THE M&M CRITIQUE OF THE MBH98 NORTHERN HEMISPHERE CLIMATE INDEX: UPDATE AND IMPLICATIONS

Stephen McIntyre
512–120 Adelaide St. West, Toronto, Ontario Canada M5H 1T1;
Ross McKitrick
Department of Economics, University of Guelph, Guelph Ontario Canada N1G2W1.

ABSTRACT

The differences between the results of *McIntyre and McKitrick* [2003] and *Mann et al.* [1998] can be reconciled by only two series: the Gaspé cedar ring width series and the first principal component (PC1) from the North American tree ring network. We show that in each case MBH98 methodology differed from what was stated in print and the differences resulted in lower early 15th century index values. In the case of the North American PC1, MBH98 modified the PC algorithm so that the calculation was no longer centered, but claimed that the calculation was "conventional". The modification caused the PC1 to be dominated by a subset of bristlecone pine ring width series which are widely doubted to be reliable temperature proxies. In the case of the Gaspé cedars, MBH98 did not use archived data, but made an extrapolation, unique within the corpus of over 350 series, and misrepresented the start date of the series. The recent Corrigendum by Mann *et al.* denied that these differences between the stated methods and actual methods have any effect, a claim we show is false. We also refute the various arguments by Mann et al. purporting to salvage their reconstruction, including their claims of robustness and statistical skill. Finally, we comment on several policy issues arising from this controversy: the lack of consistent requirements for disclosure of data and methods in paleoclimate journals, and the need to recognize the limitations of journal peer review as a quality control standard when scientific studies are used for public policy.

1. INTRODUCTION

The Northern Hemisphere temperature index of *Mann et al. [1998, "MBH98"],* together with its extension in *Mann et al. [1999],* was adopted by the Intergovernmental Panel on Climate Change *[IPCC, 2001]* as the canonical temperature history of the Northern Hemisphere. It is the authority for claims that the 1990s were the warmest decade of the millennium and its influence on the public's attitude towards climate change and climate change policy has been enormous, and was recently reinforced by its usage in the Arctic Climate Impact Assessment *[ACIA, 2004].*

In *McIntyre and McKitrick [2003*, "MM03"], we attempted to replicate the results of MBH98 and encountered many data and methodological problems, some of which had a significant effect on the central MBH98 conclusions concerning the uniqueness of the late-20th century climate. In response to MM03 and subsequent submissions and correspondence to *Nature*, Mann et al. have provided new information about MBH98, including an extensive archive of data and methods at the Supplementary Information (the "Corrigendum SI") to *Mann et. al., [2003, 2004a, 2004c*, the "Corrigendum"], an extensive archive of data and methods at a University of Virginia FTP site, *[Mann, 2002-2004]* and various written responses to our work *[Mann et al. 2003, 2004a, 200b, 2004d].* Unfortunately Mann et al. have refused to provide the source code used to generate their results, other than the limited (but essential) programs used for tree ring principal components (PCs). They have also refused to provide supporting calculations for the individual calculation steps in MBH98, especially the controversial step from 1400-1450 (the "AD1400 step"). We made unsuccessful appeals to both *Nature* and the U.S. National Science Foundation, which funded MBH98, to compel release of this material.

Because of this obstruction, not all the problems in MBH98 can be resolved. However, we believe that we have sufficient information in hand to:

   (1) completely reconcile the differing results of MM03 and MBH98;
   (2) establish the non-robustness of MBH98;
   (3) reject the temperature reconstruction in MBH98.

The results presented here do not contradict the results of MM03, but are a logical development of the issues first raised therein.

In order to establish our results, we have attempted to emulate all aspects of MBH98. Much of the controversy in the response to our first article on MBH98 pertained to the accuracy of our emulation. We have followed all published information on the MBH98 procedures, and any remaining differences likely cannot be addressed without disclosure of the actual MBH98 code. However, none of the points established herein are affected by the remaining secrecy surrounding MBH98 computational details (for full details of our emulation, including R code, see the Supplementary Information). We anticipate that there will be critical interest in the emulation itself and we will address these matters in a separate paper.

One of the points of view advocated in this article is that individual data series matter in the MBH98 results. We disagree with the view that problems with individual series simply get washed out in a multiproxy study. In the context of the MBH98 methodology this optimistic assumption is untenable.

Section 2 explains the sources of difference between MM03 and MBH98. Section 3 considers the issue of the robustness of MBH98 results. Section 4 discusses particular issues concerning bristlecone pine and cedar proxies, which are central to the matters in this paper. Section 5 deals with some remaining counter-arguments from Mann *et al.* and Section 6 offers concluding comments.

2. RECONCILING MM03 AND MBH98

Differences between MM03 results and MBH98-type results can be reconciled through variations in the handling of only two series, the Gaspé "northern treeline" series and the first principal component (PC1) from the North American proxy roster (NOAMER). The changes are illustrated in Figure 1 below.
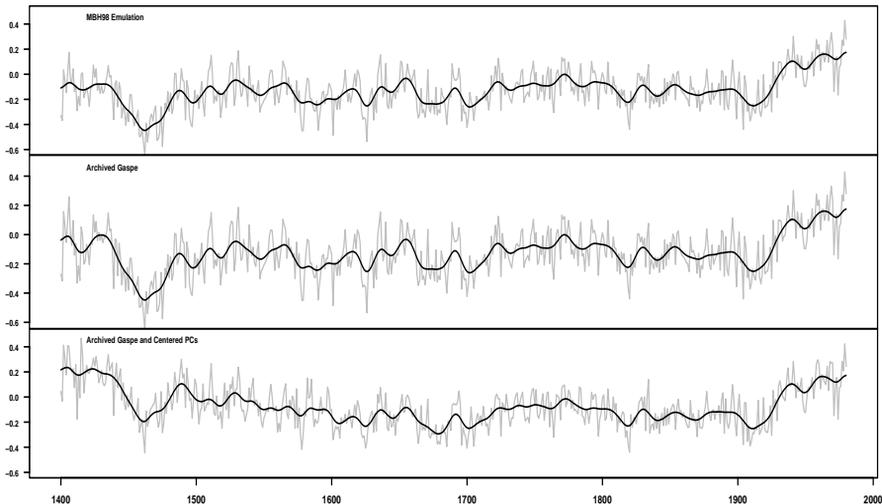

Figure 1. NH Temperature Index. Top panel: MBH98 emulation; middle panel: using
     archived Gaspé version; bottom panel: using centered PC algorithm.

The top panel shows our emulation (version 3) of MBH98, which implements new information from the Corrigendum SI. Relative to the MBH98 reconstruction, it has a Reduction of Error ("RE") statistic of 0.83 in the 1400–1901 period ($R^2 – 0.68$) – both values lower than our version 2 emultion without the Corrigendum SI information.

   The middle panel ("Archived Gaspé") shows the effect of merely using the version of the Gaspé series archived at WDCP, rather than the version as modified by MBH98, accounting for a material change in the early 15th century. The only difference between the two series is the extrapolation of the first four years in MBH98. Under MBH98 methods, a series had to be present at the start of a calculation step in order to be included in the interval roster. In only one case in the entire MBH98 corpus was this rule broken – where the Gaspé series was extrapolated in its early portion, with the convenient result of depressing early 15th century results. This extrapolation was not disclosed in MBH98, although it is now acknowledged in the Corrigendum *[Mann et al., 2004c]*. In MBH98, the start date of this series was misrepresented; we discovered the unique extrapolation only by comparing data as used to archived data. There are other considerations making this unique extrapolation singularly questionable. The Gaspé series is already included in the NOAMER principal components network (as cana036) and thus appears twice in the MBH98 data set, and the extrapolation,

curiously, is only applied to one of the columns. The underlying dataset is based on only one tree up to 1421 and only 2 trees up to 1447. *Jones and Mann [2004]* point to the need for "circumspect use" of tree ring sites with few early examples. The early portion of the series fails standard minimum signal criteria [e.g. *Wigley et al. 1984*] and indeed fails the data quality standards Mann et. al. themselves listed elsewhere. The early portion of the series was not used by the originating authors *[Jacoby and d'Arrigo, 1989; D'Arrigo and Jacoby, 1992]*, whose analysis only begins effective 1601. In fact, *Jones and Mann [2004]* do not use the Gaspé series as an individual proxy and only use the Jacoby-d'Arrigo northern treeline composite when it is adequately replicated after 1601.

The bottom panel ("Archived Gaspé and Centered PCs") shows the additional effect of using conventional (centered) PC methods and is virtually identical to MM03. MBH98 had stated that they used "conventional" principal components (PC) calculations. A conventional PC calculation applies standardization in which the columns are centered by subtracting their mean, which is done for the calculations in the bottom panel. Stepwise PC series are used (extending the NOAMER network back to AD1400), thereby avoiding the main criticism leveled against MM03.

Once again, MBH98 contained a misrepresentation, this time about their PC method. After the University of Virginia FTP site was made publicly available following MM03, by examining PC series archived there and, by examining source code for PC calculations,[1] we were able to determine that MBH98 had not carried out a "conventional" PC calculation, but had modified the PC algorithm, by, among other things, subtracting the 1902-1980 mean, rather than the 1400-1980 column mean, prior to PC calculations, so that the columns were no longer centered on a zero mean in the 1400–1980 step. By this procedure, series are more decentered, and their variance more inflated, the larger is the difference between the series mean and the mean of the 20th century subset. The effect of this transformation would have been mitigated if they had carried out a singular value decomposition on the covariance matrix, but they carried it out on the de-centered data matrix. We have shown elsewhere that this method re-allocates variance so that the PC algorithm then strongly over-weights hockey stick-shaped proxies and that it is so efficient in mining a hockey stick shape that it nearly always produces a hockey-stick shaped PC1 even from persistent red noise [*McIntyre and McKitrick, 2005;* discussed in *Muller, 2004*].

