GeoChronR – an R framework to model and analyze age-uncertain paleogeoscientific timeseries.

Nicholas McKay[1], Julien Emile-Geay[2], and Deborah Khider[3]

[1]School of Earth and Sustainability, Northern Arizona University, Flagstaff, AZ 86011
[2]University of Southern California, Los Angeles, CA
[3]Information Sciences Institute, University of Southern California, Marina del Rey, CA

Correspondence: Nicholas McKay (Nicholas.McKay@nau.edu)

Abstract. Chronological uncertainty is a hallmark of the paleosciences. While many tools have been made available to researchers to quantify age uncertainties suitable for various settings and assumptions, disparate tools and output formats often discourage integrative approaches. In addition, associated tasks like propagating age model uncertainties to subsequent analyses, and visualizing the results, have received comparatively little attention in the literature and available software. Here we describe GeoChronR, an open-source R package to facilitate these tasks. GeoChronR is built around emerging data standards for the paleosciences (Linked Paleo Data, or LiPD), and offers access to four popular age modeling techniques (Bacon, BChron, Oxcal, BAM). The output of these models is readily stored in LiPD, enabling age uncertain correlation, correlation, principal component, and spectral analyses. Five full-fledged use cases using illustrate how to use GeoChronR to facilitate these tasks, and to visualize the results in intuitive ways.

1 Introduction

1.1 Background

Quantifying chronological uncertainties, and how they influence the understanding of past changes in Earth systems, is a unique and fundamental challenge of the paleogeosciences. Without robust error determination, it is impossible to properly assess the extent to which past changes occurred simultaneously across regions, accurately estimate rates of change or the duration of abrupt events, or attribute causality – all of which limit our capacity to apply paleoscientific understanding to modern and future processes. The need for better solutions to both characterize uncertainty, and to explicitly evaluate how age uncertainty impacts the the interpretation of records of past climate, ecology or landscapes, has been long recognized (Noren et al., 2013; National Academies of Sciences, Engineering, and Medicine, 2020). In response to this need, the paleogeoscience community has made substantial advances toward improving geochronological accuracy by:

1. Improving analytical techniques that allow for more precise age determination on smaller and context-specific samples (Brown et al., 1989; Eglinton et al., 1996; Fifield, 2000; Eggins et al., 2005; Santos et al., 2010)

2. Refining our understanding of how past changes in the Earth system impact the age accuracy, for example: improvements to the radiocarbon calibration curve (Reimer et al., 2011, 2013, 2020) and advances in our understanding of spatial variability in cosmogenic production rates used in exposure dating (Balco et al., 2009; Masarik and Beer, 2009).

3. Dramatic improvement in the level of sophistication and realism in age-depth models used to estimate the ages of sequences between dated samples (e.g. Ramsey, 2009; Parnell et al., 2008; Blaauw, 2010; Blaauw and Christen, 2011b).

Over the past 20 years, these advances have been widely adopted in the paleogeosciences, albeit partially so. Indeed, despite the progress made in quantifying uncertainty in both ages determinations and age models, few studies have formally evaluated how chronological uncertainty may have affected the inferences made from them. For instance, whereas the aforemention algorithms have been broadly used, the overwhelming majority of these studies calculate the single best-estimate model (often a median or mean), use this model to put measured paleoclimatic or paleoenvironmental data on a timescale, and then proceed to analyze the record with little to no reference to the uncertainties generated by the age modeling exercise. In addition, few studies have evaluated sensitivity to the choice of age modeling technique or choice of parameters, so that the typical discussion of chronological uncertainties remain qualitative.

This paradigm is beginning to change. In recent years, a handful of studies have taken advantage of approaches that generate ensembles of age models to evaluate how the results of their analyses and conclusions vary given differences between ensemble members (e.g., Haam and Huybers, 2010; Tierney et al., 2013; Deininger et al., 2017; McKay et al., 2018; Bhattacharya and Coats, 2020). By using each ensemble age model to create a time-uncertain ensemble records, and then carrying that ensemble through the analysis, the precise impact of age uncertainty can be formally evaluated. This approach, of course, does not address all aspects of uncertainty, but it can provide key insight into which results are robust to chronological uncertainty, and which are not.

Despite its potential to substantially improve uncertainty quantification for the paleogeosciences, this framework is not widely utilized. The majority of studies utilizing this approach have been regional (e.g., Tierney et al., 2013; Deininger et al., 2017; McKay et al., 2018; Bhattacharya and Coats, 2020) or global-scale (e.g., Shakun et al., 2012; Marcott et al., 2013; Kaufman et al., 2020a) syntheses. Some primary publications of new records incorporate time-uncertain analysis into their studies (e.g., Boldt et al., 2015; Falster et al., 2018), but this remains rare. We suggest that there are several reasons for the lack of adoption of these techniques:

1. For synthesis studies, the necessary geochronological data are not publicly available for the vast majority of records. Even when they are available, the data are archived in diverse and unstructured data formats. Together, this makes what should be a simple process of aggregating and preparing data for analysis prohibitively time-consuming;

2. For studies of new and individual records, few tools for ensemble analysis are available, and those that are require a degree of comfort with coding languages and scientific programming that is rare among paleogeoscientists;

3. There is a disconnect between age-model development and time-uncertain analysis. Published approaches have utilized either simplified age-modeling approaches (e.g, Haam and Huybers, 2010; Routson et al., 2019), or specialized approaches not used elsewhere in the community [Marcott et al. (2013);Tierney et al. (2013); update with newer].

Extracting the relevant data from commonly-used age-modelling algorithms, creating time-uncertain ensembles, then reformatting those data for analysis in available tools typically requires the development of extensive custom codes. GeoChronR presents an integrative approach to facilitate this work.

1.2 Design principles

GeoChronR was built to lower the barriers to broader adoption of these emerging methods, particularly for Quaternary records, for which a variety of chronostratigraphic methods are available: radiometric dating ($^{14}\mathrm{C}$, $^{210}\mathrm{Pb}$, U/Th), exposure dating, layer-counting, flow models (for ice cores), etc. The primary uncertainty quantification device is age ensembles, regardless of how they were produced. As such, GeoChronR's philosphy and methods are expected to be more broadly applicable than Quaternary problems.

GeoChronR provides an easily-accessibly, open-source and extensible software package of industry-standard and cutting-edge tools that provides users with a single environment to create, analyze, and visualize time-uncertain data. GeoChronR is designed around emerging standards in the paleogeosciences that connects users to growing libraries of standardized datasets formatted in the Linked PaleoData format (McKay and Emile-Geay, 2016), including thousands of datasets archived at the World Data Service for Paleoclimatology (WDS-Paleo) and lipdverse.org, those at the LinkedEarth wiki, and Neotoma (Williams et al., 2018) via the neotoma2lipd package (McKay, 2020). GeoChronR reuses existing community packages, for which it builds a standardized interface, with LiPD as input/output format. Central to the development of the code and documentation was two workshops carried in out in 2016 and 2017 at Northern Arizona University (33 total participants). The workshops principally gathered early career researchers with $> 50\%$ participation of underrepresented minorities. Exit surveys were conducted to gather feedback, and to suggest improvements and extensions which were integrated into subsequent versions.

1.3 Outline of manuscript

This manuscript describes the design, analytical underpinnings and most common use cases of GeoChronR. Section 2 describes the integration of age modelling algorithms with GeoChronR. Section 3 details the methods implemented for age uncertain analysis. Section 4 goes through the principles and implementation of age-uncertain data visualization in GeoChronR, and section 5 provides five real-world examples of how GeoChronR can be used for paleoscientific workflows.

2 Age Uncertainty Quantification in GeoChronR

GeoChronR does not introduce any new approaches to age uncertainty quantification; rather, it integrates existing, widely-used packages while streamlining the acquisition of age ensemble members. Fundamentally, there are two types of age models used in the paleogeosciences: tie-point and layer-counted. Most of the effort in age uncertainty quantification in the community has been focused on tie-point modelling, where the goal is to estimate ages (and their uncertainties) along a depth profile given chronological estimates (and their uncertainties) at multiple depths downcore. Over the past 20 years, these algorithms have progressed from linear or polynomial regressions with simple characterizations of uncertainty (Heegaard et al., 2005; Blaauw, 2010) to more rigorous techniques, particularly Bayesian approaches: as of writing, the three most widely used algorithms are Bacon (Blaauw and Christen, 2011a), BChron (Parnell et al., 2008), and OxCal (Ramsey, 2008), which are all Bayesian age-deposition models that estimate posterior distributions on age-depth relationships using different assumptions and methodologies. Trachsel and Telford (2017) reviewed the performance of these three algorithms, as well as a non-Bayesian approach (Blaauw, 2010), and found that the three Bayesian approaches generally outperform previous algorithms, especially when appropriate parameters are chosen (although choosing appropriate parameters can be challenging). Bacon, BChron and Oxcal all leverage Monte Carlo Markov Chain (MCMC) techniques to sample the posterior distributions, thereby quantifiying age uncertainties as a function of depth in the section. GeoChronR interfaces with each of these algorithms through their R packages (Blaauw et al., 2020; Martin et al., 2018) (cite other R Packages), streamlining input and output. %data input and the extraction of the age ensembles from the MCMC results for further analysis.

In addition to working with ensembles from tie-point age models, GeoChronR connects users to probabilistic models of layer-counted chronologies. BAM (Comboul et al., 2014) was designed to probabilistically simulate counting uncertainty in banded archives, such as corals, ice cores, or varved sediments, but can be used to crudely simulate age uncertainty for any record, and is useful when the data or metadata required to calculate an age-depth model are unavailable. Here we briefly describe the theoretical basis and applications of each of the four approaches integrated in GeoChronR.

2.1 Bacon

The Bayesian ACcumulatiON (Bacon) algorithm (Blaauw and Christen, 2011a) is one of the most broadly used age-modelling techniques, and was designed to take advantage of prior knowledge about the distribution and autocorrelation structure of sedimentation rates in a sequence to better quantify uncertainty between dated levels. Bacon divides a sediment sequence into a parameterized number of equally-thick segments; most models use dozens to hundreds of these segments. Bacon then models sediment deposition, with uniform accumulation within each segment, as an autoregressive gamma process, where both the amount of autocorrelation and the shape of the gamma distribution are given prior estimates. The algorithm employs an adaptive Markov Chain Monte Carlo algorithm that allows for Bayesian learning to update these variables given the age-depth constraints, and converge on a distribution of age estimates for each segment in the model. Bacon has two key parameters: the shape of the accumulation prior, and the segment length, which can interact in complicaed ways (Trachsel and Telford, 2017). In our experience, the segment length parameter has the greatest impact on the ultimate shape and amount of uncertainty simulated by Bacon, as larger segments result in increased flexibility of the age-depth curve, and increased uncertainty between dated levels. Bacon is written in C++ and R, with an R interface. More recently, the authors released an R package “rbacon” (Blaauw et al., 2020), which GeoChronR leverages to provide access to the algorithm. Bacon will optionally return a subset of the MCMC accumulation rate ensemble members with high *a posteriori* probabilities, which GeoChronR uses to form age ensemble members for subsequent analysis.

2.2 BChron

BChron (Haslett and Parnell, 2008; Parnell et al., 2008) uses a similar approach, using a continuous Markov monotone stochastic process coupled to a piecewise linear deposition model. This simplicity allows semi-analytical solutions that make BChron computationally efficient. BChron was originally intended to model radiocarbon-based age-depth models in lake sedimentary cores of primarily Holocene age, but its design allows broader applications. In particular, modeling accumulation as additive independent gamma increments is appealing for the representation of hiatuses, particulary for speleothem records, where accumulation rate can vary quite abruptly between quiescient intervals of near-constant accumulation (Parnell et al., 2011; Dee et al., 2015; Hu et al., 2017). The downside of this assumption is that BChron is known to exaggerate age uncertainties in cases where sedimentation varies smoothly (Trachsel and Telford, 2017).

