The Holocene 8,4 (1998) pp. 455–471

High-resolution palaeoclimatic records for the last millennium: interpretation, integration and comparison with General Circulation Model control-run temperatures

P.D. Jones,[1] K.R. Briffa,[1] T.P. Barnett[2] and S.F.B. Tett[3]

([1]Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK; [2]Scripps Institution of Oceanography, UCSD, La Jolla, CA 92093, USA; [3]Hadley Centre for Climate Prediction and Research, Meteorological Office, Bracknell RG12 2SY, UK)

Received 16 October 1997; revised manuscript accepted 3 February 1998

Abstract: Palaeoclimatology provides our only means of assessing climatic variations before the beginning of instrumental records. The various proxy variables used, however, have a number of limitations which must be adequately addressed and understood. Besides their obvious spatial and seasonal limitations, different proxies are also potentially limited in their ability to represent climatic variations over a range of different timescales. Simple correlations with instrumental data over the period since AD 1881 give some guide to which are the better proxies, indicating that coral- and ice-core-based reconstructions are poorer than tree-ring and historical ones. However, the quality of many proxy time series can deteriorate during earlier times. Suggestions are made for assessing proxy quality over longer periods than the last century by intercomparing neighbouring proxies and, by comparisons with less temporally resolved proxies such as borehole temperatures. We have averaged 17 temperature reconstructions (representing various seasons of the year), all extending back at least to the mid-seventeenth century, to form two annually resolved hemispheric series (NH10 and SH7). Over the 1901–91 period, NH10 has 36% variance in common with average NH summer (June to August) temperatures and 70% on decadal timescales. SH7 has 16% variance in common with average SH summer (December to February) temperatures and 49% on decadal timescales, markedly poorer than the reconstructed NH series. The coldest year of the millennium over the NH is AD 1601, the coldest decade 1691–1700 and the seventeenth is the coldest century. A Principal Components Analysis (PCA) is performed on yearly values for the 17 reconstructions over the period AD 1660–1970. The correlation between PC1 and NH10 is 0.92, even though PC1 explains only 13.6% of the total variance of all 17 series. Similar PCA is performed on thousand-year-long General Circulation Model (GCM) data from the Geophysical Fluid Dynamics Laboratory (GFDL) and the Hadley Centre (HADCM2), sampling these for the same locations and seasons as the proxy data. For GFDL, the correlation between its PC1 and its NH10 is 0.89, while for HADCM2 the PCs group markedly differently. Cross-spectral analyses are performed on the proxy data and the GFDL model data at two different frequency bands (0.02 and 0.03 cycles per year). Both analyses suggest that there is no large-scale coherency in the series on these timescales. This implies that if the proxy data are meaningful, it should be relatively straightforward to detect a coherent near-global anthropogenic signal in surface temperature data.

Key words: Palaeoclimatic reconstruction, proxy climate data, short-term climatic change, high-resolution records, general circulation models, GCM, coupled general circulation models, CGCM, control-run temperatures.

 0959-6836(98)HL258XX

Introduction

Key questions in climatology today include the following. Has climate warmed over the present century due to the net effect of anthropogenic increases in greenhouse gases and sulphate aerosols? If so, what is the magnitude and pattern of the warming (e.g. Santer *et al.*, 1995; 1996; Mitchell *et al.*, 1995; Hegerl *et al.*, 1996; Tett *et al.*, 1996)? Generically referred to as Climate Change Detection (the detection and attribution of climate change to human influences) this issue imparts a particular importance to the study of high-resolution (interannual-to-decadal timescale) reconstructions of late-Holocene climate variability. Reconstructions provide evidence about both the magnitude and patterns of past climate change against which twentieth-century changes can be judged.

In this article, following earlier attempts to investigate the temperature information in selected palaeoclimatic proxies, both in terms of hemispheric averages and spatial patterns (Bradley and Jones, 1993; Barnett *et al.*, 1996), we produce new annual-resolution time series of average hemispheric temperatures for the 'summer' season for both hemispheres. We also compare the principal modes of space/time variability in this group of reconstructions, with similarly defined variability contained in millennial-long control run integrations of Coupled General Circulation Models (CGCMs). Reliable proxy reconstructions are needed to assess CGCM-based simulations of the past 1000 years, both of the inherent natural variability in unforced CGCMs (Barnett *et al.*, 1996; Bradley, 1996) and of simulations with assumed changes to past climatic forcing from solar variations and explosive volcanic eruptions (e.g. Rind and Overpeck, 1994; Lean *et al.*, 1995).

Two recent key papers have brought together information from various proxies and argue for a thorough re-examination of our ideas about the last millennium, particularly with regard to the reality of the so-called 'Medieval Warm Period' (Hughes and Diaz, 1994) and the 'Little Ice Age' (Bradley and Jones, 1993). These works build on and extend earlier pioneering attempts to synthesize climatic history (Lamb, 1977; Williams and Wigley, 1983). Today, it is the search for evidence of anthropogenic climate change that drives such work (Mann *et al.*, 1998).

The need for a better understanding of climate history throughout the world has stimulated a growing research effort and led to tremendous advances in some palaeoclimatic fields and in some regions, helped in no small way by the coordinating research activities such as IGBP's PAGES (PAst global chanGES) program (PAGES, 1995). There is now a need, however, for a wider appreciation and a greater understanding of the potential limitations of some proxies, particularly in the context of attempts to compare palaeoclimatic reconstructions and synthesize results, both continentally and hemispherically.

This paper begins with a review of temperature proxies, specifically considering the spatial limitations and timescale dependence of high-resolution temperature proxies. Our focus on temperatures is justified not only because these are the most widespread data reconstructed, which in turn determines how we characterize the past, but also because, in the climate change detection issue, temperature is the most important variable (Santer *et al.*, 1996). Annually resolved reconstructions from 17 sites around the world, where the reconstructions are readily available, are compared with instrumental data to assess whether some proxies are better than others. Intercomparisons of proxy quality across palaeoclimatic disciplines are rarely made (although Hughes and Diaz, 1994, and Cook, 1995, are exceptions). With an increasing range of proxy climatic data becoming more generally available it is important that uninitiated palaeoclimatic data users (e.g. climate modellers) be made aware of the problems we discuss here.

Averages of the ten NH and the seven SH reconstructions are compared with instrumental data. Comparisons are then made using similar sampling strategies applied to two of the available 1000-year integrations of unforced CGCMs to assess their lower-frequency variability. Finally, some conclusions and recommendations are drawn.

Limitations of palaeoclimatology

Annually resolved proxies

Palaeoclimatology is a diverse science encompassing both an extensive variety of proxy sources and a large range of timescales (Bradley, 1985). Rather than discussing the individual disciplines, we limit this study depending on the timescale resolved and restrict discussion to proxies providing information about past temperatures. There are two distinct fields of palaeoclimatic research:

(a) proxies recording time-series information at annual-to-decadal timescales during the Holocene, generally providing the greatest information over the last 1000 to 2000 years (corresponding to PAGES Stream 1 focus (PAGES, 1995));
(b) markedly poorer temporally resolved proxies, providing smeared snapshots of the past for periods of the order of a millennium (e.g. the mid-Holocene at about 6 K radiocarbon years BP and the Younger Dryas, 11–10 K BP) (equivalent to PAGES Stream 2 focus).

In (a) the strength of the proxy climate 'signal' is usually assessed by regression against instrumental temperatures. The quantitative relationship between the proxy and instrumental data is determined for a 'calibration' period with, ideally, some temperature data withheld to assess the veracity of the relationship with independent data (the 'verification' period) (Briffa, 1995). If the statistical measures are considered adequate the principle of uniformitarianism (e.g. Bradley, 1985) is invoked to reconstruct temperatures from the earlier variability of the proxy.

In (b) the quality of the proxy climatic signal is determined by assemblages of the proxy (e.g. pollen, fauna) in today's climate from a number of locations using transfer functions (Berglund *et al.*, 1996) or the mutual climatic range (Atkinson *et al.*, 1987). Assemblages in earlier periods can then be assessed. An alternate approach is that the proxy provides a direct measure (e.g. isotopic ratios) which are presumed to have a temperature dependence. The greatest limitation in (b) is the extent to which the precise dating of the past assemblages can be established when, for example, this is generally inferred from a number of radiocarbon dates (see Briffa and Atkinson, 1997, for a general discussion).

As our interest here is with the highest possible temporal resolutions, we do not consider (b) any further, focusing only on those temperature-responding proxies which have been used in (a).

Spatial limitations

All high-frequency proxy climatic indicators have spatial limitations that restrict their applicability. Historical archives can only be consulted where written evidence has been preserved. Similarly, natural archives from trees, corals and ice cores are limited to specific regions. Table 1 briefly summarizes regions of the world where temperature reconstructions have been or are potentially available. Table 1 includes instrumental data, with the beginning years of records (see Jones and Briffa, 1992; Parker *et al.*, 1994; Allan *et al.*, 1996; and references therein for further information on the history of instrumental temperature recording). Knowledge concerning the earliest instrumental temperature records in a region is important in all areas of palaeoclimatology, as it is these data that allow the reliability of each proxy record to be directly assessed.

The limitations from Table 1 are relatively obvious and unless new proxy methods are developed these limitations will remain.

Table 1 Spatial limitations (temperature/annually resolved)

Proxy variable	Spatial extent
Instrumental data	Europe from early 1700s, most other coastal regions during nineteenth century. Continental interiors by 1920s, Antarctica by late 1950s.
Contemporary, written historical records (annals, diaries, phenology etc.)	Europe, China, Japan, Korea, Eastern North America. Some potential in Middle East, Turkey and southern Asia and Latin America (since 1500s).
Tree-rings (widths)	Trees growing poleward of 30° or at high elevations in regions where a cool season suspends growth.
Tree-rings (density)	Coniferous trees in NH north of 30°N or at high elevations. From limited application with SH conifers, appears to add little to ring-width information.
Lake varves	High-latitude or high-elevation lakes in regions with high sedimentation rates and pronounced seasonal change.
Ice cores	Greenland, coastal parts of Antarctica, small high-latitude and high-elevation ice caps where little horizontal flow occurs between the top and base of the core.
Ice cores (melt layers)	Coastal Greenland and high-latitude and high-elevation ice caps, where temperatures rise above freezing for a few days each summer.
Marine varves	Marine basins where high sedimentation rates occur. Best location in Gulf of California.
Corals (growth and isotopes)	Tropics (± 30°) where shallow seas promote coral-reef growth. Most studies have concentrated on the massive species *Porites*, and *Pavona* spp. in the Pacific and Indian Oceans and *Montastrea* in the Atlantic.

Apart from a very few high-elevation ice cores, almost all tropical (within 30° latitude of the equator) reconstructions are derived from corals, with even these are limited to shallow seas. Mid- to high-latitude regions provide greater numbers of reconstructions and a greater variety of proxy sources (Bradley, 1985; Bradley and Jones, 1993). In the Southern Hemisphere, it is essential to make good use of the few land areas available.

Timescale limitations

Here we are principally concerned with the ability of a given temperature reconstruction to reproduce faithfully the information that would have been recorded by an instrumental temperature site. Although we can assess the strength of a reconstruction by regression this is generally possible only over the last 100 years. Necessary adjustments to raw data (e.g. removal of growth trends in trees during standardization (Fritts, 1976), allowance for the thinning of ice layers with depth and ice flow and changes of water depth as corals grow) can potentially limit the credibility of a reconstruction for the assessment of longer timescale variations (Briffa, 1995). For example, a 1000-year tree-ring reconstruction may explain 60% of summer temperature variance over the last 100 years. We can be confident in the reconstruction of year-to-year and decade-to-decade variability and in comparing the average of the nineteenth and twentieth or the twelfth and thirteenth centuries, for example. It requires considerable faith, however, in the 'standardization' procedure to be able to compare the twelfth and the twentieth centuries (see Briffa *et al.*, 1996, and Cook, 1995, for more discussion on this aspect with regard to tree-rings and corals).

The timescale limitations addressed here should not be confused with limitations due to age (e.g. living trees or corals only available in a particular region back to 1600) or a specific seasonal response (e.g. temperature-sensitive trees respond primarily to temperatures during the period when the ring is formed). We refer to these as temporal and seasonal limitations (see Jones and Bradley, 1992). The seasonal limitation in some reconstructions may never be overcome but the primary age limitation of proxy variables may often be extended in some regions provided overlapping historical or even subfossil samples are preserved (e.g. wood in lakes or in peat; Briffa *et al.*, 1992a; 1995; Roig *et al.*, 1996). Using extended records pieced together from such samples puts a great emphasis, however, on the appropriateness of standardization procedures and the validity of assumptions of uniform sample selection and response, all of which potentially limit the usefulness of the reconstruction on the longest timescales (Briffa *et al.*, 1996).

