Clim. Past Discuss., 2, 1001–1049, 2006
www.clim-past-discuss.net/2/1001/2006/

Climate of the Past Discussions

Climate of the Past Discussions is the access reviewed discussion forum of *Climate of the Past*

Millennial temperature reconstruction intercomparison and evaluation

M. N. Juckes[1], M. R. Allen[2], K. R. Briffa[3], J. Esper[4], G. C. Hegerl[5], A. Moberg[6], T. J. Osborn[3], S. L. Weber[7], and E. Zorita[8]

[1]British Atmospheric Data Centre, Rutherford Appleton Laboratory, UK
[2]Atmospheric, Oceanic and Planetary Physics, University of Oxford, UK
[3]Climatic Research Unit, University of East Anglia, UK
[4]Swiss Federal Research Institute for Forest, Snow and Landscape, Bern University, Switzerland
[5]Dept. Earth and Ocean Sciences, Duke University, NC, USA
[6]Department of Meteorology, Stockholm University, Sweden
[7]Royal Netherlands Meteorological Institute (KNMI), De Bilt, The Netherlands
[8]GKSS Research Centre, Geesthacht, Germany

Received: 26 September 2006 – Accepted: 13 October 2006 – Published: 26 October 2006

Correspondence to: M. N. Juckes (m.n.juckes@rl.ac.uk)

Abstract

There has been considerable recent interest in paleoclimate reconstructions of the temperature history of the last millennium. A wide variety of techniques have been used. The interrelation between the techniques is sometimes unclear, as different studies often use distinct data sources as well as distinct methodologies. Recent work is reviewed with an aim to clarifying the import of the different approaches. A range of proxy data collections used by different authors are passed through two reconstruction algorithms: firstly, inverse regression and, secondly, compositing followed by variance matching. It is found that the first method tends to give large weighting to a small number of proxies and that the second approach is more robust to varying proxy input. A reconstruction using 18 proxy records extending back to AD 1000 shows a maximum pre-industrial temperature of 0.25 K (relative to the 1866 to 1970 mean). The standard error on this estimate, based on the residual in the calibration period is 0.149 K. Two recent years (1998 and 2005) have exceeded the estimated pre-industrial maximum by more than 4 standard errors.

1 Introduction

The climate of the last millennium has been the subject of much debate in recent years, both in the scientific literature and in the popular media. This paper reviews reconstructions of past temperature, on the global, hemispheric, or near-hemispheric scale, by Jones et al. (1998) [JBB1998], Mann et al. (1998) [MBH1998], Mann et al. (1999) [MBH1999], Huang et al. (2000) [HPS2000], Crowley and Lowery (2000) [CL2000], Briffa et al. (2001) [BOS2001], Esper et al. (2002) [ECS2002], Mann and Jones (2003) [MJ2003], Moberg et al. (2005) [MSH2005], Oerlemans (2005) [OER2005], Hegerl et al. (2006a) [HCA2006].

Climate variability can be partitioned into contributions from internal variability of the climate system and response to forcings, with the forcings being further partitioned in

CPD

2, 1001–1049, 2006

Millennial temperature

M. N. Juckes et al.

EGU

natural and anthropogenic. The dominant change in forcing in the late 20th century arises from human impact in the form of greenhouse gases (primarily carbon dioxide, methane and chloro-fluoro carbons: Mitchell et al., 2001 [IPCC2001]). The changes in concentration of these gases in the atmosphere are well documented and their radiative properties which reduce, for a given temperature difference, radiative loss of heat to space from the mid and lower troposphere (for carbon dioxide, this was first documented by Arrhenius, 1896) are beyond dispute.

However, there remains some uncertainty on two issues: firstly, how much of the observed climate change in the 20th century is due to greenhouse forcing as opposed to natural forcing and internal variability; secondly, how significant, compared to past natural changes, are the changes which we now observe and expect in the future?

Uncertainty in the answer to the first question remains because of uncertainty about the precise amplitude of climate response to a given forcing and uncertainty about the amplitude of internal variability on centennial time scales. The second question reflects the uncertainty in the response of the climate system to a given change in forcing. In the last century both the variations in forcing and the variations in response have been measured with some detail, yet there remains uncertainty about the contribution of natural variability to the observed temperature fluctuations. In both cases, investigation is hampered by the fact that estimates of global mean temperature based on reliable direct measurements are only available from 1856 onwards (Jones et al., 1986).

Climate models are instrumental in addressing both questions, but they are still burdened with some level of uncertainty and there is a need for more detailed knowledge of the behaviour of the actual climate on multi-centennial timescales both in order to evaluate the climate models and in order to address the above questions directly.

The scientific basis for proxy based climate reconstructions may be stated simply: there are a number of physical indicators which contain information about the past environmental variability. As these are not direct measurements, the term proxy is used. Jones and Mann (2004) review evidence for climate change in the past two millennia and conclude that the 20th century has seen the greatest temperature change

CPD
2, 1001–1049, 2006

within any century in this period. The coolest centuries appear to have been the 15th, 17th, and 19th centuries.

This paper reviews different contributions and evaluates the impact of different methods and different data collections used. Section 2 discusses recent contributions, which have developed a range of new methods to address aspects of the problem. Section 3 discusses the technique used by MBH1998/9 in more detail in the context of criticism by McIntyre and McKitrick (2003) (hereafter MM2003). Section 4 presents some new results using data collections from 5 recent studies. Appendices provide information on the regression techniques and statistical tests used.

2 A survey of recent reconstructions

This section gives brief reviews of recent contributions, displayed in Fig. 1. Of these, 5 are estimates of the Northern Hemisphere mean temperature (MBH1999, HPS2000, CL2000, MSH2005, HCA2006), 2 of the Northern Hemisphere extra tropical mean temperature (BOS2001, ECS2002) and 3 of the global mean temperature (JBB1998, MJ2003, OER2005). HPS2000 represents land temperatures only. The ECS2002 curve is based on extratropical data but has here been calibrated to match the variance of the hemispheric mean temperature. There are also differences with respect to the seasons represented: the BOS2001 reconstruction represents growing season temperatures (April to September) over land areas poleward of 20° N. The graph labelled HCA2006 is their "CH-blend (Dark ages)" reconstruction representing the mean over land areas north of 30° N. All, except the inherently low resolution reconstructions of HPS2000 and OER2005, have been smoothed with a 21 year running mean. The time series in Fig. 1 have been centred to have zero mean in the AD 1900 to 1960 period. The ECS2002 series was obtained as an uncalibrated index and has here been scaled by 1.73 to match the variance of Northern Hemisphere temperature. With the exception of HPS2000 and OER2005, the reconstructions use partly overlapping data, so they cannot be viewed as independent from a statistical viewpoint. In addition

CPD
2, 1001–1049, 2006
Millennial temperature
M. N. Juckes et al.

to exploiting a range of different data sources, the above works also use a range of techniques. The subsections below cover different scientific themes, ordered according to the date of key publications. Some reconstructions which do not extend all the way back to AD 1000 are included because of their importance in addressing specific issues.

2.1 High-resolution paleoclimate records

Jones et al. (1998) [JBB1998] present the first annually resolved hemispheric reconstructions of temperatures back to AD 1000, using a composite of 10 standardised proxies for the northern hemisphere and 7 for the southern, with variance damped in the early part of the series to allow for the lower numbers of proxies present (6 series extend back to AD 1000) following Osborn et al. (1997). The composites are scaled by variance matching (Appendix A) against the hemispheric annual mean summer temperatures for 1931–1960. It is also shown that there are strong large scale coherencies in the proxy data which are not reproduced in climate model temperature fields. An evaluation of each individual proxy series against instrumental data from 1881 to 1980 shows that tree-rings and historical reconstructions are more closely related to temperature than those from coral and ice-cores.

With regard to the temperatures of the last millennium, the primary conclusion of JBB1998 is that the twentieth century was the warmest of the millennium. There is clear evidence of a cool period from 1500 to 1900, but no strong "Medieval Warm Period" [MWP] (though the second warmest century in the Northern Hemisphere reconstruction is the 11th). The MWP is discussed further in Sect. 2.4 below.

JBB1998 also raise concerns about the homogeneity of some of the proxies on longer timescales (see Sect. 3.5 below). This is an important issue, since many climate reconstructions (all those reviewed here except HPS2000 and OER2005) rely on constant empirical relationship between temperature anomalies and the proxy indicators.

2.2 Climate field reconstruction

Mann et al. (1999) published the first reconstruction of the last thousand years northern hemispheric mean temperature which included objective error bars, based on the analysis of the residuals in the calibration period. The authors concluded not only that their estimate of the temperature over the whole period AD 1000 to AD 1860 was colder than the late twentieth century, but also that 95% confidence limits of their estimate were below the last decade of the twentieth century. The methods they used were presented in MBH1998 which described a reconstruction back to AD 1400.

MBH1998 use a collection of 415 proxy time indicators, substantially more than used in Jones et al. (1998), but many of these are too close geographically to be considered as independent, so they are combined into a smaller number of representative series. The number of proxies also decreases significantly back in time: only 22 independent proxies extend back to AD 1400, and, in MBH1999, 12 extend back to AD 1000 (7 in the Northern Hemisphere). MBH1998 and MBH1999 have been the subject of much debate (see Sect. 3) since the latter was cited in IPCC2001, though the IPCC conclusions[1] were weaker than those of MBH1999.

MBH1998,1999 also differ from Jones et al. (1998) in using spatial patterns of temperature variability rather than a hemispheric mean temperature time series. The aim is to exploit proxies which respond to climate anomalies (e.g. rainfall changes) and so might have an indirect link to temperature changes even when there is no direct link

[1] Mitchell et al. (2001) concluded that "The 1990s are likely to have been the warmest decade of the millennium in the Northern Hemisphere, and 1998 is likely to have been the warmest year", where "likely" implies a greater than 66% probability. Since 2001 it has been recognised that there is a need to explicitly distinguish between an expression of confidence, as made by the IPCC in this quote, which should include expert assessment of the robustness of statistical methods employed, and simple citation of the results of statistical test. In the language of Manning et al. (2004) we can say that MBH1999 carried out statistical tests which concluded that the 1990s have been the warmest decade of the millennium with 95% likelihood, while IPCC2001, after assessing all available evidence had a 66% confidence in the same statement.

with local temperatures.

Different modes of atmospheric variability are evaluated through an Empirical Orthogonal Function [EOF] analysis of the time period 1902 to 1980, expressing the global field as a sum of spatial patterns (the EOFs) multiplied by Principal Components (PCs – representing the temporal evolution). Earlier instrumental data are too sparse to be used for this purpose: instead they are used in a validation calculation to determine how many EOFs should be included in the reconstruction. Time series for each mode of variability are then reconstructed from the proxy data using a least squares inverse regression. The skill of the regression of each PC is tested using the 1856 to 1901 validation data. For the reconstruction prior to 1450 AD it is determined that only one PC can be reconstructed with any accuracy.

The reconstructed temperature evolution (Fig. 1) is rather less variable than that of Jones et al. (1998). The most substantial differences occur in the 17th and 19th centuries, when the MBH1999 reconstruction is about 0.3 K warner than that of JBB1998. The overall picture is of gradual cooling until the mid 19th century, followed by rapid warming.

2.3 Borehole temperatures

Huang et al. (2000) [HPS2000] estimate northern hemisphere temperatures back to 1500 AD using measurements made in 453 boreholes (their paper also presents global and southern hemisphere results using an additional 163 southern hemisphere boreholes). The reconstruction is included here, even though it does not extend back to AD 1000, because it has the advantage of being completely independent of the other reconstructions shown. Temperature fluctuations at the surface propagate slowly downwards, so that measurements made in the boreholes at depth contain a record of past surface temperature fluctuations. HPS2000 used measurements down to around 300 m. The diffuse nature of the temperature anomaly means that short time scale fluctuations cannot be resolved. Prior to the 20th century, the typical resolution is about 100 years.

