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

Michael Döring\textsuperscript{1,2} and Markus C. Leuenberger\textsuperscript{1,2}
\textsuperscript{1}Climate and Environmental Physics, University of Bern, Bern, Switzerland\textsuperscript{2}Oeschger Centre for Climate Change Research (OCCR), Bern, Switzerland

\textbf{Correspondence:} Michael Döring (doering@climate.unibe.ch)

Received: 3 July 2017 – Discussion started: 17 July 2017
Revised: 30 April 2018 – Accepted: 12 May 2018 – Published: 12 June 2018

\textbf{Abstract.} Greenland past temperature history can be reconstructed by forcing the output of a firm-densification and heat-diffusion model to fit multiple gas-isotope data ($\delta^{15}$N or $\delta^{40}$Ar or $\delta^{13}$N\textsubscript{excess}) extracted from ancient air in Greenland ice cores using published accumulation-rate (Acc) datasets. 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, Acc) 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 and 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\textsubscript{excess} data, which led as well to a robust agreement between modelled and measured data. The obtained final mismatches follow a symmetric standard-distribution function. For the study on synthetic data, 95 % of the mismatches compared to the synthetic target data are in an envelope between 3.0 to 6.3 permeg for $\delta^{15}$N and 0.23 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 fitting the North Greenland Ice Core Project (NGRIP) $\delta^{15}$N data for two Dansgaard–Oeschger events using the presented approach, leading to results comparable to other studies.

\section{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 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 Karlén (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 these 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 showed 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 (de Beauville et al., 2017; Crausbay et al., 2017; Pál et al.,...
(2016) and human societies (Holmgren et al., 2016; Lespez 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), Cuffey and Clow (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 water-stable isotopes of oxygen and hydrogen $(\delta^{18}O_{\text{ice}}, \delta D_{\text{ice}})$ from ice-core watersamples assuming a constant (and mostly linear) relationship between temperature and isotopic composition due to fractionation effects during ocean evaporation, cloud formation and snow and ice precipitation (Johnsen et al., 2001; Stuiver et al., 1995). This method provides a rather robust tool for reconstructing past temperature for times where large temperature excursions occur when an adequate relationship is used (Dansgaard-Oeschger events, glacial-interglacial transitions; Dansgaard et al., 1982; Johnsen et al., 1992). Also, in the Holocene where Greenland temperature variations are comparatively small, seasonal changes of precipitation as well as evaporation conditions at the source region may contribute to water-isotope-data variations (Huber et al., 2006; Kindler et al., 2014; Werner et al., 2001). 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 firm layer of an ice sheet (Huber et al., 2006; Kindler et al., 2014; Kobashi et al., 2011; Orsi et al., 2014; Severinghaus et al., 1998, 2001). 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 firm processes (Leuenberger et al., 1999; Mariotti, 1983; Severinghaus et al., 1998). To robustly reconstruct the surface temperature for a given drilling site, the use of firm models describing gas and heat diffusion throughout the ice sheet is necessary to decompose the gravitational from the thermal-diffusion influence on the isotope signals.

This work addresses two issues relevant for temperature reconstructions based on nitrogen and argon isotopes. 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 firm-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 in addition to $\delta^{15}N$ for all relevant gas-isotope quantities ($\delta^{15}N$, $\delta^{40}Ar$, $\delta^{15}N_{\text{excess}}$) 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 a 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 water-stable isotopes ($\delta^{18}O_{\text{ice}}, \delta D_{\text{ice}}$), 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. Capron et al., 2010; Huber et al., 2006; Landais et al., 2004). To our knowledge, there are only two other reconstruction methods independent from water-stable 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 $\delta^{13}N_{\text{excess}}$ to calculate firm-temperature gradients, which are later temporally integrated from past to future over the time series of interest using the firm-densification and heat-diffusion model from Goujon et al. (2003). Additionally Orsi et al. (2014) use a linearized firm-model approach together with $\delta^{15}N$ and $\delta^{40}Ar$ data to extract surface-temperature histories. The method presented here can be used when no $\delta^{40}Ar$ data are available, which is often the case because $\delta^{40}Ar$ is a more analytically challenging measurement and is not as commonly measured as $\delta^{15}N$ and further allows a comparison among solutions obtained from any of the available isotope quantities.

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 questions in this study are what the minimal mismatch in $\delta^{15}N$ for Holocene-like data we can reach is and what the implication for the final temperature mismatches is. 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 $\delta^{15}N_{\text{excess}}$ datasets, for which of course the original driving temperature histories are unknown.
Figure 1. Schematic illustration of the presented gas-isotope fitting algorithm. The algorithm is implemented in four steps: step 1: first-guess input calculation; step 2: iteratively Monte Carlo based input change (indicated by the open half-cycles); step 3: signal complementation with high-frequency information; step 4: final correction. In contrast to the synthetic data study on Holocene-like data where the accumulation input Acc(t) was fixed, for the proof of concept on glacial data, the accumulation and temperature input was iteratively changed in parallel indicated by the grey variables Acc,g,0 and Acc,csc,fin. For the glacial study, only steps 1 and 2 were used.

2 Methods and data

2.1 Reconstruction approach

The problem that we deal with is an inverse problem, since the effect, observed as $\delta^{15}N$ variations, is dependent on its drivers, i.e. temperature and accumulation-rate changes. Hence, the temperature that we like to reconstructed depends on $\delta^{15}N$ and accumulation-rate changes. To solve this inverse problem, the firm-densification and heat-diffusion model (from now on referred to as firm model), which is a non-linear transfer function of temperature and accumulation rate to firm states and relates to $\delta^{15}N$ values, is run iteratively to match the modelled and measured $\delta^{15}N$ values (or other gas species). The automated procedure is significantly more efficient and less time consuming than a manual approach. The Holocene temperature reconstruction is implemented by the following four steps (Fig. 1):

- Step 1: a prior temperature input (first guess) is constructed, which serves as the starting point for the optimization.
- Step 2: a long-term 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 millennia) as well as firm-column
height changes (temperature and accumulation-rate
dependent) that drive the gravitational background signal
delta^{15}N.

- Step 3: the long-term temperature solution is comple-
mented by superimposing short-term information di-
rectly 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.

- Step 4: the gained temperature solution is further cor-
corrected using information extracted from the mismatch
between the synthetic target and modelled delta^{15}N time
series.

The functionality of the presented inversion algorithm is
schematically displayed in Fig. 1. It guides the reader
through chapters and documents where variables, listed in
Table 1, are in use. In the following, a detailed description of
each step is given.

2.1.1 Step 1: prior input

The starting point of the optimization procedure is the first
guess. To construct the first-guess temperature input T_{g,0}(t),
a constant temperature of -29.6 °C is used for the complete
Holocene section, which corresponds to the last value of the
temperature spin-up (Fig. 2b).

2.1.2 Step 2: Monte Carlo type input generator –
generating long-term solutions

During the second step of the optimization, the prior tem-
perature input T_{g,0}(t) from step 1 is iteratively (i) changed
following a Monte Carlo approach. The basic idea of the
Monte Carlo approach is to generate smooth temperature in-
puts T_{mc,j}(t) by superimposing low-pass-filtered values P_{j}
of uniformly distributed random values P_{r,j} on the prior
input T_{mc,j-1}. Then, the new input is fed to the firm model and
the mismatch D_{\delta^{15}N_{mc,j}} (with X \equiv \delta^{15}N_{mc,j}) between the
modelled \delta^{15}N_{mc,j} (here X_{mod}), calculated from the model
output, and the synthetic \delta^{15}N_{syn} (here X_{target}) is computed
every time step (i) of the target data \delta^{15}N_{syn} according to

\[ D_X = \frac{1}{n} \sum_{i=1}^{n} |D_{X,i}| = \frac{1}{n} \sum_{i=1}^{n} |X_{target,i} - X_{mod,i}|. \]  (1)

(Note: if not otherwise stated, all mismatches in this study
labelled with “D” are calculated similar to Eq. 1.)

\( D_{\delta^{15}N_{mc,j}} \) serves as the criterion which is minimized
during the optimization in step 2. If the mismatch \( D_{\delta^{15}N_{mc,j}} \) de-
creases compared to the prior input \( T_{mc,j-1}, D_{\delta^{15}N_{mc,j-1}} \),
the new input is saved and used as the new guess
\( (T_{g,j} = T_{mc,j}) \). This procedure is repeated until conver-
gence is achieved, leading to the final long-term temperature
\( T_{mc,fin}(t) \). Table 2 lists the number of improvements and iter-
ations performed for the different synthetic datasets.

