Global-scale temperature patterns and climate forcings over the past six centuries: A comment.

Stephen McIntyre[1], Ross McKitrick[2]

[1]Toronto Ontario Canada; [2]Department of Economics, University of Guelph, Guelph Ontario Canada N1G2W1.

We show that the methodology used in Mann et al. (1998) for computing their northern hemisphere temperature index was flawed. We show that the distinctive hockey-stick shape of their temperature reconstruction is primarily due to tree ring sites that exhibit non-linear or non temperature-related 20th century growth and to a questionable step in their principal components algorithm that overweights these series. Problems with early segments of other "key indicators" are identified. Without these problematic series and without their biased principal component analysis, Mann et al. are not entitled to conclude that the 20th century was uniquely warm based on their data and methods.

In a recent paper[1] (herein MM03), we developed an updated and corrected version of the climate proxy data set used by Mann et. al.[2] (MBH98) to compute a Northern Hemisphere temperature index. The most significant changes were the replacement of obsolete versions of proxy data used in MBH98 with current versions from the World Data Center for Paleoclimatology (WDCP) and the use of conventional principal component analysis to reduce networks of tree ring chronologies to regional aggregates using the maximum period in which all sites were available. Applying the methodology of MBH98 to the new data yielded a Northern Hemisphere temperature index in which the values in the 15th century exceeded those in the late 20th century, thereby contradicting the conclusions in MBH98 of a unique 20th century climate warming. In their response,[3] Mann et. al. (2003) objected to our exclusion of some proxies in the 15th century, highlighting three (of 22) long series: a ring width chronology from the site at Twisted Tree Heartrot Hill ("Twisted Tree") in northern Canada; the first principal component (PC1) of earlywood and latewood ring widths from a roster of 10 sites in southwestern United States and Mexico ("SWM") studied by Stahle et al.;[4] and the PC1 of ring widths and some densities from 70+ North American sites ("NOAMER") partly overlapping the SWM network. Mann et. al. (2003) shows that the early 15th century portion of their Northern Hemisphere temperature index would, like ours, exceed the late 20th century without these series. Hence the robustness of the MBH98 conclusions hinges on the quality and availability of these three series in the 15th century. We show herein that they are all problematic.

The differences in early 15th century values result almost entirely from the NOAMER PC1. Neither the Twisted Tree series nor the SMW PC1 has a material effect on early 15th century values, although another Canadian tree ring series from Gaspé, Quebec does affect 15th century values.

MBH98 stated that the North American network (and other networks) of tree ring sites was represented through conventional principal components analysis.[5] But inspection of their Fortran software[6] shows that, in fact, the tree ring site chronologies were first transformed by subtracting the 1902-1980 mean, then dividing by the 1902-1980 standard deviation and again by the 1902-80 detrended standard deviation; then the principal components were computed using singular value decomposition on the transformed data rather than on the covariance or correlation matrix. WDCP tree ring chronologies are already scaled to a mean index value of 1000 so no transformation is needed to carry out conventional principal components analysis. For series in which the 1902-1980 mean is similar to the 1400-1980 mean, the subtraction of the 1902-1980 mean has little impact on weightings. But for those series where the 1902-1980 mean is different than the 1400-1980 mean, the variance of the transformed versions of these series are inflated relative to other series and they become overweighted in the principal components algorithm, which selects according to a variance maximization rule. The overweighting in MBH98 is increased further by using singular value decomposition on the transformed data rather than the covariance matrix. The result is that the principal components calculated in MBH98 are dominated by series that rise (or fall) in the 20th century i.e. with a hockey-stick shape.

For example, Sheep Mountain CA (ca534) exhibits the distinct "hockey stick" shape of the final MBH98 Northern Hemisphere temperature index, while another NOAMER site, Mayberry Slough AR (ar052), has a growth peak in the early 19th century (Figure 1). The MBH98 algorithm assigns 390 times the weight to Sheep Mountain compared to Mayberry Slough in the first PC.[7] Ironically, the Sheep Mountain CA series drops sharply in the 1980s (not shown), after the calibration period.

FIGURE 1. Two NOAMER site chronologies used in MBH98. Though both series are full-length, the first series is given 390 times the weight of the second series. The

transformation of data prior to calculation in the MBH98 software causes series whose mean shifts at the end of the series to be over-weighted.

