Climate
of the Past

Comprehensive uncertainty estimation of the timing of Greenland warmings in the Greenland ice core records

Eirik Myrvoll-Nilsen[1], **Keno Riechers**[1,2], **Martin Wibe Rypdal**[3], **and Niklas Boers**[1,2,4]

[1]Potsdam Institute for Climate Impact Research, Potsdam, Germany
[2]Earth System Modelling, School of Engineering & Design, Technical University of Munich, Munich, Germany
[3]Department of Mathematics and Statistics, The University of Tromsø – The Arctic University of Norway, Tromsø, Norway
[4]Department of Mathematics, Global Systems Institute, University of Exeter, Exeter, UK

Correspondence: Eirik Myrvoll-Nilsen (myrvoll@pik-potsdam.de) and Niklas Boers (n.boers@tum.de)

Abstract. Paleoclimate proxy records have non-negligible uncertainties that arise from both the proxy measurement and the dating processes. Knowledge of the dating uncertainties is important for a rigorous propagation to further analyses, for example, for identification and dating of stadial–interstadial transitions in Greenland ice core records during glacial intervals, for comparing the variability in different proxy archives, and for model-data comparisons in general. In this study we develop a statistical framework to quantify and propagate dating uncertainties in layer counted proxy archives using the example of the Greenland Ice Core Chronology 2005 (GICC05). We express the number of layers per depth interval as the sum of a structured component that represents both underlying physical processes and biases in layer counting, described by a regression model, and a noise component that represents the fluctuations of the underlying physical processes, as well as unbiased counting errors. The joint dating uncertainties for all depths can then be described by a multivariate Gaussian process from which the chronology (such as the GICC05) can be sampled. We show how the effect of a potential counting bias can be incorporated in our framework. Furthermore we present refined estimates of the occurrence times of Dansgaard–Oeschger events evidenced in Greenland ice cores together with a complete uncertainty quantification of these timings.

1 Introduction

The study of past climates is based on proxy measurements obtained from natural climate archives such as cave speleothems, lake and ocean sediments, and ice cores. Paleoclimate reconstructions derived from proxies suffer from 3-fold uncertainty. First, the proxy measurement itself involves the typical measurement uncertainties. Second, the interpretation of proxy variables, such as isotope ratios in terms of physical variables, such as temperature, is often ambiguous, and typically no one-to-one mapping can be established between the measured proxies and the climatic quantities of interest. Third, the age has to be measured alongside the proxy variable. In most cases an age model can be inferred that provides a quantitative relationship between the depth in the archive under consideration and the corresponding age. Such age models are also subject to uncertainties.

This study is exclusively concerned with the dating uncertainties of so-called layer counted archives, where the dating is performed based on counting periodic signals in the proxy archives such as annual layers arising from the impact of the seasonal cycle on the deposition process (e.g., Rasmussen et al., 2006). This type of archive comprises varved lake sediments, ice cores, banded corals, tree rings, and some speleothems (Comboul et al., 2014). Using the example of the NGRIP ice core (North Greenland Ice Core Project members, 2004) and its associated chronology, the GICC05 (Vinther et al., 2006; Rasmussen et al., 2006; Andersen et al., 2006; Svensson et al., 2008), we present here a statistical approach to generate ensembles of age models

that may in turn be used to propagate the age uncertainties to any subsequent analysis of the time series derived from the NGRIP record. Our method can be directly adapted to other layer counted archives.

Layer counting assesses the age increments along the axis perpendicular to the layering, whose summation yields the total age. In turn, also the errors made in the counting process accumulate such that in chronologies obtained from counting annual layers, the absolute age uncertainty grows with increasing age (e.g., Boers et al., 2017).

Most importantly, dating uncertainties make it challenging to establish an unambiguous temporal relation between signals recorded in different, and possibly remote, archives. Therefore, it is often not possible to decipher the exact temporal order of events and distinguish causes from consequences across past climate changes. For example, abrupt Greenland warmings known as Dansgaard–Oeschger (DO) events (Dansgaard et al., 1993; Johnsen et al., 1992) evidenced in ice core records from the last glacial are accompanied by changes in the east Asian monsoon system, which are apparent from Chinese speleothem records (e.g., Zhou et al., 2014; Li et al., 2017). However, since the dating uncertainties exceed the relevant time scales of these abrupt climate shifts, a clear order of events cannot be determined. This prevents us from deducing if in the context of DO events, the abrupt Greenland warming triggered a hemispheric transition in the atmosphere, or vice versa, or if these changes happened simultaneously as part of a global abrupt climatic shift (Corrick et al., 2020).

For the quantification of dating uncertainties in radiometrically dated archives, there exist well established generalized frameworks. One example is the Bayesian Accumulation Model (Blaauw and Christeny, 2011) which models the sediment accumulation rate as a first order autoregressive process with gamma distributed innovations. Other methods or software include OxCal (Ramsey, 1995, 2008) and BChron (Haslett and Parnell, 2008; Parnell et al., 2008). Contrarily, the uncertainties of layer counted archives are targeted systematically only by few studies. Comboul et al. (2014) present a probabilistic model, where the number of missed and double-counted layers are expressed as counting processes shaped by corresponding error rates. However, this approach requires knowledge about these rates and further does not account for any uncertainty associated with them. An alternative Bayesian approach for quantifying the dating uncertainty of layer counted archives is presented in Boers et al. (2017), where the uncertainty is shifted from the time axis to the proxy value. However, this approach does not allow for generation of ensembles of chronologies as required for uncertainty propagation.

Even though dating uncertainties are conveniently quantified for many archives, many studies ignore these uncertainties and instead draw inference from "average" or "most likely" age scales, as already highlighted by McKay et al. (2021). This involves the risk of losing valuable informa-

tion, as shown, for example, in Riechers and Boers (2020). In some cases, rigorous propagation of uncertainty may yield results that qualitatively differ from results obtained by using the "average" or "best fit" age model. In this context, McKay et al. (2021) propose to apply the respective analysis to an ensemble of possible age scales to ensure the uncertainty propagation, in line with the strategy proposed by Riechers and Boers (2020).

We focus on the layer counted part of the GICC05 chronology, a synchronized age scale for several Greenland ice cores (Vinther et al., 2006; Rasmussen et al., 2006; Andersen et al., 2006; Svensson et al., 2008). It was obtained by counting the layers of different Greenland ice cores and synchronizing the results using matchpoints. While the recent part of the chronology is compiled from multiple cores, the older part (older than 15 kyr b2k) is based exclusively on the layer counting in the NGRIP core. We introduce a new method to generate realistic age ensembles for the NGRIP core, which conveniently represent the uncertainty associated with the GICC05.

Originally, the dating uncertainty of the GICC05 is quantified in terms of the maximum counting error (MCE). The MCE increases by 0.5 years for every layer which is deemed uncertain by the investigators during the counting process:

$$\text{MCE}(z) = 0.5 N_{\text{u}}(z), \tag{1}$$

where N_{u} is the number of uncertain layers down to depth z. In contrast to certain layers, which can be identified unambiguously in the records, uncertain layers are less pronounced and therefore it seems less certain that these signals truly correspond to physical layers. The accumulation of uncertain layers results in high values for the MCE for the older parts of the core (MCE = 2.6 kyr at 60 kyr b2k estimated age). However, it seems highly unlikely that all uncertain layers are consistently either true layers or no layers, which is why we think that the MCE is an overly careful quantification of the age uncertainty, as already suggested by Andersen et al. (2006). One might, alternatively, be tempted to treat the uncertain layers as a Bernoulli experiment with N_{u} repetitions and a probability of one half for each uncertain layer to be a true layer. However, this would neglect any sort of bias in the assessment of the uncertain layers and would lead to unrealistically small uncertainties, since over- and undercounting practically cancel each other out in this Bernoulli type interpretation (see for instance Andersen et al., 2006; Rasmussen et al., 2006).

The method presented here abandons the notion of certain and uncertain layers. Instead, we separate the GICC05 chronology into contributions that can be captured by deterministic model equations and corresponding residuals. We construct a new age–depth model by complementing the deterministic part with a stochastic component designed in accordance with the statistics of the residuals. This model can be used to generate age–depth ensembles in a computationally efficient manner. In turn, these ensembles facilitate un-

certainty propagation to subsequent analysis. The model parameters are tuned with respect to the GICC05 chronology that includes every uncertain layer as half a layer.

The outline of this paper is as follows: Section 2 gives a description of the data used for this study. Section 3.1 introduces our statistical model for the dating uncertainties, and details about how we incorporate physical processes and how we deduce the noise of the model from the statistics of the residuals. In Sect. 3.2 we show how we can formulate our model in terms of a hierarchical Bayesian modeling framework that allows for the physical and noise components to be estimated simultaneously. That section also details how one can use the resulting posterior distributions of the model parameters to obtain a full description of the posterior distributions of the dating uncertainties using a sample-based approach. We finally demonstrate how a potential counting bias could be incorporated by the model, and how that would affect the results. In Sect. 4 we show how our model can be used to obtain a full description of the dating uncertainties of abrupt warming events, which takes into account the dating uncertainties as well as the uncertainties in determining their exact position in the noisy data. Further discussion and conclusions are provided in Sect. 6.

2   NGRIP ice core data

We use the Greenland Ice Core Chronology 2005 (GICC05) (Vinther et al., 2006; Rasmussen et al., 2006; Andersen et al., 2006; Svensson et al., 2008) as defined for the NGRIP ice core together with the corresponding $\delta^{18}O$ proxy record (North Greenland Ice Core Project members, 2004; Gkinis et al., 2014; Ruth et al., 2003). An analogous analysis for the Ca^{2+} proxy record is presented in Appendix B. The final age of the layer counted part of the GICC05 is 59 944 yr b2k, and we consider data up to 11 703 yr b2k. The following Holocene part of the record is excluded since it is governed by a substantially different climate than the last glacial interval (Rasmussen et al., 2014). For the considered period, the NGRIP record is available at 5 cm resolution and thus equidistant in depth, but not in time. In total, the data comprises $n = 18\,672$ data points of the form $(z_k, y_k, x_k), k \in \{0, 1, \dots, n-1\}$, where z_k denotes the kth depth, y_k the corresponding age as indicated by the GICC05, and x_k the measured proxy value.

The GICC05 is based on counting annual layers which are evident in multi-proxy continuous flow measurements from the NGRIP, DYE3, and the GRIP ice cores. While the measurements from DYE3 and GRIP only facilitate layer counting up to ages of 8.2 and 14.9 kyr b2k, respectively, the NGRIP core allowed the identification of annual layers up to an age of 60 kyr b2k. The uncertainty of the GICC05 has been quantified as follows: whenever the investigators were uncertain about whether or not a signal in the data should be considered an annual layer, half a year was added to the

cumulative number of layers while simultaneously adding ± 0.5 yr to the age uncertainty. The total age uncertainty determined by the number of all uncertain layers up to a given depth is termed the maximum counting error (MCE). The MCE amounts to a relative age uncertainty of 0.84 % at the onset of the Holocene and 4.34 % at the end of the layer counted section of the core.

$\delta^{18}O$ values from Greenland ice cores are interpreted as a qualitative measure of the site temperature at the time of precipitation (Jouzel et al., 1997; Gkinis et al., 2014). We include these data in our study since our modeling approach will make use of the relation between atmospheric temperatures and the amount of precipitation, which in turn affects the thickness of the annual layers. In addition we use the division of the record into Greenland stadial and interstadial phases as presented in Table 2 of Rasmussen et al. (2014). We label the depths at which stadial–interstadial transitions occur by $z_1^*, \dots, z_p^*$ and the corresponding ages by $y_1^*, \dots, y_p^*$. Figure 1 shows the measured $\delta^{18}O$ values as a function of the GICC05 time scale, together with the Greenland stadial and interstadial onsets.

3   Methods

3.1   Age–depth model

We assume that depths $z = (z_1, \dots, z_n)^\top$ and proxy values $x = (x_1, \dots, x_n)^\top$ are measured accurately and hence treat them as deterministic variables. In contrast, we consider the ages $y = (y_1, \dots, y_n)^\top$ as dependent stochastic variables and will in the following establish a model to map the independent depths and stable isotope concentrations onto ages, in a way that reflects the uncertainties inherent to the dating. The model will be supplemented with information on the prevailing climate period.

