Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia

Michael E. Mann*[†], Zhihua Zhang*, Malcolm K. Hughes[‡], Raymond S. Bradley[§], Sonya K. Miller*, Scott Rutherford[¶], and Fenbiao Ni[‡]

*Department of Meteorology and Earth and Environmental Systems Institute, Pennsylvania State University, University Park, PA 16802; [‡]Laboratory of Tree-Ring Research, University of Arizona, Tucson, AZ 85721; [§]Department of Geosciences, University of Massachusetts, Amherst, MA 01003-9298; and [¶]Department of Environmental Science, Roger Williams University, Bristol, RI 02809

Communicated by Lonnie G. Thompson, Ohio State University, Columbus, OH, June 26, 2008 (received for review November 20, 2007)

Following the suggestions of a recent National Research Council report [NRC (National Research Council) (2006) *Surface Temperature Reconstructions for the Last 2,000 Years* (Natl Acad Press, Washington, DC).], we reconstruct surface temperature at hemispheric and global scale for much of the last 2,000 years using a greatly expanded set of proxy data for decadal-to-centennial climate changes, recently updated instrumental data, and complementary methods that have been thoroughly tested and validated with model simulation experiments. Our results extend previous conclusions that recent Northern Hemisphere surface temperature increases are likely anomalous in a long-term context. Recent warmth appears anomalous for at least the past 1,300 years whether or not tree-ring data are used. If tree-ring data are used, the conclusion can be extended to at least the past 1,700 years, but with additional strong caveats. The reconstructed amplitude of change over past centuries is greater than hitherto reported, with somewhat greater Medieval warmth in the Northern Hemisphere, albeit still not reaching recent levels.

climate change | global warming

Knowledge of climate during past centuries can both improve our understanding of natural climate variability and help address the question of whether modern climate change is unprecedented in a long-term context (1, 2). The lack of widespread instrumental climate records before the mid 19th century, however, necessitates the use of natural climate archives or "proxy" data such as tree-rings, corals, and ice cores and historical documentary records to reconstruct climate in past centuries. Many previous proxy data studies have emphasized hemispheric or global mean temperature (3–14), although some studies have also attempted to reconstruct the underlying spatial patterns of past surface temperature changes at global (15, 16) and regional (6, 17, 18) scales.

Most attempts to reconstruct hemispheric temperatures have used some variant on the "composite plus scale" (CPS) methodology (10), in which proxy data (such as tree rings, ice cores, or corals) considered to be sensitive to past surface temperature variations are standardized and centered, potentially weighted, and then composited to form a regional or hemispheric series. The resulting series is then regressed or simply scaled against the target instrumental series (e.g., the Northern Hemisphere mean annual temperature series) to yield a reconstruction of hemispheric or global mean temperature. Variants on this approach have used proxies selected specifically for their retention of low-frequency variability (9, 19) or have included low-resolution (decadal- or centennial-scale) proxies that might be well suited to reconstructing low-frequency climate variability (11) but which are consequently difficult to calibrate robustly against the instrumental record (20). More recently, Hegerl *et al.* (13) use a weighted composite of proxy temperature series, but scaling is accomplished by a so-called "error-in-variables" (EIV) regression method ("total least squares") to allow for errors in both predictors (i.e., proxy composite) and predictand (i.e., the instrumental hemispheric mean temperature series). Lee *et al.* (21) have recently compared a number of variants on the CPS approach.

Distinct from the CPS methods are the so-called climate field reconstruction (CFR) approaches, which instead assimilate proxy records into reconstructions of the underlying spatial climate patterns (15–18, 20, 22–26). Hemispheric or global means, as well as particular indices of interest, can be computed directly from the spatial pattern reconstructions. The CFR method offers the obvious advantage over the CPS method in that the spatial pattern of changes are available, e.g., for comparisons with model-predicted patterns of past climate change (27, 28). Most CFR methods make use of nonlocal "reduced space" relationships between predictors (e.g., sparse early instrumental measurements or longer-term proxy climate data) and predictand (the full spatial field targeted for reconstruction) through the use of large-scale covariance information (15, 17, 18, 29, 30). Such methods are useful in the context of proxy-based climate reconstruction because they take advantage of climate information embedded in remote proxies in the reconstruction of a continuous climate field. CFR approaches depend more heavily on assumptions about the stationarity of relationships between proxy indicators and large-scale climate patterns than do simpler methods (e.g., CPS method) that are based on the assumption of local statistical relationships but make greater use of the available statistical information. The two approaches are therefore complementary.

If one replaces the spatial field in the CFR approach by a single time series (e.g., hemispheric mean surface temperature), then the approach reduces to an EIV variant on multivariate regression. When a regularized ("Reg") version of the expectation-maximization (EM) is used (RegEM, see ref. 31), the procedure has the added benefit of providing an explicit regularization scheme to protect against statistical overfitting. Previous tests with synthetic proxy climate data derived from long model simulations demonstrate that this EIV procedure yields a very similar hemispheric mean reconstruction to that obtained from the application of the associated spatially-explicit CFR procedure when an estimate of only a single time series, e.g., the hemispheric mean temperature, is sought (32).

