Mann, M.E., R.S. Bradley, and M.K. Hughes, 2003: "Note on paper by McIntyre and McKitrick in Energy and Environment," University of East Anglia Climatic Research Unit website. Available online at:
http://www.cru.uea.ac.uk/~timo/paleo/EandEPaperProblem_03nov03.pdf

NOTE ON PAPER BY MCINTYRE AND MCKITRICK IN "ENERGY AND ENVIRONMENT"

Michael E. Mann, Raymond S. Bradley, Malcolm K. Hughes

The recent paper by McIntyre and McKitrick (*Energy and Environment*, 14, 751-771, 2003) claims to be an "audit" of the analysis of Mann, Bradley and Hughes (*Nature*, 392, 779-787, 1998) or "MBH98". An audit involves a careful examination, using the same data and following the exact procedures used in the report or study being audited. McIntyre and McKitrick ("MM") have done no such thing, having used neither the data nor the procedures of MBH98. Thus, it is entirely understandable that they do not obtain the same result. Their effort has no bearing on the work of MBH98, and is no way a "correction" of that study as they claim. On the contrary, their analysis appears seriously flawed and amounts to a gross misrepresentation of the work of MBH98. The standard protocol for scientific journals receiving critical comments on a published paper is to provide the authors being criticized with an opportunity to review the criticism prior to publication, and offer them the chance to respond. Mann and colleagues were given no such opportunity.

It seems clear that MM have made critical errors in their analysis that have the effect of grossly distorting the reconstruction of MBH98. Key indicators of the original MBH98 network appear to have been omitted for the early period 1400-1600, with major consequences for the character of the MM reconstruction of Northern Hemisphere temperatures over that interval.

MM do not list the number of indicators in *their* putative revision of the MBH network (which is based on a complicated combination of original data from MBH98 and data substituted from other sources). The reader must do a considerable amount of detective work, based on scrutiny of the Tables in their pages 20-23 and the indicated data links, to determine just what data have been eliminated from the original MBH network. A preliminary attempt to do this already demonstrates that their deletion of key early proxy information produces anomalous warming in the 15th century at odds with the reconstructed cold conditions of the period by MBH98 and virtually all other published Northern Hemisphere temperature reconstructions.

MM appear to have eliminated key proxy indicators from the MBH98 network by the following actions:

1) MM (see their Figure 4) describe the substitution of a shorter version (available back to mid 16th century) of one of the Jacoby et al (1989) Northern Treeline ring width series for the longer version (available back to mid 15th century) used by MBH98.

2) MM appear to eliminate, without any justification, the entire dataset of 70 Western North American (WNA) tree-ring series available between 1400 and 1600. This dataset, as several other regional tree-ring data networks, is represented by MBH98 in terms of a smaller number of representative Principal Component (PC) time series for each interval. The authors eliminate all of these data by not following (see technical point "b" later on in this document) the procedure of MBH98 of calculating the PC series *separately* for all intervals used in their stepwise reconstruction approach (i.e., first for 1400-1980, then for 1450-1980, then 1500-1980, and so on). The leading pattern of variance in this data set exhibits conditions from 1400-1800 that are dramatically colder than the mid and late 20th century, and a very prominent cooling in the 15th century in particular. The original individual proxy data used by MBH, including all of the WNA data have been available since May 2000 on the public ftp site provided by Mann and colleagues:

ftp://holocene.evsc.virginia.edu/pub/MBH98/TREE/ITRDB/NOAMER/, and on the NOAA Paleoclimatology website: ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/northamerica/usa

3) MM appear to eliminate the entire dataset of Stahle and coworkers of Southwestern U.S./Mexican late wood ring width measurements prior to the 17th century (12 back to 1500, 6 back to 1400) under the same false procedural premise described in (2). Once again, the data were available at the Mann et al public ftp site:
ftp://holocene.evsc.virginia.edu/pub/MBH98/TREE/STAHLE/SWM/

We have not determined yet, just how many important indicators were eliminated from the MBH98 dataset by the various subjective substitutions described by MM on pages 20-23. However, we have confirmed that elimination of the critical datasets (1)-(3) alone from the MBH98 network, which account for the overwhelming majority of proxy data used during the interval AD 1400-1500, yields the spurious result obtained by MM. (see Figure 1).