As another illustration of the effect of the MBH98 algorithm, we generated 70 series of stationary red noise and compared the PC1 from the MBH98 algorithm to that of a standard software package. To generate the data we took the 70 NOAMER sites available back to 1400 and fitted a lag 1 autoregression model to each. The coefficients ($\beta^1,...,\beta^{70}$) were all of magnitude less than 1. Then we generated 70 random vectors $a_t^1,...,a_t^{70}$ of length 1081, using the AR1 series $a_t^i = \beta^i a_{t-1}^i + e_t^i$, i=1,…,70. Each series was initialized at zero and the e_t^i vectors were N(0,1). The first 500 values were then dropped from each series, yielding 70 vectors of stationary red noise, each of length 581. The (conventional) first principal component from these 70 series, after smoothing, exhibits the expected stationary sawtooth pattern (Figure 2, dashed line). The MBH98 method, in which the 70 series were first shifted to the mean over rows 503-581, yields the distorted PC shown in Figure 2 (solid line). The reason for the hockey stick-shape is that some of the underlying vectors randomly trail up or down at the end of their length, and these are selected for high weighting by the MBH98 method. Ten simulations were carried out and a hockey stick shape was observed in every simulation.

FIGURE 2. First principal component from 70 stationary red noise series, smoothed with Lowess (f=0.1). Dashed line: Conventional method (princomp command in the programming package R). Solid line: MBH98 algorithm, which standardizes on a short segment (1902-1980) prior to applying singular value decomposition.

Figure 3 shows the NOAMER PC1 for the AD1400 network. The top panel is according to MBH98, the second panel is the simple mean of the proxies and the third panel (MM04) is the output of a conventional principal components algorithm[8]. The second and third panels closely resemble one another (correlation 0.92). The hockey stick shape of the top panel is not reflected in either the mean or the conventional PC1.

FIGURE 3. Top panel: PC1 of the post-1400 NOAMER tree ring network, computed by MBH98 using short-segment standardization. Second panel: simple mean of proxies. Third panel: PC1 as computed by standard software without short-segment standardization. Bottom panel: Unreported PC1 computed by MBH after censoring Graybill-Idso high-altitude series.

The fourth panel is an unreported alternate version of the NOAMER PC1 from the MBH98 data archive[9]. It too bears a striking resemblance to the MM04 PC1 (correlation 0.95), but it was calculated by MBH themselves using the same method as in the top panel after excluding 20 high-altitude series. 15 of these series are on the list of the top-weighted 16 sites and account for essentially all the behaviour of the NOAMER PC1 in the AD1400 network as calculated by MBH98. Yet these high-altitude sites are known to be unsuitable temperature proxies. 19 of the 20 are from a single researcher, Donald Graybill, who attempted to detect CO_2 fertilization, by selecting high-altitude sites with cambial dieback ("strip bark") formation.[10] 12 of them are specifically discussed in Graybill and Idso (1993)[11], which stated that their 20th century ring growth rates cannot be attributed to local or regional temperature. LaMarche et al[12] earlier stated that temperature does not account for 20th century tree ring growth rates at Sheep Mountain or Campito Mountain, both high-weighted sites. Mann et al. (1999) themselves[13] pointed out, with reference to these proxies: "A number of the highest elevation chronologies in the western U.S. do appear, however, to have exhibited long-term growth increases that are more dramatic than can be explained by instrumental temperature trends in these regions."

The principal components algorithm of MBH98 selects these 20 series for dominant weighting, and effectively nullifies the contributions of the other 50 sites. The NOAMER PC1 thus gets its hockey stick shape from the Graybill-Idso sites, which exhibit a nonclimatic response and/or a nonlinear response to 20th century temperature. Since MBH98 states (p. 780) that their method requires the assumption that proxies exhibit a linear response to temperature, the Graybill-Idso sites, explicitly acknowledged as problematic in Mann et al (1999) (ref. [13]), should have been disqualified as contributors to the NOAMER PC1 in MBH98, let alone as the main determinants of its shape.

