University of Potsdam
Institute of Earth and Environmental Science

Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research
Section Periglacial Research
Potsdam

Bachelor's Thesis

^{10}Be concentrations in an ice core from Akademii Nauk (Russian Arctic) for validation of the age-depth relationship: development of a sample scheme

Luisa von Albedyll

Potsdam, September 2015

Email:	luisa.von.albedyll@gmx.de
Registration number:	766121
Degree Program:	Geosciences
Persons in Support:	Dr. Diedrich Fritzsche
	Prof. Dr. Ulrike Herzschuh

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und eigenhändig sowie ohne unerlaubte fremde Hilfe und ausschließlich unter Verwendung der aufgeführten Quellen und Hilfsmittel angefertigt habe. Diese Arbeit wurde in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt oder veröffentlicht.

Potsdam, den

.....

Abstract

This bachelor's thesis aims to develop a sample scheme for the Akademii Nauk ice core from the Russian Arctic in order to reconstruct the long-term variations of the ^{10}Be concentration. These long-term variations are assumed to vary globally simultaneous and are therefore used to synchronize different ^{10}Be records for the purpose of dating ("wiggle matching"). This is done in order to validate an existing age-depth relationship of this ice core covering a time span of around 3 000 years. For this purpose a numerical computer model has been designed that simulates different modes of systematic sampling on high-resolution age–^{10}Be concentration records. Two sample modes have been simulated: One mode that takes discrete, short samples and another mode that takes continuous, long samples. The model takes into account the restrictions of the Akademii Nauk ice core, for example dating uncertainties and occasional missing parts of the core. All data sets are smoothed using a 22 years running mean. This averaging smooths out most of the meteorological differences, reduces aliasing effects and still enables the reconstruction of the decadal variations. For a comparison between the original and the reconstructed curve, three complementary evaluation techniques have been applied: a mathematical neighborhood, the Pearson correlation coefficient, and the cosine similarity. The model's results show clearly that continuous sampling performs better than taking only short samples leaving big gaps in between. This can be related to the fact that when sampling continuously no information is lost and aliasing effects are attenuated by "averaging" via melting of long samples. Regarding the sample length, the closest match of reconstructed and original curve is achieved for sample lengths of 4–8 years, with a slight preference for lengths of 8 years. Another optimum is reached for sample lengths of 22 years. The results give a clear hint that aliasing effects bias the reconstructed record. Apart from that, even when parts are missing in the ice core, a good reconstruction is still possible applying continuous sampling. For further sampling I suggest to identify a prominent part of the ^{10}Be long-term concentration variations and fill up the gaps between the already existing data points with continuous samples within the identified part.

Zusammenfassung

Ziel dieser Bachelorarbeit ist es, ein Beprobungsschema für den Eisbohrkern von Akademii Nauk aus der russischen Arktis zu entwickeln, welches es ermöglicht, langzeitlichen Variationen von ^{10}Be zu rekonstruieren. Da von langzeitlichen Veränderungen angenommen wird, dass sie ausreichend global und gleichzeitig geschehen, kann man den Kurvenverlauf anderen ^{10}Be Konzentrations-Alter Datensätzen zuordnen und so den Tiefen ein Alter zu ordnen („Wiggle Matching“). Im vorliegenden Fall soll ein bereits bestehendes Altersmodell, welches eine Zeitspanne von ungefähr 3 000 Jahren umfasst, validiert werden. Zu diesem Zweck ist ein numerisches Computermodell entwickelt worden, welches verschiedene Arten des systematischen Beprobens an zeitlich hochaufgelösten ^{10}Be Datensätzen testet. Zwei verschiedene Modi wurden dabei getestet: zum einen diskretes Beproben mit sehr kurzen Proben und zum anderen kontinuierliches Beproben mit aneinander angrenzenden Proben. Das Modell berücksichtigt dabei die speziellen Einschränkungen des Akademii Nauk Kernes, wie z.B. Datierungsunsicherheiten und die Tatsache, dass vereinzelt Teile fehlen. Alle Datensätze werden mit Hilfe eines gleitenden Durchschnitts auf 22 Jahre geglättet. Dadurch wird gewährleistet, dass die durch Witterung und lokale Gegebenheiten induzierten Unterschiede zwischen verschiedenen Kernen ausgeglichen und Aliasing-Effekte reduziert werden. Außerdem ermöglicht eine Mittlung über 22 Jahre immer noch die Rekonstruktion der langzeitlichen Variationen auf einer Zeitskala von Jahrzehnten. Zum Vergleich von Original- und rekonstruierter Kurve wurden drei Evaluationsstrategien angewendet, die sich gegenseitig ergänzen: eine mathematische Umgebung, der Korrelations-Koeffizient und die Cosinus-Ähnlichkeit. Die Ergebnisse des Modells zeigen deutlich, dass kontinuierliches Beproben bessere Ergebnisse produziert als diskretes Bepoben. Das liegt hauptsächlich daran, dass durch das kontinuierliche Beproben keine Informationen verloren gehen und Aliasing-Effekte bereits durch das Aufschmelzen, was einem Mitteln gleich kommt, reduziert werden. Bezüglich der Probenlänge ergibt sich die höchste Übereinstimmung zwischen Original- und rekonstruierter Kurve für Probenlängen von 4–8 Jahren, mit einer leichten Tendenz zu 8 Jahren. Darüber hinaus erreicht man auch mit Proben mit einer Länge von 22 Jahren eine sehr gute Rekonstruktion. Die Ergebnisse weisen darauf hin, dass Aliasing-Effekte die Rekonstruktion verzerren. Zusätzlich zeigt sich, dass eine Rekonstruktion auch dann möglich ist, wenn der Kern unvollständig ist. Für zukünftiges Beproben schlage ich vor, einen Bereich der langzeitlichen Variationen zu identifizieren, der viele charakteristische Variationen hat, und in der entsprechenden Tiefe die bereits bestehenden Datenpunkte mit kontinuierlicher Beprobung aufzufüllen.

Contents

List of Figures

Definitions

For a better understanding, the following terms are defined for this thesis in advance so that this chapter might be used as a glossary while reading. The terms explained here are printed italic in the text when first mentioned. For a visualization see figure 3.1 and figure 4.1.

A **sampling method** describes the rules on which the selection of samples is based on. There are probabilistic and non probabilistic sample methods. For example, systematic sampling (taking samples at a constant spacing) is a probabilistic sampling method.

In this thesis **sample mode** describes whether sampling is discrete or continuous. The discrete sample mode **"single point sampling"** refers to samples with almost no extent. The ^{10}Be concentration at one exact point is measured. In tangible terms, single point samples have the length of the smallest unit that was chosen in the computer model.
In the continuous sample mode **"'multi year sampling"**, long, extended samples are taken that are lined-up without any gaps in between. So, the end point of a sample is the start point of the next sample.

The **sample period** T_s is the temporal unit that is between two samples. The respective inverse is the sample frequency in time f_s. For example, taking a sample every 10 years results in a sample period of $T_s = 10$ years. This term can be applied for both sample modes. For multi year sampling the sample period describes the temporal unit between the center points of two samples.

Since sample period can be used for both sample modes, sample step and sample length are introduced to specify sample period within these two modes.
A **sample step** s refers to the intended distance between two single point samples.
The **sample length** l is used to describe the intended length of a continuous sample, in other words, its extent.

In this thesis the sample period normally is biased by adding a random number to simulate a dating uncertainty. In order to distinguish between the originally intended sample period and the biased one, the latter is defined as **sample interval**. A sample interval is calculated by adding sample period and the random number that symbolizes the dating uncertainty.

A **sample run** refers to one complete sampling of the original record with a specified sample mode and sample period.

A **sample scheme** includes all relevant information of sampling. It could be also understood as a sample strategy. It consists of the target of sampling and the sample framework, as well as the sampling method, sample mode, sample period and important information about restrictions of sampling and practical issues [Harris and Jarvis, 2011].

1 Introduction

Due to recent man-made climate change, research on former paleoclimate changes is nowadays one of the most important aims in climatology. Ice cores are one of the best archives for studying paleoclimate because they provide information about climate forcing factors (e.g. greenhouse gases, aerosols and volcanic dust, solar irradiance) and also about the climate response (e.g. temperature, precipitation, atmospheric circulation) [Beer, 2001]. However, a complete interpretation of the climate proxies stored in ice requires reliable dating with a high temporal resolution. Usually, dating of ice cores from regions with high accumulation rates like Greenland is based on the seasonal variation of parameters like $\delta\ ^{18}O$ or visual ice properties [Berggren et al., 2009]. If these seasonal signals are missing, modeling accumulation rates, snow compaction and ice flow can provide least a rough estimation of the age-depth relationship. Apart from that, stratigraphic dating based on signals that are considered to be globally simultaneous, for example volcanic eruptions, can validate the other dating methods [Raisbeck et al., 1998].
In the present case of the ice core from Akademii Nauk dating with the above mentioned methods led to contradictory results [Fritzsche, 2015 personal communication]. More details of dating the Akademii Nauk ice core are discussed in Opel et al. (2013). Therefore, another method called "wiggle matching" or "curve fitting" is applied to validate one of the existing age-depth models. Wiggle matching is based on the variations of the cosmogenic radionuclide ^{10}Be. For this, ^{10}Be concentrations are measured and synchronized to an already dated curve by visual or mathematical methods resulting in an age for the measured ice core [Muscheler et al., 2014].
For successful wiggle matching sampling needs to be planned carefully because a lot of specific characteristics of the ice core can falsify the sampling results. For example the periodicity of ^{10}Be, the partly absence of reliable yearly patterns in the Akademii Nauk ice core, blurring of the signal due to summer melting and occasionally missing parts of the ice core are likely to biases the ^{10}Be record. The design of a *sample scheme* that allows the reconstruction the variations of ^{10}Be concentrations under the above mentioned conditions is the aim of this bachelor's thesis.

1.1 Theoretical Background of ^{10}Be Production, Transportation and Deposition

As indicated above, wiggle matching is based on the variations of ^{10}Be concentration. The sources of these variations are numerous. Some sources generate more local variations, whereas others have a global influence. Since wiggle matching uses records from different sites, it is only successful when it is based on the global variations. In order to extract them when sampling, it is crucial to understand the origin of the global variations and the factors that might introduce a more local component.
Therefore, in the following the processes that lead to the variations in the ^{10}Be concentration are

discussed: namely the production, transportation, deposition and storage of ^{10}Be [Masarik, 2009].

^{10}Be is a radioactive isotope of $^{9}_{4}$Be (half life 1.51 $\pm 0.06 \times 10^4$ years) that is mainly **produced** in the atmosphere [McHargue and Damon, 1991]. Today the global mean production rate is 0.018 $\frac{atoms}{cm^2 \cdot s}$ [Beer, 2000].
For the sake of completeness, it should be mentioned that there is also in situ production of ^{10}Be in terrestrial or extraterrestrial material as rocks and meteorites. This is not subject of this thesis and therefore ^{10}Be is only referred to as the cosmogenic radionuclide that is produced in the atmosphere.

^{10}Be is formed by spallation, a naturally occurring nuclear fission, of oxygen and nitrogen with high-energy, secondary neutrons. These neutrons are formed by cosmic rays mainly coming from outside of our solar system [Masarik, 2009], [McHargue and Damon, 1991]. Since the cosmic ray intensity is inferred to be constant, it is excluded as source for variations of ^{10}Be concentration in an archive [Masarik, 2009]. However, not all cosmic rays reach the earth's atmosphere, because of magnetic fields shielding the earth.
Both, the magnetic field embedded in the solar wind and the geomagnetic field deflect the primary flux of charged galactic particles that lead to a reduction of ^{10}Be production [Bard et al., 2000], [Peristykh and Damon, 1998]. So, a high solar activity with strong solar winds as well as a high geomagentic field intensity result in a low ^{10}Be production rate and vice versa. These two factors arise steady variations in the ^{10}Be concentration.

However, the shielding effect of the geomagnetic field is not globally constant, but depends on the latitude. Since in high latitudes the geomagnetic field lines are curved and almost perpendicular to the surface, they allow high-energy particles to travel mostly unrestricted. This results in a higher production rate and in an independence of the geomagnetic field modulations in polar regions [Finkel, R.C. and Nishiizumi, K., 1997]. Conversely, when the influence of the geomagnetic field is low, the solar modulation effect is strongest [Beer, 2000].
Additionally, the ^{10}Be production drops with lower altitude because particles lose energy after producing generations of particles on their way through the atmosphere [Heikkilä et al., 2013], [Masarik, 2009].

In contrast to the well understood processes of production **transportation and deposition** of ^{10}Be still bear great uncertainties and are likely to intrude local components to the signal. On the other hand atmospheric mixing during transport can smooth out local production differences.

The influence of transport depends mainly on the place of origin. Two thirds of the ^{10}Be production occurs in the stratosphere and one third in the troposphere. The dynamic of these two parts of the atmosphere is very different and so is the particle transport. In the troposphere extensive convective activities take place due to decreasing temperature with increasing height. This leads to a short residence time (some weeks) of ^{10}Be in the troposphere that does not allow a global mixing of the air. Therefore, one must assume that ^{10}Be produced in the troposphere reflects more a local than a global signal, because differences introduced by the geomagnetic field are not smoothed out.

On the other hand, the higher stability of the stratosphere leads to much longer transport ways and residence times of about one to two years. Since atmospheric mixing is assumed to be complete in the stratosphere within a year, one considers the latitudinal differences in production as smoothed out for ^{10}Be produced in the stratosphere. Putting these two parts together, this results in the rough guess that the larger part of the production differences are smoothed out [Heikkilä et al., 2013]. However, when comparing records one should not neglect the fact that there are latitude differences and analyze the origin of the air masses.

Not only transport but also deposition of ^{10}Be can falsify the signal from a global to a more local one. Deposition occurs as wet deposition via precipitation (80-90 %) or as dry deposition (10-20 %) depending on the amount of precipitation [Masarik, 2009]. During wet deposition the cosmogenic nuclid is scavenged by submicron particles as e.g. sulfates that act as condensation nuclei for the formation of cloud droplets and eventually coagulate to form precipitation [Masarik, 2009]. During dry deposition particles settle down directly from the air, depending on the aerodynamic resistance of the surface, namely surface properties and local winds. Sudden changes in the fraction of wet to dry deposition might produce concentration changes that are only introduced by local climate changes and can rise the noise in the signal. But Masarik (2009) has shown that this is a minor factor and the original production signal is still preserved. For a detailed discussion on transport and deposition of ^{10}Be see also Heikkilä et al. (2011).
As a consequence, the deposition process determines the concentration because areas with higher precipitation consequently show lower ^{10}Be concentrations. Typical ^{10}Be concentrations in glacial ice range between e.g. 10^5 $\frac{atoms}{g}$ (Dome Fuji, Antarctica [Horiuchi et al., 2007], [Horiuchi et al., 2008]) and 2 $\times 10^4$ $\frac{atoms}{g}$ (Greenland, Dye-3 [Beer et al., 1990] and Akademii Nauk [Fritzsche, 2015 personal communication]). Since decay counting is not feasible due to the long half life, measuring such low concentrations requires very sensitive detection techniques as accelerator mass spectrometry (AMS) [Beer, 2001].

Finally **postdepositional altering** of the ^{10}Be concentrations in storage due to summer melting and consequential percolation must be considered. The effects of melting on ^{10}Be are not well investigated, therefore only assumptions can be made. At first, it is known that ^{10}Be is quite often part of a cloud condensation nuclei and therefore located to some extent in the interior of ice grains. This makes percolation of ^{10}Be less likely [Pohjola et al., 2002]. Secondly, Weiler et al. (2005) found for the Akademii Nauk ice core that at least the averaged ion concentrations of conservatives species like sulphate are representative for long-time changes in the atmospheric concentration. Since, ^{10}Be is often attached to sulphate [McHargue and Damon, 1991], it can be deduced that possible changes due to percolation are compensated by applying a running mean. Weiler et al. (2005) suggests a 5 years running mean to smooth out percolation effects for the conservative species.

To sum up, the ^{10}Be concentration variation is composed of a global production signal modulated by solar and geomagnetic shielding and a local system signal, influenced by meteorological

and local factors as latitude and altitude [Steinhilber et al., 2012]. Various studies have found that after averaging to 20-40 years the fraction of the system effects remains only to be 20-30 % of the complete ^{10}Be signal [Abreu et al., 2013], [Horiuchi et al., 2008], [McCracken, 2004], [Steinhilber et al., 2012], [Vonmoos et al., 2006]. For instance, Steinhilber et al. (2012) used a principal component analysis applied on ice core and tree ring data that have been averaged with a 22 running mean and found that only 30 % of the data might be due to system effects. Apart from the fact that averaging helps to suppress system effects, McCracken et al. (2004) made clear that variations that are relevant for wiggle match dating ($> 30\%$) are large compared to the meteorological influence.
Therefore it is concluded that for wiggle matching one needs to extract the global production signal and suppress the system effects.

1.2 Solar Cycles in the ^{10}Be Archives

After having identified the sources of the variations that are relevant for wiggle matching, the character of these variations is studied.
The variations of the production signal are composed of periodic and non-periodic variations. Especially the periodic variations associated with the solar cycles are a strong evidence for the fact that despite all uncertainties the overall global production signal is preserved in the ^{10}Be records of natural archives [Heikkilä et al., 2013].

Periodicities in the ^{10}Be archives are induced by a periodic solar modulation or by periodic changes of the magnetic field. A helpful tool to distinguish between both effects is the span of the periodicities. Long periodicities (e.g. 1000-2000 years) are commonly associated with variations in the geomagnetic field [Beer, 2001]. Short term fluctuations (decades to centuries) in the ^{10}Be concentration are assigned to the solar modulations like for example the prominent 11-year Schwabe cycle (respectively the 22 year Hale cycle), the 90-year Gleissberg cycle or the 205-year DeVries cycle [Beer, 2001]. Primarily these periodicities reflect changes in solar parameters e.g. the sunspot number, but since the ^{10}Be production rate is so strongly correlated with the solar winds, these periodicities are also present in the ^{10}Be archives with a time lag of about a year corresponding to the mean residence time of ^{10}Be in the air [Beer, 2001]. The solar cycles do not have a fixed periodicity. For example the 11-year Schwabe cycle varies between 7–18 years [Beer, 2000]. This decadal variation is quite dominant in the archive. During a typical 11 years solar cycle the ratio between production at solar minimum and solar maximum is 1.34, whereas the non-periodic variations fluctuate on longer time scales [Masarik, 2009].

The non-periodical variations are mainly caused by an almost complete absence of sunspots, the so-called grand solar minima. This results in ^{10}Be values 30–50% above the modern value [Bard et al., 2000] for approximately 60–100 years. Prominent grand solar minima are e.g. the Maunder minima (1645–1715) and the Dalton minimum (1790–1830). It is still part of the discussion

whether the 11-year cycle is still present, weaker or changed during these times [Beer, 2001], [Peristykh and Damon, 1998].

1.3 Wiggle Matching as Dating Tool

Furthermore, it is important to know the technique of wiggle matching to understand its requirements before sampling.

The general idea of using a ^{10}Be record as dating tool relies on the fact that the production signal stored in the local ^{10}Be archives is global and therefore comparable to different coring sites. Meeting these requirements a ^{10}Be record of unknown age is synchronized with a ^{10}Be record of known age, called calibration or reference curve, by fitting it to the shape of characteristic parts of the reference curve. This technique, called wiggle matching, is already widely used in ^{14}C dating in order to improve the accuracy of ^{14}C ages. There are different approaches to wiggle matching. First of all, a simple visual, but subjective matching is possible. Mostly a peak-to-peak correlation strategy is used, skipping few peaks with unclear relations (e.g. [Horiuchi et al., 2008]) Secondly, ad hoc statistical methods based on least squares are applied [Christen and Litton, 1995]. Thirdly Bayesian mathematics provides a complex, but well-founded way to solve the problem [Christen and Litton, 1995]. Bayesian analysis differs from classical statistics in that concepts of uncertainty and probability are interrelated and allows in that manner to implement prior information to the matching process [Walker, 2005]. Fourthly, computer programs as OxCalc that are designed for curve matching using statistical methods are available online [Bronk Ramsey, 1998]. Finally a recent study made by Muscheler et al. (2014) presents a new approach based on Monte Carlo sampling.

Especially for visual peak-to-peak correlation it is important that the original shape of the curve is preserved as good as possible. Therefore, an appropriate temporal resolution depending on the length of the whole record as well as the needed temporal accuracy. The approximate time span of this ice core of 3 000 years should be matched with other records based on long-term variations of time scales of the grand solar minima with a typical duration of 60-100 years [Beer, 2000]. In order to detect them at least a temporal resolution of 20 years seems to be appropriate.

1.4 Research Question and Outline of the general Approach

After an introduction to the theoretical basics of wiggle matching based on ^{10}Be concentrations, the focus of this thesis is summarized in the research question:
What sample scheme is most suitable in order to estimate the long-term variations of ^{10}Be concentration in the ice core of Akademii Nauk for validating the age-depth relationship?

This question is broken down into:

- What must be included in the **sample scheme**?
- How are the **long-term variations of ^{10}Be concentration** characterized and how to extract them from a high temporal resolution record?
- Which restrictions need to be discussed when sampling this particular **ice core from Akademii Nauk**?
- In which time interval does the **existing age-depth relationship** need to be validated?

In the following, these questions are partly discussed based on literature, but the central aspect of the sample scheme, the **question concerning *sample mode* (discrete or continuous sampling) and *sample period*** is so complex, that a numerical computer model that simulates sampling has been developed to answer these questions. In order to run this model the following steps have been conducted:

- Use mathematics to set up a model that conducts sampling with different sample modes and periods.
- Modify the model to reflect sampling the Akademii Nauk ice core as realistic as possible.
- Simulate sampling on three high resolution ^{10}Be concentration data sets.
- Transfer the best sample mode and period to the conditions of the Akademii Nauk ice core.

In the discussion, the results of the computer simulations are interpreted in regard to the best sample mode and period. In the following, the gap between these theoretical findings and the practical implementation of them is discussed and finally, the sample scheme developed in this study is compared to the approaches of other authors.

2 Study Site and Ice Core Parameters

The Akademii Nauk ice core was cored at 80.52°N, 94.82°E, on Severnaya Zemlya, an archipelago located in the Russian Arctic that is covered with considerable ice caps. The core was drilled close to the summit of the cold glacier Akademii Nauk [Fritzsche et al., 2005] in an elevation of about 750 m a.s.l. In contrast to Greenland drilling projects the site is located in the percolation zone assuming annual infiltration of melt water and even rain due to high insolation and temperatures above 0°C in July and August. There are no extensive studies on the percolation depth, but it is assumed that infiltration deeper than two or three annual layers is unlikely due to ice lenses stopping deeper penetration [Opel et al., 2009]. Furthermore, a modern accumulation rate of 0.46 m water equivalent was investigated based on stable water isotope analyses by Fritzsche et al. (2005). Due to the high precipitation it is assumed that most of the deposition of ^{10}Be takes place in form of wet precipitation [Heikkilä et al., 2013].

