Novel automated inversion algorithm for temperature reconstruction using gas isotopes from ice cores

Michael Döring[1,2*] and Markus Leuenberger[1,2]

[1]Climate and Environmental Physics, University of Bern, Switzerland
[2]Oeschger Centre for Climate Change Research (OCCR), Bern, Switzerland

Correspondence to: Michael Döring (doering@climate.unibe.ch)

Keywords: temperature reconstruction, ice core, nitrogen isotope, argon isotope, inverse model, firn model

Abstract. Greenland past temperature history can be reconstructed by forcing the output of a firn densification and heat diffusion model to fit multiple gas isotope data ($\delta^{15}N$ or $\delta^{40}Ar$ or $\delta^{15}N_{excess}$) extracted from ancient air in Greenland ice cores. We present here a novel methodology to solve this inverse problem, by designing a fully automated algorithm. To demonstrate the performance of this novel approach, we begin by intentionally constructing synthetic temperature histories and associated $\delta^{15}N$ datasets, mimicking real Holocene data that we use as "true values" (targets) to be compared to the output of the algorithm. This allows us to quantify uncertainties originating from the algorithm itself. The presented approach is completely automated and therefore minimizes the "subjective" impact of manual parameter tuning leading to reproducible temperature estimates. In contrast to many other ice core based temperature reconstruction methods, the presented approach is completely independent from ice core stable water isotopes, providing the opportunity to validate water isotope based reconstructions or reconstructions where water isotopes are used together with $\delta^{15}N$ or $\delta^{40}Ar$. We solve the inverse problem $T(\delta^{15}N)$ by using a combination of a Monte Carlo based iterative approach and the analysis of remaining mismatches between modelled and target data, based on cubic spline filtering of random numbers as well as the laboratory determined temperature sensitivity for nitrogen isotopes. Additionally, the presented reconstruction approach was tested by fitting measured $\delta^{40}Ar$ and $\delta^{15}N_{excess}$ data, which leads as well to a robust agreement between modelled and measured data. The obtained final mismatches follow a symmetric standard distribution function. For the synthetic data study, 95 % of the mismatches compared to the synthetic target data are in an envelope between 3.0 permeg to 6.3 permeg for $\delta^{15}N$ and 0.23 K to 0.51 K for temperature (2σ, respectively). In addition to Holocene temperature reconstructions, the fitting approach can also be used for glacial temperature reconstructions. This is shown by high quality fitting of NGRIP $\delta^{15}N$ data for two Dansgaard-Oeschger events using the presented approach, leading to results comparable to other studies.

1 Introduction

Holocene climate variability is of key interest to our society, since it represents a time of moderate natural variations prior to anthropogenic disturbance, often referred to as a baseline for today's increasing greenhouse effect driven by mankind. Yet, high resolution studies are still very sparse and therefore limit the investigation of decadal and partly even centennial climate variations over the course of the Holocene. One of the first studies about changes in the Holocene climate was conducted in the early 1970s by Denton and Karle´n (1973). The authors investigated rapid changes in glacier extents around the globe potentially resulting from variations of Holocene climatic conditions. Mayewski et al. (2004) used this data as the base of a multiproxy study identifying rapid climate changes (so called RCCs) globally distributed over the whole Holocene time period. Although not all proxy data are showing an equal behaviour in timing and extent during the quasi-periodic RCC patterns, the authors found evidence for a highly variable Holocene climate controlled by multiple mechanisms, which significantly affects ecosystems (Pál et al., 2016; Beaulieu et al., 2017; Crausbay et al., 2017) and human societies (Holmgren et al., 2016 ; Lespez, L. et al., 2016). Precise high resolution temperature estimates can contribute significantly to the understanding of these mechanisms. Ice core proxy data offer multiple paths for reconstructing past climate and temperature variability. The studies of Cuffey et al. (1995; 1997) and Dahl-Jensen et al. (1998) demonstrate the usefulness of inverting the measured borehole temperature profile for surface temperature history estimates for the investigated drilling site using a coupled heat- and ice-flow model. Because of smoothing effects due to heat diffusion within an ice sheet, this method is unable to resolve fast temperature oscillations and leads to a rapid reduction of the time resolution towards the past. Another approach to reconstruct past temperature is based on the calibration of stable water isotopes of oxygen and hydrogen ($\delta^{18}O_{ice}$, δD_{ice}) from ice core water samples assuming a constant (and mostly linear) relationship between temperature and water isotopic composition due to fractionation effects during ocean evaporation, cloud formation and snow and ice precipitation (Stuiver et al., 1995; Johnsen et al., 2001). This method provides a rather robust tool for reconstructing past temperature for times where large temperature excursions occur (Dansgaard-Oeschger events, Glacial-Interglacial transitions (Dansgaard et al., 1982; Johnsen et al., 1992)). However, in the Holocene where Greenland temperature variations are comparatively small, seasonal changes of precipitation as well as of evaporation conditions at the source region contribute possibly more to water isotope data variations (Werner et al., 2001; Huber et al., 2006; Kindler et al., 2014;). A relatively new method for ice core based temperature reconstructions uses the thermal fractionation of stable isotopes of air compounds (nitrogen and argon) within a firn layer of an ice sheet (Severinghaus et al., 1998; Severinghaus et al., 2001; Huber et al., 2006; Kobashi et al., 2011; Kindler et al., 2014). The measured nitrogen and argon isotope records of air enclosed in bubbles in an ice core can be used as a paleothermometer due to (i) the stability of isotopic compositions of nitrogen and argon in the atmosphere at orbital timescales and (ii) the fact that changes are only driven by firn processes (Mariotti, 1983; Severinghaus et al., 1998; Leuenberger et al., 1999). To robustly reconstruct the surface temperature for a given drilling site, the use of firn models describing gas and heat diffusion throughout the ice sheet is necessary for decomposing the gravitational from the thermal diffusion influence on the isotope signals.

This work addresses two issues relevant for nitrogen and argon isotope based temperature reconstructions. First, we introduce a novel, entirely automated approach for inverting gas isotope data to surface temperature estimates. For that, we force the output of a firn densification and heat diffusion model to fit gas isotope data. This methodology can be used for many different optimization tasks not restricted to ice core data. As we will show, the approach works besides $\delta^{15}N$ for all relevant gas isotope quantities ($\delta^{15}N$, $\delta^{40}Ar$, $\delta^{15}N_{excess}$) and for Holocene and glacial data as well. Furthermore, the possibility of fitting all relevant gas isotope quantities, individually or combined, makes it possible for the first time to validate the temperature solution gained from one single isotope species by comparison to the solution calculated from other isotope quantities. This approach is a completely new method which enables the automated fitting of gas isotope data without any manual tuning of parameters, minimizing any potential "subjective" impacts on temperature estimates as well as working hours. Also, except for the model spin-up, the presented temperature reconstruction approach is completely independent from stable water isotopes ($\delta^{18}O_{ice}$, δD_{ice}), which provides the opportunity to validate water isotope based reconstructions (e.g. Masson-Delmotte, 2005) or reconstructions where water isotopes are used together with $\delta^{15}N$ or $\delta^{40}Ar$ (e.g. Landais et al., 2004; Huber et al., 2006; Capron et al., 2010). To our knowledge, there are only two other reconstruction methods independent from stable water isotopes that have been applied to Holocene gas isotope data, without a priori assumption on the shape of a temperature change. The studies from Kobashi et al. (2008a, 2017) use the second order parameter $\delta^{15}N_{excess}$ to calculate firn temperature gradients, which are later temporally integrated from past to future over the time series of interest using the firn densification and heat diffusion model from Goujon et al. (2003). Additionally Orsi et al. (2014) use a linearized firn model approach together with $\delta^{15}N$ and $\delta^{40}Ar$ data to extract surface temperature histories. As both methods rely on $\delta^{15}N$ together with $\delta^{40}Ar$, they do not offer the possibility to validate one isotope based solution against the other. Also these two approaches can only be applied to ice cores where both isotope quantities are measured together with a sufficient precision.

Second, we investigate the accuracy of our novel fitting approach by examining the method on different synthetic nitrogen isotope and temperature scenarios. The aim of this work is to study the uncertainties emerging from the algorithm itself. Furthermore the focal question in this study is: what is the minimal mismatch in $\delta^{15}N$ for Holocene like data we can reach and what is the implication for the final temperature mismatches. Studying and moreover answering these questions makes it mandatory to create well defined $\delta^{15}N$ targets and related temperature histories. It is impossible to answer these questions without using synthetic data in a methodology study. The aim is to evaluate the accuracy and associated uncertainty of the inverse method itself to then later apply this method to real $\delta^{15}N$, $\delta^{40}Ar$ or $\delta^{15}N_{excess}$ datasets, for which of course the original driving temperature histories are unknown.

2. Methods and data

2.1 Firn densification and heat diffusion model

The surface temperature reconstruction relies on firn densification combined with gas and heat diffusion (Severinghaus et al., 1998). In this study, the firn densification and heat diffusion model, from now on referred to as firn model, developed by Schwander et al. (1997) is used to reconstruct firn parameters for calculating synthetic $\delta^{15}N$ values depending on the input time series. It is a semi-empirical model based on the work of Herron and Langway (1980), Barnola et al. (1991), and implemented using the Crank and Nicholson algorithm (Crank, 1975) and was also used for the temperature reconstructions by Huber et al. (2006) and Kindler et al. (2014). Besides surface temperature time series, accurate accumulation rate data is needed to run the model. The model then calculates the densification and heat diffusion history of the firn layer and provides parameters for calculating the fractionation of the nitrogen isotopes for each time step, according to the following equations:

$$\delta^{15}N_{grav}\,(z_{LID}, t) = \left(e^{\frac{\Delta m \cdot g \cdot z_{LID}(t)}{R \cdot \overline{T}(t)}} - 1\right) \cdot 1000 \qquad (1)$$

$$\delta^{15}N_{therm}(t) = \left[\left(\frac{T_{surf}(t)}{T_{bottom}(t)}\right)^{\alpha_T} - 1\right] \cdot 1000 \qquad (2)$$

$$\delta^{15}N_{mod}(t) = \delta^{15}N_{grav}(t) + \delta^{15}N_{therm}(t) \qquad (3)$$

$\delta^{15}N_{grav}(t)$ is the component of the isotopic fractionation due to the gravitational settling (Craig et al., 1988; Schwander, 1989) and depends on the lock-in-depth (LID) $z_{LID}(t)$ and the mean firn temperature $\overline{T}(t)$ (Leuenberger et al., 1999). g is the acceleration constant, Δm the molar mass difference between the heavy and light isotopes (equals 10^{-3} kg for nitrogen) and R the ideal gas constant. z_{LID} is defined as a density threshold ρ_{LID}, which is slightly sensitive to surface temperature, following the formula from Martinerie et al. (1994), with a small offset correction of 14 kg m^{-3} to account for the presence of a non-diffusive zone (Schwander et al., 1997):

$$\rho_{LID}(kg \cdot m^{-3}) = \frac{1}{\frac{1}{\rho_{ice}} - 6.95 \cdot 10^{-7} \cdot \overline{T} - 4.3 \cdot 10^{-5}} - 14 \qquad (4)$$

where

$$\rho_{ice}(kg \cdot m^{-3}) = 916.5 - 0.14438 \cdot \overline{T} - 1.5175 \cdot 10^{-4} \cdot \overline{T}^2 \qquad (5)$$

The thermal fractionation component of the $\delta^{15}N$ signal (Severinghaus et al., 1998) is calculated using Eq. (2), where $T_{surf}(t)$ and $T_{bottom}(t)$ stand for the temperatures at the top and the bottom of the diffusive firn layer. In contrast to $T_{surf}(t)$ which is an input parameter for the model, $T_{bottom}(t)$ is calculated by the model for each time step. The thermal diffusion constant α_T was measured by Grachev and Severinghaus (2003) for nitrogen (see Eq. (6)), and closely matches the value used by Leuenberger et al. (1999) based on measurements of Boersma-Klein and De Vries (1966):

$$\alpha_T = \left(8.656 - \frac{1323\ K}{\overline{T}}\right) \cdot 10^{-3} \qquad (6)$$

The firn model used here behaves purely as a forward model, which means that for the given input time series the output parameters (here finally $\delta^{15}N_{mod}(t)$) can be calculated, but it is not easily possible to construct from measured isotope data the related surface temperature or accumulation rate histories. The goal of the presented study is an automatization of this inverse modelling procedure for the reconstruction of the rather small Holocene temperature variations.

2.2 Measurement, input data and time scale

Accumulation rate data: Besides surface temperatures, accumulation rate data is needed to drive the firn model. In this study we use the original accumulation rate, reconstructed in Cuffey and Clow (1997) produced using an ice flow model adapted to the GISP2 location, but adapted to the GICC05 chronology (Rasmussen et al., 2008; Seierstad et al., 2014). Originally, the accumulation rate used to feed the ice flow model was optimised in order to match the time scale from Meese et al. (1994) for the Holocene, based on annual layer counting. Seierstad et al. (2014) transferred the GISP2 chronology to the GICC05 reference timeframe using multiple match points to the NGRIP and GRIP ice cores, both already on GICC05. We used these match points and modified the GISP2 ages in between match points linearly in order to match exactly the GICC05 duration for the considered interval duration. This way, the detailed GISP2 annual layer counting information is kept, but is only stretched/compressed in time. This was done for all intervals in between two match points. The accumulation data were then re-calculated accordingly as obviously this is needed in order to keep the same total amount of ice accumulated at the GISP2 site. From the three accumulation rate scenarios reconstructed in Cuffey and Clow (1997) and adapted here to the GICC05 chronology, the intermediate one is chosen (red curves in Fig. S01). Since the differences between the scenarios (Fig. S01) are not important for the evaluation of the reconstruction approach, they are not taken into account for this study.

$\delta^{18}O_{ice}$ data: Oxygen isotope data from the GISP2 ice core water samples measured at the University of Washington's Quaternary Isotope Laboratory is used to construct the surface temperature input of the model spin-up (12 yr to 35 kyr b2k, see Sect. 2.3.1) (Grootes et al., 1993; Meese et al., 1994; Steig et al., 1994; Stuiver et al., 1995; Grootes and Stuiver, 1997).