With respect to the other "key indicators", the Twisted Tree series as used by MBH98 does not begin until 1459 so it is irrelevant to the pre-1450 interval in any case. However, there is no justification for not using the updated version which we used. MBH98 cites Jacoby and d'Arrigo (1989)[14] as a source for 11 northern treeline sites. Jacoby and d'Arrigo obtained additional samples from several sites in or about 1992 and carried out further quality control, which is reflected in the site chronologies archived at the WDCP in the 1990s. MBH98 used older 1989 versions, some of which differ from the final archived versions. The version of the Twisted Tree series used in MBH98 was the older, unarchived edition which ended in 1976 and included a pre-16th century portion based on a single tree. The archived version, which we used, ends in 1992 and only commences in 1529, when 3 trees become available.

The presence or absence of the third "key" indicator, the SWM network, has virtually no impact on early 15th century results. Nevertheless, our exclusion of the 15th century portion of this series is justified on quality control grounds. MBH98 attributed the network to Stahle and Cleaveland (1993) (see ref. [4]) but have recently stated (ref. [19]) that it derives from Stahle et al. (1998). For each site in the SWM network, MBH98 use two series: earlywood and latewood widths, although Stahle et al. (1998) do not use latewood widths. Of the SWM sites listed in the original MBH98 Supplementary Information (SI), only 2 are available before 1450. In the FTP data[15] actually used in MBH98, there are 3 sites (6 series) extending back prior to 1450. Two of these sites (4 series) have identical values for the first 120 years for earlywood widths and the first 125 years for latewood widths, each differing thereafter, suggesting that the data are either spliced versions of different sites or different editions of the same site. Either way at least one of the two sites is clearly ineligible pre-1450, leaving only two potentially eligible sites – an insufficient number to constitute a network in usual MBH98 usage. One of them is Spruce Canyon CO, which is already used in the NOAMER roster and should therefore be dropped from the SWM group. The data for the remaining SWM site, Cerro Barajas, as used in MBH98, includes physically impossible negative values in the early portion of the series, which are not present in the version archived at WDCP. The unsuitability of the SWM PC1 in the AD1400-1450 period is clearly established and we note that Stahle et al. themselves do not apply their network prior to 1706.

Another series used (but not collected) by Jacoby and d'Arrigo, the St. Anne River (Gaspé, Canada)[16] ring width chronology, matters to the reconstruction. This series was used twice by MBH98, as an individual proxy and in the NOAMER principal components roster (as series cana036). The underlying data set commences in 1404 (MBH98, without disclosure, extrapolated the first four years), and is based on only one tree up to 1421 and 2 trees up to 1447. The early portion of the series therefore fails the quality control criteria claimed by MBH98 themselves as well as standard minimum signal criteria[17], and was not used by Jacoby and d'Arrigo, whose analysis began effective 1600. Its 20th century shape also exhibits nonclimatic factors or a nonlinear temperature response. The seemingly innocuous pre-1404 extrapolation permitted inclusion of this series by MBH98 in their AD1400 roster, where it has a material impact. Removing the undisclosed extrapolation, and noting the small number of trees up to 1447, justifies our not commencing it until the AD1450 roster. The updated version of the Jacoby and d'Arrigo treeline site at Sheenjek River, Alaska, introduced a pre-1450 segment not used in MBH98, so using it leaves the number of proxies in the AD1400 roster unchanged.

To demonstrate the materiality of the above analysis, we have reproduced the MBH98 methodology as closely as we could[18]. Figure 4 shows the original MBH98 Northern Hemisphere temperature index (top panel) and our emulation using MBH98 data, including their flawed PC series (second panel). The overall replication is reasonably good, with correlation between the two of 0.89 and the unique 20th century warming evident in both. Remaining differences may be due to slight differences in the sequences of rosters used for PC calculation and proxy calibration (which cannot be unambiguously reconstructed from available information), or to slight scaling differences due to the unavailability of the particular edition of the instrumental temperature data used in MBH98. The third panel applies the following three changes (sequential versions are available in the SI): (1) removing the unreliable pre-1500 portion segment of the SWM PC1– a step with virtually no impact; (2) using a conventional principal components algorithm on the sites listed in the MBH98 SI, calculating the NOAMER PC1 and PC2 back to 1400; and (3) commencing the Gaspé series in the AD1450 roster. The bottom panel includes changes (1-3) and: (4) uses the updated Sheenjek River series; (5) other data updates and adjustments as discussed in references [1] and [19], of which only the update to the Tasmania temperature reconstruction is material to the AD1400 portion.[20]

Figure 4. Top panel is Northern Hemisphere temperature index from MBH98. Second panel is authors' emulation using publicly disclosed data and methods (correlation=0.89). Third panel is authors' reconstruction without short-segment standardization and without the extrapolation of the Gaspé series. Bottom panel applies other data updates and corrections. Methodology for bottom three panels is identical. Note particularly the change in relative height of the 15th century portion of the graph.

