Simple noise estimates and pseudoproxies for the last 21k years

Oliver Bothe[1], Sebastian Wagner[1], and Eduardo Zorita[1]

[1]Helmholtz Zentrum Geesthacht, Institute of Coastal Research, 21502 Geesthacht, Germany

Correspondence: Oliver Bothe (ol.bothe@gmail.com)

Abstract. Climate reconstructions are means to extract the signal from uncertain paleo-observations, i.e. proxies. It is essential to evaluate these to understand and quantify their uncertainties. Similarly, comparing climate simulations and proxies requires approaches to bridge the, e.g., temporal and spatial differences between both and address their specific uncertainties. One way to achieve these two goals are so called pseudoproxies. These are surrogate proxy records within, e.g., the virtual reality of a
5 climate simulation. They in turn depend on an understanding of the uncertainties of the real proxies, i.e. the noise-characteristics disturbing the original environmental signal. Common pseudoproxy approaches so far concentrated on data with high temporal resolution from, e.g., tree-rings or ice-cores over the last approximately 2,000 years. Here we provide a simple but flexible noise model for potentially low-resolution sedimentary climate proxies for temperature on millennial time-scales, the code for calculating a set of pseudoproxies from a simulation and, for one simulation, the pseudoproxies themselves. The noise model
10 considers the influence of other environmental variables, a dependence on the climate state, a bias due to changing seasonality, modifications of the archive (e.g., bioturbation), potential sampling variability, and a measurement error. Model, code, and data should allow to develop new ways of comparing simulation data with proxies on long time-scales. Code and data are available at https://doi.org/10.17605/OSF.IO/ZBEHX.

1   Introduction

15 Proxy-records and derived reconstructions are our only observationally based information about past climates before the period covered by human observations, i.e., documentary or instrumental evidence. Climate reconstruction methods statistically process the information in the proxy records to extract the recorded climate signal. However, part of the variability in the proxy records is not caused by the climate. This non-climatic variability, i.e., the proxy noise, may cause biases and uncertainties in the resulting climate reconstructions. Evaluating the quality and reliability of reconstructions and of proxy-records requires
20 an understanding of the noise in the proxy-records. Only this knowledge allows us to obtain reliable estimates of the errors in reconstructed properties.

Some aspects of statistical climate reconstruction methods can be evaluated in so-called pseudoproxy experiments. In these experiments, the reconstruction methods are mimicked in the controlled conditions provided by climate simulations with Earth System Models. However, for these tests surrogate proxy records have to be produced which are compatible with the climate simulated by these models – the pseudoproxies. In testing the reconstruction methods, pseudoproxies then replace the real
25 paleo-observations in the method and the virtual climate of the Earth System simulation stands in for the real climate. For a

useful test of reconstruction methods, the pseudoproxies should be as realistic as possible, with statistical properties similar to the real proxies. This is achieved by contaminating the climate variables simulated by the Earth System Model with statistical noise with a certain amplitude and statistical characteristics. These properties ideally are based on estimates of a realistic or at least plausible noise to successfully mimic the behavior of real-world proxies.

5   Studies of the climate of the past 2,000 years regularly use such pseudoproxy approaches mimicking annually resolved proxies such as dendroclimatogical ones. Smerdon (2012) reviews the approach of using pseudoproxy-experiments to evaluate reconstruction methods with a focus on the last millennium. Such methods basically originated in Mann and Rutherford's (2002) paper focussing on climate-field reconstructions. The review emphasizes the essential contribution of pseudoproxy-experiments to our understanding of past climates and evaluating our methods of studying past climates. Most studies using 10 pseudoproxies concentrated on the last few millennia. Few studies considered periods further in the past (e.g., Laepple and Huybers, 2013; Dolman and Laepple, 2018).

     There are various approaches to obtain such pseudoproxies. On the one hand we can try to obtain comprehensive representations of the proxy-system, i.e., we use forward models of the proxies under consideration (compare Laepple and Huybers, 2013; Dolman and Laepple, 2018; see also, e.g., Dee et al., 2015; Dee et al., 2018; Evans et al., 2013). Secondly, we can try to 15 formulate a mathematically tractable expression of the proxy error [Dolman et al., in preparation]. A third way of formulating the proxy noise is to use a simple estimate of a plausible non-climatic error in proxy-records. The different approaches can be very general or specific for certain proxy types.

     A recipe for calculating pseudoproxies may include a variety of potential error estimates not only within the assumed proxy-system but also in the relation between the 'observed' data and time, i.e. the anchoring of the data in time. These errors include 20 different sources of errors related to laboratory uncertainties like measurement errors and reproducibility, local disturbances, dating uncertainty, time resolution, serial autocorrelation, and all possibly dependent on the overall climate state. Further uncertainty includes habitat preferences, seasonal biases, the variability in the relation between sensor and environment, long term changes in this relation, long term modifications of the archive, sampling variability and sampling disturbances, and not least generally erroneous assumptions on the researcher's side on the relation between recording sensor and environment, i.e., 25 the calibration relation.

     Regarding dating/age uncertainty, there are various approaches to dealing with it (e.g., Breitenbach et al., 2012; Carré et al., 2012; Anchukaitis and Tierney, 2013; Comboul et al., 2014; Goswami et al., 2014; Brierley and Rehfeld, 2014; Rehfeld and Kurths, 2014; Kopp et al., 2016; Boers et al., 2017) of which a number try to transfer the dating uncertainty towards the proxy-record-uncertainty (e.g., Breitenbach et al., 2012; Goswami et al., 2014; Boers et al., 2017). As our interest is less in a 30 probabilistic description and rather in how we can capture the error in a time-series, we take a simple approach to include an error-term resulting from dating uncertainty.

     Besides evaluating reconstruction methods, a plausible estimate of noise within the proxies also can assist in comparison studies between model-simulations and the proxy-records. This helps our understanding about past climate changes by consolidating information from all available sources, i.e., proxy records and model simulations. The lack of high-quality observations 35 with small uncertainty is always going to hamper efforts to assess the quality of model-simulations of past climates. Such com-

parisons have to rely on the paleo-observations from proxies, and even the highest-quality proxy-records have an irreducible amount of uncertainty.

Most often data-model-comparisons use the model reality as base of the comparisons. In the case of proxies, the comparison is between a, e.g., temperature reconstruction and the model. The alternative is to compare both in the proxy-space using a proxy-representation of the model-climate. Pseudoproxies or a recipe how to compute them may be part of an interface between the data on the one side and the model simulations on the other side.

Inferences about past climates from proxy-data base on observations on an archive that accumulated a property of a system. This (property of the) system recorded, i.e. was sensitive to, an environmental process at some date. From the recording stage to our inference there are multiple sources of error to our inference.

Evans et al. (2013) describe a simple modelling framework from environment across sensor and archive towards the observation. Each stage in this process adds uncertainty. Evans et al. simplify the full process through time to three stages: sensor, archive, and measurement. For example, the sensor 'tree' records the environmental influences in its archive 'wood', and we can make measurements on this archive, e.g., in form of tree-ring counts and widths etc. On top of this one could use additional stages for the environment and the final reconstruction, however, we can include the associated uncertainties in any of the three stages proposed by Evans et al.

In this paper, we present a formal but still simple approach to describe the noise present in proxy records, which can also be applied to the generation of pseudoproxies for timescales longer than the last few millennia, i.e. including also proxies with coarser time resolutions than interannual and afflicted by larger degrees of dating uncertainty. Thereby this work extends on previous pseudoproxy-approaches, which often concentrated on well dated proxy-systems affected by fewer sources of uncertainty. The following presents a set of assumptions on proxy noise and estimates for some of the mentioned error sources. We further provide pseudoproxies based on these assumptions for the TraCE-21ka simulation (Liu et al., 2009), which cover the last 21,000 years. We concentrate on proxies which are subject to some kind of sedimentary process. The manuscript assets at https://osf.io/zbehx/ also include example code and the calculated pseudoproxy data. Thereby the manuscript provides for one simulation the data to make an informed comparison with real proxies and the data to evaluate reconstruction techniques. Code and assumptions enable any interested researcher to produce similar pseudoproxies for their simulation of interest. We consider the measurement error, local changes to the original proxy-recording (compare, e.g., Laepple and Huybers, 2013), the basic climate state, a potential bias, and a simple estimate of the effect of dating uncertainty. All noise expressions are coded in a way to flexibly allow for different colors and types of noise.

2   Input Data

We use the summer (June, July, August; JJA) mean temperature at each grid-point of the TraCE-21ka simulation (Liu et al., 2009). To date, this is the only available interannual transient Earth System Model simulation covering the last 21,000 years. Specific technical considerations, e.g., related to freshwater pulses and sea-level adjustments lead to some artefacts in the

simulation output data fields. A brief description of the simulation can be found at http://www.cgd.ucar.edu/ccr/TraCE/, and He (2011) describes the simulation in more detail in his Ph.D.-dissertation.

The presented code uses only one grid-point at 0E, 42.68N. Figures generally show results for this grid-point as well. This choice is arbitrary. Since this is indeed a grid-point on land, the simulation output at this grid-point has the benefit of representing a rather smooth evolution of temperature over the last 21,000 years compared to, e.g., a marine grid-point affected

5 by the freshwater forcing of He (2011, compare also Liu et al. (2009)). On the other hand, this implies the disadvantage of featuring less extreme climate variations to be captured in a subsequent pseudoproxy. The document assets provide Figures equivalent to those in this document, which show the output for a grid-point at 11.25W, 42.68N in the North Atlantic off the coast of the northern Iberian peninsula.

On multi-millennial time-scales we have to consider changes in the insolation. Global insolation data is calculated using the

10 R (R Core Team, 2017) package palinsol (Crucifix, 2016).

We use for most noise-processes simple Gaussian noise. However, as the code is flexible, the user can easily change this.

3   Considerations and Results

In defining what we consider as noise, we first have to state the signal which we assume the proxy records. That is, do we assume, the proxy records local or regionally accumulated signals? Here, we take the signal of interest to be local, that is

15 non-local influences enter the noise term and are not part of the signal. In addition, there are further local factors which affect the recording of the signal but are not part of the signal of interest.

The appendix provides tables (Tables A1 to A3) summarising the considered parameters and noise models in the various steps of the following considerations.

In the following, we distinguish between different sources of errors related to the concept of sensor, archive and measure-

20 ments of Evans et al. (2013). Each section contains a discussion of the implications of the respective error term. Afterwards we discuss the results of applying the respective step in the framework to the output of the TraCE-21ka simulation.

3.1   Assumptions on essential error sources 1: Sensor

3.1.1   Noise

The sensor, e.g. an organism or a physical or biogeochemical process, reacts to multiple parts of its environment. Researchers

25 interest often is only in one of the environmental variables.

