Bayesian Analysis (2025) **TBA**, Number TBA, pp. 1–20

Synchronization of Layer-Counted Paleoclimatic Proxy Archives Using a Bayesian Regression Modeling Framework

Eirik Myrvoll-Nilsen [*,†]

Keno Riechers[†,‡] and Niklas Boers[‡,†,§]

Abstract. Layer-counted proxy records from paleoclimatic archives are subject to considerable dating uncertainties. These uncertainties originate from irregularities in the archive's deposition process that result, in turn, in errors during the layer counting process. Dating uncertainties can be quantified by assuming a probabilistic model for the relationship between the depth of a sample in the proxy archive and the age of that sample. However, systematic biases in counting or depositional processes can cause the counted chronology to deviate substantially from the true age and possibly corrupt the age model. By synchronizing a given chronology with other, independently dated archives, one can constrain the dating uncertainties and correct potential biases. This can be done by matching the chronology to tie-points obtained by identifying characteristic events which were recorded simultaneously by different archives or with independent methods. However, updating the counted age–depth relationship under the consideration of tie-points is not straightforward and no generally accepted method is presently available for layer-counted archives. A key requirement for such a method is that it should include an appropriate uncertainty-sensitive interpolation between tie-points. Using a Gaussian model to represent a potential bias, we show how tie-points and their uncertainties can be incorporated into a previously suggested Bayesian modeling framework to reflect the general uncertainties of a counted chronology. Both the uncertainty inherent to the tie-points and the age-correlation between the data from different depths in the archive are consistently represented in this approach. We demonstrate the methodology in two applications: first, using synthetic data, and second, applying the methodology to data from the NGRIP ice core, an iconic paleoclimate proxy archive.

MSC2020 subject classifications: Primary 62F15, 62P12; secondary 86A08.

Keywords: age–depth models, model synchronization, INLA, paleoclimate.

1 Introduction

Measurements of meteorological quantities such as surface temperature cannot be obtained directly for the long-term past beyond ∼370 years ago (Parker et al., 1992), when

[*]Department of Mathematics and Statistics, UiT – The Arctic University of Norway, Tromsø, Norway, eirik.myrvoll-nilsen@uit.no

[†]Complexity Science, Potsdam Institute for Climate Impact Research, Potsdam, Germany

[‡]School of Engineering & Design, Technical University of Munich, Munich, Germany

[§]Global Systems Institute and Department of Mathematics, University of Exeter, Exeter, UK

 https://doi.org/10.1214/25-BA1509

the first instrumental measurements started. Instead, preserved physical or chemical characteristics stored in natural climate archives and known to correlate with climate variables of interest are the only available means for quantitative analysis of the climate earlier than the mid 17th century (Mann and Jones, 2003; Mann et al., 2008; Ruddiman, 2014, Chapter 3). Such proxy archives include for example ice core records, cave speleothems, corals, and lake or marine sediments.

A prerequisite for the study of past climate is that each proxy measurement can be assigned an accompanying age. Together, the ages assigned to a series of proxy measurements define a chronology for a given proxy record. Along the 'growth'-axis of an archive, the age y of the deposited material increases with increasing depth z. However, this age–depth relationship $y(z)$ is typically nonlinear and may be very irregular. The dating of paleoclimate proxy records is inevitably subject to uncertainties that should be considered in any subsequent analysis. Rigorous uncertainty propagation is particularly important in the study of the spatio-temporal patterns and causal relationships between past climatic events, in particular if the analysis is based on records from multiple climate archives (see e.g. Adolphi et al., 2018; Corrick et al., 2020; Svensson et al., 2020; Myrvoll-Nilsen et al., 2022; Riechers and Boers, 2021).

Bayesian age–depth models have proven successful in quantifying the dating uncertainty of proxy records, and different techniques have been developed for different types of archives (e.g. Parnell et al., 2008; Bronk Ramsey, 2008; Blaauw and Christen, 2011; McKay et al., 2021). Here we are concerned exclusively with layer-counted age–depth models, which are based on identifying and counting annual deposition layers. The cumulative sum of layers with respect to a reference point (often the present is identified with the top of the archive) defines the age of the record at a certain depth. During the counting process, errors may arise from the physical absence of discernible annual layers, double layers formed in a single year, or from inaccuracies in the counting itself. The latter can for example occur when a signal within the record is mistakenly identified as an annual layer or when an actual annual layer is overlooked (see e.g. Andersen et al., 2006; Comboul et al., 2014).

Layer-counted archives bear the particularity that dating errors accumulate and hence the absolute uncertainty can grow significantly towards older ages. In contrast, the relative age uncertainty between two depths in the same archive typically remains moderate. As a result, the dating uncertainties of samples from different depths are correlated across the entire archive. These correlations complicate the mathematical treatment of the dating uncertainties and there exist only a few theoretical approaches to this issue (e.g. Comboul et al., 2014; Boers et al., 2017). In particular, the Greenland Ice Core Chronology 2005 (GICC05), the common chronology for the majority of deep Greenland ice core records, on which a plethora of paleoclimate studies is based, does not specify the joint dating uncertainties for the entire chronology. Instead, maximum bounds on the uncertainty of individual ages are reported (Vinther et al., 2006; Rasmussen et al., 2006; Andersen et al., 2006; Svensson et al., 2008).

Here we first detail a previously developed fully Bayesian model to generate ensembles of chronologies for the NGIRP ice core that represent the joint dating uncertainty of all data points (Myrvoll-Nilsen et al., 2022). We then address the incorporation of so-called tie-points in the model — points in the record, where external information on the record's age is available. We generalize the model for the application to arbitrary layer-counted archives and demonstrate how it can be synchronized to a given set of tie-points under full consideration of dating uncertainties.

The outline of this paper is as follows. Section 2 briefly introduces tie-points and motivates their importance in constraining layer-counted chronologies. Section 3 details how a statistical model can be used to quantify the dating uncertainties and simulate an ensemble of plausible chronologies. In Section 4 we show how tie-points and their associated uncertainty can be incorporated by defining another model for the discrepancy between the layer-counted chronology and the tie-points. Section 5 demonstrates the framework for obtaining an ensemble of synchronized chronologies, first using synthetic data and tie-points, then with a real data example using physically derived tie-points. Further discussion and conclusions are provided in Section 6.

2 Biases and tie-points

In any layer-counted chronology, there is the risk of an unknown systematic bias in the counting or deposition process. For example, investigators may mistakenly, yet consistently, interpret weak signals as annual layers that have in fact a different origin. Biases can cause the counted chronology to deviate from the true age substantially (see e.g Andersen et al., 2006, for a discussion on counting bias in ice cores). This risk may be mitigated by the incorporation of tie-points that constrain the record's age at specific depths (e.g. Adolphi et al., 2018; Svensson et al., 2020). Constraining errors by incorporating the information from such tie-points is particularly relevant in view of the cumulative nature of the uncertainties in layer-counted paleoclimate proxy archives.

Significant global or regional events in the past have left characteristic features in proxy records from different, potentially distant archives. If such features can be identified in both the layer-counted archive at hand and supplementary, independently dated paleoclimate records, they can serve as tie-points between the different chronologies. For example, Adolphi et al. (2018) used the signature of the so-called Laschamp event as a tie-point between the layer-counted GICC05 and supplementary radiocarbon based chronologies. Also, ash layers stemming from large volcanic eruptions may sometimes be used as tie-points (Cook et al., 2022).