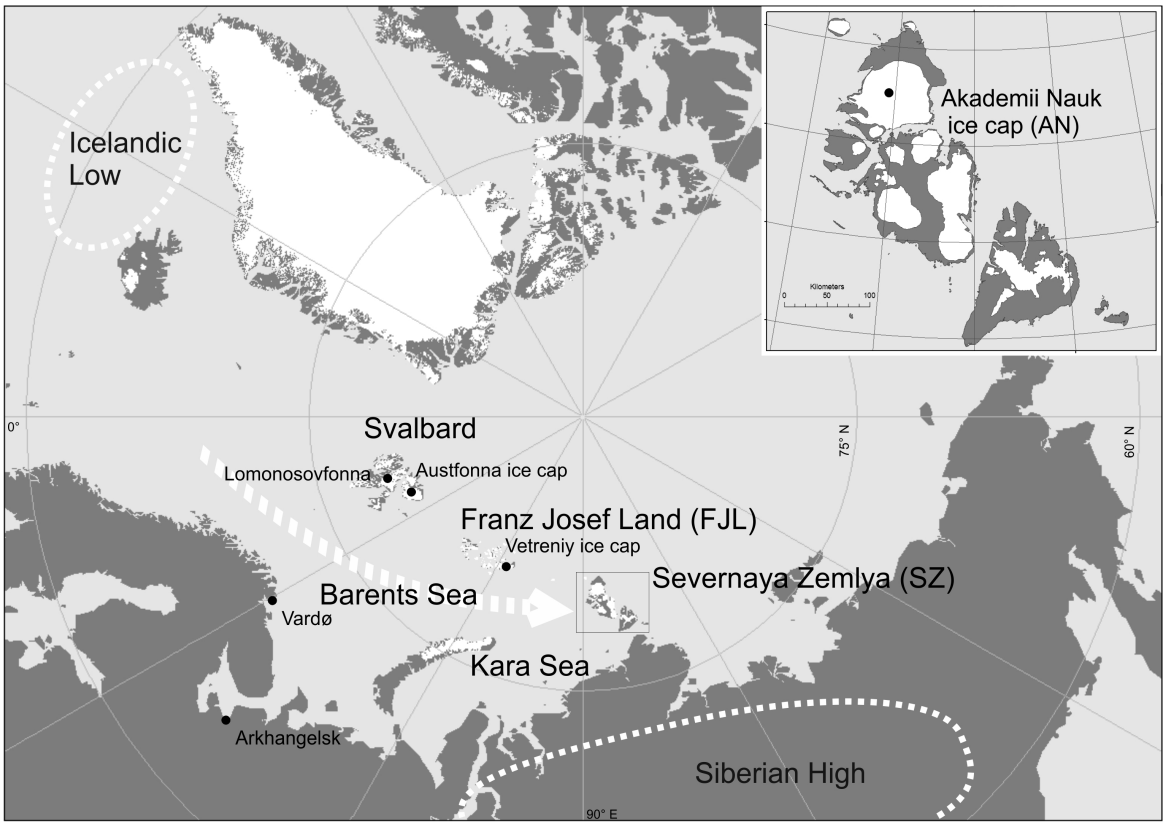

Figure 2.1: **Map of the Arctic with focus on Severnaya Zemlya.** In the inset a detailed map of Severnaya Zemlya is shown with the location of Akademii Nauk, where the ice core was taken 1999-2001 [Fritzsche et al., 2005]. Moreover the main atmospheric circulation system influencing the study site are displayed. The Icelandic Low and Siberian High are marked, as well as the prevailing air transport way (reprinted from Opel et al. (2009) with permission).

Even though the influence of system effects on the ^{10}Be archive is generally estimated to be less than 30 %, it is important to take the atmospheric conditions into account [Abreu et al., 2013]. For example, dramatical changes in the fraction of wet to dry deposition may result in a shift in the ^{10}Be concentration that is not applicable to production rate changes [Heikkilä et al., 2013]. In fact, Severnaya Zemlya is located at a transition zone, namely between the Atlantic influenced and the more continental eastern parts of Siberia. Most of the 400 mm precipitation is caused by cyclones coming from the Kara Sea with southern and south-western winds. However, the origin of the air-masses passing the site changes within the year (for details see [Opel et al., 2009]). Opel et al. (2013) found possible shifts in the atmospheric circulation patterns in the last centuries, but there is no evidence that the amount of precipitation has changed dramatically through the Holocene [Miller et al., 2010].

Ice Core Parameters

The core is stored at $-30°$C, cut in pieces of around 100 cm, in Bremerhaven. Many melt-layers are visible in the core, as well as a lot of air bubbles. This suggests that there has been a lot of melting, but most melt layers consists not only of refrozen melt water, but also ice originated from compacted snow [Fritzsche et al., 2005]. The ice is mostly free of visible impurities as dust or sediment particles, an observation that was confirmed after melting the ice. This is important, because there is a positive correlation between dust content and ^{10}Be concentration that could bias the results [Merchel 2015, personal communication]. Occasionally, short parts of the ice core (1-100 cm) are not available for ^{10}Be sampling due to the partly bad core quality. Furthermore, yearly markers as stable-isotope layers or visual differences are not completely reliable for the entire ice core, but at least for the youngest 275 years [Fritzsche et al., 2005].

3 Methods

3.1 Numerical Simulation of Sampling

In order to answer the aspect of the research question about the best sample mode (discrete or continuous) and period, a computer model has been developed that simulates sampling on high resolution age–^{10}Be concentration data sets.
The modeling has been done with Matlab©, a numerical computing environment, that allows data analysis, implementation of algorithms, and plotting of data.

Preparation of the Input Data

Different sample modes and periods have been tested on three ^{10}Be records consisting of age–^{10}Be concentration sets, namely from NGRIP [Berggren et al., 2009], Dye-3 [Beer et al., 1990], [Berggren et al., 2009] and Dome Fuji [Horiuchi et al., 2008], [Horiuchi et al., 2007]. They have been chosen because they have a high temporal resolution and cover a similar time span as the Akademii Nauk ice core (see table 3.1).
Before these data sets have been used in the model, they had to be **interpolated** on a regular temporal grid of a temporal resolution of 0.1 years. This has been necessary because the data is not completely regular distributed and on a lower temporal resolution, but the model is designed to work with regular spaced data. In the following, not only years, but also data points will be discussed. A transformation is simply done by multiplying with (dividing by) 10. For instance, an interval of 33 data points corresponds to 3,3 years.

	location and altitude	**temporal resolution of ^{10}Be record**	**time span**	**reference**
NGRIP	Greenland $75°6'N$ $42°19'W$ 2917 m a.s.l.	1 year	1995–1389 AD	[Berggren et al., 2009]
Dye-3	Greenland $65°11'N$ $43°49'W$ 2479 m a.s.l.	0.4–2.3 years	1986–1424 AD	[Beer et al., 1990], [Berggren et al., 2009]
Dome Fuji	Antarctica $77°30'S$ $37°30'E$ 3810 m.a.s.l.	mean: 9 years range: 1–22 years	1876–695 AD	[Horiuchi et al., 2008] [Horiuchi et al., 2007]

Table 3.1: input data sets of the computer model

Sampling Process

In general sampling is described mathematically as [Mari et al., 1999]:

$$X(i,j) = D_x(i,j) * S(m) \text{ with } \quad S(m) = \begin{cases} 1 & , \quad m = nT_s \\ 0 & , \quad \text{other} \end{cases} \tag{3.1}$$

where m and n are integers, S the sample function, and $X(i,j)$ is the output data set consisting of values of the input data set $D_x(i,j)$ located at multiples of the sample period T_s. For near-point measurements all information in between the sample points is not considered. Equation 3.1 lists all components that are needed to simulate sampling:

- an output data set, that hold the results of the sampling, consisting of age ($\mathbf{X}$(i,j)) and ^{10}Be concentrations ($\mathbf{Y}$(i,j)).
- an input data set $\mathbf{D}$ that is sampled, likewise consisting of age ($\mathbf{D}_x$) and ^{10}Be concentrations ($\mathbf{D}_y$)
- a sample function $\mathbf{S}$

The **output data sets** ($\mathbf{X}$(i,j), $\mathbf{Y}$(i,j)) are $\{n \times 11\}$ matrices. This is due to the fact that every sample run, so one complete sampling of a record, is reproduced for 11 different starting points (p). These starting points are 10 data points (1 year) apart from each other. This was performed to rule out biases due to interference phenomena between the sample period and the natural periodicity in the data set. The output $\mathbf{X}$(i,j) contains all information of the "location" for sampling expressed as age beginning from the youngest. The output matrix $\mathbf{Y}$(i,j) is built up likewise, containing the corresponding ^{10}Be concentrations.

In order to simplify Matlab© calculations, the **input data sets** ($\mathbf{D}_x$, $\mathbf{D}_y$) are transformed as well to $\{n \times 11\}$ matrices. Consequently, the input matrix $\mathbf{D}_x$ is built up of elven times the same column vector containing information about age. The input matrix $\mathbf{D}_y$ is built up of eleven times the same column vector containing the corresponding information about ^{10}Be concentrations.

The structure of input and output matrices should not be mixed up with the **sample function S** that is also a $\{n \times 11\}$ matrix, but whose entries refer to the index i of the input matrices. Each entry of S consists of three parts that are summed up.

- the endpoint of the last interval
- the sample period
- a random number expressing the dating uncertainty.

The first entry of S "endpoint of the last interval" points to the recursive character of the sampling. This guarantees that there are no accidental gaps in the sampling.

A *sample interval* is defined as the sum of the sample period and the dating uncertainty.

The sample period describes the amount of samples per time unit. Different periods between 10 and 220 data points have been tested.

And lastly, a dating uncertainty $\{n \times 11\}$ matrix $\mathbf{R}$ has been created based on the random number function of Matlab©. $\mathbf{R}$ contains equally distributed integer numbers between -10 and 10. By adding or subtracting those random numbers to the sample period, varying sample intervals are simulated as they occur in reality when dating is uncertain.

$$\mathbf{R} = \begin{pmatrix} r_1 & r_1 & ... & r_1 \\ r_2 & r_2 & ... & r_2 \\ ... & ... & ... & ... \\ r_n & r_n & ... & r_n \end{pmatrix} \qquad \mathbf{R}_{example} = \begin{pmatrix} 4 & 4 & ... & 4 \\ 1 & 1 & ... & 1 \\ -3 & -3 & -3 & -3 \\ ... & ... & ... & ... \end{pmatrix} \tag{3.2}$$

with $r_i \in \mathbb{Z}$ and $-10 \leq r \leq 10$

This general introduction to the composition of the sample function S is specified in the following when explaining the two different samples modes, *single point sampling* and *multi year sampling* (figure 3.1).

Single point sampling has been conducted as systematic sampling, but not continuously. Samples are treated as data points with almost no extension. In tangible terms, the samples have the length of the smallest unit that is present in the data, so 0.1 years. For single point sampling, the sample period is referred to as *sample step* s that has been chosen to be $s = \{20, 30, 40, ..., 70, 100, 110\}$ data points. For example, a sample run tests sampling with samples, covering 0.1 years, that are taken every 3 years (s=30).

So, in the case of single point sampling, each entry of S_{single} is expressed by the sum of the endpoint of the last interval, the sample step s (=sample period), and the dating uncertainty r. As starting point, so the first entry of S, $p = \{10, 20, ..., 110\}$ has been defined.

$$\mathbf{S}_{single} = \begin{pmatrix} 10 & 20 & ... & 110 \\ 10 + s + r_1 & 20 + s + r_1 & ... & 110 + s + r_1 \\ 10 + 2s + r_1 + r_2 & 20 + 2s + r_1 + r_2 & ... & 110 + 2s + r_1 + r_2 \\ ... & ... & ... & ... \\ 10 + (n-1)s + \sum_{z=1}^{z=n} r_z & 20 + (n-1)s + \sum_{z=1}^{z=n} r_z & ... & 110 + (n-1)s + \sum_{z=1}^{z=n} r_z \end{pmatrix} \tag{3.3}$$

Finally, sampling is performed by picking out the values form the input matrices that correspond to the entries of S_{single} (see figure 3.2). This is expressed by

$$X_{single}(i,j) = D_x(S_{single}) \qquad Y_{single}(i,j) = D_y(S_{single}) \tag{3.4}$$

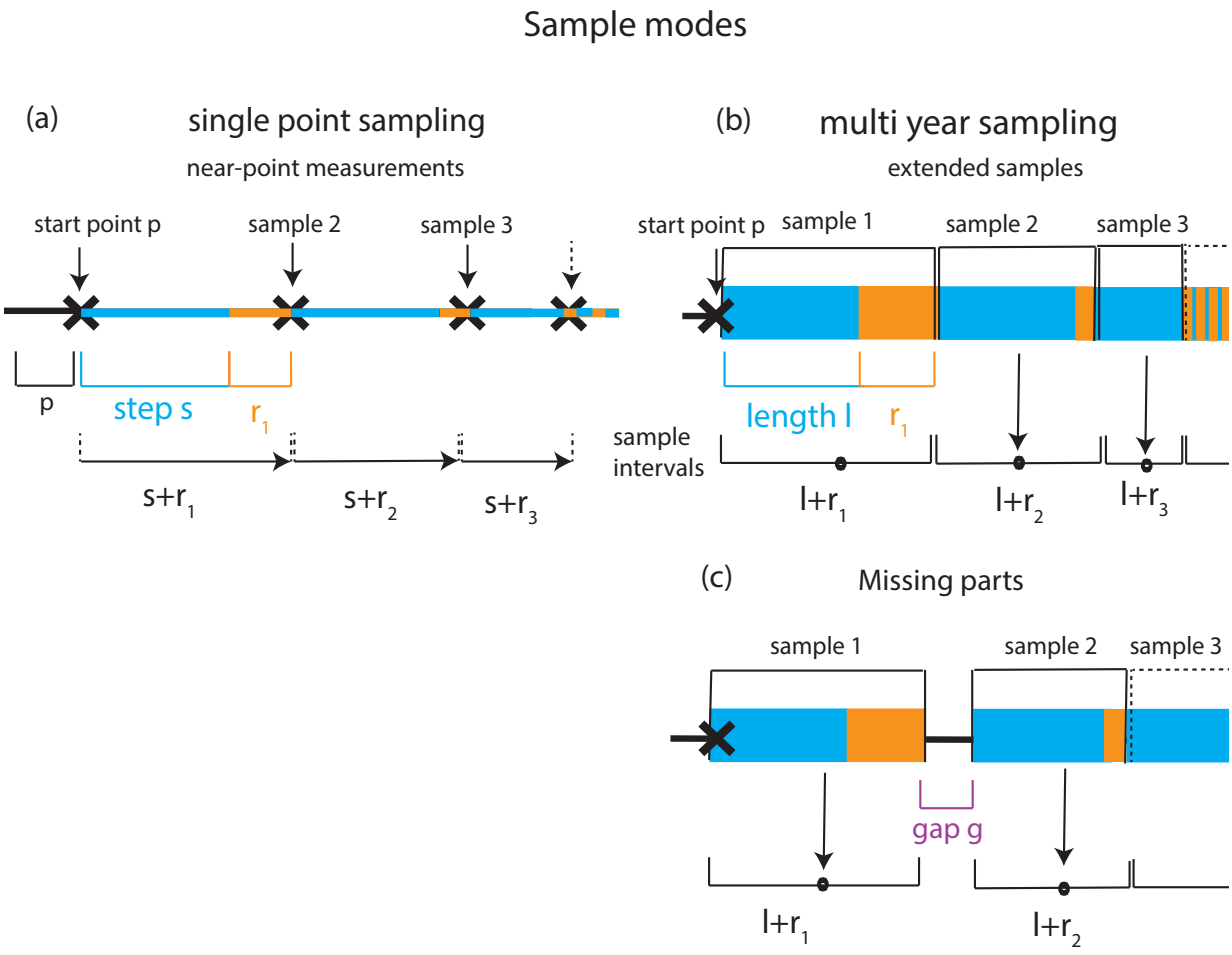

Figure 3.1: **Schematic representation of the sample modes**. (a) On the left side single point sampling is displayed. Sampling starts at the starting point p where the first sample is taken as a near-point measurement (indicated by black crosses). After one sample interval, composed of the sample step s and the uncertainty r_1, the next sample is taken. (b) Multi year sampling is shown on the left side. Unlike single point sampling, a sample, that has the length of the whole sample interval, is taken. The average concentration of this whole sample is assigned to the center point of the sample (indicated by the point in the middle of the sample intervals). (c) A subtype of multi year sampling is sampling with "missing parts" that simulates a gap g after a sample interval.
In all these sample modes the sample intervals are not constant because the dating uncertainty r_n varies between -10 and 10. In this example $r_1 = 4, r_2 = 1, r_3 = -3$. The third sample interval is an example of an interval that is shorter than the intended sample period because r_3 is negative. This is indicated by the dashed line.

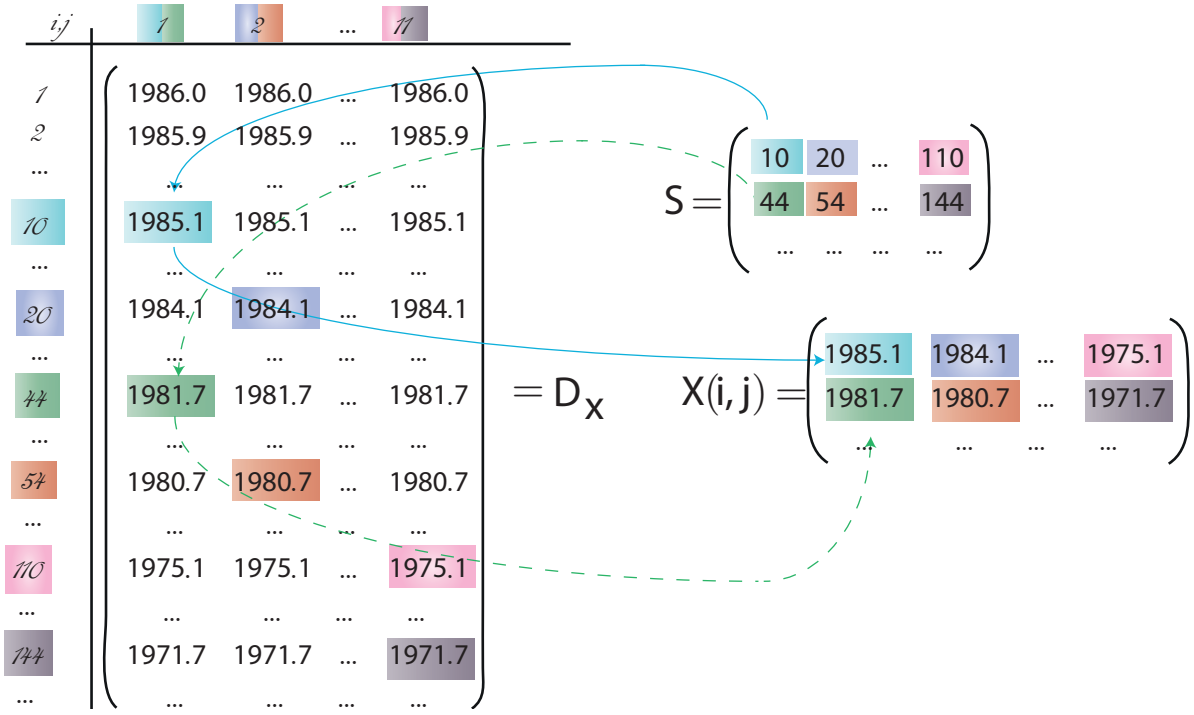

Figure 3.2: **Visualization of the functionality of the sample matrix S** expressed by $X(i,j) = D_x(S)$. An example is given of a sample run with Dye-3 data. The sample period is 30 and the first random number added is r_1=4. D_x is built up of 11 identical row vectors containing age informations in years AD. The italic numbers at the left and the top of the matrix describe the indices of this matrix. The entries of sample matrix S refer to these indices of D_x. The values of D_x corresponding to these indices are picked and stored in the output matrix X(i,j). For example, the first entry "10" in column 1 refers to the entry in row 10, column 1 of the D_x matrix. The entry is 1985.1 and that is why one can find this entry in the output matrix X(i,j). Although this is expressed for age values, this can transferred to ^{10}Be concentrations as well: $Y_{single}(i,j) = D_Y(S)$

The **multi year sampling** differs in that point from the single point sampling that it was designed to be continuous sampling with samples of a specific length of $l = \{20, 30, ..., 110, 220\}$. Since the *sample length* determines the amount of samples per time unit, it is understood as the sample period of multi year sampling.

When applying multi year sampling the average of one sample is assigned to its center point. These center points are stored in the output matrix $\mathbf{X}_{Multi}(i,j)$. The center points are calculated using the same three components of S as mentioned above.

So, first of all, the next sample interval $(i+1)$ is calculated by adding sample length l (=sample period) and dating uncertainty r_{i+1}. Then, the half of the last sample interval $(\frac{l+r_i}{2})$ and the half of the next sample interval $(\frac{l+r_{i+1}}{2})$ are added to the center point of the last interval. Consequently, $S_{multi\ X}$ is given by

$$\mathbf{S}_{multi\ X} = \begin{pmatrix} 10+\frac{l+r_1}{2} & 20+\frac{l+r_1}{2} & ... & 110+\frac{l+r_1}{2} \\ 10+l+r_1+\frac{l+r_2}{2} & 20+l+r_1+\frac{l+r_2}{2} & ... & 110+l+r_1+\frac{l+r_2}{2} \\ ... & ... & ... & ... \\ 10+(n-1)l+\sum_{z=1}^{z=n-1} r_z+\frac{l+r_n}{2} & 20+(n-1)l+\sum_{z=1}^{z=n-1} r_z+\frac{l+r_n}{2} & ... & 110+(n-1)l+\sum_{z=1}^{z=n-1} r_z+\frac{l+r_2}{2} \end{pmatrix} \tag{3.5}$$

In a next step, $\mathbf{X}_{multi}(i,j)$ is calculated by

$$\mathbf{X}_{multi}(i,j) = D_x(S_{multi\ X}) \tag{3.6}$$

(for visualization see figure 3.2). The output matrix $\mathbf{Y}_{multi}(i,j)$ is filled with average values, so the mean (my) of all y-values building up a sample interval is calculated. For this purpose, at first all points of an interval are identified (table 3.2, column 2). Then the corresponding ^{10}Be concentrations are picked out of $\mathbf{D}_y(i,j)$ (table 3.2, column 3). Finally the mean of the concentrations is calculated.

	S:index of point in interval		D:^{10}Be concentration	mean
first point:	$p+(n-1)l+\sum_{z=1}^{z=n-1} r_z$	–>	$D_y(p+(n-1)l+\sum_{z=1}^{z=n-1} r_z)$	mean = my
next point:	$p+1+(n-1)l+\sum_{z=1}^{z=n-1} r_z$	–>	$D_y(p+1+(a-1)l+\sum_{z=1}^{z=-1} r_z)$	
...	...	–>	...	
last point:	$p+nl+\sum_{z=1}^{z=n} r_z-1$	–>	$D_y(p+nl+\sum_{z=1}^{z=n} r_z-1)$	

Table 3.2: steps of the calculation for my: All points of an interval are identified (column 1 and 2), the corresponding ^{10}Be concentrations are picked out (column 3) and averaged to my.

