University of Massachusetts Amherst

From the SelectedWorks of Raymond S Bradley

April, 2003

Optimal Surface Temperature Reconstructions Using Terrestrial Borehole Data

M. E Mann
S. Rutherford
Raymond S Bradley, *University of Massachusetts - Amherst*
M. K Hughes
F. T Keimig

Available at: https://works.bepress.com/raymond_bradley/22/

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 108, NO. D7, 4203, doi:10.1029/2002JD002532, 2003

Correction published 5 June 2004

Optimal surface temperature reconstructions using terrestrial borehole data

Michael E. Mann,[1] Scott Rutherford,[1] Raymond S. Bradley,[2] Malcolm K. Hughes,[3] and Frank T. Keimig[2]

Received 14 May 2002; revised 4 November 2002; accepted 14 January 2003; published 3 April 2003.

[1] We derive an optimal Northern Hemisphere mean surface temperature reconstruction from terrestrial borehole temperature profiles spanning the past five centuries. The pattern of borehole ground surface temperature (GST) reconstructions displays prominent discrepancies with instrumental surface air temperature (SAT) estimates during the 20th century, suggesting the presence of a considerable amount of noise and/or bias in any underlying spatial SAT signal. The vast majority of variance in the borehole dataset is efficiently retained by its two leading eigenvectors. A sizable share of the variance in the first eigenvector appears to be associated with non-SAT related bias in the borehole data. A weak but detectable SAT signal appears to be described by a combination of the first two eigenvectors. Exploiting this eigendecomposition, application of optimal signal estimation methods yields a hemispheric borehole SAT reconstruction that is largely consistent with instrumental data available in past centuries, and is indistinguishable in it smajor features from several published long-term temperature estimates based on both climate proxy data and model simulations. *INDEX TERMS:* 1620 Global Change: Climate dynamics (3309); 3307 Meteorology and Atmospheric Dynamics: Boundary layer processes; 3344 Meteorology and Atmospheric Dynamics: Paleoclimatology; *KEYWORDS:* Boreholes, Little Ice Age, surface temperature, signal detection, global warming

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

1. Introduction

[2] Subsurface terrestrial borehole temperature profiles can be used to obtain an estimate of ground surface temperature (GST) changes back in time [*Harris and Chapman*, 1997; *Pollack et al.*, 1998; *Huang et al.*, 2000; *Harris and Chapman*, 2001]. An advantage claimed for these estimates over those from traditional proxy climate data (e.g., tree rings, corals, ice cores, and historical documentary records), is that they do not require calibration against independent surface temperature data. Previous estimates of Northern Hemisphere mean GST trends from boreholes indicate a net temperature change of 0.8°C to 1.0°C [*Pollack et al.*, 1998; *Huang et al.*, 2000; *Harris and Chapman*, 2001] from AD 1500 to present. If such trends in GST are, as is argued in these studies, representative of past trends in surface air temperatures (SATs), then they are in conflict with proxy-based estimates of hemispheric SAT changes [*Bradley and Jones*, 1993; *Jones et al.*, 1998; *Mann et al.*, 1998; *Mann et al.*, 1999; *Mann*, 2000; *Crowley and Lowery*, 2000; *Briffa et al.*, 2001], which typically estimate a net warming of about 0.5°C. The proxy reconstructions appear consistent with model-based, theoretical estimates of Northern Hemisphere mean SAT trends over the past millennium [*Free and Robock*, 1999; *Crowley*, 2000; *Gerber et al.*, 2003]. The model estimates, however, depend on the assumed radiative forcing history and the climate sensitivity to this forcing, both of which are uncertain. It has been argued [*Huang et al.*, 2000] that the borehole reconstructions, because they indicate larger temperature trends over the past five centuries, imply a higher climate sensitivity than the other proxy-based reconstructions (though if the assumed forcing history is reliable, it is difficult to rectify this history with the substantial trend in the borehole estimates well before the 20th century when anthropogenic forcing was negligible, and only modest, natural radiative forcing variations were likely to play any role [see, e.g., *Crowley*, 2000]). It is thus important to resolve apparent discrepancies between competing estimates of past temperature change both to resolve uncertainties in the past climate trends, and to better constrain the roles of natural and anthropogenic climate factors in governing past changes.

[3] Contributions to the borehole thermal profile from factors unrelated to GST changes, such as subsurface fluid flow, vertical and lateral inhomogeneities in bedrock proper-

[1]Department of Environmental Sciences, University of Virginia, Charlottesville, Virginia, USA.
[2]Department of Geosciences, University of Massachusetts, Amherst, Massachusetts, USA.
[3]Laboratory of Tree-Ring Research, University of Arizona, Tucson, Arizona, USA.

Copyright 2003 by the American Geophysical Union.
0148-0227/03/2002JD002532

ties, and variable topography can lead to potentially large errors in individual borehole GST estimates. If such errors are random, they can be reduced (though not eliminated) by averaging many borehole profiles within a particular region. However, other systematic errors or biases that are involved in the attempt to infer past large-scale SAT changes from borehole GST histories cannot be eliminated by simple averaging.

[4] The overwhelming majority of Northern Hemisphere borehole data come from regions that experience seasonal snow cover. The snow cover partially insulates the ground from cold-season air temperature and fluctuations therein, providing a potential insensitivity of the underlying ground temperature to cold winter air mass outbreaks (and implying a warm-season bias in borehole GST estimates, the degree of which depends on extent and duration of winter snow cover). Little, if any imprint, of the cooling associated with cold air outbreaks is recorded by a ground surface buried beneath a sufficiently thick seasonal snow cover layer. The accumulated influence of such outbreaks on winter mean SAT is considerably greater than the quite modest (on the order of a degree C or less) SAT trends sought from borehole reconstructions. In regions where midwinter snow cover has increased over the past few centuries (which could potentially be associated with either warmer or colder winters, depending on the details of air mass influence in the region), borehole GSTs may therefore exhibit a spurious apparent long-term warming (i.e., colder conditions back in time) due to an increasing incidence of insulating winter snow cover in more recent centuries. Such trends in snow cover are, in fact, consistent with inferred climate anomalies during the height of the "Little Ice Age." Recent work [*Keigwin and Pickart*, 1999; *Shindell et al.*, 2001], suggests that this period was associated with a tendency for the negative phase of the North Atlantic Oscillation (NAO)/ Arctic Oscillation (AO) pattern, implying both colder and dryer conditions over most of the continental interior of North America and Eurasia [e.g., *Thompson and Wallace*, 2001]. Though the actual past trends in winter snow cover are not known, such indirect inferences are nonetheless suggestive of the possibility of spurious apparent cooling recorded in borehole GST estimates in earlier centuries, when the ground surface may have been subjected to increased exposure to the winter atmosphere. Such influences may indeed largely explain the anomalously cold conditions suggested by borehole temperature reconstructions for the 16th–19th centuries relative to other estimates. General Circulation Model (GCM) experiments are in progress to estimate the potential impacts of changes in seasonal snow cover variability in explaining differential trends in SAT and ground temperature.

[5] Other sources of bias could potentially also be significant. Spring ground warming may be damped and delayed by melting and evaporation of accumulated winter snow cover, and melting of permafrost can lead to complications from related latent heat considerations. Anthropogenic land use changes may further contaminate the surface temperature signal contained in ground temperatures, particularly during the 20th century [*Skinner and Majorowicz*, 1999]. Additional uncertainty arises in the process of estimating the component of the borehole temperature profile associated with past GST changes, which involves estimating the steady state "background" geothermal flux, which must be removed from the borehole temperature profile to yield a "transient" profile indicative of past GST changes [*Harris and Chapman*, 1997; *Pollack et al.*, 1998; *Huang et al.*, 2000; *Harris and Chapman*, 2001]. An incorrect separation of the background and transient components could lead to a systematic bias in the estimated rates of ground temperature change. All of these influences will compromise the relationship between borehole-based GST estimates and the large-scale trends in annual SAT that are of interest.

