



AS WE SEE IT

Effects of sample size in dendroclimatology

Ulf Büntgen^{1,2,*}, Willy Tegel^{3,*}, Karl-Uwe Heussner⁴, Jutta Hofmann⁵,
Raymond Kontic⁶, Tomáš Kyncl⁷, Edward R. Cook⁸

¹Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland

²Oeschger Centre for Climate Change Research, 3012 Bern, Switzerland

³Institute for Forest Growth IWW, University of Freiburg, 79085 Freiburg, Germany

⁴German Archaeological Institute DAI, 14195 Berlin, Germany

⁵Jahrringlabor Hofmann, 72622 Nürtingen, Germany

⁶Labor Dendron, 4057 Basel, Switzerland

⁷Moravian Dendro-Labor, 61600 Brno, Czech Republic

⁸Tree-Ring Laboratory, Lamont-Doherty Earth Observatory, Palisades, New York 10964, USA

ABSTRACT: Uncertainty related to the rate and magnitude of predicted anthropogenic climate change highlights the need to enhance our understanding of past natural fluctuations in the Earth's climate system. This task emphasizes the importance of high-resolution palaeoclimatic records that cover industrial and pre-industrial times. Annually resolved and absolutely dated tree-ring measurements are a key input for cross-disciplinary research. Ambiguity due to paucity of data, however, characterizes many tree-ring data analyses. By utilizing nearly 12 000 living and historical ring width series from European fir *Abies alba* Mill., we demonstrate how massive sample replication can generate robust estimates of past growth rates, which may help reduce methodological and statistical constraints associated with many traditional tree-ring studies.

KEY WORDS: Climate change · Dendroclimatology · European fir · Forest growth · Palaeoclimatology · Tree rings

— Resale or republication not permitted without written consent of the publisher —

1. MOTIVATION

The information captured in environmental proxy records may range from interannual extremes to millennial trends, and the accurate preservation of all resolvable frequency-dependent scales of variation in these time series is still a pressing challenge in palaeoclimatology and palaeoecology. Disentangling externally forced climatological signals from ecological and physiological background noise in biological proxy archives such as tree rings is complicated because biotic and abiotic processes and their observed properties often operate at different spatio-temporal scales (Frank et al. 2007a, Cook & Pederson 2011). Such non-climatic processes can even behave nonlinearly when critical thresholds are surpassed (Ims et al. 2008, Johnson et al. 2010).

Annually resolved tree-ring chronologies based on living trees and relict materials of different age classes can potentially be used to unravel climatological from non-climatological influences on forest growth. The expected outcome from such composite records is, however, often affected by insufficient sample replication over time and diminished site control of the relict data. Such limitations can affect the signal-to-noise ratio of the mean tree-ring chronology (Wigley et al. 1984, Jones et al. 1998, 2001), and translate into artificial variance changes over time (Osborn et al. 1997, Frank et al. 2007b). Associated biases are known to reduce our ability to track the intensity and frequency of historical extremes from palaeoclimatic records (Battipaglia et al. 2010, Büntgen et al. 2011a). Such limitations also influence longer-term amplitude changes (Esper et al.

*Email: buentgen@wsl.ch, tegel@dendro.de

2005, Büntgen et al. 2006, Esper & Frank 2009a), and thus create difficulties for any timescale-dependent reconstruction of climate variability from tree rings (Osborn & Briffa 2000). Limitations of dendroclimato-logical and dendroecological studies are often also related to the obligatory tree-ring detrending and index calculation procedures applied (Briffa et al. 1992, Cook et al. 1995, Cook & Peters 1997, Esper et al. 2003, Büntgen et al. 2008, Melvin & Briffa 2008, Esper & Frank 2009b, Briffa & Melvin 2011). Consequently, many tree-ring studies have not adequately preserved spatiotemporally consistent variations across the full spectrum of interannual to centennial and longer timescales (Wilson et al. 2007, Büntgen et al. 2010a,b, Esper et al. 2012).

Here we demonstrate that greatly enhanced sample replication can assist in overcoming some of the above-stated methodological and statistical constraints. We suggest that the use of datasets of sufficient size can improve the signal-to-noise ratio in those tree-ring chronologies for which one dominant climatic control is not readily apparent or even expected. We also propose that not only living trees from temperate mesic forests but also historical, sub-fossil, and archaeological materials from similar environments can provide dendroclimatological and dendroecological opportunities if the sample size is sufficiently large.

