The influence of sampling design on tree-ring-based quantification of forest growth

CHRISTOPH NEHRBASS-AHLES[1]*, FLURIN BABST[1,2], STEFAN KLESSE[1,3], MAGDALENA NÖTZLI[1], OLIVIER BOURIAUD[4], RAPHAEL NEUKOM[1,3,5], MATTHIAS DOBBERTIN[1] and DAVID FRANK[1,3]

[1]*Swiss Federal Research Institute WSL, Zürcherstrasse 111, Birmensdorf CH-8903, Switzerland, [2]Laboratory of Tree-Ring Research, University of Arizona, 1215 E Lowell St, Tucson, AZ 85721, USA, [3]Oeschger Centre for Climate Change Research, University of Bern, Zähringerstrasse 25, Bern CH-3012, Switzerland, [4]National Forest Inventory, Forest Research and Management Institute, ICAS, 128 Bulevardul Eroilor, Voluntari O77190, Romania, [5]Department of Geography, University of Zurich, Zurich CH-8057, Switzerland*

Abstract

Tree-rings offer one of the few possibilities to empirically quantify and reconstruct forest growth dynamics over years to millennia. Contemporaneously with the growing scientific community employing tree-ring parameters, recent research has suggested that commonly applied sampling designs (i.e. how and which trees are selected for dendrochronological sampling) may introduce considerable biases in quantifications of forest responses to environmental change. To date, a systematic assessment of the consequences of sampling design on dendroecological and-climatological conclusions has not yet been performed. Here, we investigate potential biases by sampling a large population of trees and replicating diverse sampling designs. This is achieved by retroactively subsetting the population and specifically testing for biases emerging for climate reconstruction, growth response to climate variability, long-term growth trends, and quantification of forest productivity. We find that commonly applied sampling designs can impart systematic biases of varying magnitude to any type of tree-ring-based investigations, independent of the total number of samples considered. Quantifications of forest growth and productivity are particularly susceptible to biases, whereas growth responses to short-term climate variability are less affected by the choice of sampling design. The world's most frequently applied sampling design, focusing on dominant trees only, can bias absolute growth rates by up to 459% and trends in excess of 200%. Our findings challenge paradigms, where a subset of samples is typically considered to be representative for the entire population. The only two sampling strategies meeting the requirements for all types of investigations are the (i) sampling of all individuals within a fixed area; and (ii) fully randomized selection of trees. This result advertises the consistent implementation of a widely applicable sampling design to simultaneously reduce uncertainties in tree-ring-based quantifications of forest growth and increase the comparability of datasets beyond individual studies, investigators, laboratories, and geographical boundaries.

Keywords: carbon cycle, climate reconstruction, climate response, CO_2 fertilization, forest productivity, growth trends, sampling bias, tree-rings

Received 26 November 2013; revised version received 5 March 2014 and accepted 1 April 2014

Introduction

The response of terrestrial ecosystems to environmental variability and extremes is one of the largest uncertainties, not only in predicting the ecological impacts of

Correspondence: *Present address: Christoph Nehrbass-Ahles, Climate and Environmental Physics, Physics Institute, University of Bern, Sidlerstrasse 5, Bern CH-3012, Switzerland and Oeschger Centre for Climate Change Research, University of Bern, Zähringerstrasse 25, Bern CH-3012, Switzerland, tel. +41 31 631 4466, fax +41 31 631 8742,
e-mail: nehrbass@climate.unibe.ch; David Frank, tel. +41 44 7392 282, fax +41 44 7392 215, e-mail: david.frank@wsl.ch

climate change, but also in constraining the magnitude of future global warming owing to feedbacks between the terrestrial carbon, nitrogen, hydrological, and vegetation cycles (Reichstein *et al.*, 2013). Understanding the magnitude and dynamics of ecosystem processes connected with ongoing global change requires long-term observational records to benchmark current phenomena with respect to long-term baselines. The impacts from extreme events such as droughts or gradual environmental change such as the response of forests to anthropogenic emissions require particularly long observational time series. These statements are punctuated by recent evidence that controlled experiments fail

to capture fundamental aspects of ecosystem functioning and responses to climate variability (Wolkovich *et al.*, 2012).

Long-term records including phenological observations (Menzel *et al.*, 2006; Rutishauser *et al.*, 2008), forest inventory and long-term monitoring data (Pan *et al.*, 2011; Nabuurs *et al.*, 2013), as well as tree-ring chronologies (Esper *et al.*, 2012; Girardin *et al.*, 2014) have provided novel insights into the responses of terrestrial ecosystems to environmental variation and extremes. Yet, dendrochronological methods offer one of the few possibilities to reconstruct and quantify climatic or ecological events years to millennia after they occurred (Esper *et al.*, 2007a; Salzer & Hughes, 2007; Cook *et al.*, 2010; Li *et al.*, 2013). In this context, there has been a resurgence of questions related to global change processes that can be both asked and answered by tree-ring records. The commonly measured tree-ring width (TRW) parameter has been used, e.g. to investigate impacts of rising atmospheric CO_2 concentrations (Gedalof & Berg, 2010; Peñuelas *et al.*, 2011), carbon sink dynamics (Knorre *et al.*, 2005; Richardson *et al.*, 2010), climate reconstruction (Jones *et al.*, 2009; PAGES 2k Consortium, 2013), forests' climate response (Carrer & Urbinati, 2004; Lévesque *et al.*, 2013), testing and validating vegetation models (Girardin *et al.*, 2011; Babst *et al.*, 2013), and ground-truthing remote sensing observations (Beck *et al.*, 2011; Poulter *et al.*, 2013).

Contemporaneously with the growing and increasingly diverse scientific community employing tree-ring datasets, recent research spotlights various sources of uncertainty that still remain in dendrochronological investigations (Frank *et al.*, 2010; Briffa *et al.*, 2013; Franke *et al.*, 2013). Sampling design (i.e. how and which trees are selected for dendrochronological sampling) has been identified as an important caveat which may significantly contribute to these uncertainties or even lead to biased results and conclusions (Cherubini *et al.*, 1998; Melvin, 2004; Bontemps & Esper, 2011; Briffa & Melvin, 2011; Brienen *et al.*, 2012; Bowman *et al.*, 2013). Herein, we use the term 'sampling bias' to refer to biases (relative to the full population) whose origin can be traced to how and which trees are selected for subsequent tree-ring investigations.

To date, a diverse array of sampling schemes has been utilized often following study-specific procedures that may not even be well defined in the corresponding literature. Typical dendrochronological sampling may include only dominant individuals (Briffa *et al.*, 1998), stacked logs from lumberyards (Tegel *et al.*, 2010), or proportionally represent stand characteristics such as tree-size distribution (Martínez-Vilalta *et al.*, 2008; see Table 1 for more examples). The ability of the most commonly applied sampling scheme – namely the sampling of dominant, large, and healthy trees – to accurately represent ecosystem responses to environmental variation is increasingly questioned (Cherubini *et al.*, 1998; Melvin, 2004; Briffa & Melvin, 2011; Brienen *et al.*, 2012).

Cherubini *et al.* (1998) demonstrated how growth variation quantified from sampling the largest diameter trees at a given point in time will show systematic, yet unpredictable differences compared to the largest diameter trees sampled at other points in time due to temporal shifts in tree-size ranking. Melvin (2004) formally defined the so-called 'Modern-Sample Bias' that may arise from sampling currently living individuals above a certain diameter threshold and stems from the premises that forests contain trees with roughly the same diameter, but varying ages (i.e. different growth rates). Furthermore, slow growing trees are less likely to die at a given age and tend to grow older than fast growing trees (Black *et al.*, 2008; Bigler & Veblen, 2009). Consequently, sampling only the largest living trees may alter tree-ring chronologies developed from common techniques such as Regional Curve Standardization (RCS; Esper *et al.*, 2003), which could be biased low during their earlier portions and high towards present (Briffa & Melvin, 2011). Recently, Brienen *et al.* (2012) performed analyses to assess the magnitude of biases that might confound the detection of CO_2 fertilization effects. This publication subdivided the 'Modern-Sample Bias' (Briffa & Melvin, 2011) into (i) the more specific 'Slow-Grower Survivorship Bias' associated with different tree longevity/growing speed; and (ii) the 'Big-Tree Selection Bias' emerging from size-dependent sampling. These two intertwined biases were tested by means of diameter at breast height (DBH) reconstructions from a *Cedrela odorata* dataset (Brienen & Zuidema, 2006) and using a stochastic simulation model for a hypothetical tree species. This assessment showed that the 'Big-Tree Selection Bias' could lead to an erroneous growth increase of up to 50% (or even substantially higher) and suggested that positive trends in TRW often attributed to CO_2 fertilization (see Huang *et al.*, 2007; Brienen *et al.*, 2012, and references therein) might be an artefact of sampling biases.

While the above investigations have suggested poorly recognized or inherent challenges to accurately interpreting results from at least one tree-ring sampling scheme, the extent, nature, or generality of these concerns remain largely unaddressed. This issue is compounded by the diverse array of existing sampling schemes, and furthermore by the fact that similar types of investigations (e.g. climate reconstruction) themselves are often based on diverging sampling designs. This is particularly relevant for investigations where existing datasets are revisited to address new research

Table 1 Categorization of applied sampling designs in recent publications. The data were directly obtained/calculated from the corresponding publication and references therein. For further information see cited reference

	Sampling design					
	Trees per site	Species per Site	No. of sites	Selection of individuals	Designation	Publication
Climate reconstruction	12–15	1	>300	Dominant	Classic	Briffa *et al.*, 1998;
	10–72	1	10	All age classes	Pseudo-Population	Esper *et al.*, 2007b;
	~14	1	~200	Archaeological, subfossil, historical findings (Part I)	Forester	Büntgen *et al.*, 2011;
	~7	1	10	Random timbers obtained from sawmills yards (Part II)	Forester	Büntgen *et al.*, 2011;
Climate response	18–27	1	2	Dominant and supressed	Ad Hoc	Meyer & Bräker, 2001;
	~125	1	2	All age classes, with low intratree competition	Pseudo-Population	Carrer & Urbinati, 2004;
	12–15	5	2	(co-)dominant	Classic	Lévesque *et al.*, 2013;
Growth trends	8–10	1	3	All social statuses	Pseudo-Population	Bascietto *et al.*, 2004;
	~4	1	135	All DBH classes, according to their abundance	Pseudo-Population	Martínez-Vilalta *et al.*, 2008;
	5	All*	11	Largest	Ad Hoc (Classic)	Cole *et al.*, 2010;
	10–15	1	8	Dominant (at least 20 m apart)	Classic	Linares & Camarero, 2012;
Forest productivity	20–32	1–2	3	A variety of DBH classes	Pseudo-Population	Knorre *et al.*, 2005;
	1–62	2	30	Dominant, mature	Classic	Beck *et al.*, 2011;
	All	All	9	Rectangular fixed plots (0.06 ha), DBH ≥ 5 cm	Fixed plot	Liu *et al.*, 2012;
Forest dynamics	2	1	29	Two size classes (above/ below DBH = 50 cm)	Ad Hoc	Biondi, 1999;
	~93	2–4	2	Logs from woodpiles in the forest, DBH ≥ 60 cm	Forester	Brienen & Zuidema, 2006;
	16–21	1	70	Random within rect. fixed plots (1.3 ha), DBH ≥ 2.5 cm	Random	Pederson *et al.*, 2008;
	All	2	5	Circular fixed plots (~0.04 ha), DBH ≥ 5 cm	Fixed Plot	Weber *et al.*, 2008;
Forest Inventory	≤14	≥2	>600	Circular fixed plot (0.04 ha), according to guidelines†	(Classic)	CNFI, 2008;
	2–3	1–3		Circular fixed plot (~0.02 ha), (co-)dominant	(Classic)	US Forest Service, 2012

*Trees from each clonal colony (genet).

†DBH ≥ 9 cm; selection: (i) top DBH tree in plot; (ii) in each of four quadrants: largest tree of leading species, second leading species and other major species; (iii) one abitrary tree.

Gray shading emphasizes the diversity of approaches in literature for how and which trees are selected for dendrochronological sampling.

questions than they were originally intended to answer, especially if complied into larger networks. Examples of such follow-up investigations include climate reconstructions based on material which was initially collected for dating archaeological and historical structures (Büntgen *et al.*, 2011) or to determine stand age in national forest inventories (DeRose *et al.*, 2013). Likewise, publicly accessible datasets (e.g. via 'The International Tree-Ring Data Bank', ITRDB) collected for a plethora of different research questions have been

employed to address CO_2 fertilization effects (Gedalof & Berg, 2010) or the response of forest growth to climate variability (Wettstein *et al.*, 2011).

An improved understanding of how the choice of sampling influences the quantification of tree or ecosystem responses to environmental variation as well as biometric forest productivity assessments is urgently needed. In this study, we (i) categorize commonly applied sampling designs from the recent literature; (ii) sample a large population of trees and collect associ-

ated metadata; (iii) replicate diverse sampling designs by retroactively subsetting the population data; and (iv) test for emergent biases that would apply to investigations of past climate, growth response to climate variability, short to long-term growth trends, and forest productivity.

