Northern European summer temperature variations over the Common Era from integrated tree-ring density records
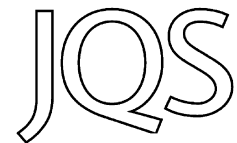
JQS

JAN ESPER,[1]* ELISABETH DÜTHORN,[1] PAUL J. KRUSIC,[2,3] MAURI TIMONEN[4] and ULF BÜNTGEN[5,6,7]
[1]Department of Geography, Johannes Gutenberg University, Mainz 55099, Germany
[2]Department of Physical Geography and Quaternary Geology, Stockholm University, Stockholm, Sweden
[3]Navarino Environmental Observatory, Messinia, Greece
[4]Finnish Forest Research Institute, Rovaniemi Unit, Rovaniemi, Finland
[5]Swiss Federal Research Institute WSL, Birmensdorf, Switzerland
[6]Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland
[7]Global Change Research Centre AS CR, Brno, Czech Republic

ABSTRACT: Tree-ring chronologies of maximum latewood density are most suitable to reconstruct annually resolved summer temperature variations of the late Holocene. The two longest such chronologies have been developed in northern Europe stretching back to the 2nd century BC, and the 5th century AD. We show where similarities between the two chronologies exist, and combine portions of both into a new summer temperature reconstruction for the Common Era. To minimize the transfer of potential biases, we assess the contribution of the candidate reconstructions' measurements, and exclude data (i) from exceptionally young and old trees, and (ii) produced by different laboratory technologies. Our new composite reconstruction reveals warmer conditions during Roman, Medieval and recent times, separated by prolonged cooling during the Migration period and Little Ice Age. Twentieth century warmth, as indicated in one of the existing density records, is reduced in the new reconstruction, also affecting the overall, millennial-scale, cooling trend over the late Holocene (−0.30 °C per 1000 years). Due to the reduced biological memory, typical for tree-ring density measurements, the new reconstruction is most suitable for evaluating the rate and speed of abrupt summer cooling following large volcanic eruptions. Copyright © 2014 John Wiley & Sons, Ltd.

KEYWORDS: dendrochronology; Finland; late Holocene; Sweden; temperature reconstruction.

Introduction

Conifer tree-rings are composed of large cells with thin walls formed in the first weeks of the vegetation period, followed by small cells with thick walls formed throughout high and late summer (Moser *et al.*, 2010). This change between earlywood and latewood structure, along with growth dormancy during winter, is the foundation for distinguishing one ring from another and for assigning each ring to a particular calendar year (Fritts, 1976). The development of high-precision radiography in the 1970s permitted use of X-ray-sensitive films to detect changes in the physical properties of wood cell structure and quantify them in the form of high-resolution density profiles in units of g cm^{-3} (Schweingruber *et al.*, 1978). In one of the first applications of this new technique to palaeoclimatic research, a wood density chronology extending back to AD 441 was developed in the 1980s (Schweingruber *et al.*, 1988), integrating samples from living trees and relict wood of *Pinus sylvestris* from the lake Torneträsk region in northern Sweden (Bartholin and Karlén, 1983). The maximum latewood density (MXD), a then novel parameter representing the microscopic area of thickest walls and smallest lumina towards the end of a ring, was found to correlate best with regional instrumental climate data (up to $\sim r = 0.8$) and subsequently used for reconstructing summer temperatures back to the 5th century AD (Schweingruber *et al.*, 1988; Briffa *et al.*, 1990, 1992).

This widely recognized reconstruction from Torneträsk (hereafter: Torn) has since been updated repeatedly by the addition of new MXD measurements from living trees, thereby advancing the most recent year of the chronology into the 21st century (Grudd, 2008). The latest of these publications (Melvin *et al.*, 2013) deals with removing trend biases that arose from combining the original, analog measurements developed in the 1980s (Schweingruber *et al.*, 1988) with digital update measurements developed in the 2000s using innovative laboratory technologies (Grudd, 2008). Doing so involved changing the mean and variance of the digital measurement series to match the characteristics of the original data, as well as treating the original and updated datasets independently in the process of detrending and chronology development (Melvin *et al.*, 2013). The recombined relict and living tree chronologies, to form a temperature reconstruction extending back to AD 441, indicated stronger 20th century warmth than previously reported, but also resulted in a better coherence between the temperature histories derived from MXD and the generally more noisier tree-ring width (TRW) data (Melvin *et al.*, 2013).

In addition to the widely cited Torn reconstructions (overviews in Büntgen *et al.*, 2011, 2012), a new pine MXD record, derived from sub-fossil wood preserved in shallow lakes from Finnish Lapland (Eronen *et al.*, 2002), has been developed providing summer temperature variations back to 138 BC (Esper *et al.*, 2012a; hereafter: N-Scan). The sub-fossil component of this reconstruction originates from an extended region north of 67°N and was linked to present by samples from living pines growing on lakeshores (Düthorn *et al.*, 2013) in northern Finland and Sweden, including two sites near Torneträsk. N-Scan includes more trees than any other MXD chronology, and has been used to describe a millennial-scale cooling trend related to long-term, orbital

*Correspondence: Jan Esper, as above.
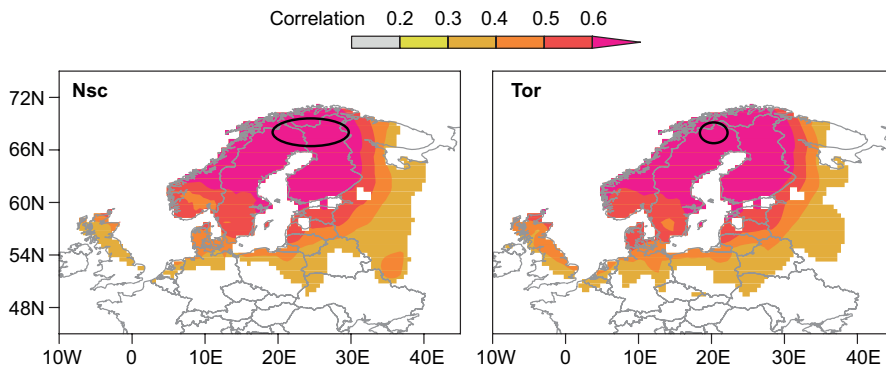E-mail: esper@uni-mainz.de

Figure 1. Correlation fields of the N-Scan and Torn MXD chronologies (black circles) with gridded (0.5° resolution) summer temperatures calculated over the 1901–2006 period. This figure is available in colour online at wileyonlinelibrary.com.

insolation changes (Esper *et al.*, 2012a). Subsequent work (Esper *et al.*, 2012b) showed that there still exist substantial differences among the various tree-ring-based climate reconstructions from Fennoscandia, indicating our poor understanding of the pre-instrumental temperature amplitude even in geographically limited regions.