2. EXAMPLE

A total of 11 873 annually-resolved and absolutely-dated ring width measurement series from living trees and historical construction timbers of European fir *Abies alba* Mill. growing in temperate low- to mid-elevation forests across France, Switzerland, Germany, and the Czech Republic have been assembled over the past 4 decades (see Büntgen et al. 2011a for details). All individual core and disc samples containing >50 rings have been averaged into 3 regional subsets (West, Mid, East) north of the Alpine arc (Fig. 1). Subset replication, without data overlap, ranges from 3517 to 4413 series (Fig. 2). The dataset covers 962 to 2007 AD, and the individual measurement series are relatively evenly distributed through-

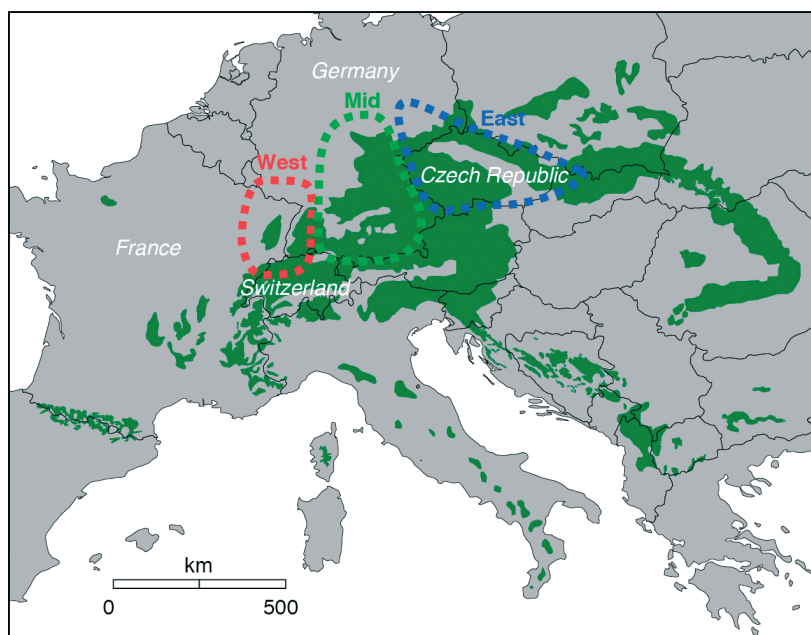


Fig. 1. Distribution of natural European fir *Abies alba* habitats, and areas from which ring width samples from living trees and historical timbers were obtained. A total of 11 873 individual samples were assembled in 3 regional subsets: West, Mid, and East, comprising 4413, 3517 and 3943 series, respectively

out this period. Similar degrees of sample replication and comparable temporal spreading of series start and end dates among the subsets are an indication of fairly synchronized settlement and building activities at the sub-continental scale. Decreasing sample replication in the 12th to 13th centuries and at the transition between the historical and living material in the early 19th century likely affects the reliability of our results during these intervals. Mean segment length (MSL: mean number of rings per sample) ranges from 81 to 85 yr and average growth rate (AGR: average ring width per sample) ranges from 1.74 to 1.84 mm (Fig. 2).

The 3 regional chronologies, without application of any tree-ring detrending, correlate significantly ($p < 0.001$) at 0.73 over their common period 1133 to 1996 AD (Fig. 3), which is replicated by at least 20 series yr^{-1} in each region. Cross-correlation of the 3 records increases to 0.81 after each ring width series was individually detrended with a 20 yr cubic smoothing spline (see Büntgen et al. 2011a for details). The raw data reflect comparable growth fluctuations at inter-annual to centennial timescales over most of the past millennium. Relatively low productivity levels are centered on ~1200, ~1450, ~1700, ~1800 AD, and in the 1970s. This last period of suggested growth reduction, however, may possibly be biased by the relative lack of randomly overlapping series in the 20th