[6] *Huang et al.* [2000] (HPS00) use a Bayesian inverse method to estimate centennial rates of ground surface temperature change from borehole profiles, integrating these back in time to obtain century-long trends in ground surface temperature trends for the past five centuries at each of 616 borehole sites over the globe. Averaging the 453 individual borehole site temperature reconstructions available over the Northern Hemisphere, HPS00 obtain a reconstruction of Northern Hemisphere mean ground temperature back to AD 1500. Since they reference the mean of their 20th century reconstruction to that of the 20th century instrumental record, any errors in their estimated rates, unless they fortuitously cancel, will lead to an increasing error in temperature estimates back in time. HPS00 argue for a favorable agreement between their mean 20th century trend and the 20th century trend in the (combined land air and sea surface) Northern Hemisphere mean annual surface temperature series (the series represents a hemispheric average of terrestrial SAT data and marine Sea Surface Temperature (SST) data [*Jones et al.*, 1999]). However, important discrepancies are evident upon comparison of their reconstruction with early instrumental SAT data, as well as other long-term proxy-based surface temperature reconstructions, as discussed below.

[7] We show that such apparent discrepancies are an artifact of the contamination of past borehole-based hemispheric temperature estimates by systematic biases such as discussed above. We describe the instrumental and borehole data used in our analysis in section 2. In section 3, we describe the methodology by which the optimal surface temperature signal in the borehole data is identified and reconstructed. We discuss and interpret the results of this analysis in section 4, and present our conclusions in section 5.

2. Data

[8] We use the same 453 Northern Hemisphere borehole temperature reconstructions as HPS00 (available at the NOAA World Data Center for Paleoclimatology; http://www.ngdc.noaa.gov/paleo/paleo.html). To insure that the information in the borehole and instrumental temperature data sets are represented in a spatially commensurate manner, and that the highly variable spatial sampling density in the individual borehole GST reconstructions does not spatially bias estimates of hemispheric means, we average the individual borehole records onto the $5° \times 5°$ spatial grid of the instrumental surface temperature data [*Jones et al.*, 1999]. The gridding process yields 94 grid points (out of a total of 1296 required to cover the Northern Hemisphere), with an average of 5 borehole records contributing to each grid point estimate, but with occupancy

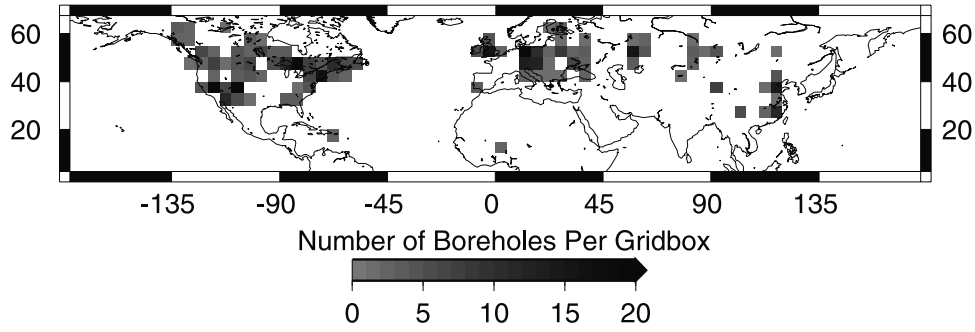


Figure 1. Distribution of the Northern Hemisphere borehole temperature reconstructions. The locations of the (94) 5° by 5° grid points sampled by the individual borehole reconstructions of HPS00 are shown, along with density of samples contributing to each grid point.

varying spatially from 1 to 22 boreholes per grid point (Figure 1). Nearly identical results are obtained through both simple arithmetic and inverse distance weight grid point estimation schemes, suggesting that the results are robust with respect to the details of the gridding process. We adopt, as approximate estimates of the borehole grid point relative sampling errors, the inverse square root of the number of borehole samples within the grid point (the relative error weights thus vary between 0.21 and 1).

[9] Spatial sampling experiments with the instrumental record suggest that the spatial sampling bias associated with estimating the full Northern Hemisphere mean temperature from the sparse, terrestrial and almost exclusively extratropical grid represented by the boreholes sites is relatively minimal during the 20th century. However, a bias in estimating past hemispheric temperature trends from this limited spatial sampling is likely if, as is suggested from previous work [*Mann et al.*, 2000; *Shindell et al.*, 2001; *Hendy et al.*, 2002; *Esper et al.*, 2002; *Mann and Hughes*, 2002; *Mann*, 2002], there are significant differences between extratropical terrestrial and full (ocean and land, tropical and extratropical) hemispheric surface temperature trends in past centuries. The implications of such potential sampling biases, in the estimate of true, full Northern Hemisphere temperature variations in past centuries, is discussed later.

[10] HPS00 use the entire 20th century for comparing instrumental and borehole warming trends. In the context of their analysis, this represents an assumption that the 20th century warming rates measured by boreholes are representative of the entire 20th century. However, the median logging date of the boreholes used in their analysis is 1978. Moreover, since there is some loss of information in the near-surface portion of the borehole profiles, even recently logged borehole records may not record the accelerated observed warming of the past couple of decades. We thus use a more conservative interval 1900–1980 as our standard for comparing borehole and instrumental data, but we also examine the sensitivity to the precise interval over which the data are compared (e.g., 1900–1960; 1900–1980; and 1900–1998).

[11] The noise and bias contributions discussed above contaminate any underlying SAT signal in the borehole data. The individual borehole GST reconstructions of HPS00 exhibit a spatial distribution of positive and negative trends that is statistically indistinguishable from a random distribution for the first three centuries (Figure 2). Such an observation is inconsistent with expectations for actual SAT data, which are known to exhibit considerable spatial coherence and large-scale correlation structure. Only for the 19th century ($p < 0.20$) and 20th century ($p < 0.05$) do the histograms of the individual borehole reconstructions show evidence of such nonrandom spatial structure. The long-term trend constructed by a simple arithmetic averaging of the individual borehole reconstructions, as presented by HPS00, thus represents (at least through 1800) a small residual of larger nearly canceling random trends, and is thus unlikely to represent a meaningful hemispheric SAT history. The century-by-century distributions of the gridded borehole data yields greater evidence of nonrandom spatial structure (Figure 2) though, as discussed below, much of this structure appears to represent bias, rather than an underlying SAT signal. Indeed [see also *Briffa and Osborn*, 2002], simply gridding the data, prior to averaging over the Northern Hemisphere, and forming an areally representative Northern Hemisphere mean estimate through a cosine latitude weighed mean of the grid point data reduces the amplitude of the long-term trend from approximately 1°C (as in HPS00) to 0.8°C, bringing the estimated change more in line with other proxy-based estimates (Figure 3), though the borehole data still suggest colder conditions in past centuries.

