Comment on "Hockey sticks, principal components, and spurious significance" by S. McIntyre and R. McKitrick.

Climate reconstructions based on proxy records require steps of standardization of the different series prior to their calibration to instrumental data. In a recent publication in GRL, McIntyre and McKitrick [2005a] suggested that the procedure applied to North American tree ring records led to a systematic bias in the famous "hockey stick" series of Northern Hemisphere temperature [Mann et al. 1998]. We show that this claim is unfounded, and that a proper standardization, independent of the reference period applied, leads to essentially the same result.

Useful Northern Hemisphere temperature reconstructions for past centuries have been available since at least 1993 [Bradley and Jones, 1993]. Since 1998, a large number of reconstructions have been published reflecting widening proxy networks and improvements in reconstruction techniques [see summary by Jones and Mann, 2004]. Of these, the pattern-based reconstruction by Mann, Bradley and Hughes [Mann et al., 1998; called MBH below] has attracted particular attention, in part because it was featured prominently in the IPCC Third Assessment Report [Houghton et al., 2001]. Because of the marked upward trend in temperature over the last 100 years following a pre-industrial period with relatively small variability, it is often referred to as the "Hockey Stick".

Recently, McIntyre and McKitrick [2005a and 2005b, subsequently called MM collectively, or MM05a and MM05b individually] have criticized a particular aspect of this reconstruction. The proxy data used in MBH include a large network of the North American tree ring data from the International Tree Ring Data Base (ITRDB). In MBH, these data are transformed through a standardization followed by Principal Component (PC) Analysis to reduce the number of predictor variables. MM claim that the standardization approach chosen by MBH biases the ITRDB information towards a "hockey stick" shape. (Because the North American records play an important role in the reconstruction of the 15th century, MM05b also conclude that the ITRDB data transformation then leads to biases in the hemispheric temperature reconstruction for that time). They present calculations based on an alternative method, which leads to results that do not exhibit a "hockey stick" shape in the first PCs. Here we show that this result is

an artifact of using only centered, but not standardized, data combined with an unchanged PC retention criterion that does not properly reflect the modification.

In MM05a, the PCs of the ITRDB data are calculated using data with the long-term mean removed. However, MM05a do not perform a division by the standard deviation. In contrast, MBH standardize relative to the 20th century (1902-1980) mean and corresponding standard deviation. MM emphasize that the "hockey stick" shape is introduced because the standardization is performed relative to a subsection rather than the full series. It turns out that the critical difference between these two approaches for both PC calculation, as well as the subsequent climate reconstruction, is not the reference period but the standardization itself.

Why should the different tree ring series from the ITRDB be standardized in the first place? Principal component analysis attempts to 'summarize' diverse data using a limited number of orthogonal principal components to capture the essence (i.e. the real information) contained in the data. The goal of this simplification in the case of the ITRDB data is to identify what the common climate signals are in the heterogeneous set of North American tree ring series. Heterogeneity is particularly present in tree ring networks that include different species, potentially different tree ring parameters (width versus cell density) and that span such geographically diverse areas as the Rocky Mountain West and Eastern North America. Division by the standard deviation (scaling) is employed in order to make the series more directly comparable. Although this is not a necessary step, it does strongly affect the resultant PCs and the number of PCs required to capture an adequate amount of the variance in the original data. In MBH, two PCs are retained. The first PC has the distinctive "hockey stick" shape. In the MM analysis, the "hockey stick" PC appears in PC4 – yet the authors choose to retain only the first two PCs. If MM had employed the step of dividing the individual series by their respective standard deviations, and thus would put all individual series onto the same 'footing', even though they center over the full series length (rather than the calibration period as in MBH), they would have captured the "hockey stick" shape with only two PCs. These results are shown in Fig. 1.

Fig. 1 shows various PC calculations for the AD 1400 network [data from MM05a supplementary information]. Fig. 1a shows PC1 (solid line) and PC2 using MBH standardization; Fig. 1b shows PC1 and PC4 (solid line) based on MM-centering without scaling. Note that the upward trending component after about 1850 that appears in MBH PC1 is in PC4 in the MM case. PC1 in Fig. 1b is the same as in Fig. 3b in MM05a.