In an effort to reconcile some of the density data from this region, we here assess the world's two longest MXD-based climate reconstructions, Torn (Melvin *et al.*, 2013) and N-Scan (Esper *et al.*, 2012a), and combine portions of both to produce a new summer temperature history for northern Europe from 17 BC to the present. The rationale for combining these data is supported by the chronologies' coherent spatial temperature fields (Fig. 1), and the desire to mitigate complications in the process of selecting proxy records for the development of large-scale temperature reconstructions (e.g. Esper *et al.*, 2002; D'Arrigo *et al.*, 2006; Mann *et al.*, 2008; Ljungqvist, 2010; Ljungqvist *et al.*, 2012). In this paper we first examine the similarities and differences between the Torn and N-Scan reconstructions, and revisit the challenge of integrating digital and analog MXD data. We describe a manner of selecting coherent MXD data, and use this subsample to produce a new temperature reconstruction for northern Scandinavia spanning the Common Era. We conclude by briefly addressing the long-term cooling trend inherent to our new reconstruction, and demonstrate the value of this record for assessing short-term cooling extremes following large volcanic eruptions.

Data and methods

Torn and N-Scan MXD data

While the Torn and N-Scan reconstructions are both derived from MXD data and share the same spatial temperature pattern, the records differ in (i) the source of their dead wood samples (lake/sub-fossil and fallen/relict), (ii) the geographical area from which samples were collected, (iii) the number of MXD measurement series and (iv) the X-ray technology used to measure MXD (details in Esper *et al.*, 2012a; Melvin *et al.*, 2013). Whereas the Torn samples all originate from locations surrounding Lake Torneträsk, in the transition zone between the boreal forest and low Arctic tundra, N-Scan integrates samples from 17 lakes and lakeshore sites spread over Finnish Lapland and the Torneträsk region. N-Scan contains substantially more dead wood MXD series (372 sub-fossil samples) than the Torn chronology (47 fallen/relict samples) making it the most well-replicated and longest MXD chronology in the world.

The 372 N-Scan sub-fossil samples were combined with 215 samples from living trees collected on the shores of four lakes located in northern Finland and Sweden. The 47 Torn relict samples were originally combined with 18 samples from living trees collected in 1980 (Schweingruber *et al.*, 1988), then updated with 35 samples from living trees in 2004

(Grudd, 2008), and again with an additional 30 living tree samples in 2010 (Melvin *et al.*, 2013). Whereas all 587 N-Scan MXD series were measured at the WSL in Birmensdorf (Switzerland) using a Walesch X-ray densitometer (Schweingruber *et al.*, 1978), the Torn chronology was produced using different technologies in different laboratories: the original 65 Torn MXD series were also measured with a Walesch densitometer in Birmensdorf, but subsequent updates were measured in Stockholm and Kiruna, Sweden, using an ITRAX digital radiographic camera. The 587 MXD series used in Esper *et al.* (2012a) are hereafter named E12, the 65 series used in Schweingruber *et al.* (1988) are named S88 and the 65 updating series used in Melvin *et al.* (2013) are named M13.

Reconstruction assessment

Coherence between the Torn and N-Scan reconstructions was assessed using Pearson correlation coefficients calculated over 100-year sliding windows. The procedure was applied to the original reconstructions as well as to first-differenced versions of the two records to evaluate common variation at both high and low frequencies. The interseries correlation (Rbar), a metric based on the correlation among all MXD measurement series in a chronology (Cook and Kairiukstis, 1990), was used to estimate the internal coherence of the reconstructions. Rbar was also calculated over 100-year windows shifted along the chronologies. Minimum least square regressions were fitted to the reconstructions to assess the long-term trends inherent to the time series over their common AD 517–2006 period, as well as the AD 517–1900, and the 138 BC to AD 1900 (N-Scan only) periods.

Detrending and chronology development

Raw MXD measurement series are characterized by an increase from ~0.5 to 0.7 g cm^{-3} over the first 30 years of a tree's lifespan, followed by a gradual decrease from ~0.7 to 0.6 g cm^{-3} through plant maturation (see Fig. 4A). This age trend has to be removed before averaging MXD series in a mean chronology, otherwise the combination of juvenile and adult MXD values, rather than variations in temperature, determines the changing chronology levels back in time (Bräker, 1981). Age trend is here removed using Regional Curve Standardization (RCS; Briffa *et al.*, 1992; Esper *et al.*, 2003), a detrending method capable of retaining the full variance frequency spectrum in the resulting chronologies (Bunde *et al.*, 2013; Franke *et al.*, 2013), from inter-annual to millennial scales (Esper *et al.*, 2012a). Additional methods that have been suggested in conjunction with RCS (Melvin and Briffa, 2008) revealed no detectable difference (Supporting information, Fig. S1).

RCS involves (i) an alignment of all MXD measurement series by biological age, (ii) calculating a mean time series of the age-aligned data, (iii) smoothing this mean curve, (iv) age-trend removal by calculating ratios between the raw MXD

data and the (smoothed) regional curve, and (v) re-dating the detrended, now dimensionless, values back to calendar years (details in Esper *et al.*, 2003). We here smoothed the mean of the age-aligned data using a 10-year spline filter (Cook and Peters, 1981), and calculated chronologies of the detrended MXD data using the arithmetic mean. When aligning the MXD measurement series by biological age, the pith offset (the number of missing innermost rings of a sample) was considered to minimize distortion of the regional curve.

RCS was applied to a combined dataset integrating S88 and M13, as well as S88 and E12 samples. In the first run, the effects of the mean and variance adjustments applied to M13 (Melvin *et al.*, 2013) on post-1700 chronology levels were assessed. Then, using only the S88 and E12 samples, the final reconstruction was produced. Combining the S88 and E12 MXD measurements required an adjustment of the S88 data to compensate for a post-1988 change of the calibration wedge used to transfer X-ray film greyscales into values of g cm^{-3}, as well as the brand of analog film used (changed from Agfa to Kodak). These procedural changes required an adjustment of +0.053 g cm^{-3} added to all S88 MXD measurement series (Fig. S2). The adjustment appeared justified, as the S88 dataset integrates several successive tree generations covering the past 1.5 millennia, allowing a robust estimation of the average residual with the E12 data (covering 2 millennia). Additional treatment of the variance of the MXD series appeared unnecessary, as the measurement's standard deviations have not changed over the past 30 years (Fig. S3).