Time scale: For the entire study the GICC05 chronology is used (Rasmussen et al., 2014; Seierstad et al., 2014). During the whole reconstruction procedure the two input time series (surface temperature and accumulation rate) are split into two parts. The first part ranges from 20 yr to 10520 yr b2k (called "Holocene section") and the second one from 10520 yr to 35000 yr b2k ("spin-up section"). The entire accumulation rate input (see Sect. 2.3.1), as well as the spin-up section of the surface temperature input, remain unchanged during the reconstruction procedure.

2.3 Reconstruction approach

The Holocene temperature reconstruction is implemented by the following four steps:

(i) A prior temperature input (first guess) is constructed, which serves as the starting point for the optimization.

(ii) A smooth solution which passes through the $\delta^{15}N$ data (here synthetic target data) is generated following a Monte Carlo approach. It is assumed that the smooth solution contains all long term temperature trends (centuries to millennial) as well as firn column height changes (temperature and accumulation rate dependent) that drive the gravitational background signal in $\delta^{15}N$.

(iii) The smooth temperature solution is complemented by superimposing high frequency information directly extracted from the $\delta^{15}N$ data (here synthetic target data). This step adds short term temperature changes (decadal) in the same time resolution as the data.

(iv) The gained temperature solution is corrected using information extracted from the mismatch between the synthetic target and modelled $\delta^{15}N$ time series.

Accumulation rate input:

The raw accumulation rate data for the main part of the spin-up section (12000 yr to 35000 yr b2k) is linearly interpolated to a 20 yr grid and low pass filtered with a 200 yr cut off period (cop) using cubic spline filtering (Enting, 1987). For the Holocene section (20-10520 yr b2k) and the transition part between Holocene and spin-up section (10520 yr to 12000 yr b2k) the raw accumulation rate data is linearly interpolated to a 1 yr grid to obtain equidistant integer point-to-point distances which are necessary for the reconstruction, and to preserve as much information as possible for this time period (Fig. S02a). Except for these technical adjustments, the accumulation rate input data remains unmodified, assuming high reliability of this data during the Holocene. The accumulation data was indeed reconstructed using annual layer counting, and a thinning model which should lead to maximum relative uncertainty of 10 % for the first 1500 m of the 3000 m ice core (Cuffey and Clow, 1997).

In order to investigate the influence of smoothing of the accumulation rate data on the model outputs, the high resolution accumulation rate dataset in the time window of 20 yr to 12000 yr (Fig. S02a) was low pass filtered with cops between 20 yr and 500 yr, and used to drive the firn model. The surface temperature input was set as constant with a value of -31 °C for this time window. Then, the deviations of the filtered from the unfiltered accumulation rates and model outputs were calculated. Figure S03 shows the absolute (I) as well as the relative deviations (II) (relative to the unfiltered scenario) as a function of the cops for the accumulation rate input data, $\delta^{15}N$, and LID model outputs. Regarding the standard deviation (1σ) of the relative accumulation deviations as a measure for the mean deviation of the filtered minus the unfiltered values show that filtering the accumulation rates leads to a mean deviation of about 20 % between the filtered and unfiltered accumulation rate data, depending on the used cop value (see Fig. S03IIa). We use the mean 99 % quantile of the same analysis (Fig. S03IIb) as a measure for the maximum deviation between the filtered and unfiltered values. The filtering clearly leads to a maximum accumulation rate deviation of about 50 %. The comparison of the related deviations in $\delta^{15}N$ and LID outputs reveals that the changes in the accumulation rates do not lead to a change in the same order for the model outputs. Indeed, the filtering of the accumulation rate data leads to deviations of less than 0.6 % and less than 1.5 % for the mean and the maximum $\delta^{15}N$ and LID deviations respectively (Fig. S03IIc,d). Therefore, it can be argued that a low pass filtering of the accumulation rates for cops between 20 yr and 500 yr does only have a small impact on the model outputs as long as the major trends are being conserved, because the filtering does not modify the mean accumulation. This result is expected due to the fact that the LID and finally $\delta^{15}N$ changes are the result of the integration of the accumulation over the whole firn column. The integration time corresponds to the age of the ice at the LID, which is the order of 200 yr for the Holocene in Greenland.

Finally, we test which fraction of the measured $\delta^{15}N$ variations can be attributed to accumulation changes. For this, we perform a sensitivity experiment (Fig. S04) where the temperature input was set as a constant value of -31 °C, and used together with the high resolution accumulation rate data (Fig. S02a) to model the LID (Fig. S04a) and $\delta^{15}N$ (Fig. S04b) values. Due to the absence of temperature changes, only the accumulation rate changes drive the evolution of the diffusive column height (LID) over time which modulates the $\delta^{15}N$ values. Next, the modelled $\delta^{15}N$ variations are compared to the $\delta^{15}N$ measurement data (Fig. S06III) (Kobashi et al., 2008b) to examine the influence of the accumulation rate changes on changes in $\delta^{15}N$ for two cases. First, for the 8.2k event, the signal amplitude in $\delta^{15}N$ is about three times higher for the measured data compared to the modelled ones (measured data: $\Delta\delta^{15}N_{8.2k,meas} \approx 60$ permeg (one permeg equals 10^{-6}); modelled data: $\Delta\delta^{15}N_{8.2k,mod} \approx 20$ permeg). The comparison of the standard deviations of the measured data with the modelled $\delta^{15}N$ data for the last 10 kyr (both quantities were normalized with their respective means), shows an even higher deviation of the measured versus the modelled variabilities by a factor of about eight (measured data: $\mathrm{std}[\delta^{15}N_{10kyr,meas} - \mathrm{mean}(\delta^{15}N_{10kyr,meas})] \approx 37$ permeg; modelled data: $\mathrm{std}[\delta^{15}N_{10kyr,mod} - \mathrm{mean}(\delta^{15}N_{10kyr,mod})] \approx 4.5$ permeg). This analysis supports our assumption that the accumulation rate history alone cannot fully explain the observed variability in $\delta^{15}N$ during the Holocene, and gives an upper limit for the contribution of the accumulation rate to the $\delta^{15}N$ signal. Therefore, the remaining part of the measured $\delta^{15}N$ variations has to be related to changes in surface temperature.

Surface temperature spin-up:

The surface temperature history of the spin-up section (Fig. S02b) is obtained by calibrating the filtered and interpolated $\delta^{18}O_{ice}$ data (Eq. (7)) using the values for the temperature sensitivity α_{18O} and offset β found by Kindler et al. (2014) for the NGRIP ice core assuming a linear relationship of $\delta^{18}O_{ice}$ with temperature.

$$T_{spin}(t) = \frac{1}{\alpha_{18O}(t)} \cdot [\delta^{18}O_{ice}(t) + 35.2\text{ ‰}] - 31.4°C + \beta(t) \quad (7)$$

The values 35.2 ‰ and -31.4 °C are modern-time parameters for the GISP2 site (Schwander et al., 1997; Grootes and Stuiver, 1997). The raw $\delta^{18}O_{ice}$ data is filtered and interpolated in the same way as the accumulation rate data for the spin-up part. The spin-up is needed to bring the firn model to a well-defined starting condition that takes possible memory effects (influence of earlier conditions) of firn states into account.

Generating synthetic target data:

In order to develop and evaluate the presented algorithm, eight temperature scenarios were constructed and used to model synthetic $\delta^{15}N$ data, which serve later on as targets for the reconstruction. From these eight synthetic surface temperature and related $\delta^{15}N$ scenarios (S1-S5 and H1-H3), three data sets (later called Holocene like scenarios H1-H3) were constructed in such a way that the resulting $\delta^{15}N$ time series are very close to the $\delta^{15}N$ values measured by Kobashi et al. (2008) in terms of variability (amplitudes) and frequency (data resolution) of the GISP2 nitrogen isotope data (Fig. S05, Fig. S06).

The synthetic surface temperature scenarios S1-S5 are created by generating a smooth temperature time series ($T_{syn,smooth}$) analogous to the Monte Carlo part of the reconstruction procedure for only one iteration step (see Sect. 2.3.2). The values for the cop used for the filtering of the random values, and the s values (standard deviation of the random values, see Sect. 2.3.2) for the first 5 scenarios can be found in table S01. The smooth temperatures (Fig. S05I) are calculated on a 20 yr grid, which is nearly similar to the time resolution of the GISP2 $\delta^{15}N$ measurement values of about 17 yr (Kobashi et al., 2008b). For the Holocene like scenarios, the smooth temperature time series were generated from the temperature reconstruction for the GISP2 $\delta^{15}N$ data (not shown here). The final Holocene surface temperature solution was filtered with a 100 yr cop to obtain the smooth temperature scenario.

Following this, high frequency information is added to the smoothed temperature histories. A set of normally distributed random numbers with a zero mean and a standard deviation (1σ) of 1 K for scenarios S1-S5 and 0.3 K for Holocene like scenarios H1-H3 is generated on the same 20 yr grid and added up to the smooth temperature time series. Finally, the resulting synthetic target temperature scenarios (Fig. S05II, Fig. S06I) are linearly interpolated to a 1 yr grid.

These synthetic temperatures are combined with the spin-up temperature and are used together with the accumulation rate input to feed the firn model. From the model output the synthetic $\delta^{15}N$ targets are calculated according to section 2.1. The firn model output provides ice age as well as gas age information. The final synthetic $\delta^{15}N$ target time series (Fig. S05III, Fig. S06II) are set intentionally on the ice age scale to mirror measured data, because no prior information is available for the gas-ice age difference (Δage) for ice core data.

2.3.1 Prior input (step 1)

The starting point of the optimization procedure is the first guess. To construct the first guess temperature input, a constant temperature of about -29.6 °C is used for the complete Holocene section, which corresponds to the last value of the temperature spin-up (Fig. S02b).

2.3.2 Monte Carlo type input generator - Generating smooth solutions (step 2)

During the second step of the optimization, the prior temperature input from step 1 is iteratively changed following a Monte Carlo approach. The basic idea of the Monte Carlo approach is to generate smooth temperature inputs by low-pass filtering uniformly distributed random values, and to superimpose this signal on the prior input. Then, the new input is fed to the firn model and the mismatch D_{mc} in between the modelled $\delta^{15}N$ signal X_{mod}, calculated from the model output, and the synthetic $\delta^{15}N$ target values X_{target} is computed.

$$D_{mc} = \frac{1}{n}\sum_{i=1}^{n}|D_i| = \frac{1}{n}\sum_{i=1}^{n}\left|X_{target,i} - X_{mod,i}\right| \quad (8)$$

D_{mc} serves as the criterion which is minimised during the optimization in step 2. If the mismatch decreases compared to the prior input, the new input is saved and used as new guess. This procedure is repeated until convergence is achieved.

Table 01 lists the number of improvements and iterations performed for the different synthetic datasets. The perturbation of the current guess $T_g(t)$ is conducted in the following way: Let $\overrightarrow{T_g} = T_g(t)$ be the vector containing the prior temperature input. A second vector $\overrightarrow{P_1}$ with the same number of elements n as $\overrightarrow{T_g}$ is generated containing n uniformly distributed random numbers within the limits of an also randomly (equally distributed) chosen standard deviation s. s is chosen from a range of 0.05-0.50 (Fig. S07II), which means that the maximum allowed perturbation of a single temperature value $T(t_0)$ is in a range of ±5 % to ±50 %. Creating the synthetic frequencies, $\overrightarrow{P_1}$ is low-pass filtered using cubic spline filtering with an equally distributed random cop (Fig. S07I) in the range of 500 yr to 2000 yr generating the vector $\vec{P}$. The new surface temperature input $\overrightarrow{T_{sm}}$ is calculated from $\vec{P}$ according to:

$$\overrightarrow{T_{sm}} = \overrightarrow{T_g}^{T} \cdot (\hat{1} + \vec{P}) \quad (9)$$

The superscript "T" stands for transposed and $\hat{1}$ is the n by 1 matrix of ones.

This approach provides a high potential for parallel computing. In this study, an eight core computer was used, generating and running eight different inputs of $\overrightarrow{T_{sm}}$ simultaneously, minimizing the time to find an improved solution. For example, during the 706 iterations for scenario S2, about 5600 different inputs were created and tried, leading to 351 improvements (see Tab. 01). Since it is possible to find more than one improvement per iteration step due to the parallelization on eight CPU's, the solution giving the minimal misfit is chosen as new first guess for the next iteration step. This leads to a decrease of the used improvements for the optimization (e.g. for S2, 172 of the 351 improvements were used). Additionally, a first gas age scale is extracted from the model using the last improved conditions, which will then be used in step 3.

