Reply to Huybers

Stephen McIntyre
512-120 Adelaide St. West,
Toronto, Ontario Canada M5H 1T1
stephen.mcintyre@utoronto.ca

Ross McKitrick
Department of Economics,
University of Guelph,
Guelph Ontario Canada N1G 2W1

Principal Components: Covariance or Correlation?

McIntyre and McKitrick [2005a, "MM05"] showed that the actual MBH98 PC method used an unreported short-centering method, which was biased towards producing a hockey stick shaped PC1 with an inflated eigenvalue. Huybers concurs with these particular findings, but argues that we "exaggerated" the MBH98 bias by comparing the MBH98 PC1 for the North American tree ring network to a covariance PC1 rather than a correlation PC1 in our Figure 3.

Tree ring chronologies (both density and ring width) are already standardized to common dimensionless units with a mean of 1. Huybers' two statistical authorities either do not recommend standardizing variance for PC analysis on series with common units [Preisendorfer 1988] or recommend the opposite (i.e. a covariance PC calculation) [Rencher 1995, p.430; see also Overland and Preisendorfer 1982; Rencher 1992]. Only Rencher [1995] even mentions the possibility of standardizing variance of networks in common units in exceptional circumstances which do not apply here.

One of Huybers' principal justifications for proposing a correlation PC1 is his observation that the covariance PC1 underweights density series, which have lower variances. But in the MBH98 network, only 2 of 70 series are density series, and both are from sites also represented in the same network with a ring width series. Indeed, the Spruce Canyon site (density series co509x and ring width series co509w) also occurs in 4 series in the MBH98 Stahle/SWM network. Accommodating these 2 density series should not be at the expense of the most appropriate treatment for the other 68.

Relative to the MBH98 PC1, the differences between the covariance PC1 and correlation PC1 are trifling and both confirm the bias reported in MM05. The MBH98 method applied to North American tree rings had the distinctive hockey stick shape (Figure 1 top left) and a very large first eigenvalue (38%), which they interpreted as evidence of a "dominant component of variance". Neither the covariance PC1 (second left) nor the correlation PC1 (third left) have a hockey stick shape and their first eigenvalue is much reduced (19%, 17% respectively). The correlation PC1 is a little higher in the 20th century than the covariance PC1, but the differences are trifling.

Huybers' approach also ignores fundamental properties of the data and introduces new and unconsidered biases:

a) Tree ring chronologies are typically autocorrelated, especially the controversial bristlecones. For autocorrelated series, the ordinary least squares sample variance (used by Huybers) is a biased (under-) estimate of the long-run variance, so the bristlecones will tend to be over-weighted this way. An "unbiased fully normalized" PC1 can be obtained using an autocorrelation-consistent variance estimator [e.g. Andrews, 1991]. This bias correction yields a result (Figure 1, bottom left) very similar to the covariance PC1.

b) Huybers argues that the correlation PC1 captured a "robust feature of the NOAMER dataset" based on its similarity to the mean of the 70 series in the AD1400 network scaled by their standard deviation (Fig 1, third right panel). If the purpose of PC analysis is merely to predict the mean, then there is no reason not to simply use the mean. As for the robustness of the feature, only 70 of 212 series in the NOAMER network extend back to AD1400. Using all 212 NOAMER series scaled by their standard deviation (Fig. 1, second right panel) yields a network mean closer to the covariance PC1 than the correlation PC1—by this criterion "full normalization" adds bias to the PC1.

The differences among these PC series can be traced to differing weights for bristlecones. Bristlecone sites are well-known examples of CO_2 fertilization and their nonclimatic biases have been extensively assessed already—see the caveat in IPCC [1996] and the review in McIntyre and McKitrick

[2005b]. Huybers is correct to acknowledge the need to assess their validity and possibly to exclude or down-weight them, but having said so it is inadequate to defer this to "future studies." We see no sense deferring to the future a remedy for what is already a well-understood source of bias.

