Response to referees ESSD-2018-137

28 March 2019

Dear referees, dear editor,

we would to like to thank you for your constructive and relevant comments on our manuscript.

In the following, we reiterate or modify our initial response and indicate the changes to the manuscript. Our answers are in red font color while referees' comments are in normal font. Where feasible, we indicate changes to the manuscript in blue.

We would like to thank you once more for your help.

On behalf of the authors,

Sincerely yours,

Oliver Bothe

List Of Changes:

Introduction

Added discussion on pseudoproxies

Extended introduction of proxy system models

Extended on differences between our approach and previous works

Data

We changed the presented grid-point in the main manuscript and the supplementary file.

We now consider annual data instead of summer data.

Considerations and Results

We added a flow chart of our procedure.

We changed our code and provide more options for the various steps of our procedure.

We changed some of the noise models to allow for (higher) autocorrelations.

Particularly we clarified the questions by the referees.

We changed the spectral power estimation to a wavelet procedure.

We now show standard deviation ratios instead of their logarithm.

We added a mathematical expression for the generalized approach.

We changed the visualisation of the ensemble data.

Conclusions

We slightly extended our outlook.

Generally

We updated the tables and the supplements.

Response to referees

Referee 1:

General Comments

In their manuscript, Bothe et al. provides a flexible approach to take into consideration noise in forward models of paleoclimate proxies, i.e. pseudoproxies. Although the need for these pseudo proxy models is increasing, I'm unclear as of how the present study adds to the growing body of methods available, including the recent SedProxy toolbox of Dolman and Laepple referenced by the author.

We appreciate the referee's concerns about the originality and usefulness of our manuscript, the chosen approach, and the data sets. We appreciate that the referee acknowledges that current research directions require and benefit from the development of pseudoproxy approaches.

Traditional pseudoproxy applications over the Common Era of the last 2,000 years started from, and mainly still rely on, simple noise-based pseudoproxies. Thus, while there is the need of complex mechanistically modelled pseudoproxies, the last 20 years have shown that paleoclimatology benefits from having access to multiple pseudproxy generating algorithms.

Pseudoproxies may help in understanding proxy systems, testing reconstruction methods, evaluating and comparing simulations among another and against proxies, and in testing data assimilation techniques.

Regarding the work by Dolman and Laepple (2018), we note that the present work developed parallel to their approach. From our point of view, our code and data complement their proxy system models from a more general point of view. Additionally, we would argue that we make differing or additional assumptions compared to Dolman and Laepple (2018).

There are some major differences to the approach of Dolman and Laepple:

- The level of complexity of the different approaches
- The generalization of our approach including non-marine proxy archives.

Generally, our approach can be seen as complementary to the one of Dolman and Laepple and used as an independent source for the generation of pseudoproxies.

We are going to argue in favor of the originality and the usefulness of our approach. Originality of research may refer to hypotheses reported, methods used, and results obtained (Dirk, 1999, www.jstor.org/stable/285800). Morgan (1985, www.ncbi.nlm.nih.gov/pmc/articles/PMC1346489) reformulates originality as "independent or creative in thought or action". Research usually does not start from zero but refers to previous work. Independence and novelty, thus, are always relative and depend on the context.

In the present case of a manuscript that provides a data set, the questions could be, following Dirk (1999), how new is the method, and how new are the data (which is the result in this case). Considering the chosen approach, we would argue that we obviously rely on previous thinking on paleoproxies and indeed part of the work relies on discussions with colleagues like Dolman, Laepple, and Weitzel, which we have to acknowledge. However, to our knowledge, there is no publication presenting a simple noise-based pseudoproxy for deglacial and longer time-scales. There is definitely no publication using our specific approach, and, to our knowledge, there is no publication providing pseudoproxies or even ensembles of pseudoproxies for full simulation output fields for easy usage.

Regarding the potential usefulness of our data, there are two things to consider. First, is our data readily usable? As referee 3 notes, this is apparently not necessarily the case and we have to try to improve on this. Second, is the data by itself of value? The data, as it is, can be used to test data assimilation methods as well as other reconstruction methods, to evaluate the TraCE-21ka simulation against old and new proxy data, and to test model-data comparison methods. Using our code allows to produce comparable proxies for other simulations and, e.g., additonally to compare these different simulations among another. Is our data worse or better suited than other pseudoproxies for these purposes? We provide an alternative, which we think worthwhile and usable.

I would not recommend the manuscript for publication in this present stage.

We appreciate the referee's comments and hope that our responses are sufficient to change this assessment.

In particular, I suggest the authors make the following points clearer in their revisions: The work seems to rely on the concept of proxy system introduced by Evans et al. (2013). A proxy system is composed of an archive, a sensor, and an observation (measurement in the present manuscript). Each components can be modeled independently to obtain the full proxy system model.

Indeed, our work relies on the concepts introduced by Evans et al. We agree, that a proxy system can be thought of as including (at least) a sensor, an archive and the observation.

If the referee implies that each has to be modelled separately, then we disagree. In our understanding, this would disqualify VSLite as proxy system model. Similarly, the approach of Thompson et al. (2011) would then not qualify either. If the referee means that it is possible to formulate models for each of the components of a proxy system in the sense of Evans et al., then we do not see their concern, as we precisely follow this approach.

In response to referee 3, we add a flowchart of our procedures, which hopefully clarifies the relation of our procedure to the conceptual ideas described by Evans et al. Furthermore, we more thoroughly introduce the concepts of a proxy system, of proxy system models, and of pseudoproxies.

Proxy systems model extend to age determination methods and I'm unsure as whether it was singled out in this manuscript.

We are unsure what the referee refers to. Thus, our following response may miss the point.

Our aim is to provide a pseudoproxy-setup that adds a noise-based error-term for the time-uncertainty to the discrete time-series of the pseudoproxy instead of providing a tuple of uncertainties for the tuple of time and data. Our rationale is that this helps in model-evaluation, model-data comparisons, and reconstruction exercises. In a sense, we single this out as we regard this to be important. In our understanding this is of importance for model-data comparison and evaluation of different model simulations.

Ideally, in order to fully represent the uncertainty in the proxy, one would want to use a proxy system model for the time axis (e,g, radiocarbon in foraminifera shells) and y-axis (e.g., Mg/Ca in foraminiferal shell). In this particular example, the archive and sensor model would be common to the x-axis and y-axis. The observation model would need to be tailored to the particular measurement.

We agree on the optimal proxy system representation. However, we explicitly aim at (i) a simplified representation and (ii) a representation that results in a time-series with associated errors where the error term also accounts for date uncertainty.

The code does not explicitly model the time axis. However, the sampling of dates and the assignment of uncertainties should allow in principle to apply an age modelling approach comparable to PRYSM 2.0 in Dee et al. (2018).

The back and forth between age uncertainties and environmental variable uncertainties in the manuscript is confusing.

We are sorry for this and we will try to make our points more clearly in the revised version of the manuscript by clarifying our terminology.

We tried to be as explicit as possible in distinguishing between the age uncertainties, the environmental uncertainties, and what we call the "effective dating uncertainty error.

I'm also unclear on how this model is fully generalizable since each type of observations made on this archive (e.g., Mg/Ca in foraminifera shells or UK37) would have specific "noise" associated with them which would need to be modeled individually.

Various sensors, various archives, and various observations, that is various proxies differ but also share common properties. In this sense we try to formulate a general model that assists in reconstruction method tests, model evaluation, and the evaluation of model-data comparisons.

We our confident, that the code for the various components is flexible enough to allow individual researchers to adapt the noise levels according to their understanding.

The authors keep referring to "non-climatic noise". Climatic noise is also included in the proxy records and is often impossible to disentangle from the other sources of noise discussed in the manuscript.

The referee is correct and we have to clarify this.

We tried to correct all instances of non-climatic noise.

Specific comments:

Introduction:

The concepts of proxy systems and proxy systems models need to be introduced earlier and a description of how the current work fits into these larger concepts need to be included.

We introduce proxy systems and proxy system models now earlier in the introduction and also relate our approach to this setting.

Please also include a discussion on how the present approach is different from the slew of studies on proxy system models and what it adds to the table.

We shortly discuss this more in the introduction.

Page 2, line 13: A proxy system is a mathematical representation of the proxy, including the error. How is this a second way. I'm also unsure how the noise is not observation specific?

We try to clarify our thinking as follows:

In our understanding there are various approaches to obtain pseudoproxies. These range from comprehensive to simplified. We can try to obtain a comprehensive representation from the environmental influences on a sensor to the measurement and implement this into a mechanistic forward model of the proxy system of interest. Such models can be more complex or they may concentrate on a core set of processes (compare the full and reduced implementations of the Vaganov-Shashkin approach to modelling tree-rings presented by, e.g., Evans et al., 2006, Tolwinski-Ward et al., 2011). That is, the first approach to obtaining pseudoproxies is process based. Other, more reduced approaches potentially ignore this mechanistic process understanding and focus on stochastic expressions of the noise that influence our inferences about past climates. Such an approach can try to formulate mathematically tractable expressions for statistical noise-terms, which represent the different processes or effects influencing the stages from the original environmental influence to our final observation and reconstruction [Dolman et al., in preparation]. Another way of producing pseudoproxies by focussing on stochastic noise expressions uses simple estimates of plausible errors. The different approaches can be very general or specific for certain proxy types. They can focus on one stage of the proxy system from environment to measurement or consider multiple stages. Indeed, all these approaches fit into the conceptual descriptions of Evans et al. (2013).

The manuscript now includes a paragraph comparable to the above initial response.

Page 2, line 30: How is a probabilistic description not a way to capture the error?

The referee is correct that our formulation is unclear. We will clarify it along the following: Our interest explicitly is to include the uncertainty from the dating in a statistical noise term for a pseudoproxy time-series. Therefore, we do not consider Bayesian or Monte Carlo methods but take a simple approach to develop an error term for the dating uncertainty.

The manuscript now includes sentences comparable to the above initial response.

Page 4, line 3: Why choose an arbitrary point on the map? Why not a place where it would be possible to have a sedimentary record in the first place (ocean or lake)?

The TraCE-21ka simulation has a low grid-resolution of about 3.75 times 3.75 degree. The chosen grid-point represents a good portion of the northern Iberian peninsula including the Pyrenees, where a lake core could be located. However, the choice is in so far arbitrary as we do not try to use the location of an available core. As this is an example for visualisation's sake, we do not see why this choice is critical. As we mentioned, the use of a land grid-point eases the readability in comparison to, e.g., the grid-point in the supplement.

We change the presented data to the grid-point at 150E, 38.97N.

Page 6, line 14: I'm rather unclear about "Bias at the reconstruction level"? Aren't all sources of noise and biases important the reconstruction level? Also how can seasonality not be considered sensor uncertainty?

We will clarify this. The referee is correct, as all sources of error affect the reconstruction, and indeed seasonality and habitat are sensor specific factors.

Our rationale here is that there are processes for which our possibly wrong attribution is a factor at the reconstruction stage and not at the sensor stage although the processes indeed are forward factors in the evolution of the record from the date of an environmental occurrence to our reconstruction.

Page 7, line 18: The change in noise level is not obvious at all.

Since we changed the grid-point, we also hope the change in the noise-level becomes more obvious. We try to provide more pointers in the text to Figures 1a.

Page 7, line 22: What three versions? Since it seems to be important, could you describe them?

We use three different amplitudes as shown in Figure 1b and also highlighted in Figure 1d.

We change "versions" to "potential amplitudes" to clarify this point.

Page 17: Why use the Lomb for comparison? It is know to have a bias in the high frequencies. The WWZ transform might be a better option for unevenly-spaced datasets.

We followed Dee et al. (2017) in using the Lomb-Scargle method. The preference of the one over the other is a matter of subjective choice in our opinion. We feel confident in using the Lomb-Scargle by the given reference and reconsidering the wider literature.

We now use the approach described by Mathias et al. (2004), which is related to the WWZ. Results and implications are comparable between Lomb-Scargle and this approach.

Page 17, line 27: In pseudo proxy experiments, something needs to be used as ground truth would it be reanalysis data, instrumental data or in the case the TracCE-21ka output. I'm unclear as how these peaks are spurious rather than "the proxy didn't capture them."

We will clarify that, in our understanding, these peaks and troughs are due to the specific forcing implementation as presented for example on http://www.cgd.ucar.edu/ccr/TraCE/.

The manuscript now includes a comment comparable to the following: First, peaks and troughs at some location are clearly attributable to the specific implementation of the forcing in the TraCE-21ka simulation (He, 2011; see also, Liu et al., 2009). That is, these signals are not realistic but due to technical decisions in the production of the simulations.

Page 19, last three lines: I agree that a process-based model would be more useful and they are fairly simple to implement. Hence, I don't understand how the noise approach presented here is useful.

Our understanding of reconstruction methods and of simulations benefits from multitudes of approaches. The benefit of this manuscript is that it provides in its assets the data. It boils the processes down to noise or bias formulations. It follows pseudoproxy approaches used for the period of the Common Era, which rely on noise. As stated above, the usefulness of a data-set or an approach often may be assessed a priori, but not always.

We hope that our clarifications in the manuscript are sufficient to convince the referee that our simplistic noise-based approach has merit. In short, we would like to stress the complementary nature of our approach to more complex set-ups and the flexibility of the noise formulations, which allow easily adapting it to changes in our understanding and thus facilitates the performance testing of different tools of paleoclimatology.

Technical corrections

I would also suggest editing the manuscript for English.

We tried to improve on our use of English.

For instance, Page 2, line 3: "as base of the comparisons". Do you mean "as a basis for comparison"?

We are sorry for this oversight.

We clarified the sentence.

Page 2, line 4: "a, eg. Temperature reconstruction and the model" is clearly missing words.

We are sorry for this oversight.

We modified the sentence.

Referee 2:

This study presents a generalizable approach to modeling sedimentary proxy systems and then shows how it works using the TraCE-21ka simulation. I think this is a good study that provides a flexible way to estimate various kinds of noise in proxies and that provides a nice set of pseudoproxies for potential use in a pseudo-reconstruction framework. I also think that this study can be useful for seeing how different uncertainties can affect proxy time series.

We would like to thank the referee for their comments and their generous evaluation of our manuscript.

I have a number of comments, corrections, and requests for clarification below:

Abstract and elsewhere: The use of e.g. and i.e. is too frequent and would be better to just re-write with words.

We hope we corrected this sufficiently in the revised version.

There are several paragraphs throughout that are just two sentences, which is a little unusual and not totally necessary, and so would be better suited to combine with surrounding paragraphs.