This observation has received a considerable amount of publicity and some observers have misunderstood the point. While we have made scripts available at our FTP site, for greater certainty, we show here the relevant R functions, which were used

---

1 See weights in ftp://holocene.evsc.virginia.edu/pub/MBH98/TREE/ITRDB/NOAMER/pca-noamer.f. This program has 588 lines of Fortran code, of which the first 323 lines are particular to the PC calculations and the final 265 lines is an SVD routine. The 1902–1980 mean is removed in lines 168–173 as follows:

```
168 c   remove 1902–19xx mean from training data
169 c
170     do i=nlow,nhigh
171        iyear = i-nlow+1
172        aprox(iyear.j)=aprox(iyear.j)-roxave(j)
173     end do
```

for these results and which show algorithmically exactly what Mann and we are doing. We emphasize that we are able to exactly replicate Mann's PC calculations with these scripts and that, in unpublished material at *Nature*, Mann has replicated our PC calculations. Although many readers may not be familiar with R *[R Development Core Team, 2004],* we strongly believe that the provision of source code in the actual language is an essential part of ensuring replicable results and that there is sufficient commonality in source codes that the following code will illuminate the issues even for persons who are unfamiliar with the language. We have also chosen to highlight source code in the running text rather than a footnote, because source code issues turn out to be an essential finding. The detrended standard deviation was calculated with the following function:

```
sd.detrend<-function(x) { t<-c(1:length(x)) ; fm<-lm(x~t); sd.detrend<-
    sd(fm$residuals); sd.detrend }
```

The MBH98 transformation was calculated with the following function:

```
mannomatic<-function(x, M=78) {N<-length(x);  xstd<- (x- mean( x[(N-
    M):N]))/sd(x[(N-M):N]);
 sdprox<-sd.detrend(xstd[(N-M):N]); mannomatic<- xstd/sdprox; mannomatic }
```

The main effect of this transformation is through the de-centering, rather than the difference between detrended and undetrended standard deviations, which, in this case, is empirically rather slight. A default value of 78 for M is used to simplify use for the 1902–1980 calibration period, which is M+1 years long.

The North American tree ring network for the AD1400 step was collected into one matrix *Tree,* in this case of dimension 581x70. A matrix *Tree.mannomatic* transformed according to MBH98 was obtained through applying the above function to the matrix as follows:

```
Tree.mannomatic<-apply(Tree,2,mannomatic)
```

By applying the svd function in R, a singular value decomposition corresponding exactly to the archived results (eigenvalues, eigenvectors and PC series) at the University of Virginia FTP site[2] was obtained:

```
PCA.mannomatic.svd<-svd(Tree.mannomatic)
```

We have reported that this algorithm nearly always yields hockey-stick shaped series from persistent red noise networks [*McIntyre and McKitrick, 2005;* also see below]. In response, some readers have expressed incredulity about whether our methods

---

2 Directory <ftp://holocene.evsc.virginia.edu/pub/MBH98/TREE/ITRDB/NOAMER/BACKTO_1400>, PCs at pc01.out, …; eigenvectors at eof01.out,…eigen.out contains percentage variance for each eigenvalue of total variance rather than actual eigenvalues.

accurately reflect actual MBH98 methods. For this specific point – the replication of tree ring PC calculations, we re-iterate that we have *exact* replication of MBH98 PCs and that the above method, however implausible it may seem on first principles, is the method that was used in MBH98 tree ring PC calculations.

While PC algorithms are related to SVD algorithms, they are not exactly the same. We were able to replicate the above results with a PC algorithm, only by specifying an uncentered option as follows:

   PCA.mannomatic.prcomp<-prcomp(Tree.mannomatic, center=FALSE)

A different protocol is used for reporting eigenvalue information in the *svd* and *prcomp* algorithms, but the results are identical, allowing for the protocol.[3]

The method which we used in MM03, MM04a and MM04b can be shown by the corresponding command, displaying two differences clearly – not using the transformed data; and the use of a *centered* method.

   PCA.centered<-prcomp(Tree, center=TRUE)

The default value is center=TRUE and the result would also have been realized by:

   PCA.centered<-prcomp(Tree)

A centered calculation is clearly what one would expect in a "conventional" calculation. A centered calculation on the de-centered matrix is a possible variation, which can be implemented through:

   PCA.mannomatic.centered<-prcomp(Tree.mannomatic, center=TRUE)

In this case, the calculation is done on the covariance matrix of the transformed data and produces an intermediate result (in terms of the hockey stick shape of the PC1). We see no advantage to the MBH98 approach of using hundreds of lines of Fortran text to carry out the above functions, thereby opening up the possibility of error, since it can be easily done in a few lines of high-level programming languages, as shown above. While the MBH98 procedure may have originated as a programming error, the Corrigendum did not admit any error and seemed to take the position that the above method was intentional (even though it was undisclosed and tends to produce hockey sticks). Be that as it may, the key difference turns out to be not the stepwise principal components method, as claimed in *Mann et al. [2003],* but the use of an uncentered algorithm on de-centered data. Together with the MBH98 use of a non-archived version of the Gaspé series (containing a unique extrapolation), this fully reconciles MM03 and MBH98 results.

---

3 PCA.mannomatic.svd$d = PCA.mannomatic.prcomp$sdev *sqrt (N-1), where N is the number of years in the calculation. PCA.mannomatic.svd$u[,k]=PCA.mannomatic.prcomp$x[,k]/(sqrt(N-1) *PCA. mannomatic.prcomp$sdev[k]).

3. EFFECT OF SLIGHT VARIATIONS ON 15TH CENTURY TEMPERATURE RESULTS

We presented these results in a slightly different form in *McIntyre and McKitrick [2004a, 2004b]*. In response to these results (and to MM03), *Mann et al. [2004a, 2004b, 2004d]* argued that they can still obtain MBH98-type results under other conditions. While we differ with Mann et al. on the issue of which methodological assumptions are "correct", if the assumptions are specified sufficiently precisely, there is surprising consensus on the actual effects. Slight variations of methods and data lead on the one hand to MM-type results (with a 15th century higher than the late 20th century) or on the other hand to MBH-type results (with a 15th century lower than the late 20th). These can be summarized as follows.

Variants on the NOAMER PC1 (After Removing the Gaspé Series Extrapolation)

- In the MBH98 de-centered PC calculation, a small group of 20 primarily bristlecone pine sites, all but one of which were collected by Donald Graybill and which exhibit an unexplained 20th century growth spurt (see Section 5 below), dominate the PC1. Only 14 such chronologies account for over 93% of the variance in the PC1,[4] effectively omitting the influence of the other 56 proxies in the network. The PC1 in turn accounts for 38% of the total variance. In a centered calculation on the same data, the influence of the bristlecone pines drops to the PC4 (pointed out in *Mann et al., 2004b, 2004d).* The PC4 in a centered calculation only accounts for only about 8% of the total variance, which can be seen in calculations by Mann et al. in Figure 1 of *Mann et al. [2004d].*

- If a centered PC calculation on the North American network is carried out (as we advocate), then MM-type results occur if the first 2 NOAMER PCs are used in the AD1400 network (the number as used in MBH98), while MBH-type results occur if the NOAMER network is expanded to 5 PCs in the AD1400 segment (as proposed in *Mann et al., 2004b, 2004d).* Specifically, MBH-type results occur as long as the PC4 is retained, while MM-type results occur in any combination which excludes the PC4. Hence their conclusion about the uniqueness of the late 20th century climate hinges on the inclusion of a low-order PC series that only accounts for 8 percent of the variance of one proxy roster.

- If de-centered PC calculation is carried out (as in MBH98), then MM-type results still occur regardless of the presence or absence of the PC4 if the bristlecone pine sites are excluded, while MBH-type results occur if bristlecone pine sites (and PC4) are included. Mann's FTP site *[Mann, 2002–2004]* actually contains a sensitivity study on the effect of excluding 20 bristlecone pine sites[5] in which this adverse finding was discovered, but the results were not reported or stated publicly and could be discerned within the FTP site only with persistent detective work.

- If the data are transformed as in MBH98, but the principal components are calculated on the covariance matrix, rather than directly on the de-centered data, the

---

4 See <ftp://holocene.evsc.virginia.edu/pub/MBH98/TREE/ITRDB/NOAMER/BACKTO_1400/eof01.out>

5 See <ftp://holocene.evsc.virginia.edu/pub/MBH98/TREE/ITRDB.NOAMER/BACKTO_1400-CENSORED>

results move about halfway from MBH to MM. If the data are not transformed (MM), but the principal components are calculated on the correlation matrix rather than the covariance matrix, the results move part way from MM to MBH, with bristlecone pine data moving up from the PC4 to influence the PC2. In no case other than MBH98 do the bristlecone series influence PC1, ruling out their interpretation as the "dominant component of variance" *[Mann et al, 2004b]*

- If no North American PC1 is used at all in the AD1400 calculations (which occurs if PC calculations are done over the maximum period in which all sites are available, as done in MM03), then MM-type results occur under both centered and decentered PC calculations, with and without bristlecone pines.

Variants on the Gaspé Series (After Applying Centered PC Method on NOAMER)

- If the archived version of the Gaspé series is used, MM-type results occur. If the early (pre-1447) portion of the site chronology with less than 3 trees is not used [see discussion in *Jones and Mann, 2004*], MM-type results occur. If the duplicate version of the Gaspé series used as an individual proxy is not used (while continuing the use of the Gaspé series in the NOAMER network with or without the extrapolation), MM-type results occur.   MBH-type results occur only if a duplicate version of the Gaspé series is used as an individual proxy and the portion of the site chronology with 1–2 trees is used and if the first four years of the chronology are extrapolated under an ad hoc procedure not otherwise used in MBH98. *Mann et al. [2004a, 2004b]* justified the extrapolation as a means of maintaining representation of northern treeline series in this interval. If representation is achieved by use of the updated version of the Sheenjek River series (which meets replication standards in the 15th century), then MM-type results occur.

Variants on the Entire Procedure

- If, as is suggested in *Mann et al. [2004a, 2004b],* no PC calculations are applied to the North American and Stahle/SWM networks and the sites are instead used as individual proxies (while otherwise carrying on with MBH98 methods), then MBH-type results are obtained regardless of whether the Gaspé series is duplicated or extrapolated. In this case, the MBH temperature reconstruction becomes little more than an index of bristlecone pine growth. However, if the bristlecone pine sites are excluded from this network, then MM-type results are obtained.

We emphasize the consensus between ourselves and Mann et al. on the results of sufficiently well-defined calculations. The PC calculations themselves are replicated between parties to complete accuracy. Differences remain in the emulations of NH temperature (given the PC series), but *Mann et al. [2003]* showed a calculation with high early 15th century results if the North American PC1 were unavailable; the comments in *Mann et al. [2004b]* about the effect of the PC4 confirm this overall agreement if assumptions are sufficiently well defined.

These results also show that the effects of individual series are not necessarily

washed out in a multiproxy method of MBH98 type, contrary to suggestions in *Zorita et al. [2003]*[6] and *von Storch et al. [2004]*[7].

