Point-by-point responses to the referees' comments

Response to Anonymous Referee # 1:

We thank the reviewer for her/his positive evaluation of our manuscript. Regarding your two main points:

1. We agree that a representation of the temporal evolution of uncertainties in terms of $\hat{x}$ is misleading, and our main point is that the sequence of probability density functions $p(x|t)$ themselves should in fact be the focus of representing this evolution. This point was already mentioned in the manuscript, but in the revised version, we have emphasized it even further.

2. The blue shadings in Fig. 2 indicate the probability densities $p(x|t)$ for each t, and not $p(t|z)$; this is also noted in the color bar label. The densities $p(x|t)$ are the final results of our approach to represent the record as probability densities (over the proxy values x) for different, error-free ages t. The dating uncertainty distribution $p(t|z)$ is used in the derivation of $p(x|t)$ via Eq. (1). Furthermore, the increase of dating uncertainties towards the past, which is apparent in Fig. 1B, is reflected in the increasing spread of $p(x|t)$ the further one goes into the past in Figs. 2C and 2D. This result is not trivial: it is a consequence of the Bayesian approach we employ. We have clarified this point further in the revised version.

Regarding the minor points in the Review:

For each annual layer, it is indeed possible that the uncertainty distributions are skewed, just as Referee #1 points out. In the original study (Andersen et al., QSR, 2006) reporting the chronology employed herein, uncertain layers are counted as $1/2 \pm 1/2$, thereby assuming a symmetric distribution. This counting should be adjusted in cases where the probability to miss a layer and the probability to count a false layer are not identical. The maximum counting error would then generally not be the same for negative and positive values, and the overall uncertainty distribution $p(t|z)$, which would be accordingly skewed, should be used when computing $p(x|t)$.

We would like to emphasize here that any functional form for $p(t|z)$ can be used in our approach. We thank the reviewer for pointing this out, and have added a corresponding sentence in the revised manuscript.

P2L2: We have corrected this in the revised manuscript.

P2L9: We agree with the reviewer that in our approach to represent dating uncertainties, abrupt transitions that actually exist in the proxy evolution will be typically smoothed out in accordance with the uncertain dating, rather than being artificially amplified. However, in traditional proxy record representations, proxy values are shown at specific time points. Ignoring the uncertainties of these time points may lead to situations where transitions appear much sharper than is actually supported by the data themselves when the dating uncertainties are considered.

P5L16: The Riemann sum, which is used to approximate the integral in this discrete setting, is defined using $r_i = (z_{i+1} - z_{i-1})/2$. This is the average of the two increments above and below the depth z_i, which the reviewer refers to as Δz_i: Setting $\Delta z_i := z_{i+1} - z_i$, we have $r_i = (z_{i+1} - z_{i-1})/2 = ((z_{i+1} - z_i) + (z_i - z_{i-1}))/2 = (\Delta z_i + \Delta z_{i-1})/2$. Taking this average is the standard approach when approximating a (continuous) integral by a discrete sum. It provides a better approximation than taking only the previous (or the following) increment. We have clarified this in the revised version.

P5L19: Thank you, this has been corrected.

P7L10: Please refer to our author comment AC1. We had, due to a typo, in fact uploaded an erroneous version of Fig. 2. In the figure attached to AC1, which is also used in the revised version, small-scale differences between Fig. 2C and Fig. 2D are clearly visible.

Figure 1A: We have, in the revised version of the manuscript, replaced the 2D figure by a 3D figure for the sketch of our method, and we appreciate the reviewer's effort in suggesting one. It might, however, be slightly misleading to use 3D cartesian coordinates because the z-axis is integrated over and thus plays a different role than the x- and t-axes. In order to avoid confusion, we have added a comment in the caption of the figure.

Figure 1B: As the reviewer notes correctly, this change in slope occurs at the transition from glacial to interglacial conditions, and reflects the substantial increase in dating uncertainties at this point, due to changing accumulation rates and increasing pressure in the ice. We have added a corresponding sentence in the revision.

Figure 2: We thank the reviewer for pointing this out to us and use identical
y-axes in the revised version of our study.

Response to Anonymous Referee # 2

We thank the reviewer for the thorough evaluation of our manuscript. Regarding her/his specific comments:

> We agree that, realistically, the probability of missing an annual layer is generally not identical to counting a false one. However, for both the NGRIP and the Suigetsu datasets, the reported dating uncertainties are based on the assumption of symmetric counting errors. It would be interesting to repeat the annual layer counting for these records, taking into account that the error distributions are most likely not symmetric. Non-symmetric errors would lead to different maximum counting errors for negative and positive values, and hence to the overall (i.e., cumulative) dating uncertainty distributions $p(t|z)$ being skewed. This skewness of $p(t|z)$ could then be used when applying our method in just the same way, as pointed out by Referee #1. Ideally, the correct $p(t|z)$ would be reported alongside with the record itself; our formalism can treat arbitrary dating uncertainty distributions, and its specific characteristics will be propagated to the final $p(x|t)$. We have added a corresponding comment in the revised version of the manuscript.

p8 lines 15-16: We agree with the reviewer. A tephra layer would indeed reset the MCE back to zero, which would be reflected in very narrow dating uncertainty distributions $p(t|z)$ around such a layer. The spread in the derived $p(x|t)$ would decrease accordingly in Figs. 2C, 2D, 4B and 4D. We have added a sentence on this point in the revision.

Abs line 10: We apologize for the incorrect usage of "calibration curve," and will replace it by "comparison curve" throughout the revised version. We will also change our suggestion to actually use the Suigetsu record as a calibration curve, since we do agree that IntCal should be used for such a purpose. In the revised manuscript, we still show how to use our Bayesian method to derive the overall uncertainties of the radiocarbon ages — including the errors from the layer-counting process — and propose that this could also be done for the IntCal calibration curve.

p3 line 31&33: In fact, we use the notation $\Delta^{14}C$ for the "inferred level of radiocarbon in the atmosphere," given in **per mil**, as deviations from the reference value of 1950, in the same way as in the associated reference (Bronk Ramsey et al., Science, 2012). We agree that, to infer these values, information about timing needs to be included. For the

4

time interval under study, the Suigetsu chronology is derived from the varve-counting process, and this temporal information is used to infer past levels of atmospheric radiocarbon (i.e. Δ^{14}C), as deviation from the 1950 reference value. Thus, this record consists of inferred Δ^{14}C values, associated with corresponding sediment depth values and ages, including age uncertainties. This is all the information we need in order to apply our Bayesian approach to derive $p(x|t)$, as shown in Fig. 4B.

p2 line 20-22: We agree with the reviewer that several studies have presented approaches to visualize dating uncertainties of paleoclimatic proxies. We already cited two studies by M. Blaauw (2010 & 1012), but will also include the 2007 reference in the revised version. In the revised manuscript, we also note more explicitly that the idea of directly visualizing dating uncertainties is not new. However, we would like to emphasize that, to the best of our knowledge, visualization of proxy record uncertainties (in the style of Blaauw et al., 2007) is not widely used in the geoscientific community that deals with layer-counted proxy archives. With our proposed approach, we are not only able to visualise the proxy record uncertainties, but also to *quantify* them in a mathematically precise sense: $p(x|t)$ yields the best estimation of x at time t, given the observed data and their uncertainties. Therefore, the derived $p(x|t)$ series can also be used for further, quantitative analyses.

p2 line 3: The reviewer is right, we have corrected this in the revision.

p2 line 19: By "bifurcation parameters", we refer to the estimated parameters of an energy-balance model in the corresponding reference. We have rephrased the sentence accordingly.

p5 line 14: A Riemann sum is the standard technique for approximating a (continuous) integral, given discrete values of the function to be integrated. We have clarified this in the revised manuscript and provide a standard, first-year calculus reference.

p5 line 19: Thank you, this has been corrected in the revised version.

A complete representation of uncertainties in layer-counted paleoclimatic archives

Niklas Boers[1,2], Bedartha Goswami[3], and Michael Ghil[1,4]

[1]Geosciences Department and Laboratoire de Météorologie Dynamique (CNRS and IPSL), École Normale Supérieure, and PSL Research University, Paris, France.
[2]Potsdam Institute for Climate Impact Research, Potsdam, Brandenburg, Germany.
[3]Institute of Earth and Environmental Science, University of Potsdam, Potsdam, Germany.
[4]Department of Atmospheric and Oceanic Sciences and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, USA.

