Towards quantifying temporal representativity errors in paleoclimatology

Dan Amrhein

September 15, 2018

Abstract

Ongoing work in paleoclimate reconstruction and model-data comparison prioritizes understanding the origins and magnitudes of paleoclimate data errors. One class of such errors arises from assumptions of proxy temporal representativity – i.e., the time scales over which climate variables are equated with paleoclimate proxy measurements. This paper addresses errors that arise when a proxy estimate of a mean value over a "measurement duration" τ_y is used to represent climate conditions over a different "target duration" τ_x. Because it is challenging to tailor proxy measurements to precise time intervals, such errors are ubiquitous in model-data and data-data comparisons, but it is not always clear how important these errors are. Moreover, because values of τ_y are often not published alongside paleo data, amplitudes of these errors may be poorly constrained.

This paper shows how time-mean representativity errors depend on τ_x, τ_y, and the spectrum of the climate signal being sampled. In some cases, particularly for small values of τ_x relative to τ_y, errors can be large relative to paleoclimatological signals of interest. As climate signal spectra become more dominated by low frequencies, the fraction of error variance decreases in observations. Comparisons reveal that errors have magnitudes that are comparable to those expected from chronological errors. Processes that smooth climate signals in archives prior to sampling, such as bioturbation, can reduce the effects of aliasing but also reduce the information content of records by destroying high-frequency information. I also show how for a paleoclimate time series, absent any other smoothing effects in records, measurement intervals shorter than the spacing between samples lead to errors. Including these sources of uncertainty will improve accuracy in model-data comparisons and data comparisions and syntheses. Reporting sampling procedures alongside published data will facilitate quantifying uncertainty and improve model-data comparisons and syntheses.

1 Introduction

Paleoclimate records provide important information about the variability, extremes, and sensitivity of Earth's climate to greenhouse gases on time scales longer than the instrumental period. As the number of published paleoclimate records has grown and the sophistication of numerical model

representations of past climates has improved, it has become increasingly important to understand the uncertainty with which paleoclimate observations represent climate variables, so that they can be compared to one another and to model output. Additionally, quantifying uncertainty is important for ongoing efforts to assimilate paleoclimate data with numerical climate models (e.g., **?***Amrhein et al.*, 2015). Estimates of time means computed from paleoclimate records can have errors arising from many different sources: biological effects, aliasing onto seasonal cycles, spatial representativity, proxy-climate calibrations, and instrument errors, to name a few. This paper focuses on errors from temporal representativity, which we define as the degree to which a measurement averaging over one time interval can be used to represent a second, target time interval. Errors in temporal representativity can result both from from biases in representativity – e.g., systematically using data from one period to represent another – and from uncertainties in the duration and age of paleoclimate observations.

Paleoclimate records are generated by discretely sampling natural archives of climate variability along a growth or accumulation axis corresponding to the passage of time. Sampled material is typically then processed so as to average properties within samples, yielding measurements that are used to represent past climate states and their evolution. Choices of sampling procedures – for instance, how much material is taken and at what intervals along a growth or accumulation axis – are made based on the time scales of interest, the amount of material available for analysis, previous sampling of the archive, and limitations from analytical and human processing costs.

The sources of errors in paleoclimate records are legion and include biological effects, chronological errors, errors and nonstationarities in calibration, instrumental errors, and seasonal biases. This work focuses on temporal representational errors, which are defined as arising when a proxy quantity is used to represent a climate variable spanning a different time duration (for instance, when a decadal-average proxy is compared to a single year of model output). If the duration constrained by the proxy is unknown or poorly constrained, then computing these errors can be a challenge. Biases refer to sampling procedures that yield a different time-average than the target average. For instance, a duration bias would occur if a measurement averaging over 200 years was

used to represent the entire LGM (nominally 4,000 years long), and there would be an additional age bias if that measurement were centered on 19,400 years ago, rather than at 21,000 years ago (the midpoint of the LGM). When these biases are uncertain – due to uncertainties in radiocarbon dating, for instance, or because information about sampling procedures is unavailable – then additional errors are introduced.

Much of the previous study of temporal representativity error has focused on aliasing, whereby variability at one frequency appears at a different frequency in discrete samples of that process. *Pisias and Mix* (1988) described consequences of aliasing in the study of deterministic peaks in climate spectra due to Milankovich forcing. *Wunsch and Gunn* (2003) described criteria for choosing sample spacing so as not to alias low-frequency variability in sediment cores, and *Wunsch* (2000) demonstrated how aliasing can lead to spurious spectral peaks in ice core records. *Beer et al.* (2012) and *von Albedyll et al.* (2017) describe how running means can reduce aliasing of solar cycle variability in ice core records. *Kirchner* (2005) showed how aliasing can bias estimates of spectral slopes. In paleoclimate, measurements are often unevenly spaced in time due to changes in archive deposition rates; *Jones* (1972) showed that aliasing is present and even exacerbated in unevenly-sampled records relative to regularly sampled ones. Bioturbation and other diagenetic processes smooth records in time (*Anderson*, 2001; *McGee et al.*, 2013) and may act as anti-aliasing filters. The present work can also be thought of as a treatment of aliasing, specifically of various frequency bands set by sampling proedures onto the zero frequency (the time mean).

This paper uses an analytical statistical model to quantify time-mean representational errors and illustrate their dependence on signal spectra and sampling time scales. Extending this result from time-mean measurements to time series demonstrates how sampling practices can lead to aliasing errors when sampling averaging intervals are different from the time between measurements, i.e. when an ocean sediment core is not sampled continuously or densely along its accumulation axis. We do not claim that time representativity error is the most important source of uncertainty in all paleoclimate records. Nevertheless, the possibly important implications of discontinuous sampling for model-data synthesis motivate this effort to understand its effects.

Variable	Meaning
t	Time
ν	Frequency
ν_{Nyq}	Nyquist frequency
ν_{cut}	Cutoff frequency
$r(t)$	Time-varying climate process
$m(t,\tau)$	Time mean of $r(t)$ over period τ centered on t
x	Target paleoclimate quantity
y	Measured paleoclimate quantity
θ	Error in representing x by y
τ_x	Averaging timescale of target quantity
τ_y	Averaging timescale of observation
τ_a	Timescale of archive smoothing
τ_s	Time interval between samples in a series
τ_0	Time series length
Δ	Measurement time offset
$\Pi(t,\tau)$	Boxcar function in time
$G(\nu,\tau)$	Heaviside function in frequency
$H(\nu)$	Power transfer function
f	Error variance fraction of time-mean observation
β	Spectral slope (times -1)

Table 1: Glossary of functions and variables. Variables denoted by a superscript i in the text denote the i^{th} value of that quantity in a times series. Fourier transformed variables are denoted by a hat (e.g., $\hat{r}(\nu)$).

The rest of the paper is as follows. Section 2 describes the statistical model for time representativity errors in time mean values and time series. Section 3 illustrates the model by applying it to the analysis of Last Glacial Maximum and Pleistocene climate properties. Implications, caveats, and future research questions appear in the Discussion. Table (1) provides a glossary of functions and variables used.

2 A statistical model for temporal representativity errors

In paleoclimatology, a common focus is computing the mean of a climate variable (sea surface temperature, for instance, or isotope ratios, or ice volume) over a particular time period (for example, a marine isotope stage). Often computing a mean is the implicit goal of binning procedures that average observations from within a time period. This section defines an analytical approach

for estimating magnitudes of errors that arise in representing a time mean by a paleoclimate observation. These errors have a compact representation in the frequency domain that allows us to understand the relative importance and interaction of sampling procedures, time uncertainty, and signal spectra in contributing to errors.

This approach is intended to be complementary to the output from proxy system models (PSMs; e.g., *Evans et al.* (2013)) that relate proxy quantities to climate variables and do not consider temporal representativity. The starting point for the model is a hypothetical climate process, $r(t)$, which the model assumes to be able to sample directly. Temporal representativity errors computed here and additional errors that are inherited from the construction of $r(t)$ from proxy observations (e.g., from instrument errors) can be added together under the approximations that the errors are independent. If they are not, more complex forward modeling of errors may be necessary; see Dolman et al. for an example for sediment cores. By considering a distribution of possible Δ values, the model can also be extended to include effects from chronological uncertainty; a demonstration of this follows in Section 3.