## Materials and methods

Study site

Our study was conducted at a long-term monitoring site in Switzerland, which is part of the Swiss 'Long-term Forest Ecosystem Research Programme' (LWF; Etzold *et al.*, 2014). The research plot is located at 480 m a.s.l. on the Swiss Plateau close to the municipality *Vordemwald* (Fig. 1a) in an uneven-aged, high forest stand. The site climate is moderate with a mean annual temperature of 8.4 °C and 1106 mm total precipitation peaking in the summer months (both computed over the recent 1978–2011 period from the nearest meteorological station located in *Wynau*; MeteoSwiss, 2013; Fig. 1b). The soil is a base-poor, temporary water stagnating Dysric Planosol (Graf Pannatier *et al.*, 2012). Monitoring parameters from the LWF programme include data on tree species, social class, and spatial position, quinquennial to annual measurements of DBH, height, and crown transparency, as well as site-specific information including meteoro-

logical data, litter fall, and soil water content since 1995 (Thimonier *et al.*, 2001).

The mean basal area (BA) of the population is 65.8 m² ha^{-1} (calculation based on DBH measurements after the 2011 growing season for a 0.5 ha subplot; see below) and is distributed over five principal species. European silver fir (*Abies alba*; ABAL) and Norway spruce (*Picea abies*, PCAB) are the dominant and codominant species, respectively. European beech (*Fagus sylvatica*; FASY) occurs frequently, but trees are small in diameter and thus contribute significantly less to the population BA than the fewer but larger sessile oak (*Quercus petraea*; QUPE) individuals (Fig. 2a). Sparsely occurring species like European larch (*Larix decidua*, LADE) as well as a very thin shrub layer do not significantly contribute to the total aboveground biomass.

For at least the past 200 years, this forest stand has been influenced by silvicultural management. Since approximately 1840, the so-called *Vorwaldsystem* (clear-cutting, but with an intermediate agricultural use of the forest floor; Bürgi & Stuber, 2003) has been practised. By 1917, the more natural regeneration-like group selection system was implemented and is still applied today (Murri, 1984). The precise management history is uncertain because the plot is small (2 ha) relative to the 56.2 ha forest division used for management plans. Over the past century, however, one can expect at least minor decadal silvicultural interferences (E. Steiner and D. Gautschi, 2012, personal communication), with the last selective thinning performed in 1995 when the permanent LWF plot was installed.
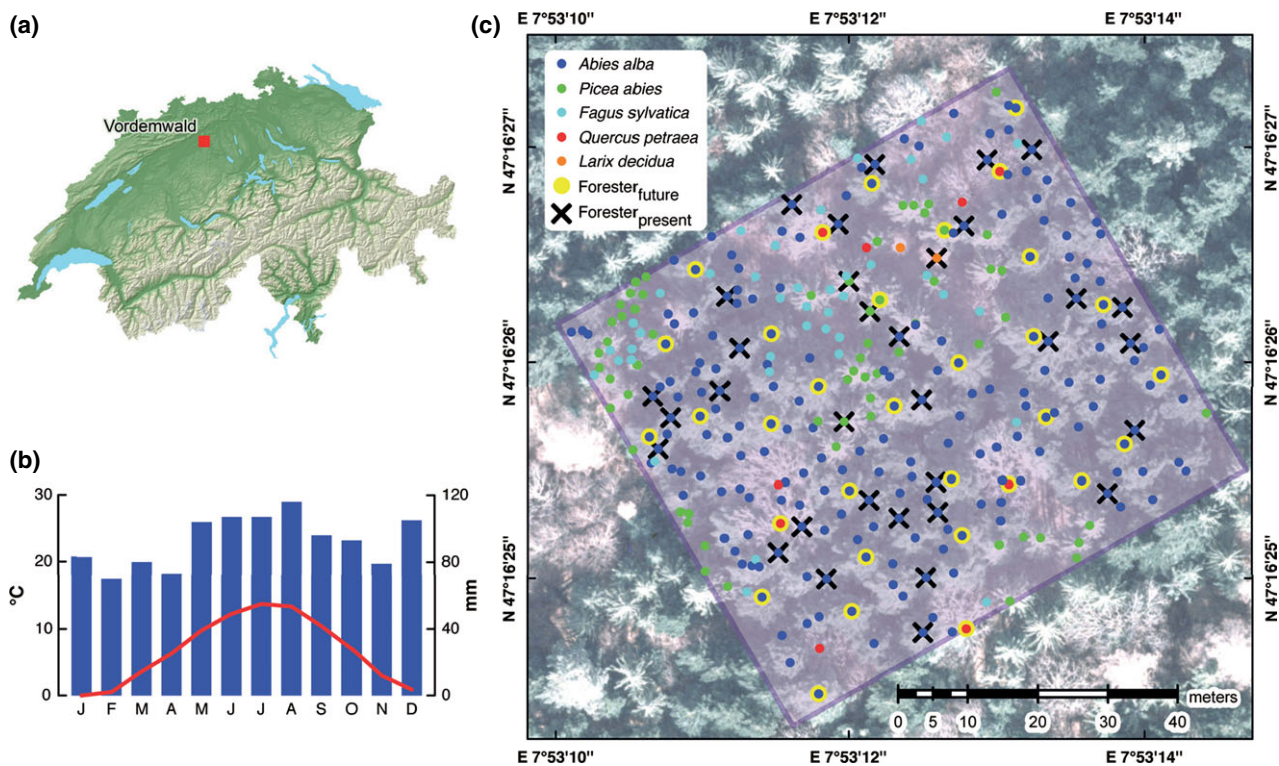
Fig. 1 (a) Location of the study site in Switzerland, (b) monthly averages of mean temperature (line) and total precipitation (bars), recorded from the nearest meteorological station (*Wynau*; MeteoSwiss, 2013) w.r.t. 1978–2011 and (c) a detailed map of the study area overlayed on an aerial photograph showing all sampled trees (dots) and those individuals selected for the 'Forester' sampling designs.
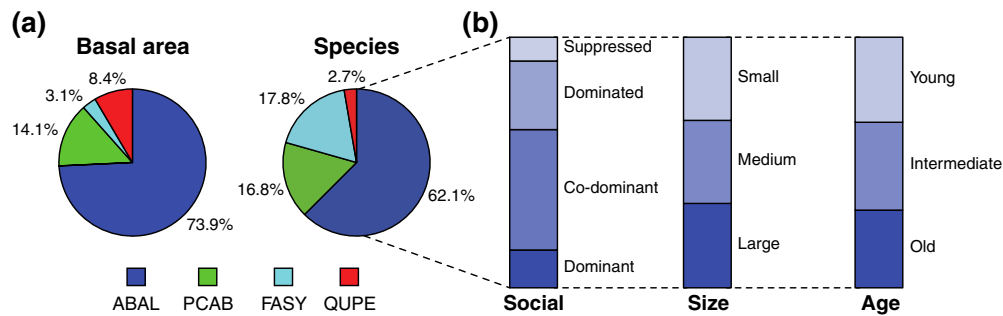
Fig. 2 (a) Distribution of total basal area and species within our study area for *Abies alba* (ABAL), *Picea abies* (PCAB), *Fagus sylvatica* (FASY), and *Quercus petraea* (QUPE). Two *Larix decidua* individuals (basal area: 0.5%, species: 0.6%) are not shown. (b) Stratification of ABAL into 3–4 classes according to social status, size, and age. See text for details.

Sample collection and measurements

To prevent callous formation over the coring wounds from influencing long-term tree diameter measurements, increment cores were taken at 0.9 m above the ground surface from every tree (dead or alive) exceeding 5.6 cm in DBH (Babst *et al.*, 2014a) within a defined 0.5 ha plot (Fig. 1c). To improve estimates of tree size and shape, two cores per tree were taken perpendicular to each other. The diameter was measured both at coring height and 1.3 m to maximize compatibility with existing monitoring data.

Samples were mounted, sanded, measured, cross-dated (visual, graphical and statistical pattern matching) and quality checked according to standard dendrochronological procedures (Speer, 2012). In case the pith was not present on a core, both samples from the respective tree were used to estimate the number of rings and distance to the pith utilizing a transparent template of concentric circles fitting to the curvature of the innermost rings. If exact dating was hampered by locally absent rings (typically the outermost rings of suppressed/stressed individuals prior to death; Wilmking *et al.*, 2012), the length of the problematic segment was measured, divided by the number of estimated years, and filled in as an average growth rate. Ultimately, samples were collected from 298 trees with a total of 35155 rings measured and dated. Locally absent rings were detected in 131 samples (25%) primarily among socially suppressed individuals with small diameters. Eight standing dead trees could not be sampled due to advanced wood decay.

Sampling designs

Subsequently, we introduce and categorize six general sampling schemes that we test in our study. While this selection is not fully encompassing, it represents the most commonly used strategies in recent literature (Table 1). The strategies are dubbed (i) Classic; (ii) Forester; (iii) Ad-hoc; (iv) Pseudo-Population; (v) Fixed Plot; and (vi) Random.

The 'Classic' sampling design involves the selection of approx. 15–20 dominant trees per site and is the most common sampling used for dendrochronological studies (Schweingruber *et al.*, 1990). This approach likely arose from the desire to sample well-established, healthy trees least affected by competition and the objective to obtain the longest possible time

series based upon the (likely incorrect) assumption that the biggest trees are also the oldest within a forest population (Black *et al.*, 2008; Bigler & Veblen, 2009; Briffa & Melvin, 2011; Brienen *et al.*, 2012). To our knowledge, the majority of tree-ring datasets published e.g. on ITRDB were collected in this manner. For our study, we performed the classical sampling by selecting 15 individuals of ABAL out of the population ('Classic$_{field}$') based upon typical field criteria (e.g. dominant, straight growth, minimal signs of injury or wounding). Further samplings were simulated, by randomly selecting 15–20 dominant trees 100 times without replacement ('Classic$_{simulated}$').

The 'Forester' sampling design was applied as early as 1904 by manually sawing sections from logs at a lumberyard (Douglass, 1919) and is still used today (Tegel *et al.*, 2010). The sampling of archaeological, subfossil and historical construction wood (Tegel *et al.*, 2010) is arguably a variant of the 'Forester' sampling design, because this material in theory originates from ancient sawmills. Assuming a selective thinning should take place at our plot, the local foresters identified 32 'future' trees (25 ABAL, 2 PCAB and 5 QUPE; Fig. 1c), with a high economic value (i.e. large DBH, straight stem, no major scars, full developed crown, healthy, etc.) and presumably a strong root system. The simulated management is geared towards enhancing the increment over the next decades to further increase the economic value of these trees. Moreover, these trees are supposed to dominate the stand and provide shelter to smaller individuals. 'Future' trees were chosen in triangular relationships with a minimum distance between the vertices of 6–7 m for ABAL and PCAB, as well as 8–12 m for the more shade intolerant QUPE. In a second step, the 'present' trees to be harvested were selected. The greatest competitor of each 'future' tree with a merchantable size was typically chosen. In total, 33 'present' trees were selected (29 ABAL, 3 PCAB and 1 LADE, Fig. 1c). Trees within a 12 m perimeter outside the plots margins were considered as well to avoid boundary effects. Two versions of the 'Forester' sampling design, namely the 'Forester$_{present}$' and the 'Forester$_{future}$' emerged from this approach.

The 'Ad Hoc' design reflects sampling schemes developed for very specific research questions. For instance, Meyer & Bräker (2001) performed a study to test for differences in climate response of dominant vs. suppressed trees by sampling

only individuals with either social status. Analogously, one could stratify a species according to its size (i.e. DBH) or roughly to its age. The 'Ad Hoc' sampling design therefore describes the exclusive sampling of trees belonging to defined strata, based on social (here: either dominant, codominant, dominated or supressed), size (small, medium or large DBH trees) or age (young, intermediate or old cambial age at coring height) classes. For our study, we retroactively stratified our dataset according to objective criteria using all data available (see below).

The 'Pseudo-Population' design refers to the selection of representative population subsets. Therefore, the population is divided into certain strata (see 'Ad Hoc' above) and trees from every stratum are sampled. Unlike the 'Ad Hoc' design, the resulting chronology is ultimately composed of samples of all strata. Examples include trees selected according to social classes (Bascietto *et al.*, 2004), size classes (Nock *et al.*, 2011) or age classes (Carrer & Urbinati, 2004). While some studies sample equal or undefined numbers of individuals from each stratum (Carrer & Urbinati, 2004; Esper *et al.*, 2007b), others collect samples proportional to the population composition and demography (Martínez-Vilalta *et al.*, 2008). In our study, this type of sampling strategy is simulated by 100 times randomly picking equal and proportional numbers (only for unequal distributions of individuals among the different strata) of samples from three to four social, size, or age classes. Identical strata as for the 'Ad Hoc' sampling design were used.