We reorganised the manuscript accordingly.

Introduction: Can you better situate the present study in the context of previous approaches to generating sedimentary proxy system models/pseudoproxies? What is unique about this approach? Is it more comprehensive than previous studies? Does it innovatively use the Evans et al. 2013 framework? Is it the first to be applied to the TraCE simulation or to generate pseudoproxies over this time frame? Etc.

We position our manuscript, data, and approach better in the larger context.

The introduction now tries to clarify the difference between our approach and previous applications, our contribution to this topic, how it relates to the framework of Evans et al. (2013), and how we use the TraCE simulation differently to other studies.

p.2 l.8 The words "The review" just after citing both Smerdon 2012 and Mann and Rutherford 2002 make it unclear which paper you're referring to.

We thank the referee for spotting this.

We clarify this.

p.2 l.12-17 I'm not sure this discussion of "three" different ways is quite right or at least I think I disagree with the framing of the issues here. For instance, the "proxy system model" framework of Evans et al. 2013 subsumes all of these. And so it's not as though using a proxy system model framework is a different approach from just estimating proxy error, it's that just estimating proxy error is usually considering only one of several issues that must be accounted for in the construction of pseudoproxies (i.e., only estimating the "sensor model" while potentially ignoring the "archive model" and the "observation model", using the terminology of Evans et al. 2013).

We try to clarify our framing as follows:

In our understanding there are various approaches to obtain pseudoproxies. These range from most comprehensive to most simplified. We can try to obtain a comprehensive representation from the environmental influences on a sensor to the measurement and implement this into a mechanistic forward model of the proxy system of interest. Such models can be more complex or they may concentrate on a core set of processes (compare the full and reduced implementations of the Vaganov-Shashkin approach to modelling tree-rings presented by, e.g., Evans et al. 2006, and Tolwinski-Ward et al., 2011). That is, the first approach to obtaining pseudoproxies is process based. Other, more reduced approaches potentially ignore this mechanistic process understanding and focus on stochastic expressions of the noise that influence our inferences about past climates. Such an approach can try to formulate mathematically tractable expressions for statistical noise-terms, which represent the different processes or effects influencing the stages from the original environmental influence to our final observation [Dolman et al., in preparation]. Another way of producing pseudoproxies by focusing on stochastic noise expressions uses simple estimates of plausible errors. The different approaches can be very general or specific for certain proxy types. They can focus on one stage of the proxy system from environment to measurement or consider more or even all stages. Indeed, all these approaches fit into the conceptual descriptions of Evans et al. (2013).

The manuscript now includes a justification comparable to the paragraph above.

p.3 l.13-15 It's not clear to me what this sentence means. "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." The different stages have different types of noise that are particular to the specific process under consideration.

Our thinking here is: Considering the reconstruction stage, our, e.g., calibration introduces additional uncertainty, which is not a priori captured by the stages sensor, archive, measurement. We can argue to include it in the measurement stage. We can also argue that these uncertainties are de facto uncertainties resulting from processes at the sensor stage or at the archiving stage. Similarly, our understanding is that the sensor model does not commonly account for all uncertainties of the environmental influences. That is, an additional environmental stage could provide weighted data of various environmental influences. These processes, however to some extent, can also be included in the sensor model or uncertainties can be assumed to mostly affect the measurement model.

The introduction now tries to clarify these points.

p.7 l.12-15 It's not clear to me what the bias term actually is here. You mention several different things like that it is dependent on insolation, or that it is scaled to be positive, or that it is randomized, or that it is scaled by an ad hoc constant. So what is it then? All of these at once? Only one at a time? Can you state this more clearly and/or perhaps show in mathematical terms what you mean for the different cases?–just having the term "Bias(t)" isn't exactly clear.

In the initial formulation, we add one bias-term, which varies with time. It is calculated dependent on insolation, it is positive but it could be negative or the sign could be randomized, and it is scaled by an ad-hoc constant. We will clarify this and provide the equations for the bias term.

Figure 6: I recommend putting the dates of the periods on this figure so it's more clear which figures correspond to what period (e.g., the deglaciation vs. the Holocene, which have very different correlation maps)

We add titles to clarify the Figure.

Figure 7: It would help the reader to briefly explain what the values imply. Logs of standard deviation ratios aren't necessarily intuitive. Also indicating the specific date ranges that you're using (as in my comment on Fig. 6) would be helpful.

We now visualize standard deviation ratios in the Figure. We also added titles.

For both Fig 6 7. The color map used here is usually for dry-wet data, but the figures aren't about hydroclimate at all. I would recommend using a different colormap so as to minimize any confusion.

We change the color scales.

Section 3.5 It would be helpful to write this generalized model down in mathematical terms, not just explain in words, so that readers can be sure what exactly you've done in producing Fig 8 or so that they can think about ways to adjust the generalized model.

We provide a mathematical formulation of the generalizations.

p.21 l.11-12 This sentence isn't clear.

We clarify the sentence as follows: If we repeat the analyses in Figures 6 and 7 for the generalized approach, differences are hardly to identify.

Section 3.5.1 Can you motivate the "modifications" you're doing here? It's not obvious to me what needs modification and why. And modifications to which approach, the full version or the generalized one? And what's the motivation for using the generalized approach vs. the full approach? I think you also need to say more clearly what approach the ensemble of pseudoproxies is based on and why you chose one relative to the other for that dataset. Would it be possible and useful to provide pseudoproxies for both approaches?

We clarify all these points in the revised version.

We now write something comparable to: In the following we present an ensemble of pseudproxies. At 144 locations we compute 500 pseudoproxy records each. For this, we make further slight modifications to the generalized approach. These adjustments relax our assumptions and result in larger differences between members of the ensemble than would be possible without the modifications. [...] Using the generalized approach provides an ensemble based on the most reduced formulation. [...] As mentioned above, these changes relax our assumptions on the effect of changes in the background climate.

Indeed, it would be also of value to provide pseudoproxy ensembles for both the full approach as well as the un-modified generalizations.

We decide not to provide an ensemble for the full approach.

p.21 l.14 Why are there 500 pseudoproxies at 144 locations? And are there 500 total or 500 for each of the 144 locations (and thus n = 500*144 pseudoproxies)?

We clarify this now in the text.

Indeed, there is an ensemble of 500 pseudoproxies at the 144 locations. We chose the 144 locations as the unique locations after screening proxy locations from a number of publications.

Fig 10. The blue lines are hard to see here.

We changed the layout of the Figure and hope it is clearer now..

p.23 l.24-25 I'd recommend un-gendering this line using "their"

We thank the referee for spotting this extremely embarrassing mistake, and want to apologize for it.

We change this.

Referee 3:

SUMMARY

Bothe, Wagner and Zorita present code to produce sediment pseudo-proxy time series, i.e. a time series of a temperature variable that originates from transient climate model output and that has been modified in several stages mimicking - statistically - the processes that affect sedimentary palaeoclimate archives. This is a timely and relevant approach and could prove useful for model data comparison in the near future with more transient paleoclimate model simulations becoming available.

We want to thank the referee for their comments and their generous evaluation of our manuscript.

GENERAL COMMENTS

- unclear aims: which properties of the data will be compared? The present formulation only allows time-mean comparisons.

We are unclear about the direction of these questions.

We will try to clarify the purpose of our pseudoproxy data.

The provided data and the code allow for the generation of pseudoproxy records for any simulation. These can be continuous or temporally sparse and regularly or irregularly sampled. The provided data particularly allows the comparison of different models against another over time, as well as testing methods for the comparison of simulation and proxy data over time. The irregular sampling however hampers the comparison of time-slices. As said, we are unclear about the direction of this question/comment.

We hope the revised manuscript clarifies our intentions for the data.

- downloading and testing the data generation is cumbersome as all parts apparently have to be manually downloaded. It would help to have a provided zip file, and a README on how to get started.

The repository allows to download folders and storage-items as zip-files. However, we provide zip-files collecting the different types of data. Note that these zip-files can become huge. A short documentation clarifies how to best approach data and code-examples. This is additionally included in the zip-files. We hope the data is now more easily accessible.

DETAILED COMMENTS

- p1 l4/following: the term "pseudoproxies" suggests that it is possible to hand the code a description of a specific sediment record (including e.g. information on the number/precision/type of dating) and all ensuing uncertainties are considered. This is not the case here, as all terms of non-climatic/insolation uncertainty considered remain statistical and non-proxy/archive specific.

We understand and respect the referee's point but after reconsidering the term and its use in the literature, we try to explain our view in using the term. The term pseudoproxy does not refer to a surrogate for a specific record. In the literature, pseudoproxy, surrogate proxy, and pseudoproxy experiments are phrases, which refer to modifications of observational data, reanalysis data, or simulation output. The applications of such modifications are not limited to either real world proxy records or real world proxy types. Such modifications in the broad sense of pseudoproxies are simply stand-ins for real world data in research enterprises of interest.

The revised manuscript explains our view of the term pseudoproxy. We add in the introduction a comment similar to the following: We clarify here our use of the term pseudoproxy. We follow the literature since Mann and Rutherford (2002). That is, a pseudoproxy simply represents a modification of observational data, reanalysis data, or simulation output. It replaces real world proxies in a certain application. The term may but does not necessarily refer to stand-ins for specific proxy records or particular proxy types. That is, the term pseudoproxy does not by itself imply that the modifications of the input data represent validly the uncertainties or characteristics of real world data. This view of the term pseudoproxy is in line with the past literature (compare, for example, Mann and Rutherford, 2002; Osborn and Briffa, 2004; von Storch et al., 2004; Jones et al., 2009; Graham and Wahl, 2011; Thompson et al., 2011; Lehner et al., 2012; Smerdon, 2012; Hind et al., 2012; Annan and Hargreaves, 2013; Kurahashi-Nakamura et al., 2014; Steiger and Hakim, 2016). These modifications may be simply by adding noise to the input data or may invoke more complex forward approaches (for example mechanistic Proxy System Models, Evans et al., 2013, see below).

-p2 l26-31: Considering dating uncertainty as purely additive white noise independent of the time axis strongly limits the suitability of the resulting time series. Autocorrelation results from the distortion of the time axis by changes in accumulation rate - which should, in a real proxy record, be captured by dating, and subsequent age modelling. Dating uncertainty represents a large component of the overall contribution to the low signal to noise ratio (c.f. Reschke, Rehfeld Laepple, Clim. Past. Discuss). The net cross-ensemble mean of the dating contribution to the final pseudoproxy uncertainty is zero in the presented formulation, as is the serial correlation of the component. Both is not appropriate. It would be beneficial to adopt (or include/prepare for) ensemblebased age models for the actual underlying proxy records; or if a simplistic solution is desired, to include the more realistic option of modeling age uncertainty by relative squeezing and stretching of the time axis.

We appreciate the referee's concerns. There are various points here. We disagree on the suitability of the resulting time-series. We are still confident that the time-series including the error terms are suitable for wide range of applications in, e.g., comparing different model simulations, model-data evaluation, and testing of reconstruction methods. The dating uncertainty in our set-up has two parts, the sampling of the dates and the sampling of the dating uncertainty on the one hand, and the associated dating error modelling on the other hand.

Indeed, not all our implementations of the dating error show notable serial correlation, however, most and particularly the main approach led to a small amount of serial correlation by making the dating uncertainty error dependent on previous errors and "measurements".

We tried to modify our code to allow for larger uncertainty and correlation in the final products. The changes do not result in notably increased serial correlations. We now provide data for this approach, but the initial approach is still included in the code. As mentioned above, we still regard the full set of pseudoproxies of immediate value for a number of tasks.

Thus, indeed, we choose to concentrate on a simplistic approach.

- p 3 l30 and following: Why is only summer seasonality considered? Is this a limitation of the pseudoproxy code?

In the initial submission, we concentrated on the modern boreal summer season since many authors attribute the sensitivity of their proxies to the summer season. We could have similarly used annual data or any other seasonal definition. The pseudoproxy code takes a time-series of any seasonality of interest. At this stage, it does not read the full annual matrix. We considered within the preparation of revisions to formulate the code more flexibly.

We now consider annual data for the simulation, and, in turn, also for the insolation bias. The code anyway allows to calculate the bias for arbitrary seasonal definitions, but we only consider one fixed definition for convenience. As noted, this is now annual.

- p4 l3: why this gridpoint? The arbitrariness of this choice somewhat illustrates that it appears difficult to use this code to include knowledge on real-life proxy datasets (e.g. sedimentation rate/dating frequency/ multi-proxy configurations).

As the referee already notes, this decision was generally arbitrary. It was made in relation to comparison with various proxies around the Iberian Peninsula and more generally Europe in another context. As we described, the visualisation for another proxy was provided in the Supplementary materials. As we state above, the aim here is not to test real world proxies but to provide data for testing methods and code to evaluate models and data. Our approach does not aim at mimicking real proxies but at providing pseudoproxies also for locations where we do not have real world knowledge.

We now consider a different grid point, which is located in the northwestern Pacific.

- Sec. 3: Please provide a graphical illustration of your pseudoproxy generation (e.g. using a graphical model).

We provide a visualisation of the procedure in the manuscript and in the manuscript assets.

- p6 l26: Autocorrelation should be considered, as several of the noise components (dating, non-local climate) are expected to be autocorrelated processes. The difficulty will be in actually estimating the true autocorrelation that should be used for the noise process.

As we note, the code provided allows for generating autocorrelated noise.

We now consider correlated noise at more instances.

- p7 l6: Is the assumption of increasing noise variability with increasing parameter variability appropriate for all noise components? It would appear that larger climatic variations might be recorded more precisely. In the absence of information whether proxy noise is smaller or larger for higher or lower climate variability this term should be reconsidered.

While larger variations may be recorded more precisely one may also assume that larger, e.g., temperature variations are associated with larger variations in other environmental components that may result in larger errors.

We introduced a switch.

- p7 l22: Can winter insolation be considered as bias?

If we understand the referees question correctly, then, yes, there is no reason why one should not consider winter insolation.

Our code does not calculate a winter insolation bias, but considering a winter insolation bias only requires changing a number of parameters defining the months of interest.

- p8 l5 and following: How do the processes and results here compare to the approach by Dolman and Laepple (2018)?

Our assumptions simplify the more complex approach of Dolman and Laepple. Our randomization could be seen as more realistic while the lack of process based assumptions make our approach less realistic.

Generally our results are not as smooth as the results of Dolman and Laepple considering the bioturbation while the assumptions on the subsampling appear to be comparable in our reading.