Bchron has several key parameters which allow a user to encode their specific knowledge about their data. In particular, the `outlierProbs` parameter is useful in giving less weigth to chronological tie points that may be considered outliers either because they create a reversal in the stratigraphic sequence or they are flagged during analysis (e.g. contamination). The `extractDate` parameter allows to set the top age of the sample. This is extremely useful for radiocarbon-based chronologies where the top age may not be accurately measured for modern samples. The `thetaMhSd`, `psiMhSd`, and `muMhSd` paraemeters constrol the Metropolis-Hastings standard deviation for the age parameters and Compund Poisson-Gamma scale and mean respectively, which influence the width of the ensemble between age control tie points. These parameters use the same default values as the official Bchron package, and we recommend that users only change them if they have good prior reason to do so.

2.3 Oxcal

The OxCal software package has a long history and extensive tools for the statistical treatment of radiocarbon and other geochronological data (Bronk Ramsey, 1995). In Ramsey (2008), age-depth modelling was introduced with three options for modelling depositional processes that are typically useful for sedimentary sequences: uniform, varve, and Poisson deposition models, labeled U-sequence, V-sequence and P-sequence, respectively. The Poisson-based model is the most broadly applicable for sedimentary, or other accumulation-based archives (e.g. speleothems), and although any sequence type can be used in GeoChronR, most users will use a P-sequence, which is the default. Analagously to segment length parameter in Bacon, the k parameter (called “eventsPerUnitLength” in GeoChronR), controls how many events are simulated per unit of depth, and has a strong impact on the flexibility of the model, as well as the amplitude of the resulting uncertainty. As the number of events increases, the flexibility of the model, and the uncertainties decrease. Trachsel and Telford (2017) found that this parameter has large impact on the accuracy of the model, and moreso than the choices made in Bacon or Bchron. Fortunately, Ramsey et al. (2010) made it possible for k to be treated as a variable, and the model will estimate the most likely values of k given a prior estimate and the data, however this calculation can greatly increase the convergence time of the model. Oxcal is written in C++, with an interface in R (Martin et al., 2018). Oxcal doesn’t calculate posterior ensembles for a depth sequence, but can optionally output MCMC posteriors at specified levels in the sequence. GeoChronR uses this feature to extract ensemble members for subsequent analysis.

2.4 Banded Age Model (BAM)

Comboul et al. (2014) is a probabilistic model of age errors in layer-counted chronologies. The model allows a flexible parametric representation of such errors (either as Poisson or Bernoulli processes), and separately considers the possibility of double-counting or missing a band. The model is parameterized in terms of the error rates associated with each event, which are intuitive parameters to paleogeoscientists, and may be estimated via replication (DeLong et al., 2013). In cases where such rates can be estimated from the data alone, an optimization principle may be used to identify a more likely age model when a high-frequency common signal can be used as a clock (Comboul et al., 2014). As of now, BAM does not consider uncertainties about such parameters, representing a weakness of the method. Bayesian generalizations have been proposed (Boers et al., 2017), which could one day be incorporated into GeoChronR if the code is made public. BAM was coded in MATLAB, Python and R, and it is this latter version that GeoChronR uses.

2.5 Storage

GeoChronR archives the outcome of all of these models using in the LiPD format (McKay and Emile-Geay, 2016). One of the primary motivations for LiPD was to facilitate age-uncertain analysis, and GeoChronR is designed to leverage these capabilities. LiPD can store multiple chronologies (called "chronData" in LiPD), each of which can contain multiple measurement tables (which house the measured chronological constraints) and any number of chronological models (which comprise both the results produced of the analysis, as well as metadata about the method used to produce those results) (figure 1). In LiPD, chronological models include up to three types of tables:

1. Ensemble tables, which store the output of an algorithm that produces age model ensembles, and a reference column (typically depth),

2. Summary tables, which describe summary statistics produced by the algorithm (e.g., median and 2σ uncertainty ranges), and

3. Distribution tables, which store age-probability distributions for calibrated ages, typically only used for calibrated radiocarbon ages.

The capability of GeoChronR to structure the output of the popular age model algorithms described in this section into LiPD is a key value proposition of GeoChronR. Once structured as a LiPD object in R, these data and models can be written out to a LiPD file and readily analyzed, shared and publicly archived.

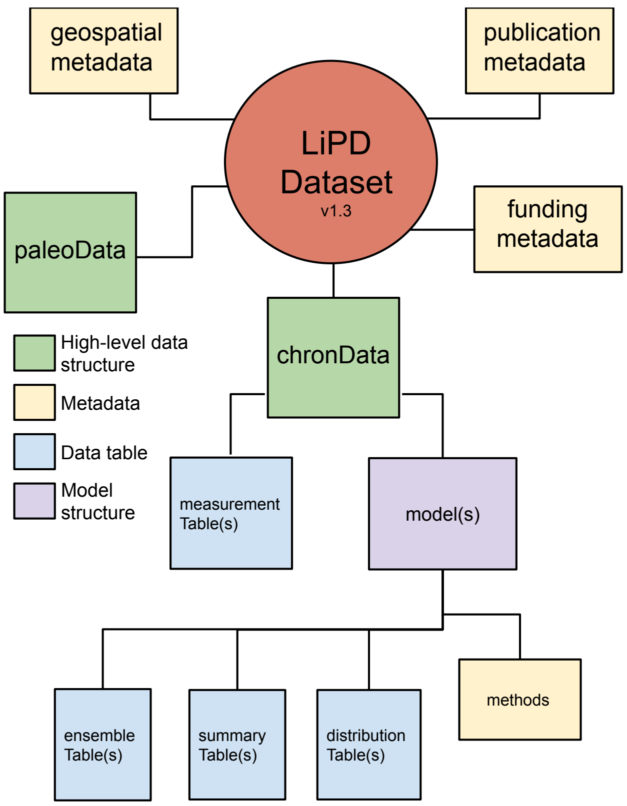

Figure 1. Schematic representation of a Linked PaleoData (LiPD) dataset, with a focus on chronological data. A LiPD dataset can contain one or more instances of all of the data objects and structures. The paleoData structure mirrors that of the chronData, but is not shown for clarity

3 Age-uncertain data analysis in GeoChronR

Some theoretical work has attempted to quantify how chronological uncertainty may affect various paleoscientific inferences (e.g. Huybers and Wunsch, 2004); however, such an approach is necessarily piecemeal, and therefore impractical given the variety of age-uncertainty structures in real-world paleogeoscientific data. Consequently, GeoChronR follows a pragmatic and broadly-used approach that leverages age ensembles, and optionally ensembles in climate proxy or paleoenvironmental data, to propagate uncertainties through all steps of an analysis. Effectively, this is done by randomly sampling from the ensemble(s) and then repeating the analysis on each sample (typically hundreds to thousands) to build an output ensemble that quantifies the impact of those uncertainties on a particular inference. These output ensembles often do not lend themselves to binary significance statistics (e.g., a p-value below 5%), but are readily used to provide quantiles that estimate probability density ranges, and can provide strong evidence for which results are robust to age and proxy uncertainty (and which are not).

Version 1.0.0 of GeoChronR has implemented ensemble analytical techniques for four of the most common analyses in the paleogeosciences: correlation, regression, spectral and principal component analyses.

3.1 Correlation

Pearson's product-moment correlation is the most common measure of a relationship between variables. Its computation is fast, lending itself to ensemble analysis, with a handful of pretreatment and significance considerations that are relevant for ensembles of paleogeoscientific data. First, correlation analysis for timeseries is built on the assumption the datasets can be aligned on a common timeline. Age-uncertain data violate this assumption. We overcome this by treating each ensemble member from one or more age uncertain timeseries as valid for that iteration, then "bin" each of the timeseries into coeval intervals. The "binning" procedure in GeoChronR sets up an interval, which is typical evenly spaced, over which the data are averaged. Generally, this intentionally degrades the median resolution of the timeseries, for example, a timeseries with 37-year median spacing could be reasonably "binned" into 100- or 200-year bins. The binning procedure is repeated for each ensemble member, meaning that between different ensembles, different observations will be placed in different bins.

Following binning, Pearson correlation is calculated and recorded for each ensemble member. In addition to the correlation coefficients, GeoChronR calculates their significance for each correlation. Paleogeoscientific timeseries are often highly autocorrelated, which can lead to spurious assessments of significance (Hu et al., 2017). Therefore, in addition to the standard Student's T-test p-value calculation, we also calculate a p-value that is adjusted for autocorrelation to reflect the reduction in degrees of freedom due to autocorrelation following Bretherton et al. (1999). However, repeating this test over multiple ensembles raises the issue of test multiplicity (Ventura et al., 2004), or "look elsewhere effect". To overcome this problem, we control for this false discovery rate using the simple approach of Benjamini and Hochberg (1995), coded in R by Ventura et al. (2004). {Julien: thoughts?} NM: I think here we need to discuss when FDR is needed, and when it's not. In my view, if you're surveying the full breadth on an ensemble, you're not searching for p-value, just trying to quantify the range. However, if the users want to subset some best-case or worst-case ensembles, and then proceed with analysis, FDR seems appropriate.

3.2 Regression

Linear regression is a commonly used tool to model the relationships between paleogeoscientific data and instrumental or other datasets. One application is calibration-in-time (Grosjean et al., 2009), whereby a proxy timeseries is calibrated to an instrumental series with a linear regression model over their period of overlap. This approach is particularly vulnerable to age uncertainties, as both the development of the relationship, and the reconstruction, are affected. GeoChronR propagates age (and optionally proxy) uncertainties through both the fitting of the ordinary least squares regression model, and the reconstruction "forecast" using the ensemble model results and age uncertainty. Like the correlation algorithm, ensemble regression uses an ensemble binning procedure that's analgous to correlation. GeoChronR then exports uncertainty structure of the modeled parameters (e.g. slope and intercept), as well as the ensemble of reconstructed calibrated data through time.

3.3 Principal Component Analysis

GeoChronR implements the age-uncertain principal component analysis (PCA) procedure introduced by Anchukaitis and Tierney (2013), with some minor modifications and additions. Like correlation and regression, PCA (or empirical orthogonal function {EOF} analysis) requires temporally aligned observations, and GeoChronR uses a binning procedure to achieve this across multiple ensembles. This differs from the implementation of Anchukaitis and Tierney (2013), who interpolated the data to a common timestep. In addition, traditional singular value decomposition approaches to PCA require a complete set of observations without any missing values. For paleoclimate data, especially when considering age uncertainty, this requirement is often prohibitive. To overcome this, GeoChronR implements multiple options for PCA analysis using the `pcaMethods` package. The default and most rigorously tested option is a probabilistic PCA (PPCA) approach that uses expectation maximization algorithms to infill missing values (Roweis, 1998). This algorithm assumes that the data and their uncertainties are normally distributed, which is often (but not always) a reasonable assumption for paleogeoscientific data. As in correlation and regression, GeoChronR propagates uncertainties through the analysis by repeating the analysis across randomly sampled age and/or proxy ensemble members to build output ensembles of the loadings (eigenvectors), variance explained (eigenvalues) and principal component timeseries. Because the sign of the loadings in PCA analyses is arbitrary and vulnerable to small changes in the input data, GeoChronR reorients the sign of the loadings for all PCs so that the mean of the loadings is positive. For well defined modes this effectively orients ensemble PCs, but loading orientation may be uncertain for lower order, or more uncertain, modes.

As in Anchukaitis and Tierney (2013), we use a modified version of Preisendorfer's "Rule N" (Preisendorfer and Mobley, 1988) to estimate which modes include more variability than those that can arise from random time series with comparable characteristics to the data. GeoChronR uses a rigorous "red" noise null hypothesis, modified from Neumaier and Schneider (2001), where following the selection of the age ensemble in each iterations, a synthetic autoregressive timeseries is simulated based on parameters fit from each dataset. This means that the characteristics of the null timeseries, including the temporal spacing, autocorrelation and, optionally, the first order trend, match those of each dataset, and vary between locations and ensemble iterations. For each iteration, the ensemble PCA procedure is replicated with the synthetic null dataset, using the same age ensemble member randomly selected for the real data. This effectively propagates the impact of age uncertainty into null hypothesis testing. Following the analysis, the distribution of eigenvalues calculated by the ensemble PCA is typically compared with the 95th percentile of the synthetic eigenvalue results in a scree plot. Only principle components whose eigenvalues exceed this threshold should be considered robust.