Correspondence to: Niklas Boers (nboers@lmd.ens.fr)

Abstract. Accurate time series representation of paleoclimatic proxy records is challenging because such records involve dating errors in addition to proxy measurement errors. Rigorous attention is rarely given to age uncertainties in paleoclimatic research, although the latter can severely bias the results of proxy record analysis. Here, we introduce a Bayesian approach to represent layer-counted proxy records – such as ice cores, sediments, corals or tree rings – as sequences of probability distributions on absolute, error-free time axes. The method accounts for both proxy measurement errors and uncertainties arising from layer-counting–based dating of the records. An application to oxygen isotope ratios from the North Greenland Ice Core Project (NGRIP) record reveals that the counting errors, although seemingly small, lead to substantial uncertainties in the final representation of the oxygen isotope ratios. In particular, for the older parts of the NGRIP record, our results show that the total uncertainty originating from dating errors has been seriously underestimated. Our method is next applied to deriving the overall uncertainties of the Suigetsu radiocarbon ~~calibration~~ comparison curve, which was recently obtained from varved sediment cores at Lake Suigetsu, Japan. This curve provides the only terrestrial radiocarbon ~~calibration~~ comparison for the time interval 12.5–52.8 kyr BP. The uncertainties derived here can be readily employed to obtain complete error estimates for arbitrary radiometrically dated proxy records of this recent part of the last glacial interval.

1 Introduction

Time series derived from paleoclimatic proxy records — such as tree rings, ice or marine sediment cores — provide the only available means for quantitative analyses of climate variability on time scales that exceed the last approximately 150 years (Mann and Jones, 2003; Mann et al., 2008). Such records yield indirect information of the past evolution of climatic observables like temperature, carbon concentration or precipitation. This information, however, is provided as a function of the depth in the record, rather than as a function of time. Depending on the type of proxy record, establishing a precise relationship between

depth and time — i.e., the age–depth model — is a highly nontrivial task (Telford et al., 2004). Widely used techniques include absolute dating methods such as radiometric dating, but also incremental dating methods, such as counting annual layers~~, also called varves~~.

In typical age modeling frameworks, the assignment of timestamps to given depths in the proxy core is irregular in time, involves possibly correlated uncertainties, and can in many cases only be performed for a subset of the proxy measurements. The interpolations that become necessary therewith further enhance the final uncertainties. Therefore, in addition to the errors associated with the measurement of a proxy value, there are also substantial uncertainties associated with its dating.

Both types of uncertainties have to be rigorously dealt with, since any conclusions drawn from the resulting time series will strongly depend on the statistical treatment of these uncertainties. ~~For example,~~ Thus, in traditional proxy record representations, proxy values are shown at specific points in time. Ignoring the uncertainties in specifying these points may lead to situations where the sharpness of transitions in the temporal evolution of the record is strongly under- or overestimated. Therefore, abrupt transitions that actually exist in the proxy evolution might be smoothed out. On the other hand, apparent abrupt transitions ~~may just be~~ might likewise be just artifacts of the age–depth model ~~construction~~ being used. Furthermore, particular care needs to be exercised when comparing proxy records obtained from different archives with independent age models. In many situations ~~,~~ — e.g., when the accumulation process of the record is taken into account ~~,~~ or for incremental dating techniques ~~,~~ — uncertainty distributions at distinct locations of a given record will not be statistically independent.

Surprisingly, out of 93 publications from the year 2008 which involve age–depth models, 65 do not specify if and how uncertainties have been accounted for in the age–depth modeling process (Blaauw, 2010). Scholz et al. (2012) have compared five recent methods to date speleothem records that estimate dating uncertainties also between dated points. For example, such uncertainties can be considered on the basis of Monte Carlo simulations, which fit ensembles of straight lines between the dated points (Scholz and Hoffmann, 2011). Alternatively, one can use a mechanistic-statistical model that combines a deterministic paleoclimate model with a statistical model of the observation process to obtain ~~bifurcation parameters~~ the optimal parameters of an energy-balance model with known error bars (Roques et al., 2014). Furthermore, Breitenbach et al. (2012) have recently proposed a methodology to treat dating uncertainty propagation that is based on repeated interpolation between dated points, and to use the resulting ensemble of age-depth models to translate the uncertainties from the age axis to the proxy axis.

Most recently, Goswami et al. (2014) introduced a Bayesian framework to treat correlated errors in radiometrically determined chronologies. A processed proxy time series resulting from this approach consists of a sequence of probability densities whose domain covers the proxy values, conditioned on prescribed age values. This sequence is represented on an absolute, error-free time axis, a setting that is helpful in many situations. For instance, quantifying rates of change in the time series during prescribed time intervals calls for such an error-free time axis. In addition, an absolute time axis is of the essence when comparing two or more proxy archives with independent chronologies and, in particular, when analyzing the synchronicity of specific events in such archives (Blaauw et al., 2010; Blaauw, 2012).

Existing approaches leading to an error-free time axis are, however, designed for radiometrically dated proxy records (Breitenbach et al., 2012; Goswami et al., 2014). They are thus not directly applicable to proxy records with layer-counting–based chronologies, like ice cores, varved sediments, banded corals, or tree rings. The nature of the chronological uncertainty in

such records is fundamentally different from that in radiometrically dated ones. Lotter and Lemcke (1999), for instance, have discussed the statistical problems arising in layer-counted chronologies in the context of annual biochemical varves.

More recently, a model based on discrete random walks was proposed (Rhines and Huybers, 2011) to account for dating errors in the GISP2 ice core proxy record (Alley et al., 1997). Comboul et al. (2014) have also carried out a thorough analysis, based on a probabilistic age model, of dating uncertainties in banded choral archives.

Inspired by the Bayesian approach for radiocarbon-dated archives of Goswami et al. (2014), we introduce here a similar approach that is specifically designed to account for the uncertainties arising in layer-counted chronologies. Our approach also relies on a Bayesian framework to propagate all uncertainties to the proxy axis, and represents the proxy record as a sequence of probability densities on a prescribed, error-free time axis.

The key observation of our approach is that the probability distribution of a proxy value x, given a calendar age t, can be expressed in terms of the probability distributions of x and t, conditioned on the measured depth in the proxy archive z:

$$p(x|t) = \frac{\int p(x|z)p(t|z)dz}{\int p(t|z)dz} \tag{1}$$

This equation reveals that $p(x|t)$ is in fact the normalized average of $p(x|z)$ over all depths z, weighted by the respective contributions $p(t|z)$, which are considered as functions of the proxy depth z. The details appear in the Methods section and are illustrated in Fig. 1A.

Ideally, precise estimates of the measurement uncertainty distribution $p(x|z)$ and the dating uncertainty distribution $p(t|z)$ would be reported together with the proxy data themselves. In practice, however, this is rarely the case, and typically both distributions are assumed to be Gaussian, in which case it would suffice to report the mean values and standard deviations for the proxy and age measurements, respectively. We will show below how the specific functional form of the uncertainty distributions strongly impacts the time series representation of the corresponding proxy record.

The idea to visualize dating uncertainties itself is not new (e.g. Blaauw et al., 2007). In the following, though, we pursue this idea further and present an approach in which the overall uncertainties are quantifed in a mathematically precise sense: $p(x|t)$ yields the best estimate of x at time t, given the observed data and their uncertainties. Therefore, the $p(x|t)$ series so derived can be used for further, quantitative analyses.

In order to emphasize the need for a rigorous handling of dating uncertainties in layer-counted proxy records, we first test and illustrate our method by applying it to $\delta^{18}O$ isotope ratios obtained from the North Greenland ice core (NGRIP) project. For this record, a layer-counted chronology exists for the past 60 kyr (Svensson et al., 2008; Rasmussen et al., 2014). Our results will reveal that the overall uncertainties due to dating errors in this record have been substantially underestimated.

Thereafter, we further generalize our approach in order to represent the increments Δx of a proxy record x between distinct time steps; doing so is important, for instance, when empirically estimating stochastic differential equation models from such records (Ditlevsen et al., 2005; Kwasniok, 2013; Krumscheid et al., 2015; Mitsui and Crucifix, 2017; Boers et al., 2017). In layer-counted chronologies, uncertainties accumulate toward the more remote past because one typically starts counting at the top of the core, i.e. at the most recent layer. Furthermore, the identification of periodic layers, such as seasonal varves, will become increasingly more uncertain due to accumulation processes and the typically decreasing quality of the record with

increasing depth. This increasing uncertainty may lead to very large absolute uncertainties for the dating of the archive's remote past. When analyzing relative changes Δx, however, only the relative dating uncertainties matter (cf. Methods).

Third, we apply our method to Δ^{14}C measurements obtained from the sediments of Lake Suigetsu, Japan, which allow for varve counting for the time interval $10.2 - 40.0$ kyr BP (Marshall et al., 2012; Schlolaut et al., 2012; Staff et al., 2013).
5   The Δ^{14}C record from Lake Suigetsu is used to calibrate measured ^{14}C ages with respect to the varve-counted chronology, and provides the only available fully terrestrial ~~calibration~~ comparison curve for radiometric dating of proxy records prior to 12.5 kyr BP (Bronk Ramsey et al., 2012).

Using our time-to-proxy method, we derive the overall uncertainties for this extensive ~~calibration~~ comparison curve, taking into account both ^{14}C measurement errors and errors stemming from the varve-counted Suigetsu chronology. Finally, ~~it is~~
10   ~~shown how the overall uncertainties~~ based on the example of the Suigetsu ~~calibration curve~~ ~~can be used~~ comparison curve, we show how to obtain complete error estimates for arbitrary radiometrically dated proxy records~~that cover the interval from 40.0 kyr BP to the present~~. We suggest, furthermore, to include such error estimates in the standard IntCal13 calibration curve (Reimer et al., 2013) itself.

