Quaternary Science Reviews 26 (2007) 3298–3310

Warmer early instrumental measurements *versus* colder reconstructed temperatures: shooting at a moving target

David Frank[a,*], Ulf Büntgen[a], Reinhard Böhm[b], Maurizio Maugeri[c], Jan Esper[a]

[a]*Swiss Federal Research Institute WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland*
[b]*Central Institute for Meteorology and Geodynamics, Hohe Warte 38, 1190 Vienna, Austria*
[c]*Istituto di Fisica Generale Applicata, via Brera, 28, 20121 Milan, Italy*

Received 28 September 2006; received in revised form 12 August 2007; accepted 24 August 2007

Abstract

Comparison of tree-ring-based warm-season temperature reconstructions and their instrumental target data reveals substantial divergence between (warmer) early instrumental measurements and (colder) proxy estimates. Here we detail this systematic misfit for the Northern Hemisphere before ~1900 and the European Alps before ~1850. Five hypotheses related to both proxy and target uncertainties are presented towards explaining this phenomenon. These include: (1) tree-ring detrending methods, (2) biological persistence in the proxy time-series, (3) uncertainties and instabilities in the growth response to given climatic parameters, (4) reduced instrumental station availability back in time, and (5) instrumental data homogeneity. We suggest that uncertainties in the choice of instrumental targets at the hemispheric scale, and instrumental data inhomogeneities at the Alpine and possibly also the hemispheric-scale are the most important factors in explaining this offset. Assessment of homogeneity at larger scales remains challenging. Attention is drawn to possible warm biases in early thermometer shelters and the relevance of proxy/target discrepancies for understanding and quantifying the amplitude of both recent anthropogenic and past natural forced climate fluctuations.

1. Introduction

To understand and quantify past variations in the earth's climate system and its forcing agents, millennial-long proxy reconstructions (e.g., Jones et al., 1998; Mann et al., 1999; Briffa, 2000; Esper et al., 2002; Moberg et al., 2005; D'Arrigo et al., 2006) and model simulations (e.g., von Storch et al., 2004; Hegerl et al., 2006) have been used to assess changes in temperature. A widely accepted conclusion from these various approaches is the unprecedented warmth within the past few decades, relative to a prolonged cooling from ~1350 to 1850, and most likely widespread warmth during medieval times (Esper et al., 2005b). Both proxy reconstruction and model parameterization are often, however, fundamentally linked with relationships to instrumental measurements during a portion of their overlapping period. Time-series of proxy data are generally fit to suitable instrumental records (targets) in a process known as calibration. Implicit to this approach is that the relationship inferred over the calibration interval—roughly the past century or so—is clearly defined, remains stable over time, and that the target is without bias in both the higher and lower frequency domains.

In this regard, recent studies have addressed the importance of properly calibrating to the 'correct' instrumental target, including exact specification of seasonality, spatial representation, and frequency domains for a realistic representation of past climate variability (Osborn and Briffa, 2000; von Storch et al., 2004; Rutherford et al., 2005; Esper et al., 2005a). Despite such general awareness, discussion still exists on proxy/instrumental misfits at regional (Wilson et al., 2005; Frank and Esper, 2005b; Büntgen et al., 2006a) to large scales (Wilson et al., 2007). In the European Alpine region, tree-ring-based reconstructions of past temperature variability (Büntgen et al., 2005, 2006a; Frank and Esper, 2005b) indicated potential discrepancies between "cooler" proxy and "warmer" instrumental data during the 19th century. Such divergence affects the magnitude of inferred past climatic changes as

*Corresponding author. Tel.: +41 44 739 2282; fax: +41 44 739 2215.
E-mail address: frank@wsl.ch (D. Frank).

0277-3791/$ - see front matter
doi:10.1016/j.quascirev.2007.08.002

the relationships defined in the calibration period directly translate into a degrees Celsius scaling for reconstructions. Any discrepancies also question our understanding of climatic variation in general. As proxy data are imperfect substitutes for their targets, proxy/instrumental comparisons allow some of the noise characteristics of proxy archives to be estimated. However, systematic and long-term discrepancies may also provide clues about the instrumental data themselves.

Herein, we focus on such early offset, at both, the large-scale prior to ~1900 and the European Alps before ~1850, by comparing evidence between warm-season temperature measurements and tree-ring-based reconstructions. This is followed by a brief accounting of five hypotheses that may explain this offset, and then by a more extensive discussion. In seeking to evaluate the hypotheses, we attempt to provide indications from research on the possible magnitudes of various factors and their likelihoods for being applicable at different spatial scales. Due to the great uncertainties, we do not belive to conclusively solve this issue, but intend to indicate research directions that should be useful to understand proxy/target offset and to discuss the paleoclimatic importance of such a solution.

2. Large-scale evidence

Four large-scale warm-season weighted tree-ring dominated temperature reconstructions (Jones et al., 1998; Briffa, 2000; Esper et al., 2002; D'Arrigo et al., 2006) have been scaled to the June–August (JJA) mean of the latest $5° \times 5°$ global temperature data set (CRUTEM3; Brohan et al., 2006), averaged over 30–90° N (Fig. 1A). JJA instrumental temperatures show a cooling from ~1850 until 1910, followed by a warming to the 1940s, a cooling to ~1970, with steeply increasing values until present. The proxy-based reconstructions (and their mean) show similar behaviours during the 20th century until their individual end years (1991–1995), but do not indicate relative warmth prior to ~1900. This divergence results in reduced (low-frequency) agreement between the reconstructions and the instrumental summer data. The observed pre-1900 proxy/target divergence is, however, restricted to the summer season, as mean annual temperatures (green in Fig. 1A) are more in line with the lower-frequency tree-ring evidence. Thus, a divergence within the instrumental data themselves is also observed, as early warm season temperature anomalies are much warmer than annual, and particularly

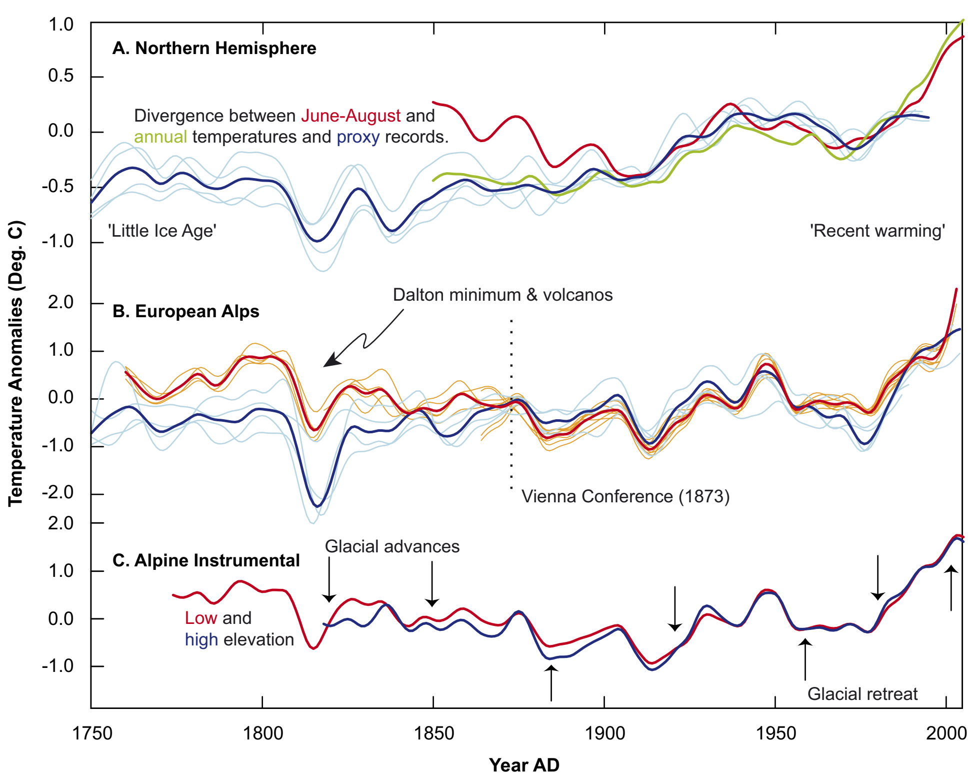

Fig. 1. Warmer early instrumental measurements and colder reconstructed temperatures: (A) comparison of large-scale tree-ring dominated warm-season temperature reconstructions (light-blue; Jones et al., 1999; Briffa, 2000; Esper et al., 2002; D'Arrigo et al., 2006) and their mean (blue), with the June–August (red) and annual (green) mean temperatures for 30–90 °N (CRUTEM3; Brohan et al., 2006). Individual reconstructions scaled to summer temperature data over the 20th century common period; (B) comparison of tree-ring-based warm-season temperature reconstructions from the European Alps (light-blue; Büntgen et al., 2005; Frank and Esper, 2005b; Büntgen et al., 2006b) and their mean (blue), with the individual warm-season temperature targets (orange; Böhm et al., 2001; Auer et al., 2007) and their mean (red). Proxy series and targets as in the original publications. Mean proxy series re-scaled to instrumental mean over the 20th century; (C) mean June–August temperature records from HISTALP (Auer et al., 2007) averaged over the entire low (red) and high (blue) elevational bands across the Greater Alpine Region. Also shown are periods of glacial advance and retreat (arrows). All series smoothed using a 15-year low-pass filter.

cold season instrumental means. Interestingly, the differences between summer and annual data largely disappear during the 20th century, when the seasons display rather synchronous fluctuations (see Parker, 1994; Jones et al., 2003; and Esper et al., 2005a for additional discussion).

