Supplementary Information
Mann et al PNAS '08

S1. Proxy Data Set Details

Dendroclimatic data included a tree ring network of 105 maximum latewood density ("MXD") gridbox (5° latitude by 5° longitude) tree-ring composite series (Briffa et al, 1998;2001; Rutherford et al, 2005), 926 tree-ring series from the International Tree Ring Data Bank (see Section 'a' below for further details), and 5 additional tree-ring based series (local temperature reconstructions and regional composite chronologies). The proxy dataset also includes (see section 'b' below) 3 marine sediment series (from two locations), 14 speleothem series (from 7 locations), 19 lacustrine series (from 12 locations), 32 ice core series (from 26 locations), 15 marine coral series (from 10 locations) and 19 historical documentary series (from 15 locations). The dataset also includes 71 European composite surface temperature reconstructions back to AD 1500 based on a composite of proxy, historical, and early instrumental data (Luterbacher et al, 2004). All proxy data were required to have temporal resolution no coarser than decadal to facilitate meaningful calibration against the instrumental record. Note that multiple series were used from a given location when more than one proxy variable was available (e.g. ice accumulation and oxygen isotopes from a particular ice core). See Table S1 and supplementary

a) Full proxy database used for annual temperature reconstructions

	NH proxies					
	Annually resolved proxies			Annually+decadally resolved proxies		
interval	Dendro	Others	Total	Dendro	Others	Total
1800-1855	889	104	993	889	147	1036
1700-1799	563	95	658	563	137	700
1600-1699	276	91	367	276	132	408
1500-1599	150	89	239	150	127	277
1400-1499	105	18	123	105	46	151
1300-1399	62	16	78	62	40	102
1200-1299	41	12	53	41	36	77
1100-1199	29	12	41	29	34	63
1000-1099	18	12	30	18	28	46
900-999	15	10	25	15	25	40
800-899	13	9	22	13	24	37
700-799	11	8	19	11	23	34
600-699	11	8	19	11	22	33
500-599	7	7	14	7	21	28
400-499	7	7	14	7	21	28
300-399	6	7	13	6	21	27
200-299	4	6	10	4	20	24
100-199	3	6	9	3	20	23
0-99	2	6	8	2	19	21

	SH proxies					
	Annually resolved proxies			Annually+decadally resolved proxies		
interval	Dendro	Others	Total	Dendro	Others	Total
1800-1855	143	22	165	143	30	173
1700-1799	78	12	90	78	17	95
1600-1699	39	9	48	39	13	52
1500-1599	22	8	30	22	12	34
1400-1499	14	8	22	14	12	26
1300-1399	8	8	16	8	12	20
1200-1299	6	7	13	6	11	17
1100-1199	5	7	12	5	11	16
1000-1099	4	7	11	4	9	13
900-999	2	7	9	2	9	11
800-899	2	6	8	2	7	9
700-799	2	6	8	2	7	9
600-699	1	6	7	1	7	8
500-599	1	6	7	1	7	8
400-499	1	4	5	1	4	5
300-399	1	4	5	1	4	5
200-299	0	4	4	0	4	4
100-199	0	4	4	0	4	4
0-99	0	4	4	0	4	4

	Global proxies					
	Annually resolved proxies			Annually+decadally resolved pro		
interval	Dendro	Others	Total	Dendro	Others	Total
1800-1855	1032	126	1158	1032	177	1209
1700-1799	641	107	748	641	154	795
1600-1699	315	100	415	315	145	460
1500-1599	172	97	269	172	139	311
1400-1499	119	26	145	119	58	177
1300-1399	70	24	94	70	52	122
1200-1299	47	19	66	47	47	94
1100-1199	34	19	53	34	45	79
1000-1099	22	19	41	22	37	59
900-999	17	17	34	17	34	51
800-899	15	15	30	15	31	46
700-799	13	14	27	13	30	43
600-699	12	14	26	12	29	41
500-599	8	13	21	8	28	36
400-499	8	11	19	8	25	33
300-399	7	11	18	7	25	32
200-299	4	10	14	4	24	28
100-199	3	10	13	3	24	27
0-99	2	10	12	2	23	25

b) Screened proxy database used for annual temperature reconstructions

	NH proxies					
	Annually resolved proxies			Annually+decadally resolved proxies		
Interval	Dendro	Others	Total	Dendro	Others	Total
1800-1855	305	98	403	305	115	420
1700-1799	225	90	315	225	107	332
1600-1699	117	86	203	117	102	219
1500-1599	62	85	147	62	100	162
1400-1499	41	14	55	41	23	64
1300-1399	18	12	30	18	20	38
1200-1299	12	9	21	12	17	29
1100-1199	7	9	16	7	17	24
1000-1099	3	9	12	3	16	19
900-999	2	7	9	2	13	15
800-899	2	6	8	2	12	14
700-799	2	6	8	2	12	14
600-699	2	6	8	2	11	13
500-599	2	5	7	2	10	12
400-499	2	5	7	2	10	12
300-399	1	5	6	1	10	11
200-299	1	5	6	1	10	11
100-199	1	5	6	1	10	11
0-99	0	5	5	0	10	10

	SH proxies					
	Annually resolved proxies			Annually+decadally resolved proxies		
Interval	Dendro	Others	Total	Dendro	Others	Total
1800-1855	50	14	64	50	14	64
1700-1799	31	9	40	31	9	40
1600-1699	13	7	20	13	7	20
1500-1599	8	6	14	8	6	14
1400-1499	5	6	11	5	6	11
1300-1399	2	6	8	2	6	8
1200-1299	1	5	6	1	5	6
1100-1199	1	5	6	1	5	6
1000-1099	1	5	6	1	5	6
900-999	0	5	5	0	5	5
800-899	0	4	4	0	4	4
700-799	0	4	4	0	4	4
600-699	0	4	4	0	4	4
500-599	0	4	4	0	4	4
400-499	0	3	3	0	3	3
300-399	0	3	3	0	3	3
200-299	0	3	3	0	3	3
100-199	0	3	3	0	3	3
0-99	0	3	3	0	3	3

	Global proxies					
	Annually resolved proxies			Annually+decadally resolved prox		
Interval	Dendro	Others	Total	Dendro	Others	Total
1800-1855	355	112	467	355	129	484
1700-1799	256	99	355	256	116	372
1600-1699	130	93	223	130	109	239
1500-1599	70	91	161	70	106	176
1400-1499	46	20	66	46	29	75
1300-1399	20	18	38	20	26	46
1200-1299	13	14	27	13	22	35
1100-1199	8	14	22	8	22	30
1000-1099	4	14	18	4	21	25
900-999	2	12	14	2	18	20
800-899	2	10	12	2	16	18
700-799	2	10	12	2	16	18
600-699	2	10	12	2	15	17
500-599	2	9	11	2	14	16
400-499	2	8	10	2	13	15
300-399	1	8	9	1	13	14
200-299	1	8	9	1	13	14
100-199	1	8	9	1	13	14
0-99	0	8	8	0	13	13

Note: Dendro = tree ring + MXD

spreadsheet ("1290proxynames.xls") for additional details and sources for proxy data used.

Table S1: Number of annually, decadally, and total proxy data available from different sources for different starting years for (a) Full proxy data set and (b) screened proxy dataset.

A. Selection of Tree ring Data

Tree-ring data included 926 tree-ring series extracted from the International Tree Ring Data Bank ("ITRDB", version 5.03: http://www.ncdc.noaa.gov/paleo/treering.html). All ITRDB tree-ring proxy series were required to pass a series of minimum standards to be included in the network: (a) series must cover at least the interval 1750 to 1970, (b) correlation between individual cores for a given site must be greater than 0.50 for this period, (c) there must be at least eight samples during the screened period 1800-1960 and for every year used.

Series that were redundant with other compilations (e.g. series used in the aforementioned MXD network) were not included. Four other series were not included because of site-specific problems (e.g. the "ar054" series because of earthquake effects after 1811). Of the remaining series, 139 had to be replaced because of format errors in the chronology file on the ITRDB (136), or because sample depth data were missing from the chronology file (57). As a result 926 series derived from data at the ITRDB were used in total. When sample depth data were absent, the raw ring-width data from ITRDB were used to recalculate the chronology using program ARSTAN (Version 6.05P), with the following settings: a) a single detrending fitting a cubic spline of 50% Variance Reduction Function at 66.6.% the length of each sample, no stabilization of variance or autoregressive modeling, indices computed as ratio, that is measurement divided by curve, and chronology calculated as bi-weight robust mean estimate. Series were required to be complete over most of the calibration interval: 25%, 50%, and 75% of the ITRDB series used end in 1979, 1984, and 1991, while 105 of the 926 series end in 1995 or later. Missing values were infilled as described in section S3 below.

Tree-ring data are derived from the means of multiple detrended sample measurement series. The more flexible the detrending function used on each sample series, the greater the risk of discarding low-frequency climate information. Conversely, stiffer detrending functions result in a more conservative detrending, retaining lower-frequency climate signal, but possibly also non-climatic information on similar time scales. Examples of such a conservative treatment would include taking the ratio of each year's tree ring value and that year's value for a fitted modified negative exponential function, of a cubic spline with a 50% variance reduction function at 2/3 the length of the measurement series. While the tree-ring data series we have used are all based on well replicated chronologies, and while the great majority of series used result from relatively conservative treatment of the kinds described above, even the most conservative standardization techniques necessarily limit the reliability of tree-ring reconstructions on timescales longer than a few centuries. Cook et al. (1995) wrote "the maximum length of recoverable climate information is ordinarily related to the lengths of the individual tree-ring series used to reconstruct the millennia-long chronology". In our case, 75% of 926 ITRDB tree-ring series have a median segment length (Figure S1) of 299 years or less. For the 40 (16) Northern Hemisphere records used in reconstructions starting at AD1200 (AD1000), 75% have a median segment length of 603 (739) years or less. Thus the capacity of the tree-ring chronologies used for periods before AD1200 to capture multi-centennial variation is greater than for the whole ITRDB data set, but is none-the-less limited. The gridded maximum density data set was developed differently by Osborn et al. (as used by Rutherford et al, 2005) with the specific goal of capturing century to multi-century variability. In addition to the problem of segment length, there are potentially severe, time-dependent biases in tree-ring series, especially at their beginning and end

(Melvin, 2004). Of the five tree-ring series included in the dataset that were not obtained directly from the ITRDB or from the authors of the gridded maximum density data set, two were standardized using the spline approach mentioned above (Cook et al’s New Zealand Reconstruction and the D’Arrigo et al Mongolian data) and three with the Regional Curve Standardization approach (Cook et al’s Tasmania reconstruction, Briffa et al’s Tornetrask data and Naurzbaev and Vaganov’s reconstruction of early summer temperature on the Taimyr peninsula). In the latter case the RCS approach was modified by whitening the records to remove persistence.

