REPLY TO "Global-scale temperature patterns and climate forcings over the past six centuries: A comment." By S. McIntyre and R. McKitrick

Michael E. Mann[1], Raymond S. Bradley[2], Malcolm K. Hughes[3]

[1]*Department of Environmental Sciences, University of Virginia, Clark Hall, Charlottesville, Virginia, 22903, USA*

[2]*Geosciences Department, University of Massachusetts, Amherst, MA 01003, USA*

[3]*Laboratory of Tree Ring Research, University of Virginia, Tucson, AZ 85721, USA*

McIntyre and McKitrick (1) (henceforth MM04) are incorrect with respect to each major point they raise, including those concerning the main data sets we used and the principal components algorithm with which we condensed proxy data in heavily represented regions. Contrary to the assertions of MM04, we here demonstrate our original reconstruction (2) is statistically robust and their criticisms of our reconstruction are without merit.

All paleoclimate reconstructions depend on the data sets used in the reconstruction. Not surprisingly, MM04 find that changing the data set leads to different results. MM04 fail to note, however, a vital step in the process of developing such reconstructions, namely validating the statistical model used to convert proxy records to estimates of past climate. This is usually done by developing estimates of the climate variable of interest, or "reconstructions", for a period covered by instrumental data not included in the training, or calibration, period of the model. Unlike our original reconstructions, and several variants described below, MM04 do not report such tests. We will show that censoring our data set as they did produces a reconstruction similar to theirs that does not pass validation tests and, unlike ours, bears no resemblance to the

instrumental record during the 19th century. We demonstrate in detail elsewhere (Mann et al, in review) that MM04's "updated" version of the data set used by us in (2) (henceforth MBH98) repeats mistakes they have made elsewhere (ref. 3, henceforth MM03) by using a censored set resulting from failure to understand the procedure used by MBH98. Their misunderstanding of the MBH98 protocol, combined with some additional unjustified substitutions of data used by MBH98 (see below), led to the elimination by MM03 of the entire North American ITRDB and Northern Treeline datasets used by MBH98 over AD 1400-1500, amounting to more than 80% of the proxy data (77/95 series) used by MBH98 prior to AD 1450 and more than 70% of the proxy data (99/139 series) between AD 1450 and 1500. MM04 effectively censor the MBH98 dataset in an identical manner to MM03, though in the case of the ITRDB data, this time through the elimination of the dominant component of variance in the dataset. It is no coincidence that it is over precisely this AD 1400-1500 interval that they obtain a dramatically different result (anomalous 15th century warmth) that conflicts not only with MBH98, but virtually all other published NH reconstructions (9). Here we demonstrate that MM04's main criticisms are without merit.

1. What is the cause of the differences between the reconstruction of MBH98 and of MM03 and MM04?

In order to simplify this reply, we make use of only the 1st eigenvector of the surface temperature field for calibration of proxy data and the associated surface temperature reconstructions. We use the proxy indicator network available back to AD 1400 for reconstructions over the AD 1400-1500 interval. The additional (2nd) eigenvector, and two additional indicators used by MBH98 from AD 1450-1500 have no significant impact on the NH mean reconstruction of interest here. To simulate what MM04 did (see also ref. 3), we eliminated

the entire North American ITRDB and Northern Treeline datasets from the MBH98 network available from AD 1400-1500 and performed the reconstruction (henceforth referred to as 'MM04c') based on the resulting network through 1971 (line 1, Figure 1a). The splice of this reconstruction from AD 1400-1500 with that of MBH98 from AD 1501-1971 provides a close approximation (henceforth referred to as 'MM04cS') to the MM03/MM04 ('MM') reconstruction (line 2, Figure 1a), and demonstrates why their attempt to reproduce our estimate fails. Slight differences of MM04cS with the precise MM result likely result from minor errors in their implementation of the methodology (Mann et al, in review). The difference between MBH98 and MM04cS is not the result of our use of infilled values for a few locations in the period 1971-80 or (for one series) for 1400-1403, as they claim (3). We show that a reconstruction calibrated using the period 1902-1971 (avoiding infilling) is effectively the same as that of MBH98 (line 3, Figure 1a). Elimination of the Stahle et al SWM/TX data challenged by MM04 (see below) also has no detectable influence on the reconstruction (line 4, Figure 1a). Thus the difference between the reconstructions of MBH98 and MM results from MM's omission or effective omission of the North American ITRDB and Northern Treeline data sets (the overwhelming majority of proxy data used by MBH98 in the early centuries of the reconstruction period).

