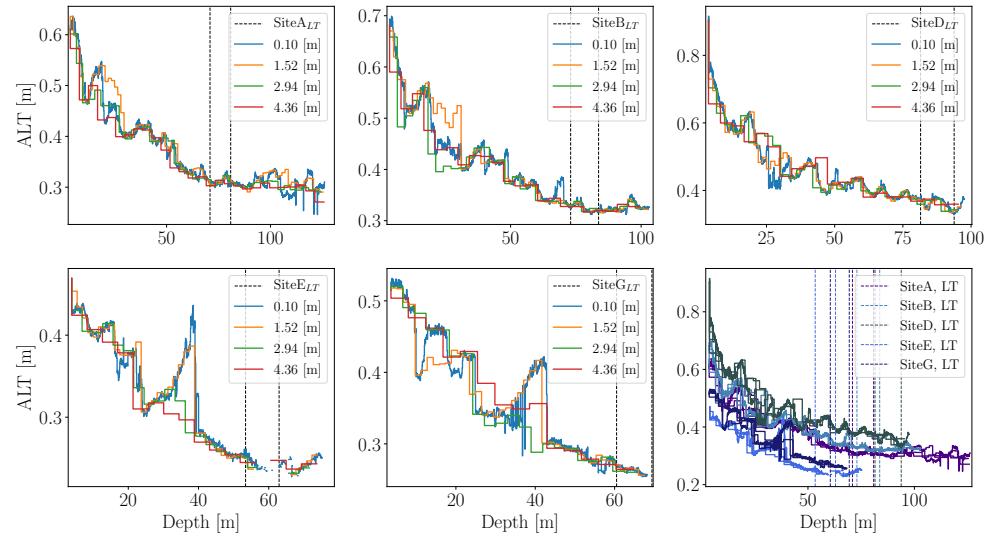


Figure 4.13: Examples of λ depth profiles given different shift lengths. All profiles are computed with a section length for spectral analysis of $l_{\text{sec}} = 5$ m.

In Figure 4.14 the λ_{LT} versus shift length can be seen, with shifts from 0.05 m to 5 m.

4.6 Splines and Interpolation

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 they are discrete samples of a continuous function. 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 4.15.

A spline is a piecewise polynomial of degree k that is continuously differentiable $k - 1$ times.

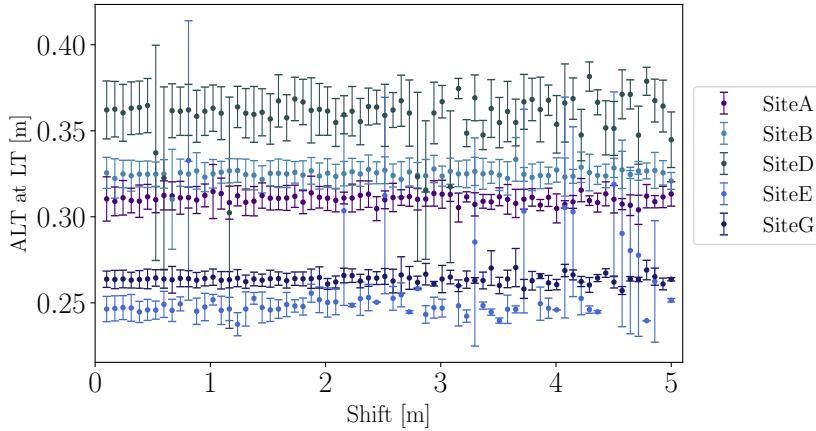


Figure 4.14: Annual layer thickness estimate at the depth section between Laki and Tambora, λ_{LT} versus different shifts of the sections used in the spectral analysis.

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, 4.2.1.1. 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.

4.6.1 Interpolation

Interpolation is a tool that can be used - and misused - to extract more information out of a given set of data. Used correctly, interpolation can reveal more information than is initially available and disclose connections not apparent at first, but used incorrectly, it can be manipulated to infer misleading correlations and lead to inaccurate conclusions. Thus it is a tool that must be used with care. Aiming to avoid incorrect deductions and inferences one should at first gain as much knowledge about the data at hand as possible. By understanding how the data have come about and gaining knowledge about the underlying physical theories a somewhat deficient data set can robustly and securely be interpolated to accommodate the needs of the analysis. In the case of this thesis, both knowledge about data gathering and the physics at

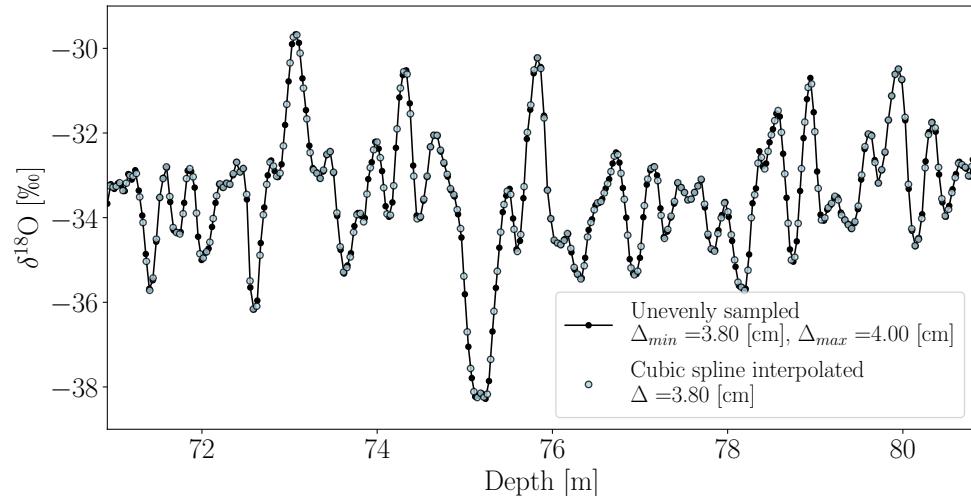


Figure 4.15: Unevenly sampled signal from Site A resampled using cubic spline interpolation to an even signal with a new sample size equal to the minimum sample size found in the raw signal.

play have been gained and thus some of the common fallacies may be avoided. The limits of the data available is due to the discrete sampling, leading to a minimum sampling of about 26 samples per meter of ice.

When considering that the depth series of 33 years between Tambora and Laki is just above 10 meters, this means that each meter of ice needs to contain at least three years on average. 26 samples per three years might not sound as a bad sampling interval, but if the goal is to show seasonality and give a best estimate of annual layer thickness, interpolation could be put to good use to be able to give better estimates of the exact placement of peaks and troughs.

4.6.2 Cubic Spline Interpolation in this Project

An in depth description of interpolation and splines can be found in Appendix 7. In general, these interpolation methods are implemented and examined in two particular sections of the analysis:

1. Cubic spline interpolation of raw, unevenly sampled data to represent even data, that can be analyzed through fast spectral transforms.
2. Cubic spline interpolation of the final back-diffused signal estimate to enhance resolution for more efficient peak detection.

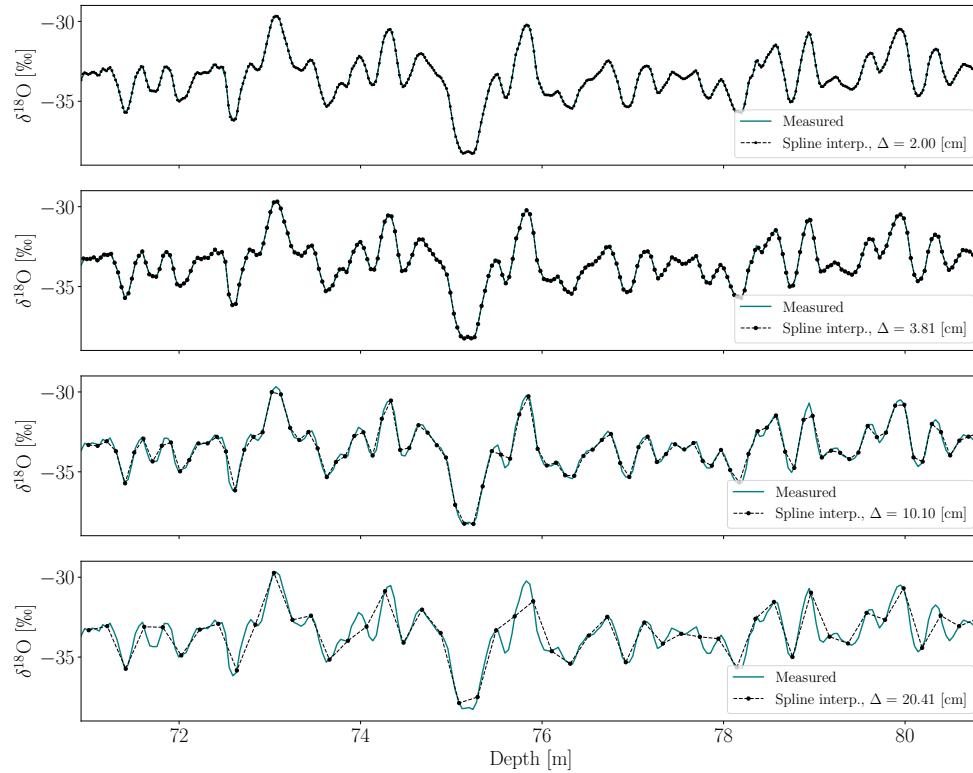


Figure 4.16: Four different resampled signals of Site A data, showing loss of information when resampling resolution is low.

Since the second interpolation is performed on data that has already been altered through the back diffusion and signal restoration process, it is the first interpolation, before back diffusion, that is the most at risk of introducing false information. To try and maintain the signal as true to its underlying samplings as possible, the first interpolation is only used if either of the uniform spectral transforms are chosen. If the interpolation is necessary, in the case of the uniform transforms, the interpolation is not chosen smaller than the smallest existing sample size in the depth section. Furthermore this interpolation is avoided when possible. This is done by primarily using the nonuniform DCT in the analyses, even though it is much slower. The effects of both interpolation methods are presented in the following chapter, in Section 5.2.2.

Chapter 5

Estimating σ from Data: Methods, Algorithms and Discussion

This chapter presents the final method and algorithm used to obtain an estimate of the optimal diffusion length, σ , for the depth sections under consideration. It consists specifically of two sections, one presenting a walk through of the constrained optimization σ estimation method, Section 5.1, and one containing a number of tests to examine the behaviour and stability of the method and algorithm, 5.2. Finally, a discussion of the further work that could be done to improve on the method is presented in Section 5.3.

Contents

5.1	σ Estimation Method	62
5.1.1	Module 1: Initialization and Back Diffusion	63
5.1.2	Module 2	66
5.1.3	Constrained Peak Detection	68
5.2	Testing and Stability	70
5.2.1	Constraints or No Constraints	70
5.2.2	Effects of Interpolations	71
5.2.3	Spectral Transforms	75
5.2.4	LT locations	78
5.3	Upgrades	79

5.1 σ Estimation Method

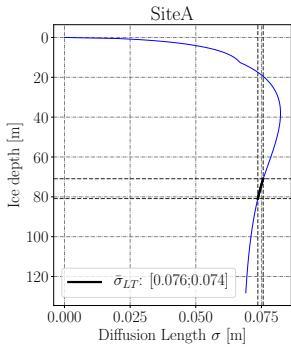


Figure 5.1: Diffusion length profile with section between Laki and Tambora highlighted.

As mentioned earlier, the diffusion length varies with depth, see Figure 5.1 for a modelled diffusion length profile for conditions at ice core drill site A. Though for the most of the final results, an assumption is made that the diffusion length along the depth section under consideration is constant, $\sigma(z) = \sigma$. This assumption is made based on the relatively small variation in diffusion length at the depths analyzed, especially compared to the uncertainty given on the final diffusion length estimates. Though this assumption is made, the subject of including a varying $\sigma(z)$ in the final analysis is presented and discussed.

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 Chapter 2.

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 5.2. 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 5.2 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 are inherent to the depth series alone, and these analyses are carried out as previously described in this thesis.

The first module could easily be used to scan the entire one dimensional σ space linearly, and from this analysis determine the optimal σ . This was first implemented, but turned out to be very slow - of course depending on the stepping size. But to speed up the search, a direct search method was

implemented.

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 5.6.

5.1.1 Module 1: Initialization and Back Diffusion

The first module of the algorithm, containing initialization and describing the general back diffusion method, is based on many of the aspects and models presented in the previous Chapters 2 and 4. Therefore the description of this module will focus on the work flow and not so much on the specific details of each process, as these have already been presented. The different chapters and sections containing the description of the methods used are referred to in the text of Figure 5.2.

The module takes an input of a measured depth series in a given interval, d , and the specifications concerning the drill site and the ice core in general. From here the work flow splits in two: one route(left) analyzing the depth series, and one(right) giving an estimate of σ at that depth, based on models.

Figures 5.3 and 5.4 illustrate the back diffusion with three different diffusion lengths, minimum and maximum theoretical estimate ($\sigma_{\text{Theo}}^{\min}$, $\sigma_{\text{Theo}}^{\max}$) and σ_{Fit} estimated from the signal's PSD, used in the Gaussian filter. The two Figures show the difference in stability for the signal. Site A exhibits a stronger reaction to a change in σ , rapidly increasing the amplitude and number of peaks counted, whereas Site B shows smaller variation in the final restored signal.

The right flow describes how the σ_0 estimate is given on the basis of the core specifications passed into the algorithm. First, a HL-density profile is modelled, based on the necessary input parameters of annual accumulation rate A_0 and drill site surface temperature T_0 , and the optional inputs of surface density ρ_0 and measured depth versus density data. Secondly, this density profile is used to compute a diffusion length profile, by the use of the Iso-CFM.

The modelled diffusion length profile is then used to find a theoretical

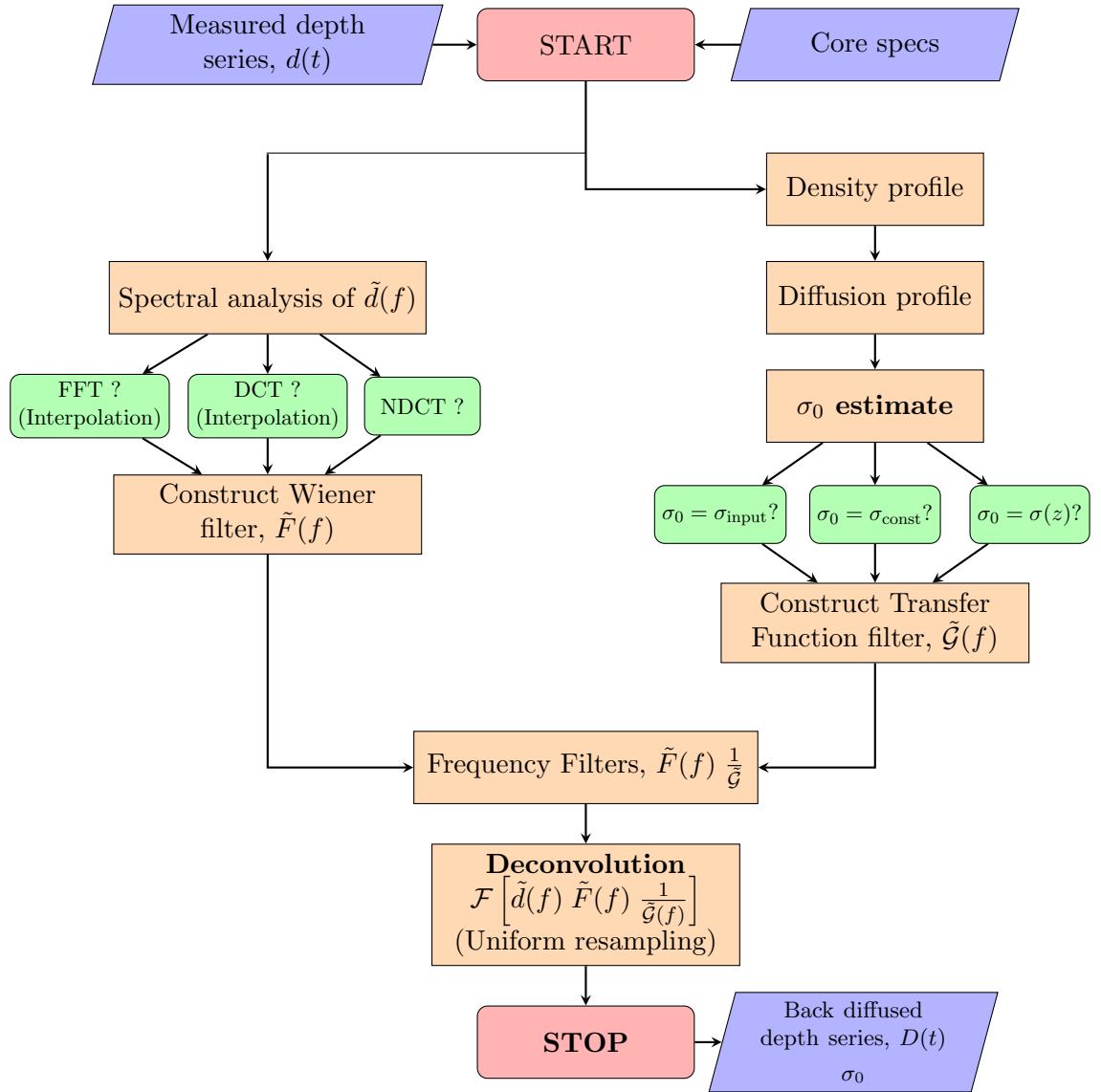


Figure 5.2: Flowchart of initialization method for back diffusion of a depth series given a diffusion length estimate. For the left part of the flow, the methods used are presented in the following sections: spectral transform and analysis is presented in Section 4.2, frequency filtering and deconvolution as means to back diffusion in Section 4.3, peak detection in Section 4.4, and spline interpolation in Section 4.6. For the right part of the flow, the theories and methods are presented in Section 2.2.

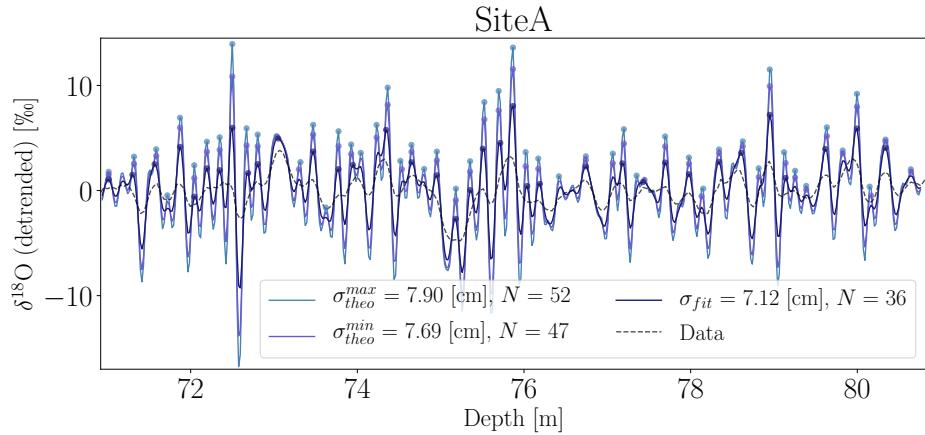


Figure 5.3: Backdiffused data for Site A, deconvolution with maximum and minimum σ_{Theo} , and σ_{Fit} estimated from the PSD of the signal.

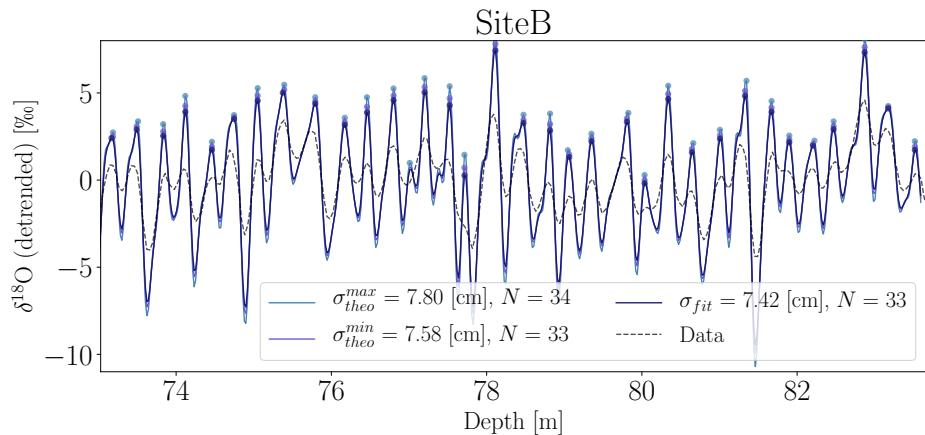


Figure 5.4: Backdiffused data for Site B, deconvolution with maximum and minimum σ_{Theo} , and σ_{Fit} estimated from the PSD of the signal.

diffusion length estimate which is used as an initial guess for the diffusion length, σ_0 . This is then used as the standard deviation in the Gaussian transfer function filter, \mathcal{G} used for deconvolution, unless a σ_0 is inputted manually. From the modelled diffusion length profile, it is also possible to choose not just a constant σ_0 , but an option for a depth-varying diffusion length, $\sigma(z)$, used in the transfer function is also available. This functionality is not analyzed in depth and would benefit from further analysis and testing.

If the varying $\sigma(z)$ is chosen, a fit is made to the theoretical diffusion length profile in the depth section corresponding to the Laki to Tambora depth. The back diffusion is then performed with a different Gaussian filter for each point, i.e. $\mathcal{G}(\sigma(z)) = \frac{1}{\sigma(z)\sqrt{2\pi}} e^{-\frac{z^2}{2\sigma^2}}$. $\sigma(z)$ is estimated from a fit to the theoretical diffusion length profile. In Figure 5.1 is shown the depth section of the diffusion length profile corresponding to that between Laki and Tambora. Figure 5.5 shows the polynomial fit made to this depth section.