[12] While HPS00 demonstrate a favorable comparison between the 20th century hemispheric mean trend in borehole (1900–1980 roughly) and instrumental temperature (1900–2000) data, a comparison of the spatial patterns in the two datasets indicates prominent discrepancies at colocated grid points (Figure 4), with more than one in four (27/94) of the borehole grid point trends exhibiting the wrong sign trend during the 20th century. A pattern correlation of $r = 0.09$ (1900–1980) indicates a weak relationship between the spatial information in the two datasets. This conclusion is largely robust with respect to the precise interval over which it is assumed the borehole trends are representative: $r = 0.11$ for the interval 1900–1998; $r = 0.15$ for the interval 1900–1960 (the spatial correlations are moderately larger if the grid point GST trends are weighted with respect to their estimated sampling errors: $r = 0.15$ for 1900–1980; $r = 0.20$ for 1900–1980; $r = 0.25$ for 1900–1960). The discrepancy is particularly large over much of North America, when boreholes show greater warming, perhaps due in part to land use changes

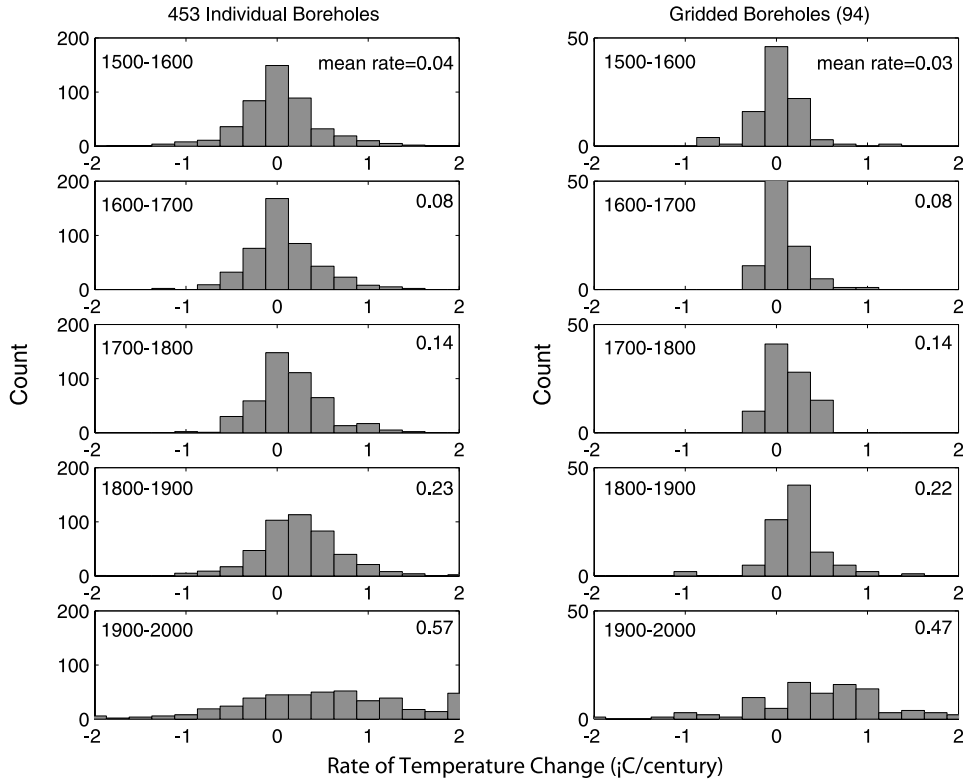


Figure 2. Histogram of distribution of trends among the (a) 453 individual and (b) gridded Northern Hemisphere borehole reconstructions for each of the past five centuries. Mean rates of change are indicated for each century in both cases.

[*Skinner and Majorowicz*, 1999]. The lack of agreement with regard to the sign of trends for the five highest latitude (Arctic) grid points underscores the added complications in interpreting annual ground surface temperature trends in regions of permafrost and/or extended seasonal snow cover [see, e.g., *Beltrami*, 1996].

[13] Typical surface temperature correlation-scale estimates [e.g., *Briffa and Jones*, 1993; *Mann and Park*, 1994] suggest at most about 15 spatial degrees of freedom at annual timescales in the surface temperature field over the Northern Hemisphere domain sampled by the boreholes, with $M_{eff} = 6$ or so degrees of freedom at decadal and longer timescales. Inspection of the patterns of anomalies shown in Figure 4 substantiates such an approximate estimate of spatial degrees of freedom for both the gridded borehole and instrumental surface temperature data. A similarity in the level of spatial coherence of trends between instrumental and borehole data thus exists despite the lack of correlation of the actual spatial patterns of change in the two datasets. The fact that the gridded borehole data are far less spatially random than the individual borehole records (Figure 2) can be attributed to an averaging out of much of the truly spatially random error in the borehole GST measurements on upscaling to the 5° by 5° grid point scale. We thus

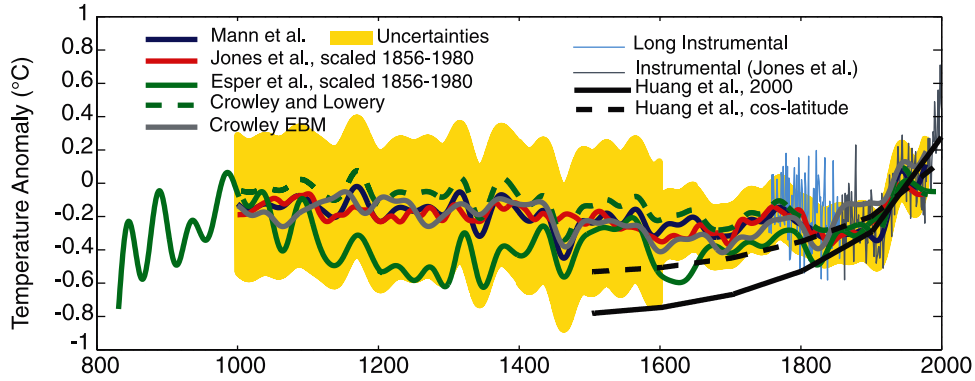


Figure 3. Comparisons between different Northern Hemisphere temperature reconstructions and instrumental record. Shown are smoothed (40 year lowpassed) reconstructions and, in the case of the Mann et al. reconstruction, the associated 95% confidence interval. Shown for comparison are the HPS00 reconstruction, and the areally weighted mean of gridded HPS00 borehole reconstructions.

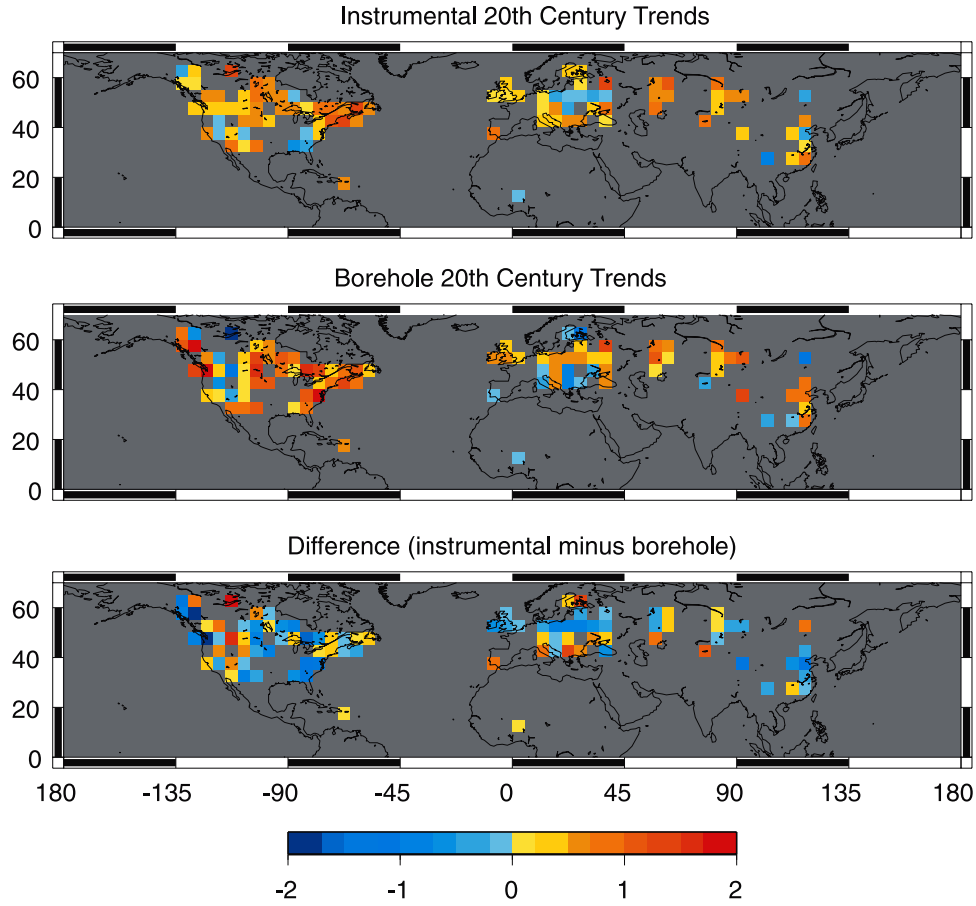


Figure 4. Comparison of linear 20th century grid point trends 1900–1980 with the (94 grid point) borehole grid mask shown in Figure 1 applied for (a) gridded borehole reconstructions, (b) instrumental surface temperature record, and (c) difference between borehole and instrumental trends. Scales are °C for the instrumental grid point data and nominal °C (based on HPS00 inversion) for the borehole grid point data.