The corresponding depth in the studied archive can then be assigned a 'corrected' age, which might deviate from the originally counted chronology. Taking into account the new constraints imposed by corrected ages, the entire chronology may be *synchronized* to a set of tie-points as illustrated in Figure 1 under consideration of all age correlations along the depth axis.

Three major challenges complicate the synchronization process. First, the dating uncertainty of the original and the supplementary chronologies must be taken into account properly. Second, the tie-points themselves are subject to considerable uncertainty themselves, because the underlying features are not sharply localized. For these two reasons, tie-points are ideally specified in terms of probability distributions. In view of the available external information, these distributions indicate the plausibility for the studied archive to have a certain age at the depth of the identified characteristic feature. Third, one must choose a suitable interpolation scheme between tie-points, which reflects possible correlations between ages and age uncertainties across the archive. This work addresses the third challenge and presupposes the availability of tie-points in a suitable format. We show how tie-points and their uncertainties can then be incorporated directly into a generalization of the theoretical framework of the Bayesian uncertainty quantification presented in Myrvoll-Nilsen et al. (2022).

3 Age–depth modeling

Before we introduce the synchronization procedure we briefly discuss how we can define a probabilistic age–depth model that captures the age and age-uncertainty correlations across the entire core. For each depth z_k in the core, a corresponding age y_k is measured along with various proxies or other covariates $\boldsymbol{w}_k$. We let $\boldsymbol{W} = (\boldsymbol{z}, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ denote the matrix with columns representing all covariates, including the depths $\boldsymbol{z} = (z_1, \ldots, z_n)$. Myrvoll-Nilsen et al. (2022) demonstrates how this information can be used in a Bayesian regression model in order to quantify the joint dating uncertainties at all depths. In essence, the approach assumes that the number of incremental layers $\Delta y_k = (y_k - y_{k-1})$ in a given depth increment Δz_k is the sum of a structured component a_k and a stochastic component ε_k

$$\Delta y_k = a_k + \varepsilon_k. \tag{1}$$

The structured component is meant to reflect large-scale processes like the thinning of annual layers down the core or the average precipitation regime, allowing different accumulation rates at different climatic periods as evident in the ice NGRIP ice core. The structured component can be specified as a linear predictor of the covariates

$$a_k := a(\boldsymbol{w}_k \mid \boldsymbol{\beta}) = \beta_0 + \sum_{i=1}^{n_\beta} \beta_i w_{ki} \tag{2}$$

with unknown effects $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{n_\beta})$.

The stochastic term represents counting errors and short-term fluctuations in the layer-formation process and is expressed using a random function of the depth z_k that depends on some parameters $\boldsymbol{\theta}$ and can capture structural correlation, such as an AR(1) process,

$$\varepsilon_k := \varepsilon(z_k \mid \boldsymbol{\theta}). \tag{3}$$

We have here, as in Myrvoll-Nilsen et al. (2022), assumed a single random effect, but this is not a requirement for the methodology we present in this paper. In this paper, we do not distinguish between certain and uncertain layers and instead treat all layers as subject to the same stochastic process, regardless of their perceived level of confidence.

Given likelihood function

$$\pi(\Delta \boldsymbol{y} \mid \boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{W}) = \prod_{i=1} \pi\left(\Delta y_i \mid \boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{W}\right), \tag{4}$$

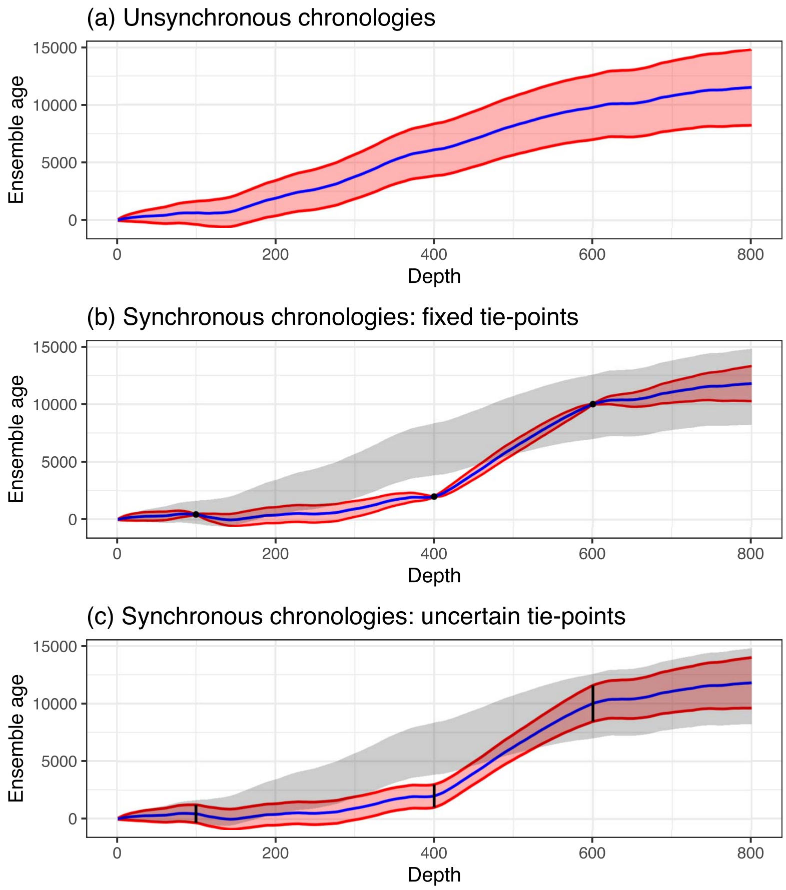

Figure 1: Illustrative example of how the dating uncertainty of a layer-counted chronology is affected by synchronization using external information expressed by tie-points as discussed in Section 2. Panel (a) shows the non-synchronized age estimate with uncertainty as a function of depth, panel (b) shows the same timescale synchronized using fixed sharp tie-points and panel (c) shows the timescale synchronized using uncertain tie-points. The blue lines represent the mean, the red area shows the 95% credible intervals, and the vertical black lines represent the 95% credible intervals of the tie-point distribution. The gray area in panels (b) and (c) provides a comparison with the 95% credible intervals of the unsynchronized timescale. The units on the x- and y-axes are arbitrary and could, for example, represent meters and years, respectively.

and an appropriately chosen prior distribution $\pi(\boldsymbol{\beta}, \boldsymbol{\theta})$ we can use a Bayesian modeling framework to infer a posterior distribution $\pi(\boldsymbol{\beta}, \boldsymbol{\theta} \mid \Delta \boldsymbol{y}, \boldsymbol{W})$ for the model parameters.

From the posterior distribution of the model parameters, we can simulate joint age increments $\Delta\boldsymbol{y}^*$ using the posterior predictive distribution

$$\pi(\Delta\boldsymbol{y}^* \mid \Delta\boldsymbol{y}, \boldsymbol{W}) = \int \pi(\Delta\boldsymbol{y}^* \mid \boldsymbol{\beta}, \boldsymbol{\theta}, \boldsymbol{W})\pi(\boldsymbol{\beta}, \boldsymbol{\theta} \mid \Delta\boldsymbol{y}, \boldsymbol{W})\mathrm{d}\boldsymbol{\beta}\mathrm{d}\boldsymbol{\theta}. \tag{5}$$

These samples reflect the inferred uncertainty and covariance structure from the observed layer-increments and covariates. Samples from the posterior predictive distribution of the age vector $\boldsymbol{y}$ can be obtained by first sampling

$$\Delta\boldsymbol{y}^* \sim \pi(\Delta\boldsymbol{y}^* \mid \Delta\boldsymbol{y}, \boldsymbol{W}) \tag{6}$$

and then take the cumulative sum $y_k^* = y_0 + \sum_{i=1} \Delta y_i^*$. We will consistently use * to denote output from the posterior predictive distribution.