2.1 Errors in time-mean values

Define a mean value of a climate variable $r(t)$ as a function of the duration τ and the time t on which that duration is centered,

$$m(t,\tau) = \int_{-\infty}^{\infty} \Pi\left(t',\tau\right) r\left(t+t'\right) dt', \tag{1}$$

where $\Pi(t,\tau)$ is the "boxcar" function centered on $t=0$ with width τ,

$$\Pi(t,\tau) = \begin{cases} 1/\tau & |t| \leq \tau/2 \\ 0 & |t| > \tau/2 \end{cases} \tag{2}$$

normalized by $1/\tau$ so that (2) gives a time average. The operation in (1) defines a moving average $m(t,\tau)$ of $x(t)$ with averaging length τ and is known as a convolution, hereafter denoted as a star,

$$m(t,\tau) = \Pi(t,\tau) \star r(t). \tag{3}$$

Our focus is on errors that arise when a mean value computed over one time period is used to represent another time period – for instance, when a time average of over 23-19 ka (thousand years ago, the nominal timing of the Last Glacial Maximum, Clark et al.) is represented by an average over 20-19 ka. To write this representation generally, say that a mean x of $r(t)$ over an interval of length τ_x centered on t is represented by an observation y that averages over a different duration τ_y centered on a different time $t+\Delta$,

$$x = m(t,\tau_x) \tag{4}$$

$$y = m(t+\Delta,\tau_y). \tag{5}$$

There are many cases in which where a paleoclimate archive becomes smoothed prior to processing, whether by bioturbation, diagenesis, residence times in karst systems downstream of caves, or other effects. These processes can be quite complex and non-constant in time; here, to gain an order-of-magnitude understanding of their effects, we assume an archive smoothing process that is a moving average over a time scale τ_a. Under such smoothing, we can then write y as a twice-smoothed function of $r(t)$,

$$y = \Pi(t,\tau_a) \star \Pi(t,\tau_y) \star r(t)$$

The error in representing x by y is

$$\theta = x - y \tag{6}$$

We will describe typical values for θ by estimating its variance, $\left\langle (\theta - \langle\theta\rangle)^2 \right\rangle$, where the angle brackets denote statistical expectation (the average value over the duration of $r(t)$). Our approach is to estimate time-mean errors by creating two time series of running time-mean values and calculating the typical deviation between them. To do this, we first make the approximation that $r(t)$ is weakly statistically stationary, meaning that its mean and variance do not change in time. (Some caveats surrounding this assumption are addressed in the Discussion.) Under the weak stationarity assumption, the mean error $\langle\theta\rangle$ is zero, and we can take the expectation by evaluating θ^2 at all the times in $r(t)$ to compute the variance,

$$\left\langle \theta^2 \right\rangle = \frac{1}{\tau_0} \int_{t_0}^{t_f} (x-y)^2 dt, \tag{7}$$

where t_0 and t_f are the starting and ending time of $r(t)$, and $\tau_0 = t_f - t_0$ is the duration of $r(t)$. Intuitively, we are estimating the error in representing s by y by computing the average squared misfit between similar pairs of offset mean values computed over the entire paleoclimate record. In practice, we do not *know* the entire paleoclimate record, but if we can estimate some of its statistics, that can be adequate to estimate $\left\langle \theta^2 \right\rangle$. Define $x' = x - \langle x \rangle$ and $y' = y - \langle y \rangle$, where $\langle x \rangle = \langle y \rangle$ for stationary $r(t)$. Then expanding the squared quantity in Equation (7) gives an expression in terms of the estimated variances $\tilde{\sigma}_x^2$ and $\tilde{\sigma}_y^2$ of x and y, respectively, and the estimated cross-covariance of x and y as a function of lag Δ, $\tilde{C}_{xy}(\Delta)$:

$$\left\langle \theta^2 \right\rangle = \frac{1}{\tau_0} \int_{t_0}^{t_f} \left(x' - y'\right)^2 dt \tag{8}$$

$$= \frac{1}{\tau_0} \left(\int_{t_0}^{t_f} x'^2 dt + \int_{t_0}^{t_f} y'^2 dt - 2 \int_{t_0}^{t_f} x' y' \, dt \right) \tag{9}$$

$$= \tilde{\sigma}_x^2 + \tilde{\sigma}_y^2 - 2\tilde{C}_{xy}(\Delta). \tag{10}$$

This expression can be usefully intepreted to reveal how errors arise in representing x by y. In the limit where $\tau_x = \tau_y$ and $\Delta = 0$, $\tilde{\sigma}_x^2 = \tilde{\sigma}_y^2 = \tilde{C}_{xy}(\Delta)$, and the estimated representativity error is zero, as we expect for the case where the measurement exactly targets the quantity of interest. If

$\tilde{C}_{xy}(\Delta) = 0$ – e.g., if Δ is so large that the measurement and target quantity are uncorrelated – then y has no skill in representing x, and the error variance is the sum of $\tilde{\sigma}_x^2$ and $\tilde{\sigma}_y^2$. At the far extreme, if there is a choice of Δ such that x and y are anticorrelated – as might happen if $r(t)$ is dominated by one or more periodic signals – then $\tilde{C}_{xy}(\Delta)$ will be negative, leading to even larger errors. Between these extremes of zero and maximum error, intermediate values of $\left\langle \theta^2 \right\rangle$ are set by timescales of sampling procedures and of the variability in $r(t)$. An understanding of these relationships arises from representing the error in the frequency domain.

2.2 Analyzing sources of error in the frequency domain

Considering errors in the frequency domain allows us to analyze the dependence of $\left\langle \theta^2 \right\rangle$ on timescales arising from 1) the underlying climate signal, $r(t)$, and 2) paleoclimate archive sampling and preservation. Using Parseval's theorem, the Fourier shift theorem, and the convolution theorem (Appendix A), denoting frequency by ν, and denoting the Fourier transform by a hat, we can rewrite (7) in the frequency domain as

$$\left\langle \theta^2 \right\rangle = \int_0^\infty \left| \hat{\Pi}(\nu, \tau_x) - e^{-2\pi i \nu \Delta} \cdot \hat{\Pi}(\nu, \tau_a) \cdot \hat{\Pi}(\nu, \tau_y) \right|^2 |\hat{r}(\nu)|^2 \, d\nu. \tag{11}$$

The second component in (11), $|\hat{r}(\nu)|^2$, is the squared magnitude of the Fourier transform of $r(t)$, and is an estimate of the power spectral density of $r(t)$. The power spectral density describes the variance contained at the frequencies in $r(t)$. The first component is a so-called power transfer function,

$$H_{\tau_s, \tau_d, \tau_y, \Delta}(\nu) = \left| \hat{\Pi}(\nu, \tau_s) - e^{-2\pi i \nu \Delta} \cdot \hat{\Pi}(\nu, \tau_a) \cdot \hat{\Pi}(\nu, \tau_y) \right|^2, \tag{12}$$

which describes how the different frequencies are weighted in the integral as a function of target averaging time scale τ_x, the data averaging time scale τ_y, the time scale of archive smoothing τ_a, and the time Δ of measurement offset. Thus, the variance of the representivity error is a weighted sum of the variance at different frequency components of $r(t)$ – that is, some of the variability in

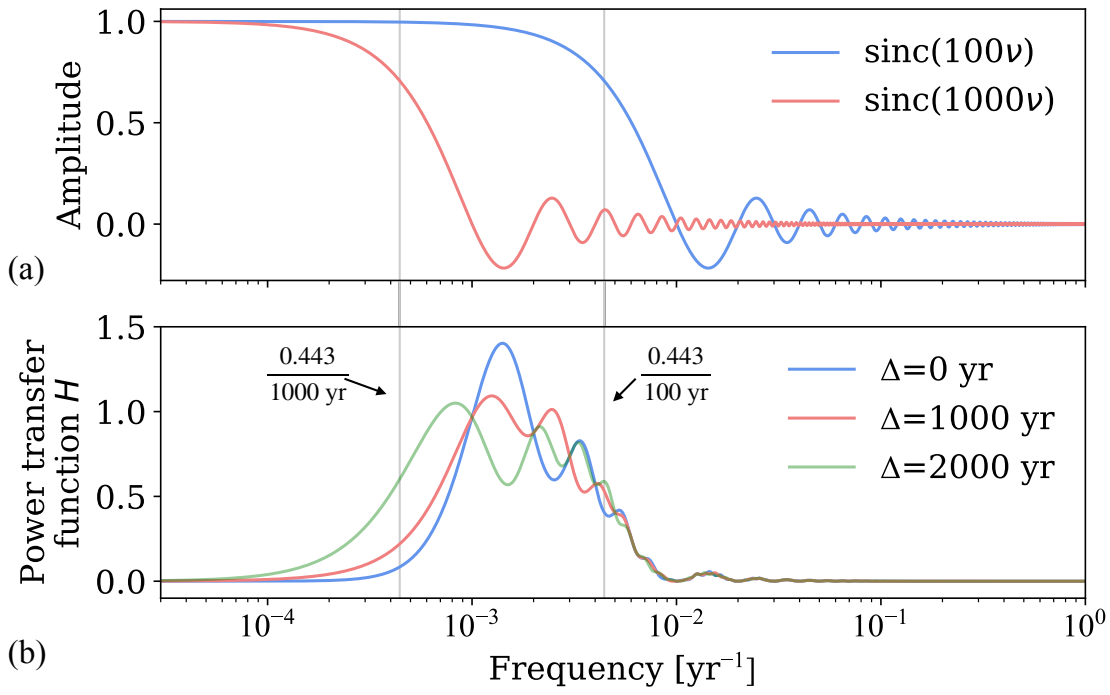

Figure 1: Variability at different frequencies in a climate signal contributes to errors in time-mean representativity errors depending on time scales. Contributions to the power transfer function H () illustrate the dependence of temporal representativity errors on different sampling time scales. Functions are plotted on a logarithmic horizontal axis and are illustrated using $\tau_x = 1000$, $\tau_y = 100$, $\tau_a = 500$, and several values of Δ. The Fourier transform of the boxcar function $\Pi(t,\tau)$ (2) is the sinc function ((13), panel a), which oscillates about 0 with the first zero crossing at $\nu = 1/\tau$. (d) When $\Delta = 0$, temporal representativity errors originate from periods lying between τ_s and $\min(\tau_d, \tau_y)$. For nonzero Δ, errors also originate from frequencies as low as $1/(2\Delta)$. These errors can be up to four times the variance of the underlying signal at some frequencies (Appendix A).

