Title: Influence of sampling and disturbance history on climatic sensitivity of temperature-limited conifers

Authors:

Miloš Rydval\textsuperscript{a,b}, email: rydval@gmail.com, tel. (00420)735872634

Daniel L. Druckenbrod\textsuperscript{c}

Miroslav Svoboda\textsuperscript{a}

Volodymyr Trotsiuk\textsuperscript{a,d,e}

Pavel Janda\textsuperscript{a}

Martin Mikoláš\textsuperscript{a}

Vojtěch Čada\textsuperscript{a}

Radek Bače\textsuperscript{a}

Marius Teodosiu\textsuperscript{f,g}

Rob Wilson\textsuperscript{b}

\textsuperscript{a} Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Praha 6–Suchdol, Prague, 16521, Czech Republic

\textsuperscript{b} School of Earth and Environmental Sciences, University of St Andrews, UK

\textsuperscript{c} Department of Geological, Environmental, & Marine Sciences, Rider University, Lawrenceville, NJ, USA

\textsuperscript{d} Swiss Federal Research Institute for Forest, Snow and Landscape Research (WSL), Birmensdorf, Switzerland.

\textsuperscript{e} Institute of Agricultural Sciences, ETH Zurich, Switzerland

\textsuperscript{f} “Marin Drăcea” National Research and Development Institute in Forestry, Romania

\textsuperscript{g} Faculty of Forestry, Ştefan cel Mare University of Suceava, Romania

This document is the accepted manuscript version of the following article:


https://doi.org/10.1177/0959683618782605
ABSTRACT: Accurately capturing medium-to-low frequency trends in tree-ring data is vital to assessing climatic response and developing robust reconstructions of past climate. Non-climatic disturbance can affect growth trends in tree-ring width (RW) series and bias climate information obtained from such records. It is important to develop suitable strategies to ensure the development of chronologies that minimize these medium-to-low frequency biases. By performing high density sampling (760 trees) over a ~40ha natural high elevation Norway spruce (Picea abies) stand in the Romanian Carpathians, this study assessed the suitability of several sampling strategies for developing chronologies with an optimal climate signal for dendroclimatic purposes. There was a roughly equal probability for chronologies (40 samples each) to express a reasonable (r=0.3-0.5) to non-existent climate signal. While showing a strong high-frequency response, older/larger trees expressed the weakest overall temperature signal. Although random sampling yielded the most consistent climate signal in all sub-chronologies, the outcome was still sub-optimal. Alternative strategies to optimise the climate signal, including very high replication and principal component analysis, were also unable to minimize this disturbance bias and produce chronologies adequately representing climatic trends, indicating that larger scale disturbances can produce synchronous pervasive disturbance trends that affect a large part of a sampled population. The Curve Intervention Detection (CID) method, used to identify and reduce the influence of disturbance trends in the RW chronologies, considerably improved climate signal representation (from r=0.28 before correction to r=0.41 after correction for the full 760 sample chronology over 1909-2009) and represents a potentially important new approach for assessing disturbance impacts on RW chronologies. Blue intensity (BI) also shows promise as a climatically more sensitive variable which, unlike RW, does not appear significantly affected by disturbance. We recommend that studies utilizing RW chronologies to investigate medium to long-term climatic trends also assess disturbance impact on those series.

KEYWORDS: disturbance detection; sampling bias; climatic signal; blue intensity; tree rings; Norway spruce; Romanian Carpathian Mountains
INTRODUCTION

The accurate representation of climatic variability in the growth trends contained in tree-ring records from climatically sensitive trees is central to assessing growth-climate response and the development of robust dendroclimatic reconstructions (e.g. Anchukaitis et al., 2017; Cook et al., 2015; Cook et al., 2016; D’Arrigo et al., 2006; Luterbacher et al., 2016; Wilson et al., 2016). The suitability of strategically sampled tree-ring chronologies for reconstructing a particular climatic variable is typically evaluated by examining the growth-climate response and the strength of this relationship. This process partly relies on the assumption that chronologies are developed from a finite number of samples that are representative of the population. In climatically sensitive stands (i.e. temperature sensitive trees at high latitude or elevation tree-line locations) it is usually assumed that when adequate measures are taken to avoid sampling trees likely affected by non-climatic influences, the common signal of the sample chronology represents the common climatic signal of the population (Hughes, 2011).