Table 2 summarizes potential timescale dependencies for the principal high-resolution palaeoclimatic indicators. We consider each indicator with respect to the information it might reliably provide on three timescales – interannual, decadal-to-century and centennial and longer. With higher-frequency variability, cross-spectral analyses between the reconstruction and the instrumental record can highlight frequencies at which agreement is best. The short length of the instrumental record means that for lower frequencies assessment is generalized and subjective. As with Table 1 we include instrumental temperature data. These *should* provide the best indicators of the recent past but instrumental improvements, changes in sites, exposures, observation times and alterations to the environment around the station (e.g. urbanization influences) all severely impair the ability with which a long record can be used to compare all years with each other (Bradley and Jones, 1985). Climatologists refer to this characteristic as homogeneity (Conrad and Pollak, 1962). Numerous methods have been developed to assess homogeneity, both of the mean and the variance, of a temperature series (e.g. Bradley and Jones, 1985; Karl and Williams, 1987; Alexandersson, 1986; Rhoades and Salinger, 1993; Peterson and Easterling, 1994; Alexandersson and Moberg, 1997). Most methods require neighbouring stations or a reference climate series (Easterling and Peterson, 1995) with which to test a candidate station. In general, assessment of homogeneity becomes more difficult the longer the record, as there will be fewer (and more distant) neighbours in the earliest years. Potential problems related to changes in exposure and observation times mean that it is harder to ensure the absolute homogeneity of daily extreme temperatures compared to monthly means.

The timescale dependence in historic and natural archives is only partly alluded to in most reconstructions. For example, it is often stated in tree-ring-based reconstructions that the standardization procedure has limited or even curtailed the usefulness of the reconstruction beyond certain timescales (e.g. Briffa *et al.*, 1988; 1990; Lara and Villalba, 1993). In dendroclimatology, therefore, the standardization procedure is crucial. The implications of this in any interpretation should be clearly stated. For historical records and ice cores, comparison with instrumental data even on the interannual timescale is often not performed. With historical records this often stems from the type of information used in the pre-instrumental period being less readily, or just not, available during the instrumental era. In most European cases the record from historical sources is blended into modern instrumental records along with early fragmentary instrumental records which could have been used to verify the longer descriptive type of information. It is argued (e.g. Pfister, 1992) that this makes best use of available information. While this might be true, any assessment of the long-term homogeneity of the reconstruction can be potentially compromised. For the other main region with long historical records (China, Japan and Korea) equivalent data, to that available in earlier centuries, is available for the twentieth century

Table 2 Timescale dependence of annual-to-decadal scale temperature proxies

Proxy variable	Interannual	Decadal-to-century	Centennial and longer
Instrumental records	Properly maintained should be 'perfect'. Able to assess changes on daily, monthly and seasonal timescales.	Site moves, observation time changes and urbanization all provide non-insurmountable problems. Should provide reality. Able to assess changing frequency of extremes.	As decadal-to-century, but the rates of change to instrumentation, sites and urbanization will make it increasingly difficult to maintain absolute levels.
Contemporary written historical records (diaries, annals etc.)	Depends on function of diary information (freeze dates, harvest dates and amounts, flowering dates, snowlines etc.). Very difficult to compare with instrumental material.	Dependent on diary length and observer age. Lower frequencies increasingly likely to be lost due to human lifespan, because of subjective nature of much of the information.	Only a few indicators are objective and might provide comparable information (snow lines, rain days, number of frosts).
Tree-rings (widths)	Generally dependent upon the growing season months. Exact calendrical dates determined by cross-dating.	Standardization method potentially compromises the interpretation on longer timescales.	Highly dependent on standardization method. Likely to have lost some variability, but difficult to assess.
Tree-rings (density)	Dependent upon late spring and summer season months. Exact calendrical dates determined by cross-dating.	As ring widths, but the standardization method can seriously influence any interpretation.	As ring widths, but even more likely to have lost lowest frequencies.
Ice core (isotopes)	Dependent on moisture source and its temperature, travel distance and time and temperature during snowfall. Rarely compared to instrumental records. Dating dependent on layer counting which will become increasingly difficult with depth.	Prone to changes in moisture source and precipitation amount. Difficult to assess the value of the record, except by comparison to other cores.	Increasingly dependent upon any flow model and layer compaction. Veracity can be assessed using other cores and 'borehole' comparisons.
Ice core (melt layers)	Dependent on summer warmth. Unable to distinguish cold years which cause no melt. Rarely compared to instrumental records. Dating dependent on layer counting which will become increasingly difficult with depth.	May not respond to full range of temperature variability. Too warm implies the whole layer will melt. Too cold and there will be no melt layers.	Increasingly dependent upon any flow model and layer compaction. Veracity can be assessed using other cores.
Corals (growth and isotopes)	Response to annual and seasonal water temperature and salinity changes. Dating dependence upon counting. Rarely cross-dated.	As coral head grows, low frequency aspects may be affected by amount of sunlight, water depth, nutrient supply etc.	Not yet achieved, except in a few cases. Veracity can be assessed by comparison with other corals.

(Wang and Zhao, 1981; Zhang, 1994). Potential problems still exist because of short instrumental records, and for temperature reconstructions because much of the historical material relates to floods and droughts.

Ice cores are almost always from remote regions, where instrumental records are often short and some distance from the location of the drilling site. In many ice cores, isotope series are used to infer that periods and years in the past were warmer or colder than conditions now. The large changes that occur on ice-age timescales are clearly indicative of dramatic increases/decreases in temperature but interpretation over the recent Holocene, particularly if we hope to use the information at the interannual timescale, requires some quantification. Improvements in calibrating ice-core isotopic records are under way, particularly in Greenland, where long temperature records at coastal sites are available (e.g. Fisher and Koerner, 1994; Fisher *et al.*, 1996) and elsewhere (Peel, 1992; Peel *et al.*, 1996; Jones *et al.*, 1993; Yao *et al.*, 1996).

Possible methods of assessing timescale limitations

With all proxies, uncertainties become greater as the timescale increases. There are two possibilities which might improve credibility of the proxy reconstructions on centennial timescales. First, as the timescale increases the number of effectively independent locations (Neff) over the surface of the Earth decreases. Jones *et al.* (1997) estimate from instrumental and long CGCM control runs for annual data on decadal timescales that Neff is of the order 20. This implies that even for relatively distant reconstructions some agreement might be expected on centennial timescales. The first possibility is, therefore, direct comparison of neighbouring records, preferably with records from different proxies, although seasonal dependencies may limit the extent to which this might be achieved. The availability of many reconstructions at the World Data Center-A for Paleoclimatology (WDC-A) in Boulder, CO., USA, now makes this increasingly feasible (see http://www.ngdc.noaa.gov/paleo/paleo.html).

The second possibility is to use the reconstruction to reproduce features of the past recorded by more slowly responding indicators. Boreholes have been used to reconstruct past temperatures (Deming, 1995) but the timescale lengthens as a function of depth and hence age. The reconstruction is estimated by inverting the temperature variations with depth to reproduce past changes in temperature at the surface. To avoid local effects with some holes, regional averages from a number of holes have begun to be analysed (Pollack *et al.*, 1996). As an alternative, annually resolved reconstructions (from trees, for example) could be used as input, and the derived diffusion with depth compared with borehole measurements. One of the only examples of this kind of approach is Beltrami *et al.* (1995). A potential drawback of such an approach is that the proxy series might respond more to growing season temperatures while the borehole temperatures are indicative of annual conditions, of which winter is generally more important.

Related opportunities exist for model simulations of glacier advances and retreats using annually resolved reconstructions as input. Raper *et al.* (1996), for example, use tree-growth reconstructions of summer temperature from northern Fennoscandia to simulate ice volume in the Storglaciären in northern Sweden. The time series of their volume estimate was shown to compare favourably with evidence of the timing of most glacier advances and retreats derived from terminal and lateral moraine positions. Some previous studies have assessed temperature changes on the basis of glacier movements, but their conclusions have largely been inferential rather than quantitative (e.g. Williams and Wigley, 1983; Villalba, 1990; Luckman *et al.*, 1997).

While the reasons for spatial limitations are easy to comprehend, the implications of temporal and particularly timescale limitations are only just beginning to be realized. The limitations are generally clear to exponents in each field but as their importance differs from proxy to proxy the relevance is sometimes overlooked. This issue becomes more important as the reconstructions produced are being widely used by climatologists and modellers. It is incumbent on all palaeoclimatologists to make users fully aware of potential problems.

Annually resolved palaeoclimatic evidence

Initial analyses of the reconstructions

Table 3 lists 17 annually resolved reconstructions of temperature, giving details of their sources. The 17 were chosen based on availability of the reconstructed data and they provide a representative sample of different proxies from widely separated regions. Additional tree-ring material is available for much of northern Europe and North America. Much is shorter in length than 300 years and here we have chosen only the longest series, concentrating on density information, where available, as this has been shown to be a very sensitive indicator of summer temperatures (Briffa *et al.*, 1996). For corals all the long available records have been used. For ice cores several more records are available, particularly for Greenland and Antarctica. We have chosen a variety of records (isotopes and melt layers), selecting those where dating is good or replication is available. Tropical ice caps (e.g. Thompson, 1996; Yao *et al.*, 1996), although potentially useful, were not considered because at their elevations they are not representative of surface conditions in the tropics (see also Hurrell and Trenberth, 1996). Chinese historical temperature reconstructions, although potentially very useful, are only available on the decadal timescale (Bradley and Jones, 1993).

The series listed in Table 3 are grouped into three latitude zones: northern and southern mid- to high-latitude (comprising trees, ice cores and historical/instrumental evidence) and tropical (all based on corals) reconstructions. All records are longer than 350 years. The three coral reconstructions are of 'annual' temperatures (i.e. the average of 12 monthly values, but not calendar years) while all the others are principally indicative of 'summer' conditions (see Table 3).

Not all the original publications describing these palaeoclimatic series state how well the proxy compares with instrumental temperatures. In Table 4 we have compared each proxy series with instrumental data averaged over the appropriate season (see Table 3), using 5° × 5° grid-box temperature anomalies (Jones, 1994; Parker *et al.*, 1994; Nicholls *et al.*, 1996). The principal reason for using the 5° × 5° grid-box data is that we are mainly interested in large-scale temperatures, as opposed to those of only local significance. In the later comparisons with the CGCM data, the models simulate large-scale as opposed to local-scale features. The implications of this are discussed in the next sections.

Correlations are shown for the common period 1881–1980 along with the ratio of the standard deviation of the proxy to the instrumental, given on the interannual and decadal (maximum number for comparison of 10) timescales. In selecting the 5° × 5°

Table 3 Abbreviation, location, palaeotype and source of the 17 palaeoclimatic series used

	Name	Abbreviation	Latitude[1]	Longitude[1]	Palaeotype[2]	Season[3]	Source	Gp[5]
1	N. Fennoscandia	NFS	68.0°N	22.0°E	D	AMJJA	Briffa *et al.* (1992b)	2
2	N. Urals	NUR	66.0°N	65.0°E	D	MJJAS	Briffa *et al.* (1995)	2
3	Jasper	JAS	52.3°N	117.0°W	D	AMJJA	Luckman *et al.* (1997)	1
4	Svalbard	SVA	79.0°N	15.0°E	M	JJA	Tarussov (1992)	1
5	Central England	ENG	52.0°N	2.0°E	T	JJA	Manley (1974)[4]	1
6	Central Europe	EUR	46.5°N	8.0°E	H	JJA	Pfister (1992)	1
7	S. Greenland	SGR	65.0°N	45.0°E	M	JJA	Kameda *et al.* (1992)	1
8	N. Tree-line	NTR	58.0°N	95.0°W	R	J-D	D'Arrigo and Jacoby (1992)	1
9	W. USA	WUS	45.0°N	115.0°W	D	MJJAS	Briffa *et al.* (1992a)	1
10	Crete	CRT	71.0°N	36.0°W	I	JJA	Fisher *et al.* (1996)	2
11	Tasmania	TAS	42.0°S	146.5°E	R	NDJFMA	Cook *et al.* (1992)	3
12	Lenca	LEN	41.5°S	72.6°W	R	DJFM	Lara and Villalba (1993)	3
13	Alerce	ALE	41.2°S	71.8°W	R	DJF	Villalba (1990)	2
14	Law Dome	LAW	66.7°S	112.8°E	I	NDJFMA	Morgan and van Ommen (1997)	1
15	Great Barrier Reef	GBR	20.0°S	150.0°E	CC	July–June	Lough and Barnes (1997)	1
16	Galapagos	GAL	0.0°	91.0°W	CI	Oct–Sept	Dunbar *et al.* (1994)	1
17	New Caledonia	NCL	22.0°S	166.0°E	CI	July–June	Quinn *et al.* (1998)	1

[1]Both these are approximate for some sites.
[2]D = density and ring widths; M = melt layers in ice core series; I = isotopic values in ice core series; R = ring widths; T = temperature measurements; H = historical temperature; CC = coral calcification; CI = coral isotopic values.
[3]Months; J–D is annual.
[4]Updated in Parker *et al.* (1992).
[5]Group membership for the cross-spectral analyses (see text) (Group 3 members are in Group 2 and Group 2 and 3 members of Group 1).