3.4 Spectral Analysis

Many research questions in the paleosciences revolve around spectral analysis: describing phase leads and lags among different climate system components over the Pleistocene (SPECMAP, Imbrie et al., 1984), the hunt for astronomical cycles over the Holocene (**?**Bond et al., 2001) or in deep time (Meyers and Sageman, 2007; Meyers, 2012, 2015), or characterizing the continuum of climate variability (Huybers and Curry, 2006; Zhu et al., 2019).

Yet, spectral analysis in the paleosciences faces unique challenges: chronological uncertainties, of course, as well as uneven sampling, which both violate the assumptions of classical spectral methods (Ghil et al., 2002).

To facilitate the quantification of chronological uncertainties in such assessments, GeoChronR implements four spectral approaches:

1. the Lomb-Scargle periodogram (VanderPlas, 2018), which uses an inverse approach to harmonic analysis in unevenly-spaced timeseries.

2. REDFIT, a version of the Lomb-Scargle periodogram tailored to paleoclimatic data (Schulz and Mudelsee, 2002; Mudelsee, 2002; Mudelsee et al., 2009). The GeoChronR uses the implementation of REDFIT from the dplR package (Bunn, 2008).

3. the wavelet-based method of Mathias et al. (2004), called `nuspectral`. This method is quite similar to the Weighted Wavelet Z-transform algorithm of Foster (1996), though it is prohibitively slow in this implementation, and the fast version using a compact-support approximations of the mother wavelet did not perform well in our tests.

4. The multi-taper method (MTM) of Thomson (1982), a mainstay of spectral analysis (Ghil et al., 2002) designed for evenly spaced timseries. GeoChronR uses the implementation of MTM in Meyers (2014), which couples MTM to efficient linear interpolation, together with various utilities to define autoregressive and power-law benchmarks for spectral peaks.

4 Visualization with GeoChronR

One of the challenges with age-uncertain analysis is that it adds at least one additional dimension to the results, which can be difficult to visualize. GeoChronR aims to facilitate simple creation of intuitive, publication-quality figures that provide multiple options for visualizing the impacts of age-uncertainty, while maintaining flexibility for users to customize their results as needed. To meet the multiple constraints of simplicity, quality and customization, GeoChronR relies heavily on the "ggplot2" package (Wickham, 2016). High-level plotting functions in GeoChronR (e.g., `plotTimeseriesEnsRibbons` and `plotPca`) produce complete figures as ggplot2 objects, that can be readily customized by adding or changing ggplot2 layers.

The figures in the Use Cases section (Section 5) are all produced directly from GeoChronR and generally fall into a few categories. The default graphical mode is used through the figures of this paper; this aesthetic is what GeoChronR produces by default. Vignettes illustrate how to customize these defaults.

4.1 Timeseries

The most common figure that users produce with GeoChronR are ensemble timeseries. GeoChronR uses two complementary approaches to visualize these ensembles. The first is the simplest, where a large subset of the ensemble members are plotted as semi-transparent lines. This approach, implemented in `plotTimeseriesEnsLines`, provides a faithful representation of the data, while the overlapping semi transparency provides a qualitative sense of the ensemble uncertainty structure. The second approach uses contours to more rigorlously visualize the structure of the time-value uncertainty space represented by the ensembles. `plotTimeseriesEnsRibbons` shows the quantiles of the ensembles at specified levels as shaded bands. This approach provides the quantitative uncertainty structure, but tends to smooth out the apparent temporal evolution of the data. Fortunately, the two approaches are complementary, and often the best approach is to quantify the ensemble distribution with ribbons in the background, and then overlap them with a handful of ensemble lines to illustrate the structure in representive ensemble members.

4.2 Maps

GeoChronR has simple mapping capabilities built in that rely on the `maps` (Becker et al., 2018) and `ggmap` (Kahle and Wickham, 2013) packages. The `mapLipd` and `mapTs` functions provide quick geospatial visualization of one or more datasets, but also serve as the basis for the visualization of ensemble spatial data produced by ensemble PCA analyses. In paleogeoscientific studies, the loadings (eigenvectors) of a PCA analysis are often portrayed as dots on a map, with a colorscale that highlights the sign and amplitude of the loadings. In ensemble PCA, the additional dimension of uncertainty in the loadings needs to be visualized as well. In GeoChronR, the median of the loadings is shown as a color, and the size of the symbol is inversely proportional to the spread of uncertainty across the ensemble. Consequently, large symbols depict loadings that are robust to the uncertainties, whereas small symbols show datasets whose loadings change substantially across the analysis. An example is shown in Section 5.4

4.3 Spectra

It is customary to plot spectra on a log-log scale, which helps separate the low powers and low frequencies. This choice also naturally highlights scaling laws (Lovejoy and Schertzer, 2013; Zhu et al., 2019) as linear structures in this reference frame. GeoChronR implements this convention by default, although the scales can be readily modified using `ggplot2`. In addition, the abscissa ($\log_{10} f$) is labeled according to the corresponding period, which are more intuitive than frequency to scientists reading the plot. To help identify significant periodicities, confidence limits can be superimposed, based on user-specified benchmarks (see 5). The `plotSpectrum` function visualizes single ensemble members (e.g. a median age model), while `plotSpectraEns` visualizes the quantiles of a distribution of age-uncertain spectra as ribbons, using the eponymous `ggplot` function. `periodAnnotate` allows to manually highlight periods of interest, layered onto an existing plot.

5 Use cases

We now illustrate the use of these tools on five use cases. The first example shows how a user might create age ensembles on different archives, and how to visualize the timing of abrupt events with appropriate uncertainty quantification. The second example walks through ensemble correlation of age-uncertain records. The third introduces the topic of age-uncertain calibration-in-time. The fourth provides an example of regional age-uncertain principal components analyses, and the fifth deals with spectral analysis. The complete details needed to reproduce these use cases are available in the RMarkdown source code for this manuscript, and are elaborated upon with additional detail in the "vignettes" included within the GeoChronR package, as well at http://lipdverse.org/geoChronR-examples/.

5.1 Creating an age ensemble

A common first task when using geoChronR is to create an age ensemble, either because the user is developing a new record, or because the age ensemble data for the record they are interested is unavailable. As described in section 2, workflows for four published age quantification programs are integrated into geoChronR. All four methods can be used simply in geoChronR with a LiPD file loaded into R that contains the chronological measurements, and the high-level functions `runBacon`, `runBchron`, `runOxcal` and `runBam`. These functions take LiPD objects as inputs, and return updated LiPD objects that include age-ensemble data generated by the respective software packages, with these data stored in appropriate tables. Typically, additional parameters are needed for to optimally run the algorithms. When these parameters are not specified, geoChronR will run in interactive mode, asking the user which variables and parameters they would like to model. These parameter choices are printed to the screen during while the program runs, or are available later with the function `getLastVarString`. By specifying these parameters, age model creation can be scripted and run in non-interactive mode. In this use case, we'll use geoChronR and BChron (Parnell et al., 2008) to calculate an age ensemble for the Hulu Cave δ^{18}O speleothem record (Wang et al., 2001), and BAM (Comboul et al., 2014) to simulate age uncertainties for the GISP2 ice core δ^{18}O dataset (Alley, 2000). The `plotChronEns` function will plot an age-depth model and uncertainties derived from the age ensemble (figure 2).

After an age ensemble has been added to a LiPD object, the user can visualize the ensemble timeseries using `plotTimeseriesEnsRi` and `plotTimeseriesEnsLines`. GISP2 δ^{18}O is plotted with age uncertainty, using both functions, in figure 3.

5.2 Correlation

Now that the user has generated age ensembles for the two datasets, they're interested to see if a correlation between the two datasets is robust to the age uncertainty modeled here. On multi-millennial timescales, the two datasets display similar features, to the extent that the well-dated Hulu Cave record has been used to support the independent chronology of GISP2 (Wang et al., 2001). In this use case, we revisit this conclusion quantitatively, and will use the age models created above, the `corEns` function in GeoChronR, to calculate the impact of age uncertainty on the correlation between these two iconic datasets. Here we calculate correlations during the period of overlap in 500 yr steps, determining significance for each pair of ensemble members while accounting for autocorrelation.

The results are visualized as a histogram of the ensemble correlation results, with color shading to highlight significance (figure 4). The two timeseries are consistently negative correlations, although 15% of the ensemble members are positive. However, only 0.3% are significant after accounting for serial autocorrelation. Given the age models and correlation parameters used here, it seems pretty unlikely that these two datasets are significantly correlated. However, it's worth noting that evaluating these results remains somewhat subjective, as there is no theoretical expectation for what fraction of ensemble correlation results should be expected to pass a significance test in age-uncertain correlation. Two timeseries, which are actually

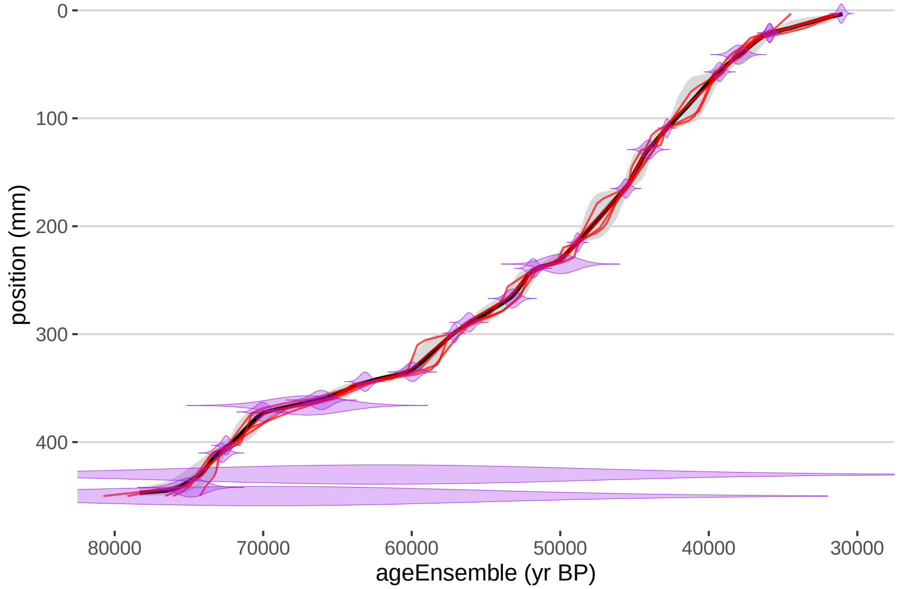

Figure 2. Age-depth model generated by BChron for a speleothem from Hulu Cave. Relative age probability distributions shown in purple. Median age-depth in black, with 50 and 95 percentile highest-density probability ranges shown in dark and light gray, respectively. Five random age-depth ensemble members shown in red.

correlated, each with it's own age-uncertainty, will commonly return some fraction of insignificant results when random ensemble members are correlated against each other. The frequency of these "false negatives" depend on the structure of the age-uncertainties and the timeseries, and will vary to some extent by random chance. One way to get a sense of the vulnerability of a timeseries to false negatives is to perform an age-uncertain correlation of a dataset with itself. It's appropriate to consider the results of this analysis a best-case scenario, and to consider the correlation results in this light. For illustration, we perform this calculation with the Hulu Cave δ^{18}O record (figure 5). The impact of age uncertainty on the correlation of this record is apparent; even when correlated against itself, only `2.1`% of the ensembles have r-values greater than 0.9, and the median correlation is `0.79`. However, all of the correlations remain significant, even after accounting for autocorrelation, indicating that age uncertainty and the structure of the timeseries does not preclude the possibility of significant correlations. Generally, it's appropriate to think of the ensemble correlation results produced by GeoChronR as a first-order estimate of the age-uncertain correlation characteristics between timeseries, rather than a binary answer to the question "Are these two datasets significantly correlated?". However, as a rule of thumb, if more than half of the ensemble correlation results are significant, it is reasonable to characterize that as a correlation that is robust to age uncertainty.