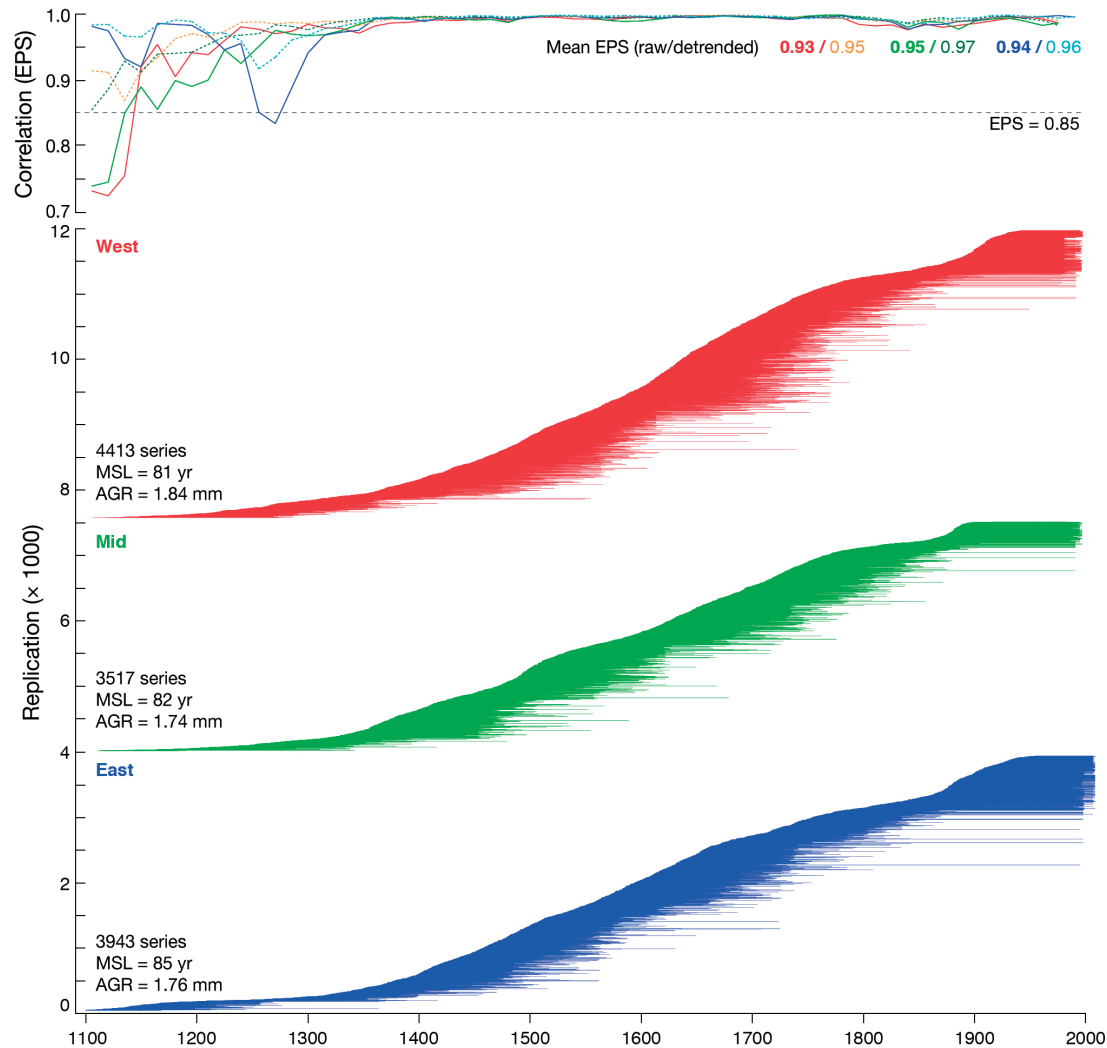


Fig. 2. *Abies alba*. Top: expressed population signal (EPS) independently computed for each of the 3 regional subsets (West: red/orange; Mid: green shades; East: blue shades) over 30 yr windows lagged by 15 yr (using raw and detrended data). Dashed black line: quality threshold of 0.85. Lower graphs: temporal sample distribution. Each horizontal bar represents one of the 11 873 individual fir measurement series sorted by region (West, Mid, East) and date of the innermost ring. MSL = Mean Segment Length; AGR = Average Growth Rate

century compared to the earlier periods (Fig. 2). While the earlier growth decreases may simply reproduce poorer climatic conditions, the 1970s depression may be driven by a complex interplay of airborne pollution and soil desiccation. Conversely, considerably higher productivity rates across all 3 regions are centered on ~1350, ~1500, ~1750, and from ~1875 to 1950 AD. Again, reduced random overlaps of living tree samples may contribute to the modern period of suggested higher growth productivity.