The 'Fixed Plot' sampling design refers to the sampling of all trees within a given area which exceed a minimum threshold diameter (e.g. 5.6 cm). This approach has a long history in studies investigating forest dynamics (Cole & Lorimer, 1994; Singer & Lorimer, 1997; Weber *et al.*, 2008; Trotsiuk *et al.*, 2012) as well as for forest productivity estimates (Graumlich *et al.*, 1989; Liu *et al.*, 2012; Babst *et al.*, 2014b). In contrast to other sampling designs, the 'Fixed Plot' design is not restricted to particular species or tree characteristics. To perform the 'Fixed Plot' sampling design, various plots (100 times each) with radii ranging from 5 to 25 m were randomly 'placed' within the study area boundaries.

The 'Random' sampling design selects individuals using a random mechanism. Even though this approach is a standard sampling scheme in various fields of environmental research (Gilbert, 1987; Zar, 1999; Elzinga *et al.*, 2001; Thompson, 2012), relatively few dendrochronological studies (again mostly in investigations of forest dynamics) performed a comparable sampling to the 'Random' design (e.g. Lorimer & Frelich, 1989; Abrams & Hock, 2006; Voelker *et al.*, 2006; Pederson *et al.*, 2008; Svoboda *et al.*, 2012). To test this sampling approach, the same numbers of trees as for the 'Fixed Plot' design (average numbers of trees included for the different radii, see above) were 100 times randomly chosen from the population (without replacement) with all samples having equal chances of being selected.

Tree-ring data processing

To simulate climate reconstruction and climate response assessments emerging from the different sampling designs,

we focused on ABAL as the dominant species in our plot. Accordingly, structural features of ABAL are shown in more detail (Fig. 2b). Information on the social status was adopted from the LWF databank (Dobbertin *et al.*, 2011), whereas DBH and age thresholds for the stratification process were based upon terciles calculated from all data available for ABAL. While ABAL individuals are unequally distributed over the four social classes, they are almost equally distributed over all strata for tree size and age.

We applied multiple standardization and homogenization techniques which are commonly used in dendrochronological research. Non-climate-related trends in the TRW measurements were removed using both, a cubic smoothing spline with a 50% frequency-response threshold of 20 years (Cook & Peters, 1981) and a RCS approach. In RCS, the single TRW series are aligned by their cambial age (i.e. biological age from the tree pith) and a mean (smoothed) growth curve is calculated. This mean curve, the so-called 'Regional Curve' (RC), represents a single population estimate for the age trend and is subsequently used to remove this site-specific trend from each single TRW curve. Lastly, the standardized series are realigned to calendar years and a mean chronology is calculated (for more extensive descriptions of RCS see Briffa *et al.*, 1992; Cook *et al.*, 1995; Esper *et al.*, 2003; Briffa & Melvin, 2011; Briffa *et al.*, 2013). The individual steps of RCS applied to the population and dominant tree datasets are shown in Fig. S1. The spline detrending (preserving mostly interannual variability) and the RCS approach (which can preserve also low-frequency variability) are typically applied in studies aiming at quantifying the climate response of forests and at climate reconstruction, respectively. The 20-year spline and RCS chronologies employed a power transformation to mitigate index value inflation (Cook & Peters, 1997) and the variance of the chronologies was stabilized to minimize variance artefacts from changes in sample replication (Osborn *et al.*, 1997; Frank *et al.*, 2007). Pith-offset estimates were used in the RCS detrending (Esper *et al.*, 2003).

Long-term changes in forest growth, such as may be expected due to CO_2 fertilization (Huang *et al.*, 2007) or long-term climate change (Villalba *et al.*, 2012) are usually assessed using either TRW (Salzer *et al.*, 2009; Andreu-Hayles *et al.*, 2011), basal area increment (BAI; Martínez-Vilalta *et al.*, 2008; Körner *et al.*, 2005) or historical DBH reconstructions in combination with allometric biomass models (Davis *et al.*, 2009; Babst *et al.*, 2014a). For this study, we focused on BAI (Biondi, 1999) for detecting trends on the single species level (here: ABAL), as TRW includes nonclimatic, spurious signals (e.g. geometrical age trend; Fritts, 1976).

TRW and BAI are closely linked to tree biomass/carbon accumulation (Bouriaud *et al.*, 2005), but do not account for differences in species-specific characteristics such as wood density or tree shape (Babst *et al.*, 2014a). Furthermore, the majority of studies investigating growth trends usually focused on single, mostly dominant species in nonmonocultural forest stands. Thus, we also quantify past forest productivity for the entire population (i.e. annual above-ground biomass increments; AGB), by reconstructing historical DBH (according to Bakker, 2005) and subsequently applying

species-specific allometric biomass models (Fiedler, 1986; Hochbichler, 2002; Cienciala *et al.*, 2005; Montero *et al.*, 2005). The model outputs were subsequently used to iteratively calculate past AGB.

Individual series were averaged into chronologies using the Tukey's biweight robust mean. Growth trends were assessed by utilizing the slope of a linear regression. To quantify emerging biases of the tested sampling designs with respect to the population signal (calculated from all data available for either ABAL or the entire population), we applied three goodness-of-fit statistics, namely (i) Pearson's r (r); (ii) Root Mean Squared Error (RMSE); and (iii) average percentage bias (pBias). For all sampling designs where we performed 100 Monte Carlo type replicates, mean values are provided, followed by their standard deviation. Most calculations, analyses, and statistics were done using the R software environment including the R software package dplR (Bunn, 2008; R Core Team, 2011).

Climatic data and analyses

To investigate the climate response of tree growth, monthly temperature, and precipitation data were obtained from the nearest $0.5° \times 0.5°$ gridbox of the CRU TS 3.10 dataset (Harris *et al.*, 2013). In addition, the Palmer Drought Severity Index (PDSI) data from the nearest $2.5° \times 2.5°$ gridbox were considered (Dai *et al.*, 2004). Standardized TRW chronologies were correlated against monthly averaged climate data over an 18-month window from May of the year prior to ring formation until current October. Correlations were computed over the 1950–2004 period ensuring robust and well-replicated data for both, tree-ring, and instrumental records.

Results

Population dataset

The TRW chronologies of all species (Fig. 3) are likely influenced by a combination of management, stand dynamics, and regional environmental variability. QUPE, the oldest species within our plot (up to 222 years old), showed relatively high growth rates until the early 19th century with a subsequent phase of slower growth between 1850 and the early 1940s, followed by a strong increase in growth. This increase is also evident in ABAL, PCAB, and to a lesser extend in FASY. Furthermore, FASY generally show low growth rates with relatively little variation. ABAL display an overall increase in TRW ($0.018 \ \text{mm yr}^{-1}$, $P < 0.001$, w.r.t. 1912–2011) with superimposed decadal-scale oscillations. PCAB growth increased steadily until approximately 1990, when growth trends in all species started to decline ($P < 0.01$, for all species) towards the present. PCAB tend to be more than 30 years younger than ABAL,
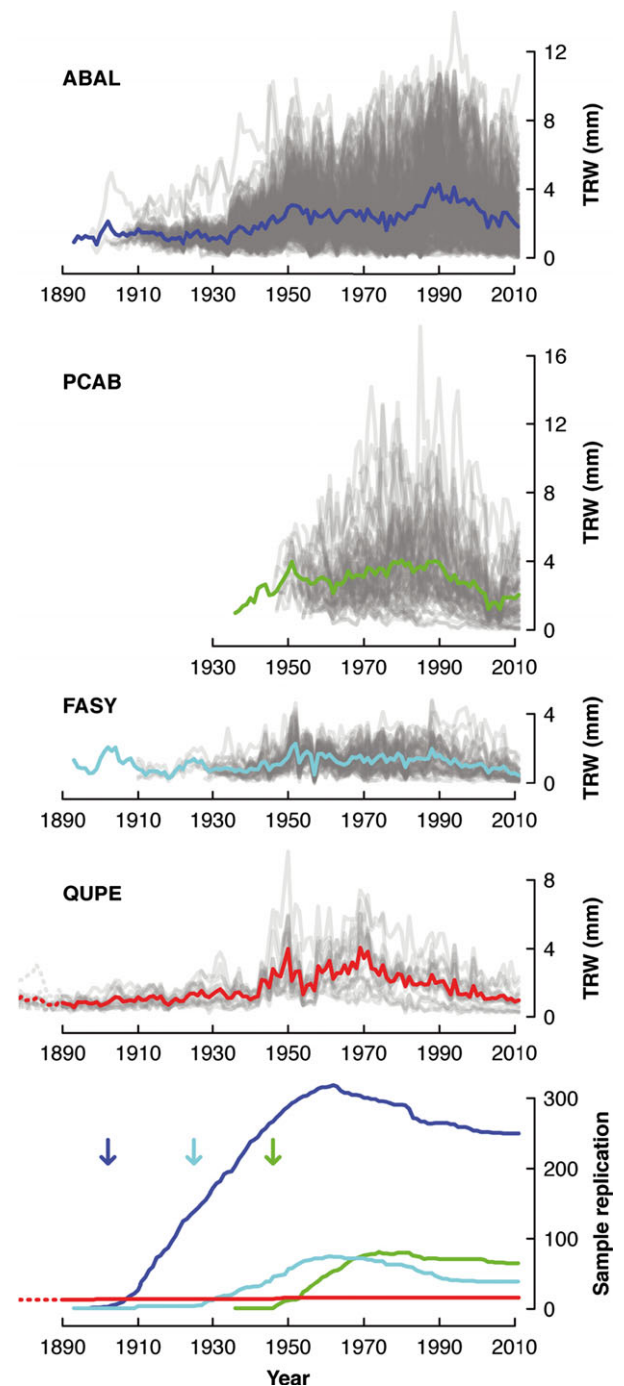
Fig. 3 Tree-ring width (TRW) chronologies from *Abies alba* (ABAL), *Picea abies* (PCAB), *Fagus sylvatica* (FASY), and *Quercus petraea* (QUPE, period of 1789–1889 truncated). Grey lines indicate the single TRW series. Data from the two *Larix decidua* are not shown due to insufficient sample size. The sample replications are shown in the lowest panel, with arrows indicating when the sample size drop below five.

but they also reach large DBH classes spanning from 6.8 to 84.0 cm, indicating that PCAB is the fastest growing species at our site.
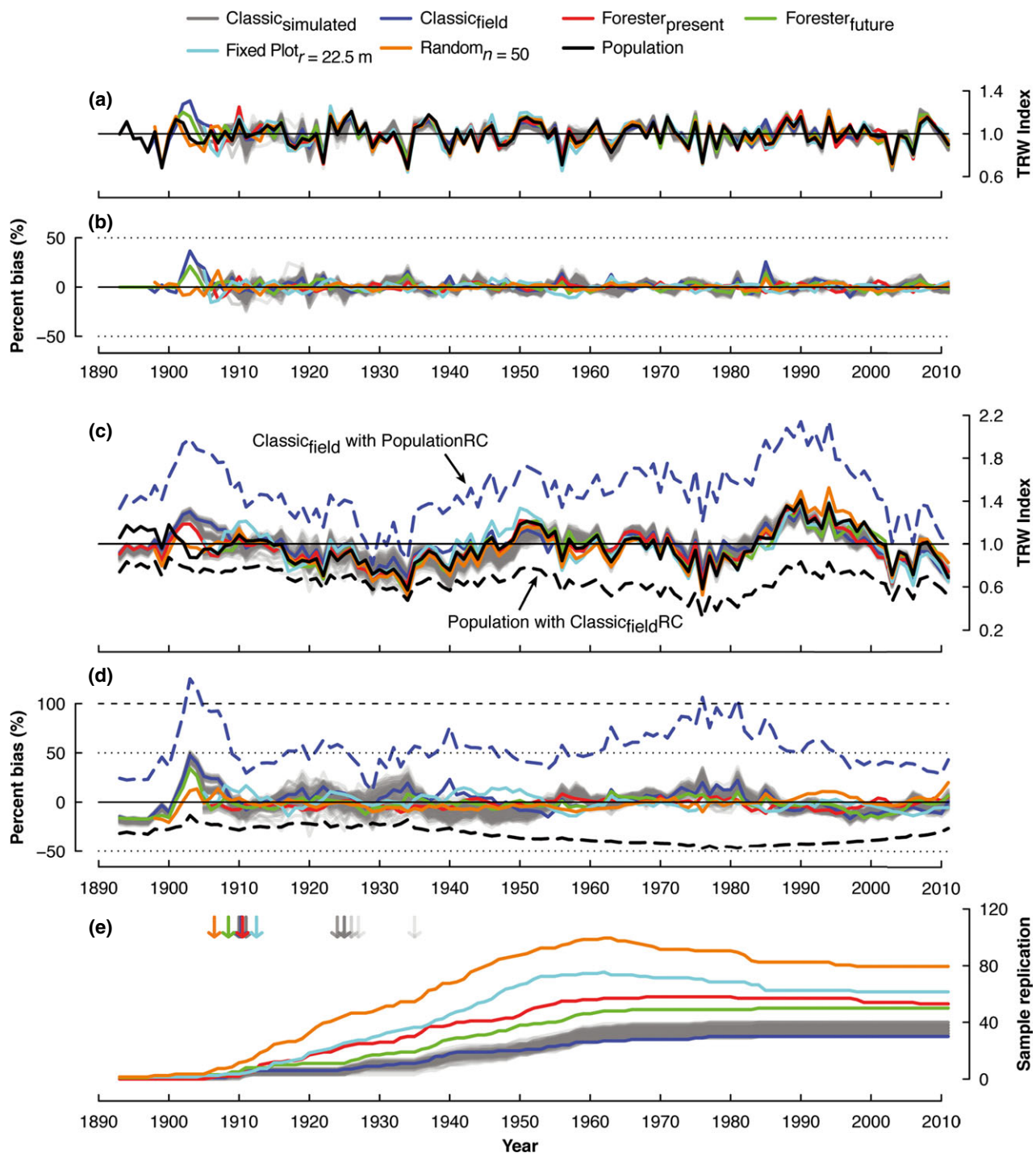
Fig. 4 Standardized tree-ring width chronologies for *Abies alba* using the 'Classic' and 'Forester' sampling designs as well as one representative example for the 'Fixed Plot' and 'Random' sampling designs detrended with (a) 20-year cubic smoothing spline and (c) a RCS approach. The corresponding average percentage deviation from the population signal (Percent bias) for each chronology is shown in (b) and (d), respectively. Dotted (dashed) lines indicate the 50% (100%) bias level. The sample replications are shown in (e), with arrows indicating when the sample replications drop below five. For the sample replication of the population chronology, see Fig. 3 (ABAL, blue line).