The "Reduction of Error" (RE) statistics between the most recent CRU Northern Hemisphere temperature averages and the four panels in Figure 4 over the 1856-1901 "verification" interval are, respectively, 0.59 (as compared to 0.69 reported in MBH98 using older data), 0.43, 0.43 and 0.42. This uses the full list of proxies, which is available only back to 1820. Using only the proxies available back to AD1400, our emulation of MBH98 yields an RE of 0.30. Using a conventional principal components algorithm, the RE falls to -0.03 and excluding the early portion of the Gaspé series, the RE drops to –1.03, indicating no reconstructive skill over this interval. The apparent reconstructive skill in MBH98 is due to retaining the Graybill-Idso series—which the originating authors warned were unreliable for temperature reconstruction—and a single tree in Gaspé. The R^2 statistic for the MBH98 series over the verification interval is 0.12, but drops to 0.016 with corrections (1-5).

Remaining methodological differences between MBH98 and us are immaterial to our conclusions, since our emulation of MBH98 shows uniquely high 20th century values of the MBH98 Northern Hemisphere temperature index, but the same reconstruction using consistent data quality criteria and conventional principal component analysis does not. We do not assert that the early 15th century was "warm", as we do not endorse MBH98 methodology for other reasons not discussed here. Thus, any inconsistency of this graph with other paleoclimate studies is irrelevant. We do assert, based on the above considerations, that the distinctive hockey-stick shape of the MBH98 temperature reconstruction is primarily due to the Graybill cambial dieback and similar tree ring sites that exhibit non-linear or non temperature-related 20th century growth and to a questionable step in their principal components algorithm that overweights these series. Without these problematic series and without their biased principal component analysis, Mann et al. are not entitled to conclude that the 20th century was uniquely warm based on their data and methods.

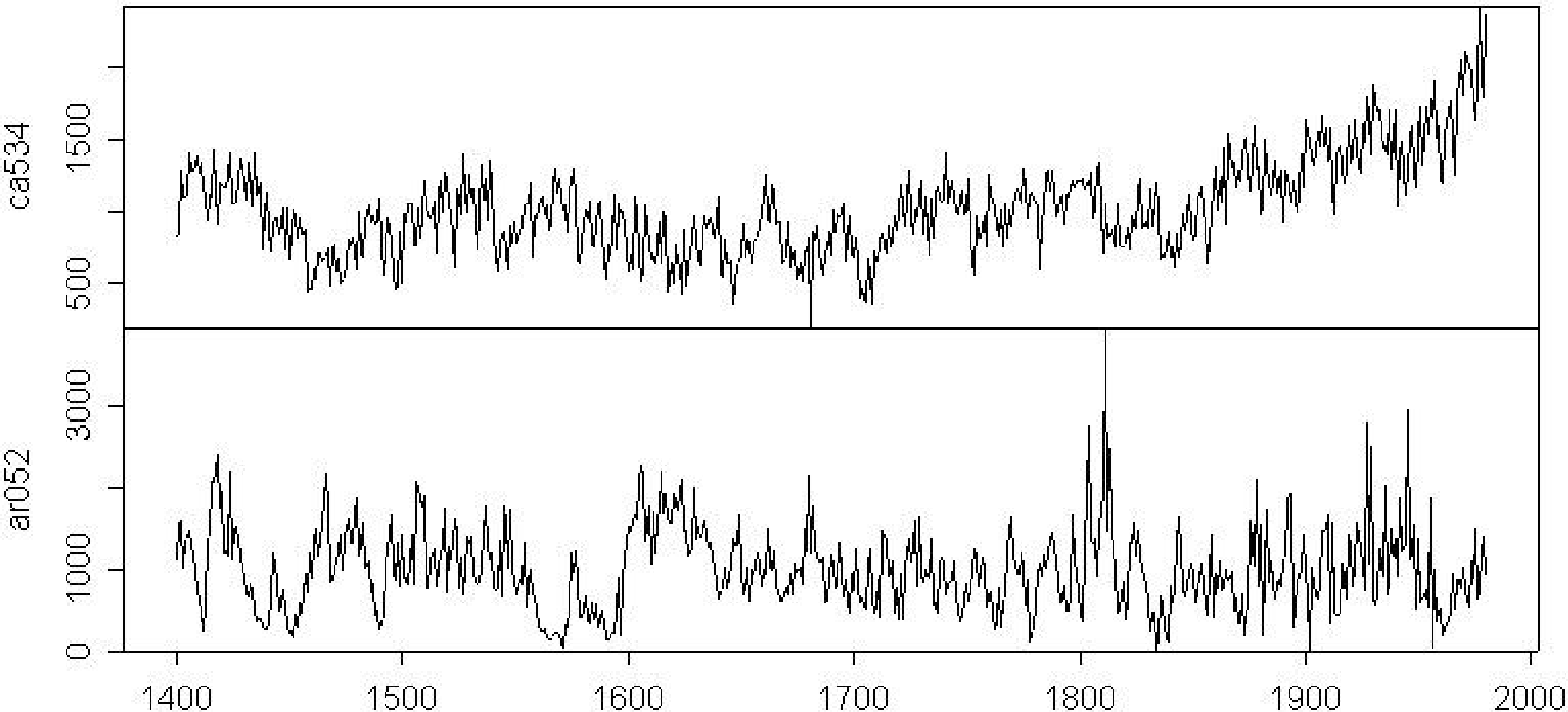
Supplementary Information accompanies the paper on *Nature*'s website (http://www.nature.com).