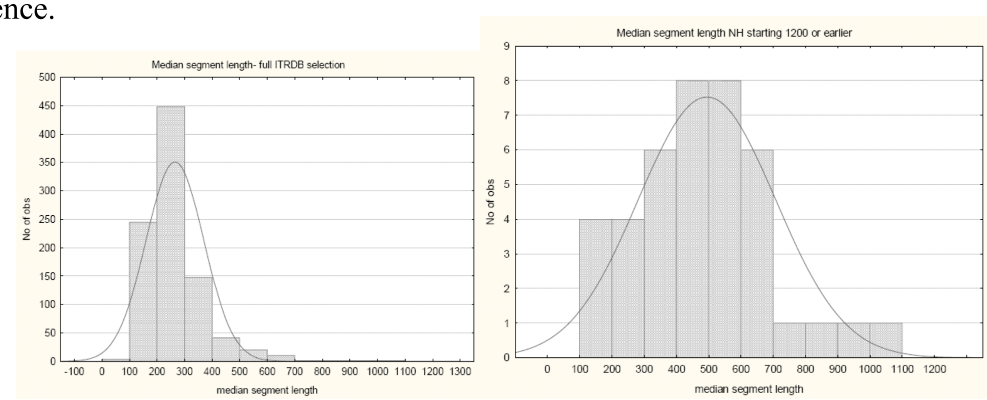


Figure S1: Histogram of the median segment lengths of the 926 ITRDB chronologies (left panel) and of those from the Northern Hemisphere starting in AD 1200 or earlier (right panel).

<u>References:</u>

Cook, E.R., Briffa, K.R., Meko, D.M., Graybill, D.A. and Funkhouser, G. 1995: The `segment length curse' in long tree-ring chronology development for palaeoclimatic studies, The Holocene, Vol. 5, No. 2, 229-237 (1995)

Melvin, T., 2004 Historical growth rates and changing climatic sensitivity of boreal conifers. Ph.D. thesis, Climatic Research Unit, East Anglia, UK.

Rutherford, S., Mann, M.E., Osborn, T.J., Bradley, R.S., Briffa, K.R., Hughes, M.K., Jones, P.D., Proxy-based Northern Hemisphere Surface Temperature Reconstructions: Sensitivity to Methodology, Predictor Network, Target Season and Target Domain, *Journal of Climate*, 18, 2308-2329, 2005.

B. Selection of Non-Tree ring Data

We selected non-tree ring proxies with the goal of obtaining data from as wide a geographic area as possible, but at the same time we tried to limit the selection to records that were reasonably well-dated and where the original analysts had shown that there was a paleoclimatic signal associated with the proxy. Given the diversity of non-tree-ring proxies, this is a much more subjective process than that which we were able to apply to the tree-ring data. One advantage of the non-tree ring proxy series used is that in most cases there is little reason to believe *a priori* that there are any problems with the series that are likely to eliminate the reliability of multi-century to millennial timescale information.

S2. Pre-Processing of Proxy Data

When records were available at sub-annual resolution, they were averaged to obtain annual mean values. To avoid aliasing bias, records with only decadal resolution were first interpolated to annual resolution, and then low-pass filtered to retain frequencies $f<0.05$ cycle/year (the Nyquist frequency for decadal sampling). The year 1995, at or shortly after which many proxy data terminate, was used as the upper limit for calibration. Due to the evidence for loss of temperature sensitivity after about 1960 (Briffa et al, 2001), MXD data were eliminated for the post-1960 interval. The RegEM algorithm of Schneider (2001) was used to estimate missing values for proxy series terminating prior to the 1995 calibration interval endpoint, based on their mutual covariance with the other available proxy data over the full 1850-1995 calibration interval. No instrumental or historical (i.e., Luterbacher et al) data were used in this procedure.

European temperature reconstructions of Luterbacher et al (2004) available at 0.5° latitude/longitude resolution and at monthly (AD 1659-1998) or seasonal (AD 1500-1658) temporal resolution were upscaled to a 5° grid and annually averaged (Jan-Dec) to yield 71 distinct annual gridbox series back to AD 1500.

References:

Luterbacher, J., D. Dietrich, Xoplaki, E., Grosjean, M. and H. Wanner, "European Seasonal and Annual Temperature Variability, Trends, and Extremes since 1500." *Science* 303: 1499-1503, 2004.

S3. Screening Procedure

To pass screening, a series was required to exhibit a statistically significant ($p<0.10$) correlation with either one of the two closest instrumental surface temperature gridpoints over the calibration interval (though see discussion below about the influence of temporal autocorrelation which reduces the effective criterion to roughly $p<0.13$ in the case of correlations at the annual timescales).

The screening process was performed on annual timescales for annually-resolved proxies, and on decadal timescales for decadally-resolved proxies. Screening was performed separately for the full available overlap interval between proxy and instrumental data (1850-1995) and the shorter calibration intervals (1896-1995 and 1850-1949) used for validation experiments. We assumed $n=144$ nominal degrees of freedom over the 1850-1995 (146 year) interval for correlations between annually resolved records (though see discussion below about influence of temporal autocorrelation at the annual timescales), and $n=13$ degrees of freedom for decadal resolution records. The corresponding one-sided $p=0.10$ significance thresholds are $|r|=0.11$ and $|r|=0.34$ respectively. For the shorter (100 year) calibration intervals, we assumed $n=98$ nominal degrees of freedom for annually resolved records, and $n=8$ degrees of freedom for decadal resolution records. The corresponding one-sided $p=0.10$ significance thresholds are $|r|=0.13$ and $|r|=0.42$ respectively. Owing to reduced degrees of freedom arising from modest temporal autocorrelation, the effective p value for annual screening is slightly higher ($p\sim0.128$) than the nominal (p=0.10) value. For the decadally-resolved proxies, the effect is negligible since the decadal timescale of the smoothing is long compared to the intrinsic autocorrelation timescales of the data.

While 484 (roughly 40%) pass the temperature screening process over the full (1850-1995) calibration interval, one would expect that no more than about 150 (less than 13%) of the proxy series would pass the screening procedure described above by chance alone. This observation indicates that selection bias, while potentially problematic when employing screened predictors (see e.g. Schneider, 2007), does not appear a significant problem in our case.

References:

Schneider, T., 2007: Uncertainty in climate sensitivity estimates, *Nature*, 446, E1-E2.

S4. Instrumental Surface Temperature Dataset

Gaps in the individual annual mean (Jan-Dec) gridbox surface temperature data available from 1850-2006 were infilled using the RegEM procedure with ridge regression described by Schneider (2001) and described in Rutherford et al (2003). A small number of grid boxes were not infilled due to the availability of too few monthly values in the raw data (see Figure S2).

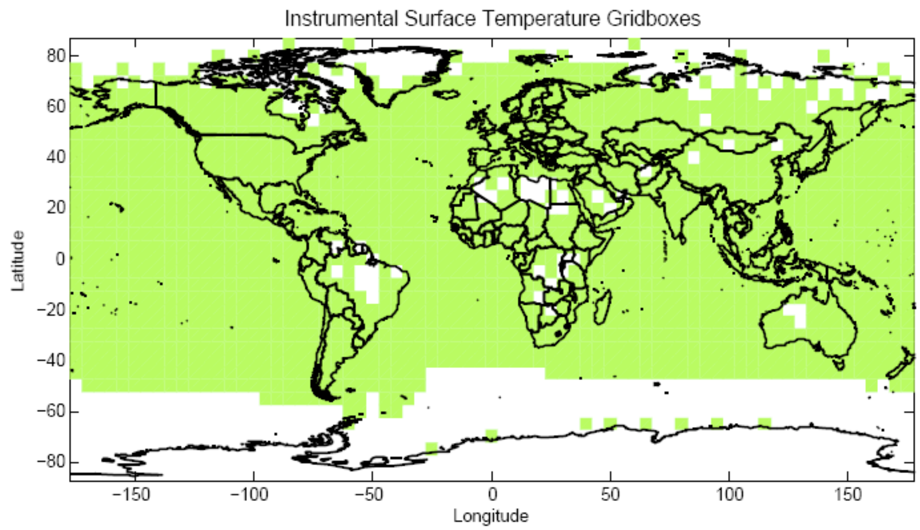


Figure S2: Distribution of infilled instrumental surface temperature gridboxes used in study.

S5. Statistical Validation Results

See supplementary spreadsheets ("eiv-validation.xls" and "cps-validation.xls") for full information from validation experiments, including the individual results of early and late validation experiments, and validation results for the networks available at the start of each century. Results are also plotted below for the specific NH, SH, and global mean validation experiments described in the manuscript.