Resulting correlation reductions between the reconstructions and warm-season instrumental data have substantial consequences on the estimated amplitude of climatic variations over the past millennium (Esper et al., 2005a), and also seemingly lead to confusion of the exact seasonality of large-scale temperature reconstructions. When calibrating against only the post 19th century data, such reconstructions tend to fit better with the summer data (e.g., Esper et al., 2005a; Wilson et al., 2007), confirming the warm season weighted response well known from regional tree-ring studies. However, inclusion of 19th century data—which possess the warm-season 'peak' not found in any tree-ring reconstruction nor annual temperature data—shifts the 'optimal' target from summer to annual.

Such comparisons are, however, problematic in long-term large-scale approaches because proxy networks currently do not well capture inter-annual variability at hemispheric-scales (Cook et al., 2004; Esper et al., 2005a). We suggest this simply results from the proxy networks not sufficiently representing the area needed to obtain a strong year-to-year signal. As comparisons are thus largely restricted to the mid to low-frequency domains, few degrees of freedom remain to robustly infer proxy/instrumental relationships. At regional-scales such restrictions are largely eliminated as tree-ring data better capture both low and high frequency variations of the target data set.

3. Regional-scale evidence

The longest and highest quality instrumental measurements worldwide can be found in Central Europe (Jones and Moberg, 2003), where several low-elevation measurement series extend to the mid-18th century and some from high-elevations to the early 19th century (Böhm et al., 2001; Auer et al., 2007). These unique observations allow annually resolved proxy records to be calibrated over exceptionally long periods (Frank and Esper, 2005b) and their temporal stability to be tested (Büntgen et al., 2006b; Carrer and Urbinati, 2006). Recent palaeoclimatic efforts in the Greater Alpine Region (GAR; 4–19° E, 43–49° N, 0–3500 m a.s.l.) have resulted in estimations of past temperature variations based on high-elevation tree-ring data, spanning the past centuries (Frank and Esper, 2005b) to millennia (Büntgen et al., 2005, 2006a). A compilation of these four reconstructions and their targets—as used in the original publications—reveals a systematic long-term divergence between warmer early instrumental measurements and colder proxy estimates prior to the mid 19th century (Fig. 1B).

This divergence is seen in four reconstructions irrespective of measurements of different tree-ring parameters, methodological approaches, and exact instrumental targets: two reconstructions are based on the maximum latewood density (MXD) parameter and two are based on ring-width (TRW) measurements; two use principal component regression to merge larger networks of living trees (Frank and Esper, 2005b), and two use averaging and include historical wood samples (Büntgen et al., 2005, 2006b). These four reconstructions are based on two independently developed networks with only a few percent of series shared. Tree-ring data were detrended using either the Regional Curve Standardization (RCS) technique (Esper et al., 2003) to preserve potential low-frequency variability or conventional individual spline fits that are a priori limited in retaining longest time-scale fluctuations (Cook et al., 1995). Tree-ring chronologies were either scaled or regressed over differing periods, using various homogenized datasets of April–September, June–September and June–August temperature means from the Greater Alpine Region (Böhm et al., 2001; Auer et al., 2007).

The instrumental data collectively describe warm conditions during the second half of the 18th century and the onset of the 19th century, followed by a depression in the 1810s–1820s—most likely a consequence of a sequence of volcanic eruptions and the Dalton solar minimum (see also Section 5.2 below)—that is part of a longer cool period until the early 20th century. The trend towards recent warming occurs in two stages, with an initial increase peaking in the 1940s, a cooling until the mid-1970s, and a second warming peaking in the 2003 extreme summer heat (Luterbacher et al., 2004; Schär et al., 2004). The various tree-ring-based growing season temperature reconstructions all display similar decadal fluctuations, but systematically 'undershoot' the early instrumental observations prior to ~1850. No such divergence is found during any other period of proxy/target overlap.

While the instrumental targets shown in Fig. 1B were taken unmodified from the individual publications, Fig. 1C shows the most recent and complete, quality controlled and homogenized HISTALP data set (Auer et al., 2007), averaged for lower and higher elevation belts (below and above 1500 m a.s.l. and representing 118 and 13 stations, respectively) across the Greater Alpine Region. These two series display a similar temperature history, although some evidence for a steeper altitudinal temperature gradient at certain times in the past (e.g., ~1885 and 1830) is displayed—making the recent warming trend slightly larger for the high-elevation network. This comparison highlights the relatively small possible uncertainties related to the representativity of instrumental data taken from the surrounding (more urban) lowlands for comparison with tree-ring-based temperature reconstructions derived from data obtained from near timberline locations. The complete absence of high-elevation station data prohibits any direct conclusions to be drawn prior to 1818 when measurements from the worldwide longest high-elevation meteorological observatory, the Grand St. Bernhard in the Swiss Alps at 7°11′E, 45°52′N, and 2472 m a.s.l., began.

The general pattern of Alpine glacier advance/retreat (Fig. 1C), as indicated by arrows, provides independent proxy evidence for decadal-scale colder/warmer conditions roughly lagged behind temperature changes by a decade or so. While the individual response time depends upon glacier size and geometry, the maximum or near maximum Holocene advance broadly reported in the 1850s (Grove, 1988; Hormes et al., 2001; Holzhauser et al., 2005) and the subsequent retreat (and end of the LIA) help provide indications for lower-frequency climate variations. However, the LIA maximum is not in line with the warmth shown in the early Alpine warm-season temperature data (Oerlemans, 2005; Vincent et al., 2005). Other proxy records, such as documentary data on summer snowfall occurrences (Pfister, 1975, 2007, pers. comm.), are also suggestive of cooler warm-season conditions, i.e. are more consistent with the tree-ring rather than the instrumental evidence.

4. Hypotheses

The observed systematic tree-ring underestimation of early instrumental temperatures calls into question our understanding of the proxy and target data used, with meaningful consequences on the interpretation of climate variability during both instrumental and pre-instrumental times. Multiple hypotheses concerning proxy and target uncertainties can be invoked to explain such patterns, these include:

(1) *Tree-ring detrending*: the removal of the biological age-trend while simultaneously retaining all relevant climatic information is ambitious. This is especially the case when retaining long-term cooling trends, as cooling conditions and the age-trend both result in narrower ring-widths in temperature sensitive tree-ring data. The RCS method overcomes such low-frequency limitations inherent to age-trend estimations derived for each individual tree-ring series, but may still harbor other biases.

(2) *Biological persistence*: annual growth variations, particularly evident in radial increments, often reflect a biological memory (via e.g., needle generations, carbohydrate reserves) to previous year conditions. Related feedbacks or physiological non-linearities may shift the spectral-properties of the inferred climatic variable from a "whiter" to "redder" process.

(3) *Climate/growth relationships*: uncertainties in the choice of the proper target season (e.g., JJA, MJJA, AMJJAS, Annual), climatic variable (e.g., precipitation, solar radiation, mean or maximum temperatures), and instabilities whereby the relative mixture of separate environmental variables that drive tree-growth varies with time, exist.

(4) *Data availability*: the quantity of instrumental data does not remain stable over time. Generally, datasets become sparser back in time and less spatially representative. Changes in data quantities directly affect the signal quality of any mean and its local variance and reduces possibilities for independent comparisons and verification.