4 Synchronization

Fitting a data set without the incorporation of external information leaves the resulting model susceptible to potential unknown structural biases in the layer counting or accumulation processes. Such biases can be revealed when comparing the estimated time scale to tie-points which are obtained by comparing the age of known climatic events to those of other records. This additional information could also be used to help constrain the dating uncertainties, producing more accurate estimates of the time scale. In this section, we demonstrate how an age–depth model of the above form can be synchronized with the chronologies of other archives by the use of tie-points at fixed depths.

For this section, and the remainder of the paper, we introduce a slight change in notation. The layer-counted data is denoted $\boldsymbol{y}_1 = (y_{1,1}, \ldots, y_{1,n})$ and the tie-points are denoted $\boldsymbol{y}_2 = (y_{2,1}, \ldots, y_{2,n})$. Both vectors are assumed to be sampled at the same depths $\boldsymbol{z} = (z_1, \ldots, z_n)$. The parameters associated with the layer-increment model defined in Section 3 are denoted $\boldsymbol{\theta}_1 = (\boldsymbol{\beta}, \boldsymbol{\theta})$. The number of observed values of $\boldsymbol{y}_2$ is typically much smaller than those of $\boldsymbol{y}_1$ which means that the amount of missing data in $\boldsymbol{y}_2$ will likely be substantial. Moreover, tie-points often come with uncertainties themselves, which must also be incorporated. These are assumed to be expressed by some known probability density function $\pi(\boldsymbol{y}_2 \mid \boldsymbol{\xi})$, where $\boldsymbol{\xi}$ denotes a set of model parameters or conditions that reflect the tie-point distribution. Assuming that the age information associated with the tie-points is indeed correct, our idea will be to impute the missing data in $\boldsymbol{y}_2$ using values obtained from statistical analysis. For this aim, we will supplement the unsynchronized model with an additional stochastic component that represents a counting bias.

Although the layer-counted data provides the best description of variation between increments, it is impossible to infer from it potential structural biases. For this we need to incorporate external information based on independently obtained tie-points $\boldsymbol{y}_2$. One simple way to incorporate this information is to condition the posterior predictive distribution, obtained from fitting the previous model to the layer-increment data, to the tie-points. A demonstration of this approach applied to Gaussian age–depth models is included in the supplementary material (Myrvoll-Nilsen et al., 2025). However, if the tie-points contradict the layer-increment model, i.e. they show that a statistically significant bias is present in the layer-increment model, then the posterior predictive model is, by itself, unable to model this discrepancy properly.

4.1 Age synchronization and bias estimation

The synchronization procedure proposed here consists of several steps as illustrated in Figure 2. To obtain a synchronous chronology we need to estimate the bias between the layer-counted process and the tie-points as described by the difference

$$\boldsymbol{\psi} = \boldsymbol{y}_2 - \boldsymbol{y}_1. \tag{7}$$

To model this bias we introduce an additional stochastic process that we will call the age discrepancy model

$$\psi_k := \psi(z_k \mid \boldsymbol{\theta}_2) \tag{8}$$

governed by some parameters $\boldsymbol{\theta}_2$. $\boldsymbol{y}_2$ can then be modeled in terms of the age increment model and age discrepancy model by

$$\boldsymbol{y}_2 = \boldsymbol{R}\Delta\boldsymbol{y}_1 + \boldsymbol{\psi}, \tag{9}$$

where $\boldsymbol{\psi} = (\psi_1, \ldots, \psi_n)^\top$ and

$$\boldsymbol{R} = \begin{pmatrix} 1 & & & \\ 1 & 1 & & \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \cdots & 1 \end{pmatrix} \tag{10}$$

represents the cumulative sums. The observed $\boldsymbol{\psi}$ will have the same amount of missing data as $\boldsymbol{y}_2$, but by fitting $\boldsymbol{\psi}$ using a Bayesian framework to the tie-points we can impute the missing values using the posterior predictive distribution. In fact, the original age–depth model and the age–discrepancy model can be fitted simultaneously to the joint data $(\Delta\boldsymbol{y}_1, \boldsymbol{y}_2)^\top$ using the following joint notation

$$\begin{pmatrix} \Delta\boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{pmatrix} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{R} & \boldsymbol{I} \end{pmatrix} \begin{pmatrix} \Delta\boldsymbol{y}_1 \\ \boldsymbol{\psi} \end{pmatrix}. \tag{11}$$

However, since the matrix $\boldsymbol{R}$ is dense, this will cause a substantial increase in computational time and memory cost, and could make the model prone to numerical instability for larger datasets. In cases where numerical stability presents an issue we suggest a simplification where we instead fit the models sequentially. A consequence of this is that the parameters of the layer-increment model, $\boldsymbol{\theta}_1$, will not be updated by the tie-point data $\boldsymbol{y}_2$. For the applications considered in this paper, we will use the simplified, sequential approach as it is computationally more straightforward and provides a sufficiently good estimate of all uncertainties involved. In cases where the age model significantly influences the age-increment model, a joint fit should be preferred to properly account for the interdependence between the models. A simulated example where a joint fit is performed using Markov chain Monte Carlo (MCMC) is included in Section 9 of the supplementary material.

In this case, in order to incorporate the uncertainty and statistical properties inferred from the layer-increment data, as well as the uncertainty associated with each tie-point, we use the following sampling based approach: After fitting a Bayesian model to the increments of layer-counted data $\Delta\boldsymbol{y}_1$, we generate samples of chronologies $\boldsymbol{y}_1^*$ from the posterior predictive distribution. If tie-points are uncertain, we generate samples from the distribution $\boldsymbol{y}_2 \sim \pi(\boldsymbol{y}_2 \mid \boldsymbol{\xi})$ and compute the age discrepancies

$$\boldsymbol{\psi} = \boldsymbol{y}_2 - \boldsymbol{y}_1^*. \tag{12}$$

For each $\boldsymbol{\psi}$, we fit the age discrepancy model to the sampled $\boldsymbol{\psi}$ using a Bayesian modeling framework. Then generate one sample $\boldsymbol{\psi}^*$ from the resulting posterior predictive distribution conditioned on $\boldsymbol{\psi}$ to impute the missing values. We add the simulated $\boldsymbol{\psi}^*$ to $\boldsymbol{y}_1^*$ to produce a sample for

$$\boldsymbol{y}_2^* = \boldsymbol{y}_1^* + \boldsymbol{\psi}^*. \tag{13}$$

This gives an uncertainty sensitive interpolation which is anchored in the given tie-points. Repeating this for all samples of $\boldsymbol{y}_1^*$ and $\boldsymbol{y}_2$ allows for an ensemble of synchronized chronologies that takes into account the uncertainty of the layer counting process, including biases, inferred from $\boldsymbol{y}_1$ as well as the uncertainties provided for the tie-points $\pi(\boldsymbol{y}_2 \mid \boldsymbol{\xi})$.