In response to a reader's suggestion, we performed a sensitivity test in which we arbitrarily increased the ring widths of all non-Graybill (50 of 70) sites by +0.5 (index units) in the first half of the 15th century, and then re-calculated the PC1 under MBH98 methodology. The purpose is to evaluate how well the added variance is retained in the final temperature index. We provide the exact script here both to describe the calculation exactly and because the results are initially very counter-intuitive and have provoked some disbelief.

```
Tree.adj<- Tree                                       #creates mirror object for testing
Tree.adj[1:50,!graybill]<- Tree.adj[1:50,!graybill]+0.5  # adds 0.5 to all non-Graybill sites (mean is 1)
Tree.adj.mannomatic<- apply(Tree.adj,2, mannomatic)   # applies MBH98 transformation to columns
PCA.adj.mannomatic<-svd(Tree.adj.mannomatic)          #svd on data matrix
PC1.adj.mannomatic<- PCA.adj.mannomatic$u [,1]        #selects PC1 from svd model
```

The results of this calculation are shown in Figure 2 together with the results from a centered calculation (all results smoothed). For a centered calculation, the increased ring widths for the first 50 years lead to an increase in the PC1 as expected. However, under the MBH98 de-centered method the *increased* ring widths for 50 non-Graybill sites in the early 15th century causes a significant *decrease* (!) in the early 15th century PC1. *Carried forward through to Northern Hemisphere temperature calculations, these increased ring widths would be construed by the MBH98 method as evidence of **colder** temperatures in the early 15th century.*

This rather perverse result nicely illustrates a problem of mechanically applying a numerical algorithm like PC analysis without regard to whether it makes sense for the underlying physical process. PC methods are indifferent to the orientation (up or down) of a series – the difference is merely the presence or absence of a negative sign. A vivid example in this context is the archived PC1 for *Mann et al. [1999]*,[8] which is upside-down as archived, but which was flipped for presentation purposes in *Mann et al. [1999]*. Under the MBH98 algorithm, the addition of the extra values in the first half of the 15th century causes the algorithm to flip the series upside-down so that they match as well as possible to the bristlecone pines, whose hockey stick pattern is

---

6 "MBH98's method yields an estimation of the value of the temperature PCs that is optimal for the set of climate indicators as a whole, so that the estimations of individual PCs cannot be traced back to a particular subset of indicators or to an individual climate indicator. This reconstruction method offers the advantage that possible errors in particular indicators are not critical, since the signal is extracted from all the indicators simultaneously."

7 "The optimized temperature fields target the whole available proxy network at a given time, so that the inclusion of a few instrumental data sets in the network should have little influence on the estimated fields, unless the instrumental records are explicitly overweighted. The advantage is that the method is robust against very noisy local records. This contrasts with direct regression methods, where the estimated temperature fields are the predictands of a regression equation. In this case a few instrumental records, highly correlated to the temperature fields, may overwhelm the influence of proxy records with lower correlations in the calibration period."

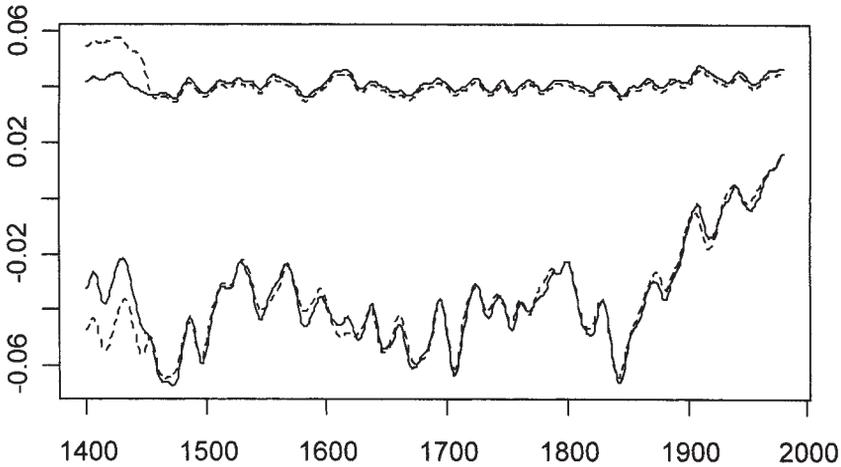8 ftp://ftp.ngdc.noaa.gov/paleo/contributions_by_author/mann1999/proxies/itrdb-namer-pc1.dat>.

Figure 2. North American AD1400 PC1. Above: PC1 using centered calculations; below: MBH98 PC1 using decentered method. Solid-base case; dashed – with arbitrary addition of 0.5 to non-Graybill sites from 1400–1450. 25-year smoothing is applied.

imprinted on the PC1. This does not occur using a centered algorithm.

4. ROBUSTNESS

The sensitivity of 15th century results to such slight variations of method and data show a fundamental instability in MBH98 results, related especially to the presence or absence of bristlecone pines and Gaspé cedars. This flatly contradicts some claims by Mann et al. about the robustness of MBH98 results. MBH98 stated the following:

> **the long-term trend in NH is relatively robust to the inclusion of dendroclimatic indicators in the network**, suggesting that potential tree growth trend biases are not influential in the multiproxy climate reconstructions. (p. 783, emphasis added)

This was stated in even stronger, and equally misleading, terms in *Mann et al. [2000]* as follows:

> We have also verified that possible low-frequency bias due to non-climatic influences on dendroclimatic (tree-ring) indicators is not problematic in our temperature reconstructions….**Whether we use all data, exclude tree rings, or base a reconstruction only on tree rings, has no significant effect on the form of the reconstruction for the period in question.** … These comparisons show no evidence that the possible biases inherent to tree-ring (alone) based studies

impair in any significant way the multiproxy-based temperature pattern reconstructions discussed here.
(http://www.ngdc.noaa.gov/paleo/ei/ei_nodendro.html, emphasis added)

The synopsis of results in Section 3 effectively disproves these claims, regardless of the point of view that one may take on questions such as whether 2 PCs or 5 PCs is "correct" for the AD1400 North American network. Each of the permutations discussed above is a sensitivity test much less draconian than excluding all tree rings. Both the Gaspé cedar series and the bristlecone pine series are obviously subsets of the dendroclimatic indicators and each has a significant effect on the 15th century results, as indeed do the specific methodological decisions (extrapolation, decentered PC calculations), which enhance the effect of these series.

Figure 3 may be helpful in illustrating exactly why these two series have such a dramatic impact on early 15th century results. The left panel is a scatterplot as follows. For each of the 22 proxies in the AD1400 roster we computed the correlation between each proxy and the temperature PC1 over the 1902–1980 interval (x-axis), and the difference between the 1902–1980 mean and the 1400–1450 mean, divided by the 1400–1980 standard deviation (y-axis). The 1902–1980 interval is the MBH98 calibration period and is the interval over which the mean is computed in the PC decentering. It can be shown that the PC weights for each proxy in the AD1400 network are closely related to the correlation with the temperature PC1. The difference of means is a measure of "MBH-ness" – series with a zero value are flat, while those with an absolute value in excess of 1 have a a hockey stick shape (sometimes upside-up and sometimes upside-down). The two points in the top right hand corner represent
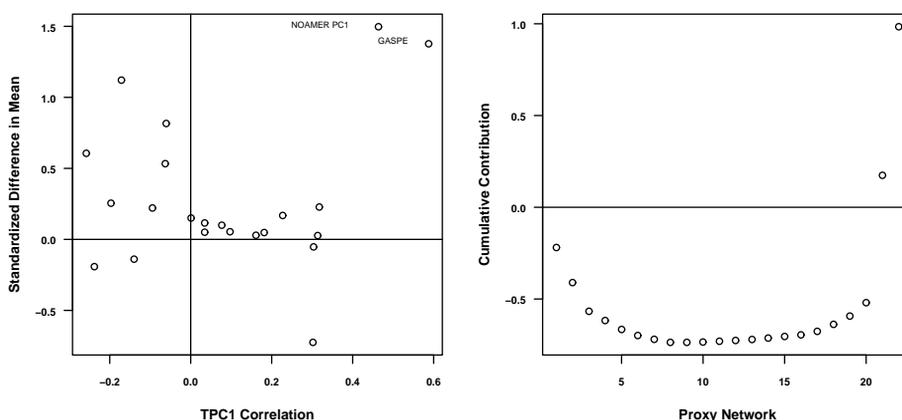

Figure 3. Left: Scatterplot with x-axis: correlation between AD1400 step proxies and the temperature PC1 in the calibration period 1902–1980; y-axis: difference between 1902–1980 mean and 1400–1450 mean divided by 1400–1980 standard deviation. right: cumulative contribution to standardized difference between 1902–1980 mean and 1400–1450 mean.

the MBH98 NOAMER PC1 and Gaspé tree ring series. Except for these two points there is an overall *negative* relationship between the difference of means and the correlation with temperature: i.e. hockey stick series fit the temperature data relatively poorly in the calibration interval. But the NOAMER PC1 and Gaspé series are such influential outliers that they reverse this pattern for the model as a whole.

In the right panel the 22 series in the AD1400 step are introduced sequentially into the multiproxy calibration model, with the Gaspé and NOAMER PC1 series introduced last. The standardized difference between the 1902–1980 mean NH temperature and the 1400–1450 mean NH temperature is computed at each step. The relatively high 1902–1980 temperature in MBH98 (i.e. the hockey stick shape) results entirely from the contributions of the two final, outlier values.

If the same calculations are carried out using centered principal components calculations and the Sheenjek River series is used as a northern treeline proxy instead of the Gaspé series, as represented in Figure 4, there are no longer two outlier series, resulting in the 1400–1450 mean temperature being higher than the 1902–1980 mean temperature.

Some consternation has been expressed by critics of MM03 that its high early 15th century values are inconsistent with other supposedly independent temperature reconstructions. However, the MM results are obtained from the same underlying proxy set as MBH98. The influence of the two outlier series can be seen in a different way in Figure 5, which shows a simple comparison of the mean of Gaspé and NOAMER PC1 against a weighted average of 6 series used in the AD1400 network
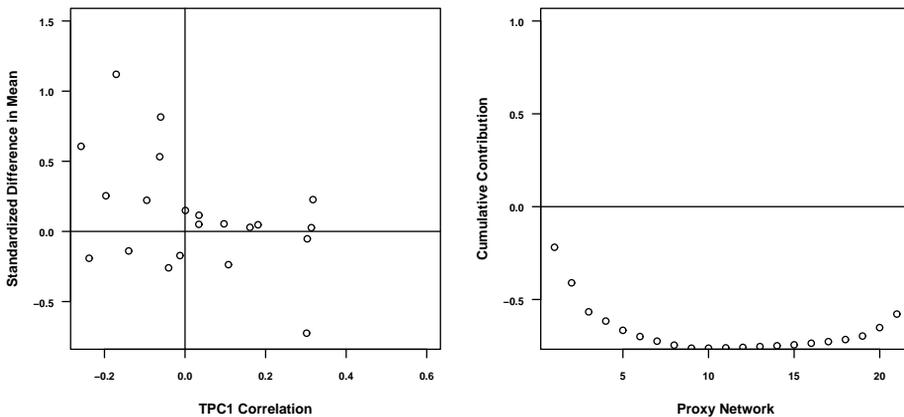