In this study, we employ a greatly expanded global multiproxy database to develop hemispheric and global mean decadal surface

---

Author contributions: M.E.M., M.K.H., and R.S.B. designed research; Z.Z., S.K.M., S.R., and F.N. performed research; M.E.M. analyzed data; and M.E.M., M.K.H., and R.S.B. wrote the paper.

The authors declare no conflict of interest.

[†]To whom correspondence should be addressed. E-mail: mann@psu.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0805721105/DCSupplemental.

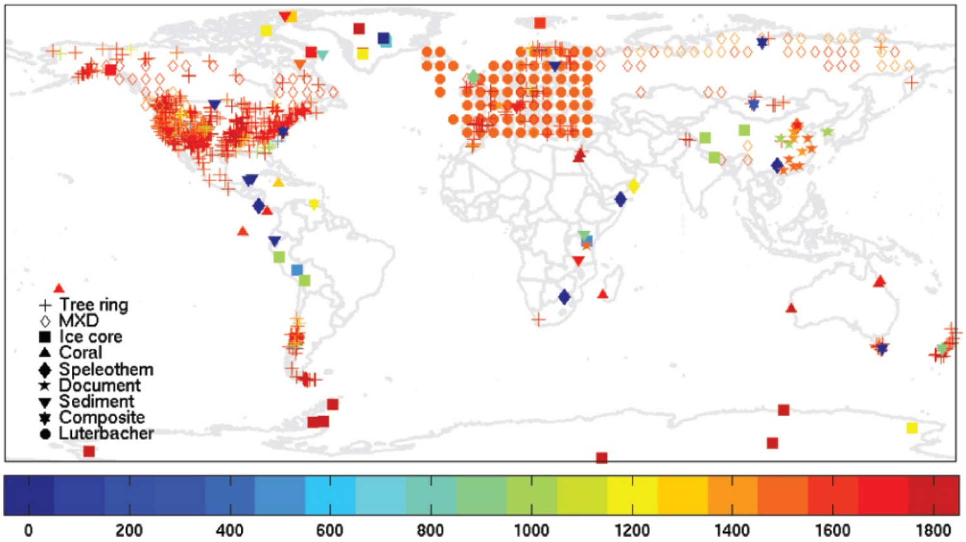

Fig. 1. Spatial distribution of proxy database (see *SI Text* for distribution of screened network and other details). Nine different proxy types are denoted with different symbols as shown in the map. Beginning dates of proxy records are represented by color scale.

temperature reconstructions for past centuries. We compare reconstructions based on the two distinct methods discussed above, the CPS approach and the RegEM-based EIV estimation procedure. Both methods have been tested and validated with long-term forced coupled model simulations (20, 32). For both methods, we perform reconstructions both with and without dendroclimatic proxies to address any potential sensitivity of our conclusions to issues that have been raised with regard to the reliability of tree-ring data on multicentury time scales (4, 11, 16, 19, 33, 34). For the CPS approach, we employ only those proxies estimated to reflect local temperature variations (based on a screening analysis over the calibration interval using instrumental grid box surface temperature data). For the EIV approach, which does not require that proxy data represent local temperature variations, results are compared by using several alternative data-selection schemes, including one that employs all available proxy records, another that employs only proxy records contained within the target hemisphere, and another that employs only the proxy data within that hemisphere that pass the temperature-screening analysis mentioned above. Results for all methods were compared by using both a "frozen" network (7–9, 13) consisting only of proxies available back to at least A.D. 1000 and a "stepwise" approach (15, 16) that makes use of the increasing numbers of proxy data that become available as time progresses. All reconstructions were validated against independent, withheld instrumental surface temperature data, and uncertainties were estimated from validation residuals. We were guided in this work by the suggestions of a recent National Research Council report (35) concerning an expanded dataset, updated data, complementary strategies for analysis, and the use of thoroughly tested statistical methods. All data used in this study and Matlab source codes for implementing the algorithms described, as well as additional supporting information are available at www.meteo.psu.edu/∼mann/supplements/MultiproxyMeans07.

Data

Proxy Data Network. We made use of a multiple proxy ("multiproxy") database consisting of a diverse (1,209) set of annually (1,158) and decadally (51) resolved proxy series [see supporting information (SI) Table S1 and Dataset S1 for further details] including tree-ring, marine sediment, speleothem, lacustrine, ice core, coral, and historical documentary series. All 1,209 series were available back to at least A.D. 1800, 460 extend back to A.D. 1600, 177 back to A.D. 1400, 59 back to A.D. 1000, 36 back to A.D. 500, and 25 back to year "0" (i.e., 1 B.C.). Our proxy database represents a significant extension of the database used in related earlier studies (4, 15, 16). The proxy network covers a broad region of the globe, including the tropics and extratropics, and terrestrial and ocean/maritime regions, with coverage that gradually decreases back in time (Fig. 1). The proxy data also reflect a broad range of seasonal windows of climate sensitivity. Further details about the proxy dataset are provided in *SI Text* and Figs. S1 and S2.

Reconstructions were performed based on both the "full" proxy data network and on a "screened" network (Table S1) consisting of only those proxies that pass a screening process for a local surface-temperature signal. The screening process requires a statistically significant ($P < 0.10$) correlation with local instrumental surface-temperature data during the calibration interval. Where the sign of the correlation could *a priori* be specified (positive for tree-ring data, ice-core oxygen isotopes, lake sediments, and historical documents, and negative for coral oxygen-isotope records), a one-sided significance criterion was used. Otherwise, a two-sided significance criterion was used. Further details of the screening procedure are provided in *SI Text*.