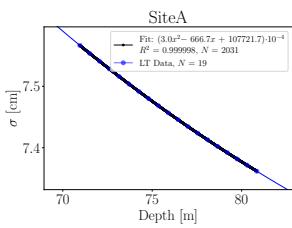


Figure 5.5: Diffusion length profile at Laki to Tambora depth, with $\sigma(z)$ fit.

The left flow shows the analysis carried out on the measured depth series, $d(t)$. First, the depth series is transformed to the frequency domain, with the user's choice of spectral transformation method. If FFT or DCT is chosen, the depth series will be interpolated to resemble equispaced data. Then the frequency series is analyzed according to Section 4.3, where fits to the signal and noise are made, that again are used to construct an optimal Wiener filter, $\tilde{F}(f)$. During this process, it is optional to choose the standard deviation of the estimated noise free signal, σ_{tot} , as the diffusion length to use in the deconvolution.

Finally, the right and the left flow are combined to construct the final restoration filter, $\tilde{R}(f) = \tilde{F}(f) \frac{1}{\tilde{\mathcal{G}}(f)}$, which is used for optimal enhancement and restoration when deconvoluting the signal to $D(t) = \mathcal{F} \left[\tilde{d}(f) \cdot \tilde{F}(f) \cdot \frac{1}{\tilde{\mathcal{G}}(f)} \right]$, and the module returns the back diffused series $D(t)$ and the diffusion length used, σ_0 .

5.1.2 Module 2: Estimating the Optimal σ

The second module contains the optimization routine to find the best estimate of σ to fulfill the constraints imposed, especially with the focus on reconstructing the signal to contain 33 peaks. It is based on a constrained direct search method which examines the one dimensional σ space. A flow chart of the module can be seen in Figure 5.6. The module takes three initialization parameters as input, the starting grid sizing Δ_σ , a small size ϵ and the number of grid points to examine N_σ , as well as the constraints imposed on the depth interval, described in the following section, 5.1.3.

The method is initialized by an input σ_0 estimate, the right flow in module 1, from either models, spectral analysis or manual input, which is used to create the first coarse $\bar{\sigma}$ grid. The method also carries out the left flow from module 1 as part of the initialization. These steps do not need to be repeated

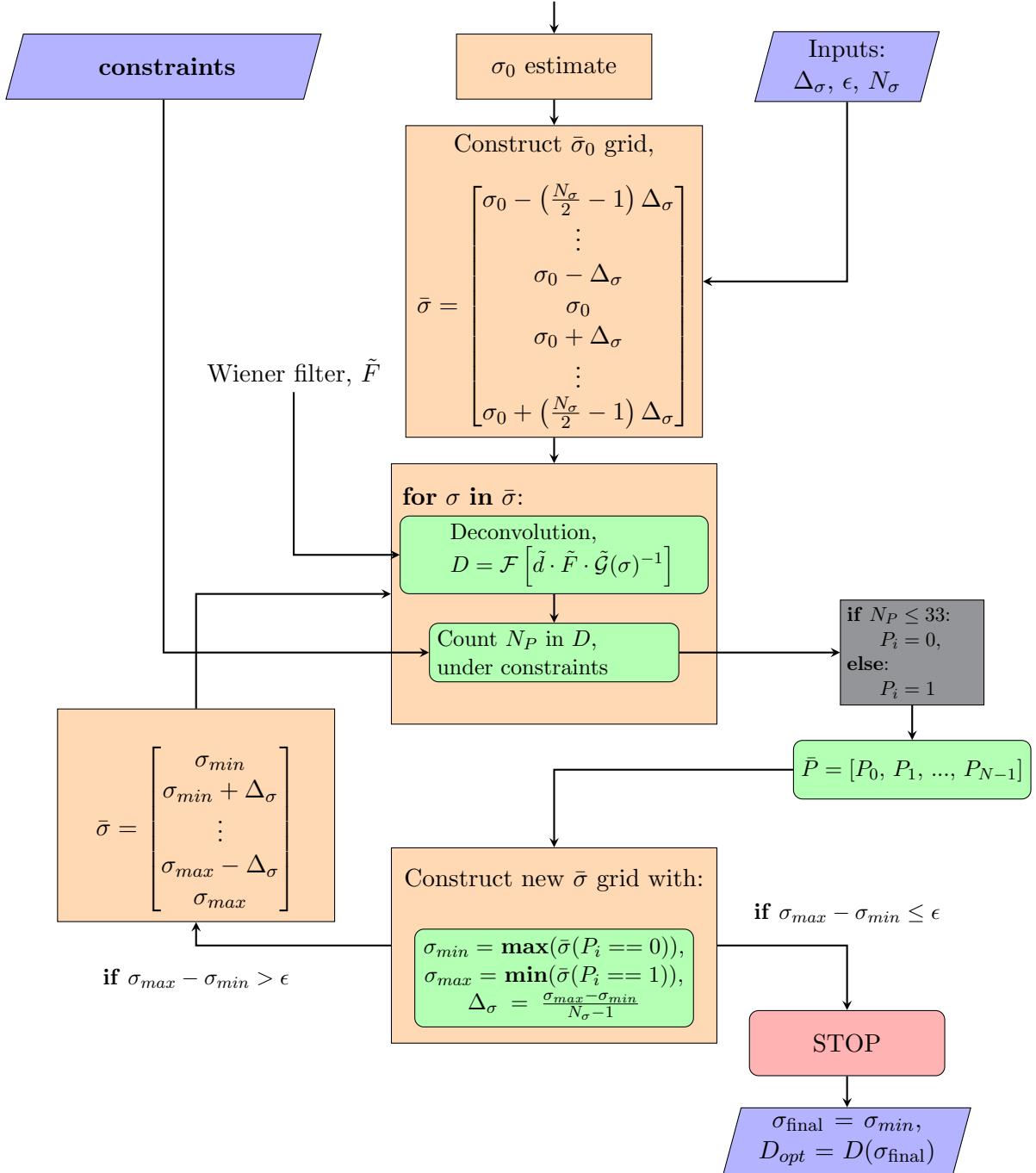


Figure 5.6: Flowchart of the module responsible for estimating the optimal σ , as described in Section 5.1.2. This module is continued from the initializations made in module 1, and contains the same deconvolution method, just utilized multiple times to search for the most optimal diffusion length, σ_{Opt} , instead of only using σ_0 .

again.

Then the module carries out the deconvolution method on the $\bar{\sigma}$ grid and uses the imposed constraints to count the number of peaks in the back diffused depth series, along with controlling that the imposed constraints are complied with. The algorithm then assigns a value of either 0 ($N_P \leq 33$) or 1 ($N_P > 33$) to the indices of a new vector \bar{P} . These values are then used to decide where to create the next search grid from, as it sets the new minimum diffusion length to $\sigma_{\min} = \max(\bar{\sigma}(P_i = 0))$ and the maximum to $\sigma_{\max} = \min(\bar{\sigma}(P_i = 1))$. This narrows down the search area and creates a finer grid. The $\sigma_{\max} - \sigma_{\min}$ test is then performed, and the deconvolution is performed once again, if $\sigma_{\max} - \sigma_{\min} > \epsilon$. If not, then the algorithm stops, presenting the final diffusion length as the last $\sigma_{\text{final}} = \sigma_{\min}$ and the final back diffused series as $D_{\text{opt}} = D(\sigma_{\text{final}})$.

The direct search method is very simple, and the main reason that this search method works is due to the investigation of the relation between number of counted peaks versus diffusion length. This relation was examined by brute force: the number of peaks versus diffusion length was computed manually from 0.01 m to 0.15 m. This showed clearly that - in the area of interest, i.e. resulting in $N_P = 33$ - the number of counted peaks increases as the diffusion length increases, see Figure 5.7.

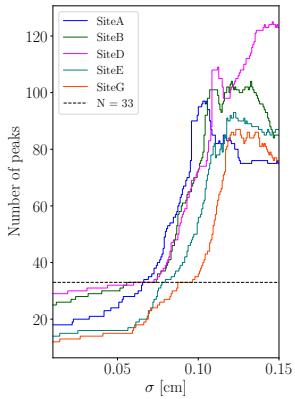


Figure 5.7: Number of peaks estimated versus diffusion length, based on diffusion lengths in the interval [0.01; 0.15] m.

The algorithm does not contain an optimized version for determining the optimal varying $\sigma(z)$, and if $\sigma(z)$ is chosen, the algorithm searches linearly through σ space, by moving the fitted $\sigma(z)$ through the diffusion length depth profile. This, along with the multiple deconvolutions with different Gaussian filters, makes this part of the algorithm very slow. This means that the further analysis does not focus on $\sigma(z)$, but future work could benefit from investigating this module further.

5.1.3 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 is fixed to $N_P = 33$, as this number of winters 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, due to the uncertainty on the exact time of deposition from the volcanic events, as discussed in Chapter 3.

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 is 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 able to detect a peak P , then a trough T , then a peak P , and so on, creating a pattern of ... $PTPTPTP\dots$. Since the deposition time of volcanic material in the ice is assumed to be Gaussian, the pattern is not restricted to starting 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 Table 5.1.

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

5.2 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) changes the resulting diffusion length estimate.

5.2.1 No constraints versus constraints

Figure 5.8 shows the depth series between Laki and Tambora events of the ice core drilled at Site B, back diffused through the algorithm described in the above. The difference between the blue and the green back diffused signals is that the blue is back diffused using the above presented constraints, and the green signal is back diffused with only the constraint $N_P = 33$. It is clear that the imposed constraints, especially the ones concerning peak distance and prominence, influence the final result of the optimal diffusion length. Particularly, the algorithm containing more constraints clearly leaves out some 'shoulders' before an actual peak in the final count, see Figure 5.8 at a depth of 77.5 m. These shoulders could be actual peaks but with the imposed constraints, they are omitted. This is something that could be further developed and examined. The appearance of these 'shoulders' could be peaks, but it is also likely that they are remnants of some noise effect that is not quite filtered in the frequency filter construction. If that is the case, then it is a positive thing that the algorithm filters them out.

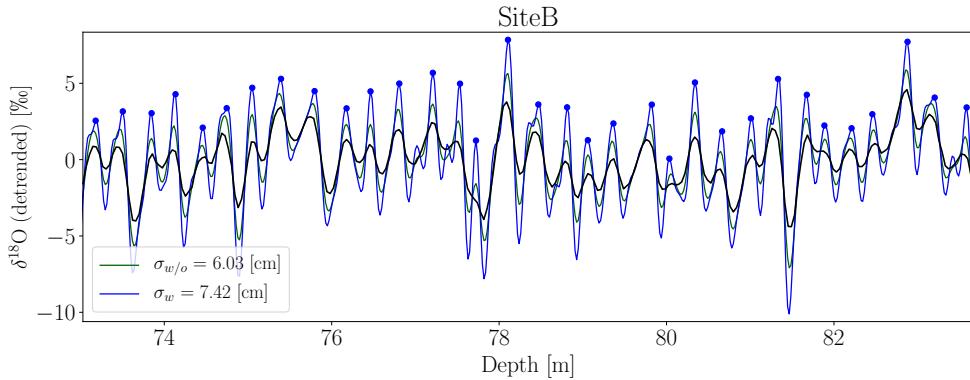


Figure 5.8: Example of how the imposed constraints affect the final diffusion length estimate for the Laki to Tambora depth section of the core drilled at Site B. The black line shows the data, the green the back diffused data using a method with less constraints, and the blue line shows the back diffused data when using the imposed constraints. The blue dots represent the peaks counted in the constrained method.

5.2.2 Effects of Interpolations

The choice of interpolation, before and after deconvolution, does affect the final result by introducing some effects not inherent in the originally measured signal. Therefore it is important to examine exactly how these interpolations influence the final σ estimate. Here that is examined by running the algorithm with different resampling sizes.

5.2.2.1 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

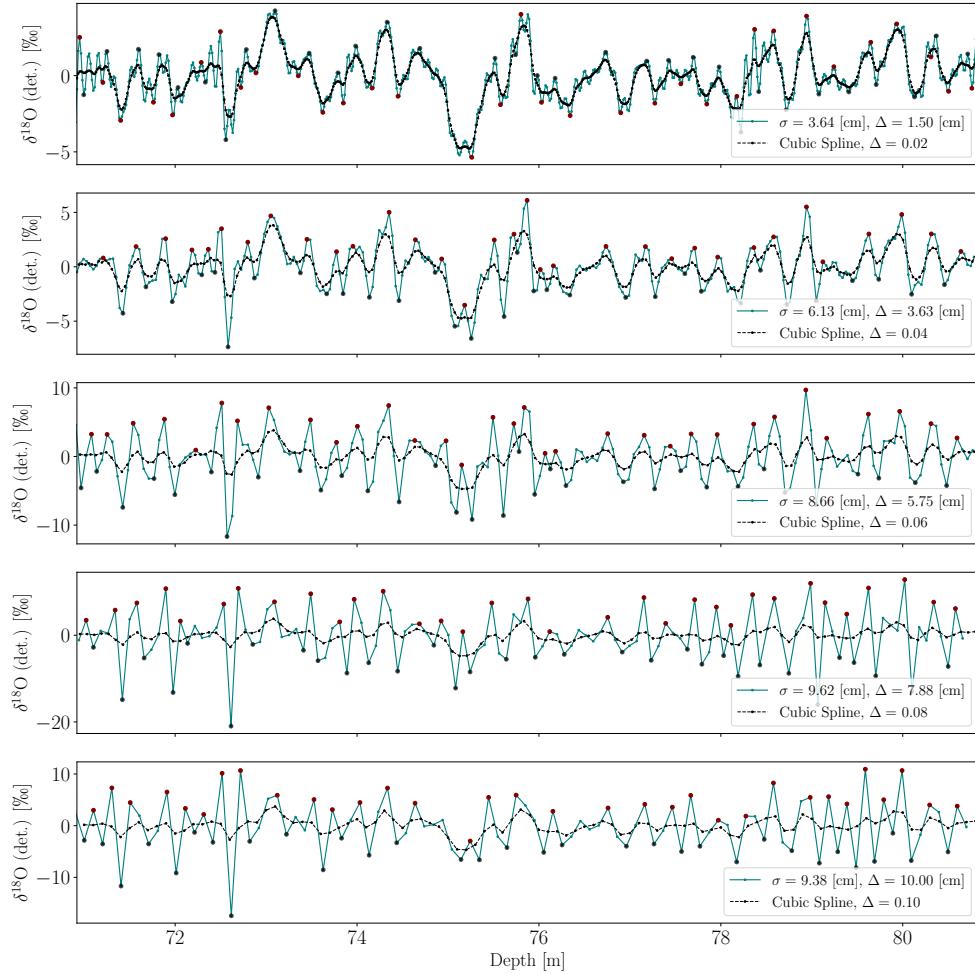


Figure 5.9: Site A, illustration of the effect of five different cubic spline resamplings before deconvolution. Original sample sizes lie in the interval [3.80; 4.00] [cm]. Resampling at smaller sample sizes show a tendency to restore some signal frequencies that are not necessarily inherent to the original signal. The resample should thus not be chosen too small as this would introduce some false signal into the results.

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 4.15 and Figure 4.16. Figure 4.16 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 $[\Delta_z^{\min}; \Delta_z^{\max}]$. 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 5.2 and an illustration of the test results can be seen in Figure 5.10.

Figure 5.10 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 small variations that the spline interpolation introduces to the signal, as can be seen in the first panel in Figure 5.9. Furthermore, Figure 5.10 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 need more enhancement by cranking up the diffusion length estimate, see final panel in Figure 5.9. 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_z^{\min} .

5.2.2.2 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 whose 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

Site	Δ_z^{\min} [cm]	Δ_z^{\max} [cm]	Δ_z^{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 5.2: 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_z^{\min} and Δ_z^{\max} .

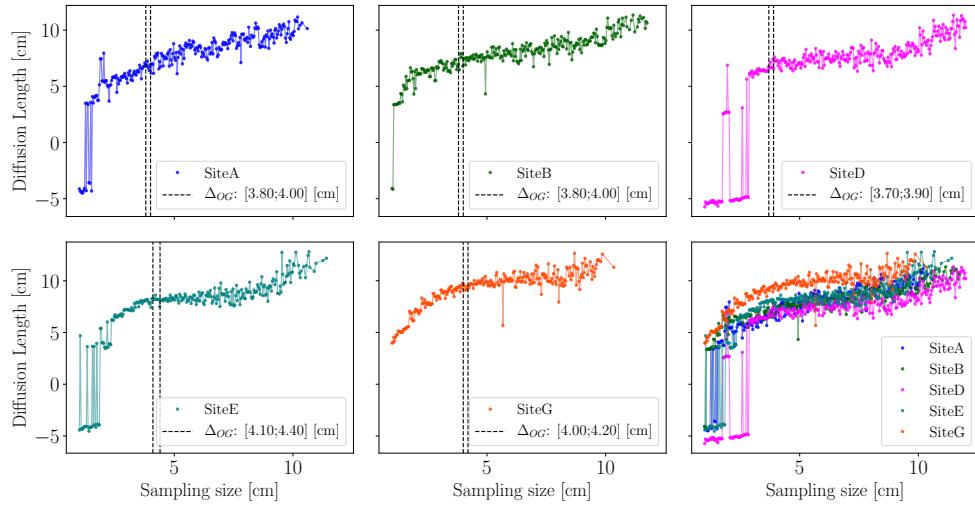


Figure 5.10: Diffusion length estimates versus resamples through cubic spline interpolation before deconvolution for Alphabet cores from sites A, B, D, E and G.

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 5.11. The effect of resampling after deconvolution on the final diffusion length estimate is illustrated in Figure 5.11.

As the actual isotopic signal is continuous and 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. In Figure 5.12 the diffusion length estimate versus the resampling size after deconvolution can be seen, and the data show 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. Based on these observations, a resample size of $\Delta_z^{\min}/2$ is chosen for interpolation after back diffusion and before peak detection. It is a compromise between choosing an interpolation that results in a higher resolution, but doing so without

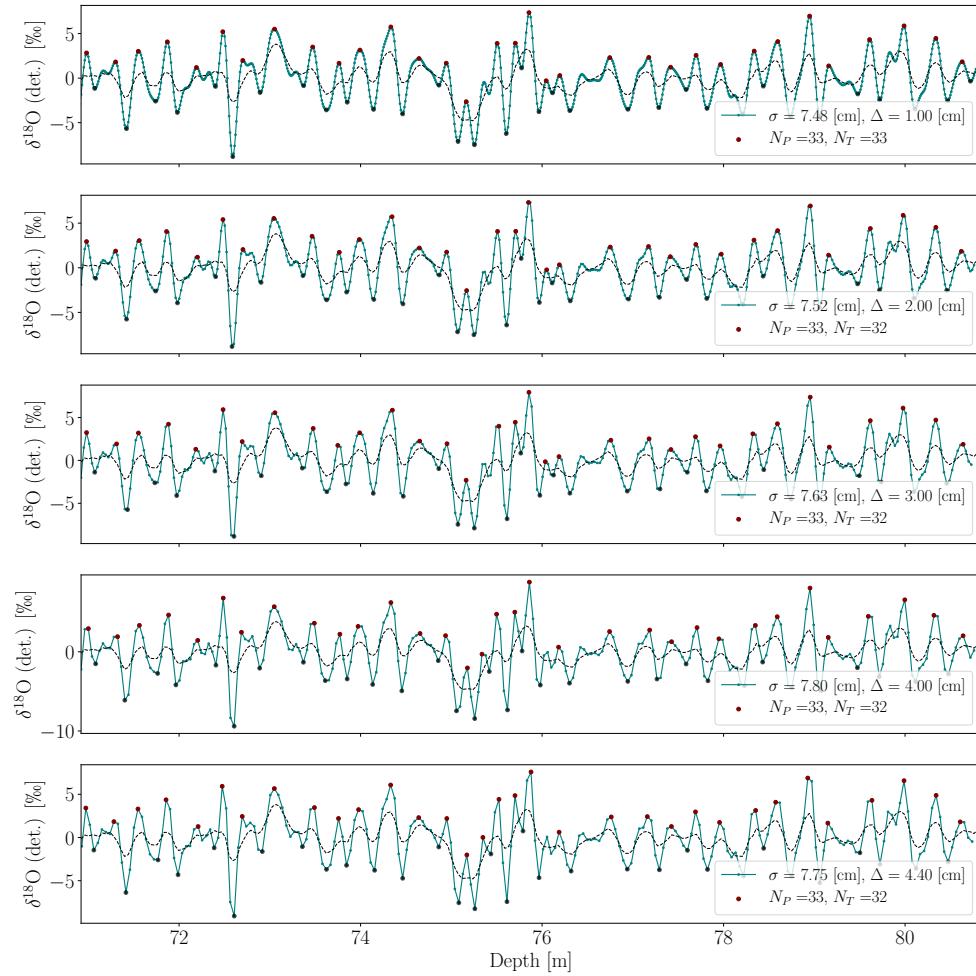


Figure 5.11: Site A, effect of cubic spline interpolation after signal has been deconvoluted. The interpolation is introduced to make peak detection more stable.

slowing down the algorithm too much.

5.2.3 Spectral Transform's Effect on Diffusion Length

The different ways of performing spectral transformation also influence the final σ estimate, not only due to the transformation itself, but for FFT and DCT, the methods demand an interpolation, which in itself influences the results.