The sensor, S, is likely a nonlinear function of the environment, $S(E)$, where $E = \{e_i\}$, with e_i being components of the environment. If our interest is only in the sensor's reaction to one variable, T,

$$S(E) \approx \widehat{S}(T, \eta_i) \qquad (1)$$

Under this assumption further components of the environment besides T contribute only noise components η_i to the reaction

30 of the sensor. These errors are not necessarily additive but can also be multiplicative or could bias the estimate. In a first step
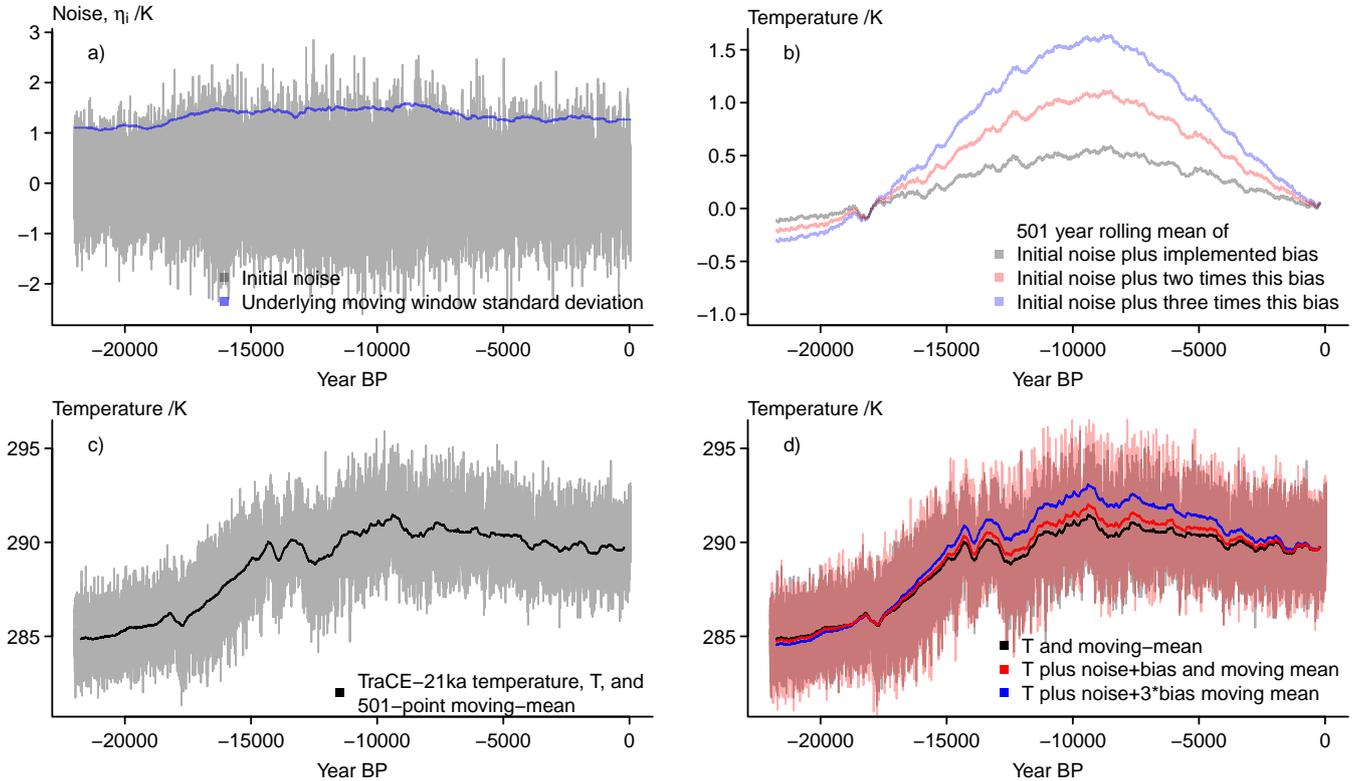
Figure 1. Visualising considered error sources at the sensor-stage: a) the initial noise and the underlying moving window standard deviations of the input signal, b) three versions of a potential bias as function of the local insolation, c) the input data and its 501-point moving mean, d) the input data and its 501-point moving mean plus noise and bias. The unsmoothed initial temperature is effectively hidden behind the unsmoothed temperature plus bias.

we, here, assume the sensor-reaction to be

$$S(E) \approx \widehat{S}(T) + \eta_i \qquad (2)$$

Any of these errors or noise-processes may show auto-correlation in either space or time or both. Any such process may, in turn, add memory to the sensor-system. Indeed this memory-effect and spatial or temporal correlations may be large. For example, if a process takes place in an environment with slowly and fast varying components, and our interest is in one of the fast components, the low frequent variations add a noise or error with high auto-correlation in time.

The sensor reacts to all, potentially high-frequent, changes in its environment. This local environment is unlikely isolated from the larger scale system. Additional noise may, thus, be due to the sensor reacting to advected environmental properties instead of "local" ones or due to the environment redistributing the sensor or the record. In the marine but also in lake domains, currents may influence the sensor, while in many domains the wind may affect the recording of the signal. Furthermore, small and large scale spatial variations of the process may affect the signal and contribute to the record. Our approach regards these

contributions as noise. All these influences may introduce spatial and, here considered to be of more importance, temporal correlations in those environmental properties which we here consider as part of the noise term. We assume that advection from other regions, i.e., currents and wind, are especially important in contributing autocorrelation to our noise process. One can see these non-local factors as noise in the archive rather than the sensor.

5   Besides simple noise, redistributions of the environmental signal may also introduce biases in our estimate of the environment. Such biases in turn are likely not fully time-constant but evolve with the environment on interannual, multi-decadal, and multi-centennial to millennial time-scales. The different time-scales result from the different time-scales of the environment. This is relevant for recent climate changes and interannual to interdecadal climate variability, but it becomes even more important for multi-millennial time-scales where we have to deal with the effects of changing seasons, glaciation, deglaciation,

10   changes in bathymetry, and lithospheric adjustments. Such biases also lead to autocorrelation in the error.

One example of such time-evolving biases are changes in the seasonality of the environmental sensor. While one can see this again as a source of uncertainty in a narrowly defined proxy-system from sensor to reconstruction, it is in the end a bias of our attribution of the measurement to one season. That is, it is a bias at the reconstruction-level rather than on the sensor level. There are other potentially erroneous attributions besides the processes' seasonality. These are the location of the process in

15   all three dimensions, e.g., the habitat of living organisms, and a generally only partially correct calibration relationship. Again, while these are environmental factors influencing the sensor and can be considered as noise here, they are mainly errors in our reconstruction-calibration-relation. This non-stationarity of our reconstruction-calibration-relation is an important source of uncertainty, although the idea that the modern relations worked over the full period of interest (e.g., Bradley, 2015) is a fundamental assumption of paleo-climatology.

20   In the following we consider two error terms. However, we assume three components of the noise to be important at the sensor level, the environmental noise, the redistribution, and the attribution errors which we here reduce to the potential biases due to changes in the seasonality. Taking all three components the sensor-record becomes

$$S(E) \approx \widehat{S}(T) + \eta_{env} + \eta_{redistr} + \eta_{season} \tag{3}$$

where we for the moment replace η_i by η_{env}. We assume in the following η_i includes both the effects of environmental

25   dependencies and of redistribution, i.e., our $\eta_i = \eta_{env} + \eta_{redistr}$. This is the first error term.

This in fact implies that we should consider auto-correlated noise-processes. However, for simplicity, the currently used version of η_i is a white noise process and thereby ignores that redistribution and other processes likely introduce temporal and spatial correlations in the errors. The code includes the noise model as a set of parameters which the user can easily change to include an autocorrelated noise model.

30   Our pseudoproxy at this point becomes, if we only modify the model-output and concentrate on one parameter T, e.g., temperature data,

$$P(x, y, t, T) = P_T = T + \eta_i \tag{4}$$

We take η_i to be

$$\eta_i = p \cdot \mathcal{N}(0, S(t)^2) \tag{5}$$

where p is a constant scaling factor , and $S(t)$ is a time-dependent standard deviation. The time-dependence mimics a dependence of the noise on the background climate variability on long time-scales. Here, we use a 1000 year moving standard deviation, i.e., $S(t_i) = \sigma(T(t_{i-499} : t_{i+500}))$. Again, the code easily allows to change the noise-model to an AR-process with innovations generated with standard deviation $S(t)$. Our formulation assumes that noise variability increases with increasing variability in the parameter T. Obviously, it could also be that noise variability reduces or reacts totally differently relative to the variability of T. The code includes a commented version where we invert the moving standard deviation about its mean.

We can consider the changes of the seasonality, η_{season}, as an orbitally influenced bias term which we compute first for the current latitude. We apply the orbital bias term as additive but one may see it as multiplicative or nonlinear effect in many cases. Therefore the code uses it after the noise term η_i. This is the second error term.

We add the bias term dependent on the latitudinal insolation. In its formulation we concentrate on summer insolation. The insolation bias is scaled to be zero in the year 0BP. The bias becomes notable at some latitudes but may be rather negligible elsewhere. The bias is scaled to be positive if the insolation is larger but this can be randomized. The bias is scaled by an ad hoc constant. Then the pseudoproxy becomes

$$P_T(t) = T(t) + \eta_i(t) + Bias(t) \tag{6}$$

3.1.2 Results

Figure 1a shows an example of the initial noise η_i. The dependence on the background state is obvious with an increase during the deglaciation. The blue line in the panel gives the underlying moving standard deviation. Indeed, Rehfeld et al. (2018) diagnose a reduction in temperature variability from the Last Glacial Maximum to the Holocene by studying centennial to millennial time-scales.

Panel b of Figure 1 compares three versions of the orbitally induced bias. We use the version with the smallest amplitude.

Panel c of the Figure presents the grid-point temperature of the TraCE-21ka simulation and a simple 501-year running mean. The comparison with Figure 1d highlights that the effect of the bias is rather small given our choice of its amplitude. Nevertheless, comparing the panels also clarifies that our implementation of the bias results in a warmer record over most of the considered time period while the record becomes slightly colder in the very early portion of the simulated data.

3.2 Assumptions on essential error sources 2: Archive

3.2.1 Noise

So far our approach describes a record of an environmental influence plus two error terms. This record becomes subsequently integrated in an archive. Afterwards, various processes may modify the archive or redistribute it. Modifications include selective
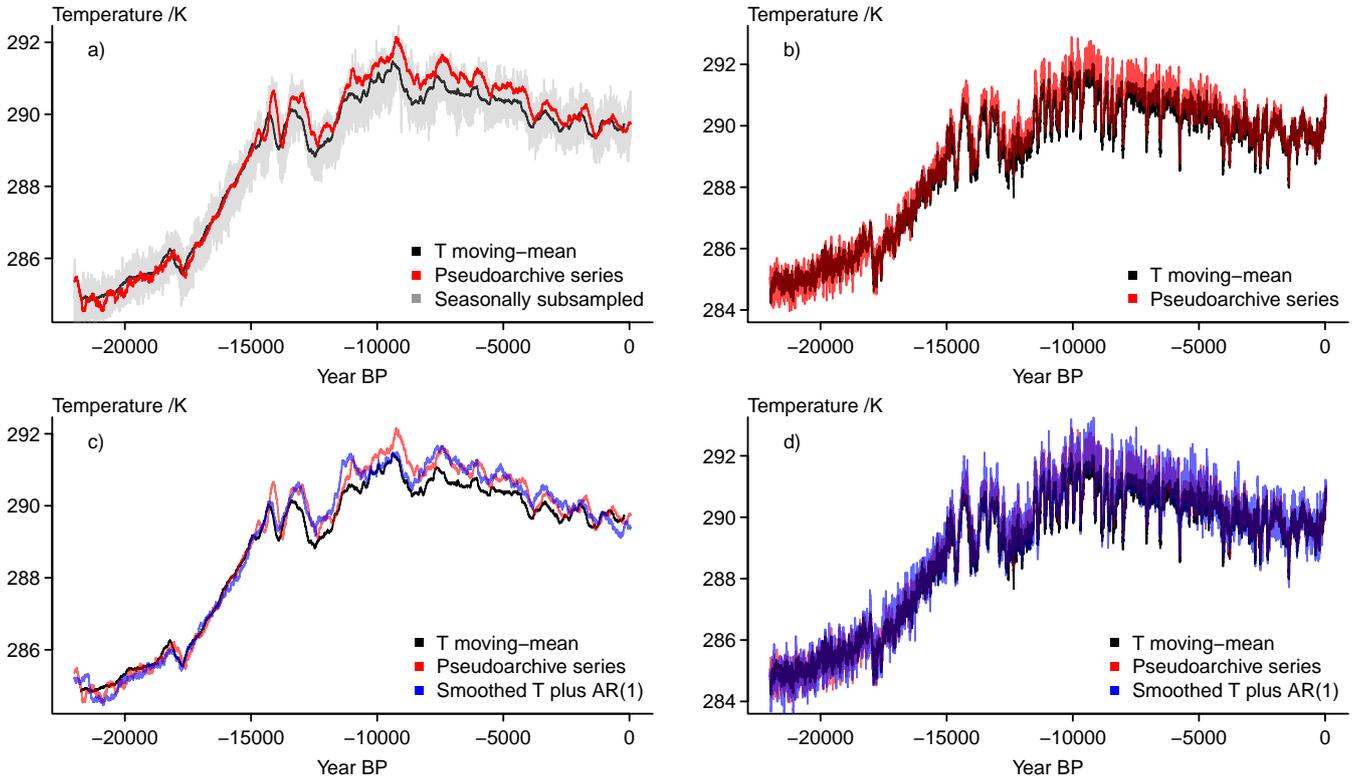
Figure 2. Visualising considered error sources at the archive-stage: a) 501-year moving mean of the input data, the pseudo-archive series with longer average smoothing lengths, and the subsampled record, b) 501-year moving mean of the input data, the pseudo-archive series with shorter average smoothing lengths, c) 501-year moving mean of the input data, the pseudo-archive series with longer average smoothing lengths, and the version with constant smoothing and added AR(1)-process, d) 501-year moving mean of the input data, the pseudo-archive series with shorter average smoothing lengths, and the version with shorter constant smoothing and added AR(1)-process.

destruction of parts of the record by processes acting all the time or by sparse random events or continually acting random processes. Examples are bioturbation or re-suspension. These processes may result either in a correlated noise in time and space or simply white noise. Other de facto white noise errors may result from our finite and random sampling of the archive. However, this may be rather part of the observational noise.