In order to motivate our modeling approach we give some general considerations about the deposition process as well as the counting process. The decisive quantity for us will be the incremental number of annual layers counted in a 5 cm depth increment of the ice core:

$$\Delta y_k = y_k - y_{k-1}. \tag{2}$$

This quantity is determined by the amount of precipitation (minus the snow that is blown away by winds) during the corresponding period, the thinning that the layers experience over time deeper down in the core, and potential errors made during the counting process. While the thinning process can be expected to happen mostly deterministically, the net annual accumulation of snow certainly exhibits stronger fluctuations. Finally, the counting error adds additional randomness. Thus, it is reasonable to regard the observed age increments Δy as a realization of a random vector ΔY which can be decomposed into a deterministic and a stochastic compo-
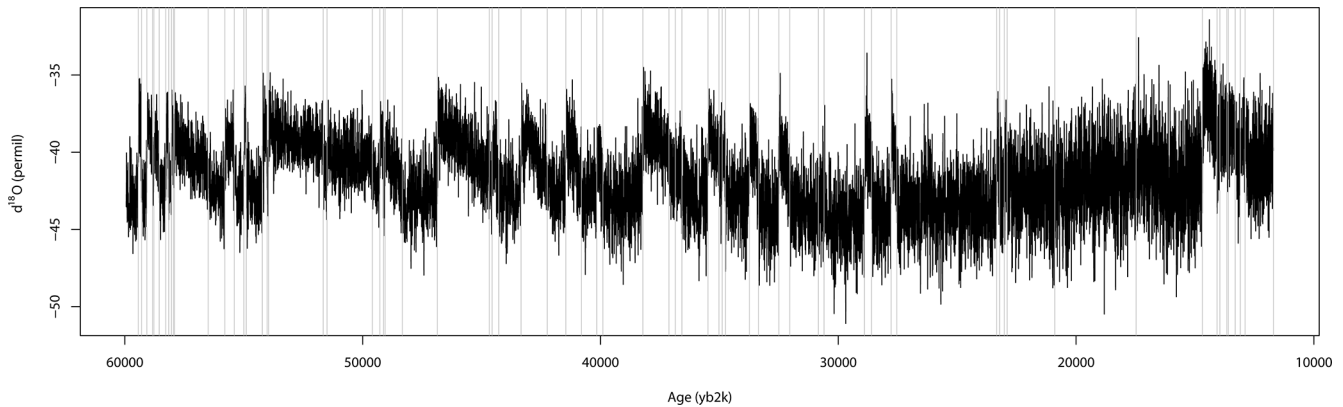
Figure 1. The measured δ^{18}O isotope values plotted against the corresponding GICC05 time scale, starting from 11 793 yr b2k. The vertical gray lines denote the transitions between Greenland stadial and interstadial periods as reported by Rasmussen et al. (2014).

nent:

$$\Delta Y_k = a(z_k) + \varepsilon_k. \tag{3}$$

Note that the number of layers within a 5 cm depth increment is not necessarily an integer number. Given that the amount of precipitation co-varies with atmospheric site temperatures we can specify

$$a(z_k) = a(z_k, x(z_k)) \rightarrow a(z_k, x_k). \tag{4}$$

Based on physical arguments and the analysis of the observed age increments Δy_k, we will propose the structural form of the deterministic part of the model and then tune the model parameters to the data. In turn, this will allow us to design the model's noise component ε in accordance with the corresponding residuals $\delta_k = \Delta y_k - a(z_k, x_k)$.

3.1.1 Linear regression

As explained above, the thickness of the counted layers, and thereby the number of layers per depth increment $\Delta z_k = z_k - z_{k-1}$, is governed by two physical factors: the amount of precipitation at the time the layer was formed and the thinning of the core due to ice flow. These processes are here assumed to follow a regression model. We take into account the thinning by implementing a second order polynomial dependency of ΔY_k with respect to the depth z_k. Choosing this non-linear function conveniently accounts for the saturation of the layer thinning evident in the NGRIP ice core. The amount of precipitation is known to co-vary with atmospheric temperatures, since by the Clausius–Clapeyron relation the moisture holding capacity of the atmosphere increases with temperatures. This is represented using a linear response to the δ^{18}O measurements. The same response is applied to the $\log(Ca^{2+})$ in the alternative analysis presented in Appendix B. Finally, we observe clear trends in the incremental layers that persist over individual stadials and interstadials. Given the consistency of these trends across the different climate periods, we decided to incorporate them into the deterministic model component. Overall, we propose a deterministic model of the form

$$a(z_k, x_k) = bz_k^2 + b_x x_k(z_k) + \sum_{i=1}^{p} \psi_i(z_k; a_i, c_i), \tag{5}$$

with

$$\psi_i(z_k; a_i, c_i) = \begin{cases} a_i + c_i z_k, & z_i^* < z_k < z_{i+1}^* \\ 0, & \text{otherwise} \end{cases}, \tag{6}$$

in order to capture the systematic features of the chronology. Here, c_i denotes the period specific slopes and a_i their corresponding offsets. For p transitions between stadials and interstadials we have to tune $2p + 2$ regression parameters, which is achieved by fitting the above model for $a(z_k, x_k)$ to the observed layer increments Δy_k given by the GICC05 time scale in a least squares approach. As explained above, the GICC05 ages contain the contribution of uncertain layers, which were counted as half a year each. Here, we abandon the distinction of certain and uncertain layers and regard the GICC05 ages as the best possible estimate of the true ages and accordingly use them directly for the optimization. The fitted model is shown in red in Fig. 22a.

3.1.2 Noise structure

After tuning the deterministic part of Eq. (3), the residuals are given by

$$\delta_k = a(z_k, x_k) - \Delta y_k. \tag{7}$$

We find the residuals to be symmetric and unimodally distributed, and apart from some degree of over-dispersion they appear to be well described by a normal distribution, as shown in Fig. 3.

Moreover, by examining the empirical autocorrelation illustrated in Fig. 3c, we observe that the residuals exhibit a fast decay of memory which is indicative of stationarity.
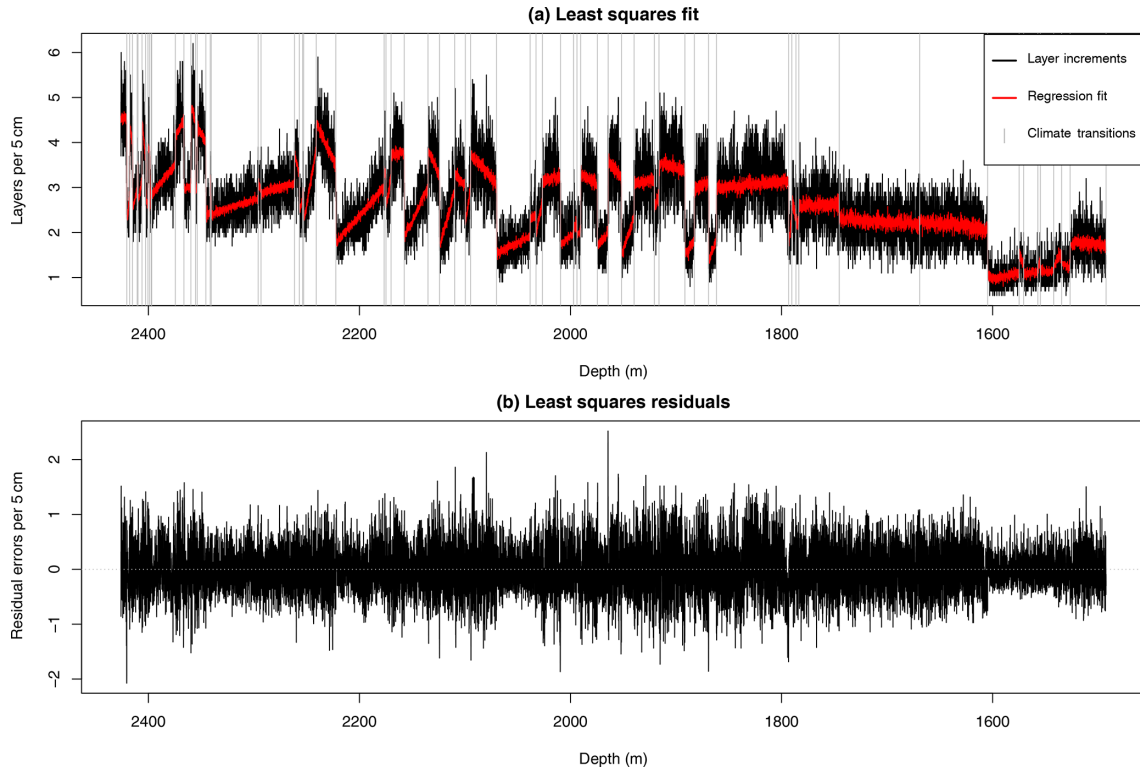
Figure 2. (a) Number of layers counted in the GICC05 time scale per 5 cm depth increments in the NGRIP ice core (black). The red line shows the fitted values from the regression model $a(z_k, x_k) = bz_k^2 + b_x x(k) + \sum_{i=1}^{p} \psi_i(z_k; a_i, c_i)$. The vertical gray lines represent the transitions between Greenland stadials and interstadials. **(b)** The residuals δ_k obtained from fitting the regression model $a(z_k, x_k)$ to the layer increments Δy_k.

This suggests that the noise can be expressed using a short-memory Gaussian stochastic process.

We explore three different models for the correlation structure of the noise ε. The first model assumes that they follow independent and identically distributed (iid) Gaussian processes:

$$\varepsilon_k \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\varepsilon^2\right). \tag{8}$$

The second model assumes that the noise can be described by a first order autoregressive (AR) process

$$\varepsilon_k = \phi\varepsilon_{k-1} + \xi_k, \tag{9}$$

where ϕ is the first-lag autocorrelation coefficient and ξ_k is a white noise process with variance $\sigma_\xi^2 = \sigma_\varepsilon^2/(1 - \phi^2)$. The third model assumes the noise follows a second order autoregressive process

$$\varepsilon_k = \phi_1\varepsilon_{k-1} + \phi_2\varepsilon_{k-2} + \xi_k, \tag{10}$$

where ϕ_1 and ϕ_2 are the first- and second-lag autocorrelation coefficients and ξ_k is a white noise process with variance:

$$\sigma_\xi^2 = \sigma_\varepsilon^2 \frac{1 - \phi_2}{(1 + \phi_2)\left((1 - \phi_2)^2 - \phi_1^2\right)}. \tag{11}$$

Note that a potential global or at least climate-regime-specific bias in the counting process, such as overseeing systematically 1 out of 10 layers, would be captured by the regression model and thus cannot be identified as a systematic error. We will investigate the influence of potential systematic errors below. Similarly, fluctuations in the physical processes can be captured by the noise model which aims to represent the counting errors. It would therefore be more accurate to interpret $a(z_k, x_k)$ as a structured component representing a part of both the physical processes and a systematic counting error which can be accounted for by linear regression, and ε_k as the fluctuations of both the physical processes and counting errors. Hence, ε_k can be considered an upper boundary on the counting uncertainty.

