Climate
of the Past
Discussions

Open Access

# *Interactive comment on* "The influence of non-stationary ENSO teleconnections on reconstructions of paleoclimate using a pseudoproxy framework" *by* R. Batehup et al.

**O. Bothe (Referee)**

ol.bothe@gmail.com

Batehup, McGregor and Gallant provide an interesting and thorough perspective on how potential non-stationary relations between ENSO and surface temperature may impact the skill of reconstructions of ENSO-variance. Thereby the manuscript "The influence of non-stationary ENSO teleconnections on reconstructions of paleoclimate using a pseudoproxy framework" is an relevant climate-simulation-based contribution to the field of Common-Era paleoclimatology.

Nevertheless, I would like to see a number of comments addressed. I also have additional suggestions. My comments are "minor" except for one.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

A more general comment to begin with: Your methods are common and simple enough. However, variance estimation from small samples and correlations from short series are rather sensitive estimates. This doesn't matter too much in your perfect proxy case, but as soon as one would consider noise-degraded pseudoproxies, the chain of statistical estimates may (or may not) result in an accumulation of uncertainties, which would quickly degrade the usefulness of any obtained reconstruction. Therefore, it would be nice to see one test-case, where the proxies are degraded by noise. However, this is certainly not a prerequisite for publication, it just might strengthen the impact of the article.

**Specific comments**

**Major comment**

There are two possibilities, either I misunderstand the description of what has been done in the PNEOF1 "experiment", or the "experiment" doesn't represent what it is meant to show.

Maybe what you do is the following: you do the running correlations, you do the EOF, and you then select proxies from the regions deemed non-stationary and also having strong associations with EOF1. If this is the case, the method does what it is meant to do.

However, I understand the description of what you do as the following: you do the running correlations, you do the EOF, and you select proxies from the regions with strong associations with EOF1. Then, you do not sample from the non-stationary regions but possibly from the weakly correlated regions, which would explain the near-total lack of skill for some methods.

Let me rephrase: do you use the EOF to sub-sample the nonstationary regions for covarying regions; or, do you use the EOF to just organise all regions for covariability?

If my understanding is correct, do you assume, and if so why, that the EOF1 represents nonstationarity?

Anyway, what does "EOF weighting < 0.1" [Line 18 on page 3871] mean? Do such values occur? (According to the color bar they don't.).

Additionally, I wonder whether the trend in the PC1 suggests some problems in the control run.

**Minor**

1. Could you please as soon as possible, i.e. before your final response to the reviewers, provide a version of Figure 10 including the MRV?

2. I think the title should mention that you reconstruct of ENSO-variance. The introduction also should clearly state it. Similarly, on some/most instances where you write "reconstructions of ENSO" it would be more appropriate to write "reconstructions of ENSO variance". Alternatively you may state early on that "reconstruction of ENSO" implies "reconstructions of ENSO variance".

Similarly at Page 3868 Line 20ff: The description of Figure 4 is, as far as I understand it, incorrect. Please be clear that it is the correlation between the running variance series and not between the Nino3.4-indices.

3. Methods:
a. Is the calibration window length implicitly meant to also be the length of the running variance windows? I ask, because you never explicitly mention the window-length for the running variances. It can't be the calibration window length, because you also use the full 499 years for calibration. So, what is the window length for the running variances?
b. Please clarify what calibration means in your setup.
c. You mainly consider skill in terms of correlation. Is correlation really the best skill-

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

measure in this case? Why?

d. On page 3860 line 10 you describe the timeseries as "June-July"-averages. I assume you mean thirteen month averages. Please clarify.

e. You mention model-drift: Is it a problem in your simulation?

f. I wonder whether it would be better to convert all correlations to Fisher-Z-scores. I do not propose to do it, I only wonder whether it might clarify the Figures.

4. There are references which should not be omitted. A paper on reconstructing ENSO-variance has to mention the work by Russon and colleagues (Russon et al., 2014, 2015). Furthermore, the general introduction of ENSO-reconstructions appears to ignore the works of Watanabe et al. (2012), Cobb et al. (2013), and Li et al. (2013). As a less important side-note I want to mention that there have been other studies dealing with non-stationarity of climate-modes in recent years.