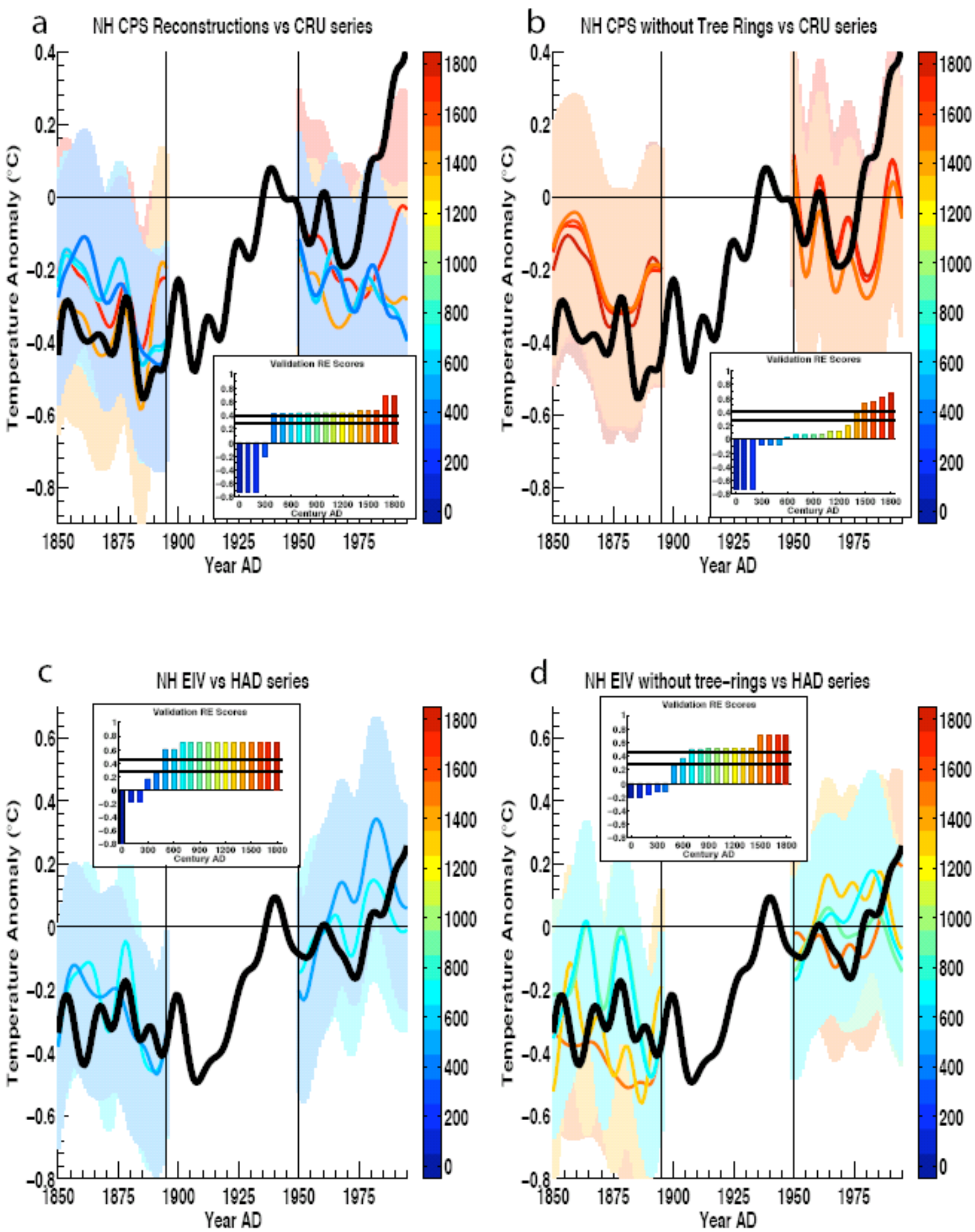


Figure S3a: Figure 2 of manuscript, but showing EIV results (c-d) for NH land+ocean

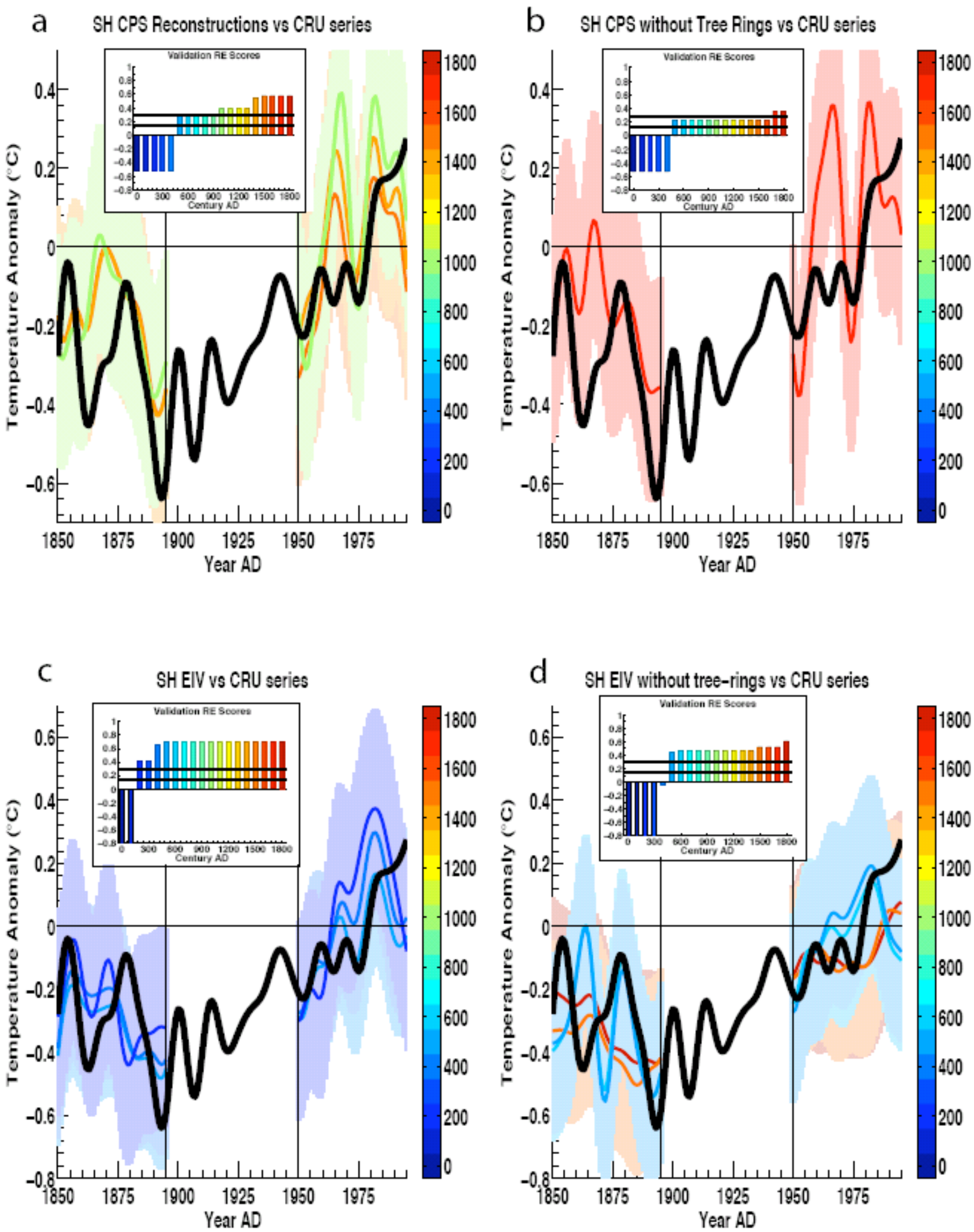


Figure S3b: As in Figure 2 of manuscript, but for SH land (a-d)

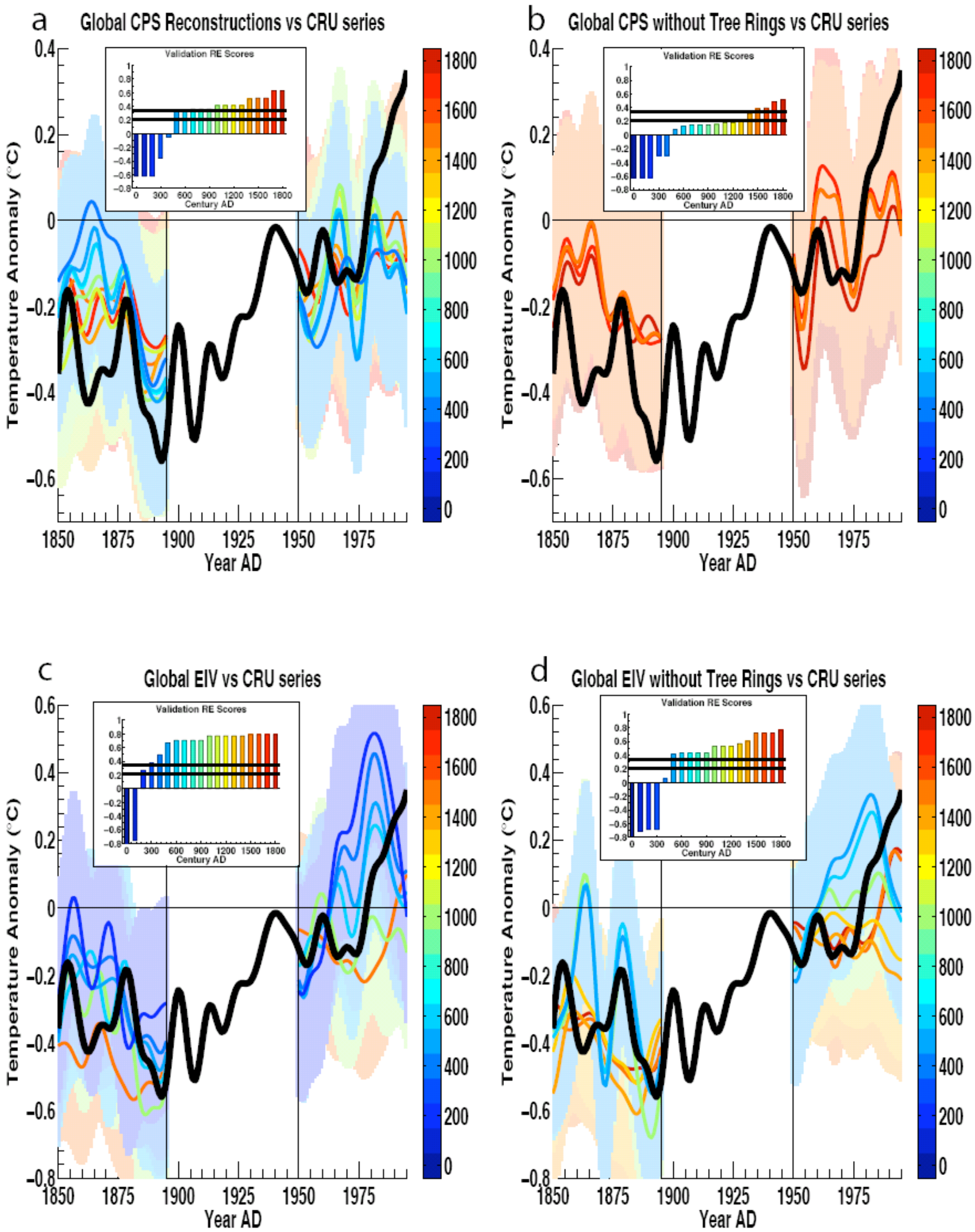


Figure S3c: As in Figure 2 of manuscript, but for global land (a-d)