Figure 4. As with Figure 3, but with centered PC calculations and Sheenjek River. Left: Scatterplot with x-axis: correlation between AD1400 step proxies and the temperature PC1 in the calibration period 1902–1980; y-axis: difference between 1902–1980 mean and 1400–1450 mean divided by 1400–1980 standard deviation. right: cumulative contribution to standardized difference between 1902–1980 mean and 1400–1450 mean. The 1400–1450 mean temperature is now higher than the 1902–1980 mean temperature.
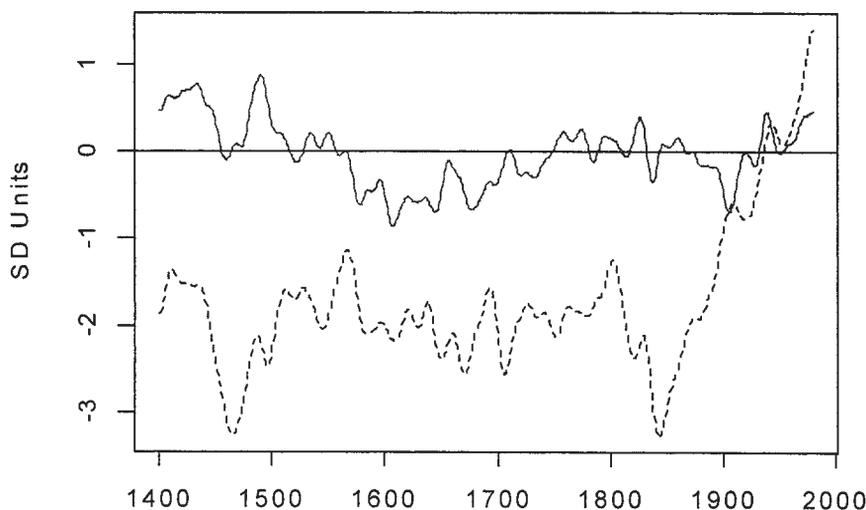
Figure 5. Solid: Weighted average of 6 MBH98 AD1400 step proxies: 4 Quelccaya
series (averaged to one series), Tornetrask temperature reconstruction and Tasmania
temperature reconstruction; Dashed – average of MBH98 NOAMER PC1 and
Gaspé. All series smoothed with 25-year Gaussian filter.

(and often used in other multiproxy studies). The weights are calculated as follows:
equal weight is given to the Tornetrask and Tasmania series, while each of the 4
constituent Quelccaya series are given a quarter-weight, reversing the sign for the
accumulation series. All series are standardized. The strong negative bias of the two
outlier series is evident, as is the closer relationship of the 6 series average to the MM-
type reconstruction.

5. BRISTLECONE PINES AND GASPÉ CEDARS
Although considerable publicity has attached to our demonstration that the PC
methods used in MBH98 nearly always produce hockey sticks, we are equally
concerned about the validity of series so selected for over-weighting as temperature
proxies. While our attention was drawn to bristlecone pines (and to Gaspé cedars) by
methodological artifices in MBH98, ultimately, the more important issue is the
validity of the proxies themselves. This applies particularly for the 1000–1399
extension of MBH98 contained in *Mann et al. [1999]*. In this case, because of the
reduction in the number of sites, the majority of sites in the AD1000 network end up
being bristlecone pine sites, which dominate the PC1 in *Mann et al. [1999]* simply
because of their longevity, not through a mathematical artifice (as in MBH98).

Given the pivotal dependence of MBH98 results on bristlecone pines and Gaspé
cedars, one would have thought that there would be copious literature proving the
validity of these indicators as temperature proxies. Instead the specialist literature only
raises questions about each indicator which need to be resolved prior to using them as

temperature proxies at all, let alone considering them as uniquely accurate stenographs of the world's temperature history.

5.1 Bristlecone Pines

There has been an undoubted increase in bristlecone pine ring widths in the 20th century. *Graybill and Idso [1993]* explicitly stated it is greater than could be explained by temperature. Ironically, *Mann et al. [1999]* (referring to the bristlecone pine sites) admits the same thing:

> 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. (p. 760)

The anomalous 20th century growth rate for bristlecone pines is illustrated in Figure 6, which compares the standardized MBH98 PC1 (dominated by bristlecone pines) to the *Briffa et al. [1992a]* North American temperature reconstruction (using tree rings from many species), which is used in MBH98 itself, as well as *Jones and Bradley [1993]*, *Jones et al. [1998]* and *Jones and Mann [2004]*. There is little visual relationship. The correlation between the two series in the MBH98 calibration period of 1902–1980 is 0; the RE statistic for the MBH98 PC1 as a predictor for the Briffa temperature reconstruction in a verification period of 1600–1901 is –7.7, showing no skill whatsoever. Thus, whatever "dominant component of variance" *[Mann et al., 2004a]* is supposedly captured in the MBH98 PC1 has apparently escaped detection in the Briffa reconstruction. The strong negative bias of the MBH98 PC1 is evident in comparison to the Briffa reconstruction. The strong negative bias of the MBH98 PC1 is also evident in periods where we have instrumental records in North America. There is no reason to believe that average temperatures in the 18th century were negative 3 standard deviation units.

Despite the reliance of MBH98 on the North American PC1, the validity of this series as a temperature proxy was not independently established in peer-reviewed literature. Co-author Hughes stated later *[Hughes and Funkhouser, 2003]* that the anomalous growth rate of bristlecone pines was a "mystery", which should have raised questions about the PC1. The strong difference between the Briffa re-construction, comprised of many species, and the MBH98 PC1 (representing only bristlecone pines) should also have raised questions about whether there may be species-particular effects related to any of the numerous unusual features of bristlecone pines.

We surveyed the literature on bristlecone pines and report here on many peculiarities pertaining to this species, which should be clearly addressed prior to relying on the MBH98 PC1 for policy purposes.

Bristlecones are famously long-lived, but despite this, do not appear to senesce *[Lanner and Connor, 2001; Connor and Lanner, 1991]*. They occur in an unusual strip bark form, where the bark in most long-lived trees dies around the circumference except for a small strip on one side. Unlike most pines, they continue to respire during the winter thereby consuming photosynthate *[Schulze et al., 1967]*.
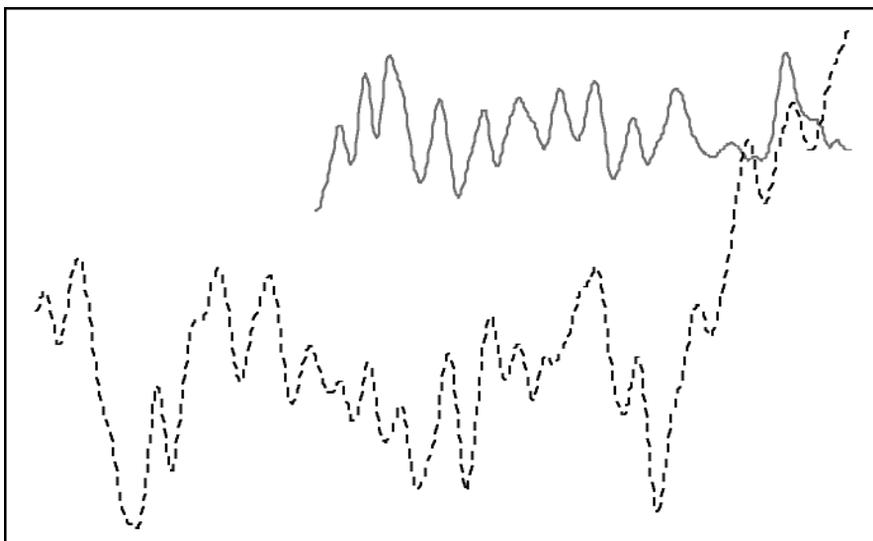
Figure 6. Solid: North American temperature reconstruction of Briffa et al. [1992a].
Dashed: the MBH98 North American PC1. All series smoothed with a 25-year
Gaussian filter. Both series are standardized by subtracting the 1902–1980 mean and
dividing by the 1902–1980 standard deviation (re-scaling is not an issue here, since
the PC calculations have already been done.)

Tree lines at bristlecone pine sites were significantly higher in the past, especially during the Holocene Optimum about 6000 years ago, but also during the medieval period *[LaMarche and Mooney, 1967; LaMarche, 1973]*. In 1957, bristlecone pines in the Campito Mountain area were not reproducing *[Billings and Thompson, 1957]*, indicating that more favourable conditions than the 1950s were required for bristlecones pine to become seedlings.

The Sheep Mountain series, ca534, is the top-weighted series in the MBH98 NOAMER PC1[1] receiving 390 times the weight of the least-weighted series. *LaMarche, Fritts, Graybill and Rose [1984]* (all authors listed for emphasis) argued that the anomalous 20th century growth in bristlecone pines was attributable to CO_2 fertilization, using Sheep Mountain as an example. In the 1980s, Donald Graybill followed up at the other bristlecone pine sites, collecting the samples which later comprised the NOAMER PC1. Graybill specifically sought out strip bark samples and reported that strip bark forms had much stronger 20th century growth than entire bark forms at the same site *[Graybill and Idso, 1993]*. More recently, *Bunn et al. [2003]* confirmed higher growth in strip bark forms than entire bark forms at sites in Montana.

---

1 ftp://holocene.evsc.virginia.edu/pub/MBH98/TREE/ITRDB/NOAMER/BACK TO_1400/ECOF01.OUT
and series identifications in ftp://holocene.evsc.virginia.edu/pub/MBH98/TREE/ITRD /NOAMER/
BACTO_1400/noamer-itrdb-ad1400

Brooks et al. [1996] also pointed out the impact of anthropogenic nitrogen on fertilization of high-altitude bristlecone pines, stating that:

> At these high-elevation catchments there has been a shift in ecosystem dynamics from an N-limited system to an N-saturated system as a result of anthropogenically-fixed N in wetfall and dryfall. Results from the Western Lakes Survey component of the National Surface Water Survey show that N saturation is a regional problem in the Colorado Front Range, with many lakes having (NO3-) concentrations greater than 10 meq/L. Foliar N to P ratios in Bristlecone Pine increase with elevation in the Colorado Front Range, indicating that at higher elevations P is translocated from foliar tissue more efficiently than N and that increasing atmospheric deposition of N with elevation is causing a change from N limitation to P limitation in the highest-elevation Bristlecone Pines.

Graybill and Idso [1993] attributed the anomalous 20th century growth of strip-bark forms to CO_2 fertilization There are some possible reasons why CO_2 fertilization may affect high-altitude strip bark forms more strongly than lower-altitude entire-bark forms and there is specific evidence for CO_2 fertilization for vegetation from the White Mountains, California, where important bristlecone pine stands are located *[Mooney et al., 1964]*. The response to changes of CO_2 concentration in controlled experiments is strongly non-linear and attenuates as CO_2 levels increase. CO_2 levels at the high altitudes of bristlecone pines (3000–3500 m) are significantly lower than at sea level and, at the lower CO_2 levels at high-altitude, the response to increased CO_2 levels is in a range with stronger response.