For the final summer temperature reconstruction we only used tree rings of a certain biological age ranging from 31 to 306 years, i.e. removed the rings ≤30 years (here termed as the 'young rings') and >306 (here termed as the 'old rings') from the combined, and adjusted S88 + E12 dataset. Sub-sample chronologies integrating MXD data constrained to certain age-bands (e.g. 1–30, 31–306 and >306 years) were calculated using the detrending program Spotty (Esper *et al.*, 2009; Fig. S4). Removal of the old rings appeared useful, as these rings are relatively rare and clustered in the second millennium AD. Removal of the young data was also useful, as these rings contain a steep MXD increase from ~0.5 to 0.7 g cm^{-3} and are clustered in the 20th century, thereby complicating proper proxy calibration and uncertainty estimation back in time (see below). All chronologies were truncated at a minimum replication of 5 MXD measurement series.

Calibration and error estimation

The MXD climate signal was estimated by calibrating the combined S88 + E12 chronology against regional June–July–August (JJA) mean temperatures recorded at the Haparanda, Karasjok and Sodankyla meteorological stations over the period 1876–2006. A split 1876–1940 calibration and 1941–2006 verification approach was used for assessing the temporal robustness of the signal (Schneider *et al.*, 2014), and the chronology transferred into JJA temperatures by scaling the record to the mean and variance of the instrumental climate data (Esper *et al.*, 2005). The resulting summer temperature reconstruction, integrating the adjusted and age-constrained MXD data from S88 and E12, is termed N-Eur.

Uncertainty estimates of N-Eur were calculated using a Monte Carlo approach that considered the declining replication of the MXD chronology back in time, together with the correlation of the less-replicated chronologies against regional JJA temperatures. The procedure involved producing pseudo-chronologies by randomly sampling $n = 5$, $n = 10$, … $n = 75$, $n = 79$ MXD series over the 1876–2006 period, and

iteratively correlating these 'restricted' chronologies with the instrumental climate data. The $n = 79$ chronology represents the (maximum) replication of the combined S88 + E12 dataset during 1876–2006 after removing all trees younger than 100 years. The process was repeated 1000 times for each replication class (5, 10, … 79), and the mean correlation of each class used to calculate calibration model standard errors. The changing standard errors, from all Monte Carlo simulations, were added to N-Eur to provide an estimate of the increasing uncertainty associated with the decreasing MXD sample replication back in time. N-Eur replication falls below five MXD series before 17 BC.

Results and discussion

Torn and N-Scan comparison

Comparison of the Torn and N-Scan temperature reconstructions reveals a significant fraction of shared variance over the past 1500 years ($r_{441–2006} = 0.65$), although the correlation declines from $r = 0.91$ in the 20th century to $r = 0.60$ in the 6th century, and down to $r = 0.43$ in the first 100 years of overlap (AD 441–540; Fig. 2A, B). Similar trends are found when comparing the first-differenced reconstructions, indicating this long-term correlation decay is unrelated to a decoupling of lower frequency variances. The changing coherency is probably a function of the changing replication through time. Both reconstructions have their greatest sample replication in the 20th century (N-Scan = 116, Torn = 61), but this quantity declines rapidly back over the past 1500 years (Fig. 2D). The N-Scan reconstruction contains 49 MXD series during the 10th century, and only 12 series during the first 100 years of overlap with Torn. Torn replication is already down to 13 series in the 10th century and falls below five series before AD 517. The influence of changing samples depths was confirmed by the stepwise reduction in replication of the N-Scan and Torn chronologies during the calibration period (1876–2006), re-developing chronologies using the reduced sub-samples (1000 times each), and iteratively calculating correlations between the sub-sample chronologies. Results show decreasing correlations as a function of sample replication. For example, sub-sample chronologies integrating 48 MXD samples correlate at 0.825, 30 samples at 0.815, 10 samples at 0.760 and five samples at 0.696.

In addition to changes in sample size, a gradual shift in the intra-reconstruction coherence of both chronologies (Fig. 2C) may also contribute to the reduced correlation between the reconstructions back in time. Yet the running 100-year RBar values exhibit minor trend, decreasing from 0.41 and 0.43 over the recent ~800 years, to 0.38 and 0.39 over the first ~800 years, in N-Scan and Torn, respectively. Similarly, the changing geographical area represented by the chronologies might add to the temporally changing correlations, as the N-Scan reconstruction integrates living trees from both Finland and Sweden, plus sub-fossil trees from Finland only. However, since the different spatial pattern of the two reconstructions' constituent time series are not reflected in the intra-reconstruction's Rbar, it appears most likely that the shrinking replication is the main reason for the reduced agreement between N-Scan and Torn back in time.