2   Data

2.1   NGRIP ice core data

15   We employ a proxy record of δ^{18}O ratios from the North Greenland Ice Core Project (NGRIP). The δ^{18}O ratios obtained from ice cores are commonly interpreted as proxies for surface air temperature variability (Johnsen et al., 1992; Dansgaard et al., 1993; Johnsen et al., 2001; Andersen et al., 2004). The layer-counted chronology of this record is the Greenland Ice Core Chronology 2005 (GICC05) (Svensson et al., 2008), which starts at 60 kyr before AD 2000, abbreviated as b2k herein. We use
20   the published version of this record, with measurement values of δ^{18}O ratios reported at a temporal resolution of 20 yr. Dating uncertainties in terms of maximum counting errors (MCE) are given for each of these time steps in the published dataset (Andersen et al., 2006; Rasmussen et al., 2006). As a result of the layer-counted dating, the MCE increase monotonically toward the past; see (Svensson et al., 2008; Rasmussen et al., 2014) and Fig. ~~1~~ 2 herein.

2.2   Suigetsu lake sediment data

25   Sediment cores obtained from Lake Suigetsu, Japan, allow for a floating, varve-counted chronology of the approximate interval $10.2 - 40.0$ kyr BP, where BP refers to the year 1950 (Marshall et al., 2012; Schlolaut et al., 2012; Staff et al., 2013). Counting errors are reported in terms of 1σ, from which we obtain the corresponding maximum counting error as $\text{MCE} = 2\sigma$ (Andersen et al., 2006). Using this chronology and combining it with speleothem data for its more recent past, a comprehensive ^{14}C record was obtained that provides a unique ~~calibration~~ comparison curve for atmospheric radiocarbon age measurements over
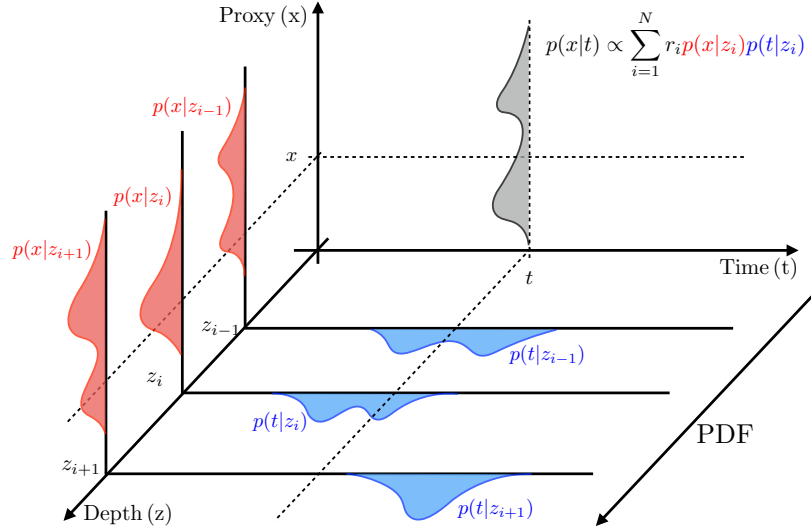30   the time interval $10.2 - 52.8$ kyr BP (Bronk Ramsey et al., 2012).

$$p(x|t) \propto \sum_{i=1}^{N} r_i p(x|z_i) p(t|z_i)$$

Figure 1. ~~Sketch of converting time axis uncertainties into proxy axis ones, and NGRIP and Suigetsu age models. A. Scheme~~ **Schematic of converting time axis uncertainties into proxy axis ones.** Schematic illustrating the key ideas of the proposed approach to represent proxy records dated by counting layers as sequences of probability densities. In this sketch, the measurements of the proxy x and the corresponding calendar age t are indicated at three different depths z_{i-1}, z_i, and z_{i+1}. The uncertainty distributions $p(x|z)$ of the proxy measurement processes are depicted in red, and those of the calendar ages $p(t|z)$ in blue. For a prescribed time t, the final probability distribution $p(x|t)$ of proxy values is given as the normalized average of $p(x|z)$, weighted by the corresponding contributions of all age uncertainty distributions $p(z|t)$~~; see the Methods section for further details. B. The age–depth relation of GICC05 for~~ Note that the ~~NGRIP ice core (solid black line) (Andersen et al., 2006; Rasmussen et al., 2006). The associated uncertainties are quantified as the maximum counting error MCE, given in years corresponding to annual layers; they are depicted as blue shading around the age–depth relation,~~ x- and additionally as t-axes play a different role than the ~~blue dashed line with scale shown on the right-hand side. In this chronology~~ z-axis, ~~uncertain layers are counted as 1/2 ± 1/2 yr (Andersen et al., 2006; Rasmussen et al., 2006), which implies that the MCE~~ since a sum is ~~half~~ taken over the ~~number of uncertain layers. C. The age–depth relation of the varve-counted segment of the Suigetsu lake deposit record. Note how the uncertainties increase monotonically for both records, because the counting process causes them~~ latter to ~~accumulate toward~~ obtain $p(x|t)$; see the ~~past~~ Methods section for further details.

We used the Δ^{14}C data, as well as the ^{14}C age measurements, together with the corresponding varve-counted time stamps from Bronk Ramsey et al. (2012). Here, Δ^{14}C refers to the inferred level of radiocarbon in the atmosphere, given in per mil, as a deviation from the reference value of 1950. We did not consider the interval $40.0 - 52.8$ kyr BP, also included in (Bronk
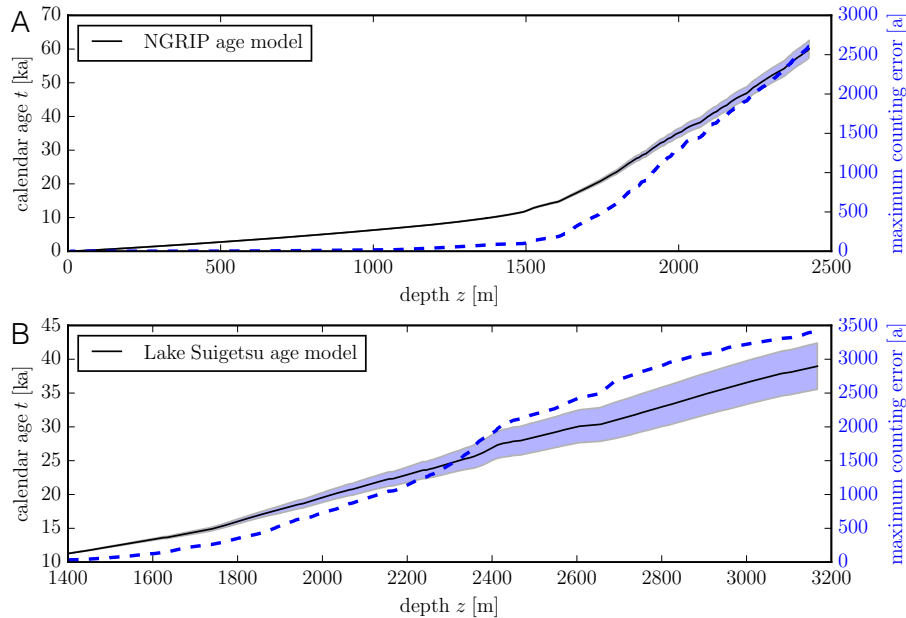
Figure 2. NGRIP and Suigetsu age models. A. The age–depth relation of GICC05 for the NGRIP ice core (solid black line) (Andersen et al., 2006; Rasmussen et al., 2006). The associated uncertainties are quantified as the maximum counting error MCE, given in years corresponding to annual layers; they are depicted as blue shading around the age–depth relation, and additionally as the blue dashed line with scale shown on the right-hand side. In this chronology, uncertain layers are counted as $1/2 \pm 1/2$ yr (Andersen et al., 2006; Rasmussen et al., 2006), which implies that the MCE is half the number of uncertain layers. Note the change in slope at the transition from glacial to interglacial conditions, which reflects the substantial increase in dating uncertainties at this point, due to changing accumulation rates and increasing pressure in the ice. B. The age–depth relation of the varve-counted segment of the Suigetsu lake deposit record. Note how the uncertainties increase monotonically for both records, because the counting process causes them to accumulate toward the past.

Ramsey et al., 2012), because the chronology for this segment is not directly varve-counted, but extrapolated from the varve chronology of the more recent time interval $10.2 - 40.0$ kyr BP.

3 Methods

In general, a proxy record consists of a set of depths z_i in the archive, as well as proxy values x_i and calendar ages t_i, which are both measured at depths z_i. In the case of interest for this study, the measurement of the calendar age is performed by counting layers that are a priori assumed to correspond to some known periodicity, such as annual layers in ice cores, sediment layers in lakes or growth rings in trees. It is already evident at this point that the uncertainties associated with the counting process are monotonically increasing in reverse time, as well as being strongly correlated.