The perturbation of the current guess \( T_{g,j} \) is conducted in
the following way: let \( T_{g,0} = T_{g,0}(t) \) be the vector contain-
ing the prior temperature input. A second vector \( (P_{r,1}) \) with
the same number of elements \( n_{mc} \) at \( T_{g,0} \) is generated, con-
taining \( n_{mc} \) uniformly distributed random numbers within the
limits of an also randomly (equally distributed) chosen stan-
dard deviation (SD). SD is chosen from a range of 0.05 to 0.50,
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, \( P_{r,1} \) is low-pass filtered
using cubic-spline filtering (Enting, 1987) with an equally
distributed random cut-off period (COP) in the range of 500
to 2000 years generating the vector \( P_{r,1} \). Hereby the low-pass
filtering of \( P_{r,1} \) reduces the amplitudes of the perturbation of
\( T_{g,0} \). The new surface temperature input \( T_{mc,1} \) is calculated from \( P_{r,1} \) according to

\[ T_{mc,1} = T_{g,0} \cdot (\hat{I} + P_{r,1}). \]  (2)

The superscript “T” stands for transposed and \( \hat{I} \) is the \( n \) by
1 matrix of ones.

This approach provides a high potential for parallel com-
puting. In this study, an eight-core computer was used, gen-
erating and running eight different inputs of \( T_{mc} \) simulta-
neously, minimizing the time to find an improved solution.
For example, during the 706 iterations for scenario S2, about
5600 different inputs were created and tested, leading to
351 improvements (see Table 2). Since it is possible to find
more than one improvement per iteration step due to the par-
allelization on eight CPUs, the solution giving the minimal
mismatch \( D_{\delta^{15}N_{mc,fin}} \) is chosen as new first-guess for the next
iteration step. This leads to a decrease of the used improve-
ments for the optimization (e.g. for S2, 172 of the 351
improvements were used). Additionally, a first gas-age scale
(\( \Delta \text{age}_{mc,fin}(t) \)) is extracted from the model using the last im-
proved conditions, which will then be used in step 3.

