Clim. Past Discuss., 7, C2198–C2203, 2012
www.clim-past-discuss.net/7/C2198/2012/

Climate of the Past Discussions

Interactive comment on "The extra-tropical NH temperature in the last two millennia: reconstructions of low-frequency variability" *by* B. Christiansen and F. C. Ljungqvist

B. Christiansen and F. C. Ljungqvist

boc@dmi.dk

Received and published: 11 January 2012

We want to thank the reviewer for this very thoughtful and thorough review. It is encouraging to see that both the two anonymous reviews and the comment by Bothe agree that our results are interesting and that the paper is well written. It is also interesting that they agree that the paper lacks details regarding the robustness of the reconstructions and that it also lacks a better comparison to previous work.

Below follow our replies to the reviewer's major comments:

1) The present paper should, as mentioned by the reviewer, be seen as a continuation of Christiansen and Ljungqvist, J. Clim., 2011 (CL2011). When writing the present

paper we wanted to focus on the results while toning down the more methodological issues which were dealt with in great details in CL2011 resulting in a lengthly paper. In the present paper we limited ourselves to the robustness to changes in the calibration interval and the general sentence "We find similar results in the present study" (as in CL2011, section 5, p 4005) while referring to CL2011 for details. In view of the comments we can now see that this was a wrong decision. In the revised version we will include a section dealing with the robustness. This section will include the results mentioned in the reply to reviewer #1 dealing with the spatial weighting of the local reconstructions and in the reply to Bothe dealing with the latitudinal distribution as well as the influence of the East Asian proxies. We will also include a discussion of the influence of annually/decadally resolved proxies as suggested by the reviewer (see also point 4 below). Most of these issues were dealt with in CL2011 and we find very similar results for the proxy compilations in the present paper, although this is of course not trivial.

We will describe the differences between the present paper and CL2011 more clearly and point out what is new in the introduction.

2) As mentioned in our reply to Bothe most kinds of cross-validation are not possible for the low-frequency diagnostics because of the limited length of the calibration period. This was also discussed in both CL2011 and Christiansen et al. 2009. In CL2011 we performed a leave-one-out cross-validation test of the correlation between the reconstruction and the NH temperature. Such a test will be included in the revised manuscript as described in the reply to Bothe. In lack of suitable cross-validation tests we have relied heavily on the ensemble pseudo-proxy method to evaluate the reconstruction method and to provide confidence levels. This method is not without assumptions though, these being that both the climate model experiment the pseudo-proxies are "realistic". Other methods that could supplement the ensemble pseudo-proxy methods in testing the low-frequency variability are Bayesian modelling and bootstrapping. However, these methods are not straightforward to implement and will not reveal the

systematic bias that may be present in the reconstructions due to a wrong reconstruction model. See the discussion in the Comment by Tingley and Li and the Reply by Christiansen (both submitted J. Clim.). We are currently working on a comparison between the different methods (Bayesian, Bootstrap, ensemble pseudo-proxies) to obtain the confidence levels. Note also that many reconstruction methods perform very well in the calibration period but still underestimate the low-frequency variability in the reconstruction period (Christiansen et al. 2009).

It is correct that the LOC method for the real-world reconstruction seems to overestimate the low-frequency variability in the calibration period: The observed extra-tropical mean falls somewhat outside the confidence intervals in the first and last years of the calibration period. It is also true that this is not seen systematically in the ensemble pseudo-proxy experiments. However, this does not mean that it can not be found by chance for a single realization. If not a chance occurrence, another possible explanation is that the overestimation is a consequence of proxies not cleanly representing local temperatures but also the NH mean itself. Such an effect could theoretically appear through teleconnections if proxies were sensitive to precipitation. This effect could also be a consequence of favoring proxies that show a strong 20th century warming in some step in the proxy selection process. While these explanations all lead to an overestimation of the low-frequency variability in the calibration period their influence on the reconstruction period depends on the detailed explanation and can be both neutral, an underestimation or and overestimation. We will discuss this in some details in the revised paper and include the observed NH mean in the figures.

2a) The discussion about the cross-over between high and low-frequency variability in LOC, that the reviewer request, is to a large extent given in our Reply to a Comment on CL2011 by A. Moberg which raised almost the same question as the reviewer. Both Comment and Reply are submitted to J. Clim. and a preprint of the Reply can be obtained from http://web.dmi.dk/fsweb/solar-terrestrial/staff/boc/reply_to_moberg.pdf. We received the Comment at the time when we were finishing the present paper and

we tried to minimize the overlap between the two manuscripts but some of the material in section 5 is also included in the Reply (which is cited there). The Reply shows from both theoretical arguments and numerical experiments how averaging (temporal and/or spatial) leads to a reduction in the noise on the reconstructions. The cut-off between exaggerated high-frequency variability and well reconstructed low-frequency variability depends on the degree of auto-correlation in the noise. With observed values of the auto-correlations in the noise (as AR1 noise or the "realistic" we use in the ensemble pseudo-proxy experiments) the 50-year smoothed values are well reconstructed with only a small amount of additive noise. We now cite the Reply more extensively in section 5 giving a brief summary of the results.

