



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

# Detection of human influence on a new, validated 1500-Year temperature reconstruction

**Citation for published version:**

Hegerl, GC, Crowley, TJ, Allen, M, Hyde, WT, Pollack, HN, Smerdon, J & Zorita, E 2007, 'Detection of human influence on a new, validated 1500-Year temperature reconstruction', *Journal of Climate*, vol. 20, no. 4, pp. 650-666. <https://doi.org/10.1175/JCLI4011.1>

**Digital Object Identifier (DOI):**

[10.1175/JCLI4011.1](https://doi.org/10.1175/JCLI4011.1)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Journal of Climate

**Publisher Rights Statement:**

Published in the Journal of Climate by the American Meteorological Society (2007)

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## Detection of Human Influence on a New, Validated 1500-Year Temperature Reconstruction

GABRIELE C. HEGERL,\* THOMAS J. CROWLEY,\* MYLES ALLEN,+ WILLIAM T. HYDE,\*  
HENRY N. POLLACK,# JASON SMERDON,@ AND EDUARDO ZORITA&

\**Division of Earth and Ocean Sciences, Nicholas School for the Environment and Earth Sciences, Duke University, Durham, North Carolina*

+ *Climate Dynamics Group, Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, United Kingdom*

# *Department of Geological Sciences, University of Michigan, Ann Arbor, Michigan*  
@ *Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York*  
& *GKSS, Geesthacht, Germany*

(Manuscript received 3 January 2006, in final form 12 June 2006)

### ABSTRACT

Climate records over the last millennium place the twentieth-century warming in a longer historical context. Reconstructions of millennial temperatures show a wide range of variability, raising questions about the reliability of currently available reconstruction techniques and the uniqueness of late-twentieth-century warming. A calibration method is suggested that avoids the loss of low-frequency variance. A new reconstruction using this method shows substantial variability over the last 1500 yr. This record is consistent with independent temperature change estimates from borehole geothermal records, compared over the same spatial and temporal domain. The record is also broadly consistent with other recent reconstructions that attempt to fully recover low-frequency climate variability in their central estimate.

High variability in reconstructions does not hamper the detection of greenhouse gas-induced climate change, since a substantial fraction of the variance in these reconstructions from the beginning of the analysis in the late thirteenth century to the end of the records can be attributed to external forcing. Results from a detection and attribution analysis show that greenhouse warming is detectable in all analyzed high-variance reconstructions (with the possible exception of one ending in 1925), and that about a third of the warming in the first half of the twentieth century can be attributed to anthropogenic greenhouse gas emissions. The estimated magnitude of the anthropogenic signal is consistent with most of the warming in the second half of the twentieth century being anthropogenic.

### 1. Introduction

Climate records over the last millennium provide observational information on natural climate variability on time scales of multiple decades and centuries. It is from this backdrop of natural climate variability that anthropogenic changes need to be distinguished (e.g., Mitchell et al. 2001; International Ad Hoc Detection and Attribution Group 2005). Reconstructions of pre-instrumental surface temperature have employed “proxy” information derived from historical records, tree rings, ice cores, marine and terrestrial sediments

(see review in Jones and Mann 2004; also Mann et al. 1999; Crowley and Lowery 2000; Cook et al. 2004; Moberg et al. 2005) and subsurface temperatures measured in boreholes (Huang et al. 2000; Harris and Chapman 2001; Beltrami 2002). A number of recent studies have used this information to estimate Northern Hemispheric temperature changes during the last 500–1000 years; these studies indicate a difference between the peak of the cold period in the seventeenth-century and the mid-twentieth-century warming that ranges from about 0.4° to 1.0°C.

Concern about the reliability of reconstructions of past temperature arises from a climate model-based evaluation of one of the reconstruction methods (von Storch et al. 2004), which suggests that calibration methods that are based on ordinary least squares regression may fail to recover some of the low-frequency

---

*Corresponding author address:* Gabriele Hegerl, Nicholas School of the Environment and Earth Sciences, Duke University, Durham, NC 27708.  
E-mail: hegerl@duke.edu

variance in the model's true hemispheric temperature. However, the magnitude of the variance loss may be smaller than estimated by von Storch et al. (Mann et al. 2005; see also Wahl et al. 2006) and depends on the properties of the noise (von Storch et al. 2006).

A further concern is that some tree-ring data, which dominate the input to most high-resolution surface proxy composites, may not adequately recover low-frequency centennial–millennial-scale variability unless standardized to preserve low-frequency information (Cook et al. 1995; Briffa et al. 2001; Esper et al. 2002). These concerns are emphasized by the discrepancy between most surface proxy reconstructions and Northern Hemispheric borehole composites (see Jones and Mann 2004). One exception involves a recent reconstruction that bases its low-frequency variability on lower-resolution sediment data from lake and ocean cores, and adjusts the variance to instrumental data during the calibration interval (Moberg et al. 2005).

In this paper we propose a simple calibration method that does not lose variance. It is based on a total least squares fit that allows for proper accounting of noise in both proxy and instrumental data during calibration, yielding a reconstruction with realistic amplitude and uncertainty ranges. When applied to model-generated “pseudoproxies” that are sampled over the same locations, and show similar correlations to local temperatures, our approach captures the full range of past climate variability and provides a reliable estimate of uncertainty. We use this technique to reconstruct a decadal record over the past 1450 yr. Over the last 500 yr, this new record is consistent with borehole results over the same time and space domains, and shows similar levels of Little Ice Age cooling and Medieval Warm Period warming as other reconstructions using data and methods that aim at fully recovering low-frequency variance.

Since the twentieth-century trend stands out less from trends in previous centuries in reconstructions with higher variance, high-variance reconstructions have sometimes been used to question the importance of anthropogenic forcing. However, natural influences on climate, such as changes in volcanism and possibly solar radiation, are responsible for a substantial fraction of past climate variations (e.g., Robock and Free 1995; Crowley 2000; Hegerl et al. 2003; Bertrand et al. 2002; Weber 2005; Stendel et al. 2005). By quantifying the influence of external forcing on these high-variance reconstructions, we find that high variability does not prevent confident detection and attribution of anthropogenic climate change.

The paper is structured as follows. Section 2 intro-

duces our reconstruction method, and section 3 demonstrates the reliability of this method with climate model data. Section 4 compares our reconstruction with other high-variance reconstructions and borehole data. A detection and attribution method is applied in section 5 to quantify the influence of external forcing on our reconstruction and other high-variance reconstructions. We summarize results in section 6.

## 2. Reconstruction of hemispheric temperature variability

Our new reconstruction of decadal Northern Hemispheric mean temperature is related to an earlier reconstruction based on a simple average approach (Crowley and Lowery 2000), but uses updated records, a modified reconstruction method, and a new calibration technique. The reconstruction consists of three individual segments. A baseline reconstruction uses 12 decadal records and covers the period to 1505. One longer, less densely sampled land temperature reconstruction, which we call CH-blend (long), is based on seven records back to A.D. 946, and CH-blend (Dark Ages) consists of five records back to A.D. 558. The three reconstructions are each based on equally long proxy records or regional proxy reconstructions based on multiple records that are distributed across the extratropical Northern Hemisphere (NH). Using only data covering the entire length of each individual reconstruction largely avoids inhomogeneity in the variability of reconstructions that arises due to changes in spatial sampling density over time (see Crowley and Lowery 2000).

### a. Construction of proxy time series

We use decadal or decadal smoothed proxy records since information about low-frequency variability is crucial to separate natural climate variability from greenhouse warming, and because fewer records are needed to reliably sample decadal and hemispheric temperature variability (global decadal mean surface temperature is estimated to have effectively 8–16 spatial degrees of freedom, see Jones et al. 1997; Zwiers and Shen 1997). Since almost all long proxies are located in the mid- to high northern land areas, the reconstruction is calibrated to 30°–90°N temperature. A second version of the reconstruction is calibrated to land and ocean regions north of 30°N, which is used for the detection study (see below) and a study estimating climate sensitivity (Hegerl et al. 2006). The tree-ring data have mostly been processed in a manner that pre-

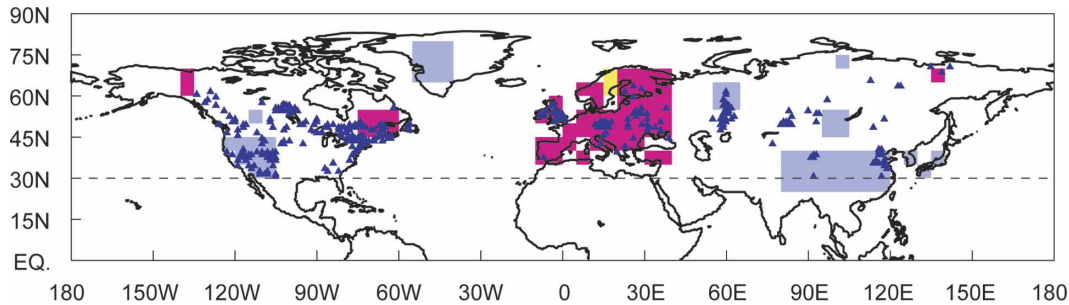


FIG. 1. Regions used for the paleoreconstruction (pink, yellow, and blue areas) and for the borehole reconstruction (blue triangles). The data shown in pink are used only for the reconstruction extending to A.D. 1505. Shaded blue regions are used in the short reconstruction and the long reconstructions to 946, of these, only the Mongolian and Alberta (Canada) records drop out in the longest reconstruction to 558. The Scandinavian record shown in yellow is used in both long and short reconstructions, as there is only a small correlation between this record and the large-scale European record that encompasses it.

serves low-frequency variance (Briffa et al. 2001; Esper et al. 2002; details are given in appendix A).

The primary reconstruction for the time interval since 1505 is based on 12 well-spaced sites, some of which are area averages based on multiple records (Fig. 1, for individual sites see appendix A). The first step of the reconstruction technique is to scale the individual proxy records to unit standard deviation, weigh them by their correlation with decadal NH 30°–90°N temperature (land or land and ocean, depending on the target of reconstruction) during the period 1880–1960, and then average them. This yields a unitless paleo time series  $\hat{\theta}_{\text{pal}}(t)$  that shows high correlations with the instrumental data during the calibration period 1880–1960 (0.97 for land only and 0.92 for land and ocean temperatures, correlations become 0.82 for land and 0.75 for land and ocean after detrending).