2. What basis is there for believing the MBH98 reconstruction is reliable?

There is a long history in high-resolution paleoclimatology of emphasizing verification of models used in climate reconstructions (4,5), by using independent data. Unlike MBH98, MM provide no assessments of the statistical reliability of the model on which their reconstruction is based. In fact, the MM04c model (line 1, Figure 1b) for the period AD 1400-1500 when applied to the interval 1854-1901 displays no statistical skill in comparisons with the instrumental data for 1854-1901 (RE = -0.76), whereas our model for the AD 1400-1500 period, applied to the

same interval (line 2, Figure 1b) exhibits considerable skill (RE = 0.51). The actual MBH98 reconstruction (i.e., line 5, Figure 1a), which employs models from an increasingly expanded proxy network in later centuries (with the full network of 112 indicators available after 1820) shows even greater verification skill (*RE* = 0.69; see MBH98). The 'reduction of error' statistic '*RE*' favored in most studies (4,5) as a diagnostic measure of reconstructive skill obeys the bounds $-\infty < RE < 1$. Negative values of *RE*, as in the case of MM04c, indicate a lack of reconstructive skill, while *RE*=0 represents the "no skill" boundary in which the reconstruction performs as well as the assignment of the calibration period mean, and substantially positive values as in MBH98 (e.g. *RE*=0.5 to 0.7 depending on the precise proxy model) suggest considerable skill. Thus, the MM04 reconstruction, unlike the MBH98 reconstruction, is unlikely to provide any useful insight into the actual climate history prior to AD 1500.

3. Is a bias introduced into our reconstruction by the principal components algorithm we used to condense data?

To examine this claim we performed a reconstruction over AD 1400-1980 using all 95 original individual proxy series available over the AD 1400-1500 interval (rather than a smaller network of 22 indicators that includes PC representations of certain sub-networks), imposing uniform weights on all series (line 1, Figure 1c). The resulting reconstruction closely resembles the MBH98 reconstruction (line 2, Figure 1c), with no indication of anomalous 15th century warmth. The model on which it is based has only a slightly lower verification score (*RE*=0.47) than the analagous MBH98 model for 22 indicators available over AD 1400-1500 (*RE*=0.51). We also show that our re-standardization of all indicators in the MBH98 network by their *detrended* standard deviation during the calibration period, prior to calibration and reconstruction did not

significantly influence the MBH98 reconstruction (line 3, Figure 1c). This latter step was motivated by the fact that 20th century trends in instrumental and proxy data typically far exceed the expectations for a 'red noise' null hypothesis (6). Normalizing by the detrended standard deviation therefore more properly weights the data series with respect to their estimated noise variance. It is inappropriate to use a red noise model in testing standardization procedures, as MM04 did, because simple spectral analyses of the actual series reveals many of them to be statistically inconsistent with an underlying red noise model. MBH98 also normalized the time series of the reconstructions of the instrumental eigenvectors to have the same variance as the actual instrumental eigenvectors during the calibration period. As MM03 did not perform this step, we have not invoked this normalization in our reproduction of their estimate (line 1, Figure 1a). We nonetheless show that this step, too, makes little difference in the MBH98 reconstruction (line 4, Figure 1c). We have thus demonstrated that MM04's claim of a bias introduced by our PCA algorithm is invalid on both theoretical and empirical grounds. The only explanation for the anomalous findings of MM03 and MM04 is their deletion of essential proxy indicators from the network, leading to a reconstruction suggesting anomalous 15th century warmth, which does not pass statistical verification and is therefore almost certainly spurious.