A relatively low mean inter-series correlation (RBAR; Wigley et al. 1984) of 0.23 was found between all series contained within each regional subset over the past millennium, which is probably the result of the broad geographic scale of the sampled subsets.

This value increases slightly to 0.31 after 20 yr spline detrending. The Expressed Population Signal (EPS; Wigley et al. 1984) is 0.94 for the raw and 0.96 for the detrended data. The EPS statistic is a summary measure of chronology uncertainty that determines how well a chronology, based on a finite number of samples, estimates the theoretical population chronology from which it was drawn (Wigley et al. 1984). The temporal expression of EPS shown here (Fig. 2), was independently computed over 30 yr windows lagged by 15 yr for each subset chronology, and is indicative of a very high level of internal signal strength, constantly ranging above the commonly applied quality threshold of 0.85. The considerable level of statistical agreement within the 3 regional chronologies sup-

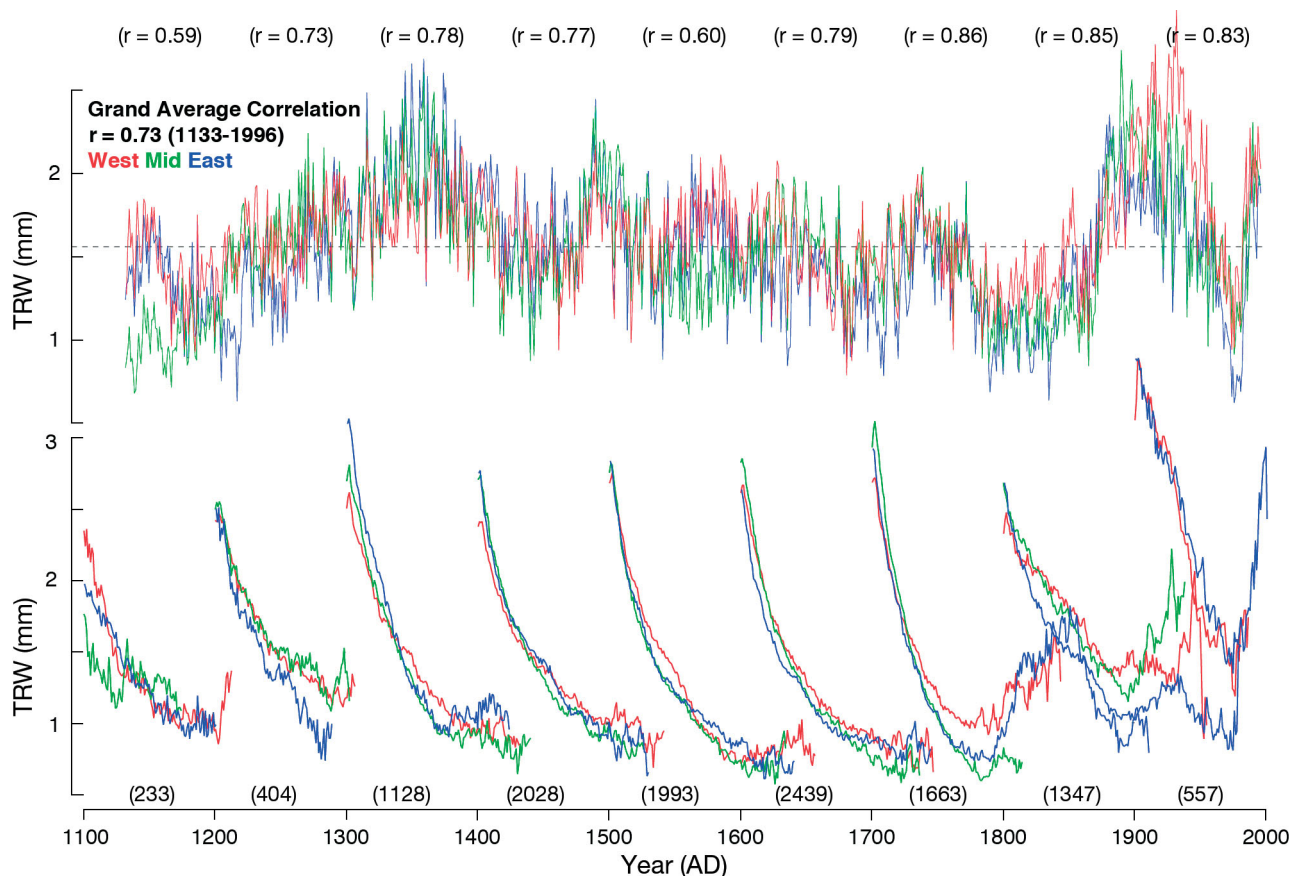


Fig. 3. *Abies alba*. Top: the 3 regional mean tree-ring width (TRW) chronologies (West, Mid, East), without detrending. These reveal significant agreement on interannual to multi-centennial timescales, indicated by their grand average correlation of 0.73 computed over the common period 1133 to 1996 AD. Cross-correlations per century are shown in parentheses. Bottom: Regional Curves (RCs), independently computed for each of the 3 subsets and for each century between 1100 and 2000 AD. We aligned all individual fir ring width measurement series by their cambial age and calculated the average growth functions per region and century. The resulting RCs were truncated at a minimum replication of 20 series, and the numbers in parentheses refer to the total number of samples with start dates within the corresponding century, i.e. that entered the step-wise calculated RCs. No RC was calculated for the 20th century in the middle region (Mid; green), because only living trees older than 100 yr were sampled in this region

ports the remarkable level of similarity in Central European fir growth rates, trends and extremes found from the French Jura and Vosges mountains in the northwest to Moravia at the border between the Czech Republic and Slovakia in the southeast (Fig. 1; see also Büntgen et al. 2011a for details). The sub-continental growth coherency obtained from the averaged raw measurement series also implies that even without detrending, i.e. statistical age trend removal, it is