- p10 l13: typo original

We want to thank the referee for spotting this.

We change this in the revised version.

- p10 and following: The measurement error will depend on the type of sampling. To which degree is the sampling of the pseudo-proxy archive consecutive, overlapping, or spot-wise?

In this implementation, we only consider a conceptual measurement error and sample the time-series at certain years in the time-series as stated in the text.

We tried to clarify this in the text.

- p13 l12: Consider also the Bayesian Age-Depth modeling methods (e.g. OxCal, Bacon etc) which provide probability density functions of the proxy records.

We mention age modelling methods now, but do not discuss them in depth. We note, that we provide in principle most informations necessary to apply methods like Bchron (Haslett and Parnell, 2008) similar to their application in PRYSM by Dee et al. (2015, 2018). Additional information for these methods could be randomized.

- Figure 5: Please provide ensemble averages that allow to assess the spectral biases due to the proxy processes more easily.

We, at this point in the manuscript, do not use ensembles but only the single estimate.

We now use a wavelet based approach and weight the spectra to provide a smoother visualisation to ease the comparison.

- Figure 10: The time series are difficult to process and compare by eye. It appears in some cases there is an amplification of the apparent signal in the pseudoproxy record. Why? Where on the globe is the SD of the pseudoproxy > the SD of the climate signal?

This is mainly due to the size of the considered bias and the amplitude of noise processes.

While we announced to provide a visualisation equivalent to Figure 7 to answer the question about the SD, we decide simply to refer to the upper panels of Figure 7. These do not show the data from Figure 10 but highlight the SD-ratios between input data and the proxies. We changed the visualisation there and now do not use the sampled interannual input data but samples of a 501-year moving average version of the input data. We also add a panel for the full sampled records.

- p28: missing section ref.

We want to thank the referee for spotting this.

We change this in the revised version.

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.~~ so called proxies. It is essential to evaluate these reconstructions 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 climate simulation. They in turn depend on an understanding of the uncertainties of the real proxies ~~, i.e.~~ including the noise-characteristics disturbing the original environmental signal. Common pseudoproxy approaches so far ~~concentrated~~ concentrate 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 pseudoproxiesthemselves~~, and one example of pseudoproxies. The noise model 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.~~for example, 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

Proxy-records and derived reconstructions are our only observationally based information about past climates before the period covered by human observations, ~~i.e.,~~ that is before we have documentary or instrumental evidence. Climate reconstruction methods statistically process the information in the proxy records to extract the recorded climate signal. However, this climate signal is potentially multivariate, and we are often only interested or able to extract the signal for one single climatic parameter.

All other imprints of climate are noise relative to this variable of interest. Furthermore, part of the variability in the proxy records is not caused by the climate ~~. This~~ but other factors influencing the original generation of the proxy-record. Thus, there are climatic and non-climatic ~~variability, i. e., the proxy noise ,~~ noise contributions to the proxy variability. This proxy noise may cause biases and uncertainties in the resulting climate reconstructions. Evaluating the quality and reliability of reconstructions and of proxy-records requires 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 for example 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~~ models—the pseudoproxies. In testing the reconstruction methods, pseudoproxies ~~then~~ eventually replace the real 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 isachieved by contaminating the climate variables simulated by~~

Our use of the term pseudoproxy follows the literature since Mann and Rutherford (2002). That is, a pseudoproxy represents a modification of observational data, reanalysis data, or simulation output. It replaces real world proxies in an application. The term does not necessarily refer to substitutes for specific proxy records or particular proxy types. That is, the term pseudoproxy does not by itself imply that the modifications of the input data represent validly the uncertainties or characteristics of real world data. This view of 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~~term pseudoproxy is in line with the past literature (compare, for example, Mann and Rutherford, 2002; Osborn and Briffa, 2004; Von Storch et al., 2004; Jones et al., 2009; Graham and Wa Modifications of the input data may be as simple as adding white or colored noise or they may invoke more complex forward approaches (for example mechanistic Proxy System Models, Evans et al., 2013, see below).

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)~~ 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~~ the work of Mann and Rutherford (2002) focussing on climate-field reconstructions. The review by Smerdon (2012) emphasizes the essential contribution of pseudoproxy-experiments to our understanding of past climates and to evaluating our methods of studying past climates. ~~Most~~ To date, most studies using 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)~~ (e.g., Laepple and Huybers, 2013; Dolman and Laepple, 2018; Dee et al., 2018).

~~There~~ 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.

In our understanding there are various approaches to obtain such pseudoproxies. ~~On the one hand we~~ These range from most comprehensive to most simplified. 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., 2018; Evans et al., 2013). Secondly, we~~ a comprehensive representation of a so called proxy system (Evans et al., 2013) from the environmental influences on a sensor to our measurement and formulate this into a mechanistic forward model of the system of interest. Such models can be very complex or they may concentrate solely on a core set of processes (compare the full and reduced implementations of the Vaganov-Shashkin approach to modelling tree-rings presented by Evans et al., 2006 That is, the first approach to obtaining pseudoproxies is process-based. Other, more reduced approaches potentially ignore this mechanistic process understanding and focus on stochastic expressions of the noise that influence our inferences about past climates. Such an approach can try to formulate ~~a mathematically tractable expression of the proxy error~~ mathematically tractable expressions for statistical noise-terms, which represent the different processes or effects influencing the stages from the original environmental conditions to our final observation [Dolman et al., in preparation, A. Dolman, personal communication, 2018, T. Laepple, personal communication, 2017]. ~~A third way of formulating the proxy noise is to use a simple estimate of a plausible non-climatic error in proxy-records. The~~ Another way of producing pseudoproxies by focussing on stochastic noise expressions uses simple estimates of plausible errors. These different approaches can be ~~very general or~~ specific for certain proxy types or very general. They can focus on one stage of the proxy system from environment to measurement or consider multiple stages.

~~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~~ All these approaches fit into the concept of a proxy system model as described by Evans et al. (2013). The idea of forward models to study the behavior of proxies and proxy systems is not new (e.g., Schmidt, 1999; Tolwinski-Ward et al., 2011; Thompson et al., 2011) but Evans et al. (2013) were the first to clearly delineate the modelling of proxy systems. A proxy system represents the biological, chemical, geological, and possibly also documentary system that translates environmental influences into an archived state on which researchers make observations. We usually refer to these observations when speaking of climate proxies. A proxy system model is a representation of how the proxy system translates the environmental influences into our observations based on our understanding. Evans et al. (2013) present a generalized concept of this modelling approach, which consists of three components: First, a sensor model reacts to the environmental influences. Second, an archive model transforms these sensor recordings into archive units. A third model translates the archive into representations of what we usually observe on an archive. For example, the sensor 'tree' records the environmental influences in its archive 'wood', and we can make measurements on this archive in form of tree-ring counts and widths etc. The full system from recording to observation is the proxy system.

Each stage in this system and its model representations adds uncertainty, and each stage omitted in a generalization also increases uncertainty. For example, the environment and the final reconstruction process can be additional stages, but we can try to include the associated uncertainties in any of the three stages proposed by Evans et al. (2013). That is, considering the reconstruction stage, the calibration introduces additional uncertainty, which is not a priori captured by the stages sensor, archive, and measurement. We can argue to include this additional source of error in the measurement stage. We can also argue that these uncertainties are de facto uncertainties resulting from processes at the sensor stage or at the archiving stage and include them there. Similarly, the sensor model does not necessarily account for all uncertainties of the environmental influences. An additional environmental stage could provide weighted data of various environmental influences (compare, e.g., Dee et al., 2018). These processes, however, can also be included in the sensor model or uncertainties can be assumed to mostly affect the measurement model. In short, inferences about past climates from proxy-data are based on observations on an archive that accumulated a property of a system. This (property of the) system was sensitive to and recorded an environmental process at some date. From the recording stage to our inference there are multiple sources of error to our inference.

The potential errors include different sources of ~~errors~~ noise 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., the calibration relation. A recipe for calculating pseudoproxies may include potential error estimates not only for parts of the assumed proxy-system but also for the relation between the 'observed' data and time, that is the anchoring of the data in time.

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 probabilistic description and rather in how we can capture the error in a~~ Our interest explicitly is to include the uncertainty from the dating in an statistical noise term for a pseudoproxy time-series~~, we~~ . Therefore, we do not consider Bayesian or Monte Carlo methods but take a simple approach to ~~include an error-term resulting from dating uncertainty~~ develop an error term for the uncertainty in the dating. We also do not include explicit age-modelling (compare, e.g., Haslett and Parnell, 2008; Blaauw and Christen, 2011; Trachsel and Telford, 2017).

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~~ or among different model-simulations. This increases our understanding about past climate changes by consolidating information from all available sources, ~~i.e.,~~ which are proxy records and model simulations. The lack of high-quality observations with small uncertainty is always going to hamper efforts to assess the quality of model-simulations of past climates. Such comparisons 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~~take place in the virtual reality of the model and use the modelled variables. In the case of proxies, the comparison is between~~a, e.g.,~~ , for example, a temperature reconstruction and ~~the~~ a 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:~~ Recent years saw an intensification in the research on forward modelling proxies for understanding proxies, testing reconstruction methods, and evaluating simulation output (see, for example, Dolman and Laepple, 2018; Dee et al., 2015, 2018; Konecky et al., 2019). Many of these approaches follow the concept of considering 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~~observations as distinct steps in the process. Still, few of these approaches consider transient time-scales beyond the late Holocene. Nevertheless, particularly the work by Dolman and Laepple (2018) and also Dee et al. (2018) allow for the calculation of different sedimentary proxies over multi-millennial time-scales based on knowledge of certain processes in the respective proxy systems.

In this paper, we adopt the conceptual sub-divisions of Evans et al. (2013) to present a formal but still simple noise based approach to describe the ~~noise present~~ disturbances masking the signal in proxy records~~, which~~. This approach can also be applied to ~~the generation of~~ produce pseudoproxies for timescales longer than the last few millennia, ~~i.e. including also~~ that is 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~~ covers the last 21,000 years. We concentrate on proxies, which are subject to some kind of sedimentary process. Thus, our work appears to be particularly similar to the proxy system model for sedimentary proxies implemented by Dolman and Laepple (2018). Dolman and Laepple (2018) also consider the long time-scales since the last glacial maximum and rely on output from the TraCE-21ka simulation for their forward modelling. Both, the present manuscript and Dolman and Laepple (2018) follow the concept outlined by Evans et al. (2013). The main difference between Dolman and Laepple (2018) and the present study is that they provide a simple process-focussed model of the proxy system, whereas we try to provide a simple characterisation of the noise in the proxy system that finally influences the proxies. The process-based formulation of Dolman and Laepple (2018) concentrates on two types of marine proxies whereas our noise-based approach tries to generalize over sedimentary proxy types. We regard both approaches as complementary and want to emphasize the value in having a multitude of methods to assess model-simulations and reconstruction methods.

Our approach contributes to the existing proxy system modelling and pseudoproxy computation applications by being an intermediate step between complex forward modelling approaches and the noise based approaches, of which the latter may ignore the proxy system workings. Our code simplifies and generalizes more complex assumptions. The noise-focus and the generalizations allow us to provide global pseudoproxy data and an ensemble of pseudoproxy data using the TraCE-21ka simulation over the time-scale of the last 21 thousand years. The manuscript assets at ~~also include example code and the~~

~~calculated pseudoproxy data~~ https://doi.org/10.17605/OSF.IO/ZBEHX provide the generated pseudoproxy data and also include sample code. 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~~ user 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)~~ annual 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.~~for example, 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~~ the Ph.D.-dissertation of He (2011) provides more details.

The presented ~~code uses only~~ results and Figures are generally for 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~~ 150°E, 38.97°N. 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 by the freshwater forcing of He (2011, compare also Liu et al. (2009))~~. On the other hand, ~~this implies the disadvantage of featuring~~ the less extreme climate variations to be captured in a subsequent pseudoproxy can be seen as a disadvantage. 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~~105°W, 45.39°S in the South Pacific.

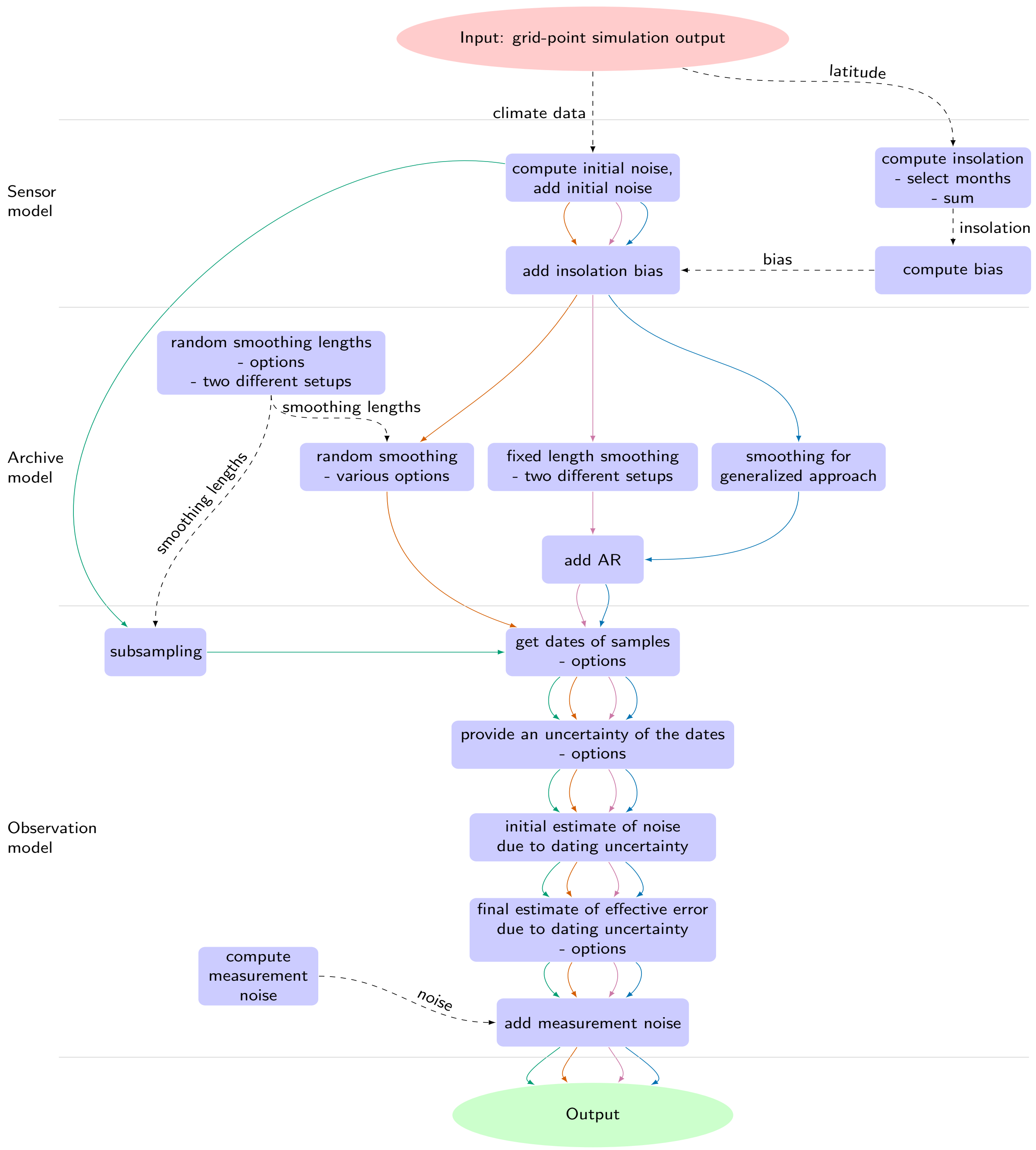
On multi-millennial time-scales we have to consider changes in the insolation caused by changes in earth's orbital elements. Global insolation data is calculated using the R (R Core Team, 2017) package palinsol (Crucifix, 2016).