Despite the significant correlation between N-Scan and Torn, the reconstructions display demonstrably different millennial-scale temperature trends over the late Holocene (Fig. 2E). While N-Scan temperature estimates decline by about −0.28 to −0.39 °C per millennium, the Torn reconstruction has a positive trend ranging from +0.09 to +0.23 °C per millennium. Esper *et al.* (2012a) suggested the elevated
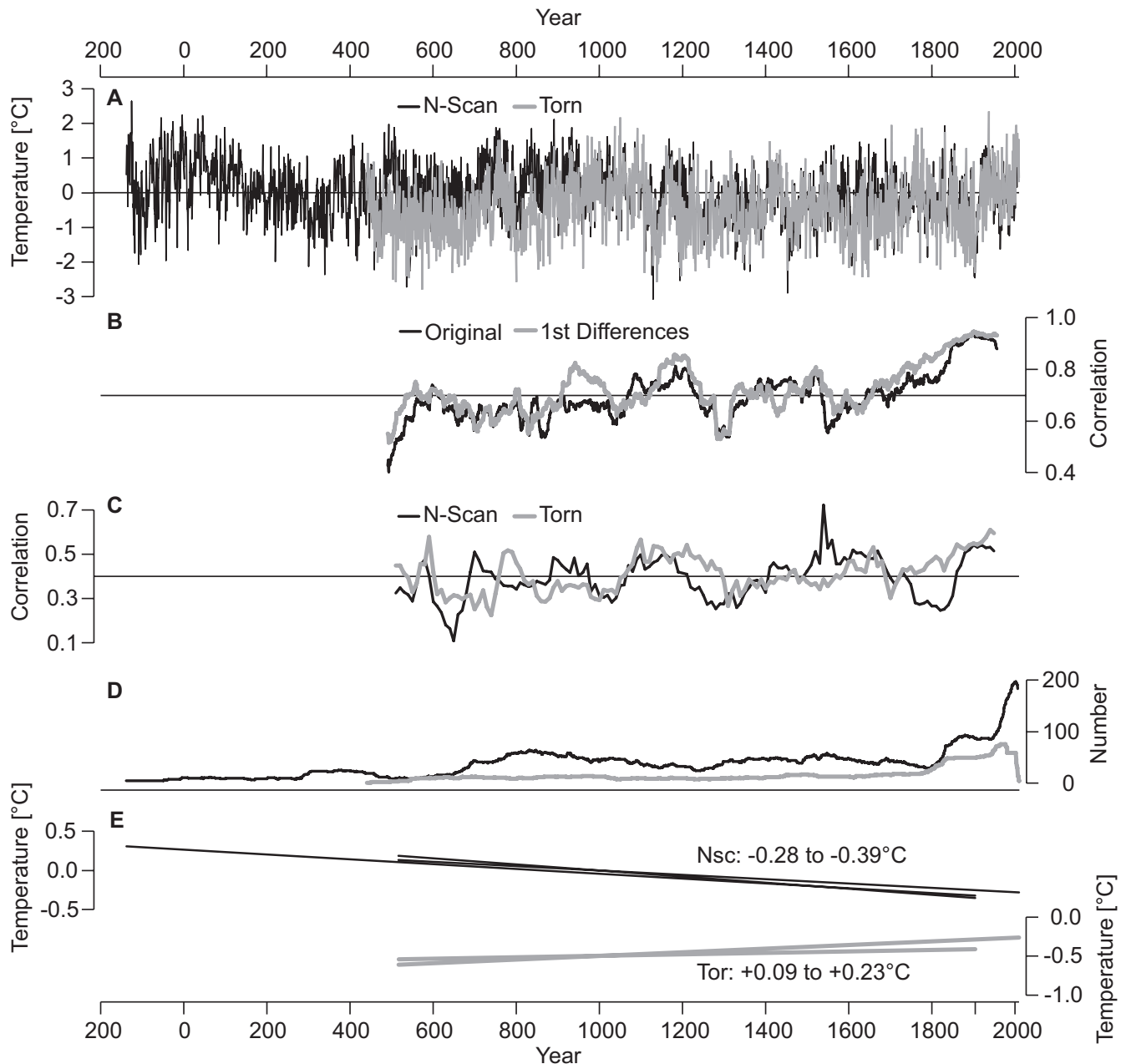
Figure 2. Comparison of the N-Scan and Torn temperature reconstructions. (A) The N-Scan and Torn reconstructions as published in Esper *et al.* (2012a) and Melvin *et al.* (2013) expressed as anomalies with respect to a 1901–2000 reference period. (B) 100-year running correlations between the original reconstructions and between the first-differenced reconstructions. (C) 100-year running interseries correlations of the N-Scan and Torn chronologies. (D) Numbers of tree samples integrated in the N-Scan and Torn chronologies. (E) Linear regressions fit to the reconstructions over the common AD 517–2006, and the AD 517–1900 and 138 BC to AD 1900 (N-Scan only) periods.

sample replication of the N-Scan chronology is a key reason for retaining an orbitally forced cooling trend in a dendro-chronological time series. The opposite, and overall positive trend in Torn, particularly the maximum + 0.23 °C per 1000 years warming rate recorded over the AD 517–2006 period, might additionally be influenced by the adjustments applied to the M13 ITRAX update series (see below).