Mann et al. (2003c) discuss the impact of changes in land use and snow cover on borehole temperature reconstructions and conclude that it results in significant errors. This conclusion has been contested by Pollack and Smerdon (2004) (on statistical grounds), González-Rouco et al. (2003) (using climate simulations) and Huang (2004) (using an expanded network of 696 boreholes in the northern hemisphere).

The overall picture of HPS2000 is of an accelerated warming trend over the last 500 years, with the 16th century being about 0.5 K below the 1900 to 1960 average. The 16th century cooling trend seen in JBB1998 and MBH1999, when smoothed to centennial resolution, is not present in HPS2000. It should be noted that the technique used to generate the bore hole estimate (Pollack et al., 1998) assumes a constant temperature prior to AD 1500. The absence of a cooling trend after this date may be influenced by this boundary condition.

2.4 Medieval Warm Period

Despite much discussion (e.g. Hughes and Diaz, 1994; Bradley et al., 2003), there is no clear quantitative understanding of what is meant by the "Medieval Warm Period" (MWP). Crowley and Lowery (2000) [CL2000] discuss the evidence for a global MWP, which they interpret as a period of unusual warmth in the 11th century. All the reconstructions of the 11th century temperature shown in Fig. 1 estimate it to have been warmer than most of the past millennium. However, a question of more practical importance is not whether it was warmer than the 12th to 19th centuries, which is generally accepted, but whether it was a period of comparable warmth to the late 20th century. MBH1999 concluded, with 95% confidence, that this was not so. CL2000 revisit the question using 15 proxy records (7 annually resolved, 3 with decadal scale variability and 5 with only centennial temporal resolution). Low-resolution (decadal and centennial) series were not used in the studies cited above, and 3 of the high-resolution series used by CL2000 are also new.

They draw attention to the spatial localization of the MWP in their proxy series: it is strong in North America, North Atlantic and Western Europe, but not clearly present elsewhere. Periods of unusual warmth do occur in other regions, but these are short and asynchronous.

Their estimate of northern hemispheric temperature over the past millennium (see Fig. 1) is close to that of MBH1999. They conclude that the occurrence of decades of temperatures similar to those of the late 20th century cannot be unequivocally ruled out, but that there is, on the other hand, no evidence to support the claims that such an extended period of large-scale warmth occurred.

MJ2003 extend the study period to the last 1800 years using a combination of low and high resolution proxies and a CVM approach (see Appendix A) with regional sub-composites and principal components to deal with local oversampling. They conclude that the late 20th century warmth is unprecedented in the last two millennia (the latter thousand years of their global reconstruction is included in Fig. 1).

Soon and Baliunas (2003) carry out an analysis of local climate reconstructions. They evaluate the number of such reconstructions which show (a) a sustained "climate anomaly" during AD 800–1300, (b) a sustained "climate anomaly" during AD 1300–1900 and (c) their most anomalous 50 year period in the 20th century. Their definition of a "sustained climate anomaly" is 50 years of warmth, wetness or dryness for (a) and (c) and 50 years of coolness, wetness or dryness in (b). It should be noted that they do not carry out evaluations which allow direct comparison between the 20th century and earlier times: they compare the number of extremes occurring in the 20th century with the number of anomalies occurring in periods of 3 and 4 centuries in the past. Both the use of sampling periods of differing length and different selection criteria make interpretation of their results problematic. They have also been criticised for interpreting regional extremes which occur at distinct times as being indicative of a global climate extremes (Jones and Mann, 2004). This issue is discussed further in Sect. 2.9 below (see also Mann et al., 2003a; Soon et al., 2003; Mann et al., 2003b). Osborn and Briffa (2006) perform a more rigorous and quantitative analysis along the lines of Soon and Baliunas (2003), using a method that by-passes the problem of proxy calibration against instrumental temperatures, and conclude that the proxy records alone show an

CPD

2, 1001–1049, 2006

Millennial temperature

M. N. Juckes et al.

EGU

unprecedented synchronous anomaly in the 20th century.

2.5 Segment length curse

Briffa et al. (2001) and Briffa et al. (2002) discuss the impact of the "segment length curse" (Cook et al., 1995; Briffa et al., 1996; Briffa, 2000) on temperature reconstructions from tree rings. Tree ring chronologies are often made up of composites of many trees of different ages at one site. The width of the annual growth ring depends not only on environmental factors but also on the age of the tree. The age dependency of growth is often removed by subtracting a growth curve from the tree ring data for each tree. This process, done empirically, will not only remove age related trends but also any environmental trends which span the entire life of the tree. Briffa et al. (2001) use a more sophisticated method (Age Band Decomposition [ABD], which forms separate chronologies from tree rings in different age bands, and then averages all the age-band chronologies) to construct northern hemisphere temperatures back to AD 1400, and show that a greater degree of long term variability is preserved (they also use density data, rather the tree ring width). The reconstruction lies between those of JBB1998 and MBH1999, showing the cold 17th century of the former, but the relatively mild 19th century of the latter.

The potential impact of the segment length limitations is analysed further by Esper et al. (2002, 2003), using "Regional Curve Standardisation" (RCS) (Briffa et al., 1992). The RCS method uses growth curves (different curves reflecting different categories of growth behaviour) obtained by compositing data from all the trees in a region with respect to age and then applying some smoothing. These smoothed composites are used to remove the growth signal from each individual series, either by subtraction or normalisation. Whereas ABD circumvents the need to use a growth curve, RCS seeks to evaluate a growth curve which is not contaminated by climate signals. The ECS2002 analysis agrees well with that of MBH1999 (and others) on short time scales, but has greater centennial variability (Esper et al., 2004). ECS2002 suggest that this may be partly due to the lack of tropical proxies in their work, which they suggest should be regarded as giving extratropical Northern Hemisphere estimate. The extratropics are known to have greater variability than the tropics. However, it has to be also noted that among the proxies used by MBH1999 (12 in total), just 2 of them are located in the tropics, both at one location (see Table 1 below).

Cook et al. (2004) study the data used by ECS2002 and pay particular attention to potential loss of quality in the earlier parts of tree-ring chronologies when a relatively small number of tree samples are available. Their analysis suggests that tree ring chronologies prior to AD 1200 should be treated with caution.

2.6 Separating timescales

Moberg et al. (2005) [MCS2005] follow BOS2001 and ECS2002 in trying to address the "segment length curse", but, rather than trying to improve the standardization of tree-ring chronologies they discard the low frequency components of the tree-ring data and replace them with information from proxies with lower temporal resolution. A wavelet analysis is used to filter different temporal scales.

Each individual proxy series is first scaled to unit variance and then wavelet transformed. Averaging of the wavelet transforms is made separately for tree ring data and the low-resolution data. The average wavelet transform of tree-ring data for timescales less than 80 years is combined with the averaged wavelet transform of the low-resolution data for timescales longer than 80 years to form one single wavelet transform covering all timescales. This composite wavelet transform is inverted to create a dimensionless temperature reconstruction, which is calibrated against the instrumental record of northern hemisphere mean temperatures, AD 1856–1979, using a variance matching method.

Unfortunately, the calibration period is too short to independently calibrate the low frequency component. The variance matching thus represents a form of cross-calibration. In all calibrations against instrumental data, the long period (multi-centennial) response is determined by a calibration which is dominated by sub-centennial variance. The MSH2005 approach makes this explicit and shows a level of centennial variability which is much larger than in MBH1999 reconstruction and similar to that in ECS2002 and also in simulations of the past millennium with two different climate models, ECHO-G (von Storch et al., 2004) and NCAR CSM ("Climate System Model") (Mann et al., 2005).

2.7 Glacial advance and retreat

Oerlemans (2005) provides another independent estimate of the global mean temperature over the last 460 years from an analysis of glacial advance and retreat. As with the borehole based estimate of HPS2000, this work uses a physically based model rather than an empirical calibration (though the physical model relies heavily on empirically determined coefficients). The resulting curve lies within the range spanned by the high-resolution proxies, roughly midway between the MBH1999 Climate Field Reconstruction and the HPS2000 borehole estimate.

Unlike the borehole estimate, but consistent with most other works presented here, this analysis shows a cooling trend prior to 1850, related to glacial advances over that period.

2.8 Regression techniques

Many of the reconstructions listed above depend on empirical relationships between proxy records and temperature. von Storch et al. (2004) suggest that the regression technique used by MBH1999 under-represents variability[2] the variability of past climate. This conclusion is drawn after a applying a method similar to that of MBH1999 to output from a climate model using a set of pseudo-proxies: time series generated

[2]This is sometimes referred to as "underestimating", which will mean the same thing to many people, but something slightly different to statisticians. In statistical parlance "estimate" implies a value derived from a formal statistical model. The estimated variability will not generally be the same as the variability of the estimated mean. The MBH1998,1999 papers provide no formal estimate of variability.

from the model output and degraded with noise which is intended to match the noise characteristics of actual proxies. Mann et al. (2005) also test reconstruction techniques against climate model simulation, though using different reconstruction techniques, and found no evidence of systematic under-representation of low-frequency variability. The debate is ongoing (Wahl and Ammann, 2006; Mann et al., 2003a,b; Soon et al., 2003; Bürger et al., 2006).

There is some uncertainty about the true nature of noise on the proxies, and on the instrumental record. The least squares estimation technique of MBH1998 effectively neglects the uncertainties in the proxy data relative to uncertainties in the temperature. Instead, HCA2006 use total least squares regression (Allen and Stott, 2003; Adcock, 1878). The method accounts for the uncertainty due to unknown noise in proxy data subject to the assumption that the instrumental noise is known. HCA2006 show that this approach leads to greater variability in the reconstruction.

Rutherford et al. (2005) take a different approach. They compare reconstructions from AD 1400 to present using a regularised expectation maximisation technique (Schneider, 2001) and the MBH1998 climate field reconstruction method and find only minor differences. Standard regression techniques assume that we have a calibration period, in which both sets of variables are measured, and a reconstruction (or prediction) period in which one variable is estimated, by regression, from the other. The climate reconstruction problem is more complex: there are hundreds of instrumental records which are all of different lengths, and similar numbers of proxy records, also of varying length. The expectation maximisation technique (Little and Rubin, 1987) is well suited to deal with this: instead of imposing an artificial separation between a calibration period and a reconstruction period, it fills in the gaps in a way which exploits all data present. Regularised expectation maximisation is a generalisation developed by Schneider (2001) to deal with ill posed problems. Nevertheless, there is still a simple regression equation at the heart of the technique.

CPD
2, 1001–1049, 2006
Millennial temperature
M. N. Juckes et al.

EGU

2.9 Natural variability and forcings

Global temperature can fluctuate through internally generated variability of the climate system (as in the El Niño phenomenon), through variability in natural forcings (solar insolation, volcanic aerosols, natural changes to greenhouse gas concentrations) and human changes. Reconstructions of variations in the external forcings for the last millennium have been put forward (Crowley, 2000), although recent studies have suggested a lower amplitude of low-frequency solar forcing (Lean et al., 2002; Foukal et al., 2004).

Analysis of reconstructed temperatures of MBH1999 and CL2000 and simulated temperatures using reconstructed solar and volcanic forcings shows that changes in the forcings can explain the reconstructed long term cooling through most of the millennium and the warming in the late 19th century (Crowley, 2000). The relatively cool climate in the second half of the 19th century may be attributable to cooling from deforestation (Bauer et al., 2003). Hegerl et al. (2003, 2006) analyse the relation between temperature reconstructions (MBH1999, CL2000, BOS2001 and ECS2002 in the first paper, BOS2001, ECS2002, MJ2003 and HCA2006 in the second) and estimated forcings (Crowley, 2000) using multiple regression. In Hegerl et al. (2006b) it is concluded that that natural forcing, particularly volcanism, explain a substantial fraction of hemispheric temperature variability on decadal and longer timescales. Greenhouse gas forcing is detectable with high significance levels in all reconstructions analysed there. Weber (2005) also analyses the relation between reconstructions (JBB1998; MBH1999; Briffa, 2000; CL2000; ECS2002; MJ2003; MSH2005) and forcings, but a timescale dependent analysis is used rather than multiple regression. It is shown that the timescale dependence of northern hemispheric temperatures on the forcings found from reconstructions is similar to that found in a climate model. The role of solar forcing is found to be larger for longer timescales, whereas volcanic forcing dominates for decadal timescales. The trend component over the period 1000 to 1850 is, however, in all reconstructions larger than the trend implied by the forcings.