possible to preserve common inter-annual to centennial growth trends over space and time if that is the goal of the study (Fig. 3). Nevertheless, it is undeniable that robust estimates of past environmental conditions, including climate fluctuations, can only be obtained through the judicious use of regional curve standardization (RCS; Briffa et al. 1992, Briffa & Melvin 2011) and/or signal free detrending (SFD; Melvin & Briffa 2008).

Regional curves (RCs), based on realigning the raw measurement series by the biological age of each tree, have been independently calculated for each region and each century (Fig. 3). The date of the innermost ring of each individual sample defines its affiliation to a specific century: series with start years between 1700 and 1799 AD were combined in the RC of the 18th century, for instance. Replication per century and region varies between 233 and 2439 series, with numbers >2000 occurring during the 15 to 18th centuries. These peak values likely correspond to synchronized settlement activities at the sub-continental scale. The resulting RCs describe spatiotemporal changes in average annual growth rates relative to cambial ages, and are, except for the 19th century, remarkably homogeneous across regions. Age-related productivity trends during almost all centuries resemble negative exponential functions

that are indicative of relatively stable old-growth forests with little disturbance, whereas the more variable regional growth trends before ~1300 AD and in the 19th century may reflect more stand dynamics and competition effects perhaps related to heightened human intervention and forest exploitation. Intense fir dying in the 1970s likely reduced competition for the surviving trees, which subsequently exhibited higher modern growth rates (Fig. 3). Some bias before ~1300 AD could also result from generally reduced sample sizes among each subset (Fig. 2). Temporal changes in growth trends throughout the past millennium may also reveal large-scale changes in forest management and resource utilization and, thus, could conceivably be used to gain insight on past harvesting intensity. Since the century-scale growth trends shown here are highly similar among the 3 regions, we assume that non-climatic factors that could have influenced tree growth and stand vigor at local-to-regional scales have been largely averaged out for most of the last 1000 yr during times with sufficient random sample overlaps. However, this hypothesis likely does not hold for the 20th century, which is a special case that requires further analysis.