The quantity we are interested in is the conditional probability distribution $p(x|t)$ of the proxy values x, given prescribed, error-free calendar ages t. The key observation is that this can be written in terms of $p(x|z)$, the probability distribution accounting for measurement uncertainties of the proxy value x at a given depth z, together with $p(t|z)$, the probability distribution of a calendar age t given a depth z, which accounts for the dating uncertainties (Goswami et al., 2014):

$$5 \quad p(x|t) = \int p(x|z)p(z|t)dz = \int p(x|z)p(t|z)\frac{p(z)}{p(t)}dz \; . \tag{2}$$

The first equality above is due to the chain rule for probabilities, and the second is due to Bayes' Theorem. Assuming so-called flat priors, i.e. a uniform distribution for the depths z, the fraction $p(z)/p(t)$ is merely a normalization constant; it can be determined by observing that, by definition, $\int p(x|t)dx = 1$ and $\int p(x|z)dx = 1$. We thus arrive at

$$p(x|t) = \frac{\int p(x|z)p(t|z)dz}{\int p(t|z)dz} \; . \tag{3}$$

10   In practical applications, only a finite number N of observations will be available. In such cases, the above~~expression~~, continuous integral will be approximated by a corresponding~~Riemann sum~~, discrete Riemann sum (Bronshtein et al., 2015) over the depths z_i at which measurements have been performed:

$$p(x|t) = \frac{\sum_{i=1}^{N} r_j p(x|z_i)p(t|z_i)}{\sum_{i=1}^{N} r_i p(t|z_i)} \; , \tag{4}$$

where $r_1 = z_2 - z_1$, $r_N = z_N - z_{N-1}$, and $r_i = (z_{i+1} - z_{i-1})/2$ for $1 < i < N$. Note that in the latter expression, r_i is the
15 average of the increments above and and below z_i: doing so provides a better approximation of the integral than taking only the previous or the following increment.

    For a prescribed calendar age t, the probability of a specific proxy value x thus involves the uncertainty distributions of x, as well as the uncertainty distributions of t, at all depths z_i (cf. Fig. 1A). Note that the probability distributions for the proxy values are derived as marginal distributions given prescribed calendar ages; the latter can thus be ~~freely chose~~chosen freely,
20 e.g., equidistantly with a desired temporal resolution.

    We note that, typically, the probability of missing a true layer is assumed to be equal to the probability of counting a false one: Thus, for the NGRIP chronology (GICC05), uncertain layers are each counted as $1/2 \pm 1/2$ yr (Andersen et al., 2006; Rasmussen et al., 20 Such an assumption will lead to a symmetric overall dating uncertainty distribution $p(t|z)$. If the probabilities for counting a false layer and missing a true one are estimated to be distinct from each other, the resulting $p(t|z)$ will be accordingly skewed.
25 In fact, our Bayesian approach for deriving $p(x|t)$ works with arbitrary functional forms for $p(t|z)$, symmetric or not.

    If a tephra layer is present in the record under study, and if the corresponding volcanic eruption can be precisely and independently dated, the MCE will be reset back to zero at that layer. This would be reflected in very narrow dating uncertainty distributions $p(t|z)$ around such a layer, and the spread in the derived $p(x|t)$ would decrease accordingly around the corresponding time.

30   In addition to $p(x|t)$, we are interested in $p(\Delta x|t, t')$, where Δx denotes the change of the proxy value x between time t and a later time t', i.e. $\Delta x = x_{t'} - x_t$. These could be adjacent time steps, in which case Δx would indicate the change of the

proxy value x per (arbitrary) unit of time. Precise estimates of the latter are crucial, for instance, in data-driven modeling of the temporal evolution of the proxy value x in terms of differential equations, either deterministic or stochastic (Ditlevsen et al., 2005; Kwasniok, 2013; Krumscheid et al., 2015; Mitsui and Crucifix, 2017; Boers et al., 2017).

As in Eq. (2) above, $p(\Delta x|t, t')$ can be expanded as

$$p(\Delta x|t,t') = \int \int p(\Delta x|z,z')p(t,t'|z,z')dzdz' \,. \tag{5}$$

For arbitrary random variables a, b, c, and d, Bayes' Theorem implies that

$$p(a|b,c,d) = \frac{p(a,b,c,d)}{p(b,c,d)} = \frac{p(c,d|a,b)p(a,b)}{p(b,c,d)}$$

and therefore

$$p(t'|t,z,z') = \frac{p(z,z'|t,t')p(t,t')}{p(t,z,z')} \,. \tag{6}$$

Inserting this into Eq. (5) and assuming a uniform prior $p(t,z,z')$, we obtain

$$p(\Delta x|t,t') = \frac{\int \int p(\Delta x|z,z')p(t'|z,z',t)dzdz'}{\int \int p(t'|z,z',t)dzdz'} \,; \tag{7}$$

here the same reasoning as for Eq. (3) applies concerning the normalization.

4 Results

4.1 Dating uncertainties in the NGRIP δ^{18}O record

The details of the NGRIP proxy record's layer-counted chronology and associated uncertainties have already been discussed at considerable length (Andersen et al., 2006; Rasmussen et al., 2006; Svensson et al., 2008). The further one goes into the past, the more the layer-counting errors accumulate, leading to $\text{MCE} = 2601$ yr at a layer-counted age of 60 kyr b2k (Fig. ~~1B~~2A). As explained in detail below, the precise dating uncertainty distribution $p(t|z)$ is typically unknown.

Given the number of uncertain layers $L = 2 \cdot \text{MCE}$, each counted as ~~1/2 ± 1/2~~ $1/2 \pm 1/2$ yr, uncorrelated counting errors would lead to a simple error estimate in terms of a normal distribution with $\sigma = \left(L(1/2 \text{ yr})^2\right)^{1/2} = \left(2 \cdot \text{MCE}\right)^{1/2}/2$. This naive estimate is, however, very unrealistic since the errors are much more likely to be correlated (Andersen et al., 2006). Several authors have suggested assuming Gaussian functional forms for $p(t|z)$, with standard deviations set to $\sigma = \text{MCE}/2$ (Andersen et al., 2006; Rasmussen et al., 2006; Svensson et al., 2008).

This assumption accounts, to some degree, for correlated errors and it gives quite conservative error estimates, but it also postulates a very specific, Gaussian form for the uncertainty distribution. In an even more conservative vein, one has to admit that nothing is known about the way the counting errors associated with uncertain layers depend on each other. This assumption would lead to a uniform distribution of dating uncertainties, namely:

$$p(t|z_i) \sim \mathcal{U}(t(z_i) - \text{MCE}, t(z_i) + \text{MCE}) \tag{8}$$

We apply our methodology to represent the NGRIP δ^{18}O record (Fig. ~~2~~3A) as a sequence of probability densities on an error-free time axis with equidistant 20-year steps. Since the proxy measurement uncertainties are not reported in the published version of this dataset, we assume for $p(x|z)$ a normal distribution, centered at the reported values and with very small $\sigma = 0.01$ per mille. The results we present in the following would not change if the proxy measurement errors were completely

5 neglected by setting $p(x|z_i) = \delta(x - x(z_i))$, the Dirac delta distribution located at the measured value $x(z_i)$.

We employ three different functional forms for the age uncertainty distribution $p(t|z)$ in order to illustrate how this choice impacts the final representation. First, the rather unrealistic choice of a normal distribution with $\sigma = \left(2 \cdot \mathrm{MCE}\right)^{1/2}/2$, which corresponds to independent dating errors, is used in Fig. ~~2~~3B. In this case, the variability of the δ^{18}O record — as seen in the blue line of Fig. ~~2~~3A, which does not account at all for dating uncertainties — is retained by $\hat{x}(t)$, the expectation value of

10 δ^{18}O with respect to $p(x|t)$ (red solid line in panels B–D of the figure). Due to the small σ imposed on $p(t|z)$ when assuming uncorrelated errors, the densities $p(x|t)$ are very localized, and hence the blue shading appears merely as single points in Fig. ~~2~~3B.

In contrast, as seen in Fig. ~~2~~3C, using a normal distribution with $\sigma = \mathrm{MCE}/2$ for the counting error distribution $p(t|z)$, as proposed in (Andersen et al., 2006; Rasmussen et al., 2006; Svensson et al., 2008), leads to substantially larger overall

15 uncertainties in $p(x|t)$. Moreover, this uncertainty grows significantly into the past, due to the accumulation of errors in the layer-counting process.

Finally, using a uniform distribution $p(t|z)$~~(,~~ cf. Eq. (8) and Fig. ~~2D)~~ 3D, yields an uncertainty range for $p(x|t)$ that is similar to the one in Fig. ~~2~~3C. The short-term variability of $\hat{x}(t)$ that is left in these two panels, however, differs substantially between the representations using a normal and a uniform distribution, respectively.

20 The assumption of uncorrelated dating errors in the layer-counted chronology GICC05 of the NGRIP records is hardly justifiable (Andersen et al., 2006; Rasmussen et al., 2006), and therefore statistically dependent dating errors should be assumed when choosing the age uncertainty distribution $p(t|z)$. The MCE of the layer-counted chronology of the NGRIP record reaches 2601 yr at 60 kyr b2k; it is hence smaller than 4%, and confirms therewith the high accuracy of the dating process by counting annual layers.