CPD
2, 1001–1049, 2006
Millennial temperature
M. N. Juckes et al.

The methods employed by Hegerl et al. (2006b) attribute about a third of the early 20th century warming, sometimes more, in high-variance reconstructions to greenhouse gas forcing. These results indicate that enhanced variability in the past does not make it more difficult to detect greenhouse warming, since a large fraction of the variability can be attributed to external forcing. Quantifying the influence of external forcing on the proxy records is therefore more relevant to understanding climate variability and its causes than determining if past periods were possibly as warm as the 20th century.

Goosse et al. (2005) investigate the role of internal variability using an ensemble of 25 climate model simulations of the last millennium made using forcing estimates from Crowley (2000). They conclude that internal variability dominates local and regional scale temperature anomalies, implying that most of the variations experienced by a region such as Europe over the last millennium could be caused by internal variability. On the hemispheric and global scale, however, the forcing dominates. This agrees with results from a long solar-forced model simulation by Weber et al. (2004). Goosse et al. (2005) make the new point, that natural variability can lead to regional temperature anomalies peaking at different times to the forcing, so that disagreements in timing between proxy series should not necessarily be interpreted as meaning there is no common forcing.

2.10 The long view

The past sections have drawn attention to the problems of calibrating temperature reconstructions using a relatively short period over which instrumental records are available. For longer reconstructions, with lower temporal resolution, other methods are available. Pollen reconstructions of climate match the ecosystem types with those currently occurring at different latitudes. The changes in ecosystem can then be mapped to the temperatures at which they now occur (e.g. Bernabo, 1981; Gajewski, 1988; Davis et al., 2003). These reconstructions cannot resolve decadal variability, but they provide an independent estimate of local low-frequency temperature variations. The results of Weber et al. (2004) and Goosse et al. (2005) suggest that such estimates of regional mean temperatures can provide some information about global mean anomalies, as they strongly reflect the external forcings on centennial and longer timescales. The Goosse et al. results also suggest that hemispheric, extra-tropical and global reconstructions should be strongly correlated on the centennial time-scale. Pollen records were also included in the CL2000 and MSH2005 reconstructions (one each), but there has, as yet, been no detailed intercomparison between the pollen based reconstructions and the higher resolution reconstructions.

3 Critics of the IPCC consensus on millennial temperatures

Mitchell et al. (2001) attempted to represent a consensus with the statement that 1990s are likely to have been the warmest decade of the millennium in the Northern Hemisphere. All temperature reconstructions described in the previous section support this view (though they disagree on other issues). Nevertheless, there are many who are strongly attached to the view that past temperature variations were significantly larger and that, consequently, the warming trend seen in recent decades should not be considered as unusual.

This section focuses on the criticisms of MBH1998, 1999 which have dominated the public debate. Though some of the critics identify the IPCC consensus with the MBH1998 work, this is not the case: the consensus rests on a broader body of work, and as formulated by IPCC2001 is less strong than the conclusions of MBH1998 (Sect. 3.2). McIntyre and McKitrick (2003) [MM2003] criticise MBH1998 on many counts, some related to deficiencies in the description of the data used and possible irregularities in the data itself. These issues have been largely resolved in Mann et al. (2004).

As noted above, the MBH1998 analysis is considerably more complex than others, and uses a greater volume of data. There are 4 main stages of the algorithm: (1) sub-sampling (using an EOF decomposition) of regions with disproportionate numbers of proxies, (2) EOF decomposition of the instrumental global temperature record, (3) regression of the temperature PCs against the proxy collection, (4) validation and estimation of uncertainty.

Stage (1) is necessary because some parts of the globe, particularly North America and Northern Europe, have a disproportionately large number of proxy records. Other authors have dealt with this by using only a small selection of the available data or using regional averages (e.g., BOS2001; HCA2006). MBH1998 use a principal component analysis to extract the common signal from the records in these densely sampled regions.

MM2003 attempted to reproduce the MBH1998 reconstruction, using the same data and method as interpreted from the published information. They attribute the failure of this attempt to errors in the MBH1998 methodology, but a major factor was their misunderstanding of the stepwise reconstruction method in relation to stage (1) (Jones and Mann, 2004; Wahl and Ammann, 2006). MBH1998 use different subsets of their proxy database for different time periods. The stepwise reconstruction allows more data to be used for more recent periods. For example, Fig. 2 illustrates how this approach applies to the North American tree ring network in stage (1). Of the total of 212 chronologies, only 66 extend back beyond AD 1400. MM2003 only calculate principal components for the period when all chronologies are present. Similarly, data from South West Mexico was omitted by MM2003 (discussed further in the supplementary material: http://www.clim-past-discuss.net/2/1001/2006/cpd-2-1001-2006-supplement.zip).

McIntyre and McKitrick (2005c) [MM2005c] revisit the MM2003 work and correct their earlier error by taking the stepwise reconstruction technique into account in stage (1). They assert that the results of MM2003, which show a 15th century reconstruction 0.5 K warmer than found by MBH1998, are reproduced with only minor changes to the MBH1998 proxy data base. Examination of the relevant figures, however, shows that this is not entirely true. The MM2005c predictions for the 15th century are 0.3 K warmer than the MBH1998 result: this is still significant, but, unlike the discredited MM2003 result, it would not make the 15th century the warmest on record.

McIntyre and McKitrick (2005a) [MM2005] continue the criticism of the techniques

CPD
2, 1001–1049, 2006
Millennial temperature
M. N. Juckes et al.

used by MBH1998 and introduce a "hockey stick index" which is defined in terms of the ratio of the variance at the end of a time series to the variance over the remainder of the series. MM2005 argue that the way in which the stage (1) principal component analysis is carried out in MBH generates an artificial bias towards a high "hockey-stick index" and that the statistical significance of the MBH results may be lower than originally estimated. This issue has been investigated at length by Wahl and Ammann (2006) in the context of the MBH1998 reconstruction (back to AD 1400). The problem identified by MM2005 relates to the "standardisation" of the proxy time series prior to the EOF decomposition. Figure 3 assess the impact of different standardisations on the reconstruction back to AD 1000, using only the proxy data used in MBH1999 for their reconstruction from AD 1000. However, rather than reproducing the MBH1999 results, the approach here is to calibrate against the AD 1856 to 1980 northern hemisphere mean temperature instead of the 1902–1980 temperature PCs. If the MBH1999 proxy records are used the orange curve (1) is obtained. Prior to AD 1500 this is close to the published MBH1999 reconstruction (black curve (6)). The relative warmth of the MBH1999 published reconstruction after AD 1500 could result from the use of greater volumes of data for more recent periods of the reconstruction. Curve (1) includes one proxy series which ends before 1980 and has been filled assuming persistence. If this series (ITRDB:FRAN010) is omitted we obtain the cyan dashed curve (2), which lies virtually on top of curve (1). MBH1999 include an adjustment of the first principal component of the North American tree-ring network. This is intended to compensate for a suspected CO_2 fertilisation effect. If the unadjusted first PC is used (purple curve (3)) the reconstruction is significantly cooler throughout the period AD 1000 to AD 1980. The yellow dashed curve (4) replaces the MBH1999 principal components with independently calculated principal components. In this calculation we have omitted 3 tree ring series (out of 28) which were extended using persistence by MBH1999, and standardised on the whole time period rather than on the calibration period. These latter changes make little impact on the result. Finally, the data used for curve (3) have also been processed with the CVM technique, producing the green curve (5). This figure shows that there is considerable sensitivity to the adjustment of the first proxy PC, but little sensitivity either to the way in which the principal components are calculated or to the omission of data filled by MBH1999. Thus, the concerns about the latter two points raised by MM2005 do not appear to be relevant here, though the sensitivity to adjustments of principal component may be a cause for concern.

The code used by MM2005 is not, at the time of writing, available, but the code fragments included in their text imply that their calculation used data which had been centred (mean removed) but had not been normalised to unit variance (standardised). This is discussed by Wahl and Ammann (2006) (see also supplementary material: http://www.clim-past-discuss.net/2/1001/2006/cpd-2-1001-2006-supplement.zip) who show that the effect of replacing the MBH1998 approach with centring and standardising on the whole time series is small, the effect of omitting the standardisation as in MM2005 is much larger: this omission causes the 20th century trend to be removed from the first principal component.

von Storch and Zorita (2005) find that the impact of the standardization modifications suggested by MM2005 on climate reconstructions to be minor. McIntyre and McKitrick (2005b) clarify their original claim, stating that the standardisation technique used by MBH1998 does not create the "hockey-stick" structure but does "steer" the selection of this structure in principal component analysis.

Briffa and Osborn (1999) and MM2005c suggest that rising CO_2 levels may have contributed significantly to the 19th and 20th century increase in growth rate in some trees, particularly the bristlecone pines, but though CO_2 fertilisation has been measured in saplings and strip-bark orange trees (which were well watered and fertilised) (Graybill and Idso, 1993, and references therein) efforts to reproduce the effect in controlled experiments with mature forest trees in natural conditions (Körner et al., 2005) have not produced positive results. MM20005c and Wahl and Ammann (2006) both find that excluding the north American bristlecone pine data from the proxy data base removes the skill from the 15th century reconstructions. MM2005c also suggest that the MBH1998 "North American proxy PC1" is a statistical outlier as far as its correlation to Northern Hemispheric, temperature T_{nh}, is concerned. Table 1 shows, however, that other tree ring series and other proxies have higher anomaly correlations with T_{nh}.

4 Varying methods vs. varying data

One factor which complicates the evaluation of the various reconstructions is that different authors have varied both method and data collections. Here we will run a representative set of proxy data collections through two algorithms: inverse regression and scaled composites. These two methods, and the different statistical models from which they may be derived, are explained in Appendix A.

Esper et al. (2005) investigated the differing calibration approaches used in the recent literature, including regression and scaling techniques, and concluded that the methodological differences in calibration result in differences in the reconstructed temperature amplitude/variance of about 0.5 K. This magnitude is equivalent to the mean annual temperature change for the Northern Hemisphere reported in the last IPCC report for the 1000–1998 period. Bürger et al. (2006) take another approach and investigate a family of 32 different regression algorithms derived by adjusting 5 binary switches, using pseudo-proxy data. They show that these choices, which have all been defended in the literature, can lead to a wide range of reconstructions given the same data. They also point out that the uncertainty is greater when we attempt to estimate the climate of periods which lie outside the range experienced during the calibration period. It is possible, however, that this question is not relevant to the reconstruction of temperatures of the past millennium: the glacier based temperature estimates of OER2005 suggest that the coldest northern hemisphere mean temperatures occurred close to the start of the instrumental record, in the 19th century, and there is a general consensus that the warmest temperatures occur at the end of the instrumental period: this would imply that the instrumental period spanned the entire range of millennial temperatures. The borehole reconstructions, however, imply that there were colder temperatures experienced in the 16th to 18th centuries. The issue is therefore

CPD
2, 1001–1049, 2006
Millennial temperature
M. N. Juckes et al.

not settled.

Several authors have evaluated composites and calibrated those composites against instrumental temperature. Many of the composites contain more samples in later periods, so that the calibration may be dominated by samples which do not extend into the distant past. Here, we will restrict attention to records which span the entire reconstruction period from to AD 1000 to AD 1980 (with some series ending slightly earlier, as discussed below). The data series used are listed in Table 1.