The general hypothesis of CO2 fertilization of tree growth has been hotly contested. Opponents [e.g. *Jacoby and D'Arrigo, 1997 and Graumlich, 1991*] have pointed to other sites which do not exhibit anomalous 20th century growth and argued that the anomalous growth effect is limited to high-altitude strip bark forms. More recently, *Schauer et al. [2001]* reported that trees with strip-bark have significantly lower pre-dawn water potentials during the driest part of the growing season relative to non-strip-barked trees. So the anomalous growth of strip bark trees may be related to effects differing from CO_2.

Whatever side one takes on the CO_2 fertilization debate is really immaterial. Even if one adopts the position of Jacoby and Graumlich – that the effect is limited to strip bark forms – since the NOAMER PC1 is comprised almost entirely of strip-bark forms it could easily be affected as a proxy, even under the limited position taken by Jacoby and Graumlich.

Mann et al. [1999] purported to adjust the NOAMER PC1 for CO_2 fertilization, by coercing the shape of the NOAMER PC1 to the Jacoby northern treeline reconstruction in the 1750–1980 period, arguing that the northern treeline series would not be affected by CO_2 levels. Once one gets into such ad hoc adjustments, many new questions need to be answered about the validity of the adjustment procedure. In the

actual *Mann et al. [1999]* adjustment, the main adjustment for "CO_2 fertilization" takes place in the 19th century rather than the 20th century, with *Mann et al. [1999]* being forced into the counterintuitive position that the effect of CO_2 fertilization was somehow stronger in the 19th century but became attenuated in the 20th century, the exact opposite of the hypothesis of *LaMarche et al. [1984]* and later *Graybill and Idso [1993].* If the differences between the northern treeline series and the bristlecone pines arise from some other factor (a couple of possibilities are discussed below), then the *Mann et al. [1999]* "adjustment" would have made the proxy record even more distorted. In MBH98, no such adjustment was made in the AD1400 period in any event. (It also appears that no such adjustment was made in *Mann and Jones [2003]* or *Jones and Mann [2004],* but we are presently unable to confirm this.)

An important influence on these trees is precipitation. *Mann and Jones [2003]* point out the need to distinguish between temperature and precipitation effects, which may have a different expression. Within bristlecone pine literature, *LaMarche and Stockton [1974]* pointed out that high-altitude bristlecone pine stands have both a lower limit and upper limit and argued that bristlecone pine growth at the lower limit was controlled by precipitation and at the upper limit by temperature. *Hughes and Funkhouser [2003]* found regional correlations among high-altitude bristlecone pine growth, which they attributed to regional climate, but still concluded that the anomalous 20th century growth was a "mystery". Even in upper limit stands, the bristlecone pine stands in the PC1 are located in semi-arid regions and the bifurcation in *LaMarche and Stockton [1974]* may be overly simplistic. Studies of actual bristlecone pine growth have shown that it is limited by soil moisture *[Fritts, 1969; Beasley and Klemmedson, 1973].* Even in higher stands, their principal botanical competition in many locations is with big sagebrush *[Wright and Mooney, 1965; Mooney et al., 1964]* with bristlecones outcompeting big sagebrush on moister dolomite substrate. This effect is vividly illustrated by Figure 2 of *Wright and Mooney [1965],* where the sharp geological contact between the dolomite and sandstone is clearly shown by the change from bristlecone pines to sagebrush at the same elevation. The same effect is also perhaps shown in the charming 19th century painting (Figure 7), where a sharp change in vegetation at the same elevation is easily observed. There is evidence that higher moisture levels in the 20th century in the American Southwest accounted for high growth rates in New Mexico *[Grissino-Mayer, 1996; D'Arrigo and Jacoby, 1991],* where two of the *LaMarche and Stockton [1974]* sites are located. The effect may extend to other locations. In the classical bristlecone pine sites of the White Mountains, where a weather station operated close to Sheep Mountain and Campito Mountain from 1954 to 1980, records show low ring widths correlate to drought, even in upper limit stands. *Mann and Jones [2003]* pointed out that precipitation proxies need to be carefully distinguished from temperature proxies and a complete demonstration that these effects have been separated in bristlecone pines is obviously required. *Williams [1996]* reported that a continuous climate record since 1951 at Niwot Ridge in the Colorado Front Range, near a bristlecone pine site, showed a decline in mean annual temperature and an increase in annual precipitation amount.

There is one other issue that needs to be canvassed and eliminated prior to reliance on bristlecone pines. The pulse in bristlecone pine growth is contemporaneous with a

Figure 7. 19th century painting showing bristlecone pines. Source: NOAA website. In the background, the pines reaching up the hill are almost certainly bristlecone pines on a dolomite substrate, with a sandstone substrate where there are no pines. Retrieved from <http://www.ngdc.noaa.gov/paleo/drought/drght_graumlich.html>

pulse in woody plant growth throughout the American Southwest, attributed to overgrazing by sheep in the late 19th Century (see Figure 8), which in turn followed the extension of the railroads *[Allen, 1998; Allen et al., 1998]*. Sheep differ from other species in that they will completely destroy grasslands by eating down to the roots, leaving barrens *[Allen, 1998]*. Although *Allen [1998]* only documented the expansion of pinyon pines and junipers into terrain formerly occupied by 19th century grasslands, Allen (2004, pers. comm.) did not exclude the possibility of a similar effect involved in anomalous 20th century growth for bristlecone pines, but was unaware of any studies on the topic. There is a published reference to the introduction of large commercial sheep flocks in the late 19th century in the White Mountains CA *[St. Andre et al. 1967]*, where the key sites of Sheep Mountain and Campito Mountain are located. The founder of the Sierra Club, John Muir, complained of the depredations of sheep in the Sierra Nevadas (adjacent to the White Mountains) as "hoofed locusts" *[Muir, 1911]*. Carl Purpus, a late 19th century botanical collector in the Sierra Nevadas, stated in 1896 that commercial flocks had cleaned out all grass to the top of Old Mt Whitney [present-day Mount Langley, which reaches 4,280 m] *[Ertter, 1988]*. Allen (pers. comm., 2004) said that there was a large commercial sheep trail at Jicarita Peak NM, another bristlecone pine site studied by *LaMarche and Stockton [1974]*. In severe high-altitude terrain, even after the departure of commercial flocks, a small population of bighorn sheep could prevent the re-establishment of grass (Leslie Thomas, Colorado Springs, landscape architect, pers. comm.) Since grass (and other herbs) compete with pines for scarce moisture, one can hardly exclude, on a priori basis, the

possibility of a connection between anomalous 20th century growth rates of bristlecone pines and a growth release following 19th century overgrazing, as experienced elsewhere in the American Southwest.

Figure 8. Sheep grazing in ponderosa pine forests and grasslands near Flagstaff, AZ, ca. 1899. Image 21a by F.H. Maude, Cline Library Special Collections, Northern Arizona University. Retrieved from
http://www.cpluhna.nau.edu/Change/grazing.htm

Finally, there may even be problems with the site chronologies as indexes of actual growth. *Cook and Peters [1997]* pointed out that conventional dendrochronology techniques resulted in a bias in 20th century results at Campito Mountain, one of the Graybill sites in the NOAMER PC1. Presumably the same effect applies to other bristlecone pine sites.

If the reader takes the (reasonable, we think) view that these unusual trees are not mystical antennae for an elusive "climate signal" missed by all other proxy indicators, then each of the above problems and issues must be dealt with systematically, prior to any reliance being placed on bristlecone pine ring widths as the dominant arbiter of world climate history.

4.2 Gaspé Cedar Series

The other critical series in MBH98 is the Gaspé cedar chronology. There are many interesting similarities between bristlecone pines and cedars. *Kelly, Cook and Larson [1992]* likened Eastern white cedars to bristlecone pines, pointing out the occurrence of strip bark forms and the apparent lack of senescence. Larson (pers. comm., 2004) stated that there is little aging effect in Ontario cliff cedars: they grow slowly when they are young and slowly when they are old.

There is virtually no site information on the Gaspé cedar series. *Sheppard and Cook [1988]* mentions the site, but is not a comprehensive publication. They commented on its "peculiar" growth spurt in the 20th century, but cautioned that there were no other northern white cedar chronologies available so it was not possible to say whether it was indicative of climatic conditions or some other influence. *Sheppard and Cook [1988]* referred to two other pending cedar studies, one in Maine and one in Michigan. The Maine series (Sag Pond) has been archived at WDCP and does not show any 20th century trend, while the Michigan series seems to have remained unpublished. *Archambault and Bergeron [1992]* published a cedar series from Lac Duparquet, Quebec, but it does not show any 20th century trend and the authors reported a positive correlation to precipitation and a negative relationship to June temperature. For Ontario cedars, *Kelly et al. [1994]* reported a strong negative relationship between temperature and observed growth rates in the 20th century.

Cedar growth is optimal under cool and moist conditions *[Kelly et al., 1994; Matthes-Sears and Larson, 1990]* and declines both in very hot and very cold weather (an upside-down U). A similar upside-down U pattern has been reported for bristlecone pines and two other conifer species *[Schoettle, 2004]*. In fact, the possibility of a quadratic ring width response to temperature has been recently posited by *D'Arrigo et al. [2004]* in connection with Twisted Tree Heartrot Hill. This possibility has far-reaching implications on the entire enterprise of estimating past temperatures from tree ring widths: with a quadratic (upside-down U) response, it is impossible to determine whether a past narrow ring width resulted from cold or hot weather.

We carried out our own comparison between gridcell temperature in the Gaspé area and Gaspé temperatures and did not find any correlation.

Cook and Peters [1997] discussed above, explored spurious end-of-sample growth bias as an artifact of tree-ring chronology de-trending. Amazingly, in addition to the Campito Mountain bristlecone pine site, their other main example was the Gaspé series (cana036). In order to eliminate this bias, the underlying tree ring chronologies would have to be re-calculated, a calculation which would have the effect of reducing its hockey-stick shape, with implications that stand alone from any of the other issues raised in this paper.

The Gaspé site was re-sampled in the early 1990s; we have seen a site chronology showing that the re-sampling did not replicate the previously reported 20th century growth spurt. However, the new data has not been published or archived, and the originating authors have refused to disclose the new data on the grounds that the older data "better" shows temperature and because their research is "mission-oriented". We have sought coordinates of the actual site in order to commission a re-sampling of the site, but we have not received this information despite repeated requests.