The distribution of the "screened" proxy network for the full interval 1850–1995 is shown in Fig. S1*B*. The rms local annual temperature correlation of the full screened network is $r = 0.39$ ($r = 0.33$) for predictors available back to A.D. 1800 (A.D. 1000). This corresponds to signal-to-noise amplitude ratios (SNRs) of SNR = 0.43 and 0.35, respectively (see ref. 32). Of the 1,209 proxy records in the full dataset, 484 (40%) pass the temperature-screening process over the full (1850–1995) calibration interval (Fig. 1; see also *SI Text* and Table S1).

Instrumental Data. We made use of the University of East Anglia (Norwich, UK) Climatic Research Unit instrumental surface-air temperature data from 1850 to 2006 (36) (www.cru.uea.ac.uk/cru/data/temperature). We used both land-only and combined land and ocean Northern Hemisphere (NH) and Southern Hemisphere (SH) annual mean series as target series for reconstruction. In each case, two different estimates of instrumental hemispheric means were used: the nominal CRUTem3v land and HadCRUT3v land+ocean hemispheric means (CRU and Had, respectively) and hemispheric means calculated from spatially in-filled versions of the grid box data (ICRU and Ihad, respectively). The global mean surface temperature series was defined as the arithmetic mean of the corresponding NH and SH series. A 1961–1990 reference period was used for all series. Further details are provided in *SI Text* and Fig. S3.

Results

Validation Exercises. We evaluated the fidelity of reconstructions through validation experiments (see *Methods*), focusing here on NH