2.3.3 Adding high frequency information (step 3)

In step 3 the missing high frequency temperature history providing a suitable fit between modelled and synthetic $\delta^{15}N$ data is directly extracted from the pointwise mismatch $D_{smooth,i}$, between the modelled $\delta^{15}N_{smooth}$ obtained in step 2 and the synthetic $\delta^{15}N$ target data. Note that for a real reconstruction, this mismatch is calculated using the measured $\delta^{15}N$ dataset instead of the synthetic one. $D_{smooth,i}$ can be interpreted in first order as the detrended high frequency signal of the synthetic $\delta^{15}N$ target values (Fig. 01c). This signal is transferred to the gas age scale provided by the firn model for the smooth temperature input to insure synchronicity in between the high frequency temperature variations extracted from the mismatch of $\delta^{15}N$ on the ice age scale and the smooth temperature solution. Additionally, the signal is shifted by about 10 yr towards modern values to account for gas diffusion from the surface to the LID (Schwander et al., 1993), which is not yet implemented in the firn model. This is necessary for adding the calculated temperature changes ΔT to the smooth signal. The ΔT values are calculated according to Eq. (10):

$$\Delta T_i = \frac{D_{smooth,i}}{\Omega_{N_2,i}}, \qquad (10)$$

using the thermal diffusion sensitivity $\Omega_{N_2,i}$ for nitrogen isotope fractionation from Grachev and Severinghaus (2003):

$$\Omega_{N_2,i} = \frac{8.656\ ‰}{\overline{T_i}} - \frac{1232\ ‰\cdot K}{\overline{T_i}^2} \qquad (11)$$

$\overline{T_i}$ is the mean firn temperature in Kelvin which is calculated by the firn model for each time point i. To reconstruct the final (high frequency) temperature input T_{hf}, the extracted short term temperature signal ΔT is simply added to the smooth temperature input T_{sm}:

$$T_{hf,i} = T_{sm,i} + \Delta T_i \qquad (12)$$

2.3.4 Final correction of the surface temperature solution (step 4)

For a further improvement of the remaining $\delta^{15}N$ and resulting surface temperature misfits, it is important to find a correction method that contains information that is also available when using measured data. The benefit of the synthetic data study is that several later unknown quantities can be calculated, and used for improving the reconstruction approach (see Sect. 3 and 4). For instance, it is possible to split the synthetic $\delta^{15}N$ data in the gravitational and thermo-diffusion parts or to use the temperature misfit, which is unknown in reality. The idea underlying the correction algorithm explained hereafter is that the remaining misfits of $\delta^{15}N$ and temperature are connected to the Monte Carlo (step 2) and high frequency part (step 3) of the reconstruction algorithm. In the present inversion framework, it is not possible to find a smooth solution which exactly passes through the $\delta^{15}N$ target data in the middle of the variance in all parts of the time series. This leads to a slightly over or underestimation of the $\delta^{15}N$ and their corresponding temperature values. For example, a slightly too low (or too high) smooth temperature estimate leads to a small increase (or decrease) of the firn column height, creating a wrong gravitational background signal in $\delta^{15}N$ on a later point in time (because the firn column needs some time to react). An additional error in the thermal diffusion signal is also created due to the high frequency part of the reconstruction, because the high frequency information is directly extracted from the deviation of the (synthetic) $\delta^{15}N$ target data and the modelled $\delta^{15}N$ data from the smooth solution of the Monte Carlo part. Therefore, this error is transferred into the next step of the reconstruction and partly creates the remaining deviations.

To investigate this problem, the deviations $D_{smooth,i}$ of the synthetic target data $\delta^{15}N_{target}$ to the smooth data $\delta^{15}N_{smooth}$ of the Monte Carlo part is numerically integrated over a time window of 200 yr (see Sect. 4), and thereafter the window is shifted from past to future in 1 yr steps resulting in a time series called IF(t). IF(t) equals a 200 yr running-mean of $D_{smooth,i}$. For t, the mid position of the window is allocated. The time evolution of IF is a measure for the deviation of the smooth solution in $\delta^{15}N$ (or temperature) from the perfect middle passage through the target data and for the slightly over and underestimation of the resulting temperature.

$$IF(t_i) = \int_{t_1}^{t_2} \left(\delta^{15}N_{target}(t) - \delta^{15}N_{smooth}(t)\right) dt \quad (13)$$

where $t_i = t_1 + \frac{t_2 - t_1}{2}$ (14)

Next, the sample cross correlation function (xcf) (Box et al., 1994) is applied to IF(t) and the remaining misfits $D_{\delta 15N,hf}$ of $\delta^{15}N$ after the high frequency part. The xcf shows two extrema (Fig. 02a), a maximum (xcf_{max}) and a minimum (xcf_{min}) at two certain lags ($lag_{max,D\delta 15N}$ at xcf_{max} and $lag_{min,D\delta 15N}$ at xcf_{min}). Now, the same analysis is conducted for IF(t) versus the temperature mismatch $D_{T,hf}$ (Fig. 02b), which shows an equal behaviour (two extrema, $lag_{max,T}$ at xcf_{max} and $lag_{min,T}$ at xcf_{min}). Comparing the two cross correlations shows that $lag_{max,D\delta 15N}$ equals the negative $lag_{min,T}$ and $lag_{min,D\delta 15N}$ corresponds to the negative $lag_{max,T}$ (Fig. 02d,e). The idea for the correction is that the extrema in $D_{\delta 15N,hf}$ with the positive lag (positive means here that $D_{\delta 15N,hf}$ has to be shifted to past values relative to IF) creates the misfit of temperature $D_{T,hf}$ on the negative lag (modern direction) and vice versa. So IF(t) yields information about the cause and allows us to correct this effect between the remaining mismatches of $\delta^{15}N$ and temperature over the whole time series. The lags are not sharp signals, due to the fact that (i) the cross correlations are conducted over the whole analysed record, leading to an averaging of this cause and effect relationship as well as that (ii) IF(t) is a smoothed quantity itself. The correction of the reconstructed temperature after the high frequency part is conducted in the following way: From the two linear relationships between IF(t) and $D_{\delta 15N,hf}$ at the two lags ($lag_{max,D\delta 15N}$ at xcf_{max}, $lag_{min,D\delta 15N}$ at xcf_{min}) two sets of $\delta^{15}N$ correction values ($\Delta\delta^{15}N_{max}$ from xcf_{max} and $\Delta\delta^{15}N_{min}$ from xcf_{min}) are calculated. Then the lags are being inverted (Fig. 02c,e) shifting the two sets of the $\delta^{15}N$ correction values to the attributed lags of the cross correlation between IF(t) and $D_{T,hf}$ (e.g. $\Delta\delta^{15}N_{min}$ to lag from xcf_{max} from the cross correlation between IF(t) and $D_{T,hf}$) therefore changing the time assignments of $\Delta\delta^{15}N_{min}(t)$ and $\Delta\delta^{15}N_{max}(t)$ to $\Delta\delta^{15}N_{min}(t+lag_{max,T})$ and $\Delta\delta^{15}N_{max}(t+lag_{min,T})$. Now, the $\Delta\delta^{15}N_{max}$ and $\Delta\delta^{15}N_{min}$ are component wise summed up leading to the time series $\Delta\delta^{15}N_{cv}(t)$. From Eq. (10) with $\Delta\delta^{15}N_{cv,i}$ instead of $D_{smooth,i}$ the corresponding temperature correction values are calculated and added to the high frequency temperature solution giving the corrected temperature $T_{corr.}$ Finally, T_{corr} is used to run the firn model to calculate the corrected $\delta^{15}N$ time series (Fig. 03). This cause and effect relationship found in the cross correlations between IF(t) and $D_{\delta 15N,hf}$, and IF(t) and $D_{T,hf}$, is exemplarily shown in Fig. 02 for scenario S1 and was found for all eight synthetic scenarios. The derived correction algorithm leads to a further reduction of the mismatches of about 40 % in $\delta^{15}N$ and temperature (see Sect. 3.2).

3. Results

3.1 Monte Carlo type input generator

Figure S08 shows the evolution of the mean misfit D_{mean} of $\delta^{15}N$ from the synthetic target versus the modelled data as a function of the applied iterations for all synthetic scenarios. One can easily see that all scenarios show a steep decline of the mismatch during the first 50 to 200 iterations followed by a rather moderate decrease, which finally leads to a constant value. During the Monte Carlo part, it was possible to reduce the misfit of $\delta^{15}N$ compared to the first guess solution by about 15 % to 75 % depending on the scenario and the mismatch of the first guess solution (see Tab. 01). This leads to a reduction of the temperature mismatches compared to the first guess temperature mismatch of about 51 % to 87 %.

Figure 01 provides the comparison between the first guess and Monte Carlo solution versus the synthetic target data for the modelled $\delta^{15}N$ (a-c) and surface temperature values (d-f) for scenario S5. Subplots (a) and (d) show the time series of the synthetic target data (black dotted line), the first guess solution (blue line) and the Monte Carlo solution (red line) for $\delta^{15}N$ and temperature. In subplots (b) and (e), the distribution of the pointwise mismatch D_i of the first guess (blue) and the Monte Carlo solution (red) versus the synthetic target data for $\delta^{15}N$ and temperature can be found. Subplots (c) and (f) contain the time series for D_i for $\delta^{15}N$ and temperature. The $D_i(\delta^{15}N)$ data is used later on to calculate the high frequency signal that is superimposed to the smooth temperature solution according to Eq. (10) and Eq. (12) (see Sect. 2.3.3). From Fig. 01 it can be concluded that the Monte Carlo part of the reconstruction algorithm (step 2) leads to two major improvements of the first guess solution. First, it is obvious that the Monte Carlo approach corrects the offsets of the first guess input, which shifts the midpoint of the distribution of $D_{mc,i}$ to zero (see Fig. 01b,e). The second improvement is that the distribution becomes more symmetric and the misfit is overall reduced (the distributions become narrower) compared to the first guess, due to the middle passage through the $\delta^{15}N$ targets. These improvements can be observed for all eight synthetic scenarios, showing the robustness of the Monte Carlo part (see Tab. 01, Fig. 01).

3.2 High frequency step and final correction

Figure 03 provides the comparison between the Monte Carlo, the high frequency and the correction parts of the reconstruction procedure for the scenarios S5. Additional data for all other scenarios can be found in table 02. The upper four plots (a-d) illustrate each reconstruction step and their effect on the modelled $\delta^{15}N$; the bottom four plots (e-h) show the corresponding results on the temperature. Plots (a) and (d) contain the time series of the synthetic $\delta^{15}N$ or temperature target (black dotted line), the high frequency solution (blue line), and the final solution after the correction part (red line). For visibility reasons, subplots (b) and (f) display a zoom-in for a randomly chosen time window of about 500 yr for the same quantities, which shows the excellent agreement in timing and amplitudes of the modelled $\delta^{15}N$ and temperature compared to the synthetic target data. Histograms (c) and (g) and subplots (d) and (h) show the distribution and the time series of the pointwise mismatches (D_i for $\delta^{15}N$, ΔT_i for temperature) between the modelled and the synthetic target data in $\delta^{15}N$ and temperature for each reconstruction step.

Compared to the Monte Carlo solution, the high frequency part leads to a large refinement of the reconstructions. For the mean $\delta^{15}N$ misfits D, the improvement between the Monte Carlo and the high frequency parts is in the range of 64 % to 76 % (see Tab. 02). This leads to a reduction of the temperature mismatches of 43 % to 67 %. The standard deviations (1σ) of the pointwise mismatches (Fig. 03c,d,g,h) in $\delta^{15}N$ and temperature after the high frequency parts are in the range of about 2.7 permeg to 5.4 permeg for $\delta^{15}N$ and 0.22 K to 0.40 K for the reconstructed temperatures depending on the scenario, which is clearly visible in the decreasing width of the histograms (subplots (c) and (g) of Fig. 03, blue against grey).

The mismatches after the correction part of the reconstruction approach show clearly a further decrease of the misfits. This means that the width of the distributions of the pointwise mismatches of $\delta^{15}N$ as well as of temperature is further reduced, and the distributions become more symmetric (long tales disappear, see histogram (c) and (g) of Fig. 03). The time series of the mismatches (subplots (d) and (h) of Fig. 03) clearly illustrate that the correction approach mainly tackles the extreme deviations (sharp reduction of extreme values occurrence in the red distribution compared to the blue distribution) leading to a further improvement of about 40 % in $\delta^{15}N$ and temperature. Finally, the 95 % quantiles of the remaining pointwise mismatches of $\delta^{15}N$ and temperature (D_i or ΔT_i) were calculated for the final solutions for all scenarios and are used as an estimate for the 2σ uncertainty of the reconstruction algorithm (see Fig. 03 and Tab. 2). The final uncertainties (2σ) are in the order of 3.0 permeg to 6.3 permeg for $\delta^{15}N$ and 0.23 K to 0.51 K for the surface temperature misfits. It is noteworthy that the measurement uncertainties (per point) of state of the art $\delta^{15}N$ measurements are in the same order of magnitude, i.e. 3 permeg to 5 permeg (Kobashi et al., 2008b), highlighting the effectiveness of the presented fitting approach. Table 03 contains the final mismatches (2σ) in Δage between the synthetic target and the final modelled data after the correction step for all scenarios and shows that with a known accumulation rate and assumed perfect firn physics, it is possible to fit the Δage history in the Holocene with mean uncertainties better than 2 yr. In other words, the uncertainty in Δage reconstruction due to the inversion algorithm alone is in the order of 2 yr.

4. Discussion

4.1 Monte Carlo type input generator

Figure S07 shows the distribution of the cop (I) and s values (II) used to create the improvements (Sect. 2.3.2) for all scenarios. The cop values are more or less evenly distributed, which shows that nearly the whole of the allowed frequency range (allowed cops were 500 yr to 2000 yr) was used to create the improvements during the iterations. In contrast, the distributions of the s values show clearly that mostly small s values are used to create the improvements, which implies that iterations with small perturbations more likely lead to an improvement than larger ones.

Figure S08 reveals a weak point of the Monte Carlo part, namely the absence of a suitable termination criterion for the optimization. The implementation until now is conducted such that the maximum number of iterations is given by the user or the iterations are terminated after a certain time (e.g. 15 h). Figure S08 shows that for nearly all scenarios it would be possible to stop the optimization after about 400 iterations, due to rather small additional improvements later on. This would decrease the time needed for the Monte Carlo part to about 10 h (a single iteration needs about 90 s). Since the goal of the Monte Carlo part is to find a temperature realisation that leads to an optimal middle passage through the $\delta^{15}N$ target data, it would be possible to use the mean difference between the $\delta^{15}N$ target and spline filtered $\delta^{15}N$ data using a certain cut off period as a termination criterion. This issue is under investigation at the moment. Another possibility to decrease the time needed for the Monte Carlo part could be an increase in the numbers of CPUs used for the parallelization of the model runs. For this study an eight core parallelization was used. A further increase in numbers of workers would improve the speed of the optimization.

4.2 High frequency step and final correction

Several analyses were conducted in order to investigate the remaining mismatches in $\delta^{15}N$ and temperature after the high frequency and the correction part of the reconstruction respectively. First, the total misfit of $\delta^{15}N$ ($D_{\delta 15Ntot}$) was separated into two fractions: gravitational ($D_{\delta 15Ngrav}$) and thermal diffusion mismatches ($D_{\delta 15Ntherm}$) of $\delta^{15}N$ (Fig. 04). Figure 04 indicates that the main fraction of the total mismatch of $\delta^{15}N$ is due to the misfit of the thermal diffusion component of the $\delta^{15}N$ signal, whereas the gravitational misfit of $\delta^{15}N$ has only a minor contribution. The ratio of the standard deviations $\sigma(D_{\delta 15Ntherm})/\sigma(D_{\delta 15Ngrav})$ is about 2.4 for the high frequency solution, and about 2.3 for the corrected signal, showing that the misfit in the thermal diffusion part is more than twice as high as in the gravitational component.