4.2 Age discrepancy model

There may be many different suitable models for the difference $\boldsymbol{\psi} = \boldsymbol{y}_2 - \boldsymbol{y}_1^*$. In this paper we focus on using a continuous 2nd order random walk (RW2) spline to describe a depth-dependent counting bias which is manifested as the discrepancy term $\boldsymbol{\psi}$. This is a continuous extension (Lindgren and Rue, 2008) of a stochastic spline model which assumes that the second-order increments are independent Gaussian variables

$$\psi_i - 2\psi_{i+1} + \psi_{i+2} \sim \mathcal{N}(0, \sigma_{\mathrm{RW2}}^2). \tag{14}$$

This model was selected since it is generally very good in situations where there is limited knowledge about the data. The model allows for the bias to change over long distances in the core, but is sufficiently smooth. There is a plethora of models that may be used instead. Which model is the most appropriate will depend on which assumptions can be made for the discrepancy and how much data is available. Once $\boldsymbol{\psi}$ has been fitted we can sample from it to produce synchronous chronologies $\boldsymbol{y}_2^*$ following (13).

We have chosen to implement this synchronization procedure using integrated nested Laplace approximations (INLAs) (Rue et al., 2009, 2017) using the R-INLA framework, available at `www.r-inla.org`. R-INLA provides a computationally efficient implementation of the RW2 model and missing values are imputed using the desired posterior predictive distribution automatically. It should be noted that using INLA in this step does not require that the layer-increments are fitted using INLA, nor that the increment model is Gaussian. As long as samples from the posterior predictive distribution is given, the discrepancies can be fitted with INLA, regardless of how the samples were obtained.

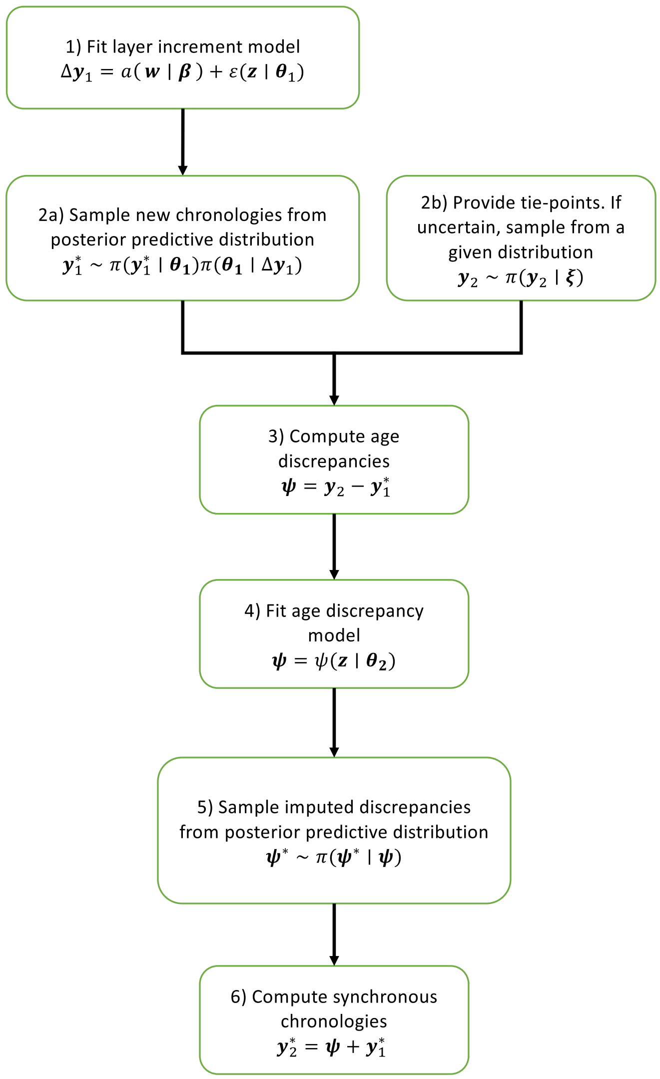

Figure 2: Diagram describing the process of fitting the layer increment and age discrepancy models to data as described in Section 4.1. The resulting model allows for the simulation of synchronous chronologies. These represent updated age–depth uncertainties, taking into account new information expressed by tie-points drawn from known probability density functions. Steps 3 to 6 are repeated for every chronology in the ensemble, while steps 1 and 2 are performed only once.

5 Examples

5.1 Simulation results

To demonstrate our procedure, we apply our model to synthetic age data with a specific structure that aims to reproduce features that we can expect from real-world applications. Specifically, the simulated core will represent depths on the interval 1–1000m at 1m resolution, leading to $n_1 = 1000$ data points. Moreover, we assume that the depth interval can be split into $p = 3$ accumulation regimes, with regime shifts at $z_0 = 1, z_1 = 250, z_2 = 750$ and $z_3 = 750$ m. We assume that the age is given by

$$y_k = y_0 + \sum_{i=1}^{k} \Delta y_i, \tag{15}$$

where

$$\Delta y_k = \beta_w z_k + \sum_{i=1}^{p} \gamma_i(z_k; b_i, c_i) + \varepsilon_k \tag{16}$$

and $\boldsymbol{\varepsilon}$ is drawn from an AR(1) process with standard deviation σ and lag-one correlation parameter ϕ. Similar to Myrvoll-Nilsen et al. (2022) we use piecewise defined functions γ_i to correspond to linear trends in different climate periods

$$\gamma_i(z_k; b_i, c_i) = \begin{cases} b_i + c_i z_k, & z_i^* < z_k < z_{i-1}^*, \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

Changing climatic conditions may impact the accumulation rates of paleoclimatic archives. For each of these periods, we assume different levels and linear trends of accumulation. In addition to depths $\boldsymbol{z}$ the covariates $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n)$ will also activate slopes c_i and intercepts b_i at the appropriate subintervals. The corresponding observed chronology starts at $y_0 = 0$, and will be constrained using $n_2 = 4$ independent tie-points. One fixed tie-point $\xi_1 = 0$ at 1 m, and three Gaussian tie-points with means $\boldsymbol{\mu_\xi} = (2200, 6800, 19000)$ yrs and standard deviations $\boldsymbol{\sigma_\xi} = (10, 50, 60)$ yrs, located at depths 200 m, 400 m and 800 m.

As a proof of concept, we generate a single synthetic chronology and then apply the uncertainty estimation procedure to the synthetic data. The hyperparameters and fixed effects used to generate the synthetic observations are listed in Table 1, along with the corresponding estimates obtained by fitting the model using `R-INLA`. The model fit of the layer increments and the dating uncertainty of the synchronized chronologies are illustrated in Figure 3a. All posterior marginal mean estimates appear to be relatively close to the true value. The true value is never outside of the 95% credible interval.

