

**Sent:** November 17, 2003 2:11 PM  
**To:** Nature  
**Subject:** Materials Complaint

Dear Sirs,

We have been carrying out a detailed examination of the paper: *Mann, M.E., Bradley, R.S. & Hughes, M.K. (1998) Global-Scale Temperature Patterns and Climate Forcing Over the Past Six Centuries*, *Nature*, **392**, 779-787, 1998 ("MBH98"). We have consulted the Nature on-line Supplementary Information, used two data files sent to us by Professor Mann's group in April 2003 (responding to a request for the FTP location of proxy data used in the above study) and, since Oct. 29, 2003, the information at the FTP site <ftp://holocene.evsc.virginia.edu/pub/>. This site is not identified through nature.com nor was it given to us upon specific request in April 2003, but is now identified by Professor Mann as the actual repository of MBH98 data.

There are a number of points in MBH98 which need substantial clarification, some of which could be resolved with through more adequate Supplementary Information. We have sought such clarification from Professor Mann without success. With reference to the policies stated at <http://www.nature.com/nature/submit/policies/index.html>, in particular item number 6, we are writing to advise you of a persistent refusal to comply with the guidelines and other issues. We draw your attention to the following issues:

- 1) There is a listing in the Nature Supplementary Information of 112 proxy series said to have been used in MBH98. In April 2003, in response to our request for an FTP site where these data could be found, an associate of Professor Mann (acting on his instructions) sent us a file (pcproxy.txt) containing 112 data series for the entire period back to 1400, together with a file apparently identifying these 112 data series. No FTP site containing these proxy series was identified to us at that time nor in subsequent correspondence with either Professor Mann nor his associate prior to October 29, 2003.

In our analysis, we identified nontrivial problems with the data in the file we had been sent. Professor Mann's associate was unable to shed any light on them. When we sought additional information from Professor Mann, he informed us that demands on his time prevented him from considering this or other inquiries.

We published our findings in *Energy and Environment* 14 (2003). Subsequently, on Oct. 29, 2003, Professor Mann stated that the errors we found in pcproxy.txt were inadvertently introduced into the file at the time it was sent to us, that we failed to consult the correct data at his University of Virginia FTP site and that our study wrongly used 112 proxy series, whereas MBH98 is based on 159 series. Upon inspecting the FTP site on October 29, 2003 (the first we learned of its existence), we identified the identical file we had been sent in April 2003 (pcproxy.txt), together with its Matlab predecessor pcproxy.mat, both dated August 8, 2002, well before our request. Sometime between October 29 and November 8, 2003, both of these files were deleted from Professor Mann's FTP site.

As to the number of proxies, MBH98 clearly refers to only 112 series, as does the on-line Supplementary Information. There is no reference to the use of 159 proxy series. We have requested a listing of these series from Mann. He has responded that the series are located on his FTP site and has refused, despite several requests, to provide more particulars. The difficulty is that the FTP site contains over 430 principal component series commencing in periods after 1400, and the relevant series are not flagged. The FTP site also includes many proxy series seemingly left out of MBH98. This level of disclosure is clearly inadequate. The deletion of archived data files in response to a controversy is also very disquieting. At a minimum, we believe that Nature should archive the 159 series now disclosed to have been used in the MBH98 calculations together with a reconciliation of the difference between the 159 series stated to have been actually used and the 112 series reported in Nature.

- 2) Some important data relating to MBH98, such as temperature principal components and EOFs, was located at Professor Mann's former University of Massachusetts FTP site. Indeed, Professor Mann's webpage on MBH98 contains no link or reference to the University of Virginia FTP site, but does contain links to this University of Massachusetts FTP site. On November 13, 2003, this FTP site was deleted. Important data pertaining to

MBH98 formerly located at the University of Massachusetts FTP site is not located at the University of Virginia FTP site.