To investigate the timing and contributions of the mismatches in $\delta^{15}N$ and temperature for scenario S1, different xcfs were calculated (Fig. S09a-d). The same analyses were conducted for all synthetic scenarios, leading to similar results. In Fig. S09a the xcf between the mismatch of total $\delta^{15}N$ ($D_{\delta 15Ntot,hf}$) and the misfit of temperature ($D_{T,hf}$) is shown. The cross correlation leads to two extrema (r_{1a}=0.70, r_{2a}=-0.55) on two certain lags (l_{1a}=-2 yr, l_{2a}=+126 yr). In subplot (b) and (c) the same analysis is conducted between the mismatch of the gravitational ($D_{\delta 15Ngrav,hf}$) component (b), and the thermal diffusion ($D_{\delta 15Ntherm,hf}$) component (c) of $\delta^{15}N$ and the temperature mismatch. It is obvious that the xcf of (a) is a combination of (b) and (c). The direct correlation on l_{1a} of (a) can be attributed mainly to the mismatch of the thermal diffusion component of $\delta^{15}N$, whereas the negative correlation on l_{2a} is due to the mismatch of the gravitational component of $\delta^{15}N$. Regarding the xcfs of (a)-(c) at a certain lag l, i.e. l = 0 yr shows that here (and on most of the other lags) the correlations between $D_{\delta 15Ngrav,hf}$ with $D_{T,hf}$ and $D_{\delta 15Ntherm,hf}$ with $D_{T,hf}$ work in opposite directions, which makes it difficult to find a way to correct the remaining temperature mismatch using only information from $D_{\delta 15Ntot,hf}$ for measurement data (when only $D_{\delta 15Ntot,hf}$ is available). The correlation on l_{1a} in (a) is weakened, whereas the lag l_{2a} is shifted to higher values because of the superposition of gravitational and thermal diffusion mismatch. Figure S09d shows also that the gravitational and thermal diffusion mismatches of $\delta^{15}N$ are not independent, but the correlations at the extrema are relatively weak (r_{1d}=0.38, r_{2d}=-0.56). The negative correlation r_{2d} is a sign for the compensation effect between the gravitational and thermal diffusion signals in $\delta^{15}N$ due to the high frequency part of the reconstruction, whereas no explanation could be found for the positive correlation r_{1d}. The symmetric behaviour of the lags for r_{1d} and r_{2d} (l_{1d} = -88 yr ≈ -l_{2d}=93 yr) suggest that r_{1d} could be an artefact of a periodic behaviour of $D_{\delta 15Ngrav,hf}$ and $D_{\delta 15Ntherm,hf}$. Figures S10a-d show the same analysis after the correction part of the reconstruction. It is evident that in all cases the extrema in the different xcfs break down due to the correction of the temperature signal, which is the consequence of the decreasing mismatches of temperature as well as of $\delta^{15}N$. The comparison of the subplots (a), (b) and (c) also shows that the remaining temperature misfits after the correction are mainly driven by the mismatches of the thermal diffusion signal of $\delta^{15}N$ with a minor contribution of the gravitational misfit.

Figures S09e-h show the cross correlations between IF(t) used for the correction of the high frequency temperature solution, and the temperature misfit (e), the mismatch of total $\delta^{15}N$ (f), the mismatch of the gravitational (g) and thermal diffusion (h) component of the $\delta^{15}N$ signal calculated from the high frequency temperature solution. For the correction, the cross correlations (e) and (f) were used (see Sect. 2.3.4 and Fig. 02). Since for measured data neither information about the temperature mismatch (the true temperature is not known) nor about the mismatch of the components of $\delta^{15}N$ (gravitational, thermal diffusion) are available, it is imperative that the symmetric behaviour between the xcf(IF(t), $D_{T,hf}$(t)) and inverted xcf(IF(t), $D_{\delta15Ntot,hf}$(t)) holds true. This criterion is fulfilled for all eight synthetic data scenarios and especially for H1-H3. The comparison of the subplots (f), (g) and (h) of Fig. S09 show the same findings as before, namely that the xcf for IF versus $D_{\delta15Ntot,hf}$ is the combination of the xcfs of IF(t) versus $D_{\delta15Ngrav,hf}$ and IF(t) versus $D_{\delta15Ntherm,hf}$, and that the major fraction of $D_{\delta15Ntot,hf}$ is contributed from $D_{\delta15Ntherm,hf}$. The advantage to use IF(t) for the correction is the symmetry between the two cross correlations, which is created by two factors. The first one is the allocation of the window mid position to the entries of IF, which leads to the symmetric behaviour of the gravitational and thermal diffusion misfits. Second, the shifting of the window in 1 yr steps creating IF(t) over the whole data set leads to an averaged information, but even more importantly, to constant dependency between the temperature and $\delta^{15}N$ mismatches. This can be used later on to fit measured data.

Additionally, the influence of the window length, used for the construction of IF(t), on the correction was analysed. The construction was conducted for different window lengths ranging from 50 yr to 750 yr (Fig. S11). Also, the correction was calculated by using only xcf_{max} or xcf_{min} of IF(t) versus $D_{\delta15N,hf}$ for correcting the temperature input. Figures S11a,b show the remaining mismatches of $\delta^{15}N$ ($D_{\delta15N,corr}$) (a), and temperature ($D_{T,corr}$) (b) after the correction as a function of the used window length for IF(t). The analysis shows that for all investigated window lengths the correction reduces the mismatches of $\delta^{15}N$ and temperature, whatever correction mode was used (calculated with xcf_{max}, xcf_{min}, or both quantities, see comparison with the blue line in (a) and (b)). Furthermore, the correction works best for window lengths in the range of 100 yr to 300 yr with an optimum at 200 yr for all cases. This indicates that the maximum mean duration effect of a $\delta^{15}N$ mismatch creating a temperature mismatch (and vice versa) is in the same range for the investigated scenarios and such small deviations (low permeg level). It is also visible that the correction using both extrema (xcf_{max} and xcf_{min}) leads to a better correction as the approach using only one quantity. This is somehow surprising because the two extrema are the result of the periodicity of IF(t), $D_{\delta15N,hf}$ and $D_{T,hf}$. An explanation for this result could be that a larger section of the temperature time series is corrected when both extrema are used for the correction, due to shifts in both directions. The correction using xcf_{max} only leads to a better fit than the one with xcf_{min}, which can be attributed to the higher correlation between IF(t) and $D_{\delta15N,hf}$. Figures S11e,f show the evolution of the lags corresponding to the two extrema for the cross correlations between

IF(t), and the $\delta^{15}N$ and temperature mismatches, respectively. The linear dependency between the lags and the window length (the lags are nearly half of the window length) is the result of the construction of IF(t), which means the averaging due to the integration in the window of this certain length and the symmetric behaviour due to the allocation of the window mid position to the entries of IF(t).

4.3 Key points to be considered for the application to real data

Benefits of the novel gas isotope fitting approach

In addition to the fitting of $\delta^{15}N$ data, the algorithm is able to fit $\delta^{40}Ar$ and $\delta^{15}N_{excess}$ data as well using the same basic concepts (Fig. S12). Here the $\delta^{40}Ar$ and $\delta^{15}N_{excess}$ data from Kobashi et al. (2008) were used as the fitting targets using the same approach. We reach final mismatches (2σ) of 4.0 permeg for $\delta^{40}Ar/4$ and 3.7 permeg for $\delta^{15}N_{excess}$, which are for both quantities below the analytical measurement uncertainty of 4.0 permeg to 9.0 permeg for $\delta^{40}Ar/4$ and 5.0 permeg to 9.8 permeg for $\delta^{15}N_{excess}$ measured data (Kobashi et al., 2008).

The automated inversion of different gas isotope quantities ($\delta^{15}N$, $\delta^{40}Ar$, $\delta^{15}N_{excess}$) provides a unique opportunity to study the differences in the gained solutions using different targets and to improve our knowledge about the uncertainties of gas isotope based temperature reconstructions using a single firn model. Next, the presented algorithm is not dependent on the firn model, which leads to the implication that the algorithm can be coupled to different firn models describing firn physics in different ways. Furthermore, an automated reconstruction algorithm avoiding manual manipulation and leading to reproducible solutions makes it possible for the first time, to study and learn from the differences in between solutions matching different targets. Finally, differences obtained by applying different firn physics (densification equations, convective zone, etc.) but the very same inversion algorithm may help to assess firn model shortcomings, resulting in more robust uncertainty estimates than it was ever possible before.

In this publication we show the functionality and the basic concepts of the automated inversion algorithm using well known synthetic $\delta^{15}N$ fitting targets. In this “perfect world scenario” the forward problem, converting surface temperature to $\delta^{15}N$, as well as the inverse problem, converting $\delta^{15}N$ to surface temperature, is completely described by the used firn model. Consequently all sources of signal noise are ignored. For the later use of the algorithm on $\delta^{15}N$, $\delta^{40}Ar$ or $\delta^{15}N_{excess}$ measured data this will not be the case anymore due to different sources of signal noise in the used measured data. As a result, differences in between temperature solutions obtained from individual targets ($\delta^{15}N$, $\delta^{40}Ar$, $\delta^{15}N_{excess}$) will become obvious. These differences will allow to quantify the uncertainties associated with different unconstrained processes. Next, we will list and discuss potential sources of uncertainties and try to provide suggestions for their handling and quantification in our approach.

Measurement uncertainty and firn heterogeneity (cm-scale variability):

Many studies have investigated the influence of firn heterogeneity (or density fluctuations) on measurements of air compounds and quantities (e.g. $\delta^{15}N$, $\delta^{40}Ar$, CH_4, CO_2, O_2/N_2 ratio, air content) extracted from ice cores resulting in cm-scale variability and leading to additional noise on the measured data (e.g., Etheridge et al., 1992; Huber and Leuenberger, 2004; Fujita et al., 2009; Capron et al., 2010; Hörhold et al., 2011; Rhodes et al., 2013, 2016; Fourteau et al., 2017). Using discrete measurement technique instead of continuous sampling methods makes it difficult to quantify these effects. However, during discrete analyses of ice core air data it is common to measure replicates for given depths, from which the measurement uncertainties of the gas isotope data is calculated using pooled-standard-deviation (Hedges L. V., 1985). Often it is not possible to take real replicates (same depth) and instead the replicates are taken from nearby depths. Hence, any potential cm-scale variability is to some degree already included in the measurement uncertainty, because each measurement point represents the average over a few centimetres of ice. This is especially the case for low accumulation sites or glacial ice samples for which the vertical length of a sample (e.g., 10-25 cm long for the glacial part of the NGRIP ice core, Kindler et al., 2014) covers the equivalent of 20 yr to 50 yr of ice at approximately 35 kyr b2k. Increasing the depth resolution of the samples would increase our knowledge of cm-scale variability, for e.g. identifying anomalous entrapped gas layers that could have been rapidly isolated from the surface due to an overlying high density layer (e.g., Rosen et al., 2014). As this variability is likely due to heterogeneity in the density profile, modelling such heterogeneities (if possible at all) may not help to better reconstruct a meaningful temperature history, but rather to reproduce the source of noise. This means that the potential cm-scale variability, in many cases, is already incorporated in the analytical noise obtained from gas isotope measurements, due to analytical techniques themselves. Assuming the measurement uncertainty as Gaussian distributed, it is easy to incorporate this source of uncertainty in the inverse modelling approach presented here. This will increase the uncertainty of the temperature according to Eq. (10).The same equation can also be used for the calculation of the uncertainty in temperature related to measurement uncertainty in general.

To answer the pertinent question of how to better extract a meaningful temperature history from a noisy ice core record, an excellent – but costly – solution is of course to use multiple ice cores. For example, a $\delta^{15}N$-based temperature reconstruction from the combination of data from the GISP2 ice core with the “sister ice core” GRIP drilled only a few kilometres apart is likely one of the best ways to overcome potential cm-scale variability. A comparison of ice cores that were drilled even closer might be even more advantageous.

Smoothing effects due to gas diffusion and trapping:

It is known that gas diffusion and trapping processes in the firn can smooth out fast signals and result in a damping of the amplitudes of gas isotope signals (e.g. Spahni et al., 2003; Grachev and Severinghaus, 2005). The duration of gas diffusion from the top of the diffusive column to the bottom where the air is closed off in bubbles is for Holocene conditions in Greenland approximately in the order of 10 yr (Schwander et al. 1997), whereas the data resolution of the synthetic targets was set to 20 yr to mimic the measurement data from Kobashi et al. (2008) with a mean data resolution of about 17 yr (see Sect. 2.3: “Generating synthetic target data”). In the study of Kindler et al. (2014) it was shown that a glacial Greenland LID leads to a damping of the $\delta^{15}N$ signal of about 30 % for a 10 K temperature rise in 20 yr. We further assume that the smoothing according to the lock-in process is negligible for Greenland Holocene conditions according to the much smaller amplitude signals and shallower LID. Yet, for glacial conditions it requires attention.

Accumulation rate uncertainties:

For the synthetic data study presented in this paper it is assumed that the used accumulation rate data is well known with zero uncertainty. This simplification is used to show the functionality and basic concepts of the presented fitting algorithm in every detail on well-known $\delta^{15}N$ and temperature targets and to focus on the final uncertainties originating from the presented fitting algorithm itself. For the later reconstruction using measured gas isotope data together with the published accumulation rate scenarios shown in Fig. S01 this will not be the case anymore. Uncertainties in layer counting and corrections for ice thinning lead to a fundamental uncertainty. Especially in the early Holocene, this can easily exceed 10 %. As the accumulation rate data is used to run the firn model, all potential accumulation uncertainties are in part incorporated into the temperature reconstruction. On the other hand, as we discussed in section 2.3, the accumulation rate variability has a minor impact compared to the input temperature on the variability of $\delta^{15}N$ data in the Holocene (see also Fig. S03, Fig. S04). The influence of these quantities, accumulation rate or temperature, on the temperature reconstruction is not equal; during the Holocene, accumulation rate variability explains about 12 % to 30 % of $\delta^{15}N$ variability. 30 % corresponds to the 8.2 kyr event and 12 % for the mean of the whole Holocene period including the 8.2 kyr event. Hence the influence of accumulation changes, excluding the extreme 8.2 kyr event, is generally below 10 % during the Holocene. If the accumulation is assumed to be completely correct then the missing part will be assigned to temperature variations. Nevertheless for the fitting of the Holocene measurement data we will use all three accumulation rate scenarios as shown in Fig. S01. The difference in the reconstructed temperatures arising from the differences of these three scenarios will be used for the uncertainty calculation as well and is most likely higher than the uncertainty arising from uncertainties due to the process of producing the accumulation rate data and from the conversion of the accumulation rate data to the GICC05 timescale.