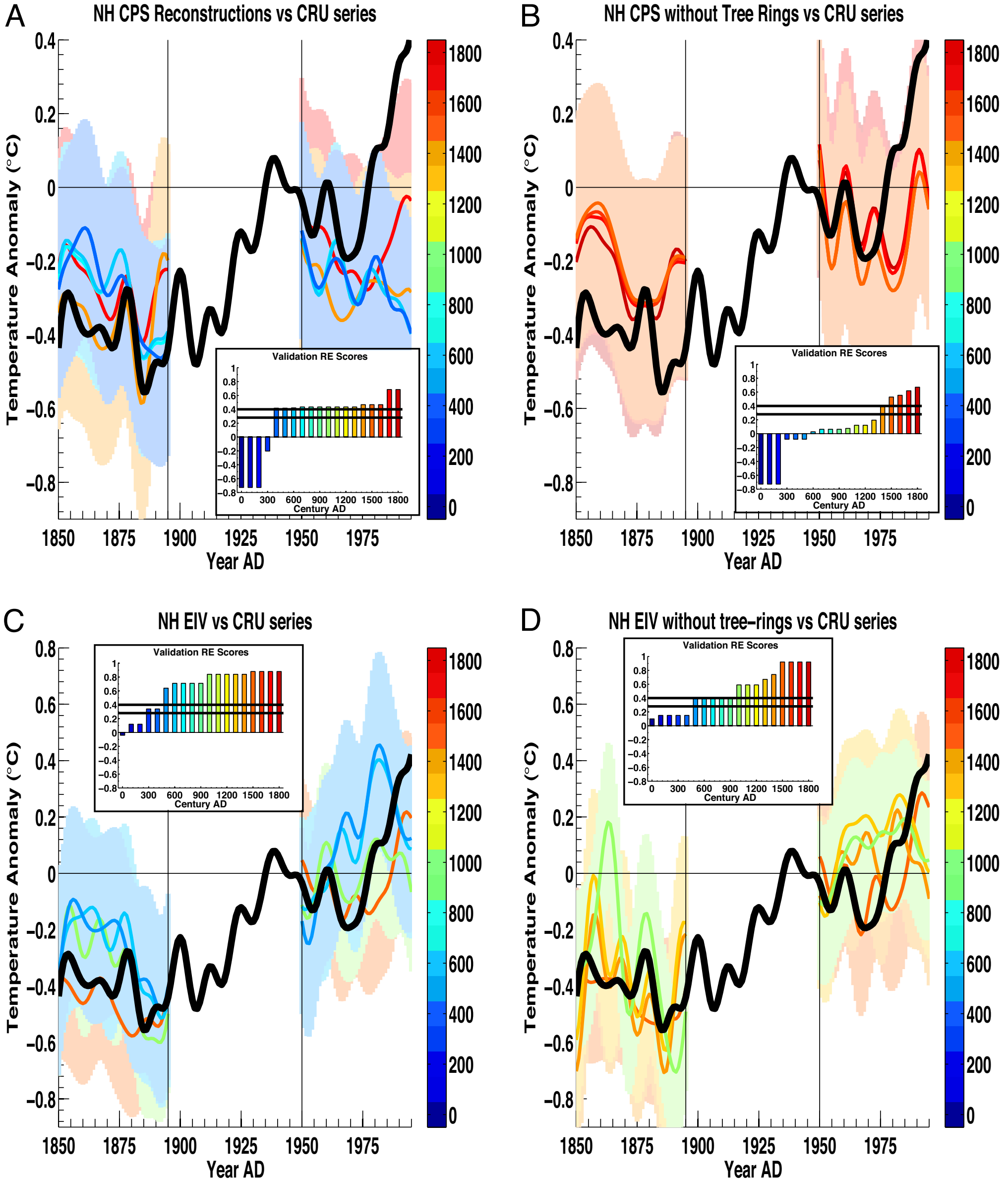

Fig. 2. Comparison of CPS: hemisphere screened (*A*); "no dendro" (*B*); and EIV full global (C) and "no dendro" (*D*). NH reconstructions (colored curves; 95% confidence intervals shown by lightly shaded regions of similar color) with decadally smoothed CRU NH land mean series (thick black curve). Reconstructions are shown over late (A.D. 1950–1995) and early (A.D. 1850–1895) validation intervals corresponding to early (A.D. 1850–1949) and late (1896–1995) calibration experiments respectively. *RE* validation scores are shown (*Insets*) as a function of the starting date of the network along with 90% and 95% significance levels based on Monte Carlo experiments. Only reconstructions that passed validation are shown. Color bar indicates network starting dates.

land temperature reconstructions (Fig. 2; see *SI Text* and Fig. S4 for NH land plus ocean, SH, and global results). The CPS and EIV methods (Dataset S2 and Dataset S3) are both observed to yield reconstructions that, in general, agree with the withheld segment of the instrumental record within estimated uncertainties based on both the early (1850–1949) calibration/late (1950–1995) validation and late (1896–1995) calibration/early (1850–1895) validation. However, in the case of the early calibration/late validation CPS reconstruction with the full screened network (Fig. 2*A*), we observed evidence for a systematic bias in the underestimation of recent warming. This bias increases for earlier centuries where the reconstruction is based on increasingly sparse networks of proxy data. In this case, the observed warming rises above the error bounds of the estimates during the 1980s decade, consistent with the known "divergence problem" (e.g., ref. 37), wherein the temperature sensitivity of some temperature-sensitive tree-ring data appears to have declined in the most recent decades. Interestingly, although the elimination of all tree-ring data from the proxy dataset yields a substantially smaller divergence bias, it does not eliminate the problem altogether (Fig. 2*B*). This latter finding suggests that the divergence problem is not limited purely to tree-ring data, but instead may extend to other proxy records. Interestingly, the problem is greatly diminished (although not absent—particularly in the older networks where a decline is observed after ≈1980) with the EIV method, whether or not tree-ring data are used (Fig. 2 *C* and *D*). We interpret this finding as consistent with the ability of the EIV approach to make use of nonlocal and non-temperature-related proxy information in calibrating large-scale mean temperature changes, thereby avoiding reliance on pure temperature proxies that may exhibit a low-biased sensitivity to recent temperature change.

The skill diagnostics (Fig. 2; see also Dataset S4) for the validation experiments indicate that both the CPS reconstructions (with the screened network) and EIV reconstruction (with the full network) produce skillful NH land reconstructions back to A.D. 400. When tree-ring data are eliminated from the proxy data network, a skillful reconstruction is possible only back to A.D. 1500 by using the CPS approach but is possible considerably further back, to A.D. 1000, by using the EIV approach. We interpret this result as a limitation of the CPS method in requiring local proxy temperature information, which becomes quite sparse in earlier centuries. This situation poses less of a challenge to the EIV approach, which makes use of nonlocal statistical relationships, allowing temperature changes over distant regions to be effectively represented through their covariance with climatic changes recorded by the network.

A skillful EIV reconstruction without tree-ring data is possible even further back, over at least the past 1,300 years, for NH combined land plus ocean temperature (see *SI Text*). This achievement represents a significant development relative to earlier studies with sparser proxy networks (4) where it was not possible to obtain skillful long-term reconstructions without tree-ring data.

There are additional caveats for results based partly on tree rings (see *SI Text* for details). The tree-ring series used before A.D. 1200 typically have median segment lengths between 500 and 700 years, limiting the reliability of information on millennial time scales. Such limitations could compromise the long-term trends in the reconstructions using these series. Interestingly, however, the long-term (i.e., pre-A.D. 1000) behavior is largely unaffected by the removal of tree-ring data, even though skillfulness in the reconstructions can no longer then be inferred for earlier centuries: the long-term CPS reconstruction is essentially unaffected back to A.D. 200, whereas the EIV reconstruction indicates only modestly different (slightly higher) temperatures before A.D. 600 (Fig. S7). Although this observation may seem paradoxical, it is the greater spatial coverage, rather than any additional low-frequency information, that leads to the greater apparent fidelity in earlier centuries of reconstructions employing tree-ring data. However, this increased fidelity is established primarily by relationships at decadal time scales, highlighting the conundrum (38) that validations with modern instrumental data are, by design, weighted toward evaluating the fidelity of high-frequency information. Although loss of low-frequency information from tree-ring data thus does not appear to impact any of the above conclusions, skillful reconstructions are not possible for the earlier centuries without the spatial coverage provided by these data.

