Clim. Past Discuss., 10, 2627-2683, 2014
www.clim-past-discuss.net/10/2627/2014/
doi:10.5194/cpd-10-2627-2014

Answers to Referee #4

Specific comments:

1.)
The extended framework allows for autocorrelation in the simulated series. However the proxy data have autocorrelation too. Why is this not included in the framework? If it is not needed for reasons that have to do with the construction of the test statistics this should be explained.

The true climate and the proxy data are actually allowed to be auto-correlated, but we need not correct for it. This is because we condition on the proxy series; that is we regard it as given (but arbitrary). This is already mentioned on page 2634 (and repeatedly in SUN12), but it will be further stressed, and also added early in Appendix A (cf. referee comment 11 below).

2.)
p2692, end of abstract: an additional reason for the inconclusive results could be the small ensemble size for the simulations; in an average over a larger ensemble the signal to noise ratio would get higher. A short comment on this should be added.

Yes, the forced part of simulated temperature variability should stand out clearer in a larger ensemble. It may nevertheless still be possible that the actual set of proxy data would not be sufficent to judge which of the two solar forcings that produce the best fit to the observations. We can add a comment about this at some relevant place in our revised text.

3.)
p2631, line 4: clarify whether these are reconstructions for local or larger-scale temperatures (I presume it is the former).

One of the fifteen reconstructions (ASIA2K) is for large-scale temperatures, but the rest are for more local temperatures. We can make this more clear, both in the introduction and in Section 3.3.

4.)
p2632, last paragraph: The calibration needed for the unbiased ranking is in other contexts, e.g. in statistical downscaling, known as 'inflated regression'. If the goal is to estimate z from τ this is fundamentally wrong, as pointed out for instance by von Storch (J. Clim 1999). In order to avoid confusion it should be pointed out that there is no conflict between the different statements on inflated regression, as the context is different.

This issue is discussed in some detail in a Remark within Section 4 of our Part 1. Actually, the paper by von Storch (J. Clim. 1999) is also cited there. Thus there is no need to repeat this discussion here, but we we can try to point out more clearly that Section 4 in Part 1 provides a discussion about this and that there is no conflict with the discussion made by von Storch.

5.)
p2639, second paragraph: It is not clear whether the instrumental error mentioned here is the error in the actual measurements, which is what the name implies, or the error in the gridded data, which I presume it is, because these are used as 'instrumental records' in the analysis. Please add a

comment to avoid confusion.

As discussed by Brohan et al. (2006) and Morice et al. (2012) (we cite both of them), errors in grid-box (or averages over several grid-boxes) temperatures are due to random errors in the actual station measurements, to various more systematic errors in the observation procedures, and to spatial sampling errors in the gridding process. The total error in the temperature data we use will thus contain some kind of combination of all different types of errors. We can add a comment to clarify this.

6.)
P2640, line 26: 'separate the climate signal from the raw data' is not precisely phrased. Standardization tries to remove non-climate-related low-frequency variability, but there is still a lot of non-climate-related high-frequency variability left in the standardized tree ring records, so it is not correct to say that the climate signal is separated (or extracted,, which I think is what the authors meant (interpreting 'separated' in the usual way would mean that the end result is the climate signal and the raw data)).

We propose to change the text on lines 24-29 to the following:

Melvin and Briffa (2008) introduced a method that allows the simultaneous estimation of the tree-ring standardization curves and the common environmental signal that is embedded in the same tree-ring records within a region. This so-called "signal-free" (SF) iterative standardization method removes the influence of the common environmental (assumed climate) signal on the standardization curve, which reduces the trend distortion that can occur near the ends of a traditionally standardized chronology. The method can be applied on both IND and RCS standardization (Melvin and Briffa 2014), but very few records have been created with this rather new technique. Three records in our collection were developed using SF in combination with RCS.

The following new reference is added:

Melvin T. M. and Briffa K.R. (2014) CRUST: Software for the implementation of Regional Chronology Standardisation: Part 1. Signal-Free RCS. Dendrochronologia 32, 7-20, doi: 10.1016/j.dendro.2013.06.002

7.)
p2642, last paragraph, p2643 first paragraph: I'm not convinced by the arguments given for selecting the size of the region for calibration. Individual tree-ring records respond to the local climate, and therefore, as pointed out by the authors, they are less closely linked to large-scale climate (the correlation gets lower). I would expect that this effect compensates the fact that large-scale temperatures have a higher ratio of forced to unforced variability in an analysis that tries to decide which of two different climate model forcings leads to better agreement with the proxies.

Although correlation maps between tree ring records and temperatures similar to those given in Fig. 1. (see next comment) are a useful guidance for the choice of area I encourage the authors to give a conceptually and if possible statistically more sound discussion of the question of which area size to choose.