$r(t)$ is erroneously aliased onto the estimate y.

We can gain understanding of how different time scales contribute to $\langle \theta^2 \rangle$ by analyzing the functional form of H. First, note that the Fourier transform of the boxcar function is a sinc function,

$$\hat{\Pi}(\nu,\tau) = \mathrm{sinc}(\tau\nu) = \frac{\sin(\pi\tau\nu)}{\pi\tau\nu}, \tag{13}$$

which converges to 1 at low frequencies and decreases to values oscillating about 0 at higher

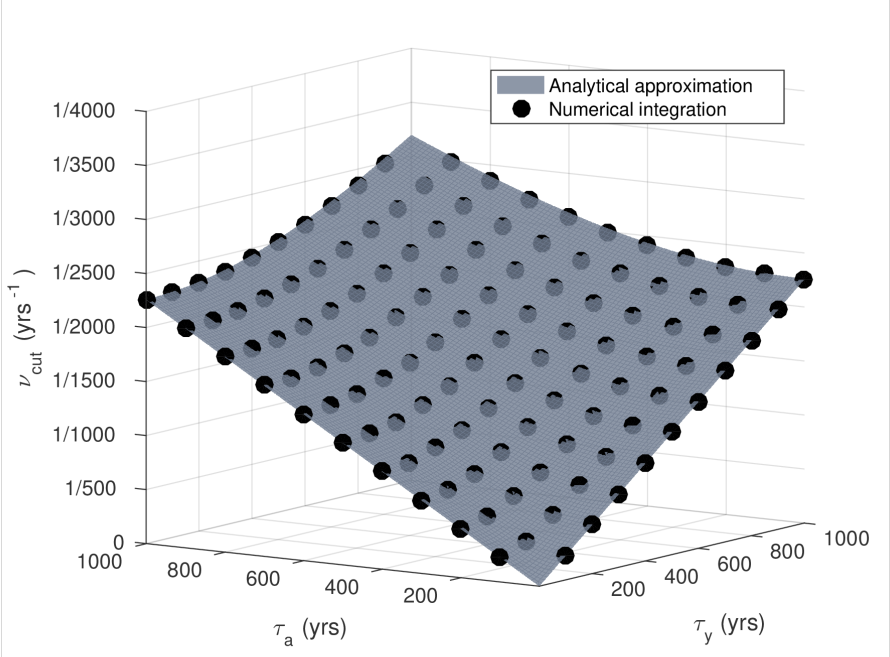

Figure 2: The cutoff frequency ν_{cut} denotes the frequency at which a filter attenuates signal power by 50% and is a measure of which frequencies are filtered by a proxy archive sampling procedure. The combined filtering effects of uniformly sampling a paleoclimate archive over an interval τ_y and archive smoothing over a length scale τ_a have a combined effect that can be observed by numerically solving for ν_{cut} in $(\mathrm{sinc}(\tau_a \nu_{cut})\,\mathrm{sinc}(\tau_y \nu_{cut}))^2 = 0.5$ and numerically integrating (dots). This solution is well-approximated by Equation (14) (grey manifold). The combined filter effects are most pronounced when τ_a and τ_y are similar; in this regime, temporal representativity errors can arise because sampling procedures over-smooth time series.

frequencies, with the first zero crossing at $1/\tau$ (Figure 1a). A common measure of filter behavior is the cutoff frequency, ν_{cut}, which is a diagnostic of the boundary between the passband and stopband of a filter. We define ν_{cut} as the frequency at which a power transfer function is 0.5, so that half the power is attenuated at that frequency. For instance, for a record filtered by a moving average of length τ, the power transfer function is $\mathrm{sinc}^2(\tau\nu)$; setting this quantity equal to 0.5 and solving for ν_{cut} yields $\nu_{cut} = 0.443\tau^{-1}$. If the paleoclimate archive was smoothed prior to sampling, then the power transfer function is $(\mathrm{sinc}(\tau_a\nu)\,\mathrm{sinc}(\tau_y\nu))^2$; setting this expression equal to 0.5 and approximating as a truncated Taylor series gives the approximate formula (Figure 2)

$$\nu_{cut} = \frac{0.443}{\sqrt{\tau_a^2 + \tau_y^2}}. \tag{14}$$

An implication is that the filtering effects of sampling and archive smoothing can be additive. In particular, choosing a sampling interval τ_y equal to τ_x – which would be idea in the absence of archive smoothing – for $\tau_a > 0$ will over-smooth a record and lead to errors because the observed quantity averages over a longer interval than the target. Finally, when there is a time offset, then nonzero values in H extend to frequencies as low as $1/2\Delta$, with sinusoidal variations in frequency contributions set by phase relationships at time scales set by Δ (Figure 1d; Appendix A). It is possible to compute additional contributions to uncertainty arising from a distribution of possible offsets, as might arise from chronological uncertainty; this approach is demonstrated in Section 3.

2.3 Implications for time-mean representational errors

Given estimates of the sampling interval, archive smoothing time scale, measurement offset, and an estimate of the signal spectrum, (11) is a closed-form expression for estimating representational errors in an estimate over a time interval τ_x. This equation can be useful for calculating errors useful for model-data and data-data comparisons and for weighting observations in the construction of temporal bin averages. It also provides a basis for some intuitive conclusions about temporal

representational errors and their implications for record sampling and uncertainty quantification:

1. Temporal representativity errors can be traced to frequency bands in the underlying signal $r(t)$ that are inadequately filtered by sampling procedures. In the limit where there is no archive smoothing and no time offset, the error in representing a mean of duration τ_x by another is roughly equal to the variance at the frequencies between $0.443\tau_x^{-1}$ and $0.443\tau_y^{-1}$. Thus, if a centennial mean is used to represent a millennial mean, to first order the expected error variance is equal to the variance in $r(t)$ at time scales between 226 and 2260 years (Figure 1). The error is the same if a centennial mean is used to represent a decadal mean.

2. The effects of archive smoothing and sampling are additive and can lead to oversmoothing. In the presence of archive smoothing, but without a time offset, the error in representing one mean by another is roughly equal to the variance at the frequencies between $0.443\tau_x^{-1}$ and $0.443(\tau_y^2+\tau_a^2)^{-\frac{1}{2}}$. Thus if a millennial mean is computed from a record that has been smoothed at millennial time scales, then the error in representing a target millennial mean is equal to the variance in $r(t)$ at time scales between 3192 and 2260 years. On the other hand, if one of τ_y and τ_a is much larger than the other, then the larger time scale will dominate; for instance, if a millennial mean is computed from a record smoothed at centennial scales, then errors arise from between 2260 and 2268 year periods, a relatively small window. Given knowledge of the smoothing time scales of a record, one can compute the ideal length $\tilde{\tau}_y$ of a sample that will minimize error in an estimate of a time mean over τ_x as $\tilde{\tau}_y = \sqrt{\tau_x^2 - \tau_a^2}$; whether this expression is useful given the complexity of archive smoothing processes will depend on the application.

3. When there is a time offset in the measurement relative to the target, additional errors result. These errors are generally small when Δ is less than τ_x or τ_y. When offsets are greater than τ_x and τ_y, measurement offsets alias variability from frequencies higher than $1/2\Delta$ onto errors.

Among the caveats for these conclusions is that the cutoff frequency is only an approximate measure of the shape of a filter. Section 3 explores these relationships further in the context of mean estimation at the Last Glacial Maximum.

2.4 Implications for time series analysis

Paleoclimate time series are sequences of time-mean values. Just as sampling, archive smoothing, and time offset errors can introduce errors in estimates of time mean properties, so too do they introduce errors in time series. Quantifying these errors can be complicated in part because because paleoclimate records are often unevenly sampled in time, owing to changes in chronologies and sampling procedures along a sediment archive. Here we adopt the machinery of the previous section to address a relatively limited question: what is the time mean that should be represented by a measurement in a time series in order to provide the most accurate discrete representation of a continuous climate process? We show that for locally constant sample spacing in the absence of archive smoothing, dense sampling (i.e., setting the averaging interval equal to the spacing between measurements) is a nearly optimal approach to minimize aliasing errors. Once again, representation in the frequency domain proves advantageous.

The sampling theorem of *Shannon* (1949) states that sampling $r(t)$ instantaneously (that is, with a very short averaging interval) at a fixed time interval τ_s unambiguously preserves signal information only when $r(t)$ does not contain any spectral power at frequencies greater than $1/2\tau_s$ (called the "Nyquist" frequency, ν_{Nyq}). When this criterion is not met, the discrete signal is corrupted by aliasing, whereby variability in $r(t)$ at frequencies greater than ν_{Nyq} appears artificially at lower frequencies in the discrete signal. To mitigate aliasing, one can either increase the sampling rate or apply a low-pass "anti-aliasing" filter to the continuous signal to attenuate power at frequencies higher than ν_{Nyq}. Sampling (non-instantaneous) time-mean values of a climate process $r(t)$ is equivalent to computing a moving average over $r(t)$; this moving average serves as an anti-aliasing filter, and can serve to reduce aliasing errors.

