Clim. Past Discuss., 2, S221–S225, 2006
www.clim-past-discuss.net/2/S221/2006/

Climate of the Past Discussions

Interactive comment on "On the verification of climate reconstructions" *by* G. Bürger and U. Cubasch

Anonymous Referee #3

Received and published: 21 July 2006

Bürger and Cubasch's reply does not address the central points of criticism in my earlier review and construes overly narrowly some specific criticisms stated as illustrations of larger problems. The reply exacerbates concerns I expressed in my earlier review (cf. point 7 below).

In direct response to BC's reply, I would like to emphasize a few points that appear to have been confusing. (Numbers refer to the numbers of the specific comments in my earlier review and in BC's reply.):

2. My criticism was not about the specific "subset" of data used. I criticized that broad methodological claims and conclusions were based on an analysis of narrow scope. It should be clear that it is a logical fallacy to draw general

conclusions—particularly negative ones, such as that about the Medieval Warm Period—based on an analysis of a narrow set of specific examples. This is particularly true if the claims are methodological but methods are inconsistently used (see, e.g., point 7 below and in the earlier review).

3. Zorita's statement that "the inference step ... was performed with non-detrended data [whereas] the calibration step was performed with detrended data" exactly illustrates the consistency problem I pointed out in my earlier review.

   If data are detrended in the calibration (estimation) step, this means that a regression model of the form

$$T = B_0 + PB_1 + B_2 t + \varepsilon, \tag{1}$$

   with a matrix T of temperatures, a matrix P of proxies, a time variable t, and parameter vectors and matrices B_0, B_1, and B_2 is estimated from the data. Here, B_0 is an intercept vector (which in least squares estimation can be omitted if the data are centered), B_1 is the matrix of regression coefficients, and B_2 is a vector of linear trend coefficients. The estimation procedure (whatever it might be) yields estimates $\hat{B}_0$, $\hat{B}_1$, and $\hat{B}_2$. A consistent inference step would then use these estimates in model (1) to estimate expected values of temperatures $\hat{T}$ given proxies P:

$$\hat{T} = \hat{B}_0 + P\hat{B}_1 + \hat{B}_2 t.$$

   However, if the inference step is "performed with non-detrended data," a regression model of the form

$$T = B_0' + PB_1' + \varepsilon' \tag{2}$$

   is implicitly assumed, for which a consistent inference step would be

$$\hat{T} = \hat{B}_0' + P\hat{B}_1'$$

CPD

2, S221–S225, 2006

Interactive Comment

with estimates $\hat{B}_0'$ and $\hat{B}_1'$. However, if the inference step is based on the estimates from detrended data (i.e., model (1)),

$$\hat{T} = \hat{B}_0 + P\hat{B}_1,$$

inconsistent estimates may result since the primed and unprimed estimates are not necessarily equal.

This is the central consistency problem I pointed out in the review.

4. My point was not that the writing needs to be expanded; it needs to be made more precise and concise.

5. My point was not narrowly to question what a “population” is; it was to show that the authors’ statements can be trivially shown to be misleading and that they miss an important necessary condition (missingness at random, MAR) for applicability of the methods they consider.

   Missingness does not refer to missingness of instrumental temperature values alone but to missingness of temperature values in general, including historical temperature values. In reconstructing historical temperature values, an important necessary condition for applicability of methods of the kind BC consider is that missingness of historical temperature values is independent of the missing historical temperature values (MAR assumption). This assumption may be questionable if missingness correlates with temperature values, as the correlation of missingness with temperature values shows. (I understand that the authors viewed the number of available grid points as a “nonsense” predictor; however, the correlation between the number of available grid points and temperatures shows that MAR is violated. The consequences of this violation need to be scrutinized.)

   The last paragraph in BC’s reply under point 5 shows that BC misunderstand the concept of MAR. MAR does not mean that the pattern of missing values (in space

CPD

2, S221–S225, 2006

EGU

or time) is random. Please consult a standard statistics text (such as that by Little and Rubin referenced earlier) for clarification of the concept.

7. Using RegEM “solely for the estimation of mean and covariance [matrices]” but using other methods for infilling is a methodological error. Missing values, mean values, and covariance matrices are related by models of the form (1) or (2). It is a methodological error to use one method for estimation of missing values and another method for estimation of mean values and covariance matrices. This error invalidates BC’s claims about the performance of RegEM. It cannot lead to consistent estimates of temperatures and covariance matrices.

9. Rescaling is a methodological error since it introduces biases in estimated temperature values even if they were unbiased before rescaling. Variance estimates have to be consistent with the underlying model (such as (1) or (2)), which means that the variance of the residuals ε or ε' must be added to the sample variance of the reconstructed temperatures $\hat{T}$ to obtain an estimate of the variance of temperatures T.

   If the same methodological error was made in other publications, it should be corrected here, not perpetuated.

   Regarding variance attenuation in RegEM or other reconstruction methods, there are two issues. First, the sample variance of any temperature reconstruction $\hat{T}$ is a biased estimate of the variance of the actual temperatures T because of the variance of the residuals in stochastic models such as (1) or (2). In the well-posed and normal case, a consistent and unbiased variance estimate can be obtained, as in the EM algorithm, by taking the residual variances into account in variances estimates. Second, the sample variance of a temperature reconstruction $\hat{T}$ has an additional bias if regularized (biased) parameter estimates are used in stochastic models such as (1) or (2) (as, for example, in principal component regression or ridge regression). This bias is more difficult to correct.

CPD

2, S221–S225, 2006

Variance biases cannot be corrected by rescaling temperature reconstructions $\hat{T}$. Rescaling gives the misleading impression of having temperature "signals" of large amplitude when in fact only the "noise" variance is large.

As stated in my original review, I expect a paper that sets out to reassess "the skill of proxy-based reconstructions of Northern Hemisphere temperature" to address carefully questions such as those of biases of variance estimates.

10. My comments were concerned with the paper under consideration, not with other papers. My criticism that the authors made general claims about well-posedness that were not sufficiently scrutinized stands unmodified.

12. The authors cannot seriously expect me and other readers to go through some third-party code and a set of parameter names specific to that code to understand what they are doing.

Interactive comment on Clim. Past Discuss., 2, 357, 2006.

CPD

2, S221–S225, 2006

Interactive Comment