

On the long-term context for late twentieth century warming

Rosanne D'Arrigo

Tree-Ring Laboratory, Lamont-Doherty Earth Observatory, Palisades, New York, USA

Rob Wilson

School of GeoSciences, University of Edinburgh, Edinburgh, UK.

Gordon Jacoby

Tree-Ring Laboratory, Lamont-Doherty Earth Observatory, Palisades, New York, USA

Received 10 June 2005; revised 23 September 2005; accepted 7 November 2005; published 7 February 2006.

[1] Previous tree-ring-based Northern Hemisphere temperature reconstructions portray a varying amplitude range between the “Medieval Warm Period” (MWP), “Little Ice Age” (LIA) and present. We describe a new reconstruction, developed using largely different methodologies and additional new data compared to previous efforts. Unlike earlier studies, we quantify differences between more traditional (STD) and Regional Curve Standardization (RCS) methodologies, concluding that RCS is superior for retention of low-frequency trends. Continental North American versus Eurasian RCS series developed prior to merging to the hemispheric scale cohere surprisingly well, suggesting common forcing, although there are notable deviations (e.g., fifteenth to sixteenth century). Results indicate clear MWP (warm), LIA (cool), and recent (warm) episodes. Direct interpretation of the RCS reconstruction suggests that MWP temperatures were nearly 0.7°C cooler than in the late twentieth century, with an amplitude difference of 1.14°C from the coldest (1600–1609) to warmest (1937–1946) decades. However, we advise caution with this analysis. Although we conclude, as found elsewhere, that recent warming has been substantial relative to natural fluctuations of the past millennium, we also note that owing to the spatially heterogeneous nature of the MWP, and its different timing within different regions, present palaeoclimatic methodologies will likely “flatten out” estimates for this period relative to twentieth century warming, which expresses a more homogenous global “fingerprint.” Therefore we stress that presently available paleoclimatic reconstructions are inadequate for making specific inferences, at hemispheric scales, about MWP warmth relative to the present anthropogenic period and that such comparisons can only still be made at the local/regional scale.

Citation: D'Arrigo, R., R. Wilson, and G. Jacoby (2006), On the long-term context for late twentieth century warming, *J. Geophys. Res.*, 111, D03103, doi:10.1029/2005JD006352.

1. Introduction

[2] Determination of how climate has varied in the past is important for evaluating the sensitivity of the earth's climate system to natural and anthropogenic forcing. High-resolution large-scale temperature reconstructions [Jacoby and D'Arrigo, 1989; D'Arrigo and Jacoby, 1993; Overpeck *et al.* 1997; Jones *et al.*, 1998; D'Arrigo *et al.*, 1999; Mann *et al.*, 1999; Briffa, 2000; Esper *et al.*, 2002a; Mann and Jones, 2003; Cook *et al.*, 2004; Moberg *et al.*, 2005] provide valuable insights into the types of natural climate changes that have occurred in the past and place recent warming into a longer-term context [Anderson and Woodhouse, 2005; Esper *et al.*, 2005b]. A great range in reconstructed amplitudes is observed, however, between the

currently existing Northern Hemisphere (NH) temperature reconstructions. One such reconstruction [Mann *et al.*, 1999] demonstrates minimal temperature amplitude (e.g., during the “Medieval Warm Period” (MWP) [Lamb, 1965] and “Little Ice Age” (LIA) [Grove, 1988]) while others [Briffa, 2000; Esper *et al.*, 2002a; Cook *et al.*, 2004; Moberg *et al.*, 2005] exhibit more pronounced variability. Moberg *et al.* [2005] considered that tree-ring records cannot reliably resolve lower frequency trends, and only used them to represent shorter-term variations. However, other studies [Esper *et al.*, 2002a; Cook *et al.*, 2004] demonstrated that tree rings can reflect longer-term trends, provided that data are appropriately processed (e.g., detrended using Regional Curve Standardization or RCS [Mitchell, 1967; Cook *et al.*, 1991; Briffa *et al.*, 1992; Becker *et al.*, 1995; Cook *et al.*, 1995; Esper *et al.*, 2002b], and/or have long individual series that can record multicentury trends. RCS allows capture of low-frequency

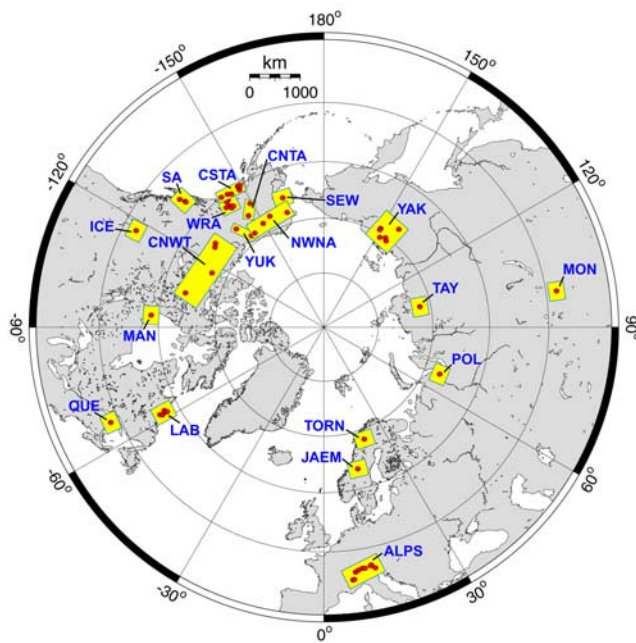


Figure 1. Location map of individual sites (red) and regional composites (yellow boxes) used to reconstruct NH temperatures over the past millennium. See Table 1 for site code description.

variance in excess of the mean length of individual samples used in chronology development [Briffa *et al.*, 1992; Cook *et al.*, 1995; Esper *et al.*, 2003, 2004]. When this method was previously applied to a tree-ring data set averaged over 14 North American and Eurasian sites [Esper *et al.*, 2002a; Cook *et al.*, 2004], pronounced “MWP” and “LIA” episodes were observed, with an approximate 1°C range over the past 1000 years [Cook *et al.*, 2004].

[3] In this paper, we develop two new tree-ring-based reconstructions of NH temperatures that address several priorities recommended for the generation of such records [Esper *et al.*, 2005b]: (1) utilization of regional proxy data processed to capture low-frequency trends, (2) need for improved coverage of millennial length records, and (3) use of nested modeling to allow systematic evaluation of uncertainties back in time. We refine the reconstructions, compared to previous efforts, by only utilizing tree-ring data that appear to portray a non-biased signal with temperature; that is, we minimized inclusion of data with mixed climate signals that also incorporate precipitation influences. Taking into account criteria that cause them to vary [Esper *et al.*, 2005a; Rutherford *et al.*, 2005], our reconstructions are compared to previous versions [Jones *et al.*, 1998; Mann *et al.*, 1999; Briffa, 2000; Esper *et al.*, 2002a; Mann and Jones, 2003; Cook *et al.*, 2004; Moberg *et al.*, 2005], as well as outputs from several climate models [Jones and Mann, 2004] to make inferences about past temperature variability, amplitude change and forcing over the past millennium.

2. Data, Methods, and Analysis

[4] Tree-ring width (and limited density [Luckman and Wilson, 2005]) data derived from living and subfossil wood

of coniferous tree species were compiled from 66 high-elevation and latitudinal treeline North American and Eurasian sites. Figure 1 shows the locations of individual sites and regional composites, identified using principal component analysis, for which raw measurements were standardized and merged to create regional, continental and hemispheric scale records. The combining of raw data for several sites within each region increased sample size for many of these composites, which is important for the successful application of RCS [Esper *et al.*, 2003]. Both standard (hereafter, STD) negative-exponential or straight-line curve fits [Cook and Kairiukstis, 1990] and RCS composites were generated for each region (Figure 2 and Table 1). Compared to previous studies [Jacoby and D’Arrigo, 1989; D’Arrigo and Jacoby, 1993; D’Arrigo *et al.*, 1999; Jones *et al.*, 1998; Mann *et al.*, 1999; Briffa, 2000; Esper *et al.*, 2002a], the North American data are much improved with new or extended millennial-length records, and updates of most of the data sets until at least the late 1990s (Figure 2 and Table 1).