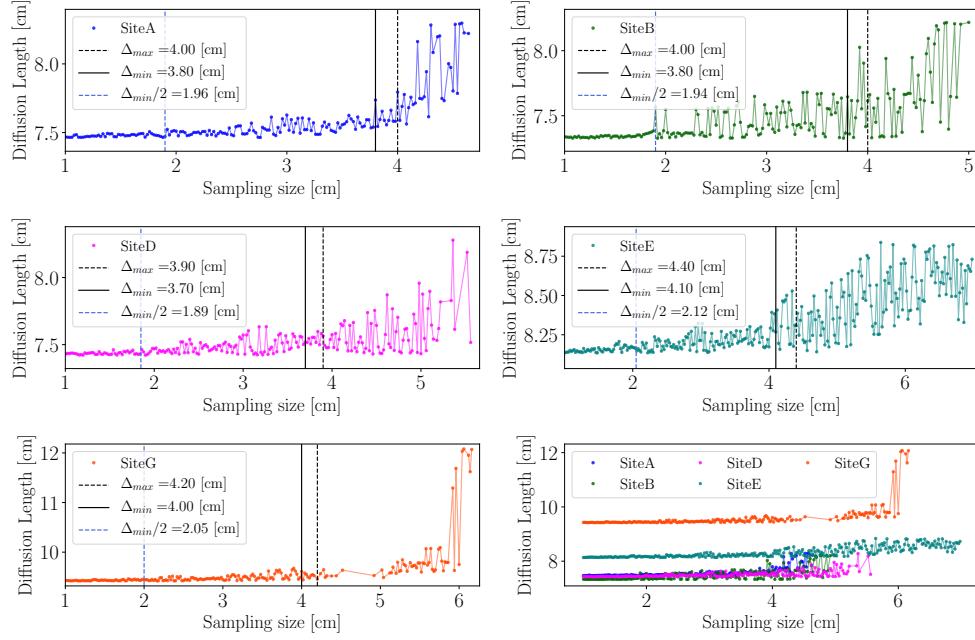


Figure 5.12: Final diffusion length estimate, given new resample by cubic spline interpolation after deconvolution for Alphabet cores from sites A, B, D, E and G. The original sample size interval is illustrated as black vertical lines.

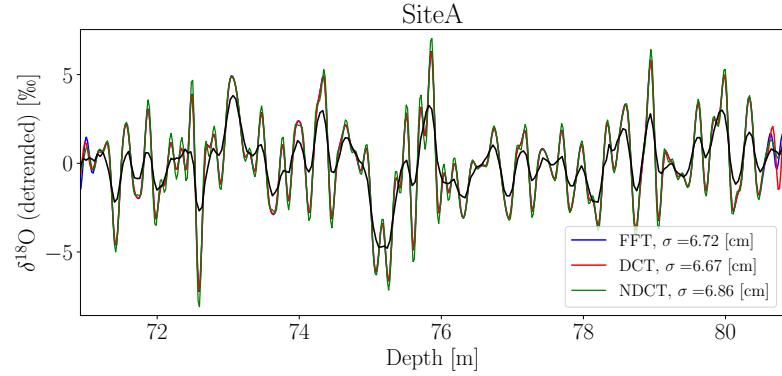


Figure 5.13: A visual example of the differences in final back diffused data when using different spectral transforms.

Figure 5.13 shows an example of the qualitative differences between using FFT, DCT or NDCT for spectral analysis. The most obvious visual

difference between the transforms is in the end sections of the interval. This might be due to some specific boundary conditions imposed on the fast Fourier and Cosine transforms. Furthermore, by careful visual inspection, it can be seen that the NDCT seems to cater to some effects of the nonuniform samplings, that the FFT and DCT do not.

There are no significant differences between the three spectral transforms, but there is some small variations. What might also be taken into consideration is, that this method might be run a large number of times - as is the case later on in this thesis - for example to estimate accuracy of the method, and one might therefore want to choose the method that gives the fastest results. Thus, the speed of the different transforms has been tested, as can be seen in Figure 5.14, with 200 separate runs where the Laki and Tambora positions have been drawn from a distribution with a standard deviation of 2 months from the estimated location. Not surprisingly, the NDCT is much slower than the DCT and the FFT, and it might prove efficient to choose either DCT or FFT if the algorithm has to run many times. For the case of this thesis, the final results have been made with the NDCT, so as not to miss any of the effects that might come from unevenly sampled data.

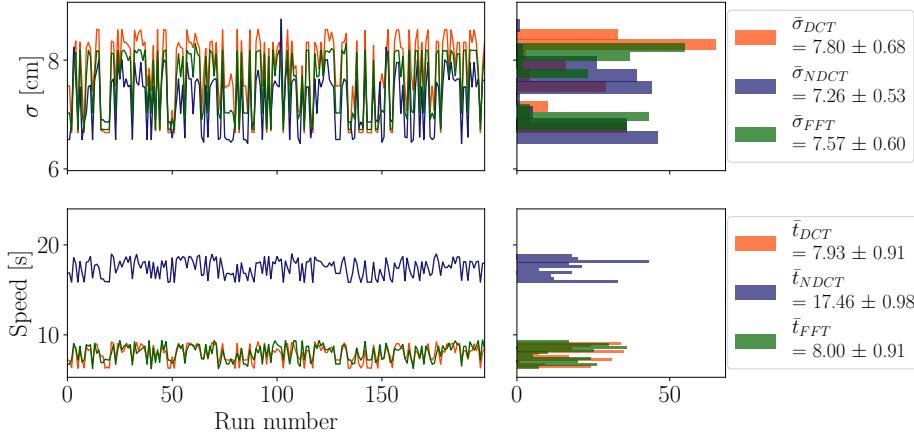


Figure 5.14: Diffusion length estimates along with speed of algorithm given the three different spectral transforms examined in this work. The volcanic event depths have been drawn from a Gaussian distribution with a width corresponding to ~ 1 month around the estimated mean depth of the event.

5.2.4 Laki and Tambora as Gaussian Distributions

As mentioned previously, the Laki and Tambora event depth locations are not exact. Thus to accommodate error in this positioning, the algorithm has been tested with Laki and Tambora locations drawn from Gaussian distributions. This was examined in four ways:

- Variation in both Laki and Tambora position, corresponding to the entire events.
- Variation in only the Tambora position while keeping a fixed section length, corresponding to the mean value $\bar{d}_L - \bar{d}_T$.
- Variation in only the Laki position while keeping a fixed section length, corresponding to the mean value $\bar{d}_L - \bar{d}_T$.
- Variation corresponding to a Gaussian distribution with a standard deviation of a depth that resembles to months in time.

The first variation of both Laki and Tambora can be seen in Figure 5.15. The method has been run 500 times with new locations drawn for each run, and the same constraints imposed each time. This results in an estimate of the diffusion length for Site A between Laki and Tambora of $\sigma = 7.32 \pm 0.67$ [cm]. But considering the lengths of the volcanic events, as seen in the ECM data, one might question this result. Due to more or less time incorporated as the depth section is widened or narrowed, some of the constraints might not be very well chosen any more. It might be that a larger depth interval could gain an extra peak or trough, so that the constraint should be $N_P = 34$ instead of $N_P = 33$. This is some work that could be continued on the algorithm, ensuring that the extra interval length increase or decrease loosens the count constraint.

The next method examines what happens when only changing the Laki or Tambora position. The results for 500 runs can be seen in Figure 5.17 and 5.16. An interesting feature shows here, and is also visible in some of the later results: the diffusion length estimates seem to not be Gaussian distributed, but concentrated around two or more different diffusion lengths. This could be an effect of the direct search algorithm, which might quantize the possible σ estimates when creating the grid. This could be examined by trying to randomize the grid creation. Furthermore, it would be interesting to

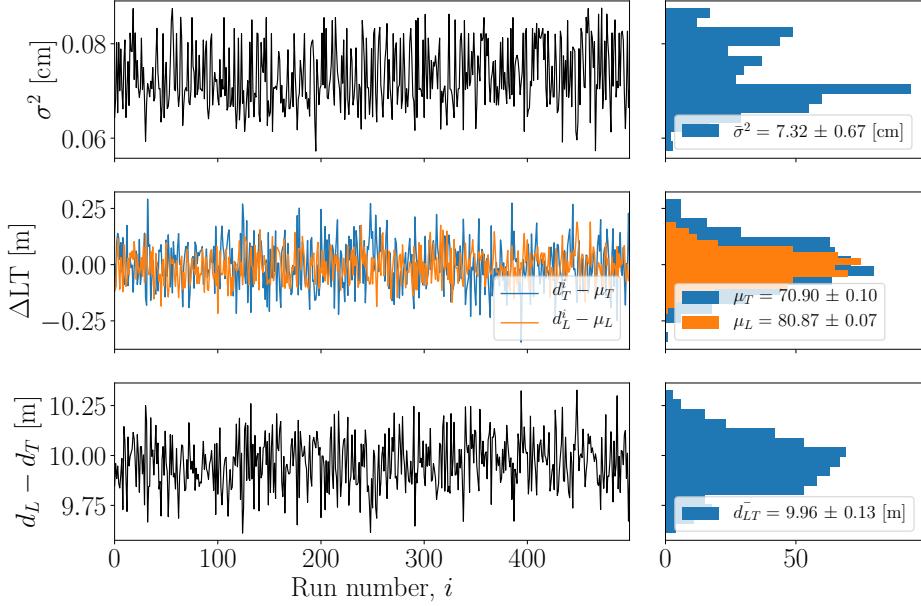


Figure 5.15: Diffusion length estimates when varying the depth locations of the volcanic events. The locations of both Laki and Tambora events have been drawn from Gaussian distributions as the ones presented in Section 3.1.

examine if there is a correlation between the positioning of the depth interval and the estimated diffusion length.

The final method is an investigation of drawing the locations of both Laki and Tambora from distributions with a standard deviation of what corresponds to two months and a mean value of where the middle of the volcanic event is estimated to be. An illustration of how much this is in depth is shown in Figure 5.18. The results can be seen in Figure ???. Again, the possible effect of the quantized grid search can be seen.

5.3 Possible Algorithm Upgrades

In future work, a number of different areas in the method and the algorithm could be upgraded and continuously developed. Some of these further work points will be presented and discussed here.

The peak detection method in the optimization module is, as men-

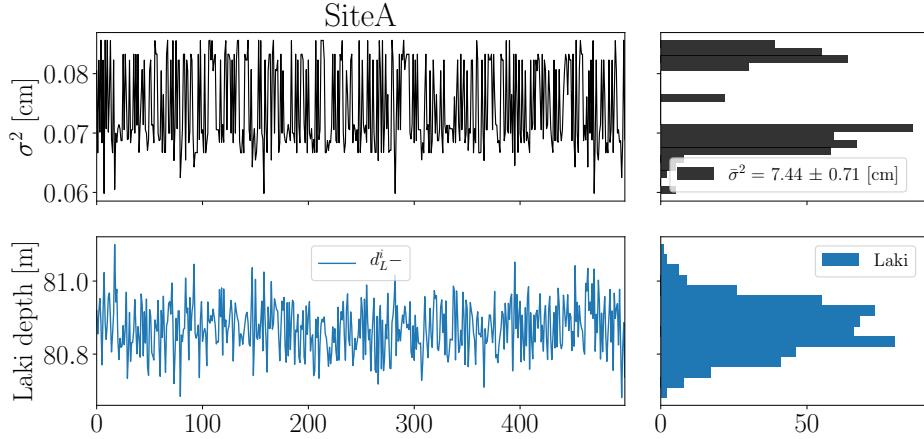


Figure 5.16: Diffusion length estimates for Site A when varying only the Laki volcanic event. The locations are drawn from Gaussian distributions as the ones presented in Section 3.1. The depth section is kept at a constant length, corresponding to the mean distance value, $\bar{d}_{\text{Laki}} - \bar{d}_{\text{Tambora}}$.

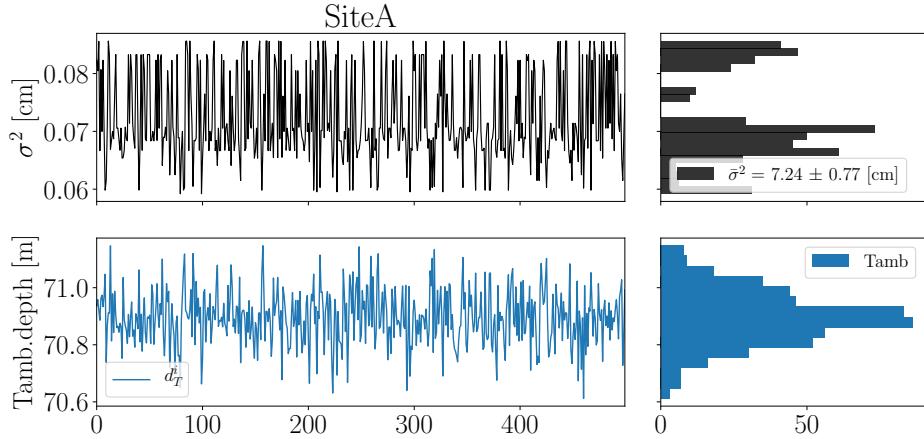


Figure 5.17: Diffusion length estimates for Site A when varying only the Tambora volcanic event. The locations are drawn from Gaussian distributions as the ones presented in Section 3.1. The depth section is kept at a constant length, corresponding to the mean distance value, $\bar{d}_{\text{Laki}} - \bar{d}_{\text{Tambora}}$.

tioned, very simply developed. There are ways that this method could be improved, for example by taking the entire yearly pattern of winter(peak)-spring-summer(trough)-fall (PSTF) into account when detecting the pattern.

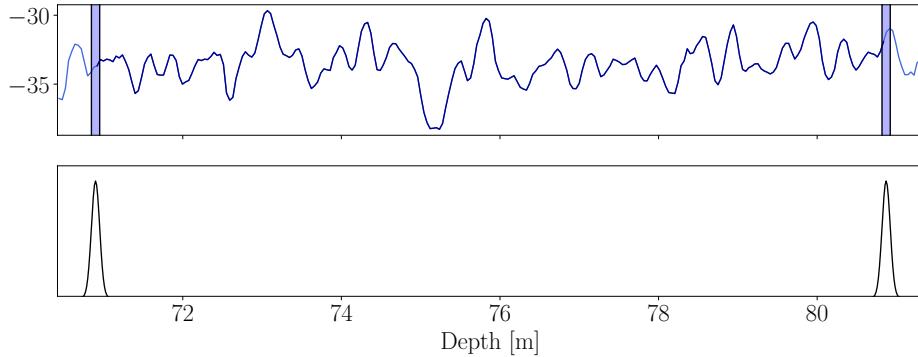


Figure 5.18: Illustration of the method implemented to manage the uncertainty of the exact depth location of the volcanic events. The method establishes a Gaussian distribution with a mean of the estimated middle of the volcanic event and a standard deviation of what corresponds to two months.

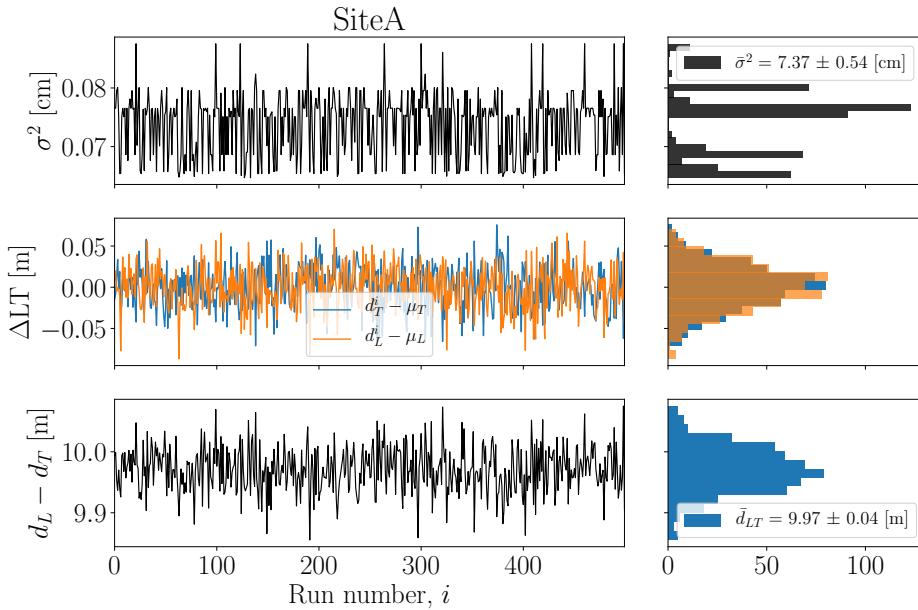


Figure 5.19: 500 runs with locations of Laki and Tambora events drawn from a Gaussian distribution with a standard deviation of two months, using NDCT as spectral transform.

This could possibly be done by investigating the derivative in the signal, and assigning S and F values to the steepest part between a P and a T. There is

also a possibility to develop a more sophisticated and maybe intelligent pattern recognition algorithm to detect yearly cycles in the signals. The pattern recognition could be based on the undiffused isotopic signals in the shallower parts of different ice cores. This would improve the certainty of detecting the right number of years in the section. By implementing more seasons, it also allows for more precise dating.

Considering the optimization routine, some issues have already presented themselves, through the distributions of the final results. These quantizations of the final results might be an effect of the search method, and could maybe be avoided by introducing a factor of randomness into the grid search. Furthermore, when shifting the estimated volcanic event positions to generate uncertainties, the optimization should also take into account if the constraints should be changed. If the locations add or subtract length from the signal, the number of peaks and troughs should also be considered to be raised or lowered. This could be worked in with the more detailed pattern recognition.

The frequency filter construction could also be investigated further. It might be allowing too much noise, or filtering some cycles away that were actually interesting. The construction could be made variable and work with the optimization module to tune and change as the constraints do. Some of these issues could be dealt with by gaining a higher resolution from the data by increasing the sampling rate in future measurements - interpolation can only get us so far without a large risk of misinterpretation.

This work has focused on estimating the maximal diffusion length that fulfills the imposed constraints. Another focus point could be on investigating all σ that fulfill the constraints. This could give a temperature range estimate instead of just one temperature estimate. Furthermore, the stability of the method at a given section might be investigated through this as well, as a section that only allows for a very limited σ range might actually show that the imposed constraints are not valid.

Furthermore, an in depth investigation and implementation of working with a variable diffusion length, $\sigma(z)$, instead of a constant σ_{const} could present a more accurate estimate of the final $\sigma_{\text{Opt}}(z)$ for the entire depth section, and not just an average. If this was implemented further, it might allow for a more thorough analysis of the variation in temperature in the depth section, instead of only examining the mean temperature of the time frame. Implementing $\sigma(z)$ could also allow for a better estimate of σ_{firn} , if the samples are unevenly

sampled, as the firn diffusion length estimate then could be made individually point by point based on both sample size and diffusion length estimate at that point.

Chapter 6

Results

This chapter contains the final results of the stability tests presented in Chapter 5 along with the final diffusion length estimates, based on the results of the stability tests and general considerations and observations made throughout the thesis. Finally these σ estimates are used to give a first estimate of the temperature for the given depth interval.

Contents

6.1	Annual Layer Thickness	85
6.2	Diffusion Lengths	86
6.2.1	σ Constraints or No constraints	86
6.2.2	Diffusion Length V N Peaks	86
6.2.3	σ v. Spectral Transforms	87
6.2.4	σ Constant or Variable	88
6.2.5	Final σ Estimates	90
6.3	Temperature Estimates from Data	90
6.3.1	Iso-CFM Possibilities	91

6.1 Annual Layer Thickness

The final annual layer thickness estimates are determined with a section length of $l_{\text{sec}} = 7$ m and a shift of $l_{\text{shift}} = 1.5$ m. From Table 6.1 it can be seen that λ is smallest for Site E and Site G, which corresponds well with those sites having the lowest accumulation rates.

Site	λ_{LT} [m]
A	0.311 ± 0.006
B	0.326 ± 0.008
D	0.354 ± 0.012
E	0.246 ± 0.005
G	0.264 ± 0.006

Table 6.1: Spectral analysis estimated annual layer thickness at the depth sections between Laki and Tambora.

6.2 Diffusion Length Estimates

Site	$\bar{\Delta}_z$ [cm]
A	3.92 ± 0.06
B	3.89 ± 0.06
D	3.77 ± 0.06
E	4.25 ± 0.07
G	4.10 ± 0.07

Table 6.2: Average sampling size in the depth between Laki and Tambora events.

Before revealing the final σ estimates, the results of the stability tests carried out are presented. This concerns the effects on the diffusion length by the different spectral transforms, the constant or variable σ , and using constraints or not. The results are presented with both the optimal found diffusion length, the afterwards corrected firn diffusion length and, for some tests, the average run time for the algorithm. In Table 6.2 the average sample sizes in the Laki to Tambora depth interval used for σ_{firn} estimates can be seen.

6.2.1 Diffusion Length Estimates if Constrained or Not Constrained

The previous chapter presented the constrained method for optimization. The results of both constrained and unconstrained optimization can be seen in Table 6.3. If the method does not impose constraints on the algorithm, the algorithm is not quite as stable, as can be seen on the average run time and the run time variances. Furthermore, the unconstrained method systematically results in a diffusion length estimate lower than the one computed through the constrained method, and generally with a higher variance.

6.2.2 Diffusion Length Estimates vs. Counted Peaks

To get an understanding of how the diffusion length affected the number of counted years, a run was made going through all diffusion lengths from 1 cm to 15 cm for all cores. Illustrations of the results can be seen in Figures 6.1 and 6.2. In Figure 6.1 it can be seen that the number of counted peaks generally increases as the diffusion length increases, until some value where the counts start being more irregular. When looking at Figure 6.2, it can be seen that for most of the cores, there exists a stable plateau of diffusion lengths that all result in $N_P = 33$, for some longer than for others. This could be used as a stability measure and a sanity check of the number of years assumed is actually likely in this section.