Table 4 Comparisons between proxy and instrumental data series

		1	2	3	4	5	6	7	8
1	NFS	0.79	100	0.65	0.80	10	0.72	1876–1975	0.74
2	NUR	0.81	100	0.94	0.92	10	1.05	1882–1980	0.82
3	JAS	0.48	96	0.66	0.45	10	0.92	1891–1982	0.42
4	SVA[9]	0.08	82	–	0.38	9	–	N/A	N/A
5	ENG	0.84	100	1.49	0.80	10	1.70	N/A	N/A
6	EUR	0.90	99	1.39	0.83	10	1.26	N/A	N/A
7	SGR	0.17	100	0.69	–0.28	10	0.65	N/A	N/A
8	NTR	0.34	75	0.64	0.87	8	1.49	1880–1974	0.73
9	WUS	0.60	100	0.57	0.79	10	0.41	1881–1982	R
10	CRT	0.30	99	1.09	0.49	10	0.62	N/A	N/A
11	TAS	0.42	100	0.70	0.58	10	0.67	1886–1989	0.57
12	LEN	0.36	82	0.85	0.55	8	0.91	1910–87	0.61
13	ALE	0.35	82	0.90	0.16	8	1.11	1908–84	0.61
14	LAW[9]	0.26	25	–	0.98	3	–	N/A	N/A
15	GBR	0.18	92	1.14	0.52	10	0.88	1906–82	0.31
16	GAL	0.39	92	0.72	0.16	10	1.08	1936–53	0.66
17	NCL	0.41	85	0.46	0.48	9	0.48	*	*

[1] Correlation coefficient calculated over 1881–1980.
[2] Number of years
[3] SD (Proxy)/SD (Observed) over 1881–1980.
[4] Correlation coefficient calculated for 10 decades over 1881–1980.
[5] Number of decades.
[6] SD (Proxy)/SD (Observed) for decades 1881–1980.
[7] Period used in calibration in original publication (where given).
[8] Correlation coefficient (interannual) given in original publication.
[9] Record never formally calibrated against instrumental temperature.
* Not known.
N/A – no formal calibration.
R – regional average, with only the component correlations given.

grid boxes we have chosen just one if the proxy series represents a small region or a number of grid boxes if the original authors were considering a large region (e.g. numbers 8 and 9 in Table 3). Correlation coefficients and the years used in the original publications are included, where these were given.

Discussion of Table 4

In general, the correlations between the various proxy series and instrumental temperatures are lower than values in the original published sources. Decadal correlations are also slightly lower than values given in Barnett *et al.* (1996). There is no single reason for the differences in the correlations achieved here and the original publications. There are three principal factors:

(1) The observed data are in most cases different from those used in the original publications. We have used 5° × 5° grid-box values and selected data for either a single grid box or a larger regional average based on a number of boxes depending on the spatial representativeness of the proxy. The series incorporate all available homogeneous station data for land areas and sea-surface temperature (SST) anomalies over the ocean. The correlations given in some of the published sources often relate to a single station (generally the nearest), whereas those listed in Table 4 almost invariably relate to averages incorporating more than one station record. The grid-box temperatures, therefore, are broader-scale indicators of temperature than individual sites. We are interested in the larger scales both in the proxy assessment and in the later comparison with the CGCMs. For coastal sites (ENG and the three ice-core series SVA, SGR and CRT), the grid-box temperatures used will have incorporated some SST data. Similarly, for the three coral sites (GBR, GAL and NCL) some land-based temperatures have been incorporated into the grid boxes used. Our correlations are, therefore, a truer indication of the proxy's usefulness on the large scale, particularly when comparisons with CGCMs are made as these will not include any sub-grid-box information.

(2) In the original sources the periods over which the correlation coefficients are calculated differ (see Table 4). Here a consistent 1881–1980 period is used. Correlation coefficients are relatively sensitive to a few outliers and it is possible that a slight difference in the analysis period can give a different result. It would perhaps be informative if running correlation coefficients (for example, using a 30-year window) or robust correlation coefficients were given in future.

(3) Related to (2) some workers quote correlation coefficients where outliers have been excluded (e.g. Lara and Villalba, 1993). This can be informative, but the true value should always be stated to avoid giving a false sense of the reliability of a proxy in earlier years (Cook *et al.*, 1992).

Although the results of Table 4 clearly provide one simple objective means of assessing the quality of palaeoclimatic reconstructions, they hide problems related to the quality and homogeneity of the proxy over the period before instrumental data (see above). For example, the EUR record appears to be an excellent temperature proxy on the evidence in Table 4, but it is a mix of historical, documentary and instrumental data with the latter dominating after 1850 (Pfister, 1992). The diary information becomes much less plentiful after the development of instrumentation in the late 1700s. Thus, even though the utmost care has been taken in its development, does the Table 4 correlation based on the last 100 years provide a true indication of its long-term usefulness? Even the quality of the Central England (ENG) record must decrease slightly in the years before about 1720, when it is based on some instrumental readings but also on documentary information and inferences from temperature proxies (such as the freezing of Dutch canals and British rivers; Manley, 1974). The most surprising correlation in Table 4 is that for ENG. The value is relatively low because the grid-box series (50–55°N, 0–5°W) incorporates up to ten station records and some SST data while ENG is based on only three inland stations, thus highlighting the potential importance of differences between the grid-box and local scales.

Rejecting any of the 17 proxies on the basis of correlation values in Table 4 (either those calculated here using 1881–1980 or those given in the original sources) is therefore not as simple as it might seem. Particularly for the ice-core and coral series, a number of arguments might be used to explain the apparently poor results.

(1) For both types of proxy, the instrumental data available (SST for corals and remote instrumental sites for ice cores) are possibly of poorer quality compared to the rest of the instrumental temperature data base because of their location. The grid boxes used generally have only one instrumental station whose local scale influences will affect the whole grid box. This was partly overcome for the ice-core sites by using several grid boxes as representative for the ice-core location. This is a much more compelling argument for the ice-core proxies than for the corals. Here, the spatial coherency of SST data (e.g. Jones *et al.*, 1997) implies that even a few values somewhere in the $5° \times 5°$ grid box should provide adequate information about the sea temperature around the coral (Parker *et al.*, 1994). The coral correlations are, therefore, particularly disappointing given the known SST coherency.

(2) The season used for comparing temperatures with both ice cores and corals might have been better chosen. In some sources the definition of the year is not given. In the SH tropics the 12 months chosen may be inappropriate. Corals grow all the year round, but may be biased to the 'summer' season (Barnes and Lough, 1996). For all the NH ice-core series, the summer (JJA) was chosen. This is clearly the correct choice for the two ice melt-layer series (SVA and SGR) as no melting occurs outside of the summer. These features may be more indicative of high summer temperatures but in these regions correlations between a July/August and a JJA average would be high anyway. It is possible that the melt layer is produced from a few extreme daily temperatures which can even occur in an otherwise cool summer. Instrumental data being collected from automatic weather stations (AWSs) can be used to assess the strength of such relationships between extreme days and the average summer temperature. Isotope series in ice cores can only record the 'temperature' when it is snowing – the 'temperature' being a function of the moisture source temperature, the present temperature and the travel time (Bradley, 1985). Again, more experimentation with instrumental data from AWSs is vital to assess both the strength of relationships with mean summer temperatures and the times of the year most snowfall occurs. Given the resources and attention devoted to ice-core studies, too few analyses of this kind have been undertaken to test the validity of what is often perceived as a simple correspondence between isotopic measurements and some temperature quantity.

(3) Dating of the annual growth rings in trees is absolute, as it is assured by cross-dating (Fritts, 1976). For both ice cores and corals exact dating can not be guaranteed, and even when multiple cores have been taken and analysed the agreement between cores can be relatively low (e.g. Barnes and Lough, 1996; Lough *et al.*, 1996; Fisher and Koerner, 1994; Fisher *et al.*, 1996). This is suggestive of a weaker common signal compared to, for example, related statistics calculated in dendroclimatology (Briffa, 1995). This implies that dating inaccuracies will be less important on decadal timescales, assuming the dating errors are random, so correlations should improve on longer timescales. However, errors may be more additive as it would seem much more likely for layers to be missed than 'false' layers identified.

If some series cannot be rejected, what general conclusions can be made from Table 4? Excluding ENG and EUR, the best proxy temperature reconstructions are clearly tree-ring-based records, particularly those incorporating density information from high-latitude and high-altitude conifers in the Northern Hemisphere. In the Southern Hemisphere, dendroclimatic records are still slightly superior to the other proxies, although the differences are smaller.

The fact that most dendroclimatic reconstructions show higher correlations with instrumental data compared to corals and ice cores may come as something of a surprise to non-palaeoclimatologists. Both ice cores and corals rely to varying extents on isotopic measurements to assess past temperatures (Fisher *et al.*, 1996; Lough *et al.*, 1996). In the past, there has been a tendency to assume that all isotope/temperature relationships explain high percentages of the climate variance. As can be seen from these few examples, this is clearly not the case. In coral-isotope studies, calibration exercises have frequently included the annual cycle clearly inflating correlation coefficients (Lough and Barnes, 1997). If coral and ice cores are to be used as reliable indicators of past temperatures greater emphasis in future work needs to be placed on determining exactly what is being measured. More detailed quantitative assessments of temperature/isotope relationships need to be undertaken using 'grid-box' scale temperatures. A recent example for ice cores is Yao *et al.* (1996).

Initial comparisons of the reconstructions

The 17 reconstructions are plotted as yearly values with 50-year Gaussian smoothed curves superimposed for each of the three zones in Figures 1 (NH > 40°N), 2 (Tropics) and 3 (SH > 40°S). All of the reconstructions have been normalized over the period 1901–50 to allow comparison on the same scale. The most striking feature of these plots is that century and longer timescale variability is considerably greater in the NH > 40°N zone (Figure 1) reconstructions than in the other two zones. In the two South American tree-ring series this may be, at least partly, due to the method used to remove growth trends in the raw ring-width data (standardization) which has also removed longer timescale climatic variability. Lara and Villalba (1993) used residuals from 128-year smoothing splines which means that there can be little variance remaining on century timescales. Longer timescale variability is, however, less apparent in Tasmania, Law Dome and in the tropical corals. In the NH > 40°N zone longer-timescale variability is relatively low in the records for ENG, EUR, CRT and WUS. Only in the last of these records can this be explained by standardization (Briffa *et al.*, 1992a). The use of the 'borehole' approach to estimating temperatures from ice cores reveals much more longer timescale variations over Greenland than is seen in the CRT record (Dahl-Jensen *et al.*, 1997).

Averaging the reconstructions by hemispheres

We have produced (Figure 4) large-scale averages of the reconstructions in each hemisphere, NH10 and SH7 and the globe (GL17, not shown). When fewer than the full number of reconstructions were available the variance of the average was reduced according to the method described in Osborn *et al.* (1998). Although this corrects the variance for the fewer series in the earlier centuries it must be remembered that for much of the time before 1500 the NH average is composed of only four series and the SH series only three. As would be expected from the earlier discussion, NH10 exhibits more century timescale variability, compared to SH7. The 50-year smoothed curve for NH10 clearly shows the twentieth century warmer than all centuries with tem-

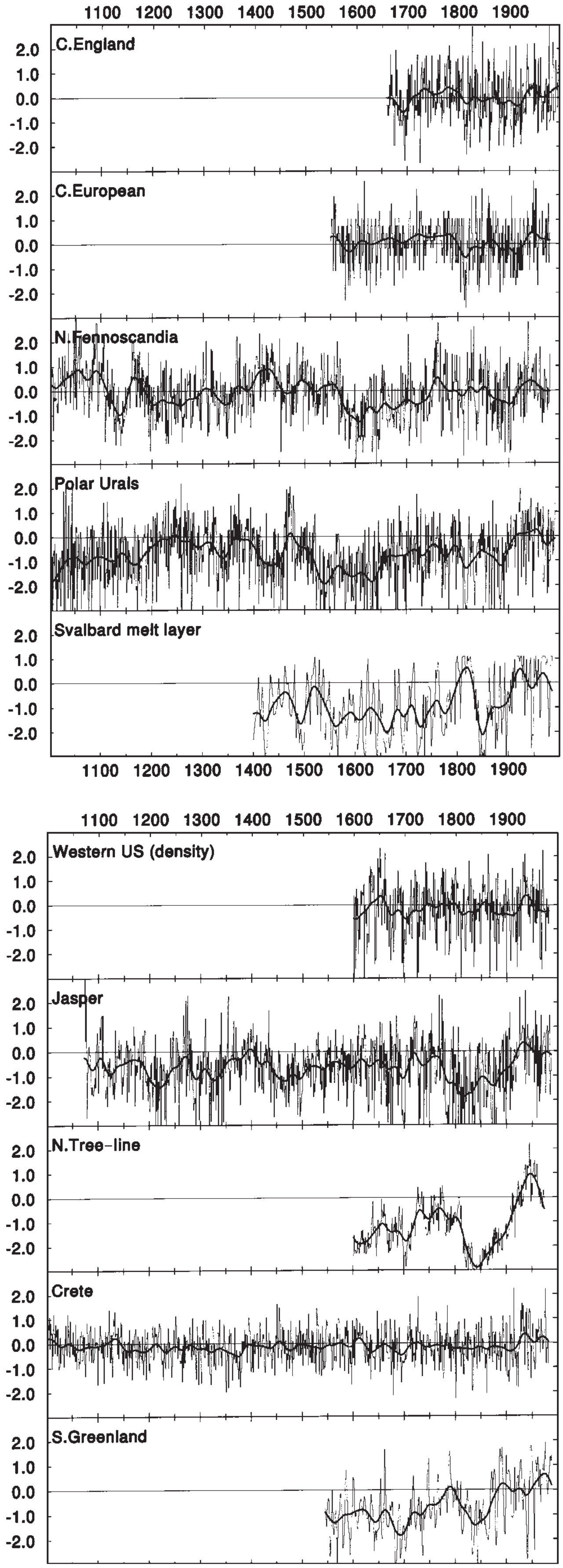

Figure 1 The ten reconstructions in NH north of 40°N, normalized using the 1901–50 period. The thick line is a 50-year Gaussian filter.