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 system records. That is, do we assume ~~,~~ that the proxy records local or regionally accumulated signals? Here, we take the signal of interest to be local, that is 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 A4) summarising the considered parameters and noise models in the various steps of ~~the following~~ our considerations.



Legend:

Paths: main approach, fixed smoothing plus AR approach, subsampled pseudoproxy, generalized approach; Dashed paths: separate components providing inputs

Nodes: Input data, Output, components of the algorithm

Figure 1. Conceptual flow of the procedures.

In the following, we distinguish between different sources of errors related to the ~~concept~~ concepts of sensor, archive, and measurements of ~~Evans et al. (2013)~~Evans et al. (2013). Figure 1 summarises our procedure. 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 and bias

~~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.~~

The sensor, ~~e.g.~~ that is for example an organism or a physical or biogeochemical process, reacts to multiple parts of its environment. Researchers' 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,

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

Under this assumption, further components of the environment besides T contribute only noise components η_i to the reaction of the sensor. ~~These errors~~ Errors due to noise are not necessarily additive but can also be multiplicative or could bias the estimate. In a first step we, here, assume the sensor-reaction to be

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

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,~~ Thus, additional noise may be due to the sensor reacting to advected environmental properties instead of ~~"local"~~ 'local' ones or due to the environment redistributing the sensor or the record. In the marine realm 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.,~~ by 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.

Besides simple noise, redistributions of the environmental signal may also introduce biases in our estimate of the environment. ~~Such biases in turn are~~ Any bias is likely not fully time-constant but ~~evolve~~ evolves 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, changes in bathymetry, and lithospheric adjustments. ~~Such~~ All of these processes may lead to biases, and 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~~ We consider this bias on the sensor level. There are other potentially erroneous attributions besides the processes' seasonality. These are the location of the process in all three dimensions, ~~e.g.~~ for example, 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~~ we consider them as noise here~~, they are mainly errors in our reconstruction-calibration-relation. This~~ . However, they reflect a non-stationarity of our reconstruction-calibration-relation~~is an important source of uncertainty, although~~ . Nevertheless, the idea that the modern relations between environment and proxy system worked over the full period of interest ~~(e.g., Bradley, 2015)~~ is a fundamental assumption of paleo-climatology (e.g., Bradley, 2015).

In the following~~we consider two error terms. However~~, we assume three components ~~of the noise~~ to be important disturbances of the signal at the sensor level~~,~~ : the environmental noise, the redistribution, and the attribution errors~~which we here reduce~~ . We reduce the latter 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~~In the following, we reduce these three components to two terms in our modifications of the input data.

3.1.2 Noise

First, we assume that η_i includes both the effects of environmental dependencies and of redistribution~~, i.e. , our~~ . That is, $\eta_i = \eta_{env} + \eta_{redistr}$. 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.~~

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

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

~~We take~~ The current version of η_i ~~to be~~

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

~~where p is a constant scaling factor, and~~ is only a weakly correlated autoregressive (AR) process of order one, which we additionally scale by an ad hoc scaling factor. It thereby only includes a small part of the potential correlations among errors due to redistribution and other processes. The innovations are sampled dependent on time and climate background from $\mathcal{N}(0, S(t)^2)$, where $S(t)$ 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~~ Our general 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~~ switch to invert the moving standard deviation about its mean or to randomize the orientation.

3.1.3 Bias

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 compute it for any latitude of interest. We apply the orbital bias term as additive but one may see it as ~~multiplicative or~~ a multiplicative or a nonlinear effect in many cases. Therefore the code uses it after the noise term η_i. ~~This~~ The bias is the second error term ~~.~~

~~We add the~~ in our formulation of modifications at the sensor level. The bias term is a scaling of the changes in annual latitudinal insolation but it is possible to choose different sub-annual time-periods of interest. The scaling is arbitrary and we refer to the provided code for details. The bias ~~term dependent on the latitudinal insolation. In its formulation we concentrate on summer insolation. The insolation bias is scaled to be~~ is zero in the year ~~0BP. The bias becomes notable at some latitudes but may be rather negligible elsewhere. The bias is scaled~~ 0 BP. We set it to be positive if the insolation is larger~~but~~ ; this can be randomized ~~. The~~ in the code. The amplitude of the bias is scaled by an ad hoc constant. ~~Then the~~ The bias becomes notable at some latitudes but may be rather negligible elsewhere. We take the bias as $Bias(t) = \beta \cdot I_n$. Where β is the scaling constant, and I_n is a normalised and shifted insolation. I_n is calculated as $I_n = ((I - \bar{I})/\sigma(I) \cdot q_{0.25} - I(t = 0BP) + 1)^u - 1$ for a chosen period. The chosen time-period influences the statistics included in the scaling. We consider the insolation since 150,000 BP. $q_{0.25}$ is the 25th percentile of the insolation data, u is generally 1, but can be sampled from $U = \{-1, 1\}$.

The pseudoproxy becomes

$$P_T(t) = T(t) + \eta_i(t) + Bias(t) \tag{5}$$

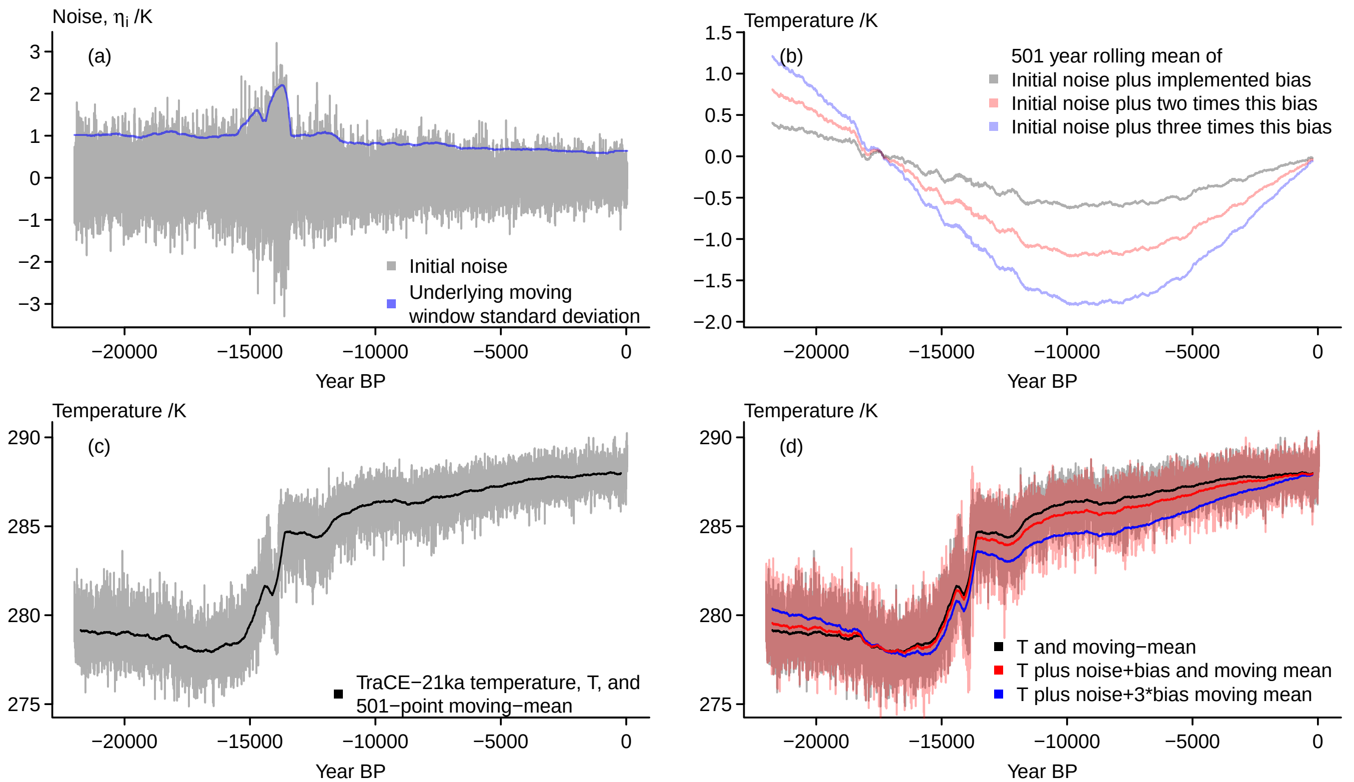


Figure 2. Visualising considered error sources at the ~~archive-stage~~sensor-stage: a) ~~501-year moving mean of the input data,~~ the ~~pseudo-archive series with longer average smoothing lengths,~~ initial noise and the ~~subsampled record, b) 501-year~~ underlying moving ~~mean~~ window standard deviations of the input ~~data~~signal, b) three versions of a potential bias as function of the ~~pseudo-archive series with shorter average smoothing lengths~~local insolation, 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~~its 501-point moving mean, 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~~ its 501-point moving mean plus noise and ~~added AR(1)-process~~bias. The unsmoothed initial temperature is effectively hidden behind the unsmoothed temperature plus bias.

3.1.4 Results

3.1.5 ~~Results~~

Figure 2a shows an example of the initial noise η_i. The dependence on the background state is ~~obvious with~~ clearly visible for the visualized grid point data. There is an increase during the deglaciation ~~. The blue line in the panel gives the underlying moving standard deviation~~and a multi-millennial reduction over the Holocene. 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 2 compares three ~~versions~~ potential amplitudes 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 2d 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~~ colder annual record over most of the considered time period while the record becomes slightly ~~colder~~ warmer 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 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.

Such modifications of the archive and sampling issues represent an important step in using inverse reconstruction methods because it is a priori not clear how the archive is generated and whether an individual measurement represents mean environmental states or relates to single events. In this context, forward models and pseudoproxy approaches of sedimentary proxies are a crucial tool in disentangling the controlling climatic environmental factors in the generation of sediment cores and their interpretation.

3.2.1 Smoothing and noise

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.~~for example, by diffusive processes or bioturbation ~~(compare, e.g., Dolman and Laepple, 2018, and their references)~~ (compare 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 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)~~ 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~~ sensor record on which we add a highly ~~correlated AR(1)-process~~autocorrelated AR-process of order one. 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~~ represent 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)~~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.~~

Indeed, the code includes options for the random smoothing lengths to depend on the mean climate or the climate variability. The provided data uses an approach where the random smoothing lengths follow an autoregressive process around a climate dependent reference smoothing length, where, considering Vardaro et al. (2009), warmer climates result in shorter smoothing intervals. The smoothed archive records are then either

$$P_T(t) = g_1(T(t) + \eta_i(t) + Bias(t), t) \tag{6}$$

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

$$P_T(t) = g_2(T(t) + \eta_i(t) + Bias(t)) + AR \tag{7}$$

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~~ represent the intra-annual seasonal uncertainty.

P_T in this case becomes

$$P_T = h(T(t), t) + h(\eta_i(t), t) + \eta_s \tag{8}$$

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

We do not include the bias term ~~here~~for the subsampled proxies. On the one hand we apply ~~this~~ the bias only for the ~~summer season~~ mean annual temperature, i.e. ~~other~~ individual seasons show different biases. While we could account for this by sampling the biases ~~of other~~ over the different 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~~ as 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~~

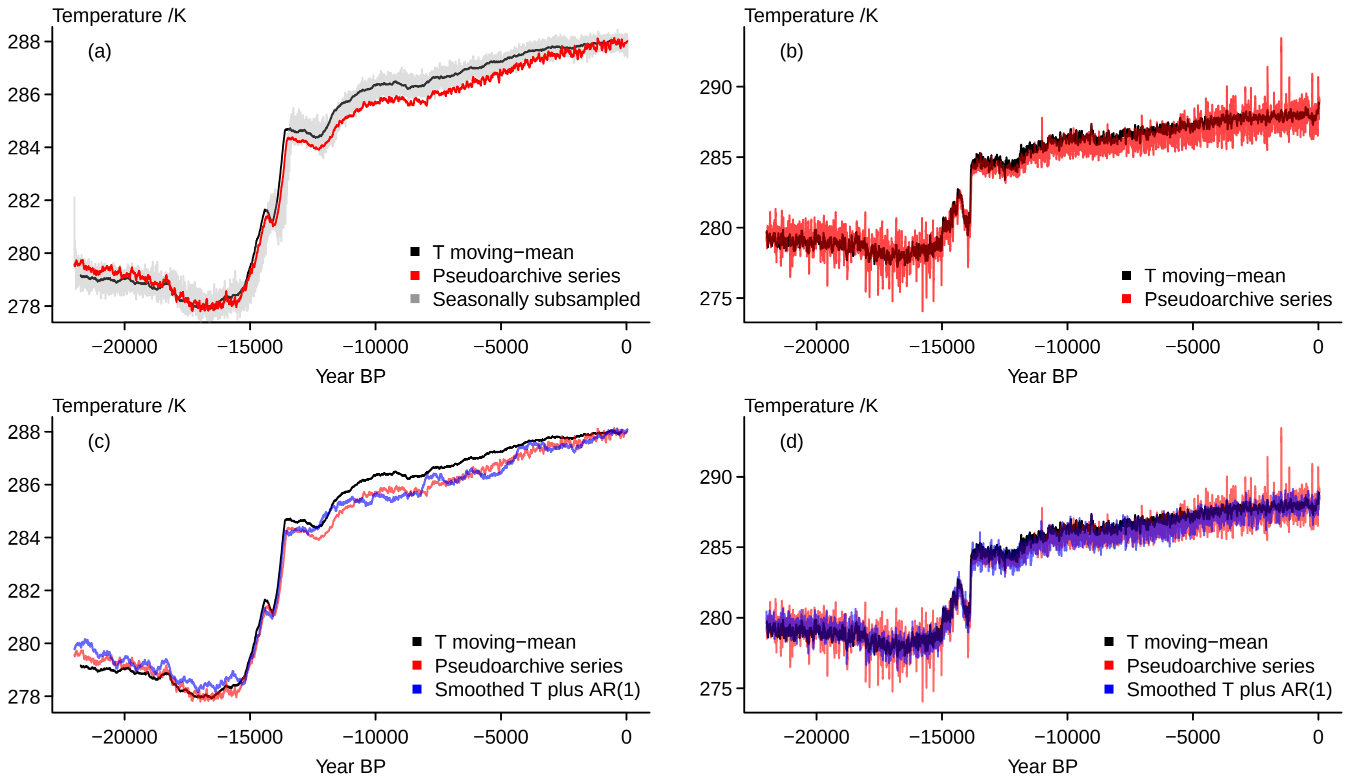


Figure 3. Visualising considered error sources at the archive-stage: a) 501-year moving mean of the input data, the pseudo-archive series with longer average smoothing lengths, and the subsampled record, b) 501-year moving mean of the input data, the pseudo-archive series with shorter average smoothing 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.