5. REFUTATION OF COUNTER-ARGUMENTS OF MANN ET AL.

We now turn to a discussion of recent arguments of *Mann et al. [2003, 2004a, 2004b, 2004d],* purporting to counter our various criticisms.

5.1 "Effective Omission" of Indicators

Mann et al. [2004a, 2004b] argued that our use of centered principal components calculations amounted to an "effective omission" of the 70 sites of the North American network. They showed that a calculation excluding the North American PC1 also resulted in MM-type results with high early 15th century values. Although the calculations to which they were referring (using centered PC methods) did *not* actually omit this network, since the results were similar to results without the PC1, Mann et al. argued that our calculations *effectively omitted* these indicators. This critique fails on several counts. First, the North American PC1 is only one of 22 series in the AD1400 step. A robust statistical method should be relatively insensitive to the presence or absence of one of 22 series. If centered principal components calculations are used, the temperature index is relatively insensitive to the presence or absence of the North American PC1. On the other hand, if de-centered principal components calculations are used, the results are very unstable to the presence or absence of the North American PC1. Robustness considerations therefore tend to support the use of a conventional centered PC method. Second, using the MBH98 decentered method, 14 bristlecone sites account for over 99% of the explained variance in the PC1. Using the terminology of Mann et al., under the decentered methodology, the other 56 sites are "effectively omitted" from the PC1, which is merely a carrier for the bristlecone pines. Using a centered methodology, the PC1 is relatively similar to the mean of all the series. Thus, we believe that it is more accurate to characterize MBH98 de-centering methods as "effectively omitting" the majority of tree ring sites.

Mann et al. [2004a, 2004b]* also argued that use of the archived version of the Gaspé series amounted to an "effective omission" of the northern treeline series. *Jones and Mann [2004]* do not use the Gaspé series at all and only use the Jacoby northern treeline series in the relatively well-replicated portion after 1601. Simply applying the quality control criteria of *Jones and Mann [2004]* should not provoke complaints about "effective omission". Moreover, we have specifically maintained the number of northern treeline series in the AD1400 step, by using the updated version of the Sheenjek River series (which could have been used in 1997). Replication in the Sheenjek River series is much superior to that of the Gaspé series, which does not meet standard quality control criteria in its early portion.

5.2 Tendency of decentered PC methods to yield PC1 hockeysticks

In *McIntyre and McKitrick [2004a, 2004b],* we pointed out that the de-centered PC method used in MBH98 tends to produce hockey-stick shaped series. We have sharpened this result considerably in *McIntyre and McKitrick [2005].* There we define a "hockey stick" as a series in which the 1902–1980 mean differs from the long-term mean by more than 1 standard deviation (–). Applying the MBH98 decentered PC method to trendless red noise with persistence properties of the North American tree ring network (modeled as fractional processes), in 10,000 simulations we found that

the 1902–1980 mean differed from the 1400–1980 mean by more than 1 σ over 99% of the time, (1.5 σ –72%; 1.75 σ –19% and 2 σ –0.2%). The hockey stick blades sloped up about half the time and down half the time, but the 1902-1980 mean is almost never within one σ of the 1400–1980 mean. PC series are unoriented so that no significance is attributed to the sign.

In their comment to the earlier version of this argument, Mann et al. [2004a, 2004b] argued that their PC series were simply linear combinations of the underlying proxies and that no pattern could be produced in the PC1 which was not in the underlying data. It is of course true that the PC series are linear combinations of the proxies, but it is evident that the de-centering process preferentially selects series with hockey-stick shapes and this over-weighting is what yields a pattern that is not representative of the underlying data. The exclusive selection of bristlecones into the PC1 should give rise to serious examination of why all other proxies are so efficiently discarded – a discussion which does not occur in MBH98.

5.3 Lack of a linear response to temperature in "key" proxies

In *McIntyre and McKitrick [2004b],* in our criticism of bristlecone pines as an arbiter of world climate, we pointed out (as above) that a linear response to temperature had not been established for these sites (as seemingly required by MBH98). *Mann et al. [2004b]* replied that:

> MM04 demonstrate their failure to understand our methods by claiming that we required that "proxies follow a linear temperature response". In fact we specified (MBH98) that indicators should be "linearly related to one or more of the instrumental training patterns[2]", not local temperatures.

We doubt the authors really believe the idea of a temperature proxy exhibiting no relationship to local temperature makes much sense. It is instructive to compare this response to the policy articulated in *Jones and Mann [2004],* which states:

> A number of other temperature reconstructions used in earlier multiproxy composites or in review papers [e.g., Jones et al., 1998; Mann et al., 1998a, 1999; Mann and Jones, 2003] are not included. This is because they are either less resolved than decadal resolution [e.g., Dahl-Jensen et al., 1998] or correlations with local grid box temperatures are not significant …

Jones and Mann [2004] do consider "climate field reconstructions" (CFRs), which appear to be similar to "instrumental training patterns" of MBH98. In this case, *Jones and Mann [2004]* argue that the CFRs should be shown to be similar to some aspect of local climate during some part of the year. This would seem to invite opportunistic use of either precipitation or temperature as a climate indicator, something for which they reproached *Soon et al. [2003].* But perhaps most telling is the comment of MBH98 co-author Hughes in *Hughes and Funkhouser [2003],* who did not attribute

the bristlecone pine growth to an "instrumental training pattern", but stated that their anomalous 20th century growth rate is a "mystery".

5.4 Insignificant Values of Verification Statistics

Mann et al. [2004a, 2004b] have argued that, regardless of how they got their results, their reconstruction with decentered PC methods and extrapolation of the duplicate Gaspé series has greater "skill" than a reconstruction with centered PC methods and use of the archived version of the Gaspé series. Most dendroclimatic reconstructions provide a suite of verification statistics, including RE, R^2, CE, sign test and product mean test *[e.g. Cook et al, 1994]*. In MBH98, only the RE statistic is reported for steps prior to the AD1820 step, including the controversial AD1400 step. Mann et al. have not provided their own results for the other verification statistics or supporting calculations from which these statistics could be calculated, and have refused requests for this information. *McIntyre and McKitrick, 2005,* using Monte Carlo simulations, shows that the MBH98 benchmark for 99% significance for the RE statistic is substantially under-stated (0.0 in MBH98 versus a Monte Carlo estimate of 0.59) and that the R^2 and other verification statistics, which were not reported in MBH98, are statistically insignificant in the AD1400 step.

Mann et al. [2004b]* contained a diatribe against the R^2 statistic. However, in other papers [e.g. *Mann and Jones, 2003],* when they were in his favour, Mann has reported R^2 statistics. In this case, we estimate the R^2 statistic as being only 0.02 – obviously well short of statistical significance and strongly indicating that even the lower level of RE significance discussed above is spurious.

5.5. "Confirmation" by other studies

Mann et al. [2003, 2004a, 2004b] argued that their results are similar to those of "independent" studies, such as *Jones, Briffa et al. [1998], Crowley and Lowery [2000], Briffa, Jones et al [2001]*, *Mann and Jones [2003]* and *Jones and Mann [2004],* calculated with different proxies and different methods. This "similarity" is typically shown by "spaghetti" diagrams supposedly illustrating the similarity, rather than through detailed analysis.

These studies are hardly "independent". If all the authors in the multiproxy articles are listed, one sees much overlapping. Mann himself was a co-author of two supposedly "independent" studies; his sometime co-author (as well as Bradley's sometime co-author) Jones was co-author of two of the others. Even *Crowley and Lowery [2000],* where there is no apparent overlap, stated that they used data supplied by Jones. This hardly amounts to "independence" in any conventional use of the term.

Many proxies are re-used in these studies, a point which *Briffa and Osborn [1999]* acknowledged, as follows:

> An uninformed reader would be forgiven for interpreting the similarity between the 1000-year temperature curve of Mann et al. and a variety of others also representing either temperature change over the NH as a whole or a large part of it (see the figure) as strong corroboration of their general validity, and, to some extent, this may well be so.

>   Unfortunately, very few of the series are truly independent: There is a
>   degree of common input to virtually every one, because there are still
>   only a small number of long, well-dated, high-resolution proxy
>   records.

Briffa's Polar Urals and Tornetrask series [*Briffa et al., 1995; Briffa et al., 1992b* respectively] are recurrent proxies as is Cook's Tasmania reconstruction [*Cook et al., 1991, 1992*]. The North American PC1, criticized here, is used as a proxy in *Mann et al. [1999]*, *Mann and Jones [2003]* and *Jones and Mann [2004]*.

Most importantly, even if such articles generate similar results to MBH98, that does not prove that MBH98 results were calculated correctly. Mann et al. have to support MBH98 on its own terms; appeals to other results are completely irrelevant.

For rhetorical purposes, agencies like the IPCC may well turn to these other studies for support, if MBH98 can no longer be used, but the prominent reliance on MBH98/99 in the Third Assessment Report is a matter of public record and cannot now be undone. If there is any lesson from our work it is that, before making prominent use of these other studies, each one needs to be proven replicable. However critical we may be of MBH98, the disclosure for nearly all the other studies is significantly worse:

- After over 20 requests, Crowley (pers. comm., Oct. 2004) supplied smoothed and transformed versions of proxy data used in *Crowley and Lowery [2000],* but stated that he could not find the actual data versions used so that these could be verified.
- A listing of the sites used in *Briffa et al. [2001]* has never been published or archived. The authors have not responded to requests for data.
- A listing of sites in *Esper et al. [2002]* is available, but the majority of site data is not publicly archived.
- Most of the data from *Mann and Jones [2003]* and *Jones and Mann [2004]* was eventually provided by Jones in July 2004. However, Jones was unable to provide the weightings used in the creation of the final results, as these were in the possession of co-author Mann.
- Of these studies, only *Jones et al. [1998]* has a relatively complete record.

None of these studies provides a careful, objective analysis of how the particular proxy records are selected from the thousands available, thereby leaving unanswered the possibility of cherry-picking. Replication is only the first step in assessment. One then has to assess the quality of the proxies actually used. For example, we have concerns about potential problems in Briffa's Polar Urals record *[Briffa et al., 1995],* which has a very significant effect on medieval values in several of these studies. We intend to address these issues in the future.

6. DISCUSSION

There are many large issues at stake in this discussion, mainly because of the powerful role a handful of published paleoclimate studies are playing in policy decisions.