Had MM produced an alternative reconstruction that stood up to rigorous verification procedures (like that of MBH98), it would be worthy of discussion. However, as a close reproduction of their reconstruction (MM04c) shows roughly the level of skill of a random reconstruction, it does not demand serious consideration. Furthermore, numerous other studies based on different data and methods (7-10), or climate modeling approaches (e.g. 11-13), lead to the same conclusion that late 20th century warmth is anomalous in the context of the past six centuries or longer. The criticisms put forth by MM04 do not stand up to scrutiny and in no way

invalidate our reconstruction, which we have demonstrated to be quite robust. Of course, we continue to seek both more and improved data, and improved techniques of reconstruction, especially for the early centuries of the second millennium AD.

Appendix: specific data details

One of the few legitimate criticisms of MBH98 by MM03 and MM04 is their identification of some errors in the details provided in the on-line "supplementary information" that accompanied the MBH98 article. We have worked closely with *Nature* to produce an archive containing (a) a corrected version of the supplementary information, (b) all data (proxy and instrumental) used by MBH98 in an easily accessible and completely documented format, and some additional methodological details useful for those seeking to reproduce our results. This archive is available publicly here:

ftp://holocene.evsc.virginia.edu/pub/Data/MANNETAL98.

Minor errors that existed in the original supplementary description provided by MBH98 are detailed elsewhere (14), but we emphasize that these are of virtually no consequence to the MBH98 hemispheric mean temperature reconstruction. For example, while it may not have been clear in the original documentation that we made use of the versions of the Central England and Central European temperatures records of Bradley and Jones' (7) (summer mean, rather than annual mean), and set two regional western North American tree-ring series equal following their termination during the mid 20th century, this has essentially no impact on our hemispheric mean reconstruction. Furthermore, as these issues (and several other series challenged by MM) involve data series that do not extend to before AD 1500, they are irrelevant to the questions at hand, and raising them represents little more than a curious distraction by MM. Related complaints by MM regarding the use by MBH98 of some limited redundant data, and challenges of particular data, such as the early Stahle et al SWM/TX data, are also irrelevant (see e.g. line 4, Figure 1a).

Moreover, MM inappropriately dismiss the quality of certain proxy data series based on their own inability to find those series in certain public data archives or in the published literature. Some of the proxy data used by MBH98 are unpublished, and were kindly provided directly to us by the original researchers. It is invalid for MM04 to impute lower quality to such data. All proxy data used by MBH98 have been provided in the public website listed earlier.

MM04 criticize the PC representation of the North American ITRDB data which was based, for the period AD 1400-1450, on an EOF analysis of the 70 constituent series which were standardized, as discussed earlier, by their detrended calibration period standard deviation. As shown earlier (line 3, Figure 1c), this criticism can be simply dismissed on the basis that use of all 70 constituent series (which are part of the total of 95 available for this period) with uniform weighting, rather than this PC representation, yields no evidence of the anomalous 15th century warmth found by MM04 (line 1, Figure 1c). Furthermore, despite the claims that the standardization procedure used to calculate the ITRDB PC#1 exaggerates the influence of one particular series, we simply note that the loadings on the 70 series associated with PC#1, by contrast, are substantially shared by more than a dozen series (14/70 series have loadings that are greater than 50% of the maximum loading of any one series). Finally, MM04 introduce problems into the important Northern Treeline dataset used by MBH98. Aside from incorrectly substituting shorter versions of the "Kuujuag" and TTHH Northern Treeline series for those used by MBH98, and introducing an extended version of another Northern Treeline series not available prior to AD 1500 at the time of MBH98, they censored from the analysis the only Northern Treeline series in the MBH98 network available over the AD 1400-1500 interval, on the technicality that it begins only in AD 1404 (MBH98 accommodated this detail by setting the values for AD 1400-1404 equal). The criticism of the use of infilled values from AD 1400-1403 can be dealt with by performing the MBH98 reconstruction from AD 1404-1500 rather than AD 1400-1500. Doing so yields essentially the same reconstruction as MBH98 (line 3, Figure 1a).

Correspondence and requests for materials should be addressed to M.E.Mann (e-mail: mann@virginia.edu).

1. McIntyre, S. and R. McKitrick. Global-scale temperatures patterns and climate forcing over the past six centuries: a comment. *Nature* **XXX**, YYY-ZZZ (2004).

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

3. McIntyre, S. & McKitrick, R. Corrections to the Mann et al (1998) Proxy Data Based and Northern Hemispheric Average Temperature Series, *Energy and Environment*, **14**, 751-771 (2003).

