Clim. Past Discuss., 6, C110–C115, 2010
www.clim-past-discuss.net/6/C110/2010/

Climate of the Past Discussions

Interactive comment on "Detecting instabilities in tree-ring proxy calibration" *by* H. Visser et al.

L. Kutzbach (Referee)

lars.kutzbach@zmaw.de

Received and published: 23 April 2010

GENERAL COMMENTS

The problem of unstable relationships between proxies and target climate variables, e.g., the "divergence problem" is currently widely and controversially discussed in the scientific literature, and since the "CRU affair" also in the wider public arena. I completely agree with the authors that dendroclimatologists should deal more explicitly with different types of uncertainty during proxy screening (their conclusion 5). Especially the occurrence of instabilities in the proxy-climate relationships leads to serious questions about the validity of paleoclimate reconstructions. Therefore, it is highly welcome that the authors focus on this problem in this article and offer new approaches to analyse this problem. The authors introduce a new approach based on stochastic response functions and the discrete Kalman filter to detect instabilities in the relationship be-

tween tree-ring proxies and target climate variables which appears very suitable for the analysis of the instability problem. I agree with the authors that the method would be very useful for screening proxies before using them for climate reconstruction. On the other hand, I disagree with the authors that the calibration time series just can be truncated if instabilities are found in the complete time series by the presented method. In my view, this will not improve the confidence in past climate reconstructions based on such proxy data. Please see the specific comments given below. The manuscript is generally well and clearly written, and only few orthographic errors (mainly concerning the placing of commas) have to be corrected (see list of technical comments below). It is also well structured; I only miss some description and/or discussion of Figure 2, which should be given in the text. Otherwise Figure 2 should be removed.

Although I do not agree with all statements of the authors (see specific comments), I think that this manuscript is a very valuable contribution to the scientific field, which would be of high interest for the readers of CP. I recommend the manuscript of Visser et al. for publication in CP after minor revisions, addressing carefully my comments or questions. Actually, I am very curious about the response of the authors regarding my more controversial comments, and I am looking forward to an interesting discussion.

SPECIFIC COMMENTS

1.) Page 229, lines 1-7: What were the initial values of the variances for the process noise and the measurement noise, respectively, for the discrete Kalman filter approach? What was the initial variance of the process noise eta(i,t) for the response function? What was the initial variance of the process noise eta(t) for the IRM trend model? In the Appendix A, the same symbol eta is used for the two different noise processes. I would find different symbols clearer. How were the initial values for the noise variances estimated? The choice of these initial noise variances can be quite critical for the performance of the model estimated by the discrete Kalmna filter. Have you performed some sensitivity test for different initial values for the noise variances? For the discrete Kalman filter approach, the process noise and the measurement noise are assumed to be independent (of each other), white, and with normal probability distributions. I think that a discussion about the appropriateness of these assumptions would be appropriate at this point. Were the model residuals checked for serial correlation and their distribution?

2.) Page 229, lines 9-11: What do you mean with “letting the data choose between being constant or time-varying. . .”. The measurement data can not “choose” what they want to be. They are what they are. Probably you mean that the model parameters can either be constant or time-varying depending on the data – but also (importantly!) on the choice of the initial noise variances by the researcher.

3.) Page 229, lines 21-24: Please find another phrasing than “. . .something which is best avoided.”. This cannot be the reason to use a model with time-varying intercept. If alpha shows time-dependent behavior, then this has to be discussed, and it can also be real. To demonstrate this situation, should not be avoided.

4.) Page 232, lines 6-9: I find this suggestion very important. Thorough testing for non-linearities should be enforced in future paleoclimate reconstruction studies. The question of non-linearities in the proxy-target relationship in the calibration period and the reconstruction period is very difficult, and more research on it would be necessary.

5.) Page 232, lines 20-23: I think that the identification of such a climate envelope would be a reasonable first step; however, the main problem of nonlinear relationships in the past would not be covered by such am envelope. The problem is that we do not know how warm it was for example in the MWP; we only see the tree-rings. If the temperatures in the MWP were above the temperatures in the second half of the 20th century, it could be that we do not see this in enhanced tree-ring widths due to nonlinearity, e.g., because trees may have grown slower under too high temperatures due to for example moisture stress (see Loehle, 2009 for a insightful discussion of the problem). In other words, not the extraordinary thick tree-rings of old trees are the problem – these could be easily excluded by the climate envelope. The problem are the

tree-rings which are in the normal width range but had been grown under hypothetically extreme warm conditions.