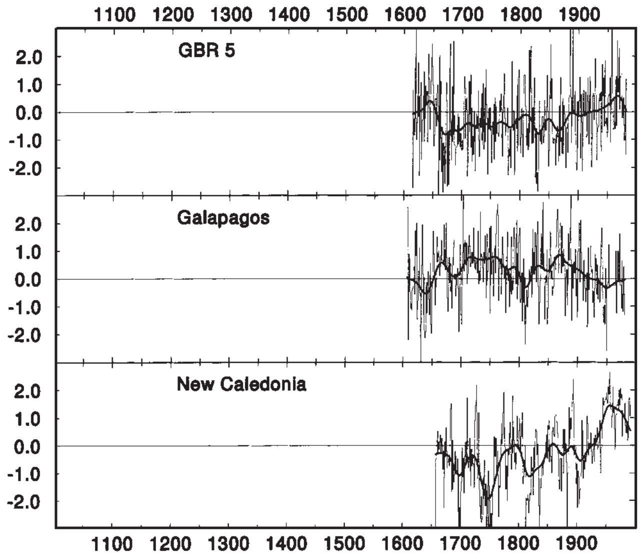

Figure 2 The three tropical reconstructions from coral, normalized over the 1901–50 period. The thick line is a 50-year Gaussian filter.

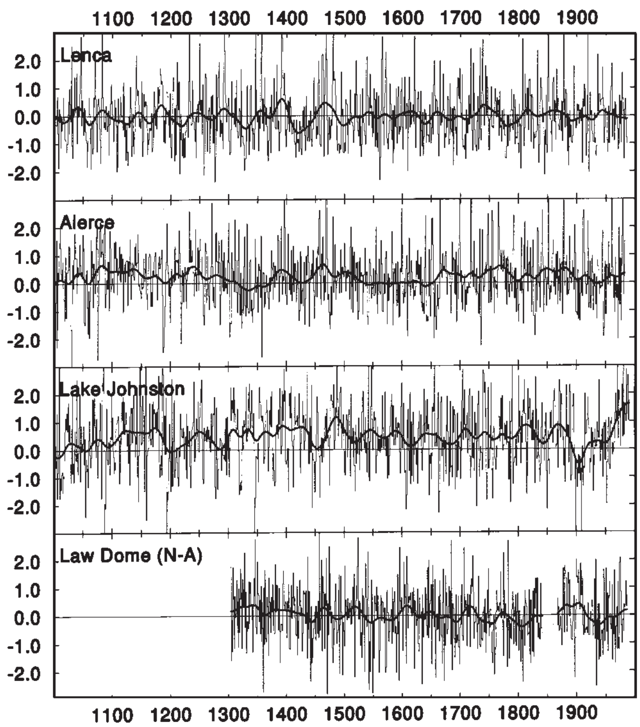

Figure 3 The four reconstructions for the SH south of 40°S, normalized over the 1901–50 period. The thick line is a 50-year Gaussian filter.

peratures only rising above the 1901–50 mean in the eleventh century. In contrast the SH7 (smoothed) is generally above the 1901–50 base for most of the millennium. Although recent (post-1950) values have been high, comparable values have been recorded for a number of periods during the millennium.

The three series (GL17, NH10 and SH7) can be compared with instrumental data in several ways. Comparison can be made with the average temperature over the hemisphere (or globe) using land-only or land + marine data (both summer season), or with the simple averages of the instrumental data where the proxy records are located (i.e. the data used in Table 4 – with different seasons at each site). Figures 5 (NH) and 6 (SH) show comparisons over the period since the 1850s. Correlations between all the series in each domain are given in Table 5. Over the 1901–91 period, NH10 has 36% of the variance in common with NH L + M averages on the interannual timescale and 70% on decadal scales.

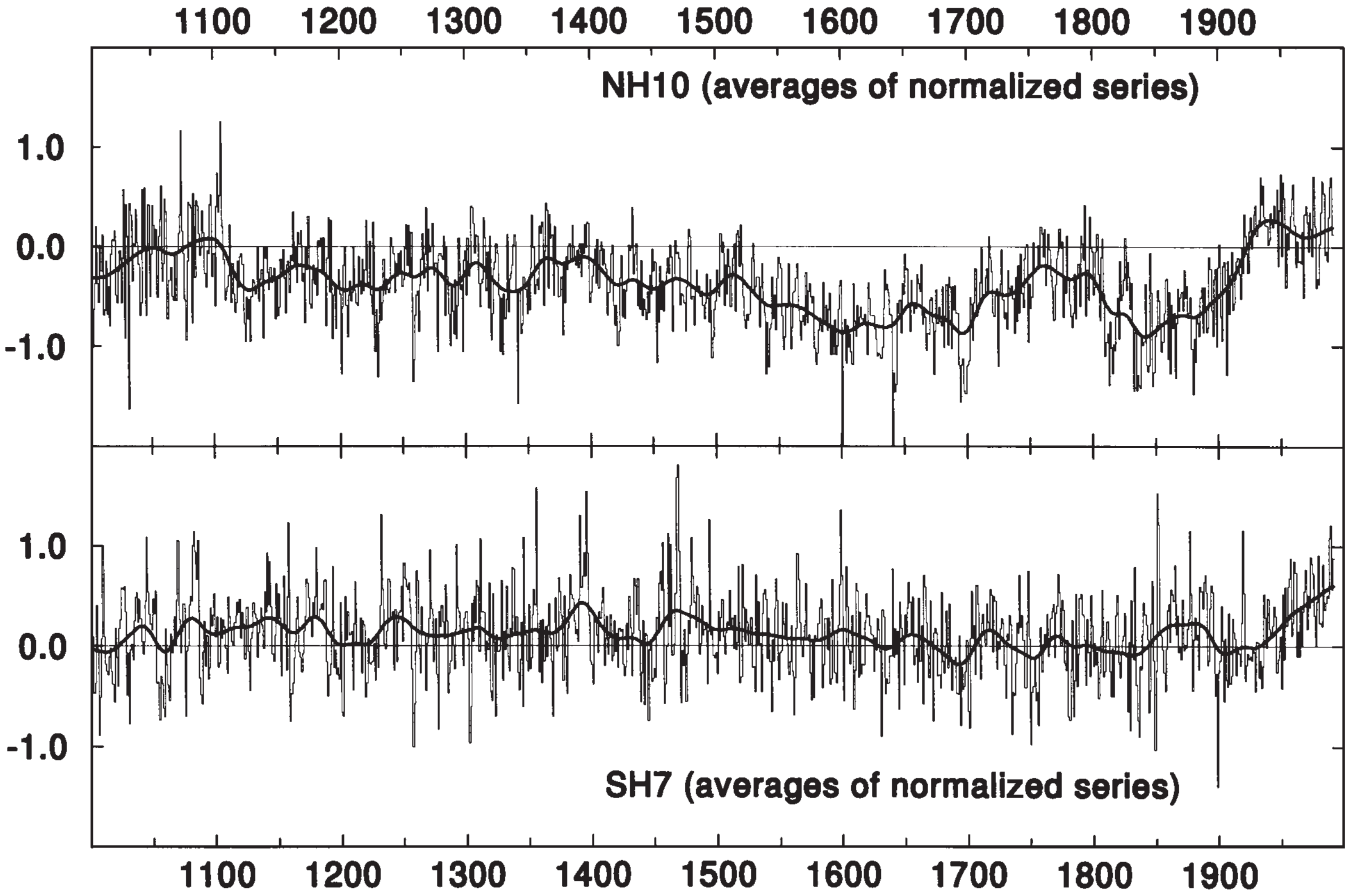

Figure 4 Averages of the reconstructions in the two hemispheres NH10 – the average of all the series in Figure 1 – and SH7 – the average of all the series in Figures 2 and 3. The thick line is a 50-year Gaussian filter.

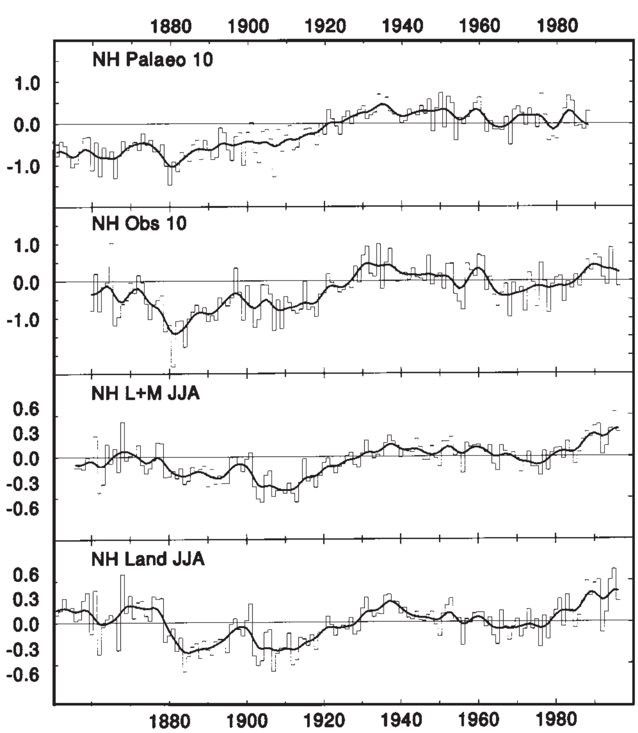

Figure 5 Comparisons of the average reconstructions (NH10) and observed data (for the same ten locations, for the whole NH for JJA – land + marine and land only). The thick line is a 10-year Gaussian filter.

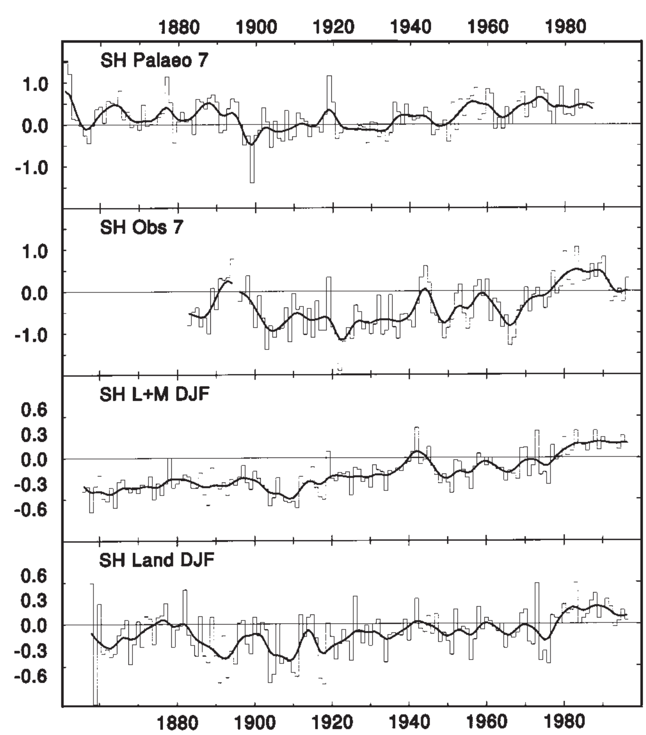

Figure 6 Comparisons of the average reconstructions (SH7) and observed data (for the same seven locations, for the whole SH for DJF – land + marine and land only). The thick line is a 10-year Gaussian filter.

Table 5 Correlations between the proxy series average (P) for each domain and observed temperatures (for the same sites (O)[1], land only (L)[2] and land + marine (L + M).[3] Correlations are calculated over the 1901–91 period using raw and ten-year Gaussian filtered data

Global (17)

	Raw				Low pass		
	O	L	L+M		O	L	L+M
P	0.75	0.61	0.69	P	0.88	0.85	0.91
O		0.72	0.77	O		0.92	0.92

NH (10)

	Raw				Low pass		
	O	L	L+M		O	L	L+M
P	0.75	0.62	0.60	P	0.88	0.84	0.86
O		0.71	0.70	O		0.90	0.87

SH (7)

	Raw				Low pass		
	O	L	L+M		O	L	L+M
P	0.44	0.26	0.42	P	0.70	0.50	0.70
O		0.53	0.66	O		0.78	0.83

[1]Observed data are the averages of the 'seasonal' values for the proxy sites. The seasons differ from site to site.
[2]For land only (Jones, 1994), data for JJA are used in the NH and DJF in the SH. For the globe the average is of the two series NH (JJA) and SH (DJF).
[3]For land + marine, data for NH is JJA and for SH is DJF. The global series is a similar average as for land only.

SH7 has 16% variance in common with SH L + M and 49% on the decadal timescale, markedly poorer than for NH10. We include the SH7 series here, but caution placing too much emphasis on its reliability.

Figure 7 shows the NH10 and SH7 series again, this time rescaled to the level and variance (based on the 1961–90 period) of the summer hemispheric (Land + Marine) temperature series. Expressed this way, the values for individual years and decades can be compared directly with temperatures experienced over the instrumental period from the 1850s. The coldest and warmest years/decades/centuries of the millennium are listed by hemisphere in Table 6. Emphasis in the discussion of Table 6 has been placed on years after 1500 when more of the proxies are available. It must again be stated that the NH average is based on only four series before this time. The coldest year (globally) of the millennium was 1601 with the coldest decade 1691–1700 and coldest century the seventeenth. The warmest global years were 1468, 1469 and 1898 (tied), with the warmest decade and century being 1981–1990 and the twentieth, respectively. The values in Table 6 only give a guide to the probable range of average summer temperatures over the last millennium. The specific values are subject to considerable uncertainty, the error of which is difficult to quantify because of the reasons discussed earlier. The sampling error estimation technique of Jones *et al.* (1997) could be used if it could be assumed each proxy was absolutely dated and of equal quality (and lower than the instrumental data) throughout its length of record. Both assumptions are difficult to verify for some proxies.