Here we will use Shannon's theorem to extend the results from the previous section to provide an expression for temporal representativity errors for individual time series measurements. To generalize to the case where records are unevenly spaced in time, we make the assumption that the sampling interval τ_s^i is locally constant: that is, for the i^{th} measurement y^i taken at time t^i, y^{i-1} was taken at time $t^i - \tau_s^i$, and y^{i+1} was taken at time $t^i + \tau_s^i$. If the sampling interval changes rapidly, conclusions from this approach might not apply. Define the moving average time series associated with y^i to be

$$y^i(t) = \Pi\left(t, \tau_y^i\right) \star \Pi\left(t, \tau_a^i\right) \star r(t) \tag{15}$$

where we have included a contribution from archive smoothing, so that its Fourier transform is

$$\hat{y}^i(\nu) = \hat{\Pi}\left(\nu, \tau_y^i\right) \cdot \hat{\Pi}\left(\nu, \tau_a^i\right) \cdot \hat{r}(\nu). \tag{16}$$

Shannon's theorem states that an accurate discrete representation of $r(t)$ results from sampling all frequencies in $r(t)$ less than or equal to the local Nyquist frequency $\nu_{Nyq}^i = 1/\left(2\tau_s^i\right)$. As such, the target value x^i for the i^{th} measurement y^i is the value of $r(t)$ sampled at t^i after filtering $r(t)$ to remove high-frequency variability. Define this filtered version of $r(t)$ time series to be the series of values $x^i(t)$. The Fourier transform of such a time series is

$$\hat{x}^i(\nu) = G\left(\nu, \tau_s^i\right) \hat{r}(\nu) \tag{17}$$

where the "ideal" transfer function $G(\nu, \tau_s)$ is the piecewise constant Heaviside function

$$G(\nu, \tau_s) = \begin{cases} 1 & \nu < 1/(2\tau_s^i) \\ 0 & \nu \geq 1/\left(2\tau_s^i\right). \end{cases} \tag{18}$$

that is "ideal" in the sense that it eliminates variability at frequencies greater than $\nu_{Nyq}^i = 1/\left(2\tau_s^i\right)$.

Then we can define the temporal representativity at the i^{th} measurement to be

$$\theta^i = x^i - y^i. \tag{19}$$

As in the previous section, we estimate the variance of θ^i by taking the expected value as if the entire record had been sampled using the local values τ_s^i and τ_y^i. Then the equivalent to (11) expressing temporal representativity error for a single measurement in a time series is

$$\left\langle \theta^{i2} \right\rangle = \int_0^\infty \left| G\left(\nu, \tau_s^i\right) - \hat{\Pi}\left(\nu, \tau_a^i\right) \cdot \hat{\Pi}\left(\nu, \tau_y^i\right) \right|^2 \left|\hat{r}\left(\nu\right)\right|^2 d\nu. \tag{20}$$

This expression of error can be useful for estimating errors and provides intution for the sources thereof:

1. Similar to 11, the error variance is a weighted integral over the power density spectrum of $r(t)$, where weights are largest approximately between ν_{cut} (Equation 14, set by sampling and archive smoothing time scales) and ν_{Nyq}^i set by the local sampling interval τ_s^i.

2. Unlike in the mean estimation case, where the power transfer function can be equal to zero, some degree of representational error is unavoidable with uniform sampling because of differences between the shape of $\mathrm{sinc}\left(\tau_y^i \nu\right)$, and the abrupt frequency cutoff specified by $G\left(\nu, \tau_x^i\right)$. Sampling a paleoclimate archive nonuniformly in time could better approximate the ideal filter and may reduce errors, but this may not be practical given the many other sources of error in paleoclimate records.

3. A scaling for the ideal sampling duration $\tilde{\tau}_y^i$ given a sample spacing τ_s^i can be found by setting $\nu_{cut}^i = \nu_{Nyq}^i$, so that to first order, the sampling procedure attenuates frequencies $\nu > \nu_{Nyq}^i$ and attenuates $\nu < \nu_{Nyq}^i$. Using (14) and doing some algebra, we obtain $\tilde{\tau}_y^{i2} = \left(0.79\tau_s^{i2} - \tau_a^{i2}\right)$; in the absence of archive smoothing, this yields $\tilde{\tau}_y^i = 0.89\tau_s^i$. Thus by this metric, sampling procedures that are "dense" or "continuous" – i.e., where $\tau_y^i = \tau_s^i$ so that no space is left

between samples – reasonably approximate the ideal anti-aliasing filter. As we will see in Section 3, the details depends on the spectrum of $r(t)$, but this analysis provides a rationale for this rule of thumb of paleoproxy sampling.

These results hold for time series whose spacing and chronologies are not changing too rapidly and where the goal is to obtain a discrete representation of a continuous process. For other objectives, such as sampling the spectrum of a signal, other approaches may apply. For instance, "burst sampling," whereby rapid observations are taken at relatively long intervals, is used in modern oceanographic procedures to estimate spectral nonstationarities in time (*Emery and Thomson*, 2014), and unevenly spaced paleoclimate observations can be leveraged to give a range of frequency information using variogram approaches (**?***Amrhein et al.*, 2015) or the Lomb-Scargle periodogram (e.g., *Schulz and Stattegger*, 1997). Even in these more flexible frameworks, however, there is a danger of aliasing, and one must carefully consider any mismatches between the frequencies sampled by an observation, and the frequencies that observation is being used to constrain.

3 Applications to the Last Glacial Maximum and Last Termination

3.1 Errors in mean values at the Last Glacial Maximum

This section illustrates dependencies of errors on time scales and signal spectra in Equation (11). We will do so in the context of the Last Glacial Maximum (LGM), the period associated with the greatest land ice extent during the last glacial period, which is typically taken to be an interval of several thousand years approximately 20,000 years ago, e.g. 23-19 ka (*MARGO Project Members*, 2009). Denote the time-mean value of a climate quantity $x(t)$ during the LGM as m_{LGM},

$$m_{\mathrm{LGM}} = \frac{1}{4000} \int_{23,000}^{19,000} x(t)\, dt.$$

To illustrate errors arising from representational errors, consider the case of a proxy measurement sampled to estimate a 1000-year time-mean value of $x(t)$ centered on 21 ka,

$$m_{21} = \frac{1}{4000} \int_{20,500}^{21,500} x(t)\, dt.$$

Given the sparsity of observations, the proxy estimate of m_{21} – i.e., dated to within the LGM, but averaging over only a subset – could reasonably be included in a compilation of LGM data. In order to quantify errors resulting from this, denote the difference as θ,

$$\theta = m_{\mathrm{LGM}} - m_{21}.$$

Taking the Fourier transform of the variance of θ and applying Parseval's Theorem (Appendix A) reveals that errors arise from aliasing, whereby temporal variability at some (not all) frequencies in $x(t)$ appears erroneously in the mean estimate of m_{21}. Indeed, the variance σ_θ^2 can be expressed as a weighted integral of the variance at different frequencies in $x(t)$. Let $P(\nu)$ be the power spectral density estimate of $x(t)$, which describes the variance in $x(t)$ as a function of frequency. Then (Appendix A),

$$\sigma_\theta^2 = \int_{-\infty}^{\infty} W(\nu, \tau_b \tau_s)\, P(\nu)\, d\nu,$$

where the weights $W(\nu, \tau_b \tau_s)$ are a function of frequency, the measurement interval (τ_s, here 1000 years), and the target duration (4000 years, as we defined for the LGM). These weights (Figure 1) have an intuitive interpretation: the error variance is (approximately) equal to the variance in $x(t)$ between 1000 and 4000 years. That is because taking a 1000-year sample effectively damps variability below 1000 years, but longer variability can still affect 1000-year averages within a 4000-year window.

The question is relevant because if we compare m_{21} to an LGM-mean estimate of $x(t)$ from a model, we might erroneously conclude that the model did not fit the data, because of errors in approximating m_{LGM} by m_{21}.

To assess the effects of these errors on estimates of paleoclimate quantities at the LGM, Figure X shows the fraction of error in the LGM-modern signal expected for different types of signal spectra.

Understanding the dependence of sampling error on τ_s, τ_b, and the signal spectrum $|\hat{x}(\nu)|^2$ is useful for quantifying errors and guiding sampling practices. To assess the effects of qualitatively different climate spectra, we will assume a power-law spectrum for $x(t)$ having the form

$$|\hat{x}(\nu)|^2 \propto \nu^{-\beta}, \tag{21}$$

where β is the spectral slope (when plotted in log-log space, $\nu^{-\beta}$ is a straight line with slope $-\beta$). Such spectra are common in climate and other systems (**?***Huybers and Curry*, 2006). Here we use as examples $\beta = 0.5$ and $\beta = 1.5$, motivated by *Huybers and Curry* (2006), who fit paleoclimate records to spectral slopes between $\beta = 0.3$ and $\beta = 1.6$. Climatological spectral features that are not described by power laws, such as peaks due to the annual cycle or Milankovich variability, can also contribute to aliasing (*Pisias and Mix*, 1988; *Wunsch*, 2000) and are not considered here.

To analyze errors, we define a sampling error fraction f to be the ratio of the error variance to the variance of the sample,

$$f = \frac{\left\langle \theta^2 \right\rangle}{\sigma_y^2} \tag{22}$$

where the numerator. (While this quantity is useful for scaling uncertainties, the denominator is the quantity of greatest interest for model-data and data-data comparisons). When performing the integrations in frequency space, a lower bound $1/\tau_0$ is required (Appendix B). For time series, τ_0 is the time series length; for problems of estimating time mean anomalies relative to the present, τ_0 is the years before present of the center of the target time interval. Here we use $\tau_0 = 21,000$ years as an estimate of the time since the LGM.