4.1 Reconstruction using a union of proxy collections

The following subsection will discuss a range of reconstructions using different data collections. The first 5 of these collections are defined as those proxies used by JBB1998; MBH1999; ECS2002; MSH2005, and HCA2006, respectively, which extend from AD 1000 to 1980 (1960 for HCA2006). These will be referred to below as, respectively, the JBB, MBH, ECS, MSH, HCA composites to distinguish them from the composites used in the published articles, which include, in some cases, additional, shorter, proxy data series. Finally there is a "Union" composite made using 18 independent northern hemisphere proxy series marked with "*" in Table 1. These series have been chosen on the basis that they extend to 1980 (the HCA composites and the French tree ring series end earlier), the southern hemisphere series have been omitted apart from the Quelcaya glacier data, Peru, which are included to ensure adequate representation of tropical temperatures. The MBH1999 North American PCs have been omitted in favour of individual series used in other studies. Finally, the Polar Urals data of ECS2002, MBH1999 and the Tornetraesk data of MSH2005 have been omitted in favour of data from the same sites used by JBB1998 and ECS2002, respectively (i.e. taking the first used series in each case). This "Union" collection contains 10 tree-ring series, 4 ice-cores, one speleothem (cave deposit) record, two sediment records and a composite record including historical data.

The composite is intended not only to average out regional anomalies but also to average out errors which might be associated with particular proxies or sets of proxies.

It is clear that the proxies are affected by factors other than temperature which are not fully understood. We are carrying out a uni-variate analysis which, by construction, treats all factors other than the one predicted variable as noise.

D'Arrigo et al. (2006) compute reconstructions from tree-rings, predominantly at high latitudes. As with Briffa et al. (1998) they find divergence between growth in these tree-rings and temperature in the late 20th century. Particular concerns have been raised about the Bristlecone Pines (which have a anomalously large positive growth anomaly in the 20th century – the Indian Garden and Methuselah walk series in Table 1) possibly due to CO_2 fertilisation (see Sect. 3) and the high latitude Eurasian trees (which have and anomalously low growth anomaly in the late 20th century – Tornetraesk, Fennoscandia, Yamal, Northern Urals in Table 1) (Briffa et al., 1998). The robustness of the Union reconstruction has been tested by creating a family of 18 reconstructions each omitting one member of the proxy collection. The standard deviation of the maximum pre-industrial temperature in this ensemble of 18 reconstructions is 0.03 K (the standard deviation of individual years is larger, 0.2 K). A chi-squared test on the distribution of the pre-industrial maxima relative to the quartiles of a Gaussian with the same standard deviation gives a value of 4.67 which is exceeded in a chi-squared distribution with 3 degrees of freedom with a probability of 20%. There is thus no significant departure from Gaussian behaviour in the maximum pre-industrial temperature, implying that no single series has excessive influence on the reconstruction.

Inspection of the impact of omitting individual series (see supplementary material: http://www.clim-past-discuss.net/2/1001/2006/cpd-2-1001-2006-supplement.zip) shows that the two with the largest impact (the Grip borehole data and the Arabian Sea sediment record) each alter the reconstruction by upto 0.1 K in the 21 year mean, which is within the uncertainty range estimated from the calibration period residual (0.15 K).

4.2 Intercomparison of proxy collections

Figure 4 shows reconstructions back to AD 1000 using composites of proxies and variance matching (CVM see Appendix A) (for the proxy principal components in the MBH collection the sign is arbitrary: these series have, where necessary, had the sign reversed so that they have a positive correlation with the northern hemisphere temperature record). The "Union" lies in the range spanned by the other reconstructions, with a temperature range about as large as in MSH and ECS. The JBB and MBH collections give a smaller range, while HCA takes an intermediate position. The "Union", however, fits the calibration period data better than any of the sub-collections. This analysis reveals that the choice of proxy records is one reason why different reconstructions show different ranges of temperature variability.

The reconstructions shown in Fig. 5 use the same data as in Fig. 4; this time using inverse regression (INVR) (Appendix A). As mentioned earlier, MBH1998 also used inverse regression, but the method used here differs from that of MBH1998 in using northern hemisphere temperature to calibrate against, having a longer calibration period, and reconstructing only a single variable instead of multiple PCs. The spread of values is substantially increased relative to the CVM reconstruction and all INVR reconstructions show larger temperature ranges than their CVM counterparts.

With INVR, only one reconstruction (ECS) shows pre-industrial temperatures warmer than the mid 20th century. The inverse regression technique applies weights to the individual proxies which are proportional to the correlation between the proxies and the calibration temperature signature. For this time series the 5 proxies are weighted as: 1.5 (Boreal); 1.8 (Polar Urals); 1.7 (Taymir); 1.3 (Tornetraesk); and 2.3 (Upper Wright) (after standardising each series to unit variance on the whole reconstruction period). Firstly, it should be noted that this collection samples North America and the Eurasian Arctic only. The hemispheric coverage is worsened by the weights generated by the inverse regression algorithm, such that nearly half the signal comes from two series located within a few degrees of each other in North America.

The MBH1999 and HPS2000 published reconstructions are shown in Figs. 4, 5 for comparison: the MBH1999 reconstruction often lies near the centre or upper bound of the spread of estimates, while the HPS2000 reconstruction is generally at the lower bound.

Much of the current debate revolves around the level of centennial scale variability in the past. The CVM results for JBB and MBH correspond to a low variance scenario, comparable to MBH1999, while other proxy collections suggest substantially larger variability. The INVR reconstructions have greater variability. It should be noted that the MBH1999 INVR result use greater volumes of data for recent centuries, so that the difference in Fig. 5 between the dashed black curve and the full green curve in the 17th century is mainly due to reduced proxy data input in the latter (there is also a difference because MBH1999 used INVR against temperature PCs rather than Northern Hemisphere mean temperature as here).

Table 2 shows the cross correlations of the reconstructions in Fig. 4, for high pass (upper right) and low pass (lower left) components of the series, with low pass being defined by a 21 year running mean. The low pass components are highly correlated.

The significance of the correlations between these six proxy data samples and the instrumental temperature data during the calibration period (1856–1980) has been evaluated using a Monte-Carlo simulation with (1) a first order Markov model (e.g. Grinstead and Snell, 1997) with the same 1-year lag correlation as the data samples and (2) random time series which reproduces the lag correlation structure of the data samples (see Appendix B2). Figure 6 shows the lag correlations. The instrumental record had a pronounced anti-correlation on the 40 year time-scale. This may be an artifact of the short data record, but it is retained in the significance calculation as the best available estimate which is independent of the proxies. The MSH, HCA and "Union" composites show multi-centennial correlations which are not present in the other data. The JBB composite clearly underestimates the decadal scale correlations, while the MSH, HCA and "Union" composites overestimate it. Results from the significance calculations are shown in Table 3. If the full lag correlation structure of the data were known, it would be true, as argued by MM2005, that the first order Markov approach generally leads to an overestimate of significance. Here, however, we have only an estimated correlation structure based on a small sample. Using this finite sample correlation is likely to overestimate long-term correlations and hence lead to an underestimate of significance. Nevertheless, results are presented here to provide a cautious estimate of significance. For the JBB composite, which has a short lag-correlation, the difference between the two methods is minimal. For other composites there is a substantial difference. In all cases the R values exceed the 99% significance level. When detrended data are used the R values are lower, but still above the 95% level. The Hegerl et al. data have only decadal resolution, so the lower significance in high frequency variability is to be expected.

Figure 7 plots the Union reconstruction, with the instrumental data in the calibration period. The composite tracks the changes in northern hemisphere temperature well, capturing the steep rise between 1910 and 1950 and much of the decadal scale variability. This is reflected in the significance scores (Table 3) which are high both for the full series and for the detrended series. The highest temperature in the reconstructed data, relative to the 1866–1970 mean is 0.25 K in AD 1091. This temperature was first exceeded in the instrumental record in 1878, again in 1938 and frequently thereafter. The instrumental record has not gone below this level since 1986. Taking σ=0.15 K, the root-mean-square residual in the calibration period, 1990 is the first year when the reconstructed pre-industrial maximum was exceed by 2 σ. This happened again in 1995 and every year since 1997. 1998 and every year since 2001 have exceeded the pre-industrial maximum by 3 σ. Two recent years (1998 and 2005) have exceeded the pre-industrial estimated maximum by more than 4 standard errors.

5 Conclusions

There is general agreement that the warmest pre-industrial temperatures of the last thousand years occurred at the start of millennium, and the coolest at some point during the 16th to 19th centuries. There is also general agreement that the warmest pre-industrial temperatures were close to the mean of the 20th century. There remains, however, a range of views about the strength of the negative anomaly in the 16th to 19th centuries: some reconstructions estimate it to have been as much as 0.6 K cooler than the 1900 to 1960 mean, others no more than 0.2 K.

The IPCC2001 conclusion that temperatures of the past millennium are unlikely to have been as warm, at any time prior to the 20th century, as the last decades of the 20th century is supported by subsequent research and by the results obtained here. Papers which claim to refute the IPCC2001 conclusion on the climate of the past millennium have been reviewed and some are found to contain serious flaws. Our study corroborates the IPCC2001 conclusions.

A major area of uncertainty concerns the accuracy of the long time-scale variability in the reconstructions. This is particularly so for timescales of a century and longer. There does not appear to be any doubt that the proxy records would capture rapid change on a 10 to 50 year time scale such as we have experienced in recent decades.

Using two different reconstruction methods on a range of proxy data collections, we have found that inverse regression tends to give large weighting to a small number of proxies and that the relatively simple approach of compositing all the series and using variance matching to calibrate the result gives more robust estimates.

A new reconstruction made with a composite of 18 proxies extending back to AD 1000 fits the instrumental record to within a standard error of 0.15 K. This reconstruction gives a maximum pre-industrial temperature of 0.25 K in AD 1091 relative to the AD 1866 to 1970 mean. The maximum temperature from the instrumental record is 0.84 K in AD 1998, over 4 standard errors larger.

The reconstructions evaluated in this study show considerable disagreement during the 16th century. The new 18 proxy reconstruction implies 21-year mean temperatures close to 0.6 K below the AD 1866 to 1970 mean. As this reconstruction only used data extending back to AD 1000, there is a considerable volume of 16th century data which has not been used. This will be a focus in future research, further information can be found on http://mitrie.badc.rl.ac.uk/.

Appendix A

Regression methods

Ideally, the statistical analysis method would be determined by the known characteristics of the problem. Unfortunately, the error characteristics of the proxy data are not sufficiently well quantified to make the choice clear. This appendix describes two methods and the statistical models which can be used to motivate them.

A1 Inverse regression (INVR)

Suppose x_{ik}, $i=1,N_{pr}$, $k=1,L$ is a set of N_{pr} standardised proxy records of length L and that we are trying to obtain an estimate $\hat{y}_i$ of a quantity y_i which is known only in a calibration period ($i \in C$).

Several "optimal" estimates of y_i can be obtained, depending on the statistical model, i.e. the hypothesised relation between the proxies and y.