25 Still, although the MCE seems small, ~~it~~ the increasing spread of $p(t|z)$ leads to strong growth in the overall uncertainties in the final representation of the δ^{18}O record we obtain; this representation is given by $p(x|t)$. For the older parts of the record, ~~these uncertainties~~ the uncertainties in $p(x|t)$ grow more and more into the past and they thus affect more and more a statistically sound estimation of proxy values; see Figs. ~~2~~3C and ~~2~~3D. This difficulty implies that great care has to be exercised when comparing and, in particular, trying to align specific events among two or more distinct proxy records.

30 Comparing the expected values $\hat{x}(t)$ of δ^{18}O in Figs. ~~2C and 2~~3C and 3D — obtained using a normal and a uniform distribution for $p(t|z)$, respectively — shows similar long-term variability, but significant differences in the short-term variability of their temporal evolution: the use of a normal dating error distribution leads to a much smoother temporal evolution of $\hat{x}$ than a uniform distribution. This apparent smoothness is, however, an artifact of employing a normal distribution for $p(t|z)$.

~~The latter functional form~~ Using a normal distribution for $p(t|z)$ is not justified by the dating process itself, since the actual
35 dependence of the dating errors is unknown. We hence argue for the use of a uniform distribution for $p(t|z)$; first, because it

provides a more accurate representation of the dating uncertainties when their dependencies are unknown, and second, because it avoids at the same time the probably spurious smoothness caused by using a normal form of $p(t|z)$. In practice, the true distribution of dating uncertainties is impossible to derive for chronologies based on counting periodic layers. Most likely, it will be a mixture of two extremes: on the one hand a normal distribution, which corresponds to weak correlations between counting errors, and on the other hand a uniform distribution, which corresponds to maximum correlations between these errors. In principle, one could use a convex linear combination of the two proposed forms, $p(t|z) = \alpha\mathcal{N} + (1-\alpha)\mathcal{U}$, properly normalized, or some other probability mixture, to obtain a desired degree of smoothness of the evolution of $\hat{x}(t)$. Of course, any a priori knowledge of uncertainties that can be derived from the dating process should be included here.

The distribution $p(x|t)$ is typically multimodal for given t and $p(t|z)$. If so, the expectation value $\hat{x}$ itself can have rather low probability, ~~cf.~~ as shown in the inset of Fig. ~~2D. To~~ 3D. We emphasize, therefore, that, in order to avoid misguided interpretations, the proxy record should ~~thus~~ be visualized in terms of $p(x|t)$ (blue shading) and not in terms of $\hat{x}$ (red solid lines). The latter is only shown here to enable a direct comparison with the traditional representation of the record as a unique, scalar time series, cf. Fig. ~~2~~3A.

If one is not interested in a representation of the proxy record itself, but rather in the representation of its relative changes, only the relative counting errors matter. In this case, our method should be applied to the one-step increments of the NGRIP $\delta^{18}O$ record, i.e. $\Delta x_i = x_{i+1} - x_i$; see Methods section. Because the relative counting errors do not accumulate over time, a very precise estimation of these *relative* changes is in fact possible, cf. Fig. ~~3~~4.

The relative representation in Fig. ~~3~~4B of the increments $\Delta x_i = x_{i+1} - x_i$ of the $\delta^{18}O$ record, using the relative counting errors between subsequent measurement points, underscores the fact that the large overall uncertainties observed for the absolute representations in Figs. ~~2~~3(C,D) are due to the accumulation of dating errors. The finding that a precise estimation of the increments is possible indicates that the short-term variability of the $\delta^{18}O$ in terms of relative changes can be estimated with a high degree of certainty. This result explains the success of attempts to derive dynamical models from paleoclimatic records (Ditlevsen et al., 2005; Kwasniok, 2013; Krumscheid et al., 2015; Mitsui and Crucifix, 2017; Boers et al., 2017) despite neglecting the dating uncertainties.

Nevertheless, the short-term variability is strongly smoothed in the absolute representation of the record (Figs. ~~2~~3C and ~~2~~3D); such smoothing correctly reflects the fact that absolute dating of the more remote parts of the record is jeopardized by accumulating counting errors. An absolute representation should be used, for instance, when comparing and aligning several proxy records on an absolute time scale, while a relative representation of increments should be used when one focuses on the high-frequency variability and relative changes of the proxy record under consideration.

4.2 Dating uncertainties in the Suigetsu ~~$\Delta^{14}C$ calibration~~ ^{14}C comparison curve

We restrict ourselves to the time interval $10.2 - 40.0$ kyr BP, for which a varve-counted chronology of the Lake Suigetsu sediment record exists (Marshall et al., 2012; Schlolaut et al., 2012; Staff et al., 2013). As for the NGRIP record, errors associated with the varve-counting accumulate toward the past: the largest error, for the varve counted part, is reported to be

$\sigma = 1707$ yr at a varve-counted age of 38964 yr; see also Fig. ~~1C~~2B. Note that, in accordance with the discussion of the NGRIP record and Fig. ~~2~~3, we set $\mathrm{MCE} = 2\sigma$. For the Suigetsu record, we choose a uniform time axis with 50 yr increments.

The Δ^{14}C series of the Suigetsu sediments, shown here in Fig. ~~4~~5A, has been used as a comparison curve to fit radiometrically dated archives — such as speleothem data (Hoffmann et al., 2010; Southon et al., 2012) — onto the Suigetsu varve chronology. As done in the previous section for NGRIP, we represent in Fig. ~~4~~5B the Suigetsu Δ^{14}C record on an error-free time axis, accounting for the reported Δ^{14}C measurement errors, as well as for the reported counting errors of the varve chronology. Due to the accumulating counting errors, the overall uncertainties in the Δ^{14}C values derived in this way become considerably larger the further one proceeds into the past.

The radiocarbon age curve in Fig. ~~4~~5C, considered as a function of the varve-counted ages, is itself a paleoclimatic proxy. We can therefore use our method to derive a representation of the overall uncertainties in this radiocarbon age ~~calibration~~ curve, originating from both radiocarbon age measurement errors and varve counting errors. This representation yields the uncertainty distribution $p(t_{\mathrm{rc}}|t)$ of the radiocarbon age t_{rc}, given the true calendar age t, as plotted in Fig. ~~4~~5D. The largest uncertainty of the radiocarbon age, quantified as the interquartile range, is 3487 yr, observed at a calendar age $t = 34\,227$ yr.

Our representation of the Suigetsu Δ^{14}C values in Fig. ~~4~~5B reveals that, despite the seemingly small varve counting errors, of less than 5%, considerable uncertainties are involved in the radiocarbon age calibration for which this record is used as a ~~standard~~ comparison curve. These uncertainties propagate ~~to~~ into the final radiocarbon age model in Fig. ~~4~~5D, which is proposed to be used to calibrate ages in other, radiocarbon-dated proxy archives. In fact, the Suigetsu comparison curve provides an important contribution to the earlier parts of the IntCal13 radiocarbon calibration curve (Reimer et al., 2013). We suggest to include the uncertainties stemming from the varve-counting process of the Suigetsu sediment record, as quantified by our Bayesian approach, into the IntCal13 calibration curve in future work.

The uncertainty distribution $p(t_{\mathrm{rc}}|t)$ of the radiocarbon ages, as derived from our method, can be directly used to obtain complete uncertainty estimates when analyzing arbitrary radiometrically dated proxy archives: Measuring a proxy variable v in a given radiocarbon-dated archive, the probability distribution of v, given a true calendar age t, can be expanded in terms of the radiocarbon age t_{rc} as

$$p(v|t) = \int p(v|t_{\mathrm{rc}})p(t_{\mathrm{rc}}|t)dt_{\mathrm{rc}} \,. \tag{9}$$

Here, $p(v|t_{\mathrm{rc}})$ is the distribution of proxy values v, measured at radiocarbon ages t_{rc} in the other archive, and $p(t_{\mathrm{rc}}|t)$ is the distribution of total uncertainties in the Suigetsu radiocarbon age model, shown in Fig. ~~4D. In order to facilitate the incorporation of these uncertainties in the estimation of a proxy variable from any radiometrically dated archive, we provide the values of $p(t_{\mathrm{rc}}|t)$ for different temporal resolutions in the online supplementary material.~~5D.

In addition, one may be interested in the probability of a true calendar age t, given a measured radiocarbon age t_{rc}. For arbitrary radiocarbon-dated proxy archives, this can be obtained from the uncertainty distribution $p(t_{\mathrm{rc}}|t)$ of the Suigetsu radiocarbon ages, via Bayes' Theorem:

$$p(t|t_{\mathrm{rc}}) = \frac{p(t_{\mathrm{rc}}|t)p(t)}{p(t_{\mathrm{rc}})}$$

Here, $p(t)$ can be assumed to be an uninformative prior, while $p(t_{\mathrm{rc}})$ reflects the uncertainty distribution of the measurement of the radiocarbon age t_{rc}.