NH Temperature Reconstructions. Although the details of the reconstructions produced for a given method and target (e.g., CPS NH land) showed some sensitivity to which proxy data sets are used, and precisely which instrumental series is used (e.g., CRU vs. ICRU), all of the individual reconstructions that pass validation fall within the uncertainties of the composite reconstruction, defined as the average of all individual reconstructions that pass validation. In other words, the various reconstructions are consistent within uncertainties. This also holds true for reconstructions resulting from the early and late calibration intervals used in validation experiments (see *SI Text* and Figs. S7–S11).

Nominally, the recent observed decadal warmth recorded in the instrumental observations exceeds the uncertainty range of the reconstructions over at least the past 1,600 years for NH land temperatures as reconstructed by CPS (Fig. S5) and the past 1,700 years for NH land plus ocean temperatures as reconstructed by EIV (Fig. S6). Because this conclusion extends to the past 1,300 years for EIV reconstructions withholding all tree-ring data, and because non-tree-ring proxy records are generally treated in the literature as being free of limitations in recording millennial scale variability (11), the conclusion that recent[‖] NH warmth likely[**] exceeds that of at least the past 1,300 years thus appears reasonably robust. For the CPS (EIV) reconstructions, the instrumental warmth breaches the upper 95% confidence limits of the reconstructions beginning with the decade centered at 1997 (2001). It is intriguing to note that the removal of tree-ring data from the proxy dataset yields less, rather than greater, peak cooling during the 16th–19th centuries for both CPS and EIV methods (see Figs. S5*a* and S6*b*, respectively), contradicting the claim (33) that tree-ring data are prone to yielding a warm-biased "Little Ice Age" relative to reconstructions using other high-resolution climate proxy indicators.

We compared the composite CPS and EIV NH mean reconstructions (Fig. 3; see also Figs. S15 and S16) with both each other and with other previously published NH reconstructions based on various combinations of proxy data and differing statistical approaches (4, 6–9, 11, 12, 19, 39–41). This comparison demonstrates rather striking agreement between alternative estimates over the past four centuries, all of which fall well within the estimated uncertainty range, with one exception: the borehole estimate from Huang *et al.*, (39) which is significantly cooler through A.D. 1800. Back to A.D. 1000, the reconstructions all agree within uncertainties with three exceptions: the Moberg *et al.* (11) reconstruction is cooler over the interval A.D. 1500–1600, the Esper *et al.* (19) reconstruction is cooler over A.D. 1050–1300, and the Huang *et al.* (39) estimate is significantly cooler over the period A.D. 1500–1800.

Peak Medieval warmth (from roughly A.D. 950-1100) is more pronounced in the EIV reconstructions (particularly for the land-only reconstruction) than in the CPS reconstructions (Fig. 3). The EIV land-only reconstruction, in fact, indicates markedly more sustained periods of warmer NH land temperatures from A.D. 700 to the mid-fifteenth century than previous published reconstructions. Peak multidecadal warmth centered at A.D. 960 (representing average conditions over A.D. 940–980) in this case corresponds approximately to 1980 levels (representing average conditions over 1960–2000). However, as noted earlier, the most recent decadal warmth exceeds the peak reconstructed decadal warmth, taking into account the uncertainties in the reconstructions.

Although the EIV and CPS reconstructions essentially agree within uncertainties back to A.D. 1000 (Fig. 3*B*), there are nonetheless some significant systematic differences in the implied long-term temperature histories over that time frame. The EIV reconstructions suggest that temperatures were relatively warm (comparable with the mean over the 1961–1990 reference period but below the levels of the past decade) from A.D. 1000 through the early 15th century, then fell abruptly. By contrast, the CPS reconstructions indicate more uniformly colder conditions, with peak Medieval warmth that does not breach the mean warmth of modern reference period (1961–1990), and a long-term, more steady decline in temperatures before 20th century warming. We could arguably take these two reconstructions as end members that bracket the possible range for peak NH mean Medieval warmth, lying somewhere between 0.4°C colder and and 0.4°C warmer than the modern reference period (1961–1990) mean, but still exceeded by the most recent decadal warming.

Before A.D. 1000, there is somewhat less agreement between the various reconstructions. In particular, reconstructions based on variants of the CPS method tend to be significantly cooler than (and outside the uncertainties of) the EIV reconstruction. Investigating the sources of these differences, we first established that the removal of the seven proxy series in our database identified *a priori* as having potentially spurious features, has no significant impact on

---

[‖] Here and elsewhere in the paper, "recent" is taken to correspond to the past decade unless noted otherwise.

[**] For the purpose of this paper, we adopt the definition of "likely" used by the Intergovernmental Panel on Climate Change (IPCC): i.e., that the probability of the assertion being true is estimated as between 66% and 90%.