[5] We reconstructed annual, rather than warm-season temperatures, as trees from selected treeline sites may integrate climate conditions during nongrowing season months [e.g., Jacoby and D’Arrigo, 1989]. Reconstructions based on seasonal versus annual temperatures should be virtually identical on multidecadal or longer timescales in any case [Esper *et al.*, 2002a; Cook *et al.*, 2004] (although note possible inhomogeneities in early instrumental summer data [Esper *et al.*, 2005a]). The Jones *et al.* [1999] gridded instrumental land temperature data set was utilized herein for calibration of the reconstructions.

[6] Regional tree-ring chronologies were assessed for signal strength and those periods represented by at least 10 tree-ring series were utilized for analysis. In all cases, this equates to an expressed population signal (EPS) statistic close to 0.85 [Cook and Kairiukstis, 1990]. We should note that replication for TORN, POL and TAY (Figure 2) did decline to eight or nine series for some select periods, but the EPS was never <0.70 . Following this strategy, six of the regional composite series have sufficient signal strength for use prior to 1000 AD. The regional chronologies were also screened by comparisons with instrumental (local and larger scale) temperature data to ensure that the temperature signal in the final reconstructions was as strong as possible and relatively unclouded by precipitation effects. In so doing, some potential data sets were discarded due to ambiguous signals. For example, we did not utilize the long bristlecone pine data sets from Colorado and California as many appear to portray a mixed precipitation and temperature signal (in addition to a purported CO_2 fertilization effect [LaMarche *et al.*, 1984]). We also did not use the Mackenzie Mountains, Boreal, Upperwright and Gotland data sets utilized by Esper *et al.* [2002a] for similar reasons, specifically that these records either (1) did not demonstrate a significant temperature signal on the local to regional scale, (2) displayed significant correlations with precipitation, or (3) were located at lower latitudes than those compiled for the present analysis.

[7] To develop the large-scale reconstructions, iterative averaging was performed to composite the regional records into continental and NH STD and RCS series, and linear regression analysis was used to calibrate these series to the

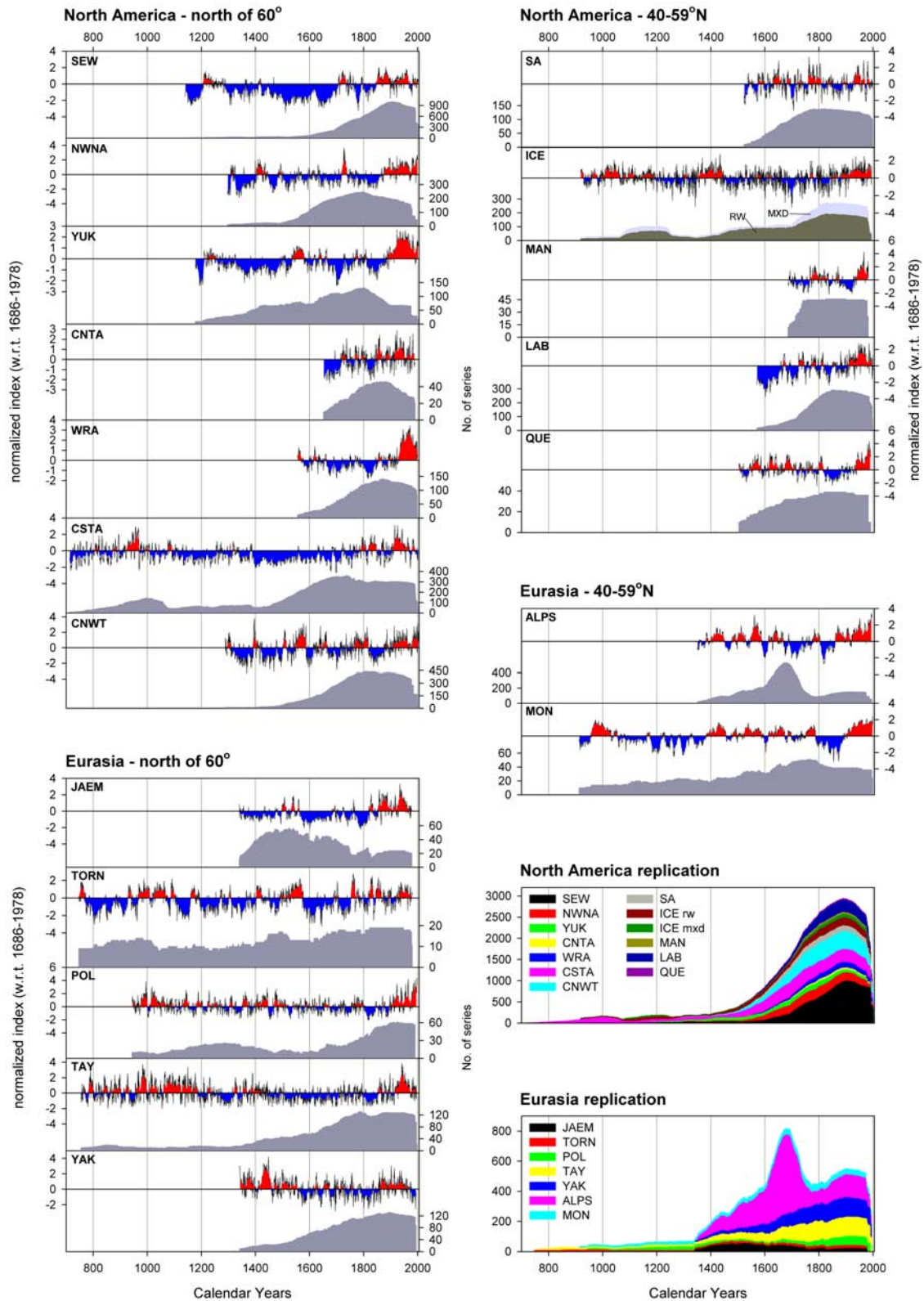


Figure 2. Individual regional composite RCS chronologies and their replication. The period shown for each chronology is that utilized for the generation of the reconstructions (see section 2 and Table 1). The time series have been loosely grouped according to latitude bands and were normalized to the common period. See Figure 1 for their locations. The bottom two panels in the right column show grouped replication plots for both North America and Eurasia.

Table 1. Detailed Information for Regional Composite Chronologies^a

Regional Grouping	Code	Full Coverage	Number of Series	MSL, years	Data Type ^b	Period > 10 Radii	Chronologies Used ^c	RCS Strategy ^d	Reference
<i>North America</i>									
Seward	SEW	978–2002	1196	209	L, S, H	1140–2002	STDPT, RCS	B	<i>D'Arrigo et al. [2005]</i>
NW North Alaska	NWNA	952–2000	294	301	L, S	1297–2000	STD, RCS	E	
Yukon	YUK	1067–2002	220	265	L, S	1177–2002	STDPT, RCSPT	E	
Central Alaska	CNTA	1556–1990	51	229	L	1652–1990	STD, RCS	A	
Wrangells	WRA	1471–1999	159	243	L	1556–1999	STD	n/a	<i>Davi et al. [2003]</i>
Coastal Alaska	CSTA	616–2002	820	242	L, S	713–2002	STD, RCS	B	<i>G. Wiles et al. (2005)^e</i>
Central NWT	CNWT	1046–2003	569	266	L, S	1288–2003	STD, RCS	D	
Southern Alaska	SA	1343–2000	143	338	L	1523–1999	STD, RCS	C	
Icefields	ICE	869–1994	374 (RW) 153 (MXD)	242/232	L, S	918–1994	STDPT, RCSPT	E	<i>Luckman and Wilson [2005]</i>
Manitoba	MAN	1650–1982	45	267	L	1686–1982	STDPT, RCS	A	
Labrador	LAB	1459–2001	371	202	L	1570–2001	STD, RCS	F	<i>D'Arrigo et al. [2003]</i>
Quebec	QUE	1404–1991	40	378	L	1504–1991	STD, RCS	E	
<i>Eurasia</i>									
Jaemtland	JAEM	1106–1978	156	159	L, S	1340–1978	STD, RCS	E	<i>Naurzbaev and Vaganov [1999]</i>
Tomtraesk	TORN	547–1980	55	309	L, S	747–1980	STD, RCSPT	A	<i>Briffa et al. [1992]</i>
Polar Urals	POL	778–1990	157	162	L, S	944–1990	STD, RCS	<i>Briffa [2000]</i>	
Taymir	TAY	513–1997	236	262	L, S	755–1997	STD, RCS	B	<i>Jacoby et al. [2000]</i>
Yakutia	YAK	1200–1994	179	277	L, S	1342–1994	STD, RCS	E	<i>Hughes et al. [1999]</i>
Alps	ALPS	986–1995	962	126	L, S, H	1350–1995	STDPT, RCS	E	<i>Wilson and Topham [2004], Nicolussi and Schiessling [2001]</i>
Mongolia	MON	262–1999	99	341	L, S	913–1999	STDPT, RCSPT	E	<i>D'Arrigo et al. [2001]</i>