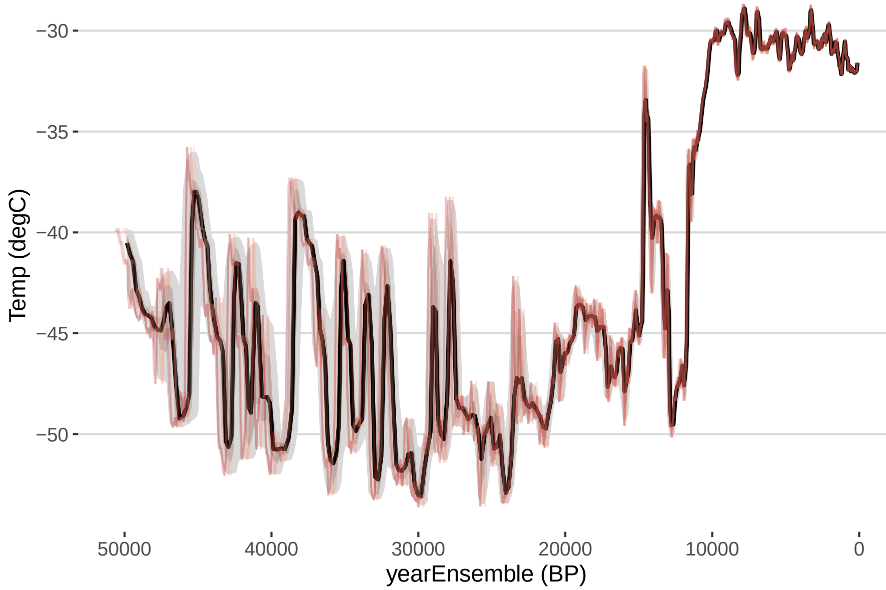

Figure 3. Impact of age uncertainty on reconstructed temperature at GISP2 over the past 50,000 years. The median ensemble member is shown in black, with the 50 and 95% highest-density probability ranges shown in dark and light gray, respectively. Five random age-uncertain temperature ensemble members shown in red.

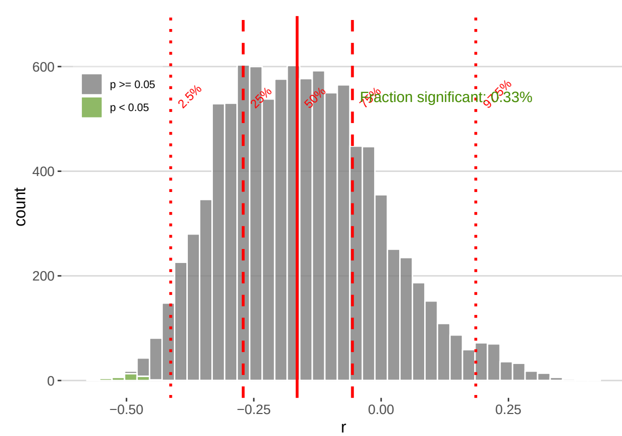

Figure 4. Distribution of age-uncertain correlation between Hulu Cave speleothem and GISP2 ice core δ^{18}O. Significant (insignificant) correlations shown in green (gray). Quantiles of the distribution shown in vertical red lines.

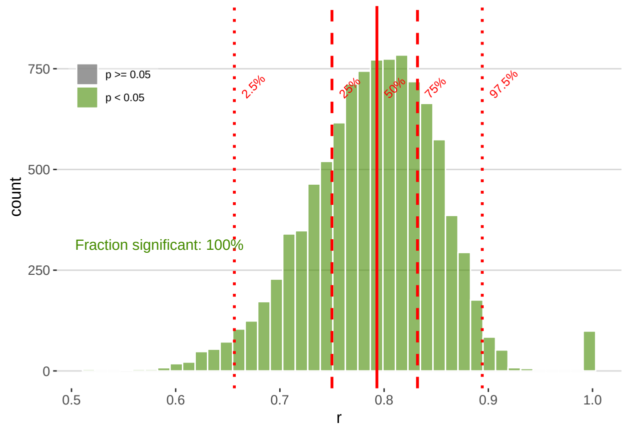

Figure 5. Distribution of age-uncertain correlation between Hulu Cave speleothem δ^{18}O and itself, treating the age uncertainty as if independent. Style as in figure
reffig:cor-hist.

5.3 Age-uncertain Calibration

A natural extension of ensemble correlation is ensemble regression. Although there are use cases where regressing one age-uncertain variable onto another is called for, here we regress an age-uncertain paleoclimate proxy onto time-certain instrumental to develop a calibration-in-time. For this use case, we reproduce the results of Boldt et al. (2015), where the authors calibrated a spectral reflectance measure of chlorophyll abundance, relative absorption band depth (RABD), to instrumental temperature in Northern Alaska. For each iteration in the analysis, a random age ensemble member is chosen and used to bin the RABD data onto a 3-year timescale. The instrumental temperature data, here taken from the nearest gridcell of the GISTEMP reanalysis product (Hansen et al., 2010), are also binned onto the same timescale, insuring temporal alignment between the two timeseries. GeoChronR then derives an ordinary linear regression model, and then uses that model to “predict” temperature values from 3-year-binned RABD data back in time. This approach propagates the age uncertainties both through the regression and prediction process.

The function `plotRegressEns` produces multiple plots that visualize the key results of age-uncertain regression, and additionally creates an overview “dashboard” that showcases the key results (figure 6). The first row of figure 6 illustrates the impact of age uncertainty on the regression modelling. In this example, the distribution of the modeled parameters, the slope and intercept of the regression equation, show pronounced modes of their distributions near 150 $^{\circ}C^{-1}$ and $^{\circ}C$, respectively, but with pronounced tails that include models with much lower slopes. This is also apparent in the scatterplot in the central panel of the top row, which illustrates the distribution of modeled relationships. Although the tendency for robust relationships are clear, models with slopes near zero also occur, suggesting that in this use case, age uncertainty can effectively destroy the relationship with instrumental data. The impact of this variability in modeled parameters, as well as the effects of age uncertainty on the

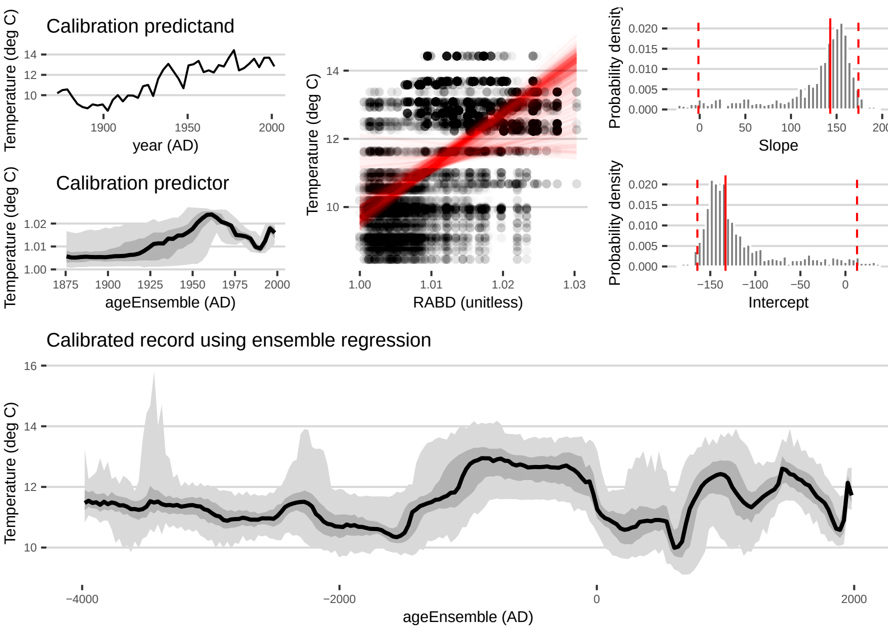

Figure 6. Results of ensemble regression. Top row, left: the calibration-in-time predictor and predictand. The predictand shows the effect of age uncertainty, the median is shown in black, and the 50 and 95% highest probability density regions are shown in dark and light gray, respectively. Top row, middle: Scatterplot showing the relation between RABD and Temperature. Points from 100 age ensemble members plotted with transparency to highligh data overlap. Ensemble regression models plotted as semi-transparent red lines. Top row, right: distribution of regression model slope (top) and intercepts (bottom). Vertical red lines show the 50th (solid), 2.5th and 97.5th (dashed) percentiles. Bottom row: reconstructed temperature using age ensembles and regression model ensembles. The median estimate is shown in black, and the 50 and 95% highest probability density regions are shown in dark and light gray, respectively.

timing of the reconstruction, are shown in the bottom panel of figure 6. The results shown here are consistent with those presented by Boldt et al. (2015), and we refer readers to that study for a full discussion of the implications of their results.

5.4 Principle Component Analysis

Thus far, the use cases have highlighted age-uncertain analyses of one or two locations, however quantifying the effects of age uncertainty can be even more impactful over large spatial datasets. Here we showcase how to use geoChronR to perform age-uncertain principle components analysis (PCA), also known as Monte Carlo Empirical Orthogonal Function (MCEOF) analysis, pioneered by Anchukaitis and Tierney (2013). When seeking to analyze a large collection datasets, the first, and often most time-intensive, step is to track down, format and standardize the data. Fortunately, the emergence of community-curated standardized data collections (e.g. PAGES2K Consortium, 2013; Emile-Geay et al., 2017; Kaufman et al., 2020b, ; Konecky et al., 2020) can greatly simplify this challenge. In this example, we examine the Arctic 2k database (McKay and Kaufman, 2014), and use GeoChronR and the LiPD Utilities to filter the data for temperature-sensitive data from the Atlantic Arctic with age ensembles relevant to the past 2000 years.

Once filtered, the data can be visualized using `plotTimeseriesStack`, which is an option to quickly plot all of the timeseries, on their best-estimate age models, aligned on a common horizontal timescale (figure 7). Although all of the datasets are relevant to Arctic temperatures over the past 2000 years, they span different time intervals, with variable temporal resolution. It is also clear that there is a lot of variability represented represented within the data, but it's difficult to visually extract shared patterns of variability. Ensemble PCA is well-suited to the modes of variability that explain the most variance within these data, while accounting for the impact of age uncertainty.

As in correlation and regression, aligning the data onto a common timescale is required for ensemble PCA. All but two of these datasets are annually resolved, and the other two have 5-year resolution, so it is reasonable to bin these data into 5 year bins. Furthermore, because many of the records do not include data before 1400 CE, we only analyze the period from 1400 to 2000 CE. The data are now prepared for the ensemble PCA calculation, following a few choices in methodology and parameters. Because the data analyzed here have variable units, and we are not interested in the magnitude of the variance only the relative variability between the datasets, we choose to use a correlation, rather than covariance, matrix. Next, we choose the number of components to estimate. After the analysis, a scree plot is used to determine the number of significant components, so we want to estimate several more components than we anticipate will be meaningful. For this use case we estimate eight components.

We now conduct the ensemble PCA, including null hypothesis testing, for 100 ensemble members. For a final analysis, 1000 ensemble members is standard, however the analysis can be time consuming and 100 members is appropriate for initial analyses. First, we plot the ensemble variance explained results for the data and the null hypothesis as a scree plot (figure 8). This represents how the variance explained by each component declines with each mode, for both the data and the null hypothesis. Due to age uncertainty, this variance explained is a distribution, which we compare to the 95th quantile of the null hypothesis ensemble. Figure 8 indicates that the first two components are clearly distinguished from the null. The third component is borderline, with the variance explained by the median of the ensemble near the null. Therefore, we will focus our investigation on the first two modes.