Such a weighted average is mathematically equivalent to a multiple regression of proxy data onto the target of reconstruction for uncorrelated individual records (our records have only modest correlations with each other), but is much less prone to overfitting than full multiple regression. This is important because the shortness of the common period of the decadal instrumental and proxy data (1880–1960, prior to that, the spatial sampling of instrumental data is very inhomogeneous), does not allow a formal separation of the period of overlap into a calibration and a validation period. Note that the only information from instrumental data used for computing  $\hat{\theta}_{\text{pal}}(t)$  is the correlation weight, which has only a small influence on the resulting time series; an unweighted, and hence fully independent, average of the sites correlates  $r = 0.96$  (0.79 detrended) with instrumental land data compared to the correlations of the weighted average of 0.97 (0.82 detrended).

### b. Calibration to Northern Hemispheric temperature

The next step is to calibrate the proxy time series  $\hat{\theta}_{\text{pal}}(t)$  against the average NH 30°–90°N (land or land and ocean) instrumental temperature record  $\hat{T}_{\text{inst}}(t)$ , using a conversion factor  $\alpha$  (hats denote the estimated time series). This is accomplished by fitting a statistical model of the form

$$T_{\text{inst}}(t) + \varepsilon_{\text{inst}}(t) = \alpha[\theta_{\text{pal}}(t) + \varepsilon_{\text{pal}}(t)], \quad (1)$$

where  $\varepsilon_{\text{inst}}$  refers to realizations of sampling error and random measurement error in instrumental data, and  $\varepsilon_{\text{pal}}$  to realizations of sampling error and nontemperature variability in proxy records. Often,  $\alpha$  is estimated by an ordinary least squares (OLS) fit, neglecting the considerable uncertainty  $\varepsilon_{\text{pal}}$ , which leads to a low bias of the scaling factor  $\alpha$  (see, e.g., von Storch et al. 2004; Allen and Stott 2003) and thus underestimates past temperature ranges in the scaled proxy time series. The loss of variance is largest if the correlation between  $\hat{T}_{\text{inst}}$  and  $\hat{\theta}_{\text{pal}}$  is low, as tends to be the case when annual data are reconstructed. To avoid this problem, some reconstructions use less rigorous calibration methods, such as adjusting to the variance of instrumental data or qualitatively scaling the low-frequency component. However, if a total least squares (TLS) fit is used, in which noise in both instrumental and proxy data is considered, the shared signal with the instrumental time series is scaled to the same amplitude (Adcock 1878; Allen and Stott 2003). This method has previously been used for estimating the amplitude of noisy fingerprints from observations in optimal fingerprinting (Stott et al. 2003). Note that if the uncertainties in the paleoreconstruction are much larger than in instrumental data, an alternative is the use of inverse

regression, neglecting error in instrumental data (Coelho et al. 2004).

The TLS fit requires an estimate of the ratio of the variances of  $\varepsilon_{\text{pal}}$  to  $\varepsilon_{\text{inst}}$ . The variance of  $\varepsilon_{\text{inst}}$  is estimated from subsampled climate model data (von Storch et al. 2004; Zorita et al. 2003) yielding an error estimate that is consistent with earlier published estimates of uncertainty associated with sampling error in decadal averaged instrumental data for the entire NH (Jones et al. 1997). Note that we assume that systematic instrumental errors are small relative to sampling errors. This may not be strictly correct for sea surface temperature data, which may be affected by systematic biases such as the correction for using canvas buckets in earlier temperature measurements (Folland et al. 2001, Brohan et al. 2006). The latter are, however, only used in the land and ocean reconstruction.

Since the variance of  $\varepsilon_{\text{pal}}$  is unknown, we account for this uncertainty by determining the scaling factor  $\alpha$  and its confidence interval for a wide range of plausible variance ratios for  $\varepsilon_{\text{pal}}$  to  $\varepsilon_{\text{inst}}$ . For each, the likelihood of the regression residual being drawn from errors with variances of  $\varepsilon_{\text{pal}}$  and  $\varepsilon_{\text{inst}}$  is determined. We then average these regression results, weighted by the likelihood of each residual, to arrive at an overall scaling  $\alpha$  and its uncertainty (details are given in appendix B). The time series  $\hat{\theta}_{\text{pal}}$  is then scaled by  $\alpha$ . The overall uncertainty for the reconstruction is derived from a combination of the uncertainty in the amplitude ( $\alpha$ ) and the range of error variances for  $\varepsilon_{\text{pal}}$  that is consistent with the calibration residual (see appendix B). Note that our approach assumes that the errors  $\varepsilon_{\text{pal}}$  and  $\varepsilon_{\text{inst}}$  are uncorrelated and normally distributed. The first assumption appears reasonable since red noise possibly present in the individual records should have been strongly reduced by hemispheric averaging in  $\hat{\theta}_{\text{pal}}$  (note the small changes in the correlation to instrumental temperature between detrended or nondetrended data), and the second is justified for hemispheric means by the central limit theorem.

The longer reconstructions are constructed in the same manner as CH-blend (short) and are calibrated to CH-blend (short) rather than to instrumental data, since the longer overlap period allows for a more accurate calibration. Due to their high correlation with CH-blend (short) ( $r = 0.97$  and  $r = 0.93$ ), this second calibration step adds only a small additional uncertainty that is accounted for in the uncertainty range (see appendix B). Note that the high correlations and overall similarity of the longer, sparser sampled reconstructions CH-blend (long) and CH-blend (Dark Ages) with the more densely sampled CH-blend (short) during the period of overlap (see Fig. 2) suggests that even a few

well-spaced decadal records can provide a surprisingly reliable estimate of decadal variability in the last five centuries.

Figure 2 shows the time series calibrated to instrumental land annual temperatures 30°–90°N. The CH-blend (short) is similar to the reconstruction used in Hegerl et al. (2003) and Crowley and Lowery (2000; see also Crowley 2000), but has enhanced amplitude (by 13% for zonal mean temperature and 7% for land temperature). Larger enhancement due to TLS calibration was found for other published records if recalibrated, yielding on the order of 20%–30% enhancement, and if annual reconstructions are attempted. Note that a prominent eleventh-century peak in the 2003 reconstruction is not as evident after balancing the composite with more data from Asia.

### 3. Testing the reconstruction method with climate model data

We have tested our reconstruction technique with model-derived synthetic proxy data (i.e., a “perfect model approach”) using the same atmosphere–ocean general circulation model (AOGCM) simulation of climate of the past few centuries that was used by von Storch et al. (2004). The simulation was driven by estimates of past volcanic, solar, and greenhouse gas forcing, primarily from Crowley (2000). The AOGCM provides dynamically consistent data with approximately realistic teleconnections, enabling a reliable estimate of the effects of sparse networks and calibration techniques on reconstructions.

We first create a synthetic instrumental (“pseudoinstrumental”) record from the simulation, using the same spatial coverage of annual mean land temperatures as in the actual instrumental record. Next, we derive a synthetic paleoreconstruction (“pseudoproxy reconstruction”) using grid boxes that represent the areas and sites used in CH-blend (see Fig. 1). These model data were perturbed by adding decadal white noise until the correlation between pseudoproxy data and the pseudoinstrumental decadal data was similar to those in CH-blend over the calibration interval. This reduces the correlation of unperturbed to perturbed decadal local model data to values of 0.1–0.7, averaging 0.45, which is similar to correlations between decadal proxy data and local temperature data. These pseudoproxies are then weighted by their correlation with pseudoinstrumental mean temperature from 30° to 90°N, averaged and calibrated to the pseudoinstrumental NH 30°–90°N mean using the TLS approach.

Figure 3 shows the mean of 10 such pseudoproxy reconstructions with different realizations of noise

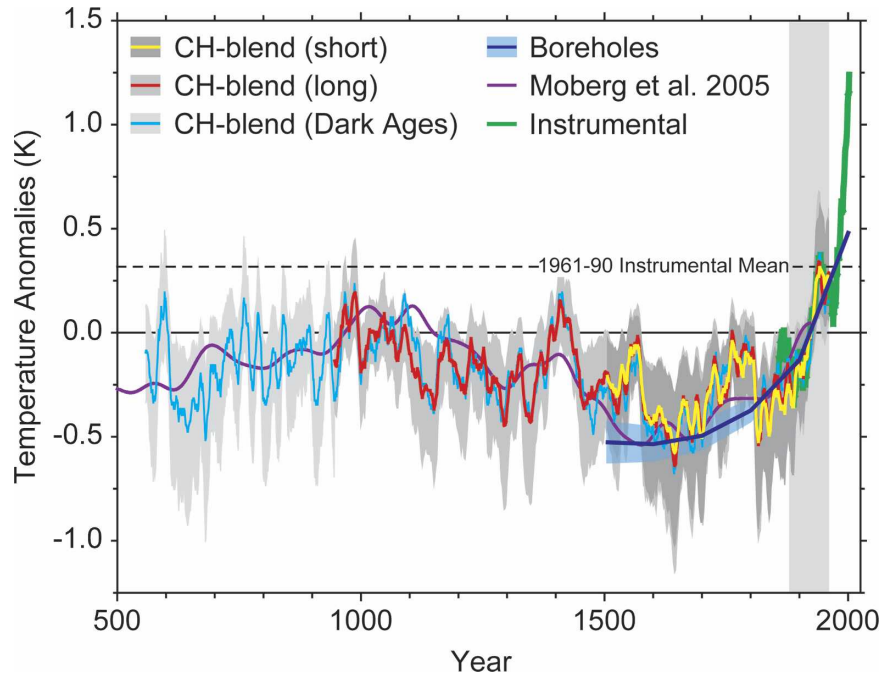


FIG. 2. Proxy reconstruction of  $30^{\circ}$ – $90^{\circ}$ N mean annual decadal averaged temperatures over land back to A.D. 558. The long time series is made from three segments covered by different amounts of data, which are kept constant within that segment. Gray shaded ranges give the 95% uncertainty bounds of decadal temperature estimates. Another recent proxy reconstruction and the borehole reconstruction over the same spatial domain are shown for comparison. Each time series is plotted relative to its 1880–1960 mean (shown by the horizontal gray shaded region). This is a departure from the commonly used reference period of the 1961–90 instrumental mean [indicated by a dotted line; e.g., Pollack and Smerdon (2004), pinned the zero crossing of the twentieth-century trend in the borehole time series to the equivalent point in the linear trend of the twentieth-century instrumental data plotted relative to the 1961–90 mean]. We now use the 1880–1960 calibration period since it demonstrates most clearly the differences in the reconstructions prior to calibration.