Tree growth is the product of a range of environmental influences that are integrated into the annual growth increment (Cook, 1985; Vaganov et al., 2006). Natural disturbance is one key element of forest ecosystem development (Attiwill, 1994). The presence of non-climatic disturbance trends in tree ring width (RW) series complicates the development of climatically sensitive tree-ring based records (e.g. Briffa et al., 1996; Gunnarson et al., 2012; Rydval et al., 2016). Yet few, if any, studies explicitly assess the influence of disturbance as a part of dendroclimatic research. A common presumption is that the effects of disturbance are either negligible or asynchronous so that their influence is canceled out through the development of a mean chronology of detrended series, or they can be minimized by applying appropriate detrending techniques in cases when such trends occur systematically (Hughes, 2011). It has been shown that larger scale intermediate and higher severity disturbances can result in synchrony of disturbance histories across the landscape on the stand level and regional spatial scales (e.g. D’Amato and Orwig, 2008; Kulakowski and Veblen, 2002; Zielonka et al., 2010). While flexible data-adaptive detrending approaches such as cubic smoothing splines (Cook and Peters, 1981) have been utilized to limit the influence of non-climatic (e.g. disturbance) trends in RW data, a detrimental side-effect of such techniques is the loss of lower frequency (i.e. multidecadal to multicentennial) climatic variability.
Numerous studies have investigated dendrochronological biases and uncertainties related to various methodological aspects of tree-ring data development including detrending (e.g. Briffa and Melvin, 2011; Cook et al., 1995; Helama et al., 2004; Melvin and Briffa, 2008; Melvin et al., 2013), sample size and signal strength (e.g. Mérian et al., 2013; Osborn et al., 1997; Wigley et al., 1984), sampling design and microsite conditions (e.g. Cherubini et al., 1998; Düthorn et al., 2013, 2015; Nehrbass-Ahles et al., 2014), and tree age (e.g. Carrer and Urbinati, 2004; Esper et al., 2008; Fish et al., 2010). In an extensive assessment of sampling design strategies, Nehrbass-Ahles et al. (2014) highlighted that many common sampling approaches used for developing representations of forest response to environmental change can induce sampling related biases. However, relatively little is known about how disturbance related growth trends affect the climate signal in tree ring series.

Time-series analysis with intervention detection (Box and Jenkins, 1970; Box and Tiao, 1975) is an evolving area for studying disturbance in RW data (Druckenbrod, 2005). A time-series based method called Curve Intervention Detection (CID) has been developed to characterize disturbance history and quantify the effects of disturbance trends on individual RW series and overall chronology structure (Druckenbrod, 2005; Druckenbrod et al., 2013). Chronology distortion and climate signal degradation, due to synchronous disturbance related growth releases as a result of systematic timber felling, was identified using the CID technique by Rydval et al. (2016) in Scots pine RW chronologies from Scotland. However, such a technique has not previously been applied to investigate trends resulting from natural sources of disturbance on the strength of the climate signal in tree-ring records.

Building on the work of Rydval et al. (2016), in this study we applied the CID method to a new forest system and species by examining RW series from an unmanaged natural closed canopy Norway spruce (Picea abies) stand in Romania (shaped by a mixed-severity natural disturbance regime with partial landscape synchronization and unperturbed by human activities - Svoboda et al., 2014) to examine the extent to which natural disturbance can affect climate signal strength in RW data. We investigate (1) whether natural disturbance can produce widespread and synchronized trends, as those resulting from human activities, that would significantly impact the expression of the climate signal in tree-ring chronologies, and (2) which sampling or data processing approach best expresses the climate signal. To this end, we firstly evaluated a set of sampling strategies by subsampling a large dataset of RW
data from a single stand according to a set of characteristics reflecting sampling strategies that are relevant in a dendroclimatic context. The application of additional data processing techniques (including disturbance trend detection and correction using the CID method, and isolation of the dominant signals through principal component analysis) were investigated in an attempt to optimize the climate signal. We applied the CID time-series analysis technique in order to characterize the disturbance history, its impact on overall chronology structure and subsequently to reduce the influence of disturbance-related trends on RW chronologies (Druckenbrod, 2005; Druckenbrod et al., 2013; Rydval et al., 2016). As an alternative to RW data, a subset of chronologies was developed from series of the blue intensity (BI) parameter (Björklund et al., 2014a; McCarroll et al., 2002; Rydval et al., 2014) to ascertain whether such data can be used to produce proxy climate records unbiased (or less biased) by the presence of disturbance trends.