The spatial and temporal results of the first two principle components are shown in figure 9. The first PC is dominated by consistently positive loadings across the North Atlantic, suggesting that this is a regionally persistent mode of variability, and indicating that none of the datasets are negatively correlated with this mode. The corresponding timeseries shows multidecadal variability, with values declining until the 18th Century, before increasing into the 20th century. Based on the region-wide coherence of the loading pattern and the similarity of the timeseries to regional temperature reconstructions (Hanhijärvi et al., 2013; McKay and Kaufman, 2014; Werner et al., 2018), the first PC likely reflects the primary pattern of regional temperature variability. Notably, the uncertainties in the PC1 timeseries, and the loadings in the spatial pattern, are generally small. This, combined with the large amount of variance explained by PC1 relative to the null hypothesis (figure 8), suggests this is a significant mode of variabity that is robust to age uncertainty. This makes intuitive sense, since the primary features of this

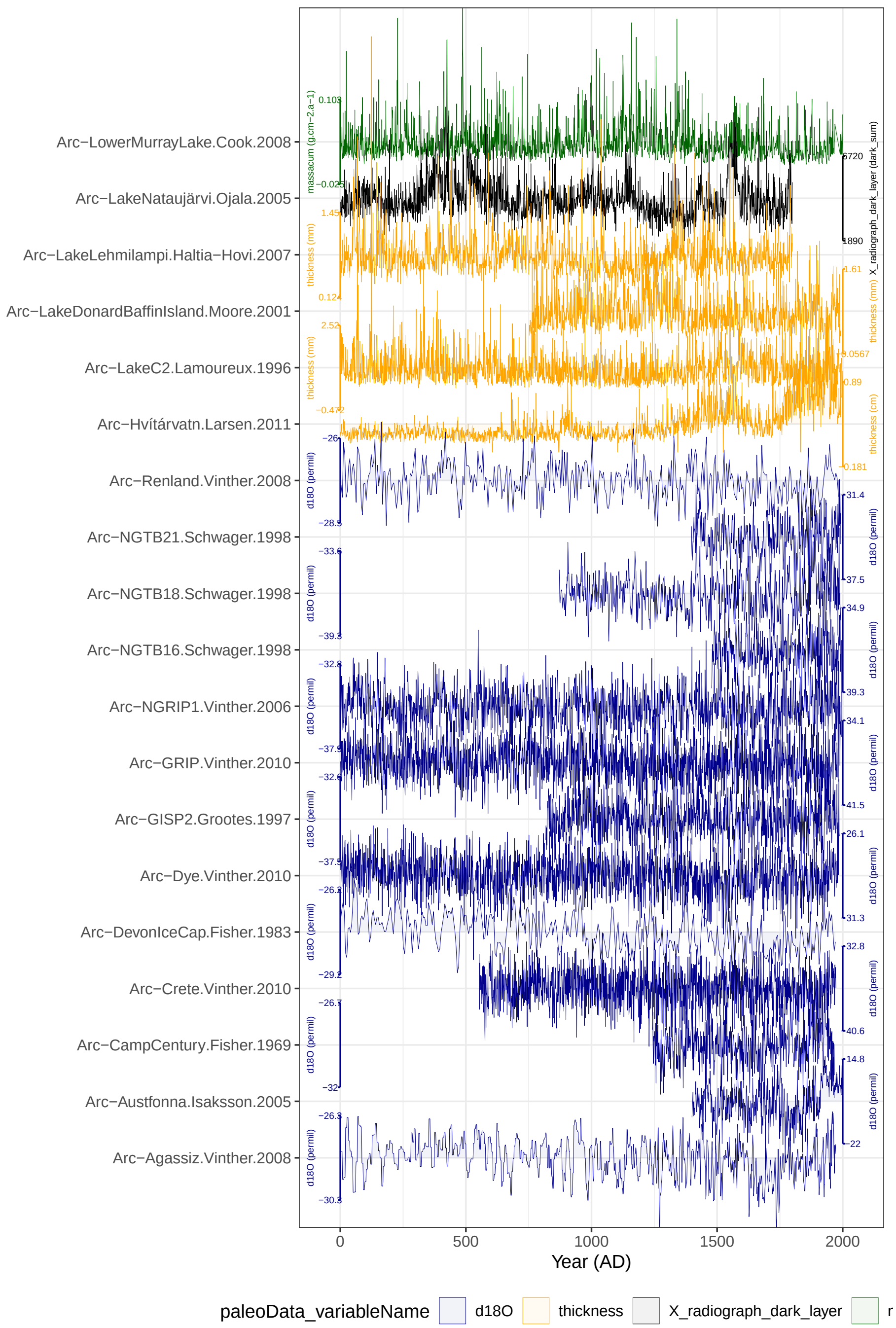

Figure 7. Temperature-sensitive timeseries from the Atlantic Arctic in [@McKayKaufman2014], arranged vertically by variable. Each timeseries is shown on its own scale to the right or left of each timeseries.

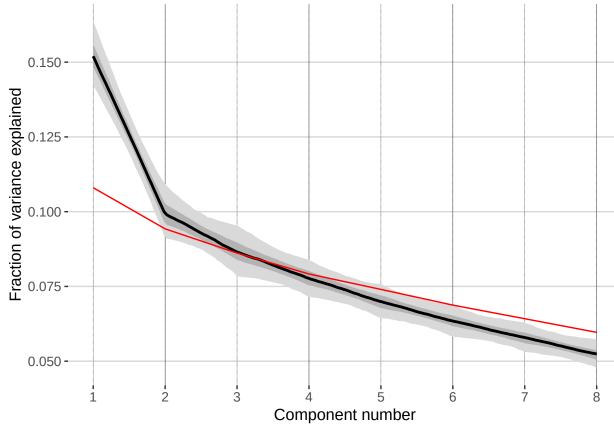

Figure 8. Ensemble PCA scree plot, showing the fraction of variance explained as a function of component (or eigenvalue) number. The ensemble PCA results for the data have uncertainty due to age-uncertainty, and themedian is shown in black, and the 50 and 95% highest probability density regions are shown in dark and light gray, respectively. The 95th percentile of the null hypothesis test is shown in red.

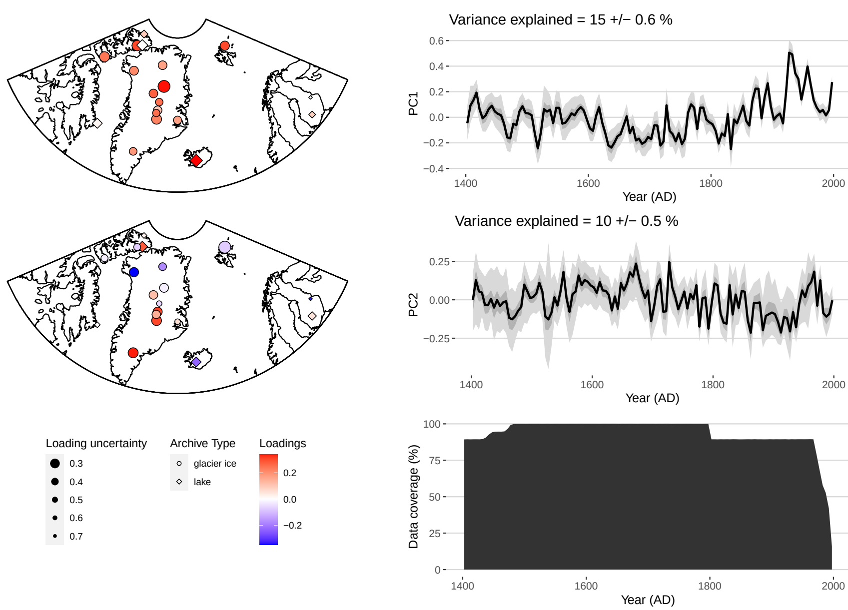

Figure 9. The spatial loading pattern (left) and timeseries (right) for the first principle component shown in the first row. The results for the second PC are shown in the second. The data density through time is shown in the bottom right corner. For the maps, the median loadings of the ensemble are shown by the color scale, and the standard deviation of the loadings across ensemble members is depicted by the size of the markers, with larger markers showing smaller uncertainties. For the timeseries plots, the median of the ensemble is shown in black, and the 50 and 95% highest probability density regions are shown in dark and light gray, respectively.

pattern are century-long trends in temperature, a timescale that substantially exceeds the age uncertainty in these data (McKay and Kaufman, 2014).

The second PC shows considerable more variability in its spatial loading pattern, and a larger impact of age uncertainty. Generally, the loadings suggest a north-south dipole over Greenland for this mode, with positive loadings present in much of southern Greenland, with negative loadings in present in much of the northern part of the region. There is a much larger impact of age uncertainty on the loadings in PC2 than in PC1, illustrated by the size of the markers on the map, which are inversely related to the standard deviation of the loadings across the ensemble PCA results, such that smaller markers indicate larger uncertainties. The PC2 timeseries includes more multidecadal variability than PC1 and is more impacted by age uncertainty. A key feature of the timeseries is a peak in values in the late 20th century, which occurs after the pronounced peak in PC1. This suggests that unlike the mid-20th century peak in warming apparent in most of the data, this later warming was dominated by contributions from southern Greenland, and counterbalanced by a decline in values in the northern Atlantic Arctic.

5.5 Spectral Analysis

To illustrate the use of spectral analysis in GeoChronR, we consider a use case where the seeks to identify the relative energy of oscillations at orbital (Milankovitch) periodicities in a deep-sea sediment core, and quantify the impact of age uncertainties on this assessment. Here we use a benthic paleotemperature record derived from the International Ocean Drilling Project core 846 (Mix et al., 1995; Shackleton, 1995), that covers the past 4.7 million years. For this assessment,s we use an updated age model that was not generated within GeoChronR, rather, the age model was created via alignment to the benthic δ^{18}O stack of Lisiecki and Raymo (2005) using the HMM-Match algorithm (Lin et al., 2014; Khider et al., 2017). HMM-Match is a probabilistic method that generates an ensemble of 1000 possible age models compatible with the chronostratigraphic constraints; this ensemble was archived as a table in the associated LiPD file.

First we use `plotTimeseriesEnsRibbons` to visualize temperature, and the impact of age uncertainty, over the past 5 million years (figure 10).

This record displays three salient features:

- a long-term cooling trend characteristic of the late Neogene and Quaternary climate.
- quasi-periodic oscillations (the legendary Pleistocene Ice Ages)
- nonstationary behavior, related to the well-known mid-Pleistocene transition from a “41k world” to a “100k world” somewhere around 0.8 Ma (Paillard, 2001).

For tractability, let us focus on the last million years, which cover the Quaternary Era. Over this interval, the time increments (Δt) are sharply peaked around 2.5 ka, spanning 0 to about 7.5 ka. From this point there are two ways to proceed: 1) use methods that explictly deal with unevenly-spaced data, or 2) interpolate to a regular grid and apply standard methods (see section @ref(sec:spec_theory)). In this use case, we will use both approaches and highlight two of the four spectral methods impemented in GeoChronR: REDFIT and MTM.