added, and two individual reconstructions with the highest and the lowest variance among these 10. This test shows that the reconstruction method reproduces the actual model  $30^{\circ}$ – $90^{\circ}$ N land temperature variability very well (correlation over the calibration interval is on average  $r = 0.97$ , and  $r = 0.68$  if using detrended time series; correlation with the model's true  $30^{\circ}$ – $90^{\circ}$ N mean ranges from  $r = 0.96$  to  $r = 0.98$  for different realizations of error over entire reconstruction). On average 97% of the variance of the true model  $30^{\circ}$ – $90^{\circ}$ N mean has been recovered. This result is based on a calibration period of 1898–1977, since the peak early-twentieth-century warming occurs later in the model and hence the shape of the time series in that period is more similar to that in the observations than the period 1881–1960, which is dominated by a linear trend (cf. Figs. 2 and 3). Note that each calibration interval is affected by different sampling error due to the extent to which the sampling regions in proxy and instrumental data capture simulated hemispheric mean change. This leads to

differences in the extent of variance recovered in the reconstruction, such as the small overestimate of cooling into the Little Ice Age seen in Fig. 3. Based on an 1880–1960 calibration period, the reconstructed variance is on average underestimated (although less so than if using an OLS technique), and it is overestimated if using a later calibration period. However, in each of these cases, as well as in the case shown in Fig. 3, the true hemispheric average remains within the estimated uncertainty range slightly more than 95% of the time, indicating that our uncertainty ranges are conservative and validating the reconstruction method.

In contrast, using OLS instead of TLS for calibration reduces the intercentury variance (filtered by applying a consecutive running 60- and 40-yr moving average) of the  $30^{\circ}$ – $90^{\circ}$ N mean on average by a factor of 0.81 for calibration of decadal data, and by a factor of 0.40 to 0.30 for annual data. This indicates a very severe reduction of low-frequency variance using OLS calibration. If the proxy data are OLS fitted to local decadal

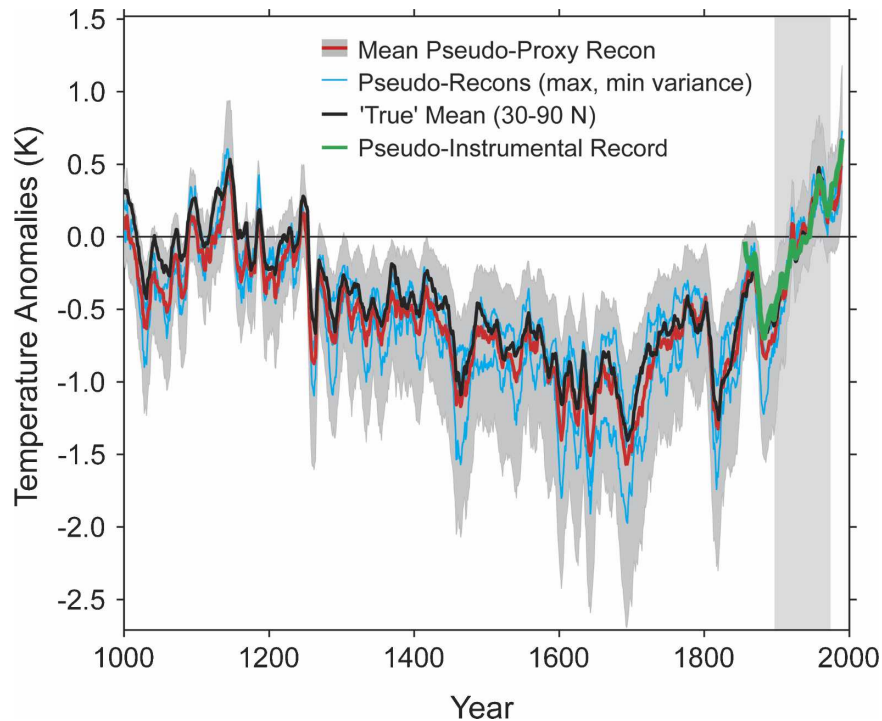


FIG. 3. Validation of our reconstruction method with climate model data perturbed by noise. The mean of 10 synthetic reconstructions and their uncertainty ranges, derived using a climate model simulation from 1000 to 1990, (red, shown with 95% error bars) is compared to the true NH 30°–90°N annual model mean (black) and the subsampled model data using instrumental coverage (green). The calibration interval is shown as vertical shaded region (note that results vary to some extent with calibration period). The case with highest and lowest variance from the 10 synthetic reconstructions (blue) is also shown. Uncertainty ranges are smaller for time periods with temperatures close to the calibration period, while periods far away from them are more affected by scaling uncertainty.

temperature and then averaged, the variance is reduced by about half.

#### 4. Comparison of reconstructed temperature history with other reconstructions based on surface and borehole data

Our reconstruction (Fig. 2) shows features seen in many other reconstructions—the broad transition from the relatively warm Middle Ages to the Little Ice Age, and the warming after the Little Ice Age to the present, along with multidecadal–centennial-scale fluctuations (see Jones and Mann 2004). The cool interval before the Middle Ages coincides with the Early Medieval Cool Period, a time of glacial advance in many regions (Grove 1988; Oerlemans 2005), which corresponds approximately to the historical interval after the collapse of the Roman Empire sometimes called the Dark Ages. Late-twentieth-century temperatures substantially exceed the maximum temperatures in the Middle Ages,

confirming that the late twentieth century is probably the warmest interval in at least 1500 yr.

Table 1 gives correlations of our reconstruction with other reconstructions, while Fig. 4 shows a comparison of our reconstruction with two other recently published reconstructions with high variance: that by Moberg et al. (2005), and by Esper et al. (2002), calibrated by Cook to NH 30°–90°N land data (Cook et al. 2004). For the comparison with Moberg et al. (2005), we have recalibrated our data to NH 0°–90°N annual means. The CH-blend data show a somewhat smaller level of cooling into the Little Ice Age and a similar warming in the Medieval Warm Period than Moberg et al. (2005), along with overall similar features to that latter reconstruction. Given the uncertainty in the amplitude of CH-blend (only amplitude uncertainty for NH 30°–90°N land is shown in Fig. 4a for simplicity, note that the amplitude uncertainty for 0°–90°N land and ocean is substantially larger) and in Moberg et al. (dotted line), the reconstructions are consistent in their esti-

TABLE 1. The 20-point smoothed correlations between different proxy reconstructions. Reconstructions are Mann et al. (1999); Esper et al. (2002), Jones et al. (2001), Moberg et al. (2005), and CH-blend (short, Dark Ages, and long, calibrated to 30°–90°N temperature over land). The right column gives the average correlation between CH-blend, Mann, Jones, and Esper.

	CH-blend	CH-DA	CH-long	Mann	Jones	Esper	Moberg	Mean and std dev
CH-blend	1.0	0.91	0.98	0.69	0.71	0.82	0.50	0.74 ± 0.07
CH-dark		1.0	0.93	0.59	0.65	0.74	0.60	0.66 ± 0.06
CH-long			1.0	0.61	0.63	0.82	0.55	0.69 ± 0.12
Mann et al.				1.0	0.70	0.52	0.44	0.60 ± 0.13
Jones et al.					1.0	0.53	0.50	0.61 ± 0.12
Esper et al.						1.0	0.28	0.52 ± 0.09
Moberg et al.							1.0	0.41 ± 0.12

mate of Little Ice Age cooling and Medieval Warm Period temperatures. Similarly, our reconstruction for 30°–90°N land data is broadly consistent with the Cook et al. (2004) calibrated version of Esper et al. (2002) for both periods. Note, however, that Esper et al. (2002) shows differences from other records in the twelfth and thirteenth centuries.

Figure 2 also shows a comparison between CH-blend and subsurface temperature histories measured in boreholes over the same spatial domain (30°–90°N land). Boreholes provide estimates of centennial-scale surface temperature changes over the past five centuries that are independent of both the instrumental record and climate proxies. Here we compare the history of temperature changes estimated from 631 extratropical boreholes (latitude >30°N, see Fig. 1 for locations) to that derived from the multiproxy reconstruction described above. The ensemble of individual borehole results has been averaged onto a 5° × 5° spatial grid to yield the area-weighted reconstruction shown in Fig. 2 (Pollack and Smerdon 2004, see also Pollack and Huang 2000), which is in good agreement with our surface proxy reconstruction for extratropical land data. An implicit assumption in the borehole verification of the new multiproxy reconstruction is that borehole reconstructions represent a valid target of verification for surface proxy reconstructions (i.e., that both track changes in the surface air temperature). An analysis of millennial simulations of air and subsurface temperatures from AOGCMs (González-Ruoco et al. 2003, 2006) shows that despite seasonal decoupling between air and ground temperatures, deep soil temperature is a good proxy for the annual surface air temperature on land.

We further compare the new multiproxy reconstruction with borehole results by calculating the subsurface temperature anomaly for each reconstruction (Fig. 4b). The temperature anomaly represents the transient warming of the subsurface at a given time as the subsurface evolves from a previous temperature field. The subsurface temperature anomalies for both the bore-

hole and CH-blend (short) reconstructions are calculated using a conductive forward model in which each reconstruction comprises a time-varying surface boundary condition (see Carslaw and Jaeger 1959). The borehole subsurface temperature anomaly shown in Fig. 4b is calculated from the mean of 631 reconstructions, each based on individually inverted borehole temperature profiles, rather than calculating an average borehole profile directly from the data. A comparison of the calculated subsurface temperature anomaly with the average observed anomaly determined by Harris and Chapman (2001) shows the two to be almost identical. Because the geothermal reconstruction extends back to only A.D. 1500, we have compared only a segment of the multiproxy reconstruction with it. To achieve isolation of the post-A.D. 1500 history, we assume a steady-state (zero heat flux) surface boundary condition prior to A.D.1500 for both the multiproxy and geothermal reconstructions. This assumption enables a direct comparison of the two reconstructions in their common time interval, 1505–1960. As can be seen in Fig. 4b, the subsurface expressions of both reconstructions are almost indistinguishable. The uncertainty envelopes of each are also shown; these were calculated by driving the model with the upper and lower uncertainty time series shown in Fig. 2. These uncertainties also encompass the results generated from the range of reasonable initialization temperatures in the conductive-forward model.

