

ENERGY & ENVIRONMENT

VOLUME 14 · NUMBER 6 · 2003

ISSN 0958-305X

©2003

MULTI-SCIENCE PUBLISHING CO. LTD.
5 Wates Way, Brentwood, Essex CM15 9TB, United Kingdom

Energy & Environment is abstracted/indexed in: Engineering Index, Pollution Abstracts, Geographical Abstracts, Environmental Abstracts, Biomass Bulletin and International Development Abstracts.

Energy & Environment is now available on-line, free to subscribers to the printed edition. For access details contact Multi-Science.

Email: mscience@globalnet.co.uk

Website: www.multi-science.co.uk

CORRECTIONS TO THE MANN et. al. (1998) PROXY DATA BASE AND NORTHERN HEMISPHERIC AVERAGE TEMPERATURE SERIES

Stephen McIntyre

512-120 Adelaide St. West, Toronto, Ontario Canada M5H 1T1;

Ross McKittrick

Department of Economics, University of Guelph, Guelph Ontario Canada N1G2W1.

ABSTRACT

The data set of proxies of past climate used in Mann, Bradley and Hughes (1998, “MBH98” hereafter) for the estimation of temperatures from 1400 to 1980 contains collation errors, unjustifiable truncation or extrapolation of source data, obsolete data, geographical location errors, incorrect calculation of principal components and other quality control defects. We detail these errors and defects. We then apply MBH98 methodology to the construction of a Northern Hemisphere average temperature index for the 1400-1980 period, using corrected and updated source data. The major finding is that the values in the early 15th century exceed any values in the 20th century. The particular “hockey stick” shape derived in the MBH98 proxy construction – a temperature index that decreases slightly between the early 15th century and early 20th century and then increases dramatically up to 1980 — is primarily an artefact of poor data handling, obsolete data and incorrect calculation of principal components.

Correspondence etc. should be directed to smcintyre25@yahoo.ca

Key words: hockey stick, multiproxy method, global temperature history, IPCC, climate change, data quality.

ACKNOWLEDGMENTS:

Helpful comments and encouragement were received from R. Carter, R. Courtney, D. Douglass, H. Erren, C. Essex, W. Kininmonth, T. Landscheidt and referees and assistance from the World Data Center for Paleoclimatology. McIntyre thanks Professor Michael Mann for supplying the data sets and information necessary for this analysis. All remaining errors are ours. No funding from any source was sought or received for this research.

1. INTRODUCTION.

In a widely cited paper, Mann, Bradley and Hughes (1998, hereafter MBH98) constructed a temperature history of the Northern Hemisphere for the period 1400-1980. The result

was the well-known “hockey stick”-shaped graph suggesting that the climate of the late 20th century is unusual compared to the centuries preceding it. This temperature history was extended to the period 1000-1399 in Mann, Bradley and Hughes (1999), who claimed that “temperatures in the latter half of the 20th century were unprecedented” and that “even the warmer intervals in the reconstruction pale in comparison with mid-to-late 20th-century temperatures”. The temperature history was given bold prominence by the Intergovernmental Panel on Climate Change (2001) where it appears in Figures 2-20 and 2-21 in Chapter 2 of the Working Group 1 *Assessment Report*, Figure 1b in the Working Group 1 *Summary for Policymakers*, Figure 5 in the *Technical Summary*, and Figures 2-3 and 9-1B in the *Synthesis Report*. Referring to this figure, the IPCC *Summary for Policymakers* (p. 3) claimed it is likely “that the 1990s has been the warmest decade and 1998 the warmest year of the millennium” for the Northern Hemisphere. The IPCC view of temperature history has in turn been widely disseminated by governments and used to support major policy decisions.¹

MBH98 applied 112 proxies and historical temperature measurements in what they called a “multiproxy approach” to construction² of a temperature index from 1400 to 1980. Although the “multiproxy” approach was apparently a novelty within the climatological community, the same algebraic and statistical methods are commonly used in economics, business and elsewhere in the social sciences, though the terminology differs from discipline to discipline.³

Upon request, Professor Mann instructed an associate to supply the collated proxy set, together with applicable weights, to the first author. When attempting to replicate MBH98 principal component (PC) calculations, an extremely low (6%) explained variance for those in the Texas-Mexico dataset was noticed, leading to a close examination of the data collation. Anomalous start years (see details below) were noticed and it was verified that these occurred only in MBH98 data and were not due to collation errors on our part. Explained variance improved significantly by moving the MBH98 data one year later, confirming that an MBH98 collation error had almost certainly occurred. We then noticed copy errors in the 1980 values for these series and stretches of identical values in other places in the database. This led to a systematic comparison of MBH98 data to original data, identifying obsolete versions and undisclosed truncation of time series. Independent calculations of the proxy principal components convinced us that those in MBH98 were erroneous we updated and corrected the database and then applied MBH98 methodology, as publicly disclosed, to construct a temperature index from 1400 to 1980. The newly calculated temperature index (see Figure 7) contradicts the MBH98 assertion of late 20th century uniqueness. We find that the particular “hockey stick” shape derived by MBH98 is primarily an artefact of poor data handling and use of obsolete proxy records.

¹See, for instance, the Government of Canada web site http://www.climatechange.gc.ca/english/issues/what_is/index.shtml.

²MBH98 refers to the index resulting from their calculation as a “reconstruction.” This is a misnomer since it is a novel index, rather than the recomputation of something previously observed. Therefore it will be referred to herein as “construction.”

³For a critique of applying stationary linear maps to nonstationary phenomena like climate see Essex and McKittrick (2002) chapter 5.

2. ERRORS AND DEFECTS IN THE MBH98 PROXY DATA BASE

The term “proxy” denotes some physical data or measurement that can potentially serve as an indirect record of local temperature conditions, including tree ring widths and densities, coral $\delta^{18}\text{O}$, $\delta^{13}\text{C}$ and calcification rates, ice core $\delta^{18}\text{O}$, melt percentages and so forth. Thirteen MBH98 series are based on instrumental temperature records and are not, strictly speaking, “proxies”. For consistency with MBH98, we will use the term “proxy” to include these series.

We will denote a proxy series with the prefix ‘#’, i.e. proxies #1—#112. Twenty-two of the 112 proxies date back as far as 1400, while all 112 are available as of 1820. Twenty-three MBH98 proxies cease to be available in the 1970s. Thirty-one of the 112 proxies are principal components (PCs) from tree ring datasets, of which 28 were PCs calculated by MBH98 themselves from 300 tree ring datasets. Three are PCs from 14 Texas-Oklahoma sites, 9 are PCs from 20 Texas-Mexico sites, 9 are PCs from 232 International Tree Ring Data Base (ITRDB) US/Canada tree ring sites, 3 are PCs from 18 South American sites and 4 are from 16 Australian sites. Inconsistently, individual US, Canadian and Mexican tree ring sites are included separately in the list of 112 proxies rather than being incorporated into the PCs for that area (see e.g. Appendix, #49, #51-61, #106.) More information about the proxies is available at the Supplementary Information web site (see Appendix).

The database used by MBH98 contains the errors and defects listed below. We detail each of these points in this section, then in Section 3 we show how correcting these errors and defects affects the calculation of the Northern Hemisphere average temperature index using MBH98 methodology.