Financial support for this research was neither sought nor received. The authors declare they have no competing financial interests.

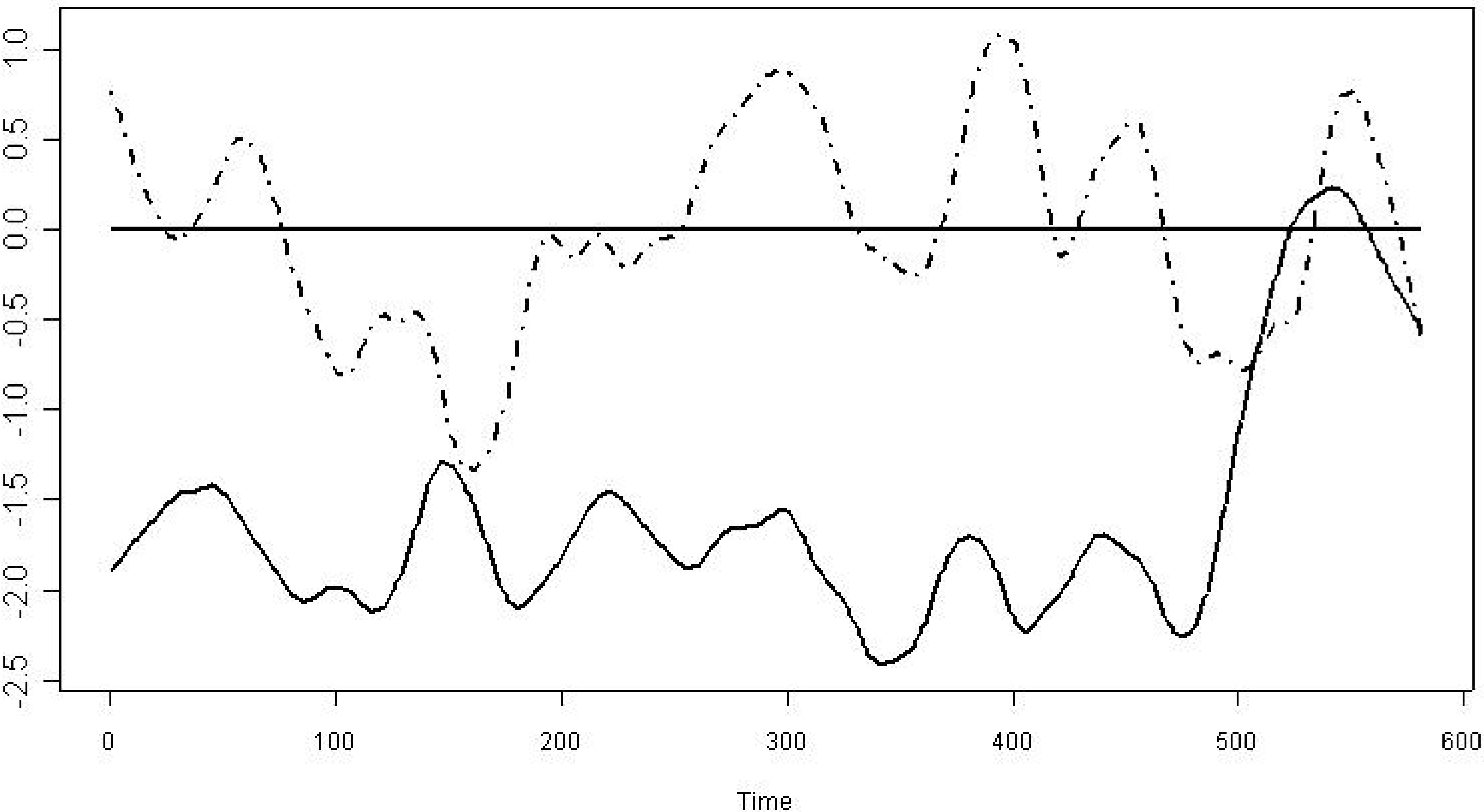
Correspondence and requests for materials should be addressed to S.M. (e-mail: stephen.mcintyre@utoronto.ca).