2.1.3 Step 3: adding short-term (high-frequency)
information

In step 3, the missing short-term temperature history provid-
ing a suitable fit between modelled and synthetic \delta^{15}N data is
directly extracted from the pointwise mismatch \( D_{\delta^{15}N_{mc,fin}}(t) \),
within the modelled \( \delta^{15}N_{mc,fin}(t) \) obtained in step 2 and the
synthetic \( \delta^{15}N_{syn} \) target. Note that for a real reconstruction,
this mismatch is calculated using the measured \( \delta^{15}N_{meas} \)
dataset instead of the synthetic one. \( D_{\delta^{15}N_{mc,fin}}(t) \) can be
interpreted in first order as the detrended high-frequency signal
of the synthetic \( \delta^{15}N_{syn} \) target. \( D_{\delta^{15}N_{mc,fin}}(t) \) is transferred to
the gas-age scale using \( \Delta \text{age}_{mc,fin}(t) \) provided by the firm-
model output for the smooth temperature input \( T_{mc,fin}(t) \).
Table 1. Used variables and acronyms with their explanations.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_T )</td>
<td>thermal-diffusion constant calculated from Eq. (12)</td>
</tr>
<tr>
<td>( \alpha_{18O} )</td>
<td>slope for ( \delta^{18}O_{\text{ice}} ) calibration (surface-temperature spin-up), Eq. (13)</td>
</tr>
<tr>
<td>Acc</td>
<td>accumulation-rate data</td>
</tr>
<tr>
<td>Acc,0</td>
<td>first-guess (prior) input accumulation-rate data</td>
</tr>
<tr>
<td>Acc,mc,fin</td>
<td>modelled accumulation-rate data from the final Monte Carlo output</td>
</tr>
<tr>
<td>( \beta )</td>
<td>intercept for ( \delta^{18}O_{\text{ice}} ) calibration (surface-temperature spin-up), Eq. (13)</td>
</tr>
<tr>
<td>COP</td>
<td>cut-off period for cubic-spline filtering</td>
</tr>
<tr>
<td>corr</td>
<td>index related to the final correction step (step 4)</td>
</tr>
<tr>
<td>CZ</td>
<td>convective zone</td>
</tr>
<tr>
<td>( D )</td>
<td>mean mismatch (general) calculated from Eq. (1)</td>
</tr>
<tr>
<td>( D(t), D_i )</td>
<td>pointwise mismatches (general)</td>
</tr>
<tr>
<td>( D_{\delta^{15}N} )</td>
<td>mean mismatch of ( \delta^{15}N )</td>
</tr>
<tr>
<td>( D_{\delta^{15}N_{\text{corr}}} )</td>
<td>mean mismatch of ( \delta^{15}N ) (( \delta^{15}N_{\text{syn}} ) vs. ( \delta^{15}N_{\text{corr}} )) calculated from the output of the final correction (step 4)</td>
</tr>
<tr>
<td>( D_{\delta^{15}N_{\text{syn}}} )</td>
<td>mean mismatch of ( \delta^{15}N ) (( \delta^{15}N_{\text{syn}} ) vs. ( \delta^{15}N_{\text{corr}} )) calculated from the output of the first-guess data</td>
</tr>
<tr>
<td>( D_{\delta^{15}N_{\text{hf}}} )</td>
<td>mean mismatch of ( \delta^{15}N ) (( \delta^{15}N_{\text{syn}} ) vs. ( \delta^{15}N_{\text{hf}} )) calculated from the output of the high-frequency step (step 3)</td>
</tr>
<tr>
<td>( D_{\delta^{15}N_{\text{mc,fin}}} )</td>
<td>mean mismatch of ( \delta^{15}N ) (( \delta^{15}N_{\text{syn}} ) vs. ( \delta^{15}N_{\text{mc,fin}} )) calculated from the final Monte Carlo output (step 2)</td>
</tr>
<tr>
<td>( Dg )</td>
<td>minimization criterion for the proof of concept on glacial data as used in Eq. (14)</td>
</tr>
<tr>
<td>( DT )</td>
<td>mean mismatch of temperature</td>
</tr>
<tr>
<td>( DT_{\text{corr}} )</td>
<td>mean mismatch of temperature (( T_{\text{syn}} ) vs. ( T_{\text{corr}} )) calculated from the output of the final correction (step 4)</td>
</tr>
<tr>
<td>( DT_{\text{syn}} )</td>
<td>mean mismatch of temperature calculated from ( T_{\text{syn}} ) vs. ( T_g )</td>
</tr>
<tr>
<td>( DT_{\text{hf}} )</td>
<td>mean mismatch of temperature (( T_{\text{syn}} ) vs. ( T_{\text{hf}} )) calculated from the output of the high-frequency step (step 3)</td>
</tr>
<tr>
<td>( DT_{\text{mc,fin}} )</td>
<td>mean mismatch of temperature calculated from the final Monte Carlo output (step 2)</td>
</tr>
<tr>
<td>( \delta^{15}N_{\text{corr}} )</td>
<td>modelled ( \delta^{15}N ) signal from the output of the final correction (step 4)</td>
</tr>
<tr>
<td>( \delta^{15}N_{\text{syn}} )</td>
<td>modelled ( \delta^{15}N ) signal (fitting target)</td>
</tr>
<tr>
<td>( \delta^{15}N_{\text{syn}},0 )</td>
<td>synthetic ( \delta^{15}N ) target (fitting target)</td>
</tr>
<tr>
<td>( \Delta )</td>
<td>gas-age – ice-age difference</td>
</tr>
<tr>
<td>( \Delta_{\text{lag},\text{mc,fin}} )</td>
<td>final gas-age – ice-age output from the Monte Carlo step (step 2)</td>
</tr>
<tr>
<td>( \Delta_{\delta^{15}N_{\text{corr}}} )</td>
<td>( \delta^{15}N ) correction values calculated from ( \Delta_{\delta^{15}N_{\text{max}}} ) and ( \Delta_{\delta^{15}N_{\text{min}}} )</td>
</tr>
<tr>
<td>( \Delta_{\delta^{15}N_{\text{syn}}} )</td>
<td>( \delta^{15}N ) correction values calculated from the linear dependency of ( \chi_{\text{max},\delta^{15}N} )</td>
</tr>
<tr>
<td>( \Delta_{\delta^{15}N_{\text{min}}} )</td>
<td>( \delta^{15}N ) correction values calculated from the linear dependency of ( \chi_{\text{min},\delta^{15}N} )</td>
</tr>
<tr>
<td>( \Delta m )</td>
<td>molar mass-difference between the heavy and light isotopes</td>
</tr>
<tr>
<td>( \Delta T )</td>
<td>high-frequency temperature signal obtained from Eq. (3) (step 3)</td>
</tr>
<tr>
<td>( \epsilon_{\delta^{15}N_{\text{hf}}} )</td>
<td>uncertainty of the ( \delta^{15}N ) data as used in Eq. (14)</td>
</tr>
<tr>
<td>( \epsilon_{\Delta \text{age}} )</td>
<td>uncertainty of the ( \Delta \text{age} ) data as used in Eq. (14)</td>
</tr>
<tr>
<td>( g )</td>
<td>gravitational acceleration</td>
</tr>
<tr>
<td>( g,0 )</td>
<td>index related to the first-guess (prior) data (step 1)</td>
</tr>
<tr>
<td>( hf )</td>
<td>index related to the high-frequency step (step 3)</td>
</tr>
<tr>
<td>( i )</td>
<td>time index</td>
</tr>
<tr>
<td>( IF )</td>
<td>“integrated factor” calculated from Eq. (6), needed for the final correction step (step 4)</td>
</tr>
<tr>
<td>( j )</td>
<td>running index for the Monte Carlo iterations (step 2)</td>
</tr>
<tr>
<td>( \text{l}_{\text{lag,max}} )</td>
<td>time lag attributed to the maximum of the sample cross-correlation function (( \chi_{\text{t}} ), (general)</td>
</tr>
<tr>
<td>( \text{l}_{\text{lag,max},\delta^{15}N} )</td>
<td>time lag attributed to the maximum of the sample cross-correlation function (( \chi_{\text{t}} ) of ( \text{IF}(t) ) vs. ( D_{\delta^{15}N_{\text{hf}}}(t) )</td>
</tr>
<tr>
<td>( \text{l}_{\text{lag,max},T} )</td>
<td>time lag attributed to the maximum of the sample cross-correlation function (( \chi_{\text{t}} ) of ( \text{IF}(t) ) vs. ( D_T(t) )</td>
</tr>
<tr>
<td>( \text{l}_{\text{lag,min}} )</td>
<td>time lag attributed to the minimum of the sample cross-correlation function (general)</td>
</tr>
<tr>
<td>( \text{l}_{\text{lag,min},\delta^{15}N} )</td>
<td>time lag attributed to the minimum of the sample cross-correlation function of ( \text{IF}(t) ) vs. ( D_{\delta^{15}N_{\text{hf}}}(t) )</td>
</tr>
<tr>
<td>( \text{l}_{\text{lag,min},T} )</td>
<td>time lag attributed to the minimum of the sample cross-correlation function of ( \text{IF}(t) ) vs. ( D_T(t) )</td>
</tr>
<tr>
<td>( \text{mc} )</td>
<td>index related to the final Monte Carlo step (step 2)</td>
</tr>
<tr>
<td>( \text{mc,fin} )</td>
<td>index related to the final Monte Carlo output (step 2)</td>
</tr>
<tr>
<td>( n )</td>
<td>number of data points of the target</td>
</tr>
</tbody>
</table>
### Table 1. Continued.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \eta_{\text{me}} )</td>
<td>length of the Holocene temperature vectors (w/o spin-off)</td>
</tr>
<tr>
<td>( \Omega_{K,j} )</td>
<td>thermal-diffusion sensitivity calculated from Eq. (4)</td>
</tr>
<tr>
<td>( P_{j} )</td>
<td>spline-filtered ( P_{T,j} )</td>
</tr>
<tr>
<td>( P_{r,j} )</td>
<td>vector containing ( \eta_{\text{me}} ) uniformly distributed random numbers</td>
</tr>
<tr>
<td>( R )</td>
<td>ideal gas constant</td>
</tr>
<tr>
<td>( \rho_{\text{ice}} )</td>
<td>ice density</td>
</tr>
<tr>
<td>( \rho_{\text{LID}} )</td>
<td>lock-in density, density threshold for calculating ( z_{\text{LID}} )</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation of the random numbers for ( P_{r,j} )</td>
</tr>
<tr>
<td>( \sigma_{d^{15}N_{\text{corr}} \text{,} , \text{me}} )</td>
<td>standard deviation of ( D_{d^{15}N_{\text{corr}}}(t) = D_{d^{15}N_{\text{me}, \text{corr}}} )</td>
</tr>
<tr>
<td>( 2\sigma_{d^{15}N_{\text{corr}, \text{50}}} )</td>
<td>95 % quantile of ( D_{d^{15}N_{\text{corr}}}(t) = D_{d^{15}N_{\text{me}, \text{corr}}} )</td>
</tr>
<tr>
<td>( \sigma_{d^{15}N_{\text{me}} \text{,} , \text{corr}} )</td>
<td>standard deviation of ( D_{d^{15}N_{\text{me}}}(t) = D_{d^{15}N_{\text{me}, \text{corr}}} )</td>
</tr>
<tr>
<td>( \sigma_{T_{\text{corr}}} )</td>
<td>standard deviation of ( D_{T_{\text{corr}}}(t) = D_{T_{\text{me}, \text{corr}}} )</td>
</tr>
<tr>
<td>( 2\sigma_{T_{\text{corr}, \text{50}}} )</td>
<td>95 % quantile of ( D_{T_{\text{corr}}}(t) = D_{T_{\text{me}, \text{corr}}} )</td>
</tr>
<tr>
<td>( \sigma_{T_{\text{me}, \text{corr}}} )</td>
<td>standard deviation of ( D_{T_{\text{me}}}(t) = D_{T_{\text{me}, \text{corr}}} )</td>
</tr>
<tr>
<td>( T_{\text{F, \text{firm}}} )</td>
<td>mean firm temperature</td>
</tr>
<tr>
<td>( T_{\text{bottom}} )</td>
<td>temperatures at the bottom of the diffusive firm layer</td>
</tr>
<tr>
<td>( T_{\text{corr}} )</td>
<td>temperature signal calculated from the final correction step (step 4)</td>
</tr>
<tr>
<td>( T_{\text{g,0}} )</td>
<td>first-guess (prior) temperature input</td>
</tr>
<tr>
<td>( T_{\text{me}} )</td>
<td>temperature signal calculated from the high-frequency step (step 3)</td>
</tr>
<tr>
<td>( T_{\text{me}, j} )</td>
<td>Monte Carlo temperature guess for iteration ( j )</td>
</tr>
<tr>
<td>( T_{\text{me, \text{fin}}} )</td>
<td>(smooth) temperature modelled from the final Monte Carlo output (step 2)</td>
</tr>
<tr>
<td>( T_{\text{surf}} )</td>
<td>surface-temperature spin-up</td>
</tr>
<tr>
<td>( w_{\text{RMSE}} )</td>
<td>mean squared errors weighted with data uncertainty as used in Eq. (14)</td>
</tr>
<tr>
<td>( xc_{\text{corr}} )</td>
<td>sample cross-correlation function, needed for the final correction step (step 4)</td>
</tr>
<tr>
<td>( xc_{\text{corr}, \text{max}} )</td>
<td>maximum of the sample cross-correlation function (general)</td>
</tr>
<tr>
<td>( xc_{\text{corr}, d^{15}N} )</td>
<td>maximum of the sample cross-correlation function of IF(t) vs. ( D_{d^{15}N_{\text{me}}}(t) )</td>
</tr>
<tr>
<td>( xc_{\text{corr, max}, T} )</td>
<td>minimum of the sample cross-correlation function</td>
</tr>
<tr>
<td>( xc_{\text{corr, min}, d^{15}N} )</td>
<td>minimum of the sample cross-correlation function of IF(t) vs. ( D_{d^{15}N_{\text{me}}}(t) )</td>
</tr>
<tr>
<td>( X_{\text{mod}} )</td>
<td>modelled data (general), can be ( d^{15}N ), ( T ) or measured data (( d^{34}Ar, d^{15}N_{\text{exc}} ))</td>
</tr>
<tr>
<td>( X_{\text{target}} )</td>
<td>fitting target (general), can be synthetic ( d^{15}N_{\text{syn}}, T_{\text{syn}} ) or measured data (( d^{34}Ar, d^{15}N_{\text{exc}} ))</td>
</tr>
<tr>
<td>( z_{\text{LID}} )</td>
<td>lock-in depth</td>
</tr>
</tbody>
</table>