- (a) unjustified truncation of 3 series;
- (b) copying 1980 values from one series onto other series, resulting in incorrect values in at least 13 series;
- (c) displacement of 18 series to one year earlier than apparently intended;
- (d) unjustified extrapolations or interpolations to cover missing entries in 19 series;
- (e) geographical mislocations and missing identifiers of location;
- (f) inconsistent use of seasonal temperature data where annual data are available;
- (g) obsolete data in at least 24 series, some of which may have been already obsolete at the time of the MBH98 calculations;
- (h) listing of unused proxies;
- (i) incorrect calculation of all 28 tree ring principal components.

(a,f) Series #10 and #11 (Central England and Central Europe air temperatures respectively) use June-July-August averages. This raises three concerns: annual data were available in the primary sources; other station temperature series used by MBH98 (#21- #31), where identified, are annual; and MBH98 claims to calculate an annual temperature index. The Central England Temperature series is truncated at 1730 rather than the available 1659 in source data, which removes a major late 17th century cold period (see Supplementary Information). Series #10 has a 1987 value which is 0.43 deg C higher than in the source data though this does not appear to affect any calculations discussed herein. Central Europe (#11) is truncated at 1550 rather

than the available 1525, which removes the warmest temperatures in the series (compare Figure 1 Top and Bottom panels). #11, which is an exceptionally long series of direct temperature information, also shows a notable lack of 20th century uniqueness. In series #100, MBH98 also crop two very high values from the start of the series. These truncations are not justified and were not disclosed by MBH98.

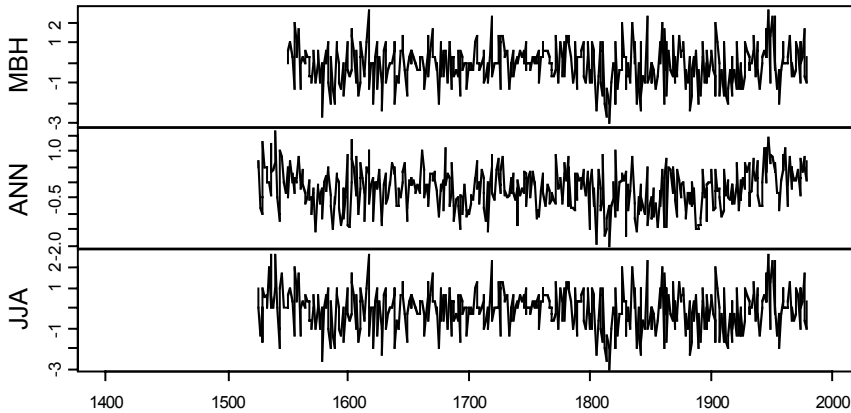


Figure 1. Temperature anomalies (C) by year for (Top panel) Central European historical air temperatures, MBH98 series #11; those differ from (Middle panel) Central European historical (Annual) from 1525-1979, which would be more relevant for inclusion in the calculation of an annual index; (Bottom panel) Central European historical (June-July-August) from 1525-1979, which matches, in the period of overlap, the incomplete record used by MBH98.

(b,c) In the MBH98 collated data set the 1980 values for series #72-#80, which are the 9 Texas-Mexico principal components computed by MBH98, are identical to 7 decimal places, an obviously impossible result (see Table 1) and therefore an error.

The 1980 values are likewise identical in the 3 Vaganov principal components (series #81-#83) and 4 of the 9 ITRDB US principal components computed by MBH98 (series #84, #90, #91 and #92); see Table 2. Interestingly, all but two of these series as collated in the MBH98 database begin in years ending in *99 or *49, rather than the apparently intended *00 and *50, and appear to have been displaced one year backward in collation. This suggests a simple clerical error, in which the series in question were copied into a file at the wrong row, then a 1980 value was filled in from an adjacent cell. Series #85-#89 commence in 1499, but lack the telltale 1980 value. The displacements will result in any extreme year in the past being one year off in these datasets, attenuating its effect in the compilation. The copy errors constitute a significant fraction of the MBH98 dataset for 1980 – the final year of the MBH98 proxy-constructed index.⁴

⁴To access the underlying data consult the supplementary information sources listed in the Appendix.

Table 1. Identical 1980 values (bold) in Texas-Mexico series in MBH98 data set.

Directory: TREE/STAHL/SWM/BACKTO_1700

MBH Series #:	#72	#73	#74	#75	#76	#77	#78	#79	#80
Record Name:	pc01.out	pc02.out	pc03.out	pc04.out	pc05.out	pc06.out	pc07.out	pc08.out	pc09.out
1976	-0.04758900	0.09825240	-0.01345320	0.01161880	0.01822490	0.03648180	0.04604640	-0.04273910	0.00526230
1977	0.02738590	-0.11581500	0.02995960	0.01370230	0.03782570	0.00327476	0.07170230	0.03729640	-0.10195200
1978	0.09249040	-0.00125138	0.08667150	0.07659540	0.02200060	0.04614070	0.03223540	0.02464170	0.02726110
1979	-0.01054950	-0.17253000	-0.00999568	-0.04078750	0.09144420	-0.00608904	-0.00508424	-0.03537360	-0.08408310
1980	0.02303040	0.02303040	0.02303040	0.02303040	0.02303040	0.02303040	0.02303040	0.02303040	0.02303040

Table 2. Identical 1980 values (bold) in Vaganov PCs and 4 ITRDB PCs in MBH98 data set.

Directory: TREE/VAGANOV/BACKTO_1750 TREE/ITRDB/NOAMER/BACKTO_1750

MBH Series #:	#81	#82	#83	#84	#90	#91	#92
Record Name:	pc01.out	pc02.out	pc03.out	pc01.out	pc07.out	pc08.out	pc09.out
1976	-0.03291460	0.02597240	0.01647480	0.00888690	0.10327000	0.18590800	0.03833640
1977	0.01170380	-0.09346030	-0.01559880	0.02153980	0.10721500	0.13304000	0.00842804
1978	0.05759400	-0.02107700	-0.13369700	0.02241670	0.11963800	0.12615900	0.01346540
1979	-0.11100000	-0.07345260	0.02438820	0.04898920	0.13681500	0.16866300	0.02801240
1980	-0.04063530	-0.04063530	-0.04063530	0.04345260	0.04345260	0.04345260	0.04345260

Table 3. Filled data series (bold) in MBH98 data set. NaN denotes a missing value, although values beyond 1980 are not relevant to the calibration interval.