interpret the predominant source of error at the grid point scale and beyond as being associated with nonrandom processes (e.g., seasonal snow cover and land-use influences) that impose intrinsically large-scale patterns of bias in the estimation of SAT trends from borehole-based GST trends. The essential distinction between borehole GST trends (which are themselves uncertain) and SAT estimates derived from those trends (which are likely to be biased for reasons discussed above) has not been given appropriate attention in previous work. We thus choose to be explicit that, while we will henceforth treat the gridded borehole GST reconstructions as estimates of grid point SAT trends, those estimates contain both random uncertainties and systematic biases that must be carefully considered in isolating the weak underlying SAT signal.

[14] The presence of nonrandom, large-scale patterns of bias in borehole-based SAT estimates makes invalid the proposition (e.g., HPS00) that an estimate of hemispheric mean temperature based on an average of hundreds of individual borehole GST reconstructions can provide any significant reduction in the uncertainty in hemispheric mean SAT estimated from the data. Rather, the average of an arbitrarily large number of borehole GST reconstructions over the Northern Hemisphere domain sampled by HPS00 can reduce the error in the hemispheric mean SAT trend estimate by no more than a factor of $M_{eff}^{1/2}$, or, roughly 2.5, relative to the RMS error at the grid point level, regardless of how many individual borehole records contribute to the estimate. The RMS error at the grid point level for the 20th century (note that the RMS errors in previous centuries might be smaller) is readily estimated as approximately 1°C from the mean amplitude of the observed differences between the colocated instrumental and borehole grid point 20th century SAT trend estimates (Figure 4, panel c). Such errors thus preclude reducing the error in a borehole hemispheric mean SAT estimate by simple averaging to less than roughly 1°C/2.5 or 0.4°C, regardless of how many individual borehole records are averaged.

[15] This intrinsic error of 0.4°C is more than 50% of the amplitude of the 20th century trend in the HPS00 reconstruction, and considerably larger than the trend in the HPS00 reconstruction for every century prior to the 20th century. A simple arithmetic mean of either individual or gridded borehole reconstructions thus cannot provide a useful estimate of past hemispheric SAT changes, as it yields trends, which are largely within the intrinsic tolerance of the estimates. A proper estimate of hemispheric surface temperature must therefore identify the weak SAT signal contained within the spatial borehole data, and separate that signal from the contributions of spatial noise and systematic

bias if a meaningful reconstruction is to be expected. We use spatial regression methods to optimally determine this surface temperature signal and its associated uncertainty, as detailed in section 3.

3. Methods

[16] We employ a spatial signal detection approach that bears a loose relationship with ''optimal detection'' approaches used in anthropogenic climate signal fingerprinting [*Mitchell et al.*, 2001]. In such ''optimal detection'' approaches, one seeks to identify, through generalized linear regression, the estimate of a target signal (as predicted by a model) in empirical data. Detection is accomplished through rotation of the empirical data, in EOF state-space, away from the direction of maximal noise (as estimated from, e.g., a control model simulation). In our approach, an independent estimate of noise is not available. Rather, we employ an EOF rotation of the information in the borehole dataset toward an independent estimate of the target spatial SAT signal from the instrumental record, based on ordinary (potentially weighted) least squares spatial regression. Once an optimal rotation is found that provides maximal (and statistically significant) agreement between the spatial information in the borehole and instrumental record during the 20th century, the associated eigenvector rotation is used to project the estimated borehole SAT signal back in time. An assumption of this approach is that the statistical relationship between borehole GST estimates and the target SAT estimates thereby determined is stationary back in time. While a similar assumption of stationarity is implicit in any attempt at borehole-based SAT reconstruction (e.g., HPS00), such an assumption is not strictly valid if the relative impacts of sources of bias noted earlier (e.g., seasonal snow cover changes) change significantly back in time. Though it is arguably most appropriate to use the borehole grid point GST reconstructions, weighted with respect to the estimated sampling errors, analyses are performed (and results cited) using both the unweighted and weighted data to test the sensitivity of the results to spatial variations in borehole grid point sampling density.

[17] The borehole GST reconstructions can be represented by an $M \times N$ data matrix $\boldsymbol{B}$, where $M = 94$ spatial grid point estimates, $N = 6$ time values (t = 1500,1600,1700,1800,1900, t_f; t_f = 1960, 1980, or 1998 depending on the 20th century borehole GST endpoint assumed). If we wish to work with the error-weighted data matrix, we divide the M rows of $\boldsymbol{B}$ by the M estimated grid point sampling errors discussed earlier. We can expand $\boldsymbol{B}$ in an empirical orthogonal eigenvector basis set using a conventional singular value decomposition (SVD) of the unnormalized, time-centered data matrix (this is equivalent to performing Principal Component Analysis (PCA) on the data covariance matrix):

$$\boldsymbol{B} = \sum_{k=1}^{N} \lambda_k \boldsymbol{u}_k^T \boldsymbol{v}_k \tag{1}$$

In this convention, λ_k^2 is the relative variance resolved by the kth eigenvector, $\boldsymbol{u}_k$ is its normalized spatial pattern (Empirical Orthogonal Function or ''EOF'') and $\boldsymbol{v}_k$ is its normalized temporal pattern (principal component or ''PC''). In this case, the sum extends over $k = 1$, N basis empirical eigenvectors. We retain only the first two eigenvectors (see Figure 5), which are the only eigenvectors determined to be statistically significant by the traditional ''Rule N'' criterion [*Preisendorfer*, 1988]. These two eigenvectors span the vast majority of the variance (first eigenvector: 82% + second eigenvector: 15% = 97% total) in the sampling-error weighted data, allowing for a useful two-dimensional state-space representation of the SAT signal, noise, and bias information in the borehole dataset [it should be noted that the first and second eigenvector each resolve a similar fraction of variance in the 20th century borehole data, 45% and 38%, respectively]. The corresponding numbers for the unweighted data are extremely similar (first eigenvector: 81% + second eigenvector: 16% = 97% of the total variance, and 46% and 43% of the variance are resolved by the first two eigenvectors, respectively, during the 20th century). Essentially the same results are obtained if the first three eigenvectors are used instead, indicating that the results of the analysis are robust with respect to precisely how many eigenvectors are retained. The eigenspectrum for the EOF decomposition, moreover, is essentially insensitive to whether the 453 individual borehole GST reconstructions, or grid point GST reconstructions, are used.