The steps done in table 3.2 are summarized in equation 3.7:

$$my = \frac{D(p+(n-1)l+\sum_{z=1}^{z=n-1} r_z + D(p+1+(n-1)l+\sum_{z=1}^{z=n} r_z) + ... + D(p+nl+\sum_{z=1}^{z=n} r_z-1)}{l+r_n} \tag{3.7}$$

For example the $my(i)$, at i=204 with l=40 and $r_i = 4$ is calculated by

$$my(204,2) = \frac{D(182)+D(182)+...+D(226)}{44}$$

Furthermore, a modification of the multi year sampling process has been developed. As indicated in chapter 2 there are missing parts in the ice core. In order to test whether results are reliable under these more difficult conditions, missing parts have been simulated for the multi year sampling. This

has been achieved by shifting the end point of the sample interval by adding a certain integer g resulting in a gap in the continuous sampling. The parameter g consists of two pieces. The first one (marked in equation 3.8 by 1) describes the length of the missing part by multiplying the maximum gap h with a random number between 0 and 1. This results in a randomly variable gap length. The second piece (marked in equation 3.8 by 2) describes how often a part is missing. This is realized by defining a rectangular function gg. This function is equal to 1 when a second random number is greater than 0.5. In all other cases it is 0. By changing 0.5 to another value, the ice core quality is introduced to the model. In short, g is defined as:

$$g = \lceil (\overbrace{h \cdot rand_1}^{1} \cdot \overbrace{gg}^{2}) \rfloor \quad \text{with} \quad gg = \begin{cases} 1 & , \quad rand_2 > 0.5 \\ 0 & , \quad \text{other} \end{cases} \tag{3.8}$$

with $rand_i \hat{=}$ random number $\in [0,1], i = 1,2$ and the ceiling and floor operation $\lceil x \rfloor$ express rounding towards the nearest integer. The maximum gap h has been set to 20 data points (=2 years).

To sum up, the results of the sampling process are the two matrices $\mathbf{X}(i,j)$ (information about age) and $\mathbf{Y}(i,j)$ (information about ^{10}Be concentrations). They represent the sampling protocol of the computer sampling.

Dealing from now on with the display of the results a new nomenclature is introduced. One column of each output matrix is joined to a $n \times 2$ matrix $k_{i,j}$ that contains information about age and corresponding ^{10}Be concentrations. The index i indicates the sample mode (S= single point sampling, M= multi year sampling) and the sample period {10, 20,..., 220} divided by 10. The index j refers to the to the starting point p. If the index j is missing, then k_i displays the mean of the results from all 11 starting points. For instance, $k_{S3,20}$ describes a curve that is reconstructed from samples taken with single point sampling in sample steps of 30 data points (3 years), beginning at starting point 20.

3.2 Evaluation of the Reconstructions

Before evaluating the results, the long term variations need to be extracted from the the strongly fluctuating data sets. This is realized by applying a 22 years running mean (moving average) to the original data sets as well as to the reconstructed curves. For details on the 22 years running mean and its calculation see appendix A.

After having smoothed the record, a **mean of comparison** is needed to compare the different sample modes and periods to a reference curve. This reference curve has been chosen to be the 22 years running mean of the original data set (for explanations see chapter 4). Three evaluation methods have been tested.

At first, a **mathematical neighborhood** B surrounding the 22 years running mean m has been defined in which a reconstructed curve $k_{i,j}$ (consisting of $x_{k_{i,j}}$ and $y_{k_{i,j}}$ values) should plot.

$$B_\delta(m) = B(\delta, m) = \{x_{k_{i,j}} \in k_{i,j} \mid \|y_{k_{i,j}} - m\| < \delta\}. \tag{3.9}$$

The parameter δ describes the size of the neighborhood. The smaller δ is chosen, the closer a reconstruction must be to the reference curve to plot within the neighborhood. The parameter has been elected for each data set separately in respect to the local extrema of the specific curve. It is defined in a way that the smallest local extrema that should be reconstructed is still visible in the record.
However, it will be shown later, that the results indicating the best sample mode and period do not depend on this parameter. For a detailed description how to define the parameter and the independence of the results of δ see appendix B.
As a result of this method, one receives the amount of samples in the neighborhood relative to the total amount of samples, expressed either on a scale of 0 to 1 or in percentage.

A second method that is often used to describe similarity is the **Pearson correlation coefficient** ρ that measures the correlation of the two data sets $k_{i,j}$ and m. It is defined as [Mudelsee, 2014]:

$$\rho(k_{i,j}, m) = \frac{cov(k_{i,j}, m)}{\sqrt{cov(k_{i,j}, k_{i,j})\ cov(m, m)}}. \tag{3.10}$$

It describes how much two data sets change together, expressed in a range between -1 and 1. Positive numbers describe a similar behavior, negative numbers mean an anti-correlation. In general, ρ describes the linear relationship between two variables.

The last method is called **cosine similarity**, a tool that compares two vectors by calculating the cosine of the angle between them. This measurement of similarity suits especially well for curves with a similar trend, but different proportions, because it neglects the actual values of the vectors and compares their inclinations. So, if two curves have the same shape, though their dimension may be different, the cosine similarity still would have a value close to 1 [Middleton, 2000]. In order to calculate the similarity, two data points of the curve $\mathbf{k}_{i,j}$ and the corresponding two data points of the 22 year running mean $\mathbf{m}$ are understood as two vectors, so that the cosine similarity, $cos(\gamma)$, is calculated by dividing the dot product by the product of the magnitudes of both vectors [Middleton, 2000]:

$$cos(\gamma) = \frac{\mathbf{k}_{i,j} \cdot \mathbf{m}}{\|\mathbf{k}_{i,j}\| \|\mathbf{m}\|} \tag{3.11}$$

In this study, the results range between 0 meaning orthogonality of the vectors, to 1 meaning exactly the same direction.

Repeatability

Since the results of this model depend on the input of the random numbers, one sample run is not representative. Therefore, in order to assure repeatability of the results, the sampling was repeated 50 times with varying random numbers. Analysis have shown that for more than 50 runs there are only negligible changes in the results. All results of the evaluations are stored, but for a more convenient treatment, the results for each sample mode and sample period are summarized after each run to:

- the mean value of all 11 starting points
- the standard deviation corresponding to the mean value of all 11 starting points
- the minimum of all 11 starting points

After 50 runs these values are summarized a second time to:

- the mean value of all mean values
- the mean value of all standard deviations
- the minimum of all minima.

4 Results

It is not convenient to collect data at an (sub-)annual temporal resolution due to likewise financial or time shortages. Therefore, an appropriate sample scheme must be designed that ensures that reconstruction is as complete as possible and free from biases. This sample scheme is set up by following the steps suggested by Harris and Jarvis (2011) and Borradaile (2003) (see figure 4.1).

At first, as **target of the sampling** the reconstruction of the long term variations has been identified. Beer (2000) distinguishes between short-term fluctuations as the 11 year solar cycle and long-term variations on a time scale of decades to centuries. This scale contains for example the great solar excursion. Consequently, these long-term variations are suitable for wiggle matching. In order to extract the long-term variations from the short-time fluctuations, a way of smoothing is necessary. It has been concluded that 22 years is a suitable time unit for long-term variations and that a running mean is the best way of smoothing (table 4.1).

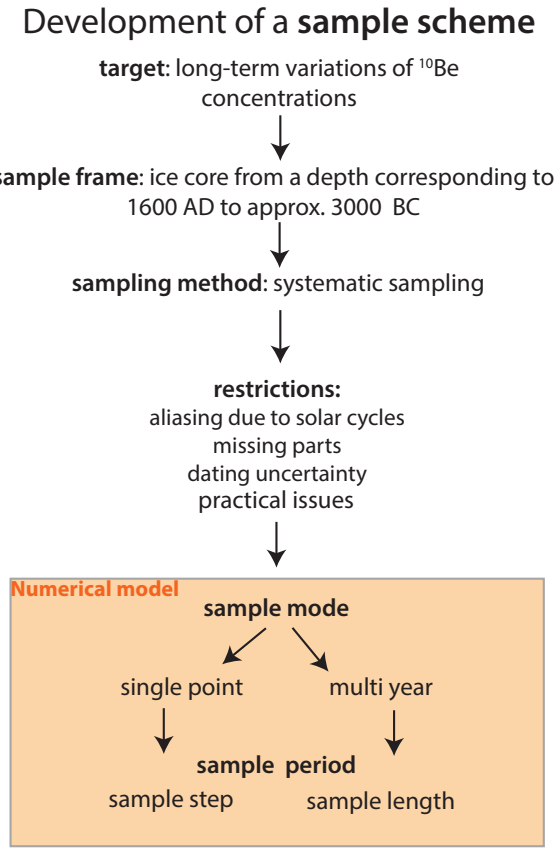

Figure 4.1: **Development of a sample scheme** by following the steps proposed by Harris and Jarvis (2011) and Borradaile (2003). The target, sample frame, sampling method and restrictions have been discussed in respect to existing literature. The sample mode and sample period have been tested using a computer model and practical issues have been introduced in accordance to personal experience of the author after taking part personally in sampling and chemical processing.

Limitations	Required Smoothing	Reference
percolation, dating uncertainties	min. 5 years	[Weiler et al., 2005], [Opel et al., 2009], section 1.1, chapter 2
system effects	min. 20-25 years to reduce influence to 20-30%	[Muscheler et al., 2014], [McCracken, 2004] section 1.1
wiggle matching	max. 20-30 years to detect grand solar minima	section 1.2
aliasing effects	11 or 22 years to eliminate the 11 (22) year periodicity	[Beer et al., 2012], section 4 Particular Restrictions

Table 4.1: Overview of the restrictions that require smoothing of a certain time span

This decision is based on the fact that data smoothed with a 22 years running mean is only slightly biased due to system effects (section 1.1). Moreover the 11 and 22 years solar cycle is suppressed (section 4 Particular Restrictions of the Sampling Process) and percolation will not influence the signal (section 1.1). On the other hand, it is still possible to resolve important wiggles for matching (section 1.3). The running mean has the great advantage that the data point density is not reduced, so even when smoothed to 22 years, there are in the most cases more than one data point per 22 years. Therefore, the 22 years running mean of the original data is interpreted as the representation of the long-term variations and used as reference curve for any further comparisons.

Secondly, the **sample frame** has been outlined. The sample frame consists of all samples that one can theoretically take. The ice core has a total length of 724 m. However, for the upper first 136 m corresponding to approximately 275 years already exists an validated age-depth model [Fritzsche et al., 2005]. Moreover, there are already 176 measured samples spanning a time range from 1947 to approximately 700 years AD (figure 4.2). Sample length and density varies a lot as it is displayed in figure 4.2. Since the age-depth relation is more or less confirmed for the time span of 1900-1600 AD [Fritzsche, 2015 personal communication], more attention is drawn to the time before 1600 to the approximate total age of 3000 years.

Thirdly, as a next step of the sample scheme, systematic sampling has been chosen as ***sampling method*** due to some reasons. Systematic sampling is a probabilistic sampling method. This has the advantage that statistical tests can be done on the data and that the results are likely to be representative of the whole ice core. Moreover, systematic sampling documents gradual changes as they are present in the data very good. It is commonly applied because it is convenient, simple and efficient when sampling large populations. One can be sure that the whole data set is represented uniformly. However, systematic sampling has the disadvantage that it can be biased strongly due to periodicity in the data. Then, attention needs to payed at aliasing effects [Borradaile, 2003], [Harris and Jarvis, 2011]. Since continuous sampling, as conducted in multi year sampling, refers to center points of samples in constant spacing, it is regarded as well as systematic sampling.

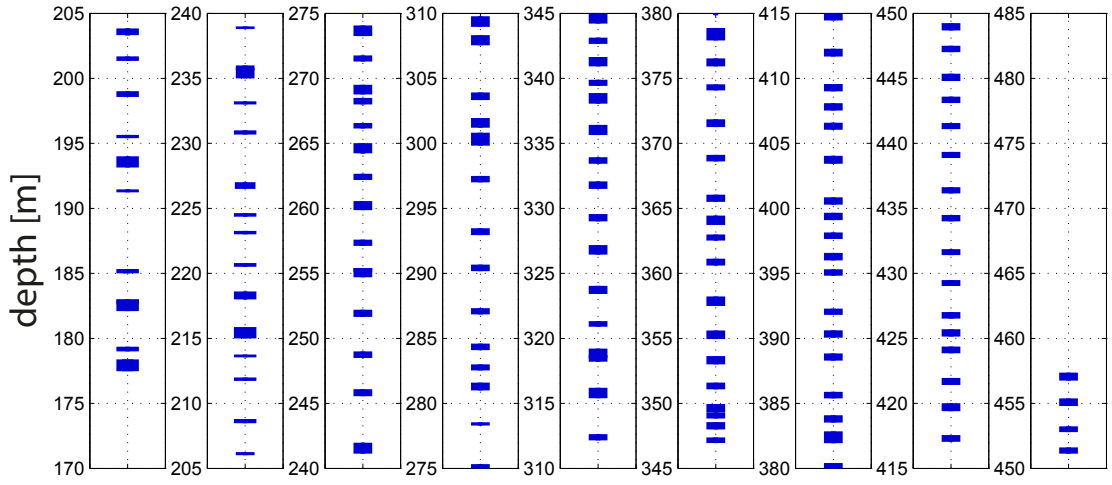

Figure 4.2: **Depth distribution of the samples that are already taken** (blue bars) before 1600 AD [Fritzsche, 2015 personal communication]. The figure displays the great variety of sample lengths and gaps between the samples. On average samples are 0.57 m long and are located 3 m apart from each other.

Particular Restrictions of the Sampling Process

When applying systematic sampling, one must consider aliasing due to the strong periodicity in the ^{10}Be concentration (section 1.2). Aliasing occurs when sampling at an interval too broad to resolve the true high frequencies in the signal [Pisias and Mix, 1988]. The Nyquist-Shannon sample theorem states that aliasing occurs when sampling a signal that has a constant bandwidth of f_0, with a sampling rate that is smaller than $2f_0$ (= Nyquist frequency f_{nyt}) [Mari et al., 1999]. Aliasing causes a shift of the variance of the high frequencies that are not detected anymore into lower frequencies. This results in a distorted signal with lower, alien frequencies between zero and the Nyquist frequency (see figure 4.3) [Wescott, 2015]. If the amplitudes of these undetected high frequencies exceed the amplitudes of lower, truly depicted frequencies, the shape of the reconstructed curve might be changed seriously.

When transferring these findings to the present problem of the ^{10}Be concentrations the following is implied: Since the periodicity of the 11 year solar cycle is not perfectly band limited, it varies between 7–18 years, it is difficult to define a clear Nyquist frequency. Assuming the typical length of 11 years, all samples taken with a frequency smaller than $f_{nyt} = \frac{2}{110}$ data points are likely to be aliased. Since the amplitude of the variations of the 11 year cycle mostly is greater than the amplitude of the long-term variations, serious changes at the shape of the curve might occur as a result.

A way to reduce the effects of aliasing is to apply a low pass filter in advance. This can be put into practice by taking longer samples, already taking an average over the whole sample length, filtering out all frequencies higher than the sample length.

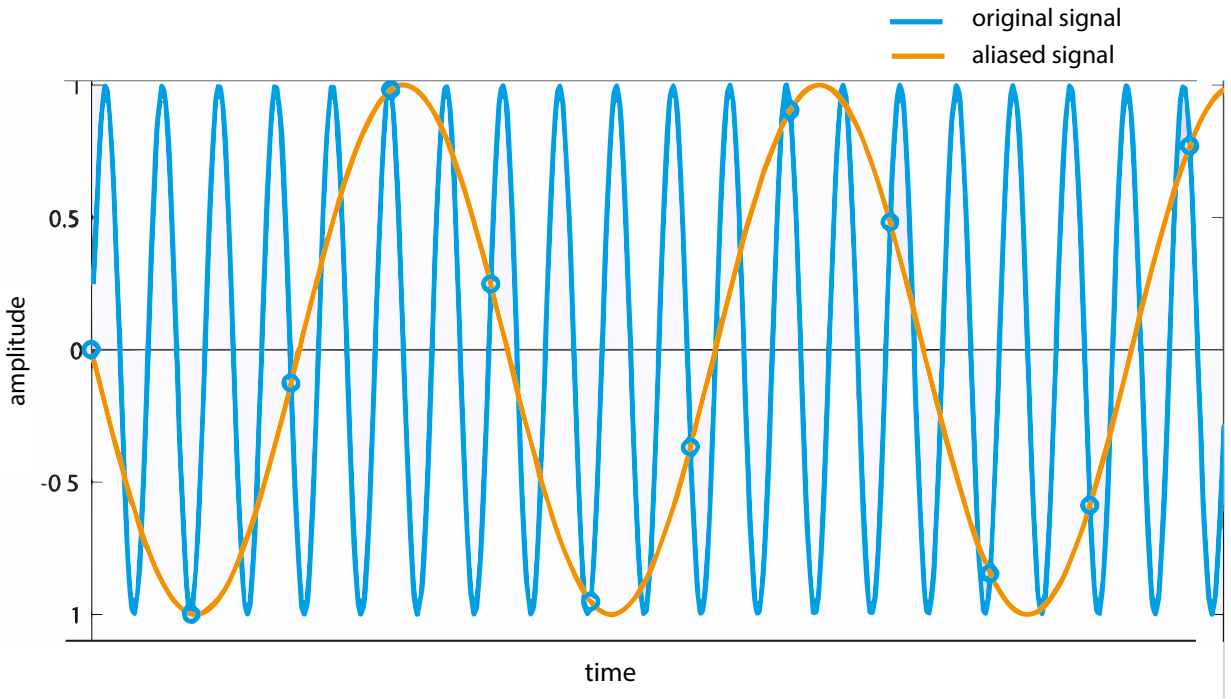

Figure 4.3: **Aliasing effects when sampling under the Nyquist frequency**. The original data (blue) has a higher frequency than the sampling (circles) resulting in a reconstructed signal (red) that has a smaller frequency than the original signal (adapted from [Burg et al., 2014]).

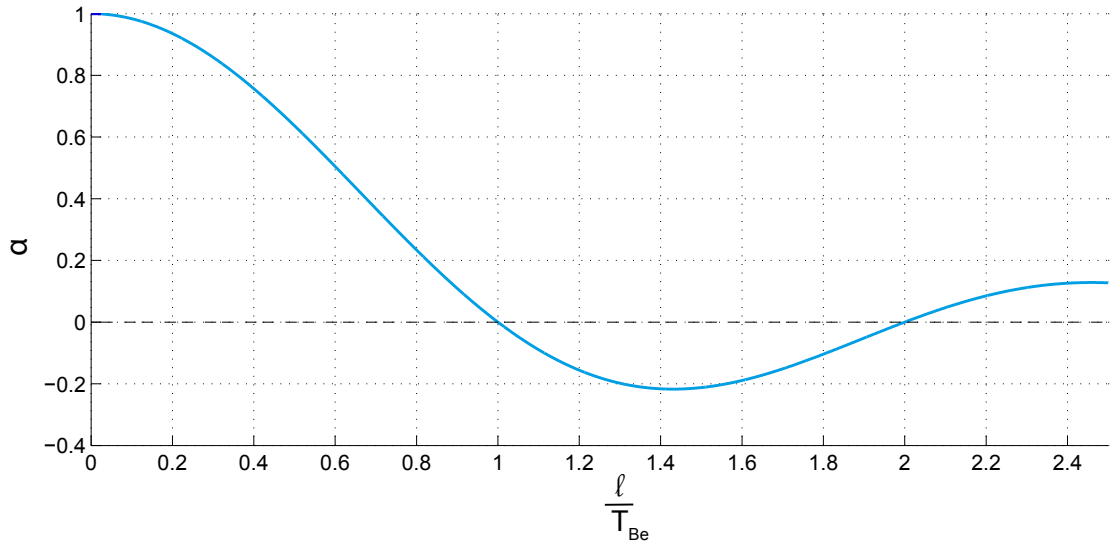

Figure 4.4: **Amplitude ratio α of reconstructed to original signal** in respect to the ratio of sample length l and original period T_{Be}. A complete suppression of the amplitude of the original signal is given for $\frac{l}{T_{Be}} = 1, 2, ...$, so for all sample lengths that are integer multiples of the sample periodicity [Beer et al., 2012].

Therefore, the high frequency components will be attenuated by sampling longer sample intervals. This can be described as well as a zero-order hold function. The attenuation is described as ratio of the amplitude of the periodicity in the sampled data relative to the amplitude in the original data by [Beer et al., 2012]:

$$\alpha = \frac{sin(\pi \frac{l}{T_{Be}})}{\pi \frac{l}{T_{Be}}} \tag{4.1}$$

where T_{Be} is the periodicity in the original data and l the sample length. Figure 4.4 shows how the amplitude of such a signal is attenuated. As a consequence, the amplitude of some periodicities can be minimized either on purpose or by accident. Consequently, from a theoretical point of view, aliasing can be suppressed with sample lengths that are multiples of the periodicity. Since the most dominant periodicities in the data are $T_{Be} = 11$ years (respective $T_{Be} = 22$ years), sample lengths of l=110, 220, 330,... data points are favorable. However, it is insecure how the varying length of the 11 year cycle, dating uncertainties and missing parts will distort this. In order to confirm this theoretical statement, the computer model has been used successfully (section 4.1).

Apart from the aspects of periodicity that bias the results, there are three additional restrictions that need to be considered when sampling the Akademii Nauk ice core. The ice core is taken in an area where summer melting occurs and does not have enough reliable yearly markers. Because of these two reasons, there is a certain dating uncertainty that has been estimated to be maximal 1 year. Moreover, occasionally parts of the core are missing and therefore, it has been tested, whether results are stable under the conditions of missing parts as well. It was estimated that at most 2 years (approx. 1 m of ice) are missing. The likelihood of missing is set to 50 % but can be modified as one likes. This value might seem really high, but 1 m is only the maximum gap. Far more often only a few centimeters are modeled to miss.

Lastly, the sample design has been reviewed from a practical perspective. In order to fulfill this at a best, I had the chance to participate in the actual sampling process and the chemical processing of the samples to understand the main limiting parameters. It was found that additional sawing for samples whose depth intervals do not correspond to the depth intervals of sample bags is very time consuming and might lead to loss of material. For chemical processing it could be investigated that the sample weight (previous studies 300-400 g) is not as relevant as it thought to be and could be discussed with the particular laboratory processing the samples [Merchel, 2015 personal communication].

4.1 Results of the Numerical Simulation

The computer model aims to identify the optimal sample mode and period. Exemplary Dye-3 data is used to display the results, but all data sets show very similar findings (NGRIP and Dome Fuji figures are in appendix C). Figure 4.5 compares single point sampling (a) with multi year sampling (b). It can be stated that both modes reconstruct the rough trend of the 22 year running mean. Moreover, at first sight, all multi year sampling curves are closer to the 22 year running mean than the single point sampling curves.