Error variance ratios in the $\beta = 1.5$ case are lower by a factor of two or more; for spectra with relatively more power at low frequencies, power in the frequency bands contributing to aliasing is relatively smaller, and errors from aliasing are less important, as also noted by *Wunsch* (1978) and

330 **?**.

331 after smoothing the record with a boxcar function with length $\tau_d = 400$ years. For τ_b greater
332 than roughly 800 years, the dependence of sampling error on τ_s and τ_b is similar to the case
333 without diagenetic smoothing. For $\tau_b < 800$ years and $\tau_s > \tau_d$, sampling errors are reduced, while
334 for $\tau_b < 800$ years and $\tau_s < \tau_d$, sampling errors increase. Notably, discontinuous sampling is
335 modestly preferable to dense sampling when τ_b is similar to τ_d and when $\tau_s < \tau_d$. Evidently when
336 the diagenetic length scale approximately equals the sampling interval, then diagenesis serves as
337 an adequate anti-aliasing filter. In this case, dense sampling is not necessary, and is even slightly
338 undesirable, as it smooths the record further and attenuates sub-Nyquist variability. In the examples
339 shown, this over-smoothing effect is small, and E_d for dense sampling differs from that for the
340 optimal strategy by a few percent. Diagenetic effects became small at roughly $\tau_b \approx 2\tau_d$, coincident
341 with the first zero crossing in $\mathrm{sinc}(\tau_b \nu)$; at that point, the additional smoothing from diagenesis
342 is relatively small, and $\mathrm{sinc}(\tau_d \nu)\,\mathrm{sinc}(\tau_b \nu)$ approaches $\mathrm{sinc}(\tau_b \nu)$ as τ_d decreases so that (**??**)
343 converges to (34).

344 Finally, we compute the same values for the time series case explored in Section X. The value
345 of τ_b that minimizes f in (34) is larger than τ_s by approximately 10% (bold lines, top panels,
346 Figure 3), meaning that slightly overlapping samples optimally reduce sampling error, though
347 dense sampling ($\tau_b = \tau_s$) is only modestly worse. In the range between $\tau_s = 10$ years and $\tau_s = 2000$
348 years, minimum ratios of error variance to sample variance are approximately 0.1 for spectral slope
349 of $\beta = 0.5$ and 0.05 for $\beta = 1.5$; there is no choice of τ_b that reduces sampling error to zero. Away
350 from minimum values, larger values of τ_b increase error monotonically because they attenuate
351 variability at sub-Nyquist frequencies, while smaller values of τ_b increase error because there is
352 less filtering of higher-frequency variability and thus increased aliasing.

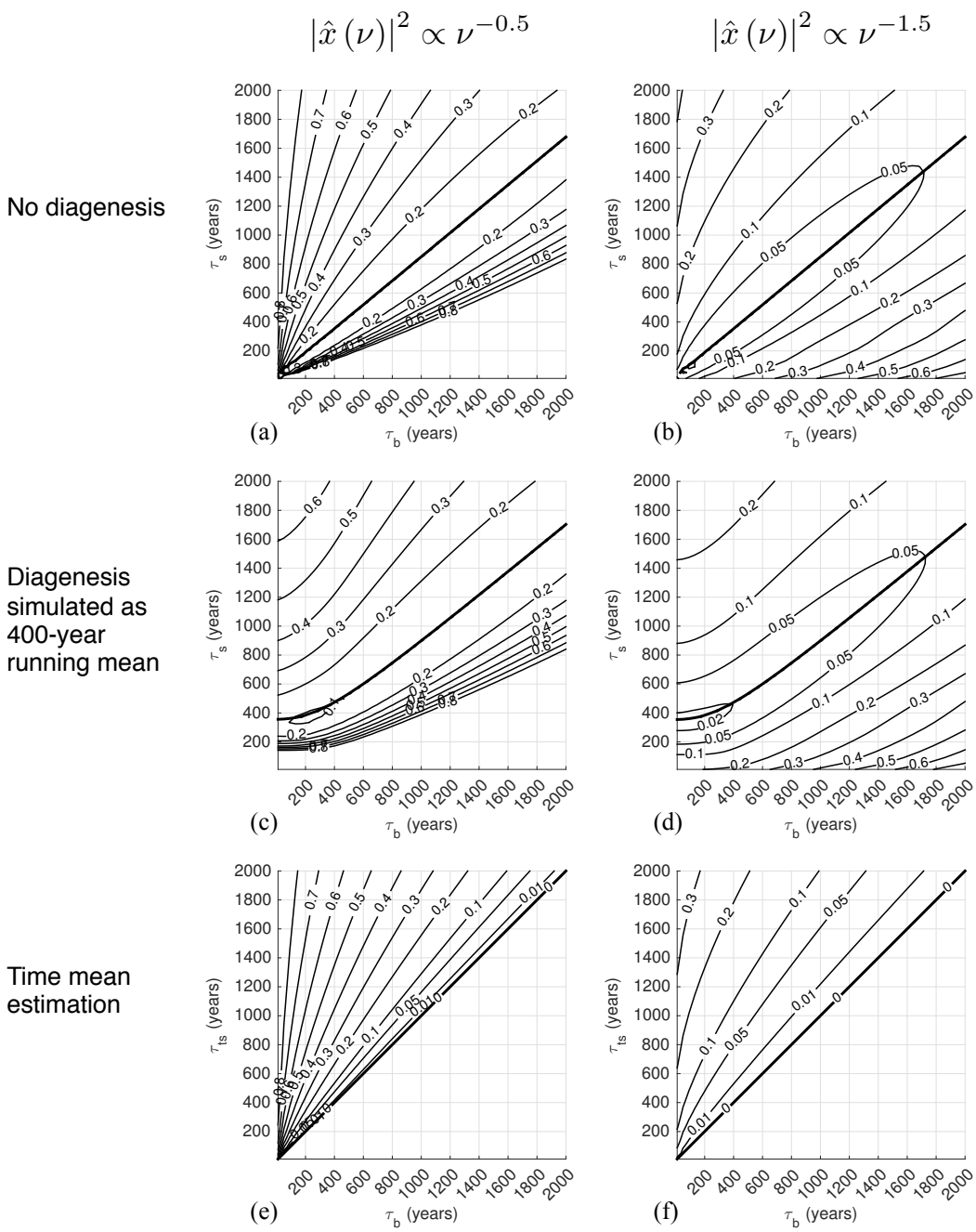

Figure 3: Sampling error variance ratios for the three cases discussed in the main text computed for power law spectra $|\hat{x}(\nu)|^2 \propto \nu^{-0.5}$ and $|\hat{x}(\nu)|^2 \propto \nu^{-1.5}$. **Top:** The ratio E (34) corresponding to the case with no diagenetic smoothing. The bold line indicates the optimal choice of the bin averaging period τ_b for each value of the sampling period τ_s. The record length is $\tau_0 = 20,000$ years. **Middle:** Same as top row but for the ratio E_d (**??**) where the record $x(t)$ was first smoothed with a moving average of length $\tau_d = 400$ years. **Bottom:** Same as top row but for the ratio E_{ts} (**??**) showing error in representing the value $m_{ts}(t)$ (**??**) of $x(t)$ averaged over an interval τ_{ts} by a measurement $m_b(t)$ averaged over a shorter interval τ_b. This quantity is zero by definition for $\tau_b = \tau_{ts}$.

3.2 Effects from chronological uncertainty

When the dating of a measurement is uncertain, the offset Δt between the target quantity s and observation y is not exactly known, and one may instead have a probability distribution function of possible offsets, $\mathbf{P}(\Delta t)$. Here we compute the first moment of error arising from chronological uncertainty. To denote that the expectation has been taken with respect to chronological uncertainty, we have denoted sampling uncertainty by an overbar.

3.3 Errors in time series of the last deglaciation

To illustrate errors from sampling procedures, we subsampled the North Greenland Ice Core Project (NGRIP; **?**) 50-year average oxygen isotope ratio time series (white line, Figure 4a) and Cariaco Basin upper 10 meter annual-average ocean temperature in the TraCE-21ka coupled model simulation of the last deglaciation (*Liu et al.* (2009); white line, Figure 4b) at lower time resolution. These time series were chosen as representative examples of the variability recorded by paleoclimate archives; they originate from different sources and represent different regions and climate variables. Both were subsampled at a sampling period of $\tau_s = 1000$ years using bin averaging periods of $\tau_b = 100$ years (discontinuous sampling) and $\tau_b = 1000$ years (dense sampling). In the subsampled NGRIP record, the discontinuously sampled record (red line) shows greater excursions than the continuously sampled record (black line) during millennial and multi-millennial variability and a different sign of change at 26 ka (Figure 4a). A comparison of spectra reveals that while the densely sampled spectrum closely resembles that of the original signal, the discontinuously sampled record has additional, spurious high-frequency variance (Figure 4c) due to aliasing. The subsampled TraCE-21ka time series show modest disagreements except for changes in the duration and timing of maximum Bølling-Allerød warming (~14.5 ka; Figure 4b). Again, signal spectra show aliased variability in the discontinuously sampled record. In both cases, aliased variability is largest at the highest frequencies in the subsampled records (near ν_{Nyq}) because aliased

variability at a frequency $\nu_0 > \nu_{Nyq}$ appears, spuriously, at a lower frequency ν_a satisfying

$$|\nu_a| = \left|\nu_0 \pm 2m\nu_{Nyq}\right| < \nu_{Nyq} \tag{23}$$

where m is an integer. For instance, given $\nu_{Nyq} = 1/2000$ yr $^{-1}$, variability at $1/1500$ yr $^{-1}$ would appear at $|1/1500 - 1/2000| = 1/3000$ yr $^{-1}$. The effect is that aliased power is reflected across ν_{Nyq}, so that power at a frequency slightly higher than ν_{Nyq} will appear slightly below it in the subsampled record, and power at a frequency much higher than ν_{Nyq} will appear at lower frequencies. Because these two signals (like most climate signals; *Hasselmann* (1976)) have more power at lower frequencies, aliased variability is greatest at the highest frequencies, which tends in turn to whiten (flatten) observed spectra (*Kirchner*, 2005).