References

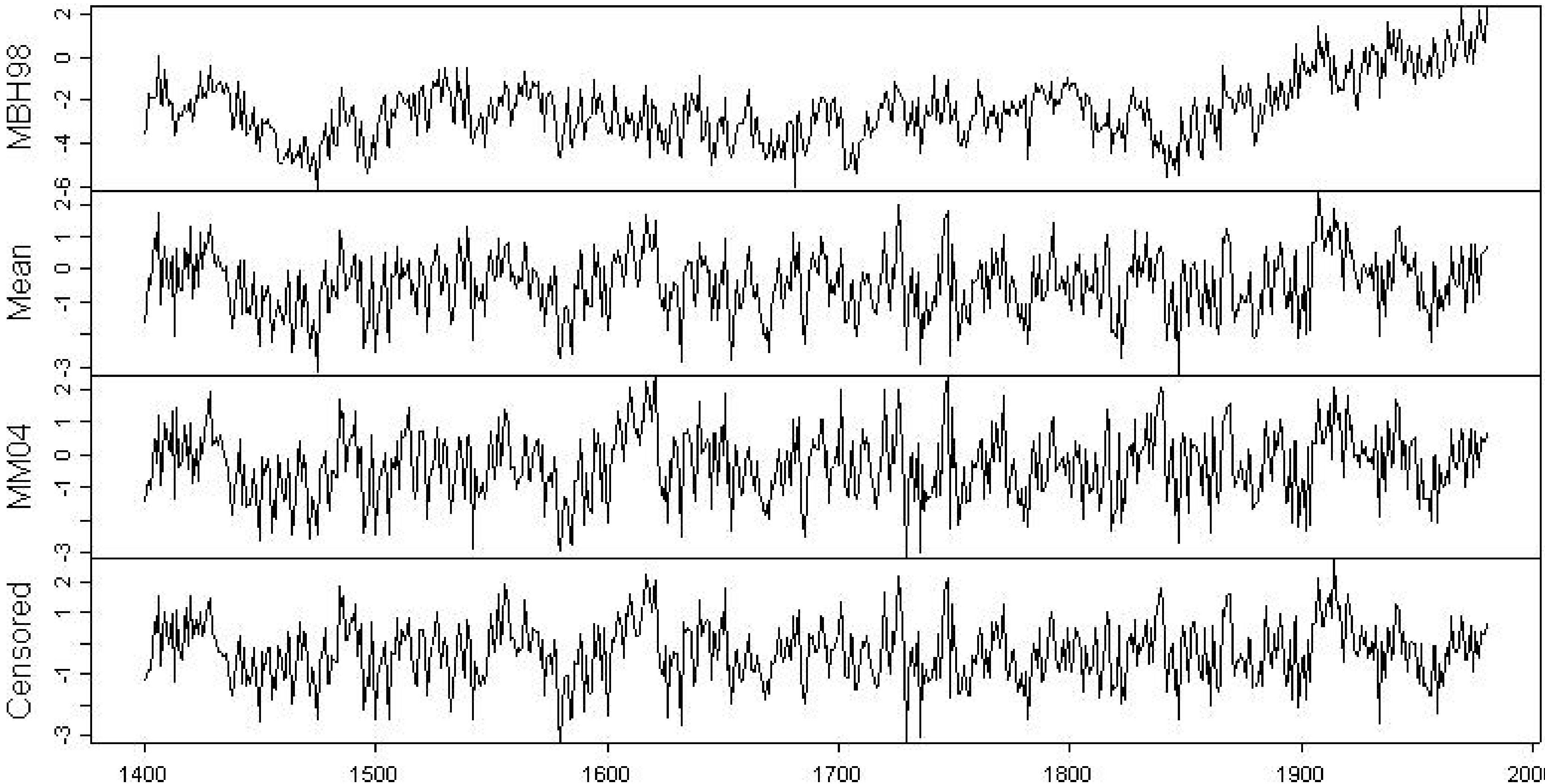
1. McIntyre, S. & McKitrick, R. (2003). *Environment and Energy* 14(6), 751-771.
2. Mann, M.E., Bradley, R.S. & Hughes, M.K. (1998). *Nature*, 392, 779-787.
3. Mann, M.E., Bradley, R.S. & Hughes, M.K. (2003). "Response to M&M" http://www.cru.uea.ac.uk/~timo/paleo/.
4. Stahle, D.W. et al. (1998). *Bulletin of the American Meteorological Society,* 79(10), 2137-2152; Stahle, D.W. and Cleaveland, M.K. (1993). *J. Climate*, 6, 129-139.
5. "Certain densely sampled regional dendroclimatic data sets have been represented in the network by a smaller number of leading principal components (typically 3–11 depending on the spatial extent and size of the data set)." Ref. [2] p. 779.
6. Professor Mann's data archive is at ftp://holocene.evsc.virginia.edu/pub/MBH98/ ("FTP"). The software is at FTP/TREE/ITRDB/NOAMER/pca-noamer.
7. The MBH98 weights are located at FTP/TREE/ITRDB/NOAMER/BACKTO_1400/eof01.out, with the sited identifications at FTP/TREE/ITRDB/NOAMER/BACKTO_1400/ noamer-itrdb-ad1400.txt.
8. This uses the command "princomp" in the statistical programming package R.
9. See FTP/TREE/ITRDB/NOAMER/BACKTO_1400-CENSORED.
10. "Another tree selection factor that is crucial to our findings involves tree form. Experience has indicated that many of the oldest five-needled pines have experienced cambial dieback to varying degrees. This appears to begin after several hundred years of growth and is progressive. These