## Climate reconstruction/standardized growth anomalies

The standardized TRW chronologies (20-year cubic smoothing spline and RCS detrending) derived from the 'Classic', 'Forester', 'Fixed Plot', and 'Random' sampling designs are shown in Fig. 4 (see Fig. S1 for methodological details). The 20-year spline detrending, which retains interannual to roughly bidecadal growth fluctuations, showed very similar variations among the sampling designs and no significant biases were detected. In contrast, the RCS approach, which can better preserve low-frequency trends, showed greater multi-decadal variability in radial growth, but also low biases relative to the population dataset chronology (i.e. typically less than ±2% for commonly used sampling designs w.r.t. 1912–2011; Table S1). For more specific, nonstandard sampling approaches, like the 'Ad Hoc' designs (Table S1), a maximum centennial pBias of −9.3% was observed.

As RCS is less frequently applied to relatively young 'living only' tree-ring datasets, we performed an initial test to understand the degree to which sampling considerations may apply for RCS chronologies typically composed of living plus relict populations (perhaps universally sampled following inconsistent or unknown criteria). Assuming that the relict/subfossil part is a representative population subset and the living part more of a 'Classic' sampling design, we provide uncertainty bounds by detrending the 'Classic$_{field}$' samples using the population RC and *vice versa*. Detrending the population samples, with the 'Classic$_{field}$' RC resulted in an average underestimation of standardized tree growth of −33.7%, whereas detrending the 'Classic$_{field}$' samples with the population RC overestimates growth by 53.8% (w.r.t. 1912–2011; dashed lines Fig. 4c and d; Table S1).

Climate response

The standardized TRW chronologies derived from different sampling designs were correlated with monthly temperature, precipitation, and PDSI to assess potential

differences in the obtained climate response (Fig. 5). We find that the sampling design has a minor influence on the climate correlations at interannual to decadal time scales. Correlations computed for all different
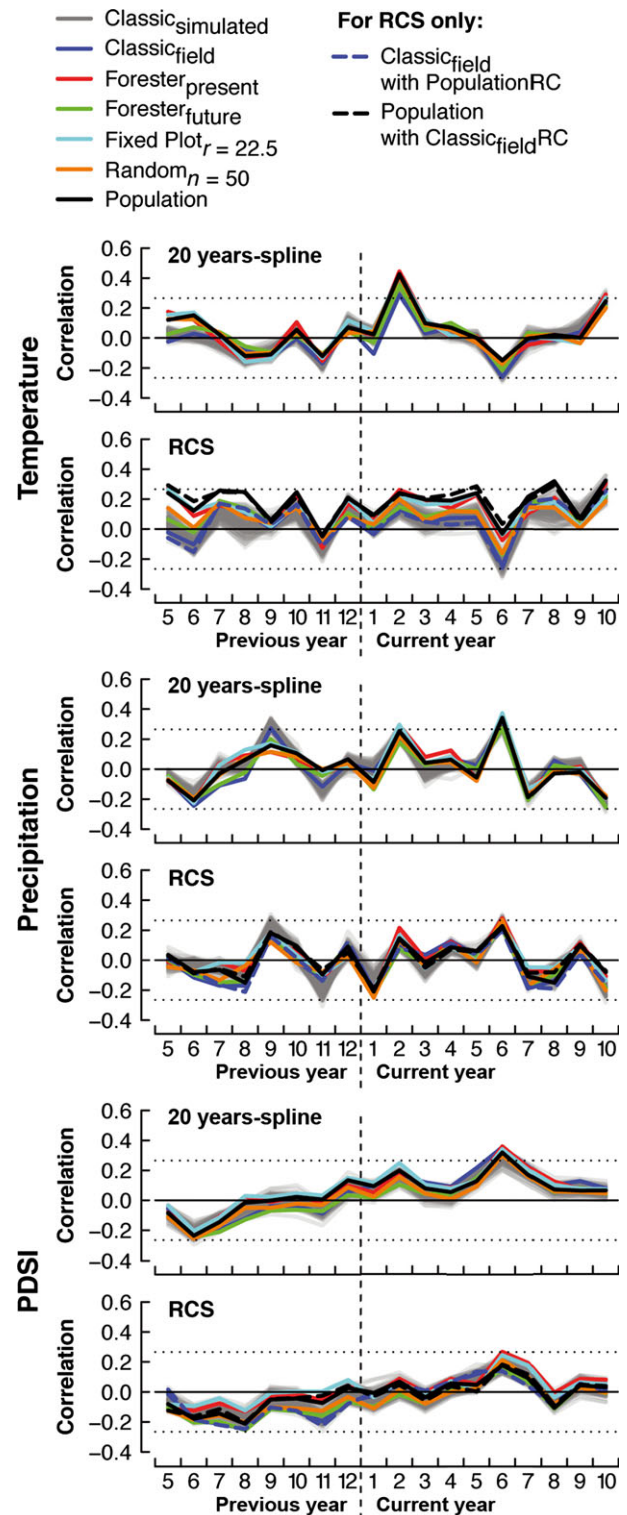

Fig. 5 Correlations between the standardized *Abies alba* tree-ring width chronologies and monthly averaged climate indices [previous May (5) to current October (10)] for temperature, precipitation and Palmer Drought Severity Index (PDSI) over the 1950–2004 period. Results are shown for chronologies obtained by applying the 'Classic' and 'Forester' sampling designs as well as one representative example for the 'Fixed Plot' and 'Random' sampling designs. For each climatic parameter, two different detrending approaches were used: Upper panels: flexible 20-year smoothing spline (20-year spline); Lower panels: Regional Curve Standardization (RCS). The 95% confidence intervals are indicated by a dotted line.

sampling designs shared the same common pattern and no significant biases were found. In terms of the climate correlations themselves, relatively few significant ($P < 0.05$) correlations for all three climate parameters were obtained – patterns typical for Central European forests growing under similar conditions (Bouriaud & Popa, 2008; Babst *et al.*, 2013). In the year of ring formation, ABAL responds positively to February temperature, June precipitation as well as PDSI, and negatively to June temperature. A positive response to September precipitation in the year prior to ring formation is the strongest lagged climatic driver observed.

Growth trends and forest productivity

The robust BAI chronologies derived from the various sampling designs revealed remarkable differences.

The 'Classic$_{field}$' sampling design overestimated the population BAI (ABAL only) on average by 196% (w.r.t. 1912–2011) with a decadal maximum in the 1980s (284% overestimation) peaking in 1985 with a 459% overestimation (Fig. 6). Furthermore, the observed trend in BAI of the population chronology over the 1940–1994 period (0.368 cm^2 yr^{-1}; $P < 0.001$) was overestimated by 346%. The magnitude of declining BAI trends (e.g. over 1994–2011) in the population chronology (-0.668 cm^2 yr^{-1}; $P < 0.001$) was similarly overestimated (visible as a steeper negative slope) by 276%. Comparable biases were obtained for the 'Classic$_{simulated}$' (309% and 137% average overestimation of positive and negative trends, respectively), the 'Forester$_{future}$' (268% and 150%), the 'Ad Hoc$_{size\_large}$' (240% and 169%) and the 'Ad Hoc$_{social\_dominant}$' (308% and 138%) sampling designs.
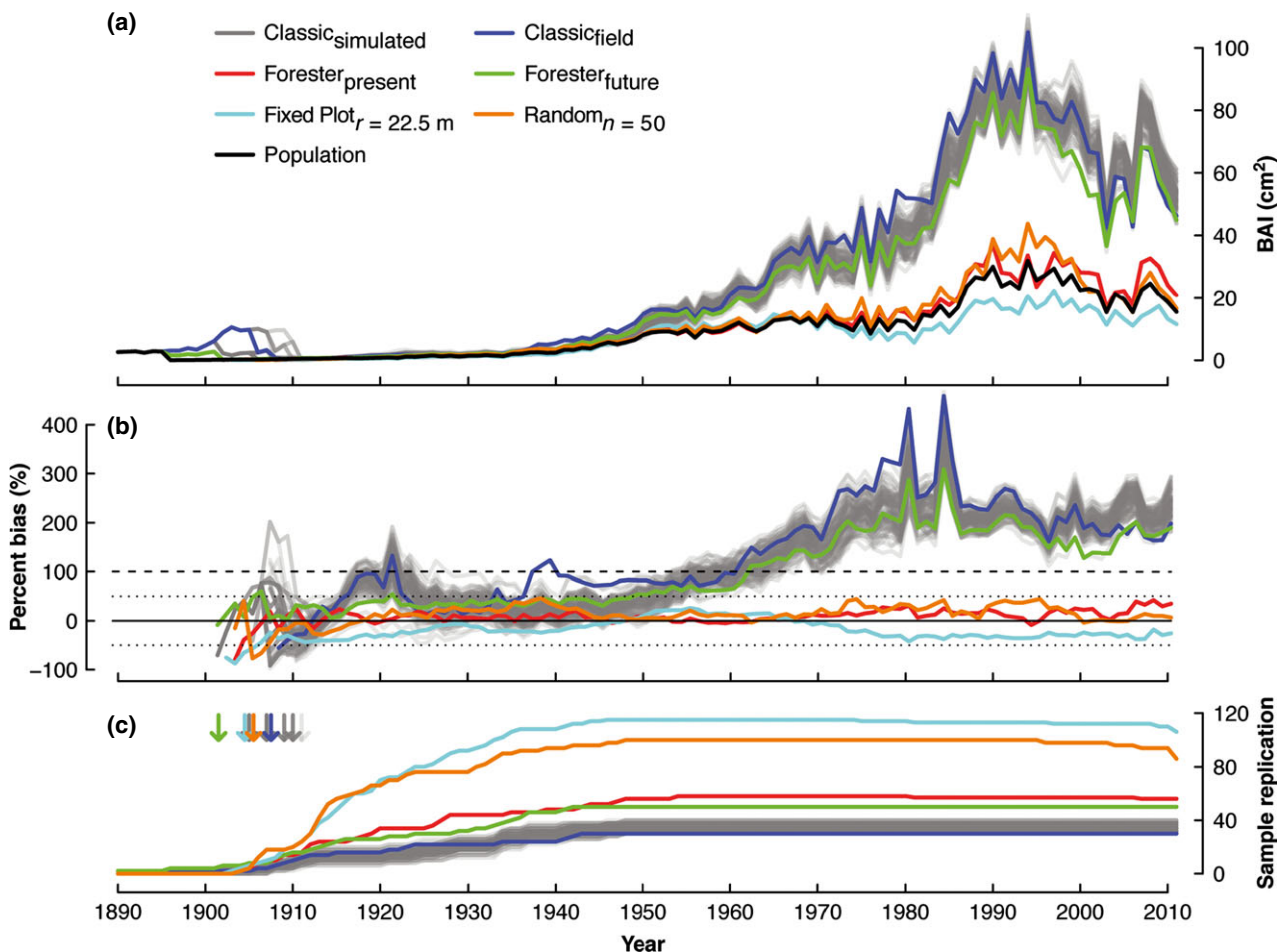
Fig. 6 (a) Robust mean basal area increment (BAI) chronologies for *Abies alba* using the 'Classic' and 'Forester' sampling designs as well as one representative example for the 'Fixed Plot' and 'Random' sampling designs. The corresponding average percentage deviation from the population signal (Percent bias) for each chronology is shown in (b). For early periods with low sample replication, biases were truncated once they exceeded 400%. Dotted (dashed) lines indicate the 50% (100%) bias level. The sample replications are shown in (c), with arrows indicating when the sample size drops below five.