METHODS

Sampling site

Samples were collected and measured from 760 high-elevation Norway spruce (Picea abies) trees (cored at breast height) located in an approx. 40 ha natural Norway spruce dominant stand (47°06'53"N,25°15'26"E) in Călimani National Park (hereafter Calimani) in the Eastern Carpathians of northern Romania (Figure 1; see Svoboda et al. (2014) for details regarding sample collection). The selected sampling site is located within an elevational range of around 1500-1650 m a.s.l., ~100-250 m below the regional timberline (approx. 1780 m a.s.l.) and ~200-350 m below the treeline (approx. 1860 m a.s.l.) (Popa and Kern, 2009) and with slope varying from around 10° to 25°. Podzols are the predominant soil type in the study region (Valtera et al., 2013). The area has a mean annual temperature of 2.1-3.1°C estimated from 0.5° gridded CRU TS3.23 temperatures (based on the period 2005-2014 and adjusted for elevation). Over the same period, temperatures have increased by approximately 1.6°C relative to the first decade of the 20th century. Mean annual precipitation is about 910 mm (2005-2014 mean).
Sample collection was performed in an area with no significant human activities in the past (including evidence from historical documentation) and subject only to natural stand dynamics and disturbance regimes (Svoboda et al., 2014). Considering the size of the sampled area and number of samples collected, this high density sampling strategy, similar to that of Nehrbass-Ahles (2014), was intended to provide a highly representative sample of the whole stand population by sampling a diverse range of tree size and age classes. This approach makes it possible to group samples and construct chronologies according to a range of characteristics.

<table>
<thead>
<tr>
<th>CHRON. (SUBSET)</th>
<th>NR. OF SERIES</th>
<th>MEAN ELEVATION</th>
<th>CHRON. LENGTH</th>
<th>EPS ≥ 0.85</th>
<th>CHRON. (BY DBH)</th>
<th>DBH RANGE (CM)</th>
<th>CHRON. (BY AGE)</th>
<th>AGE RANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLOT-ALL</td>
<td>760</td>
<td>1578</td>
<td>1673-2009</td>
<td>1744-2009</td>
<td>110-925</td>
<td>31-337</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLOT-1</td>
<td>40</td>
<td>1523</td>
<td>1731-2009</td>
<td>1888-2009</td>
<td>DBH-1</td>
<td>110-235</td>
<td>AGE-1</td>
<td>31-82</td>
</tr>
<tr>
<td>PLOT-3</td>
<td>40</td>
<td>1602</td>
<td>1743-2009</td>
<td>1906-2009</td>
<td>DBH-3</td>
<td>265-280</td>
<td>AGE-3</td>
<td>85-87</td>
</tr>
<tr>
<td>PLOT-4</td>
<td>40</td>
<td>1616</td>
<td>1772-2009</td>
<td>1905-2009</td>
<td>DBH-4</td>
<td>280-300</td>
<td>AGE-4</td>
<td>87-89</td>
</tr>
<tr>
<td>PLOT-5</td>
<td>40</td>
<td>1588</td>
<td>1768-2009</td>
<td>1820-2009</td>
<td>DBH-5</td>
<td>300-320</td>
<td>AGE-5</td>
<td>89-92</td>
</tr>
<tr>
<td>PLOT-7</td>
<td>40</td>
<td>1633</td>
<td>1712-2009</td>
<td>1835-2009</td>
<td>DBH-7</td>
<td>340-360</td>
<td>AGE-7</td>
<td>97-111</td>
</tr>
<tr>
<td>PLOT-8</td>
<td>40</td>
<td>1516</td>
<td>1673-2009</td>
<td>1897-2009</td>
<td>DBH-8</td>
<td>360-380</td>
<td>AGE-8</td>
<td>112-118</td>
</tr>
<tr>
<td>PLOT-9</td>
<td>40</td>
<td>1514</td>
<td>1700-2009</td>
<td>1903-2009</td>
<td>DBH-9</td>
<td>380-400</td>
<td>AGE-9</td>
<td>118-122</td>
</tr>
<tr>
<td>PLOT-10</td>
<td>40</td>
<td>1606</td>
<td>1763-2009</td>
<td>1835-2009</td>
<td>DBH-10</td>
<td>400-415</td>
<td>AGE-10</td>
<td>122-127</td>
</tr>
<tr>
<td>PLOT-12</td>
<td>40</td>
<td>1587</td>
<td>1763-2009</td>
<td>1906-2009</td>
<td>DBH-12</td>
<td>430-450</td>
<td>AGE-12</td>
<td>134-142</td>
</tr>
<tr>
<td>PLOT-14</td>
<td>40</td>
<td>1552</td>
<td>1705-2009</td>
<td>1861-2009</td>
<td>DBH-14</td>
<td>480-500</td>
<td>AGE-14</td>
<td>148-156</td>
</tr>
<tr>
<td>PLOT-17</td>
<td>40</td>
<td>1511</td>
<td>1752-2009</td>
<td>1903-2009</td>
<td>DBH-17</td>
<td>550-590</td>
<td>AGE-17</td>
<td>176-201</td>
</tr>
</tbody>
</table>