Kim M. Cobb, Niko Westphal, Hussein R. Sayani, Jordan T. Watson, Emanuele Di Lorenzo, H. Cheng, R. L. Edwards, and Christopher D. Charles, 2013, Highly Variable El Niño–Southern Oscillation Throughout the Holocene, Science, 339 (6115), 67-70, DOI:10.1126/science.1228246

Jinbao Li, Shang-Ping Xie, Edward R. Cook, Mariano S. Morales, Duncan A. Christie, Nathaniel C. Johnson, Fahu Chen, Rosanne D'Arrigo, Anthony M. Fowler, Xiaohua Gou Keyan Fang, 2013, El Niño modulations over the past seven centuries ,Nature Climate Change, 3, 822–826, doi:10.1038/nclimate1936

Russon, T., Tudhope, A. W., Collins, M. and Hegerl, G. C. (2015), Inferring changes in ENSO amplitude from the variance of proxy records. Geophys. Res. Lett., 42: 1197–1204. doi: 10.1002/2014GL062331.

T. Russon, A. W. Tudhope, G. C. Hegerl, A. Schurer, and M. Collins, 2014: Assessing the Significance of Changes in ENSO Amplitude Using Variance Metrics. J. Climate, 27, 4911–4922. doi: http://dx.doi.org/10.1175/JCLI-D-13-00077.1

Watanabe, M., J.-S. Kug, F.-F. Jin, M. Collins, M. Ohba, and A. T. Wittenberg (2012), Uncertainty in the ENSO amplitude change from the past to the future, Geophys. Res. Lett., 39, L20703, doi:10.1029/2012GL053305.

5. On the better performance of MRV.

a. I think this would be an interesting test for what happens when noise is introduced in the proxies.

b. You stress the overall better performance of MRV. One may argue that this is unsurprising as it uses the variance from the beginning, in this sense the comparison may be called biased. However, on the other hand, I am not sure it is true. In a real world scenario: what is the more important skill-metric, the RMSE, where MRV performs consistently worse, or the correlation? Is there possibly another better suited skill-metric? You mention the potential need for re-scaling but don't do it. Why not, I would think implementing an MRVPS (MRV plus scaling) should be easy enough. There may be reasons, but as you so far don't show the MRV series (e.g., in Figure 10), the reader is unable to assess this.

c. Similarly, page 3875 line 13ff: Isn't the damping an expression of especially large variance loss for MRV, at least if the term is used as commonly employed for reconstructions?

d. Page 3877 line 21: You write MRV excelled, but MRV also showed large RMSE. I think that should be mentioned. Can you estimate how noise/uncertainties in the proxies would affect this feature.

e. Page 3879 line 16/17: MRV is the most robust in your perfect-proxy setting.

6. With respect to the discussion of CPS on Page 3873 Line 15ff. Why do you single out CPS here? It is not really worse than RMV or EOF.

7. On the discussion of the tropical supposedly non-stationary grid-points on page 3865: Don't these mainly represent the wide range of internally varying differences in the evolution of ENSO-events which are potentially not captured by a simple stochastic process? I am not so much thinking of CP vs. EP ENSO, but different evolutions of

one of these flavors. I think this could be discussed more extensively.

You make a similar point on page 3870, line 15ff. So, it's not only about different flavors but about the large variability in how events evolve. The statistical process captures not necessarily all dynamical variability.

8. On Figure 10: The regressions and the correlations show rather weak relations between the two series. Anyway, why should the standard deviation of the running correlations of the proxies with the target and the variance reconstruction be related in any way. Put differently: what do we learn from the correlation of the variance of a measure with the square-root of the variance of running correlations of a measure with another measure?

The number of effective degrees of freedom of the time series appears to be small, are the correlations even significant?

9. Page 3871 Line 9: Which scenario do you mean? That non-stationarities potentially influence the skill? Well, that is "likely", but its effect appears to be not so important according to your results?

10. In your plots you give explained variances for running variances. How much variance is captured by the stationary sample-variance-distribution?

11. The discussion in the beginning of the last paragraph on page 3872 is not really relevant, is it? The case of a 500 year calibration window is not realistic, so the discussion should focus on the skill differences between 91yr and 31yr.