^aMSL, mean segment length.^bL, living; S, subfossil or snag material; and H, historical material.^cIn select cases, a power transform (PT) was applied to correct for data biases [Cook and Peters, 1997]. This bias was assessed by correlation and residual analysis against both local and large-scale temperature series.^dOwing to differing populations in the TR data, the data sets were often grouped into “common” populations [Esper et al., 2003]. No one strategy is appropriate for all data sets, and careful evaluation of each composite data set was made. Strategy codes are as follows: A, one regional curve; B, multiple regional curves, related to growth trend type, i.e., those series traditionally detrended using negative exponential or linear regression functions; C, multiple regional curves, related to growth level differences, i.e., data divided equally into groups of low and high growth rates; D, mixture of B and C; E, multiple regional curves: data divided “horizontally” into separate living and subfossil (and historical) subgroups; F, separate regional curves for each constituent chronology. The relevant original references for each data set are listed. Note that for the Alps, different data sets were utilized for the STDPT and RCS chronology versions. The STDPT is a highly replicated expanded data set of that detailed by Wilson and Topham [2004]. RCS was not possible, however, with these data due to significant differences in mean RW between the historical and living data sets which resulted in a highly biased RCS chronology. Instead, for the Alps RCS chronology, the Nicolussi and Schiessling [2001] long pine RCS chronology was used. No STD version of this chronology exists.^eG. Wiles et al., Tree-ring evidence for a medieval warm period along the southern coast of Alaska, manuscript in preparation, 2005.

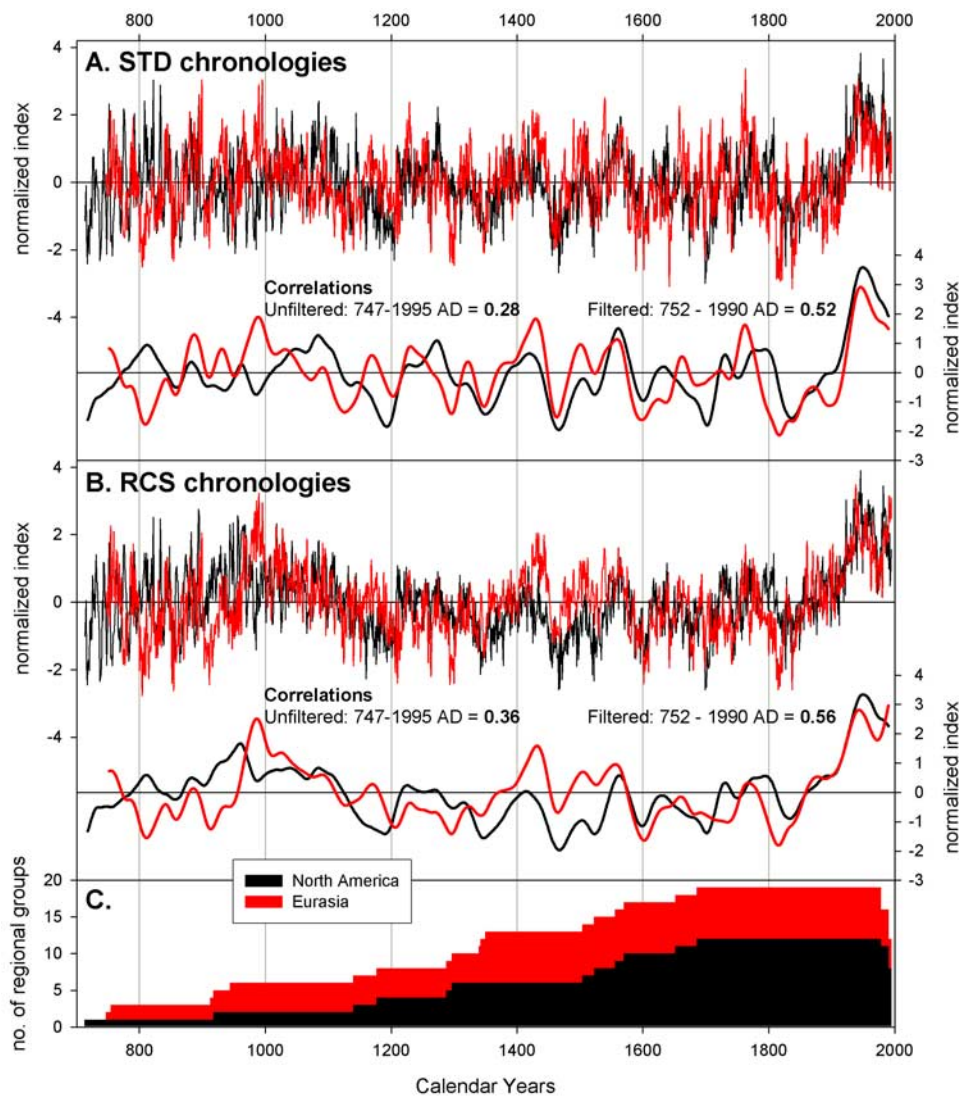


Figure 3. Comparison between continental large-scale mean width chronologies. (a) STD chronologies. (b) RCS chronologies. The time series have been normalized to the full period. Smoothed series are 20-year splines. Taking into account the autocorrelation of the series, the filtered correlations are not significant at the 95% confidence level. They are shown only to give a guide of coherence.

instrumental record. A nested approach, which accounts for the decrease in the number of chronologies back in time, was used to generate the longest possible reconstructions [Cook *et al.*, 2002]. This procedure entails normalizing the tree-ring series to the common period of all series in each nest and then averaging the series together to create a nest mean. To develop the final reconstructions, the mean and variance of each nested reconstructed time series were scaled to that of the most replicated nest (1686–1978) and the relevant sections for each nest spliced together (with all nests, the length of the final record spans from 713 to 1995). This approach stabilizes the variance of the final time series. For each nest, separate average time series were generated for North America and Eurasia (Figure 3), and these continental scale time series were averaged to produce a final large-scale hemispheric mean that was not biased to one particular continent due to varying number of series. This process, undertaken iteratively as each TR series

left the data matrix, resulted in 21 series upon which calibration and verification were made separately. Full period calibration was made over 1856–1978 (the common period of the tree-ring and temperature data), while verification was made over the period 1898–1937 after appropriate calibration using the combined 1856–1897/1938–1978 period. This nonstandard approach was employed to ensure as great a range in the predictand data as possible for calibration. Calibration and verification statistics typically employed to validate dendroclimatic reconstructions were then used to evaluate the reliability of the reconstructions [Cook and Kairiukstis, 1990]. As well as utilizing the commonly used Pearson's correlation and reduction of error (RE) statistics, we also used the coefficient of efficiency (CE). Both RE and CE are measures of shared variance between the actual and modeled series, but are usually lower than the calibration r^2 . A positive value for either statistic signifies that the regression model has some skill.