3. UTILITY AND OUTLOOK

The value of our European fir compilation is expressed by the extremely high level of common signal strength both within and between the regional subsets. Sub-continental and frequency-independent growth coherency demonstrates the palaeoclimatological and palaeoecological value of living fir trees in combination with historical, subfossil and archaeological fir material from Central European temperate forests (Fig. 1). Nevertheless, the application of proper detrending procedures remains necessary prior to any valid climatological and/or ecological interpretation.

Our example is particularly encouraging, as comparable datasets have already been collected and dendrochronologically investigated, i.e. sampled, cross-dated, and absolutely dated to calendar years. Moreover, it should be noted that the high signal strength measured by the EPS of each regional chronology was achieved even though the samples represent a broad geographical area and comprise an undoubtedly large variety of ecological site conditions and types. This study, in contrast to more targeted stand-level sampling strategies commonly applied in dendroclimatology and dendroecology,

confirms the ability of massive regional replication, i.e. with a very high sample size per year, to account for and largely suppress site inhomogeneity.

More generally, we found a high level of common growth variability in tree-ring samples from non-extreme sites collected by non-standard sampling protocols, in line with the outcome of the so-called random update sampling strategy (Tegel et al. 2010). We therefore propose supplementing unsystematically collected historical material with randomly sampled recent beams and timbers from sawmills and lumberyards. Such a combination would be ideal, in order to avoid statistical over-fitting during the proxy/target calibration interval. The resulting signal degradation in the updated portion of the hybrid chronology should conform more closely to the level of signal strength in the ancient portion of the tree-ring chronology. Such composites that include modern material, ideally scattered over the same region from where the relict data originates, will likely also enhance the development of independent and species-specific regional master chronologies for more reliable dating of historical timbers, wooden artifacts, subfossil remains, and archaeological finds.

Massive replication assists in overcoming methodological and statistical constraints in tree-ring chronology estimation (Büntgen et al. 2011b). Similar benefits could be expected for the development of other climate proxy records and also for more general bio-ecological approaches that utilize either long time series or apply space-for-time substitutes. While massive replication is highly desirable, the importance of an adequate sample size over time, as here, must also be emphasized. When this is the case, artificial changes in variance caused by varying sample size in long composite chronologies can be largely avoided or robustly corrected for (Frank et al. 2007b). This would enable more believable extreme year analyses to be performed over exceptionally long timescales when considering samples of different age classes and when utilizing appropriate detrending methods to best capture year-to-year variability (Battipaglia et al. 2010, Büntgen et al. 2011a).

Massive replication of the kind described here should produce more reliable regional tree-ring chronologies of fir and other tree species for the study of past environmental changes through the application of the RCS method (Briffa et al. 1992, Esper et al. 2003). RCS and comparable detrending techniques can preserve centennial or longer timescale variability due to external forcings such as climate in excess of the segment lengths of the individual tree-ring series being used (Cook et al. 1995). In addition, vast

replication allows for the rigorous assessment of the accuracy of RCS chronology (Briffa & Melvin 2011), including the results of the latest methodological innovations, such as signal-free and multiple-RCS methods of chronology development (see Briffa & Melvin 2011 for details). This will result in better reconstructions of the overall range of climate variability in millennium-long chronologies (Esper et al. 2005, Büntgen et al. 2006, Esper & Frank 2009a). Replication-induced biases may also impact the so-called 'divergence' and 'end-effect' issues (Cook et al. 1995, Cook & Peters 1997, Melvin & Briffa 2008, Cook & Pederson 2011), which are not exclusively caused by climate-driven growth response shifts (Wilson et al. 2007, Büntgen et al. 2008, Esper & Frank 2009b).

Events causing long-lasting and continent-wide political, socio-economic and cultural changes, such as the second plague pandemic (the 'Black Death'), which likely reduced the population of Europe by 40 to 60% from 1347 AD onwards, or the widespread destruction of settlements during the Thirty Years' War in the first half of the 17th century, may have left discernable fingerprints in the continent's dendrochronological record (Büntgen et al. 2011b). Historical tree-ring compilations, if replicated well enough, may thus offer new possibilities for data analysis and interpretation that extend beyond traditional disciplinary boundaries. For example, the utilization of thousands or even hundreds of thousands of historical felling dates of trees could create new proxy archives for tracing spatiotemporal dynamics in historical settlement and building activities (Baillie 2002).