[18] We find the optimal expression of the pattern of the 20th century surface temperature trend signal within the borehole grid point reconstructions through regressing the spatial information in the two borehole EOF patterns against that in the instrumental record,

$$\hat{\boldsymbol{I}} = a\hat{\mathbf{u}}_1 + b\hat{\mathbf{u}}_2 + e, \tag{2}$$

where $\hat{\boldsymbol{u}}_1$ and $\hat{\boldsymbol{u}}_2$ are the spatially centered, normalized spatial eigenvectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, $\hat{\boldsymbol{I}}$ is the spatially centered, normalized pattern of trend in the 20th century instrumental record, a and b are regression coefficients (partial correlation coefficients in this context) and $\boldsymbol{e}$ is the residual error term. If $\boldsymbol{B}$ is the sampling error-weighted data matrix, then the procedure represents a weighted linear regression (wherein the statistical sampling error in the borehole grid point estimates is taken into account in the regression, but no sampling error is assumed for the instrumental grid point estimates, an assumption quite consistent with the relative magnitudes of sampling errors in the two data sets).

[19] The residuals $\boldsymbol{e}$ from this regression are normally distributed (i.e., they pass a χ-squared normality test at $\alpha = 0.05$) and not serially correlated (p-value for the lag-1 autocorrelation >0.3) so that statistical inferences (e.g., the standard error estimates for the regression coefficients) from a traditional analysis of variance are valid. The angle of rotation in two-dimensional EOF space between the direction of maximum variance in the borehole data (the EOF #1 axis) and the optimal estimate of the target-warming pattern is given by

$$\phi = tan^{-1}(b/a). \tag{3}$$

The optimal estimate of the surface temperature signal component of the borehole data matrix $\boldsymbol{B}$ is then simply given by the rotation in two-dimensional EOF space:

$$\underline{\boldsymbol{B}} = cos\phi\lambda_1 \boldsymbol{u}_1^T \boldsymbol{v}_1 + sin\phi\lambda_2 \boldsymbol{u}_2^T \boldsymbol{v}_2, \tag{4}$$

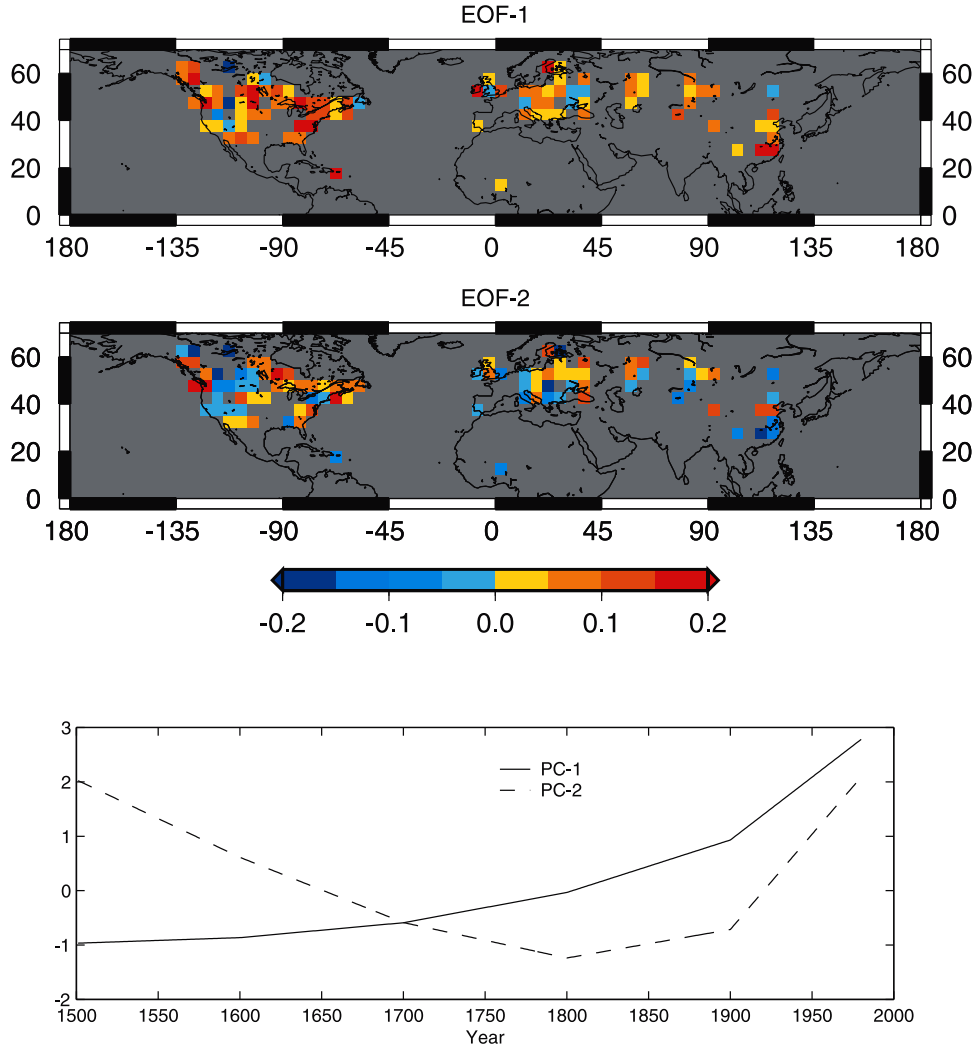


Figure 5. Spatial and temporal patterns for the two leading eigenvectors of the error-weighted gridded borehole data. The spatial patterns (EOF #1 and EOF #2) are shown in Figure 5a and 5b, while their trends from 1500–1980 are described by their corresponding Principal Component (PC) time series (Figure 5c). The patterns have been unweighted by the associated inverse sampling errors for purpose of display.

where $\underline{\boldsymbol{B}}$ is an $M \times N$ matrix containing the estimated SAT signal in $\boldsymbol{B}$. If $\boldsymbol{B}$ was the error-weighted data matrix, then the SAT reconstruction is obtained by dividing the rows of $\underline{\boldsymbol{B}}$ through by the grid point sampling errors. An areally weighted average of the M rows of $\underline{\boldsymbol{B}}$ provides the optimal estimate of the Northern Hemisphere mean surface temperature signal, with the residual variance interpreted as a combination of noise and bias. The uncertainties in the optimal reconstruction are readily determined from the standard errors in the regression coefficients in equation (2) which translate to uncertainties in the rotation described by equation (4).

[20] If the contribution of noise and bias were small, one would expect the optimal projection to yield $\Phi \approx 0$ (i.e., the surface temperature signal would lie along EOF #1 axis, the direction of maximum variance in the borehole data). We find, however, that the optimal rotation angle is significantly different from zero (i.e., it possesses a significant component along the EOF #2 axis). Using the standard period 1900–1980 to estimate the instrumental trend pattern $\hat{\boldsymbol{I}}$ in (2), based on the weighted borehole grid point data, we find an optimal angle of projection of $\Phi \approx 60°$ degrees in EOF space, associated with a weaker correlation ($r = a = 0.12$) against the EOF #1 pattern than the EOF #2 pattern ($r = b = 0.22$) [the procedure gives $a = 0.08$ and $b = 0.14$ for the unweighted borehole data]. The resultant correlation between the instrumental trend pattern and the borehole data projected onto the corresponding optimal rotation axis, though modest ($r = 0.25$), is statistically significant at the $p = 0.01$ level (we assume a one-sided significance criterion, since a negative spatial correlation between the borehole and instrumental data would be rejected as unphysical). The associated resolved spatial variance, though modest ($r^2 = 0.06$) is significantly greater than that resolved by the un-rotated gridded borehole data (i.e., the square $r^2 = 0.04$ of the spatial pattern correlation $r = 0.20$ cited earlier). If $\hat{\boldsymbol{I}}$ is determined from the 1900–1998 instrumental record (i.e., the full 20th century, as assumed in the comparisons shown

by HPS00), the optimal projection is essentially along the EOF #2 axis ($\Phi \approx 66°$). If $\hat{\boldsymbol{I}}$ is determined from the 1900–1960 interval, the optimal signal projection lies roughly along the angle $\Phi \approx 34°$ in EOF space. For no choice of instrumental interval, however, is the optimal angle of projection near $\Phi = 0$. An analysis of the unweighted data yields similar results ($\Phi \approx 60°$, $\Phi \approx 60°$, and $\Phi \approx 15°$, respectively). The Northern Hemisphere temperature reconstructions resulting from the three different choices of optimal projection interval, for both weighted and unweighted data, as shown later, all lie essentially within the uncertainties estimated for the standard (1900–1980) case. An analysis of the unweighted data yields similar results.