		No Constraints	Constraints
Site A	σ_{opt} [cm]	6.69 ± 0.85	7.88 ± 0.66
	σ_{firn} [cm]	6.58 ± 0.87	7.79 ± 0.66
	t [s]	19.85 ± 10.39	8.84 ± 1.47
Site B	σ_{opt} [cm]	5.98 ± 0.24	7.31 ± 0.18
	σ_{firn} [cm]	5.86 ± 0.24	7.21 ± 0.18
	t [s]	11.66 ± 4.46	8.68 ± 1.08
Site D	σ_{opt} [cm]	4.35 ± 0.53	6.94 ± 0.24
	σ_{firn} [cm]	4.24 ± 0.32	6.84 ± 0.24
	t [s]	14.83 ± 8.50	9.19 ± 0.72
Site E	σ_{opt} [cm]	6.20 ± 0.20	8.15 ± 0.11
	σ_{firn} [cm]	6.07 ± 0.21	8.05 ± 0.11
	t [s]	5.44 ± 1.82	6.97 ± 0.56
Site G	σ_{opt} [cm]	8.55 ± 0.27	9.35 ± 0.26
	σ_{firn} [cm]	8.46 ± 0.28	9.27 ± 0.26
	t [s]	29.99 ± 3.81	7.19 ± 0.48

Table 6.3: Optimal and corrected firn diffusion length estimates with either the non-constrained or the constrained method.

6.2.3 Diffusion Length Estimates vs. Spectral Transform Methods

In Table 6.4 the estimated diffusion lengths, both optimal and firn, are presented along with the average run time of the method. The error is estimated through drawing the Laki and Tambora event locations from a Gaussian distribution with standard deviation of two months. For all sites, the results of all three methods are almost within each other's margins. This might point to choosing one of the faster back diffusion methods to estimate the diffusion length with, if it is used for a quantitative measure. If it is needed to reconstruct a single data series one time, it might be preferable to use the NDCT.

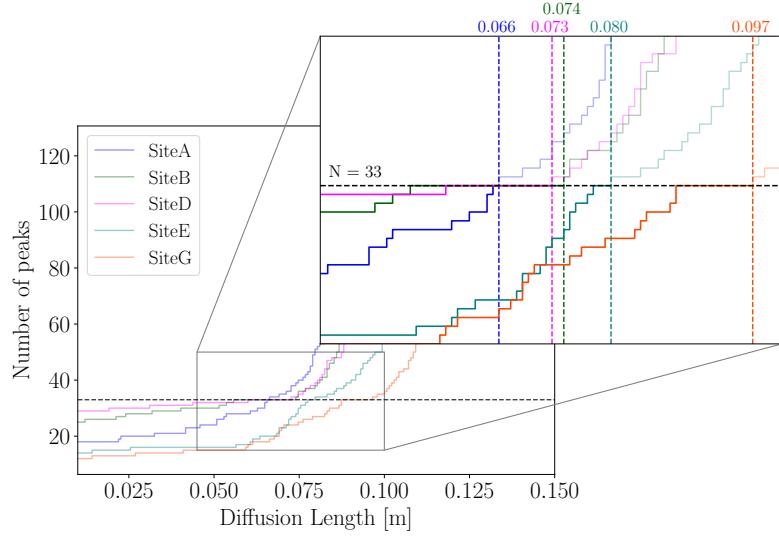


Figure 6.2: A zoom-in of the N peaks v. diffusion length plot in Figure 6.1. Specifically in focus are the maximal diffusion lengths corresponding to $N_{\text{peaks}} = 33$.

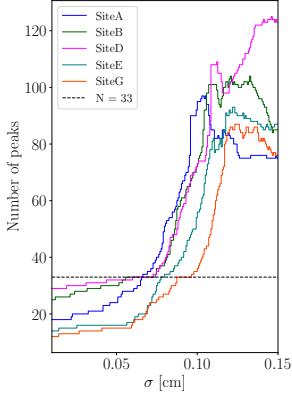


Figure 6.1: Number of peaks estimated given diffusion length, based on diffusion length in the interval $[0.01; 0.15]$ m.

6.2.4 Diffusion Length Estimates if Diffusion Length Constant or Variable

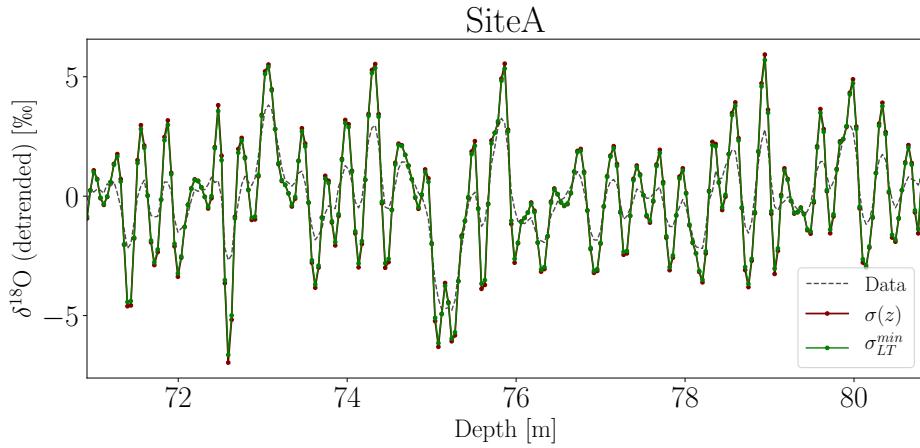


Figure 6.3: Backdiffused data, deconvolution with $\sigma_{\text{constant}} = \sigma_{LT}^{\min}$ and $\sigma(z)$.

Deconvolution with a variable diffusion length has not been quantitatively examined, but a qualitative example of the difference between deconvolution with a constant or variable Gaussian filter can be seen in Figure 6.3. It

			FFT	DCT	NDCT
Site A	σ_{opt}	[cm]	7.57 ± 0.60	7.80 ± 0.68	7.26 ± 0.53
	σ_{firn}	[cm]	7.48 ± 0.61	7.71 ± 0.69	7.17 ± 0.54
	t	[s]	8.00 ± 0.91	7.93 ± 0.91	17.46 ± 0.98
Site B	σ_{opt}	[cm]	7.11 ± 0.40	7.30 ± 0.20	7.36 ± 0.21
	σ_{firn}	[cm]	7.01 ± 0.40	7.21 ± 0.20	7.27 ± 0.22
	t	[s]	8.64 ± 0.51	8.41 ± 0.82	19.06 ± 0.55
Site D	σ_{opt}	[cm]	7.00 ± 0.41	6.96 ± 0.28	7.21 ± 0.27
	σ_{firn}	[cm]	6.91 ± 0.41	6.86 ± 0.29	7.12 ± 0.27
	t	[s]	9.24 ± 0.73	9.20 ± 0.69	19.55 ± 1.03
Site E	σ_{opt}	[cm]	8.07 ± 0.01	8.15 ± 0.11	8.21 ± 0.14
	σ_{firn}	[cm]	7.97 ± 0.01	8.05 ± 0.11	8.11 ± 0.14
	t	[s]	7.28 ± 0.36	7.03 ± 0.56	16.61 ± 0.54
Site G	σ_{opt}	[cm]	9.38 ± 0.32	9.35 ± 0.25	9.44 ± 0.24
	σ_{firn}	[cm]	9.29 ± 0.32	9.27 ± 0.25	9.36 ± 0.24
	t	[s]	7.53 ± 0.49	7.24 ± 0.49	16.45 ± 0.34

Table 6.4: Diffusion length estimates resulting in $N_{\text{peaks}} = 33$ based on different spectral transform methods, namely the FFT, DCT and NDCT presented in earlier chapters and described in Appendix 7. Along with the optimal diffusion length, the actual firn diffusion length is presented - corrected for sampling diffusion, ice diffusion and thinning. The computational time of the back diffusion process given the different spectral transforms is also presented.

is mostly presented here to emphasize that using either a constant or a variable σ does not result in a large difference in the restored signal. The change in theoretically estimated $\sigma(z)$ at the depths under consideration here is relatively small, resulting in a small difference from the constant σ . It is thus concluded that the speed is prioritized in the σ estimations and the variable $\sigma(z)$ method is abandoned.

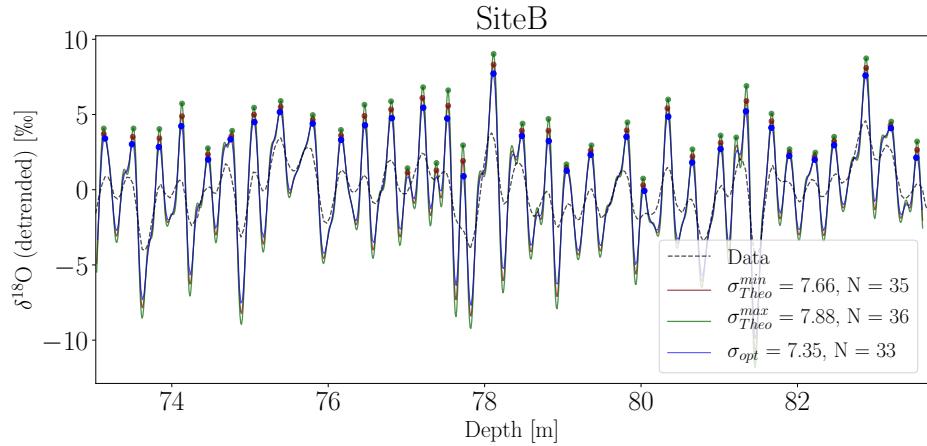


Figure 6.4: Example of back diffused data from Site A, both deconvoluted with maximum and minimum expected σ and with the final σ_{opt} .

6.2.5 Final σ Estimates Based on Previous Conclusions

Figure 6.4 shows the optimal diffusion length estimate resulting in $N_P = 33$ for Site B, along with the restorations using the theoretical diffusion length estimates, σ_{min} and σ_{max} , in the Gaussian filter. The theoretical estimates are not resulting in an unlikely number of peaks, but the optimization algorithm still estimates a lower diffusion length, resulting in the wanted 33 peaks.

In Table 6.5 the final optimized diffusion length estimates can be seen, along with the theoretical diffusion length interval and the σ estimated from the fit to the PSD of the signal.

6.3 Final Temperature Estimates from Optimal Estimated σ

The temperature estimates are, as described previously, made based on a steady state model, with a constant accumulation rate. The estimate is then made numerically through a Newton-Raphson scheme from `scipy.optimize.newton`, finding the roots of $\sigma_{\text{model}} \cdot \frac{\rho_{\text{CO}}}{\rho_{\text{ice}}} - \sigma_{\text{data}}$.

The final temperature estimates can be seen in Table 6.6 and the underlying distributions in Figure 6.5. Again in these distributions can the quantization of the results be seen, resulting in non-Gaussian distributions.

		σ_{final}	σ_{Fit}	σ_{Theo}
Site A	σ_{opt} [cm]	7.37 ± 0.54	7.12	—
	σ_{firn} [cm]	7.27 ± 0.55	—	[7.69; 7.90]
Site B	σ_{opt} [cm]	7.35 ± 0.22	7.42	—
	σ_{firn} [cm]	7.26 ± 0.22	—	[7.58; 7.80]
Site D	σ_{opt} [cm]	7.21 ± 0.28	7.56	—
	σ_{firn} [cm]	7.12 ± 0.28	—	[7.36; 7.54]
Site E	σ_{opt} [cm]	8.22 ± 0.15	7.81	—
	σ_{firn} [cm]	8.12 ± 0.15	—	[8.85; 9.13]
Site G	σ_{opt} [cm]	9.46 ± 0.24	8.21	—
	σ_{firn} [cm]	9.38 ± 0.24	—	[8.85; 9.10]

Table 6.5: Final diffusion length estimates, based on conclusions made previously in different tests, along with theoretical and fit estimates.

	Site A	Site B	Site D	Site E	Site G
T_0 [°C]	-29.41	-29.77	-28.3	-30.37	-30.1
$\bar{T}_{\text{StSt}}^{\text{Opt}}$ [°C]	-31.04 ± 2.02	-30.46 ± 0.83	-30.00 ± 1.05	-30.80 ± 0.48	-25.93 ± 0.70
$\bar{T}_{\text{StSt}}^{\text{Firn}}$ [°C]	-31.41 ± 2.07	-30.81 ± 0.85	-30.35 ± 1.07	-31.14 ± 0.49	-26.18 ± 0.71

Table 6.6: Steady state temperature estimates based on the final firn diffusion length estimates found. T_0 is the temperature used to generate the theoretical diffusion length and density profiles, and originates from [?, add. text]

The final temperatures have been presented along with the assumed temperature, T_0 , which has been used to generate modelled density and diffusion length profiles previously in this work and is the steady state assumed surface temperature at the sites.

6.3.1 Further Possibilities with the Iso-CFM

These temperature estimates are made on a steady state assumption, that is, that the annual accumulation rate, A_0 , and surface temperature, T_0 , are at a steady state. This of course is not an optimal assumption, but it is a useful stepping stone to sanity check the results and from where to continue further temperature analysis and estimations with less rigid models.

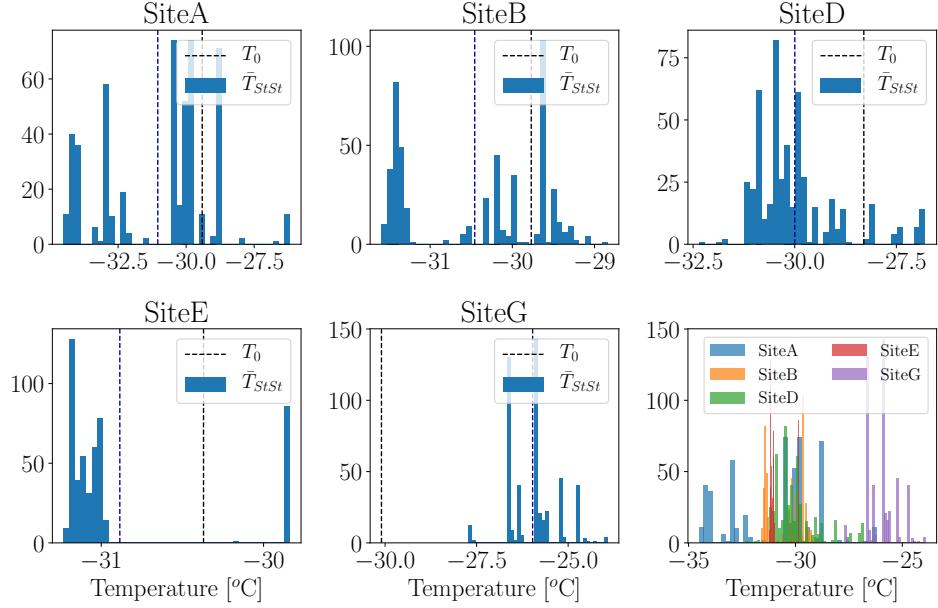


Figure 6.5: Steady state temperature distributions

An obvious easy next step would be to investigate different accumulation rates, instead of just computing for a single A_0 . This will need an adjusting of the λ used in the general method and might give insight into a more accurate A_0 estimate, still under steady state assumptions.

In further analysis and work, it would be interesting to utilize the Iso-CFM to model non-steady state solutions and use these to estimate temperatures from more realistic models. This could be models implementing seasonal changes in both temperature and accumulation rate, or statistical analysis of random variations.

Chapter 7

Conclusion and Outlook

In this final chapter of the thesis, the main finding will be summed up, along with a short presentation of the final results. Lastly an outlook section presents the next steps the conclusions point to be taken, in future research.

Contents

7.1	Conclusions	93
7.2	Outlook	95

7.1 Conclusions

As the need for more accurate and precise models for the future increases, so does the need for better and wider understanding of the past. This work set out to investigate a new method to expand our knowledge of how to reconstruct paleotemperatures best. Through a series of methodical computational and theoretical analyses, a model for estimating the temperature of an ice core time series was developed.

The starting idea was to utilize the fact that diffusion length is temperature dependent, and to use signal analysis, theoretical knowledge of diffusion processes in ice and constrained pattern recognition to estimate the diffusion length σ empirically in a water isotopic depth series. Throughout the thesis was presented the different methods used and the mathematical and theoretical foundations on which they were developed. A variety of methods were tested, discussed and further developed to finally generate a generic method.

The final method reconstructs the measured signal by back diffusion, attempting to reconstruct the part of the signal that was washed out by diffusion, using σ as a tuning parameter of the diffusion process. The method finds the optimal diffusion length by choosing the σ which results in the known number of years in a given section. In this thesis, the two volcanic events of Laki in 1783 and Tambora in 1815 have been used for specific dating of the examined water isotopic depth series, but the method is not restricted to volcanic markers nor to isotopic depth series. Any series exhibiting annual cycles and affected by diffusion could be used.

This back diffusion method results in final optimal diffusion length estimates, σ_{opt} , which are presented in Table 7.1 along with the firn diffusion estimates, σ_{firn} , corrected for sampling and ice diffusion.

			σ_{final}
Site A	σ_{opt}	[cm]	7.37 ± 0.54
	σ_{firn}	[cm]	7.27 ± 0.55
Site B	σ_{opt}	[cm]	7.35 ± 0.22
	σ_{firn}	[cm]	7.26 ± 0.22
Site D	σ_{opt}	[cm]	7.21 ± 0.28
	σ_{firn}	[cm]	7.12 ± 0.28
Site E	σ_{opt}	[cm]	8.22 ± 0.15
	σ_{firn}	[cm]	8.12 ± 0.15
Site G	σ_{opt}	[cm]	9.46 ± 0.24
	σ_{firn}	[cm]	9.38 ± 0.24

Table 7.1: Final diffusion length estimates, based on conclusions made previously in different tests.

From the optimal σ a steady state temperature was estimated. These results can be seen in Table 7.2, both the temperature estimated from σ_{Opt} and σ_{firn} . For further studies, it is evident to use the diffusion length estimates and using them in more complex temperature estimate models, and not just a steady state solution.

		Site A	Site B	Site D	Site E	Site G
T_0	[°C]	-29.41	-29.77	-28.3	-30.37	-30.1
$\bar{T}_{\text{StSt}}^{\text{Opt}}$	[°C]	-31.04 ± 2.02	-30.46 ± 0.83	-30.00 ± 1.05	-30.80 ± 0.48	-25.93 ± 0.70
$\bar{T}_{\text{StSt}}^{\text{Firm}}$	[°C]	-31.41 ± 2.07	-30.81 ± 0.85	-30.35 ± 1.07	-31.14 ± 0.49	-26.18 ± 0.71

Table 7.2: Steady state temperature estimates based on the final firm diffusion length estimates found. T_0 is the temperature used to generate the theoretical diffusion length and density profiles, and originates from [?, add. text]

7.2 Outlook

Using the estimated σ_{firm} results and the general conclusions from this thesis, an apparent next step would be to examine the paleotemperature in the given depth section more thoroughly. To be able to verify the efficiency of the developed methods in this thesis, a more extensive and complex examination of the temperature would be needed. Along with this, implementing $\sigma(z)$ and investigating how it affects the final temperature estimates can also be used for more precise and accurate temperature estimation.

Additionally, the method would benefit from being tested on multiple new ice cores, and on new depth sections, i.e. between other known volcanic events. This could also help verifying the chosen constraints, as the new cores would have a different annual layer thickness profile. This could also lead to an investigation on the choice of constraints. These constraints could be further developed and more seasonality could be implemented, and maybe the constraints should be tuned differently for different ice cores.

As a byproduct of this work, an interesting feature presented itself, namely the relationship between the number of peaks versus the diffusion length used in back diffusion. If it is possible to estimate the most stable number of peaks (i.e. where a change in diffusion length does not dramatically change the number of counted peaks), is could be possible to date a depth series with no dating markers.

Finally, the in depth analysis of stability of this method considering different computational tools points to the research field being more careful about the technical data analysis. Especially considering interpolation, which again points to more careful and higher resolution data measurements, and

considering choice of spectral transforms, where this work concludes that the NDCT restores best.

The method has room for improvement, especially some of the simpler assumptions like the constraints and the optimization routine could benefit from further development. Moreover, the work carried out in this thesis leaves room for additional examination and development, but lays a good foundation to be used as a stepping stone in future research.

Bibliography

- [1] S. Aciego, B. Bourdon, J. Schwander, H. Baur, and A. Forieri. Toward a radiometric ice clock: uranium ages of the Dome C ice core. *Quaternary Science Reviews*, 30(19):2389–2397., 2011.
- [2] N. Ahmed, T. Natarajan, and K. R. Rao. Discrete Cosine Transform. *IEEE Transactions on Computers*, C-23(1):90–93, 1974.
- [3] G. Audi, F. G. Kondev, Meng Wang, W. J. Huang, and S. Naimi. The NUBASE2016 evaluation of nuclear properties. *Chinese Physics C*, 41(3):1–138, 2017.
- [4] Eugeni Barkan and Boaz Luz. High precision measurements of $^{17}\text{O}/^{16}\text{O}$ and $^{18}\text{O}/^{16}\text{O}$ ratios in H_2O . *Rapid Communications in Mass Spectrometry*, 19(24):3737–3742, 2005.
- [5] JM Barnola, P Pimienta, D Raynaud, and YS Korotkevich. CO₂-climate relationship as deduced from the Vostok ice core: a re-examination based on new measurements and on a re-evaluation of the air dating. *Tellus*, B 43(2):83–90, 1991.
- [6] T. Blunier and J. Schwander. Gas enclosure in ice: age difference and fractionation. *Physics of ice core records*, pages 307–326, 2000.
- [7] G Bruun. z-Transform DFT filters and FFTs. *IEEE Transactions on Acoustics, Speech and Signal Processing (ASSP)*, 26(1):56–63, 1978.
- [8] Emilie Capron, A Landais, PC Tzedakis, E Bard, T Blunier, D Dahl-Jensen, T Dokken, R Gersonde, F Parrenin, M Schulz, B Vinther, and C Waelbroek. Dating and synchronizing paleoclimatic records over the last interglacial. *PAGES news*, 21(1):26–27, 2013.