Directory: TREE/MANNETAL97	45	46	51	52	54	56	58
MBH98 Series #:	cpatagonia.dat	mpatagonia.dat	treeline1.dat	treeline10.dat	treeline2.dat	treeline4.dat	treeline6.dat
Record Name:	pc01.out	pc02.out	pc03.out	pc01.out	pc02.out	pc03.out	pc04.out
1970	16.79999900	0.85000002	1.27000000	0.61400002	1.10500000	0.99900001	1.35900000
1971	15.50000000	-1.10000000	1.40900000	0.46000001	1.41199990	1.42200010	1.30300000
1972	15.90000000	0.17000000	1.25700000	0.83399999	1.38800000	1.22200000	1.38800000
1973	15.50000000	0.63999999	1.10700000	0.56199998	1.19700000	1.07100000	1.46000000
1974	16.70000100	-0.43000001	1.13300000	1.10400000	1.14400010	1.13500000	1.62899990
1975	14.90000000	-0.43000001	0.93199998	1.10400000	1.36600010	1.22400000	1.61300000
1976	14.70000000	-0.43000001	1.16100000	1.10400000	1.36600010	1.22400000	1.17600000
1977	16.10000000	-0.43000001	1.58500000	1.10400000	1.36600010	1.22400000	1.57300000
1978	15.10000000	-0.43000001	1.58500000	1.10400000	1.36600010	1.22400000	1.57300000
1979	15.10000000	-0.43000001	1.58500000	1.10400000	1.36600010	1.22400000	1.57300000
1980	15.10000000	-0.43000001	1.58500000	1.10400000	1.36600010	1.22400000	1.57300000
1981	15.10000000	NaN	NaN	NaN	NaN	NaN	NaN
1982	15.10000000	NaN	NaN	NaN	NaN	NaN	NaN

Table 4. Filled data (bold) in MBH98 data.

Directory: TREE/ITRDB/SOAMER/BACKTO_1600	93	94	95	96	97	98	99
MBH98 Series #:	pc01.out	pc02.out	pc03.out	pc01.out	pc02.out	pc03.out	pc04.out
Record Name:	pc01.out	pc02.out	pc03.out	pc01.out	pc02.out	pc03.out	pc04.out
1970	-0.05519220	0.03191820	-0.00840994	0.07964660	-0.03334510	0.00628749	0.03972250
1971	0.04456720	0.02654390	0.06869810	0.05834090	0.03159730	0.01390980	-0.00292208
1972	-0.03087120	0.03992700	0.00302668	0.15582700	0.07014980	0.03358720	-0.07598700
1973	-0.02466770	0.11485700	-0.05301170	0.18438500	0.04514380	-0.04919020	-0.05758820
1974	0.03531060	0.07091270	0.00376018	0.11299600	0.01402680	-0.00682486	-0.10635300
1975	0.04918980	0.07842340	-0.02821910	0.16178501	0.02186560	0.02133480	0.00791537
1976	0.04792550	0.07830090	-0.02856930	0.16103400	0.08604140	0.05941720	-0.05302600
1977	0.04792550	0.07830090	-0.02856930	0.16103400	0.08604140	0.05941720	-0.05302600
1978	0.04792550	0.07830090	-0.02856930	0.16103400	0.08604140	0.05941720	-0.05302600
1979	0.04792550	0.07830090	-0.02856930	0.16103400	0.08604140	0.05941720	-0.05302600
1980	0.04792550	0.07830090	-0.02856930	0.16103400	0.08604140	0.05941720	-0.05302600

(d) MBH98 insert extrapolated, interpolated or copied values during the critical calibration period into 19 series. We refer to these as “fills” hereafter. In the data set provided to the authors, the following 17 series contain end-of-sample fills for one or more years including 1980: #6, #45, #46, #50-#52, #54-#56, #58, #93-#99. Series #53 was filled for 4 years at its beginning and series #3 for 16 years in the calibration period. In the case of #3, MBH98 inexplicably replaced available source values for 1962-64 with filled values. For examples see Tables 3, 4 and 5.

Table 5. Filled series (bold) in data for MBH98 PCs.

MBH98 Series #:	51	54	56	58	53
Record Name:	ak031	ak032	cana157	cana153	cana036
1400	NA	NA	NA	NA	0.723
1401	NA	NA	NA	NA	0.723
1402	NA	NA	NA	NA	0.723
1403	NA	NA	NA	NA	0.723
1404	NA	NA	NA	NA	0.723
1405	NA	NA	NA	NA	0.874
1406	NA	NA	NA	NA	1.026
1407	NA	NA	NA	NA	1.029
1408	NA	NA	NA	NA	1.203
1409	NA	NA	NA	NA	1.055
1970	1.270	1.105	0.999	1.359	1.376
1971	1.409	1.412	1.422	1.303	1.554
1972	1.257	1.388	1.222	1.388	1.463
1973	1.107	1.197	1.071	1.460	1.618
1974	1.133	1.144	1.135	1.629	1.483
1975	0.932	1.366	1.224	1.613	1.743
1976	1.161	1.366	1.224	1.176	1.577
1977	1.585	1.366	1.224	1.573	1.583
1978	1.585	1.366	1.224	1.573	1.851
1979	1.585	1.366	1.224	1.573	1.618
1980	1.585	1.366	1.224	1.573	2.204

Series #50 is especially noteworthy. The values of series #50 for the entire period from 1962 to 1982 are copied from series #49 (see Table 6). Although MBH98 attribute both series #49 and #50 to Fritts and Shao (1992), series #49 is actually derived from Briffa et al. (1992).

Table 6. Final 20 values in MBH98 Series #49 and #50 are identical.

Directory: TREE/MANNETAL97		
MBH98 Series #:	49	50
Record Name:	trd.dat	trw.dat
1958	0.38000000	0.34000000
1959	-0.15000000	0.45000000
1960	0.28000000	0.02000000
1961	0.12000000	0.55000010
1962	-0.03999999	-0.03999999
1963	0.60000020	0.60000020
1964	-0.77999970	-0.77999970
1965	-0.80000010	-0.80000010
1966	0.28999999	0.28999999
1967	-0.23000000	-0.23000000
1968	-0.94999999	-0.94999999
1969	0.91000030	0.91000030
1970	0.31999999	0.31999999
1971	0.11000000	0.11000000
1972	-0.02000000	-0.02000000
1973	-0.01000000	-0.01000000
1974	-0.07999998	-0.07999998
1975	-0.68000010	-0.68000010
1976	-0.09000004	-0.09000004
1977	0.15000010	0.15000010
1978	-0.14000000	-0.14000000
1979	0.02000000	0.02000000
1980	-0.23999999	-0.23999999
1981	-0.01000000	-0.01000000
1982	0.05999999	0.05999999

These fills are neither required nor justified statistically and exceed MBH98 disclosure. There is no disclosure of the extent of data filling or its potential impact on the constructed temperature index in the text of the *Nature* article and, their supplementary web page (http://www.ngdc.noaa.gov/paleo/ei/data_supp.html) says only “Small gaps have been interpolated. If records terminate slightly before the end of the 1902-1980 training interval, they are extended by persistence to 1980.” Inconsistently, however, series #11, #102, #103, #104, #106 and #112 terminate prior to 1980 but were not filled in the MBH98 dataset. The fills in 1980 are pervasive: at least 30 (and up to 36) proxies in 1980 have values arising from copy errors or extrapolation.

(e) Geographical mislocations and missing attributions occur in the MBH98 data. For example, MBH98 use 11 precipitation series, for which they cite Jones and Bradley (1992) (hereafter “JB92”). JB92 (Table 13.3) lists 17 precipitation series, of which 12 are digitally published at the World Data Center for Paleoclimatology

(<http://www.ngdc.noaa.gov/paleo/paleo.html>, hereinafter denoted WDCP). In only two MBH98 precipitation series (#35 and #37) did the correlation between JB92 and MBH98 data exceed 0.9, permitting a reasonably secure identification of locations; other correlations were less than 0.5 excluding the possibility of identification. The JB92 series for Paris, France (48.8N, 2.5E) can be identified with MBH98 series #37 both from the high correlation and the identity of starting date (see Figure 2, which graphs both these series). However, MBH98 series #37 is located at the grid-box centred at 42.5N, 72.5W near Boston, Massachusetts.