3.2 Simultaneous Bayesian modeling

So far, we have fitted only the structural model component. This enabled us to investigate statistical properties of the residuals and formulate corresponding noise model candidates. Fitting both the linear regression model and the different noise models can in principle be performed in two stages: First, the linear regression model is fitted to the layer increments using the method of least squares. Thereafter, the
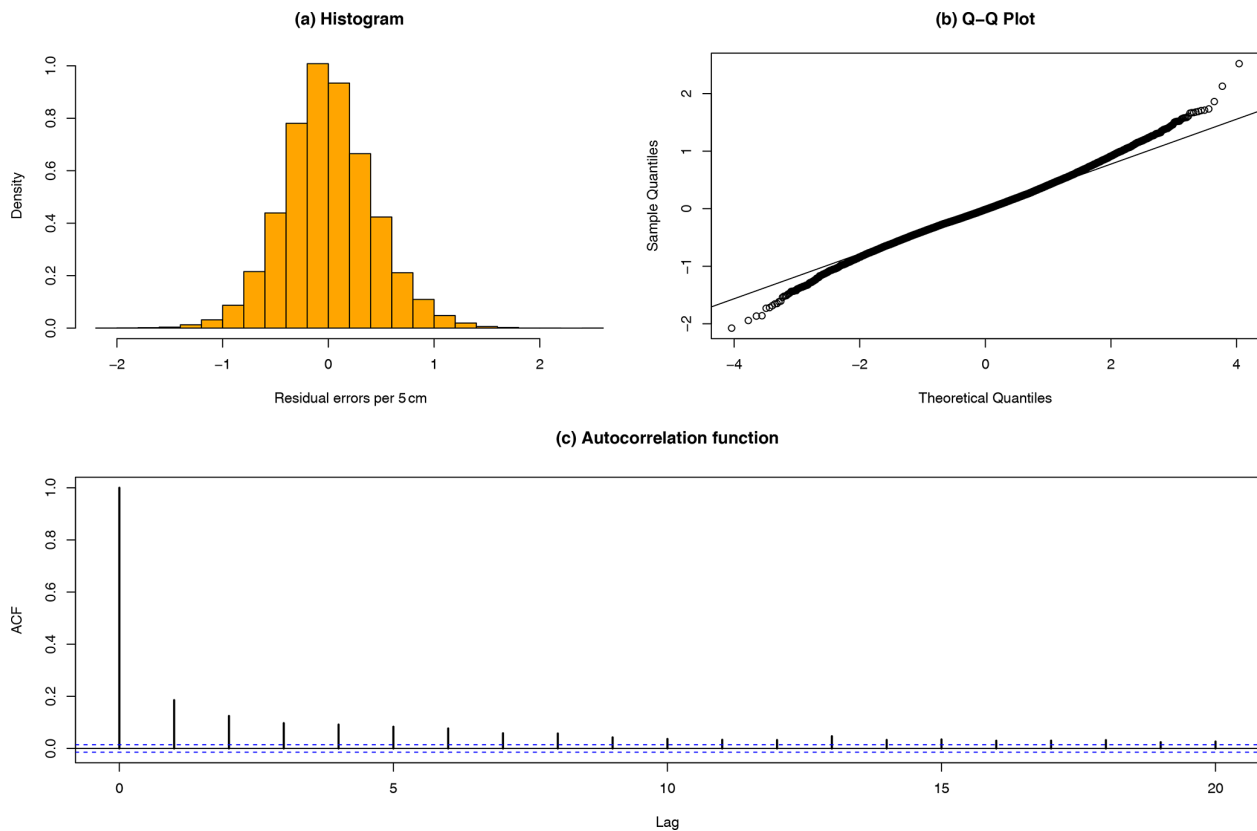
Figure 3. (a) Histogram of the residuals obtained from least squares fit of the regression model $a(z_k, x_k)$. **(b)** The corresponding quantile–quantile plot. We observed key properties of symmetry and unimodality indicating Gaussianity. The quantile–quantile plot indicates a small over-dispersion, but overall the data seem to be consistent with a normal distribution. **(c)** The autocorrelation function of the residuals from the least squares fit of the linear regression model, up to a maximum of 20 lags. A fast memory decay can be inferred.

fitted values are subtracted and the selected noise model is fitted to the residuals. However, this approach has the disadvantage that some variation that may in reality be caused by the noise process ε_k may have been attributed to the structured component $a(z_k, x_k)$ and removed before fitting the noise model. We therefore introduce here a Bayesian approach that enables us to estimate all model parameters simultaneously. The Bayesian approach has three key advantages over the least squares fitting of the structured component: first, it treats the noise and the structured component equally. Second, it returns the joint posterior probability of all model parameters which indicates the plausibility of a certain parameter configuration in view of the data. The posterior probability distribution can be regarded as an uncertainty quantification of the model's parameter configuration. Third, in the Bayesian parameter estimation, prior knowledge and constraints on the parameter can be incorporated via a convenient choice of the so-called prior distributions.

In general terms, let $\mathcal{D}$ denote some observational data and $\boldsymbol{\theta}$ denote parameters that shape a model which is assumed to reasonably describe the process that generated the data. Then Bayes' theorem can be used to deduce the posterior probability density of the parameters $\boldsymbol{\theta}$ given the data $\mathcal{D}$:

$$\pi(\boldsymbol{\theta}|\mathcal{D}) = \frac{\pi(\mathcal{D}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\mathcal{D})}. \tag{12}$$

In our case the GICC05 age $\boldsymbol{y} = (y_1, \ldots, y_n)$, or more precisely their increments $\Delta \boldsymbol{y}$, represent the observational data $\mathcal{D}$ assumed to be generated from the model defined by Eq. (3). There are $2p + 2$ parameters $\boldsymbol{\beta} = (b_2, b_x, a_1, c_1, \ldots, a_p, c_p)$ for the structured component alone and the noise adds another one to three parameters, depending on the choice of the noise structure. Thus the set of model parameters reads

$$\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\psi}), \tag{13}$$

where $\boldsymbol{\psi} = \sigma_\varepsilon$ if the residuals are assumed to follow an iid Gaussian distribution, $\boldsymbol{\psi} = (\sigma_\varepsilon, \phi)$ if they are assumed to follow an AR(1) process, and $\boldsymbol{\psi} = (\sigma_\varepsilon, \phi_1, \phi_2)$ if they are assumed to follow an AR(2) process. For any given parameter configuration, the likelihood is for all three choices of the

noise structure defined by a multivariate Gaussian distribution,

$$\pi(\Delta y | \boldsymbol{\theta}) = (2\pi)^{-n/2} |\boldsymbol{\Sigma}|^{-1} \exp\left\{-\frac{1}{2}(\Delta y \right.$$
$$\left. -\boldsymbol{a})^\top \boldsymbol{\Sigma}^{-1}(\Delta y - \boldsymbol{a})\right\}, \tag{14}$$

where $\boldsymbol{a} = (a(z_1, x_1), \ldots, a(z_n, x_n))^\top$ and the entries of the autocovariance matrix $\boldsymbol{\Sigma}$ are given by the autocovariance function $\Sigma_{ij} = \gamma(|i - j|)$ of the assumed noise model. For the iid model, the autocovariance function is simply σ_ε^2 if $i = j$, and zero otherwise, resulting in a diagonal covariance matrix. For the AR(1) model the autocovariance function is

$$\gamma(k) = \frac{\sigma_\varepsilon^2}{1 - \phi^2} \phi^{|k|}. \tag{15}$$

The autocovariance function of the AR(2) model is specified by the difference equation,

$$\gamma(k) = \phi_1 \gamma(k - 1) + \phi_2 \gamma(k - 2), \tag{16}$$

with initial conditions:

$$\gamma(0) = \left(\frac{1 - \phi_2}{1 + \phi_2}\right) \frac{\sigma_\varepsilon^2}{(1 - \phi_2)^2 - \phi_1^2} \tag{17}$$

$$\gamma(1) = \frac{\phi_1}{1 - \phi_2} \gamma(0). \tag{18}$$

A benefit of having the likelihood follow a Gaussian distribution is that it can be evaluated easily and samples can be obtained efficiently, despite the large number of parameters. Finally, we define convenient priors for the model parameters. For the parameters of the structured model component $\boldsymbol{\beta}$ we choose vague Gaussian priors, with variances that safely cover all reasonable parameter configurations. For the noise parameters $\boldsymbol{\psi}$ we restrict the scaling parameter σ_ε to be positive, and the autoregressive coefficients such that they define a stationary model. These constraints are embedded into the model by adopting suitable parameterizations. The scaling parameter σ_ε is assigned a gamma distribution through the parameterization $\kappa = \log(1/\sigma_\varepsilon^2)$. For the lag-one correlation parameter in the AR(1) model we assume a Gaussian prior on the logit transformation $\rho = \log((1 + \phi)/(1 - \phi))$. For the AR(2) model we instead assign priors on the logit transformation of the partial autocorrelations $\psi_1 = \phi_1/(1 - \phi_2)$ and $\psi_2 = \phi_2$, using penalized complexity priors (Simpson et al., 2017).

In principle, the joint posterior density can then be sampled by using a Markov chain Monte Carlo (MCMC) algorithm (e.g., Goodman and Weare, 2010). However, to solve Eq. (12) more efficiently, we formulate the problem in terms of a latent Gaussian model and then use integrated nested Laplace approximations (INLAs) (Rue et al., 2009, 2017) to compute the joint and marginal posterior distributions (for details see Appendix A).

The posterior distribution of $\boldsymbol{\theta}$ enables us to generate ensembles of different realizations of the random variable $\boldsymbol{Y}$, i.e., of the age increments that correspond to the fixed depth increments Δz. In a two-stage Monte Carlo simulation, first a value for $\boldsymbol{\theta}$ is randomly sampled from the posterior $\pi(\boldsymbol{\theta}|\Delta y)$. Second, the noise ε is sampled according to the noise model using noise parameters sampled in the first step. An ensemble generated in this fashion simultaneously reflects the uncertainty enshrined in the stochastic process and the model thereof as well as the uncertainty about the model parameters. Each realization of age increments yields a corresponding possible chronology according to

$$y_k = y_0 + \sum_{i=1}^{k} (a(z_i, x_i) + \varepsilon_i), \tag{19}$$

where y_0 is the number of reported layers up to the depth z_0. Figure 4 shows the 95 % credible intervals obtained from age ensembles for the three different noise models with respect to the GICC05 age. Each ensemble comprises 10 000 realizations of $\boldsymbol{Y}$. In this plot we notice a significant increase of uncertainty going from the iid to the AR(1) model. This is intuitive as when more memory is added to the model, the variation increases.

However, going from AR(1) to AR(2) adds only moderate additional uncertainty. We therefore argue that an AR(1) process is sufficient in terms of modeling the correlation structure of the residuals. The same is observed when using $\log(Ca^{2+})$ as a proxy instead (see Fig. B1). All computations henceforth are carried out with the AR(1) noise model.

3.3 Incorporating an unknown counting bias

When originally quantifying the uncertainty of the GICC05 chronology, a concern was that the layer counting was potentially biased in the sense that layers were consistently overcounted or missed. As highlighted by Andersen et al. (2006), there is no way to quantify a potential bias based on the data. Here, we investigate the influence that such a bias would have on our model, assuming a given maximum bias strength. To capture the effect of a systematic bias in the layer counting we introduce a scaling parameter η such that

$$\Delta Y_k = \eta(a(z_k, x_k) + \varepsilon_k). \tag{20}$$

Given that we have no knowledge about the size of the bias, η must be regarded as a random variable whose distribution can only be estimated by experts a priori. Originally, biases on the order of 1 % in the counting performed by different investigators have been observed (Rasmussen et al., 2006; Andersen et al., 2006). Here we assume
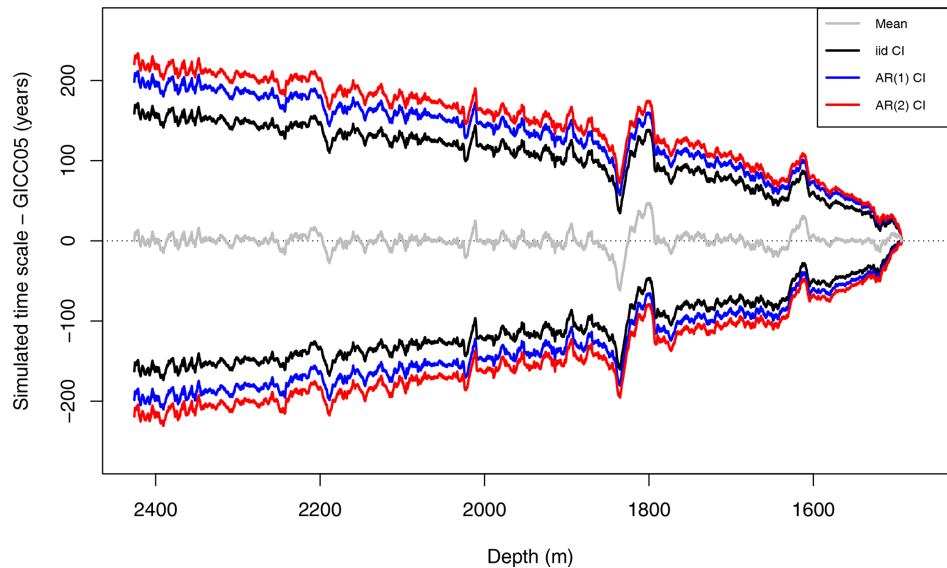
$$\eta \sim \mathcal{U}(1 \pm \Delta \eta), \tag{21}$$

Figure 4. The 95 % credible intervals of the dating uncertainty distribution when δ^{18}O is used as the proxy covariate. The GICC05 time scale has been subtracted and the noise is modeled using iid (black), AR(1) (blue), and AR(2) (red) noise models. Only the posterior marginal mean computed using AR(1) distributed noise is included (gray) since it is very similar to the mean obtained using other noise assumptions.

meaning the layer counters are just as likely to systematically over-count as to under-count, on a maximum rate of $\Delta\eta$. While the expectation for the age increments $E(\Delta Y)$ remains unchanged as long as $E(\eta) = 1$, their variance will grow due to the additional uncertainty. Figure 5 shows the 95 % credible intervals for potentially biased chronology ensembles generated under the assumption that $\Delta\eta$ equals 0 %, 2 %, or 4 %.