The results in Figures 5 and 6 and Table 5 indicate that the NH10 series is much more reliable than SH7. We now consider this series in terms of our existing knowledge of the last millennium, remembering we are considering a series that is principally indicative of summer temperatures. The course of change since 1600 with the cold seventeenth, less cold eighteenth and cold nineteenth centuries has been noted by a number of earlier stidies (e.g. Williams and Wigley, 1983; Grove, 1988; Mann *et al.*, 1998) and is indicative of two phases of the 'Little Ice Age', with the seventeenth being more severe over Eurasia and the nineteenth more severe over North America (Luckman *et al.*, 1997). Prior to this the temperature level between 1000 and 1400 was relatively stable with some variability on the 50-year timescale. After 1400 summer temperatures slowly cooled for 200 years reaching their lowest values in the early seventeenth century. While the 'Little Ice Age' cooling is clearly evident in Figure 7 we can only concur with Hughes and Diaz (1994) that there is little evidence for the 'Medieval Warm Period', although it is variably quoted as occurring between 900 and 1200 (e.g. Lamb, 1977).

With our proxies being principally summer responding we might expect the 'Little Ice Age' cooling to be greater on an annual basis if the winters were more anomalously cold. For example, there is evidence for longer and more frequent freezing of major European rivers and lakes and of the Baltic Sea from documentary sources (Lamb, 1977). Such a differing seasonal response argument cannot be made for the 'Medieval Warm Period', although the fact that we have only four series before 1400 and the timescale limitations described earlier caution against dismissing the feature. Evidence from borehole temperatures from ice cores in Greenland clearly indicate a period warmer than today between 900 and 1200 (Dahl-Jensen *et al.*, 1997).

Comparisons of the space-time variability of the palaeoclimatic data and co-located estimates from two CGCMs

Principal Components Analysis

Using a correlation matrix calculated over the period 1660–1970 (assuming average values for the few missing values for Law Dome in the 1840s/1850s), a Principal Components Analysis (PCA) was performed. Seven PCs have eigenvalues greater than unity and their eigenvalues and cumulative explained variances are listed in Table 7. An eigenvalue greater than unity is a somewhat arbitrary choice, which we make for comparison with the CGCMs later. Apart from PC1 there is little separation between the remaining six PCs. The PC weights indicate that PC1 groups together, with the same sign, all the NH >40°N reconstructions, with four of the seven SH reconstructions having the opposite sign. The PC1 weights are plotted in Figure 8a. The amplitude series of PC1 is highly correlated with NH10, (r = 0.92 over 311 years) despite explaining only 13.6% of the total variance of the dataset. SH7 is uncorrelated with PC1, but significantly correlated with PC3 (0.38) and PC5 (0.48).

The PC analyses have been repeated using results from two of the millennial-long unforced CGCM control runs available (from the Hadley Centre, HADCM2; Johns *et al.*, 1997; Tett *et al.*, 1997; and the Geophysical Fluid Dynamics Laboratory, GFDL; Stouffer *et al.*, 1994). In comparisons with the CGCM results it must be remembered that the model data are from unforced control integrations while the proxy series should contain all possible (natural and anthropogenic) forcings. For both models, series for the same 17 locations were extracted and averaged for the seasons corresponding to the palaeoclimatic series. For GFDL, all 1000 years available were used, while for HADCM2 we took the last 1000 years of the 1438 completed at this time. The standard deviations of each of the series for the two models are compared with those

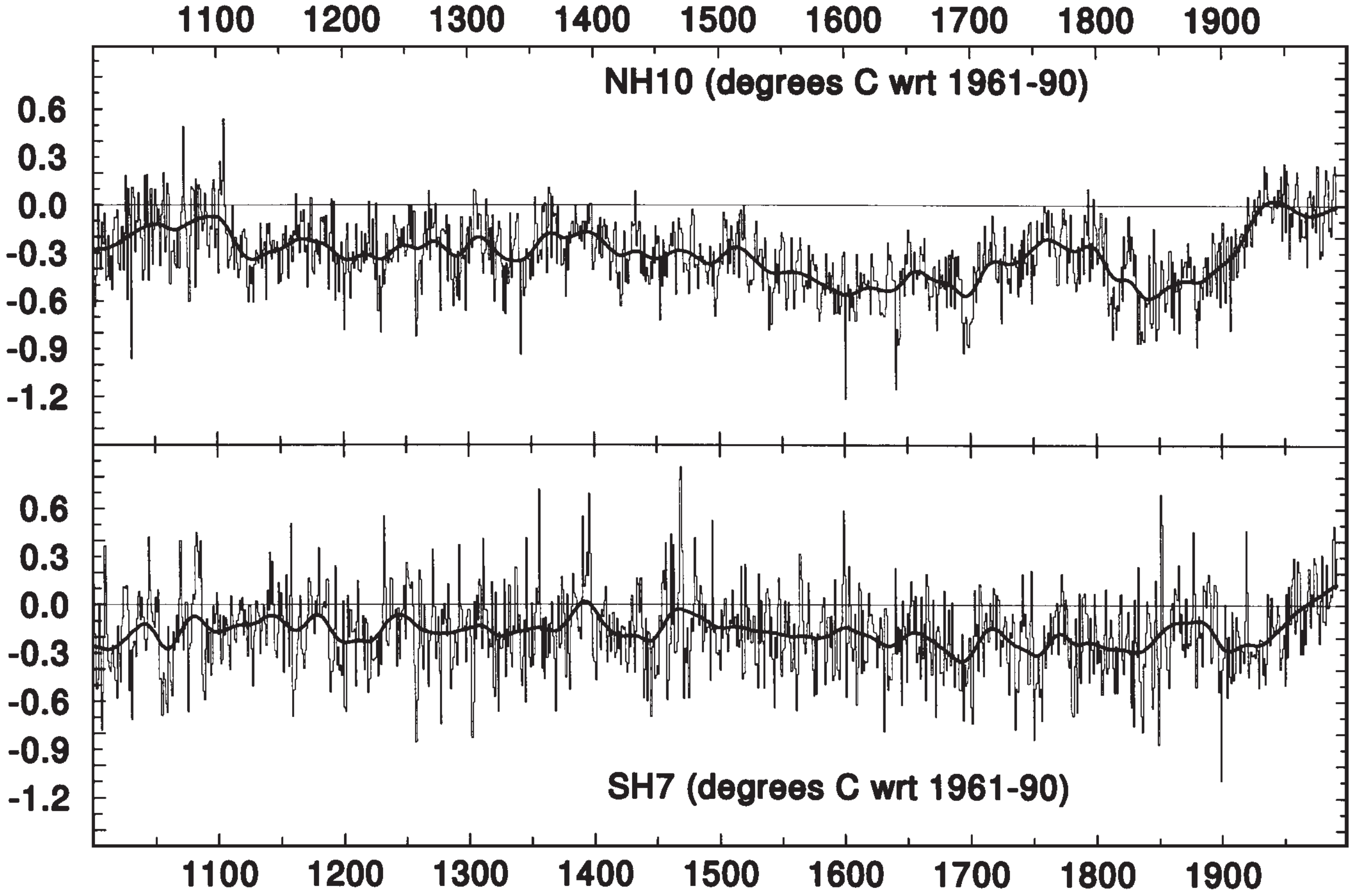

Figure 7 NH10 and SH7 rescaled so the annual values can be compared with summer (JJA or DJF) temperatures for the whole hemisphere based on land + marine data for 1961–90. The thick line is a 50-year Gaussian filter.

of the palaeoclimatic series and observations in Table 8. The model estimates compare favourably with both the proxy and observed temperature data.

For both HADCM2 and GFDL, seven PCs also had eigenvalues greater than unity. The eigenvalues and their cumulative explained variances are listed, together with the palaeoclimatic series in Table 7. PC1 for both models is not as distinct as for the proxy data, there being relatively little separation amongst the top seven eigenvalues. As another simple comparison we show the map of the weights of PC1 for GFDL (Figure 8b) and HADCM2 (Figure 8c). The GFDL map compares more favourably with the proxy pattern in Figure 8a than does the HADCM2 pattern. Time series of the seven PCs from both CGCMs (not shown) exhibit markedly lower levels of century timescale variability than for the palaeoclimatic series. Some low levels of century timescale variability are apparent in the first two PCs for GFDL, but the main features of HADCM2 are on the decadal scale. If the CGCM data are normalized and averaged together to produce equivalent values of NH10 and SH6 (see footnote 1 to Table 7), the GFDL model is similar to the palaeoclimatic data, with a correlation between NH10 and PC1 of 0.89. The HADCM2 PCA apportions the variance markedly differently with NH10 correlated with PC1 (0.45), PC2 (0.64), PC3 (0.33) and PC5 (0.42).

Therefore, in two regards, the NH10/PC1 correlation and the century timescale variability in PC1 and PC2, the GFDL model shows some features of the palaeoclimatic data, while HADCM2, with its dominant tropic-wide variation (see Tett *et al.*, 1997) is quite different. The quality of the coral series do not really allow an adequate assessment to be made of the behaviour of the El Niño/Southern Oscillation (ENSO) phenomenon in the two models. It is generally believed that ENSOs simulated by HADCM2 impact too large an area of the tropics (Tett *et al.*, 1997), but that ENSOs simulated by GFDL are probably too weak (Knutson *et al.*, 1997).

Cross-spectral analyses

In this section we examine, in frequency space, the relationships between the various reconstructions by estimating the cross spectrum, coherency squared and phase within groups of proxies. The groups were defined based on the length of common record. An identical analysis was performed on the equivalent series extracted from the 1000-year control run of the GFDL GCM (Stouffer *et al.*, 1994). Our basic goal was another method of comparison of the space-time structure of the near-surface temperature fields of the proxy and model data to see if the model's natural variability resembled that estimated from the proxy data. Since results obtained here are highly relevant to the anthropogenic signal detection problem, we will concentrate on time scales of order 30–50 years, i.e. time averages required for substantial anthropogenic signal/noise ratios (Barnett *et al.*, 1998). The marked dissimilarity between the PCs of HADCM2 and the palaeoclimatic data indicate using HADCM2 in the cross-spectral analyses would not, at this time, be useful.

Methods

Group 1 (see Table 3 for the makeup of the groups) included all of the proxies, so each member had a time series of length 236 years (1750–1985). Groups 2 and 3 had lengths of 386 (1600–1985) and 1086 (900–1985) years respectively. Groups 1, 2 and 3 had 17, 16 and 6 members respectively. Occasionally a series was 'padded' with a few zeros at the beginning and/or end to ensure each group member was of the same length.

Each series was forced to have zero mean over the analysis period and then submitted to cross-spectral analysis. This was

Table 6 Coldest and warmest summers, decades and centuries (1000–1991). All values are with respect to the 1961–90 period

Individual Summers (NH)			
Warmest		Coldest	
Year	Value	Year	Value
1106	0.54	1601	–1.21
1074	0.49	1641	–1.16
1103	0.27	1032	–0.96
1950	0.27	1342	–0.94
1976	0.26	1695	–0.93
Decades (NH)			
Warmest		Coldest	
Decade	Value	Decade	Value
1931–40	0.05	1691–1700	–0.68
1941–50	0.04	1831–40	–0.68
1981–90	0.04	1601–10	–0.64
1951–60	0.01	1641–50	–0.61
1101–10	–0.02	1811–20	–0.61
Centuries (NH)			
Warmest		Coldest	
Century	Value	Century	Value
1901–1991	–0.08	1601–1700	–0.52
1001–1100	–0.16	1801–1900	–0.48
1301–1400	–0.23	1501–1600	–0.40
1101–1200	–0.25	1401–1500	–0.31
1201–1300	–0.30	1701–1800	–0.30
Individual Summers (SH)			
Warmest		Coldest	
Year	Value	Year	Value
1469	0.86	1899	–1.09
1468	0.78	1849	–0.87
1356	0.72	1257	–0.85
1396	0.70	1750	–0.84
1851	0.69	1302	–0.83
Decades (SH)			
Warmest		Coldest	
Decade	Value	Decade	Value
1461–1470	0.19	1691–1700	–0.45
1391–1400	0.16	1751–1760	–0.41
1981–1990	0.12	1051–1060	–0.39
1971–1980	0.08	1011–1020	–0.35
1171–1180	0.04	1441–1450	–0.35
Centuries (SH)			
Warmest		Coldest	
Century	Value	Century	Value
1901–1991	–0.11	1601–1700	–0.25
1301–1400	–0.12	1701–1800	–0.22
1101–1200	–0.13	1801–1900	–0.20
1401–1500	–0.14	1001–1100	–0.19
1201–1300	–0.16	1501–1600	–0.16

accomplished by first computing the lagged cross covariance between all possible members of a specific group for lags up to 50, 50 and 150 years for groups 1–3 respectively. This combination of record length and spectral bandwidth gave cross-spectral estimates with 10–15 degrees of freedom. The coherency squared and phase were estimated from the complex Fourier transform of the cross covariance functions in the standard manner (cf. Jenkins and Watts, 1968). The raw cross spectra were smoothed with a

Table 7 Eigenvalues and cumulative percentage variances explained in a PCA of the 17 palaeoclimatic time series over 1660–1970 and equivalent 16[1] series from the 1000-year integrations of the GFDL and HC unforced control runs

	Palaeoclimate		GFDL		HC	
	λ[2]	Σ%[3]	λ	Σ%	λ	Σ%
1	2.31	13.60	1.71	10.68	1.99	12.41
2	1.59	22.93	1.42	19.56	1.90	24.30
3	1.49	31.68	1.37	28.13	1.74	35.15
4	1.37	39.73	1.22	35.77	1.57	44.99
5	1.19	46.73	1.20	43.27	1.34	53.34
6	1.14	53.40	1.08	50.01	1.16	60.59
7	1.03	59.47	1.05	56.56	1.03	67.05

[1]For the PCA with the GCM data the grid box for palaeosites LEN and ALE was the same, so the site was included once only. The resulting SH average for the GCMs is based on six rather than seven locations.
[2]Eigenvalues.
[3]Cumulative explained variance.