(5) *Data homogeneity*: observational data may contain much non-climatic noise caused by station relocation, changes in instruments and their screens, changes in recording times and observers, algorithms for the calculation of means, and other changes at the site of measurement. Errors from these sources are often neither randomly distributed in time nor in direction.

5. Discussion

5.1. Tree-ring detrending

It has been long recognized that the necessary removal of all age-related growth-trends via the tree-ring detrending process can dramatically alter the long-term trends and spectral properties of tree-ring chronologies and thus the resulting climate reconstructions (e.g., Fritts, 1976; Jacoby and D'Arrigo, 1989; Esper et al., 2002). Conventional removal of the age-trend using individual series fits does not allow climatic fluctuations longer than the median tree age to be preserved (Cook et al., 1995). This is particularly limiting where a chronology is composed of living and ancient (e.g. from buildings or moraines) material that together make a chronology significantly longer than its constituent segments. Age-related detrending methods, including the RCS (Briffa et al., 1992; Esper et al., 2003) and Age-Banding (Briffa et al., 2001) techniques help overcome this segment-length constraint, but themselves may not be entirely free from other biases (Melvin, 2004; Helama et al., 2005). Subtle population differences between living and relict populations potentially impart long-term trends (Esper et al., 2003), or the preferred selection of, for example, the oldest or largest living trees in a stand may yield biases related to tree mortality and growth rates (Melvin, 2004).

Inspection of differences between studies that used both age-related and conventional detrending methods (e.g., Wilson et al., 2005; D'Arrigo et al., 2006) or analysis of common features observed in independent data sets (or data set fractions) (e.g., Esper et al., 2002; Büntgen et al., 2005; Frank and Esper, 2005b; and Fig. 1B above) may provide indications that inferred climatic fluctuations are robust. However, it is unfortunate that the 'correct' detrending solution generally cannot be gleaned from calibration/verification statistics nor from first principles (Wilson et al., 2007). Efforts to include detrending uncertainties in climatic reconstructions, by estimation of the scatter about reasonable possible detrending choices, is an area where some quantification is possible (see Esper et al., 2007 for details).

The independence of the various Alpine reconstructions shown in Fig. 1B, including the utilization of both individual spline fits, and RCS techniques, suggests that the longer-term trends displayed are reasonably robust with respect to maximum wavelength allowed by the conventional detrending applied and conversely suggest that any possible RCS-related biases to be minimal. The central tendency that all temperature reconstructions undershoot the early instrumental data remains. At hemispheric levels, some data overlap cannot be disregarded for any large-scale studies shown in Fig. 1A, however, a new independent reconstruction by Wilson et al. (2007) that only utilizes tree-ring data not included in previous large-scale approaches displays similarly cool conditions during the 19th century, and similarly undershoots early warm-season large-scale instrumental data. Based on these independent findings, we suggest that the array of currently applied detrending methods in combination with the various tree-ring compilations utilized, are unlikely to fully account for such a systematic divergence at either the Alpine or hemispheric scales.

5.2. Biological persistence

As tree growth may integrate effects from previous year climatic and ecological conditions (e.g., Fritts, 1976; D'Arrigo et al., 1992; Frank and Esper, 2005a; Kagawa et al., 2006), it is plausible that biological-induced autocorrelation results in different signal-to-noise ratios in inter-annual or multi-decadal frequencies and may "redden" the spectral characteristics of proxy data. When these data are calibrated, the low-frequency variability may subsequently be overestimated. The effects of this spectral shifting are similar (but opposite) to those described by von Storch et al. (2004) and Osborn and Briffa (2004), where low-frequency variability may be underestimated when white noise degrades the proxy signal.

Such autocorrelation most likely results from the utilization of abundant carbohydrates stored towards the end or even after the growing season, or from longer-term benefits (or detriments) in mobilizing resources from root and needle growth following good (or poor) years (Tranquillini, 1964; Kozlowski and Pallardy, 1997). While approaches to remove the biological autocorrelation via autoregressive modeling may be successful in yielding high correlations with year-to-year climate variations, it is clear that significant low-frequency variability is lost during most such approaches, making them limited in addressing longer-term climate change issues (Cook, 1985). We thus encourage caution in producing the so-called pre-whitened or residual chronologies when addressing all kinds of longer term, global change related processes. Attempts to explicitly consider biological autocorrelation by utilization of different regression models for various frequency domains (Osborn and Briffa, 2000) did not reveal significant differences between regression coefficients for the lower and higher frequencies using MXD data from the European Alps (Büntgen et al., 2006a).

TRW tends to be more prone to biological autocorrelation and site-ecological differences in comparison to MXD measurements (e.g., Frank and Esper, 2005a). In this regard, the early 19th century is of particular interest in the European Alps and Northern Hemisphere. This period is characterized by a sequence of larger volcanic eruptions between 1808 and 1815, most likely resulting in an accumulated aerosol summer cooling effect (Robock, 2000; and references therein), in concert with reduced solar radiation during the Dalton minimum (Eddy, 1976), producing a distinct negative radiative-forcing anomaly. For the European Alps, the most pronounced anomaly follows one year after the Tambora/Indonesia event in April 1815 (Sigurdsson and Carey, 1989; Oppenheimer, 2003) and approximately seven years after a large eruption with an unknown location (Dai et al., 1991), with the year 1816 becoming known as 'the year without summer' (Harrington, 1992; Robock, 1994).

Early Alpine instrumental measurements show average JJA temperatures from 1813 to 1816 to be exceptionally cold, with a warmer 1817–1820 interlude prior to the return of the cool 1821 summer (Fig. 2). This distinctive temperature depression followed a series of warmer summers in the first decade of the 19th century (Luterbacher et al., 2004; Casty et al., 2005). In the wider AMJJAS seasonal window, often regarded as the optimal response interval for MXD, 1816 is substantially colder than the preceding years and 1821 warmer than the JJA anomaly. Importantly, temperature sensitive MXD records tend to show this same pattern, whereas for TRW records, 1821 is generally the coldest summer reconstructed. Note that for illustration purposes in Fig. 2 the offset between the tree-ring and instrumental data has been removed by scaling (see caption for details). The variations displayed by TRW and JJA temperatures suggest that in this case the TRW response is exaggerated towards colder conditions, as the trees had not fully been able to recover from possible damage and stress following the sequence of severe summers from 1813 to 1816, and likely utilized carbohydrates for purposes other than radial expansion.

While the early 1800s serve as an extreme test, the general characteristics of trees' responses are likely still applicable during less climatically severe periods. Radial growth variations of different species might also tend to display characteristic patterns, with for example, higher *lag-1* autocorrelations found for *Pinus cembra* than *Larix decidua* in the Alpine network used by Frank et al. (2005). Additional research efforts in the timing of carbon allocation via tracer studies (Kagawa et al., 2006) and methods to better quantify the spectral properties of tree-ring and other proxy data (e.g., Osborn and Briffa, 2000; Mohammad and Moberg, 2007) will allow a better differentiation of biological and climatic persistence.

5.3. Climate/growth relationships

The great variety of environmental variables that can affect tree-growth may confound the isolation of a single meteorological variable as a suitable reconstruction target. Sampling locations that approach the natural distribution

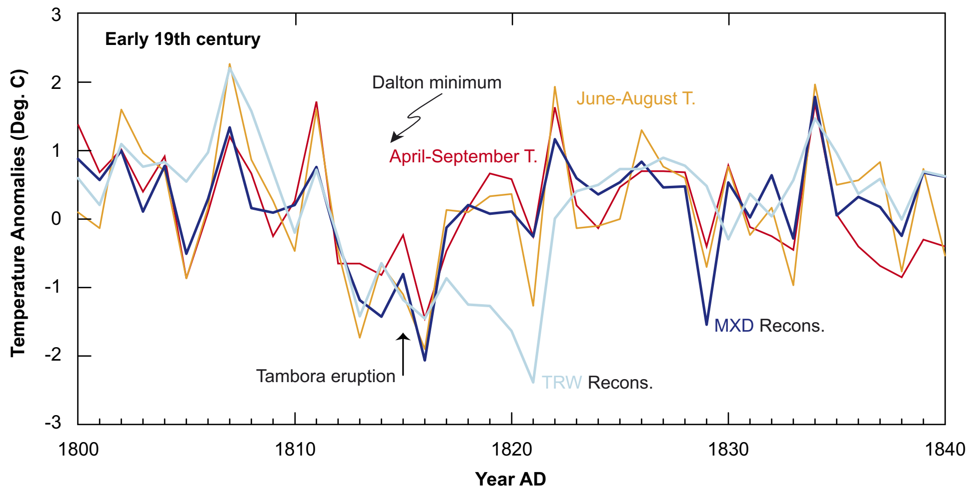

Fig. 2. Comparison of regional-scale Alpine tree-ring based warm-season temperature reconstructions using means of various ring width (light-blue) and density (blue) records (Büntgen et al., 2005, 2006b; Frank and Esper, 2005b), with the June–August (orange) and April–September (red) temperature targets from HISTALP (Auer et al., 2007). Mean proxy series are scaled against the instrumental target means over the 1800–1840 period. Lag-1 autocorrelations of the ring width and density proxies are 0.73 and 0.23. Lag-1 autocorrelations of the June–August and April–September target are 0.02 and 0.18. Correlations between the ring width (density) and June–August (April–September) temperature data are 0.62 (0.78). All values are calculated over the 1800–1840 period.

