Clim. Past Discuss., 11, C2493–C2502, 2015
www.clim-past-discuss.net/11/C2493/2015/

Climate of the Past Discussions
Open Access

Interactive comment on "The Paleoclimate reanalysis project" *by* S. A. Browning and I. D. Goodwin

S. A. Browning and I. D. Goodwin

stuart.browning@mq.edu.au

Received and published: 24 November 2015

In response to the main comments by Referee 3 we have revised the manuscript to clarify the way in which forcings are accounted for in PaleoR, expanded the discussion on uncertainty, and updated one of the SST proxies that was incorrectly represented. We have also provided a detailed justification for our decision to excluded proxy data from time periods where their climate signal is ambiguous.

1 General comments The authors chose to reconstruct past climates using an analogue approach. While this is a viable approach that is computationally inexpensive, it also is quite limited in the specific setup chosen for this study. In particular, the influence of external forcings is not taken into account in the reconstruction. This is *a fundamental limitation of the analogue approach set up chosen for this study that should be discussed more prominently in the manuscript. Also the choice to use all available years for analogue selection seems problematic in the light of studying teleconnections in case these are in fact non-stationary. With changes in teleconnections as with changes in external forcings, BMA might be selected for the wrong reasons and thus not be representative in locations away from proxy information. If BMA spread is representative of reconstruction uncertainty, then BMA spread could be used to identify such cases (also see the comment on BMA spread below).*

Referee 3 makes several comments here: (1) that forcing is not accounted for and that this is a limitation of the study, (2) including all available years in the selection process is problematic due to non stationarity issues as the best matching analogues (BMA) might not be representative of locations away from the proxy data, and (3) the BMA spread could be used to identify such cases.

(1) Our approach implicitly accounts for forcing, and this is an advantage, not a limitation because we are not imposing a priori assumptions on the analogue selection. The LME produces a realistic response to external forcing (Otto-Bliesner et al. 2015; BAMS), however this forcing is superimposed on internal variability that no model can simulate in a manor that correlates to calendar years. PaleoR effectively reorders the simulated climate so it is temporarily consistent with proxy evidence. If we consider the past millennium in terms of 3 main forcing epochs: the MCA (950-1250; net positive forcing), the LIA (1600-1800; net negative forcing) and the anthropocene (1850-2000; increasing positive forcing and late 20th century ozone depletion), then we find that PaleoR does select the majority of BMA from the respective multi-centennial epochs: i.e. MCA BMA are drawn from the LME MCA.

(2) Likewise, PaleoR does not rely on any a priori teleconnection patterns. The PaleoR spatial field results from the spatial array of proxy data and is not biased towards individuals or pairs of remote proxies. Restricting the analogue pool for certain time periods would require developing a set of assumptions and imposing those on the ana-

CPD

11, C2493–C2502, 2015

Interactive Comment

logue selection process; to avoid this we use all available years. As described above, the majority of BMA are drawn from the correct epochs and are not chosen from epochs with substantially different forcings.

(3) This statement is correct; BMA spread is increased in regions of increased uncertainty. A more detailed discussion on uncertainty has been included in Section 4.2.

There is some ambiguity in the description of the method as to whether the reconstruction was performed on yearly or decadal data. As far as I understand, the results are only presented as decadal averages, but the actual analogue selection is performed using yearly data. This seems somewhat inconsistent, in particular as the rationale for using decadal averages is to mitigate dating issues with the proxy data. In the lights of increasing the signal-to-noise ratio in the final output (decadal averages), I suggest the authors select best analogues on decadal data. This would also clarify how the 50BMA of yearly data are combined to decadal averages (given that there is no temporal consistency between the 50BMA of year 1 and year 2, . . . estimating the variability of decadal averages from the pool of 50 x 10 BMA is not trivial).

Referee 3 is recommending that the analogue selection should be performed using decadally averaged proxy data, which is how we have done it. The method section (2.1 and 2.3) clearly states that analogue selection is based on P, which contains the decadal averaged normalised proxy signals, so we are not sure where this ambiguity arises.

Another arbitrary choice that needs further justification (or revision) is to limit the selection of analogues to robust climate signals in the proxies that exceed +/- 0.5 sigma. Thereby, the authors seem to suggest that there is information in knowing that a proxy time series is anomalous (at a given space-time location) whereas there is no information in knowing that the anomaly is close to zero. This unjustifiable criterion may lead to the fact that BMA are selected that exhibit strong signals at proxy locations where the proxy data indicate no anomalies and thus BMA are selected that are in fact incon-

CPD
11, C2493–C2502, 2015

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

sistent with the proxy data. I strongly suggest the authors revise their choice to exclude small proxy anomalies from the BMA selection process.

This choice is far from arbitrary; our primary motivation for this decision is that excluding proxy data from time periods where their climatic signal is ambiguous is a logical and conservative way to account for proxy uncertainty–this motivation is clearly articulated in Section 2.1 and should not require further justification. Referee 1 also raised this comment and we have reproduced the same response below.