5 Conclusions

We have introduced statistically rigorous representations of layer-counted proxy records as sequences of probability distributions on error-free time axes. In such records, the calendar age is determined by counting layers, which are assumed to correspond to a known periodicity, such as annual layers in ice cores, varved sediment layers, banded corals, or growth rings in trees. Our approach, which is rooted in Bayesian statistics, takes into account the uncertainties in both proxy measurement and dating. Such an unambiguous representation of uncertainties is crucial, for instance, when comparing proxy records obtained from different archives. Our results indicate that the effects of dating uncertainties in paleoclimatic archives have been strongly underestimated, and emphasize the urgent need for an adequate statistical representation of proxy records with immanent dating uncertainties in order to prevent misleading interpretations: The common representation of such proxy records as time series on the mean or median values of the age distributions induces a strong bias, suggesting a much higher degree of accuracy than is actually warranted.

First, we illustrated our method by applying it to the δ^{18}O record obtained from the NGRIP ice core. A representation of this record on an absolute time scale (Fig. ~~2~~3) revealed that the statistically dependent, cumulative errors associated with the layer-counted dating, although seemingly small, lead to growing overall uncertainties the further one digs into the past.

This finding calls for great caution when aligning proxy records from distinct sources on a common, absolute time scale. On the other hand, a representation of the increments of the δ^{18}O record between subsequent time steps (Fig. ~~3~~4) is possible with a comparably high degree of certainty, because in this case only the relative counting errors from one measurement point to the next are relevant. These increments and their representation are of interest if one focuses on the short-term variability. In particular, modeling the associated time series in terms of differential equations, whether deterministic or stochastic, is relatively insensitive to such errors.

Second, we applied our method to the radiocarbon ~~calibration~~ comparison curve recently derived from annually layered sediment cores of Lake Suigetsu, Japan. To date, this is the only available terrestrial radiocarbon record that can be used to calibrate other radiometrically dated proxy archives for ages prior to 12.5 kyr BP.

The proposed method allowed us to propagate the uncertainties associated with the layer-counted Suigetsu chronology, as well as the ones associated with the radiocarbon measurements, to the final radiocarbon ~~calibration~~ comparison curve. The resulting sequence of probability distributions represents the overall uncertainties of the ~~calibration curve. This sequence~~ comparison curve. Since the Suigetsu comparison curve forms a crucial part of the standard IntCal13 radiocarbon calibration curve, this sequence of uncertainty distributions — as well as the corresponding ones for other radiocarbon proxy curves — can and should be used to obtain more accurate uncertainty estimations in arbitrary radiocarbon-dated proxy archives~~for which the Suigetsu radiocarbon calibration curve is employed~~. Furthermore, our approach allows one to compute the uncertainty distribution of the true calendar age, given a radiocarbon age estimate.

12

Author contributions. NB conceived the study, carried out the analysis, and prepared the manuscript. All authors discussed the results, drew conclusions, and contributed to editing the manuscript.

5   *Competing interests.* The authors declare that they have no competing financial interests.

References

Alley, R. B., Shurman, C. A., Meese, D. A., Gow, A. J., Taylor, K. C., Cuffey, K. M., Fitzpatrick, J. J., Grootes, P. M., Zielinski, G. A., Ram, M., Spinelli, G., and Elder, B.: Visual-stratigrapic dating of the GISP2 ice core: Basis, reproducability, and application, Journal of Geophysical Research, 102, 26 367–26 381, 1997.

Andersen, K. K., Azuma, N., Barnola, J.-M., Bigler, M., Biscaye, P., Caillon, N., Chappellaz, J., Clausen, H. B., Dahl-Jensen, D., Fischer, H., Flückiger, J., Fritzsche, D., Fujii, Y., Goto-Azuma, K., Grønvold, K., Gundestrup, N. S., Hansson, M., Huber, C., Hvidberg, C. S., Johnsen, S. J., Jonsell, U., Jouzel, J., Kipfstuhl, S., Landais, A., Leuenberger, M., Lorrain, R., Masson-Delmotte, V., Miller, H., Motoyama, H., Narita, H., Popp, T., Rasmussen, S. O., Raynaud, D., Rothlisberger, R., Ruth, U., Samyn, D., Schwander, J., Shoji, H., Siggard-Andersen, M.-L., Steffensen, J. P., Stocker, T., Sveinbjörnsdóttir, a. E., Svensson, A., Takata, M., Tison, J.-L., Thorsteinsson, T., Watanabe, O., Wilhelms, F., and White, J. W. C.: High-resolution record of Northern Hemisphere climate extending into the last interglacial period, Nature, 431, 147–151, doi:10.1038/nature02805, http://www.ncbi.nlm.nih.gov/pubmed/15356621, 2004.

Andersen, K. K., Svensson, A., Johnsen, S. J., Rasmussen, S. O., Bigler, M., Rüthlisberger, 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 Science Reviews, 25, 3246–3257, doi:10.1016/j.quascirev.2006.08.002, 2006.

Blaauw, M.: Methods and code for 'classical' age-modelling of radiocarbon sequences, Quaternary Geochronology, 5, 512–518, doi:10.1016/j.quageo.2010.01.002, http://linkinghub.elsevier.com/retrieve/pii/S1871101410000038, 2010.

Blaauw, M.: Out of tune: the dangers of aligning proxy archives, Quaternary Science Reviews, 36, 38–49, doi:10.1016/j.quascirev.2010.11.012, http://linkinghub.elsevier.com/retrieve/pii/S0277379110004087, 2012.

Blaauw, M., Christen, J. a., Mauquoy, D., van der Plicht, J., and Bennett, K. D.: Testing the timing of radiocarbon-dated events between proxy archives, The Holocene, 17, 283–288, doi:10.1177/0959683607075857, http://hol.sagepub.com/cgi/doi/10.1177/0959683607075857, 2007.

Blaauw, M., Wohlfarth, B., Andrés Christen, J., Ampel, L., Veres, D., Hughen, K. A., Preusser, F., and Svensson, A.: Were last glacial climate events simultaneous between greenland and france? a quantitative comparison using non-tuned chronologies, Journal of Quaternary Science, 25, 387–394, doi:10.1002/jqs.1330, 2010.

Boers, N., Chekroun, M. D., Liu, H., Kondrashov, D., Rousseau, D.-D., Svensson, A., Bigler, M., and Ghil, M.: Inverse stochastic-dynamic models for high-resolution Greenland ice-core records, Earth System Dynamics Discussions, 2017, 1–28, doi:10.5194/esd-2017-8, http://www.earth-syst-dynam-discuss.net/esd-2017-8/, 2017.

Breitenbach, S. F. M., Rehfeld, K., Goswami, B., Baldini, J. U. L., Ridley, H. E., Kennett, D. J., Prufer, K. M., Aquino, V. V., Asmerom, Y., Polyak, V. J., Cheng, H., Kurths, J., and Marwan, N.: COnstructing Proxy Records from Age models (COPRA), Climate of the Past, 8, 1765–1779, doi:10.5194/cp-8-1765-2012, http://www.clim-past.net/8/1765/2012/, 2012.

Bronk Ramsey, C., Staff, R. A., Bryant, C. L., Brock, F., Kitagawa, H., van der Plicht, J., Schlolaut, G., Marshall, M. H., Brauer, A., Lamb, H. F., Payne, R. L., Tarasov, P. E., Haraguchi, T., Gotanda, K., Yonenobu, H., Yokoyama, Y., Tada, R., Nakagawa, T., Tobergte, D. R., and Curtis, S.: A complete terrestrial radiocarbon record for 11.2 to 52.8 kyr B.P., Science (New York, N.Y.), 338, 370–374, doi:10.1017/CBO9781107415324.004, http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed{&}id=23087245{&}retmode=ref{&}cmd=prlinks{%}5Cnpapers2://publication/doi/10.1126/science.1226660, 2012.

Bronshtein, I. N., Semendyayev, K. A., Musiol, G., and Mühlig, H.: Handbook of Mathematics, Springer, 6th edn., doi:10.1007/978-3-662-46221-8, 2015.

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, Climate of the Past, 10, 825–841, doi:10.5194/cp-10-825-2014, 2014.

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

Ditlevsen, P. D., Kristensen, M. S., and Andersen, K. K.: The recurrence time of Dansgaard-Oeschger events and limits on the possible periodic component, Journal of Climate, 18, 2594–2603, doi:10.1175/JCLI3437.1, 2005.

Goswami, B., Heitzig, J., Rehfeld, K., Marwan, N., Anoop, A., Prasad, S., and Kurths, J.: Estimation of sedimentary proxy records together with associated uncertainty, Nonlinear Processes in Geophysics, 21, 1093–1111, doi:10.5194/npg-21-1093-2014, http://www.nonlin-processes-geophys.net/21/1093/2014/, 2014.

Hoffmann, D. L., Beck, J. W., Richards, D. A., Smart, P. L., Singarayer, J. S., Ketchmark, T., and Hawkesworth, C. J.: Towards radiocarbon calibration beyond 28 ka using speleothems from the Bahamas, Earth and Planetary Science Letters, 289, 1–10, doi:10.1016/j.epsl.2009.10.004, 2010.