S6. Supplementary Reconstruction Results

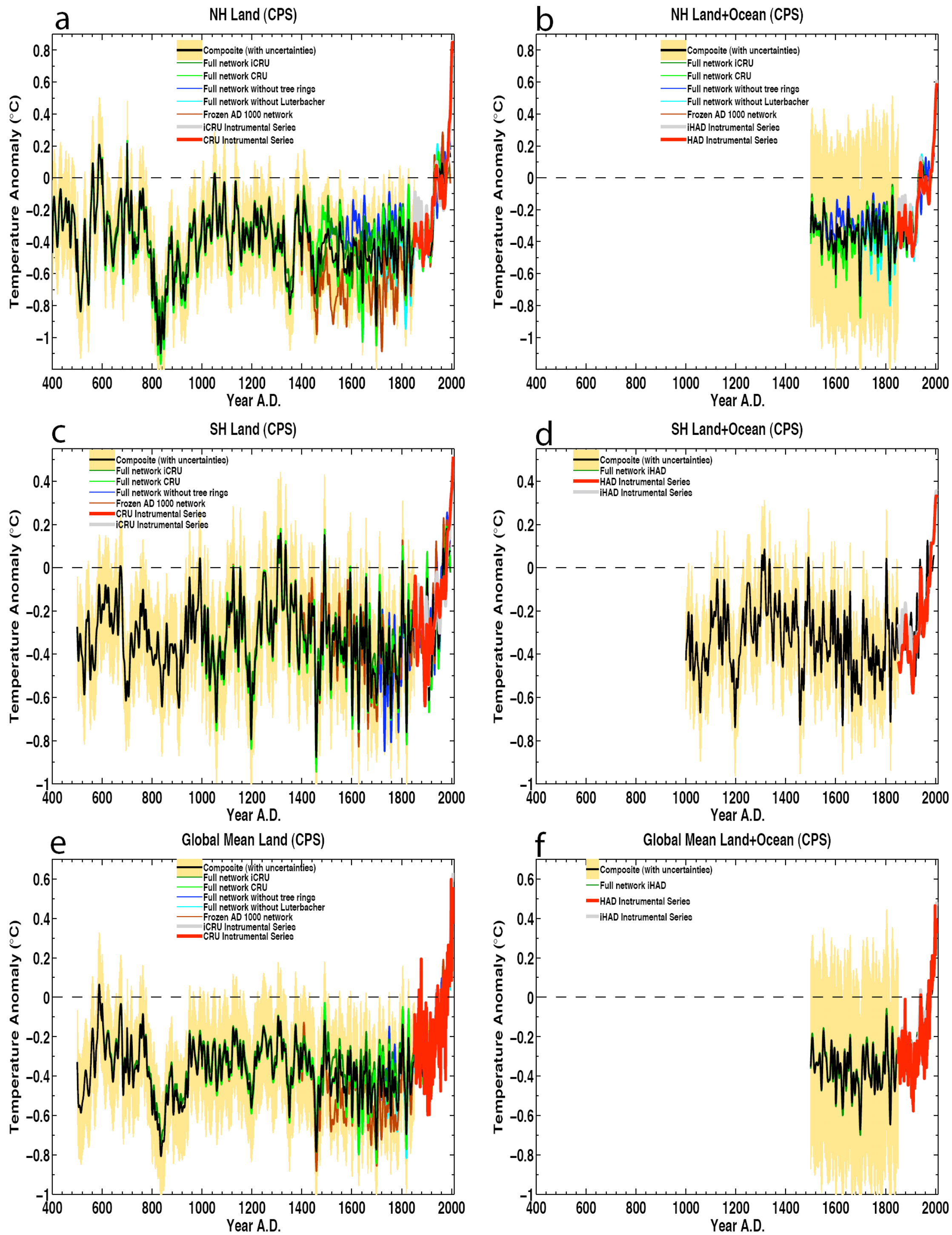


Figure S4: Comparison of NH mean (*a*: land; *b*: combined land+ocean), SH land/land+ocean mean (*c*-*d*) and Global mean (*e*-*f*) Surface Temperature Reconstructions based on CPS method, and using the various proxy networks and target instrumental series described in text. Only reconstructions that passed validation are shown. Shown also in each panel also are composites of the multiple series shown, and their associated 95% confidence intervals. Decadally-smoothed CRU land and Had Land+ocean instrumental series (1850-2006) are shown for comparison.

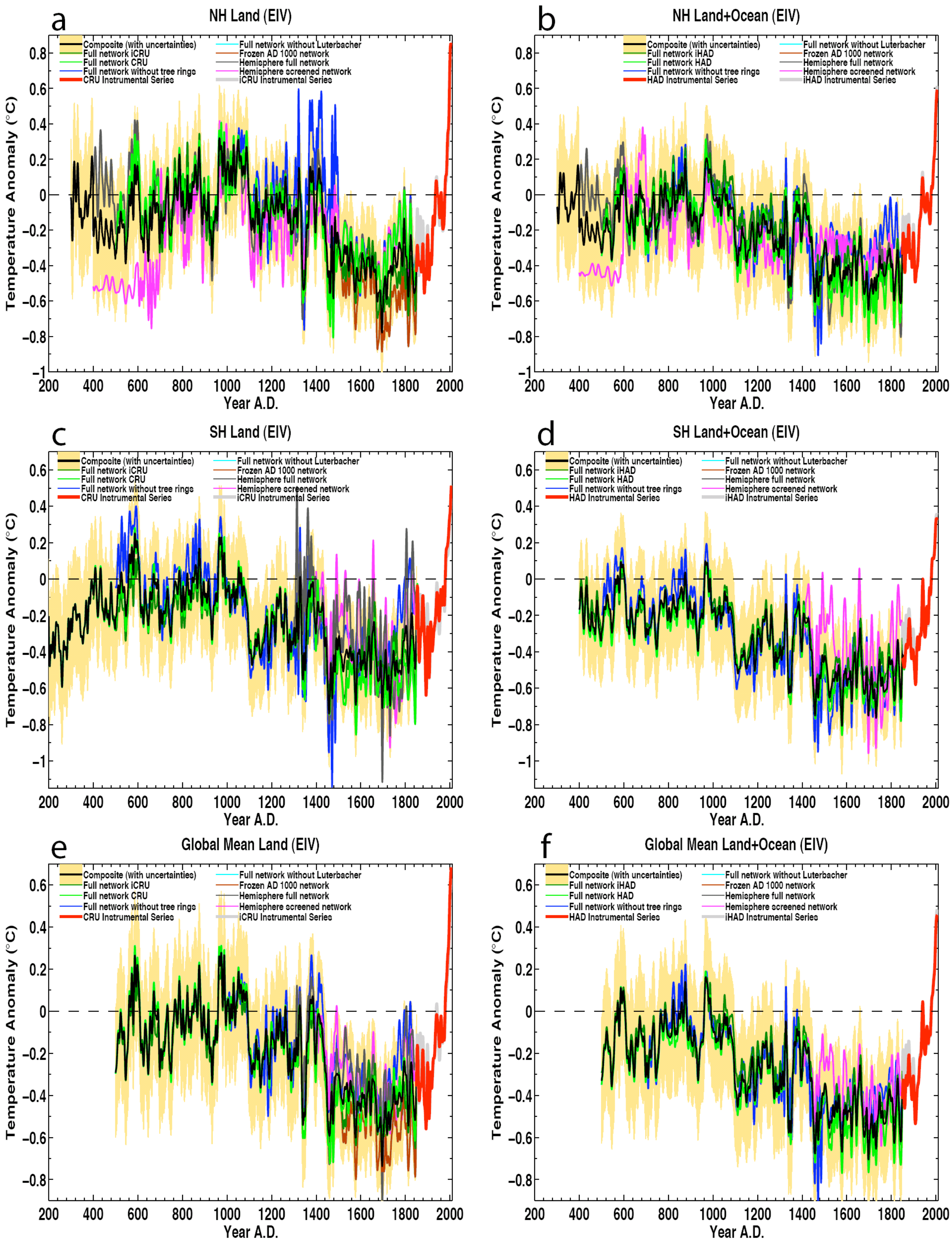


Figure S5: Comparison of NH mean (*a*: land; *b*: combined land+ocean), SH land/combined land+ocean mean (*c*-*d*) and Global mean (*e*-*f*) Surface Temperature Reconstructions based on EIV method, and using the various proxy networks and target instrumental series described in text. Details are as in Figure S4.