The biased moving average already shows the differences between the target temperature and the pseudoproxy-record (compare Figure 2. The pseudo-archive-series in Figure 3a shows this more clearly. Here we use a randomized smoothing interval. Differences are less visible for shorter random smoothing intervals (compare Figure 3b).

Further panels of Figure 3 add the constant smoothing archive approximations which we modify by an additional highly correlated AR-process (Figure 3c and d). This procedure randomly amplifies, dampens, or inverts certain biases in the presented case. That is, while the simple random smoothing ~~emphasizes~~ may emphasize 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.,~~ that is they are symmetric about their date. The time varying smoothing tries ~~to more realistically mimic~~ more realistically to imitate a bioturbation approach (compare Dolman and Laepple, 2018, and their references) and thus provides a shift in the series.

Figure 3a also shows the seasonally subsampled pseudo-archive-proxy. The data ~~does ignore~~ ignores the bias term and the resulting series is by construction symmetric around the original data, ~~i.e. the~~ our target. Nevertheless, there are pronounced deviations from the ~~orginial~~ original 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 (not shown).

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.,~~ to 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~~ . The sampling may lead to additional uncertainties due to disturbances of the archive, and the dating of our samples is a profoundly uncertain process.

~~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.~~

3.3.1 ~~Measurement error~~

3.3.2 Measurement error

Prior to dealing with errors due to dating uncertainty, we take an additional noise term to ~~mimic~~ represent measurement errors and apply this for each date to account for the potentially imperfectly measured series. The term includes not only the ~~uncertainties in~~ errors introduced by 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~~ has further to account for the accidental handling of the records in the laboratory, ~~e.g.~~ for example, 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~~ While we just mentioned various 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.

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

$$M_T = P_T + \eta_M \tag{9}$$

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

$$P_{P_T} = P_T(t = \{t_1, ..., t_N\}) \tag{10}$$

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. We do not distinguish between different sampling techniques. We simply sample the records at certain dates and add the described noise term.

~~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~~

3.3.4 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~~error. For assessing reconstruction methods and simulations, it ~~, however,~~ would be beneficial to be able to include dating uncertainty within the ~~proxy-uncertainty~~proxy-error. That is, if we consider proxies as tuples of data and date, we have to transform the uncertainty of the date into an error-term for the data. In the following we ~~have to~~ distinguish between the dating uncertainty, ~~i.e.~~ that is 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~~

proxy-record error (e.g., Breitenbach et al., 2012; Goswami et al., 2014; Boers et al., 2017). Ensemble and Bayesian age-depth modelling approaches also allow to infer an additional error term (e.g., Haslett and Parnell, 2008; Blaauw and Christen, 2011). However in the present application, 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. Nevertheless, since we provide sample dates and random sampling uncertainties, the application of age modelling to the pseudoproxies is in principle possible (e.g., following the approach of Dee et al., 2015, 2018).

The code includes several variations of ~~doing this~~our estimation of an effective dating error. These reflect different amounts of dependence between subsequent samples. ~~The following general approach is common to all . We~~ In all variants, 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~~The following general approach is common to all variations of our procedure: First, we sample ~~random dating~~ uncertainties in time for each sample date. We take these as dating uncertainty standard deviations. These uncertainties can be sampled fully randomly or dependent on the available smoothing interval data from the archive stage. 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

$$\epsilon_D = \mathcal{N}(\overline{P_{T_D}}, \sigma_D^2) \tag{11}$$

where

$$\overline{P_{T_D}} = \overline{P_T(t_S = \{t_{i-2\sigma_{dating}}, ..., t_i, ..., t_{i+2\sigma_{dating}}\})} - P_T(t = t_i) \tag{12}$$

is the mean over the region of influence and

$$\sigma_D^2 = E[(P_T(t_S) - \overline{P_{T_D}})^2] \tag{13}$$

is the variance of the distribution.

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

$$P_{P_T}(t_1, ..., t_N) = g(T(t) + \eta_i(t) + Bias(t), t)(t_1, ..., t_N) + \epsilon_D(t_1, ..., t_N) \tag{14}$$

Alternative formulations of the pseudoproxy become

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

or

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

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

$$\epsilon_{D_i} = \rho \cdot (\epsilon_{\xi_{D_{i-1}}} + (P_{P_{T_{i-1}}} - P_{P_{T_i}})) + \epsilon_{\xi_{D_i}} \tag{17}$$

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

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

~~where $\rho = 0.9$ in our code and~~ where $\epsilon_{\xi_{D_i}}$ are the random innovations for date i. Our initial choice of $\rho = 0.9$ 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,

$$\rho(t) = 1 - (t_i - t_{i-1})/(2 \cdot \sigma_d(i-1)) \tag{19}$$

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

Alternative simple formulations may include different noise processes ~~, e.g.,~~ like noise generated from Gamma-distributions. ~~Furthermore, the~~ The available smoothing interval data ~~could inform the~~ can inform the sampled dating uncertainty. We could ~~also~~ further use this information to provide a deterministic, ~~i.e.~~ not random, error for each sampled date, ~~i.e. taking~~ that is we could take 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

$$M_T = P_{P_T} + \eta_M \tag{20}$$

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. None of these are included in the current version.

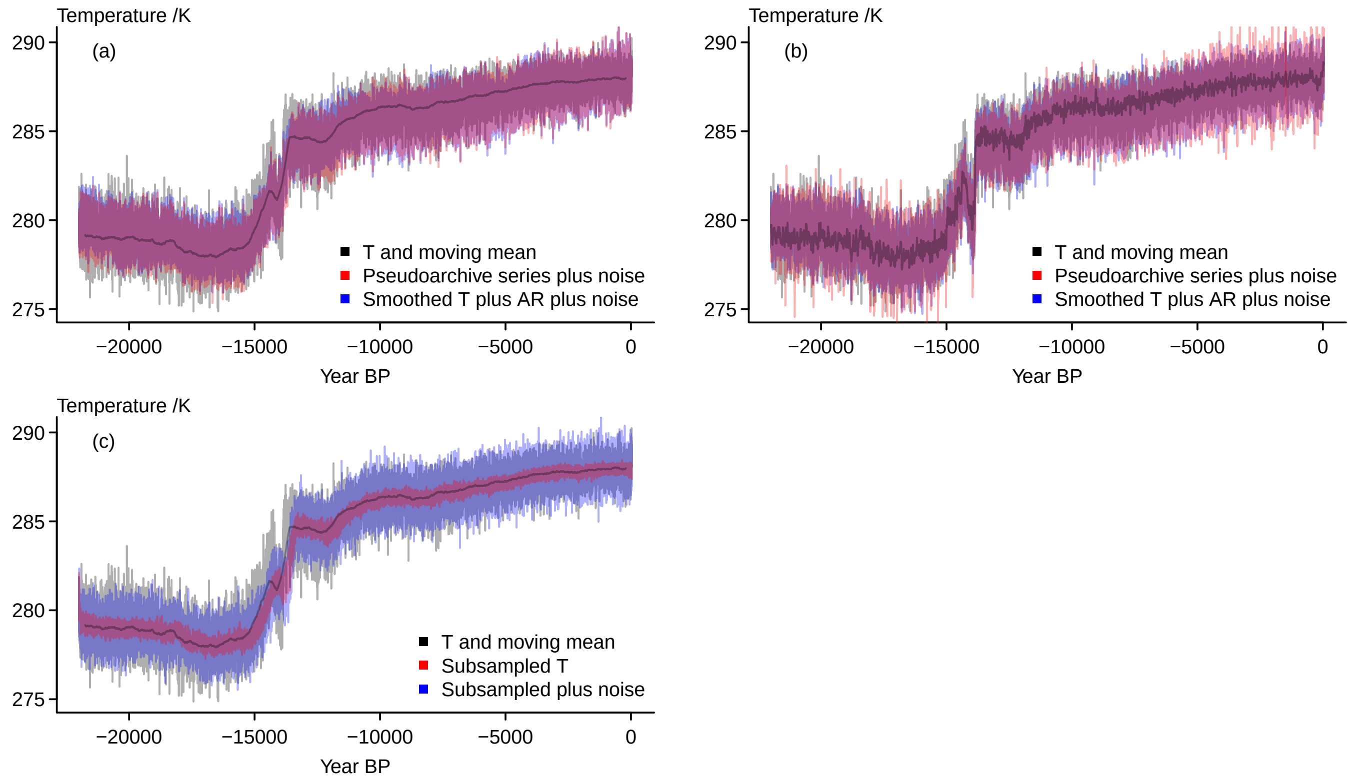


Figure 4. Visualising considered error sources at the measurement-stage for the full series: a) 501-year moving mean of the input data, the pseudo-archive series with longer average smoothing lengths and the constant smoothing plus AR series with added measurement noise, b) 501-year moving mean of the 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.

3.3.5 Results

~~Figure 4 show~~

Figure 4 shows versions of an archived proxy plus interannual measurement noise~~, i.e. they~~. The panels give an impression of how a proxy would look from measurements on a perfectly annually sampled archive. The final amplitude of the noisy proxy is ~~smaller~~ generally slightly smaller for all versions of our pseudoproxies than the amplitude of the interannual variations for the chosen location~~for all three versions , simple smoothing ,~~. This may be different at other locations. The different versions of the smoothing and of the smoothing plus AR ~~(~~approaches are shown in Figure 4a ~~), different smoothing and different smoothing plus AR (Figure 4b), and seasonally subsampled (~~and b, respectively. Figure 4c ~~). This may be different at other locations~~plots the seasonally subsampled pseudoproxy. The final ~~version generally preserves~~ versions of the pseudoproxies generally preserve previously included biases.

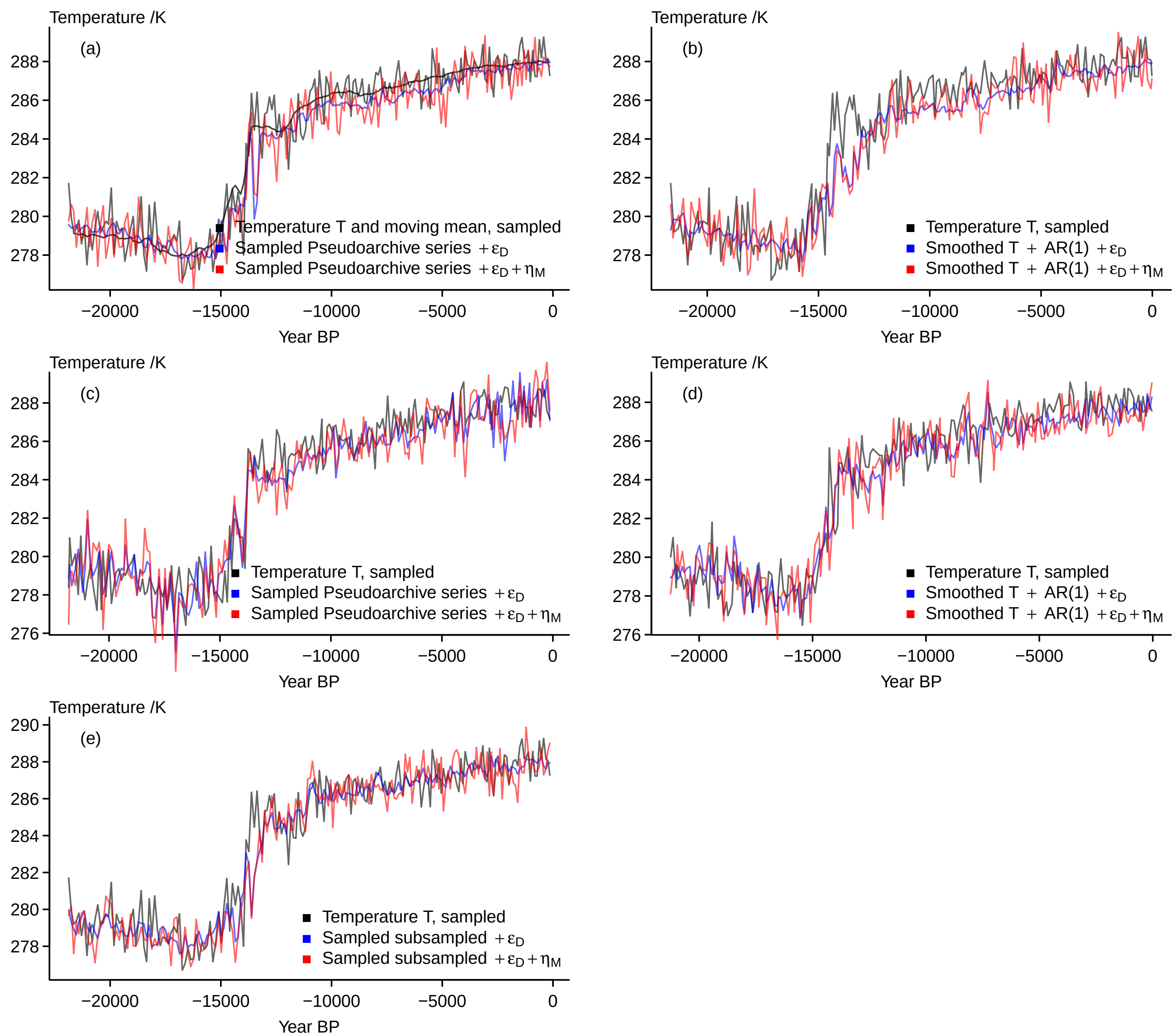


Figure 5. Visualising the sampled records: a) Input data and its 501-year moving mean, the pseudo-archive series with longer average smoothing lengths plus the effective dating error and plus the effective dating error and measurement noise, b) input data and its 501-year moving mean, the constantly 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.