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

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

Krumscheid, S., Pradas, M., Pavliotis, G. A., and Kalliadasis, S.: Data-driven coarse graining in action: Modelling and prediction of complex systems, Physical Review E, 92, 042 139, 2015.

Kwasniok, F.: Analysis and modelling of glacial climate transitions using simple dynamical systems, Philosophical Transactions of the Royal Society A, 371, 1–22, 2013.

Lotter, A. F. and Lemcke, G.: Methods for preparing and counting biochemical varves, Boreas, 28, 243–252, doi:10.1111/j.1502-3885.1999.tb00218.x, 1999.

Mann, M. E. and Jones, P. D.: Global surface temperatures over the past two millennia, Geophysical Research Letters, 30, 1820, doi:10.1029/2003GL017814, 2003.

Mann, M. E., Zhang, Z., Hughes, M. K., Bradley, R. S., Miller, S. K., Rutherford, S., and Ni, F.: Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia., Proceedings of the National Academy of Sciences of the United States of America, 105, 13 252–7, doi:10.1073/pnas.0805721105, http://www.pnas.org/content/105/36/13252.full, 2008.

Marshall, M., Schlolaut, G., Nakagawa, T., Lamb, H., Brauer, A., Staff, R., Ramsey, C. B., Tarasov, P., Gotanda, K., Haraguchi, T., Yokoyama, Y., Yonenobu, H., and Tada, R.: A novel approach to varve counting using μXRF and X-radiography in combination with thin-section microscopy, applied to the Late Glacial chronology from Lake Suigetsu, Japan, Quaternary Geochronology, 13, 70–80, doi:10.1016/j.quageo.2012.06.002, 2012.

Mitsui, T. and Crucifix, M.: Influence of external forcings on abrupt millennial-scale climate changes: a statistical modelling study, Climate Dynamics, 48, 2729–2749, doi:10.1007/s00382-016-3235-z, http://dx.doi.org/10.1007/s00382-016-3235-z, 2017.

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., R??thlisberger, R., Fischer, H., Goto-Azuma, K., Hansson, M. E., and Ruth,

U.: A new Greenland ice core chronology for the last glacial termination, Journal of Geophysical Research Atmospheres, 111, 1–16, doi: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.,
5   Vallelonga, P., Vinther, B. M., Walker, M. J. C., Wheatley, J. J., and Winstrup, M.: A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: Refining and extending the INTIMATE event stratigraphy, Quaternary Science Reviews, 106, 14–28, doi:10.1016/j.quascirev.2014.09.007, http://dx.doi.org/10.1016/j.quascirev.2014.09.007, 2014.

Reimer, P. J., Bard, E., Bayliss, A., Beck, J. W., Blackwell, P. G., Bronk Ramsey, C., Buck, C. E., Cheng, H., Edwards, R. L., Friedrich, M.,
10   Grootes, P. M., Guilderson, T. P., Haflidason, H., Hajdas, I., Hatté, C., Heaton, T. J., Hoffmann, D. L., Hogg, A. G., Hughen, K. A., Kaiser, K. F., Kromer, B., Manning, S. W., Niu, M., Reimer, R. W., Richards, D. A., Scott, E. M., Southon, J. R., Staff, R. A., Turney, C. S. M., and van der Plicht, J.: IntCal13 and Marine13 Radiocarbon Age Calibration Curves 0–50,000 Years cal BP, Radiocarbon, 55, 1869–1887, doi:10.2458/azu_js_rc.55.16947, http://researchcommons.waikato.ac.nz/handle/10289/8955{%}5Cnhttp://researchcommons.waikato.ac.nz/handle/10289/8955{%}5Cnhttps://journals.uair.arizona.edu/index.php/radiocarbon/article/view/16947, 2013.

15 Rhines, A. and Huybers, P.: Estimation of spectral power laws in time uncertain series of data with application to the Greenland Ice Sheet Project 2 δ18O record, Journal of Geophysical Research, 116, D01 103, doi:10.1029/2010jd014764, http://dx.doi.org/10.1029/2010JD014764, 2011.

Roques, L., Chekroun, M., Cristofol, M., Soubeyrand, S., and Ghil, M.: Parameter estimation for energy balance models with memory, Proceedings of the Royal Society A, A470:20140, http://www.environnement.ens.fr/IMG/RCCSG{_}rev(1).pdf, 2014.

20 Schlolaut, G., Marshall, M. H., Brauer, A., Nakagawa, T., Lamb, H. F., Staff, R. A., Bronk Ramsey, C., Bryant, C. L., Brock, F., Kossler, A., Tarasov, P. E., Yokoyama, Y., Tada, R., and Haraguchi, T.: An automated method for varve interpolation and its application to the Late Glacial chronology from Lake Suigetsu, Japan, Quaternary Geochronology, 13, 52–69, doi:10.1016/j.quageo.2012.07.005, 2012.

Scholz, D. and Hoffmann, D. L.: StalAge - An algorithm designed for construction of speleothem age models, Quaternary Geochronology, 6, 369–382, doi:10.1016/j.quageo.2011.02.002, 2011.

25 Scholz, D., Hoffmann, D. L., Hellstrom, J., and Bronk Ramsey, C.: A comparison of different methods for speleothem age modelling, Quaternary Geochronology, 14, 94–104, doi:10.1016/j.quageo.2012.03.015, 2012.

Southon, J., Noronha, A. L., Cheng, H., Edwards, R. L., and Wang, Y.: A high-resolution record of atmospheric 14C based on Hulu Cave speleothem H82, Quaternary Science Reviews, 33, 32–41, doi:10.1016/j.quascirev.2011.11.022, 2012.

Staff, R. A., Nakagawa, T., Schlolaut, G., Marshall, M. H., Brauer, A., Lamb, H. F., Bronk Ramsey, C., Bryant, C. L., Brock, F., Kitagawa, H.,
30   Van der Plicht, J., Payne, R. L., Smith, V. C., Mark, D. F., Macleod, A., Blockley, S. P. E., Schwenninger, J. L., Tarasov, P. E., Haraguchi, T., Gotanda, K., Yonenobu, H., and Yokoyama, Y.: The multiple chronological techniques applied to the Lake Suigetsu SG06 sediment core, central Japan, Boreas, 42, 259–266, doi:10.1111/j.1502-3885.2012.00278.x, 2013.

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
35   chronology, Climate of the Past, 4, 47–57, doi:10.5194/cp-4-47-2008, http://www.clim-past.net/4/47/2008/, 2008.