It is evident that the bias uncertainty contributes substantially to the age uncertainty. This is expected since a relatively small counting bias yields a large absolute error at a possible age of 60 kyr b2k, which in turn exceeds the uncertainty contribution of the noise by far. We also observed that one would need a maximum error rate of $\Delta\eta \approx 4$ % in order for the uncertainty to approach the maximum counting error at the end of the layer counted core segment.

4   Examples of applications

4.1   Dating uncertainty of DO events

Both the NGRIP δ^{18}O and log(Ca^{2+}) records are characterized by prominent abrupt shifts from low to high values, including the so-called Dansgaard–Oeschger (DO) events (Johnsen et al., 1992; Dansgaard et al., 1993). These jumps are interpreted as sudden warming events in Greenland, which took place repeatedly during the last glacial period. In order to explain the physical relationship between these abrupt Greenland warming events and apparently concomitant abrupt climate shifts evidenced in other archives from different parts of the planet, it is crucial to disentangle the exact temporal order of these events. This requires a rigorous

treatment of the uncertainties associated with the dating of DO events in Greenland ice core records.

Rasmussen et al. (2014) provide a comprehensive list of Dansgaard–Oeschger events and other stadial–interstadial transitions, indicating depths from the NGRIP ice core at which they occur, and the corresponding GICC05 age. They report the visually identified event onsets and provide uncertainty estimates in terms of data points along the depth axis and the respective MCE associated with the estimated event onset depth. This assessment was later refined by Capron et al. (2021) using the algorithm for detecting transition onsets designed by Erhardt et al. (2019). Here, we present a rigorous combination of the depth and age uncertainties, which complicate the exact dating of abrupt warming events.

First, we adopt the Bayesian transition onset detection designed by Erhardt et al. (2019) to estimate the onset of the abrupt warming transitions in the proxy records with respect to the depth in the core. By Z^* we denote a continuous stochastic variable that represents the uncertain onset depth and by x^* we denote a selected data window of the proxy record enclosing the transition. For each transition this yields a posterior distribution $\pi(Z^*|x^*)$ over potential transition onset depths, assuming a linear transition from low to high proxy values perturbed by AR(1) noise. For inference we adopt the methodology of INLA as it is particularly suited for such models, granting us a significant reduction in computational cost over traditional MCMC algorithms. The application of the transition onset detection to the onset of GI-11 is presented in Fig. 6a, with the resulting posterior marginal distribution for Z^* illustrated in Fig. 6b. Note that the dating method for the transitions is sensitive to the choice
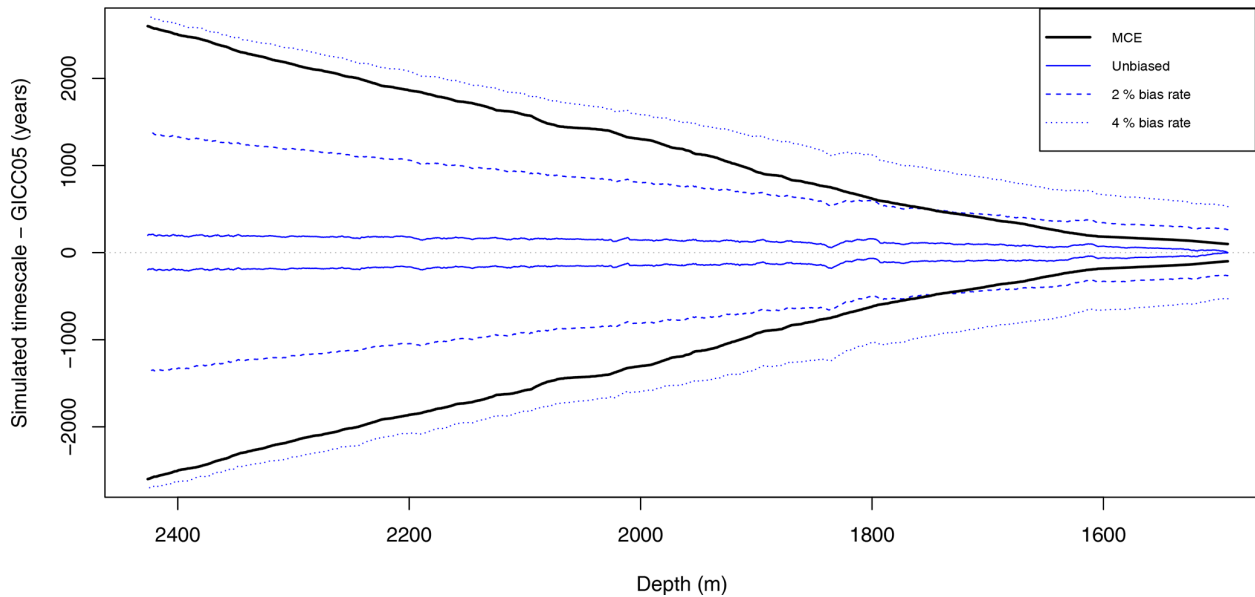
Figure 5. The 95 % credible intervals of the difference between estimated dating and the GICC05 time scale compared with the maximum counting error (solid black). Dating uncertainties in this case include a bias expressed by a stochastic scaling parameter drawn from a uniform $\mathcal{U}(1\pm\Delta\eta)$ distribution. The solid blue line represents the unbiased $\Delta\eta = 0$ dating uncertainty, while the dashed and dotted blue lines represent the biased cases of $\Delta\eta = 2\,\%$ and $\Delta\eta = 4\,\%$, respectively. The differences in dating uncertainty between iid, AR(1), and AR(2) models are dwarfed by the uncertainty introduced by the unknown bias; hence, only the AR(1) distributed residuals are shown here.

of the data window. In Appendix D we detail how the fitting windows for the linear ramps are optimally chosen. The selected data windows are listed in Table D1. Furthermore, the Bayesian transition detection fails in some cases where the transition amplitudes are small. We successfully derive posterior marginal distribution for a total of 29 events, whose summary statistics are listed in Table 1.

Each potential transition onset depth z^* yields a distribution over potential transition onset ages Y^*. This uncertainty, denoted by $\pi(Y^*|\Delta \boldsymbol{y}, z^*)$, is determined by linearly interpolating the age ensemble members generated according to Sect. 3.1 based on the observed layer increments $\Delta \boldsymbol{y}$. The posterior distribution for the transition onset date for a given DO event thus reads

$$\pi\left(Y^*|\Delta \boldsymbol{y}, \boldsymbol{x}^*\right) = \int \pi\left(Y^*|\Delta \boldsymbol{y}, Z^*\right) \pi\left(Z^*|\boldsymbol{x}^*\right) \mathrm{d}Z^*. \quad (22)$$

Technically, we create an ensemble of potential onset ages in the following way: First, we generate an ensemble of 10 000 samples from the posterior distribution of the transition onset depth,

$$z_r^* \sim \pi(Z^* \mid \boldsymbol{x}^*), \quad r \in [1, 10\,000],$$

to represent the onset depth uncertainty. Second, for each onset depth sample z_r^* we produce a simulation of a chronology from the corresponding age uncertainty:

$$y_r^*|z_r^* \sim \pi(Y^*|\Delta \boldsymbol{y}, z_r^*).$$

Thus, for both proxies ($\delta^{18}O$ and Ca^{2+}) and for each event we obtain 10 000 possible values for the transition onset age whose distribution corresponds to the posterior distribution expressed in Eq. (22). The posterior marginal mean and 95 % credible intervals for each event are reported in Table 1. This table hence gives the timing of the transitions together with the full uncertainties, stemming from the transition onset detection and the dating of the record. The estimated dating uncertainties are presented visually in Fig. 7, where they are also compared to the onset age obtained by Rasmussen et al. (2014), as well as the best estimates from Buizert et al. (2015) and Capron et al. (2021) which are presented in Table 2.

Although the GICC05 ages y^* of abrupt warming events and subevents as reported by Rasmussen et al. (2014) fall within our estimated 95 % credible intervals for all transitions, there are some transitions where there is a notable difference between the estimated posterior marginal mean $E(Y^*)$ and the reported y^*. This can be partially explained by the fact that Rasmussen et al. (2014) uses a lower 20-year temporal resolution, and that they determine the onset from three different ice cores and two different proxies ($\delta^{18}O$ and Ca^{2+}), whereas our assessment is based on the univariate NGRIP proxy records only. With $E(Y^*) - y^* \sim 120$ the difference is most prominent in the GI-11 transition, whose simulated ages are represented in the histogram in Fig. 6c. This discrepancy is caused by the difference between our estimated onset depth Z^* and z^* from our linear ramp model fit shown in Fig. 6a. Our estimated onset depth $E(Z^*)$ dif-

Table 1. Linear ramp model fits for the NGRIP depth of 29 abrupt warming events, as well as the full dating uncertainty. Data include the depth z^* and dating y^* from Rasmussen et al. (2014), as well as the posterior marginal mean and 95 % credible intervals for the estimated onset depth Z^* and age Y^*.

Event	z^* (m)	Z^* mean (m)	Z^* 95 % CI (m)	y^* (yr b2k)	Y^* mean (yr b2k)	Y^* 95 % CI (yr b2k)
GI-1d	1574.80	1574.97	(1574.85, 1575.11)	14 075	14 078.53	(14 020.02, 14 136.57)
GI-1e	1604.64	1604.52	(1604.47, 1604.57)	14 692	14 689.13	(14 621.21, 14 758.25)
GI-2.2	1793.19	1793.87	(1793.63, 1794.08)	23 340	23 383.98	(23 269.8, 23 497.12)
GI-3	1869.12	1869.22	(1868.95, 1869.56)	27 780	27 789.43	(27 661.18, 27 916.42)
GI-4	1891.57	1891.67	(1891.32, 1892.05)	28 900	28 909.82	(28 777.57, 29 042.34)
GI-5.2	1951.65	1952.02	(1951.98, 1952.06)	32 500	32 524.76	(32 383.73, 32 662.2)
GI-6	1974.55	1974.40	(1974.36, 1974.44)	33 740	33 734.54	(33 590.28, 33 874.96)
GI-7b	1997.04	1997.25	(1997.20, 1997.31)	35 020	35 029.43	(34 881.88, 35173.95)
GI-7c	2009.44	2009.76	(2009.70, 2009.82)	35 480	35 502.41	(35 352.63, 35 649.46)
GI-8c	2070.02	2069.91	(2069.78, 2070.07)	38 220	38 216.54	(38 058.91, 38 372.02)
GI-9	2099.61	2099.65	(2099.64, 2099.66)	40 160	40 163.26	(40 001.66, 40 322.67)
GI-10	2124.03	2124.38	(2124.26, 2124.47)	41 460	41 484.34	(41 320.07, 41 647.05)
GI-11	2157.49	2159.20	(2159.02, 2159.37)	43 340	43 469.75	(43 300.83, 43 637.19)
GI-12c	2222.30	2222.27	(2222.16, 2222.39)	46 860	46 860.6	(46 685.25, 47 035.35)
GI-13b	2253.84	2254.11	(2254.05, 2254.17)	49 120	49 135.6	(48 957.52, 49 313.66)
GI-13c	2256.89	2257.39	(2257.26, 2257.54)	49 280	49 313.32	(49 134.13, 49 491.73)
GI-14b	2295.90	2296.00	(2295.81, 2296.17)	51 660	51 666.83	(51 482.6, 51 849.78)
GI-14c	2340.38	2340.03	(2339.93, 2340.13)	53 960	53 943.9	(53 753.12, 54 131.19)
GI-14d	2341.38	2341.55	(2341.52, 2341.59)	54 020	54 028.81	(53 837.72, 54 215.99)
GI-14e	2345.52	2345.65	(2345.59, 2345.70)	54 220	54 233.81	(54 042.35, 54 421.9)
GI-15.1	2355.34	2355.35	(2355.33, 2355.36)	55 000	55 002.51	(54 810.67, 55 190.97)
GI-15.2	2366.32	2366.56	(2366.47, 2366.64)	55 800	55 824.2	(55 630.79, 56 013.95)
GI-16.1b	2397.35	2397.36	(2397.28, 2397.47)	57 960	57 959.99	(57 762.56, 58 153.23)
GI-16.1c	2398.78	2398.67	(2398.58, 2398.75)	58 040	58 035.39	(57 838.32, 58 228.46)
GI-16.2	2402.55	2402.30	(2402.25, 2402.34)	58 280	58 266.31	(58 068.78, 58 459.55)
GI-17.1a	2409.78	2409.50	(2409.37, 2409.66)	58 780	58 765.24	(58 567.02, 58 960.27)
GI-17.1b	2410.65	2411.26	(2411.04, 2411.45)	58 840	58 876.59	(58 677.75, 59 072.46)
GI-17.1c	2415.01	2414.83	(2414.77, 2414.89)	59 080	59 069.05	(58 870.83, 59 265.27)
GI-17.2	2420.44	2420.70	(2420.64, 2420.76)	59 440	59 465.5	(59 266.65, 59 662.2)

fers from the value z^* reported by Rasmussen et al. (2014) by approximately 1.6 m. This discrepancy propagates into an accordingly large difference in the age estimation of the transition onset. This demonstrates the importance of incorporating proper estimation and uncertainty quantification of the onset depth. Although the absolute uncertainty added from determining the onset depth can be considered negligible compared with the much larger age–depth uncertainty, there can still be a noticeable shift in the estimated onset age propagated from the estimation of the onset depth.