Having worked extensively with a range of proxy data, we understand many of the challenges required to extract a climate signal from a proxy record. Proxies can be non-linear and inevitably contain a component of non-climatic noise; while they can be very good at recording anomalous conditions, we question the interpretation of mean signals in most proxy-based studies. When a strong signal is not present in a proxy record for a given time period, we can conclude with confidence that the proxy did not record a strong anomaly: it is another matter–requiring additional evidence–to conclude a strong anomaly did not occur (see Mann et al Nat Geosci (2012) for a related discussion). During testing and development we did experiment with including all proxy records. However, at each timestep proxies displaying non-anomalous (mean) conditions tended to dominate the analogue selection process at the expense of proxies displaying anomalous conditions. This is because there are many more analogues for mean conditions then there are for anomalous conditions. When proxies displaying ambiguous signals were included in the analogue selection the 50 BMA ended up being no different to a random selection, and PaleoR had no skill at all. It is therefore not appropriate–at present–for this method to include (at each timestep) proxies that show ambiguous signals.

The authors claim that the spread of the 50 BMA can be used as an estimate for the uncertainty of the reconstruction. This is an interesting concept, that is not pursued in the paper. One way to investigate whether the above assertion has some merit would be to compare maps of BMA spread and pseudoproxy correlation, as one would *expect larger spread in areas with smaller correlation. Additionally, one may analyze the correlation between time series of BMA spread and the ensemble mean error. If the spread is indicative of the uncertainty, then there should be correlation between spread and error.*

We have included a more detailed explanation of uncertainty to the manuscript in Section 4.2. We do find increased uncertainty in regions with weaker correlations in the pseudoproxy experiment.

Finally, the evaluation of the method should be improved to build confidence in the reconstruction dataset. The authors perform in-sample validation for the first part of the evaluation (Figure 2). I strongly suggest that the approach is changed such that the proxies that are evaluated are not actually included for the selection of BMA to get an understanding of how well the reconstruction works in places where we do not have proxy data. Given that the BMA approach is computationally cheap, this should be easily achievable.

Referee 3 proposes an interesting experiment that we have performed on an earlier version of PaleoR and included in Browning (2014). While the in-sample validation tests how well PaleoR performs at locations where there are constraining data, an out-of-sample validation tests how well PaleoR performs in regions where there is no constraining data. The important point here is that these experiments test two different properties, so we should not replace one with the other. The out-of-sample validation, in Browning (2014), showed unsurprisingly that PaleoR performs well in regions with sufficient proxy data redundancy, but performs poorly–in terms of reconstructing a proxy signal–in regions where there are no proxy data to constrain it. Another point to consider in such an experiment is we would be testing the ability of PaleoR to reconstruct an inherently noisy proxy signal not an actual climate signal. Even if PaleoR were reconstructing the climate signal perfectly we would only expect correlations of r=0.5-0.7 with each proxy record. When compounded uncertainties are taken into consideration such an experiment would be unlikely to return correlation scores of greater than r=0.2-0.4. The skill of PaleoR as a function of proxy data coverage is more comprehensively explored in the pseudoproxy experiment where it is assessed for every grid point (Section 3, see also Section 4.2 and 4.3 for discussions on proxy density); as such we would prefer not to add an additional test to the manuscript, as it would be somewhat redundant.

2 Specific comments Sec. 2.3: The authors chose to select the 50 BMA without providing convincing arguments for this choice. I suggest to add a Figure to illustrate the BMA ensemble reliability (e.g. the spread to error ratio) for various choices of ensemble size.

The requested figure was produced as part of a series of sensitivity experiments we performed to determine the optimum number of analogues; an in-depth discussion on this is provided in Browning (2014) and Goodwin et al (2013). The key findings of this study are summarised in Section 2.3 and the relevant citations have been included. We do not feel the issue is sufficiently important to justify reproducing this figure in the current manuscript.

Sec. 3.1: The discussion of LME vs. reconstruction correlation is not satisfactory as basically the correlation due to external forcing (LME mean) is compared with overstated correlation (see major comment above) of the best estimate of the internal variability (reconstruction). In particular, the statement "LME provides a realistic simulation of internal climate variability that is temporally inconsistent with most proxy evidence" is troubling. To assess whether LME's internal variability is temporally inconsistent with the proxy evidence one may rather select the best matching member (out of 10 LME ensemble members) and compare the correlation of such a reconstruction to the BMA reconstruction. Such an approach could also be used illustrate the benefits of pooling all available years for BMA selection.

The discussion in Section 3.1 illustrates that it is unreasonable to expect a model simulation to track stochastic internal variability in a temporally consistent manor–for that we

CPD
11, C2493–C2502, 2015

need to include a data constraint (proxy data). The correlations between PaleoR and the proxy data are not overstated, they are exactly what they are: a measure of how well each data constraint (proxy record) forces PaleoR to track its temporal evolution at its location (see comment above). As stated in the manuscript, when comparing LME to proxy data (Figure 2b) we get similar results whether we use the LME ensemble mean or the individual ensemble members. Mean correlation scores for each individual LME simulation range from r=0.15 to r=0.17 compared the LME ensemble mean of r=0.19 (as shown in Figure 2b). Considering the cascading uncertainties associated with comparing model simulations to noisy proxy data the correspondence between LME and the proxy archive is remarkably good. In acknowledgement of this we have modified the statement to: “LME provides a realistic simulation of internal climate variability that only broadly consistent with most proxy evidence”

Sec. 3.3: In the statement L26ff, the authors seem to suggest that all differences between PaleoR index reconstructions the comparison indices indicate periods of non-canonical behavior the comparison indices cannot reproduce. This is based on the unsupported assumption that PaleoR is superior to the comparison indices. Please rephrase.