Inverse regression follows from the model

$$\beta_i y_k + \mathcal{N} = x_{ik}$$

where $\mathcal{N}$ is a noise process, independent between proxies. It follows that optimal estimate for the coefficients β_i are

$$\hat{\beta}_i = \frac{\sum_{k\in C} x_{ik} y_k}{\sum_{k\in C} y_k^2}.$$

Given these coefficients, the optimal estimate of the y_k (minimising the expected error variance) outside the calibration period is

$$\hat{y}_k = \frac{\sum_i \hat{\beta}_i x_{ik}}{\sum_i \hat{\beta}_i^2}.$$

A2 Composite plus variance matching (CVM)

This method starts out from the hypothesis that different proxies represent different parts of the globe. A proxy for the global mean is then obtained as a simple average of the proxies:

$$\overline{x_k} = N_{pr}^{-1} \sum_i x_{ik}.$$

The statistical model is then:

$$\overline{x_k} = \beta y_k + \mathcal{N}.$$

An optimal estimate of β is easily derived as $\hat{\beta} = \sum_{k \in C} x_k y_k / \sum_{k \in C} y_k^2$. However, $y_k^* = \hat{\beta}^{-1} x_k$ is not an optimal estimate of y_k: because of the added noise, y_k^* is generally an underestimate of y_k. To correct for this we should use:

$$\hat{y}_k = \beta^{-1} \overline{x_k} \sqrt{\left(\frac{\beta^2 \sigma_y^2}{\beta^2 \sigma_y^2 + \sigma_{\mathcal{N}}^2} \right)},$$

where σ_y^2 and $\sigma_{\mathcal{N}}^2$ are the expected variance of y and $\mathcal{N}$, respectively. This leads to an estimate:

$$\hat{y}_k = \overline{x_k} \left(\frac{\sigma_y}{\sigma_x} \right).$$

This is known as the variance matching method because it matches the variance of the reconstruction with that of observations over the calibration period.

Appendix B

Statistical tests

B1 Tests for linear relationships

The simplest test for a linear relationship is the anomaly correlation (also known as: Pearson Correlation, Pearson's product moment correlation, R^2, product mean test):

$$R = \frac{\overline{y'x'}}{\sqrt{\overline{y'^2}\,\overline{x'^2}}}$$

where the over-bar represents a mean over the data the test is being applied to, and a prime a departure from the mean (Pearson, 1896).

Significance calculations depend on having an accurate estimate of the number of degrees of freedom, n, in a sample. Ideally, if the noise affecting all the x and y values is independent, n is simply the number of measurements. This is unlikely to be the case, so an estimate of n is needed. The Monte-Carlo approach is more flexible: a large sample of random sequences with specified correlation structures is created, and the frequency with which the specified R coefficient is exceeded can then be used to estimate its significance.

B2 Lag correlations

Following Hosking (1984), a random time series with a specified lag correlation structure is obtained from the partial correlation coefficients, which are generated using Levinson-Durbin regression.

It is, however, not possible to generate a sequence matching an arbitrarily specified correlation structure and there is no guarantee that an estimate of the correlation structure obtained from a small sample will be realizable. It is found that the Levinson-Durbin regression diverges when run with the lag correlation functions generated from the Jones et al. (1986) northern hemisphere temperature record and also that from the HCA composite. This divergence is avoided by truncating the regression after n=70 and 100 years, respectively, for these two series. The sample lag-correlation coefficients are, in any case, unreliable beyond this point. Truncating the regression results in a random sequence with a lag correlation fitting that specified up to the truncation point and then decaying.

Appendix C

Abbreviations

Table C1 shows a list of abbreviations used in this paper.

Acknowledgements. This work was funded by the Netherlands Environment Assessment Agency (RIVM) as part of the Dutch Scientific Assessment and Policy Analysis (WAB) programme. Additional funding was provided as follows: from the UK Natural Environment Research Council for M. N. Juckes, from the Swedish Research Council for A. Moberg.

References

Adcock, R. J.: A problem in least squares, The Analyst, 5, 53–54, 1878. 1013

Allen, M. R. and Stott, P. A.: Estimating signal amplitudes in optimal fingerprinting, Part I: Theory, Clim. Dyn., 21, 477–491, 2003. 1013

Arrhenius, S.: On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground, Philosophical Magazine, 41, 237–276, 1896. 1003

Bartholin, T. and Karlén, W.: Dendrokronologi i Lapland, Dendrokronologiska Sällskapets Medelanden, 4, 3–16, 1983. 1038

Bauer, E., Claussen, M., Brovkin, V., and Huenerbein, A.: Assessing climate forcings of the Earth system for the past millennium, Geophys. Res. Lett., 30(6), 1276, doi:10.1029/2002GL016 639, 2003. 1014

Bernabo, J. C.: Quantitative estimates of temperature changes over the last 2700 years in Michigan based on pollen data, Quat. Res., 15, 143–159, 1981. 1015
Boninsegna, J. A.: South American dendroclimatological records, Climate Since A.D. 1500, edited by: Bradley, R. S. and Jones, P. D., Routledge, 446–462, 1992. 1038
Bradley, R. S., Hughes, M. K., and Diaz, H. F.: Climate in medieval time, Science, 302, 404–405, 2003. 1008
Briffa, K. R.: Annual climate variability in the Holocene: interpreting the message of ancient trees, Quart. Sci. Rev., 19, 87–105, 2000. 1010, 1014
Briffa, K. R. and Osborn, T. J.: Seeing the wood from the trees, Science, 284, 926–927, 1999. 1019
Briffa, K. R., Jones, P. D., Bartholin, T. S., Eckstein, D., Schweingruber, F. H., Karlén, W., Zetterberg, P., and Eronen, M.: Fennoscandian summers from A.D. 500: temperature changes on short and long time scales, Clim. Dyn., 7, 111–119, 1992. 1010, 1038
Briffa, K. R., Jones, P. D., Schweingruber, F. H., Shiyatov, S. G., and Cook, E. R.: Unusual twentieth-century summer warmth in a 1000-year temperature record from Siberia, Nature, 376, 156–159, 1995. 1038
Briffa, K. R., Jones, P. D., Schweingruber, F. H., Karlén, W., and Shiyatov, S. G.: Tree-ring variables as proxy-climate indicators: Problems with low-frequency signals, NATO ASI Ser, Ser. I, vol. 41, edited by: Jones, P. D., Bradley, R. S., and Jouzel, J., pp. 9–41, Springer-Verlag, New York, 1996. 1010
Briffa, K. R., Schweingruber, F. H., Jones, P. D., Osborn, T. J., Shiyatov, S. G., and Vaganov, E. A.: Reduced sensitivity of recent tree-growth to temperature at high northern latitudes, Nature, 391, 678–682, 1998. 1022
Briffa, K. R., Osborn, T. J., Schweingruber, F. H., Harris, I. C., Jones, P. D., Shiyatov, S. G., and Vaganov, E. A.: Low-frequency temperature variations from a northern tree ring density network, J. Geophys. Res., 106(D3), 2929–2941, doi:2000JD900617, 2001. 1002, 1010
Briffa, K. R., Osborn, T. J., Schweingruber, F. H., Jones, P. D., Shiyatov, S. G., and Vaganov, E. A.: Tree-ring width and density data around the Northern Hemisphere: Part 1, local and regional climate signals, The Holocene, 12, 737–757, 2002. 1010
Bürger, G., Fast, I., and Cubasch, U.: Climate reconstrcution by regression – 32 variations on a theme, Tellus A, 58(2), 227–235, 2006. 1013, 1020
Chbouki, N.: Spatio-temporal characteristics of drought as inferred from tree-ring data in Morocco, Ph.D. thesis, University of Arizona, Tucson, Arizona, 1992. 1038

CPD
2, 1001–1049, 2006
Millennial temperature
M. N. Juckes et al.

EGU

Cook, E. R., Bird, T., Peterson, M., Barbetti, M., Buckley, B., D'Arrigo, R., Francey, R., and Tans, P.: Climatic change in Tasmania inferred from a 1089-year tree-ring chronology of subalpine Huon pine, Science, 253, 1266–1268, 1991. 1038

Cook, E. R., Briffa, K. R., Meko, D. M., Graybill, D. S., and Funkhouser, G.: The 'segment length curse' in long tree-ring chronology development for paleoclimatic studies, The Holocene, 5(2), 229–237, 1995. 1010

Cook, E. R., Esper, J., and D'Arrigo, R. D.: Extra-tropical Northern Hemisphere land temperature variability over the past 1000 years, Quat. Sci. Rev., 23, 2063–2074, 2004. 1011

Cronin, T. M., Dwyer, G. S., Kamiya, T., Schwede, S., and Willard, D. A.: Medieval Warm Period, Global Planet. Change, 36, 17–29, 2003. 1038

Crowley, T. J.: Causes of climate change over the past 1000 years, Science, 289, 270–277, 2000. 1014, 1015

Crowley, T. J. and Lowery, T.: How Warm Was the Medieval Warm Period?, comment on "Man-made versus natural climate change", Ambio, 29, 51–54, 2000. 1002, 1008

Dahl-Jensen, D., Mosegaard, K., Gundestrup, N., Clow, G. D., Johnsen, S. J., Hansen, A. W., and Balling, N.: Past temperatures directly from the Greenland Ice Sheet, Science, 282, 268–271, 1998. 1038

D'Arrigo, R., Wilson, R., and Jacoby, G.: On the long-term context for late twentieth century warming, J. Geophys. Res., 111, D03103, doi:10.1029/2005JD006352, 2006. 1022

Davis, B., Brewer, S., Stevenson, A., and Guiot, J.: The temperature of Europe during the Holocene reconstructed from pollen data, Quat. Sci. Rev., 22, 1701–1716, 2003. 1015

Esper, J., Cook, E. R., and Schweingruber, F. H.: Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability, Science, 295, 2250–2253, 2002. 1002, 1010

Esper, J., Cook, E. R., Krusic, P. J., Peters, K., and Schweingruber, F. H.: Tests of the RCS method for preserving low-frequency variability in long tree-ring chronologies, Tree-Ring Research, 59, 81–98, 2003. 1010

Esper, J., Frank, D. C., and Wilson, R. J. S.: Climate reconstructions - low frequency ambition and high frequency ratification, EOS, 85, 113–120, 2004. 1010

Ferguson, C. W. and Graybill, D. A.: Dendrochronology of bristlecone pine: a progress report, Radiocarbon, 25, 287–288, 1983. 1038

Fisher, D. A., Koerner, R. M., Kuivinen, K., Clausen, H. B., Johnsen, S. J., Steffensen, J. P., Gundestrup, N., and Hammer, C. U.: Inter-comparison of ice core and precipitation records from sites in Canada and Greenland over the last 3500 years and over the last few centuries in detail using EOF techniques, NATO ASI Ser, Ser. I, vol. 41, edited by: Jones, P. D., Bradley, R. S., and Jouzel, J., pp. 297–328, Springer-Verlag, New York, 1996. 1038

Foukal, P., North, G., and Wigley, T.: A stellar view on solar variations and climate, Science, 306, 68–69, 2004. 1014

Gajewski, K.: Late Holocene climate changes in eastern North America estimated from pollen data, Quat. Res., 29, 255–262, 1988. 1015

González-Rouco, F., von Storch, H., and Zorita, E.: Deep soil temperature as proxy for surface air-temperature in a coupled model simulation of the last thousand years, Geophys. Res. Lett., 30(21), 2116, doi:10.1029/2003GL018264, 2003. 1008

Goosse, H., Renssen, H., Timmermann, A., and Bradley, R. S.: Internal and forced climate variability during the last millennium: a model-data comparison using ensemble simulations, Quat. Sci. Rev., 24, 1345–1360, 2005. 1015

Graybill, D. A. and Idso, S. B.: Detecting the aerial fertilization effect of atmospheric CO2 enrichment in tree-ring chronologies, Global Biogeochem. Cycles, 7, 81–95, 1993. 1019

Grinstead, C. and Snell, J.: Introduction to Probability, American Mathematical Society, Providence, Rhode Island, second revised edn., 1997. 1024

Grudd, H., Briffa, K. R., Karlén, W., Bartholin, T. S., Jones, P. D., and Kromer, B.: A 7400-year tree-ring chronology in northern Swedish Lapland: natural climatic variability expressed on annual to millennial timescales, The Holocene, 12, 657–666, 2002. 1038