MXD data selection

In the attempt to correct for the systematic differences between data developed on the ITRAX and Walesch instruments, Melvin *et al.* (2013) relied on an adjustment that ultimately reduced the variance and raised the mean of the ITRAX MXD series, thus producing the M13 dataset (Fig. 3). While lifting the M13 update above the original S88 data during the post-1800 period of overlap might well be justified with regard to biological growth rates (the S88 data are a collection of much older trees than the M13 data), consideration of the adjusted M13 instead of the S88 living trees in an RCS run results in a substantial increase of recent chronology values (Fig. 3D). Whether such effects are mitigated or even amplified by the use of split RCS standardization (Melvin *et al.*, 2013) is beyond the scope of this paper, although it appears preferable to use only the E12 living trees for extending N-Eur into the 21st century. Doing so reduces replication in the 19th and 20th centuries, which is, however, still larger than at any other time, and mitigates the millenni-al-scale trend differences advanced by the inclusion of the adjusted M13 data. Our decision not to join the adjusted ITRAX M13 data with the Walesch E12 data does not mean we consider M13 unreliable for reconstruction purposes (Torn and N-Scan correlate at 0.91 during the 20th century), but rather to avoid unnecessary complication when integrat-ing data developed by different techniques.
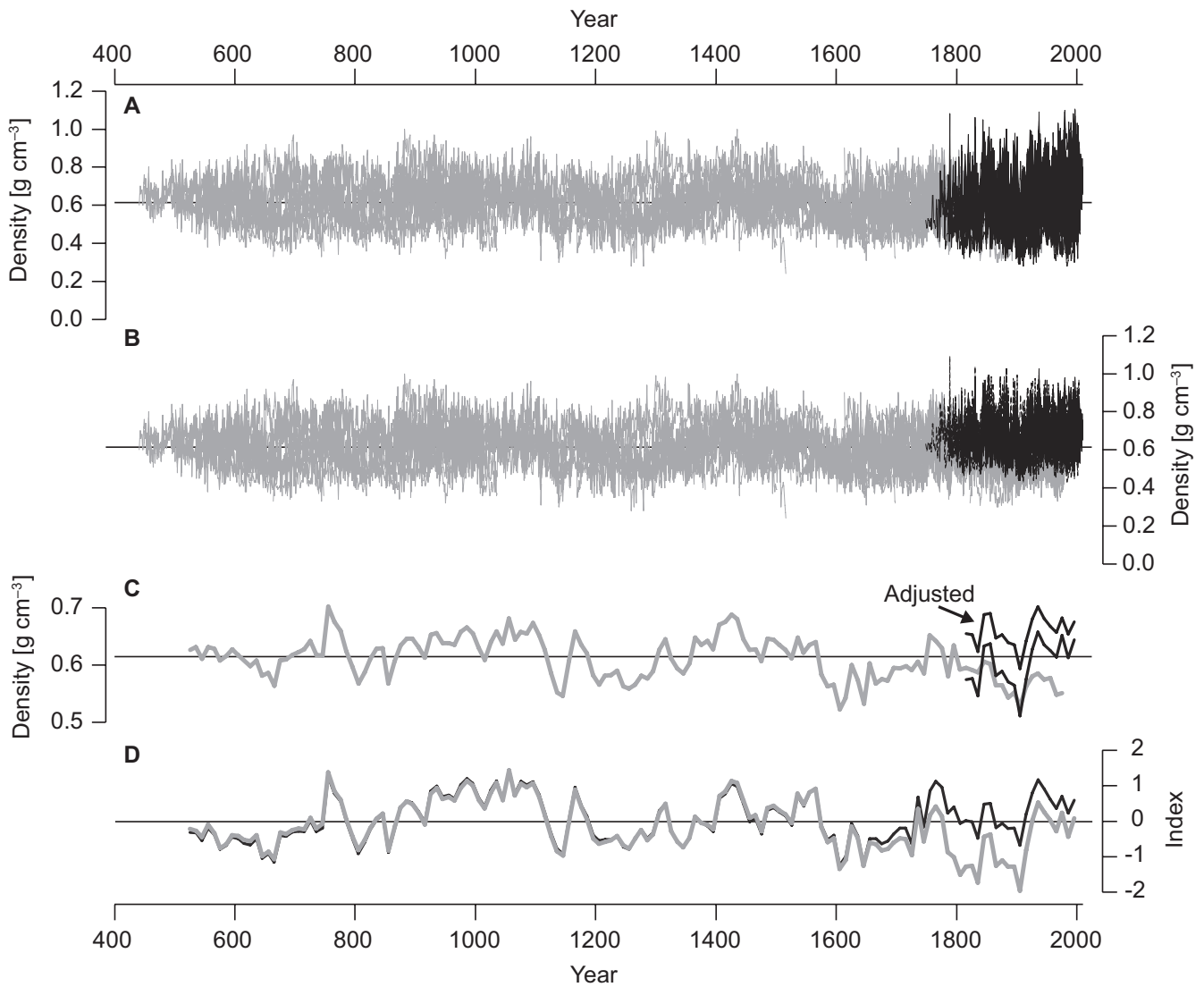
Figure 3. Torneträsk original and update data. (A) Sixty-five original MXD series from living and relict pine samples (S88; grey curves) shown together with the 65 update MXD series from just living trees (M13; black curves). The original data were measured using a Walesch X-ray densitometer; the update data were measured using an ITRAX digital radiographic camera. (B) Same as in (A), but after adjustment of the mean and variance of the ITRAX data as detailed in Melvin et al. (2013). (C) Mean chronologies of the original (grey) and updated (black) MXD data at decadal resolution. Top back curve is the mean of the adjusted update data. (D) RCS detrended chronologies of a combined dataset integrating the 47 original relict series together with the 65 living tree update series (grey), as well as the same 47 relict series together with the 65 update series after adjustment of their means and variances (black). Chronologies were normalized over the AD 517–1700 period.

To further restrain potential trend biases in the N-Eur reconstruction we removed tree rings older than 306 years and younger than 30 years from the combined S88 + E12 dataset (Fig. 4). Tree rings >306 years are relatively rare (3918 of 98 740 rings) and largely restricted to the S88 data, which include some of the oldest trees found in the Torneträsk region (up to 620 years). As these old rings are also concentrated in the second millennium AD, not meaningfully contributing to an assessment of long-term changes between the first and second millennium AD (but rather level off such variance in an RCS approach), the data were excluded from the final reconstruction. Truncation at a biological age of 306 years appeared reasonable as replication falls below 50 series beyond this age limit, and the regional curve becomes heteroscedastic (Fig. 4A, B). Reasons for not considering tree rings ≤30 years result from the growing mismatch between the S88 and E12 regional curves towards younger ages, even after adjustment of the S88 data (Fig. S2A), as well as the documented sharp MXD increase in juvenile rings, from ∼0.5 to 0.7 g cm^{-3}, in which climate signal age effects (Carrer and Urbinati, 2004; Esper

et al., 2008) are probably increased. The correlation between the central 31–306 age-band, and the young 1–30 age-band chronologies declines to $r = 0.63$ ($n = 979$ years; Fig. 4C), and removing the young rings again balances replication as these data are culminating in the 20th century calibration period (79 series in 1992; the green curve in Fig. 4D).