5   Because we focus on sedimentary proxies, we argue that the archiving process foremost is a filter of variability above a certain frequency level, e.g., by diffusive processes or bioturbation (compare, e.g., Dolman and Laepple, 2018, and their references). Dependent on the system in question this may only affect the very high frequencies but for other systems it may extend to multi-decadal or even centennial to millennial frequencies. On top of this smoothing of the archive, there may be additional noise as the smoothing function is unlikely homogeneous. We assume such a filtering to be the fundamental

10   modification of the record in the archive, and, thus, only consider this process in our archive modelling.

Inspired by the simple proxy forward formulations of Laepple and Huybers (2013; see also Dolman and Laepple, 2018), we produce five different versions of the archived pseudoproxy-series. The first and second series are simple running averages of the recorded proxy on which we add a highly correlated AR(1)-process. The two versions differ in the length of the averaging window, the AR-coefficients, and the standard-deviations of the innovations. The versions three and four similarly differ in the amount of average smoothing, but we use random window lengths for each date. The rationale for the two different smoothing lengths is to mimic both strongly and only slightly smoothed proxies.

The fifth version aims to mimic the behavior of proxies when researchers use only a small part of an available proxy, e.g., pick only a certain number of a samples. An example is the simple forward formulation for Mg/Ca proxies by Laepple and Huybers (2013; see also Dolman and Laepple, 2018).

Smoothing lengths and random factors in this approach could depend on the background climate. We choose not to consider this possibility, but one can easily include a time-dependent standard-deviation of the innovations here.

The smoothed archive records are then either

$$P_T(t) = g_1(T(t) + \eta_i(t) + Bias(t), t) \tag{7}$$

where $g_1(t)$ is the time dependent filter, or

$$P_T(t) = g_2(T(t) + \eta_i(t) + Bias(t)) + AR \tag{8}$$

where g_2 is the constant smoothing and we add an AR-process to account for the inhomogeneities in the smoothing.

The fifth version of the pseudoproxies subsamples over the random filter interval and adds a noise term to mimic a seasonal uncertainty. That is, we sample n years within the filter interval, and take the mean over the temperature and the noise for these years. We add another noise term to mimic the intra-annual seasonal uncertainty.

P_T in this case becomes

$$P_T = h(T(t), t) + h(\eta_i(t), t) + \eta_s \tag{9}$$

where $h(t)$ represents the sub-sampling and η_s the intra-annual noise.

We do not include the bias term here. On the one hand we apply this bias only for the summer season temperature, i.e. other seasons show different biases. While we could account for this by sampling the biases of other seasons or even months in producing $h(t)$ or η_s, we prefer to keep our model simpler. Excluding the bias term may be interpreted in terms of the seasonal subsampling cancelling out the bias. In reality any cancellation would not result in a convergence on the simulated climate state but more likely on a recorded value between the biased and the 'true' climate. The coded version of the sub-sampling still includes the bias-term as a comment.

3.2.2 Results

Already the biased moving average shows the differences between the target temperature and the pseudoproxy-record. The pseudo-archive-series in Figure 2a shows this more clearly. Here we use a randomized smoothing interval. Differences are less visible for shorter random smoothing intervals (compare Figure 2b).

Further panels of Figure 2 add the constant smoothing archive approximations which we modify by an additional highly correlated AR-process (Figure 2c and d). This procedure randomly amplifies, dampens, or inverts certain biases in the presented case. That is, while the simple random smoothing emphasizes the bias, the AR-procedure overlies this bias with additional millennial-scale variations.

The panels highlight an apparent offset between the randomly smoothed archive series, the constantly smoothed archive series, and the smoothed input data. The smoothed version of the input data as well as the constant filtering use a centered approach, i.e., they are symmetric about their date. The time varying smoothing tries to more realistically mimic a bioturbation approach (compare Dolman and Laepple, 2018, and their references) and thus provides a shift in the series.

Figure 2a also shows the seasonally subsampled pseudo-archive-proxy. The data does ignore the bias term and the resulting series is by construction symmetric around the original data, i.e. the target. Nevertheless, there are pronounced deviations from the orginial data. Considering only the deviations from the target temperature moving mean (not shown) highlights that this approach is notably more noisy than the filtered data but preserves pronounced longer term excursions of the input data.

3.3 Assumptions on essential error sources 3: Measurements

3.3.1 Noise

The archiving represents also a transformation from time-units to archive-distance-units, e.g., depths, rings, distances. The proxy becomes a tuple of date and data. Now the dates are uncertain as each data-point includes information from different original dates due to the smoothing function, the sampling introduces uncertainties, and our dating of our samples is a profoundly uncertain process.

3.3.2 Measurement error

Prior to dealing with dating uncertainty, we take an additional noise term to mimic measurement errors and apply this for each date to account for the potentially imperfectly measured series. The term includes not only the uncertainties in our assumed methods of measuring the proxies and the methods' potential to make mistakes. This "true" measurement error may result in biases due to limits of what our methods can detect or systematic offsets due to a laboratory-specific, potentially erroneous, approach to the measurement. Potential offsets imply that we should generally expect a certain amount of auto-correlation in this noise. The term further has to account for the accidental handling of the records in the laboratory, e.g., influences from storage or from other processing of the samples and the data, which may result in autocorrelated errors if these influences have a systematic component.

Thus, it is not necessarily the case that we can consider inter-laboratory reproducibility as white noise. However, the intra-laboratory repeatability is likely indeed a white random process. We also assume repeatability and reproducibility to be part of our measurement error term.

While there are obviously many reasons to assume autocorrelation in this error-term, we, here, only provide a white noise term for the measurement noise. Again, the code allows to modify this.
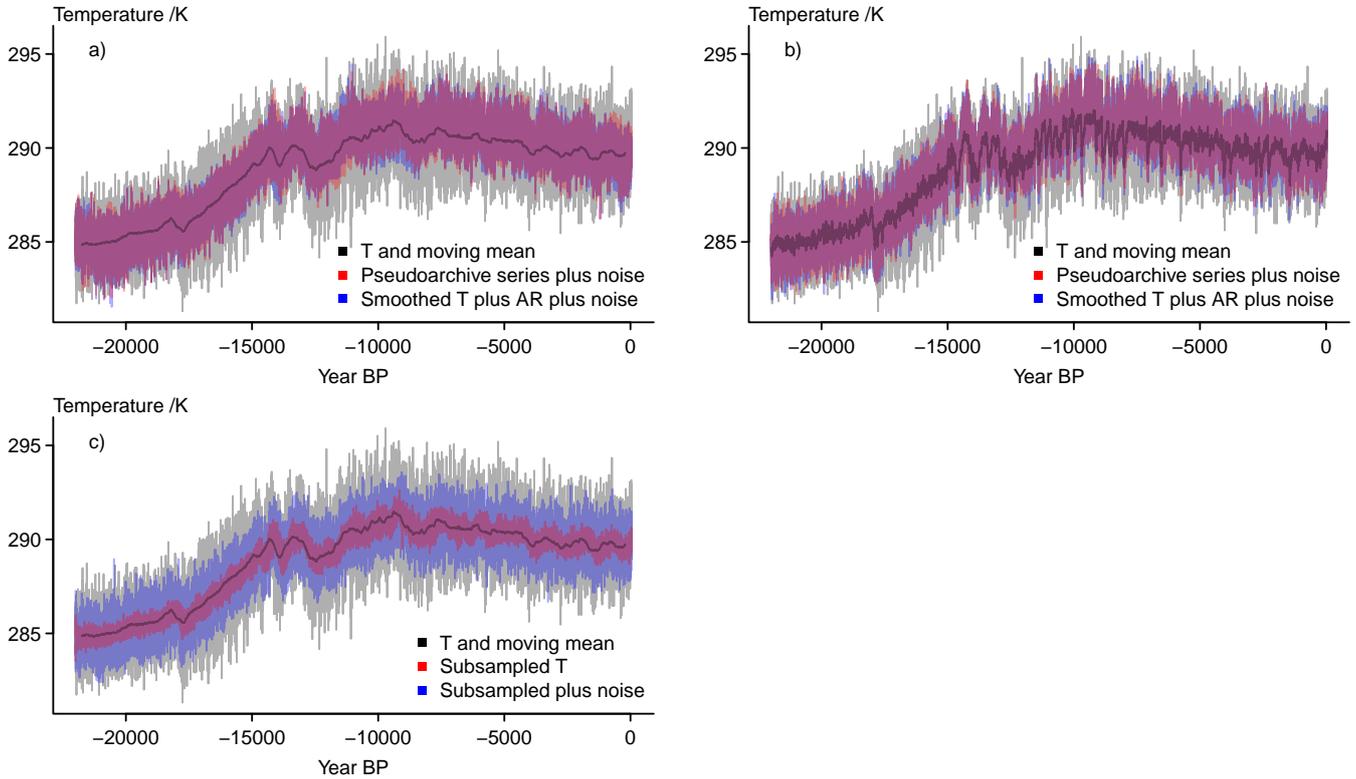
Figure 3. Visualising considered error sources at the measurement-stage for the full series: a) 501-year moving mean of the input data, the pseudo-archive series with longer average smoothing lengths and the constant smoothing plus AR series with added measurement noise, b) 501-year moving mean of the input data, the pseudo-archive series with shorter average smoothing lengths and the constant smoothing plus AR series with added measurement noise, c) 501-year moving mean of the input data, the subsampled record, and the subsampled record with added measurement noise.