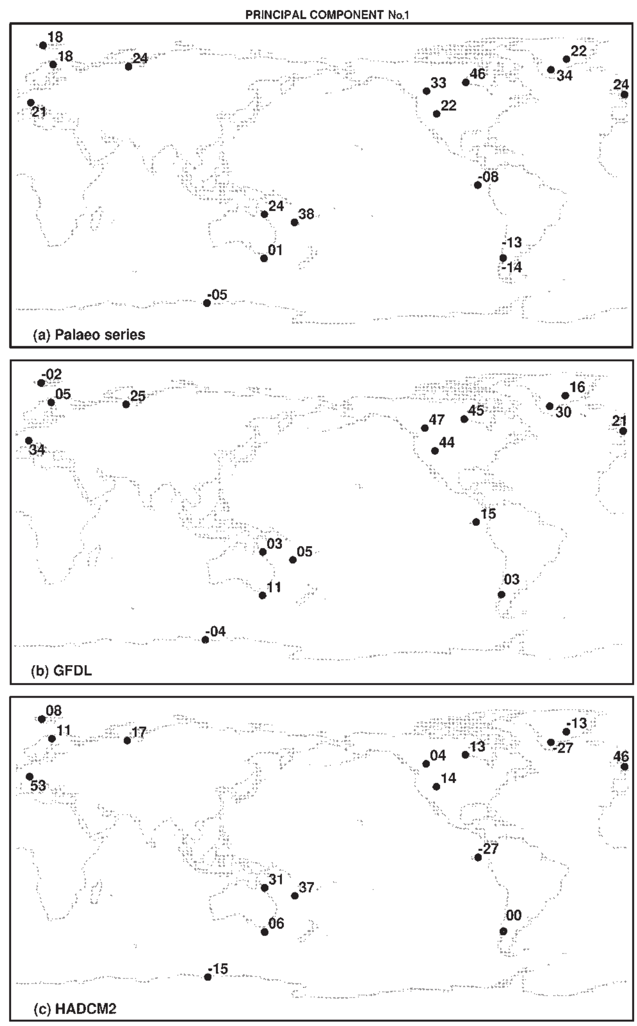

Figure 8 Maps of PC1 loadings of the 17 palaeoclimatic data sites (a), GFDL (b), and HADCM2 (c) models sampled in the same way as the palaeoclimatic data. Loadings are plotted multiplied by 100.

Hanning filter prior to coherency estimation. Given the record length and spectral bandwidth, the 95% confidence limits on the coherency squared for groups 1–3 were 0.42, 0.26 and 0.28 respectively (cf. Munk *et al.*, 1959).

An identical analysis was performed on the near-surface

Table 8 Standard deviations of the palaeoclimatic (calculated over the same years as in columns 3 and 6 of Table 4) and two GCM data series. The number of cases for the observations and palaeoclimatic series is given in Table 4. For the model the numbers are 1000 (interannual) and 100 (decaal)

		SD_{Obs}	SD_{Pal}	SD_{HC}	SD_{GFDL}
a) Interannual					
1	NFS	0.97	0.63	0.97	0.91
2	NUR	1.18	1.10	0.90	1.02
3	JAS	0.76	0.50	0.79	1.02
4	SVA	0.78	–	0.43	0.62
5	ENG	0.55	0.82	0.91	0.76
6	EUR	0.76	1.06	1.19	1.33
7	SGR	0.70	0.48	0.77	0.61
8	NTR	0.83	0.54	0.81	1.14
9	WUS	0.68	0.39	1.06	1.06
10	CRT	0.80	0.87	0.87	0.89
11	TAS	0.56	0.39	0.43	0.55
12	LEN	0.62	0.53	0.75	0.64
13	ALE	0.65	0.59	0.75	0.64
14	LAW	0.41	–	1.19	1.08
15	GBR	0.30	0.41	0.29	0.43
16	GAL	0.47	0.34	0.61	0.26
17	NCL	0.39	0.18	0.24	0.35
b) Decadal					
1	NFS	0.39	0.28	0.36	0.32
2	NUR	0.57	0.60	0.28	0.30
3	JAS	0.27	0.25	0.27	0.31
4	SVA	0.47	–	0.20	0.23
5	ENG	0.20	0.34	0.31	0.25
6	EUR	0.39	0.49	0.40	0.45
7	SGR	0.31	0.20	0.24	0.23
8	NTR	0.34	0.51	0.29	0.40
9	WUS	0.43	0.18	0.36	0.34
10	CRT	0.65	0.40	0.33	0.47
11	TAS	0.33	0.23	0.20	0.17
12	LEN	0.19	0.18	0.29	0.20
13	ALE	0.21	0.23	0.29	0.20
14	LAW	0.26	–	0.41	0.59
15	GBR	0.16	0.14	0.09	0.14
16	GAL	0.13	0.14	0.19	0.10
17	NCL	0.27	0.13	0.10	0.12

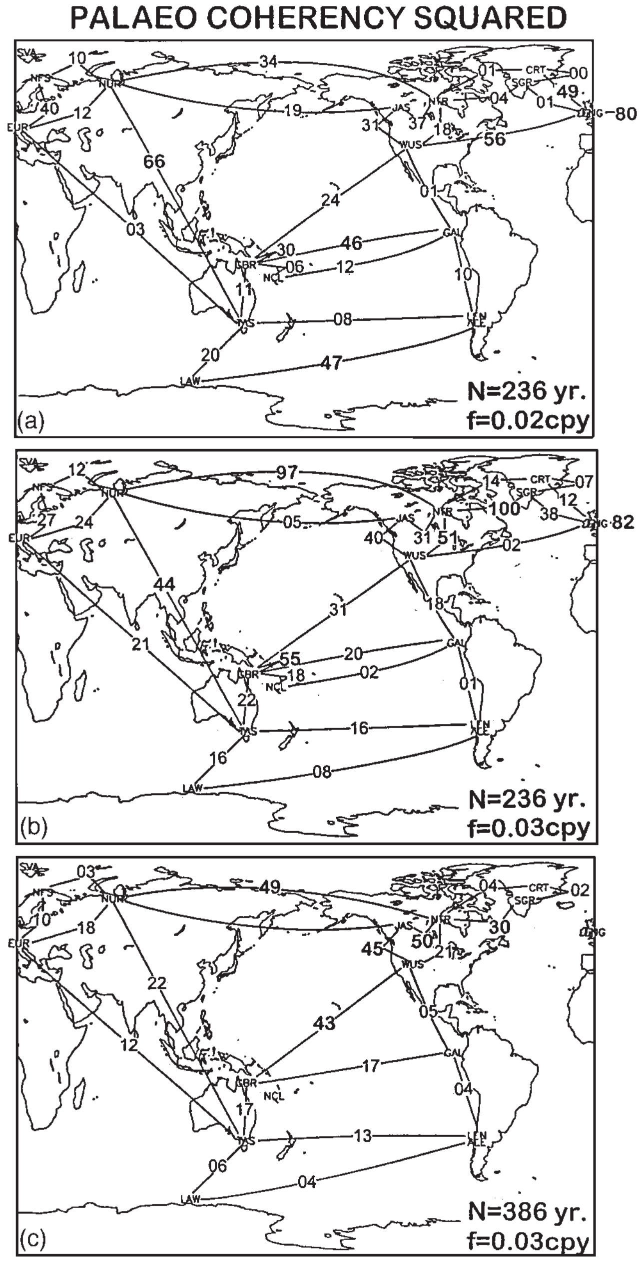

Figure 9 Coherency (squared) for the proxy data, f=0.02 cpy (top), f=0.03 cpy (middle), both using group 1 over the 236 years 1750–1985. The lower panel is for f=0.03 cpy for the 386 years 1600–1985.

temperature field from the 1000-year GFDL model control run. The monthly model data were sampled at the same locations that the proxy records are supposed to represent. Temporal consistency was ensured by using model data also at the same times of the annual cycle represented by the proxies (cf. Table 3). Finally, model groups 1–2 were formed by using the first 236 or 386 years of the model data to give equivalent lengths to the proxy series. Group 3 was taken to be 1000 years long, since the control run was not long enough to form a 1086-year series comparable to palaeo group 3.

Proxy cross spectra

The following discussion will focus on the covariation between different proxies on the 30–50-year timescale, since this is of most importance in anthropogenic signal detection. The coherency between proxy data that seem particularly good proxies for temperature (cf. Table 4) is shown in Figure 9 for group 1 for frequencies 0.02 (a) and 0.03 (b) cpy (cycles per year) and (c) group 2 for 0.03 cpy (periods of 50, 33 and 33 years respectively). Other details of the cross spectra have been discussed elsewhere (e.g. Table 4; Barnett *et al.*, 1996; 1998). The main information in this illustration is summarized as follows.

(1) Surprisingly, there is little coherency between proxies, even those that are geographically quite close. This is particularly clear for the Greenland ice-core series (SGR and CRT) and the GBR and NCL corals at f = 0.02 cpy (Figure 9a), but this is also the case among some of the proxies for northern Europe and North America. The highest coherency is found between the proxies for EUR and ENG.

(2) A few remarkably strong teleconnections are implied in the frequency band of 0.02 cpy. This is evidenced by the value of 0.66 between TAS and NUR and between longitudes of Australia and South America. In general, however, there does not appear to be anything approaching a pattern of coherent, near-global temperature change implied by the results.

(3) The sensitivity of the above two results to frequency is illustrated in Figure 9b. Nearly co-located stations still appear to lack coherency. The largest teleconnections generally suffer serious degradation with the small change in frequency. Other teleconnections appear where little or no coherency values

had been before, e.g. NUR now has a coherency squared with NTR of 0.97, compared to the 0.34 for frequency 0.02 cpy.

(4) The sensitivity of the above results to record length was tested by comparing them with those obtained from the analysis of group 2 (386 years long). This is shown in Figure 9c. Inspection of this illustration shows that increasing record length generally reduces coherency squared nearly everywhere. Hence, the results shown in the upper two panels of Figure 9 are not stable when the record length is increased by 150 years.

(5) Analysis of group 3 (not shown) reproduced the same result found in (4) above. In fact, only the two South American tree ring series shows marginal coherency in group 3.

In summary, the cross-spectral analysis of the proxy data does not suggest the presence of any large-scale pattern of near-surface temperature change on timescales of 30–50 years. Even data from nearly co-located positions generally have little relation with one another on these timescales. This result, in turn, suggests the proxy data contain a large amount of uncorrelated noise and/or that they represent considerably more than simply air temperature. If the proxy data are meaningful, then the spectrum of natural variability they imply suggests detection of a near-global anthropogenic signal in the near surface temperature field, if one exists, should be relatively straightforward. This is so because, if one believes the proxy data, no such coherent pattern of variability has existed in nature over the last 300–400 years.

Model cross spectra

Similar analyses were performed on the GFDL control integration. Again, we concentrate on the spatial covariability since the levels of variance in the model for different timescales have been described earlier (Jones *et al.*, 1997). The results for the CGCM are summarized in Figure 10 which is directly comparable with Figure 9. Careful study of this illustration and the coherency squared values for different record lengths, frequencies etc. lead to the following conclusions.

(1) The model shows similar regional decorrelation as seen in the proxy data. Of course, proxy locations that fall within the same GCM grid cell have coherency of 1 as expected. However, once separations of even a few grid cells are involved, the coherency falls dramatically.

(2) CGCM shows some teleconnections. For example the connection between the western Pacific and the Galapagos seems reliable on physical grounds due to the El Niño/Southern Oscillation (ENSO) phenonemon. Other apparently strong teleconnections seen in the proxy data analysis are not present in the model data (cf. the EUR/ENG coherency).

(3) The sensitivity of the results to frequency (Figure 9b) shows that, while the relationships in the North Atlantic remain, most of the other estimates of coherency change substantially. The North Atlantic relationships are stronger in the model than in the proxy data.

(4) Increasing the record length by 150 years substantially reduces all the levels of covariability. This is essentially the same result found for the proxy data.