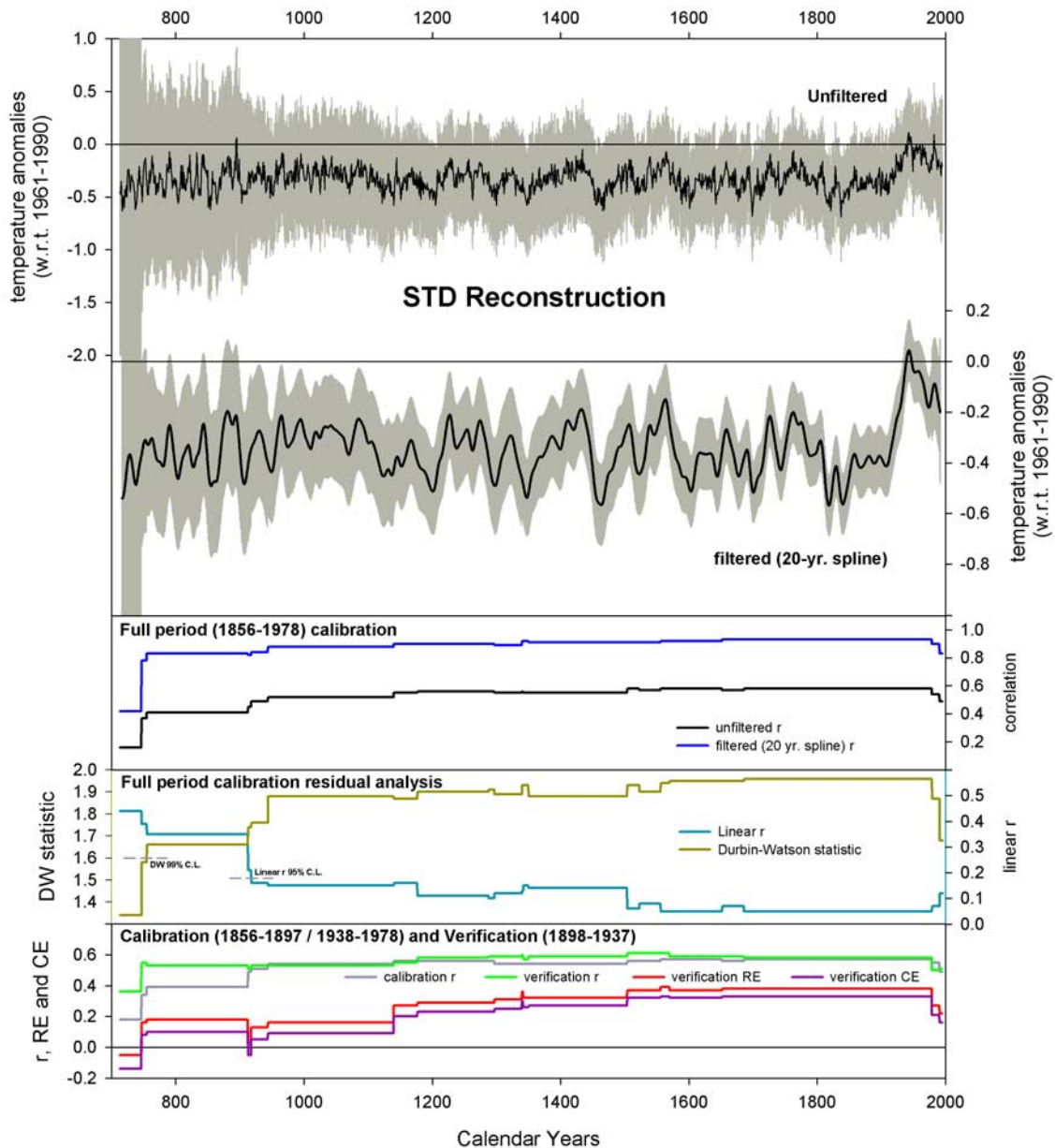


Figure 4a. STD NH reconstruction with nested modeling [Cook *et al.*, 2002] results. Top two panels show the final reconstruction with associated error bars. Bottom three panels present full period (1856–1978) calibration and associated residual series analyses along with verification (1898–1937) after calibrating over the combined 1856–1897/1938–1978 period.

CE is the more rigorous verification test. To test the robustness of the decadal to long-term signal in the reconstructed nested series, stringent assessment of the regression model residuals was also employed using the Durbin-Watson statistic (a test for residual autocorrelation) and by calculating the linear trend of the regression residual time series. As the modeled temperature signal is predominantly at timescales $> \sim 20$ years [Cook *et al.*, 2004; Esper *et al.*, 2005a], it is particularly important to identify models that have significant trends in the model residuals, as they would therefore not portray long-term variability in a robust manner. Calibration trials using a variety of data types (e.g., land only; land and sea combined) showed that, although correlations were stronger with the land/sea data

sets, significant autocorrelation was noted for all nested model residuals (presumably related to the higher autocorrelation of sea surface temperatures compared to land). Calibration against land only temperatures, despite coherence being weaker, resulted in less residual problems and better verification. Final calibration was therefore made against extratropical (20°N – 90°N) land-only mean annual (January–December) temperatures. The standard error of the regression estimate (standard deviation of the regression residuals) from the full period calibration was used to generate the 2 sigma error bars and this was also adjusted (inflated) to account for the change (decrease) in explained variance in each nest. We should note, however, that these calculated uncertainties are optimistic estimates as they do

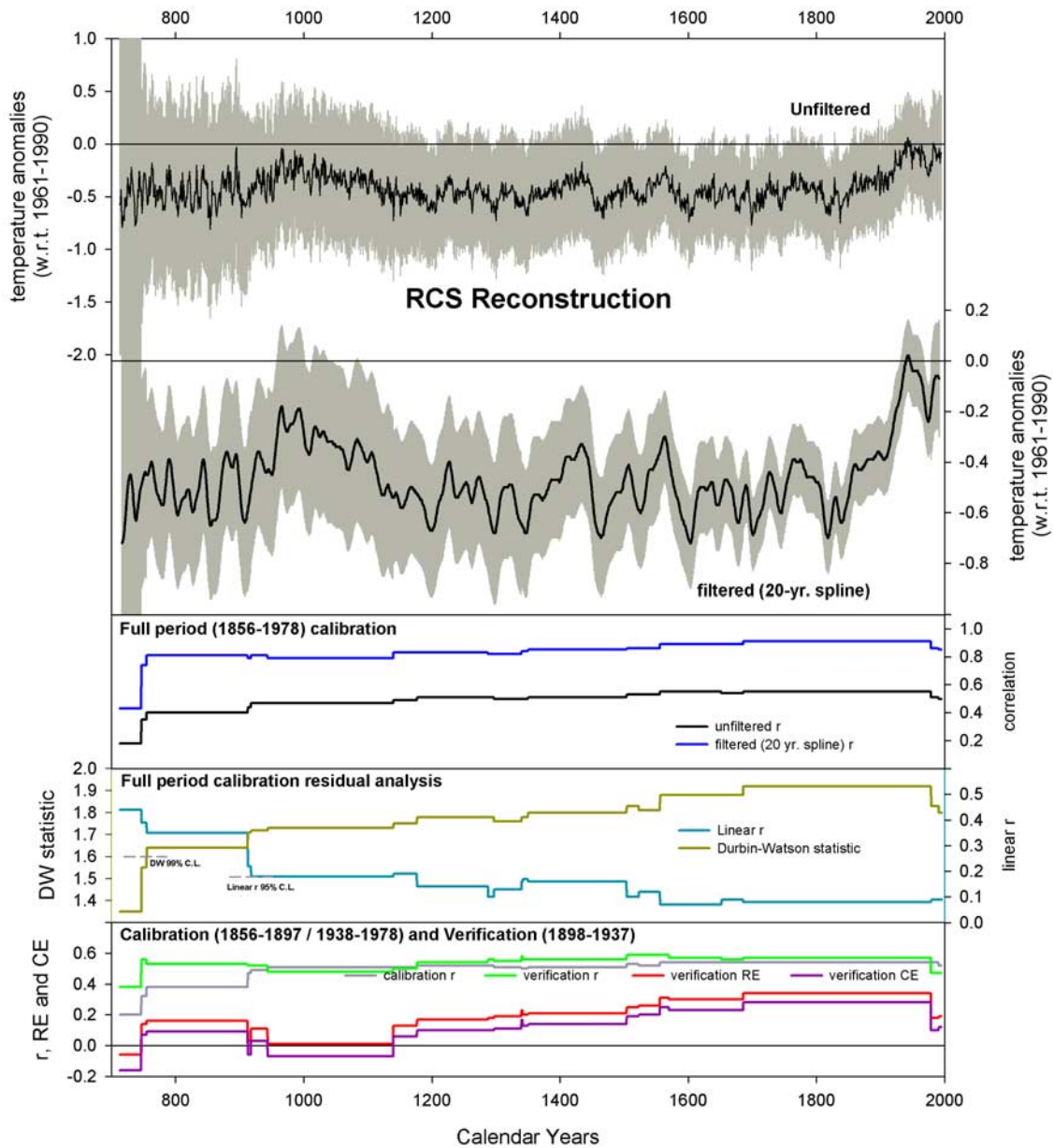


Figure 4b. As in Figure 4a but for the RCS reconstruction.

not incorporate the additional uncertainty of the regression coefficients, the weaker signal strength in the early periods of the TR chronologies (despite the use of EPS) and the fact that calibration did not include the post mid-1980s divergence (see section 3).