S7. Sensitivity Analyses (NH temperatures)

A. Sensitivity to Use of Tree-Ring Data in Early Centuries

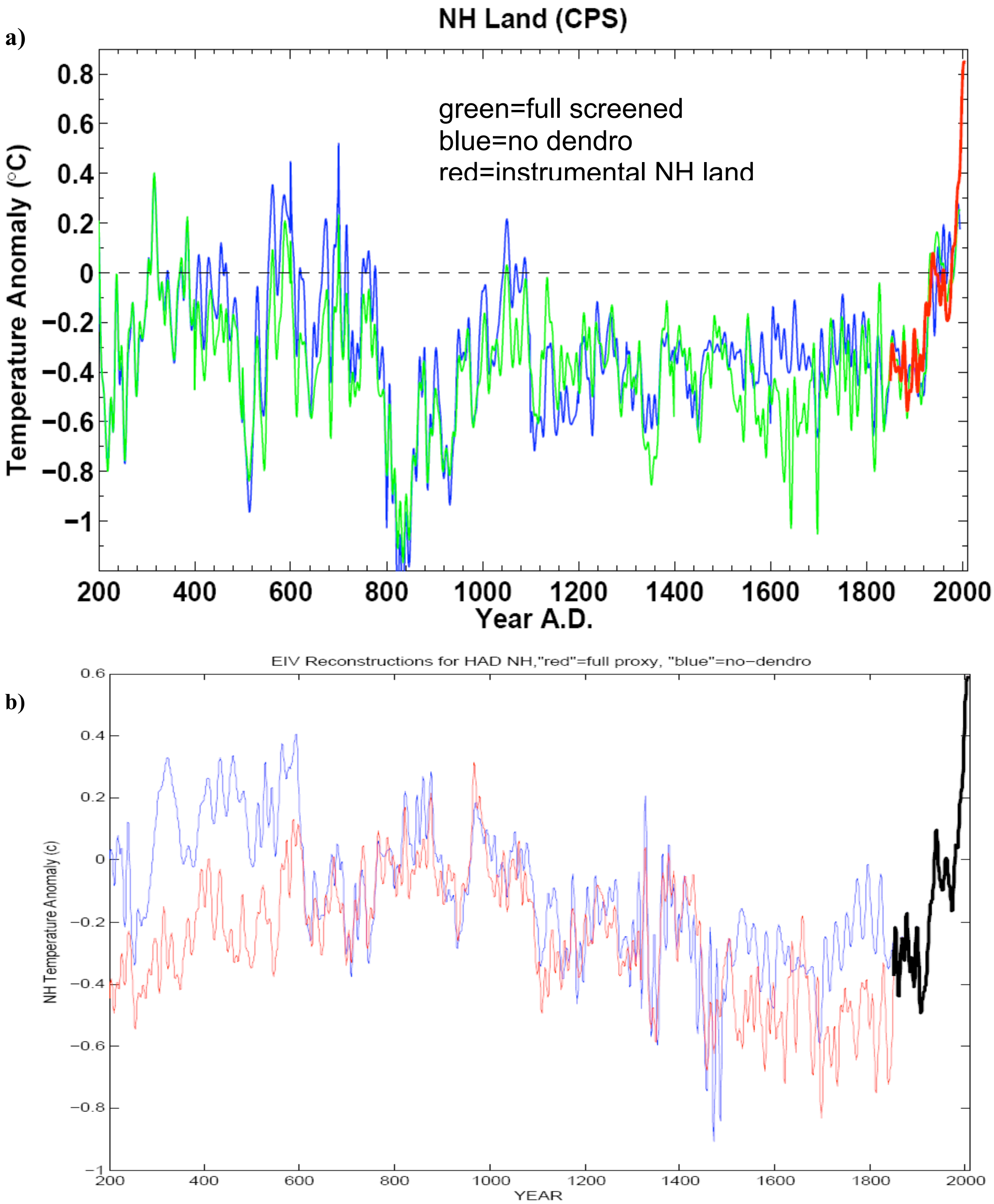


Figure S6: Comparison of long-term (a) CPS NH land and (b) EIV NH land+ocean (full global proxy network) with and without using tree-ring data [note e.g. from Fig S5B that the "full proxy network" ("no dendro" network) reconstructions are not skillful prior to AD 500 (prior to AD 700). The point of this comparison is simply that the anomalousness of recent warming in the long-term context of the reconstructions is not dependent on whether or not tree-ring data have been used.]

B. Potential Data Quality Problems

In addition to checking whether or not potential problems specific to tree ring data have any significant impact on our reconstructions in earlier centuries (see Figure S6), we also examined whether or not potential problems noted for several records (see supplementary database spreadsheet file mentioned in section S1 for details) might compromise the reconstructions. These records include the 4 Tijander et al (2003) series used (see Figure S8) for which the original authors note that human effects over the past few centuries unrelated to climate might impact records (the original paper states "*Natural variability in the sediment record was disrupted by increased human impact in the catchment area at AD 1720*." and later, "*In the case of Lake Korttajarvi it is a demanding task to calibrate the physical varve data we have collected against meteorological data, because human impacts have distorted the natural signal to varying extents"*). These issues are particularly significant because there are few proxy records, particularly in the temperature-screened data set (see Figure S8), available back through the 9th century. The Tijander et al (2003) series constitute 4 of the 15 available Northern Hemisphere records prior to that point.

In addition there are 3 other records in our database with potential data quality problems, as noted in the database notes:
Benson et al (Mono Lake): *"Data after 1940 no good—water exported to CA"*
Isdale 2001 (fluorescence): *"anthropogenic influence after 1870"*
McCulloch 2003 (Ba/Ca) : *"anthropogenic influence after 1870"*

We therefore performed additional analyses as in Figure S6, but instead comparing the reconstructions both with and without the above 7 potentially problematic series, as shown below.

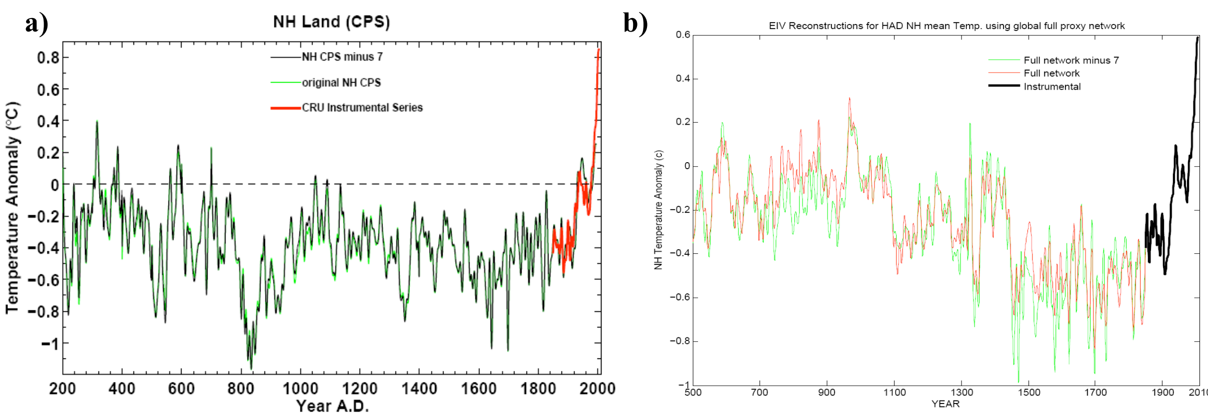


Figure S7: Comparison of long-term (a) CPS NH land and (b) EIV NH land+ocean reconstructions (full global proxy network) both with and without the 7 potentially problematic series discussed above.

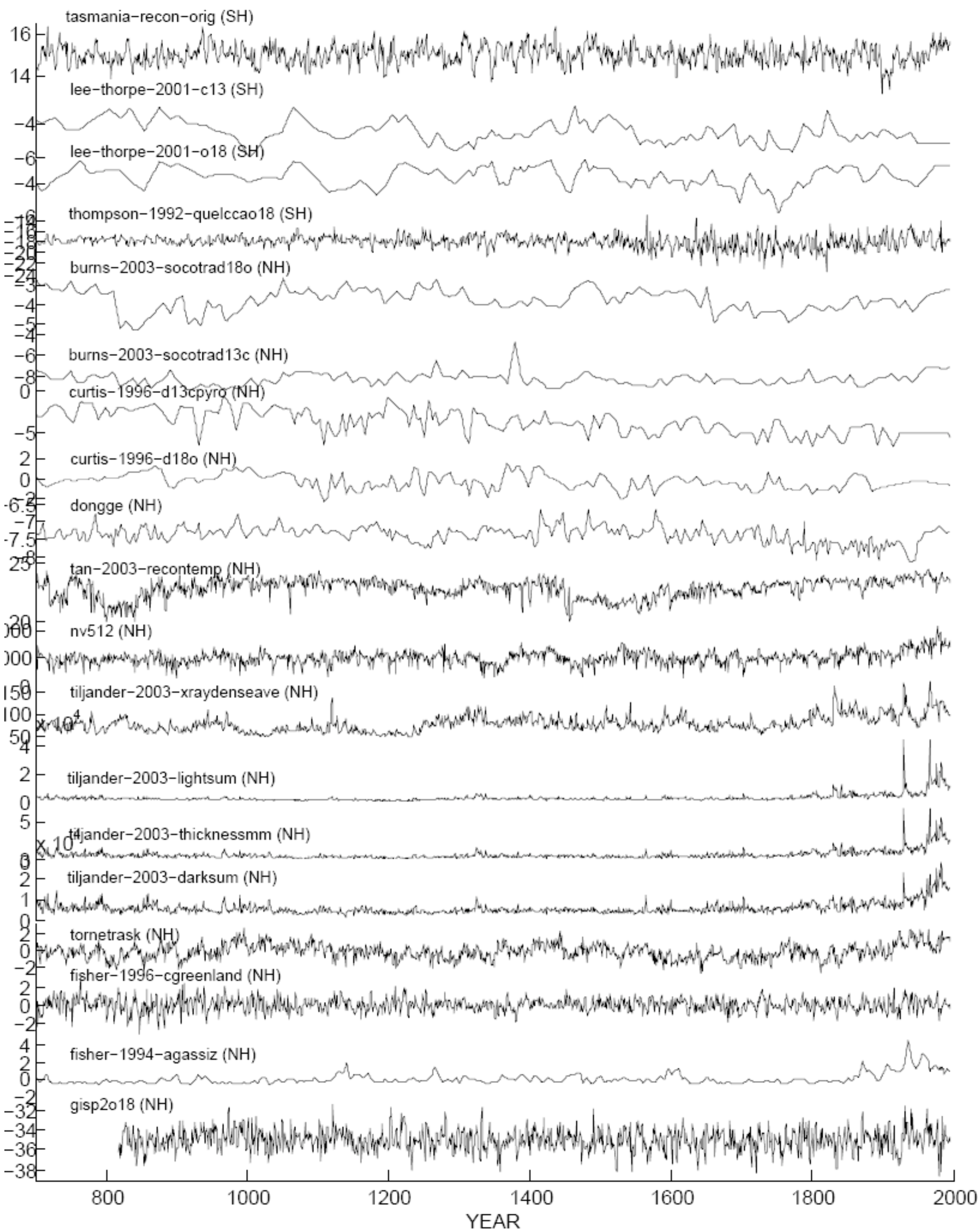


Figure S8: Plots of the 19 (decadally-smoothed) proxy records that pass the screening procedure back to at least AD 818, including 4 records from the Southern Hemisphere (labeled as 'SH') and 15 records from the Northern Hemisphere (labeled as 'NH').