We agree that our motivation here is not very precise. We will follow the advice and at least try to give a conceptually more sound motivation. We are not able, however, to provide any clearer discussion from a statistical point of view. Rather, we identify this question as one where more

research is needed, and we can point out that we welcome future efforts made to better judge how to select regions optimally for this kind of study. We assume that there is a similar need in both data assimilation and detection and attribution studies.

8.)
page 2643, first paragraph: basing the correlation maps on first differences seems fundamentally wrong and is not consistent with the argument given. It is true that correlation are influenced by strong trends and it is advisable to remove this effect. This can be done in a straightforward way by de-trending the data. Although using first differences also removes the effect of trends the resulting time series are in principle the derivatives of the original series, and the correlations measure the link between the derivatives. It remains to be seen whether the standard correlation maps look similar, but even if so there is no justification for using first differences.

These correlations are influenced by both strong and weak trends and by both linear and nonlinear trends. The form for detrending mentioned by the referee is suitable for linear trends, but here we wanted to be protected also against nonlinear trends, which could have been caused by a more general nonstationarity. An often used class of models with nonstationarity is the ARIMA class of models, and for such models linear trend and nonstationarity are simultaneously eliminated by first forming a new series of first (or higher order) differences. The resulting cross-correlations will be different, of course, but their relative magnitudes were not expected to be very different, and this was considered enough. Furthermore, empirical comparisons (using both original data, linearly detrended data and first differenced data) confirm that the form and size of the regions in this case does not depend crucially on the choice of detrending method. If the editors express a wish for it, we can easily include a set of figures showing this. However, we do not want to abstain from using first differences, since we would then be open for criticism by other scientists who do not trust trend linearity. We can additionally remark that first differences have since long been used in tree-ring research to study (by means of a sign test) whether or not sufficient similarity exists between actual and estimated data (Fritts, H.C., 1976, Tree Rings and Climate, Academic Press).

9.) page 2468, lines 13-15: This comment is a bit surprising as GHG and orbital forcing are both small for the period 1000 AD to 1850 AD.

The referee presumably means page 2648. Yes, indeed, the GHG and orbital forcings are small in the period 1000-1850, but they are not zero. Moreover, the orbital forcing has different trends in summer and winter and in the northern vs the southern hemisphere – thus causing the global annual-mean forcing to be very small. However, in the NH summer there is a decreasing forcing trend over the period due to the change in the Earth's orbit. As most of our proxy series reflect NH summer summer temperatures, it cannot be excluded that orbital forcing can have had an effect on several proxy data series used here. Phipps et al. (J. Climate, 15 September 2013, p. 6915.) use simulations with only GHG forcing, GHG+orbital, GHG+orbital+solar, GHG+orbital+solar+volcanic forcing and compare with proxy data. They discuss the possibility that orbital forcing has caused a cooling trend that is seen in many NH tree-ring series (though Esper et al., Nature Climate Change, 8 July 2012, discussed that tree-ring data may not capture this effect to a full extent). Clearly, there are ongoing discussions about these aspects. We did not mean this to be a crucial comment in our text, we merely wanted to say that the available simulation ensemble does not allow answering any question regarding how much of the ensemble-mean result is affected by the role of GHG and orbital forcing. We will attempt to clarify this when we revise the text.

10.)
Page 2652, line 29: There should be no question mark as this sentence is a statement not a question (the fact that even native speakers sometimes use a question mark in this sense doesn't make it better).

Agreed - We will replace the question mark with a period.

11.)
Comment 1.) applies here again, i.e. why is there no autocorrelation in η ?

Presuming that this comment refers to page 2655 (near lines 5-6) in Appendix A, the answer is that we allow such autocorrelation. We will add a reminder; cf. our reply to comment 1.

12.)
The comments on the forcing are confusing. If I understand correctly the effect of the forcing is contained in δ and η, so these terms are not uncorrelated with the forcing as in SUN12.

Presuming that this comment refers to the description of the statistical model in Section A1, we agree that it could be formulated more clearly, so we will make an effort to achieve this.

13.)
page 2655, part on correlation statistic: I did not fully understand this part and it seems several important details in linking the different equations are not properly explained. Please check this part carefully and explain intermediate steps in sufficient detail so the line of argument can be followed by the typical reader of Climate of the Past.

We can and will add some more detail, to make it easier for the reader.

14.)
Page 2657, part on D^2 test statistic: Again the explanation of this part is rather short.

We can and will add some more detail, to make it easier for the reader.

15.)
page 2658: why are there two types of outer brackets (curly ones and big standard ones)?

For covariances, Cov{…, …}, which have two components with a comma in between, we have used curly outer brackets, but not for variances, Var(…). However, we have not been quite consistent. Two formulas on Page 2656 do perhaps not motivate their curly brackets. Also, the curly brackets Page 2658, line 19, should be bigger, of the same size as in line 24. We will be more consistent in the revised text.

Stockholm, November 2, 2014

Anders Moberg, Rolf Sundberg, Håkan Grudd, Alistair Hind