(5) Analysis of even longer records, e.g. group 3, lowered all coherency squared values to insignificant levels. Apparently, both the model and the proxy data are characterized by highly non-stationary behaviour. Alternatively, the apparently significant values of coherency discussed above are merely spurious. However, given the number of statistically significant values compared to the total number computed, this scenario is unlikely.

In summary, the cross-spectral analysis of the GFDL CGCM control integration bears a remarkable similarity in its statistical

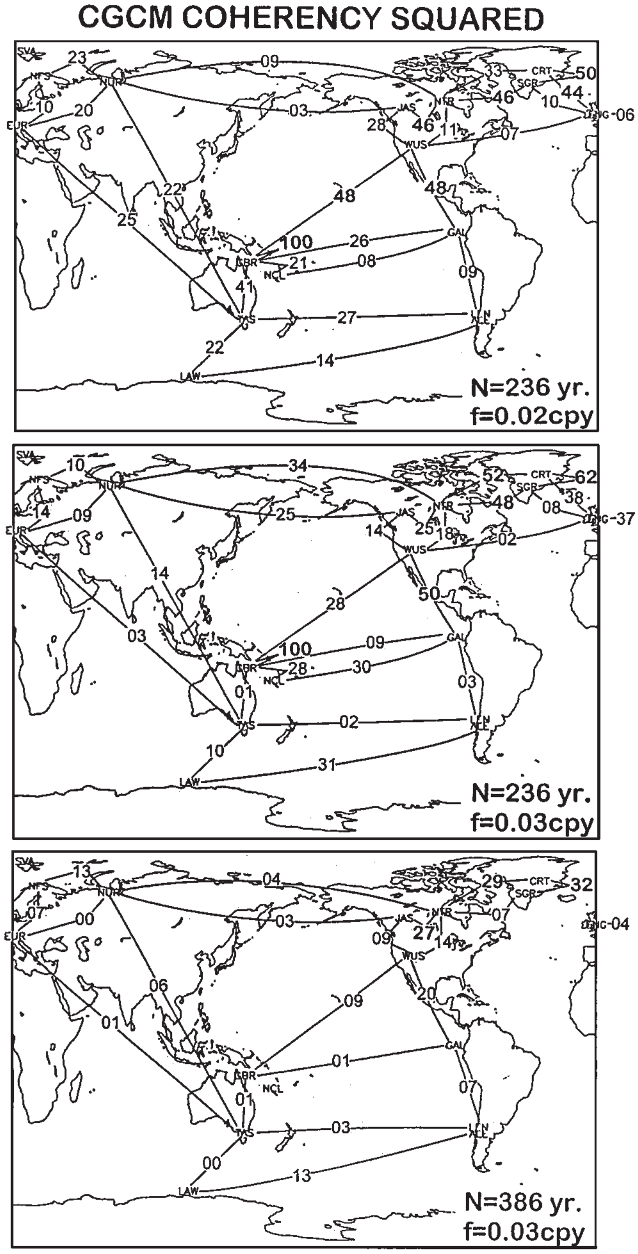

Figure 10 As Figure 9 but for 236 (model years 765–1000) and 386 (model years 615–1000) years of the GFDL 1000-year control-run integration.

properties to that obtained from the proxy data. In view of this similarity it appears the spatial structures from the control integration can be used to represent the spatial structures of naturally occurring variations in near-surface air temperature. This conclusion depends strongly on the meaningfulness of the proxy data. The results of Table 4 suggest that most bear a reasonable similarity to the instrumental data, at least over the last 100 years, so the conclusion has support.

Summary

Palaeoclimatic reconstructions representing past temperatures with annual resolution are not very plentiful. In this analysis 17 series have been brought together, but, even if large spheres of influence are assumed for each proxy, large areas, particularly in 'tropical' regions (between ±40° of latitude), are not represented. Many of the reconstructions are also limited seasonally in that they are most representative of summer or growing season conditions. Of greater importance to the use of an array of palaeoclimatic reconstructions is that each is probably limited in its ability to reproduce past temperature variations faithfully on the longest of timescales. This limitation varies from proxy to proxy and it is virtually impossible to quantify the degree to which this has occurred because instrumental series are not long enough. Two possible methods are suggested for assessing the importance of this issue: (i) intercomparison of adjacently located proxies; (ii) the use of more slowly responding proxies.

Each of the 17 'reconstructions' has been compared with observed temperatures from the $5° \times 5°$ grid-box data set for the appropriate season, using either a single or multiple grid box(es), making assumptions as to the best choice if none were given in the original source. In the cases where the original source includes a correlation coefficient with instrumental data, the values calculated here are all lower. The differences relate to the use of the grid-box data, which are obviously not site-specific. The use of grid-box data in this assessment is justified by our interest in variability on the large scale, comparable with that provided by CGCMs.

The correlations, calculated here on interannual and decadal timescales over 1881–1980, indicate that the highest similarity with grid-box temperatures is apparent in the instrumental/historical series and all the NH tree-ring-based reconstructions (all with r values >0.7). Correlations are markedly lower than this for SH tree-rings, coral and ice-core parameters. Although this analysis may provide an objective means of assessing quality and hence rejecting some reconstructions as being of little use, there are several reasons why some of the reconstructions may appear poor partly because of instrumental data quality in some regions. Also, 17 is too few reconstructions from which to allow widespread rejection.

All 17 reconstructions were normalized over a common interval and averages for the globe (GL17), NH (NH10) and SH (SH7) were produced. Compared with instrumental temperatures, NH10 has nearly 40% variance in common on the interannual timescale and 70% on the decadal. The variance in common for SH7 is markedly lower. PCA of the 17 reconstructions over the 1660–1970 period showed seven PCs with eigenvalues greater than one, with PC1 correlated with NH10 to 0.92 over the full 311-year period. For both series the warmest century is the twentieth. The more reliable NH10 series shows clear evidence of cooler centuries between 1500 and 1900, particularly the seventeenth and nineteenth centuries which are indicative of a two-phase 'Little Ice Age'. From the few reconstructions used prior to 1500 there is little evidence for the 'Medieval Warm Period'.

Similar analyses to this were undertaken using 1000 years of data extracted, from the unforced control integrations of HADCM2 and GFDL, for the same seasons. The first two GFDL PCs showed some century timescale variability and PC1 correlated with the equivalent NH10 for GFDL at 0.89. The HADCM2 PCs showed only decadal scale variability with equivalent NH10 correlating weakly with four of the first seven PCs. Cross-spectral analysis reveals that the variability in the GFDL control integration bears some similarity to that evident in the proxy data, implying that this model can be used to represent the spatial structure of naturally occurring variability in surface air temperatures.

Conclusions from this study for palaeoclimatologists and users of their data

Below we summarize a number of major implications, from this study, for the use of palaeoclimatic data in the climate change detection context:

- the twentieth century is the warmest of the millennium;
- more palaeoclimatic reconstructions are required;
- all proxy series have potential timescale limitations that are not generally appreciated outside of each palaeoclimatic field;
- all proxies have an optimal seasonal response and the period over which an assessment has been made should be clearly stated in the original publication;
- the quality of each proxy series as an indicator of past temperature needs to be explicitly quantified, through a calibration/verification exercise with instrumental data;
- possible additional limitations in the quality of proxy series in the earliest years need to be carefully assessed;
- compared to corals and ice cores, tree-ring and historical reconstructions show greater agreement with instrumental data over the last 100 years, both in terms of simple correlations and coherency on different timescales;
- in using proxy data to assess models in the detection context, larger-scale variability on the 30–50-year timescale is most important;
- the marked variability of coherency with timescale for all the palaeoclimatic series needs to be studied further; where possible greater emphasis needs to be placed on this aspect in future studies;
- more millennial-long CGCM integrations need to be studied and compared with proxy climatic data;
- we strongly endorse the importance of current initiatives to develop data banks of reconstructions suitable for studies such as that described here (http://www.ngdc.noaa.gov/paleo/paleo.htmp).

Acknowledgements

This work was supported by the US Department of Energy, Atmospheric and Climate Research Division (DE-FG02-86ER60397), the UK Department of the Environment (EPG 1/1/14) and by NOAA's Office of Global Programs' Climate Change Data and Detection Program Element, and the US DoE's Office of Health and Environmental Research. K.R. Briffa is currently funded by the EC Environment and Climate programme (ENV4-CT95-0127). S.F.B. Tett is supported by the UK Department of the Environment, Transport and the Regions (PECD 7/12/37). The HC model data were supplied through the Climate Impacts LINK project, UK Department of the Environment, Transport and the Regions (EPG 1/1/16). The authors thank all their palaeoclimatic colleagues for making their reconstructions available and Ray Bradley, Ed Cook, Tom Crowley and Janice Lough for insightful comments.

References

Alexandersson, H. 1986: A homogeneity test applied to precipitation data. *J. Climatology* 6, 661–75.

Alexandersson, H. and **Moberg, A.** 1997: Homogenization of Swedish temperature data Part 1: Homogeneity test for linear trends. *Int. J. Climatol.* 17, 25–34.

Allan, R.J., Lindesay, J.A. and **Parker, D.E.** 1996: *El Niño Southern Oscillation and climatic variability*. Melbourne: CSIRO Publishing, 405 pp.

Atkinson, T.C., Briffa, K.R. and **Coope, G.R.** 1987: Seasonal temperatures in Britain during the past 22 000 years, reconstructed using beetle remains. *Nature* 325, 587–92.

Barnes, D.J. and **Lough, J.M.** 1996: Coral skeletons: storage and recovery of environmental information. *Global Change Biology* 2, 569–82.

Barnett, T.P., Hegerl, G.C., Santer, B.D. and **Taylor, K.** 1998: The potential effect of GCM uncertainties and internal atmospheric variability on anthropogenic signal detection. *J. Climate* 11, 659–75.

