Review

Manuscript: Joint inversion of proxy system models to reconstruct paleoenvironmental time series of heterogeneous data
Authors: G. J. Bowen, B. Fisher-Femal, G.-J. Reichart, A. Sluijs and C. H. Lear
Manuscript Number: CP-2018-178

General comments

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.

Specific comments

1) *Section 2.3 Environmental models*

In S2.3 (Section 2.3) Eq. 5 is a stationary discrete-time first-order autoregression (AR1) model for the random walk disturbances ϵ_Y. According to S2.3, for BWT and $\delta^{18}O_{sw}$ this AR1 model is run at a time step of 50 kyr (site 806) and 1 kyr (site 1123 and U1385). For Mg/Ca_{sw}, the manuscript states "1 Myr time steps from 80 Ma to present", but what about for the higher-resolution sites (for Mg/Ca_{sw})? 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 clarified. 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 ϵ_Y, Y, modelled proxy time series e.g. $\delta^{18}O_f$, 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:

$$\epsilon_Y(t) = \mathrm{N}[e^{-\alpha\tau}\epsilon_Y(t-1),\ \frac{\sigma^2}{2\alpha}(1-e^{-2\alpha\tau})]$$

where $\tau = t_i - t_{i-1}$. For a continuous-time AR model, the parameters are not a function of the sampling intensity (Tomasson, 2015).

2) *Section 2.4 Model inversion*

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:

$$\mathbf{y}_t = \mathbf{y}_{t-1} + \epsilon_t \tag{1a}$$

$$\epsilon_t = \begin{bmatrix} \mathbf{A}^{(1)} & \mathbf{B} \\ \mathbf{C} & \mathbf{A}^{(2)} \end{bmatrix} \epsilon_{t-1} + \mathbf{e}_t \tag{1b}$$

where $\mathbf{A},\ \mathbf{B},\ \mathbf{C}$ are 2×2 matrices, $\mathbf{e} \sim \text{MVNormal}(\mu, \mathbf{\Sigma})$, $\mathbf{y}_t = \begin{bmatrix} \delta^{18}O_{sw}^{(1)} \\ BWT^{(1)} \\ \delta^{18}O_{sw}^{(2)} \\ BWT^{(2)} \end{bmatrix}$, and the superscripts denote sites 1 and 2. In this example, I've ignored $MgCa_{sw}$.

The matrices $\mathbf{A}^{(1)}$ and $\mathbf{A}^{(2)}$ control the intra-site auto- and cross-correlation in the disturbances, while $\mathbf{B}$ and $\mathbf{C}$ control inter-site cross-correlation in the disturbances. In inverting the above model, some matrices (or their elements) could be set to zero. 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.

3) *Section 3.2 Time series properties*

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.

S3.2 paragraph 3: So if $\delta^{18}O_{sw}$ and BWT are generated at 1 kyr time steps, and the sampling resolution of $\delta^{18}O_f$ 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?

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"?

4) *Section 3.3 Model properties*

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?

S3.3 paragraph 3 The phrase "double-count uncertainty associated with correlated parameters" is not an elegant mathematical explanation.

5) *Section 3.4 Derivative analyses*

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.

Technical corrections

p4 L13: "sampling resolution between 1 per 110 and 1 per 1700" years. Clarify for δ^{18}O, Mg/Ca, or both?

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?

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

Eq. 2 and 3: For clarity, can you make all the "functions" with round brackets e.g. $BWT(t_{MgCaf}[i])$. Change the outer brackets too i.e. $\{\alpha_1 \times \alpha_2 \times BWT(\cdot)\}$. Keep square brackets for distributions e.g. $N[\cdot]$, as you have done.

Eq. 5 Say what Y can be e.g. $Y(t)$ can be $MgCa_{sw}(t)$ or $\delta^{18}O_{sw}(t)$ or $BWT(t)$.

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

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

Figures

- 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.

- Fig. 5: Which inversion did these distributions come from e.g. site 806? (include in caption)

- Fig 6: same comment as Fig. 5.

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

Figure 9: There is a positive relationship between ΔBWT and $\Delta\delta^{18}O_{sw}$ in the two Miocene states (mentioned in the last sentence in S3.4). I think adding some straight lines to mark this, and not inverting the y-axis here would help the reader.

References

Garreta, V., Miller, P., Guiot, J., Hely, C., Brewer, S., Sykes, M., Litt, T., 2010. A method for climate and vegetation reconstruction through the inversion of a dynamic vegetation model. Climate Dynamics 35, 371–389. doi:10.1007/s00382-009-0629-1.

Tomasson, H., 2015. Some computational aspects of gaussian CARMA modelling. Statistics and Computing 24, 375–387. doi:10.1007/s11222-013-9438-9.