5 Discussion

Motivated by the relationship between air temperature on the one hand and the water-holding capacity and thus precipitation on the other hand, we assumed a linear dependency of the number of incremental layers per 5 cm on respective values of $\delta^{18}O$, which serves as an air temperature proxy. The only other climate proxy variable which is available from the NGRIP ice core at the same resolution over the same time period is the Ca^{2+} particle concentration (Ruth et al., 2003). From visual inspection one can already see that the negative logarithm of the Ca^{2+} mass concentration record shows high covariability with the $\delta^{18}O$ record (see, for example, Fig. 1 of Rasmussen et al. (2014)). Changes in the Ca^{2+} concentrations are interpreted as changes in the local and hemispheric atmospheric circulation (e.g., Ruth et al., 2007; Schüpbach et al., 2018; Erhardt et al., 2019), which would also affect the amount of precipitation over Greenland. Above, we have focused on the $\delta^{18}O$; and analogous analysis for $-\log Ca^{2+}$ is presented in Appendix B. The results of the $-\log Ca^{2+}$-based chronology modeling are in very good agreement with the ones based on δ^{18}, which corroborates our methodology.

Adopting Gaussian noise models in principle allows for a negative modeled number of incremental layers in a 5 cm incremental core segment, which does not seem plausible from a physical perspective. To avoid this, for comparison one could log-transform the age increments and then apply the modeling procedure. The model output would then have to be transferred back by taking the exponential. However,

Table 2. Greenland interstadial transitions observed in the NGRIP record. Data include the median onset NGRIP depth and age for the δ^{18}O and Ca^{2+} proxy records as reported in Capron et al. (2021) and the δ^{18}O record presented in Buizert et al. (2015).

Transition	Capron et al. (2021)			Buizert et al. (2015)	
	Depth (m)	δ^{18}O onset (yr b2k)	Ca^{2+} onset (yr b2k)	Depth (m)	δ^{18}O onset (yr b2k)
GI-1e	1604.89	14 700	14 708	1604.05	14 628
GI-2.2	1794.49	23 400	23 428		
GI-3	1869.35	27 788	27 797	1869.00	27 728
GI-4	1891.77	28 911	28 912	1891.27	28 838
GI-5.2	1952.26	32 528	32 540	1951.66	32 452
GI-7c	2010.12	35 510	35 507	2009.62	35 437
GI-8c	2070.22	38 231	38 239	2069.88	38 165
GI-10	2124.46	41 482	41 494	2123.98	41 408
GI-11	2159.33	43 471	43 366	2157.58	43 297
GI-12c	2222.71	46 887	46 896	2221.96	46 794
GI-14e	2345.73	54 235	54 233	2345.39	54 164
GI-15.1	2355.41	55 006	55 038	2355.17	54 940
GI-15.2	2366.71	55 831	55 831	2366.15	55 737
GI-16.2	2402.53	58 279	58 298	2402.25	59 018
GI-17.1c	2415.01	59 080	59 095	2414.82	59 018
GI-17.2	2420.98	59 480	59 483	2420.35	59 386

it turns out that the log transformation induces systematic deviations of the mean model output age from the reported GICC05 age. With the chances of negative layer increments being fairly small (less than 5 %), overall the original model outperforms the log-transformed model. A more detailed discussion on the log-transformed modeling approach is given in Appendix C.

From Fig. 2b we observe that the variance of the residuals increases slightly with increasing depth down in the core. This heteroskedasticity can be incorporated into our latent Gaussian model by assuming that the variance depends on the core depth in some predefined way. As this implementation is rather technical we consider it beyond the scope of the current paper. Regardless, assuming constant variance appears to be a good first order approach.

6 Conclusions

We have developed a general statistical framework for quantifying the age–depth uncertainty of layer counted proxy archives. In these records the age can be determined by counting annual layers that result from seasonal variations, which in turn impact the deposition process. By counting these layers one can assign time stamps to the individual proxy measurements. However, there is a non-negligible uncertainty associated with this counting process. Proper quantification of this uncertainty is important since it carries valuable information and the error propagates to further analyses, e.g., dating of climatic events, determining cause and effect between such events, and model-data comparisons. Originally, the uncertainty of the GICC05 is quantified in terms

of the maximum counting error (MCE), defined as half the number of uncertain layers. However, since this method assumes that uncertain layers are either true or false, we believe this to be an overly conservative estimate, giving too high uncertainty for deeper layers.

In our approach we express the number of layers per depth increment as the sum of a structured component and a stochastic component. The structured component represents physical layer thinning, a positive temperature–precipitation feedback, and persisting trends over individual stadials and interstadials. The stochastic component takes into account the natural variability of the layer thickness and the errors made in the counting process. After fitting the structured component in a least squares manner, we find the residuals to be approximately stationary, Gaussian distributed, and to exhibit short-range autocorrelation. These summary statistics motivate us to employ Gaussian white noise, or an autoregressive process of first or second order, as the stochastic part of the age–depth model.

After defining the structure of the model, we estimate all model parameters simultaneously in a hierarchical Bayesian framework. The resulting joint posterior distribution on the one hand serves as a quantification of the parameter uncertainty in the model and on the other hand allows to generate chronology ensembles that reflect the uncertainty in the age–depth relationship of the NGRIP ice core. The dating uncertainties obtained from this approach are significantly smaller than those from the MCE. We also find that our estimates do not deviate much from the GICC05 in terms of best estimates for the dating.

Additional information that may help to further constrain the uncertainties, such as tie points obtained via cosmogenic
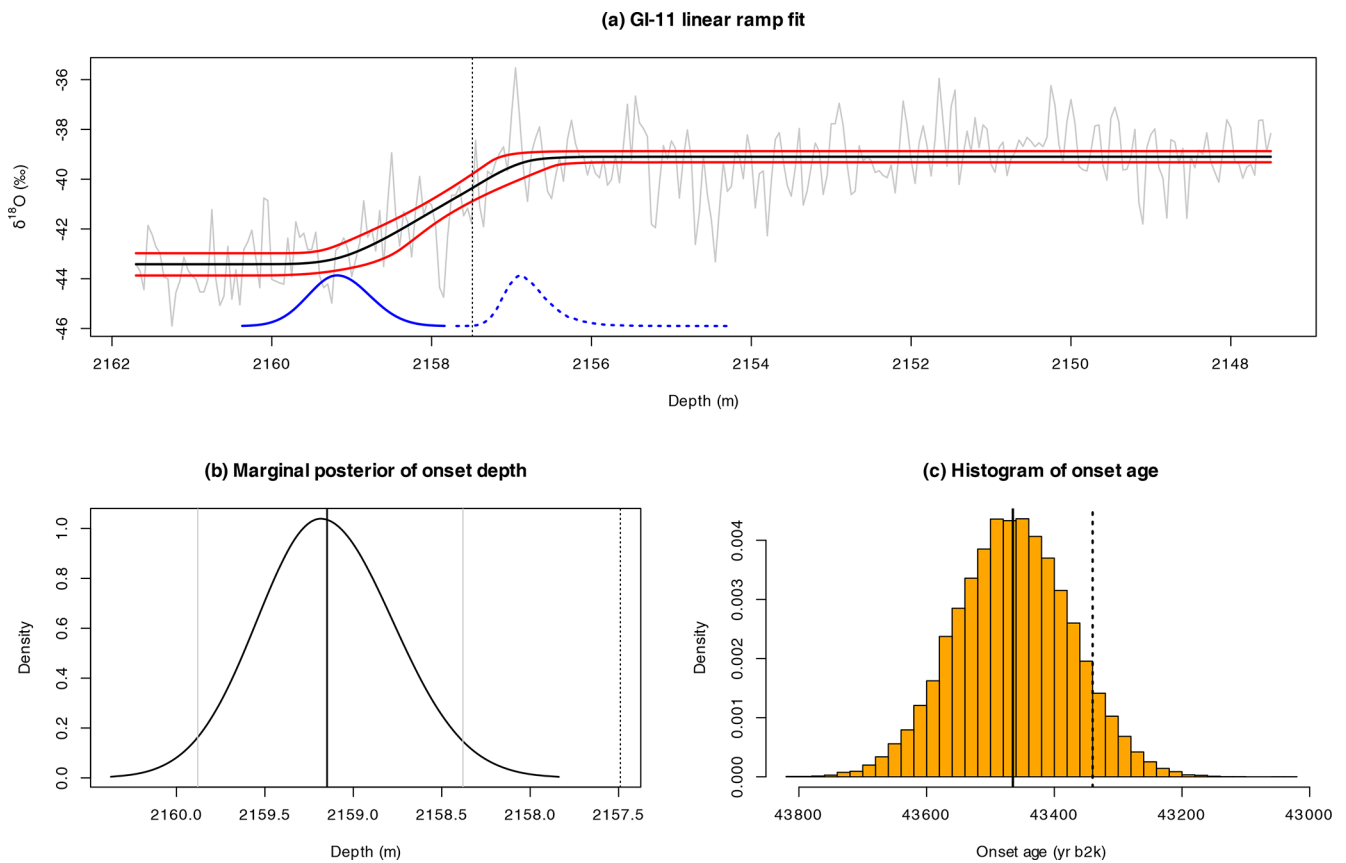
Figure 6. (a) The recorded δ^{18}O of the NGRIP data $\boldsymbol{x}^*$ (gray) as a function of depth for the GI-11 transition. The black line represents the posterior marginal mean of the linear ramp model fitted using INLA. The enclosing red lines represent the corresponding 95 % credible intervals. The blue curves at the bottom illustrate the (unscaled) posterior distributions of the onset (solid) and end point (dotted) of the transition. The vertical dotted line represents the onset depth z^* as reported in Rasmussen et al. (2014). **(b)** The posterior marginal distribution of the onset depth Z^* following a linear ramp model fit. The solid vertical lines represent the posterior marginal mean (black) and 95 % credible intervals (gray). The dotted vertical line represents z^*. **(c)** Histogram describing the complete dating uncertainty of the transition to GI-11, taking into account the uncertainty of the NGRIP depth of the onset as well as the dating uncertainty at this depth. The solid vertical black line represents the mean of these samples, $E(Y^*)$, and the dotted vertical line represents the GICC05 onset age y^* as reported in Rasmussen et al. (2014).

radionuclides (Adolphi et al., 2018), will be fed into the model in future research.

