Testing the influence of biased EOF centering: a comment to "Hockey sticks, principal components and spurious significance" by S. McIntyre and R. McKitrick

Hans von Storch and Eduardo Zorita

GKSS Research Centre, Geesthacht, Germany

We analyse a climate simulation of the last millennium to check whether the “artificial hockey stick” (AHS) effect introduced by biased centering has a significant bearing on the performance of historical climate reconstructions. The “hockey stick” shaped reconstructions of the northern hemisphere temperature has been a contested icon in climate science since a it was advanced by the IPCC as likely temperature history, against which the recent warming trends should be evaluated [*Intergovernmental Panel on Climate Change*, 2001]. This reconstruction has by now faced a number of challenges on different grounds. One of challenges was brought forward by McIntyre and McKitrik [2005] (hereafter MM05), who had noted that the original code contained an uncommon mathematical procedure. In this study we examine whether this uncommon procedure, which under certain circumstances can result in “artificial hockey sticks”, would affect the final result of the reconstruction.

The statistical method behind the “hockey stick” temperature reconstruction [*Mann et al,,* 1998] (hereafter MBH98) is based on an inverted regression method, which maps proxy data on the Northern Hemisphere temperature field. The proxy-data are irregularly distributed over the globe, and in some regions their spatial density (mainly for dendroclimatological data) is high. To avoid overweighting these regions, a Principal Component Analysis (PCA) was used to condense these spatially clustered proxy data into a few principal components. MM05 noted that MBH98 normalized their data unconventionally prior to the PCA, by centering the timeseries relative to the instrumental-period mean, 1902-1980, instead of relative to the whole available period. Why this was done is unclear, and it is certainly not a wise measure. It is, however, not entirely uncommon in climate sciences.

MM05 performed a Monte Carlo study with a series of independent red-noise series; they centered their 1000 year-series relative to the mean of the last 100 years, and calculated the

PCs based on the correlation matrix. It turned out that very often the leading PCs show a hockey stick pattern, even if the data field was by construction free of such structures. This finding was recently confirmed by others (*F. Zwiers, pers. comm*.). The paradox in the AHS effect is that the true covariance matrix is a unity matrix, so that no real structures will steer the eventual selection of the eigenvectors. However, in the biased centering approach, those timeseries with largest differences between their 1000-1901 mean and 1902-1980 mean will tend to contribute more strongly to the leading PCs, thus producing an artificial hockey-stick shape. The MBH98 algorithm, however, involves several other steps and it is not clear if the AHS-effect carries any relevance for the final temperature reconstructions.

We had previously used multicentennial climate simulations with the coupled models ECHO-G and HadCM3 to test the MBH98 reconstruction method [*von Storch et al.*, 2004]. The models were driven by estimations of past solar irradiance, radiative effects of volcanic eruptions and concentrations of greenhouse gases for the past 1000 years. In that test simulated grid-point temperatures, collocated with the frozen-in complete proxy network of MBH98 and degraded with statistical noise, played the role of pseudo-proxies. In the von Storch et al. [*2004*] test, the PCA of the pseudoproxies was not included, as the resolution of the model is coarse enough to prevent excessive clustering. Thus the AHS effect could not play a role in that analysis.

We have redone the test, this time including the previous PCA of the pseudoproxies. The PCA was applied to the pseudoproxies after AHS-centering or after centering relative to the millennial mean. The PCA was carried out with annual values in three areas separately (North America, South America, and Australia). The numbers of PCs retained for the subsequent steps was decided from the eigenvalue spectrum following accepted rules [*North et al.*, 1982], but these (within a reasonable range) had only a very minor influence on the results. Two types of noise were tested: white noise and red-noise. A guideline for the amount of added white noise is the local correlation r between real proxies and nearby temperature observations, which usually lies in the range r=0.3-0.7 [*Jones and Mann* , 2004]. This corresponds to a noise variance between 85% and 50% of the total variance [*Storch et al.*, 2004]. The level of noise at centennial timescales, or alternatively the steepness of the spectrum of a AR-1 noise, is very uncertain, so that only rough guesses can be used.

Figure 1 shows the result of these pseudo-reconstructions for one realization of the white noise (top panel, with noise variance 50%) and one realization of the red noise (bottom panel, high-frequency noise variance 50% and with 1-year lag autocorrelation of α=0.8) : both PCAcenterings has a small relevance for the final result and the differences are within the uncertainty range (Figure 1). The conclusion is essentially the same for all realizations and other constructions of noise. For instance, white noise with r=0.7 yields a standard deviation of the differences of 0.006K; r=0.4 yields 0.007K; red noise with α=0.5 and r=0.7 (r=0.4) yields 0.01K (0.02K); red noise with α=0.8 and r=0.7 (r=0.4) yields 0.02K(0.03K). Therefore, the differences increase slightly with the amount and redness of the noise, but they remain small, even in the case of high and red noise with a steep red spectrum.