Convective zone variability:

Many studies have shown the existence of a non-diffusive zone at the top of the diffusive firn column, called convective zone (CZ). The CZ is formed by strong katabatic winds and pressure gradients between the surface and the firn (e.g. Kawamura et al., 2006, 2013; Severinghaus et al., 2010). The existence of a CZ changes the gravitational background signal in $\delta^{15}N$ and $\delta^{40}Ar$ as it reduces the diffusive column height. The presented fitting algorithm was used together with the two most frequently used firn models for temperature reconstructions based on stable isotopes of air, the Schwander et al. (1997) model which has no CZ build in (or better a constant CZ of 0 m) and the Goujon firn model (Goujon et al., 2003) (which assumes a constant convective zone over time, that can easily be set in the code). This difference between the two firn models only changes significantly the absolute temperature rather than the temperature anomalies as it was shown by other studies (e.g., Guillevic et al., 2013, Fig. 3). In the presented work, we show the results using the model from Schwander et al. (1997), because the differences between the obtained solutions using the two models are negligible besides a constant temperature offset. Also, noteworthy is that there is no firn model at the moment which uses a dynamically changing CZ. Indeed, this should be investigated but requires additional intense work. Additionally, the knowledge of the time evolution of CZ changes for time periods of millennia to several hundreds of millennia (in frequency and magnitude) is too poor to estimate the influence of this quantity on the reconstruction. In principle it is possible to cancel out the influence of a potentially changing CZ by using $\delta^{15}N_{excess}$ data for temperature reconstruction, as due to the subtraction of $\delta^{40}Ar/4$ from $\delta^{15}N$ the gravitational term of the signals is eliminated. From that point of view it will be interesting to compare temperature solutions gained from $\delta^{15}N_{excess}$ fitting with the solutions based on $\delta^{15}N$ or $\delta^{40}Ar$ alone. This can offer a useful tool for quantifying the magnitude and frequency of CZ changes in the time interval of interest.

It is known that for some very low accumulation rate sites in areas with strong katabatic winds (e.g. "Megadunes", Antarctica) extremely deep CZs can occur, which are potentially able to smooth out even decadal-scale temperature variations (Severinghaus et al., 2010). For this its deepness would need to be of several dozens of meters, which is highly unrealistic even for glacial Summit conditions (Guillevic et al., 2013, see discussion in Annex A4, p. 1042) as well as for the rather stable Holocene period in Greenland for which no low accumulation and strong katabatic wind situations are to be expected.

4.4 Proof of concept for glacial data

For glacial conditions the task of reconstructing temperature (with correct frequency and magnitude) without using $\delta^{18}O_{ice}$ information is much more challenging due to the highly variable gas age - ice age differences (Δage) between stadial and interstadial conditions. Here, contrary to the rather stable Holocene period, the Δage can vary by several hundreds of years. Also the accumulation rate data is more uncertain than for the Holocene. To prove that the presented fitting algorithm also works for glacial conditions we inverted the $\delta^{15}N$ data measured for the NGRIP ice core by Kindler et al. (2014) for two Dansgaard-Oeschger events, namely DO6 and DO7. Since the magnitudes of those events are higher and the signals are smoother than in the Holocene we only had to use the Monte Carlo type input generator (see Sect. 2.3.2) for changing the temperature inputs. To compare our results to the $\delta^{18}O_{ice}$ based and manually calibrated values from Kindler et al. (2014) we use the ss09sea06bm time scale (NGRIP members: Andersen et al., 2004; Johnsen et al., 2001) as it was done in the Kindler et al. publication. For the model spin-up we use the accumulation rate and temperature data from Kindler et al. (2014) for the time span 36.2 kyr to 60 kyr. The reconstruction window (containing DO6 and DO7) is set to 32 kyr to 36.2 kyr. As the first guess (starting point) of the reconstruction we use the accumulation rate data for NGRIP from the ss09sea06bm time scale together with a constant temperature of about -49 °C for this time window. As minimization criterion D for the reconstruction we simply use the sum of the mean squared errors of the $\delta^{15}N$ and Δage mismatches weighted with their uncertainties (wRMSE) according to the following equation instead of the mean $\delta^{15}N$ misfit alone as used for the Holocene (Eq. (8)).

$$D = \sqrt{\mathrm{wRMSE}(\delta^{15}\mathrm{N})} + \sqrt{\mathrm{wRMSE}(\Delta\mathrm{age})} \tag{15}$$

$$= \sqrt{\frac{1}{N}\sum_i \left[\frac{\delta^{15}\mathrm{N}_{\mathrm{meas,i}} - \delta^{15}\mathrm{N}_{\mathrm{mod,i}}}{\varepsilon_{\delta^{15}\mathrm{N,i}}}\right]^2} + \sqrt{\frac{1}{M}\sum_j \left[\frac{\Delta\mathrm{age}_{\mathrm{meas,j}} - \Delta\mathrm{age}_{\mathrm{mod,j}}}{\varepsilon_{\Delta\mathrm{age,j}}}\right]^2}$$

Here $\varepsilon_{\delta^{15}N,i}$ and $\varepsilon_{\Delta age,j}$ are the uncertainties in δ^{15}N and Δage for the measured values i or j (Δage match points: Guillevic, M. (2013), p.65, Tab. 3.2) and N, M the number of measurement values. We set $\varepsilon_{\delta^{15}N,i}$ = 20 permeg for all i (Kindler et al., 2014) and $\varepsilon_{\Delta age,j}$ = 50 yr for all j. The relative uncertainties in Δage can easily reach up to 50 % and more in the Glacial using the ss09sea06bm time scale which results in a pre-eminence of the δ^{15}N misfits over the Δage misfits (10 % to 20 % when using GICC05 time scale, Guillevic (2013), p. 65 Tab. 3.2). Due to this issue we have to set Δage uncertainties to 50 yr to make both terms equally important for the fitting algorithm. In Fig. S13 we show preliminary results. The δ^{15}N and Δage fitting (a, b) and the resulting gained temperature and accumulation rate solutions (c, d) using the presented algorithm are completely independent from $\delta^{18}O_{ice}$ which provides the opportunity to evaluate the $\delta^{18}O_{ice}$ based reconstructions. In this study the algorithm was used in three steps. First, starting with the first guess (constant temperature), the temperature was changed as explained before. The accumulation rate was changed in parallel to the temperature allowing a random offset shift (up and down) together with a stretching or compressing (in y direction) of the accumulation rate signal over the whole time window (32 kyr to 36.2 kyr). This first step leads to the “Monte Carlo Solution 0” (MCS0) which provides a first approximation and is the base for the next step. For the next step, we fixed the accumulation rate and let the algorithm only change the temperature to improve the δ^{15}N fit (MSC1). Finally, we allow the algorithm to change the temperature together with the accumulation rate using the Monte Carlo type input generator for both quantities. This allows to change the shape of the accumulation rate data. This final step can be seen as a fine tuning of the gained solutions from the steps before. The obtained mismatches in δ^{15}N and Δage of all steps are at least of the same quality or better than the $\delta^{18}O_{ice}$ based manual method from Kindler et al. (2014) (see Tab. S02). The gained temperature solutions show a very good agreement in timing and magnitude compared to the reconstruction of Kindler et al. (2014). Also the accumulation rate solutions show that the accumulation has to be reduced significantly compared to the ss09sea06bm data to allow a high quality fit of the δ^{15}N and Δage target data, a result highly similar to Guillevic et al. (2013) and Kindler et al. (2014). The mismatches in δ^{15}N and Δage of the final MCS FIN solution show a 15 % smaller misfit in δ^{15}N (2σ) and an about 31 % smaller misfit for Δage (2σ) compared to the Kindler et al. (2014) solution. Keeping in mind that the used approach is completely independent from $\delta^{18}O_{ice}$ strengthens the functionality and quality of the presented gas isotope fitting approach also for glacial reconstructions. As this section contains a proof of concept of the presented automated gas isotope fitting algorithm on glacial data, preliminary results and ongoing work were shown here. Furthermore as the presented fitting algorithm was developed and tested in first order for Holocene like data, it is highly probable that the functionality of the algorithm using glacial data will be further extended and adjusted in future studies.

5. Conclusion

A novel approach is introduced and described for inverting a firn densification and heat diffusion model to fit small gas isotope data variations as observed throughout the Holocene. From this new fitting method, it is possible to extract the surface temperature history that drives the firn status which in turn leads to the gas isotope time series. The approach is a combination of a Monte Carlo based iterative method and the analysis of remaining mismatches between modelled and target data. The procedure works fully automated and provides a high potential for parallel computing for time consumption optimization. Additional sensitivity experiments have shown that accumulation rate changes have only a minor influence on short term variations of $\delta^{15}N$, which themselves are mainly driven by high frequency temperature variations. To evaluate the performances of the presented approach, eight different synthetic $\delta^{15}N$ time series were created from eight known temperature histories. The fitting approach leads to an excellent agreement in timing and amplitudes between the modelled and synthetic $\delta^{15}N$ and temperature data. The obtained, final mismatches follow a symmetric, standard distribution function. 95 % of the mismatches compared to the synthetic data are in an envelope in between 3.0 permeg to 6.3 permeg for $\delta^{15}N$ and 0.23 K to 0.51 K for temperature, depending on the synthetic temperature history scenarios. These values can therefore be used as a 2σ estimate for the reconstruction uncertainty arising from the presented fitting algorithm itself. For $\delta^{15}N$ the obtained final uncertainties are in the same order of magnitude as state of the art experimental measurement uncertainty. The presented reconstruction approach was also successfully applied to $\delta^{40}Ar$ and $\delta^{15}N_{excess}$ measured data. Moreover, we have shown that the presented fitting approach can also be applied to glacial temperature reconstructions with minor algorithm modifications. Based on the demonstrated flexibility of our inversion methodology, it is reasonable to adapt this approach for reconstructions of other non-linear physical processes.

Competing interests

The authors declare that they have no competing financial interests.

Acknowledgements

This work was supported by the SNF grants (SNF-159563 and SNF-172550). We would like to thank Drs. Takuro Kobashi, Philippe Kindler and Myriam Guillevic for helpful discussions about the ice core data and model peculiarities.

References

NGRIP members: Andersen, K. K., Azuma, N., Barnola, J.-M., Bigler, M., Biscaye, P., Caillon, N., Chappellaz, J., Clausen, H. B., Dahl-Jensen, D., Fischer, H., Flückiger, J., Fritzsche, D., Fujii, Y., Goto-Azuma, K., Grønvold, K., Gundestrup, N. S., Hansson, M., Huber, C., Hvidberg, C. S., Johnsen, S. J., Jonsell, U., Jouzel, J., Kipfstuhl, S., Landais, A., Leuenberger, M., Lorrain, R., Masson-Delmotte, V., Miller, H., Motoyama, H., Narita, H., Popp, T., Rasmussen, S. O., Raynaud, D., Rothlisberger, R., Ruth, U., Samyn, D., Schwander, J., Shoji, H., Siggard-Andersen, M.-L., Steffensen, J. P., Stocker, T., Sveinbjörnsdóttir, A. E., Svensson, A., Takata, M., Tison, J.-L., Thorsteinsson, T., Watanabe, O., Wilhelms, F. and White, J. W. C.: High-resolution record of Northern Hemisphere climate extending into the last interglacial period, Nature, 431(7005), 147–151, doi:10.1038/nature02805, 2004.

Barnola, J.-M., Pimienta, P., Raynaud, D. and Korotkevich, Y. S.: CO2-climate relationship as deduced from the Vostok ice core: a re-examination based on new measurements and on a re-evaluation of the air dating, Tellus B, 43(2), 83–90, doi:10.1034/j.1600-0889.1991.t01-1-00002.x, 1991.

Beaulieu, J. L. de, Brugiapaglia, E., Joannin, S., Guiter, F., Zanchetta, G., Wulf, S., Peyron, O., Bernardo, L., Didier, J., Stock, A., Rius, D. and Magny, M.: Lateglacial-Holocene abrupt vegetation changes at Lago Trifoglietti in Calabria, Southern Italy: The setting of ecosystems in a refugial zone, Quat. Sci. Rev., 158, 44–57, doi:10.1016/j.quascirev.2016.12.013, 2017.

Boersma-Klein, V. and De Vries, A. E.: The influence of the distribution of atomic masses within the molecule on thermal diffusion, Physica, 32, 717–733, doi:10.1016/0031-8914(66)90004-8, 1966.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C.: Time Series Analysis - Forecasting and Control., 1994.

Capron, E., Landais, A., Chappellaz, J., Schilt, A., Buiron, D., Dahl-Jensen, D., Johnsen, S. J., Jouzel, J., Lemieux-Dudon, B., Loulergue, L., Leuenberger, M., Masson-Delmotte, V., Mayer, H., Oerter, H. and Stenni, B.: Millennial and sub-millennial scale climatic variations recorded in polar ice cores over the last glacial period, Clim. Past Discuss., 6(1), 135–183, doi:10.5194/cpd-6-135-2010, 2010.

Craig, H., Horibe, Y. and Sowers, T.: Gravitational separation of gases and isotopes in polar ice caps., Science (80-.)., 242(4886), 1675–8, doi:10.1126/science.242.4886.1675, 1988.