- [9] H.B. Clausen, N.S. Gundestrup, S.J. Johnsen, R. Bindschadler, and J. Zwally. Glaciological Investigations in the Crête Area, Central Greenland: A Search for a new Deep-Drilling Site. *Annals of Glaciology*, 10:10–15, 1988.
- [10] H.B. Clausen and C.U. Hammer. The Laki and Tambora Eruptions as Revealed in Greenland Ice Cores from 11 Locations. *Annals of Glaciology*, 10:16–22, 1988.
- [11] Jihong Cole-Dai, David Ferris, Alyson Lanciki, Joël Savarino, Mélanie Baroni, and Mark H. Thiemens. Cold decade (AD 1810-1819) caused by Tambora (1815) and another (1809) stratospheric volcanic eruption. *Geophysical Research Letters*, 36(22), 2009.
- [12] J. W. Cooley and J. W. Tukey. An Algorithm for the Machine Calculation of Complex Fourier Series. *AMS: Mathematics of Computation*, 19:297–301, 1965.
- [13] W. Dansgaard. Stable isotopes in precipitation. *Tellus*, 16(4):436–468, 1964.
- [14] MD Ellehøj, HC Steen-Larsen, SJ Johnsen, and MB Madsen. Ice-vapor equilibrium fractionation factor of hydrogen and oxygen isotopes: experimental investigations and implications for stable water isotope studies. *Rapid Communications in Mass Spectrometry*, 27(19):2149–2158, 2013.
- [15] Adolf Fick. Über Diffusion. *Annalen der Physik*, 94(1):59–86, 1851.
- [16] V. Gkinis, S. B. Simonsen, S. L. Buchardt, J. W.C. White, and B. M. Vinther. Water isotope diffusion rates from the NorthGRIP ice core for the last 16,000 years - Glaciological and paleoclimatic implications. *Earth and Planetary Science Letters*, 405:132–141, 2014.
- [17] Vasileios Gkinis, Christian Holme, Emma C Kahle, C Max Stevens, Eric J Steig, and Bo M Vinther. Numerical experiments on firn isotope diffusion using the Community Firn Model. *Journal of Glaciology*, in review(x):1–54, 2019.
- [18] Vasileios Gkinis, Christian Holme, Emma C Kahle, Max C Stevens, Bo M Vinther, and Eric J Steig. Numerical experiments on firn isotope diffusion with the Community Firn Model. 2021.

- [19] Vasileios Gkinis, Trevor J Popp, Sigfus J Johnsen, Thomas Blunier, M Bigler, S Schüpbach, E Kettner, and Sigfus J Johnsen. Water isotopic ratios from a continuously melted ice core sample. *Atmospheric Measurement Techniques*, 4(4):2531–2542, 2011.
- [20] Vasileios Gkinis, Bo M. Vinther, Trevor J. Popp, Thea Quistgaard, Anne Katrine Faber, Christian T. Holme, Camilla Marie Jensen, Mika Lanzky, Anine Maria Lütt, Vasileios Mandrakis, Niels Ole Ørum, Anna Sofie Pedersen, Nikol Vaxevani, Yongbiao Weng, Emilie Capron, Dorthe Dahl-Jensen, Maria Hörhold, Tyler R. Jones, Jean Jouzel, Amaëlle Landais, Valérie Masson-Delmotte, Hans Oerter, Sune O. Rasmussen, Hans Christian Steen-Larsen, Jørgen Peder Steffensen, Árný Erla Sveinbjörnsdóttir, Anders Svensson, Bruce Vaughn, and James W.C. White. A 120,000-year long climate record from a NW-Greenland deep ice core at ultra-high resolution. *Scientific Data*, 8(1):1–10, 2021.
- [21] I. J. Good. The Interaction Algorithm and Practical Fourier Analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, 20(2):361–372, 1958.
- [22] C Goujon, JM Barnola, and C Ritz. Modeling the densification of polar firn including heat diffusion: application to close-off characteristics and gas isotopic fractionation for Antarctica and Greenland sites. *JGR: Atmospheres*, 108(D24):4792–4792, 2003.
- [23] W. D. Hall and H. R. Pruppacher. The Survival of Ice Particles Falling from Cirrus Clouds in Subsaturated Air. *Journal of the Atmospheric Sciences*, 33(October):1995–2006, 1976.
- [24] C. U. Hammer. Acidity of Polar Ice Cores in Relation to Absolute Dating, Past Volcanism, and Radio-Echoes. *Journal of Glaciology*, 25(93):359–372, 1980.
- [25] M. M. Herron and C. C. Jr. Langway. Firn densification: an Empirical Model. *Journal of Glaciology*, 25(93), 1980.
- [26] Christian Holme, Vasileios Gkinis, and Bo M. Vinther. Molecular diffusion of stable water isotopes in polar firn as a proxy for past temperatures. *Geochimica et Cosmochimica Acta*, 225:128–145, 2018.
- [27] S. J. Johnsen. Stable isotope homogenization of polar firn and ice. *Isotopes and impurities in snow and ice*, (1):210–219, 1977.

- [28] S. J. Johnsen, D. Dahl-Jensen, N. Gundestrup, J. P. Steffensen, H. B. Clausen, H. Miller, V. Masson-Delmotte, A. E. Sveinbjörnsdóttir, and J. White. Oxygen isotope and palaeotemperature records from six Greenland ice-core stations: Camp Century, Dye-3, GRIP, GISP2, Renland and NorthGRIP. *Journal of Quaternary Science*, 16((4)):299–307, 2001.
- [29] Sigfus Johnsen, Henrik Clausen, Kurt M Cuffey, Georg Hoffmann, Jakob Schwander, and Timothy Creyts. Diffusion of stable isotopes in polar firn and ice: the isotope effect in firn diffusion. *Physics of Ice Core Records*, 121-140, pages 121–140, 2000.
- [30] Sigfús J Johnsen, Henrik B Clausen, Kurt M Cuffey, Georg Hoffmann, Jakob Schwander, and Creyts Timothy. Diffusion of stable isotopes in polar firn and ice: the isotope effect in firn diffusion. *Physics of Ice Records*, pages 121–140, 2000.
- [31] J. Jouzel, R. B. Alley, K. M. Cuffey, W. Dansgaard, P. Grootes, G. Hoffmann, S. J. Johnsen, R. D. Koster, D. Peel, C. A. Shuman, M. Stievenard, M. Stuiver, and J. White. Validity of the temperature reconstruction from water isotopes in ice cores. *Journal of Geophysical Research: Oceans*, 102(C12):26471–26487, 1997.
- [32] Emma C. Kahle, Christian Holme, Tyler R. Jones, Vasileios Gkinis, and Eric J. Steig. A Generalized Approach to Estimating Diffusion Length of Stable Water Isotopes From Ice-Core Data. *Journal of Geophysical Research: Earth Surface*, 123(10):2377–2391, 2018.
- [33] Kara D. Lamb, Benjamin W. Clouser, Maximilien Bolot, Laszlo Sarkozy, Volker Ebert, Harald Saathoff, Ottmar Möhler, and Elisabeth J. Moyer. Laboratory measurements of HDO/H₂O isotopic fractionation during ice deposition in simulated cirrus clouds. *Proceedings of the National Academy of Sciences of the United States of America*, 114(22):5612–5617, 2017.
- [34] F. Lambert, B. Delmonte, J.-R. Petit, M. Bigler, P.R. Kaufmann, M.A. Hutterli, T.F. Stocker, U. Ruth, J.P. Steffensen, and V. Maggi. Dust-climate couplings over the past 800, 000 years from the EPICA Dome C ice core. *Nature*, 452((7187)):616–619., 2008.
- [35] C.C. Langway, H.B. Clausen, and C.U. Hammer. An Inter-Hemispheric Volcanic Time-Marker in Ice Cores from Greenland and Antarctica. *Annals of Glaciology*, 10:102–108, 1988.

- [36] June Yub Lee and Leslie Greengard. The type 3 nonuniform FFT and its applications. *Journal of Computational Physics*, 206(1):1–5, 2005.
- [37] M. Majoube. Fractionation factor of ^{18}O between water vapour and ice. *Nature*, 226(5252):1242, 1970.
- [38] J. Makhoul. A Fast Cosine Transform In One And Two Dimensions. *IEEE Transactions on Signal Processing*, ASSP-28(1):27–34, 1980.
- [39] Alan Marshall. Did a volcanic eruption in Indonesia really lead to the creation of Frankenstein? *The Conversation*, January(24), 2020.
- [40] Farokh Marvasti. *Nonuniform Sampling*. 1993.
- [41] Clare D. McGillem and George R. Cooper. *Continuous and Discrete Signal and System Analysis*. 1984.
- [42] L. Merlivat. Molecular diffusivities of H₂ 16O, H D 16O, and H₂ 18O gases. *Journal of Chemical Physics*, 69(6):2864–2871, 1978.
- [43] L. Merlivat and G. Nief. Fractionnement isotopique lors des changements d'état solide-vapeur et liquide-vapeur de l'eau à des températures inférieures à 0°C. *Tellus*, 19(1):122–127, 1967.
- [44] J Moore and J Paren. A New Technique For Dielectric Logging of Antarctic Ice Cores. *Journal de Physique Colloques*, 48(C1):155–160, 1987.
- [45] J. C. Moore and S. Fujita. Dielectric properties of ice containing acid and salt impurity at microwave and low frequencies. *Journal of Geophysical Research*, 98(B6):9769–9780, 1993.
- [46] J. C. Moore, J. G. Paren, and R. Mulvaney. Chemical evidence in polar ice cores from dielectric profiling. *Annals of Glaciology*, 14:195–198, 1990.
- [47] DM Murphy and T Koop. Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Quarterly Journal of the Royal Meteorological Society*, 131(608):1539–1565, 2005.
- [48] Members NGRIP. High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Quaternary research.*, 82(1):147–151, 2014.
- [49] Clive Oppenheimer. Climatic, environmental and human con sequences of the largest known historic eruption: Tambora volcano (Indonesia) 1815. *Progress in Physical Geography*, 27(2):230–259, 2003.

- [50] Marc-Antoine Parseval des Chênes. Mémoire sur les séries et sur l'intégration complète d'une équation aux différences partielles linéaire du second ordre, à coefficients constants. *Mémoires présentés à l'Institut des Sciences, Lettres et Arts, par divers savants, et lus dans ses assemblées. Sciences, mathématiques et physiques*, 1:638–648, 1806.
- [51] S V Patankar. PatankarNumericalHeatTransferandFluidFlow, 1980.
- [52] Sandra Petersen. About EastGRIP.
- [53] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, 2007.
- [54] René O. Ramseier. Self-diffusion of tritium in natural and synthetic ice monocrystals. *Journal of Applied Physics*, 38(6):2553–2556, 1967.
- [55] S. O. Rasmussen, K. K. Andersen, S. J. Johnsen, M. Bigler, and T. McCormack. Deconvolution-based resolution enhancement of chemical ice core records obtained by continuous flow analysis. *Journal of Geophysical Research D: Atmospheres*, 110(17):163–169, 2005.
- [56] Sune O Rasmussen, A Svensson, and M Winstrup. State of the art of ice core annual layer dating. *Past Global Changes Magazine*, 22(1):26–27, 2014.
- [57] Diego Ruiz-Antolin and Alex Townsend. A nonuniform fast fourier transform based on low rank approximation. *SIAM Journal on Scientific Computing*, 40(1):A529–A547, 2018.
- [58] Claude E. Shannon. Communication in the Presence of Noise. *Proceedings of the IRE*, 37(1):10–21, 1949.
- [59] Gabriele Steidl, Daniel Potts, and Manfred Tasche. Fast Fourier Transform For Nonequispaced Data: A Tutorial. In *Modern Sampling Theory*, volume 2, chapter Chapter 1, pages 1–23. 2001.
- [60] Jón Steingrímsson and Keneva Kunz. Fires of the earth: the Laki eruption, 1783–1784. *University of Iceland Press.*, December, 1998.
- [61] C. Max Stevens, Vincent Verjans, Jessica M. D. Lundin, Emma C. Kahle, Annika N. Horlings, Brita I. Horlings, and Edwin D. Waddington. The Community Firn Model (CFM) v1.0. *Geoscientific Model Development*, 13(9):4355–4377, 2020.

- [62] R. Stone. Iceland ' s Doomsday Scenario ? *Science*, 306(5700):1278–1281, 2004.
- [63] K. C. Taylor, C. U. Hammer, R. B. Alley, H. B. Clausen, D. Dahl-Jensen, A. J. Gow, N. S. Gundestrup, J. Kipfstuh, J. C. Moore, and E. D. Waddington. Electrical conductivity measurements from the GISP2 and GRIP Greenland ice cores, 1993.
- [64] Thorvaldur Thordaldson and Stephen Self. Atmospheric and environmental effects of the 1783–1784 Laki eruption: A review and reassessment. *Journal of Geophysical Research*, 108(D1 40111):159–171, 2004.
- [65] B. Tian and Q. H. Liu. Nonuniform fast cosine transform and chebyshev PSTD algorithms. *Progress in Electromagnetics Research*, 28:253–273, 2000.
- [66] Tad J Ulrych and Thomas N Bishop. Maximum entropy spectral analysis and autoregressive decomposition. *Reviews of Geophysics*, 13(1):183–200, 1975.
- [67] B. M. Vinther, H. B. Clausen, S. J. Johnsen, S. O. Rasmussen, K. K. Andersen, S. L. Buchardt, D. Dahl-Jensen, I. K. Seierstad, M. L. Siggaard-Andersen, J. P. Steffensen, A. Svensson, J. Olsen, and J. Heinemeier. A synchronized dating of three Greenland ice cores throughout the Holocene. *Journal of Geophysical Research Atmospheres*, 111(13), 2006.
- [68] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R.J. Nelson, Eric Jones, Robert Kern, Eric Larson, C. J. Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, Aditya Vijaykumar, Alessandro Pietro Bardelli, Alex Rothberg, Andreas Hilboll, Andreas Kloeckner, Anthony Scopatz, Antony Lee, Ariel Rokem, C. Nathan Woods, Chad Fulton, Charles Masson, Christian Häggström, Clark Fitzgerald, David A. Nicholson, David R. Hagen, Dmitrii V. Pasechnik, Emanuele Olivetti, Eric Martin, Eric Wieser, Fabrice Silva, Felix Lenders, Florian Wilhelm, G. Young, Gavin A. Price, Gert Ludwig Ingold, Gregory E. Allen, Gregory R. Lee, Hervé Audren, Irvin Probst, Jörg P. Dietrich, Jacob Silterra, James T.

- Webber, Janko Slavič, Joel Nothman, Johannes Buchner, Johannes Kulick, Johannes L. Schönberger, José Vinícius de Miranda Cardoso, Joscha Reimer, Joseph Harrington, Juan Luis Cano Rodríguez, Juan Nunez-Iglesias, Justin Kuczynski, Kevin Tritz, Martin Thoma, Matthew Newville, Matthias Kümmeler, Maximilian Bolingbroke, Michael Tartre, Mikhail Pak, Nathaniel J. Smith, Nikolai Nowaczyk, Nikolay Shebanov, Oleksandr Pavlyk, Per A. Brodtkorb, Perry Lee, Robert T. McGibbon, Roman Feldbauer, Sam Lewis, Sam Tygier, Scott Sievert, Sebastiano Vigna, Stefan Peterson, Surhud More, Tadeusz Pudlik, Takuya Oshima, Thomas J. Pingel, Thomas P. Robitaille, Thomas Spura, Thouis R. Jones, Tim Cera, Tim Leslie, Tiziano Zito, Tom Krauss, Utkarsh Upadhyay, Yaroslav O. Halchenko, and Yoshiki Vázquez-Baeza. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272, 2020.
- [69] Lijia Wei, Ellen Mosley-Thompson, Paolo Gabrielli, Lonnie G. Thompson, and Carlo Barbante. Synchronous deposition of volcanic ash and sulfate aerosols over Greenland in 1783 from the Laki eruption (Iceland). *Geophysical Research Letters*, 35(16), 2008.
 - [70] S. Weissbach, A. Wegner, T. Opel, H. Oerter, B. M. Vinther, and S. Kipfstuhl. Spatial and temporal oxygen isotope variability in northern Greenland-implications for a new climate record over the past millennium. *Climate of the Past*, 12(2):171–188, 2016.
 - [71] Joe Wheatley. *Automated Bayesian Layer Counting of Ice Cores*. Phd, University of Sheffield, 2015.
 - [72] M. Winstrup, A. M. Svensson, S. O. Rasmussen, O. Winther, E. J. Steig, and A. E. Axelrod. An automated approach for annual layer counting in ice cores. *Climate of the Past*, 8(6):1881–1895, 2012.
 - [73] C. S. Zerefos, V. T. Gerogiannis, D. Balis, S. C. Zerefos, and A. Kazantzidis. Atmospheric effects of volcanic eruptions as seen by famous artists and depicted in their paintings. *Atmospheric Chemistry and Physics*, 7(15):4027–4042, 2007.
 - [74] Xianfeng Zhao, Bingbing Xia, and Yi Deng. Strengthening QIM-based watermarking by non-uniform discrete cosine transform. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5284 LNCS:309–324, 2008.

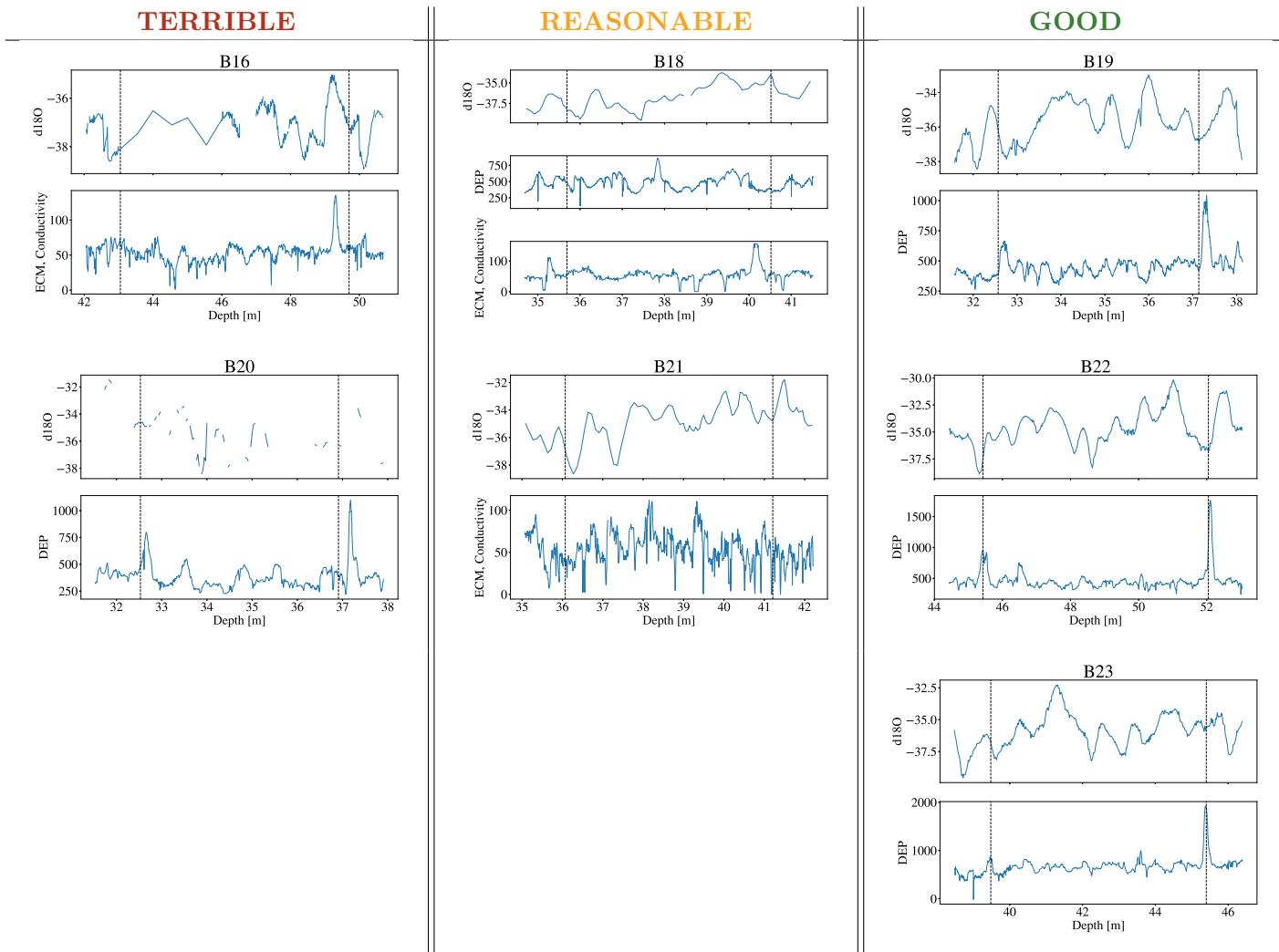
Appendices

APPENDIX II: Data - AWI B-cores

AWI B-Cores: Core Specifications

		B19	B22	B23
d_L	[m]	80.85	83.7	93.8
d_T	[m]	70.9	73.0	81.5
A_0 W.E.	[m]	0.307	0.327	0.365
A_0 I.E.	[m]	0.282	0.300	0.335
T_0	[°C]	-29.41	-29.77	-28.3

Table 1:



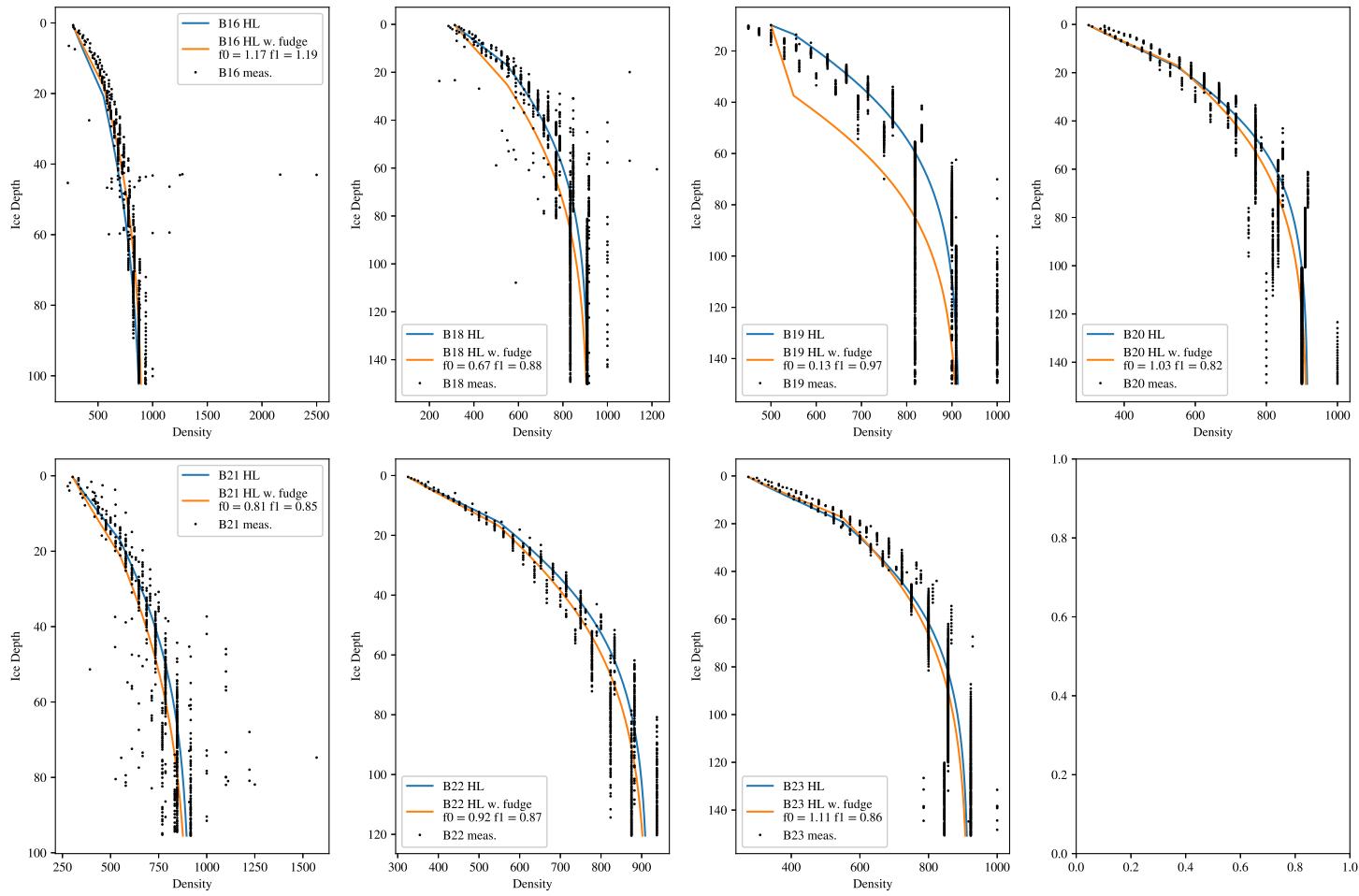


Figure 1:
AWI B-cores density measurements and HL-modeled densities.

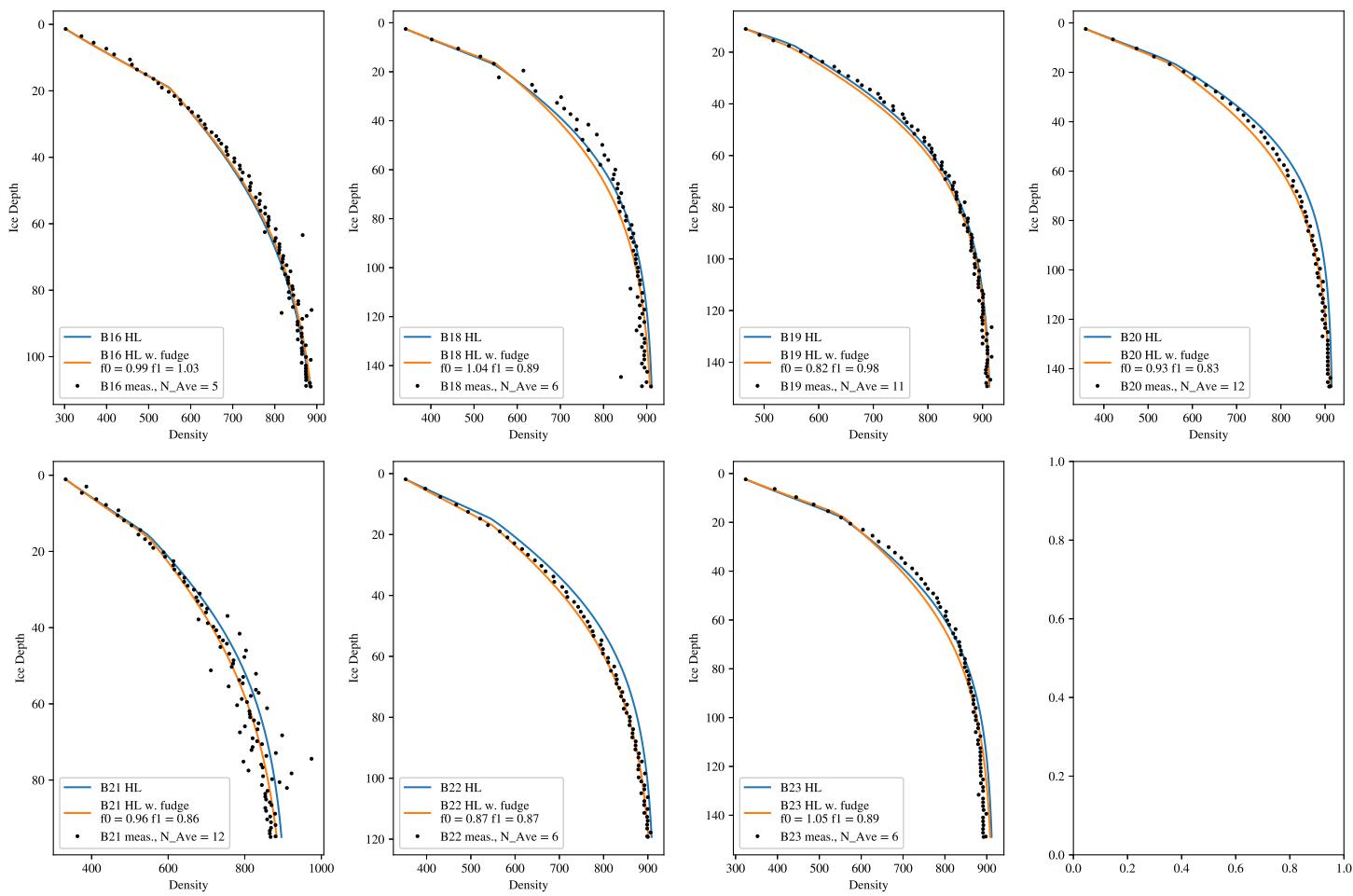


Figure 2: AWI B-cores averaged density measurements and related HL-modeled densities.

APPENDIX III: Data - Alphabet Cores

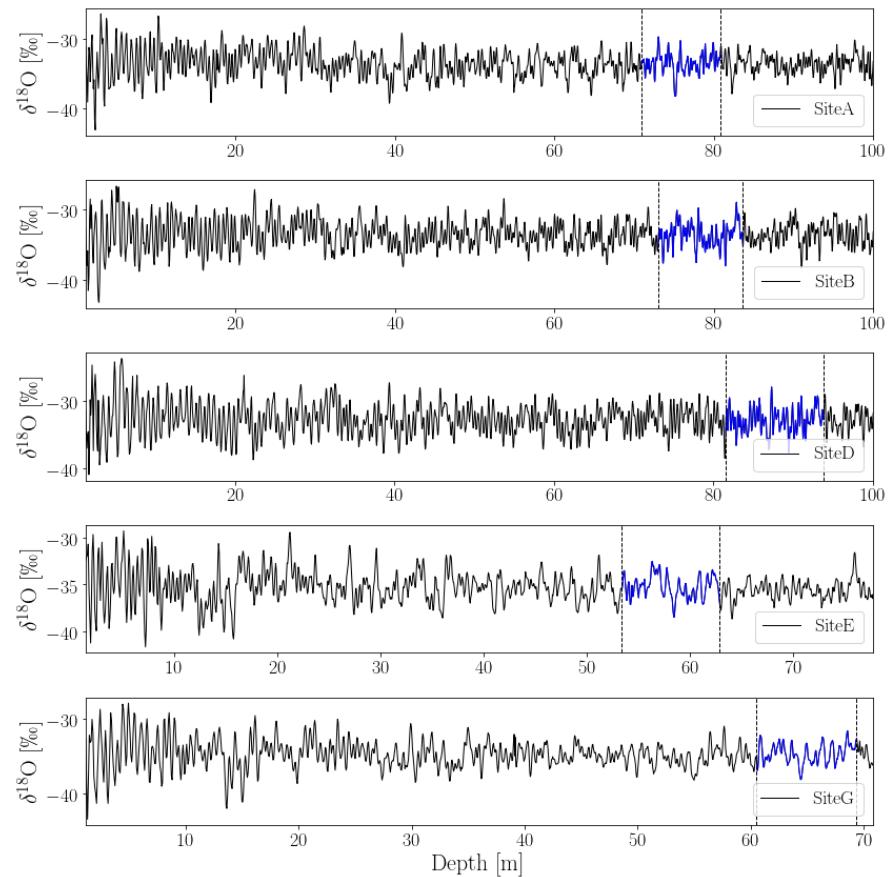


Figure 3: All six Alphabet cores under examination, with Laki to Tambora depth section highlighted.

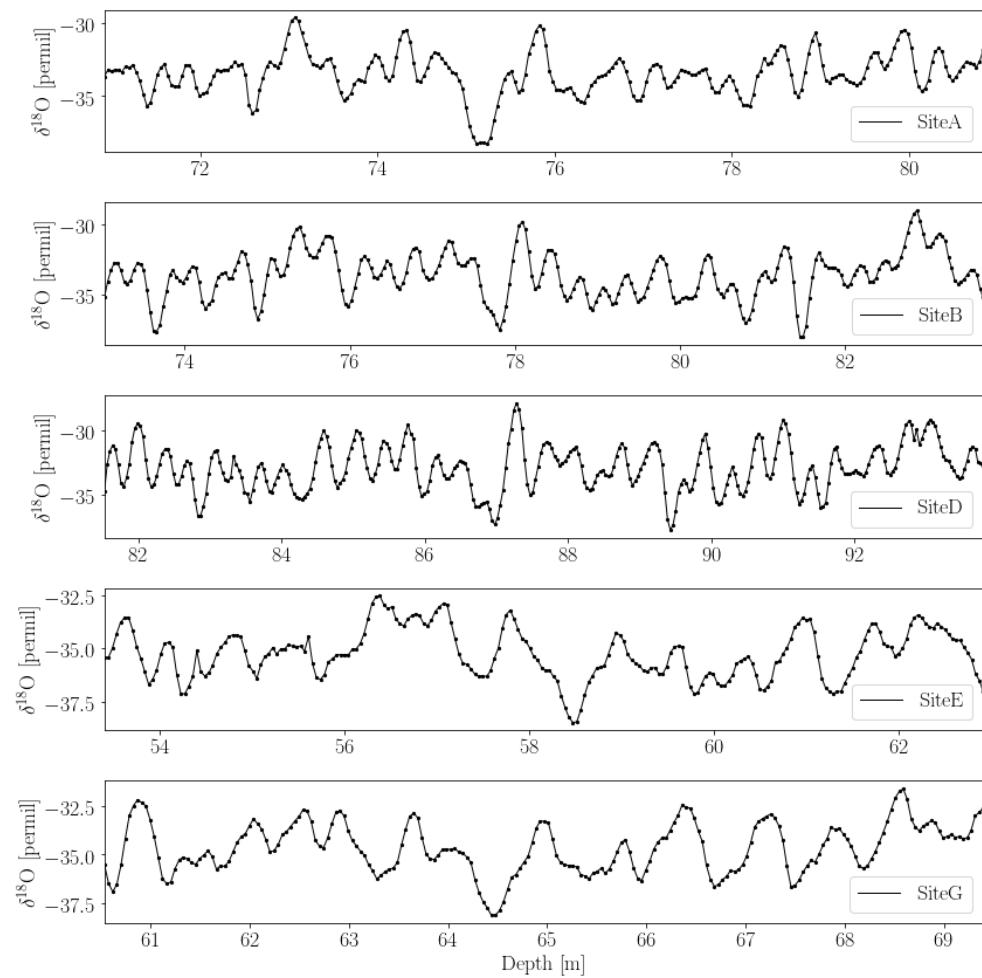


Figure 4: All six Alphabet cores, showing only the depth series concerning the Laki to Tambora section.

Alphabet Cores: Core Specifications

		Site A	Site B	Site D	Site E	Site G
d_L^{CH}	[m]	80.85	83.7	93.8	62.95	69.4
d_T^{CH}	[m]	70.9	73.0	81.5	53.4	60.5
A_0 W.E.	[m]	0.307	0.327	0.365	0.225	0.251
A_0 I.E.	[m]	0.282	0.300	0.335	0.206	0.230
T_0	[°C]	-29.41	-29.77	-28.3	-30.37	-30.1
ρ_0	[kg m ⁻³]	343.0	355.0	350.0	325.0	-
z_0		0.55	0.55	0.825	0.675	-
s_L^{CH}	[cm]	30.0	45.0	55.0	35.0	35.0
s_T^{CH}	[cm]	65.0	55.0	70.0	40.0	75.0
ρ_L	[kg m ⁻³]	836.0	841.0	857.0	786.0	807.0
ρ_T	[kg m ⁻³]	812.0	816.0	839.0	749.0	778.0
d_L^{TQ}	[m]	80.87	83.82	93.95	62.9	69.38
d_T^{TQ}	[m]	70.90	73.01	81.55	53.43	60.48
s_L^{TQ}	[cm]	34.0	30.0	30.0	45.0	35.0
s_T^{TQ}	[cm]	40.0	57.0	55.0	48.0	82.0

Table 2: Core specifications for core drilled at Site A. d describes depth of event, A describes accumulation rate, T describes temperature, ρ describes density at given depth and s describes the width of a given event. Subscripts L and T stands for volcanic events Laki and Tambora, respectively, subscript 0 describes initial surface condition, and superscripts CH and TQ represents the original and corrected values for depth and width of events.

APPENDIX IV: Spectral Transforms

Discrete and Fast Fourier Transform

The definition of the continuous Fourier transform and its inverse was presented in the above. The Fourier transform is as seen a way of representing the function under consideration as an infinite sum of periodic components. When the function is discrete, so will the Fourier transform be, and the integral is replaced with a sum. This gives us the Discrete Fourier Transform (DFT) which transforms the signal into a sum of separate components contributing at different frequencies. The DFT is dependent on the sampling interval, Δ , and we can describe our discrete signal X as a function of N discrete time steps $t_k = k \cdot \Delta$, where $k = 0, 1, \dots, N - 1$:

$$X_k \equiv X(t_k) \quad (1)$$

This sample size is supposed to be representative for the entire discrete function, if the function continues beyond the N sampled points. When sampling discretely at interval Δ , there will be a special frequency, the Nyquist critical frequency, defined through the sampling size as:

$$f_{NQ} \equiv \frac{1}{2\Delta}. \quad (2)$$

This frequency is of great importance in transformation of discrete signals. If the continuous signal is sampled at an interval Δ is bandwidth limited to frequencies smaller in magnitude than f_{NQ} , $\tilde{X}(f) = 0$ for $|f| \geq f_{NQ}$ - i.e. the transformed function has only non-zero values inside the Nyquist interval, $\tilde{X}(-f_{NQ}), \dots, \tilde{X}(f), \dots, \tilde{X}(f_{NQ})$. This means that the function is completely determined since we have all information about the signal contained in our available frequency space.

On the other hand, which is much more likely, if the continuous signal consists of frequencies both inside and outside the Nyquist interval, then all spectral information outside of this range will be falsely interpreted as being inside this range. Thus a wave inside the interval with a frequency of f_n will have a number of wave siblings outside of the interval, with frequencies of $k \cdot \frac{1}{\Delta} f_n$, k being integers, which will be aliased into the Nyquist interval and give rise to an increased power at the frequency f_n .

When analyzing an already measured discrete signal, this might give rise to some headache. What can be done is to assume that the signal has been sampled competently and then assume that the Fourier transform is zero outside of the Nyquist interval. After the analysis it will then be possible to determine if the signal was indeed competently sampled, as the Fourier series will

go to zero at f_{NQ} given a correct assumption, and go to a fixed value, if the sampling was not done competently.

Now with the basics of understanding the limits of frequency transform of a discretely sampled signal, it is possible to estimate the DFT of the signal $X_k \equiv X(t_k)$. Since the Fourier transform is a symmetric transformation it is easiest to assume that N is even.

Since the input information is of size N we should expect only to sample the frequency transform $\tilde{X}(f)$ at only discrete values of f in the range between the upper and lower critical Nyquist frequencies, $-f_{NQ}$ to f_{NQ} :

$$f_n \equiv \frac{n}{N\Delta}, \quad n = -\frac{N}{2}, \dots, \frac{N}{2} \quad (3)$$

This will indeed actually give rise to $N + 1$ values, since 0 will be in the interval as well, but the limit frequencies are actually not independent, but all frequencies between are, which reduces it to N samples.

Now the integral from Equation 4.3 needs to be estimated as a sum:

$$\tilde{X}(f_n) = \int_{-\infty}^{\infty} X(\tau) e^{2\pi i f_n \tau} d\tau \approx \sum_{k=0}^{N-1} X_k e^{2\pi i f_n t_k} \Delta = \Delta \sum_{k=0}^{N-1} X_k e^{2\pi i k \frac{n}{N}} \quad (4)$$

The Discrete Fourier Transform is thus defined as:

$$\tilde{X}_n \equiv \sum_{k=0}^{N-1} X_k e^{2\pi i k \frac{n}{N}} \quad (5)$$

This gives the approximate relation between the DFT estimate and the continuous Fourier transform $\tilde{X}(f)$ when sampling at size Δ as:

$$\tilde{X}(f_n) \approx \Delta \tilde{X}_n \quad (6)$$

The inverse DFT is given as:

$$X_n \equiv \frac{1}{N} \sum_{k=0}^{N-1} X \tilde{X}_n e^{-2\pi i k \frac{n}{N}} \quad (7)$$

Computation of the DFT can be very slow and tiresome, since it involves complex multiplication between a number of vectors and matrices. If we write Equation 5 as $\tilde{X}_n = \sum_{k=0}^{N-1} N - 1 W^{nk} X_k$, where W is a complex number $W \equiv e^{2\pi i / N}$. This shows that the vector X_k must be multiplied with a complex matrix which (n,k)th component consists of the constant W to the power

of nk . This matrix multiplication evidently leads to a process of $O(N^2)$. Fortunately, a number of different algorithms implementing a wide range of different theories from complex number arithmetic and prime-factoring to group and number theory ([12],[21], [7] and others) have been developed for fast and efficient computation of the discrete Fourier transform. One of these is called the Fast Fourier Transform (FFT), which can reduce the computations to just $\mathcal{O}(N \log_2 N)$. In this thesis the FFT used is the one implemented in the `scipy.fft` Python package [68], which is based on the works of [12]. See said article for implementation details. One important thing about this specific algorithm is that for the algorithm to function most efficiently, the number of points computed in the frequency space must be of a power of 2, following the use of base \log_2 .

Nonuniform Discrete Fourier Transform

All FFT algorithms evaluate the DFT definitions from Eqs. 5 to 7 in fast and efficient ways. But one key assumption for these methods is that the data under examination are equispaced, i.e. uniformly distributed, based on the summation in Eq. 5. The computations thus expect uniform data as input and returns uniform data as output. Unfortunately this is not always the case for data collected in physical experiments. In this case the basic assumptions for the calculations of both DFT and FFT are flawed.

The most general form of a nonuniform transform would be the one that takes non-equispaced data as input and also returns non-equispaced transforms as output. Firstly we wish to create a nonuniform discrete Fourier transform (NUDFT) that transforms a sequence of N complex numbers X_0, \dots, X_{N-1} to a different sequence of M complex numbers, $\tilde{X}_0, \dots, \tilde{X}_{M-1}$. The one-dimensional NUDFT computes the transformed vector $\tilde{\mathbf{X}} = (\tilde{X}_0, \dots, \tilde{X}_{M-1})^T$, with entries computed as the sum

$$\tilde{X}_k = \sum_{n=0}^{N-1} X_n e^{-2\pi i p_n f_k}, \quad 0 \leq k \leq M - 1. \quad (8)$$

The values X_0, \dots, X_{N-1} are sample values, p_0, \dots, p_{N-1} are sample positions and f_0, \dots, f_{M-1} are frequencies. The NUDFT vector $\tilde{\mathbf{X}}$ is found by computing M sums with each N terms. This meaning that the computational cost will be of order $\mathcal{O}(M \cdot N)$, and if $M = N$ then of $\mathcal{O}(N^2)$. The NUDFT reduces to the DFT if the points are equispaced, $p_n = \frac{n}{N}$, and the frequencies are integers, $f_k = k$, and can be computed at the cost of the FFT, $\mathcal{O}(N \log_2 N)$. In the literature there are many who have presented different ways to develop

a fast NUDFT ([59], [40], [57], [36] among others), generally referred to as NUFFT or NFFT. In this work though, the main focus is on the discrete cosine transform, and the NFFT methods will not be described in depth

Discrete Cosine Transform

The Fourier transform in any of its many forms is designed to process complex-valued signals, always producing a complex-valued spectrum, even for signals that were strictly real-valued. The real-valued or complex-valued part of the Fourier spectrum is on their own not enough to represent the full information of the signal, since neither the cosine nor the sine functions (corresponding to the real and the complex parts of the spectrum respectively), constitute a complete set of basis functions. Nonetheless, a purely real-valued signal has a symmetric Fourier spectrum, meaning that it is only necessary to compute half the number of spectral coefficients, without losing any signal information. Since the signals analyzed in this thesis are strictly real, one way to use this knowledge to improve on the works of this project is to consider a different, less expensive, purely real spectral transform. The cosine transform [2] seems to do the trick: it uses only cosine functions as basis functions and operates with only real-valued signal and spectral coefficients, and have properties similar to the Fourier transform.