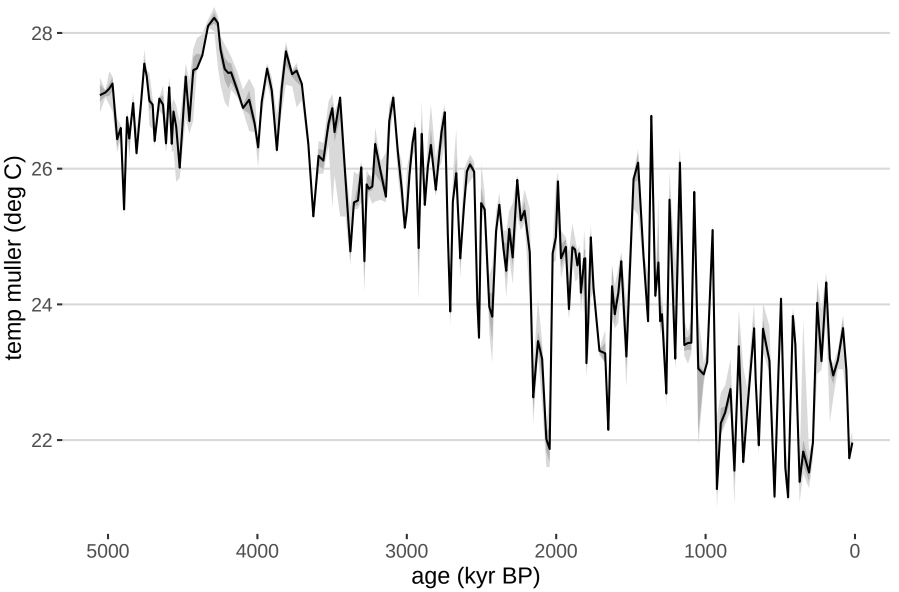

Figure 10. Temperature reconstruction from IODP 846

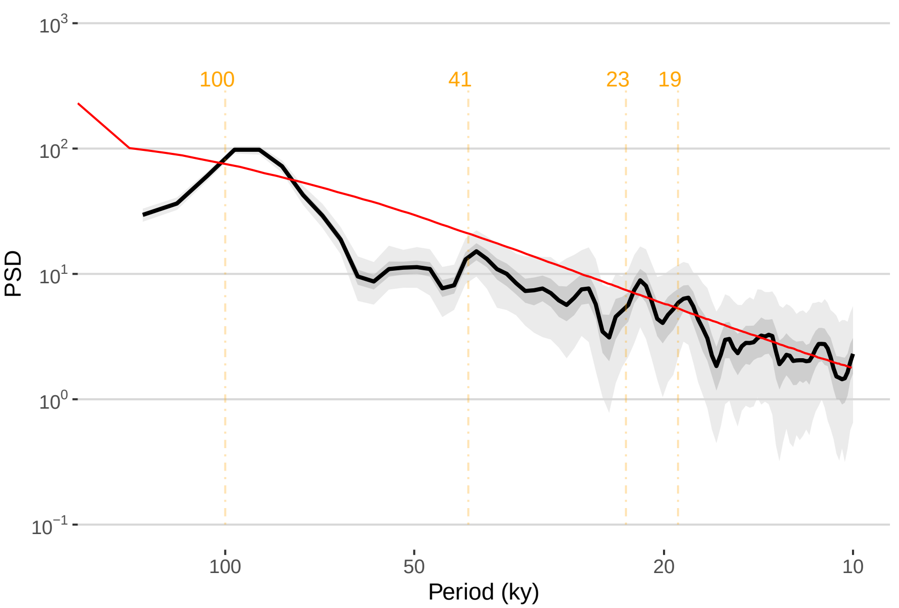

Figure 11. REDFIT spectrum for IODP 846. Median spectrum shown in black, with the 50 and 95% highest-density probability ranges shown in dark and light gray, respectively. Periodicities of interest (in kyr) shown in orange.

We use the `computeSpectraEns` option to calculate the spectra for 1000 ensemble members using the REDFIT approach (figure 11). It is clear that the data contain significant energy (peaks) near, but not exactly at, the Milankovitch periodicities (100, 41, 23, and 19 kyr).

These periodicities, particularly those associated with eccentricity (100 kyr) and precession (23 and 19 kyr), rise above the null hypothesis (the 95% quantile from an autoregressive process of order one, see Mudelsee et al. (2009)). The obliquity periodicity is relatively weak, reaching just below the AR(1) benchmark.

The Lomb-Scargle periodogram used by REDFIT is a common way to deal with unevenly-spaced timeseries, but like all periodograms, it is inconsistent: the variance (uncertainty) of spectral density at each frequency does not decrease with the number of observations. This is mitigated somewhat with the application of Welch's Overlapping Segment Averaging, but their parameter choice is fairly arbitrary. On the other hand, MTM (Thomson, 1982) is an optimal estimator, which is consistent (the more obervations, the better constrained the spectral density). Formally, MTM optimizes the classic bias-variance tradeoff inherent to all statistical inference. It does so by minimizing speactral leakage outside of a frequency band with half-bandwidth equal to pf_R, where $f_R = 1/(N\Delta t)$ is the Rayleigh frequency, Δt is the sampling interval, N the number of measurements, and p is the so-called *time-bandwidth product* (Ghil et al., 2002). p can only take a finite number of values, all multiples of 1/2 between 2 and 4. A larger p means lower variance (i.e. less uncertainty about the power), but broader peaks (i.e. a lower spectral resolution), synonymous with more uncerainty about the exact location of the peak. So while MTM might not distinguish between closely-spaced harmonics, it is much less likely to identify spurious peaks, especially at high frequencies. In addition, a battery of formal tests have been devised with MTM, allowing us to ascertain the significance of spectral peaks under reasonably broad assumptions. We show how to use this "harmonic F-test" below.

However, classic MTM can only handle evenly-spaced data. Since the data are close to evenly-spaced, is is reasonable to interpolate them using standard methods. Both interpolation and MTM are implemented with the (astrochron Meyers, 2014) package, which GeoChronR employs.

To this we can add the periods identified as significant by MTM's F ratio test. GeoChronR estimates this by computing the fraction of ensemble members that exhbit a significant peak at each frequency. One simple criterion for gauging the level of support for such peaks given age uncertainties is to pick out those periodicities that are identified as significant above a certain threshold (say, more than 50% of the time).

For consistency with REDFIT, we define the null as an AR(1) process fit to the data, but GeoChronR supports two other nulls: a power-law null and a fit to the spectral background (Mann and Lees, 1996). Both follow the `astrochron` implementation.

You may notice a few differences between the REDFIT estimate (figure 11 and the MTM estimate (figure 12). First, this ensemble of spectra exhibits a clear power law behavior from periods of 5 to 100 ky, which in this log-log plotting convention manifests as a linear decrease. This is part of the well-documented continuum of climate variability (Huybers and Curry, 2006; Zhu et al., 2019), which is conspicuously absent from the Lomb-Scargle (REDFIT) estimate, known to be extremely biased in its estimate of the spectral background.

Secondly, the MTM version with this time-bandwidth product is sharper than REDFIT, with more well-defined peaks, particularly for the obliquity period (41 ky), which clearly exceeds the 95% confidence limit. Here it is helpful to take a step

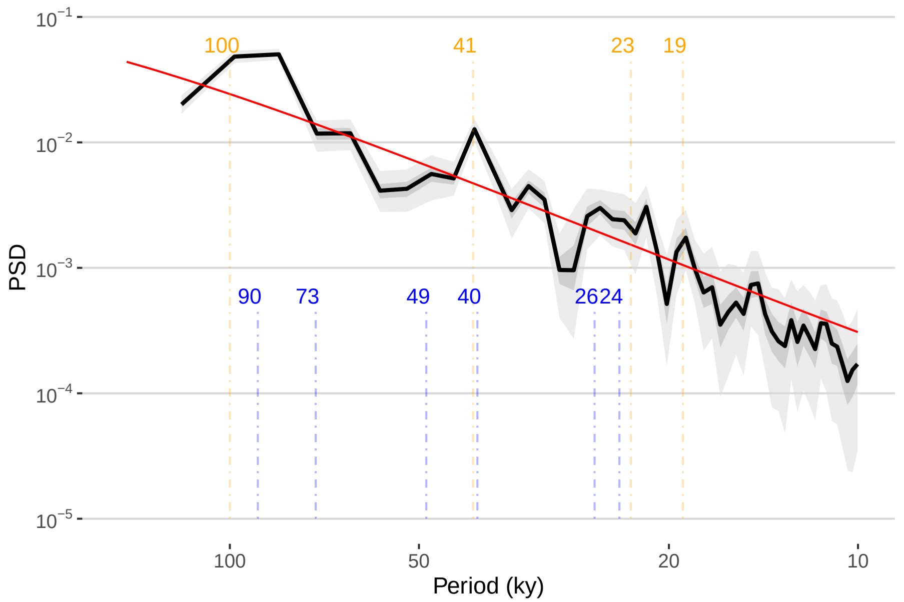

Figure 12. MTM spectrum for IODP 846. Median spectrum shown in black, with the 50 and 95% highest-density probability ranges shown in dark and light gray, respectively. Periodicities of interest (in kyr) shown in orange, and the periodicities identified as significant by the F ratio test are shown in blue.

back and contemplate our null hypothesis of AR(1) background, and the possibility that we might be underestimating the lag-1 autocorrelation, hence making the test too lenient. More importantly, the presence of scaling behavior (power law decrease) suggests that this should be a more appropriate null against which to test the emergence of spectral peaks. In GeoChronR, this can be done by specifying `mtm_null = "power_law` in the function call.

Using a power law null hypothesis makes a few cycles appear non significant, but many remain (not shown). However, carrying out a test simultaneously at many periodicities is bound to affect assesments of significance, via the multiple comparisons problem (Vaughan et al., 2011). In addition, sedimentary processes (and many processes in other proxy archives) tend to smooth out the signal over the depth axis, making comparisons at neighboring frequencies highly dependent (Meyers, 2012). One solution is to use predictions made by a physical model about the frequency and relative amplitude of astronomical cycles (Meyers and Sageman, 2007). This approach, however, is not be applicable to all spectral detection problems. Ultimately, the user must think deeply about the null hypothesis and the most sensible way to test it. Readers are invited to consider the literature for a deeper exploration of these questions (e.g., Vaughan et al., 2011; Meyers, 2012, 2015; Meyers and Malinverno, 2018).

As with all statistical analyses in the paleosciences, there are no universal solutions or parameter choices. The approaches implemented in GeoChronR, especially with default choices, are best considered as exploratory tools, that provide insight into the impacts of age uncertainty on power spectra, and to help users to tailor their null hypotheses to their scientific questions.

6 Conclusions

GeoChronR provides user-friendly access to common age-uncertain analysis tools in the paleogeosciences, along with intuitive visualization of the results. Although the focus has been on simplicity and ease of use, GeoChronR also has the underlying infrastructure to support customized analyses for users seeking to address a more nuanced or complex question. Nevertheless, in many ways the paleoscience community is only scratching the surface with age-uncertain analysis, and we look forward to working with the community to extend and expand this open-source package. Looking forward, we suggest that the next major direction for age-uncertain analysis is philosophical, not technical. Thus far, the community (and GeoChronR) has focused on quantifying the range of possibilities presented by age-uncertainty, not on developing approaches to constrain which ensemble members are most likely. Theoretically, additional information from nearby records, forcings or covariance structures could do so (Werner and Tingley, 2015), but little has been done to date.

GeoChronR is open-source community-sofware, and has benefitted substantially from multiple contributors and input from early adopters and workshop participants. We welcome feedback and strongly encourage contributions and enhancements, via the GitHub issue tracker.

Code availability. All of the code used in GeoChronR is open and available at https://github.com/nickmckay/geochronr, and we welcome contributions and extensions to the package. The Rmarkdown code used to create this manuscript is available at https://github.com/nickmckay/geochronr-paper.

Data availability. All of the data used in this paper are publicly archived, and available as LiPD files at http://lipdverse.org/geoChronR-examples/.

Competing interests. The authors declare no competing interests.

Disclaimer. The software described here is provided “as is”, without warranty of any kind, express or implied, including but not limited to the warranties of merchantability, fitness for a particular purpose and noninfringement. In no event shall the authors or copyright holders be liable for any claim, damages or other liability, whether in an action of contract, tort or otherwise, arising from, out of or in connection with the software or the use or other dealings in the software.

Acknowledgements. This project was funded by the US National Science Foundation Geoinformatics Program (EAR-1347221). We thank the participants of the GeoChronR workshops for their contributions to the testing and design, which dramatically improved this package.

References

Alley, R. B.: Ice-core evidence of abrupt climate changes, Proceedings of the National Academy of Sciences, 97, 1331–1334, 2000.

Anchukaitis, K. J. and Tierney, J. E.: Identifying coherent spatiotemporal modes in time-uncertain proxy paleoclimate records, Climate dynamics, 41, 1291–1306, 2013.