limits, however, tend to most closely meet this objective (Fritts, 1976; Körner, 1998), which is why tree-ring data only from sites near the thermal timberline should be used for simple, yet robust temperature reconstructions. Nevertheless, the growing season duration is likely somewhat plastic in accordance with environmental conditions, and trees are additionally not necessarily inclined to obey the Gregorian calendar. Thus, a reasonably well-calibrated and verified reconstruction could be developed using different environmental targets (e.g., mean JJ or JJA temperatures). Fig. 2 highlights some of the similarities in seasonal instrumental targets whereby the JJA and AMJJAS means correlate with each other at 0.78 over the period 1800–1840. The relationship jumps to 0.95 when comparing the more closely related JJAS and JJA windows. The optimal target season selected based on calibration results might, thus, not be easy to define for a given time period and additionally might even vary over time.

The above examples and the majority of tree-ring and other proxy studies conducted to date have focused on the most-readily available mean temperature target. However, some dendroclimatic evidence suggests closer relationships between tree-growth and maximum temperatures, in comparison to mean temperatures (Wilson and Luckman, 2003; Büntgen et al., 2007; Frank et al., 2007b; Wilson et al., 2007; see also D'Arrigo et al., 2007). Maximum temperatures more closely represent daytime conditions—including sunlight availability for photosynthesis—than mean temperatures. The negative correlation between summer precipitation and temperature, and the positive correlation between sunshine hours and daytime temperatures are both examples of data that are correlated in the inter-annual domains, yet may show differing long-term trends and spectral properties (Auer et al., 2007). The influence upon growth of such correlated factors and the selection of only one instrumental parameter as a reconstruction target may be responsible for additional uncertainties in proxy/target relationships.

In addition to uncertainties in the proper season and/or variable, the relative influence of environmental factors on tree-growth may shift over time. Several recent studies indicated temporal instability in the amount of explained variance between tree-growth and temperature, including studies that have shown an apparent decrease in the ability of TRW and MXD sites, primarily from the higher northern latitudes, to track recent warming (Briffa et al., 1998; Cook et al., 2004; D'Arrigo et al., 2007 and references therein). However, any exceptional inabilities of large-scale reconstructions to track temperatures during the recent warming seem to be of secondary importance to the choice of a summer or annual instrumental target when data prior to the 20th century are considered (Fig. 1A). Evidence from the Alpine region suggests that several tree-ring sites may now be less temperature sensitive due to increased late summer drought stress in the wake of temperature increases during the past decades (Büntgen et al., 2006b; Carrer and Urbinati, 2006). Such instabilities are potentially species specific, more common to TRW than MXD, and may have profound implications in the utilization of some tree-ring data as temperature proxies. It should be noted that the reconstructions presented in Fig. 1B passed tests for temporal stability in the climate signal. Furthermore, approaches considering, for example, individual monthly data are more likely to reveal instabilities than those where the target variable is a seasonal average.

Field studies that focus on the timing of intra-annual cell formation, growth and lignification (Rossi et al., 2006, 2007; Fonti et al., 2007) and a concentration on obtaining data from only the most ecologically appropriate sites will yield a better understanding of climate related changes captured in tree-ring data. Additional efforts to identify growth limitations either via empirical studies or via forward modelling approaches (e.g., Anchukaitis et al., 2006; Evans et al., 2006) could be helpful to identify the most relevant target seasons, parameters, and potential instabilities therein. Further efforts to develop high-quality instrumental data sets including a variety of parameters (e.g., maximum temperatures plus daily or weekly resolution) are required to most accurately quantify and assess the meteorological metrics and periods critical to tree-growth.

5.4. Data availability

It should be evident from the above discussion, that knowledge of tree-ring or more generally proxy-based estimates of past climatic and/or ecological conditions is fundamentally coupled with relations quantified and inferred from analyses with instrumental data. The European Alps are perhaps unique in this respect, as it is arguably the only region worldwide where high-elevation temperature sensitive tree-ring data can be compared with local instrumental stations from similar elevations over more than 150 years.

A reduction in instrumental data coverage back in time is, however, also evident for the Alpine region, and even more valid on larger-scales. In the GAR, for example, 36 stations have temperature data back until 1850, but only 16 extend prior to 1800 (Auer et al., 2007). Similarly for the NH land and sea areas, about 80% of $5^\circ \times 5^\circ$ grid-boxes in the most recent version of the CRU data set (Brohan et al., 2006) contain data in 1970s, with this figure dropping to ~40% around 1900 and <20% in 1860, with most of the early global land data concentrated in Europe (Fig. 3). In 1850, the land dataset represents about 5% of the total possible grid-boxes for the NH. Also notable, is the decline towards present in the grid-box percentage since the 1970s. Such decline results in a reduction in the portion of the

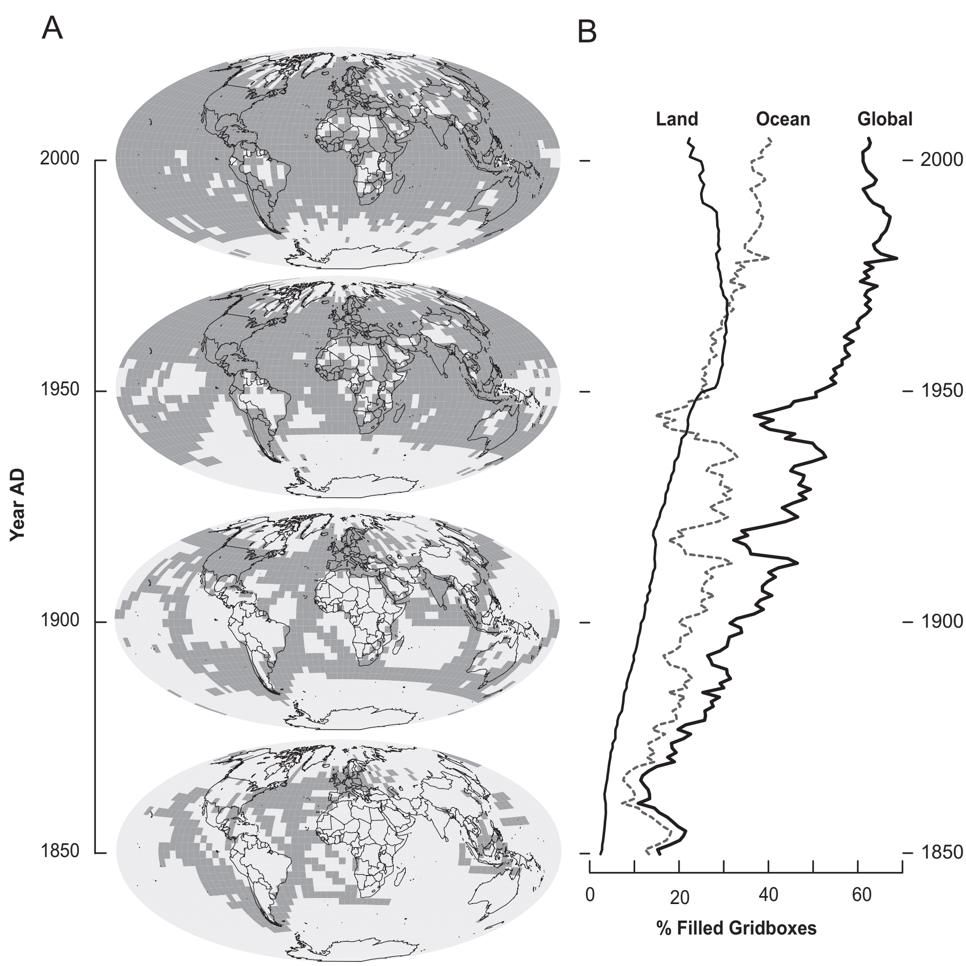

Fig. 3. Spatial and temporal coverage of the HadCRU3 data set (Brohan et al., 2006): (A) maps showing the filled gridboxes in June 1850, 1900, 1950, and 2000; (B) time-series of the percentage of filled gridboxes for the global, land, and ocean areas during June. The coverage in the sea-only gridboxes was estimated by the difference between the HadCRUT3 and the CRUTEM3 data sets, where as the land coverage is simply represented by the CRUTEM3 data set. The percentage of land (ocean) coverage is over- (under-) estimated using this approach as a coastal land station may "fill" a gridbox that is predominately over the sea. Seasonal patterns in the data coverage, for example, tend to show greater sea traffic in the Northern and Southern Hemispheres during the boreal and austral summers.

climate system observed, while simultaneously both increasing the weight of single stations in explaining the surrounding climatic variations and decreasing the number of inter-station comparisons available to properly identify and correct measurement errors.