4. Discussion

[21] The combination of signal, noise, and bias in the borehole GST reconstructions, as shown above, can efficiently be represented by a two-dimensional phase space spanned by the two leading eigenvectors of the data set. This representation allows us to determine the optimal SAT signal in the borehole GST reconstructions in a relatively straightforward manner. The instrumental surface temperature signal in the borehole data for the standard case (1900–1980 projection interval) is seen to be weakly detectable (i.e., at the $p = 0.01$ level) and is aligned more closely with the second eigenvector than with the first. This latter observation implies that much of the non-SAT related noise/bias noise variance is associated with the first eigenvector. The first eigenvector closely resembles, in its temporal shape the HPS00 reconstruction, and projects strongly onto hemispheric-mean temperature (Figure 5). Thus, a sizable component of the HPS00 hemispheric mean temperature reconstruction is interpreted as noise/bias, which as discussed above, we speculate may result in large part from changes in seasonal snow cover. However, we emphasize here that we seek simply seek to isolate the true SAT signal in borehole temperature reconstructions. Determining the ultimate combination of factors leading to the inferred patterns of bias in the data will require further research.

[22] The second eigenvector, by contrast, exhibits a more variable temporal history, with a negative trend from 1500 to 1800, slight increase from 1800 to 1900, and a large positive trend from 1900 to the present. The more heterogeneous spatial pattern associated with the second eigenvector, moreover, nearly vanishes in the hemispheric mean (Figure 5). Thus, the small component of signal that is aligned with the EOF #1 axis nevertheless dominates the optimally estimated hemispheric-mean SAT signal. In contrast, the signal contribution from the EOF #2 pattern, associated with a spatially variable pattern of warming and cooling, largely cancels in the hemispheric mean. The optimal borehole SAT reconstruction thus exhibits significantly less cooling back in time than, for example, the HPS00 reconstruction because the estimated SAT signal projects in large part onto a pattern of zero hemispheric-mean temperature change.

[23] The optimal borehole SAT reconstruction for the standard case (1900–1980 projection interval) exhibits a hemispheric mean trend back in time that is considerably reduced in amplitude relative not only to the HPS00 mean reconstruction, but even to the smaller trend resulting from the areally weighted mean of the gridded HPS00 borehole GST reconstructions (Figure 6). A similar conclusion is reached regardless of the precise projection interval, though the reconstruction based on the 1900–1960 projection interval yields a trend that is only slightly reduced relative to the areally weighted gridded borehole mean reconstruction. While the analysis of the weighted borehole data is favored, we note that the results are relatively similar for the unweighted data. Use of the weighted data changes the reconstruction for the standard projection interval (1900–1980) little, but brings the reconstructions corresponding to other two projection intervals (1900–1960 and 1900–1998) into better common agreement (Figure 6b), and into better agreement with other long-term hemispheric temperature estimates. As the surface temperature signal is only weakly detected in the borehole data, the borehole temperature reconstruction necessarily exhibits sizable uncertainties. The optimal borehole hemispheric temperature reconstruction for the standard case is well within mutual uncertainty estimates of the *Mann et al.* [1999] proxy-based annual Northern Hemisphere surface temperature reconstruction which is shown both as a smoothed annual record, and as piecewise continuous century-scale trends (with associated uncertainties), to allow for a proper comparison with the centennial resolution borehole estimates (Figure 6c). The HPS00 estimate, however, is inconsistent not only with both of these reconstructions, within estimated uncertainties, but also with the available instrumental Northern Hemisphere mean annual temperature series back through the mid 18th century (Figure 6c) (The extension of the instrumental Northern Hemisphere annual mean temperature record prior to the mid 19th century is based on a composite of the available instrumental SAT series in Europe and North America (13 available back to 1820, 9 back to 1777, 6 back to 1762, and 4 back to 1753) standardized to have the same mean and standard deviation as the full instrumental NH record of *Jones et al.* [1999] over an 1856–present overlap interval. The correlation with the full series during the overlap interval is $r = 0.45$.)

[24] There has been considerable recent discussion [e.g., *Esper et al.*, 2002; *Briffa and Osborn*, 2002; *Mann and Hughes*, 2002] of apparent discrepancies between various proxy-based reconstructions of Northern Hemisphere mean temperature spanning roughly the past millennium. Many of these differences appear to be associated simply with the fact that different reconstructions emphasize different seasons and latitudinal ranges in the underlying proxy data and target region for reconstruction [see *Briffa and Osborn*, 2002; *Mann and Hughes*, 2002; *Mann*, 2002]. A comparison of various different reconstructions was shown in Figure 3 including reconstructions of full Northern Hemisphere annual mean temperatures based either on explicit calibration of diverse (tropical and extratropical) proxy data calibrated against the full Northern Hemisphere instrumental surface temperature record [*Mann et al.*, 1999; *Crowley and Lowery*, 2000] or the modeled surface temperature response of the full Northern Hemisphere to estimated radiative forcing changes [*Crowley*, 2000]. Shown also in Figure 3 are records that implicitly represent extratropical (and primarily warm-season) Northern Hemisphere temperature variations [*Jones et al.*, 1998, *Esper et al.*, 2002], but