Moreover, biases did not diminish with increasing sample replication (Figure S2), which is typically cited as a metric for chronology quality and robustness. Interestingly, the 'Forester$_{present}$' sampling design yielded growth variations which were close to the population BAI (13% overestimation w.r.t. 1912–2011), with a maximal overestimation of 19% in the most recent decade.

In contrast, the 'Pseudo-Population$_{social}$' sampling design with equal representation among the four strata led to an average bias of −30% for the BAI chronologies (w.r.t. 1912–2011; Table 2), whereas all other 'Pseudo-Population' designs with nearly equally sized strata and the 'Pseudo-Population$_{social}$' with proportional representation resulted in smaller biases. The 'Fixed plot' sampling design imparted slight

Table 2 Influence of sampling design on basal area increment (BAI) chronologies. Goodness-of-fit statistics relate to the 1912–2011 period. If not further declared, all correlations coefficients (r) and trends are significant at the 99.9% confidence level ($P < 0.001$). Statistics for the resampling types are given as mean $\pm$ SD

Sampling design	Subdivision	BAI n	r	RMSE (cm^2)	pBias (%)	pBiasTrend (%)
Classic	Simulated	30–40	0.97 ± 0.01	28.8 ± 1.7	181.2 ± 12.1	233.7 ± 13.8
	Field	30	0.96	30.5	196.4	234.4
Forester	Future	50	0.98	24.3	153.6	192.7
	Present	58	0.99	2.6	13.4	19.5
Ad Hoc$_{social}$	Dominant	54	0.98	28.9	182.0	235.9
	Co-Dominant	172	0.99	3.6	24.1	25.1
	Dominated	92	0.17	11.9	−74.9	−95.4*
	Suppressed	22	−0.18	13.1	−88.1	−†
Ad Hoc$_{size}$	Large	124	0.98	22.2	142.2	173.3
	Medium	122	0.99	3.4	−19.1	−27.3
	Small	106	0.09	12.5	−81.6	−†
Ad Hoc$_{age}$	Old	122	0.98	1.8	1.8	−14.0
	Intermediate‡	108	0.99	1.2	−4.9	−5.3
	Young§	118	0.96	3.7	5.2	42.5
Pseudo−Population	Social	88	0.90 ± 0.07	5.5 ± 1.7	−30.2 ± 11.9	−39.8 ± 16.4
	Social (prop.)**	88	0.98 ± 0.01	2.5 ± 1.2	−1.5 ± 14.1	−2.9 ± 18.2
	Size	90	0.98 ± 0.01	2.4 ± 0.9	−10.1 ± 7.6	−13.2 ± 11.7
	Age	90	0.98 ± 0.02	2.9 ± 1.6	0.0 ± 17.9	−0.4 ± 24.5
Fixed Plot	5.0 m	6 ± 2	0.69 ± 0.30	11.2 ± 6.8	17.7 ± 68.2	26.9 ± 91.4*
	7.5 m	14 ± 6	0.79 ± 0.22	8.9 ± 7.6	6.8 ± 67.2	4.9 ± 73.7*
	10.0 m	24 ± 10	0.87 ± 0.15	7.1 ± 5.6	6.0 ± 51.2	1.7 ± 59.5*
	12.5 m	36 ± 12	0.91 ± 0.11	5.1 ± 2.6	−3.1 ± 30.1	−9.7 ± 40.5
	15.0 m	50 ± 16	0.93 ± 0.05	4.2 ± 1.5	−11.1 ± 18.6	−19.6 ± 25.9
	17.5 m	70 ± 18	0.95 ± 0.03	3.4 ± 1.3	−8.5 ± 14.7	−17.3 ± 16.6
	20.0 m	90 ± 20	0.96 ± 0.02	3.1 ± 1.1	−10.5 ± 10.4	−20.3 ± 13.0
	22.5 m	114 ± 22	0.97 ± 0.02	3.0 ± 1.0	−12.2 ± 7.8	−22.3 ± 9.1
	25.0 m	138 ± 16	0.97 ± 0.01	3.0 ± 0.9	−13.2 ± 7.5	−23.6 ± 8.2
Random		6	0.70 ± 0.33	10.9 ± 6.5	8.6 ± 71.6	7.4 ± 90.6*
		14	0.79 ± 0.24	8.3 ± 3.5	−2.0 ± 49.9	−3.7 ± 66.0*
		24	0.89 ± 0.15	6.7 ± 3.3	4.6 ± 41.7	4.8 ± 53.7*
		36	0.93 ± 0.11	4.9 ± 2.7	−2.2 ± 31.4	−3.9 ± 40.0
		50	0.96 ± 0.06	3.9 ± 1.9	−2.9 ± 24.1	−3.4 ± 31.7
		70	0.98 ± 0.01	3.4 ± 1.7	0.9 ± 20.9	1.1 ± 27.4
		90	0.98 ± 0.02	2.7 ± 1.5	0.0 ± 16.1	−0.2 ± 22.1
		114	0.99 ± 0.01	2.4 ± 1.1	−1.2 ± 13.9	−2.2 ± 18.7
		138	0.99 ± 0.01	1.9 ± 0.9	1.2 ± 11.0	1.6 ± 15.4

*$P < 0.05$.
†No significant trend.
‡w.r.t. 1913–2011.
§w.r.t. 1921–2011.
**Samples selected proportional to their abundance within the population.

systematic biases towards an underestimation of the population signal (e.g. −13% for a radius of 25 m w.r.t. 1912–2011). An average plot radius of 22.5 m was needed to hold the standard deviations for pBias and pBiasTrend calculated from all 100 chronologies below 10% (Table 2). The 'Random' design performed very well with typically less than 3% in pBias (w.r.t. 1912–2011) for chronologies consisting of more than 12 trees (i.e. $n \geq 24$; Table 2).

A BAI chronology consisting only of living individuals of ABAL resulted in a centennial overestimation of 12% in growth rates with a decadal maximum for the 1980s (19%) compared to our population chronology, which includes also dead individuals (in 2011). Similar biases as presented above for BAI were also found for TRW (Table S2), with evidence for even larger errors depending on the applied methodologies (Table S3).

The above results suggest that, among all tested sampling designs, only the 'Fixed Plot' and the 'Random' sampling designs are capable of reconstructing stand level forest productivity when information on forest density and structure is provided. Given application of these designs, the key issue for forest productivity estimations is perhaps less a question of sampling design, but rather linked with sufficient sample replication. Accordingly, we quantify the spread around the population chronology (all species) resulting from circular plots of different radii ('Fixed Plot') and number of randomly selected trees ('Random'). Results for fixed plots with 5, 10, ..., 25 m radii are shown in Fig. 7 with a summary graph comparing the biomass estimated for 2011 using both, the 'Fixed Plot' and 'Random' approaches (Figure S3). Both sampling designs show the same patterns and no significant biases were detected. Increasing numbers of samples lead to lower statistical dispersion around the population AGB. An average plot radius of 22.5 m was needed to hold the median average deviation (MAD) in 2011 below 0.1 kg m^{-2} (cf. Fig. 7c).

Discussion

Implications for global change research

The retroactive implementation of six common tree-ring sampling schemes suggested that the vast majority of tree-ring datasets produced to date would yield biased results if used to assess the magnitude and variability of long-term changes in tree growth. This finding appears crucial given the increasingly important role of tree-ring data in many aspects of global change research including terrestrial biosphere modelling (Keenan *et al.*, 2012), understanding of carbon allocation within forest ecosystems (Brüggemann *et al.*,

2011), or revealing the impacts of climate extremes (Reichstein *et al.*, 2013). The extent to which sampling design may bias climate reconstructions (Frank *et al.*, 2010; Briffa & Melvin, 2011) and/or be overcome or amplified by analytical methods (Bontemps & Esper, 2011; Briffa *et al.*, 2013; Cecile *et al.*, 2013) remains less clear.

Our findings challenge an important dendrochronological paradigm, namely that a subset of samples is typically considered to be representative for the entire population. In particular, studies using BAI (and TRW) to quantify growth trends are strongly affected by the choice of sampling design. While many previous researchers (e.g. Schweingruber *et al.*, 1990; Briffa *et al.*, 1998; Frank & Esper, 2005a,b) may have adopted or even advocated the 'Classic' sampling design as being sufficient for climate response analysis and climate reconstructions (with limitations; Table 3; Briffa & Melvin, 2011), this study has shown that the suitability of this sampling design for assessing long-term changes in growth is highly questionable.

Recent literature has presented controversial evidence for both increasing (e.g. Cole *et al.*, 2010; Koutavas, 2013) or decreasing growth trends (e.g. Nock *et al.*, 2011; Linares & Camarero, 2012) e.g. due to CO_2 fertilization effects or regionally increasing drought severity. Our study suggests, also in line with previous evidence by Brienen *et al.* (2012), that biases introduced by the choice of sampling protocols might be sufficiently large to contribute to uncertainties among these studies and possibly even drive erroneous conclusions. We found a maximum bias of 459% overestimation for a single year in BAI chronologies (Fig. 6), with overall biases for common, but likely inappropriate sampling designs, more typically in the order of 150–200% (cf. Table 2).

The source of the sampling biases that we observe appears to be driven by multiple intertwined ecological effects combined with the ways that sampling is performed. Firstly, and most trivially, within a given forest stand, there is a distribution of tree characteristics that can affect individual tree growth such as tree size, tree age, climate, management, genetics, canopy ascension or competition. Simply by sampling larger trees, it is evident that the population mean will be overestimated. Secondly, the growth rate of trees can be size dependent with trees often growing disproportionally rapidly with increasing size, consistent with metabolic scaling theory (Stephenson *et al.*, 2014). Thus, disproportionally increasing growth rates over time can be expected by sampling larger trees. Thirdly, even if we control for tree size, we still find that the growth rate of the sampled trees was biased with respect to the population mean. This latter bias can be directly quantified
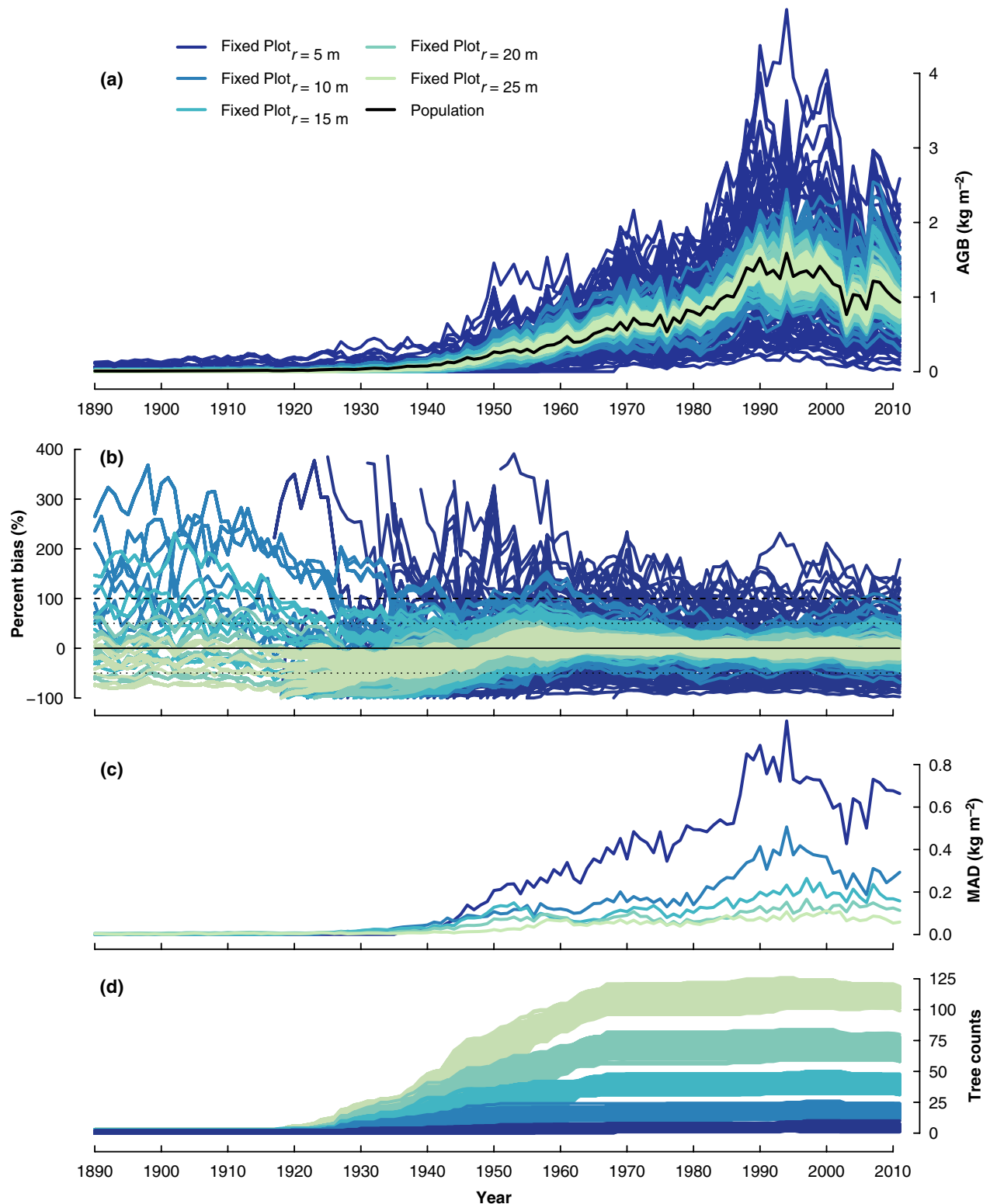
Fig. 7 (a) Above-ground biomass increment (AGB) for the 'Fixed Plot' sampling design using 5, 10, ..., 25 m radii (blue to green spectrum). Each circular plot size was randomly resampled 100 times. The corresponding average percentage deviation from the population signal (Percent bias) for each chronology is shown in (b). For early periods with low sample replication, biases were truncated once they exceeded 400%. Dotted (dashed) lines indicate the 50% (100%) bias level. As a measure of spread, the median absolute deviation (MAD) for the 100 different AGB estimations are given in (c) and the sample replication is shown in (d).