Figure 5 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 generally slightly 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 5a) highlights that estimates for past climates may well fall within the range of the original interannual temperature variability but may nevertheless 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.

~~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).~~

3.4 General Results

Figures 2 to 4 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 4 includes the effect when we hypothetically add measurement noise at every date. Under our assumptions this noise is still smaller than or only as large as the original interannual variability but, including biases, mean estimates may be ~~close to the edge~~ outside 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~~ AR-process seems to cancel the orbital bias more strongly in Figure 3. Figure 5 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 ??~~

Figure 6 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 (?). The package follows Press et al. (?) and normalises the spectra by dividing them by two times the variance of the data. We omit this normalisation here~~power spectral densities computed from a wavelet based approach similar to the Weighted Wavelet Z-transform of Foster (1996). The approach is described by Mathias et al. (2004) and McKay and colleagues provide a compiled version at https://github.com/nickmckay/nuspectral (last accessed, 11 March 2019) (Nick McKay et al.). Due to the length of computation, we do not show the density for the full 22,040 year input data but only for a record sampled every ten years. Results may be specific for the chosen grid-point.

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~~ Spectral densities for the regularly sampled original temperature data in Figure ~~??~~6a 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 ?? the Figure suggests that the sample spectral estimates are rather similar. The equivalent Figure~~ results in prominent differences for ~~another grid-point in the document assets shows larger differences over~~ multi-centennial to millennial periods. On the other hand, differences are smaller for the irregularly sampled input temperature data but still notable for millennial periods. However, there is an offset between the irregularly sampled data and the regularly sampled input data.

Spectra for full and late records of the various pseudoproxies are generally similar to the irregularly sampled input data spectra (Figure 6b-f) but the offset to the input data can be smaller than in Figure 6a. Differences between sampled late and full records are often largest at intermediate millennial periods. Deviations are largest for the subsampled pseudoproxy approach at long periods (Figure 6f) but they become also notable for the constant smoothing approaches at shorter periods in the centennial band (Figure 6c,e). This is mainly due to the characteristics of 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~~spectra for the constant smoothing, which show an increase in power spectral density for shorter and longer periods. That is, the constant smoothing

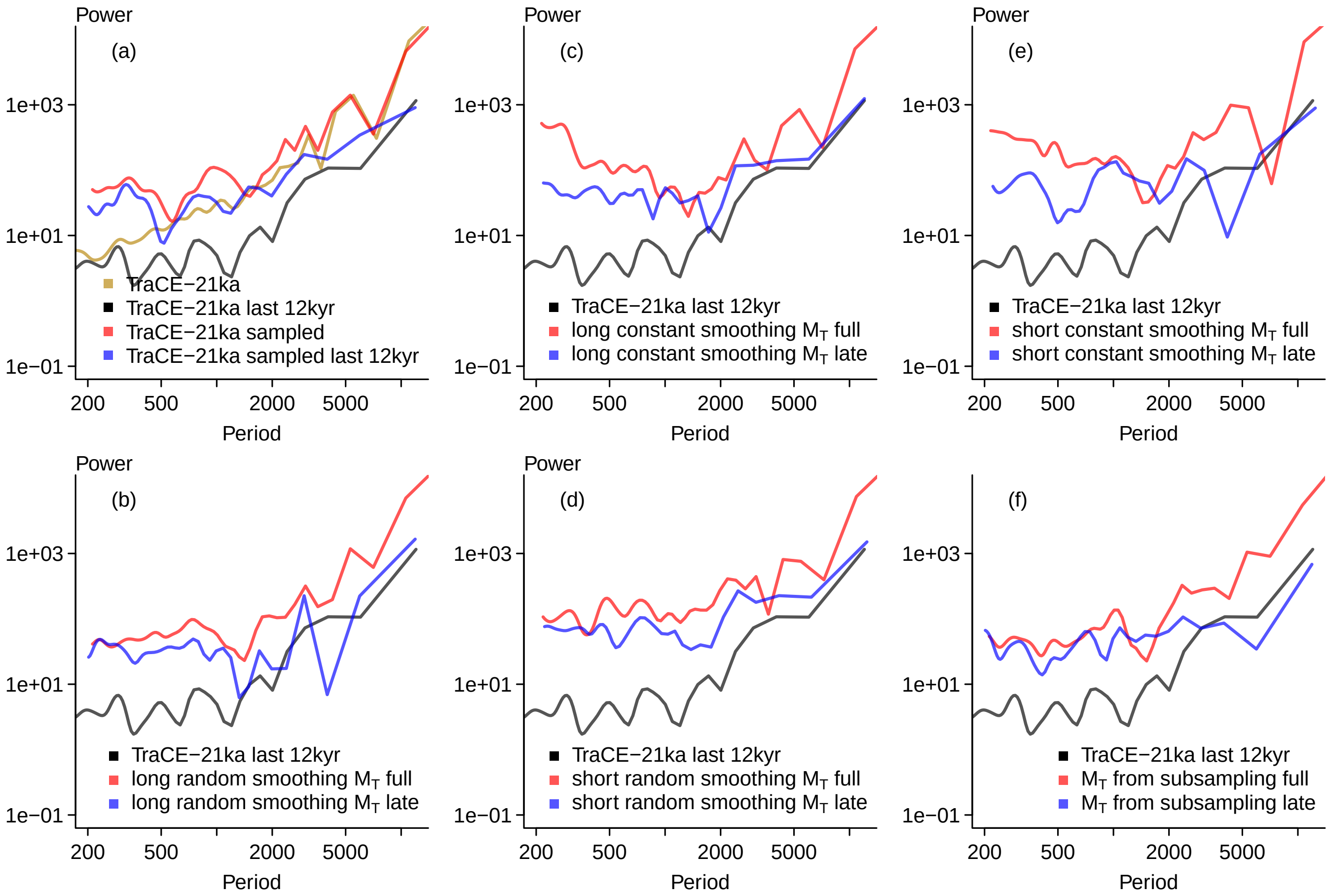


Figure 6. Wavelet based power spectral densities (Mathias et al., 2004; Nick McKay et al.). Densities are weighted following Mathias et al. (2004) to smooth the records for ease of comparison. Lines are for records split up by first 10k years of the records and the last 12k years of the records. Input data refers to the input data at 10 year intervals. 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).

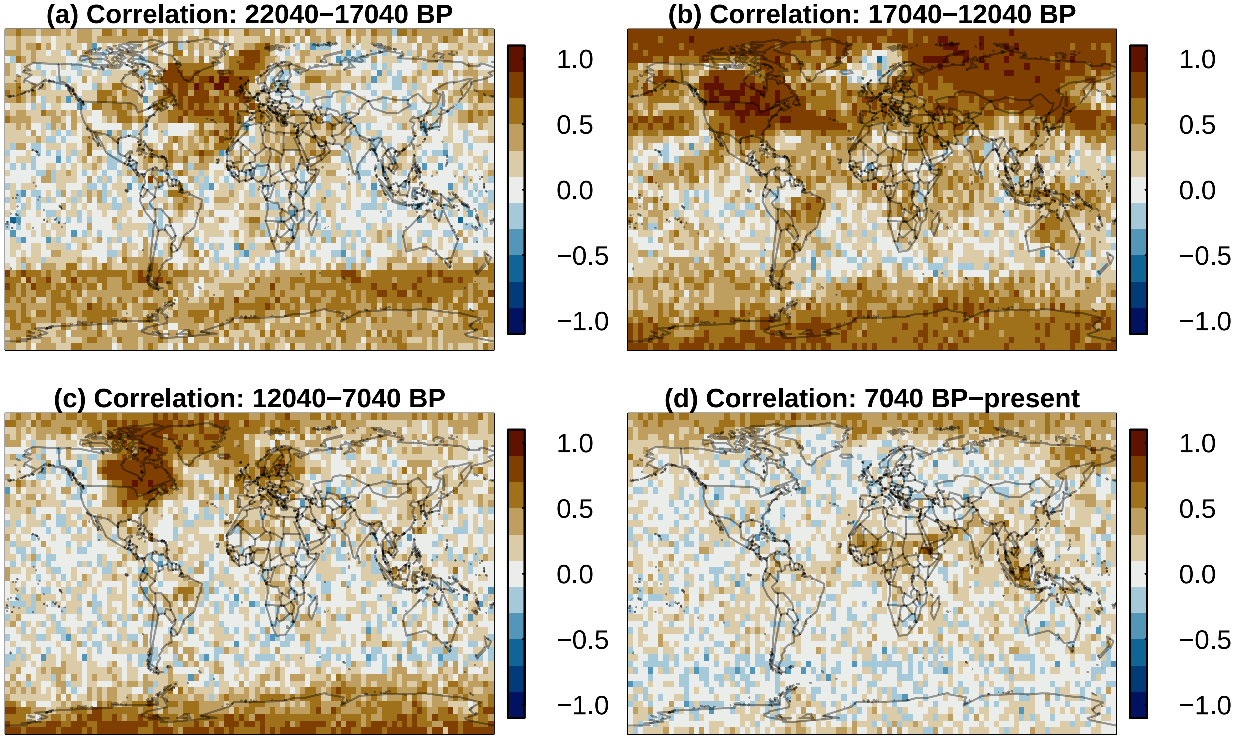


Figure 7. Point by point correlation maps between input data and the smoothed record plus AR(1)-process plus effective dating error and measurement noise for the sample dates within the first (~~top left~~a), second (~~top right~~b) and third (~~bottom left~~c) subsequent 5,000 year windows of the record and the samples within the remaining years (~~bottom right~~d).

full period spectra remind of grey noise spectra. Despite these differences and the apparent offset to the input data spectra, the irregularly sampled spectra for all cases are rather similar.

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~~ (Supplementary document 1 Figure 1 ~~(~~at https://doi.org/10.17605/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~~ peaks and troughs at some location are clearly attributable to the specific implementation of the forcing in the TraCE-21ka simulation ~~results in occasionally spurious peaks and troughs for some locations~~(He, 2011; see also Liu et al., 2009). That is, these signals are not realistic but due to technical decisions in the production of the simulations. 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~~

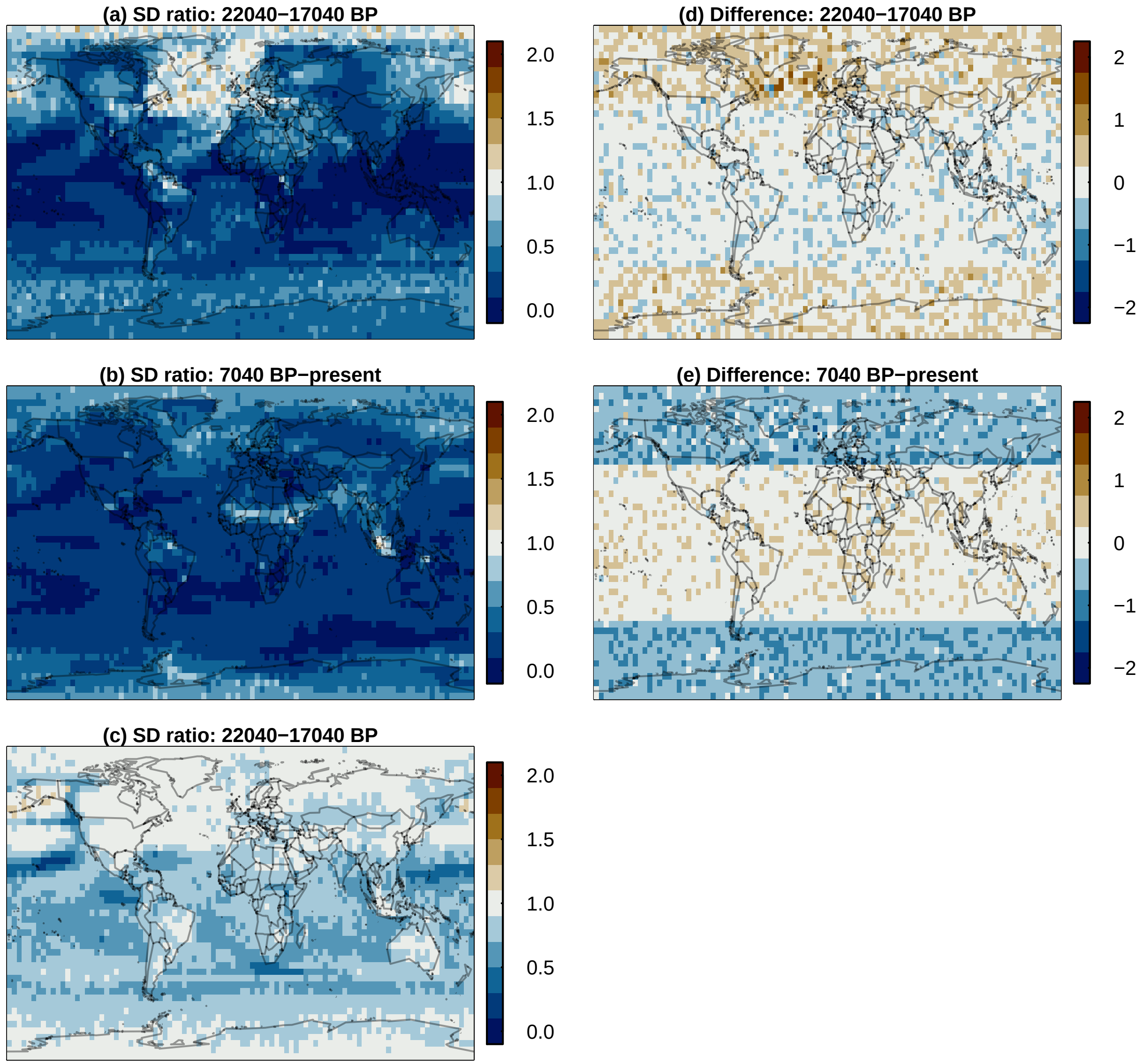


Figure 8. ~~Top~~Left, ~~logarithm of~~ standard deviation ratios of the sampled sampled 501-year moving mean input data relative to the smoothed record plus AR(1)-process and the effective dating error and the measurement noise for the samples in the first 5,000 years of record ~~on the left and~~ (a), the last 7,040 years of the record ~~on~~ (b), and the ~~right~~full record (c). ~~Bottom~~Right, 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~~ (d) and the last 7040 years of the record ~~on the right~~(e).