N-Eur summer temperature reconstruction

The 31–306 year age-band chronology correlates at 0.76 with 131 years of regional instrumental JJA temperatures (Fig. 5A). Split calibration ($r_{1876-1940} = 0.78$) and verification ($r_{1941-2006} = 0.75$) indicates the climate signal is temporally robust with no sign of divergence (D'Arrigo et al., 2008; Esper et al., 2010). The correlation is slightly smaller than that obtained using all data of the Torn and N-Scan reconstructions ($r_{1876-2006} = 0.78$), which is likely a consequence of N-Eur's lower replication (93 series in the 20th century) compared with a non-constrained chronology (180 series in the 20th century). Nonetheless, the N-Eur correlation reported here is not representative of the pre-instrumental period, as
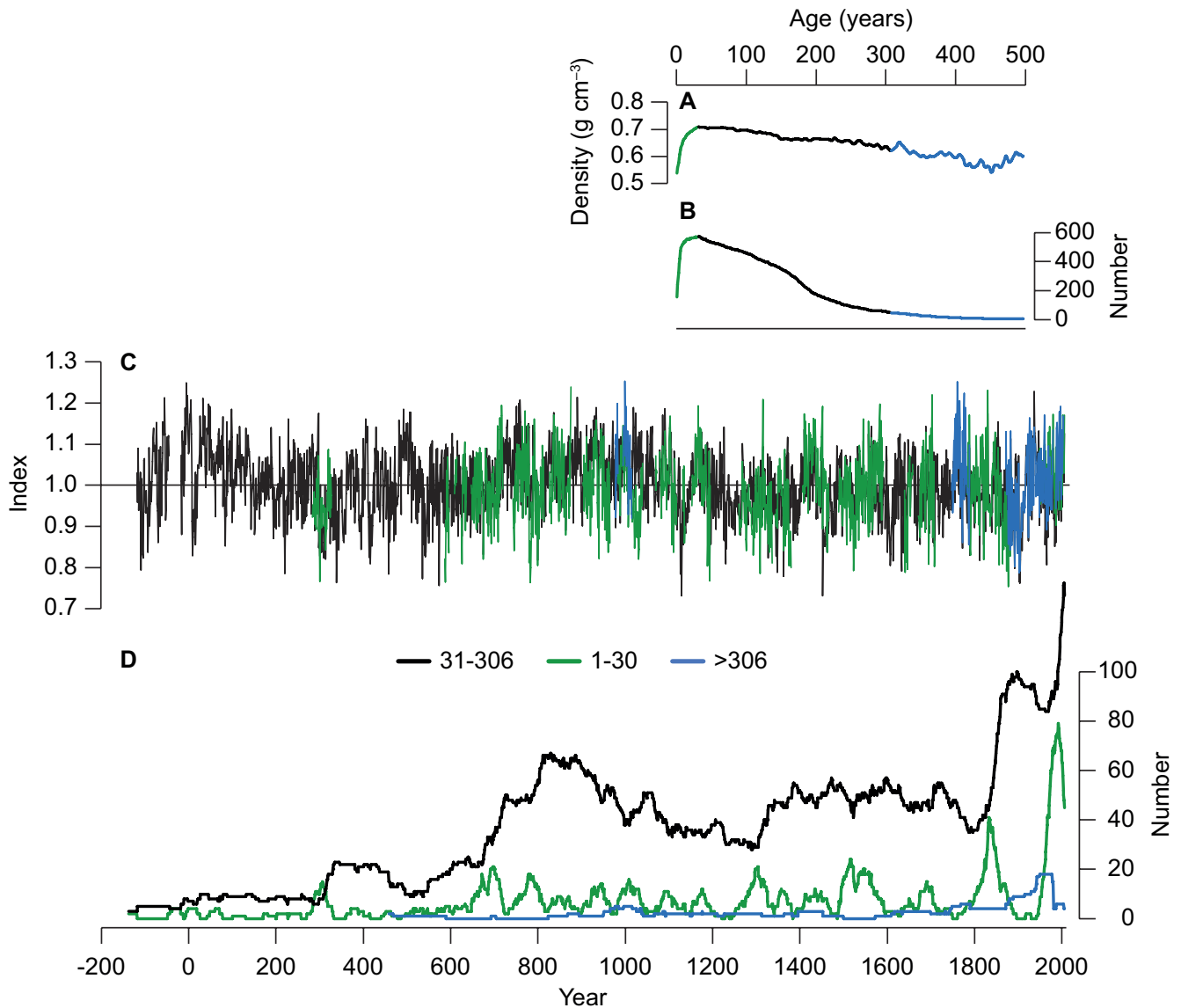
Figure 4. N-Eur age-band chronologies. (A) Regional curves of the N-Eur MXD data of the age-bands 1–30 (green), 31–306 (black) and >306 years (blue). Bottom panel (B) shows the replication curves across all age-bands. (C) RCS detrended age-band chronologies, and (D) their replication over the past 2000 years. After truncation at $n < 5$ MXD series, the chronologies cover 2097 years (age band 31–306), 979 years (age band 1–30) and 197 years (age band >306). Correlations with the 31–306 age-band chronology are 0.83 for the old chronology and 0.63 for the young chronology. This figure is available in colour online at wileyonlinelibrary.com.

the reconstruction's replication quickly decreases to 52 series in the 10th century and eight series in the 1st century AD. This decline in sample size is addressed by performing calibration experiments using differently replicated chronologies. These Monte Carlo tests produced correlations ranging from 0.69 (± 0.04) for chronologies integrating 79 MXD series to 0.63 (± 0.14) for chronologies integrating five MXD series, and were used to add variable calibration model standard errors to the N-Eur reconstruction back in time (Fig. 5B, C). The standard error increases from 0.76 °C in the 20th century to 0.92 °C in the first 100 years of the reconstruction.

N-Eur indicates a long-term cooling trend of −0.30 °C per 1000 years over the 17 BC to AD 2006 period, in line with results detailed for the N-Scan reconstruction (Esper et al., 2012a), although the latter record has been transferred into temperatures using linear regression instead of scaling (Esper et al., 2005). This overall negative trend is contrary to expectations based on the assumption that relict and sub-fossil wood, lying on the ground and in lakes for centuries, would lose density due to natural decaying processes. The assumption stems from the observation that older discs are

lighter than younger ones. Superimposed on the long-term cooling trend are centennial-scale warm episodes, including the Roman Warm Period until ∼AD 150, the Medieval Warm Period from ∼AD 700 to 1200, and the 20th century warm period (Kullman, 2013). These warm phases are separated by cooler conditions during the Migration period and the Little Ice Age in the first and second millennium AD, respectively. Summer temperature variations at this temporal scale, i.e. after smoothing with a 100-year spline filter, reveal fluctuations that range by ∼1.5 °C between warm and cold episodes.