Table 3 Comparison of sampling designs for four of the main types of dendrochronological investigations. Symbols indicate our suitability assessment for a sampling design in terms of minimizing biases related to study design (+, suitable; ○, suitable within certain limitations; and −, unsuitable). See text for further explanations

Sampling design	Climate reconstruction	Climate response	Growth trends	Forest productivity
Classic	○	+	−	−
Forester	○	+	○	−
Ad Hoc	○	+	−	−
Pseudo-Population	+	+	○	−
Fixed Plot	+	+	+	+
Random	+	+	+	+

(cf. 'Big-Tree Selection Bias'; Brienen *et al.*, 2012) in our dataset and is responsible for between 7% and 109% overestimation of BAI depending upon DBH (see Figure S4).

Variability in detrended TRW series tends to be rather insensitive to the sampling design on a centennial timescale, especially if only preservation of high-frequency variation is sought. Accordingly, no notable biases were detected in our climate response analysis, which supports earlier findings from dominant vs. suppressed (Meyer & Bräker, 2001) and young vs. old trees (Esper *et al.*, 2008). The choice of detrending remains very important as the removal of age-related trends in TRW is among the more subjective and uncertain elements in tree-ring research (Frank *et al.*, 2010). This is illustrated by the differences in bidecadal vs. long-term variability preserved in the 20-year spline and RCS detrended chronologies (Fig. 4a vs. Fig. 4c). The similarities in all RCS chronologies can be traced to a compensation of biases between the data subsets and their respective RC (Figure S1). Yet, there is no guarantee that the RCS chronologies are a faithful estimate of long-term changes in tree growth (Briffa & Melvin, 2011), with similar considerations also applying to BAI and AGB estimates (see below). Initial tests of RCS-standardization variants suggested large biases in chronologies composed of trees from different generations (e.g. living and relict/sub-fossil trees) if the sampling is not consistent over the length of the chronology (Briffa & Melvin, 2011; Esper *et al.*, 2012). Unfortunately, this condition is nearly impossible to fulfil and compatible sampling of the population(s) is difficult to demonstrate. Further investigations are needed to understand and quantify possible influences of changing sampling designs in long tree-ring chronologies.

Our Monte Carlo type resampling approach (Fig. 7, Figure S2) quantified how increasing tree counts can improve chronology confidence (typically assessed via the Expressed Population Signal; EPS; Wigley *et al.*, 1984; Mérian *et al.*, 2013), but did not mitigate systematic biases in the chronology mean. It is thus insufficient to core large numbers of trees to obtain growth signals which are representative for the entire population. Nor is it justifiable to appeal to a high EPS value as a measure of data reliability while contemporaneously disregarding the underlying sampling design.

Performance of sampling designs

We found no advantages of the 'Classic' sampling design apart from easier dating of dominant trees and reduced field efforts. This approach seems sufficient, yet not recommended (see below), for studies addressing only growth and/or short-term climate variability (Table 3).

Both 'Forester' sampling designs showed exceptionally large differences in tree growth rates. Tree selection in this sampling scheme is subjective and likely changes with management practices. We suspect that this design generally leads towards an overestimation of contemporary growth rates, as it favours large merchantable individuals of high economic value, while smaller individuals often contribute to local coarse woody debris (E. Steiner and D. Gautschi, local foresters, 2012, personal communication).

As already suggested earlier (e.g. Cherubini *et al.*, 1998), sample collection from all size classes ('Pseudo-Population$_{size}$') is favourable over the 'Classic' approach. Yet, our findings illustrated that the 'Pseudo-Population' design could yield systematic biases due to nonsystematic effects and an inhomogeneous distribution of individuals over the different strata. Thus, different strata should be sampled proportionally to their abundance (cf. Table 2; Martínez-Vilalta *et al.*, 2008) which requires biometric data collected for the entire population. The 'Pseudo-Population' sampling design, nevertheless remains susceptible to biases originating from stratification methodologies.

Babst *et al.* (2013) noted that many tree-ring investigations do not regularly collect auxiliary (i.e. tree dimensions) or site-specific (e.g. topographic and

demographic) data, which hampers their use for biomass assessments. Our study suggests, however, that even provided the necessary biometric and metadata, most of these datasets would likely yield substantially biased estimates of long-term growth changes. Arguably, the 'Fixed Plot' approach defines and thus captures a site and its population and appears to be the most broadly applicable sampling design. An assessment of the 'Fixed Plot' scheme resulted in minor biases towards an underestimation as evident from our BAI chronologies (Table 2). This bias was attributable to geometric limitations in our implementation of Monte Carlo replications. As 'Fixed Plot' radii increase, the corners of our rectangular population become neglected and chronologies are increasingly weighted towards trees in the population centre. If all species were included, biases decreased along with the patchy species distributions over the plot (Fig. 1c). At our site, a radius of 22.5 m (including an average of 57 trees) seems to be a good trade-off between signal strength and sampling efforts (Fig. 7). Thus, the sampling by Babst et al. (2014a,b), who suggested to include a minimum of 32–60 trees in managed forests across Europe, appears reasonable. The applicability of these guidelines, however, relies on the vast majority of trees within a fixed plot being datable. This may not always be the case, especially in natural stands with a complex demography or disturbance history. Hence, the requirements in terms of plot size and sample replication to obtain reliable AGB estimates will need further testing in other forest ecosystems.

The 'Random' sampling scheme was not subject to geometric restrictions in our assessments and yielded the best population fit for growth trend analysis and climate reconstruction. Furthermore, the 'Random' design enables the acquisition of a broad spatial population signal depending upon the study area and fraction of the total population that are sampled. Implementing this design over a wide area likely reduces biases imparted by the choice of a plot centre ('Fixed Plot'), but decreases possibilities to investigate site dynamics and growth disturbances. A truly random sampling design may be challenging to implement under field conditions. Analogous to the 'Fixed Plot' design, sampling approx. 50 trees seems to be a good trade-off for managed forests in Central Europe.

Both, the 'Random' and 'Fixed Plot' sampling designs included trees with diameters equal to or greater than 5.6 cm. Such a low diameter threshold helps to accurately represent the tree stand and mitigate biases (e.g. the 'Big Tree-Selection Bias'). At the same time, a greater number of trees will likely be required to develop chronologies of a given length.

Limitations and the way forward

The applicability of our results of this case study to other forest types and geographic regions remains to be quantified. In addition, the here found biases refer to a distinct area of approx. 0.5 ha which was assumed to be the 'entire' population. We thus expect that the 'true' uncertainties may in fact be larger, particularly as subsampling approaches are limited by our population size.

It needs to be emphasized that while the historical trends in tree growth may not be biased with respect to the population dataset, the population data set itself has limitations for addressing the forest growth on longest time scales. Care was taken to sample living and dead material, yet trees have been removed both from management and natural processes over the past century. The study area was shaped by several generations of foresters applying differing management practices with elusive impacts on tree growth (E. Steiner and D. Gautschi, local foresters, 2012, personal communication). Many individuals which originally belonged to this population were likely removed for economic exploitation during past thinnings (Badeau et al., 1996), smaller individuals died (and decayed) due to self-thinning or fungal infections (Cherubini et al., 2002) or parts of the population may have died in nonrandom ways (Black et al., 2008; Bigler & Veblen, 2009). Thus, we encounter a 'fading record' phenomenon (Swetnam et al., 1999), whereby the reliability of the tree-ring data in estimating tree growth becomes lower further back in time and thus must underestimate the true standing wood volume prior to a certain year. Furthermore, the natural mortality processes or management interventions are not uniform across tree-size and tree-age classes. For example, slow growing trees may die with increased frequency due to suppression towards their juvenile phases (Wyckoff & Clark, 2002), but those that do not succumb to competition may become the oldest trees in a stand (Black et al., 2008; Bigler & Veblen, 2009). Merging tree-ring records with inventory assessments, maximizing the data collected from stumps and logs, and possibly also model-data fusion approaches should be useful to provide increasingly reliable longer term annually resolved estimates of absolute forest growth.

Uncertainties in the quantification of tree biomass increment (Fig. 7), even in the most recent years, certainly originate from allometric models that were built for different stands/environments. However, we suspect that these uncertainties likely do not alter the relative biases introduced by the different sampling designs. After sampling schemes have been standardized to

yield nonbiased quantification of historical DBH with respect to the population chronology (Fig. 7), we emphasize that further development of allometric equations will be needed to improve estimates of variability in terrestrial carbon allocation (Nickless *et al.*, 2011; Goodman *et al.*, 2012).

The generality of our conclusions should be tested for different forest characteristics including (i) unmanaged and naturally grown forests; (ii) investigations at a species' distribution limits; (iii) the dominance of one growth limiting climate factor (sites classically used for climate reconstructions); (iv) the presence of subfossil or relict wood, or combinations thereof. The influence of sampling design on other tree-ring parameters (e.g. maximum latewood density, isotope ratios or wood anatomical features) must also be explored.

In drawing specific attention to significant biases (>200% in commonly employed BAI chronologies), our study encourages sampling designs beyond the 'Classic' design and provides impetus for further work aiming at reducing uncertainties in tree-ring and other proxy-based investigations. Thereby, we do not question the fundamental value of existing tree-ring datasets, but rather call for caution regarding analyses and conclusions that may be influenced by the underlying sampling designs.

The widespread adoption of the 'Fixed Plot' and 'Random' sampling designs as standard strategies would be beneficial, especially for national forest inventories (e.g. CNFI, 2008; US Forest Service, 2012), which often tend to select trees for sampling in an arguably very similar manner to the 'Classic' sampling design (Table 1). Existing inventories or monitoring projects (Ferretti, 2013) often define fixed plots and assign unique numbers to each individual, thereby providing an excellent basis for either the 'Fixed Plot' or 'Random' sampling design. The common implementation of these widely applicable sampling designs is expected to greatly reduce uncertainties and biases in tree-ring-based quantification of forest growth and to simultaneously increase the comparability of tree-ring datasets beyond individual studies, investigators, laboratories, and geographical boundaries. As we find substantial impacts in this initial study, we also hope to contribute to interdisciplinary discussions about persistent biases in research outcomes (Ioannidis, 2005; Sarewitz, 2012) and help to catalyse solutions for improving estimates of global change phenomena.

Acknowledgements

References

Abrams MD, Hock WK (2006) Annual growth rings and the impact of Benlate 50 DF fungicide on citrus trees in seasonally dry tropical plantations of northern Costa Rica. *Forest Ecology and Mangement*, **227**, 96–101.

Andreu-Hayles L, Planells O, Gutiérrez E, Muntan E, Helle G, Anchukaitis KJ, Schleser GH (2011) Long tree-ring chronologies reveal 20th century increases in water-use efficiency but no enhancement of tree growth at five Iberian pine forests. *Global Change Biology*, **17**, 2095–2112.

Babst F, Poulter B, Trouet V *et al.* (2013) Site- and species-specific responses of forest growth to climate across the European continent. *Global Ecology and Biogeography*, **22**, 706–717.

Babst F, Bouriaud O, Alexander R, Trouet V, Frank D (2014a) Towards consistent measurements of carbon accumulation: a multi-site assessment of biomass and basal area increment across Europe. *Dendrochronologia*, **201**, 1289–1303.