**Table 1**: Site and chronology descriptive information. PLOT represents chronologies developed according to sample location (i.e. plot-based), DBH chronologies are composed of samples grouped according to diameter at breast height, and AGE represents chronologies with samples grouped according to tree recruitment age. With the exception of the PLOT-ALL chronology, all other chronologies were developed using 40 samples. (EPS = expressed population signal, Wigley et al., 1984).

**Data analysis**

Sampled cores were mounted and glued on wooden mounts and subsequently surfaced with a blade to enhance the visibility of ring boundaries. To help determine tree recruitment age (i.e. the number of rings at coring height), pith-offset was estimated using an acetate sheet with concentric circles. However, the method of sample collection specifically focused on minimizing pith-offset and so the majority of samples included the pith. Ring width was then measured using a LINTAB traversing...
measuring stage coupled with TsapWin (RINTECH, Germany) measuring software to a precision of 0.01 mm. Sample crossdating was performed using standard dendrochronological approaches (Stokes and Smiley, 1968) and crossdating of measured series was checked with CDendro (Larsson, 2015).

**Disturbance detection and correction**

Curve Intervention Detection (CID) is a time-series intervention detection method based on the work of Druckenbrod (2005) and Druckenbrod et al. (2013). The method was used here to objectively identify and remove disturbance trends from individual RW series following the procedure described in Rydval et al. (2016), where it was used to identify and correct for growth release trends due to logging-related disturbance in Scottish Scots pine (*Pinus sylvestris*) samples. In this study, both growth release and growth suppression trends were detected and removed. Prior to the CID procedure, a constant of 1 mm was added to all measurements to avoid the possibility of losing tree-ring information during the disturbance removal procedure (Rydval et al., 2016). As part of the CID procedure, RW measurement series were first power transformed (Cook and Peters, 1997) and then detrended by fitting a negative exponential or linear function. Disturbance trends were identified as outliers from 9-30 year running mean distributions based on the residual series of each detrended RW series and autoregressive model estimates. The identified release / suppression trend was removed by subtracting a curve (Warren, 1980) fitted to the series from the point where the initiation of the disturbance-related trend was identified. The procedure was repeated until no further outliers were detected. The disturbance-corrected series were then re-expressed in the original (non-detrended) measurement format so that both the corrected and uncorrected series could then be detrended in the same way. For a detailed description of the method refer to Rydval et al. (2016). CID (ver. 1.05) was used in these analyses and is included in the supplemental materials as Matlab code files. A freely-available executable (ver. 1.07) is available using the Matlab compiler. Contact Daniel Druckenbrod (ddruckenbrod@rider.edu) as this version is dependent on operating system and Matlab release version. These time-series methods are a work in progress, but we also welcome other researchers to experiment with this tool to detect and isolate disturbance events in ring-width series.
**RW chronology development**

Two sets of chronologies were developed with the first set composed of series prior to disturbance correction using the CID method (i.e. uncorrected for the influence of disturbance – pre-CID) and the second set using series after correcting for disturbance trends with CID (post-CID). Using ARSTAN (Cook and Krusic, 2005), both sets of RW series were power transformed to stabilize series variance (Cook and Peters, 1997) and detrended by subtracting negative exponential or negatively sloping linear functions. The mean chronology index was calculated using Tukey’s robust bi-weight mean to reduce the influence of outlier values (Cook and Kairiukstis, 1990). Variance stabilization of the mean chronology, due to changing replication, was then performed according to the procedure described in Osborn et al. (1997).