3. Results

[8] Following *Esper et al.* [2005a] and calibration trials against a variety of seasonal temperature data sets, final calibration was made against extratropical (20°N – 90°N) land-only annual (January–December) temperatures: a logical result as the proxy records are extratropical and land based. For the final NH STD and RCS reconstructions (Figures 4a, 4b, and 5), 33% and 30% of the temperature variance was accounted for, respectively. These values are relatively low owing to little coherence between the recon-

structions and instrumental data at interannual timescales [*Cook et al.*, 2004; *Esper et al.*, 2005a]. If the time series are smoothed with a 20-year smoothing spline [*Cook and Kairiukstis*, 1990], the explained variance increases to 87% and 84%, respectively. However, fidelity of this signal decreases with time (Figures 4a and 4b). Although the reconstructions cover 713–1995, verification shows that the STD and RCS NH reconstructions are statistically robust back to 918 AD (five regional composite series) and 1117 AD (eight regional composite series), respectively. Before these dates, CE values become negative and linear trends are noted in model residuals. RE values do however remain positive back to 747 for both reconstructions, suggesting some signal fidelity when at least one record is utilized from each continent. The weaker results for the RCS reconstruction highlight the noisier nature of this detrending method [*Briffa et al.*, 1992; *Cook et al.*, 1995].

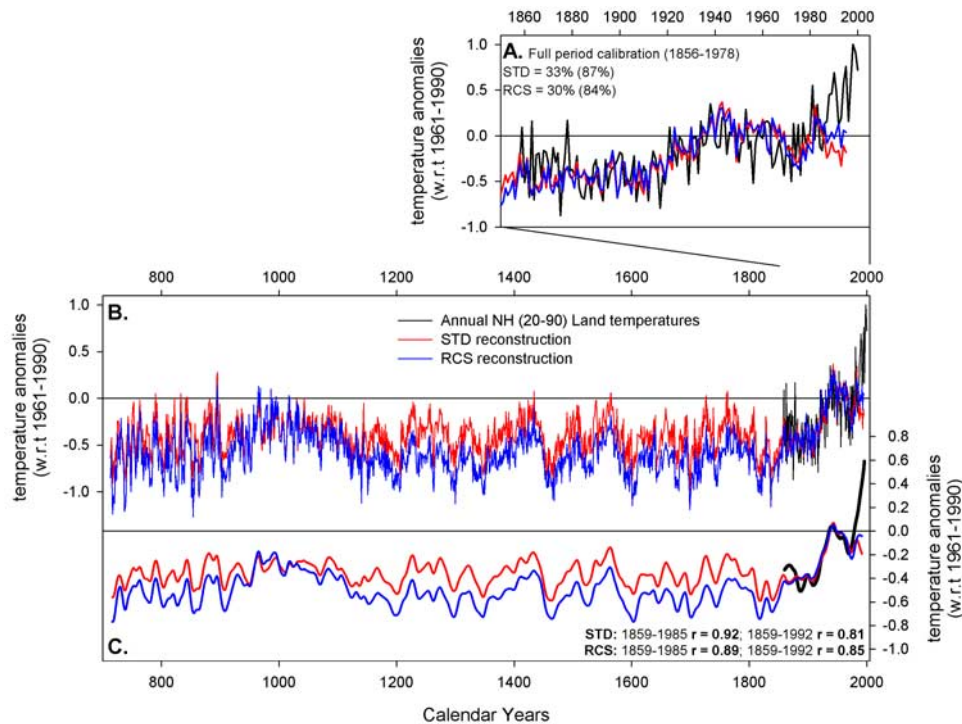


Figure 5. Comparison of STD and RCS NH reconstructions with mean annual land (20°N–90°N) temperatures. (a) Calibration period values denote the variance explained from the calibration. Values in parentheses are for 20-year smoothed filtered versions. (b) Full length time series. Although regression-based techniques can systematically underestimate low-frequency trends [Von Storch *et al.*, 2005], we have partially overcome this problem by scaling (with respect to 1856–1978) the calibrated reconstructed time series to the instrumental record [Esper *et al.*, 2005a]. (c) As in Figure 5b but using smoothed (20-year) time series. The values show correlations between the smoothed reconstructions and the instrumental records for the extended periods 1859–1985 and 1859–1992. To reduce potential end effect biases of the smoothed series, 3 years were truncated from the ends of the time series before correlation analysis.

We thus advise caution in assessing trends in RCS series due to potential biases that may occur during standardization [Esper *et al.*, 2003; K. Briffa, University of East Anglia, personal communication, 2005]. However, it is difficult to assess long-term trends between the STD and RCS reconstructions owing to the restricted length of the instrumental data. Therefore the weaker RCS results do not necessarily mean that the extra low-frequency information is biased. Furthermore, it is important to emphasize that only the RCS reconstruction enables an assessment of long-term temperature trends in excess of the mean segment length [Cook *et al.*, 1995]. Depending on the segment length structure in the individual regional composites (see Table 1), it is not possible to reconstruct low-frequency variations longer than some fraction (approximately one third) of the mean segment length when utilizing individual series detrending methods.

[9] The Eurasian and North American composites, despite high variability between the regional composite chronologies, cohere surprisingly well, especially for the RCS detrended series (Figure 3). This coherency implies common forcing, presumably related to external (solar, volcanic and anthropogenic) influences. However, although generally similar, there are also significant differences (Figure 3). The Eurasian RCS composite shows

high-index values ~ 1000 and the mid-twentieth century, with prolonged low-index values from 1100–1350 and 1600–1900, punctuated by higher values circa 1400–1550. The North American RCS record shows higher-index values in the twentieth century which exceed levels at ~ 950 . Between these highs, prolonged low-index values are noted throughout much of the last millennium. Interestingly, the period of lowest-index values in the North American series (~ 1500) coincides with a period (1400–1550) in the Eurasian series when inferred conditions would have been warmer.

[10] Clear differences are also observed between the STD and RCS reconstructions on the hemispheric scale (Figures 4a, 4b, and 5). Overall, significantly more low-frequency information is captured using RCS. The NH RCS reconstruction shows warming around the “MWP” (~ 950 –1100) and overall cooling from ~ 1100 –1400, with an extended period overlapping the “LIA” from ~ 1450 –1850 [Grove, 1988]. Late twentieth century warming exceeds peak MWP conditions by 0.67°C when comparing decadal averages (960–969 (reconstruction) = -0.12°C versus 1991–2000 (instrumental) = 0.55°C ; the reconstruction was scaled [Esper *et al.*, 2005a] to the instrumental data to calculate these results (Figure 4b)). By comparison, peak twentieth century warmth for the period covered only by the