In considering such changes in the spatial and temporal coverage of the earth's surface significant progress utilizing instrumental data and/or GCM simulations have been made in the quantification of uncertainties and sampling errors (e.g., Jones et al., 1997; Folland et al., 2001; Brohan et al., 2006; Rayner et al., 2006). Such errors depend upon the time-scale of investigation; for 21-year smoothed annual data covering the extra-tropics, the 95% confidence limits are about 0.4 °C, with errors increasing for annual and monthly values (Brohan et al., 2006). Due to the relative high station density, uncertainties are generally lower in the GAR compared to most other parts of the world, e.g., Asia, Africa and South America. We note that while replication changes commonly affect the qualities of proxy data (e.g., Frank et al., 2007a), this consideration is mitigated by the great number of tree-ring data sets in this recent time period in the Alps, and generally small replication changes during this recent and short time period (relative to tree-ring chronologies) for both the Alps and the NH. However, as noted above, the limited spatial coverage of current large-scale temperature reconstructions remains a critical issue.

Besides the spatial component of meteorological data availability back in time, a particular reduction when considering higher elevations is most evident outside the European Alps and prior to the 20th century (Diaz and Bradley, 1997). In the Alps, the altitudinal difference between stations tends to be more important than the horizontal distance in determining how well they correlate (Böhm et al., 2001). Notably, the higher elevation stations (>1500 m a.s.l.), generally corresponding to the elevational band of the tree-ring data, tend to detect more of the "background" climate, with local variations (i.e., noise) being larger at lower elevations (Böhm et al., 2001). See Böhm et al. (2001) and Auer et al. (2007) for more detailed discussions related to elevation effects and their fingerprints on instrumental measurements.

The combined uncertainty resulting from data reductions is considerable and might contribute to understanding the discrepancies between proxy and instrumental datasets. It is generally assumed that such uncertainties are symmetric and do not contain biases, but this might not always be the case. For example, the ratio of "land" to "ocean" grid cells as represented by the differences in coverage between the HadCRUT3 and CRUTEM3 datasets varies with time; around 1880 this ratio is about 0.8 and in the 1960s is about 1.4 (Fig. 3B). Brohan et al. (2006), interestingly discuss how for smoothed global averages, the uncertainties in the 1850s are actually smaller than in 1920! This has to do with specific homogenization biases that depend, not only when, but also where data are available. In this case, the tropical measurements that enter the global dataset in the early 20th century result in increased uncertainties (see below) even though the spatial coverage increases. It is notable, that Brohan et al. (2006) have provided confidence limits where the mean temperature estimate is not centered between the upper and lower confidence limits.

5.5. *Data homogeneity*

The inability to keep all aspects of temperature data acquisition constant in time requires a careful homogenization of instrumental data to eliminate non-climatic noise (e.g., Peterson et al., 1998). An idealized random distribution in the direction and size of data errors is complicated by the occurrence of systematic biases in instrumental measurements and their surroundings that might result in artificially induced long-term trends. The change in raw measurement values from the necessary homogenization applied can vary greatly in amplitude and even direction depending upon individual station histories, available metadata and its interpretation, and methodology.

A comparison of the relevant Alpine grid-cells from the ALPCLIM (Böhm et al., 2001) and CRU (Jones et al., 1999) data showed a factor of two differences in the annual linear trend from 1890 to 1998 with values of 1.1 and 0.55 °C, respectively (Böhm et al., 2001). Such differences likely largely stem from the two different homogenization approaches that have either relied extensively on metadata due to their regional nature or more statistically based approaches covering larger regions. Even for efforts confined to a similar region, differences due to the interpretation of metadata and researchers' homogenization criteria may be considerable. See Brunetti et al. (2006) for discussion of differences between the HISTALP (Auer et al., 2007) and an Italian national dataset of the same area. On the hemispheric-scale, the great number of stations largely prohibits individual examination of series and the collection and analysis of relevant metadata (see Peterson et al., 1998).

The amount of inhomogeneities does not remain constant in time. For example, the annual rate, magnitude and frequency of detected outliers constantly increases prior to ~1980 for the Greater Alpine Region (Auer et al., 2007), with a sharp decrease in data quality in the late 19th century. This decline in data quality corresponds to measurements performed before greater station network standardization following the Vienna Conference in 1873, when the predecessor of the World Meteorological Organization set up rules and guidelines for instrumental observations (Maugeri et al., 2002a, b). Data gaps are generally more episodic, typically as a function of political and economic circumstances. Alpine instrumental uncertainties prior to the 19th century are characterized by scarcer metadata, the large quantity of inhomogeneities and errors in almost all the existing records, and lower station density. Nevertheless, for a given time slice, the case of Alpine temperatures has to be regarded as less sensitive to homogeneity problems compared to temperature records elsewhere. The rather high station network density in relation to spatial correlation decay lengths (even towards the earliest recording period and at higher elevations), the relative richness of metadata, and also its division into a patchwork of national and sub-national data providers make the danger of non-detectable systematic shifts less likely in comparison to homogenization attempts in, for example, many parts of Asia, Africa and South America, and particularly in hemispheric-scale approaches.

At larger spatial scales, the relevance of common systematic biases in the instrumental record increases as these errors are reinforced, rather than mitigated, by the averaging process. The most important sources for systematic errors concern changes in thermometers and their shelters (Parker, 1994; Nørdli et al., 1997), changes in sea-surface temperature measurement practice and instrumentation (Folland and Parker, 1995; Folland et al., 2001; Folland, 2005; Rayner et al., 2006; Kent et al., 2007), with other examples including urban warming (Karl et al., 1988; Arnfield, 2003) and station movements (Peterson et al., 1998; Böhm et al., 2001). The degree of homogenization necessary to correct for such features is often unclear as entire networks may contain the same bias leaving few possibilities to identify and correct such inhomogeneities.

Perhaps of particular relevance to the offset between proxy data and early warm season measurements, regional studies from Sweden (Klingbjer and Moberg, 2003; Moberg et al., 2003) and more broadly Scandinavia (Nørdli et al., 1997), Ireland and Scotland (Jones and Lister, 2004), the United States (Chenoweth, 1993), Australia (Nicholls et al., 1996), and Spain (Brunet et al., 2006) suggested that early warm-season measurements might be subject to systematic warm biases due to insufficient sheltering from direct or reflected radiation. The detailed comparison on the early instrumental shelters for Scandinavia by Nørdli et al. (1997) suggests warm-season biases up to 0.4 °C (including measurements from the so-called free standing Swedish shelter used at some stations until the 1960s). Based on relationships with other measured climate variables including cloud cover and comparisons with other long instrumental records, Moberg et al. (2003) suggested that prior to 1860 summer temperatures were 0.5–0.8 °C cooler than those measured in Sweden. Brunet et al. (2006) present comparisons between the early Moutsouris and modern Stevenson screen for two stations in Spain where the early screens are positively biased by more than 1.3 °C on an annual basis for maximum temperatures and negatively biased by less than 0.3 °C for minimum temperatures. Interestingly, they also provide the dates that Stevenson screens were installed at 20 Spanish stations: these range from 1894 to 1915. Unfortunately, a comprehensive global compilation of such relevant metadata is missing. The benchmark study in this regard is, however, that from Parker (1994). He presents a global summary of the thermometer exposures representative for different countries at different time periods and indicates that the period from the 1870s to 1910s was characterized by large changes in thermometer exposures. According to this study, prior to the late 19th century warm-season data from the extra-tropics are likely positively biased by about 0.2 °C as a result of exposure changes. In addition to these mid to high latitude exposure changes, it is also suggested that tropical temperatures prior to the 1930s were likely positively biased based on comparisons between Stevenson screens and the so-called tropical shed exposures and between land and nearby sea-surface temperature data (Parker, 1994). Importantly, most of the sheltering biases result in greater seasonality with positive biases during summer and little or negative biases during winter.