Bristlecone impact can be seen directly by comparing the MBH98 PC1 (top left), which is weighted almost entirely from bristlecones, with an unreported PC1 from Mann's FTP site (top right), which Mann obtained by applying MBH98 PC methodology while excluding 20 bristlecone sites [ftp://holocene.evsc.virginia.edu/pub/MBH98/TREE/ITRDB/NOAMER/BACKTO_1400-CENSORED/pc01.out]. Given the tendency of the MBH98 method to yield hockey stick shaped PC1s [MM05], it is remarkable that this PC1 does not have a hockey stick shape. The correlation PC1 without bristlecones (bottom right) is virtually identical, showing that the actual PC method has little impact in this case once the bristlecones are removed. By trying to accommodate 2 density series, Huybers ends up inflating the weight of the very proxies in controversy.

In our previous Figure 3, we tried to illustrate the difference in a critical network between the actual short-centered MBH98 PC method and the covariance PC1, because we believed that this was the most appropriate implementation of the reported MBH98 method (said only to be "conventional") in a network already standardized to dimensionless units, a view which we still hold. Huybers argues that this exaggerates the impact. We disagree for all the reasons set out above. In addition, Huybers' own Figure 3 relies on a questionable rhetorical effect. Instead of centering his graph on the 1400-1980 interval used for calculations, Huybers re-centered his plotted data on the 1902-1980 mean, a method surely inconsistent with his program of "full" normalization and rather ironic in the context of MM05 criticism of short-segment centering. When centered on the 1400-1980 interval (as in our Figure 1), all PC versions other than the MBH98 PC1 are quite similar. (See also Supplementary Figure 1, re-doing Huybers' Figure 3 centered on 1400-1980.)

We re-emphasize that our comparison between the MBH98 method and a covariance PC1 was not presented as an attempt to "remove the bias in MBH98's method", and that we take no position on the relative merits of using a mean, a covariance PC1, or even using PC analysis at all, in paleoclimate work. The onus for demonstrating validity of a statistic as a temperature proxy rests entirely with its advocate. Any valid climate reconstruction should not depend on whether a correlation matrix or covariance matrix is used in tree ring PC analysis. No variation on PC methodology can overcome the problems of using bristlecones as a temperature proxy.

The RE Benchmark

Huybers did not dispute our estimate of the MBH98 cross-validation R^2 statistic (~0.0; pers.comm.); but he also argued that the MBH98 RE statistic was significant. If a proxy index has a true relationship to temperature, it is statistically impossible for its cross-validation R^2 to be ~0.0. Hence there is an apparent contradiction between the two statistics which Huybers did not resolve.

We had argued that the apparent MBH98 RE significance was "spurious". Huybers argued that our demonstration of spuriousness was flawed because we used an inverse regression method to fit the simulated PC1s to NH temperature. He proposed scaling the calibration period variance of the simulated series to the observed variance. Using this method, he reported a 99% benchmark of 0.0, arguing that this implied statistical significance for MBH98.

The variance re-scaling step called for by Huybers was not mentioned in MBH98. However, recently-released code (ftp://holocene.evsc.virginia.edu/pub/MANNETAL98/METHODS/multiproxy.f) shows that MBH98 included a re-scaling step. If simulations are done with variance rescaling on a simulated tree ring PC1 without constructing a full network of other proxies and calculating a NH temperature index, we agree that the 99% quantile is ~0. However, Huybers' simulations do not fully emulate MBH98 methodology, as his simulations did not replicate the effect of a proxy network.

We did new simulations in which we took 1000 simulated PC1s saved from the simulations described in MM05; for each PC1 in turn, we made a "proxy network" of 22 series with the other 21 being white noise (replicating the 22 series of the MBH98 AD1400 network). We then used MBH98 methodology on the proxy network, including inverse regression of the proxies. After calculating the reconstructed temperature principal component (RPC), we scaled the variance of the RPC to the "observed" variance of the temperature principal component prior to calculating a NH average, from which we calculated an RE statistic. The 99% quantile was 0.54, down slightly from 0.59 as found in MM05a.