### Table 2. Summary for the Monte Carlo approach: mismatch \( D_{\text{g,0}} \) between the modelled \( d^{15}N \) (or temperature) values using the first-guess input and the synthetic \( d^{15}N \) (or temperature) target for each scenario. \( D_{\text{me}} \) is the mismatch between the modelled \( d^{15}N \) (or temperature) using the final Monte Carlo temperature solution and the synthetic \( d^{15}N \) (or temperature) target for each scenario.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
</tr>
</thead>
<tbody>
<tr>
<td>( d^{15}N_{\text{me}, \text{per meg}} ) (permeg)</td>
<td>13.3</td>
<td>48.4</td>
<td>27.0</td>
<td>23.3</td>
<td>22.4</td>
<td>23.8</td>
<td>24.1</td>
<td>23.8</td>
</tr>
<tr>
<td>( d^{15}N_{\text{me, \text{fin}}} ) (permeg)</td>
<td>11.3</td>
<td>12.4</td>
<td>12.7</td>
<td>11.9</td>
<td>11.5</td>
<td>5.8</td>
<td>6.9</td>
<td>8.2</td>
</tr>
<tr>
<td>( d^{15}N ) improvement (permeg)</td>
<td>2.0</td>
<td>36.0</td>
<td>14.3</td>
<td>11.4</td>
<td>10.9</td>
<td>18.0</td>
<td>17.2</td>
<td>15.6</td>
</tr>
<tr>
<td>( d^{15}N ) improvement (%)</td>
<td>15.0</td>
<td>74.4</td>
<td>53.0</td>
<td>48.9</td>
<td>48.7</td>
<td>75.6</td>
<td>71.4</td>
<td>65.5</td>
</tr>
<tr>
<td>No. improvements</td>
<td>119</td>
<td>351</td>
<td>152</td>
<td>108</td>
<td>174</td>
<td>223</td>
<td>173</td>
<td>325</td>
</tr>
<tr>
<td>No. used improvements</td>
<td>89</td>
<td>174</td>
<td>103</td>
<td>74</td>
<td>102</td>
<td>129</td>
<td>112</td>
<td>193</td>
</tr>
<tr>
<td>No. iterations</td>
<td>2103</td>
<td>706</td>
<td>620</td>
<td>656</td>
<td>637</td>
<td>1636</td>
<td>1027</td>
<td>2086</td>
</tr>
<tr>
<td>No. tried solutions</td>
<td>16,824</td>
<td>5648</td>
<td>4960</td>
<td>5248</td>
<td>5096</td>
<td>13,088</td>
<td>8216</td>
<td>16,688</td>
</tr>
<tr>
<td>Execution time (h)</td>
<td>52.6</td>
<td>17.7</td>
<td>15.5</td>
<td>16.4</td>
<td>15.9</td>
<td>40.9</td>
<td>25.7</td>
<td>52.2</td>
</tr>
<tr>
<td>( D_{T_{\text{g,0}}} ) (K)</td>
<td>1.24</td>
<td>5.24</td>
<td>2.45</td>
<td>2.09</td>
<td>2.17</td>
<td>2.34</td>
<td>2.38</td>
<td>2.32</td>
</tr>
<tr>
<td>( D_{T_{\text{me}}} ) (K)</td>
<td>0.61</td>
<td>0.69</td>
<td>0.70</td>
<td>0.64</td>
<td>0.64</td>
<td>0.32</td>
<td>0.39</td>
<td>0.46</td>
</tr>
<tr>
<td>Temp. improvement (K)</td>
<td>0.63</td>
<td>4.55</td>
<td>1.75</td>
<td>1.45</td>
<td>1.53</td>
<td>2.02</td>
<td>1.99</td>
<td>1.86</td>
</tr>
<tr>
<td>Temp. improvement (%)</td>
<td>50.8</td>
<td>86.8</td>
<td>71.4</td>
<td>69.4</td>
<td>70.5</td>
<td>86.3</td>
<td>83.6</td>
<td>80.2</td>
</tr>
</tbody>
</table>
Figure 2. (a) Used accumulation-rate input time series divided in a Holocene and a spin-up section, with time resolution in the Holocene section (20 to 10,520 years b2k) of 1 year. The time resolution for the transition between the Holocene and the spin-up section (10,520 to 12,000 years b2k) is 1 year as well. This is in opposition to the rest of the spin-up section which has a time resolution of 20 years. (b) Surface-temperature spin-up calculated from $d^{18}O_{ice}$ calibration. Time resolution equals the accumulation-rate spin-up section. The first-guess surface temperature input is simply a constant value.

This is needed to insure synchronity between the high-frequency temperature variations $\Delta T(t)$ extracted from the mismatch $D_{15N_{fin}}(t)$ on the ice-age scale and the smooth temperature solution $T_{fin}(t)$. Additionally, the signal is shifted by about 10 years towards modern values to account for gas diffusion from the surface to the lock-in depth (Schwander et al., 1993), which is not yet implemented in the firm model. This is necessary for adding the calculated short-term temperature changes $\Delta T(t)$ to the smooth signal $T_{fin}(t)$. The $\Delta T$ values are calculated according to Eq. (3):

$$\Delta T_i = \frac{D_{15N_{fin}}(i)}{\Omega_{N_i}}$$

(3)

using the thermal-diffusion sensitivity $\Omega_{N_i}$ for nitrogen-isotope fractionation from Grachev and Severinghaus (2003):
\[ \Omega_{N_{i},i} = \frac{8.656 \% c}{T_i} - \frac{1232 \% c \cdot K}{T_i^2}, \]  
(4)

\[ T_{i} \] is the mean fi rm temperature in Kelvin which is calculated by the fi rm model for each time point \( i \). To reconstruct the fi nal (high-frequency) temperature input \( T_{bfi}(t) \), the extracted short-term temperature signal \( \Delta T(t) \) is simply added to the long-term temperature input \( T_{mce,fin}(t) \):

\[ T_{bfi,i} = T_{mce,fin,i} + \Delta T_i. \]  
(5)

2.1.4 Step 4: fi nal correction of the surface temperature solution

For a further improvement of the remaining \( \delta^{15}N \) and resulting surface-temperature misfits \( (D_{\delta^{15}N_{mce},i}(t), D_{dyn}(t)) \), it is important to fi nd a correction method that contains information that is also available when using measured data. The benefi t of the synthetic data study is that several later-unknown quantities can be calculated and used for improving the reconstruction approach (see Sects. 3 and 4). For instance, it is possible to split the synthetic \( \delta^{15}N_{mce} \) data in the gravitational and thermodiffusion 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 \( \delta^{15}N \) \( (D_{\delta^{15}N_{mce},i}(t)) \) and temperature \( (D_{dyn}(t)) \) 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 fi nd a long-term solution \( \delta^{15}N_{mce,fin} \) (or \( T_{mce,fin} \)) which exactly passes through the \( \delta^{15}N_{mce} \) (or \( T_{mce} \)) target in the middle of the variance in all parts of the time series. This leads to a slight over- or underestimation of \( \delta^{15}N_{mce,fin}(t) \) and their corresponding temperature values \( T_{mce,fin}(t) \). For example, a slightly too low (or too high) smooth temperature estimate \( T_{mce,fin} \) leads to a small increase (or decrease) of the fi rm-column height, creating a wrong gravitational background signal in \( \delta^{15}N_{mce,fin} \) on a later point in time (because the fi rm 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 (step 3), because the high-frequency information is directly extracted from the deviation of the synthetic target \( \delta^{15}N_{syn}(t) \) and the modelled \( \delta^{15}N_{mce,fin}(t) \) from the fi nal long-term solution \( T_{mce,fin}(t) \) 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 \( D_{\delta^{15}N_{mce,fin},i}(t) \) of the synthetic target data \( \delta^{15}N_{syn} \) to \( \delta^{15}N_{mce,fin} \) of the Monte Carlo part are numerically integrated over a time window of 200 years (Sect. 4, Supplement Sect. S3), and thereafter the window is shifted from past to future in 1-year steps resulting in a time series called \( IF(t) \). \( IF(t) \) equals a 200-year running mean of \( D_{\delta^{15}N_{mce,fin},i}(t) \). For \( t \), the middle position of the window is allocated. The time evolution of \( IF(t) \) is a measure for the deviation of the long-term solution \( \delta^{15}N_{mce,fin}(t) \) (or \( T_{mce,fin}(t) \)) from the perfect middle passage through the target data \( \delta^{15}N_{syn}(t) \) (or \( T_{syn}(t) \)) and for the slight over- and underestimation of the resulting temperature.

\[ IF(t) = \frac{1}{200} \int_{t-100}^{t+100} (\delta^{15}N_{syn}(t) - \delta^{15}N_{mce,fin}(t)) \, dt \]

\[ IF(t) = \frac{1}{200} \int_{t-100}^{t+100} D_{\delta^{15}N_{mce,fin},i}(t) \, dt \]  
(6)