We follow two different strategies for synchronizing the chronologies. First, we assume that the counting is not corrupted by a bias. Hence, the information provided by the tie-points may be used to further constrain the model as is, i.e. without the addition of the age-discrepancy component. Heuristically, we condition the unsynchronized age–depth models on the ages indicated by the tie-points as shown in Figure 3b. Second,

Parameter	True	Posterior marginal mean	95% CI
σ	3	2.93	(2.68, 3.224)
ϕ	0.8	0.786	(0.745, 0.824)
β_w	0.5	0.397	(0.143, 0.651)
b_1	10	9.87	(7.86, 11.9)
c_1	0.01	0.0111	(−0.00211, 0.0245)
b_2	20	20.4	(17.9, 22.9)
c_2	0.02	0.0195	(0.0147, 0.0243)
b_3	0	−3.38	(−14.7, 7.87)
c_3	0.03	0.0344	(0.0215, 0.0474)

Table 1: Results for fitting an appropriate layer-increment model to a synthetic example with known parameters as presented in Section 5.1. True values of fixed effects and hyperparameters used to generate the simulated data, estimated posterior marginal mean and 95% credible intervals.

we consider the possibility that the counting procedure was corrupted by a bias. We thus use a continuous RW2 model to represent this bias and use the sampling scheme described in Section 4 to obtain an ensemble of synchronized chronologies. We produce $1,000$ simulations using each approach. The results from the first approach is illustrated in Figure 3b, which shows the mean and 95% credible intervals of the difference between the ensemble of synchronized chronologies and the synthetic reference time scale $\boldsymbol{y}$. Similarly, the results using the RW2 synchronization model is included in Figure 3c. We observe increased uncertainties for the second approach, due to additional uncertainties in the RW2 component. The first approach is justified if the age distributions associated with the tie-points fall well within the uncertainty bands of the unsynchronized age–depth model. In that case, we may regard the tie-points as additional constraints on the already well-tuned model. If, however, the age-distribution at different tie-points cannot be reconciled with the uncertainties of the unsynchronized model, the possibility of a corrupting counting bias must be considered and the model should be extended by the age-discrepancy component.

5.2 NGRIP ice core data

In this section, we apply our synchronization approach to the same data as studied in Myrvoll-Nilsen et al. (2022), namely the δ^{18}O record from the North Greenland Ice Core Project (NGRIP) (North Greenland Ice Core Project members, 2004; Gkinis et al., 2014). This record covers the last 120,000 years and is one of the most intensely studied records in paleoclimate research. The δ^{18}O ratio is commonly interpreted as a proxy for the local temperature at the time of precipitation and can be used to estimate temperature variations in Greenland during this time (Jouzel et al., 1997; Gkinis et al., 2014). The NGRIP record is defined on a layer-counted chronology, the Greenland ice core chronology 2005 (GICC05) (Vinther et al., 2006; Rasmussen et al., 2006; Andersen et al., 2006; Svensson et al., 2006). We consider the pre-Holocene interval 59,944–11,703 years before 2000 CE (yr b2k) which is available at a 5 cm depth resolution and comprises n = 19,475 points of data. This spatially equidistant resolution translates into

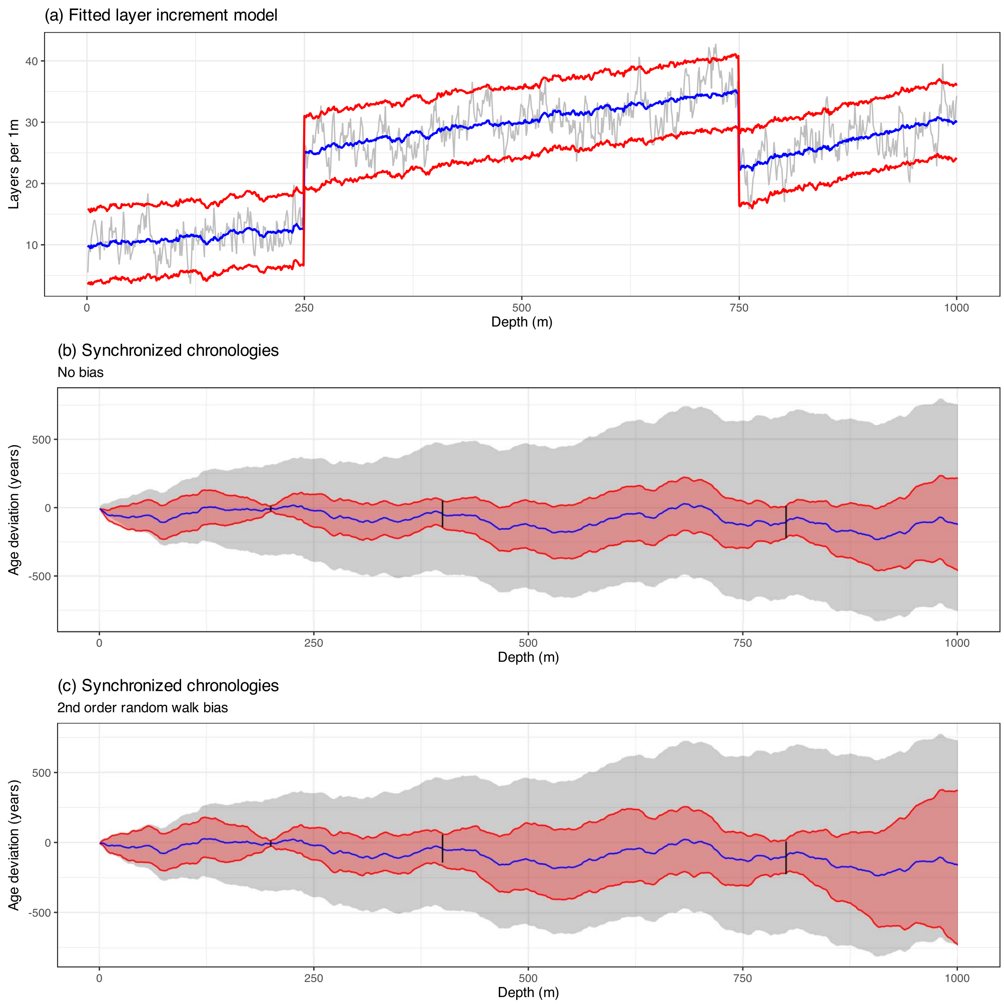

Figure 3: Results for the simulated example presented in Section 5.1 Panel (a): Layer-increment model fit to synthetic data (gray). Posterior marginal mean (blue) and 95% credible intervals (red). Panels (b) and (c): The resulting age–depth model is synchronized to four tie-points: one fixed and three with uncertainty expressed by Gaussian distributions. The posterior marginal mean is shown in blue and the 95% credible intervals of the simulated chronologies are indicated in shaded red. The cumulative sum of the observed layer increments has been subtracted. Black vertical lines represent the 95% credible intervals of the tie-point distributions, and the gray area represent the 95% credible intervals of the unsynchronized timescale. Panel (b) Assumes that there is no structural bias in the layer increments and that tie-points can be considered realizations of the age–depth model, and are obtained by conditioning the free variables on the observed tie-points. Panel (c) assumes that there is structural bias causing a discrepancy between the age inferred from the observed layer increments and the true timescale, and models this discrepancy using a 2nd order random walk model.