4. Fritts, H.C. *Tree Rings and Climate*. London, Academic Press: xii, 567 pp (1976).

5. Cook, E. R., Briffa, K. R. & Jones, P. D. Spatial Regression Methods in Dendroclimatology: A Review and Comparison of Two Techniques. *International Journal of Climatology* 14, 379-402 (1994).

6. Mann, M.E., & Park, J. Oscillatory Spatiotemporal Signal Detection in Climate Studies: A Multiple-Taper Spectral Domain Approach , *Advances in Geophysics*, 41, 1-131 (1999).

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

8. Jones, P.D., T.J. Osborn, & Briffa, K.R. The Evolution of Climate over the Last Millenium *, Science*, **292,** 662-667 (2001).

9. Mann, M.E. *et al*, On Past Temperatures and Anomalous Late 20th Century Warmth, *Eos*, **84**, 256-258 (2003).

10. Jones, P.D. & Mann, M.E. Climate Over Past Millennia, *Reviews of Geophysics*, in press, (2004).

11. Zorita, E., F. Gonzalez-Rouco, & Legutke, S. Testing the Mann et al. (1998) approach to paleoclimate reconstructions in the context of a 1000-yr control simulation with the ECHO-G Coupled Climate Model, *J. Climate, 16,* 1378-1390 (2003).

12. Gerber, S., F. Joos, P. Brügger, T. F. Stocker, M. E. Mann, S. Sitch, and M. Scholze, Constraining temperature variations over the last millennium by comparing simulated and observed atmospheric CO2, *Climate Dynamics*, *20*, 281-299, 2003.

13. Bauer, E., M. Claussen, and V. Brovkin, Assessing climate forcings of the earth system for the past millennium, *Geophys. Res. Lett*., 30 (6), doi: 10.1029/2002GL016639, 2003.

14. Mann, M.E., Bradley, R.S. & Hughes, M.K. Global-scale temperature patterns and climate forcing over the past six centuries, *Climatic Change* (in review).

Figure 1 Alternative NH reconstructions and instrumental NH record.