Referee 1 also objected to this statement so it has been rephrased.

Sec. 4.1: the statement on L5 lacks support. From the manuscript at hand I cannot see that analogue selection should be superior to regression-based reconstruction methods, let alone proper data assimilation approaches.

We have added a reference to Steiger et al (2013) who shows that data assimilation approaches consistently out-perform regression-based approaches to paleoclimate reconstruction. The Bhend et al (2012) reference also provides a review of a range of experimental paleoclimate data assimilation approaches. The last comment is curious given that data assimilation approaches used for meteorological applications are inappropriate for paleoclimate data and the “proper” approach for assimilating paleoclimate

CPD
11, C2493–C2502, 2015

and model data is, at present, undetermined.

Sec. 4.3: One way to account for proxy uncertainty is to include an estimate of proxy error into the standardization of the proxy / model differences used for analogue selection.

Including an estimate of uncertainty is appropriate for methods such as Kalman filtering, where one might seek to adjust the model simulation to a specific (uncertain) value. However, it is not be appropriate in our approach as proxy uncertainty is implicitly accounted for in the analogue selection because we do not force an exact match between proxy values and model values, we seek only the nearest value that is also consistent with other proxy evidence and the modeled climate state. Therefore, including an estimate of uncertainty, as is applied in most studies, would only add additional uncertainty to the analogue selection. Instead, we take a conservative approach to proxy uncertainty by evaluating the proxy signal for each time period and excluding proxies with ambiguous signals–as described in Section 2.1.

Figures 2 and 3: Correlations with tropical SST proxies are surprisingly low or negative (Figure 2a), whereas the pseudo-proxy analysis suggests that the reconstruction should actually produce strong correlations (Figure 3c). This is somewhat puzzling, please discuss.

Negative correlations are expected for proxy data that have an inverse relationship to their respective climate variables–this is mentioned in the figure caption. However, in response to this comment we re-examined all SST proxy data and identified that the Goodkin et al (2008) Bermuda coral SST proxy (number 90) had been incorrectly inverted, as the Sr/Ca record had already been converted to SST equivalent values by the authors. PaleoR and all figures have been recalculated. This correction should improve PaleoR; as a result there are some minor changes in the reconstructed fields and indices.

Figure 4: The scales for the PaleoR reconstruction and the comparison indices differ.

CPD
11, C2493–C2502, 2015

This is misleading and very bad form and should be changed. Also, the reason why the variability in PaleoR indices is considerably smaller than in the comparison indices should be discussed.

The reason for a reduction in amplitude is fairly straightforward: averaging 50 BMA years introduces a component of noise that reduces the amplitude of the final signal. Most of the comparison indices have been centered and scaled relative to the instrumental period to give a value that is directly comparable to observations. PaleoR indices have not been centered or scaled relative to the instrumental period and are therefore not directly comparable to observations. PaleoR anomalies are calculated relative to the LME modeled climate and not the observational period. Although this is explicitly stated in the manuscript at the end of section 2.3 we have added some additional text to the Figure 4 caption to clarify these points. Plotting on a separate axis is not at all misleading as it clearly illustrates the magnitude of the variance reduction whilst allowing a clear comparison with the patterns of variation in other indices. It would be relatively straightforward to center and scale PaleoR derived indices to conform to the comparison indices, however that is not the objective of this study and we feel it would be more misleading than presenting the data as is.

Figure 5: The comparison of two different time slices seems somewhat arbitrary. Here I would prefer composites for individual indices if such an analysis is deemed necessary at all.

The decision to highlight specific time periods instead of composites was made for several reasons: (1) to provide an example of how the spatial fields conform to the index time periods, (2) illustrate that spatial fields can provide much more information than is available from just an index, and (3) demonstrate that multiple climate variables, such as SST and SLP are dynamically consistent. Hopefully the readers can also appreciate that the ability to examine multivariate spatial patterns associated with specific time periods can be extremely useful in trying to resolve conflicting proxy signals, and/or time periods when PaleoR indices might differ from previously published reconstructions.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

Composites on the other hand are useful for exploring specific research questions. PaleoR composites have been used effectively in Browning (2014) and form the basis of some of our (as yet) unpublished research. We contemplated using Figure 5 to display composites, as suggested by Referee 3, however we decided that highlighting specific time periods was more appropriate, and might encourage interested parties to go online and investigate other time periods of interest. Our preference therefore is to retain the current figures.

3 Technical corrections P4163L15: low signal-to-noise P4173L6: PaleoR NAO and the P4174L2: expected to yield

These have been addressed.

Interactive comment on Clim. Past Discuss., 11, 4159, 2015.

CPD
11, C2493–C2502, 2015
Interactive Comment