Babst F, Bouriaud O, Papale D *et al.* (2014b) Above-ground woody carbon sequestration measured from tree rings is coherent with net ecosystem productivity at five eddy-covariance sites. *New Phytologist*, **201**, 1289–1303.

Badeau V, Becker M, Bert D, Dupouey JL, Lebourgeois F, Picard JF (1996) Long-term growth trends of trees: ten years of dendrochronological studies in France. In: *Growth Trends in European Forests* (eds Spiecker H, Mielikäinen K, Köhl M, Skovsgaard JP), pp. 167–181. Springer, Berlin.

Bakker JD (2005) A new, proportional method for reconstructing historical tree diameters. *Canadian Journal of Forest Research*, **35**, 2515–2520.

Bascietto M, Cherubini P, Scarascia-Mugnozza G (2004) Tree rings from a European beech forest chronosequence are useful for detecting growth trends and carbon sequestration. *Canadian Journal of Forest Research*, **34**, 481–492.

Beck PSA, Juday GP, Alix C *et al.* (2011) Changes in forest productivity across Alaska consistent with biome shift. *Ecology letters*, **14**, 373–379.

Bigler C, Veblen TT (2009) Increased early growth rates decrease longevities of conifers in subalpine forests. *Oikos*, **118**, 1130–1138.

Biondi F (1999) Comparing tree-ring chronologies and repeated timber inventories as forest monitoring tools. *Ecological Applications*, **9**, 216–227.

Black BA, Colbert JJ, Pederson N (2008) Relationships between radial growth rates and lifespan within North American tree species. *Ecoscience*, **15**, 349–357.

Bontemps JD, Esper J (2011) Statistical modelling and RCS detrending methods provide similar estimates of long-term trend in radial growth of common beech in north-eastern France. *Dendrochronologia*, **29**, 99–107.

Bouriaud O, Popa I (2008) Comparative dendroclimatic study of Scots pine, Norway spruce, and silver fir in the Vrancea Range, Eastern Carpathian Mountains. *Trees*, **23**, 95–106.

Bouriaud O, Bréda N, Dupouey JL, Granier A (2005) Is ring width a reliable proxy for stem-biomass increment? A case study in European beech. *Canadian Journal of Forest Research*, **35**, 2920–2933.

Bowman DMJS, Brienen RJW, Gloor E, Phillips OL, Prior LD (2013) Detecting trends in tree growth: not so simple. *Trends in Plant Science*, **18**, 11–17.

Brienen RJW, Zuidema PA (2006) Lifetime growth patterns and ages of Bolivian rain forest trees obtained by tree ring analysis. *Journal of Ecology*, **94**, 481–493.

Brienen RJW, Gloor E, Zuidema PA (2012) Detecting evidence for CO2 fertilization from tree ring studies: the potential role of sampling biases. *Global Biogeochemical Cycles*, **26**, GB1025.

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: *Dendroclimatology* (eds Hughes MK, Swetnam TW, Diaz HF), pp. 113–145. Springer, Dordrecht.

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.

Briffa KR, Schweingruber FH, Jones PD, Osborn TJ, Shiyatov SG, Vaganov EA (1998) Reduced sensitivity of recent tree-growth to temperature at high northern latitudes. *Nature*, **391**, 678–682.

Briffa KR, Melvin TM, Osborn TJ *et al.* (2013) Reassessing the evidence for tree-growth and inferred temperature change during the Common Era in Yamalia, northwest Siberia. *Quaternary Science Reviews*, **72**, 83–107.

Brüggemann N, Gessler A, Kayler Z *et al.* (2011) Carbon allocation and carbon isotope fluxes in the plant-soil-atmosphere continuum: a review. *Biogeosciences*, **8**, 3457–3489.

Bunn AG (2008) A dendrochronology program library in R (dplR). *Dendrochronologia*, **26**, 115–124.

Büntgen U, Tegel W, Nicolussi K *et al.* (2011) 2500 years of European climate variability and human susceptibility. *Science*, **331**, 578–582.

Bürgi M, Stuber M (2003) Agricultural use of forests in Switzerland 1800–1950. Field crops and forestry in alternation, forest fruits and resin. *Schweizerische Zeitschrift für Forstwesen*, **154**, 360–375.

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

Cecile J, Pagnutti C, Anand M (2013) A likelihood perspective on tree-ring standardization: eliminating modern sample bias. *Climate of the Past Discussions*, **9**, 4499–4551.

Cherubini P, Dobbertin M, Innes JL (1998) Potential sampling bias in long-term forest growth trends reconstructed from tree rings: a case study from the Italian Alps. *Forest Ecology and Management*, **109**, 103–118.

Cherubini P, Fontana G, Rigling D, Dobbertin M, Brang P, Innes JL (2002) Tree-life history prior to death: two fungal root pathogens affect tree-ring growth differently. *Journal of Ecology*, **90**, 839–850.

Cienciala E, Černý M, Apltauer J, Exnerová Z (2005) Biomass functions applicable to European beech. *Journal of Forest Science*, **51**, 147–154.

CNFI (2008) Canada's national forest inventory ground sampling guidelines. Version 5.0. Canadian Forest Service, Ottawa. Available at: https://nfi.nfis.org/documentation/ground_plot/Gp_guidelines_v5.0.pdf (accessed 11 November 2013).

Cole WG, Lorimer CG (1994) Predicting tree growth from crown variables in managed northern hardwood stands. *Forest Ecology and Management*, **67**, 159–175.

Cole CT, Anderson JE, Lindroth RL, Waller DM (2010) Rising concentrations of atmospheric CO_2 have increased growth in natural stands of quaking aspen (Populus tremuloides). *Global Change Biology*, **16**, 2186–2197.

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

Cook ER, Peters K (1997) Calculating unbiased tree-ring indices for the study of climatic and environmental change. *The 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 paleoclimatic studies. *The Holocene*, **5**, 229–237.

Cook ER, Anchukaitis KJ, Buckley BM, D'Arrigo RD, Jacoby GC, Wright WE (2010) Asian monsoon failure and megadrought during the last millennium. *Science*, **328**, 486–489.

Dai A, Trenberth KE, Qian T (2004) A global dataset of Palmer Drought Severity Index for 1870–2002: relationship with soil moisture and effects of surface warming. *Journal of Hydrometeorology*, **5**, 1117–1130.

Davis SC, Hessl AE, Scott CJ, Adams MB, Thomas RB (2009) Forest carbon sequestration changes in response to timber harvest. *Forest Ecology and Management*, **258**, 2101–2109.

DeRose RJ, Wang SY, Shaw JD (2013) Feasibility of high-density climate reconstruction based on forest inventory and analysis (FIA) collected tree-ring data. *Journal of Hydrometeorology*, **14**, 375–381.

Dobbertin M, Hug C, Schwyzer A, Borer S (2011) Aufnahmeanleitung für die Sanasilva-Inventur und Kronenansprachen auf den LWF - Flächen V8/11. Swiss Federal Research Institute WSL. Available at: http://www.wsl.ch/dienstleistungen/inventare/sanasilva/LWF_WSI_manual_8-11_web.pdf (accessed 11 November 2013).

Douglass AE (1919) *Climatic Cycles and Tree-Growth*. Carnegie Institution of Washington, Washington, DC.

Elzinga CL, Salzer DW, Willoughby JW, Gibbs JP (2001) *Monitoring Plant and Animal Populations*. Blackwell Science, Malden.

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 Research*, **59**, 81–98.

Esper J, Büntgen U, Frank DC, Nievergelt D, Liebhold A (2007a) 1200 years of regular outbreaks in alpine insects. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 671–679.

Esper J, Frank D, Büntgen U, Verstege A, Luterbacher J, Xoplaki E (2007b) Long-term drought severity variations in Morocco. *Geophysical Research Letters*, **34**, L17702.

Esper J, Niederer R, Bebi P, Frank D (2008) Climate signal age effects – Evidence from young and old trees in the Swiss Engadin. *Forest Ecology and Management*, **255**, 3783–3789.

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

Etzold S, Waldner P, Thimonier A, Schmitt M, Dobbertin M (2014) Tree growth in Swiss forests between 1995 and 2010 in relation to climate and stand conditions: recent disturbances matter. *Forest Ecology and Management*, **311**, 41–55.

Ferretti M (2013) Concepts and design principles adopted in the International Cooperative Program on the Assessment and Monitoring of Air Pollution Effects on Forests (ICP Forests). In: *Forest Monitoring Methods for Terrestrial Investigations in Europe with an Overview of North America and Asia* (eds Ferretti M, Fischer R), pp. 91–104. Elsevier, Oxford.

Fiedler F (1986) Die dendromasse eines hiebsreifen fichtenbestandes. *Beiträge für die Forstwirtschaft*, **20**, 171–180.

Frank D, Esper J (2005a) Characterization and climate response patterns of a high-elevation, multi-species tree-ring network in 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, Esper J, Cook ER (2007) Adjustment for proxy number and coherence in a large-scale temperature reconstruction. *Geophysical Research Letters*, **34**, L16709.

Frank D, Esper J, Zorita E, Wilson R (2010) A noodle, hockey stick, and spaghetti plate: a perspective on high-resolution paleoclimatology. *Wiley Interdisciplinary Reviews: Climate Change*, **1**, 507–516.

Franke J, Frank D, Raible CC, Esper J, Brönnimann S (2013) Spectral biases in tree-ring climate proxies. *Nature Climate Change*, **3**, 360–364.

Fritts HC (1976) *Tree Rings and Climate*. Academic Press, London.

Gedalof Z, Berg AA (2010) Tree ring evidence for limited direct CO_2 fertilization of forests over the 20th century. *Global Biogeochemical Cycles*, **24**, GB3027.

Gilbert RO (1987) *Statistical Methods for Environmental Pollution Monitoring*. Van Nostrand Reinhold, New York.

Girardin MP, Bernier PY, Raulier F, Tardif JC, Conciatori F, Guo XJ (2011) Testing for a CO_2 fertilization effect on growth of Canadian boreal forests. *Journal of Geophysical Research*, **116**, G01012.

Girardin MP, Guo XJ, De Jong R, Kinnard C, Bernier P, Raulier F (2014) Unusual forest growth decline in boreal North America covaries with the retreat of Arctic sea ice. *Global Change Biology*, **20**, 851–866.

Goodman RC, Phillips OL, Baker TR (2012) Tropical forests: tightening up on tree carbon estimates. *Nature*, **491**, 527–527.

Graf Pannatier E, Dobbertin M, Heim A, Schmitt M, Thimonier A, Waldner P, Frey B (2012) Response of carbon fluxes to the 2003 heat wave and drought in three mature forests in Switzerland. *Biogeochemistry*, **107**, 295–317.

Graumlich LJ, Brubaker LB, Grier CC (1989) Long-term trends in forest net primary productivity: cascade Mountains, Washington. *Ecology*, **70**, 405–410.

Harris I, Jones PD, Osborn TJ, Lister DH (2013) Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology*, **34**, 623–642.

Hochbichler E (2002) Vorläufige ergebnisse von biomasseninventuren in buchen und mittelwaldbeständen. In: *Inventur von Biomasse- und Nährstoffvorräten in Waldbeständen* (eds Dietrich HP, Raspe S, Preuhsler T), pp. 37–46. Forstliche Forschungsberichte, München.

Huang JG, Bergeron Y, Denneler B, Berninger F, Tardif J (2007) Response of forest trees to increased atmospheric CO_2. *Critical Reviews in Plant Sciences*, **26**, 265–283.

Ioannidis JPA (2005) Why most published research findings are false. *PLOS Medicine*, **2**, e124.

Jones PD, Briffa KR, Osborn TJ *et al.* (2009) High-resolution palaeoclimatology of the last millennium: a review of current status and future prospects. *The Holocene*, **19**, 3–49.

Keenan TF, Baker I, Barr A *et al.* (2012) Terrestrial biosphere model performance for inter-annual variability of land-atmosphere CO2 exchange. *Global Change Biology*, **18**, 1971–1987.

Knorre AA, Kirdyanov AV, Vaganov EA (2005) Climatically induced interannual variability in aboveground production in forest-tundra and northern taiga of central Siberia. *Oecologia*, **147**, 86–95.

Körner C, Asshoff R, Bignucolo O *et al.* (2005) Carbon flux and growth in mature deciduous forest trees exposed to elevated CO2. *Science*, **309**, 1360–1362.

Koutavas A (2013) CO2 fertilization and enhanced drought resistance in Greek firs from Cephalonia Island, Greece. *Global Change Biology*, **19**, 529–539.

Lévesque M, Saurer M, Siegwolf R, Eilmann B, Brang P, Bugmann H, Rigling A (2013) Drought response of five conifer species under contrasting water availability suggests high vulnerability of Norway spruce and European larch. *Global Change Biology*, **19**, 3184–3199.

Li J, Xie SP, Cook ER *et al.* (2013) El Nino modulations over the past seven centuries. *Nature Climate Change*, **3**, 822–826.