For the discrete version of the cosine transform, DCT, and its inverse, IDCT, a number of different definitions have been proposed, but for this work, the originally formulations by [2] are used. These are often referred to as "The DCT" and "The IDCT", and other times as DCT-II and DCT-III. The entries of the computed discrete cosine transform vector, $\tilde{X}_0, \dots, \tilde{X}_{M-1}$, of a real-valued signal of N data points, X_0, \dots, X_{N-1} , is computed as

$$\tilde{X}_k = 2 \sum_{n=0}^{N-1} X_n \cos \left(\frac{\pi(2n+1)k}{2N} \right), \quad 0 \leq k < M. \quad (9)$$

To orthonormalize the base functions, $\phi_k(n)$, the coefficients are multiplied by a scaling factor f :

$$f = \begin{cases} \frac{1}{\sqrt{2N}}, & \text{if } k = 0 \\ \frac{1}{\sqrt{4N}}, & \text{otherwise} \end{cases} \quad (10)$$

so that the base functions, $\phi_k[n] = 2f \cos\left(\frac{\pi(2n+1)k}{2N}\right)$, meet the condition:

$$\sum_{n=0}^{N-1} \phi_k[n] \phi_l[n] = \delta_{lk}. \quad (11)$$

The inverse of the DCT, the so called DCT-III, is defined, unnormalized, as:

$$X_k = \tilde{X}_0 + 2 \sum_{n=1}^{N-1} \tilde{X}_n \cos\left(\frac{\pi n(2k+1)}{2N}\right), \quad 0 \leq k < N \quad (12)$$

and orthonormalized:

$$X_k = \frac{\tilde{X}_0}{\sqrt{N}} + \sqrt{\frac{2}{N}} \sum_{n=1}^{N-1} \tilde{X}_n \cos\left(\frac{\pi n(2k+1)}{2N}\right), \quad 0 \leq k < N \quad (13)$$

Only when the DCT-III is orthonormalized is it exactly the inverse of the orthonormalized DCT-II. If they are both unnormalized, the DCT-III is the inverse of the DCT-II up to a factor $2N$. As with the DFT, the DCT can directly be computed at a cost of $\mathcal{O}(N \cdot M)$, and can also be reduced to $\mathcal{O}(N \log N)$. The fast DCT algorithm(FCT) used here is based on [38] as it is implemented in the `scipy.fft.dct` package[68].

Nonuniform Discrete Cosine Transform

Again, as with the FFT, the FCT works under the key assumption that data is equispaced. Though when data is nonuniform, the DCT is described as:

$$\tilde{X}_k = 2 \sum_{n=0}^{N-1} X_n \cos\left(2\pi f_k \left(p_n + \frac{1}{2N}\right)\right), \quad 0 \leq k < M-1 \quad (14)$$

with, in the most general case, nonuniformly spaced signal, p_0, \dots, p_{N-1} , data and frequency data, f_0, \dots, f_{M-1} . The inverse of NDCT, the INDCT, is computed as:

$$X_k = \frac{\tilde{X}_0}{\sqrt{N}} + \sqrt{\frac{2}{N}} \sum_{n=1}^{N-1} \tilde{X}_n \cos\left(\left(p_n + \frac{1}{2N}\right) 2\pi f_k\right), \quad 0 \leq k < N-1 \quad (15)$$

It is possible to develop algorithms with the computational cost of $\mathcal{O}(N \log N)$ for NDCT and INDCT as it is for the NDFT ([65], [74]) but it has showed to be out of the scope of this project and has not been implemented. This of course slows down the final optimization algorithm, as it requires a number of spectral transformations. In Section ?? it is described how the final algorithm has been designed to minimize the use of NDCT, and thus speeding up the final computations.

APPENDIX V: Splines and Interpolations

Existence, Uniqueness and Conditioning

Considering any attempt to create an interpolant to fit a number of data points, the questions of uniqueness and existence is a matter of matching the data points with the number of parameters in the interpolant. If there are too few parameters, the interpolant does not exist, as it will not pass through all data points. If there are too many, the interpolant will not be unique. Formally this can be described through a system of linear equations.

For any data set consisting of (t_i, y_i) , $i = 1, \dots, m$ points, an interpolant can be chosen from a function space spanned by some suitable set of basis functions, $\phi_1(t), \dots, \phi_n(t)$. The interpolant can then be described as a linear combination of these basis functions:

$$f(t) = \sum_{j=1}^n x_j \phi_j(t) \quad (16)$$

The interpolant can then be found by determining the parameters x_j by requiring that the interpolant f must pass through the M data points (t_i, y_i) :

$$f(t_i) = \sum_{j=1}^n x_j \phi_j(t_i) = y_i, \quad i = 1, \dots, m \quad (17)$$

This can of course also be written compactly in matrix form as a system of linear equations:

$$\mathbf{A}\mathbf{x} = \mathbf{y} \quad (18)$$

In this equation \mathbf{A} is the $m \times n$ basis matrix, which entries consists of the value of the n basis functions evaluated at the m data points, $a_{ij} = \phi_j(t_i)$, the m vector \mathbf{y} consists of the known data values y_i , and the n vector \mathbf{x} consists of the unknown, to be determined, parameters x_j .

From linear algebra we know, that if we choose the number of basis function ot be equal to the number of data points, $n = m$, the basis matrix will be square, and thus - given the matrix is nonsingular - the system will be determined, and the data points can be fit exactly. Though in some problems it is beneficial to choose the system to be either overdetermined (less parameters than data points, the data cannot be fit exactly) or underdetermined (more parameters than data points, giving freedom to allow satisfaction of additional properties or conditions).

So the existence and uniqueness of an interpolant is given by the non-singularity of the basis matrix, be it square or not and the conditioning of the matrix

points to the parameters' sensitivity to perturbations. An ill-conditioned basis matrix will lead to high sensitivity in the parameters, but this problem can still be approximately solvable through Gaussian elimination with partial pivoting, but this solution will mean that the coefficients may be poorly determined.

Polynomial Interpolation

The most common way to determine an interpolant is through polynomials. Denoting a set of all polynomials of degree at most k , $k \geq 0$ as \mathbb{P}_k , it can be seen that this set forms a vector space of dimension $k+1$. The basis functions that span this vector space can be chosen to be composed of a number of different functions and this choice has a great influence on both the cost of computation and manipulation of the interpolant, and the sensitivity of the parameters, i.e. the conditioning of the basis matrix.

Considering n data points it is obvious to choose $k = n - 1$ so that the dimension of the vector space matches the number of data points. The maybe most natural choice of basis for \mathbb{P}_{n-1} is one that consists of the first n monomials¹,

$$\phi_j(t) = t^{j-1}, \quad j = 1, \dots, n. \quad (19)$$

Thus any given polynomial $p_{n-1} \in \mathbb{P}_{n-1}$ will be of the form

$$p_{n-1}(t) = x_1 + x_2 t + \dots + x_n t^{n-1}. \quad (20)$$

In this basis the system of $n \times n$ linear equations will be of the form

$$\mathbf{A}\mathbf{x} = \begin{bmatrix} 1 & t_1 & \dots & t_1^{n-1} \\ 1 & t_1 & \dots & t_1^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_1 & \dots & t_1^{n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \mathbf{y}. \quad (21)$$

This type of matrix with geometric progression, i.e. the columns are successive powers of some independent variable t is called a Vandermonde matrix. When using the monomial basis and using a standard linear equation solver to determining the interpolants coefficients requires $\mathcal{O}(n^3)$ work and often results in ill-conditioned Vandermonde matrices \mathbf{A} , especially for high-degree polynomials. This ill-conditioning is due to the monomials of higher and higher degree being more and more indistinguishable from each other. This

¹Roughly speaking, a polynomial with only one term.

makes the columns of \mathbf{A} nearly linearly dependent, resulting in almost singular matrices, and thus highly sensitive coefficients. For high enough n , the Vandermonde matrix becomes efficiently singular, to computational precision at least, though, as mentioned, this can be worked around, but requires some additional computational work.

Piecewise Polynomial Interpolation and Splines

The amount of work needed to solve the system as well as the conditioning of the system can be improved by using a different basis all together. Some different bases superior to the monomial that are worth mentioning are the Lagrange basis functions, the Newton basis functions and the orthogonal polynomials. But for this thesis we take a step further into the interpolation theory, as the choice of basis functions might not be enough to work around some of the problems connected with fitting a single polynomial to a large number of data points(i.e. oscillatory behaviour in the interpolant, nonconvergence or issues around the boundaries).

These practical and theoretical issues can be avoided through the use of piecewise polynomial interpolation, with the advantage that a large number of data points can be fitted with low-degree polynomials.

When turning to piecewise polynomial interpolation of the data points (t_i, y_i) , $i = 1, \dots, n$, $t_1 < t_2 < \dots < t_n$, a different polynomial is chosen for each subinterval $[t_i, t_{i+1}]$. Each point t_i , where the interpolant changes is called knots or control points. The simplest piecewise interpolation is piecewise linear interpolation, where each knot is connected with a straight line. If we consider this simple example it appears that by eliminating the problems of nonconvergence and unwanted oscillatory behaviour, the smoothness of the interpolant is sacrificed. This might be true for this simplistic example but since there are a number of degrees of freedom in choosing each piecewise polynomial interpolant, the smoothness can be reintroduced by exploiting a number of these measures. One way of doing this is by demanding knowledge of both the values and the derivatives of the interpolant at each data point. This just adds more equations to the system, and thus to have a well-defined solution, the number of equations must match the number of parameters. This type of interpolation is known as Hermite interpolation. The most common choice for this interpolation, to still maintain simplicity and computational efficiency, is cubic Hermite interpolation. This introduces a piecewise cubic polynomial with n knots, and thus $n - 1$ interpolants each with 4 parameters to fit, leading to $4(n - 1)$ parameters to be determined. Since each of the $n - 1$ cubics must match the data points at each end of the subinterval, it results in $2(n - 1)$

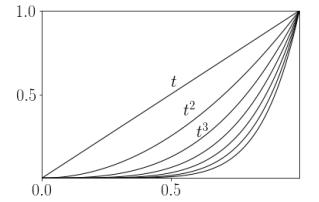


Figure 5: Illustration of the first eight monomials.

equations, and requiring the derivative to be continuous, i.e. match at the end points, an additional of $n - 2$ equations are taken in. This leads to a system consisting of $2(n - 1) + (n - 2) = 3n - 4$ equations to fit to the $4n - 4$ parameters. This leaves n free parameters, meaning that a cubic Hermite interpolant is not unique and the remaining free parameters can be used to accommodate further or additional constraints that might be around the problem at hand.

Cubic Spline Interpolation

A spline is a piecewise polynomial of degree k that is continuously differentiable $k - 1$ times.

One way of using the remaining free parameters is by introducing *splines*. A cubic spline is, given the spline definition, a piecewise cubic polynomial, a polynomial of degree $k = 3$, and must then be $k - 1 = 2$ times differentiable. Thinking back on the Hermite cubic, we were left with n free parameters. By demanding continuity of also the second derivative, we introduce $n - 2$ new parameters, leaving only 2 final parameters to be free. These 2 remaining parameters can be fixed through a number of different requirements, e.g. by forcing the second derivative at the endpoints to be zero, which leads to the *natural* spline.

The Hermite and spline interpolations are useful for different cases. The Hermite cubic might be more appropriate for preserving monotonicity if it is known that the data are monotonic. On the contrary, the cubic spline may enforce a higher degree of smoothness as it takes the second derivative into account as well.

For the case of this study, cubic spline interpolation is used to either evenly redistribute slightly unevenly sampled data or to enhance resolution for more precise peak detection. The general method for cubic spline interpolation used here is described in the following.

Assuming the original depth array \mathbf{d} is distributed as $d_{i-1} < d_i < d_{i+1}$ with $i = 0, \dots, n - 1$ has a minimum sampling distance as Δ_{\min} we define the new sampling distance for the new depth array $\hat{\mathbf{d}}$ as $\Delta = \Delta_{\min}$ - again assuming that $\hat{d}_{j-1} < \hat{d}_j < \hat{d}_{j+1}$ with $j = 0, \dots, \hat{n} - 1$. This makes it possible to define the first and last value of the new array as

$$\hat{d}_0 = \Delta \lceil \frac{d_0}{\Delta} \rceil \quad (22)$$

$$\hat{d}_{\hat{n}-1} = \Delta \lfloor \frac{d_{n-1}}{\Delta} \rfloor. \quad (23)$$

From this the number of values in the new array, \hat{n} , can be determined as

$$\hat{n} = 1 + \frac{\hat{d}_{\hat{n}-1} - \hat{d}_0}{\Delta}, \quad \hat{n} \in \mathbb{Z}. \quad (24)$$

Thus our new depth array will be given as

$$\hat{\mathbf{d}} = \hat{\mathbf{d}}_0 + j \cdot \Delta, \quad j = 0, \dots, \hat{n} - 1. \quad (25)$$

The original data are then used to define a cubic spline interpolation function to which the redistributed depth data points can be matched. For this part of the data analysis the `SciPy.interpolate` Python (REFERENCE) package with `SciPy.interpolate.CubicSpline` for the cubic spline interpolation.

APPENDIX VII: Diffusion Illustrated Through Synthetic Data

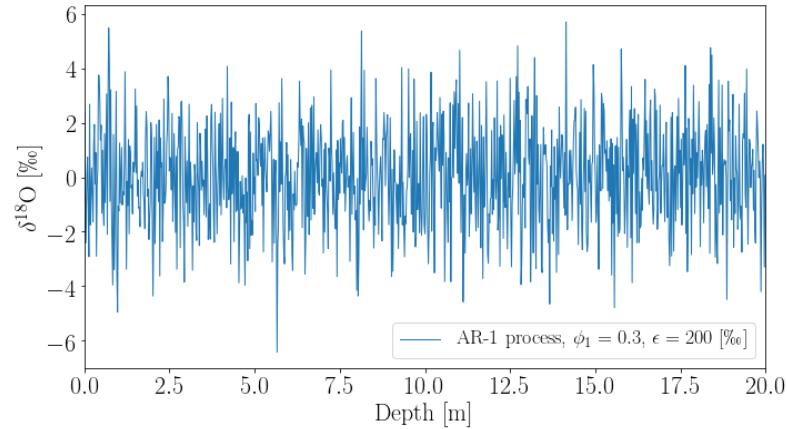


Figure 6: Illustration of modelled auto-regressive process of first order (AR-1).

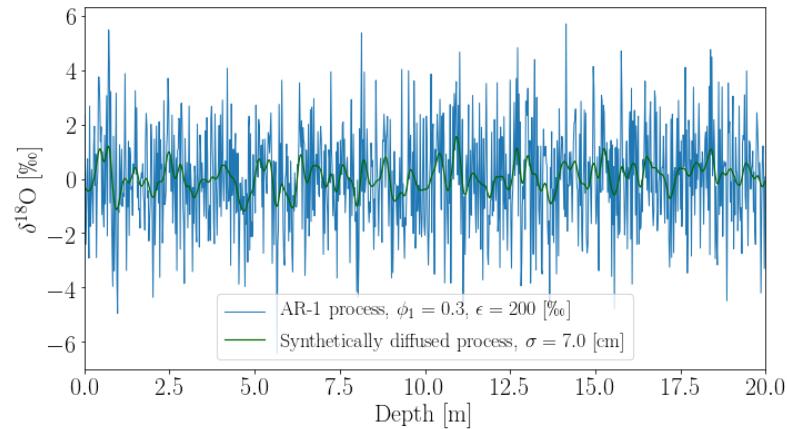


Figure 7: Illustration of AR-1 process and the same process diffused as in firn.

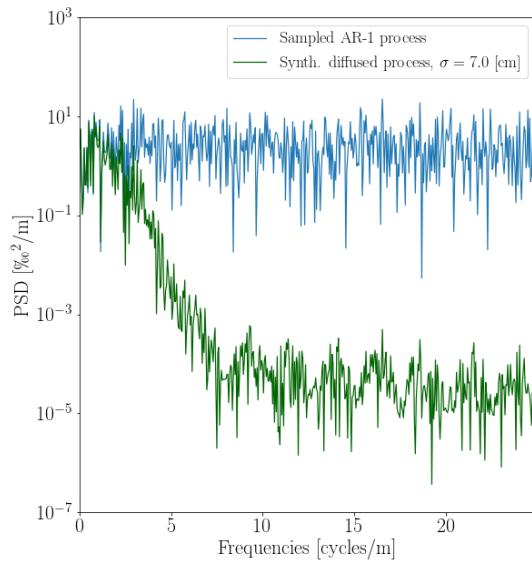


Figure 8: AR-1 process, raw and diffused, illustrated in the spectral domain.

APPENDIX IIX: Herron Langway Empirical Model

Sorge's law(REFERENCES) assumes that the relation between snow density ρ and depth h is invariant with time, given a constant snow accumulation and temperature. Furthermore, annual layer thinning by plastic flow is ignored.

Densification of firn, which can be described as the proportional change in air space, is linearly related to change in stress due to the weight of the overlying snow(REFERENCES):

$$\frac{d\rho}{\rho_i - \rho} = \text{const. } \rho dh \quad (26)$$

By integration, this implies a linear relation between $\ln \left[\frac{\rho}{\rho_i - \rho} \right]$ and h .

When considering real data, analysis shows that $\ln \left[\frac{\rho}{\rho_i - \rho} \right]$ vs h . plots have two linear segments(EXAMPLE), corresponding to the first and second stages of densification, with separation of segments at $\rho = 0.55$ and $\rho = 0.8$. These segments on the plots will yield two different slopes with slope constants:

$$C = \frac{d \ln \left[\frac{\rho}{\rho_i - \rho} \right]}{dh}, \rho < 0.55 \quad (27a)$$

$$C' = \frac{d \ln \left[\frac{\rho}{\rho_i - \rho} \right]}{dh}, 0.55 < \rho < 0.8 \quad (27b)$$

To find the densification rate, $\frac{d\rho}{dt}$, substitute $\frac{dh}{dt} = \frac{A}{\rho} \rightarrow dt = \frac{\rho}{A} dh$ and use the differentiation $\frac{\partial}{\partial t} \left[\ln \left[\frac{x(t)}{k-x(t)} \right] \right] = \frac{k \frac{dx}{dt}}{(k-x(t))x(t)}$

$$\begin{aligned} C &= \frac{\rho}{A} \frac{d \ln \left[\frac{\rho}{\rho_i - \rho} \right]}{dt} \\ &= \frac{\rho}{A} \frac{\rho_i}{\rho(\rho_i - \rho)} \frac{d\rho}{dt} \\ &= \frac{1}{A} \frac{\rho_i}{\rho_i - \rho} \frac{d\rho}{dt} \end{aligned}$$

leading to

$$\frac{d\rho}{dt} = \frac{CA}{\rho_i} (\rho_i - \rho) \quad (28a)$$

$$\frac{d\rho}{dt} = \frac{C'A}{\rho_i} (\rho_i - \rho) \quad (28b)$$

To continue from here two assumptions are made. The first is that the temperature and the accumulation rate dependencies may be separated, and that they thereby have no inter-correlation. The second is that the rate equations may be written as:

$$\frac{d\rho}{dt} = k_0 A^a (\rho_i - \rho), \rho < 0.55 \quad (29a)$$

$$\frac{d\rho}{dt} = k_1 A^b (\rho_i - \rho), 0.55 < \rho < 0.8 \quad (29b)$$

where k_0 and k_1 are Arrhenius type rate constants which are only temperature dependent, and a and b are constants determining the significance of the accumulation rate and are dependent on the densification mechanisms.

a and b may be determined by comparing slopes for densification at different sites of nearly equivalent conditions as:

$$a = \frac{\ln \left(\frac{C_1}{C_2} \right)}{\ln \left(\frac{A_1}{A_2} \right)} + 1 \quad (30)$$

and equivalently for b, with C'_1 and C'_2 .

k_0 and k_1 can be estimated by observing values of k at different temperatures and plotting $\ln(k)$ versus temperature - a so-called Arrhenius plot(REFERENCES) - to find A and E_a in equations:

$$k = Ae^{-\frac{E_a}{k_B T}} = Ae^{-\frac{E_a}{RT}} \quad (31)$$

$$\ln(k) = \ln(A) - \frac{E_a}{R} \frac{1}{T}$$

leading to values of k_0 and k_1 of:

$$k_0 = 11e^{-\frac{10160}{RT}} \quad (32a)$$

$$k_1 = 575e^{-\frac{21400}{RT}} \quad (32b)$$

Depth-density and depth-age calculations

Assuming that temperature, annual accumulation rate and initial snow density are known, the following calculations can be made:

- Density at depth h , $\rho(h)$
- Depth at pore close-off, $\rho = 0.55$
- Depth-age relationship from surface to pore close-off (stage 1 and 2).