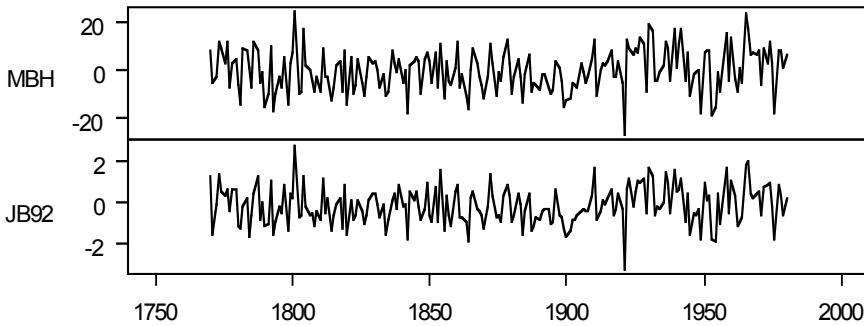


Figure 2. (Top) Station precipitation with erroneous location given as 42.5N, 72.5W in MBH98 (series #37); (Bottom) the record corresponds instead to that of JB92 Paris, France (48.8N, 2.5E) JB92 is scaled to a 1901-1950 reference period (i.e. subtracting the 1901-50 mean and dividing by the 1901-50 standard deviation). MBH98 appear to have a scale error by a factor of 10.

Two MBH98 precipitation series are in India and derive from an unreported source, since no Indian locations are listed in JB92. The other 7 MBH98 precipitation series derive either from unreported sources, from the 5 JB92 series not digitally published at WDCP or have been heavily transformed in collation. Two of the MBH98 temperature grid-box series had no locational counterparts in JB92 (Table 13.1): series #26 (52.5N, 17.5E grid-box) and series #29 (62.5N, 7.5E grid-box). In addition, MBH98 series #20 (Central Greenland ice core) is materially mislocated to the north and west. On comparison with source data, it can be seen that MBH98 have also reversed the geographical locations of series #46 and #47.

(g) Digitally published versions at the World Data Center for Paleoclimatology (WDCP, <http://www.ngdc.noaa.gov/paleo/paleo.html>) supercede the versions used by MBH98 for the following 24 series: #1, #2, #3, #6, #7, #8, #9, #21, #23, #27, #28, #30, #35, #37, #43, #51, #52, #54, #55, #56, #58, #65, #105 and #112. A listing of FTP sources is provided in the Appendix and details for each of the above series, including comparisons of different data editions, is provided in the Supplementary Information. (Since many datasets used by MBH98 remain digitally unpublished, this listing is only from datasets where a comparand was identified.) For the purposes of this study, it is immaterial whether the MBH98 datasets were obsolete as at the time of publication of MBH98 or whether they have become obsolete subsequently. However, at least some

datasets used by MBH98 were already obsolete in 1998. In response to an inquiry about series #51- #61, WDCP confirmed that the updated versions for four of the series were available as early as 1991-1992. [WDCP, pers. comm., Sept. 2003].

In some cases, the differences between MBH98 and updated series were isolated; in other cases, the differences were systematic. As an example of relatively isolated differences, MBH98 series #28 corresponds closely to a Z-transformation (subtracting the mean, dividing by the standard deviation) of the JB92 Leningrad series for most of its history, but there are major and puzzling discrepancies in the 1760s, including a discrepancy of over 4 degrees C in 1764 (see Figure 3). As with the Central European temperature series (and other long temperature series), the 20th century values are not unique.

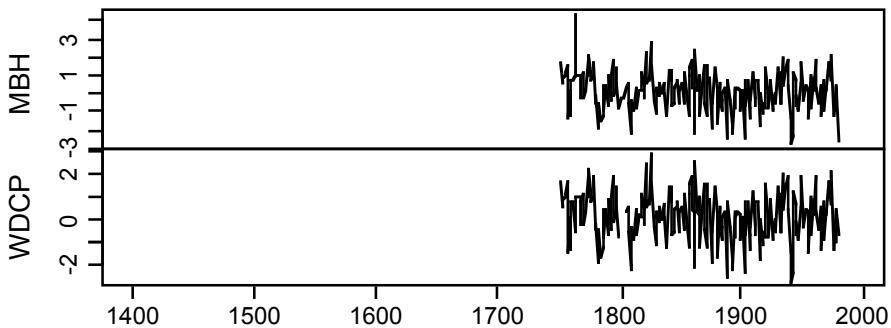


Figure 3. (Top) Station temperature for grid box 75.5N, 32.5E from MBH98 (series #28); (Bottom) Leningrad temperature from JB92 at WDCP (anomaly from 1951-1970).

As an example of systematic differences, MBH98 series #56 (Twisted Tree, Heartrot Hill, a northern treeline ring width series) used an early version of site data with values only up to 1975 and with MBH98 fills from 1976 to 1980. The updated version now at WDCP has data up to 1992 (see Figure 4) and differs quite dramatically from the MBH98 series. The MBH98 version of series #56, like MBH98 versions of many northern treeline series (#51-#58, #60-#61) shows an *increased* ring width index in the 1902-1980 period. However, in the WDCP series, there is a dramatic and sustained *reduction* in ring widths in the 1980s, with a complete reversal of the increases in the first decades of the century. This pattern occurs in other series updated into the 1990s (series #51 and #54) and was apparent by 1984 in the northerly series #59 (Hornby Cabin) (see Supplementary Information). The later edition of #56, presumably for quality control reasons, discontinued some early estimates made in the first edition.

(h) Five series purportedly in the multiproxy network (fran003, ital015, ital015x, spai026 and spai047 in the MBH98 list "ITRDB -Miscellaneous") cannot be located in either the MBH98 collated set or the proxy PC compilations.

(i) Of the 112 proxies in MBH98, 28 are principal components calculated by MBH98 from International Tree Ring Data Base (ITRDB) site chronologies stored at WDCP for the sites listed in MBH98 Supplementary Information (see http://www.ngdc.noaa.gov/paleo/ei/data_supp.html) for the following five different

regions: Texas-Oklahoma, Texas-Mexico, North America, South America and Australia-New Zealand. The principal component calculations have two types of problems: first, MBH98 does not establish consistent rules for inclusion or exclusion of sites in regional aggregates and, second, the MBH98 principal components fail to maximize explained variances.

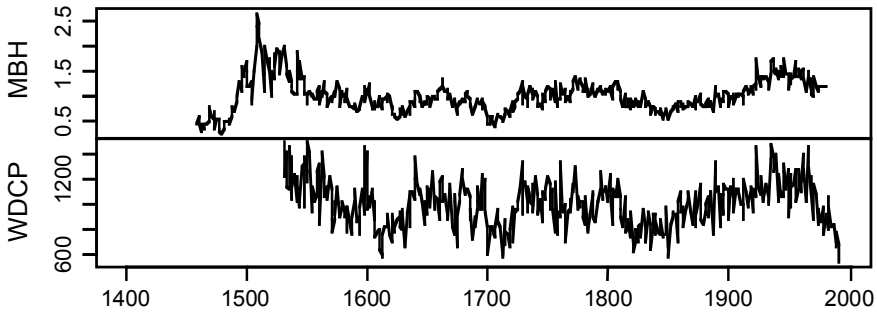


Figure 4. (Top) Twisted Tree, Heartrot Hill (northern treeline) ring width index from MBH98 (series #56) (1459-1975 plus 5 fills at end); (Bottom) Twisted Tree, Heartrot Hill ring width index from WDCP (1530-1992). MBH have divided WDCP values by 1000. Neither series is Z-transformed.