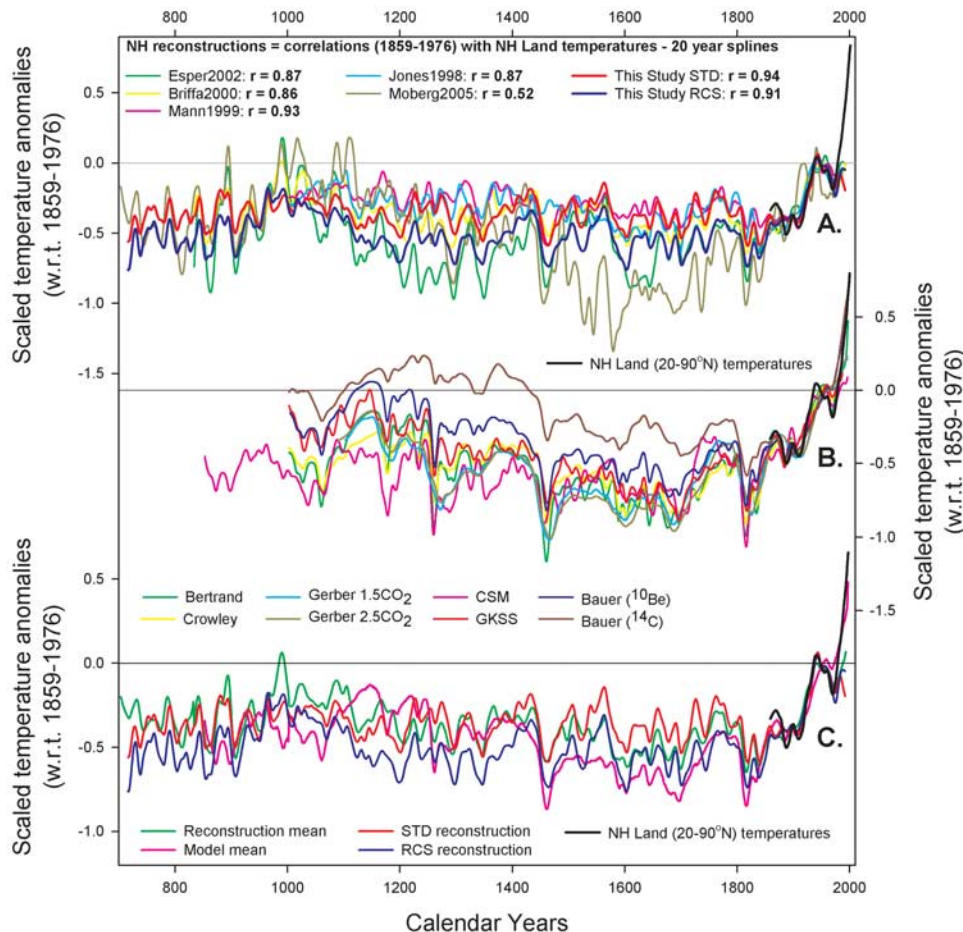


Figure 6. (a) Comparison of STD and RCS NH reconstructions with previous reconstructions. (b) Model-based estimates of NH temperatures for the last millennium [Jones and Mann, 2004]. (c) Comparison of mean series of the previously developed reconstructions and models with the STD and RCS series. The reconstruction and model time series were normalized to the common period and averaged. All smoothed series in this figure were scaled to the smoothed instrumental NH temperature series over the period 1859–1976.

proxy data (1937–1946, 0.17°C) exceeds peak MWP conditions by 0.29°C .

[11] Taking into account differences in calibration method and target season [Esper et al., 2005a; Rutherford et al., 2005] that affect interpretation, we compare our reconstructions with previous reconstructions after scaling [Esper et al., 2005a] all the records to land-only annual temperatures (20°N – 90°N) (Figure 6a; see also auxiliary material¹ Figure 1). Strong similarities are not surprising owing to some overlapping data (especially in Eurasia) with prior studies. Our reconstructions fall within the middle range of sensitivity of series published thus far (Table 2). The long-term trends of the STD reconstruction most closely match the Mann et al. [1999] and Jones et al. [1998] series, whereas the RCS reconstruction compares best with the Esper et al. [2002a] and Cook et al. [2004] series. This observation validates the hypothesis [Esper et al., 2004] that one reason for the relative lack of long-term variability in

the work of Mann et al. [1999] was their use of standard detrending procedures that removed low-frequency variation. Besides differences attributable to RCS and other factors noted above, solar-forced thermohaline circulation changes [Bond et al., 2001] and their preferred impact on higher-latitude climate may partly account for greater variability in reconstructions with an extratropical emphasis [Esper et al., 2002a].

Table 2. Coldest and Warmest Decades (Anomaly Values in Parentheses) Calculated Over 1000–1979 for Each Reconstruction After They Have Been Scaled (With Respect to 1856–1978) to NH Land Only (20°N – 90°N) Mean Annual Temperatures^a

	Coldest	Warmest	Amplitude
This Study STD	1813–1822 (−0.74)	1938–1947 (0.20)	0.94
This Study RCS	1600–1609 (−0.97)	1937–1946 (0.17)	1.14
Esper et al. [2002a]	1345–1354 (−1.18)	1950–1959 (0.15)	1.34
Briffa [2000]	1813–1822 (−0.80)	1951–1960 (0.10)	0.90
Mann et al. [1999]	1458–1467 (−0.68)	1957–1966 (0.10)	0.79
Jones et al. [1998]	1693–1702 (−0.77)	1929–1938 (0.06)	0.83
Moberg et al. [2005]	1576–1585 (−1.33)	1104–1113 (0.23)	1.56

^aThe difference between these two values is defined as the amplitude.

¹Auxiliary material is available at <ftp://ftp.agu.org/apend/jd/2005JD006352>.

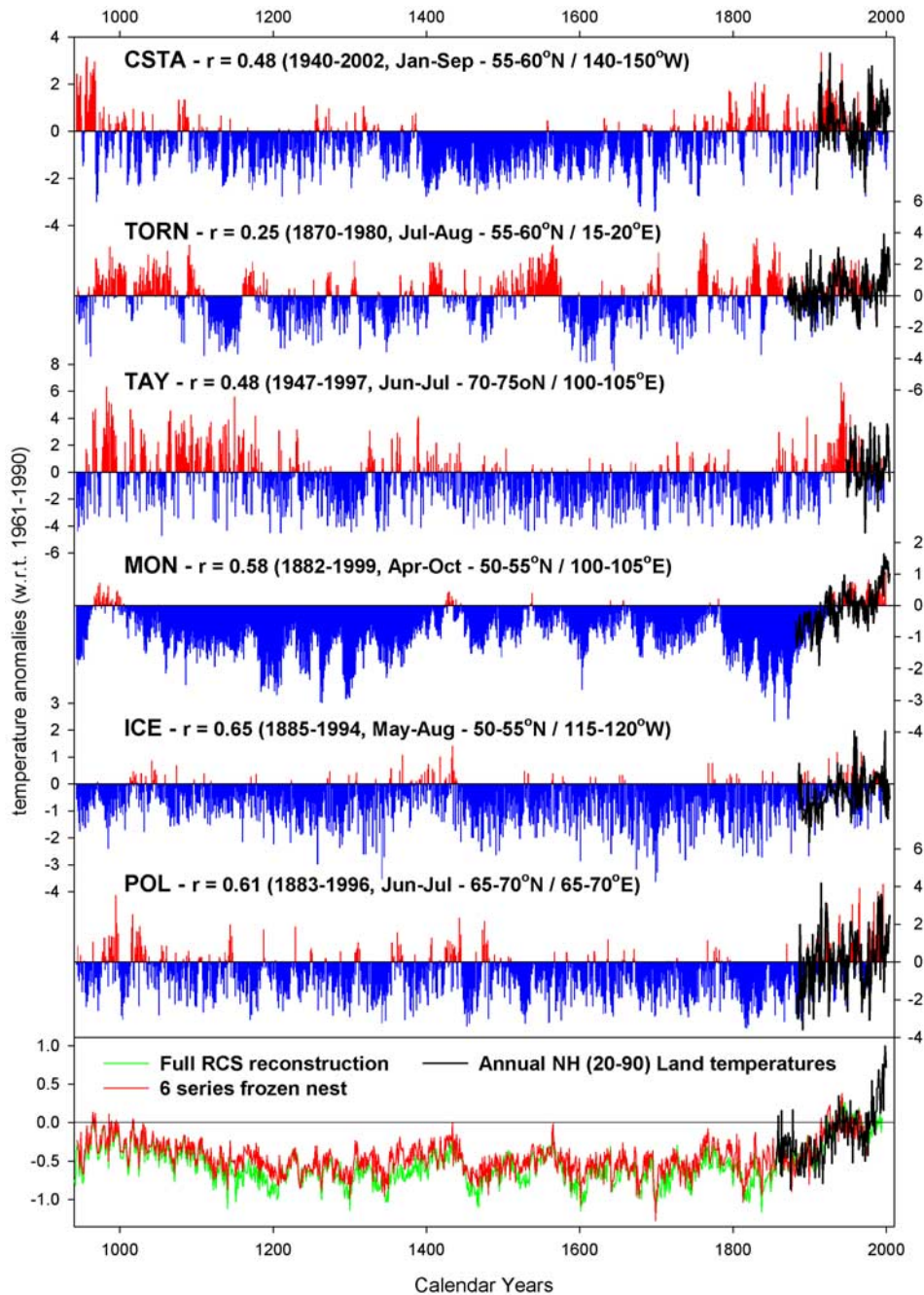


Figure 7. Six longest (>1000 years) chronologies after they have been scaled against the optimum season of their local Jones gridded data. The spatial area of the grids used is detailed. The bottom panel compares the mean of the six chronologies with the RCS reconstruction after they have both been scaled to mean annual land (20°N–90°N) temperatures.