One of the biggest concerns regarding the layer counting is that of a potential counting bias. Such a systematic error cannot be corrected after the counting and, therefore, we investigate how a potential unknown counting bias increases the uncertainty of the presented age–depth model. If such a counting bias is restricted to ± 4 % we obtain total age uncertainties comparable to the estimates based on the MCE. Finally, we apply our method to the dating of DO events. Using a Bayesian transition onset detection we are able to combine the uncertainty of the onset depth with the corresponding age uncertainty, and to give a posterior distribution that entails the complete dating uncertainty of each transition onset. We find that previous estimates of the DO onsets reported in Rasmussen et al. (2014) are well within our estimated uncertainty ranges both in terms of depth and age. However, the

dating uncertainties of the abrupt warming event onsets are all considerably smaller than the MCE, even when accounting for the additional uncertainty associated with the onset depth.

In theory, it should be possible to apply this approach to other layered proxy records as well. However, there are some requirements that need to be fulfilled for this approach to be applicable. The first condition is that a potential layer thinning can be adequately expressed by a regression model. In our results we find the residuals to follow a Gaussian process, but it should be possible for the model to be adapted such that it supports other distributions for the residuals as well. However, depending on the model, if the residuals exhibit too long memory then this could lead to the simulation procedure having an infeasibly high computational cost. Moreover, if there are many effects in the regression model there needs to be sufficient data to achieve proper inference.
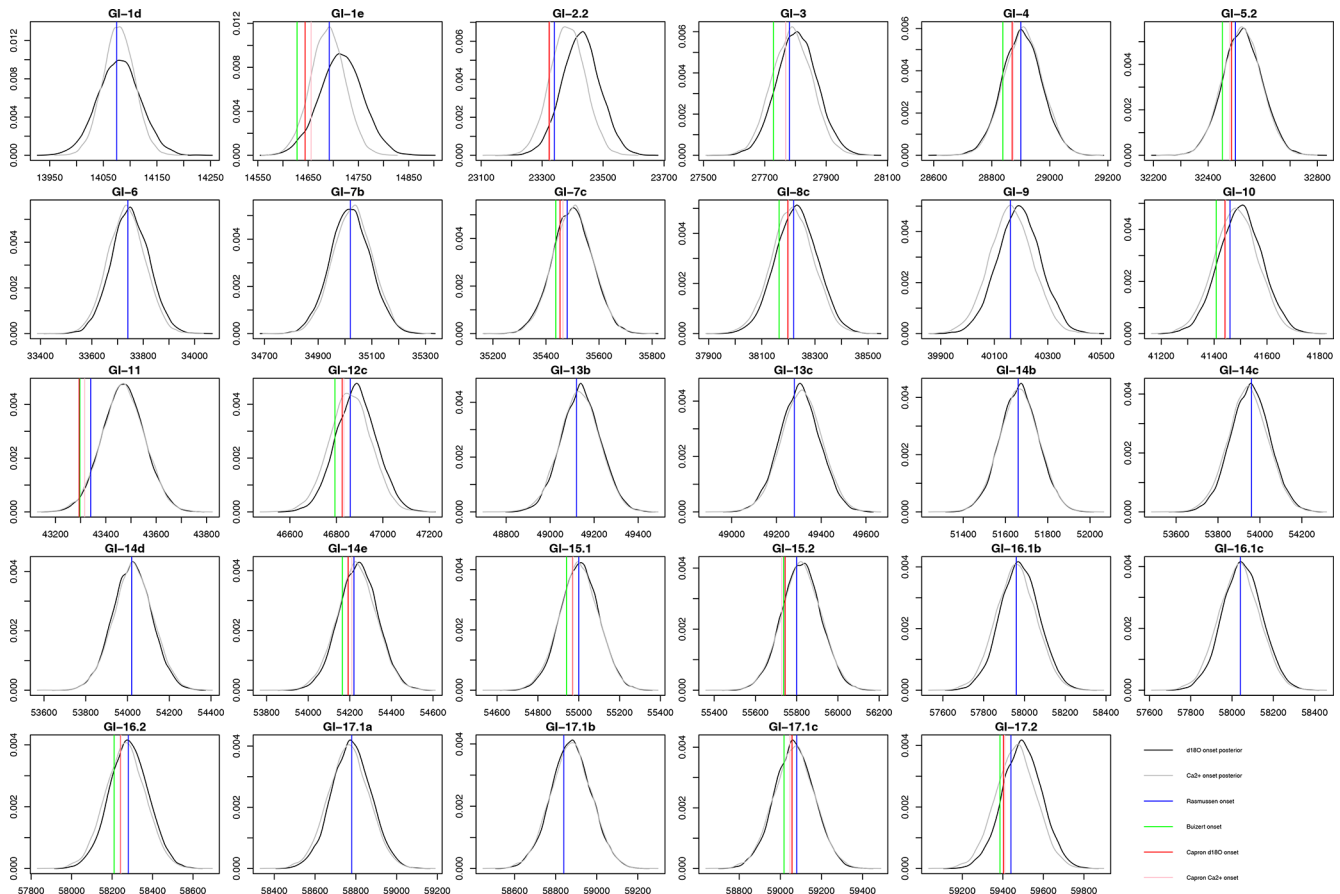
Figure 7. Posterior distributions of the onset ages of abrupt climate transitions for the NGRIP $\delta^{18}O$ (black) and Ca^{2+} (gray) proxy records. The results are compared with ages reported by Rasmussen et al. (2014) (blue), Buizert et al. (2015) (green), and Capron et al. (2021) (red and pink for the $\delta^{18}O$ and Ca^{2+} record, respectively). The Ca^{2+}-based posterior distributions rely on modeled chronologies which are presented in Appendix B and which are not shown in the main text.

Appendix A: Latent Gaussian model formulation

In this study we consider different Gaussian models for the noise component, including independent identically distributed (iid) and first and second order autoregressive (AR) models. These models all exhibit the Markov property, meaning there is a substantial amount of conditional independence. So-called Gaussian Markov random fields are known to work really well with the methodology of integrated nested Laplace approximations (INLAs), which will grant a substantial reduction in computational cost in obtaining full Bayesian inference. However, this requires formulating our model into a latent Gaussian model where the data, here $\mathcal{D} = (\Delta y_1, \ldots, \Delta y_n)^\top$, depend on a set of latent Gaussian variables $\boldsymbol{X} = (X_1, \ldots, X_N)^\top$ which in turn depend on hyperparameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)^\top$. This class of models constitutes a subset of hierarchical Bayesian models and is defined in three stages.

The first stage defines the likelihood of the data and how they depend on the latent variables. For the data and models

used in this study we assume a direct correspondence between an observation y_i and the corresponding latent variable X_i, which is achieved using a Gaussian likelihood with some negligible fixed variance and mean given by the linear predictor,

$$\eta_k = E(\Delta y_k) = bz_k^2 + b_x x_k + \sum_{i=1}^{p} \psi_i(z_k; a_i, c_i) + \varepsilon_k(\boldsymbol{\theta}).$$

Here, $\boldsymbol{\beta} = (b_0, b, b_x, \{a_i\}, \{c_i\})$ are known as fixed effects, even though they are indeed stochastic variables in the Bayesian framework. The noise variables $\varepsilon_k(\boldsymbol{\theta})$ are referred to as random effects since they depend on hyperparameters $\boldsymbol{\theta}$. The hyperparameters are $\boldsymbol{\theta} = \sigma_\varepsilon$ if we assume the residuals follow an iid Gaussian process, $\boldsymbol{\theta} = (\sigma_\varepsilon, \phi)$ if they follow an AR(1) process, and $\boldsymbol{\theta} = (\sigma_\varepsilon, \phi_1, \phi_2)$ if they follow an AR(2) process. All random terms in the predictor, and the predictor itself, are included in the latent field $\boldsymbol{X} = (\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\varepsilon})$. The latent field is assigned a prior distribution in what is the second stage of defining a latent Gaussian model. For latent Gaus-

sian fields this prior is multivariate Gaussian:

$$X \mid \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta})).$$

Specifically, we assume vague Gaussian priors for $\boldsymbol{\beta}$, while the prior for $\boldsymbol{\varepsilon}(\boldsymbol{\theta})$ is either an iid, AR(1), or AR(2) process. The predictor $\boldsymbol{\eta}$ is then a Gaussian with a mean vector corresponding to the linear regression $\boldsymbol{a}(\boldsymbol{\beta})$ and a covariance matrix given by the covariance structure of the assumed noise model.

The third and final stage of the latent Gaussian model definition is to specify a prior distribution on the hyperparameters. We use the default prior choices included in the `R-INLA` package, which means that for all models of $\boldsymbol{\varepsilon}$ considered in this paper the scaling parameter σ_ε^2 is assigned a log-gamma distribution through the transformation $\kappa = 1/\sigma_\varepsilon^2$. When the residuals follow an AR(1) distribution we assume a Gaussian prior on the additional lag-one correlation parameter using a logit transformation $\rho = \log((1 + \phi)/(1 - \phi))$. For the AR(2) residuals we instead assign penalized complexity priors (Simpson et al., 2017) on the partial autocorrelations $\psi_1 = \phi_1/(1 - \phi_2)$ and $\psi_2 = \phi_2$, also using a logit transformation.

Inference is obtained by computing the posterior marginal distributions:

$$\pi(X_k | \mathcal{D}) = \int \pi(X_k | \boldsymbol{\theta}, \mathcal{D}) \pi(\boldsymbol{\theta} | \mathcal{D}) \mathrm{d}\boldsymbol{\theta}$$

and

$$\pi(\theta_k | \mathcal{D}) = \int \pi(\boldsymbol{\theta} | \mathcal{D}) \mathrm{d}\boldsymbol{\theta}_{-k}.$$

The notation θ_k refers to the kth hyperparameter, and $\boldsymbol{\theta}_{-k}$ refers to all except the kth hyperparameter. These integrals can be approximated efficiently using `R-INLA`, and the resulting posterior marginal distributions are included in Fig. A1. The estimated hyperparameter posterior marginal means and credible intervals (given in parentheses) are $\sigma_\varepsilon^{\text{iid}} \approx 0.427$ (0.423, 0.431) for iid residuals, $\sigma_\varepsilon^{\text{AR}}(1) \approx 0.428$ (0.423, 0.434) and $\phi^{\text{AR}}(1) \approx 0.194$ (0.178, 0.217) for AR(1) residuals, and $\sigma_\varepsilon^{\text{AR}}(2) \approx 0.429$ (0.423, 0.438), $\phi_1^{\text{AR}}(2) \approx 0.180$ (0.158, 0.210), and $\phi_2^{\text{AR}}(2) \approx 0.108$ (0.089, 0.134) for AR(2) residuals.
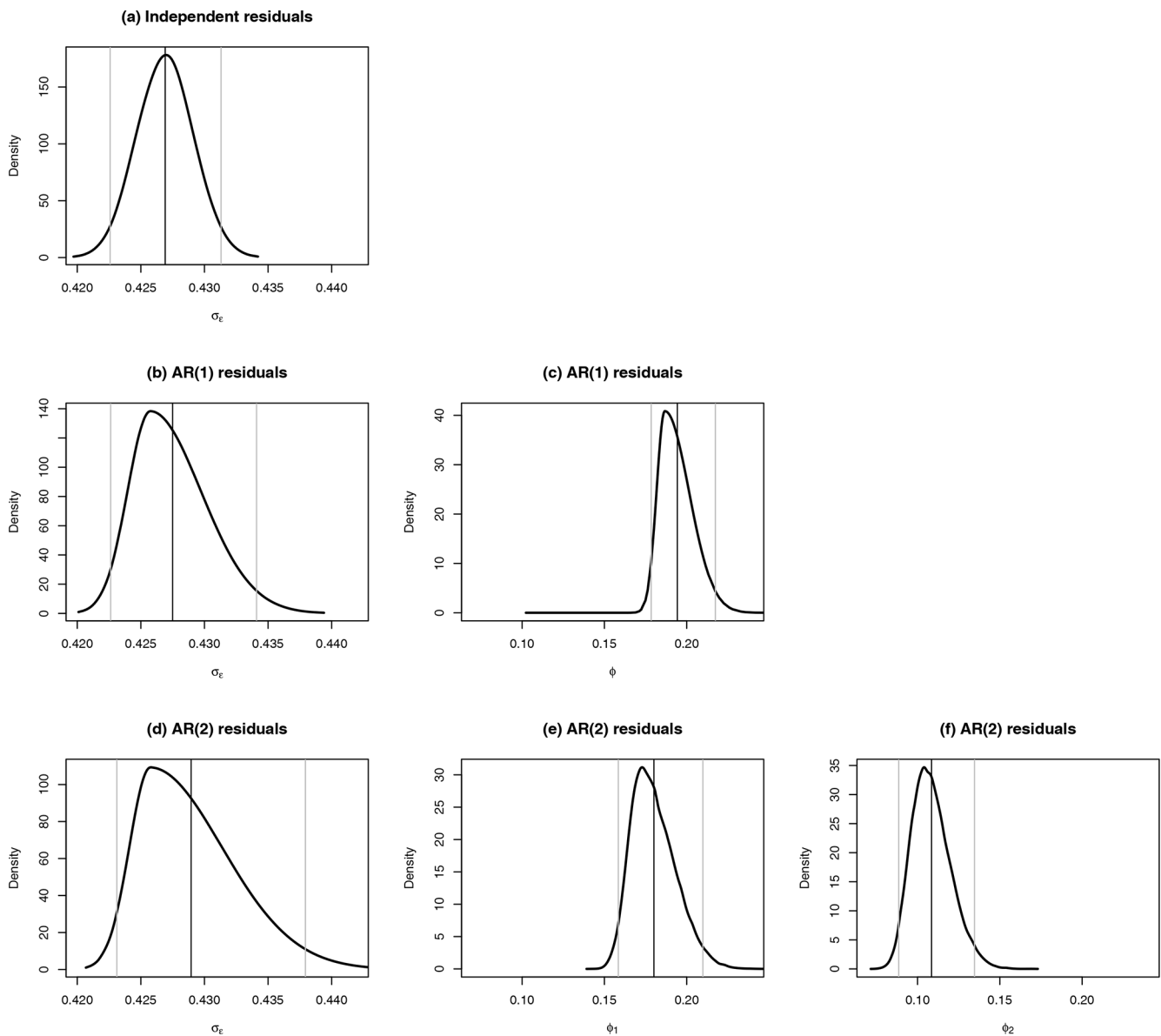
Figure A1. The posterior marginal distributions obtained by fitting the model with INLA. Panel **(a)** shows the density of σ_ε using iid distributed residuals. Panels **(b)** and **(c)** show the densities of σ_ε and ϕ using AR(1) distributed residuals. Panels **(d)**–**(f)** show the densities of $\sigma_\varepsilon, \phi_1$, and ϕ_2 using AR(2) distributed residuals. The vertical lines represent the mean (black) and 95 % credible intervals (gray).