Next, the sample cross-correlation function (x cf ) (Box et al., 1994) is applied to \( IF(t) \) and the remaining misfits \( D_{\delta^{15}N_{mce},i}(t) \) of \( \delta^{15}N \) after the high-frequency part. The x cf shows two extrema (Fig. 3a), a maximum (x cf max) and a minimum (x cf min) at two certain lags \( (lag_{max,\delta^{15}N} \) at x cf max, \( \delta^{15}N \) and \( lag_{min,\delta^{15}N} \) at x cf min, \( \delta^{15}N \)). Now, the same analysis is conducted for \( IF(t) \) versus the temperature mismatch \( T_{dyn}(t) \) (Fig. 3b), which shows an equal behaviour (two extrema, \( lag_{max,\delta^{15}N} \) at x cf max, \( \delta^{15}N \) and \( lag_{min,\delta^{15}N} \) at x cf min, \( \delta^{15}N \)). Comparing the two cross correlations shows that x cf max, \( \delta^{15}N \) equals the negative \( lag_{min,\delta^{15}N} \) and x cf min, \( \delta^{15}N \) corresponds to the negative \( lag_{max,\delta^{15}N} \) (Fig. 3d, e). The idea for the correction is that the extrema in the cross-correlation \( IF(t) \) versus \( D_{\delta^{15}N_{mce},i}(t) \) with the positive lag (positive means here that \( D_{\delta^{15}N_{mce,fin},i}(t) \) has to be shifted to past values relative to \( IF(t) \)) creates the misfit of temperature \( \delta N_{mce,fin}(t) \) on the negative lag (modern direction) of \( IF(t) \) versus \( T_{dyn}(t) \), and vice versa. So, \( IF(t) \) yields information about the cause and allows us to correct this effect between the remaining mismatches, \( D_{\delta^{15}N_{mce},i}(t) \) and \( \delta N_{mce,fin}(t) \), 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 and 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 \( D_{\delta^{15}N_{mce},i}(t) \) at the two lags \( lag_{max,\delta^{15}N} \) at x cf max, \( \delta^{15}N \) and \( lag_{min,\delta^{15}N} \) at x cf min, \( \delta^{15}N \), two sets of \( \delta^{15}N \) correction values \( (\Delta^\delta^{15}N_{mce,fin}(t) \) from x cf max, \( \delta^{15}N \) and \( \Delta^\delta^{15}N_{mce,fin}(t) \) from x cf min, \( \delta^{15}N \)) are calculated. Then, the lags are inverted (Fig. 3c, e), shifting the two sets of the \( \delta^{15}N \) correction values to the attributed lags of the cross correlation between \( IF(t) \) and \( \delta N_{mce,fin}(t) \) (e.g. \( \Delta^\delta^{15}N_{mce,fin}(t) \) to lag from x cf max, \( \delta^{15}N \)), from the cross correlation between \( IF(t) \) and \( T_{bfi}(t) \), therefore changing the time assignments of \( \Delta^\delta^{15}N_{min}(t) \) and \( \Delta^\delta^{15}N_{max}(t) \) to \( \Delta^\delta^{15}N_{min}(t + lag_{max,\delta^{15}N}) \) and \( \Delta^\delta^{15}N_{max}(t + lag_{max,\delta^{15}N}) \). Now, the \( \Delta^\delta^{15}N_{max}(t) \) and \( \Delta^\delta^{15}N_{min}(t) \) are summed up component-wise, leading to the time series \( \Delta^\delta^{15}N_{mce,fin}(t) \). From Eq. (3) with \( \delta^{15}N_{mce,fin}(t) \), the corresponding temperature correction values are calculated and added to the high-frequency temperature solution \( T_{bfi}(t) \), giving the corrected temperature \( T_{cor}(t) \). Finally, \( T_{cor}(t) \) is used to run the firm model to calculate the

corrected $\delta^{15}N_{\text{corr}}(t)$ time series. This cause-and-effect relationship found in the cross correlations between $\text{IF}(t)$ and $D_{\delta^{15}N_{\text{grav}}}(t)$, and $\text{IF}(t)$ and $D_{\text{TL}}(t)$, is exemplarily shown in Fig. 3 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).

2.2 Firn-densification and heat-diffusion model

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, developed by Schwaneder 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) and 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 are 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:

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

$$
\delta^{15}N_{\text{therm}}(t) = \frac{T_{\text{surf}}(t) - T_{\text{bottom}}(t)}{1} \cdot 1000
$$

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

$\delta^{15}N_{\text{grav}}(t)$ is the component of the isotopic fractionation due to the gravitational settling (Craig et al., 1988; Schwaneder, 1989) and depends on the lock-in depth (LID) $z_{\text{LID}}(t)$ and the mean firn temperature $T(t)$ (Leuenberger et al., 1999). $g$ is the gravitational acceleration, $\Delta m$ the molar mass difference between the heavy and light isotopes (equal to $10^{-3}$ kg mol$^{-1}$ for nitrogen) and $R$ the ideal gas constant. $z_{\text{LID}}$ is defined as a density threshold $\rho_{\text{LID}}$, 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 (Schwaneder et al., 1997):

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

where

$$
\rho_{\text{ice}}(\text{kg m}^{-3}) = 916.5 - 0.14438 \cdot T - 1.5175 \cdot 10^{-4} \cdot T^2.
$$
The thermal-fractionation component of the $\delta^{15}N$ signal (Severinghaus et al., 1998) is calculated using Eq. (8), where $T_{\text{surf}}(t)$ and $T_{\text{bottom}}(t)$ stand for the temperatures at the top and the bottom of the diffusive firm layer. In contrast to $T_{\text{surf}}(t)$, which is an input parameter for the model, $T_{\text{bottom}}(t)$ is calculated by the model for each time step. The thermal-diffusion constant $\alpha_T$ was measured by Grachev and Severinghaus (2003) for nitrogen (Eq. 12):

$$\alpha_T = 8.656 - \frac{1323 K}{T} \cdot 10^{-3}. \quad (12)$$

The firm model used here behaves purely as a forward model, which means that for the given input time series the output parameters (here, finally $\delta^{15}N_{\text{mod}}(t)$) 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.

2.3 Measurement, input data and timescale

2.3.1 Timescale

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 to 10 520 years b2k (called the “Holocene section”) and the second one from 10 520 to 35 000 years b2k (“spin-up section”). The entire accumulation-rate input, as well as the spin-up section of the surface-temperature input, remains unchanged during the reconstruction procedure.

2.3.2 Accumulation-rate data

Besides surface temperatures, accumulation-rate data are needed to drive the firm model. In this study, we use the original accumulation rates, reconstructed in Cuffey and Clow (1997), produced using an ice-flow model adapted to the Greenland Ice Sheet Project Two (GISP2) location but adapted to the GICC05 chronology (Rasmussen et al., 2008; Seierstad et al., 2014). A detailed description of the adaption procedure can be found in Sect. S1 of the Supplement. The raw accumulation-rate data for the main part of the spin-up section (12 000 to 35 000 years b2k) are linearly interpolated to a 20-year grid and low-pass filtered with a 200-year COP using cubic-spline filtering (Enfing, 1987). For the Holocene section (20–10 520 years b2k) and the transition part between Holocene and spin-up section (10 520 to 12 000 years b2k), the raw accumulation-rate data are linearly interpolated to a 1-year 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. 2a). Except for these technical adjustments, the accumulation-rate input remains unmodified, assuming high reliability of these data during the Holocene. The accumulation data were reconstructed using annual layer counting, and a thinning model which should lead to maximum relative uncertainty of 10% for the first 1500 of the 3000 m ice core (Cuffey and Clow, 1997). 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. S1 in the Supplement). Since the differences between the scenarios are not important for the evaluation of the reconstruction approach, they are not taken into account for this study.

Additionally, two sensitivity experiments were conducted (see Sect. S2 in the Supplement) in order to investigate (i) the influence of low-pass filtering of the high-resolution accumulation rates on the model outputs and (ii) the possible contribution of the accumulation-rate variability to the $\delta^{15}N$ data during the Holocene. The first experiment shows that filtering the accumulation rates with cut-off periods in the range of 20 to 500 years has nearly no influence on the modelled $\delta^{15}N$ or lock-in depth as long as the major trends are being conserved. The second experiment leads to the finding that the accumulation-rate variability explains about 12 to 30% of $\delta^{15}N$ variability. A total of 30% corresponds to the 8.2 kyr event and 12% to 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 most parts of the Holocene.

2.3.3 $\delta^{18}O_{\text{ice}}$ data

Oxygen-isotope data from the GISP2 ice-core-water samples measured at the University of Washington’s Quaternary Isotope Laboratory are used to construct the surface-temperature input of the model spin-up (12 to 35 kyr b2k, Grootes et al., 1993; Grootes and Stuiver, 1997; Meese et al., 1994; Steig et al., 1994; Stuiver et al., 1995; data availability: Grootes and Stuiver, 1999). The raw $\delta^{18}O_{\text{ice}}$ data are filtered and interpolated in the same way as the accumulation-rate data for the spin-up part.

2.3.4 Surface-temperature spin-up

The surface-temperature history of the spin-up section (Fig. 2b) is obtained by calibrating the filtered and interpolated $\delta^{18}O_{\text{ice}}$ data (Eq. 13) using the values for the temperature sensitivity $\alpha_{T_{\text{air}}}$ and offset $\beta$ found by Kindler et al. (2014) for the North Greenland Ice Core Project (NGRIP) ice core assuming a linear relationship of $\delta^{18}O_{\text{ice}}$ with temperature.

$$T_{\text{spin}}(t) = \frac{1}{\alpha_{T_{\text{air}}}(t)} \cdot \delta^{18}O_{\text{ice}}(t) + 35.2 \%e - 31.4 \degree C + \beta(t) \quad (13)$$
The values 35.2 °C and –31.4 °C are modern-time parameters for the GISP2 site (Grootes and Stuiver, 1997; Schwander et al., 1997). The spin-up is needed to bring the firm model to a well-defined starting condition that takes possible memory effects (influence of earlier conditions) of firm states into account.