[12] Figure 6b plots several large-scale NH climate models [Jones and Mann, 2004]. Although amplitudes are quite variable, they all show the same general trend, with warming around the twelfth to thirteenth centuries and cooling ~1450–1850. When scaled [Esper *et al.*, 2005a] and compared to the instrumental record, recent warming appears unprecedented over the last 1000 years in both the models and reconstructions. The STD and RCS reconstructions are compared to separate mean series of other NH reconstructions and models (Figure 6c). The RCS recon-

struction compares better with the model mean than the STD series ($r = 0.57$ versus 0.49 , respectively). This suggests that although the STD reconstruction is perhaps superior from a calibration/verification point of view (Figures 4a and 4b), it is systematically biased in the frequency domain as it does not portray long-term variability adequately. We emphasize, however, that this comparison must be considered with caution, as there is considerable model uncertainty [Jones and Mann, 2004]. There are also important differences between the model mean and RCS

series. The RCS reconstruction shows cooler conditions from ~1100–1400. The models, however, express warmer conditions in the twelfth century, suggesting that the warmest phase of the MWP was later than in the reconstructions.

[13] To highlight the uncertainties in the investigation of the significance of the MWP versus twentieth century warming, Figure 7 shows the six longest (>1000 years) chronologies after they have been scaled against the optimum season of their local Jones gridded data. The figure clearly shows that the recent period does not look particularly warmer compared to the MWP; MON is the only exception in this regard. However, the mean of these six series (Figure 7, bottom), which compares well with the RCS reconstruction, clearly places recent warming well above reconstructed conditions of the MWP. This therefore highlights a bias/artifact in the full RCS reconstruction (and likely in many of the other reconstructions) where the MWP, because it is expressed at different times in the six long records, is “averaged out” (i.e., flattened) compared to the recent period which shows a much more globally consistent signal. This observation not only emphasizes the problem of using such a small data-set, and calibrating during a period where the global signal is more coherent (and therefore resulting in more optimistic calibration/verification results), but that the reconstruction of one single large-scale parameter (in this case annual temperatures) does not provide any valid spatial climatic information. However, in light of this observation, although we cannot make any robust conclusions about mean MWP conditions compared to the present (unless one looks at the individual regional records), we can confidently state that the global warming “fingerprint” is globally more homogenous than warming during the MWP.

4. Discussion and Conclusions

[14] We have presented STD and RCS NH temperature reconstructions for the past 1250 years. In so doing, we have addressed several recommended priorities [Esper *et al.*, 2005b] for the development of large-scale reconstructions (see above). On the basis of the above comparisons and analyses, we conclude that the RCS reconstruction is superior to the more traditional STD method with regards to the ability to retain low-frequency (centennial to multi-centennial) trends. The NH RCS reconstruction displays pronounced variability, including significant “MWP” and “LIA” departures. An apparent decrease in recent temperature sensitivity for many northern sites [Jacoby and D'Arrigo, 1995; Briffa *et al.*, 1998] is evident in our reconstructions, with divergence from instrumental temperatures after ~1986 (Figure 5). There are several hypotheses for this divergence [Jacoby and D'Arrigo, 1995; Briffa *et al.*, 1998; Vaganov *et al.*, 1999; Barber *et al.*, 2000; Wilson and Luckman, 2003; D'Arrigo *et al.*, 2004; Wilmking *et al.*, 2005], none of which appear consistent for all NH sites. Although we calibrated to the common 1856–1978 period, valid calibration using a reduced data set would be possible until the mid-1980s (Figure 5). After this period, however, the divergence between the tree-ring and instrumental data results in weakening of calibration results and failed verification statistics.

[15] Model comparisons show reasonable coherence over the last 600 years with the RCS reconstruction (Figure 6). Proxy reconstructions, however, show an earlier peak in MWP warmth compared to models, possibly reflecting that this was a spatially complex, highly variable period [Jones and Mann, 2004] and that not enough proxy records yet exist for this time. It is also possible that the models are themselves biased in some way (e.g., although they incorporate external (solar, volcanic, anthropogenic) forcings, they do not take into account internal atmosphere-ocean dynamics [Jones and Mann, 2004]). Taken at face value, our reconstruction indicates that MWP conditions were nearly 0.7°C cooler than those of the late twentieth century. These results suggest how extreme recent warming has been relative to the natural fluctuations of the past millennium. This conclusion, however, must be taken cautiously. First, there is significant divergence between reconstructed and actual temperatures since the mid-1980s, which, until valid reasons for this phenomenon have been found, can only question the ability of tree-ring data to robustly model earlier periods that could have been similarly warm (or warmer) than the present. Second, there are presently only very few millennial length records available for direct comparison between the recent period and the MWP, and these records show trends which are not necessarily coherent over the latter interval, resulting in a “flattening” of MWP conditions compared to recent warming in our reconstruction. Ultimately, many long records from new NH locations and updating of existing records (mainly in Eurasia) to the present are required. Successful modeling of paleoclimate data with the high temperatures of the late 1990s is essential if we are to make robust, definitive conclusions about past temperature amplitudes and variability.

[16] **Acknowledgments.** We thank the National Science Foundation Earth System History program and the NOAA Climate Change and Detection program for funding this research. All authors contributed equally to this work. We thank G. Wiles, B. and C. Buckley, B. Luckman and K. Nicolussi, and contributors to the International Tree-Ring Data Bank for collecting or providing data, and E. Cook and J. Esper for informative discussions. We gratefully acknowledge the U.S. Parks Service and Parks Canada. Lamont-Doherty Earth Observatory contribution 6829.

References

- Anderson, D., and C. Woodhouse (2005), Let all the voices be heard, *Nature*, **433**, 587–588.
- Barber, V., G. P. Juday, and B. Finney (2000), Reduced growth of Alaska white spruce in the twentieth century from temperature-induced drought stress, *Nature*, **405**, 668–672.
- Becker, M., G. Bert, J. Bouchon, J. Dupouey, J. Picard, and E. Ulrich (1995), Long-term changes in forest productivity in northeastern France: The dendroecological approach, in *Forest Decline and Atmospheric Deposition Effects in the French Mountains*, edited by G. Landemann and M. Bonneau, pp. 143–156, Springer, New York.
- Bond, G., *et al.* (2001), Persistent solar influence on North Atlantic climate during the Holocene, *Science*, **294**, 2130–2136.
- Briffa, K. (2000), Annual climate variability in the Holocene: Interpreting the message from ancient trees, *Quat. Sci. Rev.*, **19**, 87–105.
- Briffa, K., *et al.* (1992), Fennoscandian summers from AD 500: temperature changes on short and long time scales, *Clim. Dyn.*, **7**, 111–119.
- Briffa, K., *et al.* (1998), Reduced sensitivity of recent tree growth to temperature at high northern latitudes, *Nature*, **391**, 678–682.
- Cook, E., and L. Kairiukstis (1990), *Methods of Dendrochronology*, Springer, New York.
- Cook, E., and K. Peters (1997), Calculating unbiased tree-ring indices for the study of climate and environmental change, *Holocene*, **7**, 361–370.
- Cook, E., T. Bird, M. Peterson, M. Barbetti, B. Buckley, R. D'Arrigo, R. Francey, and P. Tans (1991), A 1089-yr temperature record for