Crank, J.: The Mathematics of Diffusion, 414 pp., Clarendon. [online] Available from: https://scholar.google.ch/scholar?q=Crank%2C+J.%2C+The+Mathematics+of+Diffusion%2C+414pp.%2C+Clarendon%2C+Oxford%2C+England%2C+%281975%29.&btnG=&hl=de&as_sdt=0%2C5 (Accessed 21 March 2017), 1975.

Crausbay, S. D., Higuera, P. E., Sprugel, D. G. and Brubaker, L. B.: Fire catalyzed rapid ecological change in lowland coniferous forests of the Pacific Northwest over the past 14,000 years, Ecology, doi:10.1002/ecy.1897, 2017.

Cuffey, K. M. and Clow, G. D.: Temperature, accumulation, and ice sheet elevation in central Greenland through the last deglacial transition, J. Geophys. Res. Ocean., 102(C12), 26383–26396, doi:10.1029/96JC03981, 1997.

Cuffey, K. M., Clow, G. D., Alley, R. B., Stuiver, M., Waddington, E. D. and Saltus, R. W.: Large Arctic Temperature Change at the Wisconsin-Holocene Glacial Transition, Science, 270(5235), 455–458, doi:10.1126/science.270.5235.455, 1995.

Dahl-Jensen, D., Mosegaard, K., Gundestrup, N., Clow, G. D., Johnsen, S. J., Hansen, A. W. and Balling, N.: Past Temperatures Directly from the Greenland Ice Sheet, Science (80-.)., 282(5387), 268–271, doi:10.1126/science.282.5387.268, 1998.

Dansgaard, W., Clausen, H. B., Gundestrup, N., Hammer, C. U., Johnsen, S. F., Kristinsdottir, P. M. and Reeh, N.: A new greenland deep ice core., Science, 218(4579), 1273–7, doi:10.1126/science.218.4579.1273, 1982.

Denton, G. H. and Karle´n, W.: Holocene climatic variations: their pattern and possible cause., Quat. Res., 3(2), 155–205, doi:10.1016/0033-5894(73)90040-9, 1973.

Enting, I. G.: On the Use of Smoothing Splines to Filter CO 2 Data, J. Geophys. Res., 92(D9), 10977–10984, doi:10.1029/JD092iD09p10977, 1987.

Etheridge, D. M., Pearman, G. I. and Fraser, P. J.: Changes in tropospheric methane between 1841 and 1978 from a high accumulation-rate Antarctic ice core, Tellus B, 44(4), 282–294, doi:10.1034/j.1600-0889.1992.t01-3-00006.x, 1992.

Fourteau, K., Faïn, X., Martinerie, P., Landais, A., Ekaykin, A. A., Lipenkov, V. Y. and Chappellaz, J.: Analytical constraints on layered gas trapping and smoothing of atmospheric variability in ice under low accumulation conditions, Clim. Past Discuss., 1–28, doi:10.5194/cp-2017-78, 2017.

Fujita, S., Okuyama, J., Hori, A. and Hondoh, T.: Metamorphism of stratified firn at Dome Fuji, Antarctica: A mechanism for local insolation modulation of gas transport conditions during bubble close off, J. Geophys. Res. Solid Earth, 114(3), doi:10.1029/2008JF001143, 2009.

Goujon, C., Barnola, J.-M. and Ritz, C.: Modeling the densification of polar firn including heat diffusion: Application to close-off characteristics and gas isotopic fractionation for Antarctica and Greenland sites, J. Geophys. Res. Atmos., 108(D24), n/a-n/a, doi:10.1029/2002JD003319, 2003.

Grachev, A. M. and Severinghaus, J. P.: Laboratory determination of thermal diffusion constants for 29N2/28N2 in air at temperatures from -60 to 0°C for reconstruction of magnitudes of abrupt climate changes using the ice core fossil-air plaeothermometer, Geochim. Cosmochim. Acta, 67(3), 345–360, doi:10.1016/S0016-7037(02)01115-8, 2003.

Grachev, A. M. and Severinghaus, J. P.: A revised + 10±4°C magnitude of the abrupt change in Greenland temperature at the Younger Dryas termination using published GISP2 gas isotope data and air thermal diffusion constants, Quat. Sci. Rev., 24(5–6), 513–519, doi:10.1016/j.quascirev.2004.10.016, 2005.

Grootes, P. M. and Stuiver, M.: Oxygen 18/16 variability in Greenland snow and ice with 10-3- to 105-year time resolution, J. Geophys. Res. Ocean., 102(C12), 26455–26470, doi:10.1029/97JC00880, 1997.

Grootes, P. M., Stuiver, M., White, J. W. C., Johnsen, S. and Jouzel, J.: Comparison of oxygen isotope records from the GISP2 and GRIP Greenland ice cores, Nature, 366(6455), 552–554, doi:10.1038/366552a0, 1993.

Guillevic, M.: PhD thesis., 2013.

Guillevic, M., Bazin, L., Landais, A., Kindler, P., Orsi, A., Masson-Delmotte, V., Blunier, T., Buchardt, S. L., Capron, E., Leuenberger, M., Martinerie, P., Prié, F. and Vinther, B. M.: Spatial gradients of temperature, accumulation and δ18O- ice in Greenland over a series of Dansgaard-Oeschger events, Clim. Past, 9(3), 1029–1051, doi:10.5194/cp-9-1029-2013, 2013.

Hedges L. V., O. I.: Statistical methods for meta-analysis, Academic Press, San Diego., 1985.

Herron, M. M. and Langway, C. C.: Firn densification: an empirical model., J. Glaciol., 25(93), 373–385, doi:10.1017/S0022143000015239, 1980.

Holmgren, K., Gogou, A., Izdebski, A., Luterbacher, J., Sicre, M. A. and Xoplaki, E.: Mediterranean Holocene climate, environment and human societies, Quat. Sci. Rev., 136, 1–4, doi:10.1016/j.quascirev.2015.12.014, 2016.

Hörhold, M. W., Kipfstuhl, S., Wilhelms, F., Freitag, J. and Frenzel, A.: The densification of layered polar firn, J. Geophys. Res. Earth Surf., 116(1), doi:10.1029/2009JF001630, 2011.

Huber, C. and Leuenberger, M.: Measurements of isotope and elemental ratios of air from polar ice with a new on-line extraction method, Geochemistry, Geophys. Geosystems, 5(10), n/a-n/a, doi:10.1029/2004GC000766, 2004.

Huber, C., Leuenberger, M., Spahni, R., Flückiger, J., Schwander, J., Stocker, T. F., Johnsen, S., Landais, A. and Jouzel, J.: Isotope calibrated Greenland temperature record over Marine Isotope Stage 3 and its relation to CH4, Earth Planet. Sci. Lett., 243(3–4), 504–519, doi:10.1016/j.epsl.2006.01.002, 2006.

Johnsen, S. J., Clausen, H. B., Dansgaard, W., Fuhrer, K., Gundestrup, N., Hammer, C. U., Iversen, P., Jouzel, J., Stauffer, B. and Steffensen, J. P.: Irregular glacial interstadials recorded in a new Greenland ice core, Nature, 359(6393), 311–313, doi:10.1038/359311a0, 1992.

Johnsen, S. J., Dahl-Jensen, D., Gundestrup, N., Steffensen, J. P., Clausen, H. B., Miller, H., Masson-Delmotte, V., Sveinbjörnsdottir, A. E. and White, J.: Oxygen isotope and palaeotemperature records from six Greenland ice-core stations: Camp Century, Dye-3, GRIP, GISP2, Renland and NorthGRIP, J. Quat. Sci., 16(4), 299–307, doi:10.1002/jqs.622, 2001.

Kawamura, K., Severinghaus, J. P., Ishidoya, S., Sugawara, S., Hashida, G., Motoyama, H., Fujii, Y., Aoki, S. and Nakazawa, T.: Convective mixing of air in firn at four polar sites, Earth Planet. Sci. Lett., 244(3–4), 672–682, doi:10.1016/j.epsl.2006.02.017, 2006.

Kawamura, K., Severinghaus, J. P., Albert, M. R., Courville, Z. R., Fahnestock, M. A., Scambos, T., Shields, E. and Shuman, C. A.: Kinetic fractionation of gases by deep air convection in polar firn, Atmos. Chem. Phys., 13(21), 11141–11155, doi:10.5194/acp-13-11141-2013, 2013.

Kindler, P., Guillevic, M., Baumgartner, M., Schwander, J., Landais, A. and Leuenberger, M.: Temperature reconstruction from 10 to 120 kyr b2k from the NGRIP ice core, Clim. Past, 10(2), 887–902, doi:10.5194/cp-10-887-2014, 2014.

Kobashi, T., Severinghaus, J. P. and Barnola, J. M.: 4 ± 1.5 °C abrupt warming 11,270 yr ago identified from trapped air in Greenland ice, Earth Planet. Sci. Lett., 268(3–4), 397–407, doi:10.1016/j.epsl.2008.01.032, 2008a.

Kobashi, T., Severinghaus, J. P. and Kawamura, K.: Argon and nitrogen isotopes of trapped air in the GISP2 ice core during the Holocene epoch (0-11,500 B.P.): Methodology and implications for gas loss processes, Geochim. Cosmochim. Acta, 72(19), 4675–4686, doi:10.1016/j.gca.2008.07.006, 2008b.

Kobashi, T., Kawamura, K., Severinghaus, J. P., Barnola, J. M., Nakaegawa, T., Vinther, B. M., Johnsen, S. J. and Box, J. E.: High variability of Greenland surface temperature over the past 4000 years estimated from trapped air in an ice core, Geophys. Res. Lett., 38(21), n/a-n/a, doi:10.1029/2011GL049444, 2011.

Kobashi, T., Menviel, L., Jeltsch-Thömmes, A., Vinther, B. M., Box, J. E., Muscheler, R., Nakaegawa, T., Pfister, P. L., Döring, M., Leuenberger, M., Wanner, H. and Ohmura, A.: Volcanic influence on centennial to millennial Holocene Greenland temperature change, Sci. Rep., 7(1), 1441, doi:10.1038/s41598-017-01451-7, 2017.

Landais, A., Barnola, J. M., Masson-Delmotte, V., Jouzel, J., Chappellaz, J., Caillon, N., Huber, C., Leuenberger, M. and Johnsen, S. J.: A continuous record of temperature evolution over a sequence of Dansgaard-Oeschger events during Marine Isotopic Stage 4 (76 to 62 kyr BP), Geophys. Res. Lett., 31(22), 1–4, doi:10.1029/2004GL021193, 2004.

Lespez, L. et al.: Rapid climatic change and social transformations Uncertainties, adaptability and resilience, in The Mediterranean Region under Climate Change A Scientific Update, p. 738, Marseille., 2016.

Leuenberger, M. C., Lang, C. and Schwander, J.: d15N measurements as a calibration tool for the paleothermometer and gas-ice age differences: A case study for the 8200BP event on GRIP ice, J. Geophys. Res., 104(D18), 22163–22170, doi:Doi 10.1029/1999jd900436, 1999.

Mariotti, A.: Atmospheric nitrogen is a reliable standard for natural 15N abundance measurements, Nature, 303(5919), 685–687, doi:10.1038/303685a0, 1983.

Martinerie, P., Lipenkov, V. Y., Raynaud, D., Chappellaz, J., Barkov, N. I. and Lorius, C.: Air content paleo record in the Vostok ice core (Antarctica): A mixed record of climatic and glaciological parameters, J. Geophys. Res., 99(D5), 10565, doi:10.1029/93JD03223, 1994.

Masson-Delmotte, V.: GRIP Deuterium Excess Reveals Rapid and Orbital-Scale Changes in Greenland Moisture Origin, Science (80-.)., 309(5731), 118–121, doi:10.1126/science.1108575, 2005.

Mayewski, P., Rohling, E., Curt Stager, J., Karlén, W., Maasch, K. A., David Meeker, L., Meyerson, E. A., Gasse, F., van Kreveld, S., Holmgren, K., Lee-Thorp, J., Rosqvist, G., Rack, F., Staubwasser, M., Schneider, R. and Steig, E. J.: Holocene climate variability, Quat. Res., 62(3), 243–255, doi:10.1016/j.yqres.2004.07.001, 2004.

Meese, D. A., Alley, R. B., Flacco, R. J., Germani, M. S., Gow, A. J., Grootes, P. M., Illing, M., Mayewski, P. A., Morrison, M. C., Ram, M., Taylor, K., Yang, Q. and Zielinski, G. A.: Preliminary depth-age scale of the GISP2 ice core. 1994.

Orsi, A. J., Cornuelle, B. D. and Severinghaus, J. P.: Magnitude and temporal evolution of Dansgaard-Oeschger event 8 abrupt temperature change inferred from nitrogen and argon isotopes in GISP2 ice using a new least-squares inversion, Earth Planet. Sci. Lett., 395, 81–90, doi:10.1016/j.epsl.2014.03.030, 2014.

Pál, I., Buczkó, K., Vincze, I., Finsinger, W., Braun, M., Biró, T. and Magyari, E. K.: Terrestrial and aquatic ecosystem responses to early Holocene rapid climate change (RCC) events in the South Carpathian Mountains, Romania, Quat. Int., doi:10.1016/j.quaint.2016.11.015, 2016.

Rasmussen, S. O., Seierstad, I. K., Andersen, K. K., Bigler, M., Dahl-Jensen, D. and Johnsen, S. J.: Synchronization of the NGRIP, GRIP, and GISP2 ice cores across MIS 2 and palaeoclimatic implications, Quat. Sci. Rev., 27(1–2), 18–28, doi:10.1016/j.quascirev.2007.01.016, 2008.

Rasmussen, S. O., Bigler, M., Blockley, S. P., Blunier, T., Buchardt, S. L., Clausen, H. B., Cvijanovic, I., Dahl-Jensen, D., Johnsen, S. J., Fischer, H., Gkinis, V., Guillevic, M., Hoek, W. Z., Lowe, J. J., Pedro, J. B., Popp, T., Seierstad, I. K., Steffensen, J. P., Svensson, A. M., Vallelonga, P., Vinther, B. M., Walker, M. J. C., Wheatley, J. J. and Winstrup, M.: A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three s ynchronized Greenland ice-core records: Refining and extending the INTIMATE event stratigraphy, Quat. Sci. Rev., 106, 14–28, doi:10.1016/j.quascirev.2014.09.007, 2014.