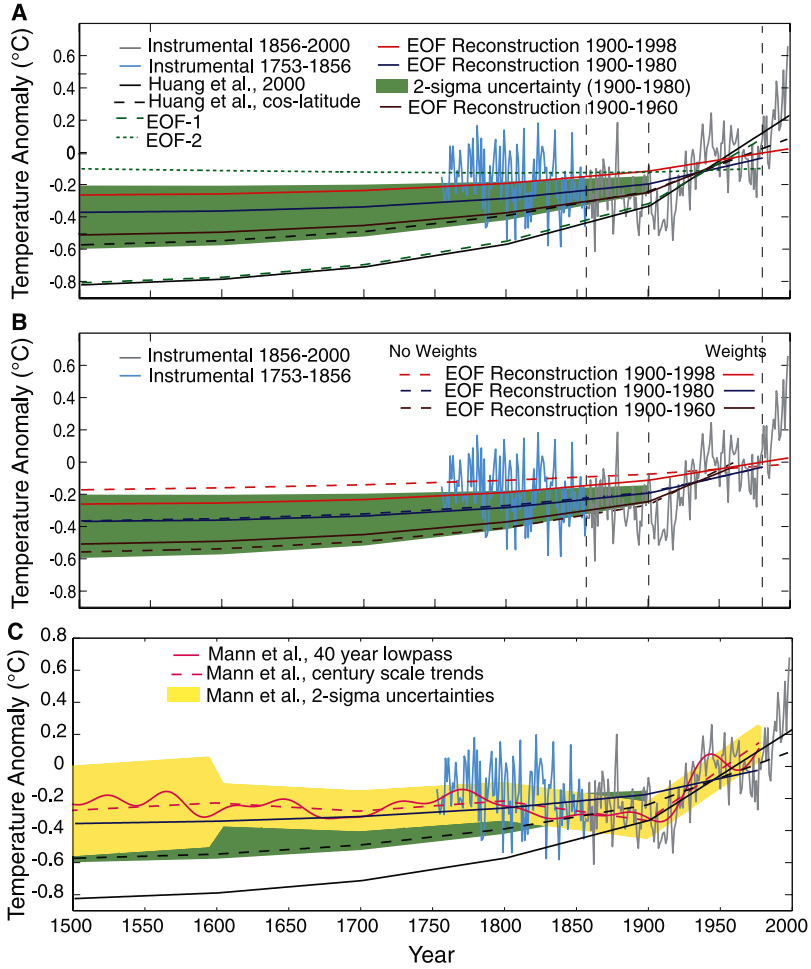


Figure 6. Borehole-based reconstructions of Northern Hemisphere mean temperature compared with instrumental Northern Hemisphere annual mean surface temperature (*Jones et al.* [1999]; light gray) extended to mid 18th century with the available instrumental data (see text; light blue-gray). (a) HPS00 Northern Hemisphere temperature reconstruction compared to areally weighted mean of gridded HPS00 borehole reconstructions, and optimal borehole hemispheric SAT estimate (using 1900–1960, 1900–1980, and 1900–1998 projection intervals). The 95% confidence interval (light green shading) is also shown for the reconstruction corresponding to the 1900–1980 projection interval. Also shown are the separate projections of the first two eigenvectors onto hemispheric mean temperature. (b) Comparison of the optimal borehole SAT estimate for the unweighted and weighted borehole grid point GST reconstructions. (c) Comparison of HPS00 reconstruction, areally weighted mean of gridded HPS00 borehole reconstructions, and the optimal borehole estimate (based on 1900–1980 target instrumental signal; 95% confidence interval also shown) with *Mann et al.* [1999] reconstruction (latter shown both as smoothed (40 year lowpassed) annual values, and piecewise continuous linear trend estimate with its associated 95% confidence interval).

have been rescaled against the annual mean full Northern Hemisphere instrumental surface temperature record of *Jones et al.* [1999] from 1856–1980 to allow direct comparison with the other reconstructions. Using such a convention, any differences between the true Northern Hemisphere estimates and the rescaled extratropical estimates prior to the mid 19th century can be interpreted as indicating a change in the relationship between warm-season extratropical and annual, full Northern Hemisphere mean surface temperature variations prior to that time. Such changes should be expected if the primary influences on hemispheric-mean or global-mean surface temperatures during the pre-anthropogenic era (e.g., solar and volcanic forcing, and internal variability) are different from those (e.g., anthropogenic forcing) dominating more recent variations. Nearly all of the proxy reconstructions are seen (Figure 3) to be internally consistent (i.e., well within the uncertainties of the Mann et al reconstruction). The *Esper et al.* reconstruction, which indicates cooler temperatures during the early 17th through early 19th centuries, is an exception. The greater cooling indicated in this latter reconstruction may result from enhanced cooling in the Northern Hemisphere continental centers during the height of the "Little Ice Age" [see, e.g., *Shindell et al.*, 2001], most likely associated, in this case, with the enhanced volcanic cooling sampled by continental warm-season sen-

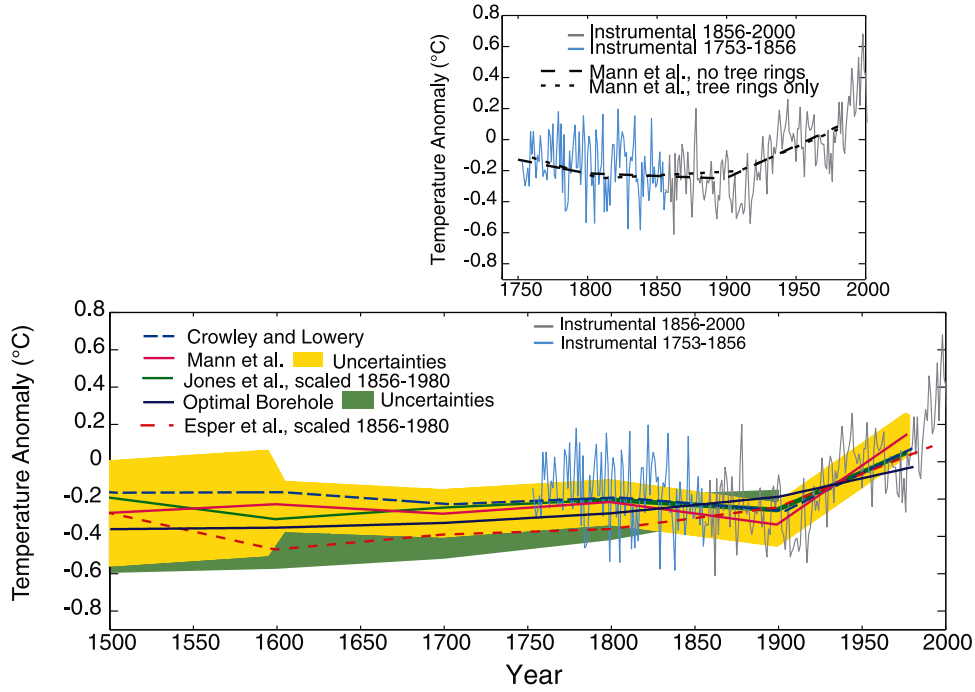


Figure 7. Comparison of optimal borehole reconstruction (based on 1900–1980 target instrumental signal; 95% confidence interval shown) with piecewise continuous linear trends fit to various proxy-based hemispheric temperature reconstructions. The *Jones et al.* [1998] and *Esper et al.* [2002] extratropical summer temperature reconstructions have been scaled to the full Northern Hemisphere 1856–1980 annual mean. All other proxy-series shown are based on original published calibrations, but aligned to a 1961–1990 reference period mean. Note that the Mann et al. only tree ring and only nontree ring reconstructions are only available back to 1750 [see *Mann et al.*, 2000].

sitive tree ring data [see *Mann*, 2002]. Such an assertion is bolstered by a striking coincidence between anomalous cold episodes (e.g., late 12th–early14th, mid 15th, early 17th, and early 19th centuries) and the periods of peak explosive volcanic forcing [e.g., *Crowley*, 2000] over the past 1000 years.

[25] It is instructive, in this context, to compare the borehole temperature reconstructions against the various other reconstructions when represented by their piecewise continuous century-scale trends in a manner directly comparable with the borehole estimates (Figure 7). Shown also in Figure 7 are the trends for the Mann et al. Northern Hemisphere temperature reconstruction based both on proxy data sets consisting of only nontree ring, and only tree ring indicators [*Mann et al.*, 2000]. In all cases, the trends in the proxy reconstructions are consistent with the optimal borehole estimates reported here within indicated uncertainties. In combination with previous regional comparisons of tree ring and borehole reconstructions [*Beltrami et al.*, 1995; *Majorowicz and Skinner*, 2001], these latter comparisons refute the argument [*Broecker*, 2001] that the use of tree ring indicators (carefully screened for low-frequency content as in *Mann et al.* [1998], and in combination with other proxies) leads to any inherent loss of resolution of century-scale climate variability relative to estimates from terrestrial borehole data. It is not necessary to analyze tree ring data with the "RCS" method (as in *Esper et al.* [2002]) to come to this conclusion [see also *Briffa et al.*, 2001].

[26] The optimal borehole hemispheric SAT reconstruction exhibits a slightly colder 16th and early 17th century, than the full Northern Hemisphere proxy reconstructions, and it is observed to be intermediate in its primary features between the *Mann et al.* full Northern Hemisphere temperature reconstruction, and the (appropriately scaled) *Esper et al.* [2002] extratropical Northern Hemisphere temperature reconstruction. This latter observation is consistent with the existence of real, though evidently modest, differences between extratropical and full Northern Hemisphere temperature trends in past centuries as discussed earlier, since the borehole data, like the *Esper et al.* [2002] tree ring data, are almost entirely extratropical in nature. Moreover since many of the borehole sites are at least partially insulated from winter continental SAT variations through seasonal snow cover, they might be expected to exhibit a similar warm-season sampling bias to the *Esper et al.* [2002] reconstruction (though, as discussed before, this bias is likely to be time dependent in the case of the borehole estimates).