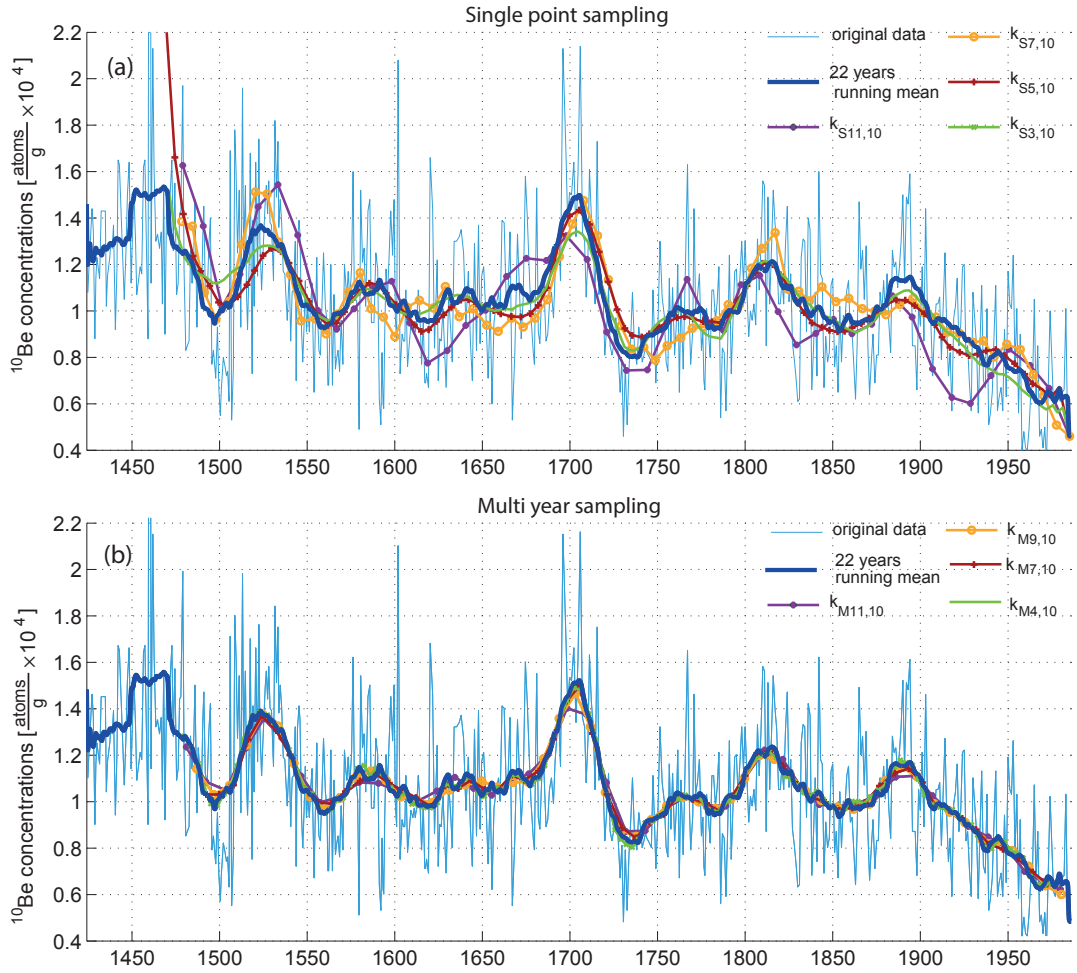

Figure 4.5: **Visual comparison of single point sampling curves (a) and multi year sampling curves (b)**. Both panels show the original data, the 22 years running mean and selected curves with starting point 10. In (a) single point sampling curves with sample steps s=110, 70, 50, 30 are displayed. In (b) multi year sampling curves with sample lengths of l=110, 90, 70, 40 are shown. The multi year sampling curves are all very close to the 22 years running mean, but the single point sampling curves differ considerably from it.

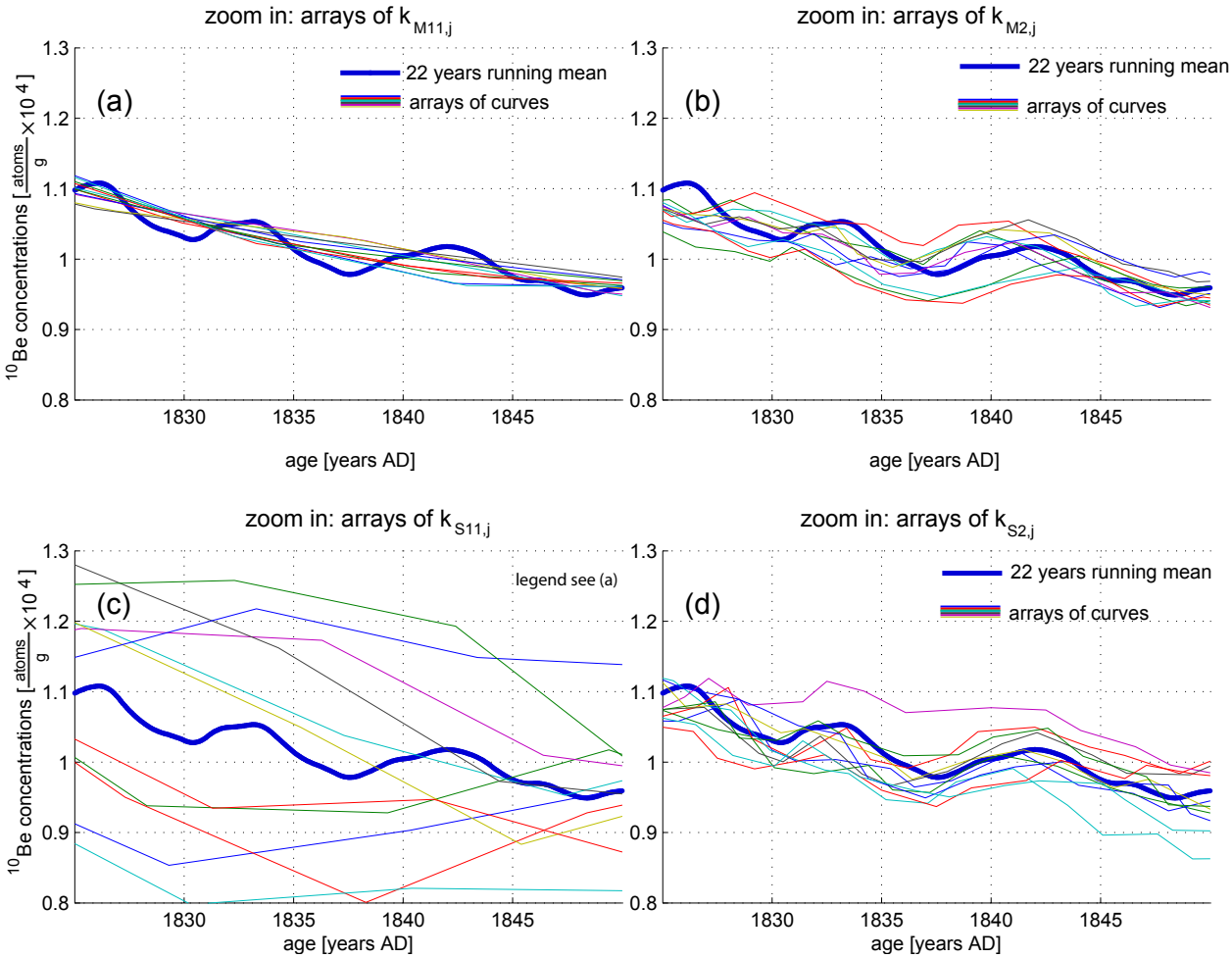

Figure 4.6: **Comparison of different starting points**. In each panel the 22 years running mean and an array of curves with 11 different starting points is displayed. Panels on the top (a, b) show multi year sampling, panels at the bottom (c, d) single point sampling. Large sample periods ($T_s = 110$ data points) are shown on the left (a, c) and small sample periods ($T_s = 20$ data points) on the right (b, d). The graphs indicate that multi year sampling curves scatter less than single point sampling curves. Longer sample lengths reduces scattering even more. Regarding single point sampling shorter sample steps minimize scattering.

Evaluating the influence of the different starting points, it was found that the deviations between the curves are minimal for multi year sampling with long sample lengths ($k_{M11,j}$). However, curves with short sample lengths ($k_{M2,j}$) show still less deviation than single point sampling with long sample intervals ($k_{S11,j}$) (see figure 4.6).

Figure 4.7 displays the borders of the neighborhood in a red line and the results of selected multi year sampling curves. The parameter δ that defines the mathematical neighborhood is calculated to be $\delta = 0.035 \ \frac{atoms}{g} \times 10^4$ for the Dye-3 data. It can be inferred that really short samples, as $k_{M1,10}$,

vary extremely, portraying not always the real shape of the curve (e.g. at 1595-1615 years AD). On the other hand, very long sample lengths, as $k_{M22,10}$, only map the long term trend of the 22 running mean and are not able to catch smaller wiggles (e.g. around 1610-1620 years AD). Sample lengths between these two extremes, as $k_{M8,10}$, provide good results.
These assumptions based on visual analysis are confirmed by the results of the evaluation methods (figure 4.8 and 4.9). Figure 4.8 allows a comparison between single and multi year sampling revealing that multi year sampling scores better for all sample lengths except length $l = 10$.

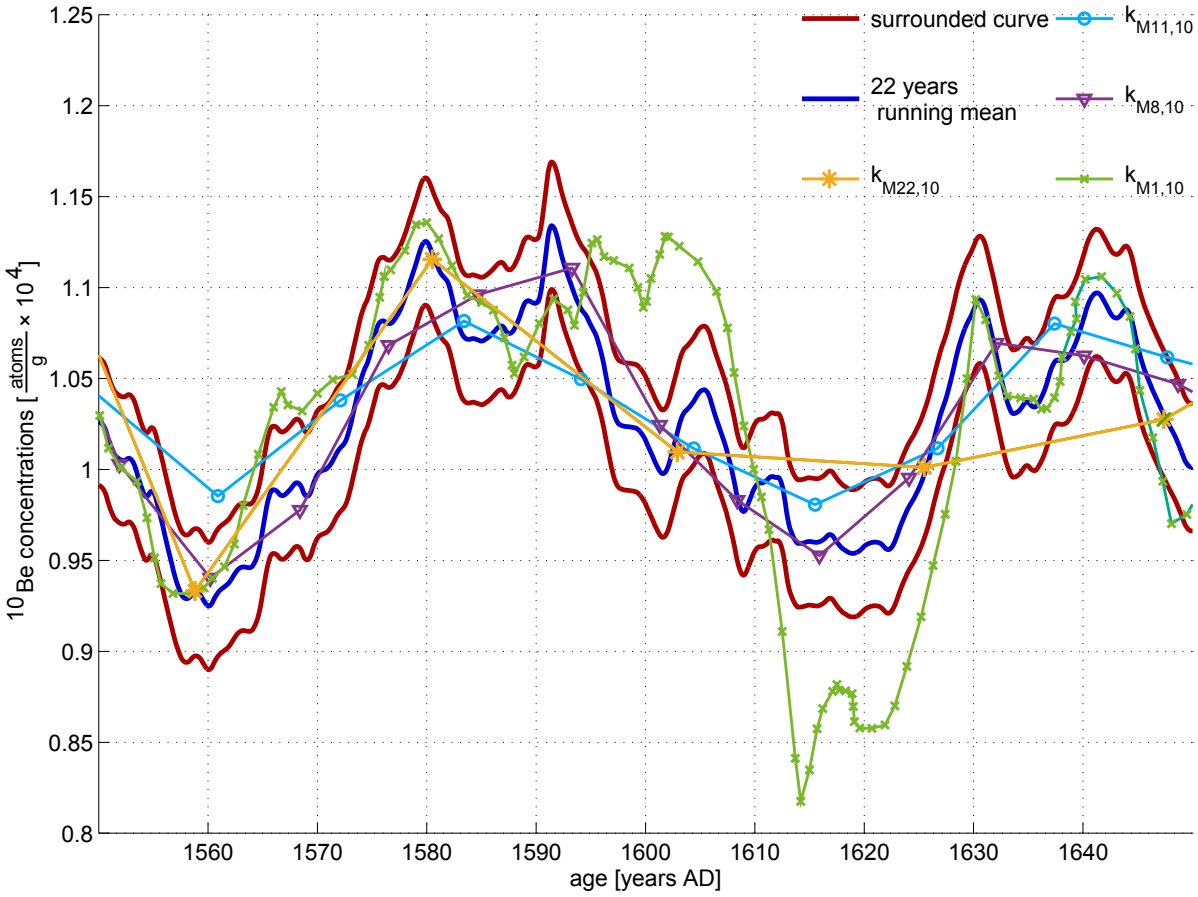

Figure 4.7: **Neighborhood B(δ,m) and selected multi year sampling curves**. The surrounded curves (red) that mark the borders of the neighborhood B(δ,m) are displayed with the 22 years running mean and selected multi year sampling curves for a time interval from 1550-1650. A visual analysis reveals that small sampling intervals as l=10 and long sample intervals as l=110 do often plot outside the neighborhood. Intermediate sample lengths as l=80 plot almost completely within the neighborhood

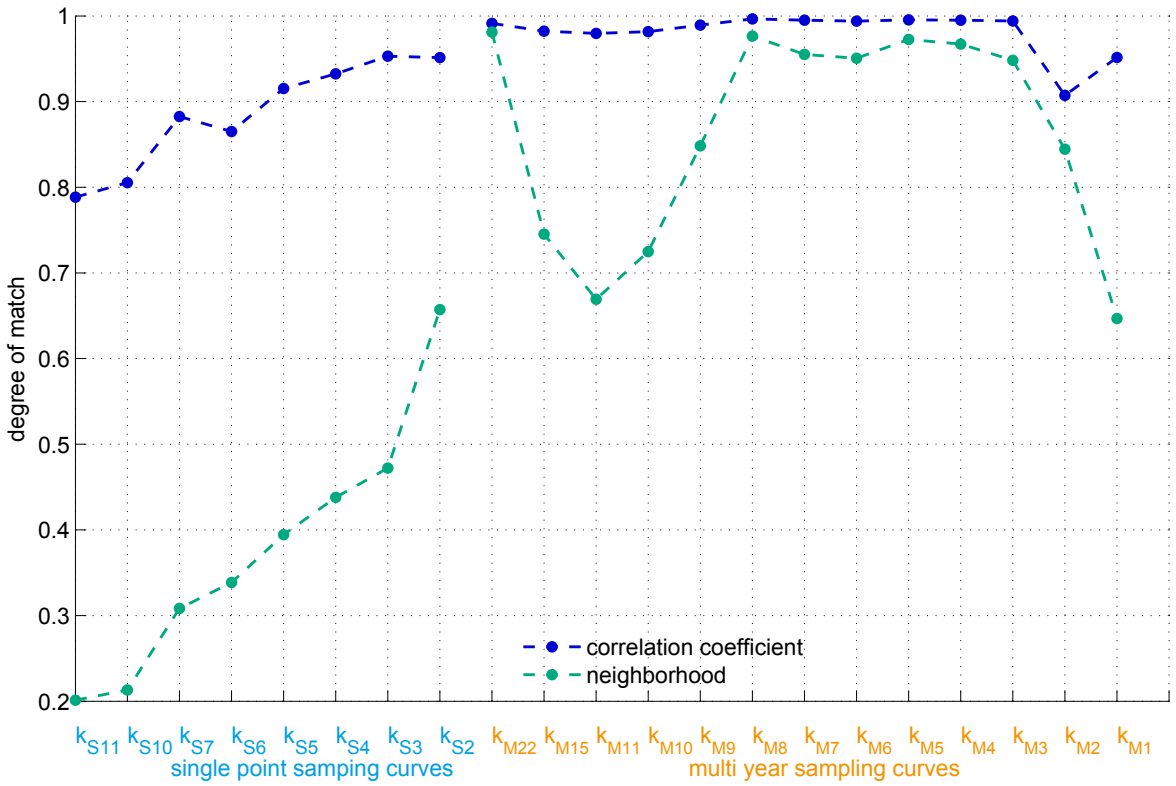

Figure 4.8: **Comparison of evaluation results for different sample periods (single point sampling and multi year sampling)**. On the x-axis the different sample modes and periods are displayed (from larger to smaller periods) and on the y-axis the corresponding mean value of 50 runs that expresses the degree of match (on a scale of 0–1) is shown. In general, it can be stated that multi year sampling receives a higher degree of match than single point intervals.

A more detailed look at the results of the multi year sampling is displayed in figure 4.9.
In part (a) and (b) of this figure the degree of match is expressed for continuous sampling and the sampling with missing parts. For each case, the average of all 50 runs is displayed along with the mean standard deviation of 50 runs as errorbars and the minimum (see section 3.1 Repeatability). The mean contains information about the behavior that is expected. The standard deviation is a parameter to determine the extent of difference between the starting points and the minimal curve displays what can be the worst reconstruction when sampling with this sample length.

The first observation that can be made is that all curves follow more or less the same pattern. For the neighborhood method a maximum $> 95\%$ is observed at a length of l=80, however, the maximum has more the shape of a plateau from l=80 to l=40. Apart from that, the sample length of l=220 appears to be another maximum of $> 95\%$. This confirms the theoretical statement made studying the aliasing effects. However, the sample length of l=110, corresponding to 11 years is a local minimum.

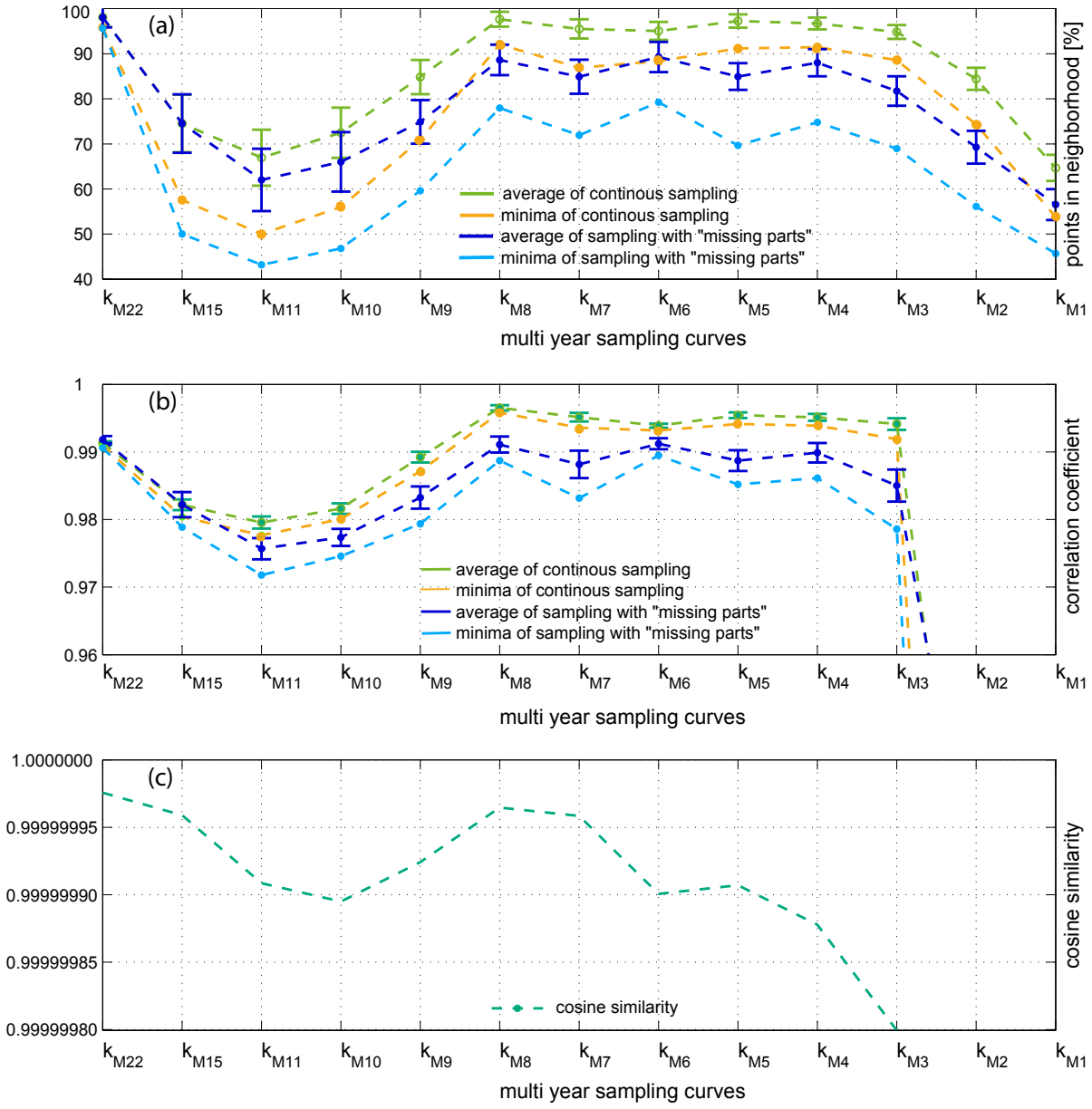

Figure 4.9: **Comparison of evaluation results for different sample lengths (multi year sampling)**. For each sample length (x-axis) the corresponding degree of match is plotted (y-axis). In (a) the percentage of points within the neighborhood B(δ,m) is displayed. In (b) the results of the correlation coefficient are shown. In (c) the results of the cosine similarity are summarized. All curves show a similar pattern with a maximum in shape of a plateau at l=80 to l=40. A second maximum is found at l=220. For values that are too low to be displayed (k_{M2} and k_{M1}) in (b) and (c) see appendix C.

Moreover, it can be stated that the standard deviations of the continuous sampling is larger for very long and very short sample lengths, indicating greater differences between starting points. The greater differences between the starting point for long sample lengths is contrary to the visual impression made before.

In general, the results of sampling with missing parts are worse than continuous sampling, but the average still reaches values greater than 85 % (neighborhood evaluation) being close or even higher than the results of the minimum of continuous sampling.

Figure 4.9 (b) summarizes the results of the correlation coefficient, revealing a very similar pattern. The values range between 0.97 and 0.995, confirming good correlations for all sample lengths.

The only real difference to figure 4.9 (a) is the extended length of the maximum plateau covering the sample lengths l=80, 70, ..., 40, 30 followed by a extreme decrease (for a better overview see figure 4.8).

The curve of the cosine similarity (figure 4.9 (c)) affirms the observations made before. Two maxima are found at k_{M22} and k_{M8}. The plateau is not as prominent as for the other two graphs. Furthermore, it must be noted, that all values are in a very small range (0.99999980 - 0.99999998) indicating even for the minimum a very high correlation.

5 Discussion

5.1 Interpretation of the Results

The first important conclusion of the computer simulation is to prefer multi year sampling in contrast to single year sampling. This statement is based on the fact that multi year sampling achieves except for very short sample lengths better results than single point sampling (figure 4.8). An explanation for this is the fact that in multi year sampling mode no information is lost and aliasing is attenuated by averaging that takes place when melting longer ice samples.

Furthermore, the evaluation shows that the best multi year length is l=80, whereas l=70, 60, 50, 40 also perform well (figure 4.9). These intermediate lengths compensate outliers, but are still at such a high temporal resolution that they are capable of following short time fluctuations of the 22 years running mean (figure 4.7). Additionally, the sample length l=220 performs nearly as good as l=80. This is consistent with the theoretical considerations of aliasing that prognoses a total suppression of the 11 year solar cycle for sample lengths that are multiples of the underlying periodicity. However, since the correlation coefficient of the sample length l=220 is lower than the ones of the sample lengths l=80, 70, ..., 40 the latter are preferred. Furthermore, the visual analysis of figure 4.7 suggests that $k_{M22,j}$ smooths out characteristic parts of the curve because the temporal resolution is too low. Another point in favor for the intermediate sample lengths is their small standard deviation. This means that the starting point is nearly irrelevant. Moreover, the minima curve confirms that even in the worst case, the above mentioned sample lengths perform best.