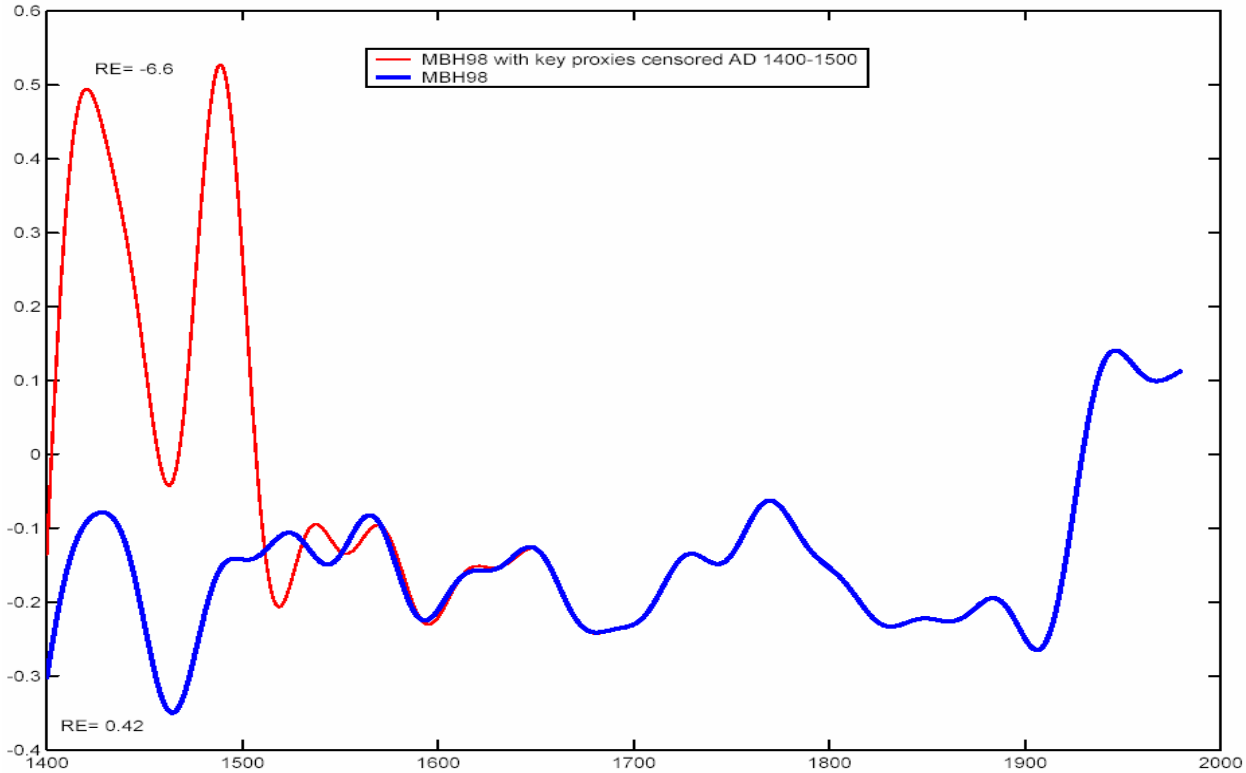

FIGURE 1. COMPARISON OF MBH98 RECONSTRUCTION (BLUE) WITH RECONSTRUCTION RESULTING FROM THE ELIMINATION OF KEY PROXY DATA SETS (1)-(3) OVER THE AD 1400-1500 INTERVAL. THIS YIELDS ESSENTIALLY THE SAME RESULT OBTAINED BY MM BY THE APPARENT ELIMINATION OF THESE DATASETS OVER THE 1400-1600 PERIOD. (BOTH SERIES HAVE BEEN SMOOTHED WITH A 40 YEAR LOWPASS FILTER).

MBH98 employed the standard statistical tool of cross-validation to verify the skill of their reconstructions. MM describe no such tests. Since increasingly sparse networks are used progressively farther back in time, a series of cross-validation experiments have to be performed to estimate the skill for different time intervals. For the AD 1400-1500 period, this involves, in MBH98, performing the reconstruction over the interval 1400-1901 based on calibration against the instrumental record over the interval 1902-1980, using the specific network of proxy indicators available for the AD 1400-1500 period. The reconstruction is then independently

compared against the instrumental record over the interval (1854-1901) not used for calibration. The skill can be described (see MBH98) by a 'Reduction of Error' statistic (RE), which is bounded by negative infinity and positive one, with substantially positive numbers indicative of predictive skill. The mean expected value for a random estimate is -1.

For the reconstruction with the data eliminated in a manner similar to that implicit in the MM approach, the RE score (-6.6) is far worse than even a typical random estimate, and such a result would have been discarded as unreliable based on the cross-validation protocol used by MBH98.

The anomalous warm values during the 15th century are the artifact of an entirely unreliable statistical estimate. By contrast, the MBH98 reconstruction indicates an RE of 0.42 for the 1400-1500 interval, indicative of significant predictive skill during that time interval.

The above discussion should be adequate to provide readers with a sense of the depth of the flaws underlying the reconstruction achieved by MM that is so at odds with at least a dozen other recently published empirical and model-based estimates of Northern Hemisphere mean temperature changes in past centuries.

There are numerous other additional, more technical problems in their approach that would appear to have rendered the MM analysis flawed irrespective of the elimination of important data. We briefly list the few most significant of these:

(a) Use of Internally Inconsistent Surface Temperature Estimates

MM appear to combine gridpoint standard deviations estimated from one version of the University of East Anglia surface temperature record, with standardized EOFs from MBH98 based on a different temperature data set. MM also appear to inconsistently use standard deviations of un-detrended data, while MBH98 had normalized their EOFs by detrended gridpoint standard deviations. MM would have avoided this problem by performing a PCA on the instrumental record themselves, rather than mixing and matching EOFs, PC series, and eigenvalues from MBH98 with standardization factors calculated from a different instrumental temperature dataset.

(b) Incorrect representation of the MBH98 proxy data set.

MBH98 calculated PCs of proxy sub-networks separately for each interval in their stepwise reconstruction. This is the only sensible approach, as it allows all data available over each sub-interval to be used (i.e., first for 1400-1980, then separately for 1450-1980, 1500-1980, and so on). This requires 159 independent time series to represent all indicators required for reconstructions of all possible sub-intervals, even though the maximum number ever used for a particular sub-interval is 112. By not following this protocol, MM appear to have eliminated in the range of 100 proxy series used by MBH98 over the interval 1400-1600.

(c) Lack of the use of an objective criterion in the determination of the number of retained instrumental PCs in the reconstruction:

Since the proxy data network developed by MM differed from that used by MBH98 for all intervals, it was inappropriate for MM to use the same instrumental temperature eigenvector subsets that had been selected by MBH98 for their reconstruction. The subsets were selected by MBH98 based on the application of an objective criterion to the specific available proxy networks available, and were optimized with respect to those networks. The basis sets used by MM have thus appear not to have been optimized with respect to the *different* proxy network they actually use.