2b) The LOC method overestimates high-frequency variability as noted by the reviewer and discussed above. We agree that the annually resolved reconstructions are not that interesting and could be removed from the plots. However, it is interesting to see how the high-frequency variability depends on the number of proxies. This is discussed on page 4002, line 22. As the discussion of the high-frequency variability will expand in the revised version (see 2a) we are not sure that it will be wise to remove the annually resolved reconstructions from all the plots. A solution could be to remove the annually resolved reconstructions from the existing figures and add a new figure showing how the high-frequency variability depends on the number of proxies.

3) We will include a figure comparing our new LOC reconstructions with the other reconstructions mentioned by the reviewer (including Hegerl et al. 2007). Such a figure was also included in CL2011. This new figure will also include the LOC reconstruction from CL2011 as suggested by Bothe.

4) The local temperatures were degraded by smoothing to the same resolution as the proxies as described on the bottom of page 3998. The influence of this procedure was tested in CL2011 (section 8, page 6029) and found to be very small. This will be described more clearly in the revised version.

As mentioned above we will include a new section dealing with the robustness of the reconstructions. This will include the suggestions from both reviewers and Bothe: influence of method of spatial averaging, the latitudinal influence and the influence of annually vs decadally resolved proxies. The figure below shows the two-millennia long reconstruction based on only annually resolved proxies together with the reconstruction from Fig. 5. We note the relatively high resemblance in the low-frequency variability although the new reconstruction is based on only 7 proxies. Also note that the high-frequency variability has increased as expected.

We will expand the description of the ensemble pseudo-proxy method and discuss its limitations. As mentioned by the reviewer such efforts to estimate the the confidence levels are probably always optimistic as not all sources of stochasticity are included. In CL2011 (section 7) the ensemble pseudo-proxy method was described in details and it was also discussed (section 8) how it can be expanded to include additional sources of uncertainty, e.g., missing values in the temperature dataset.

Minor comments: We will consider all the comments in the revised version.

Interactive comment on Clim. Past Discuss., 7, 3991, 2011.

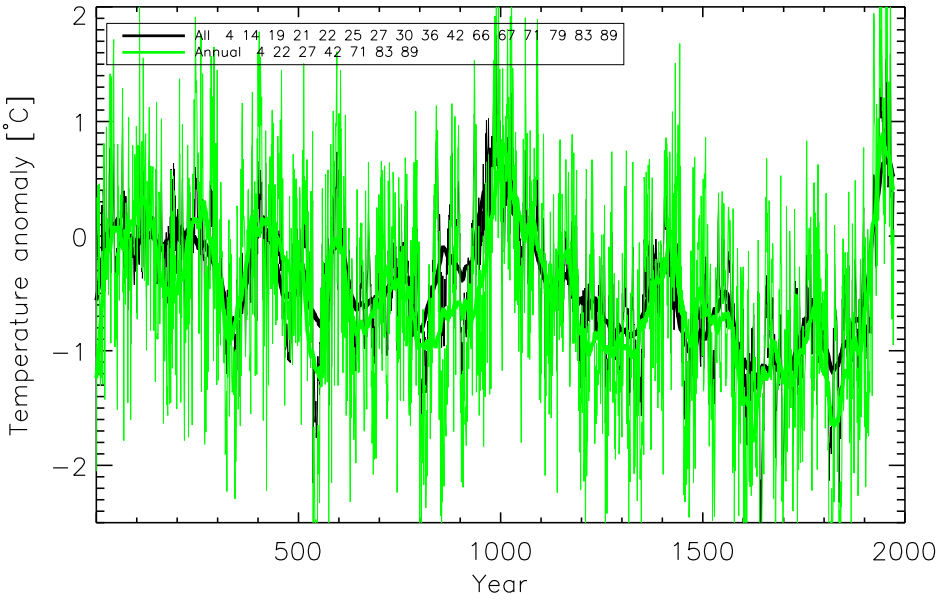

Fig. 1. The reconstruction from Fig. 5 in the submitted paper (black curves) together with a reconstruction based on only the annually resolved proxies (green curves). All other conditions are similar.