### 2.3.5 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 as targets for the reconstruction. From these eight synthetic surface-temperature and related $\delta^{15}N$ scenarios (S1–S5 and H1–H3), three datasets (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. (2008b) in terms of variability (amplitudes) and frequency (data resolution) of the GISP2 nitrogen-isotope data (Figs. 4, 5).

The synthetic surface-temperature scenarios S1–S5 are created by generating a long-term temperature time series ($T_{\text{syn,smooth}}$) analogous to the Monte Carlo part of the reconstruction procedure for only one iteration step (see Sect. 2.1). The values for the cut-off period used for the filtering of the random values, and the SD values (standard deviation of the random values; see Sect. 2.1) for the first five scenarios can be found in Table 3. The long-term temperatures (Fig. 4f) are calculated on a 20-year grid, which is nearly similar to the time resolution of the GISP2 $\delta^{15}N$ measurement values of about 17 years (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-year cut-off to obtain the long-term temperature scenario.

Following this, high-frequency information is added to the long-term 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-year grid and added up to the long-term temperature time series. Finally, the resulting synthetic target-temperature scenarios (Figs. 4II, 5I) are linearly interpolated to a 1-year grid.

These synthetic temperatures are combined with the spin-up temperature and are used together with the accumulation-rate input to feed the firm model. From the model output, the synthetic $\delta^{15}N$ targets are calculated according to Sect. 2.2. The firm-model output provides ice-age as well as gas-age information. The final synthetic $\delta^{15}N$ target time series ($\delta^{15}N_{\text{syn}}$) are set intentionally on the ice-age scale to mirror measured data, because no prior information is available for the gas-age–ice-age difference (Δage) for ice-core data.

### 3 Results

#### 3.1 Monte Carlo type input generator

Figure 6 shows the evolution of the misfit $D_{\delta^{15}N_{\text{me,j}}}$ between the synthetic target data ($\delta^{15}N_{\text{syn,j}}$) and the modelled output $\delta^{15}N_{\text{me,j}}$ of the Monte Carlo part (step 2) as a function of the applied iterations (j) 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 $D_{\delta^{15}N_{\text{me}}}$ compared to the first-guess solution $D_{\delta^{15}N_{\text{L0}}}$ by about 15 to 75% depending on the scenario and the mismatch of the first-guess solution (see Table 2). This leads to a reduction of the temperature mismatches $D_{T_{\text{me}}}$ compared to the first-guess temperature $D_{T_{\text{L0}}}$ mismatch of about 51 to 87%.

Figure 7 provides the comparison between the first-guess (g,0; step 1) and Monte Carlo (mc,fin; step 2) solution versus the synthetic target data (syn) for the modelled $\delta^{15}N$ (Fig. 7a–c) and surface-temperature values (Fig. 7d–f) for scenario S5. Subplots (a) and (d) show the time series of the synthetic target (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 $D_{t}$ of the first-guess (blue) and the Monte Carlo solutions (red) versus the synthetic target data for $\delta^{15}N$ ($D_{\delta^{15}N}$) and temperature ($D_{T}$) can be found. Subplots (c) and (f) contain the time series for $D_{\delta^{15}N_{\text{me}}}$ and $D_{T_{\text{me}}}$. The $D_{\delta^{15}N_{\text{me,fin}}}(t)$ data (red) are used to calculate the high-frequency signal that is superimposed to the long-term temperature solution $T_{\text{mc,fin}}$ according to Eqs. (3) and (5) (see Sect. 2.1, step 3). From Fig. 7, 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 (g,0), which shifts the midpoint of the distributions of $D_{\delta^{15}N_{\text{me}}}$ and $D_{T_{\text{me}}}$ to zero (see blue against red in Fig. 7b, e). The second improvement is that the distributions become more symmetric and the misfit is overall reduced (the distributions become narrower) comp-
Figure 4. (I) S1–S5: synthetic smooth temperature scenarios for the construction of the target-temperature data. (II) S1–S5: synthetic target surface-temperature scenarios. (III) S1–S5: corresponding synthetic δ¹⁵N target time series.
pared to the first guess due to the middle passage through the $\delta^{15}$N$_{syn}$ targets. These improvements can be observed for all eight synthetic scenarios, showing the robustness of the Monte Carlo part (see Table 2, Fig. 7).

3.2 High-frequency step and final correction

Figure 8 provides the comparison between the Monte Carlo (mc-fin; step 2), the high-frequency (hf; step 3) and the correction (corr; step 4) parts of the reconstruction procedure for the scenario S5. Additional data for all other scenarios can be found in Table 4. The upper four plots (Fig. 8a–d) illustrate each reconstruction step and their effect on the modelled $\delta^{15}$N; the bottom four plots (Fig. 8e–h) show the corresponding results on the temperature. Plots (a) and (e) contain the time series of the synthetic $\delta^{15}$N$_{syn}$ or $T_{syn}$ target (syn; black dotted line), the high-frequency solution (hf; blue line) and the final solution after the correction part (corr; red line). For visibility reasons, subplots (b) and (f) display a zoom-in for a randomly chosen time window of about 500 years 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 ($D_{\delta^{15}N}$ for $\delta^{15}$N; $D_T$ 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 $D_{\delta^{15}N}$, the improvement between the Monte Carlo and the high-frequency parts is in the range of 64 to 76 % (see Table 4). This leads to a reduction of the temperature mismatches $D_T$ of 43 to 67 %.

The standard deviations (1σ) of the pointwise mismatches (Fig. 8c, d, g, h)
in $\delta^{15}N$ and temperature after the high-frequency parts are in the range of about 2.7 to 5.4 permeg (one permeg equals $10^{-6}$) for $\delta^{15}N$ and 0.22 to 0.40 K for the reconstructed temperatures depending on the scenario, which is clearly visible in the decreasing width of the histograms (Fig. 8c and g, 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 $D_{\delta^{15}N}$ as well as $D_T$ is further reduced, and the distributions become more symmetric (long tails disappear; see histograms c and g; red against blue of Fig. 8). The time series of the mismatches (Fig. 8d and h) 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 ($2\sigma_{\delta^{15}N_{corr.95}}$, $2\sigma_{T_{corr.95}}$) of the remaining pointwise mismatches of $\delta^{15}N$ and temperature ($D_{\delta^{15}N}$ or $D_T$) were calculated for the final solutions for all scenarios and are used as an estimate for the $2\sigma$ uncertainty of the reconstruction algorithm (see Fig. 8c, g and Table 4). The final uncertainties (2$\sigma$) are of the order of 3.0 to 6.3 permeg for $\delta^{15}N$ and 0.23 to 0.51 K for the surface temperature misfits. It is noteworthy that the measurement uncertainties (per point) of state-of-the-art $\delta^{15}N$ measure-

### Table 4. 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 and the synthetic $\delta^{15}N$ (or temperature) target. $\sigma$ is the standard deviation of the pointwise mismatches $D_t$. The 95% quantiles ($2\sigma_{\delta^{15}N_{corr.95}}$ or $2\sigma_{T_{corr.95}}$) of the pointwise $\delta^{15}N$ (or temperature) mismatches are used as an estimate for the $2\sigma$ uncertainty for the final solution (values in bold).