Even when parts are missing, continuous sampling still results in 85 % points within the defined neighborhood and a correlation coefficient of ρ=0.99 (figure 4.9). Due to this high degree of match, it has been concluded that reconstruction of the long-term variations is possible even when the ice core is occasionally incomplete. Missing parts in the ice core change the shape of the curve slightly, so that the sample length of l=60 reaches the best results. Since reality is somewhere between a complete ice core and a core with occasionally missing parts, it is deduced that all sample lengths between l=80, 70, 60, 50, and 40 are advisable with a slight preference for l=80. The selection of more than one sample length leaves some decision-making scope for practical issues.

The results of the cosine similarity resemble more or less the course of the other curves. Although the shape of the maximum is less pronounced, the findings of the cosine similarity shall not restrict the above discussed range of optimal sample lengths any further, because the results are all in a very small range.

5.2 Bridging the Gap between Theory and Practice

Single point sampling is modeled to be a near-point measurement. In the present case this refers to a sample length of 0.1 years, but a further interpolation decreases the sample size of single point sampling. Such a short sample interval is unlikely to be taken in reality and must be neglected from the beginning, because percolation makes longer sample intervals necessary. Therefore, for a more detailed look, single point sampling must be changed fundamentally from near-point measurements to sampling of one or two year intervals with varying gaps between the samples resulting in a hybrid form between the two sample modes.

In this context, it is useful to analyze the results of these intervals of l=10 and l=20 in the multi year sampling mode. Such an analyze is helpful to identify certain problems in advance, that might occur in both sample modes. It can be stated that in multi year sampling mode these very short sample lengths achieve very bad results. An explanation for the bad performance could be the large ratio between sample length l and dating uncertainty r. The ratio is $\frac{l}{r} \geq 0.5$, so the effective sample lengths vary between 50 % and 150 % for l=20 and 10 % and 200 % for l=10 of the original value. Such a high degree of variability produces outliers that are hard to fit in even after averaging. Thus, single point sampling with basic intervals of l=10 or l=20 might probably produce quite low results except the concept of the dating uncertainty is changed.

For this case and in general, the nature of the dating uncertainties needs to be investigated better. At first, the random distribution of the dating uncertainty does not meet reality because variations in the sample interval are rather systematic in reality, as the major source of them is a change in the accumulation rate [Pisias and Mix, 1988]. This leads to a positive correlation between dating uncertainties and the sample length. This would result in smaller dating uncertainties for shorter sample lengths producing better results for those. Although further studies could improve this point of the model, it is questionable if the dating uncertainty could fall beneath 1 year due to the unknown effects of percolation.

In accordance with practical issues a few considerations must be added to the results of the computer model. First of all, the model is based on an age-concentration relationship because it allows an easy transfer between different records. However, this ice core has only a few reliable indicators for yearly markers. Thus, results of this study must be transferred from age into depth taking the accumulation rate and layer thinning into account. Since accurate changes of these factors with depths are not completely known, sampling will result in varying time intervals. Therefore, further studies must be done on depth-concentration sets to identify possible sources of errors.

When transforming age to depth, this results in depth intervals that are different from the depth intervals in which the core is stored (sample bags are $\approx 1\ m$, see chapter 2). Hence, extra sawing for the extraordinary sample lengths elongates the sampling process and rises the probability of measurement errors and loss of material. Therefore, in practice, multiples of the sample bag length are preferred. A second practical limitation is the weight of the samples that should be between 300

and 400 g. Samples of 8 years (roughly 4 m length, minimum 1 cm diameter) exceed this weight by far, when no additional cutting is done. However, after having studied the chemical treatment of the samples, there seems to be no chemical, and only practical reasons for this restriction [Merchel, 2015 personal communication]. Consequently, for further sampling heavier samples should be considered after a consultation with the laboratory facility processing the samples.

Moreover, the transferability of the results to the Akademii Nauk ice core are discussed. Primarily, it is assumed that the results depend on the periodicity in the data that is global and therefore the results of this study can be transfered to any other ice core. This is confirmed by the fact that the three tested different records show the same results. However, if there is a dependency on weather conditions or e.g. the scale of the ^{10}Be concentrations, the Dye-3 data set is the one that is the most similar to the Akademii Nauk ice core. Snow accumulation is high at both sites, resulting in low ^{10}Be concentrations between 0.1-4 (Akademii Nauk) and 0.4-5 $\frac{atoms}{g} \times 10^4$ (Dye-3) deposited by mainly wet deposition. However, Akademii Nauk is located far more north than Dye-3 (closer to NGRIP) but at much lower altitude (750 m) than both other records (> 2479 m). Since none of the three records suffers from summer melting, computer simulations should be done on ice core records from study sites with summer melting, e.g. Svalbard, to test the specific influence of percolation.

As mentioned above, for calculating depth intervals thinning of annual layers has to be taken into account. Opel et al. (2013) summarized that it is assumed that the layer thickness can be explained by a Nye model that has been modified by a growth term. This knowledge can be used when calculating the 8 year intervals. For more convenience in sampling one could change the calculated depth intervals within the range of the dating uncertainty of 1 year. This will still require a lot of extra sawing. Therefore another idea is proposed that takes advantage of the fact that all sample lengths between 8 and 4 years perform quite well in the computer model. Since layers thin out of more than 50 %, one could consider taking gradually (or in steps) "longer" samples in sense of age using the full range from 4 to 8 years. This would result in almost constant depth, but changing time intervals. This suggestion is coherent with the sampling of the NGRIP ice core before 1780 AD that was sampled equidistantly leading to intervals covering 0.4-2.3 years [Berggren et al., 2009]. This requires a temporal griding afterwards, but this has been done for other ice cores before (e.g. GRIP [Yiou et al., 1997]) Consequently, a further aim is to test in the computer model whether mixing sample lengths from 8 to 4 years still produces good results.

There is already a considerable amount of data points for reconstruction, but no overall confidence. Since Walker (2005) states that wiggle matching works best over a short, but well described prominent section of the reference curve, it seems to be a good idea to concentrate the sampling on a short period. A suitable part contains well structured features, in concrete terms, considerable variations of the ^{10}Be concentration, because this adds further confidence to the matching approach [Abreu, 2009], [Ruth et al., 2007]. Therefore, I suggest to identify a time interval that contains such prominent wiggles (e.g. 1400-1550) and "fill it up" by continuous sampling in order to reconstruct a secure dating point.

5.3 Comparison to other Studies

Wiggle matching for dating or correlation of ice cores has only been used by a few researchers. In the following their approaches are compared to each other in regard to the sample mode, anti-aliasing strategies, sampling method, treatment of data and evaluation methods.
According to the **sample mode** Yiou (1997) stresses that for an effective study of the ^{10}Be production signal a continuous sequences is essential. The findings of this model clearly point towards a higher success for continuous sampling. Therefore, they are easily to put into context of the strong majority of studies that use continuous sampling. For instance, Raisbeck et al. (1998) sampled the last 7000 years of Vostock ice core in continuous 50 cm intervals. This interval length was also chosen by Horiuchi et al. (2008) for the Dome Fuji ice core corresponding to 10 years to reconstruct the time span from 1900 to 700 years AD. A similar time interval (9 years) reached Raisbeck et al. (2007) by sampling EPICA Dome C ice core in continuous 11 cm samples. Records from Greenland (GRIP data extended with the GISP2 data), that are geographically closer to this study site, are used for wiggle matching by Muscheler et al. (2014). The GRIP ice core has been sampled in constant intervals of 55 cm [Yiou et al., 1997] corresponding to a mean temporal resolution of 5 years for 304-9315 years BP [Vonmoos et al., 2006].
On the other hand, Henderson working with an ice core from Franz Joseph Land (Russian Arctic) for his dissertation is the only author that was found using a not continuous sample mode [Henderson, 2002]. He matched 15 samples each containing a time interval of around 7–8 years (3 m) covering a total time span of about 500 years. However, it remains unsure whether this sampling can be applied reliably for every ice core and every time period because the grand solar excursions used for matching have sometimes shorter durations than the time intervals chosen by Henderson. So important peaks might miss and lead to misinterpretations.

There are quite opposing opinions on the question whether ^{10}Be records are biased seriously by **aliasing** or not. On the one hand, authors working with Antarctic ice cores or tree rings (^{14}C) as Alley, Ramsey and Blauuw reported that they regard those biases as no major source of error. They argue that aliasing is lost in noise because normal sample lengths are normally greater than 10 years, and later even averaged to 25-100 years [Alley, Ramsey, Blauuw, 2015 personal communication]. Mudelsee (2014) agrees that aliasing is not a major problem because he assumes that high-frequencies are not preserved in ice cores due to diffusion-like processes working as low-pass filter. Moreover, he claims that due to uneven sample intervals and timescale errors aliasing is reduced, because in this case a Nyquist frequency is not well defined and therefore aliasing is less systematic. This is partly confirmed by a numerical simulation done by Pisias and Mix (1988). Although they have found that aliased peaks appear even when using sample intervals that vary randomly, they could show that the magnitude of the aliased peaks was smaller. They concluded that “sufficiently random sampling interval perhaps induced by long-period variations in the sedimentation rate, might reduce the aliasing problem” [Pisias and Mix, 1988]. Sedimentation rate can easily transfered

in this context to accumulation rate. Additionally, Wunsch (2000) adds that "melting multiyear segments of the core [...] is a very effective filter" that might reduce aliasing. But he limits his statement to sample lengths that are integers of years, because he assumes that fractions of years can produce considerable leakage.

On the other hand, Beer et al. (2012) stress that large differences between original and reconstructed signal in respect to their amplitude and phase that arise from sampling can lead to incorrect conclusions, especially for solar physics and atmospheric processes. This is supported by McCracken et al. (2004) and McCracken (2003, 2004) that discuss pseudo-random variability of up to 35 % of the variations in the ^{10}Be concentration as a consequence of the unresolved 11 year variation.
The findings of this numerical modeling support in principle Beer's point, because it is assumed that the differences of up to 25 % between the sample lenghts l=220, 150, ..., 40 are associated with aliasing effects. Primarily, the curve in figure 4.9 follows the trend suggested by the attenuation in shape of a sinc function. As expected there is a maximum at k_{M22}, but k_{M11} is a local minimum. This could be explained by a shift of the sinc relationship towards smaller sample lengths. Such a shift might have different sources. Pisias and Mix (1988) state that uneven sample intervals as they are present in this model lead to a wider distribution of variance associated with the aliased peaks over a wider frequency band. This might lead to such changes, as well as the fact that the solar cycle is not constantly 11 years, but varies between 7–18 years [Beer, 2000]. Furthermore, it can not be excluded and should be tested whether smoothing to another value than 22 years, for example 11 years change the shape of the curve. To sum up, in this study aliasing is regarded as an error source, even though the effects are minimized by melting and uneven sample intervals due to dating uncertainties. A detailed study of original and sampled data in the frequency domain might help to quantify the effects of aliasing.

Since aliasing is a result of systematic sampling, the **sampling method** is questioned as well [Harris and Jarvis, 2011]. As mentioned above, a lot of studies use systematic sampling, but most of them sample with constant depth intervals resulting in uneven time intervals (e.g. [Raisbeck et al., 2007], [Berggren et al., 2009]). As discussed above, these irregular sample intervals might already attenuate aliasing effects. But there is also the possibility to choose another sampling method that introduces even a higher degree of randomness to the system.

Sampling methods are divided into two groups: probabilistic and non-probabilistic methods. Non-probabilistic sample techniques select samples on the researcher's choice and are not suitable for conducting statistical analyses. This is especially relevant for quality research, for instance to get a first impression [Harris and Jarvis, 2011]. Henderson (2002) seemed to have applied this method, because he states that his samples have been "selected judiciously". As already indicated, this can work, but because there are no other studies using this method, no reliable guarantee of success is given.
On the other hand, there are probabilistic sampling methods that are based on probability theory. Systematic sampling is an example that is commonly applied. But there are more probabilistic

methods, for example (simple) random sampling that selects samples based on a random algorithm without replacement. A specification of this technique is the stratified random sampling when random sampling is done in a given interval that was defined by systematic rules. This method is suggested by Ramsey [2015, personal communication] and Borradaile (2003) to reduce the influence of aliasing. For multi year sampling it is already discussed that this reduces the magnitude of the aliased frequencies [Pisias and Mix, 1988], but for single point sampling no reliable results are present. Therefore, further studies need to be done to test whether the advantages of suppressing aliasing can outweigh the problems concerning small sample lengths.
A subsequent variation of these main forms is cluster sampling that is based on two stages. At first the relevant area is identified, for example via non-probabilistic methods. Then, this area is sampled, e.g. with a probabilistic method [Borradaile, 2003]. This method was applied on the Akademii Nauk ice core before by identifying time spans of prominent peaks in the ^{10}Be concentration that have been sampled. This method produced reliable results for the time span 1900-1600 [Fritzsche, 2015 personal communication], but could not completely confirm earlier ages. Reasons for this must be evaluated before further applying cluster sampling.

In this study **smoothing** was done with a 22 years running mean to reduce disturbing factors as system effects. There are a lot of different approaches of how to smooth data, depending on the aim and the temporal range of the studies. Working groups apply low pass or band pass filter ([Beer, 2000], [Berggren et al., 2009], [Miyahara et al., 2004], [Abreu et al., 2013]) removing the frequency range that is associated with the 11 year cycle. Numerous studies also apply binomial filters ([Muscheler et al., 2014]; [Vonmoos et al., 2006]) or Gaussian filters [Bard et al., 2000]. These filters work similar to weighted running means. For instance, a study based on similar data sets as this one by Muscheler et al. 2014 applied a 61 pt binomial filter that is similar to a 25 years running mean in order to "minimize[s] weather-related noise in the ^{10}Be to a large extent" [Muscheler et al., 2014]. However, Berggren et al. (2009) who worked with Dye-3 and NGRIP data used a band pass filter to process the data. Muscheler et al. (2014) summarize this discussion by stating that "results do not depend on the filter method as long as the typical solar variations on multi-decadal to centennial time scales are passed" [Ibid.]. For this work, it was decided to follow the example of McCracken et al. (2004). These authors, being aware of aliasing effects, used a 22 running mean to show that there is a good correlation of 87 % (91 % without outliers) between the Dye-3 (annual resolution) and South Pole (7-8 year resolution) data. Since this study works with similar sample lengths and a similar time frame, this was decided to be reasonable. This choice was further strengthened by a recent study of Steinhilber et al. (2012) who applied a 22 year running average as well for smoothing their data sets of various temporal resolutions for a principal component analysis.

In this study, **interpolation** was performed linearly, in order to reach a higher temporal resolution. So far mentioned, two different ways are used by other scientists. Ruth et al. (2007) applied a spline interpolation to achieve a temporal resolution of two years. On the other side, for example

Muscheler et al. (2014) and Abreu et al. (2012) used a linear interpolation to reach a higher temporal resolution. Since none of the above mentioned methods takes into account the seasonal variation pattern of ^{10}Be e.g. summer peaks due to intrusions of ^{10}Be rich air from the stratosphere [Pedro et al., 2011], there is probably no real difference between the two methods. In order to make the model more realistic, such a seasonal pattern function could be applied. However, since summer melting occurs at Akademii Nauk the seasonal pattern might be transformed to a even more complex structure due to percolation.

Finally, the **evaluation methods** should be discussed. As already indicated above, the cosine similarity method produces similar results as the two other methods, but the maximal discrepancy between the maximum and minimum is so small (about 0.059 ‰ of the mean value; standard deviation: 0.018 ‰) that the method is considered subordinate to the other evaluation methods. Apart from that, the correlation coefficient is used a lot by other researchers (e.g. [Abreu, 2009], [Horiuchi et al., 2008], [McCracken, 2004], [Muscheler et al., 2014]) to compare different curves in the context of ^{10}Be. A method comparable to the evaluation by the neighborhood method was not found. However, since it produces the same results as the correlation coefficient method, it was approved as well. The combination of these three complementary methods has the advantage that they evaluate different parameters of the curve and so a failure of all is less likely.

6 Conclusion and Outlook

A sample scheme has been set up for the Akademii Nauk ice core in order to reconstruct the long-term variations of the ^{10}Be concentration. As target of the sampling the reconstruction of the ^{10}Be concentration variations has been defined. The sample framework has been outlined as the core fraction older than 1600 AD. The systematic sampling method has been tested further in a numerical computer model in order to identify the the optimal sample mode and period. Important characteristics of the Akademii Nauk ice core as dating uncertainties due to percolation and the partly absence of yearly markers and occasionally missing core parts have been implemented into the model. The following conclusions are drawn:

- Multi year sampling achieves a higher degree of match than single point sampling because no information is lost and the strong variations of ^{10}Be concentration during the 11 year solar circle is already suppressed to a certain extent. An optimum in sample length is reached for sample lengths of 80, 70, 60, 50 and 40 data points, among which l=80 is slightly better than the other options. A reconstruction with samples of the length l=220 also produces very good results, but this sample length has been subordinated to the other sample lengths because a visual analysis shows that such long samples smooth out important variations of the curve. These findings were confirmed by all three evaluation methods. Moreover, these optima are even stable when taking into account that the ice core occasionally has missing parts.

- A gap between practice and theory of the model was identified for single point sampling. This sample mode was modeled insufficiently realistic enough, because sample sizes were chosen too small (0.1 years) than percolation effects require. Furthermore, the implementation of the uncertainty is under consideration. The model assumes a randomly distributed uncertainty, though it is more likely that it varies systematically. In fact, a positive correlation of uncertainty and sample lengths is possible. This might improve the results of the smaller sample lengths of l=20 and l=10 data points because the large ratio of uncertainty to sample length is suspected to be the reason for the particularly bad results.

- The specific suggestion of this study is to sample in 8 year intervals. The corresponding depth needs to be calculated taking the accumulation rate and layer thinning into account. However, since such a pure calculation does not consider practical issues, it was proposed to test whether sampling in equidistant steps using the whole range of the appropriate sample lengths would produce reliable results as well.

  Since there is already a lot of data for the ice core, and only some “save points” are missing, I suggest to identify a prominent extremum with strong variations of the ^{10}Be concentration in another records, estimate the corresponding depth in the Akademii Nauk ice core and fill up the remaining gaps in the record with continuous samples. This approach is likely to result in a save “anchor point” for the existing record.

- The approach of this study is consistent with a lot of other studies using wiggle matching for dating. Most authors use continuous sampling, interpolate their data linearly and smooth by applying moving averages or binomial filters. Some authors also use low-pass or bandpass filters, but this does not influence the results.

- The influence of aliasing is subject of a controversial discussion. One the one hand, researchers think that dilution effects in ice and irregular spaced sample intervals reduce aliasing effects to a negligible minimum. On the other hand, this study provides hints that that sampling can bias a record strongly due to unresolved frequencies of the 11 solar cycle.

- For further research on this topic I consider a modification of the single point sampling method to longer sample intervals. This might be combined with a change of the sampling method to a stratified random sampling to reduce aliasing. A more detailed study on the nature of uncertainty needs to be done as well to produce more realistic results. For further investigations on aliasing effects in the sampled data, I suggest to carry out a frequency analysis of the original and sampled data.

References

[Abreu, 2009] Abreu, J. (2009). *10Be in polar ice cores and 14C in tree rings: Separation of production and system effects, comparison with past climate changes and implications for solar physics*. PhD thesis, ETH Zürich.

[Abreu et al., 2013] Abreu, J. A., Beer, J., Steinhilber, F., Christl, M., and Kubik, P. W. (2013). 10 Be in Ice Cores and 14 C in Tree Rings: Separation of Production and Climate Effects. *Space Science Reviews*, 176(1-4):343–349.

[Bard et al., 2000] Bard, E., Raisbeck, G., Yiou, F., and Jouzel, J. (2000). Solar irradiance during the last 1200 years based on cosmogenic nuclides. *Tellus B*, 52(3):985–992.

[Beer, 2000] Beer, J. (2000). Long-term indirect indices of solar variability. *Space Science Reviews*, 94(1/2):53–66.

[Beer, 2001] Beer, J. (2001). Ice core data on climate and cosmic ray changes. In Kirkby, J., editor, *Workshop on Ion-Aerosol-Cloud Interactions*. pp. 3-11.

[Beer et al., 1990] Beer, J., Blinov, A., Bonani, G., Finkel, R., Hofmann, H., Lehmann, B., Oeschger, H., Sigg, A., Schwander, J., Staffelbach, T., Stauffer, B., Suter, M., and Wölfli, W. (1990). Use of 10Be in Polar ice to trace the 11-year cycle of solar activity. *Nature*, (347):164–167.

[Beer et al., 2012] Beer, J., McCracken, K. G., and Steiger, R. v. (2012). *Cosmogenic radionuclides: Theory and applications in the terrestrial and space environments*. Physics of earth and space environments. Springer, Heidelberg and New York.

[Berggren et al., 2009] Berggren, A.-M., Beer, J., Possnert, G., Aldahan, A., Kubik, P., Christl, M., Johnsen, S. J., Abreu, J., and Vinther, B. M. (2009). A 600-year annual 10 Be record from the NGRIP ice core, Greenland. *Geophysical Research Letters*, 36(11).

[Borradaile, 2003] Borradaile, G. J. (2003). *Statistics of earth science data: Their distribution in time, space, and orientation*. Springer, Berlin and New York. pp 1-17.

[Bronk Ramsey, 1998] Bronk Ramsey, C. (1998). Probability and Dating, Radiocarbon, 40 (1) 461-474. *Radiocarbon*, (40):461–474.

[Burg et al., 2014] Burg, J., Romney, J., and Schwartz, E. (2014). Digital Sound and Music, Concepts, Applications and Science. http://digitalsoundandmusic.com/5-3-4-mathematics-and-algorithms-for-aliasing; accessed: 17/08/2015.

[Christen and Litton, 1995] Christen, J. A. and Litton, C. D. (1995). A bayesian approach to wiggle-matching. *Journal of Archaeological Science*, 22(6):719–725.

References

[Finkel, R.C. and Nishiizumi, K., 1997] Finkel, R.C. and Nishiizumi, K. (1997). Beryllium 10 concentrations in the Greenland Ice Sheet Project 2 ice core from 340 ka. *Journal of Geophysical Research*, 102(C12):26,699–26,706.

[Fritzsche et al., 2005] Fritzsche, D., Schütt, R., Meyer H., Miller, H., Wilhelms, F., Opel, T., and Savatyugin, L. (2005). A 275 year ice core record from Akademii Nauk ice cap, Severnaya Zemlya, Russian Arctic. *Annals of Glaciology*, (42):361–366.

[Harris and Jarvis, 2011] Harris, R. and Jarvis, C. (2011). *Statistics for Geography and Environmental Science*. Pearson Education Limited, Edinburgh gate. pp 82-113.

[Heikkilä et al., 2013] Heikkilä, U., Beer, J., Abreu, J. A., and Steinhilber, F. (2013). On the Atmospheric Transport and Deposition of the Cosmogenic Radionuclides (10Be): A Review. *Space Science Reviews*, 176(1-4):321–332.

[Henderson, 2002] Henderson, K. (2002). *An ice core paleoclimate study of windy dome, Franz Josef Land (Russia): Development of a recent climate history for the barents sea* . PhD thesis, Graduate School of the Ohio State University.

[Horiuchi et al., 2007] Horiuchi, K., Ohta, A., Uchida, T., Matsuzaki, H., Shibata, Y., and Motoyama, H. (2007). Concentration of 10Be in an ice core from the Dome Fuji station, Eastern Antarctica: Preliminary results from 1500 to 1810 yr AD. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 259(1):584–587.