## 5. Implications for attribution of climate change to causes

Changes in external radiative forcing need to be considered in order to understand the origin of past climate variability. We thus apply a detection and attribution analysis in order to quantify the role of external forcing, including greenhouse gas forcing, in our reconstruction and in other reconstructions with substantial century-scale variability. Such an analysis based on paleoreconstructions from several centuries enables a better sepa-



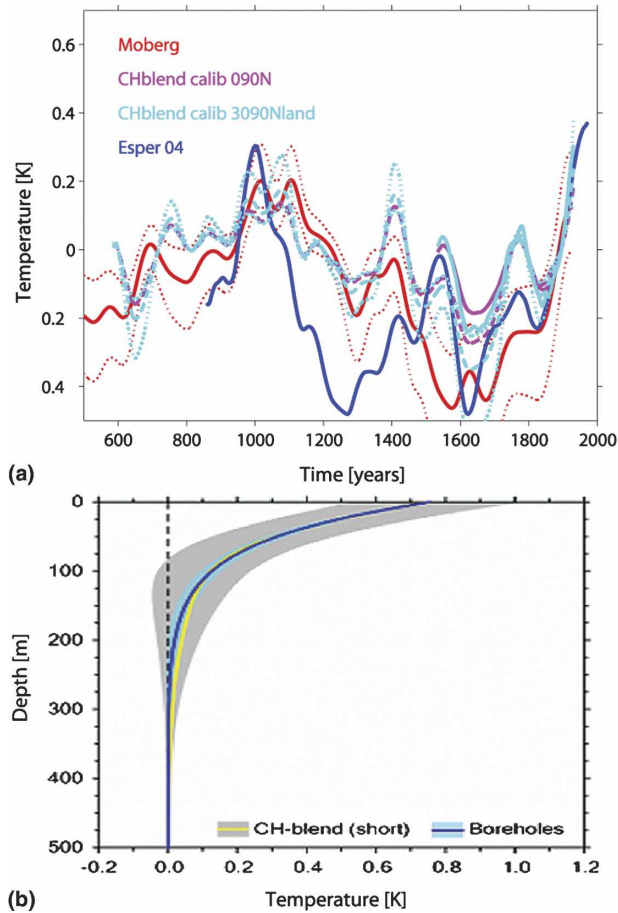


FIG. 4. Comparison of CH-blend with two other reconstructions and boreholes. (a) Comparison of CH-blend (short: solid, long: dashed) and Esper et al. (after Cook et al. 2004) time series have been smoothed by applying a 60- and then a 40-yr moving average (this yields nearly identical smoothing as that used by Moberg et al. 2005). For CH-blend land data, we also show the 95% uncertainty range in the amplitude of the reconstruction (dotted, see Fig. 2); for Moberg et al.'s reconstruction, the 95% range is shown by red dots. All reconstructions have been centered on the period 1860 to 1920 for a quantitative comparison with Moberg (which ends in 1925). Esper and CH-blend are shortened due to the application of the filter. (b) Comparison of two forward modeled subsurface temperature anomalies driven respectively by the composite extratropical borehole reconstruction and the CH-blend (short) reconstruction.

ration of the climate response to natural forcing, particularly between solar and volcanic forcing, both of which are spuriously correlated over the twentieth century (e.g., North and Stevens 1998), and between solar and anthropogenic forcing (e.g., Cubasch et al. 1997). Long reconstructions also provide an opportunity to estimate internal climate variability independently from climate model data. The dependence of detection results on climate model control simulations is an im-

portant uncertainty in the attribution of climate change to causes (Mitchell et al. 2001).

We use a multifingerprint detection and attribution method (Hegerl et al. 2003; see also Hegerl et al. 1997; Tett et al. 1999) using fingerprints for solar, volcanic, and a combination of greenhouse gas and aerosol forcing. Fingerprints for external forcing are derived from simulations with a two-dimensional Energy Balance Model (EBM) and additionally from the ECHAM Hamburg Ocean Primitive Equation (ECHO-G) AOGCM used above. The EBM is the same as used in Hegerl et al. (2003), it has a realistic land-sea distribution and seasonal cycle. It is driven with solar forcing variations (“sol”), greenhouse gas plus sulfate aerosol forcing (“anthro”), and forcing by explosive volcanism (“vol”, for forcing details see appendix C). All climate model data are sampled over the same spatial and temporal domain that is represented by the proxy reconstruction (see Table 2, note that we use zonal mean CH-blend data since these are simulated more realistically in the EBM). The influence of external forcings on the proxy reconstruction,  $T_{\text{proxy}}$ , is assessed by estimating a scaling factor  $\beta$  to best match the time-dependent fingerprints  $f(t)$  to observations (note that this scaling factor is determined by an OLS approach since the EBM provides noise-free fingerprints):

$$T_{\text{proxy}}(t) = \beta_{\text{sol}} \times f_{\text{sol}}(t) + \beta_{\text{vol}} \times f_{\text{vol}}(t) + \beta_{\text{anthro}} \times f_{\text{anthro}}(t) + \varepsilon_{\text{noise}}(t). \quad (2)$$

We estimate the amplitudes  $\beta$  of the fingerprints and the residual climate variability  $\varepsilon_{\text{noise}}$  simultaneously from the proxy reconstruction from 1270 on, since forcing data are less reliable prior to that time. We show results mainly for CH-blend (long) for this analysis; results for CH-blend (short) are similar. Errors in the overall magnitude of forcing or climate model response will not affect detection results. Random errors in the size of individual volcanic eruptions tend to influence results only slightly (see Hegerl et al. 2006). Note also that we do not rotate fingerprints in order to optimize signal-to-noise ratios as often done analysis of instrumental records (e.g., Hegerl et al. 1997; Stott et al. 2003).

A fingerprint is detected in the reconstruction if its scaling factor  $\beta$  is significantly larger than variations in  $\beta$  that occur due to internal climate variability alone. The residual variability  $\varepsilon_{\text{noise}}$  is used as our main estimate of internal climate variability that determines variations in  $\beta$  due to internal climate variability. The estimated residuals are shifted in time by varying increments to increase the number of sample time series for internal climate variability and thus to improve the power of the detection test (see Hegerl et al. 2003). The

TABLE 2. Detection and attribution results for a range of high-variance paleoclimatic records of the last millennium. The “Y” or “N” indicates that the response to external forcing is, or is not, detectable, with the number in brackets giving the best-guess scaling factor for the model’s fingerprint. Here “ghg+aer” denotes the anthropogenic fingerprint, which is a combination of greenhouse gas forcing and aerosol forcing from 1900 on, and an asterisk “\*” marks that a result is sensitive to details of the analysis. The bottom row gives the standard deviation of the (decadally smoothed) residual and, in parentheses, the amount of decadal variance explained by external forcing. The 5%–95% of the estimated contribution of external forcings to the early twentieth-century trend (1901–50) are also given (marked by “20thc,” except for Moberg, which ends too early. Note that the nondetection of anthropogenic forcing in Moberg from 1000 on is largely due to an increase in residual variability, possibly partly due to forcing error.

Record	Briffa et al. (2001)	Esper et al. (2002)	CH-blend	Moberg et al. (2005)	Moberg
analysis period	1402–1940	1400–1960			
represents	20°–90°N land	20°–90°N land	1270–1960	1270–1925	1001–1925
the following:	growing season	growing season	30°–90°N all annual	0°–90°N all annual	0°–90°N all annual
Volcanic	Y (0.9)	Y (1.0)	Y (1.5)	Y (1.1)	Y(1.4)
Solar	N (–0.1)	N (–0.2)	N (0.5)	N (Y periods)	Y(2.2)
Ghg+aer	Y (1.1)	Y (1.9)	Y (1.0)	Y (1.3)*	N
20thc Ghg+aer	10%–99%	50%–100%	22%–52%	—	—
20thc volcanic	18%–50%	18%–51%	16%–39%	—	—
20thc solar	<22%	<15%	<16%	—	—
20thc internal	13%	2%	30%	—	—
Residual std	0.11 (57%)	0.17 (60%)	0.10 (70%)	0.13 (61%)	0.18 (52%)

number of effective degrees of freedom in these samples is limited due to autocorrelation. To estimate this effect, the autocorrelation and thus the decorrelation time of 20-yr offset residuals has been estimated (reasonable changes in offset period do not affect results). The effective sample size is then estimated from the number of samples separated by the decorrelation time, and further reduced by the number of signals fitted to the data (three for the multisignal approach). This often yields less than 10 degrees of freedom. However, all results given here are robust to reasonable changes in estimated degrees of freedom [for the shortest Briffa et al. (2001) time series the residual from using the unscaled solar and volcanic signals is used to preserve more degrees of freedom].

We have assessed the robustness of our attribution results by using the internal variability in the ECHO-G simulations (Zorita et al. 2003; von Storch et al. 2004) rather than residual variability in proxy reconstructions to estimate the significance of scaling factors for fingerprints. We concatenated the control simulation variability and the difference between both forced AOGCM simulations [adjusted by a factor of  $\sqrt{2}$  to account for doubled variance in a difference between simulations]. These data are then treated similarly as the noise residual. Generally, the model has somewhat more variance than the residual from all proxy reconstructions (Table 2, ECHO-G has a standard deviation of 0.18 K for 30°–90°N decadal averaged temperature, note that residual variability from reconstructions has been corrected for bias in variance due to fitting three fingerprints), suggesting that the use of this climate

model’s internal variability is conservative in detection and attribution approaches.