<table>
<thead>
<tr>
<th>Scenario</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>H1</th>
<th>H2</th>
<th>H3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{\delta^{15}N_{hf}}$ (permeg)</td>
<td>2.7</td>
<td>3.6</td>
<td>4.3</td>
<td>3.2</td>
<td>3.5</td>
<td>2.1</td>
<td>2.5</td>
<td>2.6</td>
</tr>
<tr>
<td>Improvement (hf vs. MC) (%)</td>
<td>76.1</td>
<td>71.0</td>
<td>66.1</td>
<td>73.1</td>
<td>69.6</td>
<td>63.8</td>
<td>63.8</td>
<td>68.3</td>
</tr>
<tr>
<td>$\sigma_{\delta^{15}N_{hf}}$ (permeg)</td>
<td>3.5</td>
<td>4.6</td>
<td>5.4</td>
<td>4.0</td>
<td>4.3</td>
<td>2.7</td>
<td>3.1</td>
<td>3.3</td>
</tr>
<tr>
<td>$D_{\delta^{15}N_{corr}}$ (permeg)</td>
<td>1.7</td>
<td>2.1</td>
<td>2.6</td>
<td>1.9</td>
<td>2.0</td>
<td>1.2</td>
<td>1.5</td>
<td>1.6</td>
</tr>
<tr>
<td>Improvement (corr vs. hf) (%)</td>
<td>37.0</td>
<td>41.7</td>
<td>39.5</td>
<td>40.6</td>
<td>42.9</td>
<td>42.9</td>
<td>48.0</td>
<td>38.5</td>
</tr>
<tr>
<td>$\sigma_{\delta^{15}N_{corr}}$ (permeg)</td>
<td>2.2</td>
<td>2.7</td>
<td>3.3</td>
<td>2.4</td>
<td>2.5</td>
<td>1.5</td>
<td>1.7</td>
<td>1.9</td>
</tr>
<tr>
<td>$2\sigma_{\delta^{15}N_{corr.95}}$ (permeg)</td>
<td>4.4</td>
<td>5.3</td>
<td>6.3</td>
<td>4.7</td>
<td>4.9</td>
<td>3.0</td>
<td>3.4</td>
<td>3.7</td>
</tr>
<tr>
<td>$D_T_{hf}$ (K)</td>
<td>0.20</td>
<td>0.32</td>
<td>0.33</td>
<td>0.25</td>
<td>0.27</td>
<td>0.18</td>
<td>0.21</td>
<td>0.22</td>
</tr>
<tr>
<td>$\sigma_{T_{hf}}$ (K)</td>
<td>0.26</td>
<td>0.40</td>
<td>0.43</td>
<td>0.32</td>
<td>0.35</td>
<td>0.22</td>
<td>0.26</td>
<td>0.27</td>
</tr>
<tr>
<td>$D_T_{corr}$ (K)</td>
<td>0.12</td>
<td>0.18</td>
<td>0.20</td>
<td>0.14</td>
<td>0.15</td>
<td>0.10</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>$\sigma_{T_{corr}}$ (K)</td>
<td>0.15</td>
<td>0.24</td>
<td>0.25</td>
<td>0.19</td>
<td>0.19</td>
<td>0.12</td>
<td>0.14</td>
<td>0.15</td>
</tr>
<tr>
<td>$2\sigma_{T_{corr.95}}$ (K)</td>
<td>0.31</td>
<td>0.48</td>
<td>0.51</td>
<td>0.38</td>
<td>0.37</td>
<td>0.23</td>
<td>0.27</td>
<td>0.30</td>
</tr>
</tbody>
</table>
Figure 7. (a–e) First-guess (g,0) versus Monte Carlo (mc,fin) $\delta^{15}$N solution for the scenario S5: (a) synthetic target $\delta^{15}$N$_{syn}$ (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 mismatches of $D_{15N}$ for the first-guess solution (blue) and the Monte Carlo solution (red) versus the synthetic target. (c) Time series of the pointwise mismatches of $D_{15N}$ for the first-guess solution (blue) and the Monte Carlo solution (red) versus the synthetic target. (d–f) First-guess versus Monte Carlo surface-temperature solution $T_{surf}$ for the scenario S5: (d) synthetic surface-temperature target $T_{syn}$ (black dotted line), first-guess temperature input (blue line) and Monte Carlo solution (red line). (e) Histogram shows the pointwise temperature mismatches $D_T$ for the first-guess solution (blue) and the Monte Carlo solution (red) versus the synthetic surface-temperature target. (f) Time series of the pointwise temperature mismatches $D_T$ for the first-guess solution (blue) and the Monte Carlo solution (red) versus the synthetic surface-temperature target.

4 Discussion

4.1 Monte Carlo type input generator

Figure 9 shows the distribution of the COPs (I) and SD values (II) used to create the improvements (Sect. 2.1, step 2) for all scenarios. The cut-off periods are more or less evenly distributed, which shows that nearly all of the allowed frequency range (500 to 2000 years) was used to create the improvements during the iterations. In contrast, the distributions of the SD values show clearly that mostly small SD values are...
used to create the improvements, which implies that iterations with small perturbations will more likely lead to an improvement than larger ones.

Figure 6 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 6 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 realization 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 number 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

To investigate the timing and contributions of the remaining mismatches in $\delta^{15}$N and temperature for scenario S1 after the high-frequency (step 3) and correction parts (step 4), different cross-correlation experiments were conducted (see

Table 5. Final mismatches $\Delta$($\Delta$age) (2\sigma) of $\Delta$age between the corrected solution and the synthetic targets for all scenarios.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>$2\sigma$ $\Delta$(\Delta age) (yr)</th>
<th>Scenario</th>
<th>$2\sigma$ $\Delta$(\Delta age) (yr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>1.14</td>
<td>S5</td>
<td>1.24</td>
</tr>
<tr>
<td>S2</td>
<td>1.60</td>
<td>H1</td>
<td>1.23</td>
</tr>
<tr>
<td>S3</td>
<td>1.98</td>
<td>H2</td>
<td>1.18</td>
</tr>
<tr>
<td>S4</td>
<td>1.41</td>
<td>H3</td>
<td>1.30</td>
</tr>
</tbody>
</table>

Figure 9. (I) Counts of the COPs and (II) counts of the SD values used to create the improvements for the smooth temperature solutions of the Monte Carlo input generator for all synthetic scenarios (S1–S5 and H1–H3). A SD value of 0.1, for example, means that the maximum allowed perturbation of one temperature value $T(t_0)$ is $\pm$ 10%.

www.clim-past.net/14/763/2018/
Sect. S3 in the Supplement). The experiments led to equal results. The major fraction of the final mismatches of $\delta^{15}N$ emerges from mismatches in the thermal-diffusion component $D^{15}N_{\text{form}}$. Also a cancellation effect between the gravitational component $D^{15}N_{\text{grav}}$ and $D^{15}N_{\text{clear}}$ of the total mismatch in $\delta^{15}N$ became obvious, affecting the calculation of $\text{lag}_{\text{max}}^{15}N$ and $\text{lag}_{\text{min}}^{15}N$ and most likely leading to a fundamental residual uncertainty in the low-permeg level for the corrected $\delta^{15}N$ data. The same analyses were conducted for all synthetic scenarios, leading to similar results.

Additionally, the influence of the window length, used for the calculation of $\text{IF}(t)$, on the correction was analysed, showing that for all investigated window lengths the correction reduces the mismatches of $\delta^{15}N$ and temperature, whatever correction mode was used (calculated with $\text{xcf}_{\text{max}}$, $\text{xcf}_{\text{min}}$ or both quantities). Moreover, the correction is most efficient for window lengths in the range of 100 years to 300 years with an optimum at 200 years for all cases.

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

4.3.1 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 $\delta^{15}N_{\text{excess}}$ data as well using the same basic concepts (Fig. 10). Here, the $\delta^{40}Ar$ and $\delta^{15}N_{\text{excess}}$ data from Kobashi et al. (2008b) were used as the fitting targets. We reach final mismatches (2$\sigma$) of 4.0 permeg for $\delta^{40}Ar/4$ and 3.7 permeg for $\delta^{15}N_{\text{excess}}$, which are for both quantities below the analytical measurement uncertainty of 4.0 to 9.0 permeg for $\delta^{40}Ar/4$ and 5.0 to 9.8 permeg for $\delta^{15}N_{\text{excess}}$ measured data (Kobashi et al., 2008b).

The automated inversion of different gas-isotope quantities ($\delta^{15}N$, $\delta^{40}Ar$, $\delta^{15}N_{\text{excess}}$) 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 firm model. Next, the presented algorithm is not dependent on the firm model, which leads to the implication that the algorithm can be coupled to different firm models describing firm 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 between solutions matching different targets. Finally, differences obtained by applying different firm physics (densification equations, convective zone, etc.) but the very same inversion algorithm may help to assess firm-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, are completely described by the used firm model. Consequently, all sources of signal noise are ignored. For the later use of the algorithm on $\delta^{15}N$, $\delta^{40}Ar$ or $\delta^{15}N_{\text{excess}}$ 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 between temperature solutions obtained from individual targets ($\delta^{15}N$, $\delta^{40}Ar$, $\delta^{15}N_{\text{excess}}$) will become obvious. These differences will allow us 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.