5. Conclusions

[27] Application of appropriate signal detection techniques to previously published borehole temperature reconstructions yields an estimated SAT signal that is consistent with proxy-based estimates of hemispheric surface temperature changes in past centuries. It is thus incorrect to argue [*Broecker*, 2001] (see also the response by *Bradley et al.*

[2001]) that information from terrestrial boreholes is in conflict with conclusions regarding past hemispheric temperature changes from other proxy-based reconstructions. Rather, reconstructions based on the restricted information contained in borehole data that is consistent with modern SAT data, reinforces conclusions based on other proxy indicators. The analysis presented in this study also highlights the importance of taking into account differences in regional sampling (e.g., extratropical only versus full Northern Hemisphere emphasis) in properly comparing estimates of Northern Hemisphere mean temperature variation in past centuries. Finally, by reconciling the estimates of past temperature change from boreholes and proxy climate indicators, our analysis tends to support low-to-medium range estimates [*Crowley*, 2000; *Gerber et al.*, 2003] of climate sensitivity.

[28] **Acknowledgments.** The authors acknowledge support for this work from the Earth Systems History Program (NOAA and NSF), and Department of Energy. We gratefully acknowledge the helpful comments of several anonymous reviewers.

References

Beltrami, H., Active Layer Distortion of Annual Air/Soil Thermal Orbits: Permafrost and Past Climates, *Permafrost Periglacial Processes*, *7*, 101–110, 1996.

Beltrami, H., D. S. Chapman, S. Archambault, and Y. Bergeron, Reconstruction of high resolution ground temperature histories combining dendrochronological and geothermal data, *Earth Planet. Sci. Lett.*, *136*, 437–445, 1995.

Bradley, R. S., and P. D. Jones, "Little Ice Age" summer temperature variations: Their nature and relevance to recent global warming trends, *The Holocene*, *3*, 367–376, 1993.

Bradley, R. S., K. R. Briffa, T. J. Crowley, M. K. Hughes, P. D. Jones, and M. E. Mann, Scope of Medieval Warming, *Science*, *292*, 2011–2012, 2001.

Briffa, K. R., and P. D. Jones, Global surface air temperature variations during the twentieth century: Part 2, Implications for large-scale high-frequency palaeoclimatic studies, *The Holocene*, 77–88, 1993.

Briffa, K. R., and T. J. Osborn, Blowing Hot and Cold, *Science*, *295*, 2227–2228, 2002.

Briffa, K. R., T. J. Osborn, F. H. Schweingruber, I. C. Harris, P. D. Jones, S. G. Shiyatov, and E. A. Vaganov, Low-frequency temperature variations from a northern tree ring density network, *J. Geophys. Res.*, *106*, 2929–2941, 2001.

Broecker, W. S., Was the Medieval Warm Period Global?, *Science*, *291*, 1497, 2001.

Crowley, T. J., Causes of Climate Change Over the Past 1000 Years, *Science*, *289*, 270–277, 2000.

Crowley, T. J., and T. Lowery, How Warm Was the Medieval Warm Period?, *Ambio*, *29*, 51–54, 2000.

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

Free, M., and A. Robock, Global Warming in the Context of the Little Ice Age, *J. Geophys. Res.*, *104*, 19,057–19,070, 1999.

Gerber, S., F. Joos, P. P. Bruegger, T. F. Stocker, M. E. Mann, and S. Sitch, Constraining Temperature Variations over the last Millennium by Comparing Simulated and Observed Atmospheric CO_2, *Clim. Dyn.*, *20*, 281–299, 2003.

Harris, R. N., and D. S. Chapman, Borehole Temperatures and a Baseline for 20th-Century Global Warming Estimates, *Science*, *275*, 1618–1621, 1997.

Harris, R. N., and D. S. Chapman, Mid-Latitude (30′–60′ N) climatic warming inferred by combining borehole temperatures with surface air temperatures, *Geophys. Res. Lett.*, *28*, 747–750, 2001.

Hendy, E. J., M. K. Gagan, C. A. Alibert, M. T. McCulloch, J. M. Lough, and P. J. Isdale, Abrupt Decrease in Tropical Pacific Sea Surface Salinity at End of Little Ice Age, *Science*, *295*, 1511, 2002.

Huang, S., H. N. Pollack, and P.-Y. Shen, Temperature Trends Over the Past Five Centuries Reconstructed from Borehole Temperature, *Nature*, *403*, 756–758, 2000.

Jones, P. D., K. R. Briffa, T. P. Barnett, and S. F. B. Tett, High-Resolution Paleoclimatic Records for the Last Millennium: Interpretation, Integration, and Comparison with Circulation Model Control-Run Temperatures, *The Holocene*, *8*, 455–471, 1998.

Jones, P. D., M. New, D. E. Parker, S. Martin, and J. G. Rigor, Surface air temperature and its changes over the past 150 years, *Rev. Geophys.*, *37*, 173–199, 1999.

Keigwin, L. D., and R. S. Pickart, Slope Water Current over the Laurentian Pan on Interannual to Millennial Time Scales, *Science*, *286*, 520–523, 1999.

Majorowicz, J. A., and W. R. Skinner, Reconstruction of the surface warming history of western interior Canada from borehole temperature profiles and other climate information, *Clim. Res.*, *16*, 157–200, 2001.

Mann, M. E., Lessons for a New Millennium, *Science*, *289*, 253–254, 2000.

Mann, M. E., The Value of Multiple Proxies, *Science*, *297*, 1481–1482, 2002.

Mann, M. E., and M. K. Hughes, Tree-Ring Chronologies and Climate Variability, *Science*, *296*, 848, 2002.

Mann, M. E., and J. Park, Global scale modes of surface temperature variability on interannual to century time scales, *J. Geophys. Res.*, *99*, 25,819–25,833, 1994.

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

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

Mann, M. E., E. Gille, R. S. Bradley, M. K. Hughes, J. T. Overpeck, F. T. Keimig, and W. Gross, Global Temperature Patterns in Past Centuries: An interactive presentation, *Earth Interactions*, *4-4*, 1–29, 2000.

Mitchell, J. F. B., D. J. Karoly, G. C. Hegerl, F. W. Zwiers, M. R. Allen, and J. Marengo, Detection of Climate Change and Attribution of Causes, in *Climate Change 2001: The Scientific Basis*, edited by J. T. Houghton, et al., pp. 695–738, Cambridge Univ. Press, New York, 2001.

Pollack, H. N., S. Huang, and P.-Y. Shen, Climate Change Record in Subsurface Temperatures: A Global Perspective, *Science*, *282*, 279–281, 1998.

Preisendorfer, R. W., *Principal Component Analysis in Meteorology and Oceanography*, Elsevier Sci., New York, 1988.

Shindell, D. T., G. A. Schmidt, M. E. Mann, D. Rind, and A. Waple, Solar Forcing of Regional Climate Change During the Maunder Minimum, *Science*, *294*, 2149–2152, 2001.

Skinner, W. R., and J. A. Majorowicz, Regional climatic warming and associated twentieth century land-cover changes in north-western North America, *Clim. Res.*, *12*, 39–52, 1999.

Thompson, D. W. J., and J. M. Wallace, Regional Climate Impacts of the Northern Hemisphere Annular Mode, *Science*, *293*, 85–89, 2001.

R. S. Bradley and F. T. Keimig, Department of Geosciences, 233 Morrill Science Center, University of Massachusetts, Amherst, MA 01003-9297, USA. (rbradley@geo.umass.edu; frank@geo.umass.edu)

M. K. Hughes, Laboratory of Tree Ring Research, University of Arizona, Tucson, AZ 85721, USA. (mhughes@ltrr.arizona.edu)

M. E. Mann and S. Rutherford, Department of Environmental Sciences, Clark Hall, 291 McCormick Road, University of Virginia, Charlottesville, VA 22903, USA. (mann@virginia.edu; srutherford@virginia.edu)