Our example also points to the potential value of massively replicated tree-ring datasets to the broad community studying vegetation and forests, including biologists, ecologists, historians, and even economists. By providing a millennium-long perspective of natural variation in growth rates of one of the most important European conifer species, indirect estimates of past fir biomass productivity may allow modern production levels to be placed against a backdrop of pre-industrial fluctuations. Central European fir growth rates described here for the 20th century exceed all previous centuries during the past millennium, although it remains to be seen if these modern peak values are biased by the lack of random data overlaps, as is the case prior to the 20th century. The extent to which this recent growth increase is unprecedented relative to even longer-term changes in forest productivity over the Holocene remains subject to further investigations. These could include different proxy reconstructions in tandem with inde-

pendent model simulations. Since tree-ring data do not provide information about species abundance, future studies of the kind outlined here should ideally be combined with high-resolution pollen evidence to provide a more complete evaluation of past vegetation dynamics.

Although we are aware that the amount of tree-ring replication used in this study contains hundreds or even thousands of times more series than most of the tree-ring studies so far published, we are convinced of its potential value for stimulating research in interdisciplinary arenas of the environmental sciences. Therefore, the development of highly replicated tree-ring datasets from as many tree species and geographic locations as possible should be encouraged and made available to the greater scientific community for analysis and study. At the same time, we are sensitive to the ethical component and fiscal complexity that may arise when data are exchanged among academic and economic partners, such as federal research institutes and private companies specializing in tree-ring dating. Nevertheless, we specifically want to encourage individual actors among different institutions, including our close colleagues, to start systematically aggregating, managing and allocating measurements not only from living trees, but also from historical, subfossil and archaeological material, in a highly transparent way, e.g. via the International Tree Ring Data Bank (ITRDB; www.ncdc.noaa.gov/paleo/treering.html). Synergetic effects beyond traditional academic and commercial economic boundaries will likely compensate for the effort.

Acknowledgements. F. H. Schweingruber, H. P. Kahle, J. Esper, and D. Frank stimulated discussion. Supported by the DFG grant #SP437/16-1, and the Eva Mayr-Stihl Foundation. E.R.C. acknowledges the NOAA OGP Climate Change and Global Change Program, grant NA10OAR4310123. Lamont-Doherty Earth Observatory publication number 7650. P. J. Krusic and 3 anonymous referees kindly commented on an earlier version of this manuscript.

LITERATURE CITED

- Baillie MGL (2002) Future of dendrochronology with respect to archaeology. *Dendrochronologia* 20:69–85
- Battipaglia G, Frank DC, Büntgen U, Dobrovolny P, Brazdil R, Pfister C, Esper J (2010) Five centuries of Central European temperature extremes reconstructed from tree-ring density and documentary evidence. *Global Planet Change* 72:182–191
- Briffa KR, Melvin TM (2011) A closer look at Regional Curve Standardization of tree-ring records: justification of the need, a warning of some pitfalls, and suggested improvements in its application. In: Hughes MK, Swetnam TW, Diaz HF (eds) *Dendroclimatology: progress and prospects*. Springer, Dordrecht, p 113–145