Appendix B: Ca^{2+} analysis

As an alternative, we perform an analogous study using the log(Ca^{2+}) as the proxy variable x_k, which is available at the same period and resolution (Ruth et al., 2003). Missing values in the Ca^{2+} data set are filled using linear interpolation. Performing the same analysis as described in Sect. 3 we produce joint samples of the chronologies. The resulting posterior marginal means and 95 % credible intervals of the dating uncertainties are illustrated in Fig. B1, where the credible intervals obtained using an iid, AR(1), and AR(2) noise model are compared. These results are consistent with those obtained using δ^{18}O as the proxy variable, shown in Fig. 4.
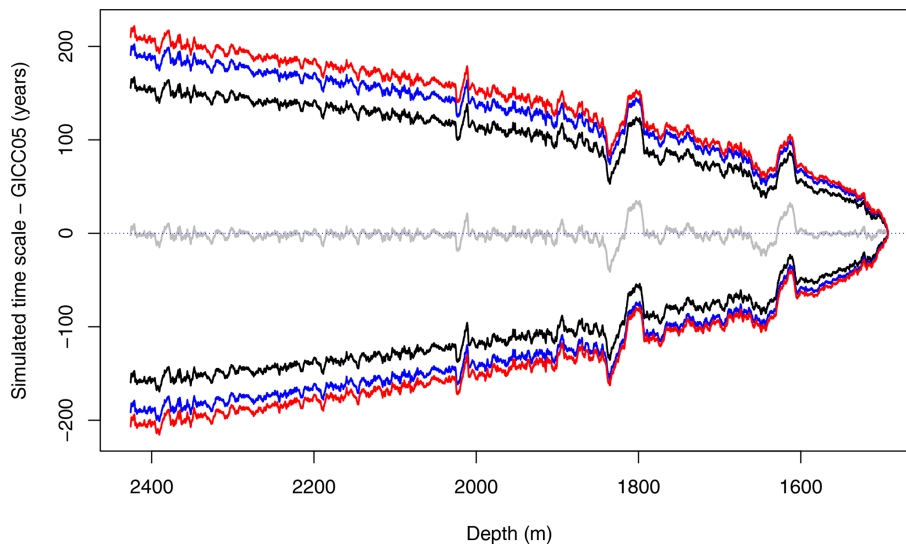
Figure B1. The 95 % credible intervals of the dating uncertainty distribution when log(Ca^{2+}) is used as the proxy covariate. The GICC05 time scale has been subtracted and the noise is modeled using iid (black), AR(1) (blue), and AR(2) (red) noise models. Only the posterior marginal mean computed using AR(1) distributed noise is included (gray) since it is very similar to the mean obtained using other noise assumptions.

Appendix C: Log-normal distribution

A possible concern with the approach presented above is that assuming a normal distribution on the layer increments assigns a non-zero probability of negative depositions, violating monotonicity. Furthermore, one might be concerned by our choice of an additive thinning function rather than a multiplicative. Both of these issues are resolved by instead assuming a log-normal regression model on the layer increments, i.e.,

$$\log \Delta y_k = a(z_k, x_k) + \varepsilon_k, \tag{C1}$$

where ε_k follows either an iid, AR(1), or AR(2) model as before. This results in the joint age variables being a multivariate process with marginals described by sums of log-normal distributions.

Simulations of the chronology can be produced as follows: First, the latent Gaussian model is fitted to the data and second, $a(z_k, x_k)$ and ε_k are sampled from the resulting posterior distributions. Then compute the sampled chronologies:

$$y_k = y_0 + \sum_{i=1}^{k} \exp(a(z_i, x_i) + \varepsilon_i). \quad \text{(C2)}$$

The resulting dating uncertainty is illustrated in Fig. C1 and is more curved than for the original model. This suggests that the normal distribution is a better fit to the observed layer increments. We will hence focus on the original Gaussian model in the analysis of this paper.
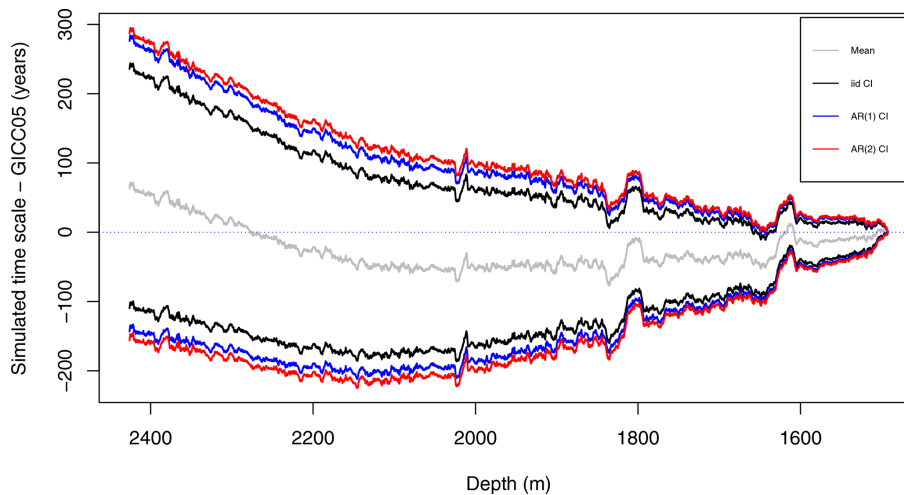
Figure C1. The 95 % credible intervals of the dating uncertainty distribution when the GICC05 time scale has been subtracted and a log-normal distribution is assumed, using iid (black), AR(1) (blue), and AR(2) (red) noise models. Only the posterior marginal mean computed using AR(1) distributed noise is included (gray) since it is very similar to the mean obtained using the other noise models.

Appendix D: Determination of the fitting windows

The estimation for the onset depth is sensitive to the choice of the data window which represents the transition. It is important to select these carefully such that the data best represent a single linear ramp function while being of a sufficient size. As some DO events are located more closely to other transitions than others, it is necessary to determine these data windows individually for each transition. As such there are indeed some transitions where it is difficult to determine a clear transition point, and a linear ramp model is not appropriate. These transitions will be omitted from our analysis. The reduction in computational cost granted by adopting the model for INLA allows us to perform repeated fits to determine the optimal data interval based on a given criteria. Specifically, we adjust both sides of the interval until we find the data window for which the fitted model yields the lowest amplitude of the AR(1) noise, measured by the posterior marginal mean of the standard deviation parameter.

We impose some restrictions on the domain of the optimal start and end points of our data interval. To achieve the best possible fit we want our interval to include both the onset and end point of the transition, which we suspect are located close to the NGRIP onset depth z^* given by Table 2 of Rasmussen et al. (2014). Unless the DO events are located too close to adjacent transitions, we assume the optimal interval always contains the points representing 1 m above to 2.5 m deeper than z^*. These are the minimum distances from the proposed onset. Similarly, we also introduce a maximum distance from z^* to be considered for the start and end points of the data interval. This is set to be 10 m above and 15 m deeper than z^*, or at the adjacent transitions if they are located closer than this. From these intervals we create a two-dimensional grid for which we perform the INLA fit for each grid point. The start and end points of the interval representing the grid point for which INLA found the lowest noise amplitude are selected. Those events that failed to provide a decent fit after the conclusion of this procedure were discarded. This left us with 29 DO events for which the results are displayed in Table D1. Although the GICC05 onset depths z^* fall outside the 95 % credible intervals for several transitions, they are still remarkably close to our best estimates, considering Rasmussen et al. (2014) used a lower 20 year resolution data set.

Table D1. The optimal interval for the data window for fitting a linear ramp model to 29 DO events, expressed in terms of depth and the corresponding index in our data. The Rasmussen et al. (2014) onset depths z^* were used as a starting midpoint in the optimization procedure.

Event	k^*	Index interval	z^* (m)	Depth interval (m)
GI-1d	1648	(1622, 1711)	1574.8	(1573.5, 1577.95)
GI-1e	2245	(2175, 2285)	1604.65	(1601.15, 1606.65)
GI-2.2	6016	(5986, 6079)	1793.2	(1791.7, 1796.35)
GI-3	7534	(7482, 7574)	1869.1	(1866.5, 1871.1)
GI-4	7983	(7929, 8023)	1891.55	(1888.85, 1893.55)
GI-5.2	9185	(9106, 9223)	1951.65	(1947.7, 1953.55)
GI-6	9643	(9572, 9683)	1974.55	(1971, 1976.55)
GI-7b	10 093	(10 068, 10 142)	1997.05	(1995.8, 1999.5)
GI-7c	10 341	(10 169, 10 404)	2009.45	(2000.85, 2012.6)
GI-8c	11 552	(11 352, 11 592)	2070	(2060, 2072)
GI-9	12 144	(12 098, 12 174)	2099.6	(2097.3, 2101.1)
GI-10	12 633	(12 563, 12 690)	2124.05	(2120.55, 2126.9)
GI-11	13 302	(13 102, 13 386)	2157.5	(2147.5, 2161.7)
GI-12c	14 598	(14 417, 14 718)	2222.3	(2213.25, 2228.3)
GI-13b	15 229	(15 212, 15 274)	2253.85	(2253, 2256.1)
GI-13c	15 290	(15 245, 15 339)	2256.9	(2254.65, 2259.35)
GI-14b	16 070	(16 054, 16 110)	2295.9	(2295.1, 2297.9)
GI-14c	16 960	(16 881, 16 983)	2340.4	(2336.45, 2341.55)
GI-14d	16 980	(16 968, 16 992)	2341.4	(2340.8, 2342)
GI-14e	17 062	(16 986, 17 185)	2345.5	(2341.7, 2351.65)
GI-15.1	17 259	(17 247, 17 272)	2355.35	(2354.75, 2356)
GI-15.2	17 478	(17 365, 17 610)	2366.3	(2360.65, 2372.9)
GI-16.1b	18 099	(18 087, 18 111)	2397.35	(2396.75, 2397.95)
GI-16.1c	18 128	(18 108, 18 146)	2398.8	(2397.8, 2399.7)
GI-16.2	18 203	(18 190, 18 215)	2402.55	(2401.9, 2403.15)
GI-17.1a	18 348	(18 315, 18 371)	2409.8	(2408.15, 2410.95)
GI-17.1b	18 365	(18 353, 18 431)	2410.65	(2410.05, 2413.95)
GI-17.1c	18 452	(18 387, 18 464)	2415	(2411.75, 2415.6)
GI-17.2	18 561	(18 531, 18 608)	2420.45	(2418.95, 2422.8)

References

Adolphi, F., Bronk Ramsey, C., Erhardt, T., Edwards, R. L., Cheng, H., Turney, C. S. M., Cooper, A., Svensson, A., Rasmussen, S. O., Fischer, H., and Muscheler, R.: Connecting the Greenland ice-core and U/Th timescales via cosmogenic radionuclides: testing the synchroneity of Dansgaard–Oeschger events, Clim. Past, 14, 1755–1781, https://doi.org/10.5194/cp-14-1755-2018, 2018.