In addition to developing an uncorrected (pre-CID) and disturbance corrected (post-CID) mean chronology from all 760 samples, the entire collection of series was also divided into 19 separate sub-plot chronologies (each including 40 series) compiled by grouping series according to 1) the plot location where the samples were collected (PLOT) (Figure 1); 2) random sample selection without replacement (RAN); 3) tree recruitment age (AGE), 4) and the diameter at breast height (DBH) (Table 1). All chronologies were truncated based on an expressed population signal (EPS – Wigley et al., 1984) cut-off of EPS ≥ 0.85.

Principal Component (PC) analysis, with varimax rotation, was applied using the IBM SPSS (v.20.0) statistical package (SPSS, 2011) to both pre-CID and post-CID location-based (PLOT) chronologies to reduce the dimensionality of the RW predictor dataset in order to extract the dominant modes of variance. Based on the temporal span of the shortest chronology (Table 1), the period 1909-2009 was used in order to include all chronologies in the analysis. Only the lowest order PC scores with an eigenvalue >1 were retained.
Blue Intensity chronology development

Similar to maximum latewood density, blue intensity (BI) represents summer growing conditions usually reflecting a (late) summer response to temperature in conifers from temperature limited locations (e.g. Björklund et al., 2014b; McCarroll et al., 2013; Rydval et al., 2014; Wilson et al., 2014). BI measurements were developed for a subset of the samples (three chronologies – PLOT-3, PLOT-7 and PLOT-10; 40 samples each) following Rydval et al. (2014). Since, unlike other conifers such as pine, Norway spruce samples do not exhibit any apparent visual colour difference between the heartwood and sapwood that would affect BI measurements, chemical treatment involving sample resin extraction was not performed. Such an approach was also considered adequate in a study by Wilson et al. (2014) examining BI data from Engelmann spruce in British Columbia. Samples surfaced with sanding paper up to 1200 grit grade were scanned using an Epson Expression 10000 XL flatbed scanner combined with SilverFast Ai (v.6.6 - Laser Soft Imaging AG, Kiel, Germany) scanning software. Scanner calibration was performed with the SilverFast IT8 calibration procedure using a Fujicolor Crystal Archive IT8.7/2 calibration target. A resolution of 2400 dpi was used for scanning. During the scanning process, samples were covered with a black cloth to prevent biases due to ambient light.

CooRecorder measurement software (Larsson, 2015) was used to measure BI from scanned images. The BI series were then inverted according to Rydval et al. (2014) to express a positive relationship with RW and instrumental temperatures and subsequently detrended by subtraction from fitted negatively sloping linear functions. The mean BI chronology was calculated and truncated (EPS = 0.85) in the same way as the RW chronologies.

Climate data

For this study, in order to allow the assessment of the longest possible temporal span of the tree-ring data, we used mean temperature series from a meteorological station in Sibiu, Romania (hereafter SIBIU) covering the period 1851-2015 (data for 1918 were unavailable and were estimated from the relevant 0.5° CRU TS3.23 grid scaled to SIBIU) located approximately 170 km to the SSW of
An additional temperature record was composited using the longest Central and Eastern European instrumental records in order to assess the whole span of the full 760 sample Calimani chronology. This Central/East European (CEU) composite covers the period 1773-2014 and includes temperature series from Prague (Czech Republic), Vienna (Austria), Kraków (Poland), Budapest (Hungary), Lviv (Ukraine) and Kishinev (Moldova). The individual instrumental series were converted to anomalies relative to 1961-1989 and combined as a simple average. To adjust for variance changes due to the changing number of series in the composite through time, the variance of the mean series was adjusted according to Osborn et al. (1997). Climate data were used to assess the strength of the climatic signal in tree ring chronologies using the Pearson's correlation coefficient (r).