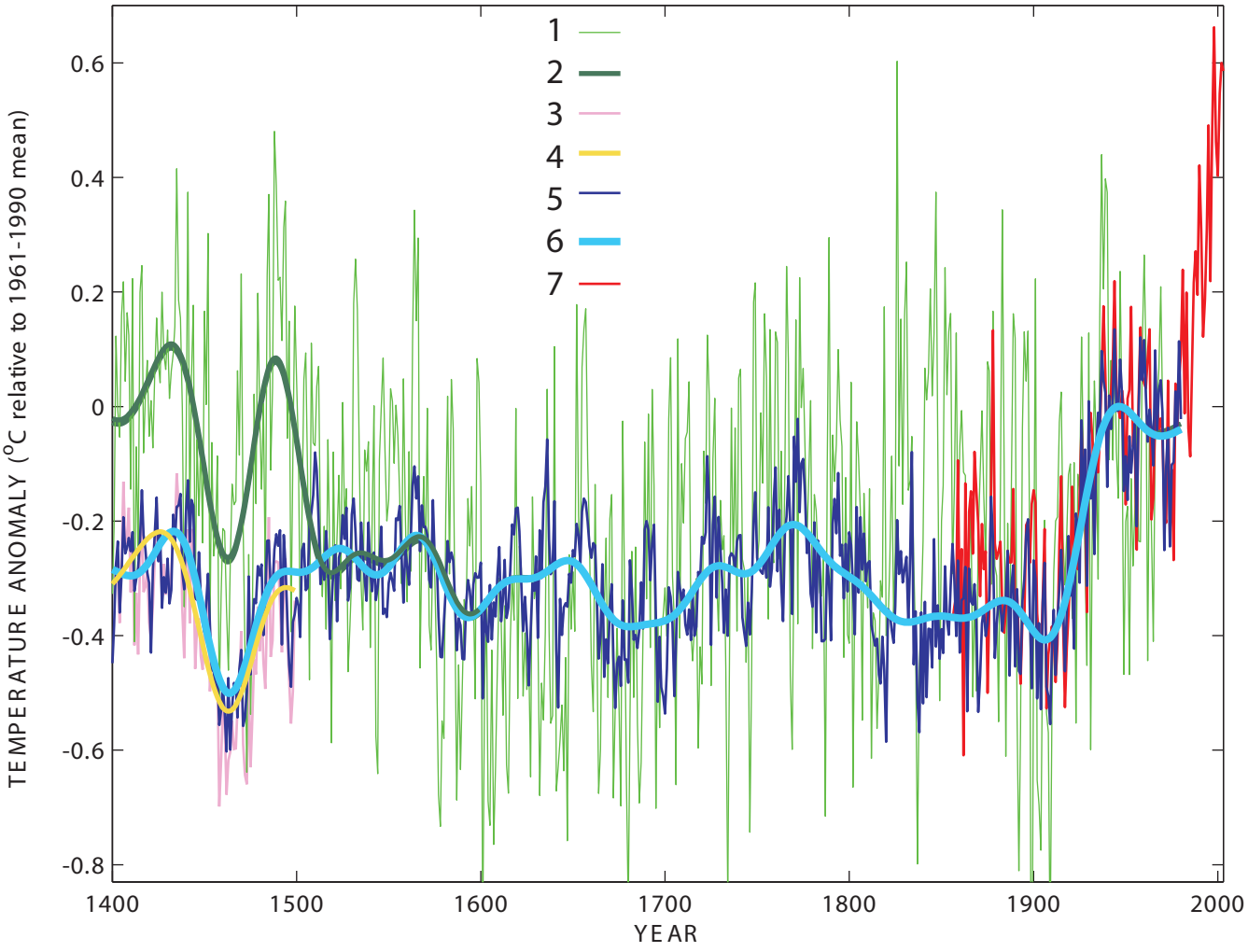
(A) Comparison of MBH98 NH reconstruction with reproduction of MM reconstruction based on elimination of key datasets. Verification results are given for comparison. Series shown include (1) MM04c reconstruction from AD 1400-1971, (2) MM04cS reconstruction from AD 1400-1971, smoothed on a 40 year timescale, (3) Alternative version of MBH98 reconstruction based on calibration period of 1902-1971, shown from AD 1404-1500, (4) Alternative version of MBH98 reconstruction based on elimination of SWM/TX data, shown from AD 1400-1500 (only 40 year smoothed series shown for clarity), (5). The actual MBH98 reconstruction, which makes use increasingly widespread proxy data in later centuries, (6) A 40-year smooth of the MBH98 reconstruction, and (7) The instrumental annual mean NH series 1856-2003. Note that MBH98 and MM04cS coincide by construction (e.g.