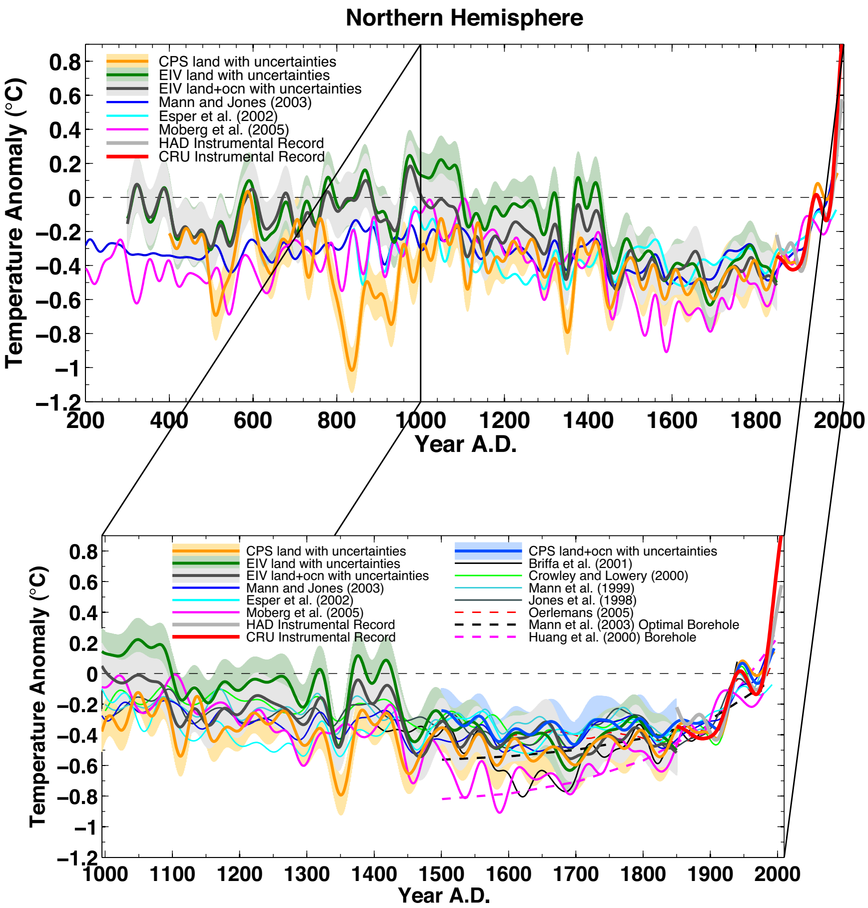

Fig. 3. Composite CPS and EIV NH land and land plus ocean temperature reconstructions and estimated 95% confidence intervals. Shown for comparison are published NH reconstructions, centered to have the same mean as the overlapping segment of the CRU instrumental NH land surface temperature record 1850–2006 that, with the exception of the borehole-based reconstructions, have been scaled to have the same decadal variance as the CRU series during the overlap interval (alternative scaling approaches for attempting to match the amplitude of signal in the reconstructed and instrumental series are examined in *SI Text*). All series have been smoothed with a 40-year low-pass filter as in ref 33. Confidence intervals have been reduced to account for smoothing.

the CPS (or EIV) reconstructions (*SI Text* and Fig. S8). However, we observed that the pronounced cooling between approximately A.D. 750 and A.D. 1000 in the current CPS reconstruction is based on prominent excursions in a relatively small number (see Fig. S9) of the 15 NH proxy series available in the screened network back through the 9th century and that the amplitude of the cooling is somewhat sensitive to the removal of individual proxy records (see Fig. S10). Analysis of synthetic "pseudoproxy" proxy networks (*SI Text*, Figs. S12–S14, and Tables S2–S4) indicates that such apparent pronounced hemispheric temperature anomalies in the reconstructions can arise as purely spurious features with the CPS approach, given such sparse networks, an artifact of the statistics of averaging a small number of noisy local temperature estimates. By contrast, we find in these experiments that the EIV reconstructions are significantly more skillful, given a particular synthetic data network. Where the two methods no longer yield reconstructions that agree within uncertainties, it is therefore likely that the EIV reconstruction is the more reliable, although with the caveat that this finding has been demonstrated only under the assumptions implicit in the pseudoproxy analyses (e.g., that proxies have a linear, if noisy, relationship with local temperature variations). For this reason, we place greatest confidence in the EIV reconstructions, particularly back to A.D. 700, when a skillful reconstruction as noted earlier is possible without using tree-ring data at all.

SH and Global Temperature Reconstructions. Conclusions for SH mean temperatures are somewhat weaker (Figs. S5 and S6), plausibly due to the relative paucity of proxy data in the SH (Fig. 1). Nominally, recent warmth appears anomalous in the context of the past 1,500 years from the CPS reconstructions, but skillful CPS reconstructions are not possible without tree-ring data before A.D. 1700, implying additional caveats as discussed above. Recent warmth exceeds that reconstructed for at least the past 1,800 years in the EIV reconstructions, and this conclusion extends back at least 1,500 years without using tree-ring data. However, the estimated uncertainties are compatible with the possibility that recent SH warmth might have been breached during brief periods in the past. Similarly, for global mean temperature, the CPS reconstruction suggests that recent warmth is anomalous for at least the past 1,500 years, but with the caveat that tree-ring data are required for a skillful long-term reconstruction. The EIV reconstruction indicates recent warmth that exceeds the reconstructed warmth (past 1,500 years with caveats related to the use of tree-ring data, and the past 1,300 years if tree-ring data are excluded), but like the SH, the uncertainties are compatible with the possibility of brief periods of similar warmth over the past 1,500 years. More confident statements about long-term temperature variations in the SH and globe on the whole must await additional proxy data collection.