C. Leverage of Individual Records (CPS approach)

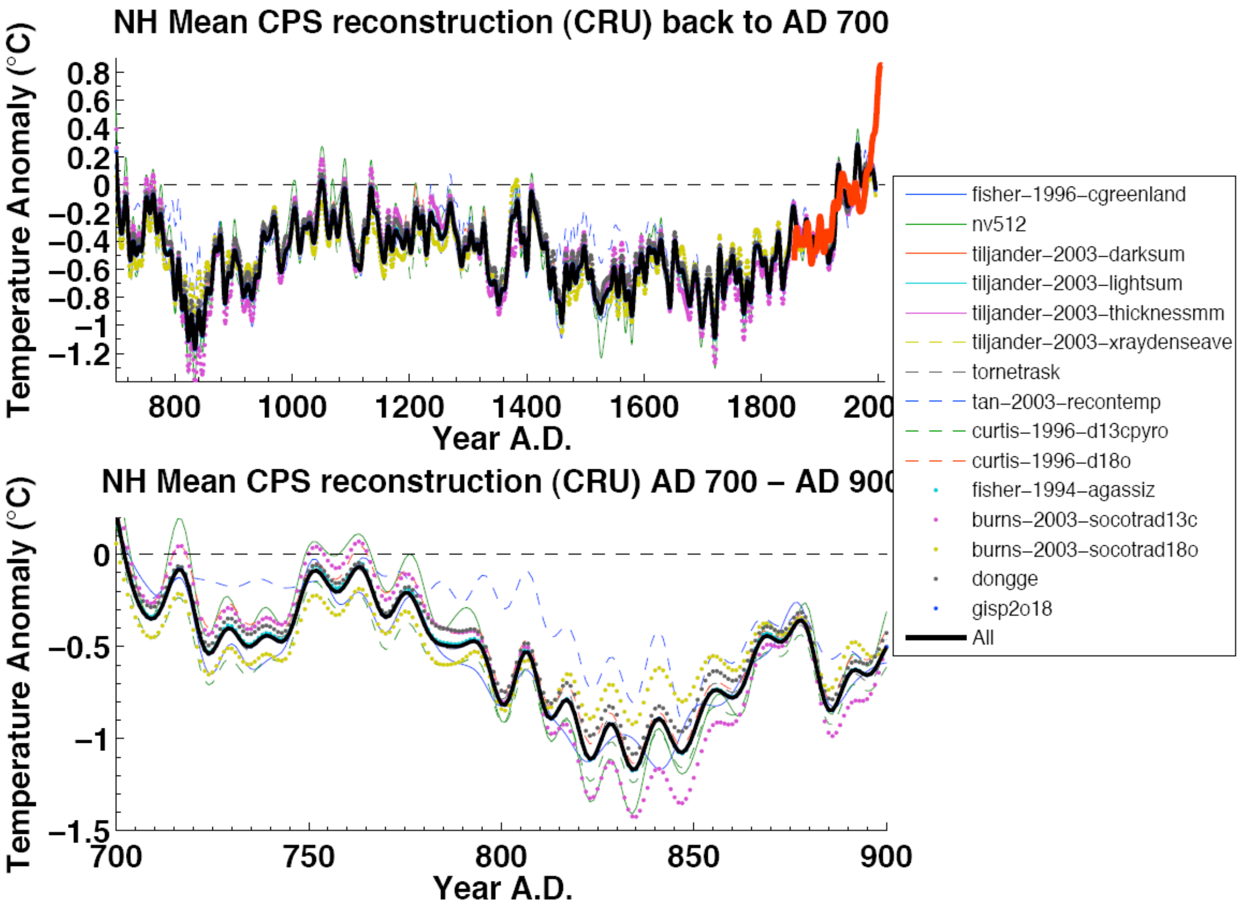
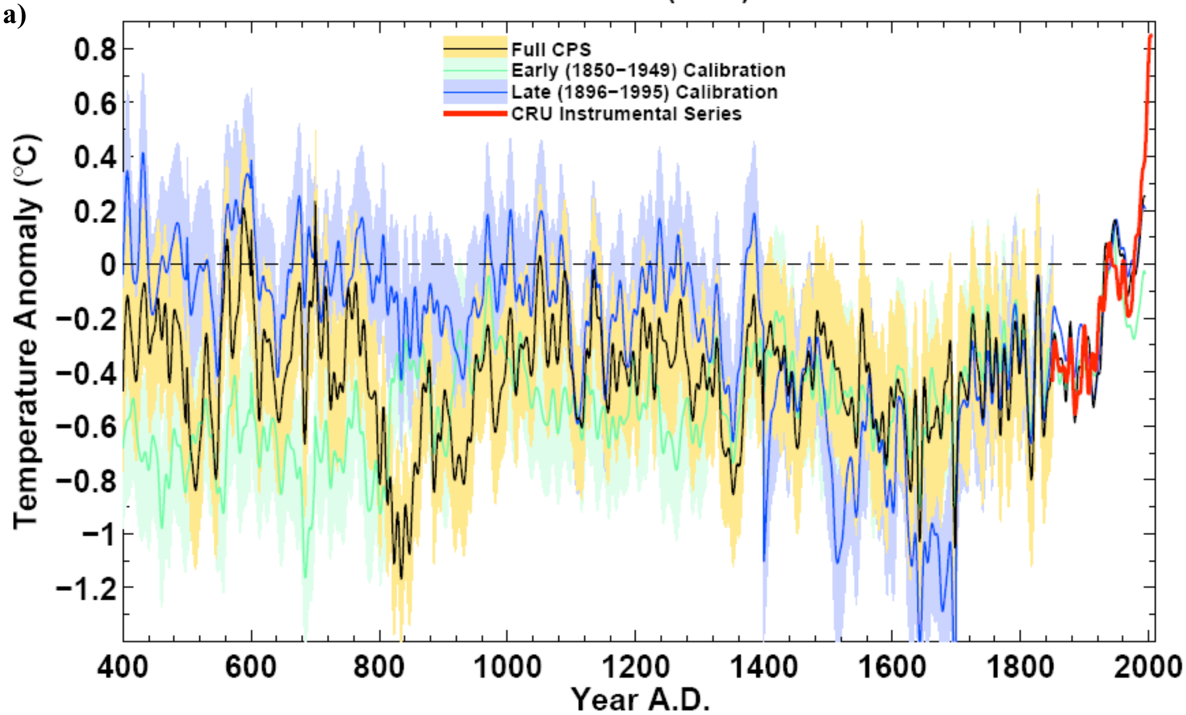


Figure S9: Plots of the NH CPS Land reconstruction withholding each of the 15 available Northern Hemisphere screened proxy records available back to the early 9th century as indicated.

D. Robustness of Reconstructions with respect to Calibration Period

a)



b)

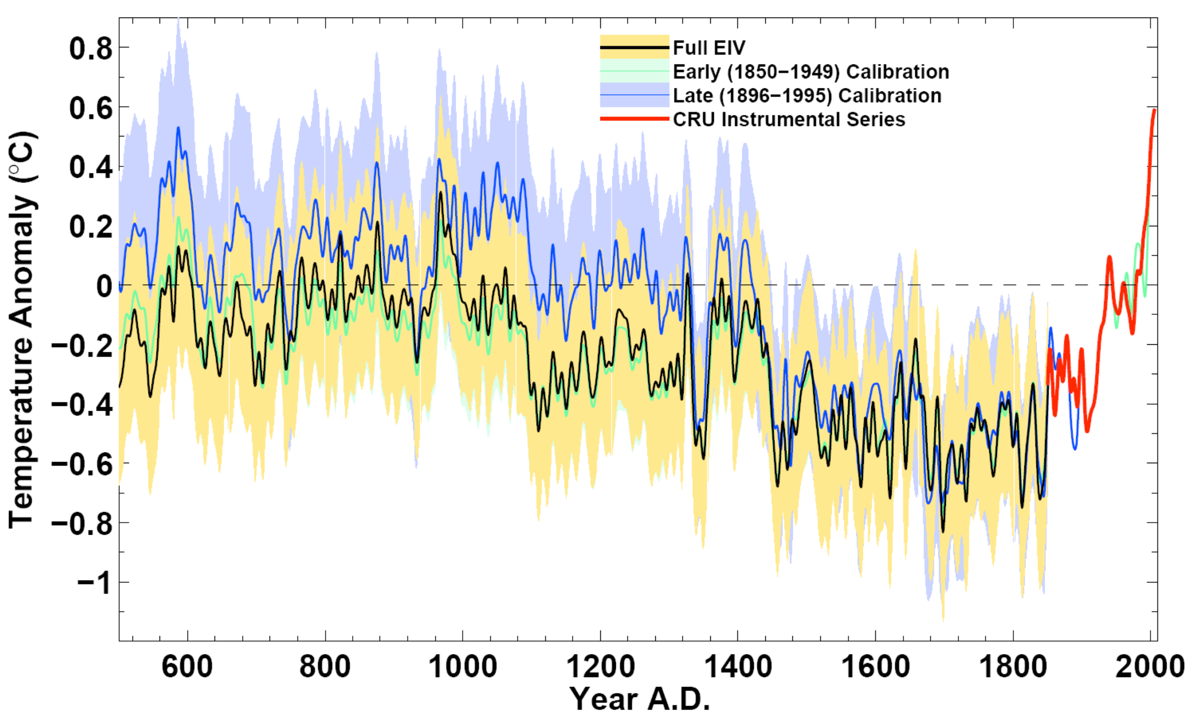


Figure S10: Plots of the NH (a) CPS Land reconstruction and (b) EIV Land+ocean (full global proxy network) that result from full period calibration, and early and late calibration intervals used in validation experiments.