[Horiuchi et al., 2008] Horiuchi, K., Uchida, T., Sakamoto, Y., Ohta, A., Matsuzaki, H., Shibata, Y., and Motoyama, H. (2008). Ice core record of 10Be over the past millennium from Dome Fuji, Antarctica: A new proxy record of past solar activity and a powerful tool for stratigraphic dating. *Quaternary Geochronology*, 3(3):253–261.

[Mari et al., 1999] Mari, J. L., Glangeaud, F., and Coppens, F. (1999). *Signal processing for geologists & geophysicists*. Editions Technip, Paris. pp 204-206, 220-222.

[Masarik, 2009] Masarik, J. (2009). *Environmental Radionuclides: Tracers and Timers of Terrestrial Processes*, volume 16, chapter 1 Origin and Distribution of Radionuclides in the Continental Environment. Radioactivity in the Environment.

[McCracken, 2004] McCracken, K. G. (2004). A phenomenological study of the long-term cosmic ray modulation, 850–1958 AD. *Journal of Geophysical Research*, 109(A12).

[McHargue and Damon, 1991] McHargue, L. R. and Damon, P. E. (1991). The global beryllium 10 cycle. *Reviews of Geophysics*, 29(2):141.

[Middleton, 2000] Middleton, G. V. (2000). *Data analysis in the earth sciences using Matlab®*. Prentice Hall, Upper Saddle River, NJ. p 84.

[Miller et al., 2010] Miller, G. H., Brigham-Grette, J., Alley, R. B., Anderson, L., Bauch, H. A., Douglas, M., Edwards, M. E., Elias, S. A., Finney, B. P., Fitzpatrick, J. J., Funder, S. V., Herbert, T. D., Hinzman, L. D., Kaufman, D. S., MacDonald, G. M., Polyak, L., Robock, A., Serreze, M. C., Smol, J. P., Spielhagen, R., White, J., Wolfe, A. P., and Wolff, E. W. (2010). Temperature and precipitation history of the Arctic. *Quaternary Science Reviews*, 29(15-16):1679–1715.

[Miyahara et al., 2004] Miyahara, H., Masuda K., Furuzawa, H., Menjo, H., Muraki Y., Kitagawa H., and Nakamura, T. (2004). Variation of theradiocarbon content in tree rings during the Spoerer Minimum. *Radiocarbon*, (46):955–958.

[Mudelsee, 2014] Mudelsee, M. (2014). *Climate time series analysis: Classical statistical and bootstrap methods*, volume 51. Springer, Heidelberg, second edition.

[Muscheler et al., 2014] Muscheler, R., Adolphi, F., and Knudsen, M. F. (2014). Assessing the differences between the IntCal and Greenland ice-core time scales for the last 14,000 years via the common cosmogenic radionuclide variations. *Quaternary Science Reviews*, 106:81–87.

[Opel et al., 2009] Opel, T., Fritzsche, D., Meyer, H., Schütt, R., Weiler, K., Ruth, U., Wilhelms, F., and Fischer, H. (2009). 115 year ice-core data from Akademii Nauk ice cap, Severnaya Zemlya: high-resolution record of Eurasian Arctic climate change. *Journal of Glaciology*, 55(189):21–31.

[Pedro et al., 2011] Pedro, J. B., Smith, A. M., Simon, K. J., van Ommen, T. D., and Curran, M. A. J. (2011). High-resolution records of the beryllium-10 solar activity proxy in ice from Law Dome, East Antarctica: measurement, reproducibility and principal trends. *Climate of the Past*, 7(3):707–721.

[Peristykh and Damon, 1998] Peristykh, A. and Damon, P. (1998). Modulation of Atmospheric 14C Concentration by the Solar Wind and Irradiance Components of the Hale and SCHWABE Solar Cycles. *Solar Physics*, (177):343–355.

[Pisias and Mix, 1988] Pisias, N. G. and Mix, A. C. (1988). Aliasing of the geologic record and the search for long-period Milankovitch cycles. *Paleoceanography*, 3(5):613–619.

[Pohjola et al., 2002] Pohjola, V. A., Moore, J. C., Isaksson, E., Jauhiainen, van de Wal, R.S.W., Martma, T., Meijer, H., and Vaikmäe, R. (2002). Effect of periodic melting on geochemical and isotopic signals in an ice core from Lomonosovfonna, Svalbard. *Journal of Geophysical Research*, 107(D4).

[Raisbeck et al., 1998] Raisbeck, G. M., Yiou, F., Bard, E., Dollfus, D., Jouzel, J., and Petit, J. R. (1998). Absolute Dating of the Last 7000 Years of the Vostok Ice Core Using 10Be. *Mineralogical Magazine*, 62A(2):1228.

[Raisbeck et al., 2007] Raisbeck, G. M., Yiou, F., Jouzel, J., and Stocker, T. (2007). Direct north-south synchronization of abrupt climate change record in ice cores using Beryllium 10. *Climate of the Past, European Geosciences Union (EGU)*, (3):541–547.

[Ruth et al., 2007] Ruth, U., Barnola, J.-M., Beer, J., Bigler, M., Blunier, T., Castellano, E., Fischer, H., Fundel, F., Huybrechts, P., Kaufmann, P., Kipfstuhl, S., Lambrecht, A., Morganti, A., Oerter, H., Parrenin, F., Rybak, O., Severi, M., Udisti, R., Wilhelms, F., and Wolff, E. (2007). "EDML1": a chronology for the EPICA deep ice core from Dronning Maud Land, Antarctica, over the last 150 000 years. *Climate of the Past*, 3(3):475–484.

[Steinhilber et al., 2012] Steinhilber, F., Abreu, J. A., Beer, J., Brunner, I., Christl, M., Fischer, H., Heikkila, U., Kubik, P. W., Mann, M., McCracken, K. G., Miller, H., Miyahara, H., Oerter, H., and Wilhelms, F. (2012). 9,400 years of cosmic radiation and solar activity from ice cores and tree rings. *Proceedings of the National Academy of Sciences*, 109(16):5967–5971.

[Vonmoos et al., 2006] Vonmoos, M., Beer, J., and Muscheler, R. (2006). Large variations in Holocene solar activity: Constraints from 10 Be in the Greenland Ice Core Project ice core. *Journal of Geophysical Research*, 111(A10).

[Walker, 2005] Walker, M. J. (2005). *Quaternary dating methods*. J. Wiley, Chichester, West Sussex, England. pp 35-37.

[Weiler et al., 2005] Weiler, K., Fischer, H., Fritzsche, D., Ruth, U., Wilhelms, F., and Miller, H. (2005). Glaciochemical reconnaissance of a new ice core from Severnaya Zemlya, Eurasian Arctic. *Journal of Glaciology*, 51(172):64–74.

[Wescott, 2015] Wescott, T. (2015). Sampling: What Nyquist Didn't Say, and What to Do About It. http://www.wescottdesign.com/articles/Sampling/sampling.pdf; accessed: 04/26/2015.

[Yiou et al., 1997] Yiou, F., Raisbeck, G. M., Baumgartner, S., Beer, J., Hammer, C., Johnsen, S., Jouzel, J., Kubik, P. W., Lestringuez, J., Stiévenard, M., Suter, M., and Yiou, P. (1997). Beryllium 10 in the Greenland Ice Core Project ice core at Summit, Greenland. *Journal of Geophysical Research*, 102(C12):26.783–26.794 // 26783.

Appendix

Appendix A: Smoothing with Moving Averages

Simple moving averages calculate a mean for each point of the curve. In order to calculate an average over 22 years, 22 needs to be among the multiples of the original temporal resolution. So, for example, samples with a temporal resolution of 2 years are easily averaged to 22 years by calculating the mean value of 11 samples. But for samples that have a temporal resolution of e.g. 7 years, a weighing term needs to be introduced to achieve an average of 22 years. This slightly changed version of the simple running mean is referred to as weighted moving average (WMA). The weighing term is applied symmetrically to the end points.

$$SMA_i = \frac{d_{i-n} + d_{i-n+1} + ... + d_{i-1} + d_i + d_{i+1} + ... + d_{i+n-1} + d_{i+n}}{2n+1} \tag{6.1}$$

$$WMA_i = \frac{b_{i-n}d_{i-n} + b_{i-n+1}d_{i-n+1} + ... + b_{i-1}d_{i-1} + b_i d_i + b_{i+1}d_{i+1} + ... + b_{i+n-1}d_{i+n-1} + b_{i+n}d_{i+n}}{\sum_{z=i-n}^{z=i+n} b_z} \tag{6.2}$$

with $1 \leq n < i$ and $n, i \in \mathbb{N}$, d_i is the ^{10}Be concentration at point i and $2n+1$ the number of samples included in the moving average. The weighing factor is b_i. For instance, regarding samples with a length of 7 years, the SMA for point i averages 21 years and is calculated by

$$SMA_i = \frac{d_{i-1} + d_i + d_{i-1}}{3}$$

The symmetrically weighted running mean averaging 22 years is calculated by

$$WMA_i = \frac{\frac{1}{14}d_{i-2} + \frac{7}{7}d_{i-1} + \frac{7}{7} * d_i + \frac{7}{7}d_{i+1} + \frac{1}{14}d_{n+2}}{\frac{1}{14} + \frac{7}{7} + \frac{7}{7} + \frac{7}{7} + \frac{1}{14}}$$

These two different ways of averaging have been compared to each other and the reference curve (22 years running mean of original data (figure 6.1)). The figures disclose that there is barely a difference between the different averages. In case there is a slight difference, the weighted moving average is closer to the 22 years running mean than the simple moving average. This visual impression is confirmed by the findings of the evaluation methods that show better results for the WMA (figure 6.2). Therefore, it was concluded to use the WMA for the following considerations. The fact that all curves are averaged to the same term allows to exclude the error source that different running means (e.g. 20, 21, 22, 24) have any influence on the evaluation.

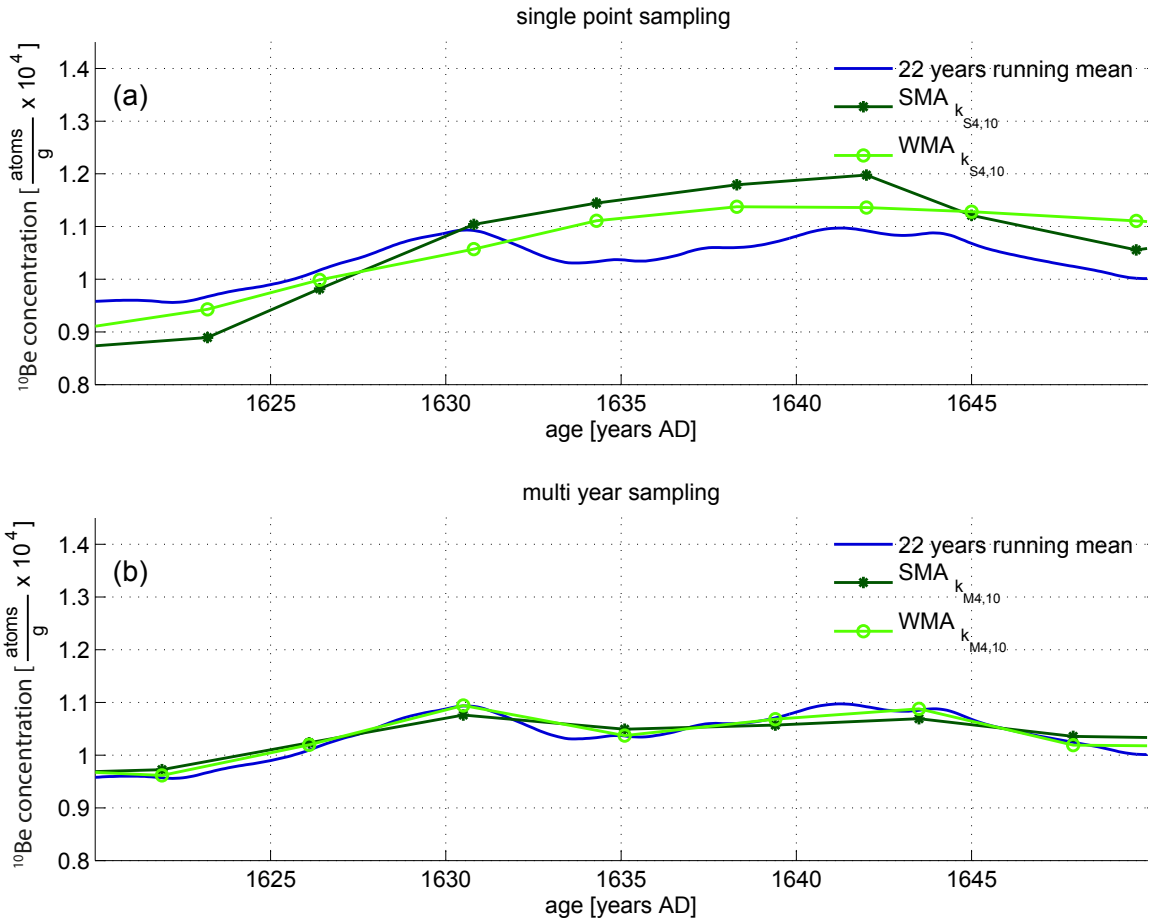

Figure 6.1: Enlarged section of the Dye-3 record for **comparison of a simple moving average (SMA) and a weighted moving average (WMA) to the reference curve (22 years running mean)**. In (a) a single point sampling curve with step width s=40 and in (b) a multi year sampling curve with sample length l=40 are chosen exemplary for comparison. There is almost no difference between the two averages, though the WMA is slightly closer to the reference curve.

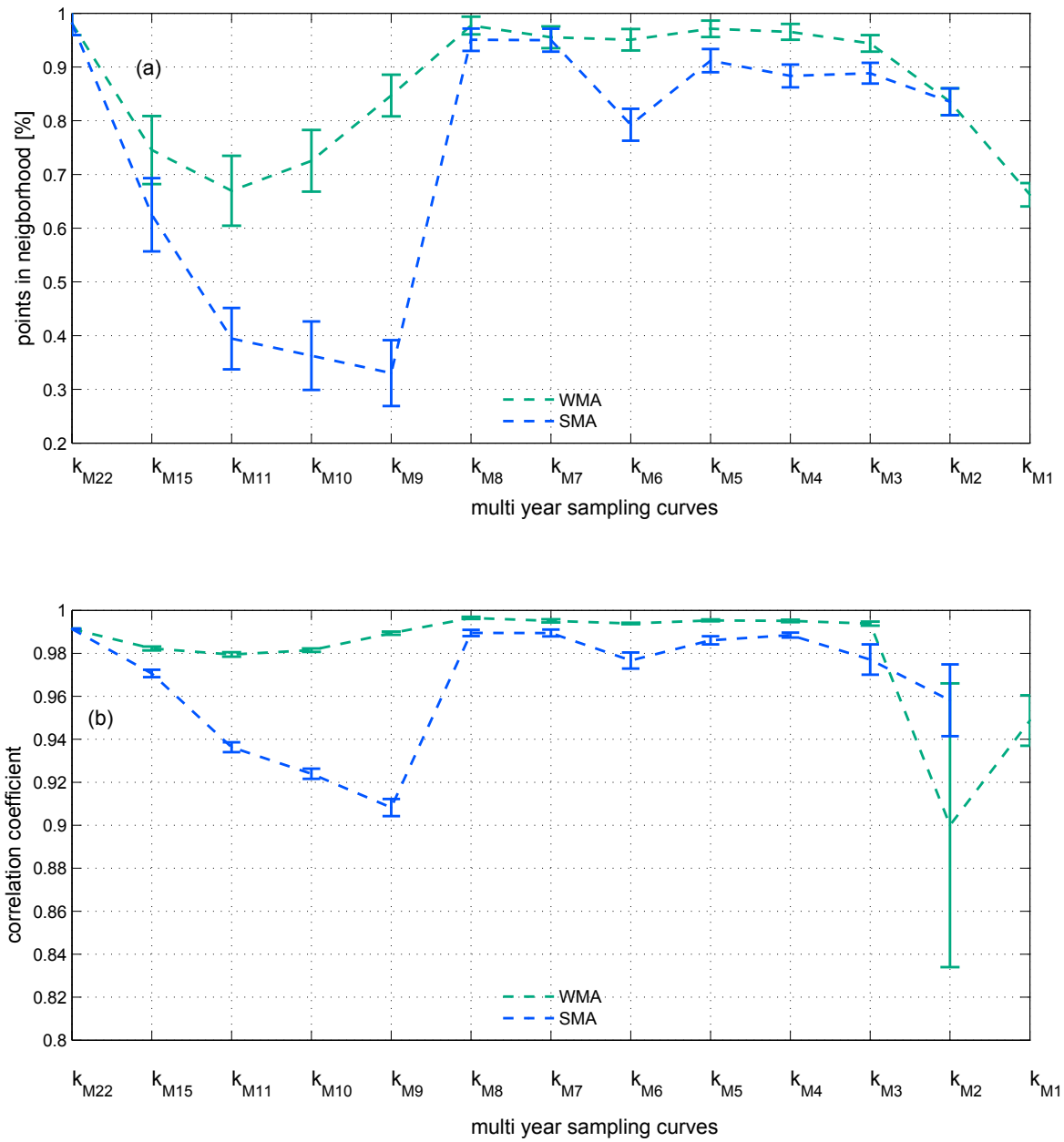

Figure 6.2: **Comparison of a simple moving average (SMA) and a weighted moving average (WMA)** with the neighborhood evaluation method and the correlation coefficient. In (a) the percentage of points within the neighborhood is displayed for the average of multi year sampling (50 runs). In (b) the correlation coefficient is shown. It can be stated that the results of the SMA are in general lower than the one of the WMA.

Appendix B: Neighborhood Evaluation

In the following it is explained how the parameter δ is determined that defines the mathematical neighborhood $B_\delta(m)$. Since δ is related to the local extrema of the curve, these local extrema have been detected by the Matlab© findpeak function in a first step. Then the ratio of peak prominence (h) to peak width (w) has been calculated for each peak (corresponding to the mean inclination of the peaks $\frac{h}{w}$). Among those, the smallest prominence to width ratio has been chosen. Next, the parameter δ was elected in a way that the general trend between two points, so a positive or negative inclination, is preserved. For this purpose at first, the time interval in that the trend must be present is selected (22 years). Then, in regard to the mean inclination the corresponding concentration difference is calculated (by $\frac{h}{w} \cdot 22$). Finally, this result is divided by 2, resulting in δ. All points that have a smaller distance to the running mean than δ are part of the neighborhood (see figure 6.3).

Later, it has been tested, whether there is an influence of δ on the results. It could be shown, that the general statements about local maxima are the same for different values of δ. As expected the degree of match decreases when a smaller value for δ is chosen.

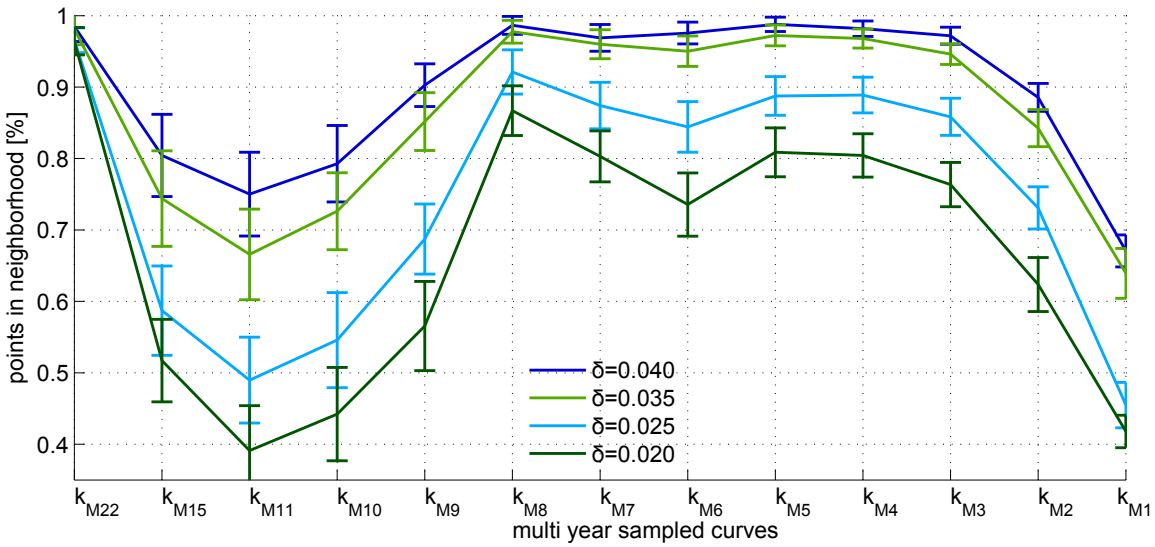

Figure 6.4: **Comparison of the influence of** δ on the results of the neighborhood evaluation method for multi year sampling. Displayed is the mean value of 50 runs and the mean standard deviation of 50 runs as errorbars for different values of δ. δ defines the size of the mathematical neighborhood. It is calculated relative to the minimal inclination in the record. For Dye-3 this results in δ=0.035 $\frac{atoms}{g} \times 10^4$. However, this figure shows that the statements about the optimal sample length are independent of δ because different values for $\delta = 0.040, 0.025, 0.020$ indicate the same local maxima, even though the degree of match is different

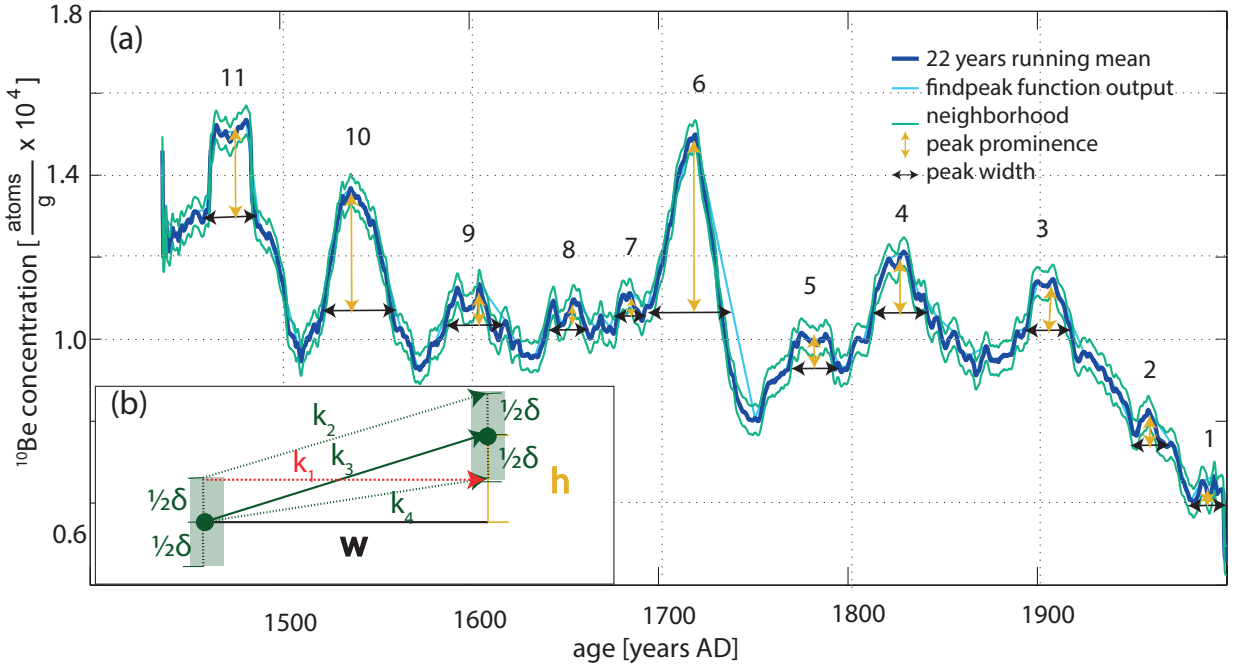

Figure 6.3: **Determination of δ defining the $B_\delta(m)$ neighborhood** exemplary done on the Dye-3 data. Local maxima (numbers 1-11) of the 22 years running mean have been identified by the Matlab findpeak function in (a). Then, the peak prominence to peak width ratio has been calculated ($\frac{h}{w}$). The smallest ratio has been selected. The general trend of this inclination (positive / negative) must be visible between two point that span an interval of 22 years. To preserve such trend, these two points must stay within a neighborhood of $\frac{h}{w} \cdot 22 \cdot \frac{1}{2}$ (see insert (b)). All curves (k_2, k_3, k_4) show the required positive inclination when connecting the two points within a neighborhood of δ. The connection k_1 is the worst case, connecting the highest point of the neighborhood of the start point with the lowest point of the neighborhood of the end point. In this case the slope is 0. Consequently, for all points within the neighborhood $B_\delta(m) = \{x_{k_{i,j}} \in k_{i,j} \mid \|y_{k_{i,j}} - m\| < \delta\}$ the inclination of the curve is positive (or zero for one point). For the Dye-3 data set the calculation results in $\delta = 0.035\ \frac{atoms}{g} \times 10^4$.