Besides these lower frequency trends, N-Eur reveals extremely warm and cold summers deviating up to +3.2 °C (mean of 2 BC, 4 BC, AD 1937) and −3.4 °C (mean of AD 574, 1130, 1453) from the 1961–1990 reference period. Coincidently, one of the three coldest years (AD 1453, −3.47 °C) is synchronous with the large eruption of the Kuwae volcano in Vanuatu (Plummer et al., 2012), suggesting some of the earlier extremes might point to previously unknown eruptions. Given the good fit with instrumental climate data, the Common Era coverage and the good replication, exceeding all other MXD chronologies, the JJA
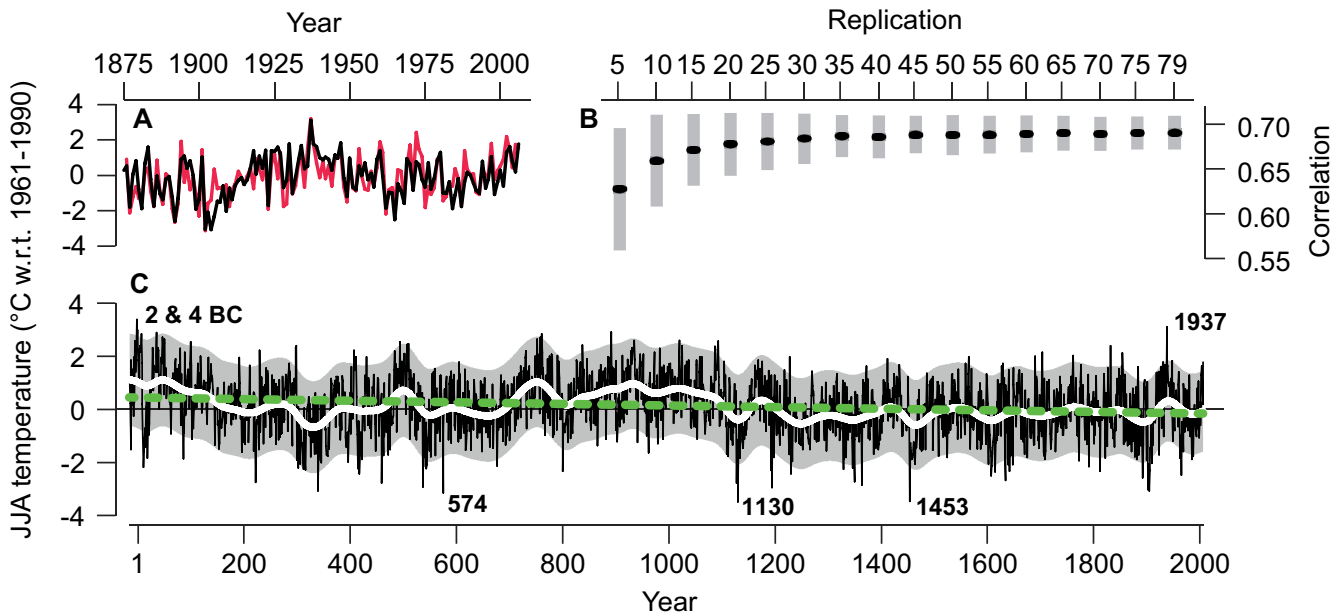
Figure 5. N-Eur temperature reconstruction. (A) Calibration of the N-Eur chronology (black) against regional JJA temperatures (red) over the 1876–2006 period. (B) Mean correlations (black) of variously replicated ($n=5$, $n=10$, … $n=79$) MXD chronologies against regional JJA temperatures. Grey bars indicate the standard deviation of correlations for each replication class 5–79 over 1000 iterations. (C) N-Eur temperature reconstruction at annual resolution (black) and after smoothing using a 100-year filter (white). Green curve is a linear regression fit to the annually resolved data over the 17 BC to AD 2006 period. Grey area indicates the smoothed standard error (2SE) of the calibration model, using variable chronology replicated error estimates (from B). Dates indicate the three coldest and warmest years over the past two millennia. This figure is available in colour online at wileyonlinelibrary.com.

temperature reconstruction presented here might be the most suitable proxy time series to estimate post volcanic cooling (Anchukaitis *et al.*, 2012; D'Arrigo *et al.*, 2013; Esper *et al.*, 2013a,b), and to compare with ice core-derived evidence of volcanic eruptions (Baillie, 2008; Sigl *et al.*, 2013).

Conclusions

The world's two longest MXD chronologies were combined into one single reconstruction of summer temperature variability in northern Europe back to 17 BC. The reconstruction explains 58% of the variance in instrumental JJA temperature fluctuations, although climate signal strength decreases to ~47% in the 10th and ~40% in the 1st century AD. The signal decay is caused by a declining sample replication back in time, a reconstruction property evaluated here by iteratively calibrating differently replicated chronologies against the instrumental temperature data.

The combination of MXD data from two independent reconstructions was challenged by changing laboratory procedures, established over the past 30 years, to measure high-resolution wood density profiles. A recently completed update of MXD data from living trees was excluded from the reconstruction, as the necessary adjustment of the mean and variance, for the fraction of data covering the most recent centuries, appeared problematic. Excluding these data also reduced the sample replication during the 20th century, a period during which the number of MXD measurement series outnumbered pre-instrumental replication. Another data reduction was invoked by excluding MXD measurements from extremely young and old tree rings, as these data do not meaningfully contribute to the reconstruction's low-frequency variance but rather complicate the determination of uncertainty estimates over the past 2000 years.

The final product, a summer temperature reconstruction derived from a combined, adjusted and age-constrained MXD dataset from northern Sweden and Finland, shows a long-term cooling trend of −0.30 °C per 1000 years over the

Common Era in northern Europe. The reconstruction has centennial-scale variations superimposed on this trend, indicating conditions during Medieval and Roman times were probably warmer than in the late 20th century. The record reveals several extreme summer temperature deviations exceeding −3 °C that might prove useful in the detection and dating of volcanic eruptions, and assessment of post-eruption cooling effects.

Supporting Information

Additional supporting information can be found in the online version of this article:

Figure S1. Comparison of RCS and Signal Free detrended chronologies.
Figure S2. Adjustment of the S88 data.
Figure S3. Variance of MXD time series.
Figure S4. RCS detrended age-band chronologies.

Abbreviations. JJA, June–July–August; MXD, maximum latewood density; RCS, Regional Curve Standardization; TRW, tree-ring width

References

Anchukaitis KJ, Cook ER, D'Arrigo RD, *et al.* 2012. Tree rings and volcanic cooling. *Nature Geosciences* **5**: 835–836.

Baillie MGL. 2008. Proposed re-dating of the European ice core chronology by seven years prior to the 7th century AD. *Geophysical Research Letters* **35**: doi: 10.1029/2008GL034755.

Bartholin TS, Karlén W. 1983. Dendrokronologi i Lappland AD 436–1981. *Dendrokronologiska Sällskapets Meddelanden* **5**: 3–16.

Bräker OU. 1981. Der Alterstrend bei Jahrringdichten und Jahrringbreiten von Nadelhölzern und sein Ausgleich. *Mitteilungen der forstlichen Bundesversuchsanstalt Wien* **142**: 75–102.

Briffa KR, Bartholin TS, Eckstein D, *et al.* 1990. A 1,400-year tree-ring record of summer temperatures in Fennoscandia. *Nature* **346**: 434–439.

Briffa KR, Jones PD, Bartholin TS, *et al.* 1992. Fennoscandian summers from AD 500: temperature changes on short and long timescales. *Climate Dynamics* **7**: 111–119.

Büntgen U, Raible CC, Frank D, *et al.* 2011. Causes and consequences of past and projected Scandinavian summer temperatures, 500–2100 AD. *PLOS ONE* **6**: e25133.