Rhodes, R. H., Faïn, X., Stowasser, C., Blunier, T., Chappellaz, J., McConnell, J. R., Romanini, D., Mitchell, L. E. and Brook, E. J.: Continuous methane measurements from a late Holocene Greenland ice core: Atmospheric and in-situ signals, Earth Planet. Sci. Lett., 368, 9–19, doi:10.1016/j.epsl.2013.02.034, 2013.

Rhodes, R. H., Faïn, X., Brook, E. J., McConnell, J. R., Maselli, O. J., Sigl, M., Edwards, J., Buizert, C., Blunier, T., Chappellaz, J. and Freitag, J.: Local artifacts in ice core methane records caused by layered bubble trapping and in situ production: A multi-site investigation, Clim. Past, 12(4), 1061–1077, doi:10.5194/cp-12-1061-2016, 2016.

Rosen, J. L., Brook, E. J., Severinghaus, J. P., Blunier, T., Mitchell, L. E., Lee, J. E., Edwards, J. S. and Gkinis, V.: An ice core record of near-synchronous global climate changes at the Bølling transition, Nat. Geosci., 7(6), 459–463, doi:10.1038/ngeo2147, 2014.

Schwander, J.: The Transformation of Snow to Ice and the Occlusion of Gases, in The Environmental Record in Glaciers and Ice Sheets, pp. 53–67., 1989.

Schwander, J., Barnola, J.-M., Andrié, C., Leuenberger, M., Ludin, A., Raynaud, D. and Stauffer, B.: The age of the air in the firn and the ice at Summit, Greenland, J. Geophys. Res., 98(D2), 2831–2838, doi:10.1029/92JD02383, 1993.

Schwander, J., Sowers, T., Barnola, J. M., Blunier, T., Fuchs, A. and Malaizé, B.: Age scale of the air in the summit ice: Implication for glacial-interglacial temperature change, J. Geophys. Res., 102(D16), 19483–19493, doi:10.1029/97JD01309, 1997.

Seierstad, I. K., Abbott, P. M., Bigler, M., Blunier, T., Bourne, A. J., Brook, E., Buchardt, S. L., Buizert, C., Clausen, H. B., Cook, E., Dahl-Jensen, D., Davies, S. M., Guillevic, M., Johnsen, S. J., Pedersen, D. S., Popp, T. J., Rasmussen, S. O., Severinghaus, J. P., Svensson, A. and Vinther, B. M.: Consistently dated records from the Greenland GRIP, GISP2 and NGRIP ice cores for the past 104ka reveal regional millennial-scale d18O gradients with possible Heinrich event imprint, Quat. Sci. Rev., 106, 29–46, doi:10.1016/j.quascirev.2014.10.032, 2014.

Severinghaus, J. P., Sowerst, T. and Alleyt, R. B.: Timing of abrupt climate change at the end of the Younger Dryas interval from thermally fractionated gases in polar ice, Nature, 39(January), 141–146, doi:10.1002/jqs.622, 1998.

Severinghaus, J. P., Grachev, A. and Battle, M.: Thermal fractionation of air in polar firn by seasonal temperature gradients, Geochemistry, Geophys. Geosystems, 2(7), n/a-n/a, doi:10.1029/2000GC000146, 2001.

Severinghaus, J. P., Albert, M. R., Courville, Z. R., Fahnestock, M. A., Kawamura, K., Montzka, S. A., Mühle, J., Scambos, T. A., Shields, E., Shuman, C. A., Suwa, M., Tans, P. and Weiss, R. F.: Deep air convection in the firn at a zero-accumulation site, central Antarctica, Earth Planet. Sci. Lett., 293(3–4), 359–367, doi:10.1016/j.epsl.2010.03.003, 2010.

Spahni, R.: The attenuation of fast atmospheric CH 4 variations recorded in polar ice cores, Geophys. Res. Lett., 30(11), 1571, doi:10.1029/2003GL017093, 2003.

Steig, E. J., Grootes, P. M. and Stuiver, M.: Seasonal Precipitation Timing and Ice Core Records, Science (80-.)., 266(5192), 1885–1886, doi:10.1126/science.266.5192.1885, 1994.

Stuiver, M., Grootes, P. M. and Braziunas, T. F.: The GISP2 δ18O Climate Record of the Past 16,500 Years and the Role of the Sun, Ocean, and Volcanoes, Quat. Res., 44(3), 341–354, doi:10.1006/qres.1995.1079, 1995.

Werner, M., Heimann, M. and Hoffmann, G.: Isotopic composition and origin of polar precipitation in presentand glacial climate simulations, Tellus Ser. B Chem. Phys. Meteorol. B, 53, doi:10.1034/j.1600-0889.2001.01154.x, 2001.

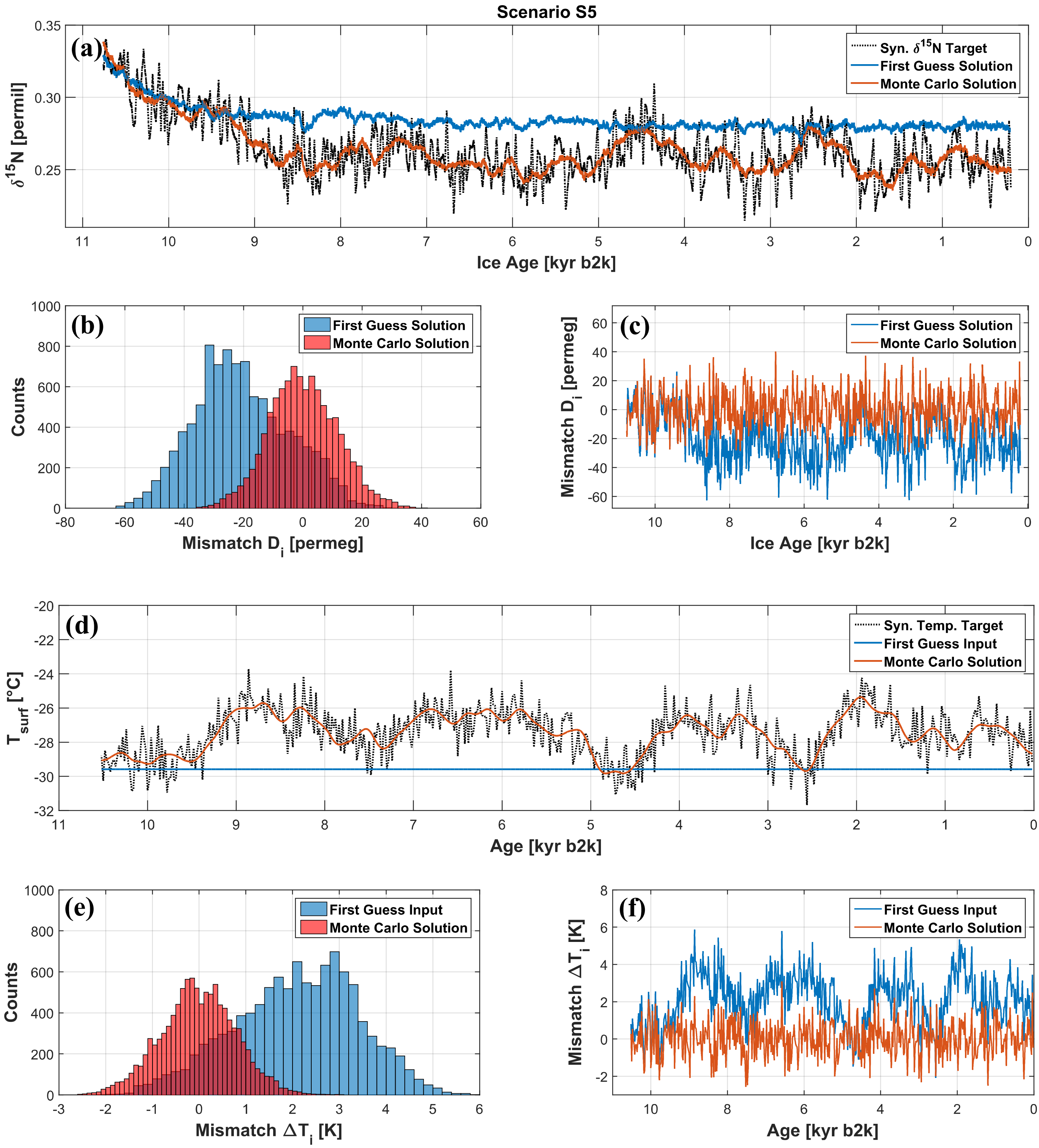

Figure 01: (a-c) First guess vs. Monte Carlo $\delta^{15}N$ solution for the scenario S5: (a) Synthetic $\delta^{15}N$ target (black dotted line), modelled $\delta^{15}N$ time series for the first guess input (blue line) and Monte Carlo solution (red line); (b) Histogram shows the pointwise $\delta^{15}N$ mismatches D_i for the first guess solution (blue) and the Monte Carlo solution (red) versus the synthetic target; (c) Time series for the pointwise $\delta^{15}N$ mismatches D_i for the first guess solution (blue) and the Monte Carlo solution (red) versus the synthetic target;
(d-f) First guess vs. Monte Carlo surface temperature solution for the scenario S5: (d) Synthetic surface temperature (black dotted line), first guess temperature input (blue line) and Monte Carlo solution (red line); (e) Histogram shows the pointwise temperature mismatches ΔT_i for the first guess solution (blue) and the Monte Carlo solution (red) versus the synthetic surface temperature target; (f) Time series for the pointwise temperature mismatches ΔT_i for the first guess solution (blue) and the Monte Carlo solution (red) versus the synthetic surface temperature target;

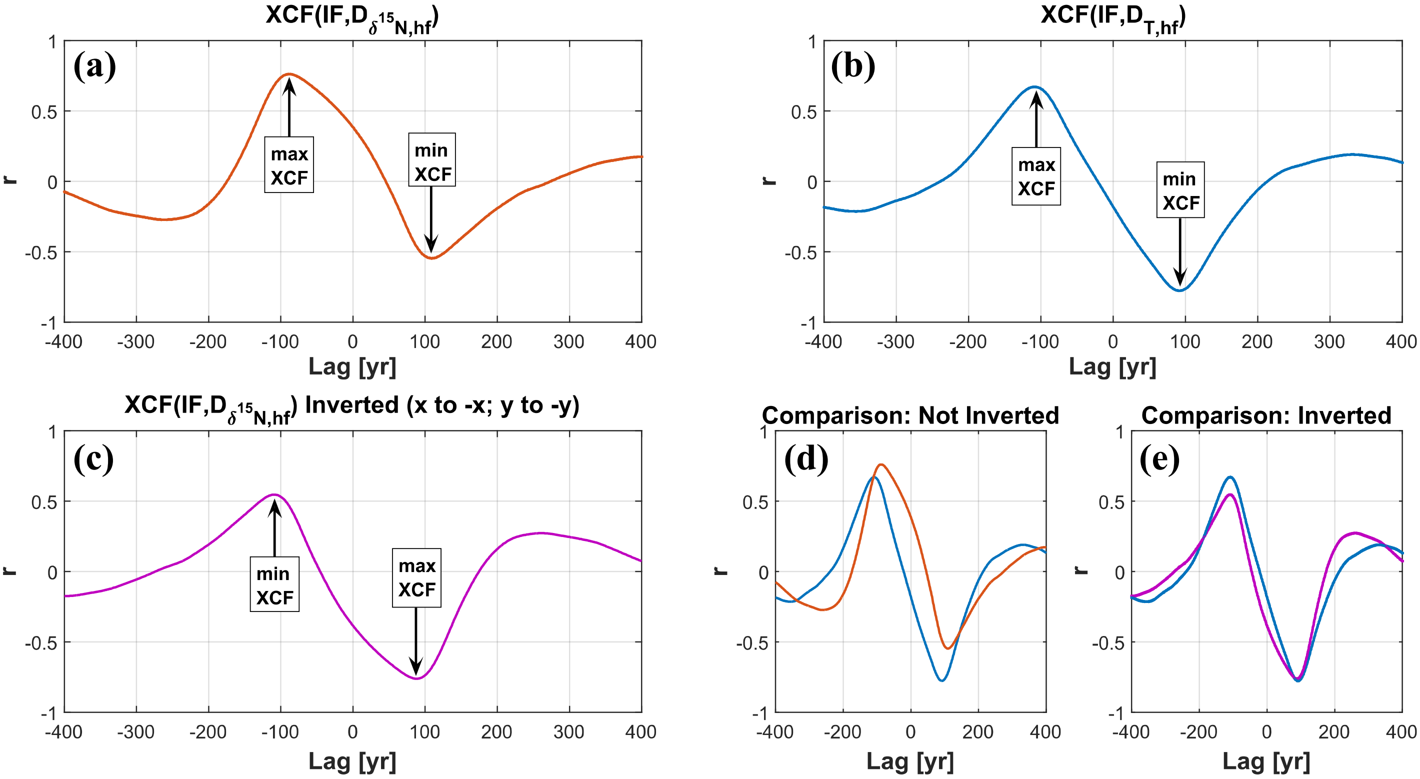

Figure 02: Scenario S1: (a) Cross correlation function (xcf) between IF and the remaining mismatch in $\delta^{15}N$ ($D_{\delta 15N,hf}$) after the high frequency part, shows two extrema: the maximum correlation (max xcf) and the minimum correlation (min xcf); (b) Cross correlation function (xcf) between IF and the remaining mismatch in temperature ($D_{T,hf}$) after the high frequency part shows two extrema: the maximum correlation (max xcf) and the minimum correlation (min xcf); (c) Inverting of (a) in x (lag) and y (correlation coefficient) direction; (d) Comparison between (a) and (b); (e) Comparison between (a) and (c); The temperature correction values are calculated from the linear dependency between IF and $D_{\delta 15N,hf}$. After shifting IF to max xcf (lag max) and min xcf (lag min) $D_{cv,\delta 15N,max}$ and $D_{cv,\delta 15N,min}$ are calculated. Next, $D_{cv,\delta 15N,max}$ and $D_{cv,\delta 15N,min}$ are inverted. That means for $D_{cv,\delta 15N,max}$ that the values of $D_{cv,\delta 15N,max}$ are shifted back (-lag max) and shifted further to lag min. After inverting, $D_{cv,\delta 15N,max}$ and $D_{cv,\delta 15N,min}$ are summed up component wise to calculate $D_{cv,\delta 15N}$. Using $D_{cv,\delta 15N}$ in Eq. (10) leads to the temperature correction values which are added to the temperature T_{hf};