Appendix C: Additional Figures

In the following additional figures with Dye-3 data are plotted. Furthermore, figures with NGRIP and Dome Fuji data are displayed. The corresponding figures from the Results chapter with Dye-3 data are labeled.

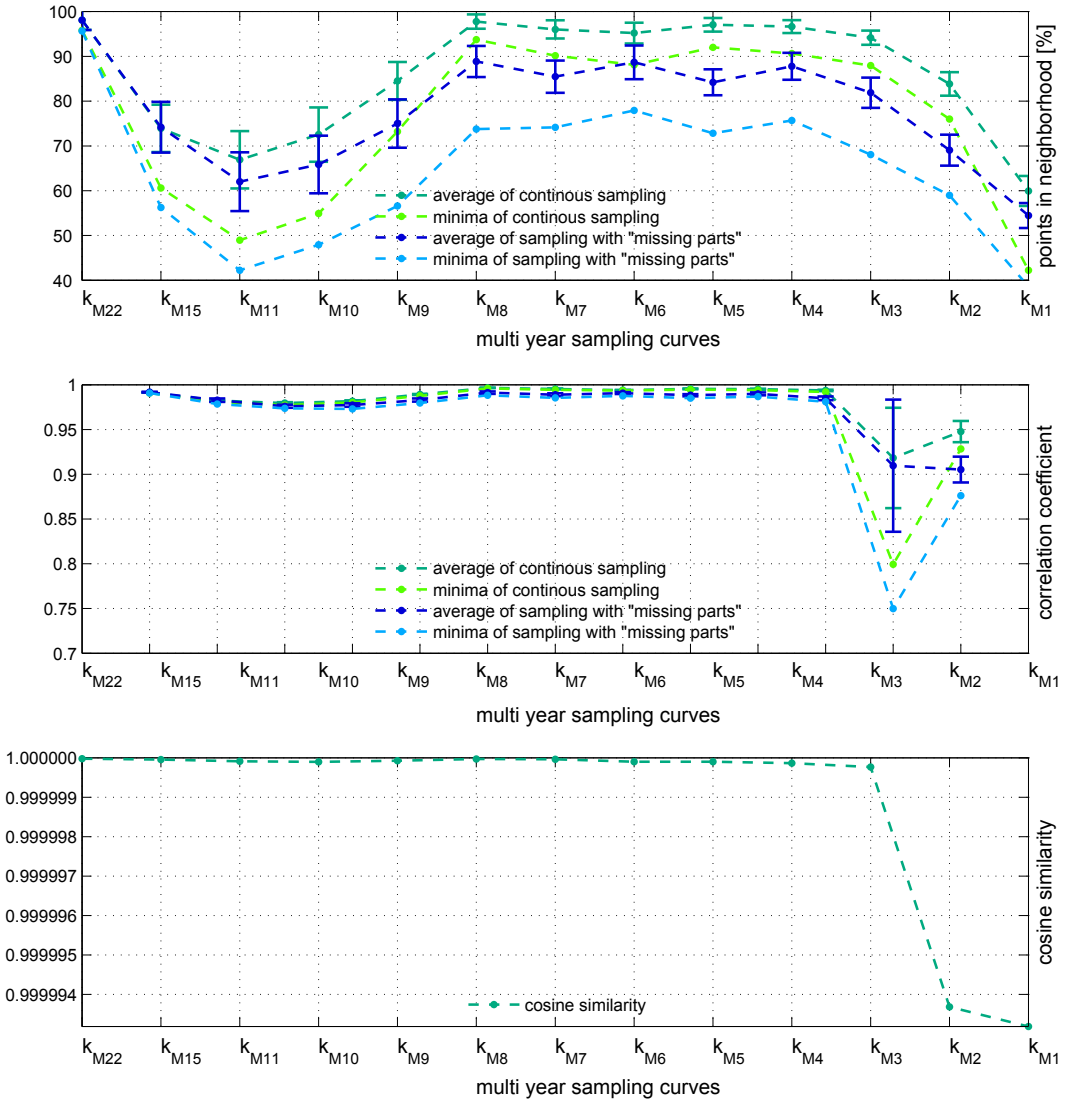

Figure 6.5: **Dye-3: Complete evaluation results for multi year sampling** (see figure 4.9). This figure contains all results (as well the results of k_{M2} and k_{M1} that were left out for figure 4.9 due to better display.) In (a) the percentage of points within the neighborhood $B(\delta, m)$ is displayed. In (b) the results of the correlation coefficient are shown. In (c) the results of the cosine similarity are summarized. For details see figure 4.9

NGRIP data set

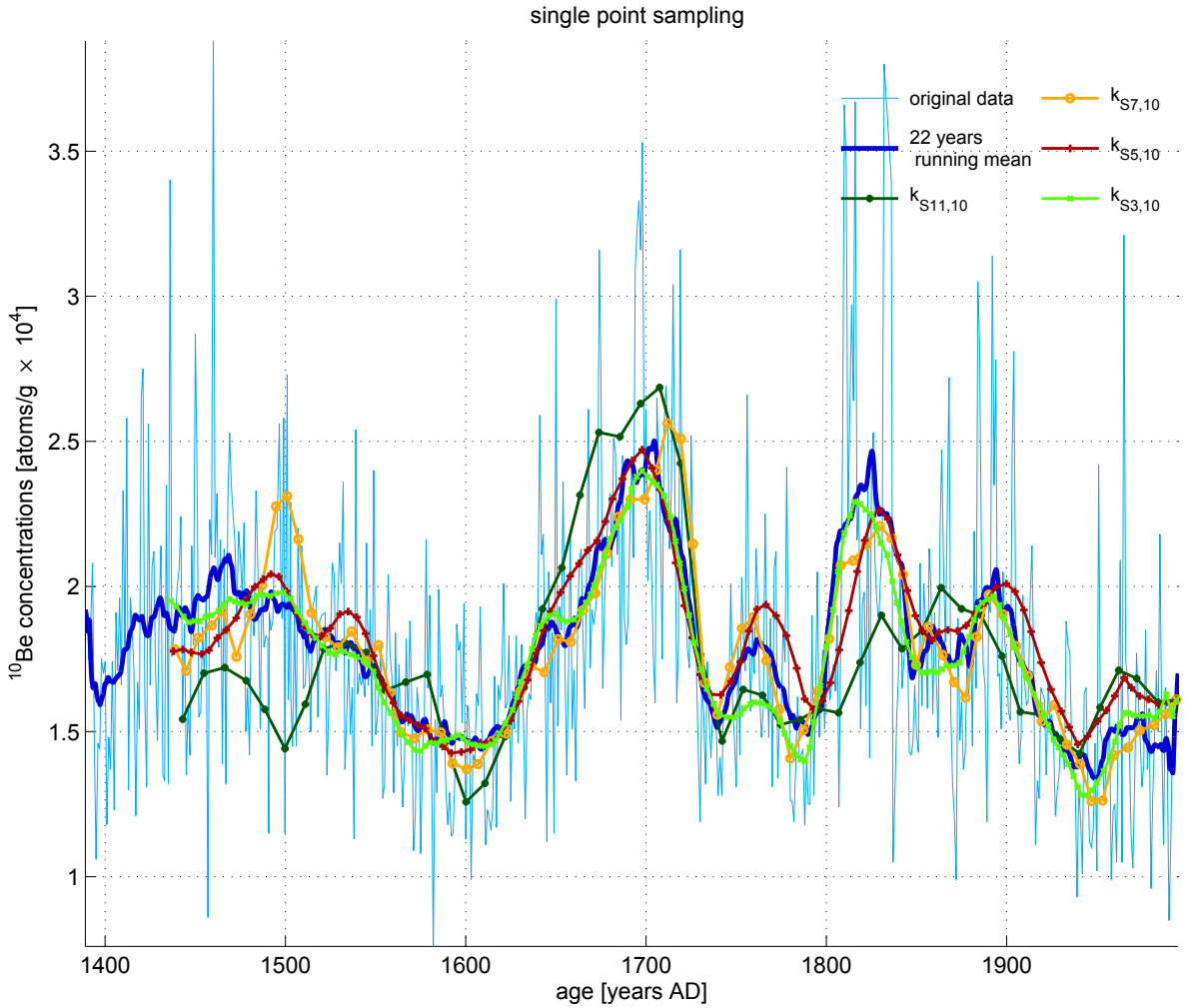

Figure 6.6: **NGRIP data: single point sampling curves** (see figure 4.5). The original data, the 22 years running mean and selected curves with starting point 10 are shown. Single point sampling curves with sample steps s=110, 70, 50, 30 are displayed. The single point sampling curves differ considerably from the 22 years running mean.

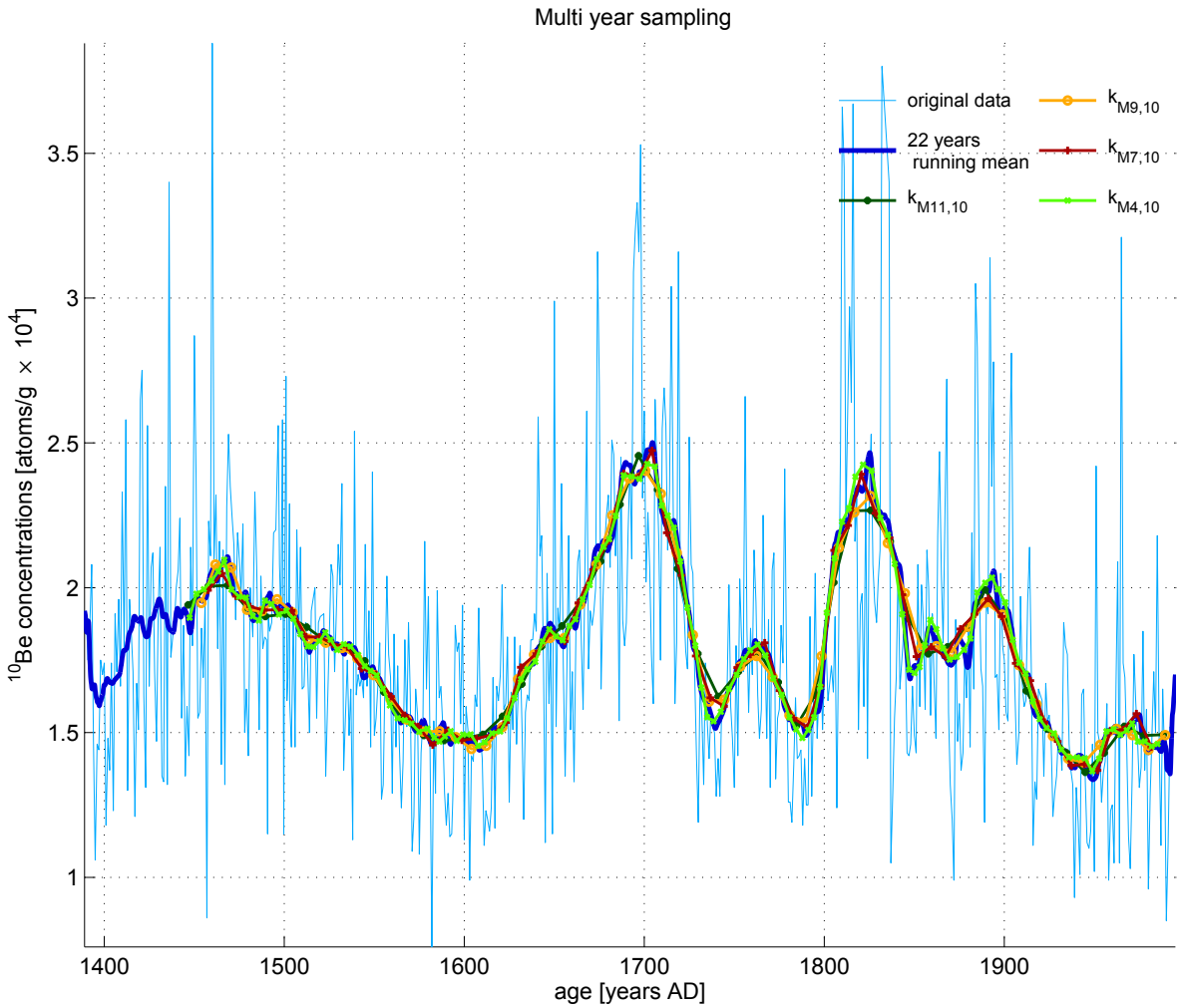

Figure 6.7: **NGRIP data: Evaluation results of multi year sampling curves** (see figure 4.5). The original data, the 22 years running mean and selected curves with starting point 10 are shown. Multi year sampling curves with sample lengths of l= 110, 90, 70, 40 are shown. The multi year sampling curves are all very close to the 22 years running mean.

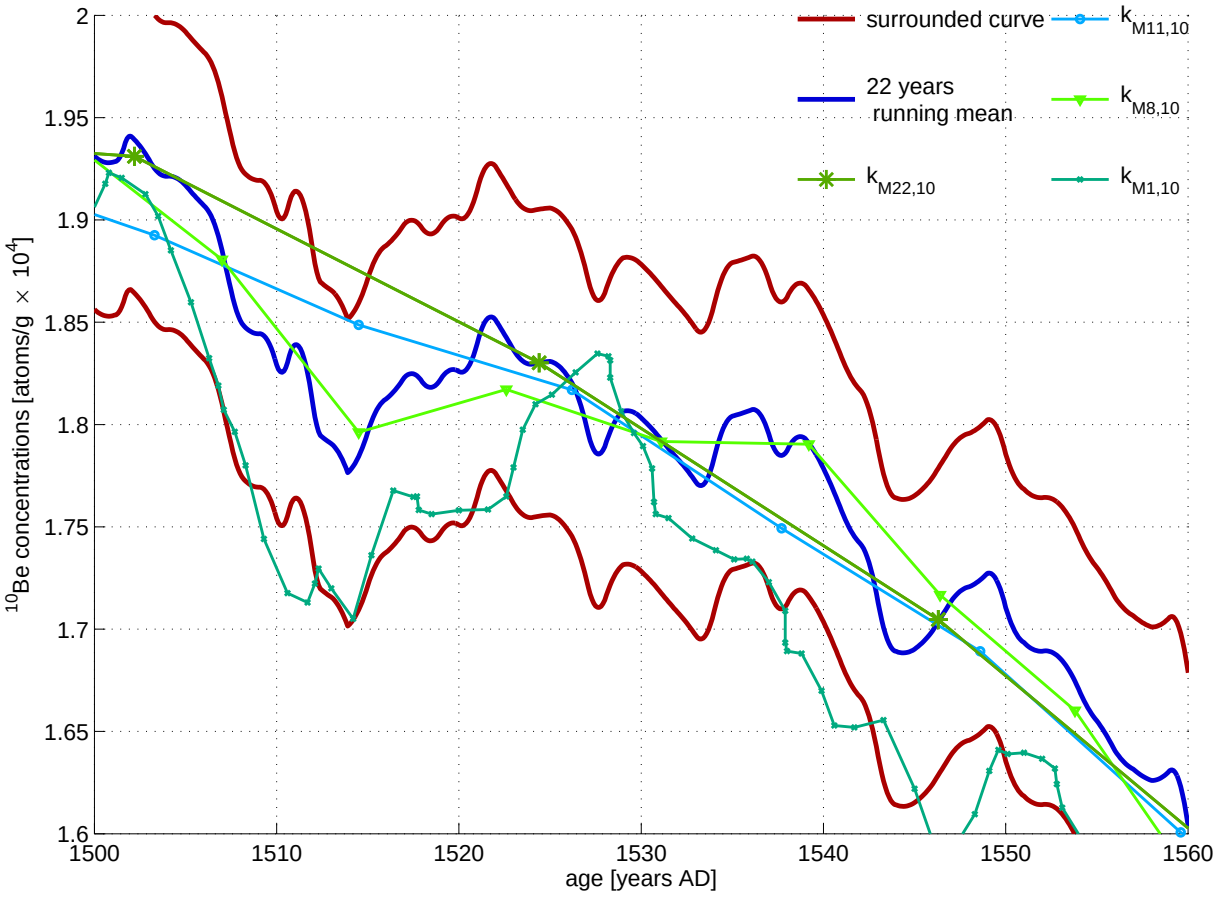

Figure 6.8: **NGRIP data: Display of the neighborhood $B(\delta, m)$ and selected multi year sampling curves** (see figure 4.7). The surrounded curve (red) that marks the borders of the neighborhood B(δ,m) is displayed with the 22 years running mean and selected multi year sampling curves for a time interval from 1550-1650. A visual analysis reveals that small sampling intervals as l=10 and long samples intervals as l=110. Intermediate sample lengths as l=70 plot almost completely within the neighborhood.

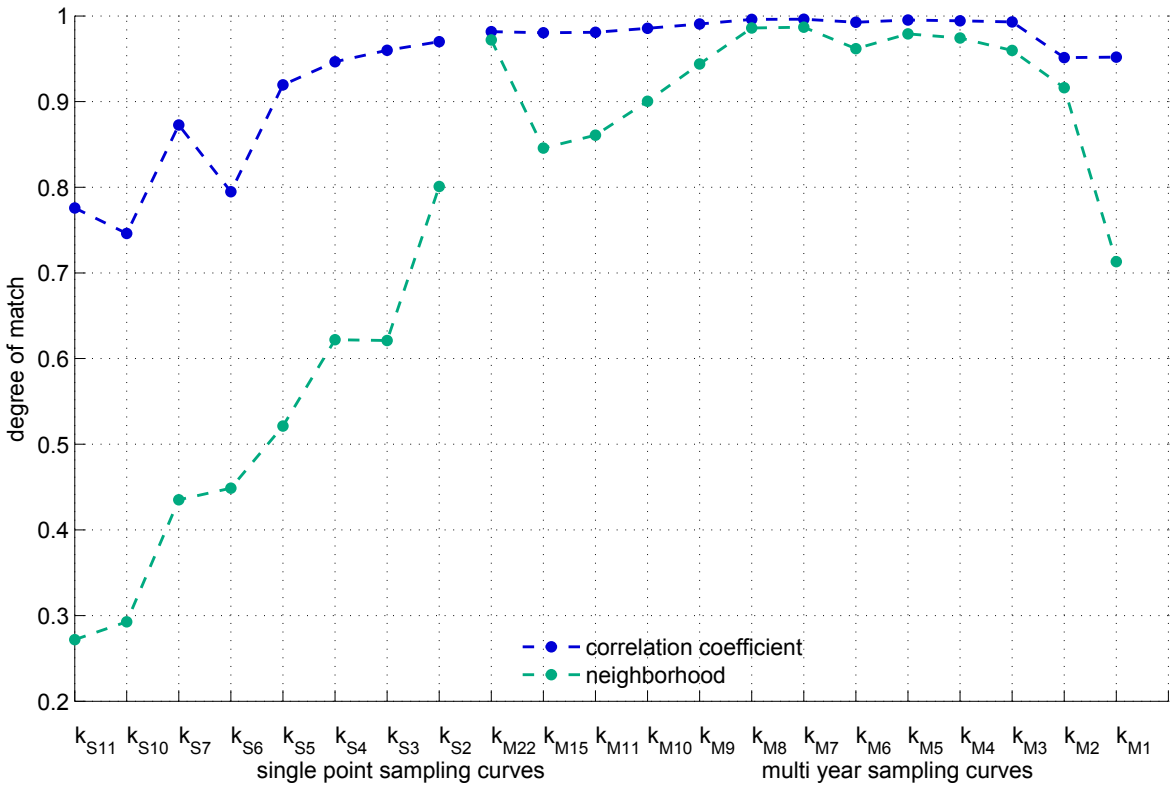

Figure 6.9: **NGRIP data: Comparison of multi year and single point sampling for neighborhood and correlation coefficient evaluation methods** (see figure 4.8). On the x-axis the different sample modes and periods are displayed and on the y-axis the corresponding mean value of 50 runs that expresses the degree of match (on a scale of 0–1) is shown. In general, it can be stated that multi year sampling receives higher matches than single point sampling.