Telford, R., Heegaard, E., and Birks, H.: All age–depth models are wrong: but how badly?, Quaternary Science Reviews, 23, 1–5, doi:10.1016/j.quascirev.2003.11.003, http://linkinghub.elsevier.com/retrieve/pii/S0277379103003160, 2004.
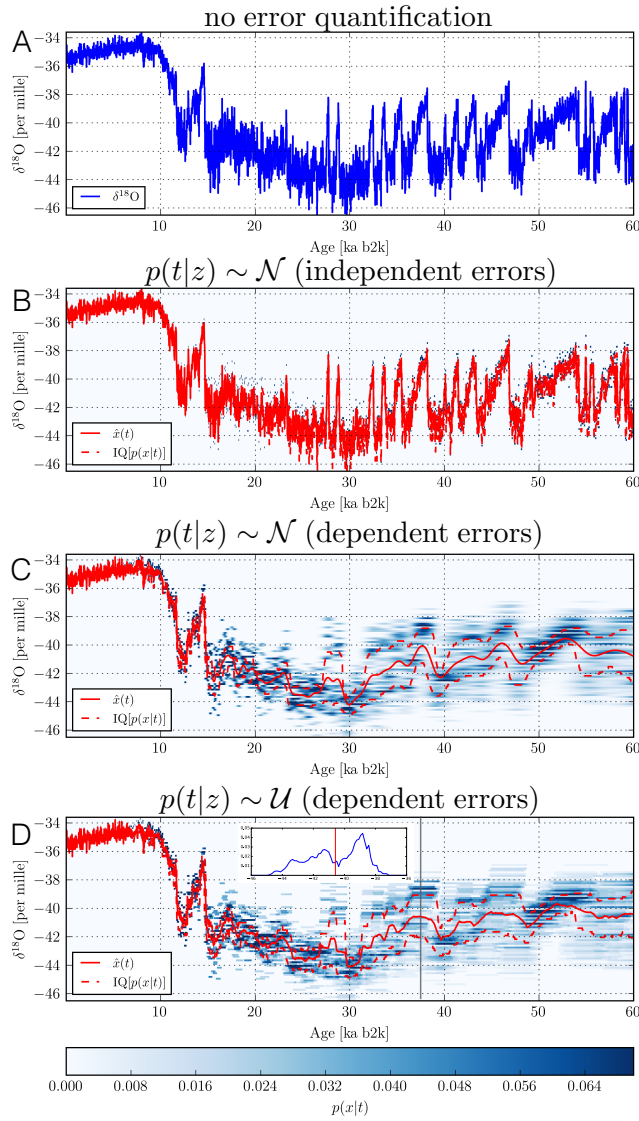
Figure 3. Representation of the NGRIP δ^{18}O record, taking into account absolute counting errors in the dating. A. Traditional representation of the record with no accounting for dating uncertainties at all: the time steps are obtained as the mean values of the corresponding dating uncertainty distributions. B. Representation of the same dataset as a sequence of probability densities, derived via Eq. (4); here we used a normal distribution with $\sigma = (2 \cdot \mathrm{MCE})^{1/2}/2$, which corresponds to uncorrelated errors, for the age uncertainty distribution $p(t|z)$. The blue shading showing the distributions $p(x|t)$ as a function of the prescribed time t is, due to the small σ for uncorrelated errors, only visible as single blue points. In this case, the time-dependent expectation value of $x = \delta^{18}$O with respect to $p(x|t)$, denoted by $\hat{x}(t)$ (red), closely resembles the traditional representation in panel A. C. Same as panel (B), but using a normal distribution with $\sigma = \mathrm{MCE}/2$ to represent the counting errors. D. Same as (B), but using a uniform distribution of width $2 \cdot \mathrm{MCE}$ to represent the counting errors. The inset shows $p(x|t)$ (blue) and $\hat{x}(t)$ (red) at $t = 37.5$ kyr b2k, also indicated by the grey vertical line in the main panel. Note that, in panels C and D, the spread of the distributions $p(x|t)$ — as quantified by the interquartile (IQ) range of $p(x|t)$ (dashed red line) — becomes wider the further one goes into the past. This increase in the spread of $p(x|t)$ is also reflected by a decrease of the high-frequency variability of its expectation values $\hat{x}(t)$.
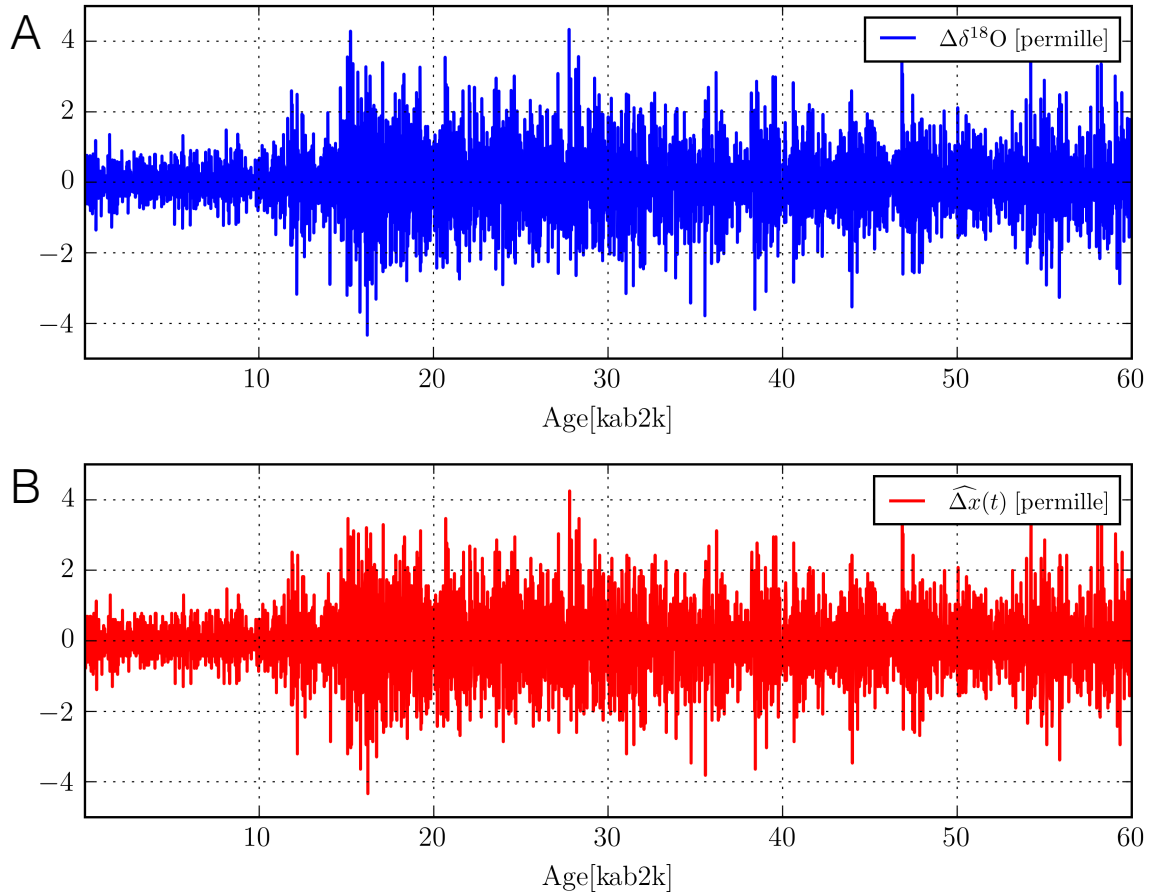
Figure 4. Representation of the one-step increments $\Delta\delta^{18}O$ of the NGRIP record, taking into account the relative counting errors. A. Traditional representation of the increments $\Delta\delta^{18}O$ between subsequent time steps; the time steps are obtained as in Fig. 23. B. Representation of the $\delta^{18}O$ increments between subsequent time steps as a sequence of probability densities, derived via Eq. (7). Note that the relative counting errors are sufficiently small to obtain very sharp, unimodal distributions $p(\Delta x|t, t+1)$. For this reason, only the expectation value $\widehat{\Delta x}(t)$ (red line) of the increments with respect to $p(\Delta x|t, t+1)$ is shown in this case.
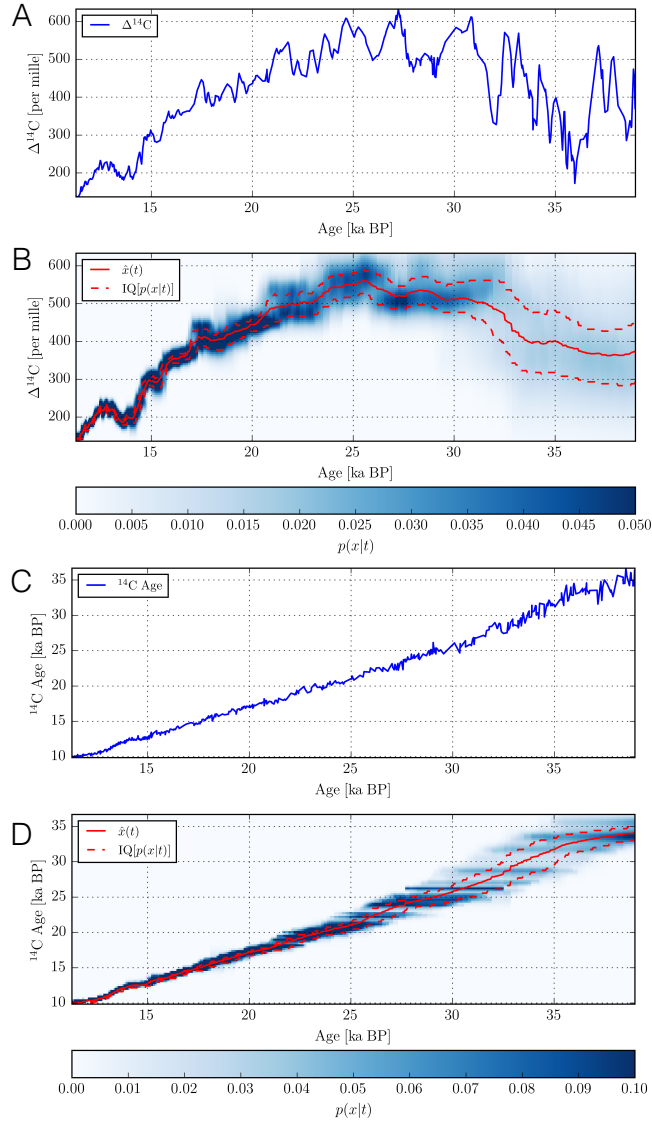
Figure 5. Representation of the Suigetsu lake sediment data, taking into account proxy uncertainties as well as absolute counting errors in the dating. A. Traditional representation of the Δ^{14}C time series, without accounting for uncertainties. The time span corresponds to the interval for which a varve-counted chronology is available. B. Representation of the Δ^{14}C from the Suigetsu lake sediment dataset as a sequence of probability densities, derived via Eq. (4). As for the NGRIP record (Fig. ~~2~~3), a uniform distribution of width 2MCE is used for the age uncertainty distribution $p(t|z)$. C. The radiocarbon (^{14}C) ~~calibration~~ comparison curve derived from the Suigetsu record, without accounting for uncertainties. D. Representation of the Suigetsu radiocarbon ~~calibration~~ comparison curve as sequence of probability densities, taking into account radiocarbon measurement errors as well as errors originating from varve counting.