**Climate
of the Past
Discussions**

Interactive comment on "Climate of the last millennium: ensemble consistency of simulations and reconstructions" *by* O. Bothe et al.

T.L. Edwards (Referee)

tamsin.edwards@bristol.ac.uk

Received and published: 8 October 2012

This manuscript is quite poorly-written and hard to understand. There are several aspects in which the authors do not convince me they have understood, or at least sufficiently addressed, the scientific and statistical pitfalls of applying rank histograms (and related analyses) to palaeoclimate data.

These comments apply to the manuscript up to and including Section 3.1.1.

General scientific issues

a) Incorrect scientific questions / inferences

- The abstract states that the simulated and reconstructed ensembles are both over-

dispersed. I think this is a consequence of using an ensemble mean as a target (lower interannual variability). I don't think ensemble mean targets, and therefore these conclusions, are valid.

- Incorrect inferences from figures - see specific comments for p2418 below

b) No discussion / addressing of correlation of reconstruction ensemble members

- Correlation across reconstructions is mentioned by Frank et al., and described as a problem by Marzban et al., but not mentioned or addressed

c) No discussion / addressing of correlation of simulation ensemble members

- As above. I'm not convinced that a forcing ensemble is a valid subject for this kind of analysis. If it is, it requires thoughtful discussion, and addressing of, the high degree of correlation between ensemble members due to the design of the ensemble (limited number of possible forcing inputs).

d) Insufficient discussion / sensitivity studies of effects of temporal correlation

- How were the bootstrap lengths chosen, and what is the sensitivity to this choice?

- How is the temporal correlation estimated (degrees of freedom / correlation coefficient) for the chiˆ2 test?

- What is the sensitivity of the result to this choice?

Other manuscript issues

2411/18 - There is no "status of truth" even when there are observations: our knowledge is always imperfect.

2411/20, 2413/8 - This is not an "objective" measure because arbitrary choices are (necessarily) made throughout.

2414/18 - Explain bootstrapping method and motivation more clearly, for those not familiar with it

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

2414/24 - Explain climatological and ensemble components of reliability

2415/11 - Add table of simulation experiments to clarify forcing combinations

2415/11 - Add figure showing forcing timeseries

2425/11 - Explain how weak and strong solar forcings are chosen - bounds of an uncertainty range?

2415/17 - Are all reconstructions annual? (Sentence implies not all are)

2415/18 - What are the different members of the reconstruction ensemble? How many are there?

2415/22 - What are the sub-ensemble members, how many are there, and how were they chosen?

2416/12 - "similar caveats" - which? What is their implication here?

2416/23 - Needs better explanation of how adding error to the target can compensate for temporal smoothing

2417/15 - Comment on / explain Fig. 1 here, especially righthand column

2417/21 - Remove all plots without observational errors added - this is not a study of the method but an application

2418/14 - Explain this result explicitly: the simulated ensemble mean is generally near the centre of the ensemble of reconstructions, while the reconstruction ensemble mean is generally near the centre of the ensemble of simulations. Comment on why this is - presumably because the target is an ensemble mean and therefore has lower interannual variability. Is this a valid target? I don't think so (see Scientific Issues above). How about repeating this using one (or each) of the individual ensemble members as a target?

2418/17 - Give the critical values for the chi^2 test

2418/17 - "highlights the problem of..."strictness" - I would say that it actually indicates a problem: for example, using the wrong number of degrees of freedom.

2418/20 - "as shown by..." - I don't understand how, when the bootstrapped envelope encompasses the target zero line.

2418/21 - "as shown by..." - explain how - is this because the grey band is slightly lower than the target? I can't see how this is worse in the 16th century.

2418/24 - The slope is within the bootstrapped intervals, therefore should not be considered a slope

2418/28 - I think this should be 3c,e not 3b,c

2419/11 - Does this chiˆ2 test include autocorrelation? If not, why use the bootstrapped intervals if these are then ignored in favour of a chiˆ2 test under an assumption of independence?

2419/14-17 - "not shown" vs "compare Fig. 3" This section is not clear as to what is shown and what is not. Rank histogram not shown? Or weak forcing ensemble not shown? Shown in Fig. 3c-f or only c+d?

2419/19 - To improve clarity move "not shown" to after "truth"

2419/19 - "into perspective" - what does this mean: confirms?

2419/23 - bootstrap "generally" and "otherwise" - be more specific: a + e vs a only?

2419/26 - "amplified picture" - what does this mean? It seems to contradict line 19.

2420/5 - "strong deviations" - presumably this is in R-Q-Q plot: be more specific, otherwise it sounds as though it is a deviation in the signal (e.g. bias) rather than quantiles (dispersion). I don't understand the rest of this sentence.

2420/8 - 50yr moving average has not been mentioned or explained before this point. I don't understand the rest of this sentence.

Figure 1 - Too small and very unclear. Split into 3 (or more). Add individual legends to each to clarify confusing colour schemes. Remove ensemble means for clarity. Why is the moving average only shown for CE? This should be full size. I think there are both solid and dashed black lines but this is not explained. Why is the solid one so thick?

Figure 2 - Remove all values of chi^2 where autocorrelation is not accounted for - this is not a study of method

Figure 3 - Some of these colours look the same to me

Throughout:

- I find the parenthetic structure for two-part sentences hard to read; suggest rewriting in separate sentences

- replace "truth" with "target"

- update Marzban et al. 2010 ref to 2011

Interactive comment on Clim. Past Discuss., 8, 2409, 2012.