As to the first problem, MBH98 do not provide justification for excluding the Texas-Oklahoma and Texas-Mexico sites from the North American compilation. Similarly puzzling are the occurrences of other sites as individual proxies rather than being incorporated into the regional PC groups. Series #106 occurs within the Texas-Mexico region; series #49-64 are all North American sites or reconstructions; series #46-47 are within the South American region and series #43 and #45 are reconstructions within the Australia-New Zealand region.

The second problem was determined indirectly as the MBH98 principal component calculations are unpublished. We collated the source data from WDCP for all sites listed in MBH98 (except, immaterially, one MBH98 US site which could not be identified in the WDCP database). The collations are available in Supplementary Information. The start dates of the MBH98 PC's are not consistent with those of available data. In 12 cases, MBH98 commenced their calculation after the date in which all records were available (e.g. Australia-New Zealand region where MBH98 commenced in 1750, although a start date of 1625 was possible.) In 16 cases, MBH98 commenced their PC index in a period *prior* to that available in the data (e.g. Texas-Mexico). Because standard PC algorithms fail in the presence of missing data, an important part of the methodology—namely how missing data were treated in the PC calculation—remains unexplained in MBH98.

We computed all 28 PCs, together with their explained variances, using a standard principal component algorithm for the maximum period in which all records were available within each region. For comparison, weighting factors for the MBH98 PCs eigenvectors were computed which maximized the explained variance of the underlying ITRDB data, and the resulting explained variance was compared to our own computations using a standard algorithm. In all cases, explained variance for the

recomputed PCs exceeded that for the MBH98 PCs (see Table 7). Indeed it was the observation of the unusually poor fit between the MBH98 Texas-Mexico PCs and the underlying ITRDB data that led to the detailed audit undertaken in this paper.

Table 7. 5 Regions in which MBH98 computed principal components. Each column shows the number of source sites listed by MBH98, the number found at WDCP; the number of PCs extracted; the dates spanned at WDCP and in the MBH98 PCs; the explained variance of each group.

REGION:	Texas- Oklahoma	Texas- Mexico	ITRDB North America	South America	Australia- NZ
# of Source Sites Listed	14	20	232	18	16
# at WDCP	14	20	231	18	16
# of MBH PCs	3	9	9	3	4
WDCP Available Period	1698-1980	1760-1977	1619-1971	1568-1972	1625-1974
MBH PC Start	1-3: 1700	1: 1400 2: 1499 3-4: 1599 5-9: 1699	1-2: 1400 3-6: 1499 7: 1599 8-9: 1749	1-2: 1600 3: 1750	1-4: 1750
MBH PC End ⁵	1980	1979	1980	1976	1976
Explained Variance: MBH	32%	6%	14%	26%	38%
Recalculated	39%	76%	40%	35%	46%

Figure 5 shows the MBH98 and re-calculated Australian PC1. The Australian PC1 is one of relatively few MBH98 series that shows anomalous 20th century behaviour and which closes on a dramatic “uptick”. The correct computation shows that this feature of this particular MBH98 series is entirely an artefact of incorrect calculation.

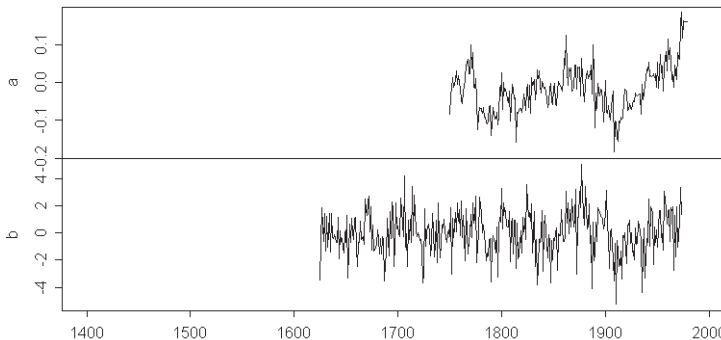


Figure 5. (a) Australia PC1 in MBH98 (series #96) graphed over time (b) PC1 for the MBH98 Australia dataset calculated using standard algorithm.

⁵Excluding filled values

3. TEMPERATURE INDEX CONSTRUCTION USING CORRECTED DATA

A corrected and updated proxy database has been developed, in which the measures outlined above were adopted, including the following:

- the most recent editions of the MBH98 series have been used where identified and available;
- arbitrary MBH98 truncations and fills have been deleted;
- correct tree ring principal component calculations have been used.

We replicated the methodology of MBH98 as closely as we could using publicly available documentation and such private assistance as we were able to obtain.

MBH98 purports to establish relationships between the proxies and 16 temperature principal components calculated from the Climate Research Unit (CRU) instrumental temperature database, using a subset of 1,082 out of 2,592 cells and the 79-year period from 1902-1980 as a calibration period. These 16 temperature principal components are referred to as TPC1—TPC16. Prior to this calculation, the CRU data was scaled cellwise. We downloaded original temperature data from CRU and gridpoint locations from the MBH98 website and calculated scaling factors for downstream use in calculation of northern hemisphere temperature averages. Four MBH98 cells contained no observations in the CRU data and were excluded from all calculations.

Following the description of MBH98 procedures in their Supplementary Information, our construction is done piecewise for each of the periods listed in Table 8, using the roster of proxies available throughout the period and the selection of TPCs for each period listed in Table 8. There are slight discrepancies between 1500 and 1750 in the number of proxies which MBH98 reported to be available and the number actually available in the MBH98 data set (see columns 2-3).

The anomalous listing of TPCs 6 and 8 in the period 1750 to 1759 is assumed to be an erroneous rendering of TPCs 7 and 9, but there is little sensitivity to this assumption. Following MBH98, the number of TPCs used in the construction decreases from 11 in the latest period to 1 in the earliest period, as shown in Table 8.

Following MBH98 procedures as publicly disclosed, for each combination of proxy roster and TPC selection, the proxies were first calibrated against the temperature PCs in the calibration period of 1902-1980 and then the temperature PCs were constructed in each period using the proxy and TPC rosters prescribed by MBH98 for the period, together with weighting factors supplied to the authors by an associate of Prof. Mann. From these constructed PCs, using MBH98 eigenvalues and eigenvectors, gridded temperature series for 1,082 cells were obtained. From the cells in the northern hemisphere (excluding the four cells with no observations and hence no scaling factor), a northern hemisphere average temperature index was calculated. We have posted scripts for this construction in Supplementary Information.

It should be noted that each of the above steps in the MBH98 northern hemisphere temperature index construction is a linear operation on the proxies. Accordingly, given the roster of proxies and TPCs in each period, the result of these linear operations is a set of proxy weighting factors, which generates the NH average temperature construction. These weighting factors are not disclosed in MBH98.