Conclusions

We find that the hemispheric-scale warmth of the past decade for the NH is likely anomalous in the context of not just the past 1,000 years, as suggested in previous work, but longer. This conclusion appears to hold for at least the past 1,300 years (consistent with the recent assessment by ref. 2) from reconstructions that do not use tree-ring proxies, and are therefore not subject to the associated additional caveats. This conclusion can be extended back to at least the past 1,700 years if tree-ring data are used, but with the additional strong caveats noted. When differences in scaling between previous studies are accounted for, the various current and previous estimates of NH mean surface temperature are largely consistent within uncertainties, despite the differences in methodology and mix of proxy data back to approximately A.D. 1000. The reconstructions appear increasingly more sensitive to method and data quality and quantity before A.D. 1600 and, particularly, before approximately A.D. 1000. Conclusions are less definitive for the SH and globe, which we attribute to larger uncertainties arising from the sparser available proxy data in the SH. Given the uncertainties, the SH and global reconstructions are compatible with the possibility of warmth similar to the most recent decade during brief intervals of the past 1,500 years. A targeted effort to recover additional high-quality, long paleoclimate proxy records from the SH could reduce these current existing uncertainties. Similarly, reducing uncertainties for the period before A.D. 1000 for the NH will require additional proxy records of sufficient length that preserve climate signal on the millennial time scale.

Methods

Detailed descriptions of the CPS and EIV procedures are provided in *SI Text*. Reconstructions were based on calibration over the full 146-year interval 1850–1995. Statistical validation was achieved by using a split calibration/verification procedure wherein data were calibrated alternatively on both the most recent (1896–1995) and oldest (1850–1949) 100-year subintervals, whereas the remaining 46 years were used to validate the reconstruction. Results from the early and late validation experiments were then averaged for the purpose of estimating skill metrics and uncertainties. So-called ''reduction of error'' (*RE*) and ''coefficient of efficiency'' (*CE*) skill scores for the decadal reconstructions were used as metrics of validation skill as in past work (20, 32). Because of its established deficiencies as a diagnostic of reconstruction skill (32, 42), the squared correlation coefficient r^2 was not used for skill evaluation. Statistical significance of *RE* and *CE* scores were estimated by Monte Carlo simulations based on the null hypothesis of first-order autoregressive ''red noise'' (20, 32). Only those reconstructions that passed validation at the $P = 0.05$ (i.e., 95% significance) level based on both metrics were retained. Uncertainties were estimated from the residual decadal variance during the validation period based (32, 42) on the average validation period r^2 (which in this context has the useful property that, unlike *RE* and *CE*, it is bounded by 0 and 1 and can therefore be used to define a ''fraction'' of unresolved variance). Although use of validation residuals for uncertainty estimation avoids the potential overfit bias associated with use of calibration residuals, an important caveat nonetheless remains that such statistical uncertainty estimates do not account for the possible degradation in the proxy record before the validation interval, which is potentially an issue in earlier centuries in particular.

ACKNOWLEDGMENTS. We are indebted to G. North and G. Hegerl for their valuable insight, suggestions, and comments and to L. Thompson for presiding over the review process for this paper. M.E.M. and Z.Z. gratefully acknowledge support from the Atmospheric Sciences program of the National Science Foundation (Grant ATM-0542356). R.S.B. acknowledges support from the Office of Science (Biological and Environmental Research), U.S. Department of Energy, Grant No. DE-FG02-98ER62604. M.K.H. and F.N. were supported by National Oceanic and Atmospheric Administration Grant NA16GP2914 from Climate Change Data and Detection.