Results using the CH-blend reconstruction show that its decadal temperature variability is largely driven by external forcing (typically 60%–75% of the variance in entire proxy reconstructions, Table 2). We clearly detect the influence of volcanism on hemispheric temperatures (see Fig. 5a), similar to previous studies (Robock and Free 1995; Crowley 2000; Hegerl et al. 2003). The EBM does not simulate changes in atmospheric dynamics associated with volcanism (Shindell et al. 2004), but these affect hemispheric-scale and annual temperature only slightly (Thompson et al. 2000). Consistent with earlier results (Hegerl et al. 2003), a response to solar forcing cannot be robustly distinguished. This can be due to either the climate response to solar forcing being small, or to low-frequency variations in solar forcing being different from estimates used here. The latter is quite possible given large uncertainties in these reconstructions (Lean et al. 2002). Note that the detection and attribution approach has only been applied to the best-guess reconstruction. This is appropriate since the detection of fingerprints is unaffected by the amplitude the reconstruction (see Mitchell et al. 2001), and random sampling errors are indirectly accounted for by the noise residual.

The fingerprint analysis also reveals that anthropogenic greenhouse gas and aerosol forcing is detectable in our reconstruction by the end of the record (<2.5% significance level). The scaling factor for the anthropogenic signal  $\beta_{\text{anthro}}$  is estimated to be more than 4.5 times larger than the standard deviation of scaling fac-

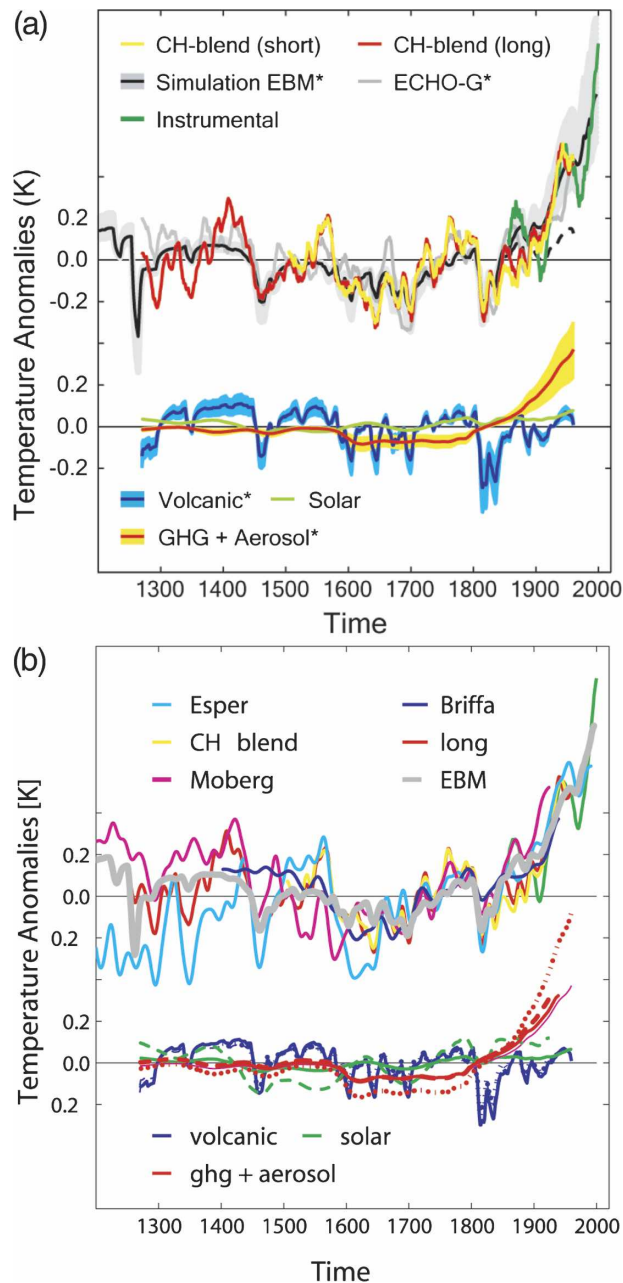


FIG. 5. Contribution of external forcing to reconstructed decadal mean temperature (a) to CH-blend and (b) to a range of high-variance reconstructions. The top of (a) shows CH-blend and the instrumental record compared to simulations with an AOGCM and an EBM forced with estimates of volcanic, solar, and anthropogenic forcing (EBM simulation with natural forcing only dashed). The simulations are scaled to best fit the reconstruction (90% confidence interval for EBM fingerprint shaded). The bottom of (a) shows an estimate of the contribution to CH-blend (long) from individual forcings (volcanism, solar forcing, and greenhouse gas and aerosol forcing combined) and the associated 90% uncertainty range for the detectable signals, which are marked by an asterisk (\*). The top of (b) compares reconstructions by Esper et al. (2002), Moberg et al. (2005), Briffa et al. (2001), and CH-blend (for details, see Table 2) with NH 30°–90°N

tors derived from the residual of the reconstruction. Sensitivity tests reveal that this result is not sensitive to omitting a small reduction in greenhouse gas forcing during the Little Ice Age (Fig. 5). The detected signal of greenhouse warming is consistent with results from the late twentieth century and thus supports the finding that most of the late-twentieth-century warming has been anthropogenic (e.g., Tett et al. 1999; Mitchell et al. 2001). The anthropogenic forcing explains about a third of the trend over the first half of the twentieth century in the reconstruction (90% confidence interval 22%–52%, Table 2, see also Fig. 5a). Additional contributions to the early-twentieth-century warming arise from a decrease in volcanism (16%–39%), internal variability (best estimate 30%, cf. Delworth and Knutson 2000), and an uncertain contribution from an increase in solar forcing that is smaller than typically estimated from data for the twentieth century alone (Stott et al. 2003).

The fingerprint of all radiative forcings combined from the ECHO-G simulations (von Storch et al. 2004; Zorita et al. 2003) can also be detected in CH-blend (long) from 1270 on (using a single model simulation as fingerprint) and CH-blend (short) (using an average of two simulations). The model fingerprint has to be substantially reduced in amplitude to match our reconstruction, despite application of a TLS fit to account for noise in the fingerprint (see Stott et al. 2003). As a sensitivity test, we have also performed the same multifingerprint detection analysis with data from the ECHO-G simulations instead of the reconstruction with similar detection results. However, the ECHO-G simulation has substantially enhanced volcanic (scaling factor of  $2.8 \pm 0.8$ ) and probably enhanced solar fingerprints (factor of  $2.7 \pm 1.7$ ) relative to the EBM fingerprints, despite both models having a similar climate sensitivity of  $\sim 2.5^\circ\text{C}$ . This difference is at least partly explained by the use of an older version of the volcanic forcing from Crowley (2000), and a stronger solar forcing.

A detection and attribution analysis of other high-

←

average temperature from an EBM simulation forced with volcanic, solar, and anthropogenic forcing combined (unscaled, gray) and instrumental data (green line), all data are smoothed removing variance below 20 yr. The bottom of (b) compares the estimated contribution from individual forcings (volcanism, solar forcing, and greenhouse gas and aerosol forcing combined; scaling factor, see Table 2) to individual records: Briffa (solid, fat), Esper (dotted), Moberg (dashed), and CH-blend (solid, thin), contributions from different forcings to reconstructions are differentiated by color (e.g., blue for volcanism). Forcing fingerprints in the bottom of (a) and (b) are centered to the period analyzed.

variance reconstructions yields similar results (Table 2), and estimated forced signals are quite consistent between reconstructions (Fig. 5b). Table 2 also shows estimates of the contribution of various forcings to the trend over the first half of the twentieth century. The detection of the anthropogenic fingerprint is robust for all results except Moberg, where significance levels of that signal are dependent on details of the analysis. The Moberg et al. (2005) reconstruction is also the only one that shows a detectable solar signal, although this is somewhat sensitive to details of the analysis. The residual decadal variance of all reconstructions is less than that of the ECHO-G model (Table 2). Thus, our results suggest that model simulated internal climate variability, from ECHO-G, is conservative.

## 6. Conclusions

We have introduced a simple method to reconstruct past climate variability that applies a total least squares approach in the calibration step that overcomes the inherent loss of variance in ordinary least squares calibration methods. This reconstruction method has been successfully validated with climate model data. A reconstruction based on this method is consistent with borehole reconstructions, when compared over the same spatial domain, and shows Little Ice Age cooling and Medieval Warm Period warming that is consistent with that in other recent reconstructions that focus on the recovery of low-frequency variance. Our results demonstrate that long, properly calibrated high-resolution proxies indeed carry low-frequency information and produce reconstructions with considerable variability between centuries.

We find that natural forcing, particularly by volcanism, explains a substantial fraction of decadal variance in our and other high-variance reconstructions. Greenhouse gas forcing is detectable with high significance level in all analyzed reconstructions except Moberg et al. (2005), which ends in 1925, and explains a considerable fraction of the early-twentieth-century warming. In contrast, solar forcing is marginally detectable. In conclusion, our results indicate that enhanced variability in the past does not make it more difficult to detect greenhouse warming, since a large fraction of the variability can be attributed to external forcing. Quantifying the influence of external forcing on proxy records is therefore more relevant to understanding climate variability and its causes than determining if past periods were possibly as warm as the twentieth century.

*Acknowledgments.* The authors wish to thank J. Kenyon for technical support, F. Zwiers for discus-

sion and many helpful suggestions, and three reviewers for their comments. NOAA, NOAA's Office of Global Programs, and the DOE Office of Science, Office of Biological and Environmental Research, and Duke University also provided assistance. Hegerl was additionally supported by NSF Grants ATM-0002206 and ATM-0296007. Zorita was partially funded by the European Research Project SO&P. Pollack and Smerdon were supported by NASA Grant GEWEC-0000-0132, and Smerdon was additionally supported by a Lamont-Doherty Postdoctoral Fellowship.

## APPENDIX A

### Records Used for the New 1500-Yr Reconstruction

The CH-blend reconstruction is composed of records from 12 sites, some of which contain multiple records (Fig. 1 shows their locations). It is termed a blended time series because it includes some data from Esper et al. (2002) that is not available elsewhere. All records obtained from Esper et al. (2002) had already been standardized using regional curve standardization (RCS) methods, which better preserve low-frequency variations in tree-ring-based reconstructions. Processing of the other records varied by source (see below). With the exception of one record (East Asia, which is decadal), all records are annual records that we have decadal smoothed.