We apply the measurement error term at the end. However, we introduce this term before dealing with the dating uncertainty since we provide proxies without dating uncertainty. The measured proxy-series becomes

$$M_T = P_T + \eta_M \tag{10}$$

In reality, we do not have a continuously sampled series, but obtain only samples at certain intervals. Assuming N samples the sampled pseudoproxy becomes

$$P_{P_T} = P_T(t = \{t_1, ..., t_N\}) \tag{11}$$

The sampling of the archive likely produces errors in the samples. We assume these are included in the measurement uncertainty. We provide at each grid-point sampled series of the pseudoproxies detailed above.
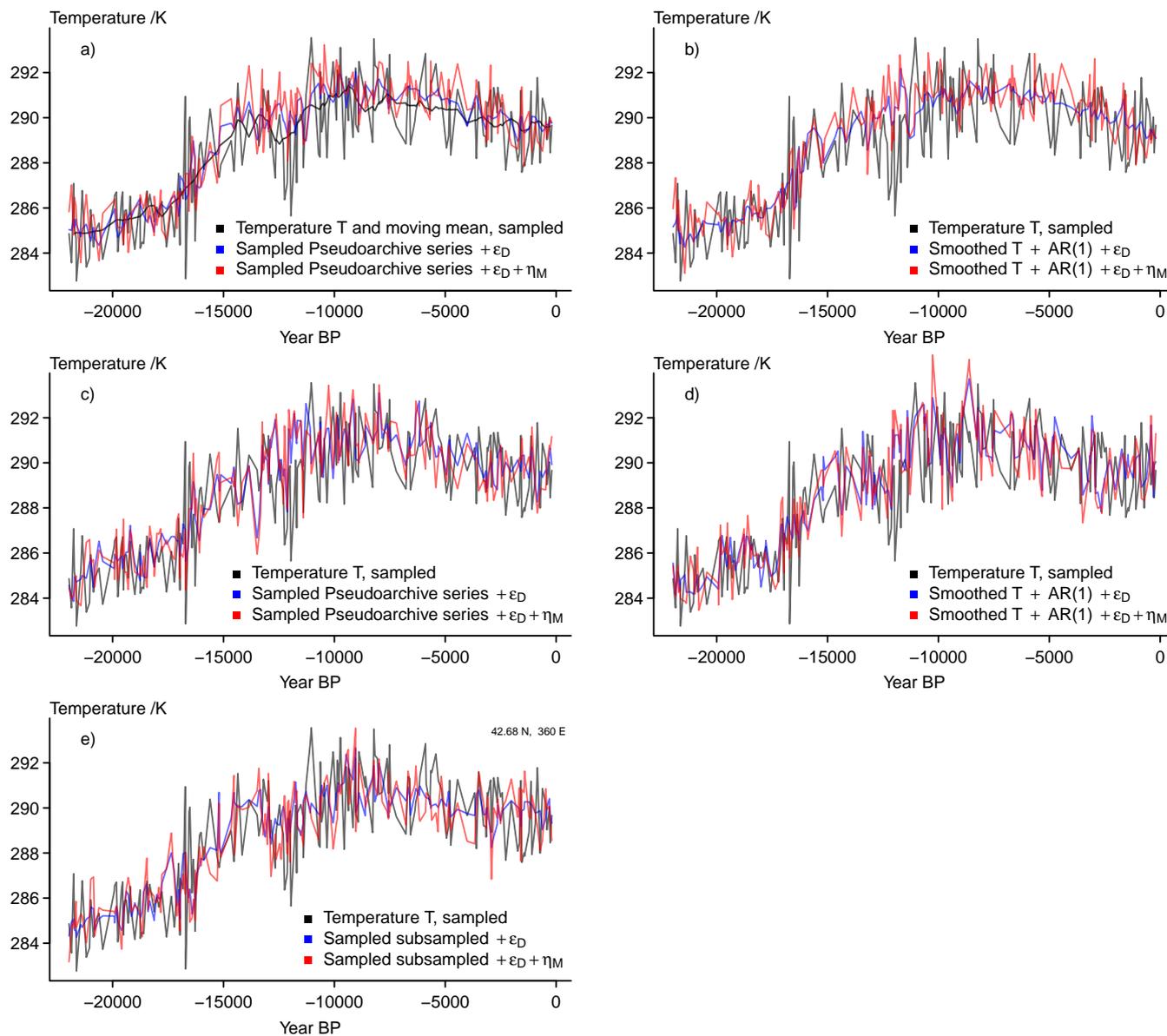
Figure 4. Visualising the sampled records: a) Input data and its 501-year moving mean, the pseudo-archive series with longer average smoothing lengths plus the effective dating error and plus the effective dating error and measurement noise, b) input data and its 501-year moving mean, the constantly smoothed record with longer smoothing length plus AR series with added effective dating error and with added effective dating error and measurement noise, c) input data and its 501-year moving mean, the pseudo-archive series with shorter average smoothing lengths plus the effective dating error and plus the effective dating error and measurement noise, d) input data and its 501-year moving mean, the constantly smoothed record with shorter smoothing length plus AR series with added effective dating error and with added effective dating error and measurement noise, e) input data and its 501-year moving mean, the subsampled record with added effective dating error and with added effective dating error and measurement noise.

3.3.3 Dating uncertainty

Dating uncertainty represents a big part of our overall uncertainty for many proxies, especially for sedimentary proxy-records. In our framework, already the smoothing function redistributes information from one date across the archive. Usually one considers this temporal uncertainty separately from the proxy-record uncertainty. For assessing reconstruction methods and simulations, it, however, would be beneficial to be able to include dating uncertainty within the proxy-uncertainty. In the following we have to distinguish between the dating uncertainty, i.e. the uncertainty that a sample is from a certain date, and the dating uncertainty error, by which we mean the potential error in our (pseudo)proxy due to the uncertain dating.

There are a number of approaches to transfer the dating uncertainty towards the proxy-record-uncertainty (e.g., Breitenbach et al., 2012; Goswami et al., 2014; Boers et al., 2017). We want to capture the error in a time-series. Thus, we take a very simple approach, which assumes that the error due to dating uncertainties is related to the climate state over the period of the dating uncertainty.

The code includes several variations of doing this. These reflect different amounts of dependence between subsequent samples. The following general approach is common to all. We only consider dependence between two subsequent samples while for real proxies the correlations may extend across larger portions of the proxy-record.

We proceed as follows: First, we sample random dating uncertainties in time for each sample date. We take these as dating uncertainty standard deviations. Then we take the effective dating uncertainty error at each sample date/depth to be a random sample from a normal distribution.

The mean of this distribution is the difference between the sample-data and the mean over the data within plus and minus two dating uncertainty standard deviations. The standard deviation of the distribution is the standard deviation of the differences between the individual data points within this interval and this mean.

The effective dating error is then

$$\epsilon_D = \mathcal{N}(\overline{P_{T_D}}, \sigma_D^2) \tag{12}$$

where

$$\overline{P_{T_D}} = \overline{P_T(t_S = \{t_{i-2\sigma_{dating}}, ..., t_i, ..., t_{i+2\sigma_{dating}}\})} - P_T(t = t_i) \tag{13}$$

is the mean over the region of influence and

$$\sigma_D^2 = E[(P_T(t_S) - \overline{P_{T_D}})^2] \tag{14}$$

is the variance of the distribution.

In the simplest formulation ignoring the dependence between subsequent dates, the sampled pseudoproxies become

$$P_{P_T}(t_1, ..., t_N) = g(T(t) + \eta_i(t) + Bias(t), t)(t_1, ..., t_N) + \epsilon_D(t_1, ..., t_N) \tag{15}$$

Alternative formulations of the pseudoproxy become

$$P_{P_T}(t_1, ..., t_N) = g(T(t) + \eta_i(t) + Bias(t))(t_1, ..., t_N) + AR_i(t_1, ..., t_N) + \epsilon_D(t_1, ..., t_N) \tag{16}$$

or

$$P_{P_T}(t_1,...,t_N) = h(T(t),t)(t_1,...,t_N) + h(\eta_i(t),t)(t_1,...,t_N) + \eta_s(t_1,...,t_N) + \epsilon_D(t_1,...,t_N) \tag{17}$$

This initial formulation of the effective dating uncertainty error ignores potential correlation between the dating errors. The most simple way to account for this makes subsequent errors dependent

$$\epsilon_{D_i} = \rho \cdot (\epsilon_{\xi_{D_{i-1}}} + (P_{P_{T_{i-1}}} - P_{P_{T_i}})) + \epsilon_{\xi_{D_i}} \tag{18}$$

This formulation has only a minor influence on the results. It is included in the code via a binary switch.

A slightly more complex formulation makes the error term at each date dependent on the previous sample's age uncertainties and mean data. Previous refers to archive units instead of time units. Then the dating error becomes

$$\epsilon_{D_i} = \rho \cdot (\epsilon_{D_{i-1}} + (P_{P_{T_{i-1}}} - P_{P_{T_i}})) + \epsilon_{\xi_{D_i}} \tag{19}$$

where $\rho = 0.9$ in our code and $\epsilon_{\xi_{D_i}}$ are the random innovations for date i. $\rho = 0.9$ can give large effective dating uncertainty errors. A switch in the code allows to use this inter-dependent error.

Another switch allows to consider the dependence between samples as a function of their dates and the dating uncertainty,

$$\rho(t) = 1 - (t_i - t_{i-1})/(2 \cdot \sigma_d(i-1)) \tag{20}$$

The time-dependent dating uncertainty for each date $\sigma_d(t)$ is generated randomly (compare above σ_D). We provide data for this case with a time-dependent $\rho(t)$.

Alternative simple formulations may include different noise processes, e.g., noise generated from Gamma-distributions. Furthermore, the available smoothing interval data could inform the dating uncertainty. We could also use this information to provide a deterministic, i.e. not random, error for each sampled date, i.e. taking a bias based on all dates influencing the selected date within the dating uncertainty.

In our current setup the age uncertainty does not depend on the measurement noise. The measurement error is added afterwards to the series including the effective dating uncertainty error. This decision is arbitrary. On the one hand a classical dating uncertainty affects the measured value. Then, also P_{P_T} above should already include the measurement error. On the other hand, the dating uncertainty affects the archived values independent of the measurement noise. Therefore we keep both independent and do not provide a dataset for the dependent case or code-switches.

The measured proxy-series becomes

$$M_T = P_{P_T} + \eta_M \tag{21}$$

The final proxy is in temperature units as is the initial input data. We ignore a separate term for potentially non-linear and climate-state dependent errors in our calibration relationship and assume the measurement noise term accounts for this as well. A separate term could be again a state-dependent Gaussian noise. It could also be a noise from a skewed distribution whose, e.g., mode depends on the background climate. On the other hand, a state-dependent bias term could simulate a mis-specified calibration relation while a time-dependent bias term could simulate a degenerative effect over time within the archived series.

3.3.4 Results

Figure 3 show versions of an archived proxy plus interannual measurement noise, i.e. they give an impression how a proxy would look from measurements on a perfectly annually sampled archive. The final amplitude of the noisy proxy is smaller than the amplitude of the interannual variations for the chosen location for all three versions, simple smoothing, smoothing plus AR (Figure 3a), different smoothing and different smoothing plus AR (Figure 3b), and seasonally subsampled (Figure 3c). This may be different at other locations. The final version generally preserves previously included biases.

Figure 4 presents a number of series sampled at $N = 200$ dates. All panels include the original temperature data sampled at these 200 dates. The Figure emphasizes how the initial temperature variability at the chosen grid-point is larger than any of our uncertainty estimates.

Our effective dating uncertainty error seldom results in large deviations from the archived record. The subsequently applied measurement error also only seldom leads to large offsets compared to either the original data or the effectively date uncertain record.

Thus, for our chosen parameter settings and the shown grid-point, the pseudoproxies fall within the range of the initial estimates. In turn, if we assume we have reliable calibration relationships, our calibrated proxy-series should also be reliable estimates of the past states.

Nevertheless, the biased estimates occasionally are only bad matches for the original data. This is also the case, but less so, for the subsampled data where we did not include the bias. Comparing the sampled pseudoproxy series to the smoothed original temperature data (compare Figure 4a) highlights that estimates for past climates may well fall within the range of the original interannual temperature variability but strongly misrepresent the mean climate represented by the sample.

Considering the effective dating uncertainty error, the discrepancies between input data and pseudoproxy are rather small for uncorrelated or weakly correlated age uncertainties. However, in the case of strong dependencies between subsequent data, pronounced biases and mismatches may occur (not shown). The assumed co-relation between two dates has a strong influence on the size of these mismatches. We show the case for a time-dependent co-relation between subsequent dates, which gives intermediately sized mismatches.

3.4 General Results

Figures 1 to 3 present the different versions of the pseudoproxies for the chosen location. Under our assumptions the influence of the orbital bias term is notable. The approaches using time-dependent smoothing or simple smoothing plus an AR-process may nearly or fully cancel the bias. This effect is less prominent for the time-dependent filter. Generally, both approaches seem to have similar effects.