~~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~~ the bias term in its current version may have only a small influence at certain latitudes. Third, our noise model shows ~~generally the largest effect~~ often larger effects in the mid latitudes and the tropics. ~~However, there~~ There is also a longitudinal dependence.

Supplementary document 2 Figure 1 (https://doi.org/10.17605/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 for a small selection of grid-points, highlighting the common lack of a clear relation besides the deglaciation.

Figure 7 ~~gives~~ provides correlation coefficients between the sampled ~~original~~ interannual 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 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 weakin the high latitudes. Correlations are largest in later periods~~In the period until present, correlations are overall weak. High latitude correlations are most notable during the deglaciation and slightly less notable during the first millennia of the Holocene. In these periods, correlations appear to be largest in areas with glacial remnants.

Figure 8 adds for the first~~and lastperiod the logarithm of~~, the last, and the full period the relative standard-deviation σ_{T21k}/σ_P in the ~~top row~~ left column and the bias $\bar{T}_{T21k} - \bar{T}_P$ in the ~~bottom row~~right column. T21k refers to the simulation, P to the pseudoproxies. For the standard deviation ratios, we use 501-year moving averages of the TraCE-21ka data. Variability is generally larger in the pseudoproxies except for the North Atlantic and the northern high latitudes in the early period, and it is larger in the pseudoproxies more or less everywhere in the late period. Over the full period, variability is notably larger mainly 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~~, it is about equal over Antarctica and wide regions of the northern Hemisphere. The variability is clearly larger in the input data only over a small region in the northern Pacific.

The overall largest bias occurs off the coast of southeastern Greenland in the early period in Figure 8. Otherwise there is a spatial separation between the mid- to high latitudes and the tropics and subtropics for both periods. The bias is more prominent in the higher latitudes where it is predominantly positive in the early period but predominantly negative in the late period. Obviously, the general latitudinal bias pattern is by construction because we construct the bias as function of latitudinal insolation.

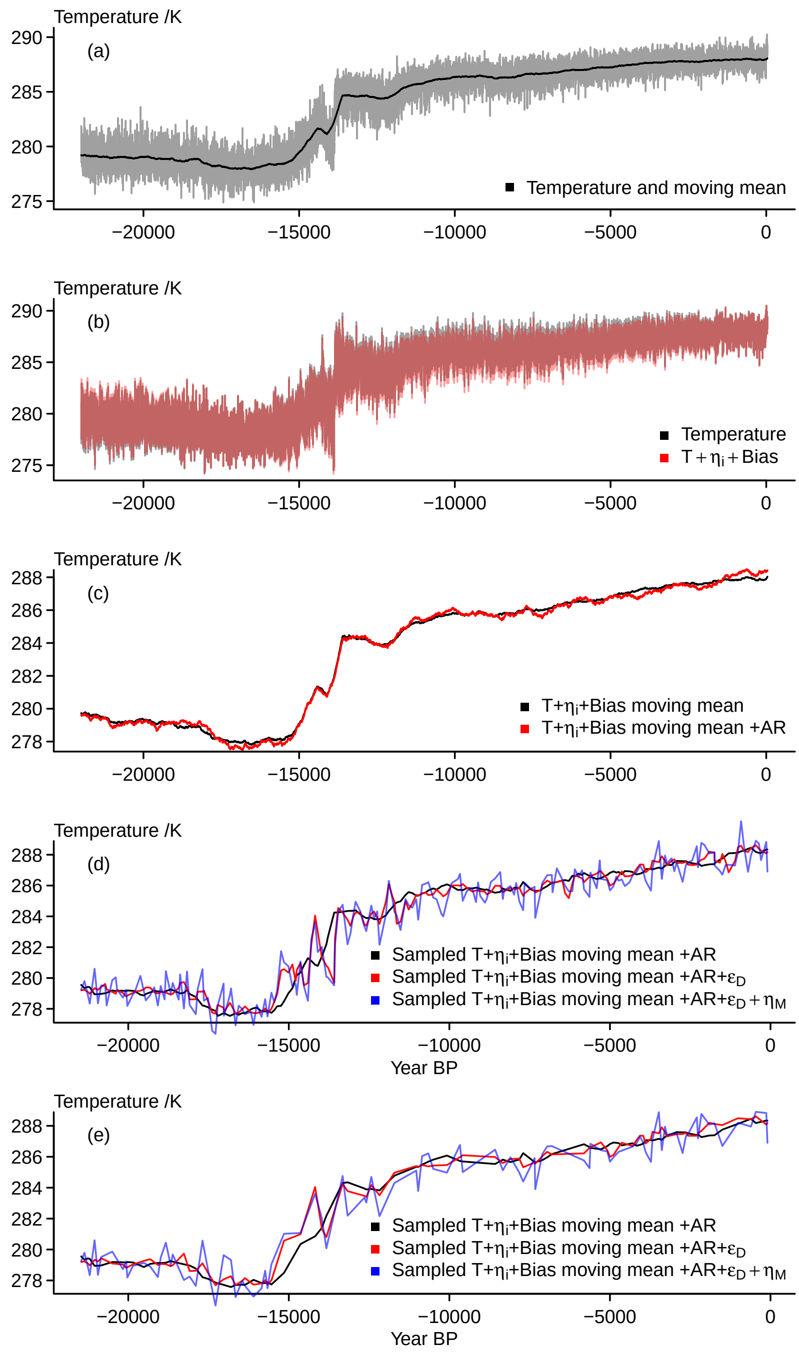


Figure 9. Visualising the simplified essence of the surrogate proxy calculations: a) input data and 501-point moving mean, b) input data plus initial noise and bias term, c) moving mean of input data plus noise plus bias and the same record plus an AR(1)-process, d) smoothed temperature plus noise plus 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.

3.5 On generalizations of the errors

3.6 ~~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~~to a higher degree. 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 evaluations and model-data comparisons. Most importantly, such a generalized approach also allows to quickly produce ensembles of pseudoproxies.

~~The~~ Following our previous assumptions, 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~~ Here, we use an AR-process of order one 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.

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 non-linearities between the climate and sensor and thus a truly non-linear pseudoproxy~~In short, the generalized pseudoproxy becomes:

$$M_T(t_1,...,t_N) = g(T(t) + \eta_i t + Bias(t)) + \epsilon_D(t_1,...,t_N) + \eta_M \tag{21}$$

where g is the smoothing, η_i is the initial noise, $Bias$ is the bias term, ϵ_D is the effective dating error, and η_M is the measurement error. This is conceptually identical to the smoothing plus AR approach presented above. Its derivation is less grounded in real proxies. The provided data differs only in the amount of autocorrelation in the noise terms.

Figure 9 summarises results for ~~our~~ the 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~~ if we take the assumptions made above ~~are~~ to be 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 7 and 8~~

While researchers may validly wish for such simplified recipes for producing pseudoproxies, using a full or at least more complex process-based approach is advisable, if it is necessary to account for effects of biology, environmental long-term changes, and other weakly constrained uncertainties. More complex approaches further allow to better mimic non-linearities between the climate and sensor and thus a truly non-linear pseudoproxy.

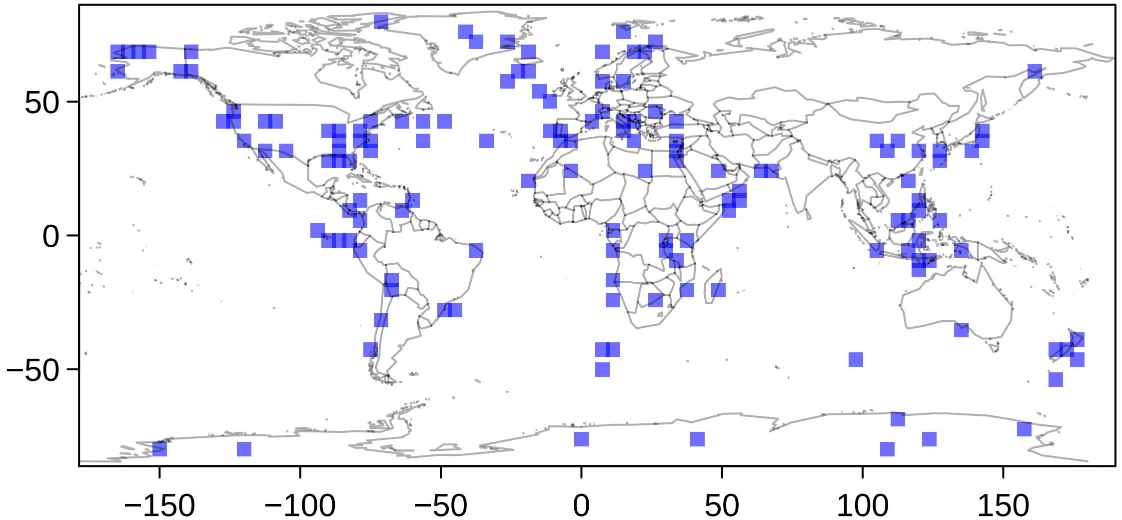


Figure 10. Map of the locations for the ensemble of surrogate proxies.

3.5.1 Ensemble of Pseudoproxies

~~We make further slight modifications to our approach to obtain~~

In the following we present an ensemble of ~~500 pseudoproxies at~~ pseudproxies. At 144 locations ~~. These~~ we compute 500 pseudoproxy records each. For this, we make slight modifications to the generalized approach. These adjustments relax our assumptions and result in larger differences between members of the ensemble than would be possible without the modifications. The 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)~~Shakun et al. (2012), Clark et al. (2012), or Marcott et al. (2013). Figure 10 shows the locations.

~~Map of the locations for the ensemble of surrogate proxies.~~ Using the generalized approach provides an ensemble based on the most reduced formulation. The provided code allows users to produce ensembles for their input data of interest.

Modifications to the code are ~~, for one~~as follows: First, we use a number of ~~parameters'~~ parameter values sampled from either uniform distributions around the otherwise fixed value or from a list of values. Second, we consider ~~the series S~~ random orientations for bias and moving standard ~~deviation~~ deviations, that is we take S as S^u where we sample u from $U = \{-1, 1\}$. ~~The~~ We provide the script for the ensemble production as supplementary example code at ~~highlights these differences~~https://doi.org/10.17605/OSF.IO/ZBEHX. As mentioned above, these changes relax our assumptions on the effect of changes in the background climate.

~~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.~~

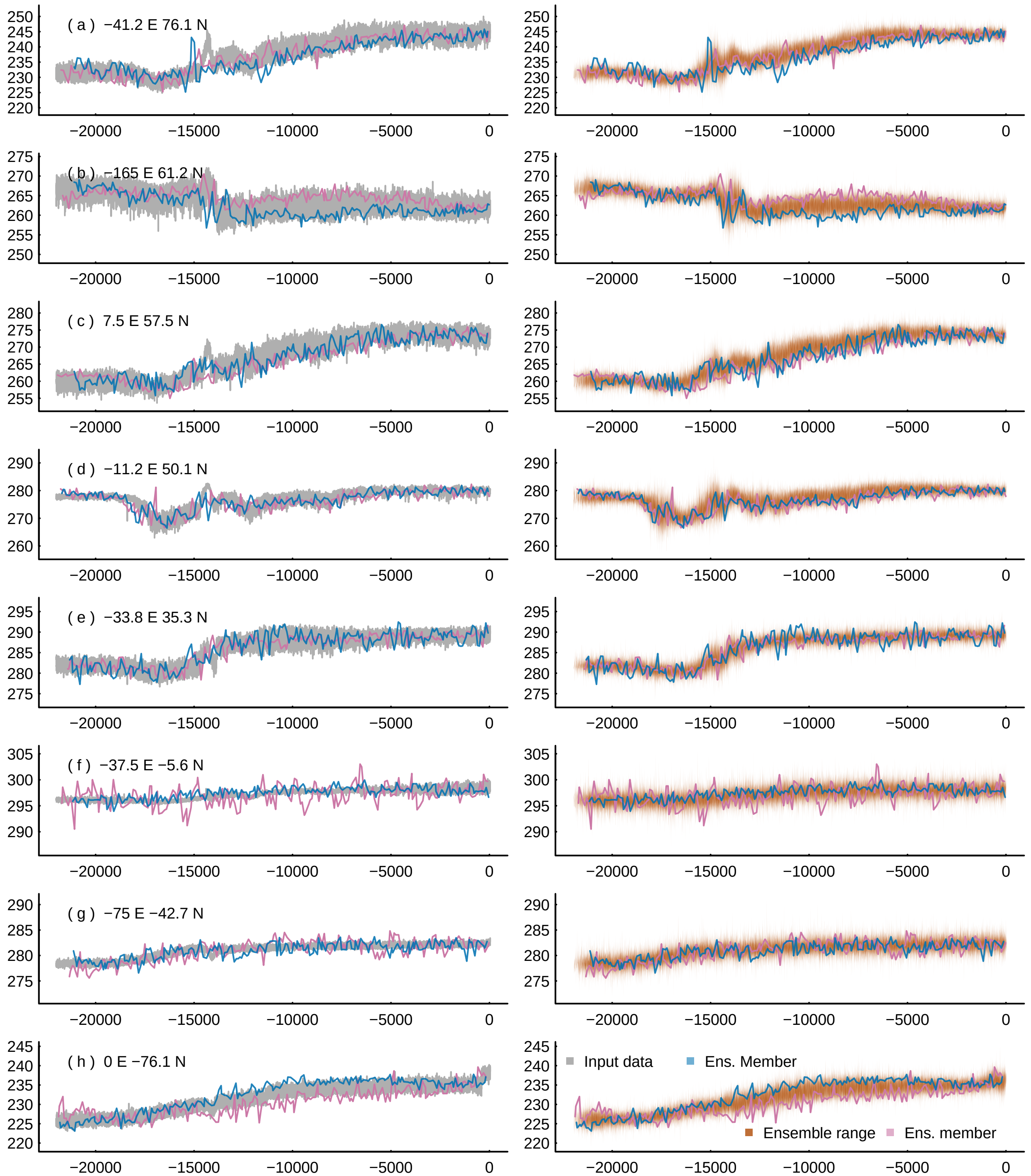


Figure 11. Visualising the surrogate proxy-ensemble at selected locations (Longitude and Latitude in top left corners of the left column panels): The left column shows the input data plotted as grey lines, and two random members of the ensemble as blue and purple lines. The right column plots the range of the ensemble transparently brown shaded, and blue and purple lines are the same two random members. The x-axes are years BP. The panel on the bottom right shows the Figure legend.

