Clim. Past Discuss.,
https://doi.org/10.5194/cp-2018-178-AC2, 2019

Climate of the Past Discussions
Open Access
EGU

Interactive comment on "Joint inversion of proxy system models to reconstruct paleoenvironmental time series from heterogeneous data" *by* Gabriel J. Bowen et al.

Gabriel J. Bowen et al.

gabe.bowen@utah.edu

Received and published: 19 May 2019

Reviewer 1 suggests several potential refinements to the model formulation used here, requests clarification of some methodological details, and offers some suggestions for refining the presentation of our work. Several of these suggestions have potential to make incremental but important improvements to our work, and in almost all cases we think the suggestions will make our manuscript presentation clearer and better. None of the suggestions imply any fundamental problem with the study, nor do we anticipate they will materially change the results or take-home messages presented in our original submission. We propose to leverage the reviewer's input to improve the manuscript,

Printer-friendly version

Discussion paper

as descried below.

The reviewer comments are shown here following the hash sign (#).

A Bayesian Hierarchical Model is used to reconstruct several environmental variables using some proxy variables, at three marine sites. The method isn't new e.g. Garreta et al. (2010), but the manuscript is useful as another example of this type of approach. The manuscript is unclear in places.

We're glad the reviewer finds value in the paper. As acknowledged in the text and citations the approach here is not without precedent. We'll add the Garreta example, another interesting one using a much more complex (dynamic vegetation) process model to our reference list in the revision.

In S2.3 (Section 2.3) Eq. 5 is a stationary discrete-time first-order autoregression (AR1) model for the random walk disturbances eY . According to S2.3, for BWT and d18Osw this AR1 model is run at a time step of 50 kyr (site 806) and 1 kyr (site 1123 and U1385). For Mg=Casw, the manuscript states "1 Myr time steps from 80 Ma to present", but what about for the higher-resolution sites (for Mg/Casw)? Also it is not clear what happens when the model time steps are different from the marine proxy time steps (which are irregular, S2.1 paragraph 2) - this point needs to be clarifed. It would be good to have a graphical depiction of the method (e.g. included in Fig. 1), with example time series (with clear time points) showing eY, Y, modelled proxy time series e.g. d18Of, and observed time series. Just show a portion of the time series, so that the time points for each time series are clear. # Further, instead of a discrete-time model, why not use a continuous-time model, which handles irregular time steps better than a discrete-time model. For example, a continuous-time time AR1 model is: (equation) For a continuous-time AR model, the parameters are not a function of the sampling intensity (Tomasson, 2015).

We thank the reviewer for this suggestion, and in preparing our revision will explore the idea of using a continuous-time AR1 model. This would have clear benefits (including eliminating the need for interpolation, discussed below) and assuming it is feasible for the study systems we'll intended to adopt this tweak in the revision. We do not anticipate that the change will have a strong (or maybe even detectable) impact on the key outcomes. We will also work with the reviewer's idea of illustrating the time-series model properties through an addition to figure 1, which we like in theory and will do our best to implement without compromising the clarity/focus of the figure.

S2.4 paragraph 3: "we conducted three different analyses ... the third inverting both records together." It's not clear what is meant by the latter phrase. An example of a discrete-time vector correlated random walk model is: (equation) . . . Exactly what model was used in "inverting both records together", together with an explanation of why it would be mathematically different from "inverting data from each site independently" needs to be included in the manuscript text. Further there are vector continuous-time series models, which might be better to use for inverting multiple time series with irregular time intervals.

The presentation here is somewhat ambiguous, and will resolve this in the revision. Our analysis remains agnostic of the correlation structure between the paleoenvironmental state variables at the two sites. They are modeled as independent time series, with no correlation term. We recognize that alternative models, such as that proposed by the reviewer, would allow incorporation of additional prior information and perhaps provide stronger process model constraints on the paleoenvironmental time series, acknowledging that they are likely not truly independent. However, the model proposed by the reviewer is just one step along a continuum of model forms that one could apply which would, at its end point, lead to a climate system model that expressed a full set of physics-based expectations for the relationship between the environmental state variables at the sites. While we acknowledge the potential value in such an analysis (which would basically become a data assimilation analysis), we are proposing and exploring a framework that lies at the other end of the continuum. Our goal is to offer a widely applicable and approachable framework in which practitioners who

Printer-friendly version

Discussion paper

already routinely develop quantitative interpretations of their data without reference to any formal statistical framework or paleoenvironmental model can begin to adopt such without compromising the data-driven nature of their interpretations or having to frame them in the context of the complexity and structural assumptions of more complex paleoenvironmental models.