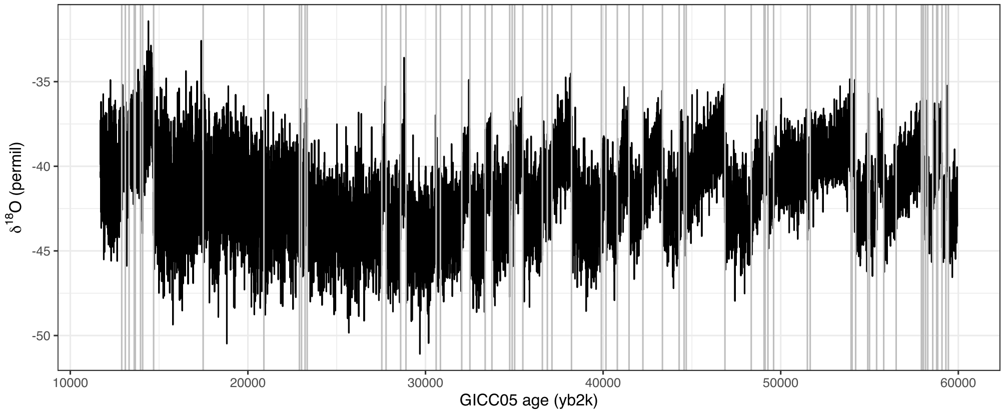

Figure 4: Real data example as discussed in Section 5.2: NGRIP δ^{18}O stable isotope ratios plotted against the corresponding GICC05 time scale, for the period from 11,000 to 60,000 years before the year 2000 AD (North Greenland Ice Core Project members, 2004; Gkinis et al., 2014). Note that following common conventions in paleoclimate studies, age increases from right to left. Gray vertical lines indicate transitions between Greenland stadial and interstadial periods as reported in Rasmussen et al. (2014).

about 7 years of resolution at the beginning of this period and approximately annual resolution at its end. Each depth z_k is accompanied by a δ^{18}O ratio and a cumulative layer count, that is, an age y_k. Figure 4 shows the NGRIP δ^{18}O ratios against the GICC05 chronology.

We adopt the same age model used in Myrvoll-Nilsen et al. (2022) based on the incremental counted layers per depth interval. The 'physical' component therein is given by

$$a(z_k \mid \boldsymbol{\beta}) = \beta z_k^2 + \beta_x x_k + \sum_{i=1}^{p} \gamma_i(z_k; b_i, c_i). \tag{18}$$

Here, x_k is the observed δ^{18}O ratio at depth z_k and $\sum_{i=1}^{p} \gamma_i(z_k; b_i, c_i)$ is a sum of piece-wise linear functions. Each $\gamma_i(z_k; b_i, c_i)$ represents an individual linear trend that prevails for one of the p climatic stadial or interstadial periods separated by pronounced climatic transitions. The timings of these transitions are given by Table 2 in Rasmussen et al. (2014) and are indicated in Figure 4 by gray lines. The fixed effects are denoted by $\boldsymbol{\beta} = (\beta, \beta_w, \{b_i\}, \{c_i\})$. The noise component $\boldsymbol{\varepsilon}$ is assumed to follow an AR(1) process

$$\varepsilon_k = \phi\varepsilon_{k-1} + \omega_k, \quad \omega_k \sim \mathcal{N}\left(0, \frac{\sigma_\varepsilon^2}{1-\phi^2}\right), \tag{19}$$

which depends on two parameters $\boldsymbol{\theta} = (\sigma_\varepsilon, \phi)$, the standard deviation σ_ε and the lag-one correlation parameter ϕ.

We fit the model to the GICC05 timescale as described in Section 3. The fit returns the posterior distributions for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$. To synchronize the GICC05 we use four uncertain

	GICC05 age (yb2k)	NGRIP depth (m)	Offset mean	95% CI
1	12,050	1502.8	−7.04	(−30.92, 11.94)
2	13,050	1532.7	3.34	(−12.37, 21.99)
3	22,050	1767.25	521.44	(336.19, 714.03)
4	42,050	2132.4	275.24	(101.51, 509.08)

Table 2: Summary of tie-points taken from Adolphi et al. (2018) and Muscheler et al. (2020) used in the real data example. The table shows their location on the GICC05 time scale and the corresponding NGRIP depth (rounded to closest 5cm). The Mean and 95% quantile intervals for the age discrepancy from the GICC05 chronology are included.

(non-Gaussian) tie-points. These were obtained by Adolphi et al. (2018) and Muscheler et al. (2020) by assessing common features in the record of cosmogenic radionuclides that are evident both in the NGRIP record and in different independently dated archives. The tie-point uncertainties are summarized in Table 2, where their location is given by their age in the GICC05 timescale and the corresponding NGRIP depth $z_{t_1}, \dots, z_{t_4}$.

From Figure 5 it is evident that tie-points 3 and 4 are significantly different from the counted time scale, suggesting a clear bias in the counting process. After sampling an ensemble of $1,000$ layer-counted chronologies from the unsynchronized age–depth model and tie-point tuples $(y_{2,t_1}, \dots, y_{2,t_4})$ from the Adolphi/Muscheler distributions at depths $z_{t_1}, \dots, z_{t_4}$, we perform Bayesian fits using the RW2 synchronization model for each ensemble member. The resulting ensemble of synchronized chronologies is illustrated in Figure 6. We observe a noticeable shift implying that the GICC05 tends to underestimate the age starting from around 1540m of depth, as already pointed out by Adolphi et al. (2018). Although the deviation from the GICC05 shrinks over time, the original counted age remains outside the 95% highest posterior density (HPD) intervals over the remaining course of the record. These results seem to confirm that there is indeed a structural bias in the layer-counting process, assuming that the tie-points provided by Adolphi et al. (2018) and Muscheler et al. (2020) are accurate.

6 Discussion

We have shown how tie-points, whose uncertainty is expressed through a probability distribution, can be incorporated into a Bayesian age–depth model for layer-counted proxy archives. We use a Bayesian model both to express the uncertainties in the observed layer increments and to express the discrepancy between chronologies sampled from an unsynchronized model and the tie-points. These can be fitted jointly using, for example, an MCMC approach. To improve computational efficiency and stability, we present a simplification where these models are fitted separately. By sampling from the fitted layer-increment and age-discrepancy models, it is possible to produce an ensemble of synchronized chronologies which align with the observed tie-points. Moreover, fitting the age-discrepancy model to an ensemble of tie-points using a Monte Carlo approach allows for the uncertainties of the tie-points to be incorporated as well. We have applied this methodology to two examples. A synthetic example where the observed layer count

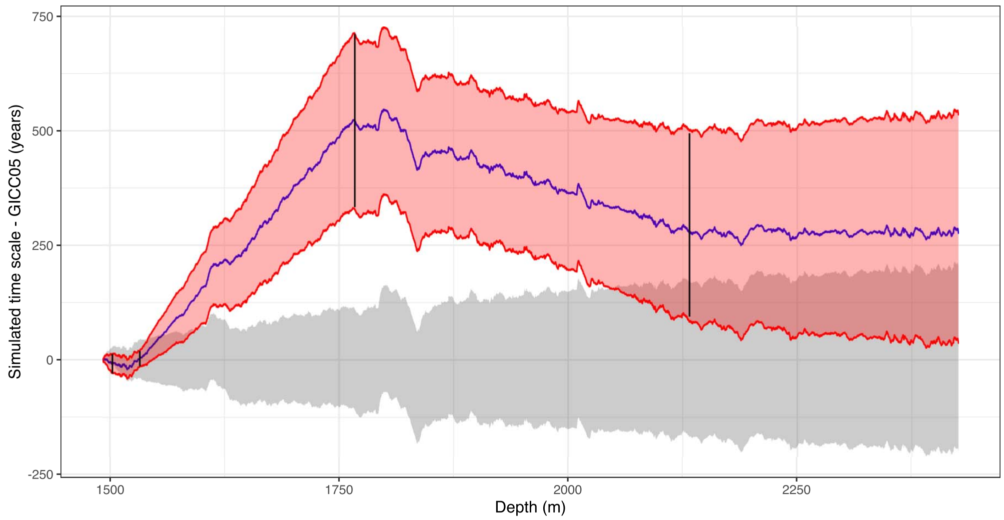

Figure 5: The estimated dating uncertainties of the GICC05 timescale synchronized to Adolphi et al. (2018) and Muscheler et al. (2020) tie-points, obtained from simulating 10,000 chronologies, under the assumption that there is no discrepancy between the age–depth model and the true chronology. The blue line represents the posterior marginal mean (with the GICC05 subtracted), the red shaded area represent the 95% highest posterior density (HPD) intervals, the vertical black lines show the 95% HPD intervals of the tie-points and the shaded gray area shows the uncertanties for the unsynchronized timescale.

is generated in a simulation and the corresponding model is synchronized to Gaussian tie-points, and a real data example where we synchronized the GICC05 chronology of the NGRIP ice core record to tie-points taken from Adolphi et al. (2018) and Muscheler et al. (2020).