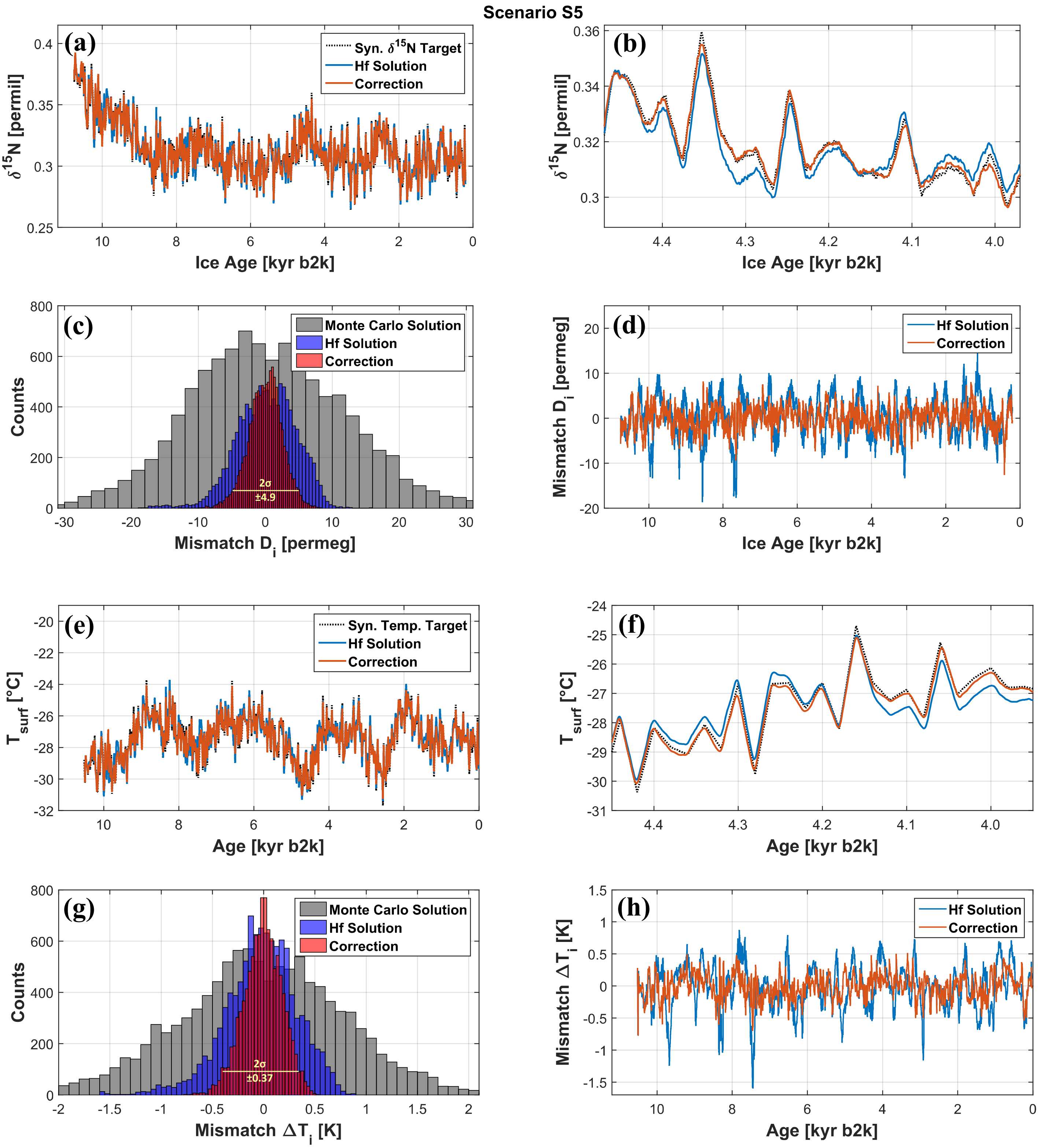

Figure 03: (a-d) $\delta^{15}N$: (a) Synthetic $\delta^{15}N$ target (black dotted line), modelled $\delta^{15}N$ time series after adding high frequency information (blue line) and correction (red line) for the scenario S5; (b) Zoom in for a randomly chosen 500 yr interval shows the decrease of the mismatch after the correction compared to the high frequency solution; (c) Histogram shows the pointwise mismatches D_i from the synthetic $\delta^{15}N$ target for the Monte Carlo solution (grey), the high frequency solution (blue) and the correction (red); The 95 % quantile is 4.9 permeg (yellow line) and used as an estimate for 2σ uncertainty of the final solution; (d) Time series for the pointwise mismatches D_i from the synthetic $\delta^{15}N$ target for the high frequency solution (blue) and the correction (red); (e-h) temperature: (e) Synthetic temperature target (black dotted line), modelled temperature time series after adding high frequency information (blue line) and correction (red line); (f) Zoom in for a randomly chosen 500 yr interval shows the decrease of the mismatch after the correction compared to the high frequency solution; (g) Histogram shows the pointwise mismatches ΔT_i from the synthetic temperature target for the Monte Carlo solution (grey), the high frequency solution (blue) and the correction (red); The 95 % quantile is 0.37 K (yellow line) and used as an estimate for 2σ uncertainty of the final solution; (h) Time series for the pointwise mismatches ΔT_i from the synthetic temperature target for the high frequency solution (blue) and the correction (red);

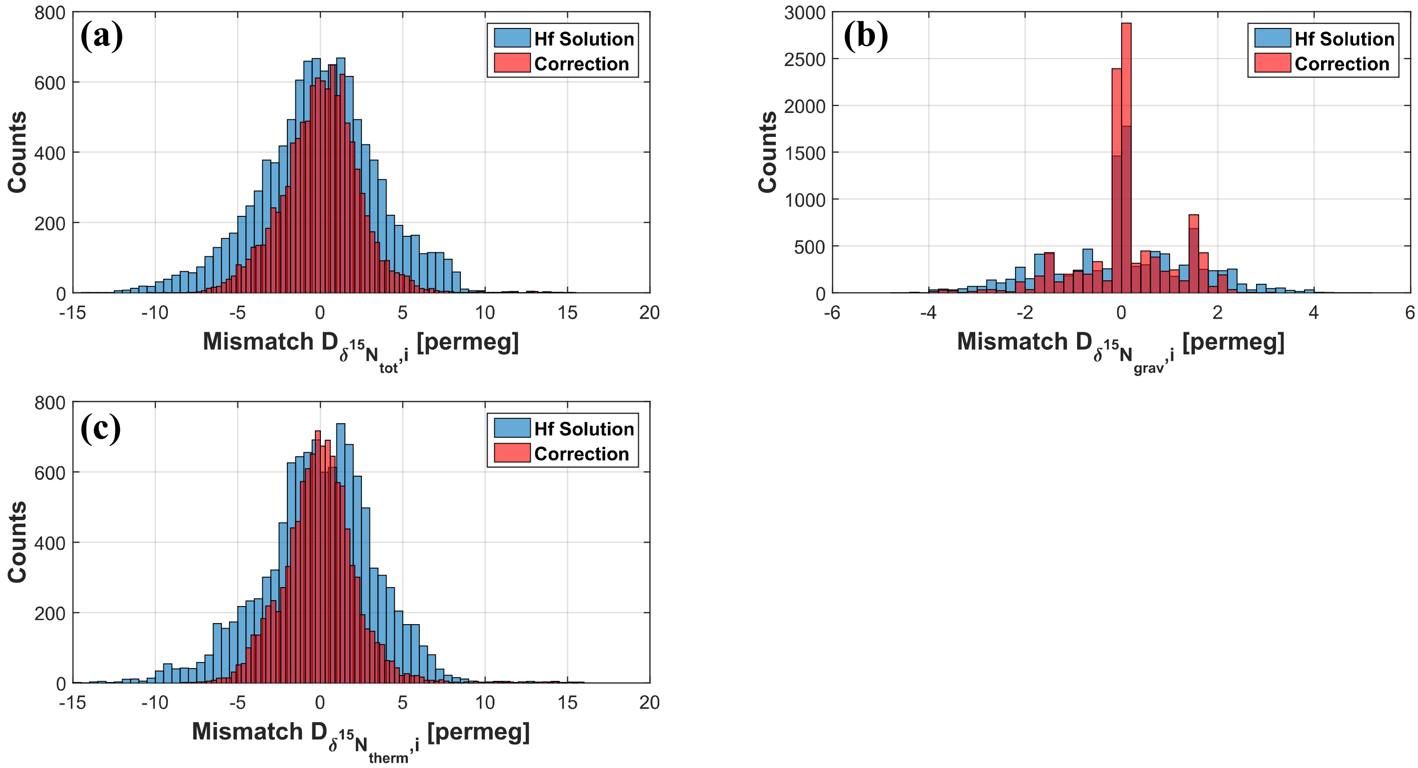

Figure 04: Histograms shows the pointwise mismatches $D_{\delta15N,i}$ in $\delta^{15}N$ between the synthetic target data and the $\delta^{15}N$ solution of the high frequency part (blue) and the correction part (red) for (a) the mismatch in total $\delta^{15}N$ ($D_{\delta15Ntot,i}$), (b) in the gravitational ($D_{\delta15Ngrav,i}$), and (c) in the thermal diffusion component ($D_{\delta15Ntherm,i}$) of $\delta^{15}N$ for the synthetic data scenario S1;

Tables

Scenario	S1	S2	S3	S4	S5	H1	H2	H3
$D_{15N,guess}$ [permeg]	13.3	48.4	27.0	23.3	22.4	23.8	24.1	23.8
$D_{15N,mc}$ [permeg]	11.3	12.4	12.7	11.9	11.5	5.8	6.9	8.2
$\delta^{15}N$ Improvement	2.0	36.0	14.3	11.4	10.9	18.0	17.2	15.6
[permeg | %]	15.0	74.4	53.0	48.9	48.7	75.6	71.4	65.5
# improvements	119	351	152	108	174	223	173	325
# used improvements	89	174	103	74	102	129	112	193
# iterations	2103	706	620	656	637	1636	1027	2086
# tried solutions	16824	5648	4960	5248	5096	13088	8216	16688
Time [h]	52.6	17.7	15.5	16.4	15.9	40.9	25.7	52.2
Comments	Week-end					Week-end		Week-end
$D_{T,guess}$ [K]	1.24	5.24	2.45	2.09	2.17	2.34	2.38	2.32
$D_{T,mc}$ [K]	0.61	0.69	0.70	0.64	0.64	0.32	0.39	0.46
Temp. Improvement	0.63	4.55	1.75	1.45	1.53	2.02	1.99	1.86
[K | %]	50.8	86.8	71.4	69.4	70.5	86.3	83.6	80.2

Table 01: Summary for the Monte Carlo approach; Mismatch D_{guess} between the modelled $\delta^{15}N$ (or temperature) values using the first guess input and the synthetic $\delta^{15}N$ (or temperature) target for each scenario; D_{mc} is the mismatch between the modelled $\delta^{15}N$ using the final Monte Carlo temperature solution and the synthetic $\delta^{15}N$ (or temperature) target for each scenario; 3 runs were conducted over weekend, which leads to a higher number of iterations;

Scenario	S1	S2	S3	S4	S5	H1	H2	H3
$D_{15N,hf}$ [permeg]	2.7	3.6	4.3	3.2	3.5	2.1	2.5	2.6
Improvement (hf vs. MC) [%]	76.1	71.0	66.1	73.1	69.6	63.8	63.8	68.3
$\sigma_{15N,hf}$ [permeg]	3.5	4.6	5.4	4.0	4.3	2.7	3.1	3.3
$D_{15N,corr}$ [permeg]	1.7	2.1	2.6	1.9	2.0	1.2	1.3	1.6
Improvement (corr vs. hf) [%]	37.0	41.7	39.5	40.6	42.9	42.9	48.0	38.5
$\sigma_{15N,corr}$ [permeg]	2.2	2.7	3.3	2.4	2.5	1.5	1.7	1.9
$2\sigma_{15N,corr,95}$ [permeg]	**4.4**	**5.3**	**6.3**	**4.7**	**4.9**	**3.0**	**3.4**	**3.7**
$D_{T,hf}$ [K]	0.20	0.32	0.33	0.25	0.27	0.18	0.21	0.22
$\sigma_{T,hf}$ [K]	0.26	0.40	0.43	0.32	0.35	0.22	0.26	0.27
$D_{T,corr}$ [K]	0.12	0.18	0.20	0.14	0.15	0.10	0.11	0.12
$\sigma_{T,corr}$ [K]	0.15	0.24	0.25	0.19	0.19	0.12	0.14	0.15
$2\sigma_{T,corr,95}$ [K]	**0.31**	**0.48**	**0.51**	**0.38**	**0.37**	**0.23**	**0.27**	**0.30**

Table 02: Summary for the high frequency (hf) and correction part (corr) of the reconstruction approach. D is the mean mismatch between the modelled $\delta^{15}N$ (or temperature) data versus the synthetic $\delta^{15}N$ (or temperature) target. σ is the standard deviation of the point wise mismatches D_i. The 95 % quantiles ($2\sigma_{15N,corr,95}$ or $2\sigma_{T,corr,95}$) of the pointwise $\delta^{15}N$ (or temperature) mismatches (D_i or ΔT_i) are used as an estimate for the 2σ uncertainty for the final solution.

Scenario:	2σ Δ(Δage) [yr]	Scenario:	2σ Δ(Δage) [yr]
S1	1.14	**S5**	1.24
S2	1.60	**H1**	1.23
S3	1.98	**H2**	1.18
S4	1.41	**H3**	1.30

Table 03: Final mismatches Δ(Δage) (2σ) of Δage between the corrected solution and the synthetic targets for all scenarios.