Gupta, A. K., Anderson, D. M., and Overpeck, J. T.: Abrupt changes in the Asian southwest monsoon during the Holocene and their links to the North Atlantic Ocean, Nature, 421, 354–357, 2003. 1038

Hantemirov, R. M. and Shiyatov, S. G. A.: A continuous multimillennial ring-width chronology in Yamal, northwestern Russia, The Holocene, 12, 717–726, 2002. 1038

Hegerl, G., Crowley, T., Allen, M., Hyde, W., Pollack, H., Smerdon, J., and Zorita, E.: Detection of human influence on a new, validated, 1500 year temperature reconstruction, J. Climate, in press, 2006a. 1002

Hegerl, G., Crowley, T., Hyde, W., and Frame, D.: Climate sensitivity constrained by temperature reconstructions over the past seven centuries, Nature, 440, 1029–1032, 2006b. 1014, 1015

Hosking, J. R. M.: Modeling persistence in hydrological time series using fractional differencing, Water Resour. Res., 20(12), 1898–1908, 1984. 1029

Huang, S.: Merging information from different resources for new insight into climate change in the past and future, Geophys. Res. Lett., 31, L13205, doi:10.1029/2004GL019781, 2004. 1008
Huang, S., Pollack, H. N., and Shen, P. Y.: Temperature trends over the past five centuries reconstructed from borehole temperatures, Nature, 403, 756–758, 2000. 1002, 1007
Hughes, M. K. and Diaz, H. F.: Was there a 'Medieval Warm Period' and if so, where and when?, Clim. Change, 26, 109–142, 1994. 1008
Hughes, M. K. and Funkhouser, G.: Frequency-dependent climate signal in upper and lower forest border trees in the mountains of the Great Basin, Clim. Change, 59, 233–244, 2003. 1038
Jones, P. D. and Mann, M. E.: Climate over past millennia, Rev. Geophys, 42, RG2002, doi:10.1029/2003RG000143, 2004. 1003, 1009, 1017
Jones, P. D., Wigley, T. M. L., and Wright, P. B.: Global temperature variations between 1861 and 1984, Nature, 332, 430–434, 1986. 1003, 1030
Jones, P. D., Briffa, K. R., Barnett, T. P., and Tett, S. F. B.: High-Resolution Paleoclimatic Records for the Last Millennium: Interpretation, Integration and Comparison with Circulation Model Control-Run Temperatures, Holocene, 8, 455–471, 1998. 1002, 1005, 1006
Körner, C., Asshoff, R., Bignucolo, O., Hättenschwiler, S., Keel, S. G., Peláez-Riedl, S., Pepin, S., Siegwolf, R. T. W., and Zotz, G.: Carbon Flux and Growth in Mature Deciduous Forest Trees Exposed to Elevated CO2, Science, 309, 1360–1362, 2005. 1019
LaMarche, V. C.: Frequency-dependent relationships between tree-ring series along an ecological gradient and some dendroclimatic implications, Tree-Ring Bull., 34, 1–20, 1974. 1038
Lara, A. and Villalba, R.: A 3620-year temperature record from Fitzroya cupressoides tree rings in southern South America, Science, 260, 1104–1106, 1993. 1038
Lean, J. L., Wang, Y. M., and Sheely Jr., N. R.: The effect of increasing solar activity on the Sun's total and open magnetic flux during multiple cycles: Implications for solar forcing of climate, Geophys. Res. Lett., 29(24), 2224, doi:10.1029/2002GL015880, 2002. 1014
Little, R. J. A. and Rubin, D. B.: Statistical analysis with missing data, John Wiley and Sons, New York, pp. 278, 1987. 1013
Lloyd, A. H. and Graumlich, L. J.: Holocene dynamics of treeline forests in the Sierra Nevada, Ecology, 78, 1199–1210, 1997. 1038
Mann, M., Ammann, C., Bradley, R., Briffa, K., Jones, P., Osborn, T., Crowley, T., Hughes, M., Oppenheimer, M., Overpeck, J., Rutherford, S., Trenberth, K., and Wigley, T.: On past

temperatures and anomalous late-20th century warmth, Eos Trans. AGU, 84(27), 256–258, 2003a. 1009, 1013

Mann, M., Ammann, C., Bradley, R., Briffa, K., Jones, P., Osborn, T., Crowley, T., Hughes, M., Oppenheimer, M., Overpeck, J., Rutherford, S., Trenberth, K., and Wigley, T.: Response (to comments on "On past temperatures and anomalous late-20th century warmth"), Eos Trans. AGU, 84(44), 473–474, 2003b. 1009, 1013

Mann, M. E. and Jones, P. D.: Global surface temperatures over the past two millennia, Geophys. Res. Lett., 30(15), 1820, doi:10.1029/2003GL017814, 2003. 1002

Mann, M. E., Bradley, R. S., and Hughes, M. K.: Global-scale temperature patterns and climate forcing over the past six centuries, Nature, 392, 779–787, 1998. 1002

Mann, M. E., Bradley, R. S., and Hughes, M. K.: Northern Hemisphere Temperatures During the Past Millennium: Inferences, Uncertainties, and Limitations, Geophys. Res. Lett., 26(6), 759–762, 1999. 1002, 1006

Mann, M. E., Rutherford, S., Bradley, R. S., Hughes, M. K., and Keimig, F. T.: Optimal surface temperature reconstructions using terrestrial borehole data, J. Geophys. Res, 108(D7), 4203, doi:10.1029/2002JD002532, 2003c. 1007

Mann, M. E., Bradley, R. S., and Hughes, M. K.: Corrigendum: Global-scale temperature patterns and climate forcing over the past six centuries, Nature, 430, 105, 2004. 1016

Mann, M. E., Rutherford, S., Wahl, E., and Ammann, C.: Testing the fidelity of methods used in proxy-based reconstructions of past climate, J. Climate, 18, 4097–4107, 2005. 1012, 1013

Manning, M. M., Petit, M., Easterling, D., Murphy, J., Patwardhan, A., Rogner, H.-H., Swart, R., and Yohe, G.: IPCC Workshop on Describing Scientific Uncertainties in Climate Change to Support Analysis of Risk and of Options: Workshop report, Intergovernmental Panel on Climate Change (IPCC), Geneva, 2004. 1006

McIntyre, S. and McKitrick, R.: Corrections to the Mann et al. (1998) Proxy Data Based and Northern Hemispheric Average Temperature Series, Energy and Environment, 14, 751–771, 2003. 1004, 1016

McIntyre, S. and McKitrick, R.: Hockey sticks, principal components, and spurious significance, Geophys. Res. Lett., 32, L03710, doi:10.1029/2004GL02, 2005a. 1017

McIntyre, S. and McKitrick, R.: Reply to comment by von Storch and Zorita on 'Hockey sticks, principal components, and spurious significance', Geophys. Res. Lett., 32, L20714, doi:10.1029/2005GL023089, 2005b. 1019

McIntyre, S. and McKitrick, R.: The M&M critique of the MBH1998 northern hemisphere climate

index: update and implications, Energy and Environment, 16, 69–100, 2005c. 1017
Mitchell, J. F. B., Karoly, D. J., Hegerl, G. C., Zwiers, F. W., Allen, M. R., and Marengo, J.: Detection of climate change and attribution of causes, in: Climate Change 2001: The Scientific Basis, edited by: Houghton, J. T., Ding, Y., Griggs, D. J., Noguer, M., van der Linden, P. J., and Xiaosu, D., Cambridge Univ. Press, New York, 2001. 1003, 1006, 1016
Moberg, A., Sonechkin, D. M., Holmgren, K., Datsenko, N. M., and Karlén, W.: Highly variable Northern Hemisphere temperatures reconstructed from low- and high-reslution proxy data, Nature, 433, 613–617, 2005. 1002, 1011
Naurzbaev, M. M. and Vaganov, E. A.: 1957-year tree-ring chronology of eastern part of Taymir, Siberian Journal of Ecology, 6, 159–168, 1999. 1038
Oerlemans, J. H.: Extracting a Climate Signal from 169 Glacier Records, Science, 308, 675–677, 2005. 1002, 1012
Osborn, T. J. and Briffa, K. R.: The spatial extent of 20th-Century warmth in the context of the past 1200 years, Science, 313, 841–844, 2006. 1009
Osborn, T. J., Briffa, K. R., and Jones, P. D.: Adjusting variance for sample-size in tree-ring chronologies and other regional mean timeseries, Dendrochronologia, 15, 89–99, 1997. 1005
Pearson, K.: Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity and Panmixia, Philosophical Transactions of the Royal Society A, 187, 253–318, 1896. 1029
Pollack, H. N. and Smerdon, J. E.: Borehole climate reconstructions: Spatial structure and hemispheric averages, J. Geophys. Res, 109(D), 11 106, doi:10.1029/2003JD004163, 2004. 1008
Pollack, H. N., Huang, S., and Shen, P. Y.: Climate change record in subsurface temperatures: A global perspective, Science, 282, 279–281, 1998. 1008
Rutherford, S., Mann, M. E., Osborn, T. J., Bradley, R. S., Briffa, K. R., Hughes, M. K., and Jones, P. D.: Proxy-based Northern Hemisphere surface temperature reconstructions: Sensitivity to method, predictor network, target season, J. Climate, 18, 2308–2329, 2005. 1013
Schneider, T.: Analysis of Incomplete Climate Data: Estimation of Mean Values and Covariance Matrices and Imputation of Missing Values, J. Climate, 14, 853–871, 2001. 1013
Serre-Bachet, F., Martinelli, N., Pignatelli, O., Guiot, J., and Tessier, L.: Evolution des températures du Nord-Est de L'Italie depuis 1500 AD. Reconstruction d'après les cernes des arbres, Dendrochronologia, 9, 213–219, 1991. 1038
Shiyatov, S. G.: The upper timberline dynamics during the last 1100 years in the Polar Ural

mountains, edited by: Frenzel, B., pp. 195–203, Fischer, Stuttgart, Germany, 1993. 1038

Soon, W. and Baliunas, S.: Proxy climatic and environmental changes over the past 1000 years, Clim. Res., 23, 89–110, 2003. 1009

Soon, W., Baliunas, S., and Legates, D. R.: Comments on "On past temperatures and anomalous late-20th century warmth", Eos Trans. AGU, 84(44), 473–476, 2003. 1009, 1013

Tan, M., Liu, T., Hou, J., Qin, X., Zhang, H., and Tieying, L.: Cyclic rapid warming on centennial-scale revealed by a 2650-year stalagmite record of warm season temperature, Geophys. Res. Lett., 30, 1617, doi:10.1029/2003GL017352, 2003. 1038

Thompson, L. G.: Ice core evidence from Peru and China, in: Climate Since A.D. 1500, edited by: Bradley, R. S. and Jones, P. D., pp. 517–548, New York, 1992. 1038

von Storch, H. and Zorita, E.: Comment on "Hockey sticks, principal components, and spurious significance: by S. McIntyre and R. McKitrick, Geophys. Res. Lett., 32, L20701, doi:10.1029/2005GL022753, 2005. 1019

von Storch, H., Zorita, E., Jones, J. M., Dimitriev, Y., González-Rouco, F., and Tett, S. F. B.: Reconstructing past climate from noisy data. Science, 306, 679–682., doi:10.1126/science.1096109, 2004. 1012

Wahl, E. R. and Ammann, C. M.: Robustness of the Mann, Bradley, Hughes Reconstruction of Surface Temperatures: Examination of Criticisms Based on the Nature and Processing of Proxy Climate Evidence, Climatic Change, in press, 2006. 1013, 1017, 1018, 1019

Weber, S. L.: A timescale analysis of the NH temperature response to volcanic and solar forcing in the past millenium, Clim. Past, 1, 9–17, 2005. 1014

Weber, S. L., Crowley, T. J., and der Schrier, G. V.: Solar irradiance forcing of centennial climate variability during the Holocene, Clim. Dyn., 22, 539–553, 2004. 1015