S3.2 paragraph 1: In statistics, the idea of smoothing (whether by frequentist or bayesian methods) stems from the idea that a time series = state variable + noise. Looking at Lear et al. (2015), the L15 reconstruction appears to have a higher variability simply because there was no smoothing employed. A better comparison here would be to create a reconstruction using both frequentist smoothing and bayesian smoothing methods, and then compare. The current comparison here seems a bit apples and oranges.

Indeed, the crux of this comparison is smoothing, and based on the reviewer's feedback we propose to emphasize this more clearly with revisions to the language in this part of the discussion. We prefer not to present this as a comparison of Bayesian and frequentist smoothing techniques, however. The crux of our paper is not to enter into the Bayesian vs. frequentist discussion. Instead, we are trying to present an alternative to the reconstruction methods used nearly ubiquitously in the (pre-Holocene paleoclimate) community (and honestly most Holocene work), which do not embrace or consider concepts such as smoothing, multi-variate proxy models, or temporal autocorrelation of environmental timeseries. Our point in this section is that 1) smoothed reconstructions are a more realistic/honest expression of the information contained in proxy timeseries records, and that 2) the method demonstrated here offers an approach to optimize the properties of the smoothed reconstruction based on the data, rather than adopting an ad hoc approach (e.g., splines or running averages with arbitrarily specified parameters) as is commonly done if and when smoothing is conducted. The comparison is apples to oranges, but we think also of value.

S3.2 paragraph 3: So if d18Osw and BWT are generated at 1 kyr time steps, and

the sampling resolution of d18Of is between 1 per 110 and 1 per 1700 years, do you generate the model time series first, at 1 ka steps, and then use Eq. 5 to “integrate" to the proxy time points (if necessary)? How is that integration done?

In our original analysis, values for proxy time points are obtained using a ‘nearest neighbor’ approach, i.e. the value at the nearest proxy time series point is used. We will clarify and discuss/justify this in the revision if we end up maintaining a discrete time series model approach, or if we adopt the continuous AR1 model this will become a moot point.

S3.2 paragraph 4: The following sentence could be worded better: “Moreover, because temporal autocorrelation of the environmental variables is considered ...". I think you are trying to say its both the autocorrelation (in the environmental states) and sample density which make the credible intervals what they are. In the next sentence, can you explain mathematically what is meant by “the strength of the proxy constraints"?

Yes, we can work on rewording/elaborating to clarify as requested.

S3.3 paragraph 2 (“These refinements reflect ...") After 800 ka, perhaps the higher proxy model variance is suggesting the environmental model is missing something? For example, what would be the effect of adding a stochastic periodic component to the process model to capture the 100 ka cycle after 800 ka?

Absolutely, this conclusion is essentially what we were suggesting here. Adapting the process model would be a good, perhaps more appropriate, alternative to adapting the data model across the 800 kya boundary, and we propose to explore this alternative in preparing our revision. This will depend a bit on whether the sampling resolution of the site 806 data is adequate to constrain the sub-100 kyr variability in this interval, in which case it makes sense to treat it as ‘signal’ (i.e. in the process model) or ‘noise’ (i.e. as done, in the error term of the data model).

S3.3 paragraph 3 The phrase “double-count uncertainty associated with correlated

parameters" is not an elegant mathematical explanation.

With all respect, and acknowledging the suggestion, we are not writing this for mathematicians but rather for paleoclimate practitioners. Here we are attempting to provide a common-language explanation of some of the contrasts between the proposed approach and common practice. This phrase may be somewhat imprecise, but we think it makes the point in a way that most readers will grasp it.

It's unclear exactly how the dotted blue line in Fig. 8a is calculated. Explain. Also statistical tests don't always need to assume independence, because there are ways of accounting for autocorrelation in a statistical test. # "The net result in this case ... some 100-200 kyr earlier using the traditional approach": would this sentence be true if autocorrelation was taken into account in the traditional approach. I'm looking for a fair comparison here. # Also for the solid blue line in Fig 8a - give details of its calculation.