4 Discussion

This paper illustrates and quantifies sampling errors associated with discontinuous sampling in paleoclimate records. Discontinuous sampling leads to aliasing errors with amplitudes determined by the sampling interval τ_s, the sample averaging bin width τ_b, and the record spectrum. In the absence of bioturbation or other diagenetic smoothing, choosing $\tau_b = \tau_s$ approximately minimizes sampling error. When τ_b is small relative to τ_s, sampling errors can be large relative to the signal variance, on the order of tens of percents. In addition, the use of a measurement with bin averaging period τ_b shorter than a time slice length τ_{ts} – for instance, at the Last Glacial Maximum, ca. 23-19 kyr ago (e.g., *MARGO Project Members*, 2009; *Caley et al.*, 2014) – can introduce errors that may explain some of the disagreement among proxy measurements within time-slice periods.

Sampling errors have the potential to be substantial for spectra and sampling procedures relevant to paleoclimatology. The errors shown in Figure 1 could arise from sampling a sediment core using bin average lengths Δ_d =1cm every sample length Δ_s =10cm for sedimentation rates of 0.1 kyr/cm, typical for high sedimentation rate regions such as the Bermuda Rise or Feni Drift. These errors have implications for model-data comparisons and paleoclimate data assimilation. For in-

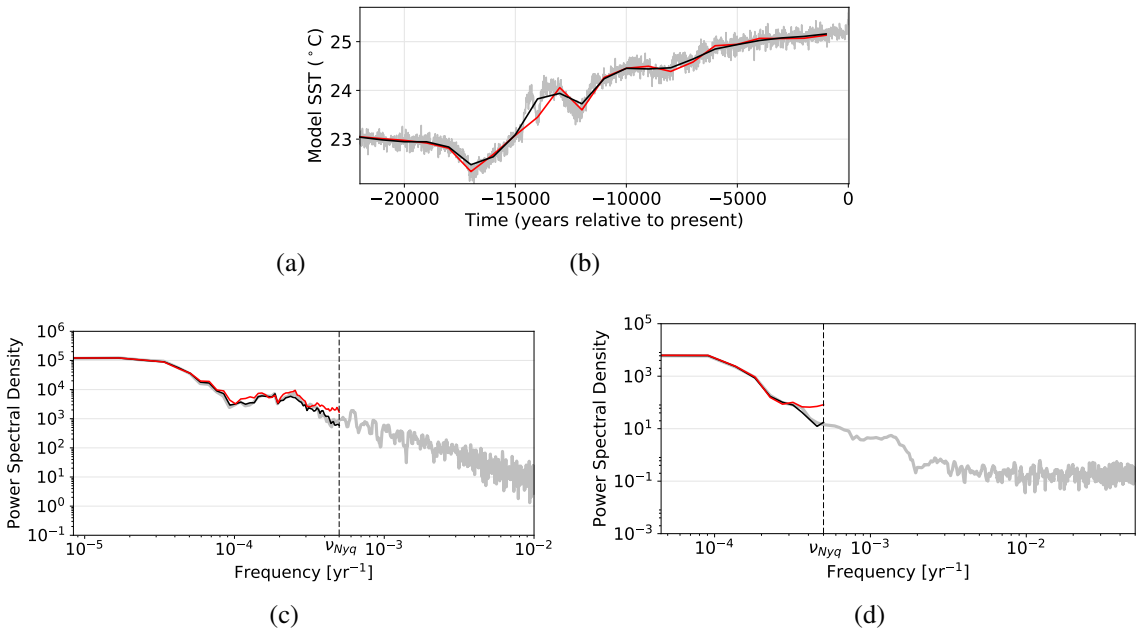

Figure 4: Effects of sampling errors illustrated by subsampling high-resolution records. (a) Comparison of the NGRIP oxygen isotope time series at 50 year resolution (grey), discretely subsampled using bin average periods of τ_b =1000 years and sampling periods of $\tau_s = 1000$ years (dense sampling, black) and using $\tau_b = 100$ with $\tau_s = 1000$ years (discontinuous sampling, red). (b) Same as (a) but for annually-averaged upper-ocean temperature at a grid box in the Cariaco Basin in the TraCE-21ka forced deglacial coupled climate simulation (*Liu et al.*, 2009). (c) Power spectral density estimates for the NGRIP record (grey), densely sampled record (black), and discontinuously sampled record (red) computed using the Thompson multitaper method. The grey envelope indicates 95% confidence intervals on the original spectrum computed using bootstrapping (*Prieto et al.*, 2009). The subsampled records have zero spectral power above the Nyquist frequency $\nu_{Nyq} = 5 \times 10^{-4}$ yrs^{-1}. (d) Same as (c) but corresponding to the TraCE-21ka time series.

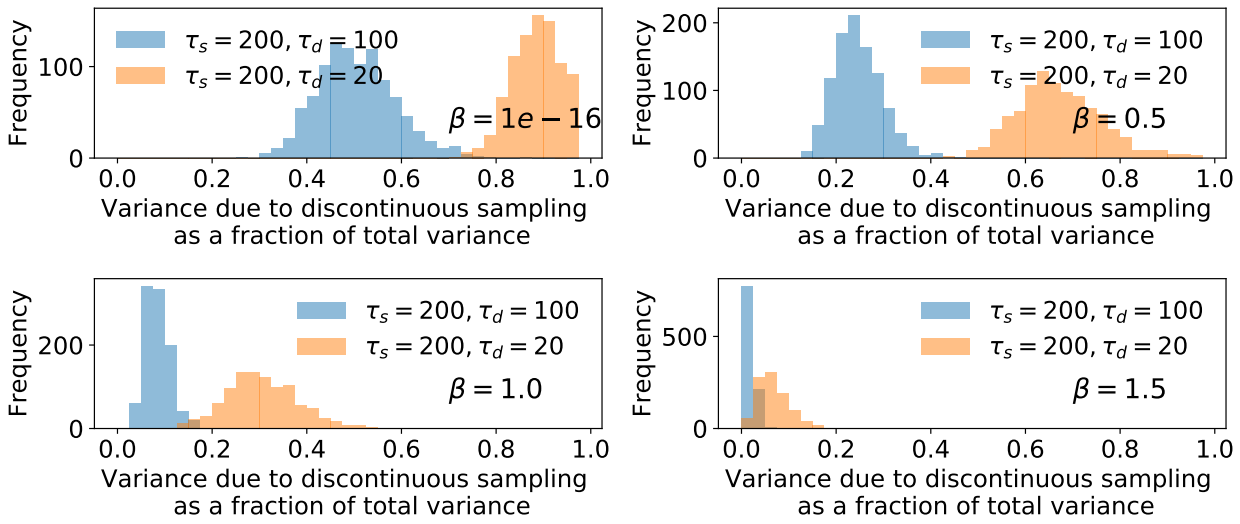

Figure 5

stance, the tendency for aliasing errors to be spectrally blue (highest at the highest frequencies in a record) should motivate considering the spectrum of model-data misfits and possibly focusing on lower-frequency model-data intercomparisons. In data intercomparisons, differing sampling strategies could contribute to disagreement among paleoclimate records obtained from different archives, while sampling errors could be correlated among related proxy quantities obtained from a single set of samples; the combination of these two effects could lead to apparently better agreement of properties within archives than between archives. Different sampling strategies can also lead to spurious lags in climate signals (Figure 4b). When τ_b is not clearly recorded in publications or published data, as is often the case, it can be difficult or impossible to quantify sampling uncertainties or devise appropriate time-averaging schemes for model-data comparisons or proxy system models.

Wavenumbers!

Estimating the signal spectrum $|\hat{x}|^2$ is a nontrivial task, particularly as it can be affected by aliasing (Figure 4; *Kirchner* (2005)). One could approximate $|\hat{x}|^2$ using spectra from other records that are more highly-resolved or were sampled densely, e.g. from a sediment core at an adjacent site, or a record believed to record similar climate variability. Alternately, measurements of archive

properties that can be made cheaply and at high resolution – such as magnetic susceptibility, wet bulk density, and other proxy properties that are routinely made on sediment cores – could prove useful for estimating $|\hat{x}|^2$ if those properties are related linearly to $x(t)$ (*Herbert and Mayer*, 1991; *Wunsch and Gunn*, 2003). Finally, power law spectra could be used as an approximate measure, as was done in the examples in Section 3..