Table 8. Intervals defining proxy groups and subset of temperature PCs used in coefficient fitting process.

Interval	No. of proxies reported available	No. found in data set	Number of Temp PCs fit to proxies	Temperature PC Identifiers
1400-1450	22	22	1	1
1450-1500	24	24	2	1,2
1500-1600	28	34	2	1,2
1600-1700	57	54	4	1,2,11,15
1700-1730	74	73	5	1,2,5,11,15
1730-1750	79	78	5	1,2,5,11,15
1750-1760	89	89	8	1-3,5,6,8,11,15
1760-1780	93	93	9	1-5,7,9,11,15
1780-1800	97	97	11	1-5,7,9,11,14-16
1800-1820	102	102	11	1-5,7,9,11,14-16
1820-1971	112	112	11	1-5,7,9,11,14-16
1972+	112	106-111	11	1-5,7,9,11,14-16

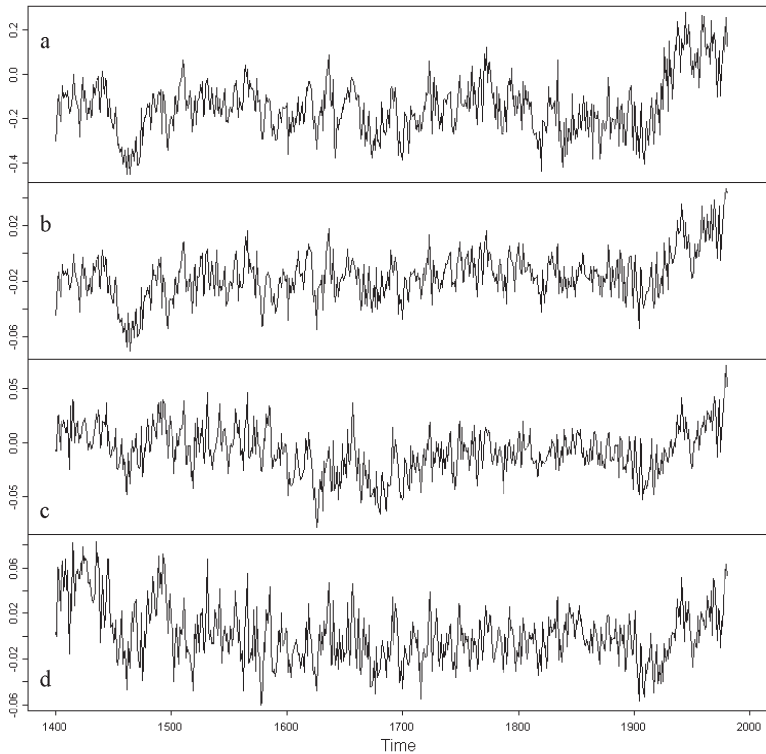


Figure 6. (a) MBH98 NH temperature series (deg C), 1400-1980, which relies heavily on (b) TPC1 from MBH98. (c) Authors' replication of TPC1 using MBH98 methods and data. (d) Authors' TPC1 using MBH98 methods but with data corrected as outlined in text.

The well-known “hockey stick” shape of the MBH98 northern hemisphere temperature index is shown in Figure 6a. It depends strongly on the temperature PC1 (Figure 6b) so we will illustrate its replication, although all TPCs were calculated and used in the NH construction. Our replication of TPC1 using the MBH98 method and data is shown in Figure 6c. Our version of TPC1 in Figure 6c is clearly similar to the calculation of MBH98 in Figure 6b (correlation 0.95 in the 20th century), indicating substantial success in replicating the MBH98 methodology, but some differences remain, possibly due to undisclosed variations in their procedures and assumptions. The TPC1 construction using corrected data is in Figure 6d, showing higher 15th century values than 20th century values, unlike the MBH98 TPC1.

Figures 7 and 8 show the impact of the corrections on northern hemisphere temperature construction. In Figure 7 the top line is the MBH98 construction (reproducing Figure 6a), while the bottom line shows the Northern Hemisphere multiproxy temperature index resulting from the application of MBH98 procedures on an updated and correctly collated assembly of the MBH98 library of proxy data. On the basis of corrected and updated data, 15th century values are higher than those in the 20th century, contradicting the MBH98 conclusion of a unique late 20th century climate. Figure 8 shows 20-year smoothed series for comparison.

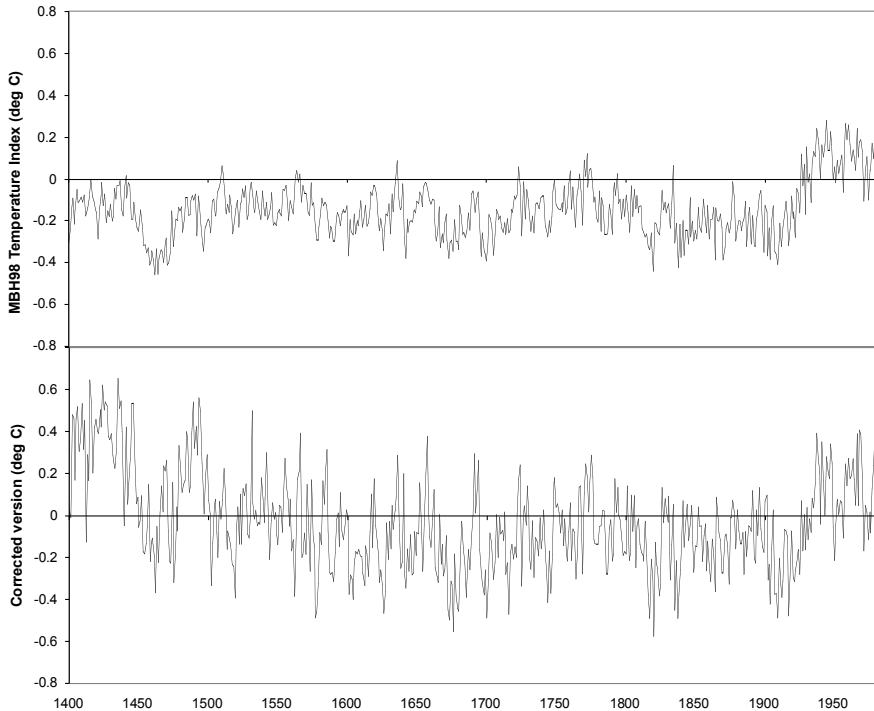


Figure 7. Temperature anomalies index (deg C) 1400-1980 for Northern Hemisphere average temperature construction from (top) Mann et. al. (1998); and (bottom) based on this work using corrected and updated data as outlined in text.