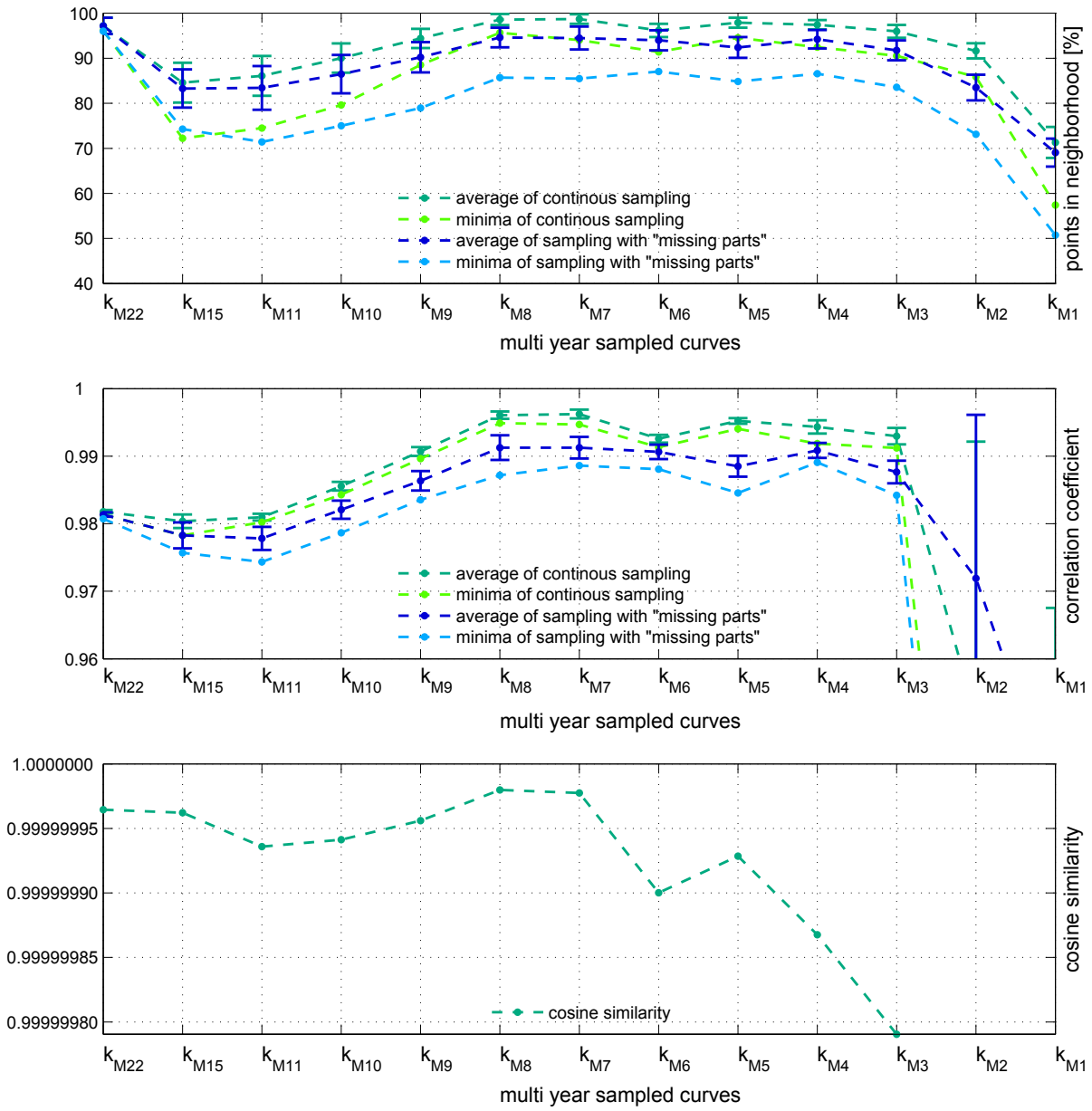

Figure 6.10: **NGRIP data: Evaluation results for multi year sampling** (see figure 4.9). For each sample length (x-axis) the corresponding degree of match is plotted (y-axis). In (a) the percentage of points within the neighborhood B(δ,m) is displayed. In (b) the results of the correlation coefficient are shown. In (c) the results of the cosine similarity are summarized. All curves show a similar pattern with a maximum in shape of a plateau at l=80 to l=40. A second maximum is at l=220.

Dome Fuji data set

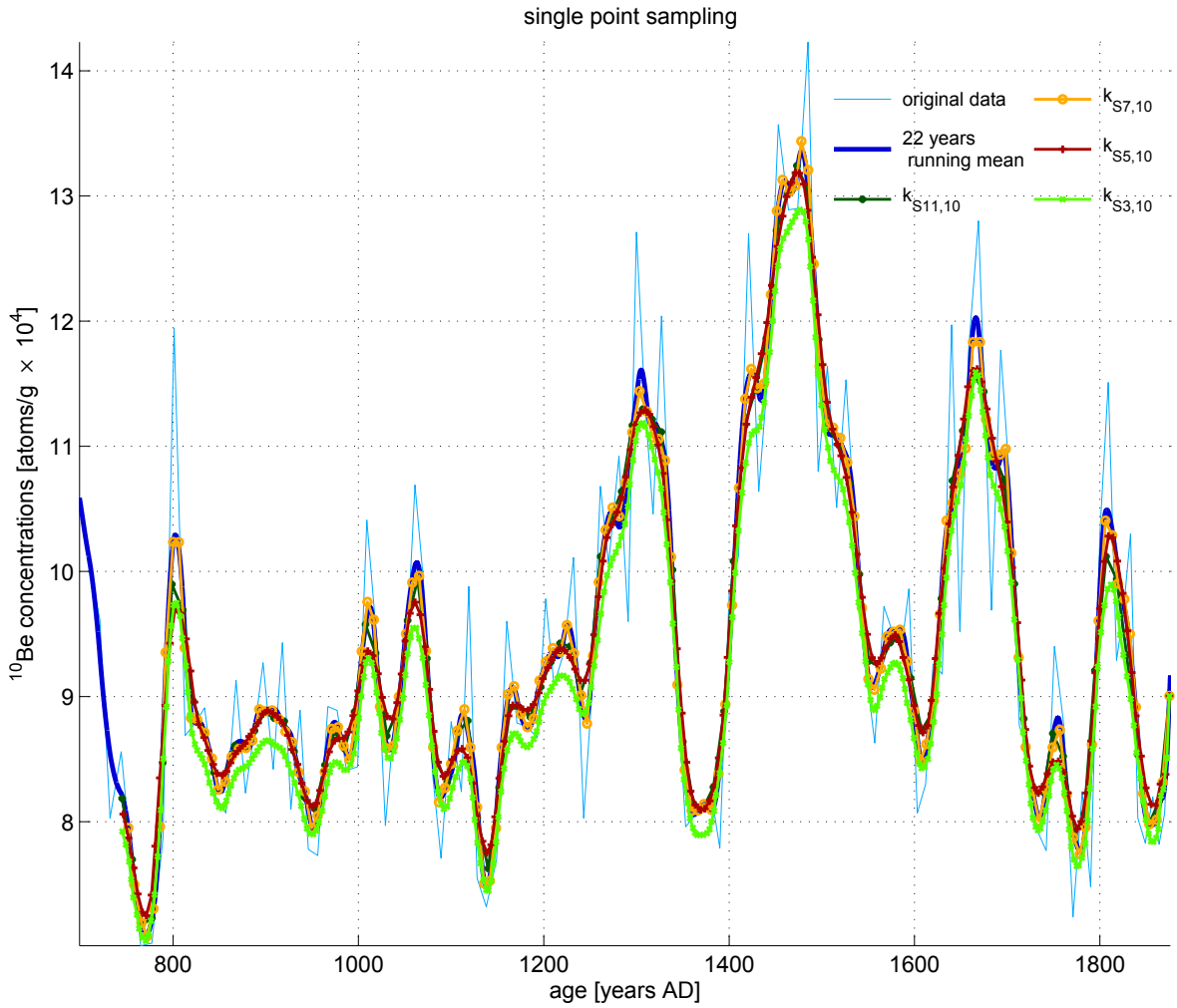

Figure 6.11: **Dome Fuji data: single point sampling curves** (see figure 4.5). The original data, the 22 years running mean and selected curves with starting point 10 are shown. Single point sampling curves with sample steps s=110,70,50,30 are displayed.The single point sampling curves differ considerably from the 22 years running mean.

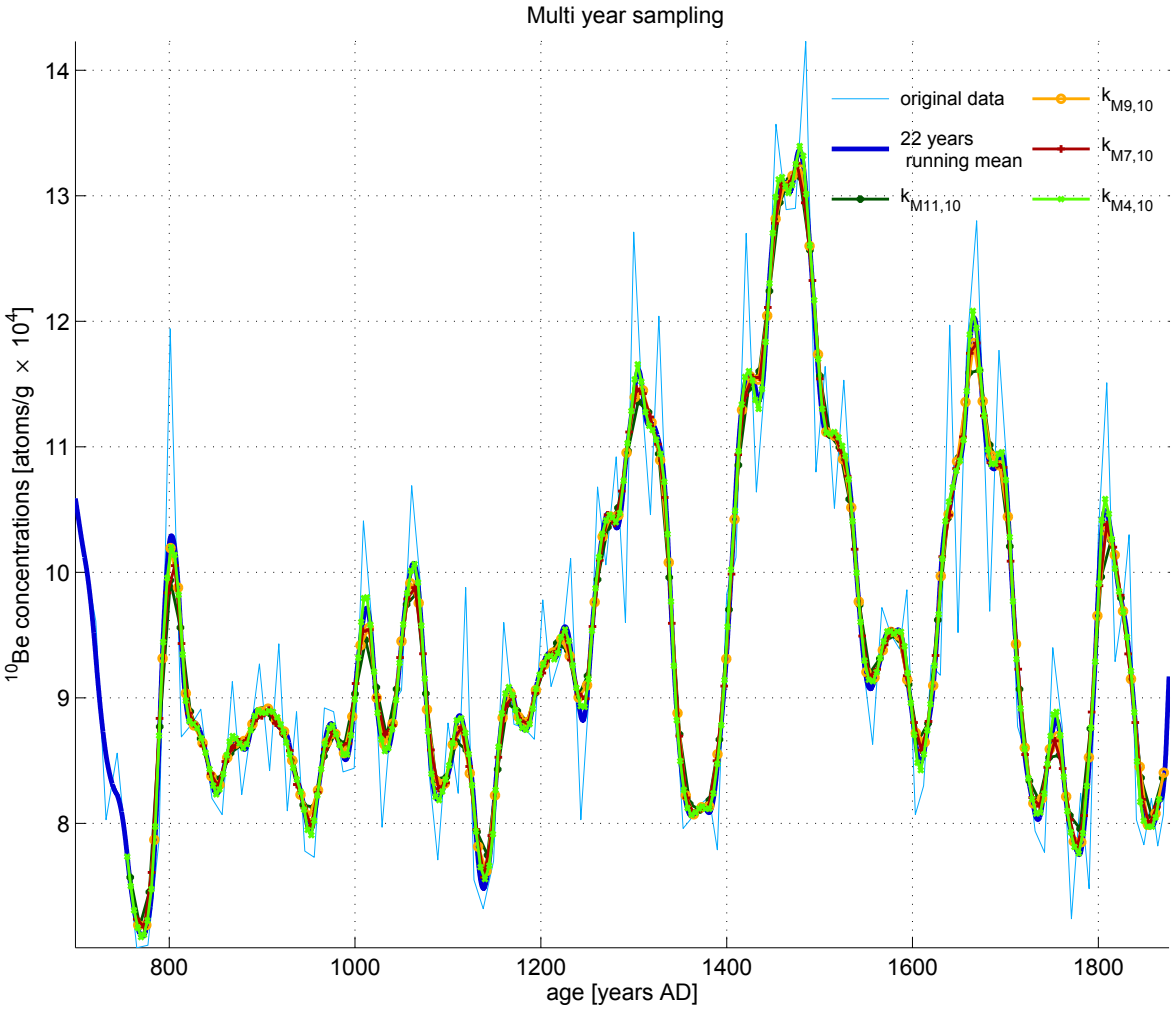

Figure 6.12: **Dome Fuji data: Evaluation results of multi year sampling curves** (see figure 4.5). The original data, the 22 years running mean and selected curves with starting point 10 are shown. Multi year sampling curves with sample lengths of l=110, 90, 70, 40 are shown. The multi year sampling curves are all very close to the 22 years running mean.

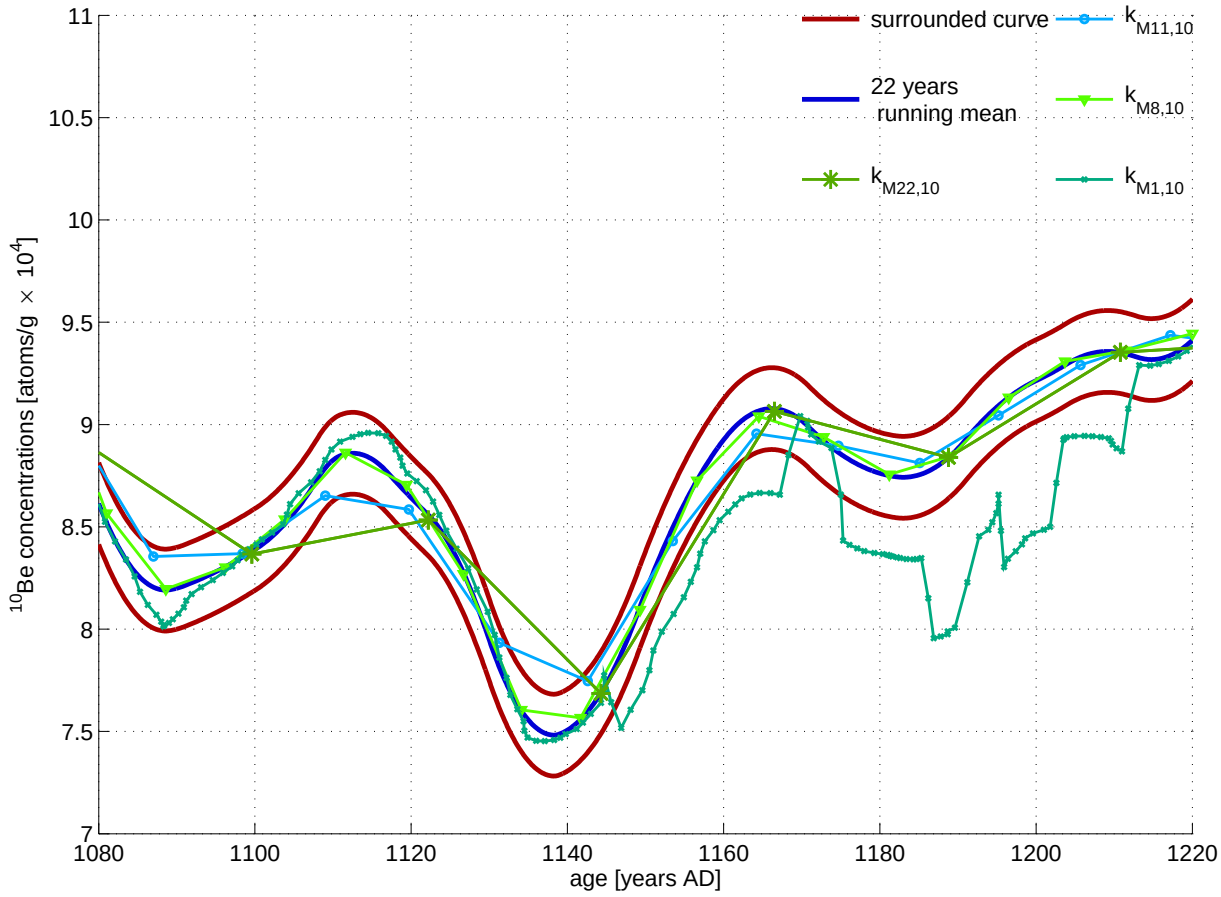

Figure 6.13: **Dome Fuji data: Display of the neighborhood $B(\delta, m)$ and selected multi year sampling curves** (see figure 4.7). The surrounded curve (red) that marks the borders of the neighborhood B(δ,m) is displayed with the 22 years running mean and selected multi year sampling curves for a time interval from 1550-1650. A visual analysis reveals that small sample intervals as l=10 and long samples intervals as l=110. Intermediate sample lengths as l=70 plot almost completely within the neighborhood.

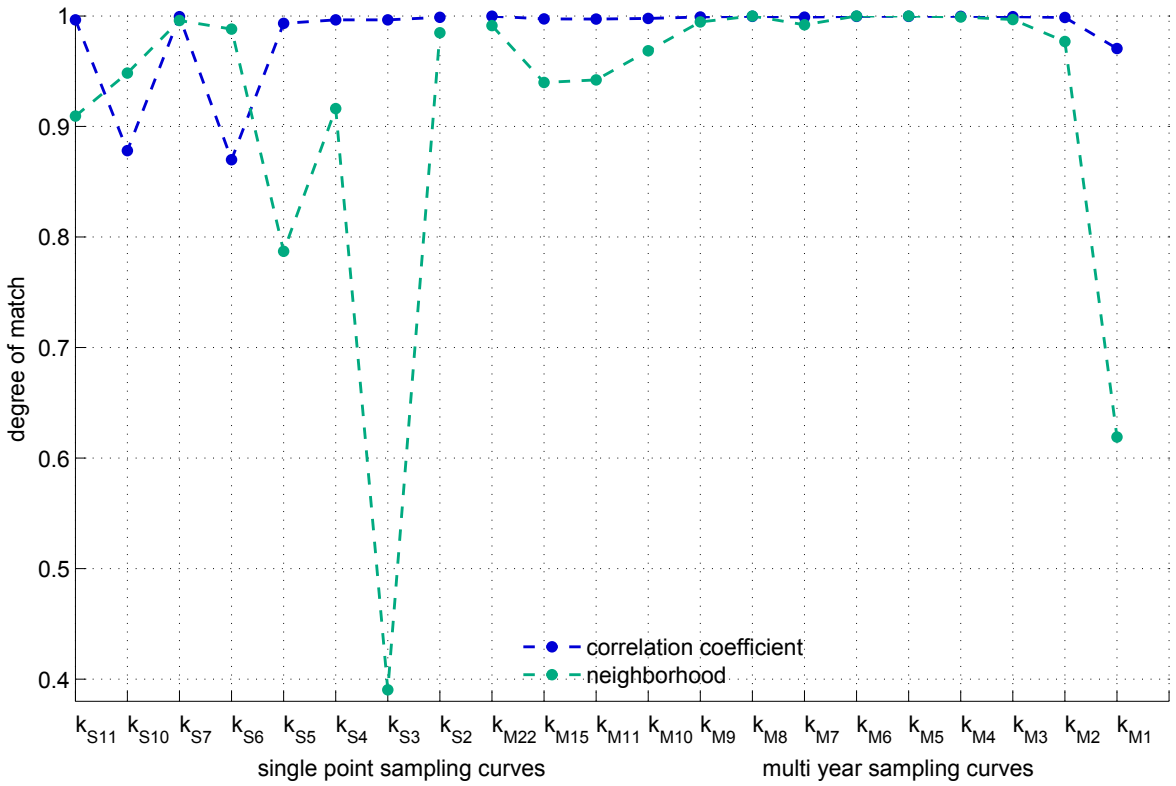

Figure 6.14: **Dome Fuji data: Comparison of multi year and single point sampling for neighborhood and correlation coefficient evaluation methods** (see figure 4.8). On the x-axis the different sample modes and periods are displayed and on the y-axis the corresponding mean value of 50 runs that expresses the degree of match (on a scale of 0–1) is shown. In general, it can be stated that multi year sampling receives a higher match than single point sampling.

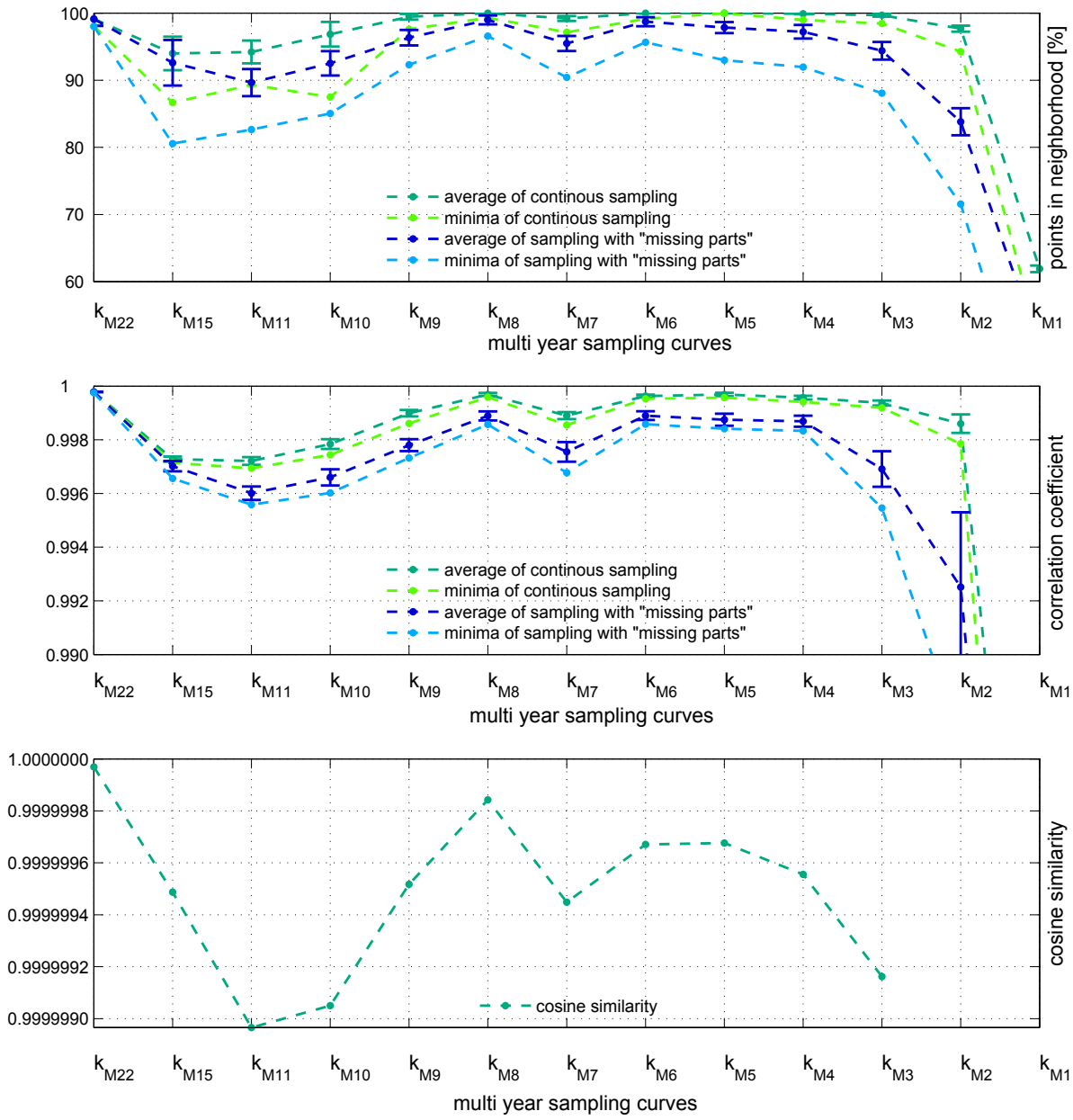

Figure 6.15: **Dome Fuji data: Evaluation results for multi year sampling** (see figure 4.9). For each sample length (x-axis) the corresponding degree of match is plotted (y-axis). In (a) the percentage of points within the neighborhood B(δ,m) is displayed. In (b) the results of the correlation coefficient are shown. In (c) the results of the cosine similarity are summarized. All curves show a similar pattern with a maximum in shape of a plateau at l=80 to l=40. A second maximum is at l=220.

Acknowledgements

Ich möchte meinem Betreuer am AWI Herrn Fritzsche herzlich dafür danken, dass er mir während meiner Bachelorarbeit spannende, praktische Einblicke in das Arbeiten mit Eisbohrkernen ermöglicht hat. Darüber hinaus danke ich ihm sehr für die Zeit, die er sich immer für meine Fragen genommen hat. Ebenso danke ich Frau Silke Merchel vom Helmholz-Zentrum Dresden-Rossendorf für die Zeit, die sie sich genommen hat, mir einen Einblick in das chemische Aufbereiten von ^{10}Be in Eis zu gewähren. Mein Dank gilt auch meiner zweiten Betreuerin Frau Ulrike Herzschuh. Bedanken möchte ich mich außerdem beim Alfred-Wegener-Institut, welches mich bei den beiden Praktika in Bremerhaven und Dresden-Rossendorf finanziell unterstützt hat.