Barnett, T.P., Santer, B.D., Jones, P.D., Bradley, R.S. and **Briffa, K.R.** 1996: Estimates of low frequency natural variability in near-surface air temperature. *The Holocene* 6, 255–63.
Beltrami, H., Chapman, D.S., Acchambault, S. and **Bergeron, Y.** 1995: Reconstruction of high-resolution ground temperature histories combining dendrochronological and geothermal data. *Earth Planet Sci. Lett.* 136, 437–45.
Berglund, B.E., Birks, H.J.B., Ralska-Jansiewiczowa, M. and **Wright, H.E.**, editors 1996: *Paleoecological events during the last 15000 years: regional synthesis of paleoecological studies of lakes and mires in Europe*, Chichester: Wiley, 764 pp.
Bradley, R.S. 1985: *Quaternary palaeoclimatology*, London: Allen and Unwin, 472 pp.
— 1996: Are there optimum sites for global paleotemperature reconstruction? In Jones, P.D., Bradley, R.S. and Jouzel, J., editors, *Climatic variations and forcing mechanisms of the last 2000 years*, Berlin: Springer, 603–24.
Bradley, R.S and **Jones, P.D.** 1985: Data bases for isolating the effects of increasing carbon dioxide concentration. In MacCracken, M.C. and Luther, F.M., editors, *Detecting the climatic effects of increasing carbon dioxide*, US Department of Energy, Carbon Dioxide Research Division, 29–53.
— 1993: 'Little Ice Age' summer temperature variations: their nature and relevance to recent global warming trends. *The Holocene* 3, 367–76.
Briffa, K.R. 1995: Interpreting high-resolution proxy climate data – the example of dendrochronology. In von Storch, H. and Navarra, A., editors, *Analysis of climate variability: applications of statistical techniques*, Berlin: Springer, 77–94.
Briffa, K.R. and **Atkinson, T.C.** 1997: Reconstructing Late-Glacial and Holocene Climates. In Hulme, M. and Barrow, E., editors, *Climates of the British Isles: present, past and future* Routledge, London, 84–111.
Briffa, K.R., Bartholin, T.S., Eckstein, D., Jones, P.D., Karlén, W., Schweingruber, F.H. and **Zetterberg, P.** 1990: A 1,400-year tree-ring record of summer temperatures in Fennoscandia. *Nature* 346, 434–39.
Briffa, K.R., Jones, P.D., Bartholin, T.S., Eckstein, D., Schweingruber, F.H., Karlén, W., Zetterberg, P. and **Eronen, M.** 1992b: Fennoscandian summers from ad 500: temperature changes on short and long timescales. *Climate Dynamics* 7, 111–19.
Briffa, K.R., Jones, P.D. and **Schweingruber, F.H.** 1988: Summer temperature patterns over Europe: a reconstruction from 1750 ad based on maximum latewood density indices of conifers. *Quaternary Research* 30, 36–52.
— 1992a: Tree-ring density reconstructions of summer temperature patterns across western North America since 1600. *J. Climate* 5, 735–54.
Briffa, K.R., Jones, P.D., Schweingruber, F.H., Karlén, W. and **Shiyatov, S.G.** 1996: Tree-ring variables as proxy climatic indicators: problems with low-frequency signals. In Jones, P.D., Bradley, R.S. and Jouzel, J., editors, *Climatic Variations and Forcing Mechanisms of the last 2000 years*, Berlin: Springer, 9–41.
Briffa, K.R., Jones, P.D., Schweingruber, F.H., Shiyatov, S.G. and **Cook, E.R.** 1995: Unusual twentieth-century summer warmth in a 1,000-year temperature record from Siberia. *Nature* 376, 156–59.
Conrad, V. and **Pollak, L.D.** 1962: *Methods in climatology*. Cambridge, Mass.: Harvard University Press, 459 pp.
Cook, E.R. 1995: Temperature histories from tree rings and corals. *Climate Dynamics* 11, 211–22.
Cook, E.R., Bird, T., Peterson, M., Barbetti, M., Buckley, B., D'Arrigo, R. and **Francey, R.** 1992: Climate change over the last millennium in Tasmania reconstructed from tree-rings. *The Holocene* 2, 205–17.
Dahl-Jensen, D., Gundestrup, N.S., Mosegaard, K. and **Clow, G.D.** 1997: Reconstruction of the past climate from GRIP temperature profile by Monte Carlo inversion. Paper presented at the 1997 Fall AGU meeting in San Francisco.
D'Arrigo, R.D. and **Jacoby, G.C. Jr.** 1992: Dendroclimatic evidence from northern North America. In Bradley, R.S. and Jones, P.D., editors, *Climate since* ad *1500*, Routledge, London, 296–311.
Deming, D. 1995: Climatic warming in North America: Analysis of borehole temperatures. *Science* 268, 1576–77.
Dunbar, R.B., Wellington, G.M., Colgan, M. and **Glynn, P.W.** 1994: Eastern Pacific sea surface temperature since 1600 ad: the $\delta^{18}0$ record of climate variability in Galapagos Corals. *Paleoceanography* 9, 291–315.
Easterling, D.R. and **Peterson, T.C.** 1995: A new method for detecting undocumented discontinuities in climatological time series. *Int. J. Climatol.* 15, 369–77.
Fisher, D.A. and **Koerner, R.M.** 1994: Signal and noise in four ice-core records from the Agassiz Ice Cap, Ellesmere Island, Canada: details of the past millennium for stable isotopes, melt and solid conductivity. *The Holocene* 4, 113–20.
Fisher, D.A., Koerner, R.M., Kuivinen, K., Clausen, H.B., Johnsen, S.J., Steffensen, J-P., Gundestrup, N. and **Hammer, C.U.** 1996: Intercomparison of ice core ($\delta^{18}0$) 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. In Jones, P.D., Bradley, R.S. and Jouzel, J., editors, *Climatic Variations and Forcing Mechanisms of the Last 2000 years*, Berlin: Springer, 297–328.
Fritts, H.C. 1976: *Tree rings and climate*. London: Academic Press, 567 pp.
Grove, J.M. 1988: The Little Ice Age. London: Methuen, 498 pp.
Hegerl, G.C., von Storch, H. Hasselmann, K., Santer, B.D., Cubasch, U and **Jones, P.D.** 1996: Detecting greenhouse gas induced climate change with an optimal fingerprint. *J. Climate* 9, 2281–306.
Hughes, M.K. and **Diaz, H.F.** 1994: Was there a 'Medieval Warm Period', and if so, where and when? *Climatic Change* 26, 109–42.
Hurrell, J.W. and **Trenberth, K.E.** 1996: Satellite versus surface estimated air temperature since 1979. *J. Climate* 9, 2222–32.
Jenkins, G.C. and **Watts, D.G.** 1968: *Spectral analysis and its applications*. San Francisco: Holden-Day, 525 pp.
Johns, T.C., Carnell, R.E., Crossley, J.F., Gregory, J.M., Mitchell, J.F.B., Senior, C.A., Tett, S.F.B. and **Wood, R.A.** 1997: The second Hadley Centre coupled ocean-atmosphere GCM: model description, spinup and validation. *Climate Dynamics* 13, 103–34.
Jones, P.D. 1994: Hemispheric surface air temperature variations – a reanalysis and an update to 1993. *J. Climate* 7, 1794–802.
Jones, P.D. and **Bradley, R.A.** 1992: Climatic variations over the last 500 years. In Bradley, R.S. and Jones, P.D., editors, *Climate since* ad *1500*, London: Routledge, 649–65.
Jones, P.D. and **Briffa, K.R.** 1992: Global surface air temperature variations over the twentieth century: Part 1, Spatial temporal and seasonal details. *The Holocene* 2, 165–79.
Jones, P.D., Marsh, R., Wigley, T.M.L. and **Peel, D.A.** 1993: Decadal timescale links between Antarctic Peninsula ice-core oxygen-18, deuterium and temperature. *The Holocene*, 3, 14–26.
Jones, P.D., Osborn, T.J. and **Briffa, K.R.** 1997: Estimating sampling errors in large-scale temperature averages. *J. Climate* 10, 2548–68.
Karl, T.R. and **Williams, C.N. Jr.** 1987: An approach to adjusting climatological time series for discontinuities and inhomogeneities. *J. Clim. App. Met.* 26, 1744–63.
Kameda, T., Narita, H., Shoji, H., Nishio, F. and **Watanabe, O.** 1992: 450-year summer temperature record from melt feature profile in South Greenland ice core. In Mikami T., editors, *Proceedings of the International Symposium on the Little Ice Age climate* Department of Geography, Tokyo Metropolitan University, Tokyo, 101–107.
Knutson, T.R., Manabe, S. and **Gu, D.** 1997: Simulated ENSO in a global coupled ocean-atmosphere model: multidecadal amplitude modulation and CO_2 sensitivity. *J. Climate* 10, 138–61.
Lamb, H.H. 1977: *Climate: present, past and future, vol. 2: climatic history and the future*. London: Methuen.
Lara, A. and **Villalba, R.** 1993: A 3,620-year temperature reconstruction from *Fitzroya Cupressoides* tree rings in southern South America. *Science* 260, 1104–06.
Lean, J., Beer, J. and **Bradley, R.** 1995: Reconstruction of solar irradiance since 1610: implications for climatic change. *Geophys. Res. Letts* 22, 3195–98.
Lough, J.M. and **Barnes, D.J.** 1997: Several centuries of variation in skeletal extension, density and calcification in massive *Porites* colonies from the Great Barrier Reef: a proxy for seawater temperature and a background of variability against which to identify unnatural changes. *J. Exp. Marine Biology and Ecology* 211, 29–67.
Lough, J.M., Barnes, D.J. and **Taylor, R.B.** 1996: The potential of massive corals for the study of high-resolution climate variation in the past millennium. In Jones, P.D., Bradley, R.S. and Jouzel, J., editors, *Climatic variations and forcing mechanisms of the last 2000 years*, Berlin: Springer, 355–71.
Luckman, B.H., Briffa, K.R., Jones, P.D. and **Schweingruber, F.H.**

1997: Summer temperatures at the Columbia Icefield, Alberta, Canada, 1073–1987. *The Holocene* 7, 375–89.

Manley, G. 1974: Central England temperatures: monthly means 1659 to 1973. *Quart. J. Roy. Meteor. Soc* 100, 389–405.

Mann, M.E., Bradley, R.S. and **Hughes, M.K.** 1998: Global temperature patterns over the past five centuries: implications for anthropogenic and natural forcing of climate. *Nature* 392, 779–87.

Mitchell, J.F.B., Johns, T.C., Gregory, J.M. and **Tett, S.F.B.** 1995: Transient climate response to increasing sulphate aerosols and greenhouse gases. *Nature* 376, 501–504.

Morgan, V. and **van Ommen, T.D.** 1997: Seasonality in late-Holocene climate from ice core records. *The Holocene* 7, 351–54.

Munk, W., Snodgrass, F. and **Tucker, M.** 1959: Spectra of low-frequency ocean waves. *Bull. Scripps. Inst. Oceanography* 7, 283–362.

Nicholls, N., Gruza, G.V., Jouzel, J., Karl, T.R., Ogallo, L.A. and **Parker, D.E.** 1996: Observed climate variability and change. In Houghton, J.T., Meira Filho, L.G., Callander, B.A., Harris, N., Kattenberg, A. and Maskell, K., editors, *The IPCC Second Scientific Assessment*, Cambridge: Cambridge University Press, 133–92.

Osborn, T.J., Briffa, K.R. and **Jones, P.D.** 1998: Adjusting variance for sample size in treering chronologies and other regional-mean timeseries. *Dendrochronologia*, in press.

PAGES 1995: *Paleoclimates of the Northern and Southern Hemispheres. The PANASH Project. The Pole-Equator-Pole Transects*. PAGES series 95–1, 92 pp.

Parker, D.E., Jones, P.D., Folland, C.K. and **Bevan, A.** 1994: Interdecadal changes of surface temperature since the late 19th century. *J. Geophys. Res.* 99, 14373–99.

Parker, D.E., Legg, T.P. and **Folland, C.K.** 1992: A new daily Central England temperature series. *Int. J. Climatol.* 12, 317–42.

Peel, D.A. 1992: Ice core evidence from the Antarctic Peninsula region. In Bradley, R.S. and Jones, P.D., editors, *Climate since* AD *1500*, London: Routledge, 549–71.

Peel, D.A., Mulvaney, R., Pasteur, E.C. and **Chenery, C.** 1996: Climatic changes in the Atlantic sector of Antarctica over the past 500 years from ice core and other evidence. In Jones, P.D., Bradley, R.S. and Jouzel, J., editors, *Climatic variations and forcing mechanisms of the last 2000 years*, Berlin: Springer, 243–62.

Pfister, C. 1992: Monthly temperature and precipitation in central Europe 1525–1979: quantifying documentary evidence on weather and its effects. In Bradley, R.S. and Jones, P.D., editors, *Climate since* AD *1500*, London: Routledge, 549–71.

Peterson, T.C. and **Easterling, D.R.** 1994: Creation of homogeneous composite climatological reference series. *Int. J. Climatol.* 14, 671–79.

Pollack, H.N., Shen, P.Y. and **Huang, S.** 1996: Inference of ground surface temperature history from subsurface temperature data: interpreting ensembles of borehole logs. *Pure and Applied Geophysics* 147, 537–50.

Quinn, T.M., Crowley, T.J., Taylor, F.W., Henin, C., Joannot, P. and **Join, Y.** 1998: A multicentury coral stable isotope record from New Caledonia. *Paleoceanography*, in press.

Raper, S.C.B., Briffa, K.R. and **Wigley, T.M.L.** 1996: Glacier change in northern Sweden from AD 500: a simple geometric model of Storglaciären. *Journal of Glaciology* 42, 341–51.

Rhoades, D.A. and **Salinger, M.J.** 1993: Adjustment of temperature and rainfall records for site changes. *Int. J. Climatol*. 13, 899–913.

Rind, D. and **Overpeck, J.T.** 1994: Hypothesized causes of decade-to-century scale climate variability: climate model results. *Quaternary Science Reviews* 12, 357–74.

Roig, F.A., Roig, C., Rabassa, J. and **Boninsegna, J.** 1996: Fuegian floating tree-ring chronology from sub-fossil *Nothofagus* wood. *The Holocene* 6, 469–76.

Santer, B.D., Taylor, K.E., Wigley, T.M.L., Penner, J.E., Jones, P.D. and **Cubasch, U.** 1995: Towards the detection and attribution of an anthropogenic effect on climate. *Climate Dynamics* 12, 79–100.

Santer, B.D., Wigley, T.M.L., Barnett, T.P. and **Anyamba, E.** 1996: Detection of climate change and attribution of causes. In Houghton, J.T., Meira Filho, L.G., Callander, B.A., Harris, N., Kattenberg, A. and Maskell, K., editors, *The IPCC Second Scientific Assessment*, Cambridge: Cambridge University Press, 407–44.

Stouffer, R.J., Manabe, S. and **Vinnikov, K.Ya.** 1994: Model assessment of the role of natural variability in recent global warming. *Nature* 367, 634–36.

Tarussov, A. 1992: The Arctic from Svalbard to Severnaya Zemlya: climatic reconstruction from ice cores. In Bradley, R.S. and Jones, P.D., editors, *Climate since* AD *1500*, London: Routledge, 505–16.

Tett, S.F.B., Johns, T.C. and **Mitchell, J.F.B.** 1997: Global and regional variability in a coupled AOGCM, 1997. *Climate Dynamics* 13, 303–23.

Tett, S.F.B., Mitchell, J.F.B., Parker, D.E. and **Allen, M.R.** 1996: Human influence on the atmospheric vertical temperature structure: detection and observations. *Science* 274, 1170–73.

Thompson, L.G. 1996: Climatic changes for the last 2000 years inferred from ice core evidence in tropical ice cores. In Jones, P.D., Bradley, R.S. and Jouzel, J., editors, *Climatic variations and forcing mechanisms of the last 2000 years*, Berlin: Springer, 281–95.

Villalba, R. 1990: Climatic fluctuations in Northern Patagonia during the last 1000 years as inferred from tree-ring records. *Quaternary Research* 34, 346–60.

Wang, S-W. and **Zhao, Z-C.** 1981: Droughts and floods in China, 1470–1979. In Wigley, T.M.L., Ingram, M.J. and Farmer, G., editors, *Climate and history*, Cambridge: Cambridge University Press, 271–88.

Williams, L.D. and **Wigley, T.M.L.** 1983: A comparison of evidence for the Late Holocene summer temperature variations in the Northern Hemisphere. *Quaternary Research* 20, 286–307.

Yao, T., Thompson, L.G., Mosley-Thompson, E., Zhihong, Y., Xingping, Z and **Lin P-N.** 1996: Climatological significance of $\delta^{18}0$ in north Tibetan ice cores. *J. Geophys. Res.* 101, 29531–37.

Zhang, D. 1994: Evidence for the existence of the Medieval Warm Period in China. *Climatic Change* 26, 289–97.