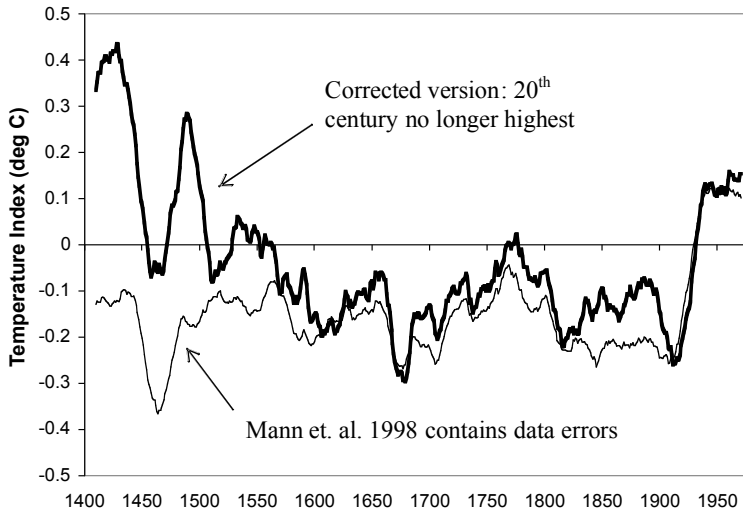


Figure 8. As Figure 7, using 20-year running mean to smooth.

4. CONCLUSIONS

The MBH98 hockey stick-shaped NH temperature index discussed here has been extremely influential in discussions of 20th century global warming. Together with a pre-1400 extension derived in Mann et. al. (1999) and a spliced instrumental temperature series, this index figured prominently in the IPCC Third Assessment Report (IPCC 2001) and numerous other publications. However, the dataset used to make this construction contained collation errors, unjustified truncation or extrapolation of source data, obsolete data, incorrect principal component calculations, geographical mislocations and other serious defects. These errors and defects substantially affect the temperature index.

Although not all of the dataset could be audited, it was possible to prepare a data base with substantially improved quality control, by using the most recent data and collating it correctly, by avoiding arbitrary filling in or truncation of data and by computing principal components using standard algorithms. Without endorsing the MBH98 methodology or choice of source data, we were able to apply the MBH98 methodology to a database with improved quality control and found that their own method, carefully applied to their own intended source data, yielded a Northern Hemisphere temperature index in which the late 20th century is unexceptional compared to the preceding centuries, displaying neither unusually high mean values nor variability. More generally, the extent of errors and defects in the MBH98 data means that the indexes computed from it are unreliable and cannot be used for comparisons between the current climate and that of past centuries, including claims like “temperatures in the latter half of the 20th century were unprecedented,” and “even the warmer intervals in the reconstruction pale in comparison with mid-to-late 20th-century temperatures” (see press release accompanying Mann et al 1999) or that

the 1990s was “likely the warmest decade” and 1998 the “warmest year” of the millennium (IPCC 2001).

REFERENCES

Briffa, K.R., P.D. Jones, and F.H. Schweingruber, 1992, Tree-Ring Density Reconstructions of Summer Temperature Patterns across Western North America since 1600, *Journal of Climate*, Vol. 5, No. 7.

Essex, C. and R. McKittrick (2002). *Taken By Storm: The Troubled Science, Policy and Politics of Global Warming*. Toronto: Key Porter.

Fritts, H.C. & Shao, X.-M. (1992), Mapping climate using tree-rings from western North America, Dendroclimatic evidence from the northern Soviet Union, in *Climate since A.D. 1500*, (eds Bradley, R.S. & Jones, P.D., 269-294, Routledge, 1992).

Intergovernmental Panel on Climate Change (2001) *Climate Change 2001: The Scientific Basis*. Cambridge: CUP.

Jones, P.D. & Bradley, R.S. (1992) Climatic variations in the longest instrumental records, in *Climate Since A.D. 1500*, (eds Bradley, R.S. & Jones, P.D., 246-268, Routledge, 1992).

Mann, M.E., Bradley, R.S. & Hughes, M.K. (1998) Global-Scale Temperature Patterns and Climate Forcing Over the Past Six Centuries, *Nature*, No. 392, pp. 779-787, 1998.

Mann, M.E., Bradley, R.S. and Hughes, M.K., (1999). Northern Hemisphere Temperatures During the Past Millennium: Inferences, Uncertainties, and Limitations, *Geophysical Research Letters*, No. 26, pp. 759-762. See also press release at <http://www.umass.edu/newsoffice/archive/1999/030399warming.html>.

APPENDIX: SUPPLEMENTARY INFORMATION SOURCES

Supplementary information for this paper, including detailed information about all 112 proxy series, the computations and data used for the Figures, are available at <http://www.climate2003.com/index.html> and <http://www.uoguelph.ca/~rmckitri/research/trc.html>.

The supporting web site for the MBH98 paper is http://www.ngdc.noaa.gov/paleo/ei/data_supp.html.

FTP References for Updated MBH98 Series. Column 1 is MBH98 series number. Column 2 is MBH98 series descriptor. Column 3 shows whether a digital update is referred to in the text. Column 4 shows the digital publication reference (see Supplementary Information). Column 5 is applicable line in multi-set FTP reference. NA- No digital publication located. NV- Digital publication located, but not compared.

Series #	MBH98 Description	Ref	Digital Publication	Line
1	Burdekin River Coral Fluorescence	Yes	ftp://ftp.ngdc.noaa.gov/paleo/coral/west_pacific/great_barrier/burdekin_2001.txt	47
2	Great Barrier Reef Coral Thickness Index	Yes	ftp://ftp.ngdc.noaa.gov/paleo/coral/west_pacific/great_barrier/aims10coreavg.txt	
3	Galapagos Urvina Bay UR-86 dO18	Yes	ftp://ftp.ngdc.noaa.gov/paleo/coral/east_pacific/urvcomp.txt	
4	Red Sea, Aqaba Core 18 dO18	No	ftp://ftp.ngdc.noaa.gov/paleo/coral/red_sea/aq18-18o.txt	
5	Red Sea, Aqaba, Core 18 dC13	No	ftp://ftp.ngdc.noaa.gov/paleo/coral/red_sea/aq18-13c.txt	
6	Espiritu Santu, Vanuatu dO18	Yes	ftp://ftp.ngdc.noaa.gov/paleo/coral/west_pacific/vanuatu_annual.txt	
7	New Caledonia dO18	Yes	ftp://ftp.ngdc.noaa.gov/paleo/coral/west_pacific/nc_published_1992-1657_qtr.txt	
8	Gulf of Chiriqui, Panama dO18	No	ftp://ftp.ngdc.noaa.gov/paleo/coral/east_pacific/secas-10yr-iso.txt	
9	Gulf of Chiriqui, Panama dC13	No	ftp://ftp.ngdc.noaa.gov/paleo/coral/east_pacific/secas-10yr-iso.txt	
10	Central England Historical	Yes	http://www.metoffice.com/research/hadleycentre/CR_data/Daily/HadCET_act.txt	
11	Central Europe Historical	Yes	ftp://ftp.ngdc.noaa.gov/paleo/climate1500ad/ch6.txt	
12	Quelccaya Summit dO18	No	ftp://ftp.ngdc.noaa.gov/paleo/icecore/trop/quelccaya/q83cor1.txt	
13	Quelccaya Summit Accum. (m)	No	ftp://ftp.ngdc.noaa.gov/paleo/icecore/trop/quelccaya/q83cor1.txt	
14	Quelccaya Ice Core 2 dO18	No	ftp://ftp.ngdc.noaa.gov/paleo/icecore/trop/quelccaya/q83summ.txt	
15	Quelccaya Ice Core 2 Accum (m)	No	ftp://ftp.ngdc.noaa.gov/paleo/icecore/trop/quelccaya/q83summ.txt	
16	Dunde Ice Core dO18	NA	NA	
17	West Greenland Ice Melt (pct)	NA	NA	
18	Svalbard Ice Melt, 5-yr avg ("pct)	NA	see ftp://ftp.ngdc.noaa.gov/paleo/climate1500ad/ch26.txt	
19	Penny, Baffin Island dO18	NA	NA	
20	Central Greenland (Stack) dO18	NA	NA	
21	Station temperature 42.5N, 92.5W	Yes	ftp://ftp.ngdc.noaa.gov/paleo/climate1500ad/ch13.txt	2370
22	Station temperature 47.5N, 2.5E	NA	NA	
23	Station temperature 47.5N, 7.5E	Yes	ftp://ftp.ngdc.noaa.gov/paleo/climate1500ad/ch13.txt	780
24	Station temperature 47.5N, 12.5E	NA	NA	
25	Station temperature 47.5N, 17.5E	NA	NA	
26	Station temperature 52.5N, 17.5E	NA	NA	
27	Station temperature 57.5N, 17.5E	Yes	ftp://ftp.ngdc.noaa.gov/paleo/climate1500ad/ch13.txt	272
28	Station temperature 75.5N, 32.5E	Yes	ftp://ftp.ngdc.noaa.gov/paleo/climate1500ad/ch13.txt	1289