Andersen, K. K., Svensson, A., Johnsen, S. J., Rasmussen, S. O., Bigler, M., Röthlisberger, R., Ruth, U., Siggaard-Andersen, M. L., Peder Steffensen, J., Dahl-Jensen, D., Vinther, B. M., and Clausen, H. B.: The Greenland Ice Core Chronology 2005, 15–42 ka. Part 1: constructing the time scale, Quaternary Sci. Rev., 25, 3246–3257, https://doi.org/10.1016/j.quascirev.2006.08.002, 2006.

Blaauw, M. and Christeny, J. A.: Flexible paleoclimate age-depth models using an autoregressive gamma process, Bayesian Anal., 6, 457–474, https://doi.org/10.1214/11-BA618, 2011.

Boers, N., Goswami, B., and Ghil, M.: A complete representation of uncertainties in layer-counted paleoclimatic archives, Clim. Past, 13, 1169–1180, https://doi.org/10.5194/cp-13-1169-2017, 2017.

Buizert, C., Cuffey, K. M., Severinghaus, J. P., Baggenstos, D., Fudge, T. J., Steig, E. J., Markle, B. R., Winstrup, M., Rhodes, R. H., Brook, E. J., Sowers, T. A., Clow, G. D., Cheng, H., Edwards, R. L., Sigl, M., McConnell, J. R., and Taylor, K. C.: The WAIS Divide deep ice core WD2014 chronology – Part 1: Methane synchronization (68–31 ka BP) and the gas age–ice age

difference, Clim. Past, 11, 153–173, https://doi.org/10.5194/cp-11-153-2015, 2015.

Capron, E., Rasmussen, S. O., Popp, T. J., Erhardt, T., Fischer, H., Landais, A., Pedro, J. B., Vettoretti, G., Grinsted, A., Gkinis, V., Vaughn, B., Svensson, A., Vinther, B. M., and White, J. W.: The anatomy of past abrupt warmings recorded in Greenland ice, Nat. Commun., 12, 2106, https://doi.org/10.1038/s41467-021-22241-w, 2021.

Comboul, M., Emile-Geay, J., Evans, M. N., Mirnateghi, N., Cobb, K. M., and Thompson, D. M.: A probabilistic model of chronological errors in layer-counted climate proxies: applications to annually banded coral archives, Clim. Past, 10, 825–841, https://doi.org/10.5194/cp-10-825-2014, 2014.

Corrick, E. C., Drysdale, R. N., Hellstrom, J. C., Capron, E., Rasmussen, S. O., Zhang, X., Fleitmann, D., Couchoud, I., Wolff, E., and Monsoon, S. A.: Synchronous timing of abrupt climate changes during the last glacial period, Science, 369, 963–969, 2020.

Dansgaard, W., Johnsen, S. J., Clausen, H. B., Dahl-Jensen, D., Gundestrup, N. S., Hammer, C. U., Hvidberg, C. S., Steffensen, J. P., Sveinbjörnsdottir, A. E., Jouzel, J., and Bond, G.: Evidence for general instability of past climate from a 250-kyr ice-core record, Nature, 364, 218–220, https://doi.org/10.1038/364218a0, 1993.

Erhardt, T., Capron, E., Rasmussen, S. O., Schüpbach, S., Bigler, M., Adolphi, F., and Fischer, H.: Decadal-scale progression of the onset of Dansgaard–Oeschger warming events, Clim. Past, 15, 811–825, https://doi.org/10.5194/cp-15-811-2019, 2019.

Gkinis, V., Simonsen, S. B., Buchardt, S. L., White, J. W., and Vinther, B. M.: Water isotope diffusion rates from the North-GRIP ice core for the last 16,000 years – Glaciological and paleoclimatic implications, Earth Planet. Sc. Lett., 405, 132–141, https://doi.org/10.1016/j.epsl.2014.08.022, 2014.

Goodman, J. and Weare, J.: Ensemble samplers with affine invariance, Comm. App. Math. Com. Sc., 5, 1–99, 2010.

Haslett, J. and Parnell, A.: A simple monotone process with application to radiocarbon-dated depth chronologies, J. R. Stat. Soc. C-Appl., 57, 399–418, https://doi.org/10.1111/j.1467-9876.2008.00623.x, 2008.

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

Jouzel, J., Alley, R. B., Cuffey, K. M., Dansgaard, W., Grootes, P., Hoffmann, G., Johnsen, S. J., Koster, R. D., Peel, D., Shuman, C. A., Stievenard, M., Stuiver, M., and White, J.: Validity of the temperature reconstruction from water isotopes in ice cores, J. Geophys. Res.-Oceans, 102, 26471–26487, https://doi.org/10.1029/97JC01283, 1997.

Li, T. Y., Han, L. Y., Cheng, H., Edwards, R. L., Shen, C. C., Li, H. C., Li, J. Y., Huang, C. X., Zhang, T. T., and Zhao, X.: Evolution of the Asian summer monsoon during Dansgaard/Oeschger events 13–17 recorded in a stalagmite constrained by high-precision chronology from southwest China, Quaternary Res., 88, 121–128, https://doi.org/10.1017/qua.2017.22, 2017.

McKay, N. P., Emile-Geay, J., and Khider, D.: geoChronR – an R package to model, analyze, and visualize age-uncertain data, Geochronology, 3, 149–169, https://doi.org/10.5194/gchron-3-149-2021, 2021.

Myrvoll-Nilsen, E.: eirikmn/dating_uncertainty: First release (v1.1.0), Zenodo [code], https://doi.org/10.5281/zenodo.6637528, 2022.

North Greenland Ice Core Project members: High-resolution record of Northern Hemisphere climate extending into the last interglacial period, Nature, 431, 147–151, https://doi.org/10.1038/nature02805, 2004.

Parnell, A. C., Haslett, J., Allen, J. R., Buck, C. E., and Huntley, B.: A flexible approach to assessing synchroneity of past events using Bayesian reconstructions of sedimentation history, Quaternary Sci. Rev., 27, 1872–1885, https://doi.org/10.1016/j.quascirev.2008.07.009, 2008.

Ramsey, C. B.: Radiocarbon calibration and analysis of stratigraphy: the OxCal program, Radiocarbon, 37, 425–430, https://doi.org/10.1017/s0033822200030903, 1995.

Ramsey, C. B.: Deposition models for chronological records, Quaternary Sci. Rev., 27, 42–60, https://doi.org/10.1016/j.quascirev.2007.01.019, 2008.

Rasmussen, S. O., Andersen, K. K., Svensson, A. M., Steffensen, J. P., Vinther, B. M., Clausen, H. B., Siggaard-Andersen, M. L., Johnsen, S. J., Larsen, L. B., Dahl-Jensen, D., Bigler, M., Röthlisberger, R., Fischer, H., Goto-Azuma, K., Hansson, M. E., and Ruth, U.: A new Greenland ice core chronology for the last glacial termination, J. Geophys. Res.-Atmos., 111, 1–16, https://doi.org/10.1029/2005JD006079, 2006.

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., Wheatley, J. J., and Winstrup, M.: A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: refining and extending the INTIMATE event stratigraphy, Quaternary Sci. Rev., 106, 14–28, https://doi.org/10.1016/j.quascirev.2014.09.007, 2014.

Riechers, K. and Boers, N.: A statistical approach to the phasing of atmospheric reorganization and sea ice retreat at the onset of Dansgaard–Oeschger events under rigorous treatment of uncertainties, Clim. Past Discuss. [preprint], 2020, 1–30, https://doi.org/10.5194/cp-2020-136, 2020.

Rue, H., Martino, S., and Chopin, N.: Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion), J. R. Stat. Soc. Series B, 71, 319–392, 2009.

Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P., and Lindgren, F. K.: Bayesian Computing with INLA: A Review, Annu. Rev. Stat. Appl., 4, 395–421, 2017.

Ruth, U., Wagenbach, D., Steffensen, J. P., and Bigler, M.: Continuous record of microparticle concentration and size distribution in the central Greenland NGRIP ice core during the last glacial period, J. Geophys. Res.-Atmos., 108, 1–12, https://doi.org/10.1029/2002jd002376, 2003.

Ruth, U., Bigler, M., Röthlisberger, R., Siggaard-Andersen, M. L., Kipfstuhl, S., Goto-Azuma, K., Hansson, M. E., Johnsen, S. J., Lu, H., and Steffensen, J. P.: Ice core evidence for a very tight link between North Atlantic and east Asian glacial climate, Geophys. Res. Lett., 34, 1–5, https://doi.org/10.1029/2006GL027876, 2007.

Schüpbach, S., Fischer, H., Bigler, M., Erhardt, T., Gfeller, G., Leuenberger, D., Mini, O., Mulvaney, R., Abram, N. J., Fleet, L., Frey, M. M., Thomas, E., Svensson, A., Dahl-Jensen, D., Kettner, E., Kjaer, H., Seierstad, I., Steffensen, J. P., Rasmussen, S. O., Vallelonga, P., Winstrup, M., Wegner, A., Twarloh, B., Wolff, K., Schmidt, K., Goto-Azuma, K., Kuramoto, T., Hirabayashi, M., Uetake, J., Zheng, J., Bourgeois, J., Fisher, D., Zhiheng, D., Xiao, C., Legrand, M., Spolaor, A., Gabrieli, J., Barbante, C., Kang, J. H., Hur, S. D., Hong, S. B., Hwang, H. J., Hong, S., Hansson, M., Iizuka, Y., Oyabu, I., Muscheler, R., Adolphi, F., Maselli, O., McConnell, J., and Wolff, E. W.: Greenland records of aerosol source and atmospheric lifetime changes from the Eemian to the Holocene, Nat. Commun., 9, 1476, https://doi.org/10.1038/s41467-018-03924-3, 2018.

Simpson, D., Rue, H., Riebler, A., Martins, T. G., and Sørbye, S. H.: Penalising model component complexity: A principled, practical approach to constructing priors, Stat. Sci., 232, 1–28, 2017.

Svensson, A., Andersen, K. K., Bigler, M., Clausen, H. B., Dahl-Jensen, D., Davies, S. M., Johnsen, S. J., Muscheler, R., Parrenin, F., Rasmussen, S. O., Röthlisberger, R., Seierstad, I., Steffensen, J. P., and Vinther, B. M.: A 60 000 year Greenland stratigraphic ice core chronology, Clim. Past, 4, 47–57, https://doi.org/10.5194/cp-4-47-2008, 2008.

Vinther, B. M., Clausen, H. B., Johnsen, S. J., Rasmussen, S. O., Andersen, K. K., Buchardt, S. L., Dahl-Jensen, D., Seierstad, I. K., Siggaard-Andersen, M. L., Steffensen, J. P., Svensson, A., Olsen, J., and Heinemeier, J.: A synchronized dating of three Greenland ice cores throughout the Holocene, J. Geophys. Res.-Atmos., 111, 1–11, https://doi.org/10.1029/2005JD006921, 2006.

Zhou, H., Zhao, J. X., Feng, Y., Chen, Q., Mi, X., Shen, C. C., He, H., Yang, L., Liu, S., Chen, L., Huang, J., and Zhu, L.: Heinrich event 4 and Dansgaard/Oeschger events 5–10 recorded by high-resolution speleothem oxygen isotope data from central China, Quaternary Res., 82, 394–404, https://doi.org/10.1016/j.yqres.2014.07.006, 2014.