In order to isolate the effects sampling errors we have ignored errors in record chronologies and assumed that proxy archives store information continuously, thereby neglecting, for example, errors due to small numbers of foraminifera in sediment cores or particle size sorting in diagenesis. While we have assumed that $x(t)$ is stationary, nonstationarity in record spectra could lead to time variations in errors, as may be evident in panels (a) and (c) of Figure 4, where sampling error variance appears to be greatest when record variance is largest. There is clear need for a comprehensive approach that can reveal interactions among the various sources of uncertainty in paleoclimate records; sampling error estimates should be taken as a lower bound on total uncertainties.

Caveats: Simplification of processes. Such smoothing would occur, for example, in the case where bioturbation homogenizes sedimentary material within a depth horizon.

This measure adequately describes variability for Gaussian $P(\Delta t)$; however, (due, for instance to the non-monotonicity of radiocarbon concentrations in the atmosphere; *Reimer et al.* (2009) $P(\Delta t)$ is in many cases non-Gaussian. As such, while it is instructive to consider effects as a function of a simple length scale, in many cases these estimates should be taken as a lower bound on uncertainty. More complete characterizations can be achieved by Monte Carlo sampling of age model uncertainty (e.g., *Tierney and Anchukaitis* (2011)). Because age model errors can be correlated between measurements – for instance, in the case where ages are interpolated between a set of "tie points" that is sampled less frequently then other observations – characterizing age model errors in time series can quickly become even more complex. Bayesian approaches to the generation of chronologies *Buck* (2004); *Buck and Millard* (2004); **?** give some insights into these cases, including pathological examples of inversions where time changes direction within a record.

Huybers and Wunsch (2004) include the effect of uncertainties in tie points in order to align multiple records of Pleistocene oxygen isotopes, and *Haam and Huybers* (2010) developed tools for estimating the statistics of time-uncertain series. The effect of time uncertainty on estimates of signal spectra is modest in some cases (Rhines and Huybers), possibly because time uncertainty acts like a smoother when taken in aggregate over an entire record (Moore and Thomson). Eventually, more complex forward models such as SEDPROXY (Laepple and Dolman 2018), which incorporate many of the effects of sampling discussed in this manuscript, may be the only way to assess the many interrelated sources of age model error and their interplay with other uncertainties.

Many practical considerations motivate the state of the art in paleoclimate sampling strategies, and in many cases the advantages of discontinuous sampling may outweigh the concerns described here. For instance, records sampled continuously cannot be used as a starting point for subsequently constructing higher-resolution records. Moreover, preservation of natural archives for subsequent analyses is important for reproducibility and for sharing scarce resources between laboratories, and may be complicated by continuous sampling. Finally, uncertainties due to discontinuous sampling may be dwarfed by other sources of error.

The hope is that the uncertainty quantification techniques described here can motivate sampling strategies for paleoclimate records and improve the fidelity of data-data and data-model comparisons. Values of τ_s, τ_b, β, etc. used here were chosen to provide illustrative examples, and independent analysis using appropriate values in (20), (**??**), and (**??**) is necessary to guide sampling or quantify uncertainty in other records. Regardless of the sampling strategy used, clearly reporting τ_s and τ_b in publications and data sets will facilitate uncertainty quantification and model-data comparisons and syntheses.

5 Appendix 1: Derivation of temporal representativity uncertainty in the frequency domain

The Fourier transform will be written using the operator $\mathscr{F}$ and by a hat, and denoting frequency by ν,

$$\mathscr{F}(x(t)) \equiv \hat{x}(\nu) = \int_{-\infty}^{\infty} x(t)\, e^{-2\pi i \nu t} dt.$$

Parseval's theorem states that the integral of a squared quantity in the time domain is equal to the integral of the squared amplitude of the Fourier transform of that quantity, so that

$$\left\langle \theta^2 \right\rangle = \int_{-\infty}^{\infty} \left(m(t,\tau_x) - m(t+\Delta t,\tau_y)\right)^2 dt \tag{24}$$

$$= \int_{-\infty}^{\infty} \left|\mathscr{F}\left[m(t,\tau_x) - m(t+\Delta t,\tau_y)\right]\right|^2 d\nu. \tag{25}$$

By the Fourier shift theorem,

$$\mathscr{F}\left[m(t+\Delta,\tau_y)\right] = e^{-2\pi i \nu \Delta}\mathscr{F}\left[m(t,\tau_y)\right]. \tag{26}$$

Then, by the linearity of the Fourier transform,

$$\left\langle \theta^2 \right\rangle = \int_{-\infty}^{\infty} \left|\hat{m}(\nu,\tau_y) - e^{-2\pi i \nu \Delta}\hat{m}(\nu,\tau_x)\right|^2 d\nu. \tag{27}$$

By the convolution theorem, convolution in the time domain is equivalent to multiplication in the frequency domain, and vice versa. Thus, the Fourier transform of a time mean as defined in Equation X is

$$\hat{m}(\nu,\tau) = \mathscr{F}\left[\Pi(t,\tau) \star r(t)\right] \tag{28}$$

$$= \hat{\Pi}(\nu,\tau)\cdot\hat{r}(\nu). \tag{29}$$

Substituting into Equation X yields

$$\left\langle \theta^2 \right\rangle = \int_{-\infty}^{\infty} \left| \hat{\Pi}(\nu, \tau_x) - e^{-2\pi i \nu \Delta} \cdot \hat{\Pi}(\nu, \tau_y) \right|^2 |\hat{r}(\nu)|^2 \, d\nu, \tag{30}$$

which states that $\left\langle \theta^2 \right\rangle$ is a weighted integral over the power spectral density $|\hat{r}(\nu)|^2$ of the climate signal $x(t)$.

To isolate the effect of a time offset Δ, consider the limit where τ_x and τ_y approach zero (corresponding to instantaneous observations in time), so that $\left\langle \theta^2 \right\rangle$ approaches

$$\left\langle \theta^2 \right\rangle = \int_{-\infty}^{\infty} \left| 1 - e^{-2\pi i \nu \Delta} \right|^2 |\hat{r}(\nu)|^2 \, d\nu. \tag{31}$$

Expanding $\left|1 - e^{-2\pi i \nu \Delta}\right|^2$ and simplifying gives

$$\left\langle \theta^2 \right\rangle = \int_{-\infty}^{\infty} (2 - 2\cos(2\pi\nu\Delta)) \, |\hat{r}(\nu)|^2 \, d\nu \tag{32}$$

so that the power transfer function is $H = 2 - 2\cos(2\pi\nu\Delta)$ and the expected error due to Δ is a cosinusoidally-weighted function of the signal power spectrum (Figure 6). H takes a minimum value of 0 at frequencies

$$\nu_{min} = 0, \frac{1}{\Delta}, \frac{2}{\Delta}, \ldots \frac{n}{\Delta}$$

for integer values of n; at these frequencies, measurements spaced by Δ in time are in phase and are therefore exactly correlated. The weights take a maximum value of 4 at frequencies

$$\nu_{max} = \frac{1}{2\Delta}, \frac{3}{2\Delta}, \frac{5}{2\Delta}, \ldots \frac{n}{\Delta} + \frac{1}{2\Delta}$$

where measurements separated by Δ are always exactly out of phase. At those frequencies, the underlying signal $r(t)$ is projected twofold onto the error, so that its variance contribution is multiplied fourfold. These variations in frequency contributions to error modulate effects from smoothing and sampling timescales, as illustrated in Figure 1.

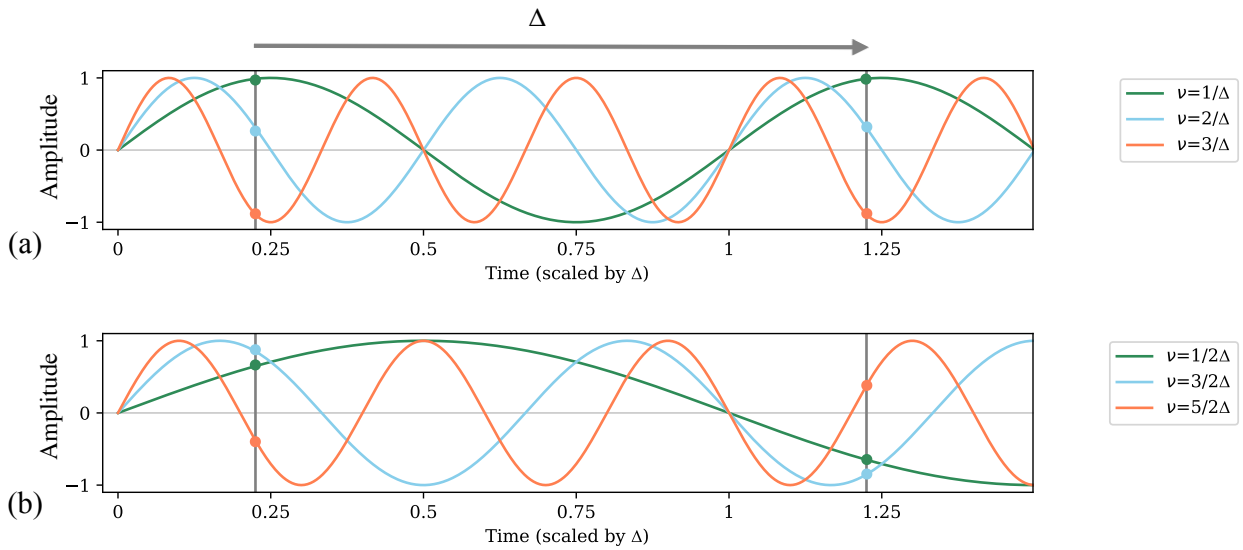

Figure 6: Illustration of the frequency dependence of errors in representing an instantaneous measurement of a hypothetical climate process $r(t)$ at a time t by another measurement $r(t+\Delta)$. Each line represents a different frequency component of $r(t)$, grey vertical lines represent sampling times, and colored circles represent values of components at those times. At frequencies $\nu = \frac{n}{\Delta}$ for $n = 0, 1, 2, \ldots$, (a), the Fourier components of $x(t)$ will be exactly in phase when sampled at a time lag Δ, so these components do not contribute to the error variance $\left\langle (r(t) - r(t+\Delta))^2 \right\rangle$. By contrast, at frequencies $\nu = \frac{n}{\Delta} + \frac{1}{2\Delta}$ (b), the Fourier components are exactly out of phase, so these components tend to contribute most to the error variance. At intermediate frequencies, contributions lie between the two extremes, leading to a cosine function of error contribution as a function of frequency (Equation 32).