Related: The second part of the paragraph (on page 3873) appears to be partially redundant.

12. Page 3876 line 2ff: Please do not just point to panels but give some more information on what we see.

13. It may help the reader if you extend a bit in line 22ff of page 3876 on what Wittenberg (2009) showed.

14. More a technical comment, but I put it here as well to emphasize it: I think Figure 8 is rather unclear, since I am not really able to distinguish the different hatchings.

**Technical comments, etc.**

General: I am surprised by the comma placement in the manuscript. However, as a non-native English-speaker, I do not annotate it.

My subjective impression is that the abstract could be shortened and could formulate the relevant points more concisely.

Page 3854 Lines 17-20: Please rephrase the sentence. (I do not really see how the "to which"-part relates to the previous sentence-structure.)

Page 3854 Line 24: What do you mean by "uniformly-spaced"?

Page 3859 Lines 21-24: I think this sentence could be clarified

Page 3860 Lines 21-22: "is used" . . . "are selected". Please clarify is/are.

Page 3862 Line 3: "Fig. 1b" -> "(Fig. 1b)"

Page 3864 Line 22: you write of the "possible" range of running correlations. Is "possible" the correct word here. Does a statistic tell us about what is possible? Change to "expected"

Page 3865 Line 22: Do you examine the "likely" - as you write - or the "potential" effects of non-stationarities?

Page 3866 Line 22: "(2005); Hegerl" -> "(2005) and Hegerl"

Page 3867 Line 16: I don't think you write in Sect. 2 how you calculate the running variance for the ENSO index (see above, what is the window-length of the running variance).

Page 3867 Line 19: "dataset is available with larger proxy networks". My impression

Interactive
Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

is, that there may be an "and" missing between available and with.

Page 3869 Line 5: "The skill metrics", I think you mean "the proportion of skill metrics"?

Page 3869 Line 9: I think you mean "Sect. 4.3 and 4.4"

Page 3870 Line 1: "In all". Should that be "For all"?

Page 3872 Line 15ff: I think the second part of this sentence is incomplete.

Page 3872 Line 18: I don't mind "Fairly good chance" but I can imagine colleagues who are rather annoyed by such a phrase.

Page 3872 Line 26: Is "would be" correct? Shouldn't it read "is" which may be qualified by "likely" or "potentially".

Page 3873 Line 7: I think the "However" is wrong.

Page 3873 Line 15: "It is noted"? Why passive construction? -> You note.

Page 3874 Line 3: Not the red line outperforms the other lines, but using 91 year calibration windows performs better than shorter windows.

Page 3874 Line 16ff: Is the sentence correct? Do you generally plot the "variance taken . . . of the correlations"?

Page 3875 Line 18: Is "on this paper" correct?

Page 3877 line 15: "highlight a case for considering" or just "highlight"?

Page 3878 line 21ff: I think the second part of this sentence is incomplete.

Page 3879 line 17: "many various"?

Page 3879/3880: My impression is that the second part of the conclusions is more or less redundant and repeats what the first part already said.

Page 3880 line 5ff: I think this perspective is not really necessary, but that's just per-

sonal taste.

Figure 1: Please rephrase "is the correlation between of the entire 499 years of TS at each grid point and the model calculated Nino 3.4 index correlation coefficients"

Figure 2: Please provide labels for the color bars of panels b,d,f.
verses -> versus

Figure 3: You write "the pseudo-reconstructions running variance". Wouldn't it be more appropriate to write the "pseudo-reconstructions of running variance"?

Figure 4 (and other Figure captions): You write "reconstruction's running variance" which is appropriate for CPS and EPC and to some extent for RMV, but shouldn't it be "reconstructed running variance".
"explaining greater than 50% of explained variance" -> "explaining greater than 50% of variance"?

Figure 6: Please rephrase "and this determines what values of proportion can be taken as larger groups have a wider range of possible non-stationarity proportions than smaller groups".

Figure 8: The hatching is, from my point of view, nearly unidentifiable.

Supplement: Please provide a clear copyright statement.

Figure S1: Could you reshape the aspect ratio for final publication?

Figure S4: Caption, last line: "with with"

————————————————————

Interactive comment on Clim. Past Discuss., 11, 3853, 2015.