Our results, derived in the artificial world of an extended historical climate simulation, indicate therefore that the AHS does not have a significant impact but leads only to very minor deviations. We suggest, however, that this biased centering should be in future avoided as it may unnecessarily compromise the final result.

Finally, we note that we have strictly addressed here the question of the PCA-centering within the MBH98 algorithm. Other concerns raised by MM05 [see, e.g., *Crok*, 2005] about the MBH methodology have been not dealt with.

Acknowledgement. This work was accomplished within the EU project SO&P.

References

Crok, M. (2005), Kyoto Protocol based on flawed statistics. *Naturwetenskap & Techniek 2*, 20-31.

Intergovernmental Panel on Climate Change (IPCC) (2001), *Climate Change 2001: The scientific basis: contributions of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, edited by J.T. Houghton et al., 881 pp, Cambridge University Press, New York.

Jones, P. D., and M. E. Mann (2004), Climate over past millennia. *Rev. Geophys.*, *42*(2)}RG2002,doi:10.1029/2003RG00143.

Mann, M.E., R.S. Bradley and M.K. Hughes (1998), Global-scale temperature patterns and climate forcing over the past 6 six centuries. *Nature 392*, 779-787.

McIntyre, M., and R. McKitrick (2005), Hockey sticks, principal components and spurious significance. *Geophys. Res. Lett. 32*, L03710, doi:10.1029/2004GL021750, 1-5.

North, G.R., T.L. Bell, R.F. Cahalan and F.J. Moeng (1982), Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev. 110*, 699-706.

von Storch, H., E. Zorita, J. Jones, Y. Dimitriev, F. González-Rouco, and S. Tett (2004), Reconstructing past climate from noisy data. *Science 306*, 679-682.

List of Figures

Figure 1: Time-filtered Northern Hemisphere annual near-surface air temperature anomalies in the last millennium (relative to the mean of the calibration period 1902-1980), as simulated with the model ECHO-G [*von Storch et al.*, 2004] (black thick line), and resulting from the application of the MBH98 algorithm to a network of pseudo-proxies taken from the same simulation and degraded with stochastic noise (thick colored lines: one noise realization close to the mean of 100 realizations). The variance of the pseudoproxies contains 50% noise (top panel: white noise; bottom panel: red noise with one-year-lag autocorrelation of 0.8). The pseudoproxies were subjected to separate PCA in North America, South America and Australia with full (1000-1980; red) or partial (1902-1980; blue) centering. Thin colored lines represent the 2σ reconstruction scatter (after time filtering) estimated from 100 realizations of the noise.

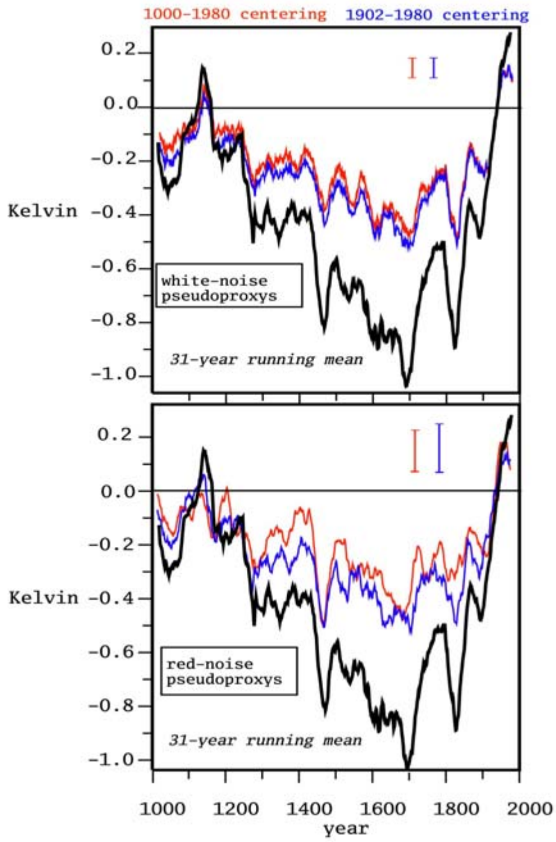

Figure 1: Time-filtered Northern Hemisphere annual near-surface air temperatures anomalies in the last millennium (relative to the mean of the calibration period 1902-1980), as simulated with the model ECHO-G [*von Storch et al.*, 2004] (black thick line), and resulting from the application of the MBH98 algorithm to a network of pseudo-proxies taken from the same simulation and degraded with stochastic noise (thick colored lines: one noise realization close to the mean of 100 realizations). The variance of the pseudoproxies contains 50% noise (top panel: white noise; bottom panel: red noise with one-year-lag autocorrelation of 0.8). The pseudoproxies were subjected to separate PCA in North America, South America and Australia with full (1000-1980; red) or partial (1902-1980; blue) centering. Thin colored lines represent the 2σ reconstruction scatter (after time filtering) estimated from 100 realizations of the noise.