- *Western United States:* This time series uses an RCS processed tree-ring composite used in Mann et al. (1999), and kindly provided by M. Hughes, and two sites generated by Lloyd and Graumlich (1997), analyzed by Esper et al. (Boreal and Upper Wright), and provided by E. Cook. The Esper analyses were first averaged. Although there are a number of broad similarities between the Esper and Hughes reconstructions, the correlation is only 0.66. The two composites were averaged.
- *Alberta, Canada:* This time series is also a composite of two different analyses of the 1997 reconstruction of Luckman et al. (1997)—one is unchanged from Luckman's paper, and the other (Athabasca) has been RCS processed by Esper et al. (2002). The correlation between these analyses is unimpressive (0.11); the records were simply averaged. Note that although the correlation with the decadal smoothed 30°–90°N instrument (land) temperatures varies greatly between the two records (0.14 for Athabasca, 0.82 for Jasper), the composite correlation is 0.84.
- *Mackenzie Delta:* The original time series (Szeicz and MacDonald 1995) provided by Esper et al. only had a

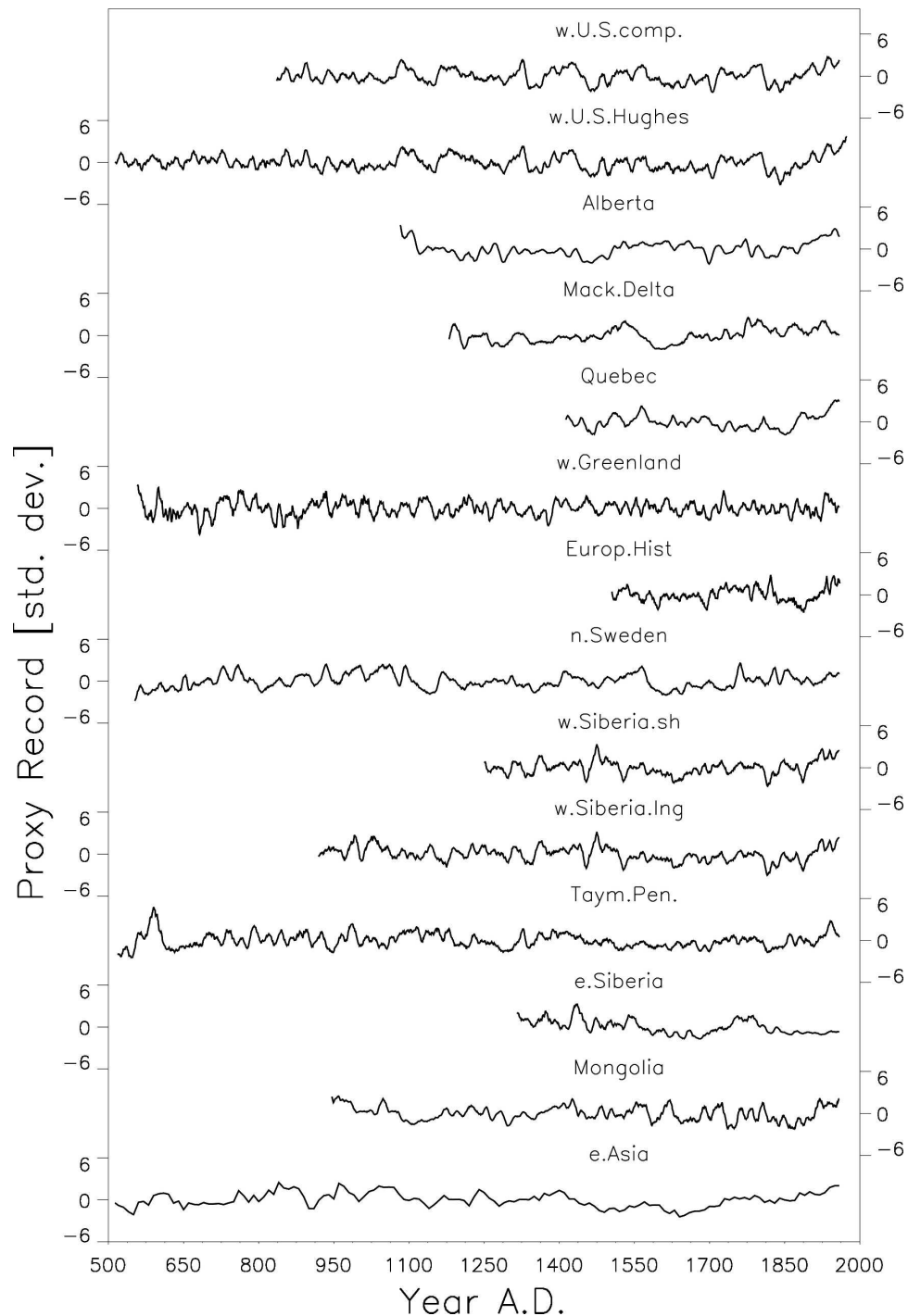


FIG. A1. Decadally smoothed or decadal proxy sites used for the reconstructions in standard deviation units. The records are explained in appendix A.

0.04 correlation with the 1880–1960 decadal average of NH temperature, which yields a very small weight if used for the hemispheric composite. We experimented with various other data from the National Geophysical Data Center (NGDC, more information

available online at <http://www.ncdc.noaa.gov/paleo/ftp-search.html>) to determine if other reconstructions for that area would yield more information for a hemispheric reconstruction. We found that proxy data for that region generally show little correlation

with hemispheric mean temperature. We nevertheless included this site for the sake of completeness and in order to include as many long sites as possible.

- *Quebec*: The situation with Esper's Quebec reconstruction is somewhat similar to what was experienced for the Mackenzie Delta time series; the correlation between their Quebec record and the 30°–90°N average is only 0.25, partly because the time series ends in ~1930. Examination of the NGDC database indicates that the original Esper et al. reconstruction appears to be from the Boniface site. A record from nearby St. Anne also shows many similarities to Boniface ( $r = 0.66$ ), extends closer in time to the present, but is also slightly shorter (the Boniface/St. Anne correlation is 0.70). Although the Boniface/St. Anne composite has a very high correlation with the 30°–90°N (land) record (0.88), inspection of shorter records from Fort Chimo and No Name Lake showed a different twentieth-century response—earlier warming and late cooling. To preclude a Quebec composite from indicating a potentially unrealistic magnitude of late-twentieth-century warmth for the whole region, we created a shorter composite of the four sites that averages records from a Fort Chino and No Name Lake composite after 1806. The new composite still shows significant warming in the twentieth century, but not as extreme as the Boniface/St. Anne sites alone. It is not claimed that this solution represents the best possible way to deal with the conflicting evidence from Quebec; the problem can best be resolved if more long records from other regions of Quebec can be uniformly stacked together without any late-century adjustments.
- *Western Greenland*: This composite is from Fisher et al. (1996).
- *European historical*: This composite was kindly provided by Luterbacher et al. (2004).
- *Northern Sweden*: This is from Grudd et al. (2002) by way of Esper.
- *Western Siberia*: In order to avoid any heavy biases of the mean composite by a number of sites from one region, the western Siberia time series is a composite of three/four time series from this region: two “polar Urals” records east of the Urals–Yamal (Briffa et al. 1995) and Mangazeja (Hantemirov and Shiyato 2002; both by way of Esper et al.) and two records from west of the Urals (Hantemirov and Shiyato 2002). The records from each side of the Urals were first averaged and then combined for the west Siberia short composite; the west Siberia long composite involved Yamal and the west Urals composite. The sites from Esper have been RCS processed.

TABLE A1. Correlations of local proxy data with decadal smoothed hemispheric mean instrumental data.

Site	30°–90°N land	30°–90°N zonal
Western U.S. composite	0.71	0.61
Western U.S. Hughes	0.78	0.65
Alberta	0.82	0.79
Mackenzie	0.04	–0.14
Quebec	0.72	0.87
Western Greenland	0.57	0.57
Europe historical	0.89	0.83
Northern Sweden	0.83	0.89
Western Siberia short	0.72	0.51
Western Siberia	0.74	0.59
Taimyr	0.87	0.87
Eastern Siberia	0.42	0.50
Mongolia	0.78	0.66
East Asia	0.84	0.90

- *Taimyr Peninsula*: This is from Naurzbaev et al. (2002) by way of Esper.
- *Eastern Siberia*: The Esper et al. (2002) composite used the Zhaschivresk time series from Schweingruber. However, this composite only went to 1708. We combined it with a ring width (by Schweingruber, available from NGDC) series from the nearby Ayan-dina River site after removing the obvious growth overprint in the early part of the younger record.
- *Mongolia*: This is from the D'Arrigo et al. (2001) study. However, the full composite illustrated in this paper is not available. We reconstructed the composite from nine records from tree-ring sites sent to the NGDC sites. The early growth part of the tree-ring series from overlapping records was removed without further removal of low-frequency variability.
- *East Asia*: This is the high-resolution record (10-yr average) from Yang et al. (2002).

Individual proxy sites are shown in Fig. A1, correlations between these and NH 30°–90°N land and land–ocean instrumental temperatures are given in Table A1.