The time-scale synchronization as proposed here addresses the concern of possibly unknown biases in the manual counting process. Such biases can cause the number of counted layers to deviate from the true age and, in turn, affect analyses performed on the layer-counted chronology. Due to the cumulative nature of the dating process via counting layers, such biases can potentially lead to severe errors in the age estimates of the deeper parts of the record. By matching the chronology to tie-points, it is possible to constrain potential systematic counting biases. Moreover, if tie-points are accurately estimated, then they can also reduce the dating uncertainty, thus making the chronology and all results based on it more reliable.

Tie-points may be very sparsely distributed. This presents a challenge with our framework, as fitting a model to the observed offsets may be challenging due to lack of data. It is important to choose an appropriate discrepancy model based on the amount of data available and on the assumptions that can be made about the process. Moreover, the synchronized chronologies are only as accurate as the tie-points used to constrain the

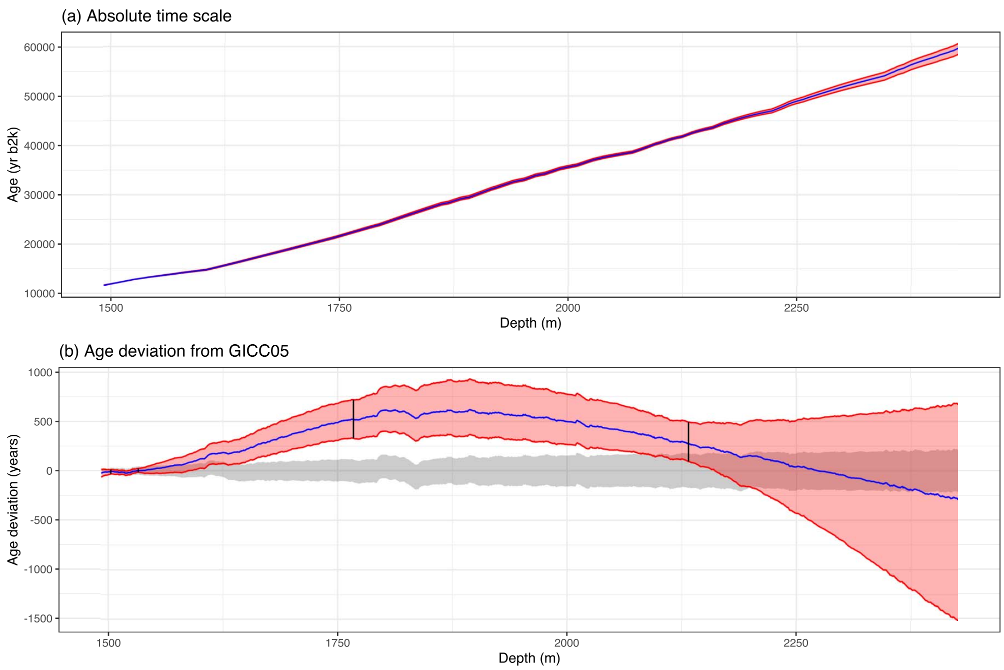

Figure 6: Similar to Figure 5, but describing the discrepancy between the age–depth model and the true time scale, as represented by the tie-points, using a 2nd order random walk model. The posterior marginal mean (blue) and 95% HPD credible intervals of the simulated chronologies, synchronized to tie-points (vertical black) with a 2nd order random walk model. The top panel (a) shows the absolute ages, and the bottom panel (b) shows the ages with the GICC05 subtracted compared to the unsynchronized time scale (shaded gray area).

uncertainties. If the tie-points are flawed, e.g. if they are biased or based on inaccurate information, then this will affect the accuracy of the synchronized chronology. It should be emphasized that the proposed methodology treats the tie-points as ground truth and prioritizes them over the previously available information. Weighing both the external tie-points as well as the unsynchronized chronology equally strong would require a different methodology.

To make our methodology accessible we have created an R package that takes care of both the fitting, synchronization and simulation processes. The package is called `bremla` (**B**ayesian **re**gression **m**odeling of **l**ayer-counted **a**rchives) and presents a user-friendly interface for the methodology presented in this paper. Currently, only the AR(1), RW1 and RW2 synchronization models are incorporated in `bremla`. Other Gaussian models are supported, but require specific implementation either by using the R-INLA framework, or by using `rgeneric`, `R-INLA`'s custom modeling framework. The package performs a full Bayesian analysis of the accumulation model using INLA, allows for an ensemble of chronologies (both synchronized and unsynchronized) to be simulated, and includes functions that plot and summarize relevant results. The package and its accompanying documentation can be downloaded from the Github repository `https://github.com/eirikmn/bremla`.

Acknowledgments

The authors would like to thank the anonymous referees, an Associate Editor and the Editor for their constructive comments that improved the quality of this paper. They would also like to thank Florian Adolphi for providing the tie-point probability distributions presented in Adolphi et al. (2018) and Muscheler et al. (2020).

Funding

This is ClimTip contribution #36; The ClimTip project has received funding from the European Union's research and innovation programme under grant agreement No. 101137601. NB acknowledges further funding by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 956170, and by the Volkswagen foundation. EMN has also received funding from the Norwegian Research Council (IKTPLUSS-IKT og digital innovasjon, project No. 332901).

Supplementary Material

Supplement to "Synchronization of layer-counted paleoclimatic proxy archives using a Bayesian regression modeling framework" (DOI: 10.1214/25-BA1509SUPP; .pdf). Provides details on how to perform naïve synchronization of Gaussian layer-increment models, and a demonstration of fitting the layer-increment and age-discrepancy models jointly using an MCMC approach.