6.) Page 233, lines 2-3: Please define mu(s) and alpha(s) here in the text.

7.) Page 233, lines 7-8: What do you mean with “2-alpha significance level” and with “1-alpha significance level”. How was alpha defined?

8.) Page 233, lines 21-23: Why different detrending methods?

9.) Page 234, lines 14-15: I think that this should be explained in a little more detail. I appears to me that this definition of “explained variance” considers only the variance explanation by the response function alpha(t) X(t), not of the whole model including the time-varying intercept. Is this right?

10.) Page 234, after line 22: I think that you should describe in this section also the results shown in Figure 2. Figure 2 is now not described nor discussed at all.

11.) Page 236, lines 13-17: I am strongly opposed to such an approach. By just omitting the part of the calibration period where the response function and the trend behaves differently than before, the problem is not solved at all in my view. It is very clear that you can loose your ability to detect instabilities if you shorten the calibration time series, but this does not mean that they do not exist anymore. On the other hand, if you are able to detect significant instabilities with the approaches presented here, then you have to expect similar problems also in the past. Thus, I would argue that the presented approach would be really very useful for screening proxies for reconstruction studies in the sense: If proxies already show significant instabilities during the calibration period, there is a strong indication that it has these instabilities also in the past, which means that it cannot be useful for unbiased reconstructions of past conditions. Proxies which can be shown to be stable in their response to climate during the calibration period can be expected with some confidence that they were also stable during the past. But of course, also this can be rather uncertain, and this uncertainty should also be discussed.

12.) Page 237, lines 10-21: See comment no. 11 above and also comment no. 5 on the climate envelope.

13.) Page 239, lines 8-11: However, I would argue that the detection of instabilities by the presented approach should be used as a strict screening criterion. Just shortening the calibration period does not solve the problem in my view. If the presented approach to identify instabilities in the proxy-climate relationship over the calibration period would be used in a strict sense and without just shortening the calibration period, the resulting screening would exclude probably more proxies than other methods used in previous studies (See also comment no. 11)

14.) Page 240, list of conclusions: I completely agree with conclusions no. 1, no. 3, and no. 5. Regarding no. 2, I have another opinion. See comments no. 11 and no. 13 above. Regarding no. 4, I think that the detection of instabilities by the presented approaches should be used as a strict screening criterion. If this is done, the new screening method would filter out more proxies than was done by methods used in previous studies. (See comment no. 13).

15.) Page 241, lines 14-21: Please explain here more in detail what RE (reduction of error statistic) and CE (coefficient of efficiency) stands for, not only referring to the literature. RE is a model performance measure calculated for the calibration period, and the coefficient of efficiency is a model performance measure calculated for a validation period. This difference should be shortly explained.

TECHNICAL COMMENTS

16.) Page 238, line 2: Insert comma after “approach”.

17.) Page 238, line 6: Remove comma before “since”.

18.) Page 238, line 13: Insert comma after “cases”.

19.) Page 238, line 15: What do you mean with “tree parameter”?

20.) Page 239, line 3: Insert comma after “series”.

21.) Page 239, line 4: Please write hear more precise: “Squared correlation coefficients” or “Coefficients of determination” would be better than “Squared correlations”.

22.) Page 240, line 16: Insert comma after “examples”.

23.) Page 240, line 20: Insert comma after “reconstructions”.

24.) Page 241, line 19: R2 is commonly named “coefficient of determination”.

25.) Page 242, line 8: I think that sigma(epsilon) should be squared.

26.) Page 242, line 9, lines 13-14. I think that it would be better to use different symbols for the noise processes disturbing the response function and the IRW trend model, respectively.

27.) Page 251, Figure 1, and page 254, Figure 3: The trend difference on the y-axis in the lower panels of the figures should have a unit, propably yrˆ-1.

28.) Page 253, caption of Figure 3, lines 3-4: Please modify to “. . .shows the time-varying response weights, and. . .”

29.) Pages 250-251, Figure 1, and Pages 253-254, Figure 3: I would combine the panels in one graph for each example.

Interactive comment on Clim. Past Discuss., 6, 225, 2010.