The apparent contradiction between verification statistics is thus fully resolved: the cross-validation R^2 of ~0.0 demonstrates that the MBH98 model is statistically insignificant; the new simulations, implementing the variance re-scaling called for by Huybers and newly-revealed in the MBH98 code, confirm our earlier finding that the seemingly high MBH98 RE statistic is spurious.

REFERENCES

Andrews, D.W.K. (1991), Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, 59, 817–858.

Huybers, P. (2005), Comment on "Hockey sticks, principal components and spurious significance" by McIntyre and McKitrick, *Geophys. Res. Let.*, 32, XXX, doi: XXXXXX.

Intergovernmental Panel on Climate Change (1996), *Climate Change 1995. The science of climate change.* Cambridge University Press, Cambridge.

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

McIntyre, S. and R. McKitrick, (2005), Hockey Sticks, Principal Components and Spurious Significance, Geophys. Res. Let., 32, L03710, doi:10.1029/2004GL021750.

McIntyre, S. and R. McKitrick (2005b), The M&M Critique of the MBH98 Northern Hemisphere Climate Index: Update and Implications, *Energy and Environment* 16, 69-99.

Overland, J.E. and R.W. Preisendorfer, (1982), A significance test for principal components applied to a cyclone climatology, *Mon. Weather Rev.,* 110, 1-4.

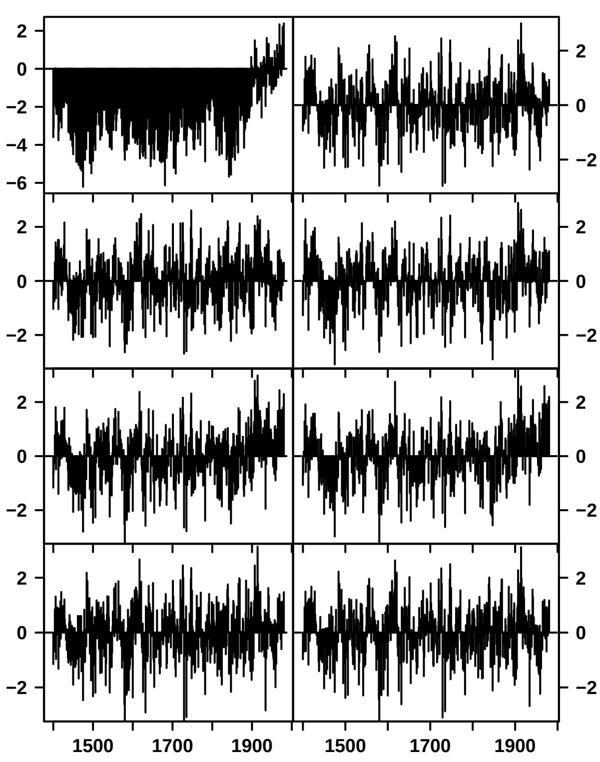
Preisendorfer, R.W. (1988), Principal Component Analysis in Meteorology and Oceanography, Elsevier.

Rencher, A.C. (1992), Interpretation of canonical discriminant functions, canonical variates and principal components, The American Statistician 46, 217-225.

Rencher, A. C., (1995). *Methods of multivariate analysis*. John Wiley and Sons.

FIGURES

Figure 1. MBH98 North American AD1400 network. **Top row:** left, MBH98 PC1 centered on 1902-1980 (as in MBH98); right, same calculation with bristlecone pines removed. **Second row:** left, covariance PC1; right, mean of all 212 series. **Third row:** left, correlation PC1; right, mean of 70 full-length series. **Fourth row:** left, autocorrelation-consistent correlation PC1; right, correlation PC1 with bristlecone pines censored.



2
0
−2
−4
−6
1500
1700
1900