4.3.2 Measurement uncertainty and firm heterogeneity (centimetre-scale variability)

Many studies have investigated the influence of firm heterogeneity (or density fluctuations) on measurements of air compounds and quantities (e.g. $\delta^{15}N$, $\delta^{40}Ar$, CH4, CO2, O2/N2 ratio, air content) extracted from ice cores resulting in centimetre-scale variability and leading to additional noise on the measured data (e.g. Capron et al., 2010; Etheridge et al., 1992; Fourtou et al., 2017; Fujita et al., 2009; Hörhold et al., 2011; Huber and Leuenberger, 2004; Rhodes et al., 2013, 2016). Using discrete measurement techniques 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 are calculated using pooled standard deviation (Hedges and Olkin, 1985). Often, it is not possible to take real replicates (same depth), and instead the replicates are taken from nearby depths. Hence, any potential centimetre-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 to 50 years of ice at approximately 35 kyr b2k. Increasing the depth resolution of the samples would increase our knowledge of centimetre-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 centimetre-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 uncer-
Figure 10. Fitting of GISP2 Holocene $\delta^{40}$Ar (a–d) and $\delta^{15}$N_{excess} (e–h) data (measurement data from Kobashi et al., 2008b): (a) measured versus modelled $\delta^{40}$Ar/4 time series. (b) Zoom-in for the same quantity as in panel (a). (c) Time series of the final mismatches $\Delta\delta^{40}$Ar/4 of the measured minus the modelled $\delta^{40}$Ar/4 data. (d) Histogram for the same quantity as in panel (c) showing an overall final mismatch ($\sigma$) of 4.0 permeg and offset (os) of $-0.5$ permeg. (e) Measured versus modelled $\delta^{15}$N_{excess} time series. (f) Zoom-in for the same quantity as in panel (e). (g) Time series of the final mismatches $\Delta\delta^{15}$N_{excess} of the measured minus the modelled $\delta^{15}$N_{excess} data. (h) Histogram for the same quantity as in panel (g) showing an overall final mismatch ($\sigma$) of 3.7 permeg and offset (os) of $-0.8$ permeg.

tainty in the inverse-modelling approach presented here. This will increase the uncertainty of the temperature according to Eq. (3). 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 30 km apart is likely one of the best ways to overcome potential centimetre-scale variability. A comparison of ice cores that were drilled even closer might be even more advantageous.

4.3.3 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. Grachev and Severinghaus, 2005; Spahni, 2003). 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 of the order of 10 years (Schwan-
der et al., 1997), whereas the data resolution of the synthetic targets was set to 20 years to mimic the measurement data from Kobashi et al. (2008b) with a mean data resolution of about 17 years (see Sect. 2.3: “Generating synthetic target data”). In the study of Kindler et al. (2014), it was shown that a glacial Greenland lock-in depth leads to a damping of the $\delta^{15}$N signal of about 30% for a 10 K temperature rise in 20 years. 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.

4.3.4 Accumulation-rate uncertainties

For the synthetic data study presented in this paper, it is assumed that the used accumulation-rate data are 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 Sect. S1 of the Supplement, this will not be the case anymore. Uncertainties in layer counting and corrections for ice thinning will lead to a fundamental uncertainty. Especially in the early Holocene, this can easily exceed 10%. As the accumulation-rate data are used to run the firm model, all potential accumulation uncertainties are in part incorporated into the temperature reconstruction. On the other hand, as we discussed in Sect. S2 of the Supplement, the accumulation-rate variability has a minor impact compared to the input temperature on the variability of $\delta^{15}$N data in the Holocene. 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. A total of 30% corresponds to the 8.2 kyr event and 12% to 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 Sect. S1 of the Supplement. 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.

4.3.5 Convective zone variability

Many studies have shown the existence of a well-mixed zone at the top of the diffusive firm column, called the convective zone (CZ). The CZ is formed by strong katabatic winds and pressure gradients between the surface and the firm (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 firm models for temperature reconstructions based on stable isotopes of air, the Schwander et al. (1997) model which has no CZ built in (or better – a constant CZ of 0 m) and the Goujon firm 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 firm 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, except a constant temperature offset. Also noteworthy is that there is no firm 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 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 metres, 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 us-
ing $\delta^{18}$O_{ice} information is much more challenging due to the highly variable gas-age – ice-age differences ($\Delta$age) between stadial and interstadial conditions. Here, contrary to the rather stable Holocene period, the $\Delta$age can vary by several hundreds of years. Also, the accumulation-rate data are 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 NGRIIP 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. 2.1.2) 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 timescale (NGRIP members, 2004; Johnsen et al., 2001) as it was done in the Kindler et al. (2014) 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 to 60 kyr. The reconstruction window (containing DO6 and DO7) is set to 32 to 36.2 kyr. As the first-guess (starting point) of the reconstruction, we use the accumulation-rate data ($\delta^{18}$O_{ice}) for NGRIIP from the ss09sea06bm timescale together with a constant temperature of about $-49^\circ$C for this time window. As minimization criterion $D_g$ for the reconstruction, we simply use the sum of the root mean squared errors of the $\delta^{15}$N and $\Delta$age 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. 1).

$$D_g = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left( \delta^{15}\text{N}_{\text{meas},i} - \delta^{15}\text{N}_{\text{mod},i} \right)^2}$$

Here, $\delta^{15}$N_{meas,i} and $\delta^{15}$N_{mod,i} are the uncertainties in $\delta^{15}$N and $\Delta$age for the measured values $i$ or $k$ ($\Delta$age match points: Guillévici, 2013, p. 65, Table 3.2) and $N, M$ the number of measurement values. We set $\delta^{15}$N_{meas,i} = 20 per mil for all $i$ (Kindler et al., 2014) and $\delta^{15}$N_{mod,i} = 50 years for all $k$. The relative uncertainties in $\Delta$age can easily reach up to 50% and more in the glacial period using the ss09sea06bm timescale which results in a pre-eminence of the $\delta^{15}$N misfits over the $\Delta$age misfits (10 to 20% when using GICC05 timescale; Guillévici, 2013, p. 65 Table 3.2). Due to this issue, we have to set $\delta^{15}$N_{meas,i} to 50 years to make both terms equally important for the fitting algorithm. In Fig. 11, we show preliminary results. The $\delta^{15}$N and $\Delta$age fitting (Fig. 11a, b) and the resulting gained temperature and accumulation-rate solutions (Fig. 11c, 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 $\gamma$ direction) of the accumulation-rate signal over the whole time window (32 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 $\delta^{15}$N fit (MCS1). 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 $\delta^{15}$N and $\Delta$age 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 Table 6). 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 suitable fit of the $\delta^{15}$N and $\Delta$age target data, a result highly similar to Guillévici et al. (2013) and Kindler et al. (2014). The mismatches in $\delta^{15}$N and $\Delta$age of the final (MCS PIN) solution show a 15% smaller misfit of $\delta^{15}$N ($2\sigma$) and an about 31% smaller misfit of $\Delta$age ($2\sigma$) 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 firm-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 firm 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 is fully automated and provides 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 δ^{15}N, which themselves are mainly driven by high-frequency temperature variations. To evaluate the performances of the presented approach, eight different synthetic δ^{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 δ^{15}N and temperature data. The obtained final mismatches follow a sym-
Table 6. Proof of concept for glacial reconstruction: $D_{gl}$ is the used minimization criterion (see Sect. 4.4).

<table>
<thead>
<tr>
<th>Solution</th>
<th>$D_{gl}$</th>
<th>Mismatch $\delta^{15}N$ (2σ) (permeg)</th>
<th>Mismatch $\delta^{15}N^*$ (2σ) (permeg)</th>
<th>Mismatch $\Delta age$ (2σ) (yr)</th>
<th>Mismatch $\Delta age^*$ (2σ) (yr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kindler et al. (2014)</td>
<td>3.6</td>
<td>44.5</td>
<td>17.9</td>
<td>256</td>
<td>101</td>
</tr>
<tr>
<td>First guess</td>
<td>7.8</td>
<td>128.7</td>
<td>63.8</td>
<td>328</td>
<td>138</td>
</tr>
<tr>
<td>MCS0</td>
<td>3.1</td>
<td>50.0</td>
<td>19.3</td>
<td>199</td>
<td>82</td>
</tr>
<tr>
<td>MCS1</td>
<td>2.9</td>
<td>44.3</td>
<td>17.6</td>
<td>200</td>
<td>84</td>
</tr>
<tr>
<td>MCS FIN</td>
<td>2.6</td>
<td>37.8</td>
<td>15.6</td>
<td>175</td>
<td>63</td>
</tr>
</tbody>
</table>

* The mean mismatches for $\delta^{15}N$ and $\Delta age$ were calculated according to Eq. (1).

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

Edited by: Emily Dearing Crampton Flood
Reviewed by: three anonymous referees

References


Dahl-Jensen, D., Mosegaard, K., Gundestrup, N., Clow, G. D., Johnsen, S. J., Hansen, A. W., and Balling, N.: Past Temperatures


Döring, M. and Leuenberger, M. C.: Synthetic and fitted d15N and temperature data and GISP2 accumulation rates (13.5–52 497.5 yr b2k) on GICC05 time scale, PANGAEA, https://doi.org/10.1594/PANGAEA.888997, 2018.


Guillevic, M.: Characterisation of rapid climate changes through isotope analyses of ice and entrapped air in the NEEM ice core, PhD thesis, University of Copenhagen, Copenhagen, Denmark, Université de Versailles Saint Quentin en Yvelines, Versailles, France, 2013.