The ability of later researchers to carry out independent due diligence in

paleoclimate is severely limited by the lack of journal policies or traditions requiring contributors to promptly archive data and methods. *King [1995]* has excellent comments on replication. In this respect, paleoclimate journal editors should consider changes taking place at some prominent economics journals. For example the *American Economic Review* now requires, as a precondition of publication, archiving data and computational code at the journal. This is a response to the critique of *McCullough and Vinod [2003],* and earlier work by *Dewald et al. [1986].* The files associated with paleoclimate studies are trivial to archive. In our view, if the public archive does not permit the replication of a multiproxy study, then it should be proscribed for use in policy formation *[McCullough and Vinod, 2003].*

In addition, we are struck by the lack of policy both in paleoclimate publications and in climate policy reports (e.g. IPCC, ACIA) regarding the reporting of results adverse to their claims. While it may be assumed that results adverse to their claims would be generally disclosed, we are unaware of any paleoclimate journal which explicitly articulates this as a requirement to authors. In contrast, for a prospectus offering securities to the public, officers and directors are required to affirm that the prospectus contains "full, true and plain disclosure", which requires the disclosure of material adverse results. In MBH98, there are a number of examples, where results adverse to their claims were not reported (and in some cases, actual misrepresentations), as listed below (most of which we have discussed passim above):

- MBH98 did not report the results adverse to their conclusions from calculations excluding bristlecone pines (contained in the BACKTO_1400-CENSORED directory).
- For steps prior to 1820, MBH98 did not report verification statistics other than the RE statistic. Unlike the above case, we cannot prove on the present record that Mann et al. had calculated these other statistics, but we consider it quite likely that these statistics were calculated and not reported. (In this case, we believe that diligent referees, even under the limited scope and mandate of journal peer review, should have requested the reporting of this information.)
- MBH98 did not report results from calculations using archived Gaspé tree ring data (which did not contain the extrapolation of early values). Again, while we cannot prove that they actually carried out calculations using the archived version, we find it inconceivable that this unique extrapolation would have been made without previously doing a calculation using the archived version. Although the Corrigendum (six years after the event) disclosed the existence of this extrapolation, it did not disclose its uniqueness or the actual effect of this previously undisclosed extrapolation, disclosure which we believe to be essential for full disclosure, since the very existence of the extrapolation had been hidden from referees and previous readers by a misrepresentation of the start date of this series.
- MBH98 incorrectly stated that conventional PC methods were used, which necessarily means centered calculations. This error in their prior disclosure should have been prominently disclosed in the Corrigendum together with its effects on PC calculations described, especially since it was at the heart of our submission then

under review at *Nature*. Mann et al. could then try to argue in that context that the effect was limited (an argument with which we obviously disagree). Instead, the prior incorrect disclosure was not mentioned at all in the printed Corrigendum; in the Corrigendum SI, the incorrect prior disclosure is not specifically mentioned; the method itself is acknowledged, but it is not prominent and even carries a denial that the method made any difference (a claim discussed at length above).

• The aggressive claims that MBH98 methods were "robust" (see discussion above) are extremely problematic. As noted above, Mann et al. had carried out a sensitivity study on the exclusion of the bristlecone pines and knew that their 15th century results were not robust to these sites. We also believe that they knew the instability regarding the Gaspé series (or else they wouldn't have done the extrapolation.) We find it difficult to understand how the claims to robustness could have made under these circumstances.

We are also struck by the extremely limited extent of due diligence involved in peer review as carried out by paleoclimate journals, as compared with the level of due diligence involved in auditing financial statements or carrying out a feasibility study in mineral development. For example, "peer review" in even the most eminent paleoclimate publications, as presently practiced, does not typically involve any examination of data, replication of calculations or ensuring that data and computational procedures are archived. We are not suggesting peer reviewers should be auditors. Referees are not compensated for their efforts and journals would not be able to get unpaid peer reviewers to carry out thorough audits. We ourselves do not have explicit recommendations on resolving this problem, although ensuring the archiving of code and data as used is an obvious and inexpensive way of mitigating the problem.

But it seems self-evident to us that, recognizing the limited due diligence of paleoclimate journal peer review, it would have been prudent for someone to have actually checked MBH98 data and methods against original data before adopting MBH98 results in the main IPCC promotional graphics..

REFERENCES

Archambault, S., and Bergeron Y., 1992. An 802-year chronology from the Quebec boreal forest. *Canadian Journal of Forest Research*, Vol. 22, pp. 674–682.

Arctic Climate Impact Assessment, 2004. *Impacts of a Warming Arctic: Arctic Climate Impact Assessment*, Cambridge University Press.

Allen, C. D., J. L. Betancourt, and T. W. Swetnam. 1998. *Landscape Changes in the Southwestern United States: Techniques, Long-term Data Sets, and Trends*. Pages 71-84, In T. Sisk, editor, *Perspectives on the Land Use History of North America: A Context for Understanding our Changing Environment*. U.S. Geological Survey, Biological Science Report USGS/BRD/BSR-1998-0003. 104 pp

Allen, C.D. 1998. Where have all the grasslands gone? Quivera Coalition Newsletter, Spring/Summer. Retrieved from Grahame, John D. and Thomas D. Sisk, editors. 2002. Canyons, cultures and environmental change: An introduction to the land-use history of the Colorado Plateau at http://www.cpluhna.nau.edu/Research/grasslands1.htm

Beasley, R.S. and J.G. Klemmedson, Recognizing Site Adversity and Drought-Sensitive Trees in Stands of Bristlecone Pine (pinus longaeva), *Economic Botany* Vol. 27, pp. 141–146.

Billings, W.D. and J.H. Thompson, 1957. Composition of a stand of old bristlecone pines in the White Mountains of California*, Ecology*, Vol. 28, pp. 158–160.

Bradley, R.S., and Jones P.D., `Little Ice Age' summer temperature variations: their nature and relevance to recent global warming trends, *The Holocene,* Vol. 3, pp. 367–376, 1993.

Briffa, K.R. and T.J. Osborn, 1999. Climate Warming: Seeing the Wood from the Trees, *Science* Vol. 284, p. 926.

Briffa, K.R., Jones, P.D. and Schweingruber, F.H., 1992a. Tree-ring density reconstructions of summer temperature patterns across western North America since A.D.1600, *Journal of Climate* Vol. 5, pp. 735–754.

Briffa, K.R., Jones, P.D., Bartholin, T.S., Eckstein, D., Schweingruber, F.H, Karlen, W., Zetterberg, P. and Eronen, M., 1992b. Fennoscandian summers from A.D.500: temperature changes on short and long timescales. *Climate Dynamics* Vol. 7, pp. 111–119.

Briffa, K.R., Jones, P.D., Schweingruber, F.H., Shiyatov, S.G. and Cook, E.R., 1995. Unusual twentieth-century summer warmth in a 1,000-year temperature record from Siberia. *Nature* Vol. 376, pp. 156–159.

Briffa, K.R., T. J. Osborn, F.H. Schweingruber, I.C. Harris, P. D. Jones, S.G. Shiyatov, and E.A. Vaganov, 2001. http://www.ngdc.noaa.gov/paleo/pubs/briffa2001/fig1.jpg" Low-frequency Temperature Variations from a Northern Tree Ring Density Network, *Journal of Geophysical Research,* 106 D3, pp. 2929–2941.

Brooks, P.D., M.W. Williams, and S.K. Schmidt, 1996. Microbial activity under alpine snowpacks. *Biogeochemistry* Vol. 32, pp. 93–113.

Bunn, A.G., R.L. Lawrence, G.J. Bellante, L.A. Waggoner, and L.J. Graumlich, Spatial variation in distribution and growth patterns of old growth strip-bark pines. *Arctic, Antarctic, and Alpine Research*, 2003, Vol. 35, pp. 323–330.

Connor, Kristina F. and Ronald M. Lanner, 1991. Effects of tree age on pollen, seed, and seedling characteristics in Great Basin bristlecone pine. *Botanical Gazette* Vol. 152, pp. 107–113.

Cook, E. R., Bird, T., Peterson, M., Barbetti, M., Buckley, B., D'Arrigo, R., Francey, R., and Tans, P. 1991. Climatic change in Tasmania inferred from a 1089-year tree-ring chronology of subalpine huon pine. *Science* Vol. 253, pp. 1266–1268.

Cook, E. R., Bird, T., Peterson, M., Barbetti, M., Buckley, B., D'Arrigo, R. and Francey, R. 1992. Climatic change over the last millennium in Tasmania reconstructed from tree rings. *The Holocene* Vol. 2, No. 3, pp. 205–217.

Cook, E.R., Briffa, K.R. and Jones, P.D. 1994. Spatial regression methods in

dendroclimatology: a review and comparison of two techniques. *International Journal of Climatology* Vol. 14, pp. 379–402.

Cook, E.R. and Peters, K. 1997. Calculating unbiased tree-ring indices for the study of climatic and environmental change. *The Holocene* Vol. 7, No. 3, pp. 359–368.

Crowley, T.J. and Lowery, T.S., 2000. How warm was the Medieval warm period? *Ambio* Vol. 29, pp. 51–54.

Dewald, W.G., J.G. Thursby, and R.G. Anderson, 1986. Replication in empirical economics: The journal of money, credit and banking project. *American Economic Review*, Vol. 76, No. 4, pp. 587–603.

D'Arrigo, R.D., and Jacoby, G.C., 1991, A 1000-year record of winter precipitation from northwestern New Mexico, USA; A reconstruction from tree-rings and its relation to El Niño and the Southern Oscillation: *The Holocene*, Vol. 1, pp. 95–101.

D'Arrigo, R.D. and Jacoby, G.C., 1992 in *Climate Since A.D. 1500* (eds. Bradley, R.S. & Jones, P.D., pp. 246–268, Routledge.

D'Arrigo, R. D., Robert K. Kaufmann, Nicole Davi, Gordon C. Jacoby, Cheryl Laskowski, Ranga B. Myneni and Paolo Cherubini, 2004. Thresholds for warming-induced growth decline at elevational tree line in the Yukon Territory, Canada. *Global Biogeochemical Cycles*, 18, GB3021, doi:10.1029/2004GB002249.

Ertter, B. , 1988. C. A. Purpus: his collecting trips in the Sierra Nevada and Owens Valley, California, 1895—1898. *In:* Plant Biology of Eastern California. Natural History of the White-Inyo Range, Symposium Vol. 2, pp. 303–309. Retrieved from University of California, Berkelyy website at <http://ucjeps.berkeley.edu/Purpus>

Esper, J., Cook, E.R. and Schweingruber, F.H., 2002. Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. *Science* Vol. 295, pp. 2250–2253.

Fritts, Harold C., 1969. Bristlecone pine in the White Mountains of California, growth and ring-width characteristics. *Papers of the Laboratory of Tree-Ring Research* No.4. Tucson: University of Arizona Press

Graumlich, L.J., 1991. Subalpine tree growth, climate, and increasing CO2: an assessment of recent growth trends. *Ecology* Vol. 72, pp. 1–11;

Graybill, D.A., and S.B. Idso. 1993. Detecting the aerial fertilization effect of atmospheric CO_2 enrichment in tree-ring chronologies. *Global Biogeochemical Cycles* Vol. 7, pp. 81–95.