Balco, G., Briner, J., Finkel, R. C., Rayburn, J. A., Ridge, J. C., and Schaefer, J. M.: Regional beryllium-10 production rate calibration for late-glacial northeastern North America, Quaternary Geochronology, 4, 93–107, https://doi.org/10.1016/j.quageo.2008.09.001, http://www.sciencedirect.com/science/article/pii/S1871101408000502, 2009.

Becker, R. A., Wilks, A. R., Minka, R. B. T. P., and Deckmyn, A.: maps: Draw Geographical Maps, https://CRAN.R-project.org/package=maps, r package version 3.3.0, 2018.

Benjamini, Y. and Hochberg, Y.: Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing, Journal of the Royal Statistical Society. Series B (Methodological), 57, 289–300, https://doi.org/10.2307/2346101, 1995.

Bhattacharya, T. and Coats, S.: Atlantic-Pacific Gradients Drive Last Millennium Hydroclimate Variability in Mesoamerica, Geophysical Research Letters, 47, e2020GL088 061, 2020.

Blaauw, M.: Methods and code for "classical" age-modelling of radiocarbon sequences, Quaternary Geochronology, 5, 512–518, 2010.

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

Blaauw, M. and Christen, J. A.: Flexible paleoclimate age-depth models using an autoregressive gamma process, Bayesian Analysis, 6, 457–474, http://projecteuclid.org/euclid.ba/1339616472, 2011b.

Blaauw, M., Christen, J. A., and Aquino L., M. A.: rbacon: Age-Depth Modelling using Bayesian Statistics, https://CRAN.R-project.org/package=rbacon, r package version 2.4.2, 2020.

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

Boldt, B. R., Kaufman, D. S., McKay, N. P., and Briner, J. P.: Holocene summer temperature reconstruction from sedimentary chlorophyll content, with treatment of age uncertainties, Kurupa Lake, Arctic Alaska, The Holocene, https://doi.org/10.1177/0959683614565929, 2015.

Bond, G., Kromer, B., Beer, J., Muscheler, R., Evans, M. N., Showers, W., Hoffmann, S., Lotti-Bond, R., Hajdas, I., and Bonani, G.: Persistent Solar Influence on North Atlantic Climate During the Holocene, Science, 294, 2130–2136, 2001.

Bretherton, C. S., Widmann, M., Dymnikov, V. P., Wallace, J. M., and Bladé, I.: The effective number of spatial degrees of freedom of a time-varying field, Journal of climate, 12, 1990–2009, 1999.

Bronk Ramsey, C.: Radiocarbon calibration and analysis of stratigraphy: The OxCal program, Radiocarbon, 37, 425–430, 1995.

Brown, T. A., Nelson, D., Mathewes, R. W., Vogel, J. S., and Southon, J. R.: Radiocarbon dating of pollen by accelerator mass spectrometry, Quaternary Research, 32, 205–212, https://doi.org/10.1016/0033-5894(89)90076-8, http://www.sciencedirect.com/science/article/pii/0033589489900768, 1989.

Bunn, A. G.: A dendrochronology program library in R (dplR), Dendrochronologia, 26, 115–124, https://doi.org/10.1016/j.dendro.2008.01.002, 2008.

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

Dee, S., Emile-Geay, J., Evans, M. N., Allam, A., Steig, E. J., and Thompson, D. M.: PRYSM: An open-source framework for PRoxY System Modeling, with applications to oxygen-isotope systems, Journal of Advances in Modeling Earth Systems, 7, 1220–1247, 2015.

Deininger, M., McDermott, F., Mudelsee, M., Werner, M., Frank, N., and Mangini, A.: Coherency of late Holocene European speleothem δ 18 O records linked to North Atlantic Ocean circulation, Climate Dynamics, 49, 595–618, 2017.

DeLong, K. L., Quinn, T. M., Taylor, F. W., Shen, C.-C., and Lin, K.: Improving coral-base paleoclimate reconstructions by replicating 350 years of coral Sr/Ca variations, Palaeogeography, Palaeoclimatology, Palaeoecology, 373, 6 – 24, https://doi.org/10.1016/j.palaeo.2012.08.019, http://www.sciencedirect.com/science/article/pii/S0031018212004956, 2013.

Eggins, S. M., Grün, R., McCulloch, M. T., Pike, A. W., Chappell, J., Kinsley, L., Mortimer, G., Shelley, M., Murray-Wallace, C. V., and Spötl, C.: In situ U-series dating by laser-ablation multi-collector ICPMS: new prospects for Quaternary geochronology, Quaternary Science Reviews, 24, 2523–2538, http://www.sciencedirect.com/science/article/pii/S0277379105002040, 2005.

Eglinton, T. I., Aluwihare, L. I., Bauer, J. E., Druffel, E. R., and McNichol, A. P.: Gas chromatographic isolation of individual compounds from complex matrices for radiocarbon dating, Analytical Chemistry, 68, 904–912, http://pubs.acs.org/doi/abs/10.1021/ac9508513, 1996.

Emile-Geay, J., McKay, N. P., Kaufman, D. S., Von Gunten, L., Wang, J., Anchukaitis, K. J., Abram, N. J., Addison, J. A., Curran, M. A., Evans, M. N., et al.: A global multiproxy database for temperature reconstructions of the Common Era, Scientific data, 4, 170 088, 2017.

Falster, G., Tyler, J., Grant, K., Tibby, J., Turney, C., Löhr, S., Jacobsen, G., and Kershaw, A. P.: Millennial-scale variability in south-east Australian hydroclimate between 30,000 and 10,000 years ago, Quaternary Science Reviews, 192, 106–122, 2018.

Fifield, L. K.: Advances in accelerator mass spectrometry, Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 172, 134–143, http://www.sciencedirect.com/science/article/pii/S0168583X00002299, 2000.

Foster, G.: Wavelets for period analysis of unevenly sampled time series, Astron. Jour., 112, 1709, https://doi.org/10.1086/118137, 1996.

Ghil, M., Allen, R. M., Dettinger, M. D., Ide, K., Kondrashov, D., Mann, M. E., Robertson, A., Saunders, A., Tian, Y., Varadi, F., and Yiou, P.: Advanced spectral methods for climatic time series, Rev. Geophys., 40, 1003–1052, 2002.

Grosjean, M., von Gunten, L., Trachsel, M., and Kamenik, C.: Calibration-in-time: Transforming biogeochemical lake sediment proxies into quantitative climate variables, Pages News, 17, 108–110, 2009.

Haam, E. and Huybers, P.: A test for the presence of covariance between time-uncertain series of data with application to the Dongge Cave speleothem and atmospheric radiocarbon records, Paleoceanography, 25, PA2209, https://doi.org/10.1029/2008PA001713, 2010.

Hanhijärvi, S., Tingley, M. P., and Korhola, A.: Pairwise comparisons to reconstruct mean temperature in the Arctic Atlantic Region over the last 2,000 years, Climate dynamics, 41, 2039–2060, 2013.

Hansen, J., Ruedy, R., Sato, M., and Lo, K.: Global surface temperature change, Reviews of Geophysics, 48, 2010.

Haslett, J. and Parnell, A.: A simple monotone process with application to radiocarbon-dated depth chronologies, Journal of the Royal Statistical Society: Series C (Applied Statistics), 57, 399–418, https://doi.org/10.1111/j.1467-9876.2008.00623.x, http://dx.doi.org/10.1111/j.1467-9876.2008.00623.x, 2008.

Heegaard, E., Birks, H. J. B., and Telford, R. J.: Relationships between calibrated ages and depth in stratigraphical sequences: an estimation procedure by mixed-effect regression, The Holocene, 15, 612–618, http://hol.sagepub.com/content/15/4/612.short, 2005.

Hu, J., Emile-Geay, J., and Partin, J.: Correlation-based interpretations of paleoclimate data – where statistics meet past climates, Earth and Planetary Science Letters, 459, 362–371, https://doi.org/10.1016/j.epsl.2016.11.048, 2017.

Huybers, P. and Curry, W.: Links between annual, Milankovitch and continuum temperature variability, Nature, 441, 329–332, http://dx.doi.org/10.1038/nature04745, 2006.

Huybers, P. and Wunsch, C.: A depth-derived Pleistocene age model: Uncertainty estimates, sedimentation variability, and nonlinear climate change, Paleoceanogr., 19, 1028–+, https://doi.org/10.1029/2002PA000857, 2004.

Imbrie, J., Hays, J., Martinson, D., Mcintyre, A., Mix, A., Morley, J., Pisias, N., Prell, W., and Shackleton, N.: The orbital theory of Pleistocene climate change: Support from a revised chronology of the marine $\delta^{18}O$ record, in: Milankovitch and Climate, edited by Berger, A., Imbrie, J., Hays, J., and Kukla, J., pp. 269–305, 1984.

Kahle, D. and Wickham, H.: ggmap: Spatial Visualization with ggplot2, The R Journal, 5, 144–161, https://journal.r-project.org/archive/2013-1/kahle-wickham.pdf, 2013.

Kaufman, D., McKay, N., Routson, C., Erb, M., Dätwyler, C., Sommer, P. S., Heiri, O., and Davis, B.: Holocene global mean surface temperature, a multi-method reconstruction approach, Scientific Data, 7, 1–13, 2020a.

Kaufman, D., McKay, N., Routson, C., Erb, M., Davis, B., Heiri, O., Jaccard, S., Tierney, J., Dätwyler, C., Axford, Y., et al.: A global database of Holocene paleotemperature records, Scientific data, 7, 1–34, 2020b.

Khider, D., Ahn, S., Lisiecki, L. E., Lawrence, C. E., and Kienast, M.: The Role of Uncertainty in Estimating Lead/Lag Relationships in Marine Sedimentary Archives: A Case Study From the Tropical Pacific, Paleoceanography, 32, 1275–1290, https://doi.org/10.1002/2016PA003057, https://doi.org/10.1002/2016PA003057, 2017.

Konecky, B. L., McKay, N. P., Comas-Bru, L., Dassié, E. P., DeLong, K. L., Falster, G. M., Fischer, M. J., Jones, M. D., Jonkers, L., Kaufman, D. S., et al.: The Iso2k Database: A global compilation of paleo-δ18O and δ2H records to aid understanding of Common Era climate, Earth System Science Data, 2020.

Lin, L., Khider, D., Lisiecki, L., and Lawrence, C.: Probabilistic sequence alignment of stratigraphic records, Paleoceanography, 29, 976–989, https://doi.org/doi:10.1002/2014PA002713, 2014.

Lisiecki, L. E. and Raymo, M. E.: A Pliocene-Pleistocene stack of 57 globally distributed benthic δ^{18}O records, Paleoceanography, 20, PA1003, https://doi.org/10.1029/2004PA001071, 2005.

Lovejoy, S. and Schertzer, D.: The Weather and Climate: Emergent Laws and Multifractal Cascades, Cambridge University Press, https://books.google.com/books?id=SeBjgLD43IIC, 2013.

Mann, M. and Lees, J.: Robust Estimation of Background Noise and Signal Detection in Climatic Time Series, Clim. Change, 33, 409–445, 1996.

Marcott, S. A., Shakun, J. D., Clark, P. U., and Mix, A. C.: A Reconstruction of Regional and Global Temperature for the Past 11,300 Years, Science, 339, 1198–1201, https://doi.org/10.1126/science.1228026, http://www.sciencemag.org/content/339/6124/1198.abstract, 2013.

Martin, H., Schmid, C., Knitter, D., and Tietze, C.: oxcAAR: Interface to 'OxCal' Radiocarbon Calibration, https://CRAN.R-project.org/package=oxcAAR, r package version 1.0.0, 2018.

Masarik, J. and Beer, J.: An updated simulation of particle fluxes and cosmogenic nuclide production in the Earth's atmosphere, Journal of Geophysical Research: Atmospheres (1984–2012), 114, http://onlinelibrary.wiley.com/doi/10.1029/2008JD010557/full, 2009.