Finally, we can represent when a climate signal $r(t)$ has been smoothed in a paleoclimate archive prior to sampling by substituting a new climate signal, $r(t)$, with a running mean applied,

$$r'(t) = \Pi(t, \tau_a) \star r(t).$$

Substituting $\hat{x}'(\nu)$ into (**??**) and applying the convolution theorem gives

$$\left\langle \theta^2 \right\rangle = \int_{-\infty}^{\infty} \left| \hat{\Pi}(\nu, \tau_x) - \hat{\Pi}(\nu, \tau_a) \cdot \hat{\Pi}(\nu, \tau_y) \cdot \right|^2 |\hat{r}(\nu)|^2 d\nu. \tag{33}$$

6 Appendix 2: Integration limits on error variance fractions

To compute the error variance fraction () this for power-law spectra we must define over a finite time series length τ_0 because when $\beta > 1$, the variance of $x(t)$ increases with the length of a time series. In frequency space, limiting time series length appears as a lower bound on integrals. Thus for the case of the time mean estimation problem, substituting () and () into () gives

$$f = \frac{\int_{\frac{1}{\tau_0}}^{\infty} H_{\tau_s, \tau_d, \tau_y, \Delta}(\nu) \, \nu^{-\beta} d\nu}{\int_{\frac{1}{\tau_0}}^{\infty} \operatorname{sinc}(\tau_y \nu) \, \nu^{-\beta} d\nu}. \tag{34}$$

For $\beta > 1$, as τ_0 increases, f decreases because the total variance (denominator) grows larger while the sampling error variance (numerator) is approximately constant because H is zero outside of a fixed frequency window whereas the sinc function is 1 at low frequencies. If all other parameters are held equal, as τ_0 increases, the denominator in (34) can become arbitrarily large and f can approach zero as aliasing errors are dwarfed by large variability at longer time scales. Thus, when using power law spectra to estimate errors, it is important to pick τ_0 to calibrate error estimates to appropriate time scales; for instance, for the LGM, an appropriate choice would be $\tau_0 = 21{,}000$ years.

Acknowledgements

Thanks to LuAnne Thompson, Greg Hakim, Lloyd Keigwin, Cristi Proistecescu, Carl Wunsch, and Thomas Laepple for useful conversations and comments. Support came from NOAA grant XXXX and an NSF postdoctoral fellowship. MATLAB code to integrate these equations and reproduce Figure 4 is available at https://github.com/amrhein/Discontinuous_Sampling.

References

Amrhein, D. E., G. Gebbie, O. Marchal, and C. Wunsch (2015), Inferring surface water equilibrium calcite δ 18 O during the last deglacial period from benthic foraminiferal records: Implications for ocean circulation, *Paleoceanography*, *30*(11), 1470–1489.

Anderson, D. M. (2001), {A}ttenuation of millennial-scale events by bioturbation in marine sediments, *Paleoceanography*, *16*(4), 352–357.

Beer, J., K. McCracken, and R. Von Steiger (2012), *{C}osmogenic radionuclides: theory and applications in the terrestrial and space environments*, Springer Science & Business Media.

Buck, C. E. (2004), Bayesian Chronological Data Interpretation: Where Now?, *LECTURE NOTES IN STATISTICS-NEW YORK-SPRINGER VERLAG-*, pp. 1–24.

Buck, C. E., and A. Millard (2004), *Tools for constructing chronologies: crossing disciplinary boundaries*, Springer Verlag.

Caley, T., D. M. Roche, C. Waelbroeck, and E. Michel (2014), {O}xygen stable isotopes during the {L}ast {G}lacial {M}aximum climate: perspectives from data–model (i{LOVECLIM}) comparison, *Climate of the Past*, *10*(6), 1939–1955.

Emery, W. J., and R. E. Thomson (2014), *{D}ata analysis methods in physical oceanography*, Newnes.

Evans, M. N., S. E. Tolwinski-Ward, D. M. Thompson, and K. J. Anchukaitis (2013), {A}pplications of proxy system modeling in high resolution paleoclimatology, *Quaternary Science Reviews*, *76*, 16–28.

Haam, E., and P. Huybers (2010), {A} test for the presence of covariance between time-uncertain series of data with application to the {D}ongge {C}ave speleothem and atmospheric radiocarbon records, *Paleoceanography*, *25*(2).

Hasselmann, K. (1976), {S}tochastic climate models part {I}. {T}heory, *Tellus*, *28*(6), 473–485.

Herbert, T. D., and L. A. Mayer (1991), {L}ong climatic time series from sediment physical property measurements, *Journal of Sedimentary Research*, *61*(7).

Huybers, P., and W. Curry (2006), {L}inks between annual, {M}ilankovitch and continuum temperature variability, *Nature*, *441*(7091), 329–332.

Huybers, P., and C. Wunsch (2004), A depth-derived Pleistocene age model: Uncertainty estimates, sedimentation variability, and nonlinear climate change, *Paleoceanography*, *19*(1), PA1028.

Jones, R. H. (1972), {A}liasing with unequally spaced observations, *Journal of Applied Meteorology*, *11*(2), 245–254, doi:10.1175/1520-0450(1972)011<0245:AWUSO>2.0.CO;2.

Kirchner, J. W. (2005), Aliasing in 1/f(alpha) noise spectra: origins, consequences, and remedies., *Physical review. E, Statistical, nonlinear, and soft matter physics*, *71*(6 Pt 2), 066,110, doi:10.1103/PhysRevE.71.066110.

Liu, Z., B. L. Otto-Bliesner, F. He, E. C. Brady, R. Tomas, P. U. Clark, A. E. Carlson, J. Lynch-Stieglitz, W. Curry, E. Brook, and Others (2009), {T}ransient simulation of last deglaciation with a new mechanism for {B}ølling-{A}llerød warming, *Science*, *325*(5938), 310–314.

MARGO Project Members (2009), {C}onstraints on the magnitude and patterns of ocean cooling at the {L}ast {G}lacial {M}aximum, *Nature Geoscience*, *2*(2), 127–132.

McGee, D., G. Winckler, J. B. W. Stuut, L. I. Bradtmiller, and Others (2013), {T}he magnitude, timing and abruptness of changes in {N}orth {A}frican dust deposition over the last 20,000 yr, *Earth and Planetary Science Letters*, *371*, 163–176.

Pisias, N. G., and A. C. Mix (1988), {A}liasing of the geologic record and the search for long-period {M}ilankovitch cycles, *Paleoceanography*, *3*(5), 613–619.

Prieto, G. A., R. L. Parker, and F. L. Vernon III (2009), {A} {F}ortran 90 library for multitaper spectrum analysis, *Computers & Geosciences*, *35*(8), 1701–1710.

Reimer, P. J., M. G. L. Baillie, E. Bard, A. Bayliss, J. W. Beck, P. G. Blackwell, C. B. Ramsey, C. E. Buck, G. S. Burr, R. L. Edwards, and Others (2009), {INTCAL}09 {AND} {MARINE}09 {RADIOCARBON} {AGE} {CALIBRATION} {CURVES}, 0–50,000 {YEARS} {CAL} {BP}, *RADIOCARBON*, *51*(4), 1111–1150.

Schulz, M., and K. Stattegger (1997), SPECTRUM: Spectral analysis of unevenly spaced paleoclimatic time series, *Computers and Geosciences*, *23*(9), 929–945, doi:10.1016/S0098-3004(97)00087-3.

Shannon, C. E. (1949), {C}ommunication in the presence of noise, *Proceedings of the IRE*, *37*(1), 10–21.

Tierney, J. E., and K. J. Anchukaitis (2011), {I}dentifying coherent spatio-temporal modes tropical {Africa}n hydrology during the last 1500 years, *Climate Dynamics, in prep.*

von Albedyll, L., T. Opel, D. Fritzsche, S. Merchel, T. Laepple, and G. Rugel (2017), 10 {B}e in the {A}kademii {N}auk ice core–first results for {CE} 1590–1950 and implications for future chronology validation, *Journal of Glaciology*, pp. 1–9.

Wunsch, C. (1978), {T}he {N}orth {A}tlantic general circulation west of 50˜{W} determined by inverse methods, *Reviews of Geophysics*, *16*(4), 583–620.

Wunsch, C. (2000), 0n sharp spectral lines in the climate record and the millennial peak, *Paleoceanography*, *15*(4), 417–424.

Wunsch, C., and D. E. Gunn (2003), {A} densely sampled core and climate variable aliasing, *Geo-Marine Letters*, *23*(1), 64–71.