Linares JC, Camarero JJ (2012) From pattern to process: linking intrinsic water-use efficiency to drought-induced forest decline. *Global Change Biology*, **18**, 1000–1015.

Liu Y, Zhang Y, Liu S (2012) Aboveground carbon stock evaluation with different restoration approaches using tree ring chronosequences in Southwest China. *Forest Ecology and Management*, **263**, 39–46.

Lorimer CG, Frelich LE (1989) A methodology for estimating canopy disturbance frequency and intensity in dense temperate forests. *Canadian Journal of Forest Research*, **19**, 651–663.

Martínez-Vilalta J, López BC, Adell N, Badiella L, Ninyerola M (2008) Twentieth century increase of Scots pine radial growth in NE Spain shows strong climate interactions. *Global Change Biology*, **14**, 2868–2881.

Melvin TM (2004) Historical growth rates and changing climatic sensitivity of boreal conifers. PhD thesis, University of East Anglia, East Anglia, 271 pp.

Menzel A, Sparks TH, Estrella N *et al.* (2006) European phenological response to climate change matches the warming pattern. *Global Change Biology*, **12**, 1969–1976.

Mérian P, Pierrat JC, Lebourgeois F (2013) Effect of sampling effort on the regional chronology statistics and climate–growth relationships estimation. *Dendrochronologia*, **31**, 58–67.

MeteoSwiss (2013) Data portal. Available at: http://www.meteoschweiz.admin.ch/web/en/services/data_portal.html (accessed 11 November 2013).

Meyer FD, Bräker OU (2001) Climate response in dominant and suppressed spruce trees, Picea abies (L.) Karst., on a subalpine and lower montane site in Switzerland. *Ecoscience*, **8**, 105–114.

Montero G, Ruiz-Peinado R, Muñoz M (2005) *Producción de Biomasa y Fijación de CO2 Por Los Bosques Españoles*. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Ministerio de Educación y Ciencia, Madrid.

Murri M (1984) Die Bewirtschaftung des Zofinger Stadtwaldes seit der Gründung des Kantons Aargau. Entwicklung von Waldfunktion und Forsteinrichtungen. Diploma thesis, Swiss Federal Institute of Technology ETH, Zurich, 260 pp.

Nabuurs GJ, Lindner M, Verkerk PJ, Gunia K, Deda P, Michalak R, Grassi G (2013) First signs of carbon sink saturation in European forest biomass. *Nature Climate Change*, **3**, 792–796.

Nickless A, Scholes RJ, Archibald S (2011) A method for calculating the variance and confidence intervals for tree biomass estimates obtained from allometric equations. *South African Journal of Science*, **107**, 1–10.

Nock CA, Baker PJ, Wanek W, Leis A, Grabner M, Bunyavejchewin S, Hietz P (2011) Long-term increases in intrinsic water-use efficiency do not lead to increased stem growth in a tropical monsoon forest in western Thailand. *Global Change Biology*, **17**, 1049–1063.

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.

PAGES 2k Consortium (2013) Continental-scale temperature variability during the past two millennia. *Nature Geoscience*, **6**, 339–346.

Pan Y, Birdsey RA, Fang J *et al.* (2011) A large and persistent carbon sink in the world's forests. *Science*, **333**, 988–993.

Pederson N, Varner JM, Palik BJ (2008) Canopy disturbance and tree recruitment over two centuries in a managed longleaf pine landscape. *Forest Ecology and Management*, **254**, 85–95.

Peñuelas J, Canadell JG, Ogaya R (2011) Increased water-use efficiency during the 20th century did not translate into enhanced tree growth. *Global Ecology and Biogeography*, **20**, 597–608.

Poulter B, Pederson N, Liu H *et al.* (2013) Recent trends in inner Asian forest dynamics to temperature and precipitation indicate high sensitivity to climate change. *Agricultural and Forest Meteorology*, **178**, 31–45.

R Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. Available at: http://www.r-project.org/ (accessed 11 November 2013).

Reichstein M, Bahn M, Ciais P *et al.* (2013) Climate extremes and the carbon cycle. *Nature*, **500**, 287–295.

Richardson AD, Williams M, Hollinger DY *et al.* (2010) Estimating parameters of a forest ecosystem C model with measurements of stocks and fluxes as joint constrains. *Oecologia*, **164**, 25–40.

Rutishauser T, Luterbacher J, Defila C, Frank D, Wanner H (2008) Swiss spring plant phenology 2007: extremes, a multi-century perspective, and changes in temperature sensitivity. *Geophysical Research Letters*, **35**, L05703.

Salzer MW, Hughes MK (2007) Bristlecone pine tree rings and volcanic eruptions over the last 5000 year. *Quaternary Research*, **67**, 57–68.

Salzer MW, Hughes MK, Bunn AG, Kipfmueller KF (2009) Recent unprecedented tree-ring growth in bristlecone pine at the highest elevations and possible causes. *Proceedings of the National Academy of Sciences*, **106**, 20348–20353.

Sarewitz D (2012) Beware the creeping cracks of bias. *Nature*, **485**, 149–149.

Schweingruber F, Kairiukstis L, Shiyatov S (1990) Primary data. In: *Methods of Dendrochronology: Applications in the Environmental Sciences* (eds Cook ER, Kairiukstis LA), pp. 23–96. Springer, Dordrecht.

Singer MT, Lorimer CG (1997) Crown release as a potential old-growth restoration approach in northern hardwoods. *Canadian Journal of Forest Research*, **27**, 1222–1232.

Speer JH (2012) *Fundamentals of Tree Ring Research*. The University of Arizona Press, Tucson.

Stephenson NL, Das AJ, Condit R *et al.* (2014) Rate of tree carbon accumulation increases continuously with tree size. *Nature*, **507**, 90–93.

Svoboda M, Janda P, Nagel TA, Fraver S, Rejzek J, Bače R (2012) Disturbance history of an old-growth sub-alpine Picea abies stand in the Bohemian Forest, Czech Republic. *Journal of Vegetation Science*, **23**, 86–97.

Swetnam TW, Allen CD, Betancourt JL (1999) Applied historical ecology: using the past to manage for the future. *Ecological Applications*, **9**, 1189–1206.

Tegel W, Vanmoerkerke J, Büntgen U (2010) Updating historical tree-ring records for climate reconstruction. *Quaternary Science Reviews*, **29**, 1957–1959.

Thimonier A, Schmitt M, Cherubini P, Kräuchi N (2001) Monitoring the Swiss forest: building a research platform. In: *Monitoraggio Ambientale: Metodologie ed Applicazioni. Atti del VIII Corso di Cultura in Ecologia, 2001* (eds Anfodillo T, Carraro V), pp. 121–132. Centro Studi per l'Ambiente Alpino, Università degli Studi di Padova, S. Vito di Cadore, Padova.

Thompson SK (2012) *Sampling*. Wiley, Hoboken.

Trotsiuk V, Hobi ML, Commarmot B (2012) Age structure and disturbance dynamics of the relic virgin beech forest Uholka (Ukrainian Carpathians). *Forest Ecology and Management*, **265**, 181–190.

US Forest Service (2012) Forest inventory and analysis. National core field guide. Volume I: Field data collection procedures for phase 2 plots. Available at: http://www.fia.fs.fed.us/library/field-guides-methods-proc/docs/2013/Core%20FIA%20P2%20field%20guide_6-0_6_27_2013.pdf (accessed 20 November 2013).

Villalba R, Lara A, Masiokas MH *et al.* (2012) Unusual Southern Hemisphere tree growth patterns induced by changes in the Southern Annular Mode. *Nature Geoscience*, **5**, 793–798.

Voelker SL, Muzika RM, Guyette RP, Stambaugh MC (2006) Historical CO2 growth enhancement declines with age in Quercus and Pinus. *Ecological Monographs*, **76**, 549–564.

Weber P, Bugmann H, Fonti P, Rigling A (2008) Using a retrospective dynamic competition index to reconstruct forest succession. *Forest Ecology and Management*, **254**, 96–106.

Wettstein JJ, Littell JS, Wallace JM, Gedalof Z (2011) Coherent region-, species-, and frequency-dependent local climate signals in northern hemisphere tree-ring widths. *Journal of Climate*, **24**, 5998–6012.

Wigley TML, Briffa KR, Jones PD (1984) On the average value of correlated time series, with applications in dendroclimatology and hydrometeorology. *Journal of Climate and Applied Meteorology*, **23**, 201–213.

Wilmking M, Hallinger M, Van Bogaert R *et al.* (2012) Continuously missing outer rings in woody plants at their distributional margins. *Dendrochronologia*, **30**, 213–222.

Wolkovich EM, Cook BI, Allen JM *et al.* (2012) Warming experiments underpredict plant phenological responses to climate change. *Nature*, **485**, 494–497.

Wyckoff PH, Clark JS (2002) The relationship between growth and mortality for seven co-occuring tree species in the southern Appalachian Mountains. *Journal of Ecology*, **90**, 604–615.

Zar JH (1999) *Biostatistical Analysis*. Prentice Hall, Upper Saddle River.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. Illustration of the Regional Curve Standardization (RCS) method as pertains to results shown in Fig. 4 and Table S1. (a) Tree-ring width (TRW) series for *Abies alba* with all individuals shown in light grey and superimposed the 'Classic$_{field}$' sampling design individuals shown in dark grey. Furthermore, mean chronologies for the Population (black) and the 'Classic$_{field}$' subset (blue) are shown. In RCS, the tree-ring time series are firstly aligned by cambial age and a so-called 'Regional Curve' (RC) of the Population (black) and 'Classic$_{field}$' subset (blue) are calculated [smoothed mean of the age-aligned data; (b)]. Secondly, differences between the individual TRW series and the RC are computed and the series are realigned to calendar year (c), mean chronologies are shown in black and blue, respectively. The mean chronologies are then rescaled to a mean of unity, the so-called 'Tree-Ring Indices' (d).

Figure S2. Influence of sample replication on the 'Classic$_{simulated}$' sampling design for *Abies alba* basal area increment (BAI) chronologies (cf. Fig. 6). A total number of 5, 10, . . ., 25 dominant trees (shades of grey) were 100 times randomly selected from the population as shown in (a). The percentage biases are shown in (b), whereby the dotted (dashed) lines indicate the 50% (100%) bias level. For early periods with low sample replication, biases were truncated once they exceeded 500%. As a measure of spread, the median absolute deviation (MAD) is presented in (c) and the sample replication (2 samples per tree) is given in (d) with arrows indicating when the sample replications drop below five series.

Figure S3. Comparison between the above-ground biomass increment (AGB) estimates for 2011 derived from the 'Fixed Plot' (blue) and 'Random' (green) sampling schemes. Radii for 'Fixed Plot' and the average number of trees included the fixed plots ($\pm$SD) increase from left to right from 5, 10, . . ., 25 m. The corresponding 'Random' design shows distributions after sampling the average numbers of trees included in the 'Fixed Plot'. The thick black line within the boxes indicates the median after resampling 100 times, the box marks the interquartile range, whiskers show the inter percentile range and open circles indicate outliers. The solid red line indicates the population value calculated using all available data.

Figure S4. The role of tree size [expressed as the diameter at breast height (DBH)] and age (at breast high) on *Abies alba* basal area increment (BAI) and tree-ring width (TRW). Each BAI and TRW value was plotted along with the corresponding tree size (a and b) and age (c and d). Histograms indicate the relative distribution of data points for each axis and are valid for the entire column or row. Data points were coloured in shades of blue according to the trees social status after the 2011 growing season (cf. Fig. 2b). Curves are smoothed estimates for the Population (black) and dominant tree subset (blue; 'Ad Hoc$_{social\_dominant}$' sampling design) and were truncated when the number of data points drops below 10 for a DBH or age. Controlling for the size or age (i.e. at a given DBH or age) indicates that the dominant tree subset is clearly associated with greater growth rates.

Table S1. Influence of sampling design on *Abies alba* Regional Curve Standardized (RCS) chronologies. Goodness-of-fit statistics relate to the 1912–2011 period. If not further declared, all correlations coefficients (r) and trends are significant at the 99.9% confidence level ($P < 0.001$). Statistics for the resampling types are given as mean $\pm$ SD.

Table S2. Influence of sampling design on *Abies alba* tree-ring width (TRW) chronologies using cross-dated values only. Goodness-of-fit statistics relate to the 1912–2011 period. If not further declared, all correlations coefficients (r) and trends are significant at the 99.9% confidence level ($P < 0.001$). Statistics for the resampling types are given as mean $\pm$ SD.

Table S3. Influence of sampling design on *Abies alba* tree-ring width (TRW) chronologies using all cross-dated values plus growth averages for noncross-datable segments. Goodness-of-fit statistics relate to the 1912–2011 period. If not further declared, all correlations coefficients (r) and trends are significant at the 99.9% confidence level ($P < 0.001$). Statistics for the resampling types are given as mean $\pm$ SD.