S8. Analyses using synthetic ‘pseudoproxy’ networks

A. comparison of EIV and CPS w/ networks of varying sparseness

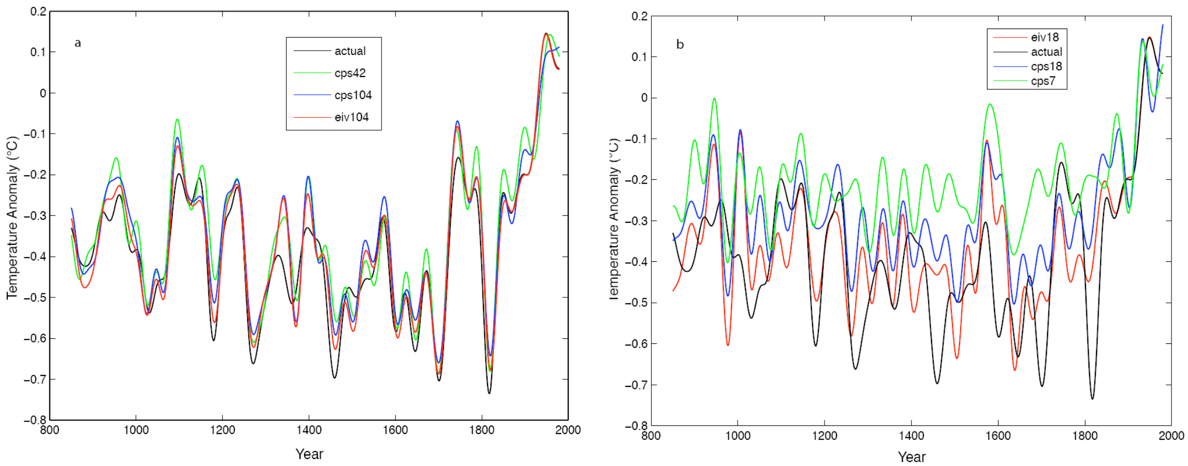


Figure S11: Comparisons of CPS and EIV NH mean reconstructions using synthetic ‘pseudoproxy’ networks calibrated over the 1856-1980 interval of a model simulation from AD 850-1999 using the NCAR CSM coupled model [see Mann et al, 2005; 2007]. Results are shown for (a) a large (104 pseudoproxy), moderate signal-to-noise ratio (SNR=0.4) proxy network and (b) a smaller (18 pseudoproxy), low signal-to-noise ratio (SNR=0.25) proxy network. The pseudoproxy networks are designed to have attributes similar to proxy networks used in the reconstructions described in the manuscript: estimated SNR and network sizes for ‘a’ and ‘b’ are similar to those available for the actual proxy networks used in the present study back to AD 1300 and AD 0 respectively. Parallel CPS analyses are performed for networks using a random subset of 40% of the pseudoproxy series (i.e. 42 of the 104 pseudoproxies for ‘a’ and 7 of the 18 pseudoproxies for ‘b’) to reflect the smaller number (approximately 40% of total) of proxies available for the ‘screened’ network used in CPS as in the present study (all series smoothed to emphasize 40 year and longer timescales).

a. SNR=0.4	**RE**	**CE**	**r^2**
CPS (42)	0.92	0.61	0.71
CPS (104)	0.94	0.70	0.76
EIV (104)	0.95	0.74	0.75
b. SNR=0.25			
CPS (7)	0.55	-1.18	0.09
CPS (18)	0.68	-0.56	0.16
EIV (18)	0.76	-0.15	0.16

Table S2: Decadal validation Scores (based on full pre-calibration interval AD 850-1855) corresponding to NH reconstructions shown in Figure S11 ‘a’ and ‘b’. Provided are the standard RE and CE scores favored for skill evaluation. r^2 scores are provided for comparison, despite their demonstrated limited utility in skill evaluation (see Mann et al, 2007 for skill metric definitions and other details).

References:

Mann, M.E., Rutherford, S., Wahl, E., Ammann, C., Robustness of Proxy-Based Climate Field Reconstruction Methods, *J. Geophys. Res.*, 112, D12109, doi: 10.1029/2006JD008272, 2007.

Mann, M.E., Rutherford, S., Wahl, E., Ammann, C., Testing the Fidelity of Methods Used in Proxy-based Reconstructions of Past Climate, *Journal of Climate*, 18, 4097-4107, 2005.

B. Alternative CPS scaling #1 – trend matching

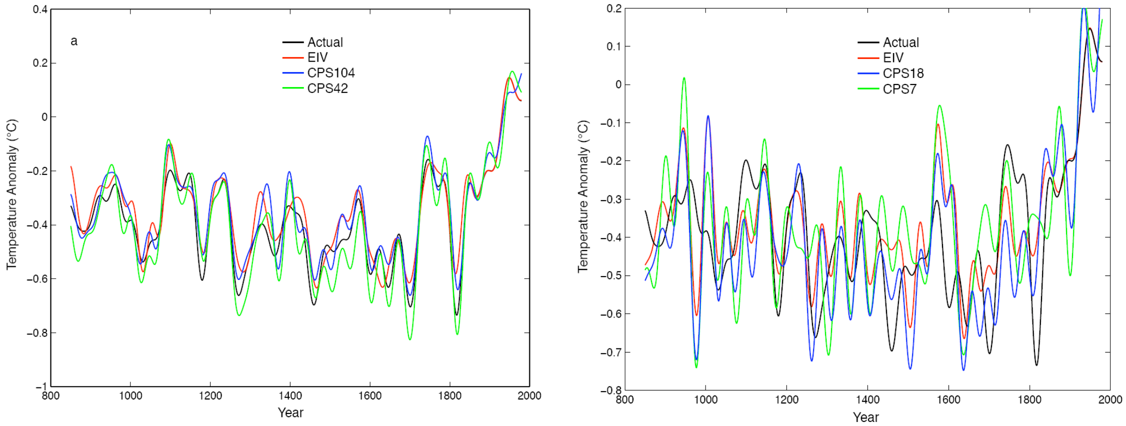


Figure S12: As in Figure S11, but where CPS scaling involves matching trend of composite to that for the target series over the (1856-1980) calibration interval.

a. SNR=0.4	**RE**	**CE**	**r^2**
CPS (42)	0.89	0.46	0.71
CPS (104)	0.94	0.69	0.76
b. SNR=0.25			
CPS (7)	0.36	-2.11	0.09
CPS (18)	0.49	-1.47	0.16

Table S3: As table S2, but where CPS scaling involves trend matching (see Figure S12)

C. Alternative CPS scaling #2 – signal amplitude matching

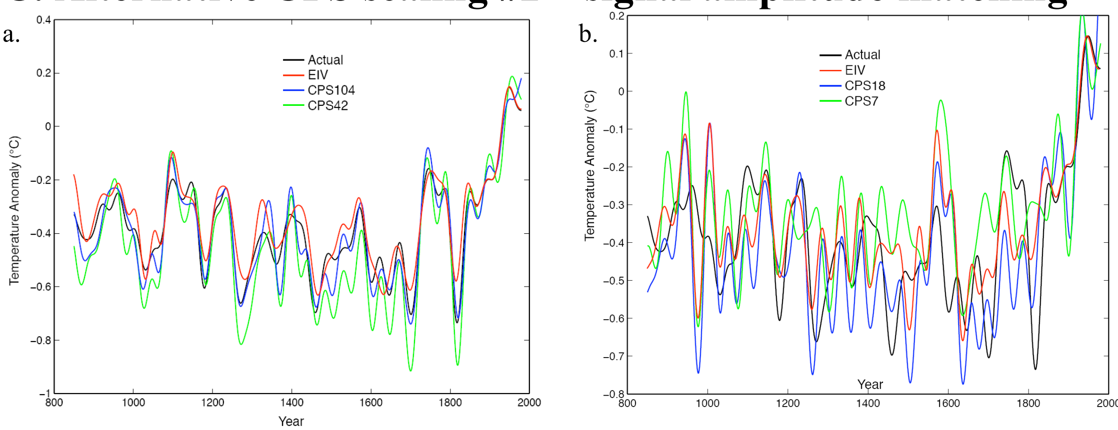
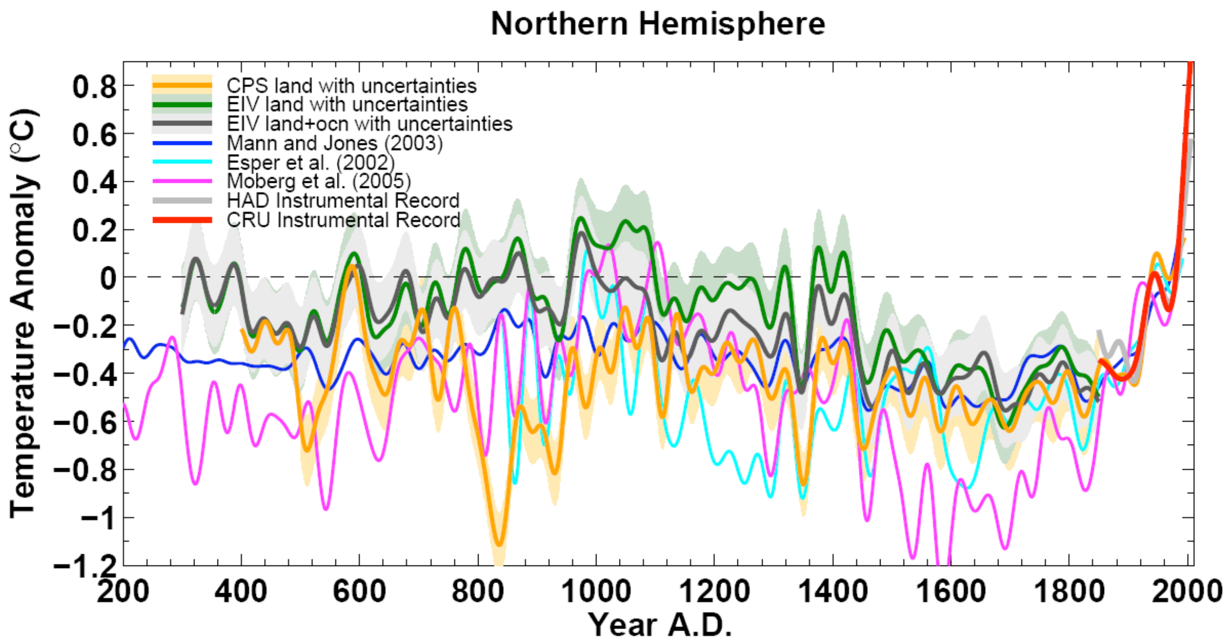


Figure S13: As in Figure S11, but where CPS scaling is based an alternative estimate of the "signal amplitude" of the reconstruction [if the calibration period correlation between the true NH mean temperature series and the proxy reconstruction is r, then the conventional (variance-matched) CPS reconstruction is rescaled by $1/r$ so that the estimated *signal component of the record* has the same decadal variance as the target series over the (1856-1980) calibration interval.

a. SNR=0.4	**RE**	**CE**	**r^2**
CPS (42)	0.88	0.44	0.84
CPS (104)	0.92	0.63	0.87
b. SNR=0.25			
CPS (7)	0.48	-1.55	0.29
CPS (18)	0.45	-1.67	0.40

Table S4: As table S2, but where CPS scaling is based on estimated signal-to-noise ratio (see Figure S13).