The homogenization necessary to account for the changes in observing practices, environments and instrumentation easily are of the magnitude of the offset between the tree-ring reconstructions and instrumental temperatures. For example, in the GAR over 1.0 °C were added to the April–September data from the Swiss data subset (Böhm et al., 2001), which is in the order of the offset as seen in Fig. 1B. Homogenization corrections found in larger-scale approaches (e.g., GHCN, Peterson and Vose, 1997; GISS, Hansen et al., 1999) are also often greater than 1 °C. Such changes seemingly place tree-ring estimates within uncertainties of the early instrumental data. The bias direction for insufficient sheltering in both the extra-tropics and tropics is in line with the tree-ring evidence. Furthermore, the larger impact in summer of these biases is consistent with the divergence between warm and cold-season instrumental observation, which makes the long-term fit between tree-ring data and annual (or even winter) temperatures more similar (Fig. 1A). The current range of estimates for the warm-season thermometer shelter bias (+0.2 to +0.8 °C) is on the same order of the offset between proxy and early warm-season instrumental data at the Alpine and Hemispheric scales. However, additional studies where simultaneous measurements are taken at present and former locations with newer and older thermometers and shelters should provide more insights into the specific factors, environmental conditions, and geographic locations that might result in larger or smaller biases.

6. Conclusions

Independent of the tree-ring dataset, measured parameter, age-trend removal, reconstruction methodology, and the calibration period, season, and approach, all of the Alpine reconstructions in Fig. 1B systematically and continuously undershoot their early instrumental target before ~1850, with a similar consensus derived from large-scale reconstructions before ~1900 (Fig. 1A). Many possible hypotheses can be invoked to explain the offset between the early instrumental warm season data and tree-ring estimates. While, full quantification of the likelihood and magnitude of each of these factors remains a central

Table 1
Summary of the factors that can potentially account for the proxy/instrumental offset and their relevance at the Alpine and Hemispheric scales

	Effect	Relevance NH	Relevance Alps		Comments
Detrending	Trend uncertainties.	+	+ +		Alps: longer comparison period (back to 1760) increases importance.
Biological autocorrelation	Decreased annual relationships and frequency dependent signals.	+	+ + +	(TRW) (MXD)	NH: Calibrations effectively only conducted in low-frequency domain.
Growth/climate (a) Target choice (b) Mixed/unstable signal	 (a) Diffuse proxy specification. (b) Error-term non-random.	 + + + + + +	 + + + + +	 (TRW) (MXD)	 NH: Seasonal differences prior to 1900 critical in instrumental data.
Station decrease	Reduced degrees of freedom and spatial representation.	+ + +	+		Early data from surrounding urban areas (Alps) and SST dominated (NH).
Homogenization	Trend uncertainties.	?	+ + + +		NH: limited application of metadata.

Abbreviations: NH—Northern Hemisphere, but refers loosely to large-scale approaches ; TRW—Tree ring-width ; MXD—Maximum latewood density. +, + +, + + +, and + + + + symbols refer to 'low', 'low-medium', 'medium-high', and 'high', respectively. ? Denotes 'unknown'.

question that cannot yet be satisfactorily addressed, the synthesis in Table 1 suggests the importance of choosing the proper instrumental target at large-scales, and homogenization at regional and possibly at hemispheric scales. If the homogenization indeed turns out to be a critical factor, the difference in timing between the observed divergence in the European Alps and from hemispheric comparisons is consistent with regional historical differences in thermometer exposures, station histories and practices, data availability, and subsequent differences in homogenization procedures performed at regional and global levels.

In many senses the homogenization procedure parallels tree-ring detrending: both techniques are required to eliminate non-climatic noise components of the respective time-series; and there appears to be no universally applicable or justifiable solution, which results in increased uncertainties in longer-term trends. Necessary homogenization and ongoing changes to the instrumental datasets, makes calibration of proxy records somewhat analogous to shooting at a moving target. This is particularly critical for hemispheric-scale approaches where strong trend differences between warm-season and annual records and little to no skill in the high-frequency domain lead to great uncertainties in proxy calibration. If the early instrumental warm season warmth turns out to be an artifact, its correction will result in an increased industrial warming trend, thereby affecting both the amplitude of reconstructed (past) temperature fluctuations and the model estimates of (future) change. Furthermore, such a correction might largely eliminate seasonal differences, thereby having implications on the paradigm of winter *versus* summer trends in perhaps not only the past, but also for projected warming.

Even though our understanding of recent climatic changes is clearly most limited during the early observational period, current political and economic developments such as reduced scientific funding and the collapse of the Eastern Bloc have also resulted (at least temporarily) in the deterioration in the land based observational network. While on-going changes to instrumentation and the automatization of meteorological observatories (e.g., Quayle et al., 1991; Nørdli et al., 1997) and more newly recognized biases in old instrumental series (Parker and Horton, 2005) require attention, the described inconsistencies between instrumental temperature series and tree-ring proxies should provide an impetus to intensify research in the early instrumental period. Efforts are currently underway to analyze and correct for early shelter biases in the GAR (Böhm et al., in prep.), with further research, including other measured climate variables such as precipitation, cloud cover, and surface air pressure and additional proxies being important to clarify the critical transition interval from proxy to instrumental evidence (e.g., Hiebl, 2006). With broader attempts, a deeper statistically and physically based understanding will provide a more seamless bridge between the climatic variability measured in recent decades and that reconstructed over the past centuries to millennia. Ultimately, an improved quantification of the exact magnitude of past climatic variation should be achieved by further investigations, thereby allowing better constraints of the earth's climate sensitivity and the role of past and present forcing agents.

Acknowledgements

We thank Rosanne D'Arrigo, Neil Roberts, and two anonymous reviewers for helpful comments. Supported by the European Union (ALP-IMP ♯EVK-CT-2002-00148 and Millennium ♯017008) and the Swiss National Science Foundation (NCCR-Climate and Euro-Trans).

References

Anchukaitis, K.J., Evans, M.N., Kaplan, A., Vaganov, E.A., Hughes, M.K., Grissino-Mayer, H.D., Cane, M.A., 2006. Forward modelling of regional-scale tree-ring patterns in the southeastern United States and the recent influence of summer drought. Geophysical Research Letters 33, L04705.

Arnfield, A.J., 2003. Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. International Journal of Climatology 23, 1–26.

Auer, I., 31 Co-authors, 2007. HISTALP—Historical instrumental climatological surface time series of the Greater Alpine REGION. International Journal of Climatology 27, 17–46. doi:10.1002/joc.1377.

Böhm, R., Auer, I., Brunetti, M., Maugeri, M., Nanni, T., Schöner, W., 2001. Regional temperature variability in the European Alps: 1760–1998 from homogenized instrumental time series. International Journal of Climatology 21, 1779–1801.

Briffa, K.R., 2000. Annual climate variability in the Holocene: interpreting the message of ancient trees. Quaternary Science Reviews 19, 87–105.

Briffa, K.R., Jones, P.D., Bartholin, T.S., Eckstein, D., Schweingruber, F.H., Karlén, W., Zetterberg, P., Eronen, M., 1992. Fennoscandian summers from AD 500: temperature changes on short and long time scales. Climate Dynamics 7, 111–119.

Briffa, K.R., Schweingruber, F.H., Jones, P.D., Osborn, T.J., Shiyatov, S.G., Vaganov, E.A., 1998. Reduced sensitivity of recent tree-growth to temperature at high northern latitudes. Nature 391, 678–682.

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

Brohan, P., Kennedy, J.J., Harris, I., Tett, S.F.B., Jones, P.D., 2006. Uncertainty estimates in regional and global observed temperature changes: a new data set from 1850. Journal of Geophysical Research 111, D12106.

Brunet, M., Saladié, O., Jones, P.D., Sigró, J., Aguilar, E., Moberg, A., Lister, D.H., Walther, A., Lopez, D., Almarza, C., 2006. The development of a new dataset of Spanish daily adjusted temperature series (SDATS) (1850–2003). International Journal of Climatology 26, 1777–1802.

Brunetti, M., Maugeri, M., Monti, F., Nanni, T., 2006. Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series. International Journal of Climatology 26, 345–381.