Fig. 1c shows PC1 and PC2 (solid) if the ITRDB-data is not only centered but also scaled by their respective standard deviation over the full data period. Note the similarity between these results with the MBH results in Fig. 1a. The difference consists of a reversed allocation of PC1 and PC2 with regard to the primary "hockey stick" shape, with part of the signal of 20th century offset already present in PC1 (and thus split between PC1 and PC2). That the first two PCs are in fact capturing nearly the same information is even more striking if PC1 and PC2 series are added (Fig. 1d). The minor amplitude difference in the 'shaft' results in a deviation of five hundreds of a degree in the subsequent climate reconstruction if the revised MM method (including scaling) is applied (Wahl and Ammann, in review).

No matter what standardization procedure is applied, a "hockey stick" shape appears in the important PCs – as PC1 in MBH, as PC4 in MM05a, and as PC2 in our revision of MM05a. While it is true that the choice by MBH of performing their standardization relative to the calibration period did force essentially all of the "hockey stick" component present in the underlying data already into the first component (rather than spreading it over several PCs), the MM-claim that a "hockey stick" outcome in the PCs is an artifact of the MBH standardization procedure is incorrect. Already, if all the raw tree ring data are averaged, there is a distinct upward shift in the last 100 or so years. But most importantly, if all proxy series are used in a framework where they are comparable, i.e. if they are not only centered but also scaled by their respective standard deviation, then the "hockey stick" pattern is prominent in the first two PCs.

We have also examined the MM approach for benchmarking the RE statistic presented in MM05a. Although the MM method generates realistic pseudoproxy series with autocorrelation (AC) structures like those of the original proxy data, these time series have nearly uniform variances, unlike those of the original proxies. PCs derived from such data generally have AC structures *unlike* those derived from the original proxies, and thus they should not be used as equivalent to the original PCs. Restoring the variances of the original proxy data to the pseudoproxy series yields PCs with AC structures like those of the original PCs. But more importantly for the benchmarking, we confirm Huybers' (2005) correction to the MM RE calculations, which rescales the variance of the fitted NH temperatures to match that of the observed values regressed against the simulated PC1s. This approach more accurately mimics the actual MBH procedure, which applies a parallel rescaling to the fitted instrumental PCs that drive the MBH climate field reconstruction process. Using our AC-correct PC1s, RE = 0.0 occurs at the 0.985 level of significance.

In summary, different standardization procedures prior to principal component analysis can change the order of PCs in which the analysis is going to extract information. However, if properly performed, all approaches that capture an acceptable amount of the variance in the underlying proxy data lead to essentially the same reconstruction results. This is also the case if no standardization is applied (or even when all tree ring series are entered individually into the climate reconstruction rather than through PCs). The "hockey stick" appears in all the summaries because it is an important part of the ITRDB network. The claim by MM that a spurious "hockey stick" climate reconstruction is introduced by data transformation is unfounded. Finally, if implemented correctly, the MBH method RE benchmark remains at 0.0 to distinguish between meaningful and meaningless reconstructions.

References:

Bradley R.S. and P.D. Jones, 1993: "Little Ice Age" summer temperature variations: Their nature and relevance to recent global warming trends, The Holocene, *3*, 367-376.

Houghton J.T., et al., 2001: *Climate Change 2001. The scientific basis*. Cambridge University Press, 881pp.

Hybers P., 2005: Comment on "Hockey sticks, principal components, and spurious significance" by McIntyre and McKitrick. *Geophys. Res. Lett. (this issue)*

Jones P.D., and M.E. Mann, 2004: Climate over past millennia. Rev. Geophys., 42, RG2002, doi:10.1029/2003RG000143, 2004.

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, 2005a: Hockey sticks, principal components, and spurious significance. Geophys. Res. Lett., *32*, L03710, doi:10.1029/2004GL021750, 2005.

McIntyre S. and R. McKitrick, 2005b: The M&M critique of the MBH98 Northern Hemisphere climate index: Update and implications. Energy and Environment, 16(1), 69-100.

Wahl E.R. and C.M. Ammann, in review: Robustness of the Mann, Bradley, Hughes Reconstruction of Surface Temperatures: Examination of Criticisms Based on the Nature and Processing of Proxy Climate Evidence. *Climatic Change,* in review.

Figure 1: Comparison of ITRDB PCs contained in the 1400 network for standardization procedures proposed by (a) MBH (PCs 1 and 2), (b) MM05a (PCs 1 and 4) and (c) MM_revised (PCs 1 and 2). Panel (d) illustrates the sum of vectors PC1 and PC2 for MBH and MM_revised calculations. Dashed lines in panel (b) and (c) illustrate the offset of the 20th century mean to the earlier centuries in PC1.