References

Adolphi, F., Bronk Ramsey, C., Erhardt, T., Lawrence Edwards, R., Cheng, H., Turney, C. S., Cooper, A., Svensson, A., Rasmussen, S. O., Fischer, H., and Muscheler, R. (2018). "Connecting the Greenland ice-core and U/Th timescales via cosmogenic radionuclides: Testing the synchroneity of Dansgaard-Oeschger events." *Climate of the Past*, 14(11): 1755–1781. 2, 3, 14, 15, 17

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. (2006). "The Greenland Ice Core Chronology 2005, 15–42 ka. Part 1: constructing the time scale." *Quaternary Science Reviews*, 25(23-24): 3246–3257. 2, 3, 11

Blaauw, M. and Christen, J. A. (2011). "Flexible paleoclimate age-depth models using an autoregressive gamma process." *Bayesian Analysis*, 6(3): 457–474. MR2843540. doi: `https://doi.org/10.1214/ba/1339616472`. 2

Boers, N., Goswami, B., and Ghil, M. (2017). "A complete representation of uncertainties in layer-counted paleoclimatic archives." *Climate of the Past*, 13(9): 1169–1190. 2

Bronk Ramsey, C. (2008). "Deposition models for chronological records." *Quaternary Science Reviews*, 27(1-2): 42–60. 2

Comboul, M., Emile-Geay, J., Evans, M. N., Mirnateghi, N., Cobb, K. M., and Thompson, D. M. (2014). "A probabilistic model of chronological errors in layer-counted climate proxies: Applications to annually banded coral archives." *Climate of the Past*, 10(2): 825–841. 2

Cook, E., Abbott, P. M., Pearce, N. J., Mojtabavi, S., Svensson, A., Bourne, A. J., Rasmussen, S. O., Seierstad, I. K., Vinther, B. M., Harrison, J., Street, E., Steffensen, J. P., Wilhelms, F., and Davies, S. M. (2022). "Volcanism and the Greenland ice cores: A new tephrochronological framework for the last glacial-interglacial transition (LGIT) based on cryptotephra deposits in three ice cores." *Quaternary Science Reviews*, 292: 107596. 3

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. (2020). "Synchronous timing of abrupt climate changes during the last glacial period." *Science*, 369(6506): 963–969. 2

Gkinis, V., Simonsen, S. B., Buchardt, S. L., White, J. W., and Vinther, B. M. (2014). "Water isotope diffusion rates from the NorthGRIP ice core for the last 16,000 years – Glaciological and paleoclimatic implications." *Earth and Planetary Science Letters*, 405: 132–141. 11, 13

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. (1997). "Validity of the temperature reconstruction from water isotopes in ice cores." *Journal of Geophysical Research: Oceans*, 102(C12): 26471–26487. 11

Lindgren, F. and Rue, H. (2008). "On the second-order random walk model for irregular locations." *Scandinavian Journal of Statistics*, 35(4): 691–700.
URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9469.2008.00610.x` MR2468870. doi: `https://doi.org/10.1111/j.1467-9469.2008.00610.x`. 8

Mann, M. E. and Jones, P. D. (2003). "Global surface temperatures over the past two millennia." *Geophysical Research Letters*, 30(15). 2

Mann, M. E., Zhang, Z., Hughes, M. K., Bradley, R. S., Miller, S. K., Rutherford, S., and Ni, F. (2008). "Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia." *Proceedings of the National Academy of Sciences*, 105(36): 13252–13257. 2

McKay, N., Emile-Geay, J., and Khider, D. (2021). "geoChronR – an R package to model, analyze, and visualize age-uncertain data." *Geochronology*, 3: 149–169. 2

Muscheler, R., Adolphi, F., Heaton, T. J., Bronk Ramsey, C., Svensson, A., van der Plicht, J., and Reimer, P. J. (2020). "Testing and improving the IntCal20 calibration curve with independent records." *Radiocarbon*, 62(4): 1079–1094. 14, 15, 17

Myrvoll-Nilsen, E., Riechers, K., Rypdal, M. W., and Boers, N. (2022). "Comprehensive uncertainty estimation of the timing of Greenland warmings in the Greenland ice core records." *Climate of the Past*, 18: 1275–1294. 2, 4, 10, 11, 13

Myrvoll-Nilsen, E., Riechers, K., and Boers, N. (2025). "Supplement to "Synchronization of layer-counted paleoclimatic proxy archives using a Bayesian regression modeling framework"." doi: `https://doi.org/10.1214/25-BA1509SUPP`. 7

North Greenland Ice Core Project members (2004). "High-resolution record of Northern Hemisphere climate extending into the last interglacial period." *Nature*, 431(7005): 147–151. 11, 13

Parker, D., Legg, T., and Folland, C. (1992). "A new daily Central England Temperature Series." *Int. J. Clim.*, 12. 1

Parnell, A. C., Haslett, J., Allen, J. R., Buck, C. E., and Huntley, B. (2008). "A flexible approach to assessing synchroneity of past events using Bayesian reconstructions of sedimentation history." *Quaternary Science Reviews*, 27(19-20): 1872–1885. 2

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. (2006). "A new Greenland ice core chronology for the last glacial termination." *Journal of Geophysical Research Atmospheres*, 111(6): 1–16. 2, 11

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. (2014). "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 Science Reviews*, 106: 14–28. 13

Riechers, K. and Boers, N. (2021). "Significance of uncertain phasing between the onsets of stadial–interstadial transitions in different Greenland ice core proxies." *Climate of the Past*, 17(4): 1751–1775. 2

Ruddiman, W. F. (2014). *Earth's Climate - Past and Future*. New York: W. H. Freeman and Company, 3rd edition. 2

Rue, H., Martino, S., and Chopin, N. (2009). "Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion)." *J R Stat Soc Series B*, 71: 319–392. MR2649602. doi: `https://doi.org/10.1111/j.1467-9868.2008.00700.x`. 8

Rue, H., Riebler, A., Sørbye, S. H., Illian, J. B., Simpson, D. P., and Lindgren, F. K.

(2017). "Bayesian computing with INLA: A review." *Annu. Rev. Stat. Appl.*, 4(1): 395–421. MR3634300. doi: https://doi.org/10.1214/16-STS576. 8

Svensson, A., Andersen, K. K., Bigler, M., Clausen, H. B., Dahl-Jensen, D., Davies, S. M., Johnsen, S. J., Muscheler, R., Rasmussen, S. O., Röthlisberger, R., Peder Steffensen, J., and Vinther, B. M. (2006). "The Greenland Ice Core Chronology 2005, 15–42 ka. Part 2: comparison to other records." *Quaternary Science Reviews*, 25(23-24): 3258–3267. 11

Svensson, A., Andersen, K. K., Bigler, M., Clausen, H. B., Dahl-Jensen, D., Johnsen, S. J., Rasmussen, S. O., Seierstad, I., Steffensen, J. P., Svensson, A., Vinther, B. M., Davies, S. M., Muscheler, R., Parrenin, F., and Röthlisberger, R. (2008). "A 60,000 year Greenland stratigraphic ice core chronology." *Climate of the Past Discussions*, 3(6): 1235–1260. MR0585238. doi: https://doi.org/10.7146/math.scand.a-11858. 2

Svensson, A., Dahl-Jensen, D., Steffensen, J. P., Blunier, T., Rasmussen, S., Vinther, B., Vallelonga, P., Capron, E., Gkinis, V., Cook, E., Kjær, H. A., Muscheler, R., Kipfstuhl, S., Wilhelms, F., Stocker, T., Fischer, H., Adolphi, F., Erhardt, T., Sigl, M., Landais, A., Parrenin, F., Buizert, C., McConnell, J., Severi, M., Mulvaney, R., and Bigler, M. (2020). "Bipolar volcanic synchronization of abrupt climate change in Greenland and Antarctic ice cores during the last glacial period." *Climate of the Past*, 16: 1565–1580. 2, 3

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. (2006). "A synchronized dating of three Greenland ice cores throughout the Holocene." *Journal of Geophysical Research Atmospheres*, 111(13): 1–11. MR0943584. doi: https://doi.org/10.1080/03461238.1987.10413831. 2, 11