Büntgen, U., Esper, J., Frank, D.C., Nicolussi, K., Schmidhalter, M., 2005. A 1052-year tree-ring proxy of Alpine summer temperatures. Climate Dynamics 25, 141–153.

Büntgen, U., Frank, D.C., Nievergelt, D., Esper, J., 2006a. Summer temperature variations in the European Alps, AD 755–2004. Journal of Climate 19, 5606–5623.

Büntgen, U., Frank, D.C., Schmidhalter, M., Neuwirth, B., Seifert, M., Esper, J., 2006b. Growth/climate response shift in a long subalpine spruce chronology. Trees Structure and Function 20, 99–110.

Büntgen, U., Frank, D.C., Grudd, H., Esper, J., 2007. Eight centuries of Pyrenees summer temperatures from tree-ring density. Climate Dynamics (in review).

Carrer, M., Urbinati, C., 2006. Long-term change in the sensitivity of tree-ring growth to climate forcing in Larix decidua. New Phytologist 170, 861–872.

Casty, C., Wanner, H., Luterbacher, J., Esper, J., Böhm, R., 2005. Temperature and precipitation variability in the European Alps since 1500. International Journal of Climatology 25, 1855–1880.

Chenoweth, M., 1993. Nonstandard thermometer exposures at US cooperative weather stations during the late nineteenth century. Journal of Climate 6, 1787–1797.

Cook, E.R., 1985. A time series analysis approach to tree-ring standardization. Ph.D. Dissertation, University of Arizona, Tucson, AZ.

Cook, E.R., Briffa, K.R., Meko, D.M., Graybill, D.A., Funkhouser, G., 1995. The 'segment length curse' in long tree-ring chronology development for palaeoclimatic studies. The Holocene 5, 229–237.

Cook, E.R., Esper, J., D'Arrigo, R., 2004. Extra-tropical Northern Hemisphere temperature variability over the past 1000 years. Quaternary Science Reviews 23, 2063–2074.

Dai, J., Mosley-Thompson, E., Thompson, L.G., 1991. Ice core evidence for an explosive tropical volcanic eruption six years preceding Tambora. Journal of Geophysical Research 96, 17,361–17,366.

D'Arrigo, R.D., Jacoby, G.C., Free, R.M., 1992. Tree-ring width and maximum latewood density at the North American treeline: parameters of climatic change. Canadian Journal of Forest Research 22, 1290–1296.

D'Arrigo, R., Wilson, R.J.S., Jacoby, G., 2006. On the long-term context for late 20th century warming. Journal of Geophysical Research 111, D03103.

D'Arrigo, R., Wilson, R., Liepert, B and Cherubini, P., 2007. On the 'Divergence Problem' in Northern forests: a review of the tree-ring evidence and possible causes. Global and Planetary Change (in press).

Diaz, H.F., Bradley, R.S., 1997. Temperature variations during the last century at high elevation sites. Climatic Change 36, 253–279.

Eddy, J.A., 1976. The maunder minimum. Science 192, 1189–1202.

Esper, J., Cook, E.R., Schweingruber, F.H., 2002. Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. Science 295, 2250–2252.

Esper, J., Cook, E.R., Krusic, P.J., Peters, K., Schweingruber, F.H., 2003. Tests of the RCS method for preserving low-frequency variability in long tree-ring chronologies. Tree-Ring Research 59, 81–98.

Esper, J., Frank, D.C., Wilson, R.J.S., Briffa, K.R., 2005a. Effect of scaling and regression on reconstructed temperature amplitude for the past millennium. Geophysical Research Letters 32, L07711.

Esper, J., Wilson, R.J.S., Frank, D.C., Moberg, A., Wanner, H., Luterbacher, J., 2005b. Climate: past ranges and future changes. Quaternary Science Reviews 24, 2164–2166.

Esper, J., Frank, D.C., Büntgen, U., Verstege, A., Luterbacher, J., Xoplaki, E., 2007. Long-term drought severity changes in Morocco, Northwestern Africa. Geophysical Research Letters 34, doi:10.1029/2007GL030844.

Evans, M.N., Reichert, B.K., Kaplan, A., Anchukaitis, K.J., Vaganov, E.A., Hughes, M.K., Cane, M.A., 2006. A forward modeling approach to paleoclimatic interpretation of tree-ring data. Journal of Geophysical Research 111, G03008.

Folland, C.K., 2005. Assessing bias corrections in historical sea surface temperature using a climate model. International Journal of Climatology 25, 895–911.

Folland, C.K., Parker, D.E., 1995. Correction of instrumental biases in historical sea surface temperature data. Quarterly Journal of the Royal Meteorological Society 121, 319–367.

Folland, C.K., Rayner, N.A., Brown, S.J., Smith, T.M., Shen, S.S.P., Parker, D.E., Macadam, I., Jones, P.D., Jones, R.N., Nicholls, N., 2001. Global temperature change and its uncertainties since 1861. Geophysical Research Letters 28, 2621–2624.

Fonti, P., Solomonoff, N., Garcia Gonzalez, I., 2007. Earlywood vessels size of Castanea sativa records temperature before their formation. New Phytologist 173, 562–570.

Frank, D., Esper, J., 2005a. Characterization and climate response patterns of a high-elevation, multi-species tree-ring network for the European Alps. Dendrochronologia 22, 107–121.

Frank, D., Esper, J., 2005b. Temperature reconstructions and comparisons with instrumental data from a tree-ring network for the European Alps. International Journal of Climatology 25, 1437–1454.

Frank, D., Wilson, R., Esper, J., 2005. Synchronous variability changes in alpine temperature and tree-ring data over the last two centuries. Boreas 34, 498–505.

Frank, D., Esper, J., Cook, E.R., 2007a. Adjustment for proxy number and coherence in a large-scale temperature reconstruction. Geophysical Research Letters 34, doi:10.1029/2007GL030571.

Frank, D., Ovchinnikov, D., Kirdyanov, A., Esper, J., 2007b. The potential for long-term climatic reconstructions in the central Altay mountains from living and relict larch. In: Haneca, K., et al. (Eds.), Tree Rings in Archaeology, Climatology and Ecology, TRACE, vol. 5, pp. 85-96.

Fritts, H.C., 1976. Tree Rings and Climate. Academic Press, London pp 567.

Grove, J.M., 1988. The Little Ice Age. Methuen & Co., 498 pp.

Hansen, J., Ruedy, R., Glascoe, J., Sato, M., 1999. GISS analysis of surface temperature change. Journal of Geophysical Research 104, 30,997–31,022.

Harrington, C.D., 1992. The Year without a Summer? World Climate in 1816. Canadian Museum of Nature, Ottawa, 576 pp.

Hegerl, G.C., Crowley, T.J., Hyde, W.T., Frame, D.J., 2006. Climate sensitivity constrained by temperature reconstructions over the past seven centuries. Nature 440, 1029–1032.

Helama, S., Timonen, M., Lindholm, M., Meriläinen, J., Eronen, M., 2005. Extracting long-period climate fluctuations from tree-ring chronologies over timescales of centuries to millennia. International Journal of Climatology 25, 1767–1770.

Hiebl, J., 2006. The early instrumental climate period in Europe: evidences from the Alpine region and Southern Scandinavia. Masters Thesis. University of Vienna, 105 pp.

Holzhauser, H., Magny, M., Zumbühl, H.J., 2005. Glacier and lake-level variations in west-central Europe over the last 3500 years. The Holocene 15, 789–801.

Hormes, A., Müller, B.U., Schlüchter, C., 2001. The Alps with little ice: evidence for eight Holocene phases of reduced glacier extent in the Central Swiss Alps. The Holocene 11, 255–265.

Jacoby, G.C., D'Arrigo, R., 1989. Reconstructed northern hemisphere annual temperature since 1671 based on high-latitude tree-ring data from North America. Climatic Change 14, 39–59.

Jones, P.D., Moberg, A., 2003. Hemispheric and large-scale surface air temperature variations: an extensive revision and an update to 2001. Journal of Climate 16, 206–223.

Jones, P.D., Lister, D., 2004. The development of monthly temperature series for Scotland and Northern Ireland. International Journal of Climatology 24, 569–590.

Jones, P.D., Osborn, T.J., Briffa, K.R., 1997. Estimating sampling errors in large-scale temperature averages. Journal of Climate 10, 2548–2568.