- Tasmania inferred from tree-rings of subalpine huon pine, *Science*, 253, 1266–1268.
- Cook, E., et al. (1995), The “segment length curse” in long tree-ring chronology development for paleoclimatic studies, *Holocene*, 5, 229–237.
- Cook, E., R. D'Arrigo, and M. E. Mann (2002), A well-verified, multi-proxy reconstruction of the winter North Atlantic Oscillation index since AD 1400, *J. Clim.*, 15, 1754–1764.
- Cook, E., J. Esper, and R. D'Arrigo (2004), Extra-tropical Northern Hemisphere land temperature variability over the past 1000 years, *Quat. Sci. Rev.*, 23, 2063–2074.
- D'Arrigo, R., and G. Jacoby (1993), Secular trends in high northern-latitude temperature reconstructions based on tree rings, *Clim. Change*, 25, 163–177.
- D'Arrigo, R., G. Jacoby, M. Free, and A. Robock (1999), Northern Hemisphere temperature variability for the past three centuries: tree-ring and model estimates, *Clim. Change*, 42, 663–675.
- D'Arrigo, R., et al. (2001), 1738 years of Mongolian temperature variability inferred from a tree-ring record of Siberian pine, *Geophys. Res. Lett.*, 28, 543–546.
- D'Arrigo, R., et al. (2003), Interannual to multidecadal modes of Labrador climate variability inferred from tree rings, *Clim. Dyn.*, 20, 219–228.
- D'Arrigo, R., R. Kaufmann, N. Davi, G. Jacoby, C. Laskowski, and R. Myneni (2004), Thresholds for warming-induced growth decline at elevational treeline in Yukon Territory, Canada, *Global Biogeochem. Cycles*, 18, GB3021, doi:10.1029/2004GB002249.
- D'Arrigo, R., et al. (2005), North Pacific-related climate variability inferred from Seward Peninsula, Alaska tree rings since A.D. 1358, *Clim. Dyn.*, 24, 227–236, doi:10.1007/s00382-004-0502-1.
- Davi, N., G. Jacoby, and G. Wiles (2003), Boreal temperature variability inferred from maximum latewood density and tree-ring width data, Wrangell Mountain region, Alaska, *Quat. Res.*, 60, 252–262.
- Esper, J., E. Cook, and F. Schweingruber (2002a), Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability, *Science*, 295, 2250–2253.
- Esper, J., F. Schweingruber, and M. Winiger (2002b), 1300 years of climatic history for western central Asia inferred from tree-rings, *Holocene*, 12, 267–277.
- Esper, J., et al. (2003), Tests of the RCS method for preserving low-frequency variability in long tree-ring chronologies, *Tree Ring Res.*, 59, 81–89.
- Esper, J., D. Frank, and R. Wilson (2004), Climate reconstructions: Low-frequency ambition and high-frequency ratification, *Eos Trans. AGU*, 85, 113, 120.
- Esper, J., D. Frank, R. Wilson, and K. Briffa (2005a), Effect of scaling and regression on reconstructed temperature amplitude for the past millennium, *Geophys. Res. Lett.*, 32, L07711, doi:10.1029/2004GL021236.
- Esper, J., R. J. S. Wilson, D. C. Frank, A. Moberg, H. Wanner, and J. Luterbacher (2005b), Climate: Past ranges and future changes, *Quat. Sci. Rev.*, 24, 2164–2166.
- Grove, J. (1988), *The Little Ice Age*, Methuen, New York.
- Hughes, M. K., et al. (1999), Twentieth-century summer warmth in northern Yakutia in a 600 year context, *Holocene*, 9, 603–608.
- Jacoby, G., and R. D'Arrigo (1989), Reconstructed Northern Hemisphere annual temperature since 1671 based on high latitude tree-ring data from North America, *Clim. Change*, 14, 39–59.
- Jacoby, G. C., and R. D'Arrigo (1995), Tree-ring width and density evidence of climatic and potential forest change in Alaska, *Global Biogeochem. Cycles*, 9, 227–234.
- Jacoby, G. C., et al. (2000), Long-term temperature trends and tree growth in the Taymir region of northern Siberia, *Quat. Res.*, 53, 312–318.
- Jones, P., and M. E. Mann (2004), Climate over past millennia, *Rev. Geophys.*, 42, RG2002, doi:10.1029/2003RG000143.
- Jones, P., et al. (1998), High-resolution paleoclimatic records for the last millennium: Interpretation, integration and comparison with general circulation model control-run temperatures, *Holocene*, 8, 455–471.
- Jones, P., et al. (1999), Climatic extremes: Approaches to creating indices from daily and monthly data, *Clim. Change*, 42, 131–149.
- LaMarche, V. C., et al. (1984), Increasing atmospheric carbon dioxide: Tree ring evidence for enhancement in natural vegetation, *Science*, 225, 1019–1021.
- Lamb, H. H. (1965), The early medieval warm epoch and its sequel, *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 1, 13–37.
- Luckman, B., and R. Wilson (2005), Summer temperatures in the Canadian Rockies during the last millennium—A revised record, *Clim. Dyn.*, 24, 131–144.
- Mann, M. E., and P. Jones (2003), Global surface temperatures over the past two millennia, *Geophys. Res. Lett.*, 30(15), 1820, doi:10.1029/2003GL017814.
- Mann, M. E., R. Bradley, and M. Hughes (1999), NH temperatures during the past millennium: Inferences, uncertainties and limitations, *Geophys. Res. Lett.*, 26, 759–762.
- Mitchell, V. (1967), An investigation of certain aspects of tree growth rates in relation to climate in the central Canadian boreal forest, *Tech. Rep. 33*, Dep. of Meteorol., Univ. of Wisconsin-Madison, Madison.
- Moberg, A., et al. (2005), Highly variable Northern Hemisphere temperatures reconstructed from low-and high-resolution proxy data, *Nature*, 433, 613–617, doi:10.1038/nature03265.
- Naurzbaev, M. M., and E. A. Vaganov (1999), 1957-year tree-ring chronology of eastern part of Taymir, *Sib. J. Ecol.*, 6, 159–168.
- Nicolussi, K., and P. Schiessling (2001), Establishing a multi-millennial *Pinus cembra* chronology for the central Eastern Alps, in *International Conference of Tree-Rings and People*, edited by M. Kaennel-Dobbertin and O. U. Bräker, pp. 251–252, Swiss Fed. Res. Inst. WSL, Davos, Switzerland.
- Overpeck, J., et al. (1997), Arctic environmental change of the last four centuries, *Science*, 278, 1251–1256.
- Rutherford, S., et al. (2005), Proxy-based Northern Hemisphere surface temperature reconstructions: Sensitivity to method, predictor network, target season, and target domain, *J. Clim.*, 18, 2308–2329.
- Vaganov, E., M. K. Hughes, A. V. Kirdyanov, F. H. Schweingruber, and P. P. Silkin (1999), Influence of snowfall and melt timing on tree growth in subarctic Eurasia, *Nature*, 400, 149–151.
- Von Storch, H., et al. (2005), Reconstructing past climate from noisy data, *Science*, 306, 679–682.
- Wilmking, M., R. D'Arrigo, G. Jacoby, and G. Juday (2005), Divergent growth responses in circumpolar boreal forests, *Geophys. Res. Lett.*, 32, L15715, doi:10.1029/2005GL023331.
- Wilson, R. J. S., and B. H. Luckman (2003), Dendroclimatic reconstruction of maximum summer temperatures from upper tree-line sites in interior British Columbia, *Holocene*, 13, 853–863.
- Wilson, R. J. S., and J. Topham (2004), Violins and climate, *Theor. Appl. Climatol.*, 77, 9–24.

R. D'Arrigo and G. Jacoby, Tree-Ring Laboratory, Lamont-Doherty Earth Observatory, 61 Route 9W, Palisades, NY 10964, USA. (rdd@ldeo.columbia.edu)

R. Wilson, School of GeoSciences, Grant Institute, University of Edinburgh, West Mains Road, Edinburgh EH9 3JW, UK. (rob.wilson@ed.ac.uk)