For Figure 11 we select ~~twenty~~ 8 locations to represent the locally diverse representations of the climate in the TraCE-21ka simulation 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 11 mainly reflects the results of previous ~~Sections~~ sections 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~~Provided Data

Tables 1 to 4 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. The data repository at https://doi.org/10.17605/OSF.IO/ZBEHX provides instructions how to access the file structures.

4 Conclusions and outlook

~~We present in this document, the associated code, and the provided data~~ This publication presents a flexible yet simple approach for describing the ~~non-climatic error~~ error originating from climatic and non-climatic sources in proxy-records over multi-millennial time-scales including the last deglaciation. The assumptions are ~~simplistic but base upon~~ relatively simple but they are based on similar assumptions for process-based 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~~ 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). The provided pseudoproxy data, and the code to compute further pseudoproxies allows the application of our pseudoproxy-approach for the evaluation of models, the comparison of models to paleo data, and the testing of reconstruction and data-assimilation methods.

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~~ face larger biases in reality than represented by our choice. ~~Every researcher should make his~~ Users should make their own choice of parameters according to ~~his~~ their 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

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_ and the ending _annual.nc

Filename	Variable-Name	Variable-Description	Category	Grid size	Ensemble size
Bothe_Trace21k_Pseudo_Proxies_					
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_ and the ending _annual.nc

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

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_ and the ending _annual.nc

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

Table 4. 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_ and the ending _annual.nc

Filename
~~samp.meas.noise.subsampled.save.nc samp.meas.noise.subsampled.save sampled seasonally subsampled + noise plus measurement noise field 96x48 1unc.date.meas.noise~~
Essence_noise.gen.dat ~~.nc~~
Essence_bias.noise.gen.dat ~~.nc~~
Essence_smooth.bias.noise.gen.dat ~~.nc~~
Essence_ar.smooth.bias.noise.gen.dat ~~.nc~~
Essence_~~sampled.nc~~ uncertain-sampled
essence_ensemble ~~.nc~~
~

do not consider spatial correlations in the noise~~generally and between different locations~~. Such correlations between locations 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).~~ for example, Breitenbach et al. (2012) or Goswami et al. (2014). Neither, does our approach as of now explicitly link to probabilistic age-modelling approaches as described by Haslett and Parnell (2008), Blaauw and Christen (2011), or Trachsel and Telford (2017).

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, to 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.

The present approach ignores a variety of possible complications. For example, ~~there is not so far a method to include~~ we currently do not consider hiatusses in the sensor. Furthermore, the dependency on the background climate is ~~minimal~~small. Nevertheless, we are confident that this approach is of value for the comparison of simulation data and proxy data over long periods, for testing reconstruction methods, and for evaluating different model simulations against each other.

Table A1. List of parameters used.

Description	Parameter	Value	Category
Season limits for insolation bias	mon1.for.insol, mon2.for.insol	~~6, 8~~ 1, 12	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~~ Switch for proportionality of initial noise	switch.orient.runsd.noise.env	0	all
Model for the initial noise	model.noise.1	c(0.3)	field, essence
~~standard~~ Standard deviation of innovations ~~of~~ for initial noise	sd.noise.1	not used	field, essence
Length of window influencing initial noise	length.window.runsd	1000	field, essence
Switch for orientation of bias	switch.orient.bias.seas	0	all
Scaling of bias term	amp.bias.seas	4	field, essence
~~mean~~ height			

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 3.6 are available from the Open Science Framework (OSF) at https://doi.org/10.17605/OSF.IO/ZBEHX/. There, one also finds sample code for computing proxies and the script for computing the ensemble at 144 locations.

Appendix A: Tables of parameters

Tables A1 to A4 summarise the considered parameters and noise models. They also clarify whether the ~~parameters~~ parameter settings are used for a global field of surrogate proxies, a more generalized approach, an ensemble calculation, or all.

Table A2. Continuation of list of parameters used.

Description	Parameter	Value	Category
Switch for smoothing variant	switch.smoothing	3	field
Secondary switch for smoothing, see code	switch.sm.2	1	field
Scaling for climate dependence of smoothing	scale.sm	1/10	field
Mean smoothing length for longer random smoothing	rand.mean.length.smooth	350	field ~~, essence~~
~~standard~~ Standard deviation for longer random smoothing	rand.sd.length.smooth	75	field ~~, essence~~
~~fixed~~ Model for longer alternative smoothing	model.smooth.1	c(0.99)	field
Model for longer alternative climate dependent smoothing	model.clim.smooth.1	c(0.9)	field
Basis long smoothing length for alternative approach	rand.length.smooth.mean.1	500	field
Standard deviation for longer alternative smoothing approaches	sd.model.smooth.1	10	field
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~~ Mean smoothing length for ~~longer~~ shorter smoothing	rand.mean.length.smooth.2	31	field ~~, essence~~
~~standard~~ Standard deviation for shorter random smoothing	rand.sd.length.smooth.2	5	field ~~, essence~~
~~fixed~~ Model for shorter alternative smoothing	model.smooth.2	c(0.7)	field
Model for shorter alternative climate dependent smoothing	model.clim.smooth.2	c(0.9)	field
Basis short smoothing length for alternative approach	rand.length.smooth.mean.2	31	field
Standard deviation for shorter alternative smoothing approaches	sd.model.smooth.2	4	field
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~~ height			

Table A3. Continuation of list of parameters used.

Description
Number of picked samples for subsampling
~~standard~~ Standard deviation of innovations for subsampling noise
~~model~~ Model of subsampling noise
1.96 ~~Sigma~~ sigma of measurement-noise
~~noise~~ Noise model for measurement noise
~~Continuation of list of parameters used. Description Parameter Value Categorynoise~~ Noise model for measurement noise for subsampled record
1.96 ~~Sigma~~ sigma for measurement noise for subsampled record
~~switch~~ Switch for correlated effective dating error
~~switch~~ Switch for weakly correlated only
~~switch~~ Switch for time dependent correlated
~~fixed~~ Fixed correlated dating error coefficient
~~mean~~ Mean of distribution of dating uncertainty
~~standard~~ Standard deviation of distribution of dating uncertainty
~~switch for proportionality of initial noise - not used~~ Switch for length of influence on dating uncertainty
~~switch for orientation of bias~~ Switch for date sampling
Switch for dating uncertainty sampling
Model for initial noise for generalized case
~~model~~ Model for initial noise for generalized case
~~standard deviation for initial noise~~ Standard deviation for AR process innovations, generalized case
~~smoothing~~ Smoothing length generalized case ~~, prescribed~~
~~alternative model for initial noise model.gen.noise c(0.7) essencealternative model for measurement noise - not used model.noise.meas.b c(0.35) essenc~~

Table A4. Continuation of list of parameters used.

Description	Parameter	Value
Ensemble size	size.ensemble	500
~~amplitude~~ Amplitude of scaling of initial noise	amp.noise.env	$U(0.4, 1.5)$
~~scaling~~ Scaling of bias	amp.bias.seas	$U(3, 10)$
~~standard~~ Standard deviation of measurement noise	lim.noise.meas	$U(0.75, 3)/1.959964$
~~ar-coefficient~~ AR-coefficient of measurement noise model	rand.model.coeff	$U(0.3, 0.8)$
~~ar-coefficient~~ AR-coefficient of initial noise model	rand.model.coeff.gen	$U(0.6, 0.8)$
~~window~~ Window of influence of background ~~climate - not~~ climate—not used	rand.width.background.sd	$U(500, 2000)$
~~window~~ Window of influence of background climate	rand.width.background.sd	1000
~~width~~ Width of window of filter influence	length.filter.uniform	~~l_{fil} is random sample from all~~ $\cancel{L = \{30}$ $\cancel{S_N = \{-1, 1\}}$ $L = \{301, 303, 305, ..., 100$

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~~ Anthropocene—Modeling a Complete Glacial Cycle. It is funded through the German Federal Ministry of Education and Research's (BMBF) Research for Sustainability initiative (FONA). We acknowledge discussions with Andrew Dolman, Thom Laepple, and Nils Weitzel as influences in our approach.

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.

Annan, J. D. and Hargreaves, J. C.: A new global reconstruction of temperature changes at the Last Glacial Maximum, Climate of the Past, 9, 367–376, https://doi.org/10.5194/cp-9-367-2013, 2013.

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

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~~PRoxY System Modeling, with applications to oxygen-isotope systems, ~~J. Adv. Model. Earth Syst., p. n/a, ,~~ Journal of Advances in Modeling Earth Systems, 7, 1220–1247, https://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, ~~,~~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., Reichert, B. K., Kaplan, A., Anchukaitis, K. J., Vaganov, E. A., Hughes, M. K., and Cane, M. A.: A forward modeling approach to paleoclimatic interpretation of tree-ring data, Journal of Geophysical Research: Biogeosciences, 111, G03 008+, https://doi.org/10.1029/2006JG000166, 2006.

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, ~~,~~ 2013.

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

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.

Graham, N. and Wahl, E.: Paleoclimate Reconstruction Challenge: Available for participation, PAGES news, 19, 71–72, https://doi.org/10.1029/2009PA001758.For, 2011.

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

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.

Hind, A., Moberg, A., and Sundberg, R.: Statistical framework for evaluation of climate model simulations by use of climate proxy data from the last millennium - Part 2: A pseudo-proxy study addressing the amplitude of solar forcing, Climate of the Past, 8, 1355–1365, https://doi.org/10.5194/cp-8-1355-2012, 2012.

Jones, P. D., Briffa, K. R., Osborn, T. J., Lough, J. M., Van Ommen, T. D., Vinther, B. M., Luterbacher, J., Wahl, E. R., Zwiers, F. W., Mann, M. E., Schmidt, G. A., Ammann, C. M., Buckley, B. M., Cobb, K. M., Esper, J., Goosse, H., Graham, N., Jansen, E., Kiefer, T., Kull, C., Küttel, M., Mosley-Thompson, E., Overpeck, J. T., Riedwyl, N., Schulz, M., Tudhope, A. W., Villalba, R., Wanner, H., Wolff, E., and Xoplaki, E.: High-resolution palaeoclimatology of the last millennium: A review of current status and future prospects, Holocene, 19, 3–49, https://doi.org/10.1177/0959683608098952, 2009.

Konecky, B., Dee, S. G., and Noone, D.: WaxPSM: A forward model of leaf wax hydrogen isotope ratios to bridge proxy and model estimates of past climate, Journal of Geophysical Research: Biogeosciences, p. 2018JG004708, https://doi.org/10.1029/2018JG004708, 2019.

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, 113, E1434–E1441, https://doi.org/10.1073/pnas.1517056113, ~~,~~ 2016.

Kurahashi-Nakamura, T., Losch, M., and Paul, A.: Can sparse proxy data constrain the strength of the Atlantic meridional overturning circulation?, Geoscientific Model Development, 7, 419–432, https://doi.org/10.5194/gmd-7-419-2014, https://www.geosci-model-dev.net/7/419/2014/, 2014.

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.

Lehner, F., Raible, C. C., and Stocker, T. F.: Testing the robustness of a precipitation proxy-based North Atlantic Oscillation reconstruction, Quaternary Science Reviews, 45, 85–94, https://doi.org/10.1016/j.quascirev.2012.04.025, 2012.

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, 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.

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, ~~,~~ 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.

~~Press, W., Teukolsky, S., Vetterling, S~~

Mathias, A., Grond, F., Guardans, R., Seese, D., Canela, M., and ~~Flannery, B~~Diebner, H. H.: ~~Numerical Recipes in C:The Art of Scientific Computing~~Algorithms for Spectral Analysis of Irregularly Sampled Time Series, Journal of Statistical Software, 11, ~~Cambridge University Press, Cambridge, Cambridge, 1994.~~ 1–27, https://doi.org/10.18637/jss.v011.i02, 2004.

Nick McKay, Matthew Graham, Feng Zhu, and Julien Emile-Geay: https://github.com/nickmckay/nuspectral, https://github.com/nickmckay/nuspectral, last accessed 2019-03-11.

Osborn, T. J. and Briffa, K. K.: The real color of climate change?, Science, 306, 621–622, https://doi.org/10.1126/science.1104416, 2004.

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, 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~~

Schmidt, G. A.: Forward modeling of carbonate proxy data from planktonic foraminifera using oxygen isotope tracers in a global ocean model, Paleoceanography, 14, 482–497, https://doi.org/10.1029/1999PA900025, 1999.

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, 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, 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.

Steiger, N. and Hakim, G.: Multi-timescale data assimilation for atmosphere-ocean state estimates, Climate of the Past, 12, 1375–1388, https://doi.org/10.5194/cp-12-1375-2016, 2016.

Thompson, D. M., Ault, T. R., Evans, M. N., Cole, J. E., and Emile-Geay, J.: Comparison of observed and simulated tropical climate trends using a forward model of coral δ18O, Geophysical Research Letters, 38, L14 706+, https://doi.org/10.1029/2011gl048224, 2011.

Tolwinski-Ward, S. E., Evans, M. N., Hughes, M. K., and Anchukaitis, K. J.: An efficient forward model of the climate controls on interannual variation in tree-ring width, Climate Dynamics, 36, 2419–2439, https://doi.org/10.1007/s00382-010-0945-5, 2011.

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

Vardaro, M. F., Ruhl, H. A., and Smith, K. L.: Climate variation, carbon flux, and bioturbation in the abyssal north pacific, Limnology and Oceanography, 54, 2081–2088, https://doi.org/10.4319/lo.2009.54.6.2081, 2009.

Von Storch, H., Zorita, E., Jones, J. M., Dimitriev, Y., González-Rouco, F., and Tett, S. F.: Recontructing past climate from noisy data, Science, 306, 679–682, https://doi.org/10.1126/science.1096109, 2004.