29	Station temperature 62.5N, 7.5E	NA	NA		
30	Station temperature 62.5N, 12.5E	Yes	ftp://ftp.ngdc.noaa.gov/paleo/climate/1500ad/ch13.txt	52	
31	Station temperature 62.5N, 42.5E	NA	NA		
32	Station Precipitation 12.5N, 62.5E	NA	NA		
33	Station Precipitation 17.5N, 72.5E	NA	NA		
34	Station Precipitation 37.5N, 77.5W	NA	NA		
35	Station Precipitation 42.5N, 2.5E	Yes	ftp://ftp.ngdc.noaa.gov/paleo/climate/1500ad/ch13.txt	3650	
36	Station Precipitation 42.5N, 7.5E	NA	NA		
37	Station Precipitation 42.5N, 72.5W	Yes	ftp://ftp.ngdc.noaa.gov/paleo/climate/1500ad/ch13.txt 3848		
38	Station Precipitation 47.5N, 2.5E	NA	NA		
39	Station Precipitation 47.5N, 12.5E	NA	NA		
40	Station Precipitation 52.5N, 12.5E	NA	NA		
41	Station Precipitation 52.5N, 2.5W	NA	NA		
42	Station Precipitation 57.5N, 7.5W	NA	NA		
43	Tasmania T-reconstruction	Yes	ftp://ftp.ngdc.noaa.gov/paleo/treering/reconstructions/tasmania/tasmania_recon.txt		
44	Java	NA	NA		
45	New Zealand T-reconstruction	Yes	ftp://ftp.ngdc.noaa.gov/paleo/climate/1500ad/ch24.txt	65	
46	cpatagonia 41S!	Yes	ftp://ftp.ngdc.noaa.gov/paleo/climate/1500ad/ch23.txt	187	
47	npatagonia 38S!	No	ftp://ftp.ngdc.noaa.gov/paleo/climate/1500ad/ch23.txt	236	
48	Upper Kolyma River, Russia T	NA			
49	Western North America T (MXD)	No	ftp://ftp.ngdc.noaa.gov/paleo/treering/reconstructions/westamerica/briffa1992/briffa1992.txt		
50	Western North America T (RW)	NV	ftp://ftp.ngdc.noaa.gov/paleo/treering/reconstructions/westamerica/readme_westamerica_recons.txt		
51	Treeline, 41.2 Alaska	Yes	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/northamerica/usa/ak031.crn		
52	Treeline, Fort Chimo PQ	Yes	ftp://ftp.ngdc.noaa.gov/paleo/treering/measurements/northamerica/canada/cana002.crn		
53	Treeline, Gaspé PQ	Yes	ftp://ftp.ngdc.noaa.gov/paleo/treering/measurements/northamerica/canada/cana036.crn		
54	Treeline, Arrigetch AK	Yes	ftp://ftp.ngdc.noaa.gov/paleo/treering/measurements/northamerica/usa/ak032.crn		
55	Treeline, Sheenjek R, Alaska	Yes	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/northamerica/usa/ak033.crn		
56	Treeline, TTHH Canada	Yes	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/northamerica/canada/cana157.crn		
57	Treeline, Mackenzie Mts, Canada	No	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/northamerica/canada/cana154.crn		

58	Treeline, Coppermine R, Canada	Yes	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/northamerica/canada/canal53.cm
59	Treeline, Hornby Cabin, Canada	No	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/northamerica/canada/canal55.cm
60	Treeline, Churchill, Canada	No	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/northamerica/canada/canal58.cm
61	Treeline, Castle Penin, Canada	No	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/northamerica/canada/canal59.cm
62	Precip-Recon- SE USA-NC	NA	NA
63	Precip-Recon- SE USA -SC	NA	NA
64	Precip-Recon - SE USA -GA	NA	NA
65	Mongolia, Tarvagatny Pass	Yes	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/asia/mong003.cm
66	Yakutia, Russia T-reconstruction	NA	NA - but compare to ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/asia/russ142w_cms.cm
67	Fennoscandia T-reconstruction	NA	NA - but compare to ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/europe/swed019x_cms.cm
68	Northern Urals T-reconstruction	NA	NA
69	USA, OK PC1	NA	NA
70	USA, OK PC2	NA	NA
71	USA, OK PC3	NA	NA
72	Mexico PC1	NA	NA
73	Mexico PC2	NA	NA
74	Mexico PC3	NA	NA
75	Mexico PC4	NA	NA
76	Mexico PC5	NA	NA
77	Mexico PC6	NA	NA
78	Mexico PC7	NA	NA
79	Mexico PC8	NA	NA
80	Mexico PC9	NA	NA
81	Vaganov12 Chronologies PC1	NA	NA
82	Vaganov 40 Chronologies - PC1	NA	NA
83	Vaganov 58 Chronologies PC1	NA	NA
84	USA PC1	NA	NA
85	USA PC2	NA	NA
86	USA PC3	NA	NA

87	USA PC4	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/asia/chim004.crn
88	USA PC5	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/asia/chim004x.crn
89	USA PC6	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/europe/fran009.crn
90	USA PC7	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/europe/fran010.crn
91	USA PC8	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/europe/fran011.crn
92	USA PC9	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/asia/indi002x.crn
93	South America PC1	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/northamerica/mexico/mexi001.crn
94	South America PC2	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/africa/morc011.crn
95	South America PC3	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/africa/morc001.crn
96	Australia PC1	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/africa/morc014.crn
97	Australia PC2	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/europe/spat011.crn
98	Australia PC3	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/europe/spat012.crn
99	Australia PC4	NA	NA	ftp://ftp.ngdc.noaa.gov/paleo/treering/chronologies/europe/swed002.crn
100	CHIN04	Yes		
101	CHIN04X	No		
102	FRAN009	No		
103	FRAN010	No		
104	FRAN011	No		
105	INDI008X	Yes		
106	MEXI001	No		
107	MORO003	No		
108	MORO007	No		
109	MORO008	No		
110	SPAI011	No		
111	SPAI012	No		
112	SWED002B	Yes		