Figure 3 includes the effect when we hypothetically add measurement noise at every date. Under our assumptions this noise is still smaller than the original interannual variability but, including biases, mean estimates may be close to the edge of the interannual variability of the original data. In these examples, the variability of the subsampled proxies is comparable to the smoothed ones after a measurement error is added. It is interesting to note that for the smaller smoothing the AR seems to
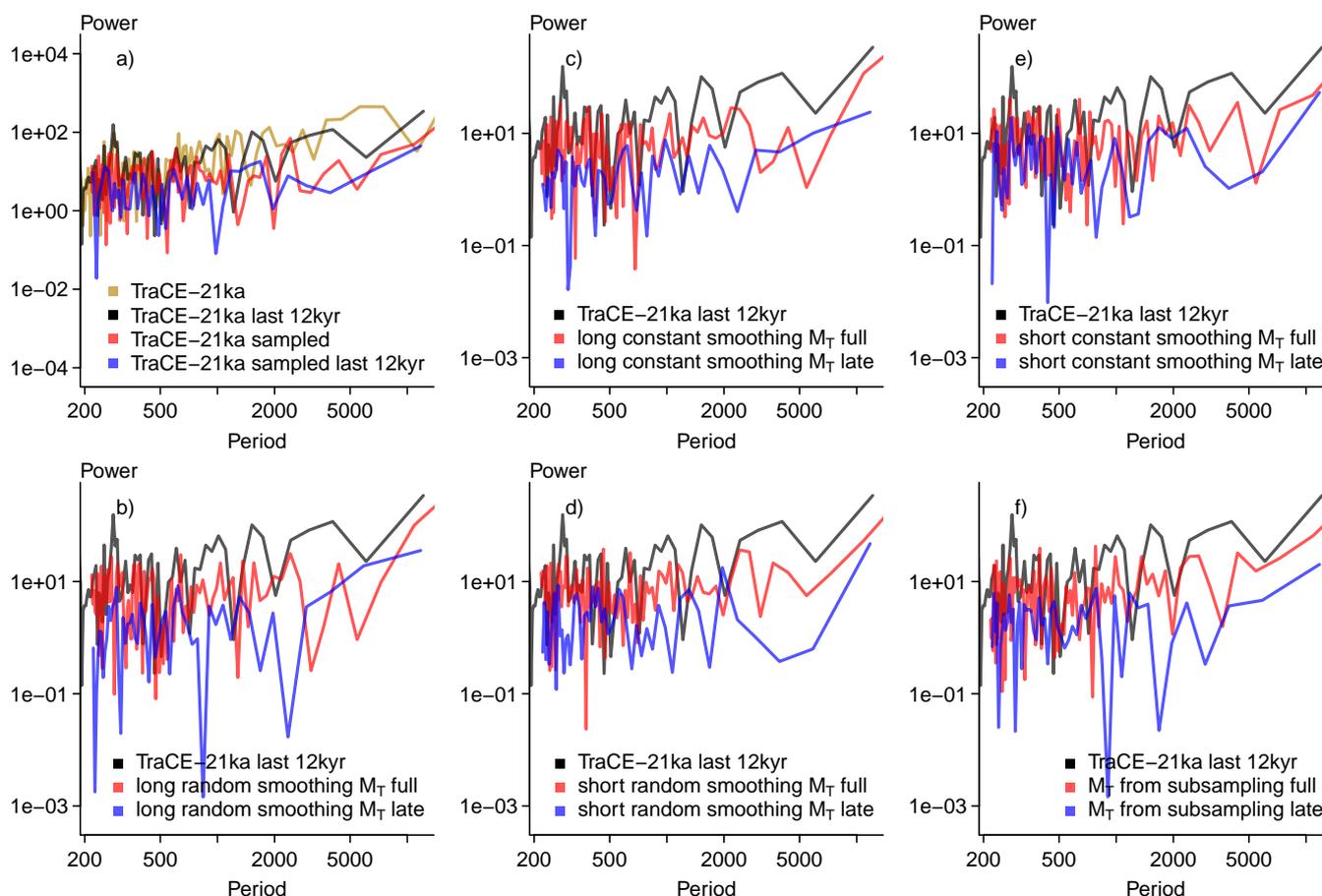
Figure 5. Lomb-Scargle periodograms of selected records split up by first 10k years of the records and the last 12k years of the records. All panels include the late input data from the TraCE-21ka simulation as black lines, red lines are in all panels for a full period record, blue lines are in all panels for the last 12k years of the version of a pseudoproxy. In addition to the input data from the TraCE-21ka simulation the panels show: a) the sampled TraCE-21ka simulation input data, b) the sampled pseudoarchive-series with long average smoothing plus the effective dating error and the measurement noise (long random smoothing M_T), c) the constantly smoothed record with a longer smoothing plus an AR(1)-process and including the effective dating error and the measurement noise (long constant smoothing M_T), d) the sampled pseudoarchive-series with short average smoothing plus the effective dating error and the measurement noise (short random smoothing M_T), e) the constantly smoothed record with a shorter smoothing plus an AR(1)-process and including the effective dating error and the measurement noise (short constant smoothing M_T), f) the subsampled data plus the effective dating error and the measurement noise (M_T from subsampling).

cancel the orbital bias more strongly in Figure 2. Figure 4 shows the data-sets sampled at $N = 200$ dates. It clarifies the error described for the interannual data. The document assets provide equivalent visualisations for another grid-point. These generally confirm the above descriptions.

3.4.1 Spectral power

Figure 5 adds a comparison of non-normalised Lomb-Scargle spectral power estimates. The Lomb-Scargle periodogram allows estimating the spectral power for records with uneven sampling. The calculation uses code based on the R (R Core Team, 2017) package lomb (Ruf, 1999). The package follows Press et al. (1994) and normalises the spectra by dividing them by two times the variance of the data. We omit this normalisation here.

The Figure shows estimates for the full records and for the data of the last twelve thousand years of the records. Spectra for the original and subsampled temperature data in Figure 5a highlight that the differentiation between full and late records does not result in large differences if we consider interannual data, and differences are also not too large if we consider the subsampled data except possibly for very long periods. While differences are larger for the subsampled uncertain final pseudoproxies in subsequent panels of Figure 5 the Figure suggests that the sample spectral estimates are rather similar. The equivalent Figure for another grid-point in the document assets shows larger differences over the full period since the record shows large millennial variability in the early part of the time series. The sampled data does not capture this millennial scale variability.

3.4.2 Global data

The supplementary assets for this manuscript include plots of selected series from our analyses at all grid-points starting from the south towards the north in supplementary document 1 Figure 1 (https://osf.io/zbehx/). These series are the input data at the grid-point, the smoothed-plus-AR-process series at the grid-point, and its subsampled version including all uncertainties.

These plots highlight three main points. First, the specific forcing implementation of He (2011; see also Liu et al., 2009) for the TraCE-21ka simulation results in occasionally spurious peaks and troughs for some locations. Furthermore there is potentially unrealistic variability at some grid-points for some periods. Second, as all our time-series are for the averages over the boreal summer season June, July, and August, our bias term does not show any large influence in high latitudes of the southern hemisphere. However, its effect is also not too large in the northern high latitudes. Third, our noise model shows generally the largest effect in the mid latitudes and the tropics. However, there is also a longitudinal dependence.

Supplementary document 2 Figure 1 (https://osf.io/zbehx/) emphasizes the regional differences in the long term climate evolutions by selecting only grid-points in equal intervals to provide a more intuitive view of the globe. Similarly Supplementary document 2 Figure 2 adds for a small selection of grid-points scatter plots of the pseudoproxy on the y-axis against the original data on the x-axis highlighting the common lack of a clear relation besides the deglaciation.

Figure 6 gives correlation coefficients between the sampled original grid-point data and the pseudoproxies including all uncertainties for the strong smoothing plus AR. The four panels show correlations for those samples within the first, second, and third 5,000 year chunks of the original data, and those samples in the remaining years. We choose to present the data
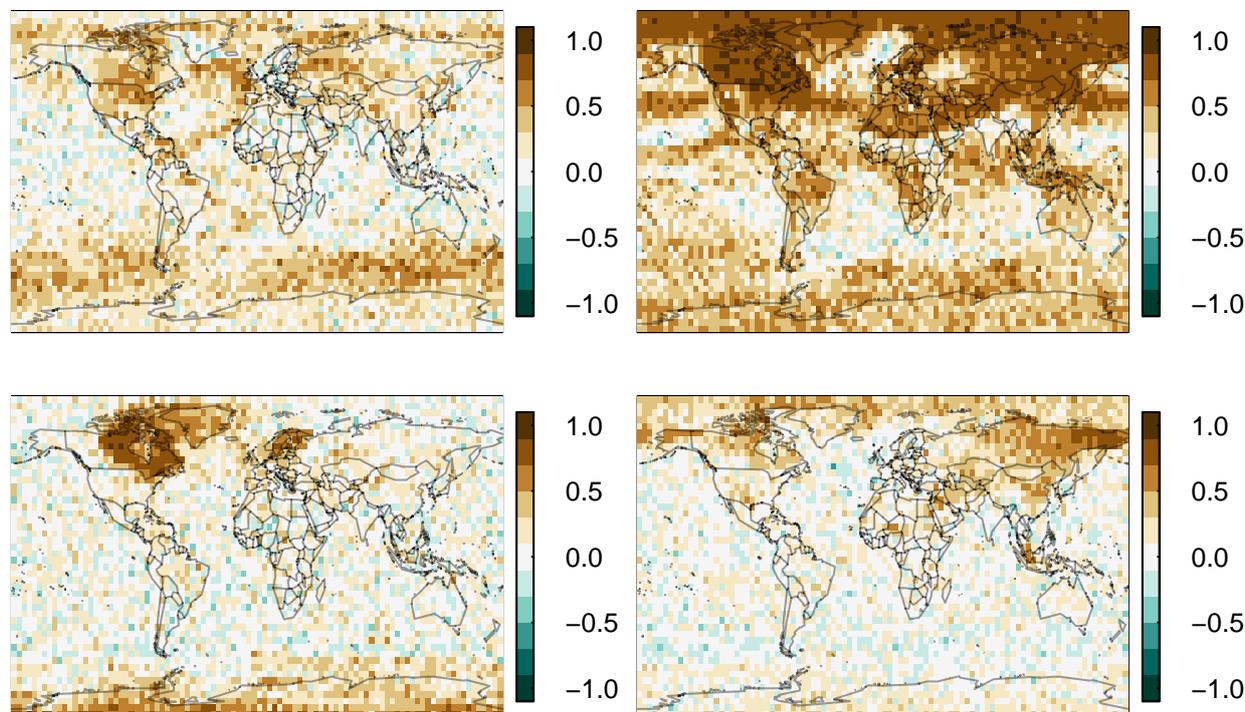
Figure 6. Point by point correlation maps between input data and the smoothed record plus AR(1)-process plus effective dating error and measurement noise for the sample dates within the first (top left), second (top right) and third (bottom left) subsequent 5,000 year windows of the record and the samples within the remaining years (bottom right).

this way to avoid detrending the data over the deglaciation interval. Relations between original data and pseudoproxies are generally weakest in the tropical belt. Except for the deglaciation in the top right panel they are also often weak in the high latitudes. Correlations are largest in later periods in areas with glacial remnants.

Figure 7 adds for the first and last period the logarithm of the relative standard-deviation σ_{T21k}/σ_P in the top row and the
10   bias $\bar{T}_{T21k} - \bar{T}_P$ in the bottom row. T21k refers to the simulation, P to the pseudoproxies. Variability is generally larger in the pseudoproxies in the tropics and the southern hemisphere midlatitudes but not elsewhere. The bias is largest over the southern oceans where the pseudoproxies may be up to 2K warmer than the original data in the early period.