S9. Comparisons of various NH reconstructions using alternative scaling conventions



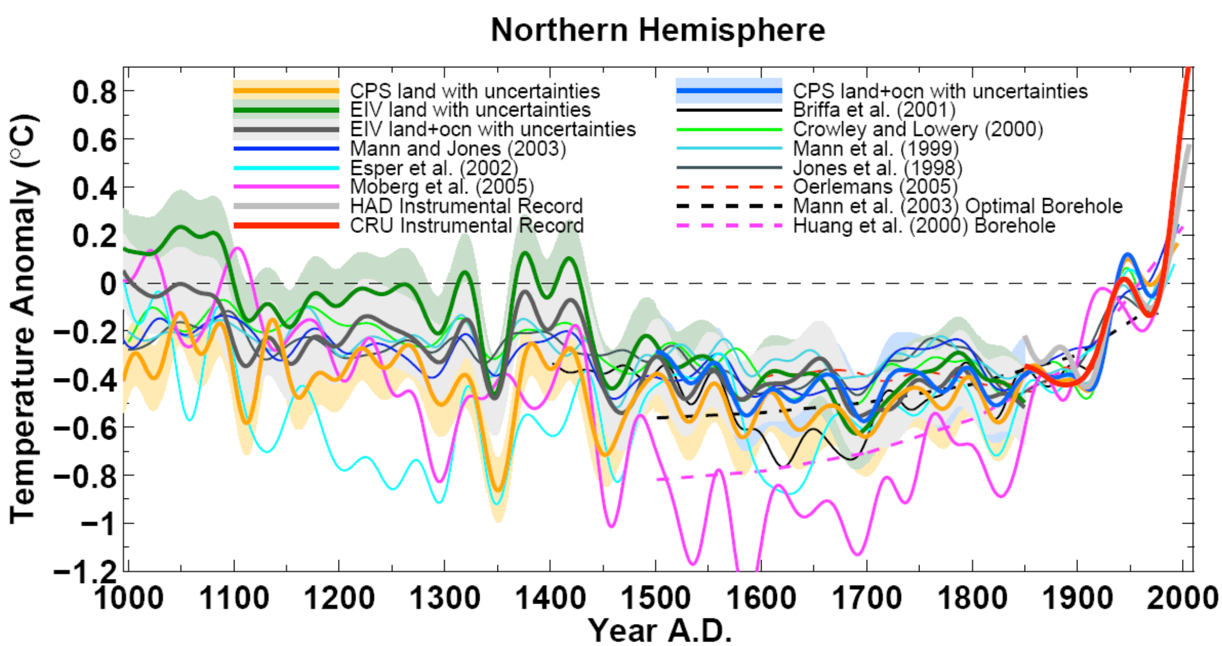
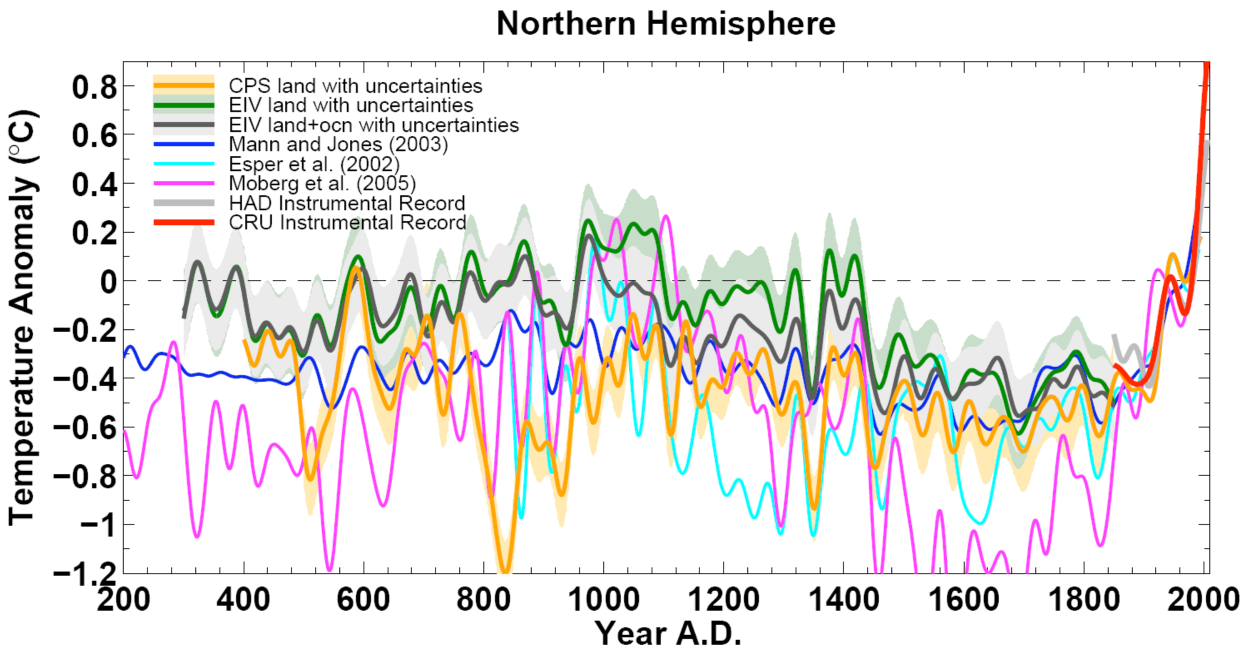


Figure S14: As in Figure 4 of paper, but where both the CPS reconstructions of this study and the other previously published reconstructions compared to (with the exception of the two borehole-based reconstructions Huang et al, 2000 and Mann et al 2003/Rutherford and Mann 2004) have been scaled to have the same trend amplitude as the CRU NH land surface temperature record during the overlap interval.



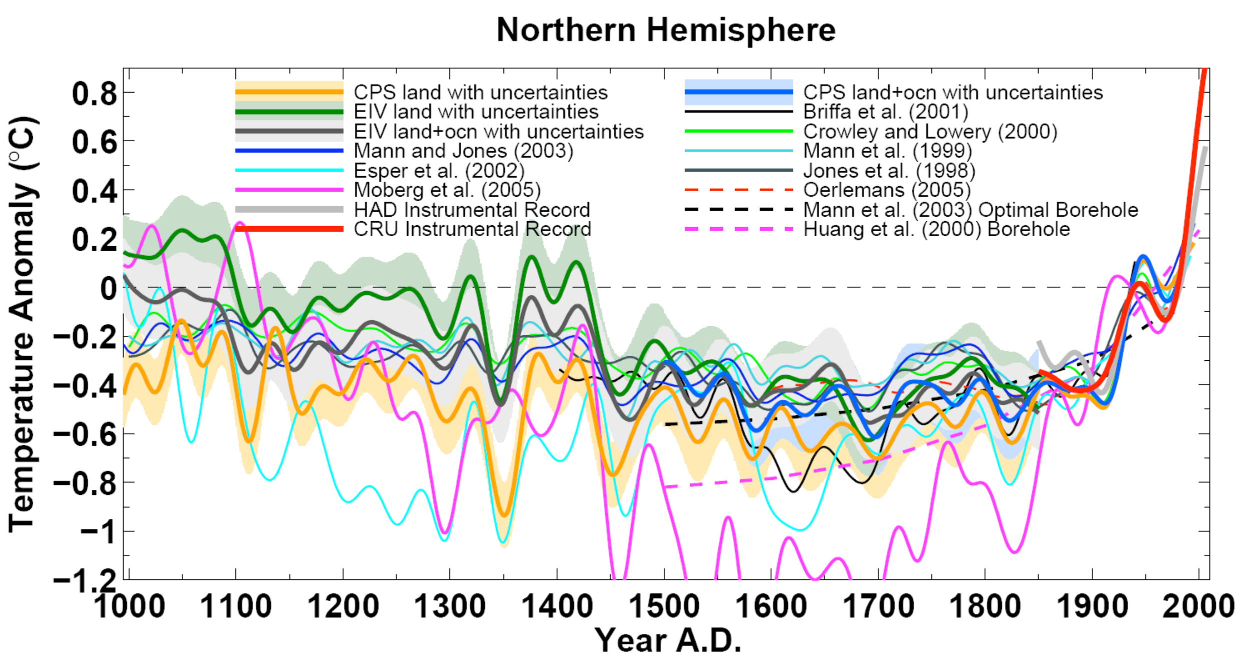


Figure S15: As in Figure 4 of paper, but where both the CPS reconstructions of this study and the other previously published reconstructions compared to (with the exception of the two borehole-based reconstructions Huang et al, 2000 and Mann et al 2003/Rutherford and Mann 2004) have been scaled based on the estimated signal amplitude indicated from the estimated signal-to-noise ratio [if the calibration period correlation between the instrumental NH mean temperature series and the proxy reconstruction is r, than the conventional (variance-matched, i.e. as in Figure 4 of paper) CPS reconstruction is rescaled by $1/r$ so that the estimated *signal component of the record* has the same decadal variance as the instrumental NH series]