1. stage of densification: Depth-density profile:

$$\rho(h) = \frac{\rho_i Z_0}{1 + Z_0} \quad (33)$$

where $Z_0 = e^{\rho_i k_0 h + \ln\left[\frac{\rho_0}{\rho_i - \rho_0}\right]}$. In this segment, the depth-density is independent of accumulation rate. The critical density depth can be calculated as:

$$h_{0.55} = \frac{1}{\rho_i k_0} \left[\ln \left[\frac{0.55}{\rho_i - 0.55} \right] - \ln \left[\frac{\rho_0}{\rho_i - \rho_0} \right] \right] \quad (34)$$

and the age at close-off depth as:

$$t_{0.55} = \frac{1}{k_0 A} \ln \left[\frac{\rho_i - \rho_0}{\rho_i - 0.55} \right] \quad (35)$$

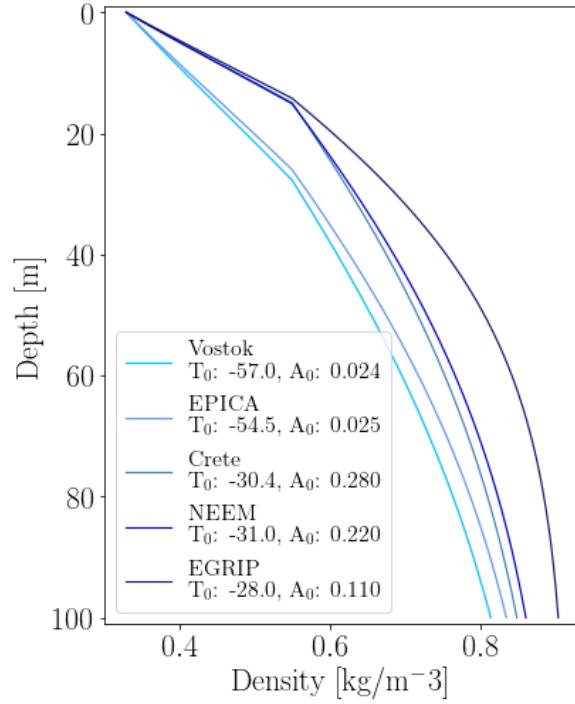


Figure 9: Density profile examples given five different initial conditions representing present day conditions at the five different ice core locations. Temperature, T_0 , is in $^{\circ}\text{C}$ and accumulation, A_0 , is in meter of water equivalent per year.

2. stage of densification: The depth-density profile

$$\rho(h) = \frac{\rho_i Z_1}{1 + Z_1} \quad (36)$$

where $Z_1 = e^{\rho_i k_1 (h - h_{0.55}) \frac{1}{A^{0.5}} + \ln \left[\frac{0.55}{\rho_i - 0.55} \right]}$. The age of firn at a given density ρ :

$$t_\rho = \frac{1}{k_1 A^{0.5}} \ln \left[\frac{\rho_1 - 0.55}{\rho_1 - \rho} \right] \quad (37)$$

An estimate of the mean annual accumulation rate can be made from the slope C' and the mean annual temperature:

$$A = \left(\frac{\rho_i k_1}{C'} \right)^2 \quad (38)$$

APPENDIX IX: Iso-Community Firn Model

Firn Diffusivity

The Iso-CFM framework contains a number of different new modules added to the CFM. A specific module for calculation of the firn diffusivity, $D(\rho(z))$, is provided, containing several different methods for the calculations of the individual parameters contained in the calculation of the diffusivity constant, using the formulation in [30, Johnsen et al., 2000]:

$$D_i(z) = \frac{m p D_{\text{air}}}{R T(z) \alpha_{s/v}^i \tau(z)} \left(\frac{1}{\rho(z)} - \frac{1}{\rho_{\text{ice}}} \right) \quad (39)$$

with $i \in \text{O}^{18}, \text{O}^{17}, \text{D}$ representing the three different types of water isotopic ratios generally examined. The different terms in 39 each describes the following:

m : molecular weight in [kg]

$R = 8.314478 \left[\frac{\text{m}^3 \text{Pa}}{\text{K mol}} \right]$: molar gas constant

T Temperature [K]

p : saturation vapor pressure over ice in [Pa]

The saturation vapor pressure over ice can be calculated in two different ways, as in [47, Murphy & Koop 2005]:

$$p = \exp \left(28.9074 - \frac{6143.7}{T} \right) \quad (40)$$

$$p = \exp \left(9.5504 - \frac{5723.265}{T} + 3.530 \ln(T) - 0.0073 T \right) \quad (41)$$

where Eq. 41 takes the temperature dependence of the latent heat of sublimation of ice into account when integrating the Clausius-Clapeyron equation. A third expression is presented in [29, Johnsen et al., 2000] as:

$$p = 3.454 \cdot 10^{12} \exp \left(\frac{-6133}{T} \right) \quad (42)$$

which will be the one used for analytical calculations of diffusion length in this project.

D_{air} : diffusivity of water vapor in air,

calculated from $P_0 = 1\text{Atm}$, $T_0 = 273.15\text{K}$, T temperature in [K] and P ambient pressure in [Atm], as in [23, Hall and Prupacher, 1976]:

$$D_{\text{air}} = 2.1 \cdot 10^{-5} \left(\frac{T}{T_0} \right)^{1.94} \left(\frac{P_0}{P} \right) \quad (43)$$

From [42, Merlivat, 1978] the additional diffusivity of water isotopes ratios for ^{18}O and ^2H vapor were defined as

$$D_{\text{air}^2\text{H}} = \frac{D_{\text{air}}}{1.0251} \quad (44)$$

$$D_{\text{air}^{18}\text{O}} = \frac{D_{\text{air}}}{1.0285} \quad (45)$$

$\alpha_{s/v}^i$: solid-to-vapour fractionation factor. $i = ^{18}\text{O}, \text{D}, ^{17}\text{O}$,

For both $\alpha_{s/v}^{18}$ and $\alpha_{s/v}^2$ there are multiple options for parameterisation of the fractionation factor. Considering $\alpha_{s/v}^{18}$, one can choose between [37, Majoube 1971] and [14, Ellehøj et al., 2013], respectively as:

$$\ln(\alpha_{s/v}^{18}) = \frac{11.839}{T} - 28.224 \cdot 10^{-3} \quad (46)$$

and

$$\ln(\alpha_{s/v}^{18}) = 0.0831 - \frac{49.192}{T} + \frac{8312.5}{T^2} \quad (47)$$

For $\alpha_{s/v}^2$ the parameterisation from [43, Merlivat and Nief, 1967], [14, Ellehøj et al., 2013] or [33, Lamb et al., 2017], respectively as:

$$\ln(\alpha_{s/v}^2) = \frac{16288}{T^2} - 9.45 \cdot 10^{-2}, \quad (48)$$

$$\ln(\alpha_{s/v}^2) = 0.2133 - \frac{203.1}{T} + \frac{48888}{T^2} \quad (49)$$

or

$$\ln(\alpha_{s/v}^2) = \frac{13525}{T^2} - 5.59 \cdot 10^{-2}. \quad (50)$$

The parameterisation of the fractionation factor related to the ^{17}O water isotopic ratios is based on [4, Barkan and Luz, 2005] as $\alpha_{s/v}^{17} = 0.529\alpha_{s/v}^{18}$.

For a comparison of the different parameterisations see [18, Gkinis et al., 2021]. The default choices in the iso-CFM modules is [37, Majoube 1971] for $\alpha_{s/v}^{18}$ and [43, Merlivat and Nief, 1967] for $\alpha_{s/v}^2$

τ : firn tortuosity

In [30, Johnsen et al., 2000] a parameterisation of the firn tortuosity was presented as:

$$\frac{1}{\tau} = \begin{cases} 1 - b_\tau \left(\frac{\rho}{\rho_{ice}} \right)^2, & \rho \leq \frac{\rho_{ice}}{\sqrt{b_\tau}} \\ 0, & \rho > \frac{\rho_{ice}}{\sqrt{b_\tau}} \end{cases} \quad (51)$$

where $b_\tau = 1.3$ and $\rho_{ice} = 917 \frac{\text{kg}}{\text{m}^3}$, implying for $\frac{1}{\tau} = 0$ a close-off density of $\rho_{co} = 804.3 \frac{\text{kg}}{\text{m}^3}$. This close-off density refers to the density at the depth where diffusive fluxes stop and $\frac{D_{air}}{D_{eff}} \rightarrow \infty$. Different parameterisations have been suggested, some defined from the total porosity, but for this project, the expression used is the one given in Eq. 51.

Diffusion Length Profiles

The iso-CFM computes a numerical solution for σ using a time-stepping scheme, as is the case for the original CFM, to estimate the most likely diffusion length profile at a given site. From each time step the CFM computes $\frac{d\rho}{dt}$ and T , and the iso-CFM uses these results to calculate the quantity $\frac{d\sigma^2}{dt}$:

$$\frac{d\sigma^2}{dt} = 2 \left(D(t) - \frac{\sigma^2}{\rho} \frac{d\rho}{dt} \right) \quad (52)$$

Eq. 52 shows that the diffusion length signal throughout the ice is a result of two processes, opposing each other: the always positive diffusivity term $D(t)$, and the densification process contributing negatively to the change over time, $-\frac{\sigma^2}{\rho} \frac{d\rho}{dt}$. After a certain depth, the densification term comes to dominate and thus the entire equation becomes negative and the value of the diffusion length is decreasing, see Figure 10.

To simplify the work of this thesis, the numerical module of the CFM and the iso-CFM has not been implemented in the final computations, and the diffusion length profiles referred to in the rest of the project are calculated through an analytical method, using equations derived from Eq. 52 analytically. A short walk-through of the derivations will be presented here

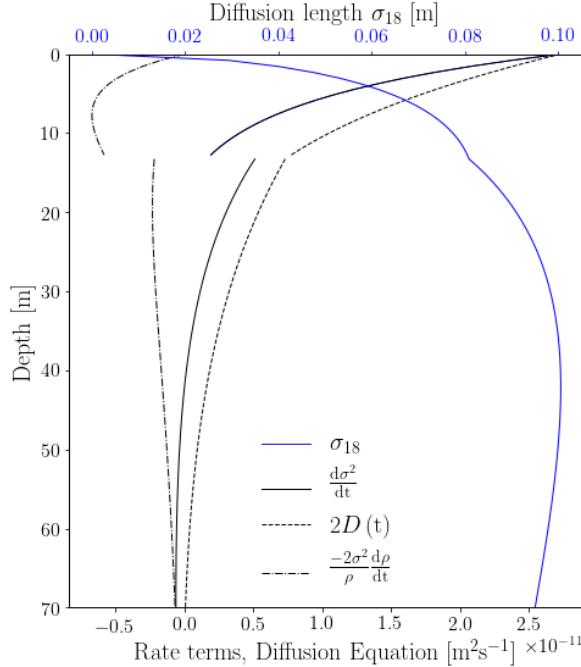


Figure 10: Contribution of the diffusion(dashed) and densification(dot-dashed) terms from Eq. 52 to the final analytical diffusion length solution (blue).

as they are described in [18, Gkinis et al., 2021]. By substitution of variables rearrangement Eq. 52 becomes:

$$\frac{d\sigma^2}{d\rho} + \frac{2\sigma^2}{\rho} = 2 \left(\frac{d\rho}{dt} \right)^{-1} D(\rho) \quad (53)$$

which can be converted to integral form:

$$\sigma^2(\rho) = \frac{1}{\rho^2} \int_{\rho_0}^{\rho} 2\rho'^2 \left(\frac{d\rho'}{dt} \right)^{-1} D(\rho') d\rho' \quad (54)$$

Then, by using the densification rate parameterisation given in [25, Herron and Langway, 1980], the expression becomes:

$$\frac{d\rho(z)}{dt} = k(T) A^\nu (\rho_{ice} - \rho(z)), \quad (55)$$

where $k(T)$ is an Arrhenius-type densification rate constant, dependent on temperature and densification zone described by:

$$k_0(T) = 0.011 \exp \left(-\frac{10160}{RT} \right), \quad \nu_0 = 1 \quad (56)$$

in the upper densification zone, $\rho < 550\rho_{\text{co}}$. In the lower densification zone, $\rho \geq \rho_{\text{co}}$, it is described as:

$$k_1(T) = 0.575 \exp\left(-\frac{21400}{RT}\right), \quad \nu_1 = 0.5. \quad (57)$$

Using the parameterization of the diffusivity coefficient from Eq. 39 and expressing the term $1/\tau = 1 - b_{\text{tau}} \left(\frac{\rho}{\rho_{\text{ice}}}\right)^2$ in densities, the diffusivity coefficient can be described as a function of density:

$$D_i(\rho) = \frac{m p D_{\text{air}}}{R T \alpha_{s/v}^i} \left(1 - b_\tau \left(\frac{\rho}{\rho_{\text{ice}}}\right)^2\right) \left(\frac{1}{\rho} - \frac{1}{\rho_{\text{ice}}}\right). \quad (58)$$

By then inserting in Eq. 54, defining $\frac{m p D_{\text{air}}}{R T \alpha_{s/v}^i} = \zeta$, and integrating the final analytical equations for the diffusion length in upper and lower densification zones can be obtained:

$$\sigma^2(\rho < \rho_{\text{co}}) = \frac{\zeta}{\rho^2 k_0 A^{\nu_0} \rho_{\text{ice}}} \left[\rho^2 - \rho_0 - \frac{b_\tau}{2\rho_{\text{ice}}^2} (\rho^4 - \rho_0^4) \right] \quad (59)$$

$$\begin{aligned} \sigma^2(\rho \geq \rho_{\text{co}}) = & \frac{\zeta}{\rho^2 k_1 A^{\nu_1} \rho_{\text{ice}}} \left[\rho^2 - \rho_{\text{Cr}} - \frac{b_\tau}{2\rho_{\text{ice}}^2} (\rho^4 - \rho_{\text{Cr}}^4) \right] \\ & + \frac{\zeta}{\rho^2 k_0 A^{\nu_0} \rho_{\text{ice}}} \left[\rho_{\text{Cr}}^2 - \rho_0 - \frac{b_\tau}{2\rho_{\text{ice}}^2} (\rho_{\text{Cr}}^4 - \rho_0^4) \right] \end{aligned} \quad (60)$$

The analytical equations have been used for creating a contour plot of the analytical solutions for σ_{18} at the close-off density, ρ_{co} .

These analytical equations are used to compute diffusion lengths to compare with the optimal diffusion length estimates computed from the raw data. One could advantageously spend some time and energy on using the iso-CFM to numerically compute the comparison diffusion lengths with different temperature and accumulation forcing to recreate a diffusion length profile corresponding to the largest likelihood at a given drill site. Since the iso-CFM do consist of many different modules all with different possibilities for parameterisation, it is outside the scope of this project to develop a iso-CFM diffusion length estimate. In-depth methodology and results from the iso-CFM can be found in [18, Gkinis et al., 2021].

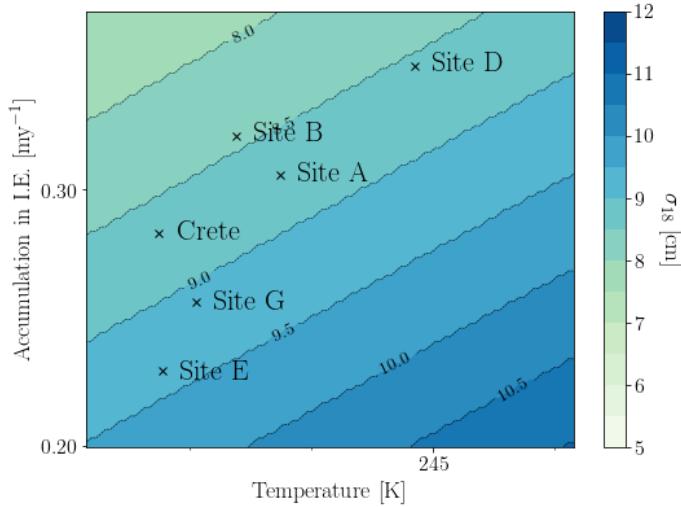


Figure 11: Crete and surrounding Alphabet cores, as their analytical solutions place them according to observed temperature and accumulation rate.

APPENDIX X: Dielectric Profiling

In this thesis has previously been presented a method for measuring electrical conductivity in ice cores, which is used for locating volcanic events. A different method was later developed to demonstrate how both acids and salts play a decisive role in the determination of the electrical behavior of ice. The dielectric response of an ice core can be used to determine the total ionic concentration of the core. For ECM the measurements are sensitive to the fluctuating distance between ice core and electrodes, and after each measurement a fresh piece of ice needs to be prepared to repeat a measurement.

A new dielectric profiling technique (DEP) was developed (REFERENCES) with the advantages over the ECM that no direct contact is needed between the electrodes and the ice, so that the ice can stay in a protective polythene sleeve and the experiment easily can be repeated on the same piece of ice. Together the ice core and the polythene sleeve creates a complete system, where the plastic acts as an electrical blocking layer.

The dielectric response is measured by a sweeping of the AF-LF frequency range for the entire ice-polythene system. At LF the conductivity of the composite system is within a few percentages of the intrinsic behavior of the ice itself. At HF-VHF frequencies it also approximates well enough (REFERENCES)

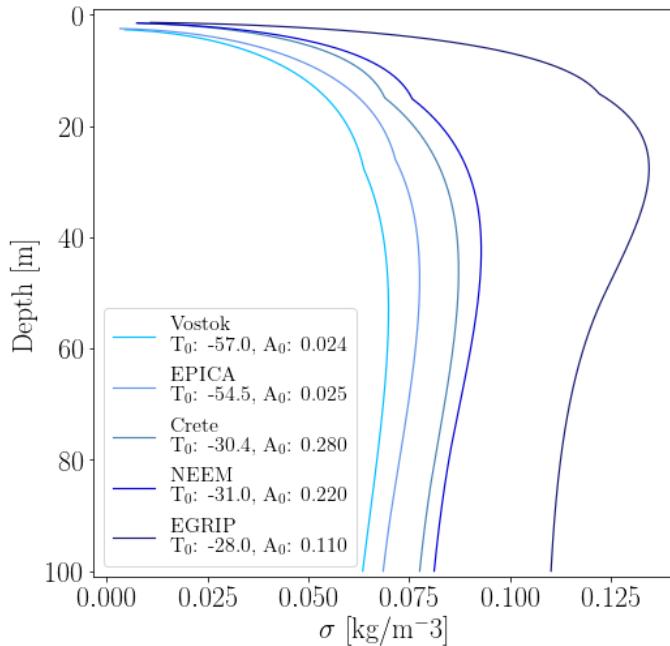


Figure 12: Analytically calculated diffusion length profile examples given five different initial conditions representing present day conditions at the five different ice core locations. Temperature, T_0 , is in °C and accumulation, A_0 , is in meter of water equivalent per year.

ENCES).

The measured dielectric parameters are the conductivity of ice at HF-VHF range, denoted σ_∞ where ∞ signifies a frequency much higher than the relaxation frequency, f_r , of the dominant dispersion in the system. Both of these parameters display clear chemical response signals which can be used either alone or in combination with other ice core analysis measurements like ECM and isotope analysis.

If the core under analysis is chemically analyzed for Na^+ , Mg^{2+} , Cl^- , SO_4^{2-} and NO_3^- , a number of important parameters, which can be used to evaluate the response of the dielectric parameters, can be calculated(REFERENCES):

- The salt parameter, which represents the total marine cation concentration calculated with the assumed marine ratios as:

$$[\text{salt}] = 1.05([\text{Na}^+] + [\text{Mg}^{2+}]) \quad (61)$$

- XSO_4 , the excess sulphate, which represents the amount the sulphate concentration is above the expected if the salt and sulphate ions were in normal sea salt ratios. Excess sulphate is essentially sulphuric acid, which is the main acidic component of the ice.
- The strong acid content of the ice has been calculated as (assuming no other ions present in significant quantities):

$$[\text{acid}] = [\text{Cl}^-] + [\text{SO}_4^{2-}] + [\text{NO}_3^-] - 1.05([\text{Na}^+] + [\text{Mg}^{2+}]) \quad (62)$$

From data, it can be seen that acid and salt concentration peaks clearly affect σ_∞ and f_r (EXAMPLES, REFERENCES). The relationship between salt and acid, and the two dielectric parameters have been derived through non-linear regression analysis. In PAPER(REFERENCES) the linear responses for the DEP at -22°C were:

$$\sigma_\infty = (0.39 \pm 0.01)[\text{salt}] + (1.43 \pm 0.05)[\text{acid}] + (12.7 \pm 0.3) \quad (63)$$

with 76.6 % variance

$$f_r = (440 \pm 11)[\text{salt}] + (612 \pm 65)[\text{acid}] + (8200 \pm 400) \quad (64)$$

with 68.4 % variance. σ_∞ is measured in $\mu\text{S}/\text{m}$, f_r in Hz and [acid] and [salt] in $\mu\text{Eq}/\text{l}$. The total ionic concentration of the ice core is strongly linked to the dielectric parameters, and a regression between the total anion concentration and the dielectric parameters gives:

$$[\text{anions}] = [\text{salt}] + [\text{acid}] = 0.022\sigma_\infty^{1.89} + 10^{-6}f_r^{1.61} - 0.2 \quad (65)$$

with 86.7 % variance.

The DEP complements the ECM technique by not only reacting to acids alone, as ECM does, but responds to both neutral salts and acids. The acid term is here associated with the DC conductivity, the same way it is also detected by ECM. The dielectric dependence on salts is consistent with the Bjerrum L defect² affecting every one or two salt ions in the ice, indicating that a large fraction of the neutral salt is incorporated into the ice lattice.³

³A Bjerrum defect is a crystallographic defect specific to ice, partly responsible for the electrical properties of ice. Usually a hydrogen bond will normally have one proton, but with a Bjerrum defect it will have either two protons (D defect) or no proton (L defect). (REFERENCES)

The sensitivity to salt concentrations allows for identifications of periods with major storms and open seas which are also important identifiers for paleo climate research, along with the volcanic eruption detection made possible through the ECM.