Büntgen U, Frank D, Neuenschwander T, *et al.* 2012. Fading temperature sensitivity of Alpine tree growth at its Mediterranean margin and associated effects on large-scale climate reconstructions. *Climatic Change* **114**: 651–666.

Bunde A, Büntgen U, Ludescher J, *et al.* 2013. Is there memory in precipitation? *Nature Climate Change* **3**: 174–175.

Carrer M, Urbinati C. 2004. Age-dependent tree-ring growth responses to climate in *Larix decidua* and *Pinus cembra*. *Ecology* **85**: 730–740.

Cook ER, Peters K. 1981. The smoothing spline: a new approach to standardizing forest interior tree-ring width series for dendroclimatic studies. *Tree-Ring Bulletin* **41**: 45–53.

Cook ER, Kairiukstis LA. 1990. *Methods of Dendrochronology – Applications in the Environmental Science*. Kluwer Publishers: Dordrecht.

D'Arrigo R, Wilson R, Jacoby G, *et al.* 2006. On the long-term context for late twentieth century warming. *Journal of Geophysical Research* **111**: 1–46.

D'Arrigo RD, Wilson R, Liepert B, *et al.* 2008. On the 'divergence problem' in northern forests: a review of the tree-ring evidence and possible causes. *Global and Planetary Change* **60**: 289–305.

D'Arrigo R, Wilson R, Anchukaitis KJ. 2013. Volcanic cooling signal in tree ring temperature records for the past millennium. *Journal of Geophysical Research* **118**: 9000–9010.

Düthorn E, Holzkämper S, Timonen M, *et al.* 2013. Influence of micro-site conditions on tree-ring climate signals and trends in Central and Northern Sweden. *Trees* **27**: 1395–1404.

Eronen M, Zetterberg P, Briffa KR, *et al.* 2002. The supra-long Scots pine tree-ring record for Finnish Lapland: Part 1, Chronology construction and initial inferences. *Holocene* **12**: 673–680.

Esper J, Cook ER, Schweingruber FH. 2002. Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. *Science* **295**: 2250–2253.

Esper J, Cook ER, Krusic PJ, *et al.* 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 DC, Wilson RJS, *et al.* 2005. Effect of scaling and regression on reconstructed temperature amplitude for the past millennium. *Geophysical Research Letters* **32**: doi: 10.1029/2004GL021236.

Esper J, Niederer R, Bebi P, *et al.* 2008. Climate signal age effects – evidence from young and old trees in the Swiss Engadin. *Forest Ecology and Management* **255**: 3783–3789.

Esper J, Krusic PJ, Peters K, *et al.* 2009. Exploration of long-term growth changes using the tree-ring detrending program Spotty. *Dendrochronologia* **27**: 75–82.

Esper J, Frank DC, Büntgen U, *et al.* 2010. Trends and uncertainties in Siberian indicators of 20th century warming. *Global Change Biology* **16**: 386–398.

Esper J, Frank DC, Timonen M, *et al.* 2012a. Orbital forcing of tree-ring data. *Nature Climate Change* **2**: 862–866.

Esper J, Büntgen U, Timonen M, *et al.* 2012b. Variability and extremes of Northern Scandinavian summer temperatures over the past two millennia. *Global and Planetary Change* **88–89**: 1–9.

Esper J, Büntgen U, Luterbacher J, *et al.* 2013a. Testing the hypothesis of post-volcanic missing rings in temperature sensitive dendrochronological data. *Dendrochronologia* **31**: 216–222.

Esper J, Schneider L, Krusic PJ, *et al.* 2013b. European summer temperature response to annually dated volcanic eruptions over the past nine centuries. *Bulletin of Volcanology* **75**: 1–14.

Franke J, Frank DC, Raible CC, *et al.* 2013. Spectral biases in tree-ring climate proxies. *Nature Climate Change* **3**: 360–364.

Fritts HC. 1976. *Tree Rings and Climate*. Academic Press: New York.

Grudd H. 2008. Torneträsk tree-ring width and density AD 500–2004: A test of climatic sensitivity and a new 1500-year reconstruction of north Fennoscandian summers. *Climate Dynamics* **31**: 843–857.

Kullman L. 2013. Ecological tree line history and palaeoclimate – review of megafossil evidence from the Swedish Scandes. *Boreas* **42**: 555–567.

Ljungqvist FC. 2010. A new reconstruction of temperature variability in the extra-tropical Northern hemisphere during the last two millennia. *Geografiska Annaler: Series A, Physical Geography* **92**: 339–351.

Ljungqvist FC, Krusic PJ, Brattström G, *et al.* 2012. Northern Hemisphere temperature patterns in the last 12 centuries. *Climate of the Past* **8**: 227–249.

Mann ME, Zhang Z, Hughes MK, *et al.* 2008. Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 13252–13257.

Melvin T, Briffa KR. 2008. A 'signal-free' approach to dendroclimatic standardization. *Dendrochronologia* **26**: 71–86.

Melvin TM, Grudd H, Briffa KR. 2013. Potential bias in 'updating' tree-ring chronologies using Regional Curve Standardization: re-processing the Torneträsk maximum-latewood-density data. *Holocene* **23**: 364–373.

Moser L, Fonti P, Büntgen U, *et al.* 2010. Timing and duration of European larch growing season along an altitudinal gradient in the Swiss Alps. *Tree Physiology* **30**: 225–233.

Plummer CT, Curran MAJ, van Ommen TD, *et al.* 2012. An independently dated 2000-yr volcanic record from Law Dome, East Antarctica, including a new perspective on the dating of the 1450s CE eruption of Kuwae, Vanuatu. *Climate of the Past* **8**: 1929–1940.

Schneider L, Esper J, Timonen M, *et al.* 2014. Detection and evaluation of an early divergence problem in Northern Fennoscandian tree-ring data. *Oikos* **123**: 559–566.

Schweingruber FH, Fritts HC, Bräker OU, *et al.* 1978. The X-ray technique as applied to dendroclimatology. *Tree-Ring Bulletin* **38**: 61–91.

Schweingruber FH, Bartholin T, Schär E, *et al.* 1988. Radiodensitometric-dendroclimatological conifer chronologies from Lapland (Scandinavia) and the Alps (Switzerland). *Boreas* **17**: 559–566.

Sigl M, McConnell JR, Layman L, *et al.* 2013. A new bipolar ice core record of volcanism from WAIS Divide and NEEM and implications for climate forcing of the last 2000 years. *Journal of Geophysical Research A* **118**: 1151–1169.