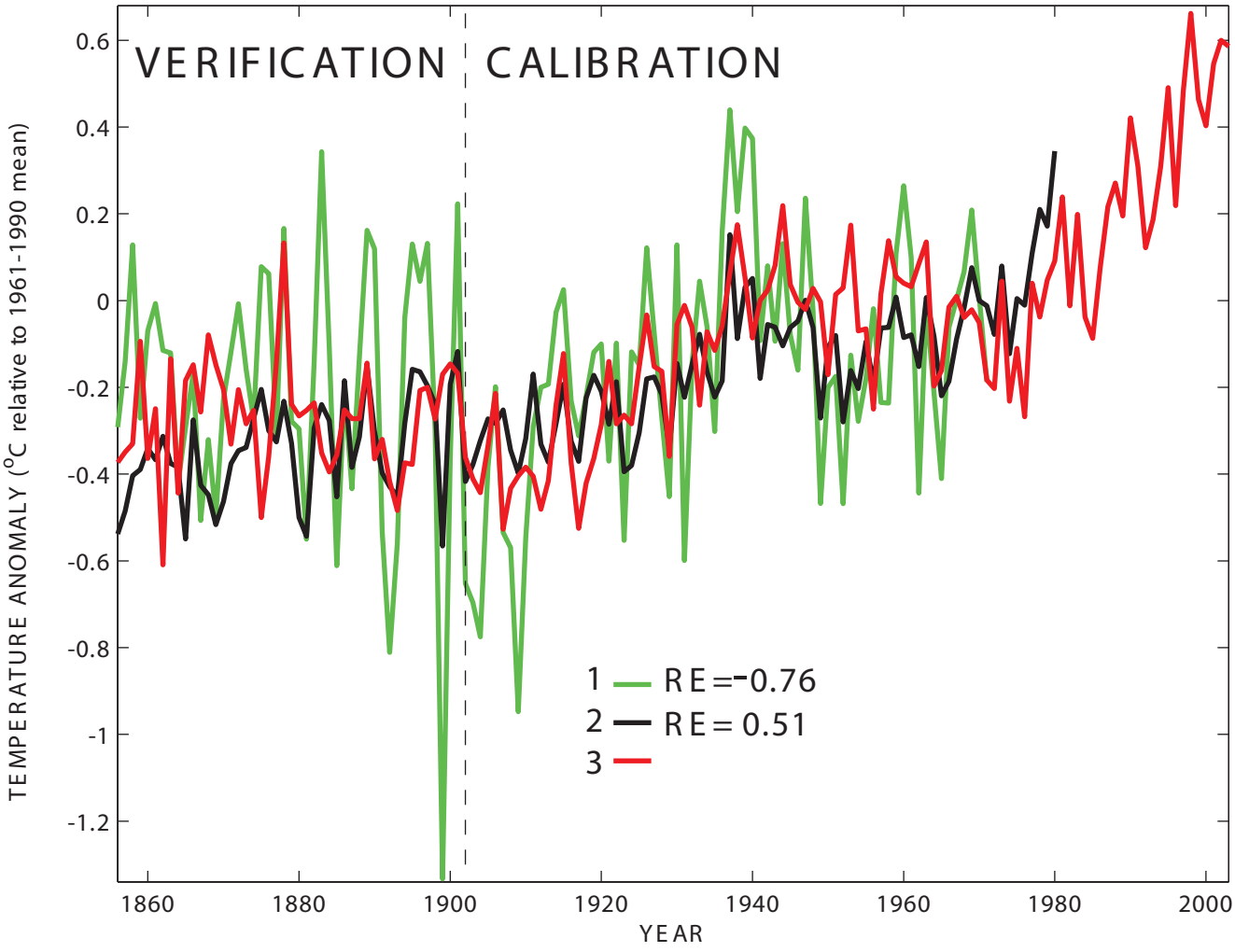
series 6 vs. series 2) after AD 1500, though the 40-year smoothing introduces some small differences prior to approximately AD 1550.