Yang, B., Bräuning, A., Johnson, K. R., and Yafeng, S.: General characteristics of temperature variation in China during the last two millennia, Geophys. Res. Lett., 29(9), 1324, doi:10.1029/2001GL014485, 2002. 1038

Table 1. A selection of millennial temperature proxy records used in previous studies and used in the analysis of Sect. 4. The primary reference for each data set is indicated by the superscript in the first column as follows: 1: (Dahl-Jensen et al., 1998), 2: Yang et al. (2002), 3: Shiyatov (1993), 4: Grudd et al. (2002), 5: Gupta et al. (2003), 6: Lloyd and Graumlich (1997), 7: Tan et al. (2003), 8: Thompson (1992), 9: Bartholin and Karlén (1983), 10: Naurzbaev and Vaganov (1999), 11: Briffa et al. (1992), 12: Hantemirov and Shiyatov (2002), 13: Briffa et al. (1995), 14: Lara and Villalba (1993), 15: Fisher et al. (1996), 16: Boninsegna (1992), 17: Cook et al. (1991), 18: Cronin et al. (2003), 19: Chbouki (1992), 20: Ferguson and Graybill (1983), 21: Hughes and Funkhouser (2003), 22: LaMarche (1974), 23: Serre-Bachet et al. (1991). In column 4 three-letter combinations indicate publications in which the data were used (JBB1998; MBH1999; ECS2002; MSH2005, and HCA2005) with the year omitted for brevity. The 18 proxy series marked with a "*" are used in the "Union" reconstruction. The type of proxy is indicated in column 6:: tree-ring [TR], tree-ring composite [TR C], tree-ring principle component [TR PC], sediment [SE], ice core [IC], multi-proxy composite [MC] and speleothem [SP]. R is the anomaly correlation between the proxy record (interpolated to annual resolution where necessary) and the Northern Hemisphere temperature record.

Name	Lat.	Lon.	Id	R	Type
GRIP: borehole temperature (degC) (Greenland)[1]	73	−38	*,MSH	0.67	[IC]
China: composite (degC)[2]	30	105	*,MSH	0.63	[MC]
Taymir (Russia)	72	102	HCA	0.60	[TR C]
Eastern Asia	35	110	HCA	0.58	[TR C]
Polar Urals[3]	65	67	ECS, MBH	0.51	[TR]
Tornetraesk (Sweden)[4]	58	21	MSH	0.50	[TR]
ITRDB [pc01]	40	−110	MBH	0.49	[TR PC]
Mongolia	50	100	HCA	0.46	[TR C]
Arabian Sea: Globigerina bulloides[5]	18	58	*,MSH	0.45	[SE]
Western Siberia	60	60	HCA	0.44	[TR C]
Northern Norway	65	15	HCA	0.44	[TR C]
Upper Wright (USA)[6]	38	−119	*,ECS	0.43	[TR]

CPD
2, 1001–1049, 2006

Millennial temperature

M. N. Juckes et al.

EGU

Table 1. Continued.

Name	Lat.	Lon.	Id	R	Type
Shihua Cave: layer thickness (degC) (China)[7]	40	116	*,MSH	0.42	[SP]
Western Greenland	75	−45	HCA	0.40	
Quelcaya 2 [do18] (Peru)[8]	−14	−71	*,MBH	0.37	[IC]
Boreal (USA)[6]	35	−118	*,ECS	0.32	[TR]
Tornetraesk (Sweden)[9]	58	21	*,ECS	0.31	[TR]
Taymir (Russia)[10]	72	102	*,ECS, MSH	0.30	[TR]
Fennoscandia[11]	68	23	*,JBB,MBH	0.28	[TR]
Yamal (Russia)[12]	70	70	*,MSH	0.28	[TR]
Northern Urals (Russia)[13]	66	65	*,JBB	0.27	[TR]
ITRDB [pc02]	42	−108	MBH	0.21	[TR PC]
Lenca (Chile)[14]	−41	−72	JBB	0.18	[TR]
Crete (Greenland)[15]	71	−36	*,JBB	0.16	[IC]
Methuselah Walk (USA) [ITRDB:CA535][20,21,22]	37	−118	*,MSH	0.14	[TR]
Greenland stack[15]	77	−60	MBH	0.13	[IC]
Morocco, Col du Zad, [ITRDB:MORC014][19]	33	−5	*,MBH	0.13	[TR]
North Patagonia[16]	−38	−68	MBH	0.08	[TR]
Indian Garden (USA) [ITRDB:NV515][20,21,22]	39	−115	*,MSH	0.04	[TR]
Tasmania[17]	−43	148	MBH	0.04	[TR]
ITRDB [pc03]	44	−105	MBH	−0.03	[TR PC]
Chesapeake Bay: Mg/Ca (degC) (USA)[18]	38	−76	*,MSH	−0.07	[SE]
Quelcaya 2 [accum] (Peru)[8]	−14	−71	*,MBH	−0.14	[IC]
France [ITRDB:FRAN010][23]	44	7	MBH	−0.17	[TR]

Table 2. Cross correlations (times 100) between CVM reconstructions from different proxy data bases (see Sect. 4.1). Lower left block correspond to low pass filtered series (21 day mean), upper right to high pass filtered. Note that the HCA proxies are smoothed composites, so the high-pass signal in the reconstruction is reduced.

	MBH	MSH	ECS	JBB	HCA	Union
MBH	–	14	28	63	22	60
MSH	63	–	28	12	7	54
ECS	71	58	–	13	34	57
JBB	66	46	47	–	13	34
HCA	76	69	85	54	–	23
Union	72	95	70	50	78	–

Table 3. R values for six CVM composites evaluated using the Northern Hemisphere mean temperature (1856 to 1980). Columns 2 and 3 show R values for the 95% significance levels, evaluated using a Monte Carlo simulation with 10 000 realisations. In columns 2, 7 and 8 the full lag-correlation structure of the data is used, in column 3 a first order Markov approach is used, based on the lag one auto-correlation. Column 4 shows the R value obtained from the data and column 5 shows the same using detrended data. Column 6 shows the standard error (root-mean-square residual) from the calibration period. Columns 7 and 8 show significance levels, estimated using Monte Carlo simulations as in column 2, for the full and detrended R values. The HCA detrended significance is low because the proxies have been smoothed, removing high frequency information.

Source	$R_{95\|h}$	$R_{95\|AR}$	R	R_{detr}	σ	Signif.	Signif. (detrended)
MBH	0.212	0.176	0.535	0.306	0.174	99.98%	99.46%
MSH	0.269	0.216	0.587	0.365	0.162	99.98%	98.98%
ECS	0.284	0.204	0.599	0.389	0.161	99.98%	98.94%
JBB	0.184	0.178	0.367	0.263	0.203	99.96%	99.36%
HCA	0.339	0.232	0.649	0.406	0.151	99.94%	97.72%
Union	0.276	0.216	0.681	0.427	0.144	99.98%	99.54%

Table C1. Acronyms used in the text.

ABD	Age Band Decomposition tree ring standardisation method
CSM	Climate System Model: A coupled ocean-atmosphere climate model produced by NCAR, http://www.cgd.ucar.edu/csm/
CFM	Climate Field Reconstruction: method for reconstructing spatial structures of past climate variables using proxy data
CVM	Composite plus Variance Matching reconstruction method
ECHO-G	Hamburg coupled ocean-atmosphere climate model
EOF	Empirical Orthogonal Component
INVR	Inverse Regression reconstruction method
IPCC	The Intergovernmental Panel on Climate Change, established by the World Meteorological Organization (WMO) and the United Nations Environment Programme (UNEP) to assess scientific, technical and socio-economic information relevant for the understanding of climate change, its potential impacts and options for adaptation and mitigation. It is open to all Members of the UN and of WMO.
ITRDB	International Tree-Ring Data Bank, maintained by the NOAA Paleoclimatology Program and World Data Center for Paleoclimatology (http://www.ncdc.noaa.gov/paleo)
MWP	Medieval Warm Period
NCAR	National Center for Atmospheric Research (http://www.ncar.ucar.edu)
PC	Principal Component
RCS	Regional Curve Standardisation tree ring standardisation method

CPD
2, 1001–1049, 2006

Millennial temperature

M. N. Juckes et al.

EGU

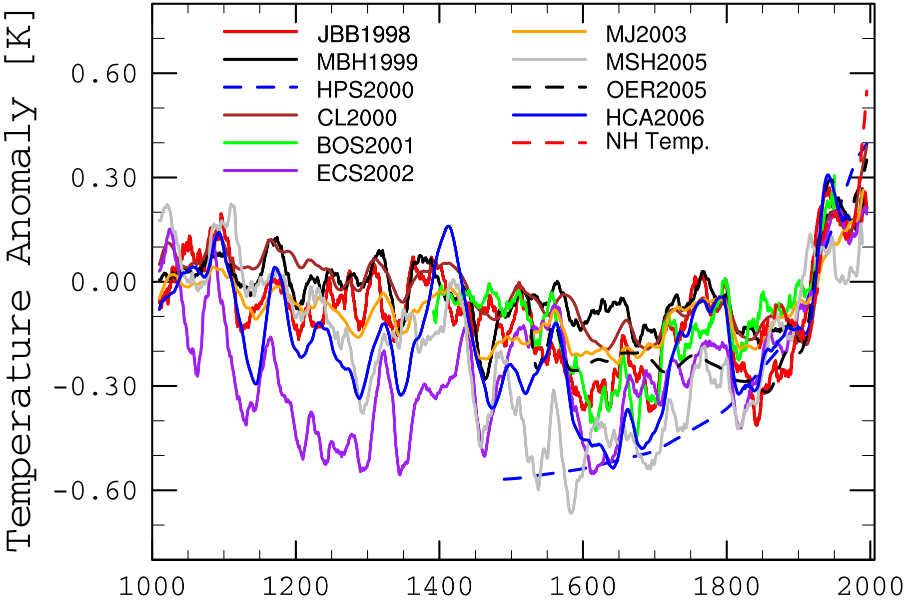

Fig. 1. Proxy based temperature reconstructions from AD 1000 to present, see text for details. With mean of 1900 to 1960 removed, 21-year running means.

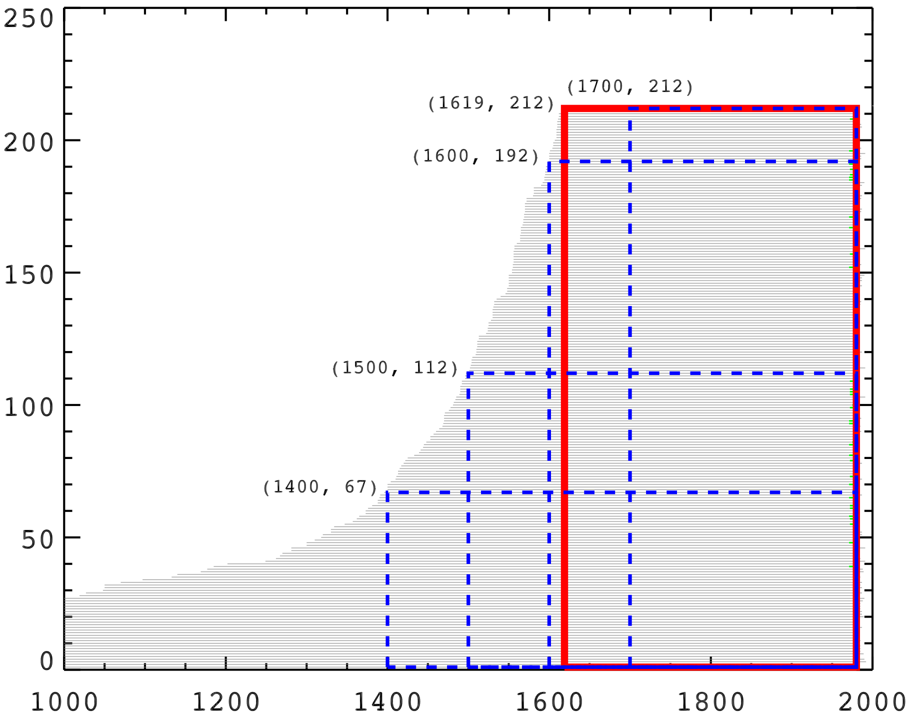

Fig. 2. Data blocks for PC calculation of North American tree ring sites by MBH1998. Each of the 212 data series is shown as a horizontal line over the time period covered. The dashed blue rectangles indicate some of the blocks of data used by MBH1998 for their proxy principal component calculation, using fewer series for longer time periods. The red rectangle indicates the single block used by MM2003, neglecting all data prior to 1619.