## APPENDIX B

### Estimating Uncertainty Ranges of Our Reconstructions

This appendix details the derivation of the uncertainty ranges of our reconstruction. As explained in the body of the paper, we apply a total least squares fit (Allen and Stott 2003), since this method can account for errors in instrumental and proxy data without loss of amplitude of the reconstruction. It is related to fitting the main axis of an ellipse through a cloud of points, after scaling the data so that the ratio of

the variance of errors ( $\varepsilon_{\text{inst}}/\varepsilon_{\text{pal}}$ , see section 2) is 1. The uncertainty  $\varepsilon_{\text{inst}}$  of the instrumental data for land 30°–90°N annual averages due to sampling is estimated from comparing true model averages from a coupled climate model simulation with the ECHO-G model (von Storch et al. 2004) to averages based on instrumental coverage over the area reconstructed and the calibration interval. Estimates of  $\varepsilon_{\text{inst}}$  varied between two simulations and time periods analyzed and range from a standard deviation of 0.034°–0.055°C (0.046°–0.056°C land and ocean combined). This is broadly similar to the error estimated in Jones et al. (1997) of 0.060°C for decadal averages over the whole Northern Hemisphere in the first half of the twentieth century. We chose standard errors  $\varepsilon_{\text{inst}}$  of 0.047°C for land records and 0.056°C for land–ocean combined records (amplitudes and uncertainty ranges vary only by a few percent if  $\varepsilon_{\text{inst}}$  is varied within the range found in subsampled ECHO-G simulations and for instrumental data). Since the true error  $\varepsilon_{\text{pal}}$  is not known, we have varied its variance from 0 to 16 times that of the instrumental error, and calculated the probability of the actual calibration residual between reconstruction and instrumental data based on an F-distribution for each variance ratio. Note that the thus determined most likely ratio of proxy to instrumental data is one that yields a residual that is close to that actual observed. To obtain our best-guess and uncertainty range, the results for every variance ratio have been integrated, weighted by the likelihood of each residual given the assumed noise ratio. In all cases, a variance ratio of 16 yielded very small probabilities, and was also assumed to be highly unlikely, so we have not sampled the tail of the F-distribution beyond a ratio of 16. Overall, this yields a best-guess amplitude  $\alpha$  and error bars that account for the uncertainty in our knowledge of the variance of  $\varepsilon_{\text{pal}}$ .

Using this approach not only yields slightly larger best-guess values of the scaling factor  $\alpha$  than using an OLS approach (see section 2), but more importantly substantially increased 97.5% upper limits for the amplitude (37% higher for land, 65% higher for zonal). This demonstrates that given the true uncertainty in paleoclimatic reconstructions, large past climate variations are much more likely to be retained with this method than with standard regression techniques.

Uncertainty of each decadal mean value of the reconstruction in any given year consists not only of the uncertainty in the scaling of the record, but also of the random sampling error and the nontemperature variability in the proxy records that is not removed by averaging. The latter is unknown, but the F values of the residuals in calibration for each ratio  $\text{var}(\varepsilon_{\text{inst}})/\text{var}(\varepsilon_{\text{pal}})$  discussed above yield the most likely variance ratio and

hence (since we can estimate the variance of  $\varepsilon_{\text{inst}}$ ) also the most likely variance of  $\varepsilon_{\text{pal}}$ . This variance is used to estimate the random sampling uncertainty of the reconstruction for the best-guess scaling factor. The sampling uncertainty for the upper 97.5% range is derived from the ratio of both errors' variance associated with a high-scaling factor, and for the lower 2.5% bar from that for the low-scaling value. Resulting errors are typically about 2–3 times as large for the high-amplitude upper-scaling limit than for the low-amplitude lower limit, and on the order of twice the instrumental error.

The amplitude error and the interdecadal noise are combined by associating the time series with upper- and lower-amplitude limits with the respective errors, and choosing the entire range covered as the reconstruction's full error range. This results in an overestimate of uncertainties, but not dramatically so (e.g., the order of 3% rather than 5% exceedances of 95% uncertainty range in the perfect model study discussed above). This uncertainty range is shown in the figures.

We have scaled only the best-sampled reconstruction CH-blend (short) from 1505 on to instrumental data. Both the long CH-blend (long) and the very long reconstruction CH-blend (Dark Ages) are calibrated to this shorter and better sampled proxy series 1505–1960. This yields a small increase in uncertainty range due to the calibration between long and short series, which is accomplished by an inverse OLS fit [the extreme case of a TLS fit with no additional noise on the target of reconstruction; see Coelho et al. (2004); this is appropriate since only the subsampled longer series exhibits sampling error when calibrated against the shorter, more densely sampled, series]. We calibrate the primary reconstruction [termed CH-blend (short), 1505–1960] to the instrumental period as discussed above. The longer, less densely sampled reconstructions [CH-blend (long), to 946, and CH-blend (Dark Ages), to 558] are then calibrated to CH-blend (short). Due to their high correlation with CH-blend (short) ( $r = 0.97$  and  $r = 0.93$ ), this adds only little additional uncertainty that is accounted for in the uncertainty range for this section of the time series.

## APPENDIX C

### External Radiative Forcing during the Past Millennium

This appendix briefly describes the estimates for external radiative forcing used in our fingerprints for the past millennium, largely referring to Hegerl et al. (2003). Greenhouse gas forcing for the instrumental period is based on the Intergovernmental Panel on Cli-

mate Changed (IPCC) radiative forcing reconstructions for well-mixed greenhouse gases. Ice core evidence suggests small multidecadal–centennial-scale changes in greenhouse gases such as CO<sub>2</sub>, N<sub>2</sub>O, CH<sub>4</sub> over the last millennium (Etheridge et al. 1996, 1998; Fluckiger et al. 1999), in addition to the anthropogenic increase in these greenhouse gases over the industrial period. We applied standard radiative forcing conversion algorithms to convert these changes to radiative forcing.

The solar forcing time series is based on a C<sup>14</sup> residual record of which has been converted to C<sup>14</sup> production rate changes using a box model (Bond et al. 2001). This record was combined with the Lean et al. (1995) solar irradiance reconstruction [note that there is considerable uncertainty in reconstructions of solar forcing, see Lean et al. (2002)].

We also applied an updated version of a global volcano reconstruction (Crowley 2000), which is based on long ice core records of volcanism that extend back to A.D. 1200, four ice cores from Greenland and five from Antarctica and is described in more detail in Hegerl et al. (2003) and a new publication (T. Crowley 2006, unpublished manuscript). The reconstruction includes a latitude profile of forcing that is either based on knowledge of the origin of the eruption or on an estimate based on the representation of the aerosols in ice cores in both hemispheres. The reconstruction of aerosols in ice cores is then calibrated to aerosol optical depth (AOD) and converted to radiative forcing perturbations using the Hansen et al. (2002)  $\Delta F = 21.0$  (AOD) conversion. Because the radiative forcing from extremely large volcanic eruptions results in larger particles that enhance the infrared warming effect (Pinto et al. 1989), we used a 2/3 power scaling between AOD and radiative forcing (RF) for eruptions exceeding 4 W m<sup>-2</sup>. This adjustment has only modest effects except for the very large 1258 eruption not included in the analysis, and the 1453 Kuwae eruption loadings over Antarctica.

#### REFERENCES

- Adcock, R. J., 1878: A problem in least squares. *Analyst*, **5**, 53–54.
- Allen, M. R., and P. A. Stott, 2003: Estimating signal amplitudes in optimal fingerprinting. Part I: Theory. *Climate Dyn.*, **21**, 477–491.
- Beltrami, H., 2002: Climate from borehole data: Energy fluxes and temperatures since 1500. *Geophys. Res. Lett.*, **29**, 2111, doi:10.1029/2002GL015702.
- Bertrand, C., M. F. Loutre, M. Crucifix, and A. Berger, 2002: Climate of the last millennium: A sensitivity study. *Tellus*, **54A**, 221–244.
- Bond, G., and Coauthors, 2001: Persistent solar influence on North Atlantic climate during the Holocene. *Science*, **294**, 2130–2136.
- Briffa, K. R., P. D. Jones, F. H. Schweingruber, S. G. Shiyatov, and E. R. Cook, 1995: Unusual 20th century summer warmth in a 1,000 year temperature record from Siberia. *Nature*, **376**, 156–159.
- , T. J. Osborn, F. H. Schweingruber, I. C. Harris, P. D. Jones, S. G. Shiyatov, and E. A. Vaganov, 2001: Low-frequency temperature variations from a northern tree ring density network. *J. Geophys. Res.*, **106**, 2929–2941.
- Brohan, P., J. J. Kennedy, L. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: A new dataset from 1850. *J. Geophys. Res.*, **111**, D12106, doi:10.1029/2005JD006548.
- Carslaw, H. S., and J. C. Jaeger, 1959: *Conduction of Heat in Solids*. 2d ed. Oxford University Press, 510 pp.
- Coelho, C. A. S., S. Pezzulli, M. Balmaseda, F. J. Doblas-Reyes, and D. B. Stephenson, 2004: Forecast calibration and combination: A simple Bayesian approach for ENSO. *J. Climate*, **17**, 1504–1516.
- Cook, E. R., K. R. Briffa, D. M. Meko, D. A. Graybill, and G. Funkhouser, 1995: The “segment length curse” in long tree-ring chronology development for palaeoclimatic studies. *Holocene*, **5**, 229–237.
- , J. Esper, and R. D’Arrigo, 2004: Extra-tropical Northern Hemisphere land temperature variability over the past 1000 years. *Quat. Sci. Rev.*, **23**, 2063–2074.
- Crowley, T. J., 2000: Causes of climate change over the past 1000 years. *Science*, **289**, 270–277.
- , and T. S. Lowery, 2000: How warm was the medieval warm period? A comment on “Man-made versus natural climate change.” *Ambio*, **29**, 51–54.
- Cubasch, U., R. Voss, G. C. Hegerl, J. Waszkewitz, and T. J. Crowley, 1997: Simulation of the influence of solar radiation variations on the global climate with an ocean-atmosphere general circulation model. *Climate Dyn.*, **13**, 757–767.
- D’Arrigo, R., and Coauthors, 2001: 1738 years of Mongolian temperature variability inferred from a tree ring width chronology of Siberian pine. *Geophys. Res. Lett.*, **28**, 543–546.
- Delworth, T. L., and T. R. Knutson, 2000: Simulation of early 20th century global warming. *Science*, **287**, 2246–2250.
- Esper, J., E. R. Cook, and F. H. Schweingruber, 2002: Low-frequency signals in long tree-ring chronologies for reconstruction of past temperature variability. *Science*, **295**, 2250–2253.
- Etheridge, D. M., L. P. Steele, R. L. Langenfelds, R. J. Fancey, J.-M. Barnola, and V. I. Morgan, 1996: Natural and anthropogenic changes in atmospheric CO<sub>2</sub> over the last 1000 years from air in Antarctic ice and firn. *J. Geophys. Res.*, **101** (D2), 4115–4128.
- , —, R. J. Fancey, and R. L. Langenfelds, 1998: Atmospheric methane between 1000 A.D. and present: Evidence of anthropogenic emissions and climatic variability. *J. Geophys. Res.*, **103**, 15 979–15 993.
- Fisher, D. A., R. M. Koerner, K. Kuivinen, H. B. Clausen, S. Johnsen, J. P. Steffensen, N. Gundestrup, and C. U. Hammer, 1996: Intercomparison of ice core  $\delta^{18}\text{O}$  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. *Climatic Variations and Forcing Mechanisms of the Last 2000 Years*, P. D. Jones et al., Eds., Springer-Verlag, 297–328.
- Fluckiger, J., A. Dallenbach, T. Blunier, B. Stauffer, T. Stocker, D. Raynaud, and J.-M. Barnola, 1999: Variations in atmospheric N<sub>2</sub>O concentration during abrupt climate changes. *Science*, **285**, 227–230.