1. Folland CK, *et al.* (2001) Observed climate variability and change. *Climate Change 2001: The Scientific Basis*, eds Houghton JT, *et al.* (Cambridge Univ Press, Cambridge, UK), pp 99–181.
2. Jansen EJ, *et al.* (2007) Palaeoclimate. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, eds Solomon S, *et al.* (Cambridge Univ Press, Cambridge, UK).
3. Bradley RS, Jones PD (1993) ''Little Ice Age'' summer temperature variations: Their nature and relevance to recent global warming trends. *Holocene* 3:367–376.
4. Mann ME, Bradley RS, Hughes MK (1999) Northern Hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophys Res Lett* 26:759–762.
5. Briffa KR, Jones PD, Schweinbruger FH, Osborn TJ (1998) Influence of volcanic eruptions on Northern Hemisphere summer temperatures over the past 600 years. *Nature* 393:450–454.
6. Briffa KR, *et al.* (2001) Low-frequency temperature variations from a northern tree-ring density network. *J Geophys Res* 106:2929–2941.
7. Jones PD, Briffa KR, Barnett TP, Tett SFB (1998) High-resolution palaeoclimatic records for the last millennium: Interpretation, integration and comparison with general circulation model control-run temperatures. *Holocene* 8:455–471.
8. Crowley TJ, Lowery TS (2000) How warm was the Medieval Warm Period? A comment on 'Man-made versus natural climate change'. *Ambio* 39:51–54.
9. Mann ME, Jones PD (2003) Global surface temperature over the past two millennia. *Geophy Res Lett* 30:1820.
10. Jones PD, Mann ME (2004) Climate over past millennia. *Rev Geophys* 42:RG2002.
11. Moberg A, Sonechkin DM, Holmgren K, Datsenko NM, Karlen W (2005) Highly variable Northern Hemisphere temperatures reconstructed from low- and high-resolution proxy data. *Nature* 433:613–617.
12. Oerlemans J (2005) Extracting a climate signal from 169 glacier records. *Science* 308:675–677.
13. Hegerl GC, Crowley TJ, Hyde WT, Frame DJ (2006) Climate sensitivity constrained by temperature reconstructions over the past seven centuries. *Nature* 440:1029–1032.
14. Juckes MN, *et al.* (2007) Millennial temperature reconstruction intercomparison and evaluation. *Clim Past* 3:591–609.
15. Mann ME, Bradley RS, Hughes MK (1998) Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* 392:779–787.
16. Rutherford S, *et al.* (2005) Proxy-based Northern Hemisphere surface temperature reconstructions: Sensitivity to methodology, predictor network, target season and target domain. *J Climate* 18:2308–2329.
17. Evans MN, Kaplan A, Cane MA (2002) Pacific sea surface temperature field reconstruction from coral $\delta 18O$ data using reduced space objective analysis. *Paleoceanography* 17:1007.
18. Luterbacher J, *et al.* (2002) Reconstruction of sea level pressure fields over the Eastern North Atlantic and Europe back to 1500. *Clim Dyn* 18:545–561.
19. Esper J, Cook ER, Schweingruber FH (2002) Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. *Science* 295:2250–2253.
20. Mann ME, Rutherford S, Wahl E, Ammann C (2005) Testing the fidelity of methods used in proxy-based reconstructions of past climate. *J Climate* 18:4097–4107.
21. Lee TCK, Zwiers FW, Tsao M (2008) Evaluation of proxy-based millennial reconstruction methods. *Clim Dyn* 31:263–281.
22. Fritts HC (1991) *Reconstructing Large-Scale Climatic Patterns from Tree-Ring Data* (Univ of Arizona Press, Tucson, AZ).
23. Mann ME, Rutherford S (2002) Climate reconstruction using 'Pseudoproxies'. *Geophys Res Lett* 29:1501.
24. Pauling A, Luterbacher J, Wanner H (2003) Evaluation of proxies for European and North Atlantic temperature field reconstructions. *Geophys Res Lett* 30:1787.
25. Luterbacher J, Dietrich D, Xoplaki E, Grosjean M, Wanner H (2004) European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* 303:1499–1503.
26. Zhang Z, Mann ME, Cook ER (2004) Alternative methods of proxy-based climate field reconstruction: Application to summer drought over the conterminous United States back to A.D. 1700 from tree-ring data. *Holocene* 14:502–516.
27. Delworth TL, Mann ME (2000) Observed and simulated multidecadal variability in the North Atlantic. *Clim Dyn* 16:661–676.
28. Shindell DT, Schmidt GA, Mann ME, Rind D, Waple A (2001) Solar forcing of regional climate change during the Maunder Minimum. *Science* 294:2149–2152.
29. Smith TM, Reynolds RW, Livezey RE, Stokes DC (1996) Reconstruction of historical sea surface temperatures using empirical orthogonal functions. *J Clim* 9:1403–1420.
30. Kaplan A, Kushnir Y, Cane MA, Blumenthal MB (1997) Reduced space optimal analysis for historical data sets: 136 years of Atlantic sea surface temperatures. *J Geophys Res* 102C:27835–27860.
31. Schneider T (2001) Analysis of incomplete climate data: estimation of mean values and covariance matrices and imputation of missing values. *J Clim* 14:853–871.
32. Mann ME, Rutherford S, Wahl E, Ammann C (2007) Robustness of proxy-based climate field reconstruction methods. *J Geophys Res* 112:D12109.
33. Broecker WS (2001) Was the Medieval Warm Period global? *Science* 291:1497–1499.
34. Mann ME, Hughes MK (2002) Tree-ring chronologies and climate variability. *Science* 296:848.
35. NRC (National Research Council) (2006) *Surface Temperature Reconstructions for the Last 2,000 Years* (Natl Acad Press, Washington, DC).
36. Brohan P, Kennedy JJ, Haris I, Tett SFB, Jones PD (2006) Uncertainty estimates in regional and global observed temperature changes: A new dataset from 1850. *J Geophys Res* 111:D12106.
37. D'Arrigo RD, Wilson R, Jacoby G (2006) On the long-term context for 20th century warming. *J Geophys Res* 111: D03103.
38. Esper J, Frank DC, Wilson RJS (2004) Climate reconstructions —Low frequency ambition and high frequency ratification. *Eos* 85:113–120.
39. Huang S, Pollack HN, Shen P-Y (2000) Temperature trends over the past five centuries reconstructed from borehole temperature. *Nature* 403:756–758.
40. Mann ME, Rutherford S, Bradley RS, Hughes MK, Keimig FT (2003) Optimal surface temperature reconstructions using terrestrial borehole data. *J Geophys Res* 108:4203.
41. Rutherford S, Mann ME (2004) Correction to ''Optimal surface temperature reconstructions using terrestrial borehole data.'' *J Geophys Res* 109:D11107.
42. Wahl ER, Ammann CM (2007) Robustness of the Mann, Bradley, Hughes reconstruction of surface temperatures: Examination of criticisms based on the nature and processing of proxy climate evidence. *Clim Change* 85:33–69.