- Briffa KR, Jones PD, Bartholin TS, Eckstein D and others (1992) Fennoscandian summers from AD 500: temperature changes on short and long timescales. *Clim Dyn* 7: 111–119
- Büntgen U, Frank DC, Nievergelt D, Esper J (2006) Summer temperature variations in the European Alps, AD 755–2004. *J Clim* 19:5606–5623
- Büntgen U, Frank DC, Wilson R, Carrer M, Urbinati C, Esper J (2008) Testing for tree-ring divergence in the European Alps. *Glob Change Biol* 14:2443–2453
- Büntgen U, Franke J, Frank D, Wilson R, Gonz  les-Rouco F, Esper J (2010a) Assessing the spatial signature of European climate reconstructions. *Clim Res* 41:125–130
- Büntgen U, Trouet V, Frank D, Leuschner HH, Friedrichs D, Luterbacher J, Esper J (2010b) Tree-ring indicators of German summer drought over the last millennium. *Quat Sci Rev* 29:1005–1016
- Büntgen U, Br  zdil R, Heussner KU, Hofmann J and others (2011a) Combined dendro-documentary evidence of Central European hydroclimatic springtime extremes over the last millennium. *Quat Sci Rev* 30:3947–3959
- Büntgen U, Tegel W, Nicolussi K, McCormick M and others (2011b) 2500 years of European climate variability and human susceptibility. *Science* 331:578–582
- Cook ER, Pederson N (2011) Uncertainty, emergence, and statistics in dendrochronology. In: Hughes MK, Swetnam TW, Diaz HF (eds) *Dendroclimatology: progress and prospects*. Springer, Dordrecht, p 77–112
- Cook ER, Peters K (1997) Calculating unbiased tree-ring indices for the study of climatic and environmental change. *Holocene* 7:361–370
- Cook ER, Briffa KR, Meko DM, Graybill DA, Funkhouser G (1995) The segment length curse in long tree-ring chronology development for palaeoclimatic studies. *Holocene* 5:229–237
- Esper J, Frank DC (2009a) The IPCC on a heterogeneous Medieval Warm Period. *Clim Change* 94:267–273
- Esper J, Frank DC (2009b) Divergence pitfalls in tree-ring research. *Clim Change* 94:261–266
- Esper J, Cook ER, Krusic PJ, Peters K, Schweingruber FH (2003) Tests of the RCS method for preserving low-frequency variability in long tree-ring chronologies. *Tree-Ring Res* 59:81–98
- Esper J, Wilson RJS, Frank DC, Moberg A, Wanner H, Luterbacher J (2005) Climate: past ranges and future changes. *Quat Sci Rev* 24:2164–2166
- Esper J, B  ntgen U, Timonen M, Frank DC (2012) Variability and extremes of northern Scandinavian summer temperatures over the past two millennia. *Global Planet Change* 88–89:1–9
- Frank D, B  ntgen U, B  hm R, Maugeri M, Esper J (2007a) Warmer early instrumental measurements versus colder reconstructed temperatures: shooting at a moving target. *Quat Sci Rev* 26:3298–3310
- Frank D, Esper J, Cook ER (2007b) Adjustment for proxy number and coherence in a large-scale temperature reconstruction. *Geophys Res Lett* 34. doi:10.1029/2007GL030571
- Ims RA, Henden JA, Killengreen ST (2008) Collapsing population cycles. *Trends Ecol Evol* 23:79–86
- Johnson DM, B  ntgen U, Kausrud K, Frank DC and others (2010) Climate change forces elevation shift in outbreak epicenter of larch budmoth. *Proc Natl Acad Sci USA* 107: 20576–20581
- Jones PD, Briffa KR, Barnett TP, Tett SFB (1998) High-resolution palaeoclimatic records for the last millennium: interpretation, integration and comparison with General Circulation Model control-run temperatures. *Holocene* 8:455–471
- Jones PD, Osborn TJ, Briffa KR (2001) The evolution of climate over the last millennium. *Science* 292:662–667
- Melvin TM, Briffa KR (2008) A ‘signal-free’ approach to dendroclimatic standardisation. *Dendrochronologia* 26:71–86
- Osborn TJ, Briffa KR (2000) Revisiting timescale-dependent reconstruction of climate from tree-ring chronologies. *Dendrochronologia* 18:9–26
- Osborn TJ, Briffa KR, Jones PD (1997) Adjusting variance for sample-size in tree-ring chronologies and other regional-mean timeseries. *Dendrochronologia* 15:89–99
- Tegel W, Vanmoerkerke J, B  ntgen U (2010) Updating historical tree-ring records for climate reconstruction. *Quat Sci Rev* 29:1957–1959
- Wigley TML, Briffa KR, Jones PD (1984) On the average of correlated time series, with applications in dendroclimatology and hydrometeorology. *J Clim Appl Meteorol* 23: 201–213
- Wilson RJS, D’Arrigo R, Buckley B, B  ntgen U and others (2007) A matter of divergence: tracking recent warming at hemispheric scales using tree-ring data. *J Geophys Res* 112. doi:10.1029/2006JD008318

Editorial responsibility: Nils Chr. Stenseth, Oslo, Norway

*Submitted: January 4, 2012; Accepted: April 26, 2012
Proofs received from author(s): June 29, 2012*