3.5 Generalization of the errors

While we already chose comparatively simple procedures for our approach to obtain pseudoproxies from a model simulation, it is likely possible to simplify these even more. Such a general expression for the error in proxies over multi-millennial time-scales may be more usable in a number of ad-hoc model-data comparisons. Most importantly, such a generalized approach also
5   allows to quickly produce ensembles of pseudoproxies.
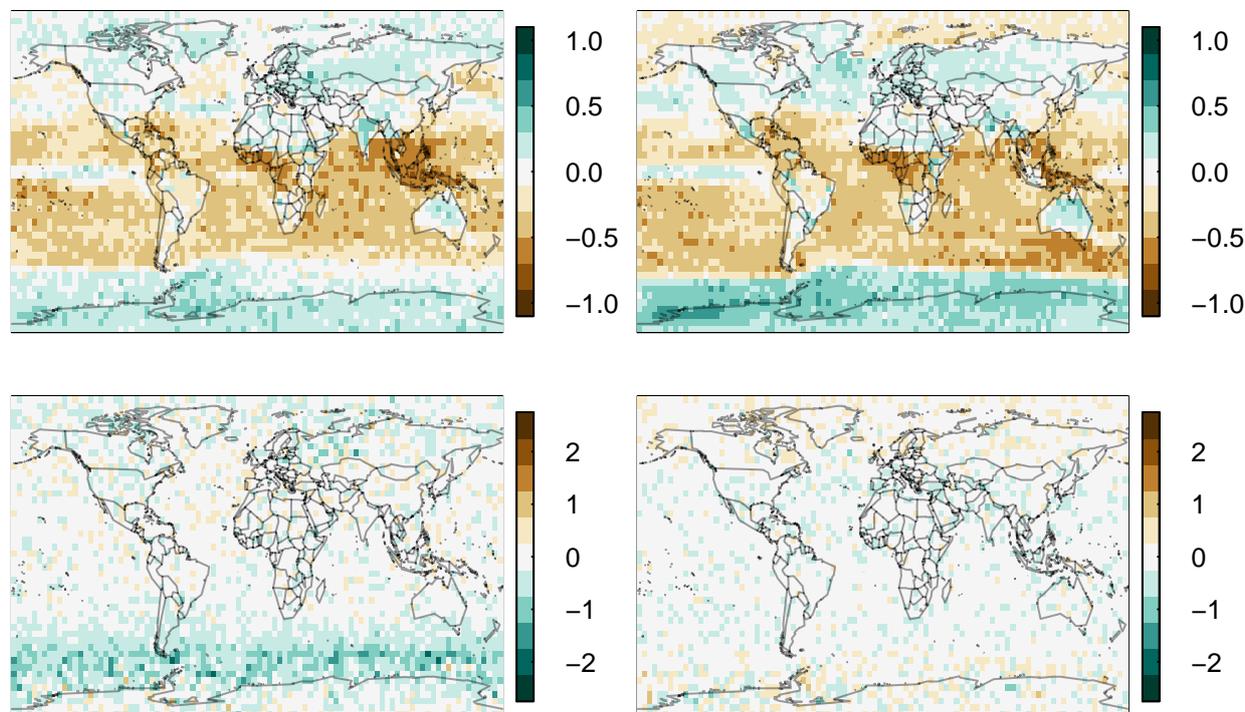
Figure 7. Top, logarithm of standard deviation ratios of the sampled input data relative to the smoothed record plus AR(1)-process and the effective dating error and the measurement noise for the samples in the first 5,000 years of record on the left and the last 7,040 years of the record on the right. Bottom, differences between the mean of the sampled input data and the mean of the smoothed record plus AR(1)-process and the effective dating error and the measurement noise for the samples in the first 5,000 years of record on the left and the last 7040 years of the record on the right.

The easiest way to obtain such a generalized error-model would be to assume a simple, potentially correlated noise model for the sensitivity of the sensor to the environment. While above we use white noise, here we indeed include an AR1-process with AR-coefficient $\phi = 0.7$. Either here or later one scales the series or adds a bias term to account for changing seasonality over multi-millennial time-scales.

10   The sum of the input data and this error are then subject to a simple moving averaging function. On top of this another simple correlated noise process mimics that the redistribution in the archive is not constant in time.

Another random component accounts for the measurement error. Thus, simple correlated noise may be enough to catch the essence of the error.

Nevertheless a full process-based approach is likely better to fully account for potential effects of biology, environmental long-term changes, orbital changes and other weakly constrained uncertainties. Such a full approach further allows for real

5   non-linearities between the climate and sensor and thus a truly non-linear pseudoproxy.
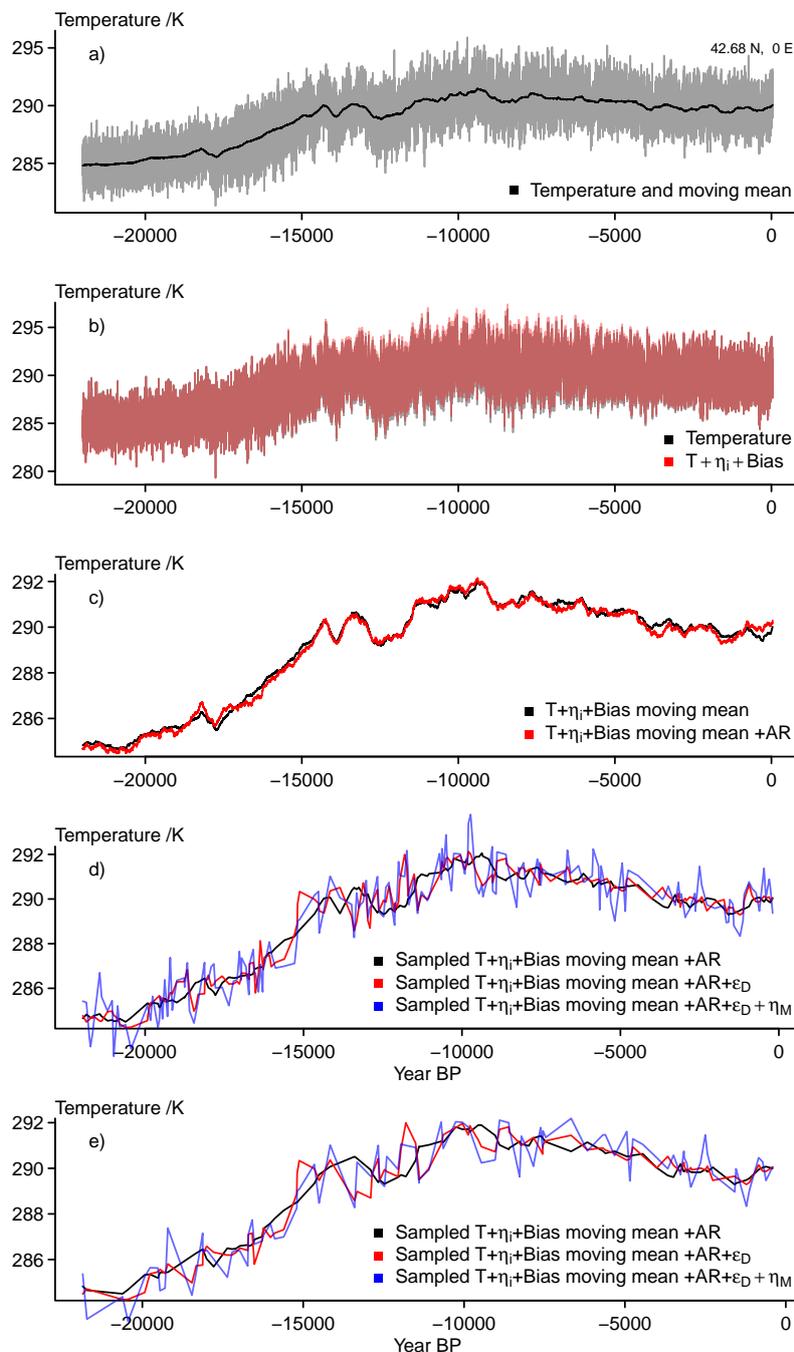
Figure 8. Visualising the simplified essence of the surrogate proxy calculations: a) input data and 501-point moving mean, b) input data plus initial noise and bias term, c) moving mean of input data plus noise plus bias and the same record plus an AR(1)-process, d) smoothed temperature plus noise plus bias plus AR-process sampled at 200 dates, this record plus the effective dating error, and this record plus the effective dating error and measurement noise, e) smoothed temperature plus noise plus bias plus AR-process sampled at 100 dates, this record plus the effective dating error, and this record plus the effective dating error and measurement noise.
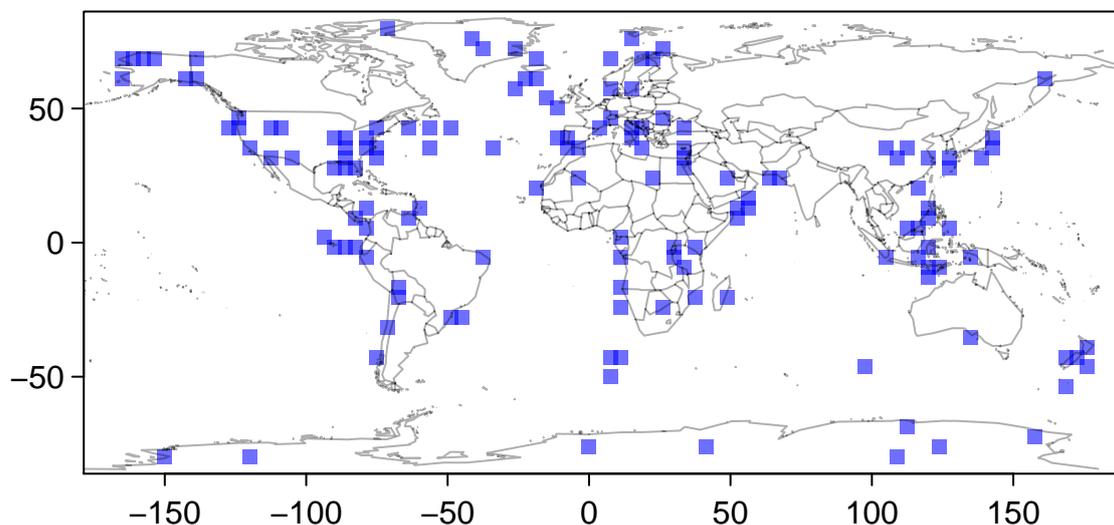
Figure 9. Map of the locations for the ensemble of surrogate proxies.

Figure 8 summarises results for our generalized approach. It clarifies that while an error may mask certain features of the past climate evolution, this simple generalized pseudoproxy-generation is unlikely to distort the proxy completely assuming the assumptions made above are approximately appropriate. Interestingly, the generalization appears to modify the input signal slightly less than the more complex approach. However, as we display slightly different data comparisons here, it is more appropriate to note that the dating uncertainty has only a minor effect compared to the initial bias and AR-process modifications and compared to the subsequent addition of the measurement noise. A global analysis of correlations and variability is hardly to distinguish from the maps presented for the more complex approach in Figures 6 and 7.

3.5.1 Ensemble of Pseudoproxies

We make further slight modifications to our approach to obtain an ensemble of 500 pseudoproxies at 144 locations. These locations are the grid-points which are close to proxies either included in Shakun et al. (2012), Clark et al. (2012) or Marcott et al. (2013). Figure 9 shows the locations.

Modifications to the code are, for one, we use a number of parameters' values sampled from either uniform distributions around the otherwise fixed value or from a list of values. Second, we consider the series S for bias and moving standard deviation as S^u where we sample u from $U = \{-1, 1\}$. The supplementary example code at https://osf.io/zbehx/ highlights these differences.