- Folland, C. K., and Coauthors, 2001: Global temperature change and its uncertainties since 1861. *Geophys. Res. Lett.*, **28**, 2621–2624.
- González-Rouco, F., H. von Storch, and E. Zorita, 2003: Deep soil temperature as proxy for surface air-temperature in a coupled model simulation of the last thousand years. *Geophys. Res. Lett.*, **30**, 2116, doi:10.1029/2003GL018264.
- González-Rouco, J. F., H. Beltrami, E. Zorita, and H. von Storch, 2006: Simulation and inversion of borehole temperature profiles in surrogate climates: Spatial distribution and surface coupling. *Geophys. Res. Lett.*, **33**, L01703, doi:10.1029/2005GL024693.
- Grove, J. M., 1988: *The Little Ice Age*. Methuen, 498 pp.
- Grudd, H., K. R. Briffa, W. Karlen, T. S. Bartholin, P. D. Jones, and B. Kromer, 2002: A 7400 year tree ring chronology in northern Swedish Lapland: Natural climate variability expressed on annual to millennial time scales. *Holocene*, **12**, 657–665.
- Hansen, J., and Coauthors, 2002: Climate forcings in Goddard Institute for Space Studies SI2000 simulations. *J. Geophys. Res.*, **107**, 4347, doi:10.1029/2001JD001143.
- Hantemirov, R. M., and S. G. Shiyato, 2002: A continuous multi-millennial ring width chronology in Yamal, northwestern Siberia. *Holocene*, **12**, 717–726.
- Harris, R. N., and D. S. Chapman, 2001: Mid-latitude (30°–60°N) climatic warming inferred by combining borehole temperatures with surface air temperatures. *Geophys. Res. Lett.*, **28**, 747–750.
- Hegerl, G. C., K. Hasselmann, U. Cubasch, J. F. B. Mitchell, E. Roeckner, R. Voss, and J. Waszkewitz, 1997: Multi-fingerprint detection and attribution of greenhouse-gas and aerosol-forced climate change. *Climate Dyn.*, **13**, 613–634.
- , T. J. Crowley, S. K. Baum, K.-Y. Kim, and W. T. Hyde, 2003: Detection of volcanic, solar, and greenhouse gas signals in paleo-reconstructions of Northern Hemispheric temperature. *Geophys. Res. Lett.*, **30**, 1242, doi:10.1029/2002GL016635.
- , —, W. T. Hyde, and D. Frame, 2006: Climate sensitivity constrained by temperature reconstructions of the last seven centuries. *Nature*, **440**, 1029–1032.
- Huang, S., H. N. Pollack, and P. Y. Shen, 2000: Temperature trends over the past five centuries reconstructed from borehole temperatures. *Nature*, **403**, 756–758.
- International Ad Hoc Detection and Attribution Group, 2005: Detecting and attributing external influences on the climate system: A review of recent advances. *J. Climate*, **18**, 1291–1314.
- Jones, P. D., and M. E. Mann, 2004: Climate over past millennia. *Rev. Geophys.*, **42**, RG2002, doi:10.1029/2003RG000143.
- , T. J. Osborn, and K. R. Briffa, 1997: Estimating sampling errors in large-scale temperature averages. *J. Climate*, **10**, 2548–2568.
- , —, and —, 2001: The evolution of climate over the last millennium. *Science*, **292**, 662–667.
- Lean, J. L., J. Beer, and R. Bradley, 1995: Reconstruction of solar irradiance since 1610: Implications for climate change. *Geophys. Res. Lett.*, **22**, 3195–3198.
- , Y.-M. Wang, and N. R. Sheeley Jr., 2002: The effect of increasing solar activity on the Sun's total and open magnetic flux during multiple cycles: Implications for solar forcing of climate. *Geophys. Res. Lett.*, **29**, 2224, doi:10.1029/2002GL015708.
- Lloyd, A. H., and L. J. Graumlich, 1997: Holocene dynamics of treeline forests in the Sierra Nevada. *Ecology*, **78**, 1199–1210.
- Luckman, B. H., K. R. Briffa, P. D. Jones, and F. H. Schweingruber, 1997: Tree-ring reconstructions of summer temperature at the Columbia Ice Field, Alberta, Canada. *Holocene*, **7**, 375–389.
- Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean, and H. Wanner, 2004: European seasonal and annual temperature variability, trends, and extremes since 1500. *Science*, **303**, 1499–1503.
- Mann, M. E., R. S. Bradley, and M. K. Hughes, 1999: Northern Hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophys. Res. Lett.*, **26**, 759–762.
- , S. Rutherford, E. Wahl, and C. Ammann, 2005: Testing the fidelity of methods used in proxy-based reconstructions of past climate. *J. Climate*, **18**, 4097–4107.
- Mitchell, J. F. B., D. J. Karoly, G. C. Hegerl, F. E. Zwiers, and J. Marengo, 2001: Detection of climate change and attribution of causes. *Climate Change 2001: The Scientific Basis*, J. T. Houghton et al., Eds., Cambridge University Press, 695–738.
- Moberg, A., D. M. Sonechkin, K. Holmgren, N. M. Datsenko, and W. Karlen, 2005: Highly variable Northern Hemisphere temperatures reconstructed from low- and high-resolution proxy data. *Nature*, **433**, 613–617.
- Naurzbaev, M. M., E. A. Vaganov, O. V. Sidorova, and F. H. Schweingruber, 2002: Summer temperatures in eastern Taimyr inferred from a 2427 year late Holocene chronology and earlier floating series. *Holocene*, **12**, 727–736.
- North, G. R., and M. Stevens, 1998: Detecting climate signals in the surface temperature record. *J. Climate*, **11**, 563–577.
- Oerlemans, J., 2005: Extracting a climate signal from 169 glacier records. *Science*, **308**, 675–677.
- Pinto, J. P., R. P. Turco, and O. B. Toon, 1989: Self-limiting physical and chemical effects in volcanic eruption clouds. *J. Geophys. Res.*, **94**, 11 165–11 174.
- Pollack, H. N., and S. Huang, 2000: Climate reconstruction from subsurface temperatures. *Annu. Rev. Earth Planet. Sci.*, **28**, 339–365.
- , and J. E. Smerdon, 2004: Borehole climate reconstructions: Spatial structure and hemispheric averages. *J. Geophys. Res.*, **109**, D11106, doi:10.1029/2003JD004163.
- Robock, A., and M. Free, 1995: Ice cores as an index of global volcanism from 1850 to the present. *J. Geophys. Res.*, **100**, 11 549–11 567.
- Shindell, D. T., G. A. Schmidt, M. E. Mann, and G. Faluvegi, 2004: Dynamic winter climate response to large tropical volcanic eruptions since 1600. *J. Geophys. Res.*, **109**, D05104, doi:10.1029/2003JD004151.
- Stendel, M., I. A. Morgensen, and J. H. Christensen, 2005: Influence of various forcings on global climate in historical times using a coupled atmosphere-ocean general circulation model. *Climate Dyn.*, **26**, 1–15.
- Stott, P. A., G. S. Jones, and J. F. B. Mitchell, 2003: Do models underestimate the solar contribution to recent climate change? *J. Climate*, **16**, 4079–4093.
- Szeicz, J. M., and G. M. MacDonald, 1995: Dendroclimatic reconstruction of summer temperatures in Northwestern Canada based on age dependent modelling. *Quat. Res.*, **44**, 257–266.
- Tett, S. F. B., P. A. Stott, M. R. Allen, W. J. Ingram, and J. F. B. Mitchell, 1999: Causes of twentieth century temperature change near the earth's surface. *Nature*, **399**, 569–572.
- Thompson, D., J. M. Wallace, and G. C. Hegerl, 2000: Annular

- modes in the extratropical circulation. Part II: Trends. *J. Climate*, **13**, 1018–1036.
- von Storch, H., E. Zorita, J. M. Jones, Y. Dimitriev, F. González-Rouco, and S. F. B. Tett, 2004: Reconstructing past climate from noisy data. *Science*, **306**, 679–682.
- , —, —, F. González-Rouco, and S. F. B. Tett, 2006: Response to comment on “Reconstructing past climate from noisy data.” *Science*, **312**, 529c.
- Wahl, E. R., D. M. Ritson, and C. M. Ammann, 2006: Comment on “Reconstructing past climate from noisy data.” *Science*, **312**, 529b.
- Weber, S. L., 2005: A timescale analysis of the NH temperature response to volcanic and solar forcing in the past millennium. *Climate Past Discuss.*, **1**, 137–214.
- Yang, B., A. Braeuning, K. R. Johnson, and S. Yafeng, 2002: General characteristics of temperature variation in China during the last two millennia. *Geophys. Res. Lett.*, **29**, 1324, doi:10.1029/2001GL014485.
- Zorita, E., J. F. González-Rouco, and S. Legutke, 2003: Testing the Mann et al. (1998) approach to paleoclimate reconstructions in the context of a 1000-yr control simulation with the ECHO-G coupled climate model. *J. Climate*, **16**, 1378–1390.
- Zwiers, F. W., and S. S. Shen, 1997: Errors in estimating spherical harmonic coefficients from partially sampled GCM output. *Climate Dyn.*, **13**, 703–716.