Mathias, A., Grond, F., Guardans, R., Seese, D., Canela, M., and Diebner, H.: Algorithms for Spectral Analysis of Irregularly Sampled Time Series, Journal of Statistical Software, Articles, 11, 1–27, https://doi.org/10.18637/jss.v011.i02, https://www.jstatsoft.org/v011/i02, 2004.

McKay, N. P.: neotoma2lipd package, https://github.com/nickmckay/neotoma2lipd/, 2020.

McKay, N. P. and Emile-Geay, J.: Technical note: The Linked Paleo Data framework : a common tongue for paleoclimatology, Climate of the Past, 12, 1093–1100, https://doi.org/10.5194/cp-12-1093-2016, http://www.clim-past.net/12/1093/2016/, 2016.

McKay, N. P. and Kaufman, D. S.: An extended Arctic proxy temperature database for the past 2,000 years, Scientific Data, 1, https://doi.org/10.1038/sdata.2014.26, 2014.

McKay, N. P., Kaufman, D. S., Routson, C. C., Erb, M. P., and Zander, P. D.: The onset and rate of Holocene Neoglacial cooling in the Arctic, Geophysical Research Letters, 45, 12–487, 2018.

Meyers, S. R.: Seeing red in cyclic stratigraphy: Spectral noise estimation for astrochronology, Paleoceanography, 27, https://doi.org/10.1029/2012PA002307, https://doi.org/10.1029/2012PA002307, 2012.

Meyers, S. R.: Astrochron: An R Package for Astrochronology, https://cran.r-project.org/package=astrochron, 2014.

Meyers, S. R.: The evaluation of eccentricity-related amplitude modulation and bundling in paleoclimate data: An inverse approach for astrochronologic testing and time scale optimization, Paleoceanography, 30, 1625–1640, https://doi.org/10.1002/2015PA002850, https://doi.org/10.1002/2015PA002850, 2015.

Meyers, S. R. and Malinverno, A.: Proterozoic Milankovitch cycles and the history of the solar system, Proceedings of the National Academy of Sciences, 115, 6363–6368, https://doi.org/10.1073/pnas.1717689115, https://www.pnas.org/content/115/25/6363, 2018.

Meyers, S. R. and Sageman, B. B.: Quantification of deep-time orbital forcing by average spectral misfit, Am J Sci, 307, 773–792, https://doi.org/10.2475/05.2007.01, 2007.

Mix, A. C., Le, J., and Shackleton, N.: Benthic foraminiferal stable isotope stratigraphy of site 846: 0-1.8 Ma, in: Proceedings of the Ocean Drilling Program. Scientific Results, vol. 138, 1995.

Mudelsee, M.: TAUEST: a computer program for estimating persistence in unevenly spaced weather/climate time series, Computers and Geosciences, 28, 69–72, https://doi.org/10.1016/S0098-3004(01)00041-3, 2002.

Mudelsee, M., Scholz, D., Röthlisberger, R., Fleitmann, D., Mangini, A., and Wolff, E. W.: Climate spectrum estimation in the presence of timescale errors, Nonlinear Processes in Geophysics, 16, 43–56, https://doi.org/10.5194/npg-16-43-2009, http://www.nonlin-processes-geophys.net/16/43/2009/, 2009.

National Academies of Sciences, Engineering, and Medicine: A Vision for NSF Earth Sciences 2020-2030: Earth in Time, The National Academies Press, Washington, DC, https://doi.org/10.17226/25761, https://www.nap.edu/catalog/25761/a-vision-for-nsf-earth-sciences-2020-2030-earth-in, 2020.

Neumaier, A. and Schneider, T.: Estimation of parameters and eigenmodes of multivariate autoregressive models, ACM Transactions on Mathematical Software (TOMS), 27, 27–57, 2001.

Noren, A., Brigham-Grette, J., Lehnert, K., Peters, S., Williams, J., Ito, E., Anderson, D., and Grimm, E.: Cyberinfrastructure for Paleogeoscience, workshop report, NSF EarthCube, 2013.

PAGES2K Consortium: Continental-scale temperature variability during the past two millennia, Nature Geosci, 6, 339–346, https://doi.org/10.1038/ngeo1797, http://dx.doi.org/10.1038/ngeo1797, 2013.

Paillard, D.: Glacial cycles: Toward a new paradigm, Rev. Geophys., 39, 325–346, https://doi.org/10.1029/2000RG000091, 2001.

Parnell, A., Haslett, J., Allen, J., Buck, C., and Huntley, B.: A flexible approach to assessing synchroneity of past events using Bayesian reconstructions of sedimentation history, Quaternary Science Reviews, 27, 1872–1885, 2008.

Parnell, A. C., Buck, C. E., and Doan, T. K.: A review of statistical chronology models for high-resolution, proxy-based Holocene palaeoenvironmental reconstruction, Quaternary Science Reviews, 30, 2948 – 2960, https://doi.org/10.1016/j.quascirev.2011.07.024, http://www.sciencedirect.com/science/article/pii/S0277379111002356, 2011.

Preisendorfer, R. W. and Mobley, C. D.: Principal component analysis in meteorology and oceanography, Developments in atmospheric science, 17, 1988.

Ramsey, C. B.: Deposition models for chronological records, Quaternary Science Reviews, 27, 42–60, 2008.

Ramsey, C. B.: Bayesian analysis of radiocarbon dates, Radiocarbon, 51, 337–360, https://journals.uair.arizona.edu/index.php/radiocarbon/article/view/3494, 2009.

Ramsey, C. B., Dee, M., Lee, S., Nakagawa, T., and Staff, R. A.: Developments in the calibration and modeling of radiocarbon dates, Radiocarbon, 52, 953–961, 2010.

Reimer, P., Baillie, M., Bard, E., Bayliss, A., Beck, J., Blackwell, P., Ramsey, C., Buck, C., Burr, G., Edwards, R., Friedrich, M., Grootes, P., Guilderson, T., Hajdas, I., Heaton, T., Hogg, A., Hughen, K., Kaiser, K., Kromer, B., McCormac, F., Manning, S., Reimer, R., Richards, D., Southon, J., Talamo, S., Turney, C., van der Plicht, J., and Weyhenmeyer, C.: IntCal09 and Marine09 Radiocarbon Age Calibration Curves, 0-50,000 Years cal BP, Radiocarbon, 51, https://journals.uair.arizona.edu/index.php/radiocarbon/article/view/3569, 2011.

Reimer, P., Austin, W. E., Bard, E., Bayliss, A., Blackwell, P. G., Bronk Ramsey, C., Butzin, M., Cheng, H., Edwards, R. L., Friedrich, M., et al.: The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0-55 kcal BP), Radiocarbon, 2020.

Reimer, P. J., Bard, E., Bayliss, A., Beck, J. W., Blackwell, P. G., Ramsey, C. B., Buck, C. E., Cheng, H., Edwards, R. L., Friedrich, M., et al.: IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP, Radiocarbon, 55, 1869–1887, 2013.

Routson, C. C., McKay, N. P., Kaufman, D. S., Erb, M. P., Goosse, H., Shuman, B. N., Rodysill, J. R., and Ault, T.: Mid-latitude net precipitation decreased with Arctic warming during the Holocene, Nature, 568, 83–87, 2019.

Roweis, S. T.: EM algorithms for PCA and SPCA, in: Advances in neural information processing systems, pp. 626–632, 1998.

Santos, G. M., Southon, J. R., Drenzek, N. J., Ziolkowski, L. A., Druffel, E. R., Xu, X., Zhang, D., Trumbore, S. E., Eglinton, T. I., and Hughen, K. A.: Blank assessment for ultra-small radiocarbon samples: chemical extraction and separation versus AMS, Radiocarbon, 52, 1322–1335, http://darchive.mblwhoilibrary.org:8080/handle/1912/4377, 2010.

Schulz, M. and Mudelsee, M.: REDFIT: estimating red-noise spectra directly from unevenly spaced paleoclimatic time series, Computers & Geosciences, 28, 421–426, https://doi.org/10.1016/S0098-3004(01)00044-9, http://www.sciencedirect.com/science/article/pii/S0098300401000449, 2002.

Shackleton, N. J.: New data on the evolution of Pliocene climatic variability, in: Paleoclimate and Evolution, with Emphasis on Human Origins, edited by Vrba, E. S., Denton, G. H., Partridge, T. C., and Burckle, L. H., pp. 242–248, Yale University Press, New Haven, CT., 1995.

Shakun, J. D., Clark, P. U., He, F., Marcott, S. A., Mix, A. C., Liu, Z., Otto-Bliesner, B., Schmittner, A., and Bard, E.: Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation, Nature, 484, 49–54, http://dx.doi.org/10.1038/nature10915, 2012.

Thomson, D. J.: Spectrum estimation and harmonic analysis, Proc. IEEE, 70(9), 1055–1096, 1982.

Tierney, J. E., Smerdon, J. E., Anchukaitis, K. J., and Seager, R.: Multidecadal variability in East African hydroclimate controlled by the Indian Ocean, Nature, 493, 389–392, http://www.nature.com/nature/journal/v493/n7432/abs/nature11785.html, 2013.

Trachsel, M. and Telford, R. J.: All age–depth models are wrong, but are getting better, The Holocene, 27, 860–869, https://doi.org/10.1177/0959683616675939, https://doi.org/10.1177/0959683616675939, 2017.

VanderPlas, J. T.: Understanding the Lomb–Scargle Periodogram, The Astrophysical Journal Supplement Series, 236, 16, https://doi.org/10.3847/1538-4365/aab766, http://stacks.iop.org/0067-0049/236/i=1/a=16, 2018.

Vaughan, S., Bailey, R. J., and Smith, D. G.: Detecting cycles in stratigraphic data: Spectral analysis in the presence of red noise, Paleoceanography, 26, https://doi.org/10.1029/2011PA002195, https://doi.org/10.1029/2011PA002195, 2011.

Ventura, V., Paciorek, C. J., and Risbey, J. S.: Controlling the Proportion of Falsely Rejected Hypotheses when Conducting Multiple Tests with Climatological Data, Journal of Climate, 17, 4343–4356, https://doi.org/10.1175/3199.1, 2004.

Wang, Y.-J., Cheng, H., Edwards, R. L., An, Z., Wu, J., Shen, C.-C., and Dorale, J. A.: A high-resolution absolute-dated late Pleistocene monsoon record from Hulu Cave, China, Science, 294, 2345–2348, 2001.

Werner, J. P. and Tingley, M. P.: Probabilistically constraining proxy age-depth models within a Bayesian hierarchical reconstruction model, Climate of the Past, 11, 533, 2015.

Werner, J. P., Divine, D. V., Charpentier Ljungqvist, F., Nilsen, T., and Francus, P.: Spatio-temporal variability of Arctic summer temperatures over the past 2 millennia, Climate of the Past, 14, 527–557, 2018.

Wickham, H.: ggplot2: Elegant Graphics for Data Analysis, Springer-Verlag New York, https://ggplot2.tidyverse.org, 2016.

Williams, J. W., Grimm, E. C., Blois, J. L., Charles, D. F., Davis, E. B., Goring, S. J., Graham, R. W., Smith, A. J., Anderson, M., Arroyo-Cabrales, J., Ashworth, A. C., Betancourt, J. L., Bills, B. W., Booth, R. K., Buckland, P. I., Curry, B. B., Giesecke, T., Jackson, S. T., Latorre, C., Nichols, J., Purdum, T., Roth, R. E., Stryker, M., and Takahara, H.: The Neotoma Paleoecology Database, a multiproxy, international, community-curated data resource, Quaternary Research, 89, 156–177, https://doi.org/10.1017/qua.2017.105, 2018.

Zhu, F., Emile-Geay, J., McKay, N. P., Hakim, G. J., Khider, D., Ault, T. R., Steig, E. J., Dee, S., and Kirchner, J. W.: Climate models can correctly simulate the continuum of global-average temperature variability, Proceedings of the National Academy of Sciences, 116, 8728, https://doi.org/10.1073/pnas.1809959116, http://www.pnas.org/content/116/18/8728.abstract, 2019.