For Figure 10 we select twenty locations to represent the locally diverse representations of the climate in TraCE-21ka and how the ensemble of pseudoproxies modifies this. The Figure provides an impression of the range of the local ensembles and of two random ensemble members around the original temperature series. The diversity of the local climates in TraCE-21ka carries over to individual pseudoproxies and their ensembles. Besides this, Figure 10 mainly reflects the results of previous Sections
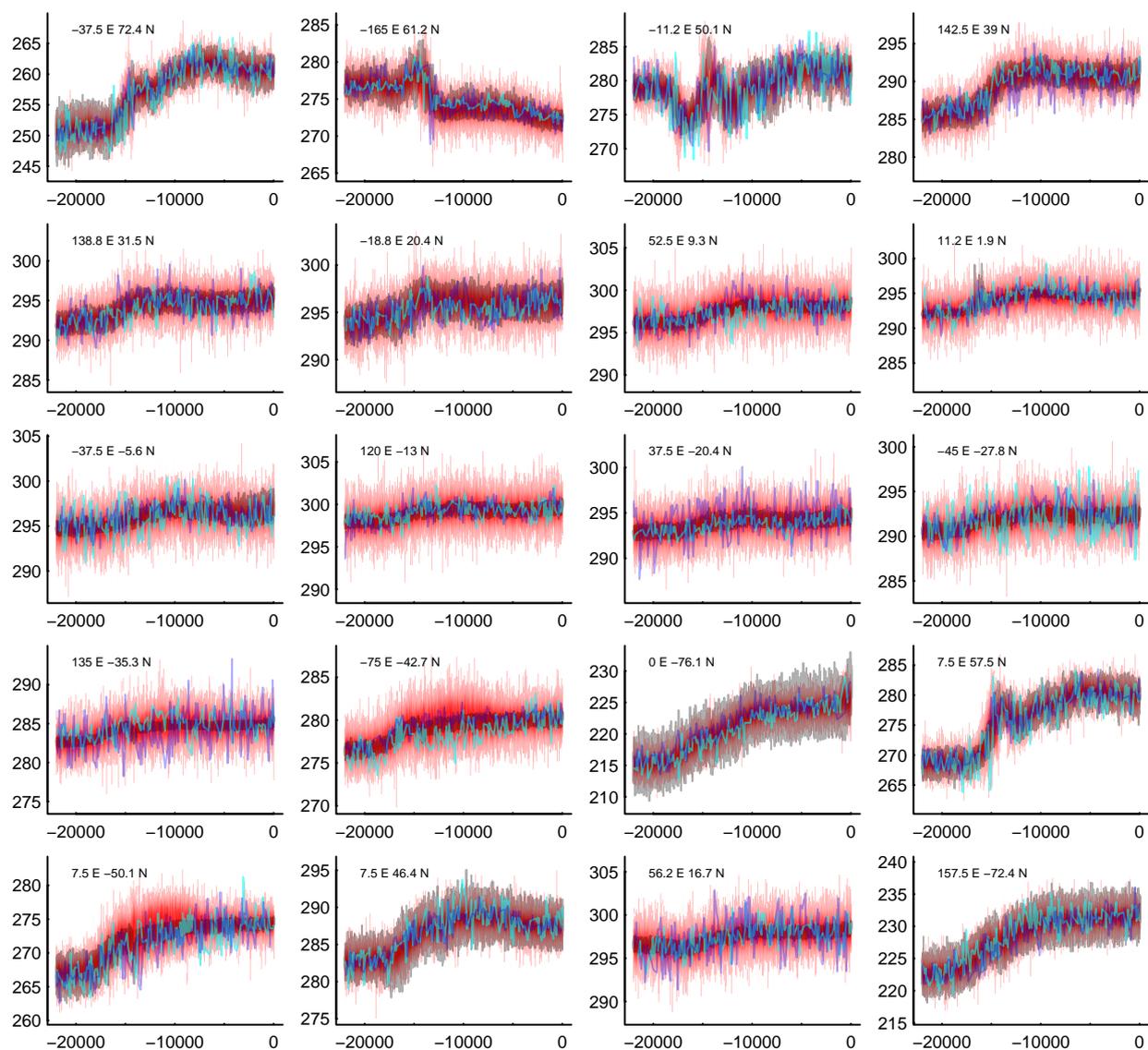
Figure 10. Visualising the surrogate proxy-ensemble at selected locations (Longitude and Latitude in top right corners of panels): Input data is plotted as grey lines, the range of the ensemble is transparent red shaded, and blue and cyan lines are two random members of the ensemble.

regarding how constrained our pseudoproxies are. However, we see commonly pseudoproxies and ensembles exceeding the variability of the original temperature data, not least because of our modifications to the selection of parameters and the orientation of the bias about its mean.

3.6 Provided Data

Tables 1 to 3 detail the provided data files. All files are in netcdf-format. These are generally gridded files on the original TraCE-21ka grid.

Only the ensembles of pseudoproxies are provided at their respective individual grid-points.

4 Conclusions

We present in this document, the associated code, and the provided data a flexible yet simple approach for describing the non-climatic error in proxy-records over multi-millennial time-scales. The assumptions are simplistic but base upon similar assumptions for proxy-system forward models.

The approach can be easily extended to compute ensembles of proxies for single locations. We chose to give one set of pseudoproxies for each grid-point of the Trace-21ka simulation and an ensemble of pseudoproxies at locations close to real proxy-locations. This simulation has a specific climatology (Liu et al., 2009) but a comparison to real proxy data may easily be achieved by only considering anomalies (as done, e.g., by Marsicek et al., 2018).

We choose only one possible set of parameters in our pseudoproxy-model, but we sample around this set for the ensemble of pseudoproxies. We choose these specific parameters to provide some disturbance to the data but not to get anywhere too far away from the original state. For example, it is quite likely that we have to deal with larger biases in reality than represented by our choice. Every researcher should make his own choice of parameters according to his assumptions on the various noise-contributions.

One can easily extend the chosen approach to even longer time-scales. Some modifications may be advisable considering the dating uncertainty to account for the likely sparser data further back in time, to better accommodate the increasing uncertainty, and especially to be more realistic in considering an effective dating uncertainty error for the pseudoproxy data. Similarly, we do not consider spatial correlations in the noise generally and between different locations. Such correlations are probably relevant for some noise-terms while they are probably less important for others.

We focused on the time-series approach and did not choose a probabilistic approach like, e.g. Breitenbach et al. (2012) or Goswami et al. (2014). There are a variety of other potential approaches how to obtain simple pseudoproxies from the model data. One such example would be to consider an envelope around the model state, select randomly a set of dates from the original data, fit a smooth through this set and then sample again around this uncertain smoothing. Similarly, Gaussian Process Models or Generalized Additive Models may be valuable means in producing pseudoproxies for paleoclimate studies over time-scales longer than the Common Era of the last 2,000 years. For example, Simpson (2018) shows the benefits of Generalized Additive Models for studies on paleoenvironmental time series.

Table 1. List of files provided, the variables included, their description, the category (full surrogate proxy field, essence field, or ensemble), and size of the ensemble. All files have the same stem Bothe_Trace21k_Pseudo_Proxies_

Bothe_Trace21k_Pseudo_Proxies_

Filename	Variable-Name	Variable-Description	Category	Grid size	Ensemble size
noise.save.nc	noise.save	initial environmental noise	field	96x48	1
bias.noise.data.save.nc	bias.noise.data.save	data + noise + bias	field	96x48	1
smooth.save.nc	smooth.save	smoothed data + noise + bias	field	96x48	1
meas.noise.smooth.save.nc	meas.noise.smooth.save	smoothed data + noise + bias plus measurement noise	field	96x48	1
ar.smooth.save.nc	ar.smooth.save	constantly smoothed plus AR-process	field	96x48	1
meas.noise.ar.smooth.save.nc	meas.noise.ar.smooth.save	constantly smoothed plus AR plus measurement noise	field	96x48	1
short.smooth.save.nc	short.smooth.save	smoothed date + noise + bias for shorter smoothing	field	96x48	1
meas.noise.short.smooth.save.nc	meas.noise.short.smooth.save	smoothed data + noise + bias plus measurement noise for shorter smoothing	field	96x48	1
short.ar.smooth.save.nc	short.ar.smooth.save	constantly smoothed plus AR-process for shorter smoothing	field	96x48	1
meas.noise.short.ar.smooth.save.nc	meas.noise.short.ar.smooth.save	constantly smoothed plus AR plus measurement noise for shorter smoothing	field	96x48	1
subsampled.save.nc	subsampled.save	seasonally subsampled data + initial noise	field	96x48	1
meas.noise.subsampled.save.nc	meas.noise.subsampled.save	seasonally subsampled + noise plus measurement noise	field	96x48	1

Table 2. Continued list of files provided, the variables included, their description, the category (full surrogate proxy field, essence field, or ensemble), and size of the ensemble. All files have the same stem Bothe_Trace21k_Pseudo_Proxies_

Filename	Variable-Name	Variable-Description	Category	Grid size	Ensemble size
sampled.nc	samp.subsampled.save, samp.meas.noise.smooth.save, samp.input.save, samp.noise.save, samp.bias.noise.data.save, samp.ar.smooth.save, samp.smooth.save, samp.short.smooth.save, samp.short.ar.smooth.save, samp.meas.noise.short.smooth.save, samp.dates.save, samp.meas.noise.ar.smooth.save, samp.meas.noise.short.ar.smooth.save	sampled versions of the various variables and the dates of the samples	field	96x48	1
dating-error.nc	unc.date.samp, unc.date.subsampled.save, unc.date.meas.noise.smooth.save, unc.date.noise.save, unc.date.bias.noise.data.save, unc.date.ar.smooth.save, unc.date.smooth.save, unc.date.short.smooth.save, unc.date.short.ar.smooth.save, unc.date.meas.noise.short.smooth.save, unc.date.meas.noise.ar.smooth.save, unc.date.meas.noise.short.ar.smooth.save	date uncertain versions of the various variables and the dating uncertainties	field	96x48	1

Earth System Science Data Discussions — Open Access

Table 3. Continued list of files provided, the variables included, their description, the category (full surrogate proxy field, essence field, or ensemble), and size of the ensemble. All files have the same stem Bothe_Trace21k_Pseudo_Proxies_

Filename	Variable-Name	Variable-Description	Category	Grid size	Ensemble size
samp.meas.noise.subsampled.save.nc	samp.meas.noise.subsampled.save	sampled seasonally subsampled + noise plus measurement noise	field	96x48	1
unc.date.meas.noise.subsampled.save.nc	unc.samp.meas.noise.subsampled.save	sampled seasonally subsampled + noise plus measurement noise plus effective dating error	field	96x48	1
Essence_gen.noise.env.nc	gen.noise.env	generalized environmental noise term	essence	96x48	1
Essence_noise.gen.dat.nc	noise.gen.dat	input data + generalized envrionmental noise	essence	96x48	1
Essence_bias.noise.gen.dat.nc	bias.noise.gen.dat	input data + generalized noise + bias term	essence	96x48	1
Essence_smooth.bias.noise.gen.dat.nc	smooth.bias.noise.gen.dat	smoothed input + noise + bias	essence	96x48	1
Essence_ar.smooth.bias.noise.gen.dat.nc	ar.smooth.bias.noise.gen.dat	smoothed input + noise + bias plus AR-process	essence	96x48	1
Essence_sampled.nc	sampled ar.smooth.bias.noise.gen.dat	sampled smoothed input + noise + bias plus AR-process	essence	96x48	1
Essence_uncertain-sampled.nc	unc.samp.ar.smooth.bias.noise.gen.dat, meas.unc.samp.ar.smooth.bias.noise.gen.dat, unc.date.samp.gen, samp.dates.save.gen	date uncertain versions of generalized data, generalized dating uncertainty, sample dates	essence	96x48	1
essence_ensemble.nc	Pseudoproxy, Dates, DateUncertainty	Surrogate proxy data, Dating, Uncertainty of Dating	ensemble	144	500

   The present approach ignores a variety of possible complications. For example, there is not so far a method to include hiatusses in the sensor. Furthermore, the dependency on the background climate is minimal. Nevertheless, we are confident

5   that this approach is of value for the comparison of simulation data and proxy data over long periods.

5   Code and data availability

The TraCE-21ka simulation data is available from www.cgd.ucar.edu/ccr/TraCE and was obtained via the Earth System Grid (www.earthsystemgrid.org/project/trace.html). Our results as described in section **??** are available from the Open Science Framework (OSF) at https://osf.io/zbehx/. There, one also finds example code for computing proxies for a single grid-point and for computing the ensemble at 144 locations.