Grissino-Mayer, Henri D., 1996. A 2129 year annual reconstruction of precipitation for northwestern New Mexico, USA. In Dean, J.S., Meko, D.M., and Swetnam, T.W., eds., *Tree Rings, Environment, and Humanity. Radiocarbon 1996*, Department of Geosciences, The University of Arizona, Tucson: pp. 191–204.

Hughes, M. K. and L. J. Graumlich. 1996. Climatic variations and forcing mechanisms of the last 2000 years. Volume 141. Multi-millenial dendroclimatic studies from the western United States. NATO ASI Series, pp. 109–124.

Hughes, M.K. and G. Funkhouser. 2003. Frequency-dependent climate signal in upper and

lower forest border trees in the mountains of the Great Basin. *Climatic Change*: Vol. 59, pp. 233–244

International Panel on Climate Change, 2001. *Climate Change 2001: The Scientific Basis*. Retrieved from http://www.grida.no/climate/ipcc_tar/.

Jacoby and d'Arrigo, 1989. *Climatic Change* Vol. 14, pp. 39–59.

Jacoby, Gordon C. and Rosanne D. D'Arrigo, 1997. Tree rings, carbon dioxide, and climatic change, *Proc. Natl. Acad. Sci. USA* 94, 8350-8353.

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

Jones, P. D., Briffa, K. R., Barnett, T. P. and Tett, S. F. B., 1998. High-resolution palaeoclimatic records for the last millennium; interpretation, integration and comparison with general circulation model control-run temperatures, The Holocene, Vol. 8, pp. 455–471.

Kelly, P.E., Cook, E.R., and Larson, D.W. 1992. Constrained growth, cambial mortality, and dendrochronology of ancient *Thuja occidentalis* on cliffs of the Niagara Escarpment: an eastern version of bristlecone pine? *Int. J. Pl. Sci*. Vol. 153, pp. 117–127.

Kelly, P.E., Cook, E.R., and Larson, D.W. 1994. A 1397-yr tree-ring chronology of *Thuja occidentalis* from cliff-faces of the Niagara Escarpment, southern Ontario, Canada. *Can. J. Forest Res*. Vol. 24, pp. 1049–1057.

King, Gary, 1995. Replication, Replication, *PS: Political Science and Politics*, with comments from nineteen authors and a response, A Revised Proposal, Proposal, Vol. 28, No. 3, pp. 443–499.

LaMarche, VC, Jr., 1973, Holocene climatic variations inferred from treeline fluctuations in the White Mountains, California. *Quaternary Research* Vol. 3, pp. 632–660.

LaMarche, V.C. and H.A. Mooney, 1967. Altithermal Timberline Advance in Western United States, *Nature* Vol. 213, pp. 980–982

LaMarche, V.C. and Stockton, C.W. 1974. Chronologies from temperature-sensitive bristlecone pines at upper treeline in western United States. *Tree-Ring Bulletin* Vol. 34, pp. 21–45.

LaMarche, V.C., D.A. Graybill, H.C. Fritts, and M.R. Rose., 1984. Increasing atmospheric carbon dioxide: tree ring evidence for growth enhancement in natural vegetation. *Science Vol.* 225, pp. 1019–1021.

Lanner, R.M. and K.F.Connor, 2001. Does Bristlecone Pine Senesce? *Experimental Gerontology,* Vol. 36, pp. 675–685.

Mann, Michael (2002-2004). MBH98 Database, ftp://holocene.evsc.virginia.edu/pub/MBH98.

Mann, M.E., Jones, P.D., Global surface temperature over the past two millennia, Geophysical *Research Letters*, Vol. 30, No. 15, p. 1820, doi: 10.1029/2003GL017814, 2003.

Mann, M.E., Bradley, R.S. and Hughes, M.K., 1998. Global-Scale Temperature Patterns and Climate Forcing Over the Past Six Centuries, Nature, Vol. 392, pp. 779–787.

Mann, M.E., Bradley, R.S. and Hughes, M.K., Northern Hemisphere Temperatures During the Past Millennium: Inferences, Uncertainties, and Limitations, Geophysical Research

Letters, Vol. 26, pp. 759–762, 1999.

Mann, M.E., E. Gille, R.S. Bradley, M.K. Hughes, J.T. Overpeck, F.T. Keimig, and W. Gross. 2000. Global temperature patterns in past centuries: An interactive presentation. *Earth Interactions* Vol. 4-4, pp. 1–29. Retrieved from NOAA website at http://www.ngdc.noaa.gov/paleo/ei, which includes additional note http://www.ngdc.noaa.gov/paleo/ei/ei_nodendro.html.

Mann, M.E., Bradley, R.S. and Hughes, M.K., 2003. Note on Paper by McIntyre and McKitrick in "Energy And Environment". Retrieved from <ftp://holocene.evsc.virginia.edu/pub/mann/EandEPaperProblem.pdf>

Mann, M.E., Bradley, R.S. and Hughes, M.K., 2004a. Reply to: Global-scale temperature patterns and climate forcings over the past six centuries: a comment, by S. McIntyre and R. McKitrick. Retrieved from website of Stephen Schneider <http://stephenschneider.stanford.edu/Publications/PDF_Papers/MannEtAl2004.pdf> .

Mann, M.E., Bradley, R.S. and Hughes, M.K., 2004b. Reply to: Global-scale temperature patterns and climate forcings over the past six centuries: a comment, by S. McIntyre and R. McKitrick, unpublished.

Mann, M.E., Bradley, R.S. and Hughes, M.K., 2004c. Corrigendum: Global-scale temperature patterns and climate forcing over the past six centuries, *Nature* Vol. 430, No. 105 (2004).

Mann, M.E., Bradley, R.S. and Hughes, M.K., 2004d. False Claims by McIntyre and McKitrick regarding the Mann *et al.* (1998) reconstruction. Retrieved from website of realclimate.org at < http://www.realclimate.org/index.php?p=8>

Matthes-Sears, U. and D.W. Larson, 1990. Environmental controls of carbon dioxide uptake in two woody species with constrasting distributions at the edge of cliffs. *Can. J. Bot.* Vol. 68, pp. 2371–2380.

McCullough, B.D. and H. D. Vinod, 2003. Verifying the Solution from a Nonlinear Solver: A Case Study, *American Economic Review* Vol. 93, pp. 873–892 with comments and replies (2004) at *American Economic Review* Vol. 94, pp. 391–403.

McIntyre, S. and R. McKitrick, 2003. "Corrections to the Mann et. al. (1998) Proxy Data Base and Northern Hemispheric Average Temperature Series" *Energy and Environment* Vol. 14, pp. 751–771.

McIntyre, S. and R. McKitrick, 2004a. Global-scale temperature patterns and climate forcings over the past six centuries: a comment. . Retrieved from http://www.uoguelph.ca/~rmckitri/research/fallupdate04/submission.1.final.pdf

McIntyre, S. and R. McKitrick, 2004b. Global-scale temperature patterns and climate forcings over the past six centuries: a comment. Retrieved from http://www.uoguelph.ca/~rmckitri/research/fallupdate04/MM.resub.pdf

McIntyre, S. and R. McKitrick, 2005. Hockey Sticks, Principal Components and Spurious Significance. In press, *Geophysical Research Letters.*

Mooney, H.A., M West and R. Brayton, 1966. Field measurements of the metabolic responses of bristlecone pine and big sagebrush in the White Mountains of California, *Botanical*

Gazette Vol. 127, pp. 105–113.

Mooney, H.A., R.D. Wright and B.R. Strain, 1964. The Gas Exchange Capacity of Plants in relation to vegetation zonation in the White Mountains of California, *Am. Mid Nat.* Vol. 72, p. 281.

Muir, John, 1911. My First Summer in the Sierra. *Atlantic Monthly.* Retrieved from <http://www.sierraclub.org/john_muir_exhibit/writings/my_first_summer_in_the_sierra/>

Muller, Richard, 2004. Global Warming Bombshell: A prime piece of evidence linking human activity to climate change turns out to be an artifact of poor mathematics. *MIT Technology Review* Retrieved from *http://www.technologyreview.com/articles/04/10/ wo_muller101504.asp*

R Development Core Team (2004), R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-00-3, <http://www.R-project.org>.

Schauer, A.J., A.W. Schoettle, and R.L. Boyce, 2001. Partial cambial mortality in high-elevation Pinus aristata (Pinaceae). *American Journal of Botany* Vol. 88, pp. 646–652.

Schoettle, A.W., 2004. Ecological Roles of Five-Needle Pines in Colorado: Potential Consequences of Their Loss, in Sniezko, Richard, Safiya Samman, Scott E. Schlarbaum and Howard B. Kriebel, eds. 2004. Breeding and genetic resources of five-needle pines: growth adaptability and past resistance; 2001 July 23–27; Medford OR, ISA, IUFRO Working Paper 2.01.15. Proceedings RMRS-P-32, Fort Collins CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Schulze, E.D., H.A. Mooney and E.L. Dunn, 1967. Wintertime photosynthesis of bristlecone pine (pinus aristata) in the White Mountains of California. *Ecology* Vol. 48, pp. 1044–1047.

Sheppard, Paul R. and Edward R. Cook, 1988. Scientific Value of Trees in Old-Growth Areas. *Natural Areas Journal,* Vol. 8, No. 1, pp. 7–12.

Soon, W., Baliunas, S., Idso, C., Idso, S. and Legates, D.R., 2003. Reconstructing climatic and environmental changes of the past 1000 years: a reappraisal. *Energy & Environment* Vol. 14, pp. 233–296.

St. Andre, G., H.A. Mooney and R. D. Wright, 1967. The Pinyon Woodland Zone in the White Mountains of California. *Am. Mid. Nat* Vol. 73, pp. 225–239

Storch, H. von, E. Zorita, J. M. Jones, Y. Dmitriev and S. F. B. Tett, 2004. Reconstructing past climate from noisy data. *Science* Vol. 306, pp. 679–682.

Wigley, T.M.L., Briffa, K.R. and Jones, P.D., 1984. On the average value of correlated time series with applications in dendroclimatology and hydrometeorology. *Journal of Climate and Applied Meteorology* Vol. 23, pp. 201–213.

Williams, M., 1996. Hot Science Soundbites for the 1996 Niwot Ridge LTER Workshop. Retrieved from <http://culter.colorado.edu:1030/Niwot/Workshops/96soundbites.html>.

Wright, R.D. and H.A. Mooney, 1965, Substrate-oriented distribution of bristlecone pine in the White Mountains of California. *Am. Mid. Nat.* Vol. 73, pp. 257–184.

E. Zorita, F. González-Rouco and S. Legutke, 2003. 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 . *Journal of Climate* Vol. 16, pp. 1378–1390.

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