We will happily elaborate/be more specific on the calculation of the 'traditional' analyses presented in the figure. In the original draft we had erred on the side of brevity in an attempt to interrupt the flow during the latter part of the manuscript. We see how this compromises the clarity of the analysis, however, and will revise to ensure the calculations are described in enough detail to be reproducible from the text alone (i.e. not requiring reference to the data analysis code, which is already publically available and fully documents the details). With respect to autocorrelation – indeed this is the crux of the difference noted in figure 8. We will try to make this clearer/provide greater emphasis in the revision. Akin to the comment above on smoothing, our point here is that the JPI framework integrates explicit treatment of time series autocorrelation, ensuring that data interpretations developed from the method reflect a robust consideration of such, unlike many analyses presented in the literature. There are other ways of achieving this, of course, and we'll be sure to better make that point in our revision.

p4 L13: "sampling resolution between 1 per 110 and 1 per 1700" years. Clarify for d18O, Mg/Ca, or both?

This summarizes across both proxies, and we will clarify in the revised text.

p5 L11-12: The Evans et al. (2013) terminology includes “sensor models", “archive models", and “observation models". Clarify which of your equations relate to which type of Evans’s models?

I personally have struggled with this, as I don’t think there is a 1:1 mapping in this case. Part of the issue is that the Evans et al conception includes a strong focus on the processes that integrate proxy responses in a biological or sedimentary medium that accumulates over time (e.g., sediment stack, incremental growth structure; archive model) and how those integration processes are sampled (observation model). These are not explicitly treated here, or in many proxy interpretations, and in some lower-resolution deep time studies may be less critical than in much of the higher-resolution shallow-time work (I’ll note, however, that I’m not sure I actually believe this. . .it is a frontier area and there are now a handful of really interesting avenues being pursued, e.g., with respect to processes such as seasonal sampling of different proxy archives and the impact of sedimentary architecture and allogenic processes on signal integration/preservation). At any rate, what we have here, in my interpretation, is primarily a sensor model, which also embeds some aspect of what would appear in archive and observations models in the proxy model error term. We will state this (more concisely that I have here!) in the revision.

p5 L17: “age estimate and uncertainty" Ambiguous wording, because as is it reads “age estimate and age uncertainty".

We will reword as suggested.

Eq. 2 and 3: For clarity, can you make all the “functions" with round brackets e.g. BWT(tMgCaf [i]). Change the outer brackets too i.e. {}. Keep square brackets for distributions e.g. N[], as you have done.

We will reformat the equations as suggested.

Eq. 5 Say what Y can be e.g. Y (t) can be MgCasw(t) or d18Osw(t) or BWT(t).

We will elaborate as suggested.

p8 L2: Clarify the phrase “stiff" time series behaviour (give a reference if possible)

We will do some literature research to see if we can come up with a more formal way to express this result. . .we were trying here to colloquially express the condition in which error variance is small and error autocorrelation large, as for the Mg/Ca_sw, which leads to long burn-in times using most methods for generating MCMC step proposals.

p12 L8: “Across all scales": Across all sites?

We intended ‘across all timescales’, but the suggested ‘across all sites’ would probably be clearer and will be adopted in the revision.

Additional figures showing the calibration datasets, with individual draws from the posterior distribution, should be included. These could go in the manuscript or supplementary material.

We can easily add these to the SI in the revision, and are happy to do so.

Fig. 5: Which inversion did these distributions come from e.g. site 806? (include in caption) # Fig 6: same comment as Fig. 5.

We will clarify that these are from the 806 analysis and also indicate the taxon to which they apply in the revised figure legends.

Fig. 7: The prior distributions in (d) and (g) don’t integrate to 1.0 e.g. 2.5 x 0.8 = 2.0. I can’t tell if all the other distributions integrate to one or not.

Thanks for catching this. . .it is a plotting error (we carried over the y-axis value of 2.5 appropriate to the prior in panel a) which we will correct in the revision.

Figure 9: There is a positive relationship between DBWT and Dd18Osw in the two Miocene states (mentioned in the last sentence in S3.4). I think adding some straight

Printer-friendly version

Discussion paper

lines to mark this, and not inverting the y-axis here would help the reader.

We will certainly un-invert the y axis as suggested (appropriate here since we are plotting d18O of seawater and not carbonate). We will also explore ways of adding lines that represent the correlation in posterior values from different states.

Interactive comment on Clim. Past Discuss., https://doi.org/10.5194/cp-2018-178, 2019.