Appendix A: Tables of parameters

Tables A1 to A3 summarise the considered parameters and noise models. They also clarify whether the parameters settings are used for a global field of surrogate proxies, a more generalized approach, an ensemble calculation, or all.

Table A1. List of parameters used.

Description	Parameter	Value	Category
Season limits for insolation bias	mon1.for.insol, mon2.for.insol	6, 8	all
Number of samples along the full record	n.samples	200	all
Scaling of initial noise amplitude	amp.noise.env	0.5	field, essence
model for initial noise	model.noise.1	c()	field, essence
standard deviation of innovations of initial noise	sd.noise.1	not used	field
Scaling of bias term	amp.bias.seas	4	field, essence
mean smoothing length for longer random smoothing	rand.mean.length.smooth	350	field, essence
standard deviation for longer random smoothing	rand.sd.length.smooth	75	field, essence
fixed longer smoothing length	fix.length.smooth	501	field, essence
Minimum allowed longer random smoothing length	min.rand.length.smooth	40	field, essence
AR-coefficient for added AR(1)-process	coeff.ar.smooth	0.999	field, essence
Standard deviation for the innovations	sd.ar.smooth	0.01	field, essence
mean smoothing length for longer shorter smoothing	rand.mean.length.smooth.2	31	field, essence
standard deviation for shorter random smoothing	rand.sd.length.smooth.2	5	field, essence
fixed shorter smoothing length	fix.length.smooth.2	31	field, essence
Minimum allowed shorter random smoothing length	min.rand.length.smooth.2	5	field, essence
AR-coefficient for added AR(1)-process	coeff.ar.smooth.2	0.9	field, essence
Standard deviation for the innovations	sd.ar.smooth.2	0.15	field, essence
number of picked samples for subsampling	n.samp.pick	30	field
standard deviation of innovations for subsampling noise	sd.noise.pick	0.5	field
model of subsampling noise	model.noise.pick	c()	field
1.96 Sigma of measurement-noise	lim.noise.meas	1.5	field, essence
noise model for measurement noise	model.noise.meas	c()	field, essence

Table A2. Continuation of list of parameters used.

Description	Parameter	Value	Category
noise model for measurement noise for subsampled record	model.seas.pick.noise.meas	c()	field
1.96 Sigma for measurement noise for subsampled record	lim.seas.pick.noise.meas	1.5	field
switch for correlated effective dating error	switch.cor.date.unc	1	all
switch for weakly correlated only	switch.weak.cor.date.unc	1	all
switch for time dependent correlated	switch.delta.cor.date.unc	1	all
fixed correlated dating error coefficient	cor.date.unc	0.9	all
mean of distribution of dating uncertainty	mean.date.unc	350	all
standard deviation of distribution of dating uncertainty	sd.date.unc	100	all
switch for proportionality of initial noise - not used	switch.orient.runsd.noise.env		field, essence
switch for orientation of bias	switch.orient.bias.seas	0	field, essence
model for initial noise for generalized case	coeff.gen.ar.smooth	0.999	essence
standard deviation for initial noise innovations, generalized case	sd.gen.ar.smooth	0.01	essence
smoothing length generalized case, prescribed		501	essence
alternative model for initial noise	model.gen.noise	c(0.7)	essence
alternative model for measurement noise - not used	model.noise.meas.b	c(0.35)	essence
alternative scaling of amplitude of initial noise - not used	amp.noise.env.2	0.25	essence

Table A3. Continuation of list of parameters used.

Description	Parameter	Value	Category
Ensemble size	size.ensemble	500	ensemble
amplitude of scaling of initial noise	amp.noise.env	$U(0.4, 1.5)$	ensemble
scaling of bias	amp.bias.seas	$U(3, 10)$	ensemble
standard deviation of measurement noise	lim.noise.meas	$U(0.75, 3)/1.959964$	ensemble
ar-coefficient of measurement noise model	rand.model.coeff	$U(0.3, 0.8)$	ensemble
ar-coefficient of initial noise model	rand.model.coeff.gen	$U(0.6, 0.8)$	ensemble
window of influence of background climate - not used	rand.width.background.sd	$U(500, 2000)$	ensemble
window of influence of background climate	rand.width.background.sd	1000	ensemble
width of window of filter influence	length.filter.uniform	l_{fil} is random sample from $L = \{301, 303, 305, ..., 1001\}$	all
orientation of bias		$sign_{orbit}$ is random sample from $S_O = \{-1, 1\}$	ensemble
proportionality of initial noise		$sign_{noise}$ is random sample from $S_N = \{-1, 1\}$	ensemble

Competing interests. The authors are not aware of any circumstances that one could see as conflicts of interest.

Acknowledgements. This work contributes to PalMod, German Climate Modeling Initiative: From the Last Interglacial to the Anthropocene
- Modeling a Complete Glacial Cycle. It is funded through the German Federal Ministry of Education and Research's (BMBF) Research for
5   Sustainability initiative (FONA).

References

Anchukaitis, K. J. and Tierney, J. E.: Identifying coherent spatiotemporal modes in time-uncertain proxy paleoclimate records, Climate Dynamics, 41, 1291–1306, https://doi.org/10.1007/s00382-012-1483-0, http://dx.doi.org/10.1007/s00382-012-1483-0, 2013.

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://www.clim-past.net/13/1169/2017/, 2017.

Bradley, R. S.: Chapter 1 – Paleoclimatic Reconstruction, pp. 1–11, Academic Press, https://doi.org/10.1016/B978-0-12-386913-5.00001-6, 2015.

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

Brierley, C. and Rehfeld, K.: Paleovariability: Data Model Comparisons, Past Global Changes Magazine, 22, 57–116, 2014.

Carré, M., Sachs, J. P., Wallace, J. M., and Favier, C.: Exploring errors in paleoclimate proxy reconstructions using Monte Carlo simulations: paleotemperature from mollusk and coral geochemistry, Climate of the Past, 8, 433–450, https://doi.org/10.5194/cp-8-433-2012, http://dx.doi.org/10.5194/cp-8-433-2012, 2012.

Clark, P. U., Shakun, J. D., Baker, P. A., Bartlein, P. J., Brewer, S., Brook, E., Carlson, A. E., Cheng, H., Kaufman, D. S., Liu, Z., Marchitto, T. M., Mix, A. C., Morrill, C., Otto-Bliesner, B. L., Pahnke, K., Russell, J. M., Whitlock, C., Adkins, J. F., Blois, J. L., Clark, J., Colman, S. M., Curry, W. B., Flower, B. P., He, F., Johnson, T. C., Lynch-Stieglitz, J., Markgraf, V., McManus, J., Mitrovica, J. X., Moreno, P. I., and Williams, J. W.: Global climate evolution during the last deglaciation., Proceedings of the National Academy of Sciences of the United States of America, 109, E1134–42, https://doi.org/10.1073/pnas.1116619109, 2012.

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

Crucifix, M.: palinsol: Insolation for Palaeoclimate Studies, https://CRAN.R-project.org/package=palinsol, r package version 0.93, 2016.

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, J. Adv. Model. Earth Syst., p. n/a, https://doi.org/10.1002/2015ms000447, http://dx.doi.org/10.1002/2015ms000447, 2015.

Dee, S. G., Russell, J. M., Morrill, C., Chen, Z., and Neary, A.: PRYSM v2.0 : A Proxy System Model for Lacustrine Archives, Paleoceanography and Paleoclimatology, https://doi.org/10.1029/2018PA003413, http://doi.wiley.com/10.1029/2018PA003413, 2018.

Dolman, A. M. and Laepple, T.: Sedproxy: a forward model for sediment archived climate proxies, Climate of the Past Discussions, pp. 1–31, https://doi.org/10.5194/cp-2018-13, https://www.clim-past-discuss.net/cp-2018-13/, 2018.

Evans, M. N., Tolwinski-Ward, S. E., Thompson, D. M., and Anchukaitis, K. J.: Applications of proxy system modeling in high resolution paleoclimatology, Quaternary Science Reviews, 76, 16–28, https://doi.org/10.1016/j.quascirev.2013.05.024, http://dx.doi.org/10.1016/j.quascirev.2013.05.024, 2013.

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

He, F.: Simulating Transient Climate Evolution of the Last Deglaciation with CCSM3, Ph.D. thesis, University of Wisconsin-Madison, http://www.cgd.ucar.edu/ccr/TraCE/doc/He_PhD_dissertation_UW_2011.pdf, 2011.

Kopp, R. E., Kemp, A. C., Bittermann, K., Horton, B. P., Donnelly, J. P., Gehrels, W. R., Hay, C. C., Mitrovica, J. X., Morrow, E. D., and Rahmstorf, S.: Temperature-driven global sea-level variability in the Common Era, Proceedings of the National Academy of Sciences,

10   113, E1434–E1441, https://doi.org/10.1073/pnas.1517056113, http://www.pnas.org/lookup/doi/10.1073/pnas.1517056113, 2016.

Laepple, T. and Huybers, P.: Reconciling discrepancies between Uk37 and Mg/Ca reconstructions of Holocene marine temperature variability, Earth and Planetary Science Letters, 375, 418–429, https://doi.org/10.1016/j.epsl.2013.06.006, http://www.sciencedirect.com/science/article/pii/S0012821X13003221, 2013.

Liu, Z., Otto-Bliesner, B. L., He, F., Brady, E. C., Tomas, R., Clark, P. U., Carlson, A. E., Lynch-Stieglitz, J., Curry, W., Brook, E., Erickson,

15   D., Jacob, R., Kutzbach, J., and Cheng, J.: Transient simulation of last deglaciation with a new mechanism for Bolling-Allerod warming., Science (New York, N.Y.), 325, 310–4, https://doi.org/10.1126/science.1171041, http://science.sciencemag.org/content/325/5938/310.abstract, 2009.

Mann, M. E. and Rutherford, S.: Climate reconstruction using 'Pseudoproxies', Geophysical Research Letters, 29, 1501+, https://doi.org/10.1029/2001gl014554, http://dx.doi.org/10.1029/2001gl014554, 2002.

20   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 (New York, N.Y.), 339, 1198–201, https://doi.org/10.1126/science.1228026, http://science.sciencemag.org/content/339/6124/1198.abstract, 2013.

Marsicek, J., Shuman, B. N., Bartlein, P. J., Shafer, S. L., and Brewer, S.: Reconciling divergent trends and millennial variations in Holocene temperatures, Nature, 554, 92–96, https://doi.org/10.1038/nature25464, http://www.nature.com/doifinder/10.1038/nature25464, 2018.

25   Press, W., Teukolsky, S., Vetterling, S., and Flannery, B.: Numerical Recipes in C: The Art of Scientific Computing, Cambridge University Press, Cambridge, Cambridge, 1994.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/, 2017.

Rehfeld, K. and Kurths, J.: Similarity estimators for irregular and age-uncertain time series, Climate of the Past, 10, 107–122,

30   https://doi.org/10.5194/cp-10-107-2014, http://dx.doi.org/10.5194/cp-10-107-2014, 2014.

Rehfeld, K., Münch, T., Ho, S. L., and Laepple, T.: Global patterns of declining temperature variability from the Last Glacial Maximum to the Holocene, Nature, https://doi.org/10.1038/nature25454, http://www.nature.com/doifinder/10.1038/nature25454, 2018.

Ruf, T.: The Lomb-Scargle Periodogram in Biological Rhythm Research: Analysis of Incomplete and Unequally Spaced Time-Series, Biological Rhythm Research, 30, 178–201, 1999.

35   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, https://doi.org/10.1038/nature10915,

565   http://dx.doi.org/10.1038/nature10915, 2012.

Simpson, G. L.: Modelling Palaeoecological Time Series Using Generalised Additive Models, Frontiers in Ecology and Evolution, 6, 149, https://doi.org/10.3389/fevo.2018.00149, https://www.frontiersin.org/article/10.3389/fevo.2018.00149/full, 2018.

Smerdon, J. E.: Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments, WIREs Clim Change, 3, 63–77, https://doi.org/10.1002/wcc.149, http://dx.doi.org/10.1002/wcc.149, 2012.