Jones, P.D., Briffa, K.R., Barnett, T.P., Tett, S.F.B., 1998. High-resolution palaeoclimatic records for the past millennium: interpretation, integration and comparison with general circulation model control-run temperatures. The Holocene 8, 455–471.

Jones, P.D., New, M., Parker, D.E., Martin, S., Rigor, I.G., 1999. Surface air temperature and its changes over the past 150 years. Reviews of Geophysics 37, 173–199.

Jones, P.D., Osborn, T.J., Briffa, K.R., 2003. Changes in the Northern Hemisphere annual cycle: implications for paleoclimatology? Journal of Geophysical Research 108, 4588, doi:10.1029/2003JD003695.

Kagawa, A., Sugimoto, A., Maximov, T., 2006. Seasonal course of translocation, storage and remobilization of 13C pulse-labeled photoassimilate in naturally growing Larix gmelinii saplings. New Phytologist 171, 793–804.

Karl, T.R., Diaz, H.F., Kukla, G., 1988. Urbanization: its detection and effect in the United States climate record. Journal of Climate 1, 1099–1123.

Kent, E., Woodruff, S., Berry, D.I., 2007. Metadata from WMO publication no. 47 and an assessment of voluntary observing ship observation heights in ICOADS. Journal of Atmospheric and Oceanic Technology 24, 214–234.

Körner, C., 1998. A re-assessment of high elevation treeline positions and their explaination. Oecologia 115, 445–459.

Klingbjer, P., Moberg, A., 2003. A composite monthly temperature record from Tornedalen in northern Sweden, 1802–2002. International Journal of Climatology 23, 1465–1494.

Kozlowski, T.T., Pallardy, S.G., 1997. Physiology of Wood Plants. Academic Press, San Diego, California, 411 pp.

Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M., Wanner, H., 2004. European seasonal and annual temperature variability, trends and extremes since 1500. Science 303, 1499–1503.

Mann, M.E., Bradley, R.S., Hughes, M.K., 1999. Northern Hemisphere temperatures during the past millennium—inferences, uncertainties, and limitations. Geophysical Research Letters 26, 759–762.

Maugeri, M., Buffoni, L., Chlistovsky, F., 2002a. Daily Milan temperature and pressure series (1763–1998): history of the observations and data and metadata recovery. Climatic Change 53, 101–117.

Maugeri, M., Buffoni, L., Delmonte, B., Fassina, A., 2002b. Daily Milan temperature and pressure series (1763–1998): completing and homogenising the data. Climatic Change 53, 119–149.

Melvin, T.M., 2004. Historical growth rates and changing climate sensitivity of boreal conifers. Ph.D. Thesis, University East Anglia, Norwich, pp. 271.

Moberg, A., Alexandersson, H., Bergström, H., Jones, P.D., 2003. Were southern Swedish summer temperatures before 1860 as warm as measured? International Journal of Climatology 23, 1495–1521.

Moberg, A., Sonechkin, D.M., Holmgren, K., Datsenko, N.M., Karlén, W., 2005. Highly variable Northern Hemisphere temperatures reconstructed from low- and high-resolution proxy data. Nature 433, 613–617.

Mohammad, R., Moberg, A., 2007. Pseudoproxies with pre-determined spectra—Part I: Method for Generating Individual Series. Climate Dynamics (in review).

Nicholls, N., Tapp, R., Burrows, K., Richards, D., 1996. Historical thermometer exposures in Australia. International Journal of Climatology 16, 705–710.

Nørdli, P.Ø., Alexandersson, H., Frich, P., Førland, E.J., Heino, R., Jonsson, T., Tuomenvirta, H., Tveito, O.E., 1997. The effect of radiation screens on Nordic time series of mean temperature. International Journal of Climatology 17, 1667–1681.

Oerlemans, J., 2005. Extracting a climate signal from 169 glacier records. Science 308, 675–677.

Osborn, T.J., Briffa, K.R., 2000. Revisiting timescale-dependent reconstruction of climate from tree-ring chronologies. Dendrochronologia 18, 9–25.

Osborn, T.J., Briffa, K.R., 2004. The real color of climate change? Science 306, 621–622.

Oppenheimer, C., 2003. Climatic, environmental and human consequences of the largest known historic eruption: Tambora volcano (Indonesia) 1815. Progress in Physical Geography 27, 230–259.

Parker, D.E., 1994. Effects of changing exposure of thermometers at land stations. International Journal of Climatology 14, 1–31.

Parker, D., Horton, B., 2005. Uncertainties in central England temperature 1878–2003 and some improvements to the maximum and minimum series. International Journal of Climatology 25, 173–1188.

Peterson, T.C., Vose, R.S., 1997. An overview of the Global Historical Climatology Network temperature database. Bulletin of the American Meteorological Society 78, 2837–2849.

Peterson, T.C., Easterling, D.R., Karl, T.R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Bohm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Foland, E.J., Hanssen-Bauer, I., Alexandersson, H., Jones, P., Parker, D., 1998. Homogeneity adjustments of in situ atmospheric climate data: a review. International Journal of Climatology 18, 1493–1517.

Pfister, C., 1975. Agrarkonjunktur und Witterungsverlauf im westlichen Schweizer Mittelland 1755–1797. Ein Beitrag zur Umwelt- und Wirtschaftsgeschichte des 18. Jahrhunderts. Lang, Bern. 229 pp.

Quayle, R.G., Easterling, D.R., Karl, T.R., Hughes, P.Y., 1991. Effects of recent thermometer changes in the Cooperative Station Network. Bulletin of the American Meteorological Society 72, 1718–1723.

Rayner, N.A., Brohan, P., Parker, D.E., Folland, C.K., Kennedy, J.J., Vanicek, M., Ansell, T., Tett, S.F.B., 2006. Improved analyses of changes and uncertainties in marine temperature measured in situ since the mid-nineteenth century: the HadSST2 data set. Journal of Climate 19, 446–469.

Robock, A., 1994. Review of year without a summer? World Climate in 1816, Climatic Change 26, 105–108.

Robock, A., 2000. Volcanic eruptions and climate. Reviews of Geophysics 38, 191–219.

Rossi, S., Deslauriers, A., Anfodillo, T., Morin, H., Saracino, A., Motta, R., Borghetti, M., 2006. Conifers in cold environments synchronize maximum growth rate of tree-ring formation with day length. New Phytologist 170, 301–310.

Rossi, S., Deslauriers, A., Anfodillo, T., Carraro, V., 2007. Evidence of threshold temperatures for xylogenesis in conifers at high altitudes. Oecologia 152, 1–12.

Rutherford, S., Mann, M.E., Osborn, T.J., Bradley, R.S., Briffa, K.R., Hughes, M.K., Jones, P.D., 2005. Proxy based northern hemisphere surface temperature reconstructions: sensitivity to methodology, predictor network, target season, and target domain. Journal of Climate 18, 2308–2329.

Schär, C., Vidale, P.L., Lüthi, D., Frei, C., Häberli, C., Liniger, M.A., Appenzeller, C., 2004. The role of increasing temperature variability in European summer heat waves. Nature 427, 332–336.

Sigurdsson, H., Carey, S., 1989. Plinian and co-ignimbrite tephra fall from the 1815 eruption of Tambora volcano. Bulletin of Volcanology 51, 243–270.

Tranquillini, W., 1964. The physiology of plants at high altitudes. Annual Reviews of Plant Physiology 15, 345–362.

Vincent, C., Le Meur, E., Six, D., Funk, M., 2005. Solving the paradox of the end of the Little Ice Age in the Alps. Geophysical Research Letters 32, L09706.

von Storch, H., Zorita, E., Jones, J.M., Dimitriev, Y., Gonzalez-Rouco, F., Tett, S.F.B., 2004. Reconstructing past climate from noisy data. Science 306, 679–682.

Wilson, R.J.S., Luckman, B.H., 2003. Dendroclimatic reconstruction of maximum summer temperatures from upper tree-line sites in Interior British Columbia. The Holocene 13, 853–863.

Wilson, R.J.S., Luckman, B.H., Esper, J., 2005. A 500-year dendroclimatic reconstruction of spring/summer precipitation from the Lower Bavarian forest region, Germany. International Journal of Climatology 25, 611–630.

Wilson, R.J.S., D'Arrigo, R., Buckley, B., Büntgen, U., Esper, J., Frank, D., Luckman, B., Payette, S., Vose, R., Youngblut, D., 2007. A matter of divergence: Tracking recent warming at hemispheric scales using tree-ring data. Journal of Geophysical Research 112, D17103, doi:10.1029/2006JD008318.