- 3) There are numerous discrepancies between the data listing at the Nature SI site and that in the newly-disclosed FTP site at the University of Virginia. Many of the proxy series used in MBH98 are principal components calculated from larger collections of tree ring site “chronologies”. The Nature SI site provides listings of the sites reported to have been used in these calculations, with rosters identified for five regions (Oklahoma, US Southwest/Mexico, North America, South America, and Australia) plus a listing of the number of series used for a Vaganov (Russian) collection. We have compared these listings with the information on the sites actually used at the FTP site and found many inconsistencies. For example, in the South American region, only 11 of the 18 sites listed in the Supplementary Information are actually used in MBH98 calculations, according to the evidence of the FTP site. Moreover there is evidence archived at Professor Mann’s FTP site of intentional exclusion of a disclosed South American site. An archived email from Hughes to Mann, dated July 29, 1997, and located at [ftp://holocene.evsc.virginia.edu/pub/MBH98/TREE/VAGANOV/ORIG/malcolm\\_29-JUL-97](ftp://holocene.evsc.virginia.edu/pub/MBH98/TREE/VAGANOV/ORIG/malcolm_29-JUL-97) states that exclusion of the site arge030 would be “better for our purposes”. This site arge030 can be shown to be Rio Cisne (through the World Data Center for Paleoclimatology) and is listed in the Nature Supplementary Information, but is excluded from the actual data in the FTP site.
- 4) We found multiple versions at the FTP site of several series, in which a secondary series is a truncated version of the full underlying data. For instance the FTP site contains a correct and complete record of the historical annual Central England Temperature data and historical annual Central Europe Temperature data. Evidence at the site indicates that MBH98 instead used summertime averages from these series, in which the first 70 and 25 years of data, respectively, were intentionally deleted. No explanation is available for this, and the Nature SI site does not indicate or explain this editing of source data.
- 5) MBH98 states that “conventional” principal components methods were used in the calculation of temperature principal components. The underlying temperature dataset contains endemic missing data; indeed, 4 temperature grid cells from the UK Climate Research Unit selected by MBH98 contain no observations at all. “Conventional” principal components algorithms fail with missing data. Accordingly, MBH98 could not have used conventional methods and a proper description of the methodology actually employed is required. We sought specific clarification on this matter from Professor Mann and were refused. Nor is any available at the FTP site. Disclosure would be trivially easy (simply posting a command file for a statistics package would suffice). Without such information it is effectively impossible for an independent researcher to replicate his calculations.
- 6) We attempted in our *Energy and Environment* article to assess the impact of the data errors and quality control defects in the proxy data, which we had identified. For the calculation of principal components for the various tree ring regions, we used a standard principal components algorithm (princomp in R), which provides values only for periods in which there are no missing values. In response, MBH98 have stated that they used something called a “stepwise principal components approach”, a phrase which is not used in the Nature article or Supplementary Information. Based on their recent comments, this “approach” apparently involves changing the proxy rosters at some points in the PC series, leading to different values of the regional PC depending on the period being analyzed. Whatever the merits of the approach, there is no disclosure of it in MBH98, leaving the reader uninformed about the existence and formulae of a fundamental methodology as well as such basic details as the number and composition of principal components used in each region on a period-by-period bases. We have requested information from Professor Mann allowing us to identify the PC proxy rosters on a period-by-period basis and were refused. He deems it sufficient to merely point to his FTP site, but having examined this site we are unable to uniquely identify the “stepwise” PC rosters or the procedures used to splice different series together.
- 7) MBH98 states that for the temperature data “the mean was removed, and the series was normalized by its standard deviation”. This is, therefore, the method we employed. In their reply to our article, MBH criticized us on the grounds that they actually used “de-trended gridpoint standard deviations” to normalize temperature data. Whether the difference is material or not, the fact that their employed methodology contradicts the description in MBH98 means that independent researchers are unable to replicate their work.

- 8) Prior to the publication of our article, we requested other particulars on the computational methodology from Professor Mann and were refused. Accordingly, we attempted to assess the impact of the data problems by following the methodology publicly disclosed in MBH98. Professor Mann then criticized us for failing to replicate previously undisclosed details of his methodology. We once again requested particulars on his methodology, including copies of the computer programs used to read in the proxy and temperature series and to produce the Northern Hemisphere temperature index—but we have been categorically refused.

The policies of Nature rightly place a burden on authors to disclose data and methods to any interested readers. We have been systematically and deliberately stymied by Professor Mann on the most elementary requests: a proper listing of his data series and the exact computational procedures used. In the process of trying to obtain this information we have concluded that the disclosure at the Nature SI site is not merely inadequate, but in some cases it contradicts what is now revealed at the University of Virginia FTP site.

Under the circumstances, we believe that the full data set and accompanying programs for MBH98 should now be included in the Nature Supplementary Information, along with an accounting of any discrepancies between what has been listed at Nature.com to date and what was actually used in MBH98.

We can make copies of all correspondence and our extensive notes on the data issues available for your inspection if needed. Thank you for your consideration.

Yours truly

Stephen McIntyre  
Toronto, Canada

Ross McKittrick  
University of Guelph  
Guelph, Canada