so-called 'strip-bark' trees can have active cambium that is only a few centimetres in width. Foliage and cones are accordingly limited. Trees of this nature were the primary focus of investigation whenever possible. A separate set of full-bark trees growing among strip-bark forms was also recently sampled in the Sheep Mountain stand for comparative purposes." Source: ref. [11].
11. Graybill, D.A., and S.B. Idso, (1993). *Global Biogeochemical Cycles*, 7, 81-95.
12. LaMarche, V. C., Jr., Graybill, D. A., Fritts, H. C., and Rose, M. R. (1984). *Science* 225, 1019–1021.
13. Mann, M.E., Bradley, R.S. and Hughes, M.K., (1999). *Geophysical Research Letters* 26, 759-762.
14. Jacoby, G.C. and D'Arrigo, R.D. (1989). *Climatic Change* 14, 39-59.
15. See ftp://holocene.evsc.virginia.edu/pub/MBH98/.
16. While the site (cana036 at WDCP) is described as a "northern treeline" site, Gaspé, Quebec is nowhere near the northern treeline and has extensive forestry operations. Its inclusion as a northern treeline proxy is puzzling.
17. Wigley, T.M.L., Briffa, K.R. and Jones, P.D. (1984). *Journal of Climate and Applied Meteorology* 23, 201-213.
18. We requested copies of the computer programs used in MBH98 in order to fully reconcile our results to MBH98, but MBH have refused to disclose these programs.
19. Mann, Bradley and Hughes, Corrigendum, *Nature* forthcoming.
20. Cook E.R., Buckley B.M., D'Arrigo R.D., and Peterson, M.J. (2000). *Climate Dynamics* 16, 79-91.