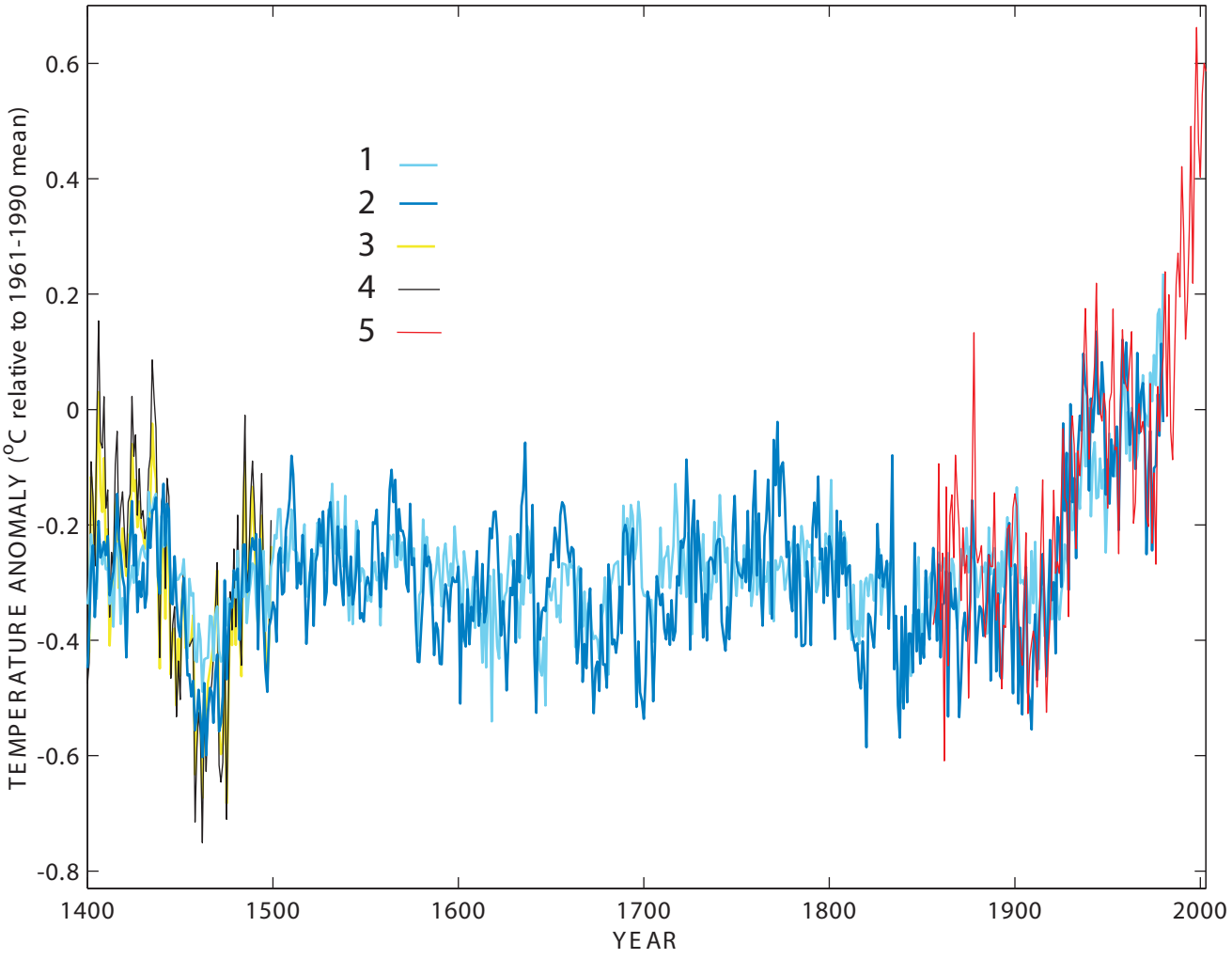
(B) Comparison during calibration and verification intervals of (1) MM04c reconstruction from 1856-1971 and (2) MBH98 reconstruction from 1854-1971 based on AD 1400-1500 proxy model. The instrumental annual mean NH series 1856-2003 (3) is shown for comparison. Vertical dashed line denotes border between calibration and verification intervals. Verification *RE* scores are indicated for (1) and (2). Note that some of the disagreement between reconstruction and instrumental series in the early half of the verification period is due increasing sampling uncertainty in the NH series shown. To avoid this complication, *RE* scores, as in MBH98, are based on a 'frozen grid' NH estimate constructed by an areal hemispheric mean of all 219 Northern Hemisphere gridbox estimates available on a nearly continuous basis throughout the verification period.

(C) Comparisons showing influences of various details of methodology on the MBH98 reconstruction. Series shown include (1) Alternative version of MBH98 reconstruction based on use of all 95 individual proxy series available from AD 1400-1500 with uniform weights, shown from AD 1400-1980, (2) MBH98 reconstruction, Alternative versions of MBH98 reconstruction (shown for AD 1400-1500 period) in which (3) indicators have not been restandardized by detrended calibration period variance and (4) time series of the reconstructed instrumental eigenvectors have not been standardized to have same variance as the corresponding instrumental eigenvectors during the calibration period. The

instrumental annual mean NH series 1856-2003 (5) is shown for comparison. Verification scores are *RE*=0.47 for (1), *RE*=0.42 for (3), and *RE*=0.38 for (4).


1
2
3
4
5
6
7
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
1400
1500
1600
1700
1800
1900
2000
YEAR
TEMPERATURE ANOMALY (°C relative to 1961-1990 mean)


VERIFICATION
CALIBRATION
TEMPERATURE ANOMALY (°C relative to 1961-1990 mean)
YEAR
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
1860
1880
1900
1920
1940
1960
1980
2000
1 RE=-0.76
2 RE= 0.51
3


TEMPERATURE ANOMALY (°C relative to 1961-1990 mean)
YEAR
1
2
3
4
5
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
1400
1500
1600
1700
1800
1900
2000