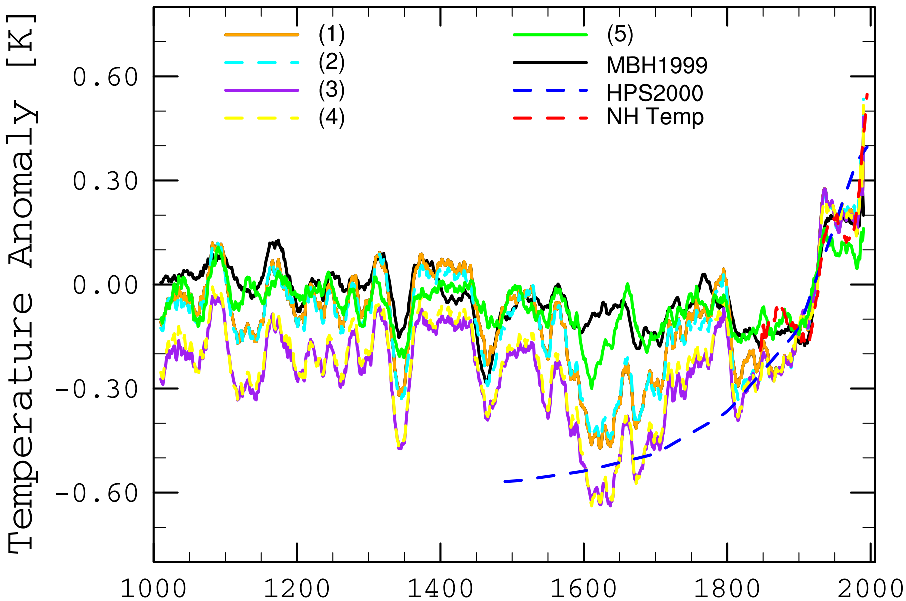

Fig. 3. (1) to (5): Reconstructions back to 1000, calibrated on 1856 to 1980 northern hemisphere temperature, using the MBH1999 proxy data collection. Also shown are: the MBH1999 NH reconstruction, the HPS2000 reconstruction and the the Jones et al. (1986) instrumental data. All data have been smoothed with a 21-year running mean. (1) to (4) are constructed using the INVR technique and differ in the details of the proxies: (1) uses the proxy series as used by MBH1999, including an adjustment to the first Principal Component of the North American tree-rings and filling, with persistence, of the last 7 years of the French tree-ring series (ITRDB:FRAN010), (2) omits the French tree-ring series, (3) as (2), but also using unadjusted PCs, (4) as (2), using recalculated PCs based on 25 series out of the 28 used by MBH1999, omitting 3 that required filling, and standardising on the whole series rather than the calibration period. (5) uses the same data as (3) with the CVM procedure.

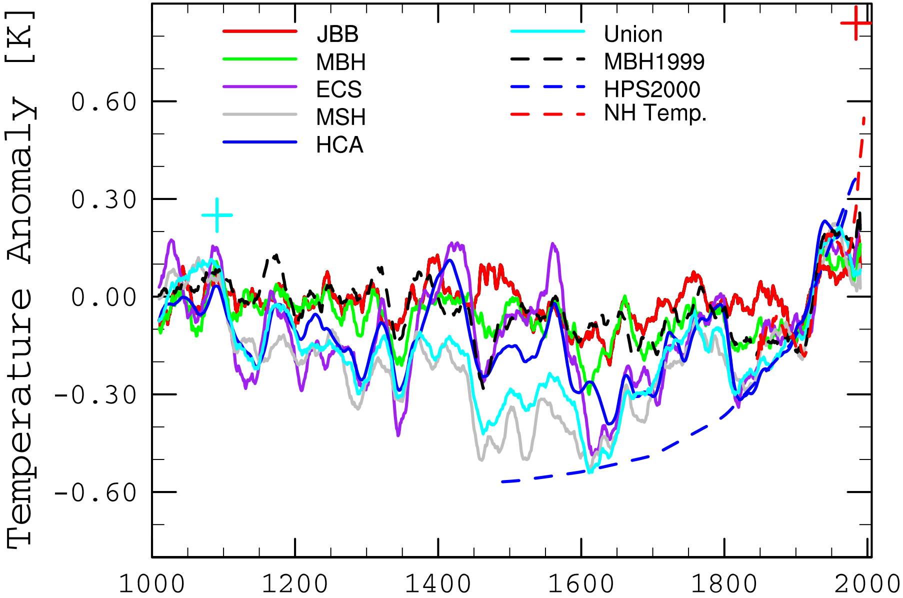

Fig. 4. Reconstruction back to AD 1000, calibrated on 1856 to 1980 northern hemisphere temperature, using CVM, for a variety of different data collections. The MBH1999 and HPS2000 NH reconstructions and the Jones et al. (1998) instrumental data are shown for comparison. Graphs have been smoothed with a 21-year running mean and centred on 1866 to 1970. The maximum of the "Union" reconstruction in the pre-industrial period (0.25 K, 1091 AD) is shown by a cyan cross, the maximum of the instrumental record (0.841 K, 1998 AD) is shown as a red cross.

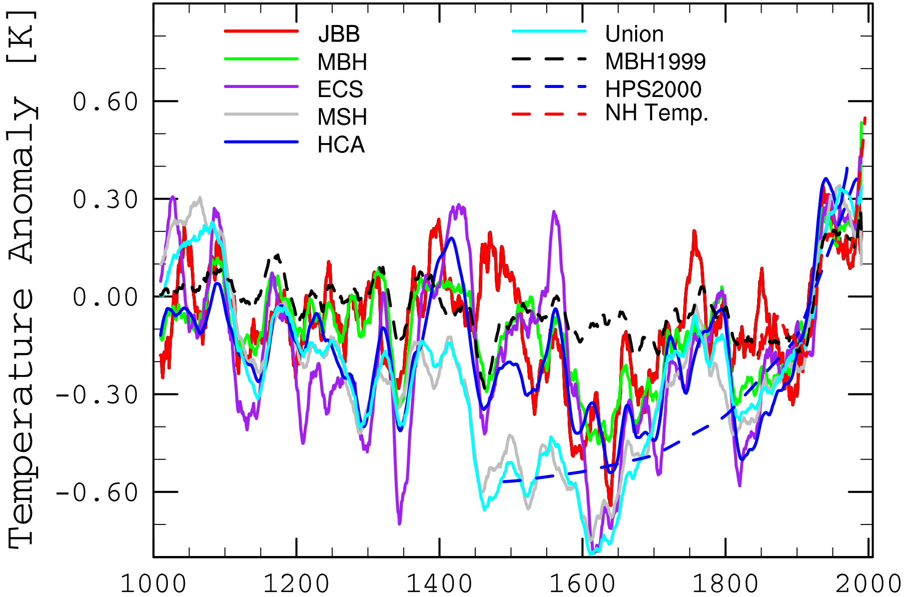

Fig. 5. As Fig. 4, except using inverse regression.

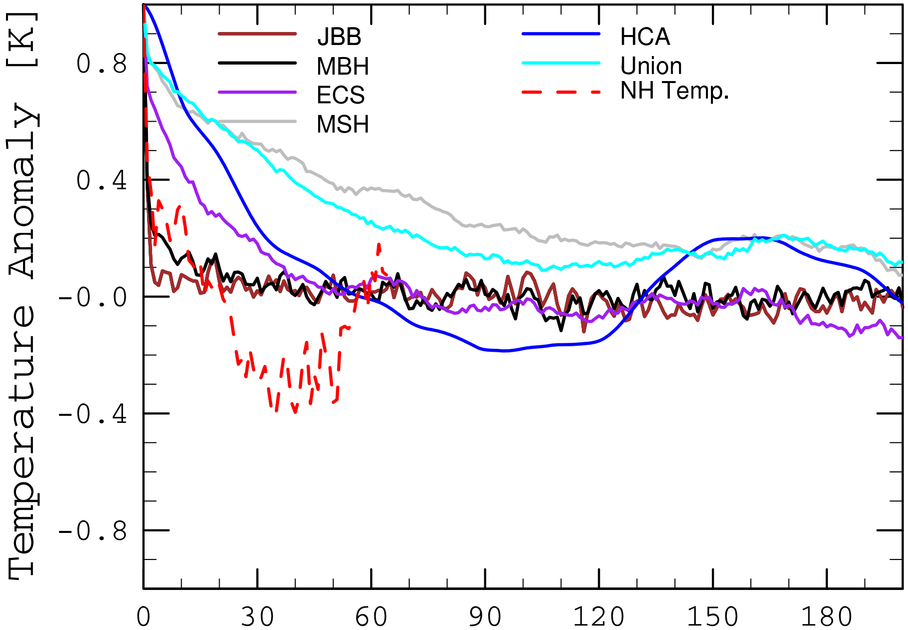

Fig. 6. Lag correlations for proxy composites and instrumental record (red).

CPD
2, 1001–1049, 2006
Millennial temperature
M. N. Juckes et al.

EGU

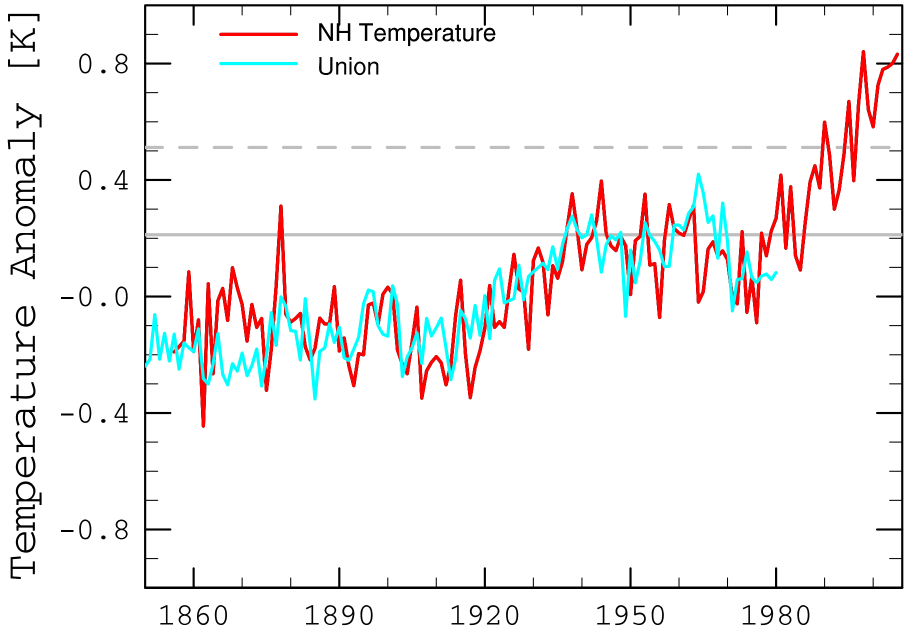

Fig. 7. The "Union" reconstruction, using "composite plus variance matching" (cyan) and the Northern Hemisphere instrumental temperature (red). The pre-industrial maximum of the reconstruction is shown by a grey line, and that value plus 2 σ as a grey dashed line.