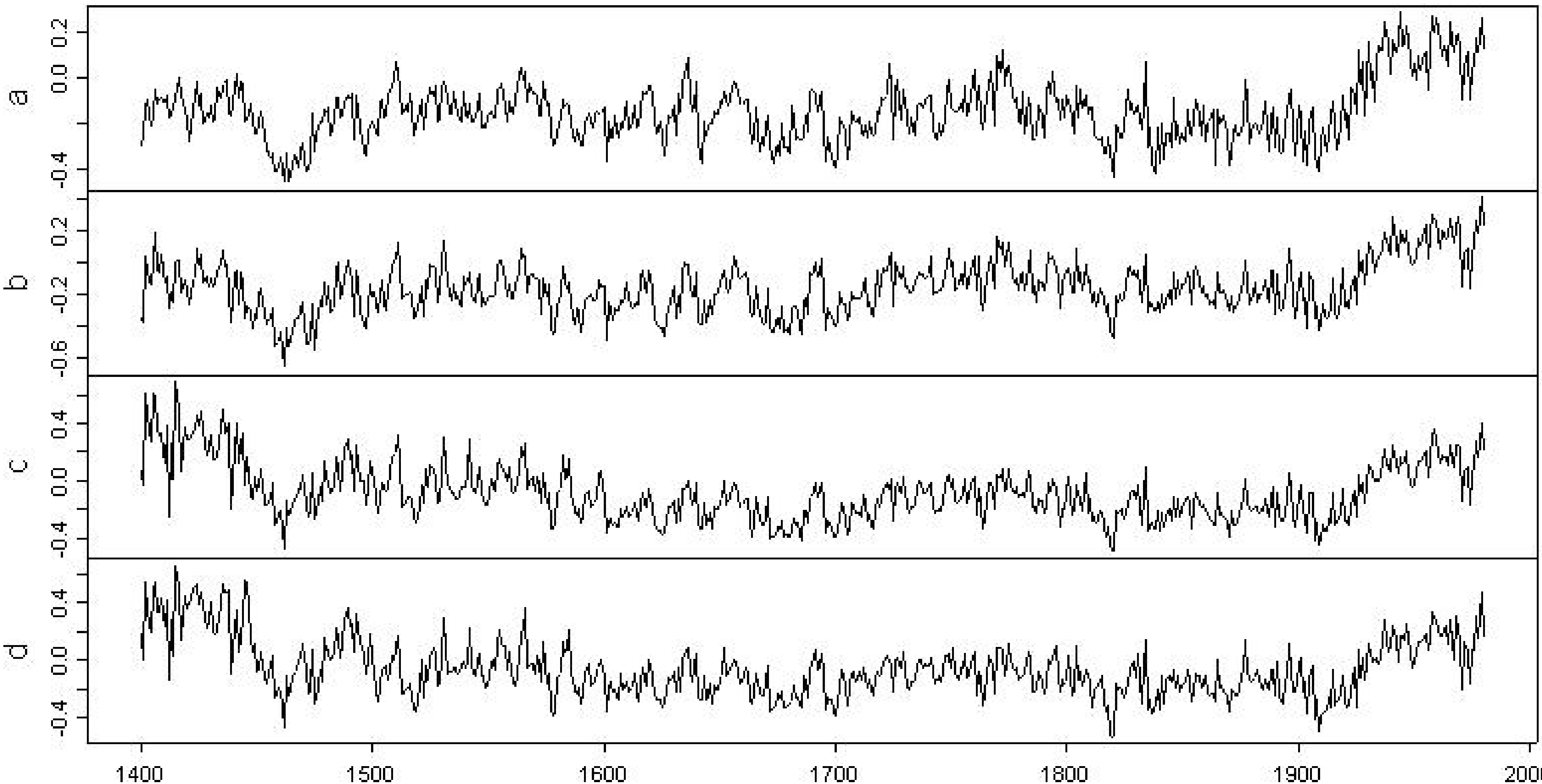


ca534
500
1500
ar052
0
1000
3000
1400
1500
1600
1700
1800
1900
2000



1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
0
100
200
300
400
500
600
Time


MBH98
Mean
MM04
Censored
2
0
-2
-4
-6
2
1
0
-1
-3
1400
1500
1600
1700
1800
1900
2000



a
0.2
0.0
-0.4
b
0.2
-0.2
-0.6
c
0.4
0.0
-0.4
d
0.4
0.0
-0.4
1400
1500
1600
1700
1800
1900
2000