RESULTS

Four sets of chronologies developed according to various sampling strategies are presented in Figure 2 with additional chronology information in Table 1. As the strongest significant chronology response was observed with June-July mean temperatures (see supplementary Figure S1), RW chronologies were assessed using this seasonal window. This seasonal response agrees with Sidor et al. (2015) who also noted a significant June-July mean temperature signal in high-elevation spruce sites in the Romanian Carpathians including Calimani.

The location-based ‘PLOT’ chronologies (Figure 2a – see Figure 1 for plot locations) displayed a large range of variability (especially before ~1960) which is also reflected in the wide range of variation in correlation between each chronology and June-July average instrumental temperatures ($r = 0.07$ to $0.46$; $r_\chi^2 = 0.26$ – Table 2). The ‘RAN’ chronologies based on random selection of samples (without replacement; Figure 2b) produced a more uniform range of variability which was also observed in the relationship between the chronologies and instrumental temperatures ($r = 0.24$ to $0.35$; $r_\chi^2 = 0.26$ – Table 2). When compared with the PLOT chronologies, the correlation range of these ‘random sample’ chronologies against instrumental temperatures was considerably narrower, although the mean correlation was virtually the same and while the very low correlations were no longer observed, the higher correlations were also no longer present. When grouped according to stem size (i.e. DBH – Figure
2c), chronologies displayed considerable variability particularly in the first half of the 20th century as well as in the most recent period (i.e. after ~1990). Chronologies composed of series from broader-stemmed (higher DBH) trees tended to correlate more weakly with instrumental temperatures (r = -0.06 to 0.28 for chronologies DBH-12 – DBH-19; see Table 3 for details), whereas trees with narrower stems (lower DBH) appeared to exhibit higher correlations (r = 0.37 to 0.47 for chronologies DBH-1 – DBH-11 excluding the weaker DBH-9 chronology; see Table 3). The chronologies grouped according to age showed a similar range of variability to the DBH-based chronologies (Figure 2d). Although not as clear, there was a tendency for younger chronologies to correlate more strongly than the oldest chronologies (Figure 2d; Table 3). However, when examining only the high frequency (inter-annual) relationship between the chronologies and temperature (1st differenced results in Table 3), there was little difference between the young and old tree chronologies and larger trees actually displayed a stronger signal than chronologies from smaller trees (r = 0.30 to 0.43, r\textsubscript{s} = 0.38 for chronologies DBH-1 – DBH-11; r = 0.42 to 0.51, r\textsubscript{s} = 0.47 for chronologies DBH-12 – DBH-19). Unsurprisingly, a strong relationship (r = 0.63) was observed between age and DBH (Figure S2), which indicates that older trees generally also tend to be larger (i.e. higher DBH) and vice-versa.

Table 2: Average correlation and correlation range of chronologies before (pre-CID) and after (post-CID) disturbance correction and first differenced chronologies developed with different sampling strategies, including samples grouped by location (PLOT), random sample selection (RAN), grouping according to diameter at breast height (DBH), and sample age (AGE), against SIBIU Jun-Jul mean instrumental temperatures. (r\textsubscript{s} represents the mean correlation ± 1 standard deviation, while r\textsubscript{range} represents the full correlation range)
Table 3: Correlations of chronologies before (pre-CID) and after (post-CID) disturbance correction and first differenced chronologies sampled using different sampling strategies, including random sample selection (RAN), grouping according to diameter at breast height (DBH), and sample age (AGE), against Jun-Jul mean instrumental temperatures from Sibiu (shading is used to aid interpretation of the results with darker shades indicating higher correlations).

A summary of the general disturbance history at Calimani is provided in Figure 3a. The results showed three major pulses or clusters of disturbance events, which affected a large proportion of the stand, detected in the 1740s, the middle of the 19th century and the 1910s followed by growth releases in the subsequent decades attributable to those disturbances. Disturbance suppression events were detected in the mid/late 18th century, although the predominant release events were more prevalent whereas suppression events did not appear to considerably affect the mean disturbance chronology. The pre- and post- correction chronologies (Figure 3b and 3c respectively) indicated a wider spread in individual pre-CID chronologies and greater deviation from the mean chronology compared to their post-CID counterparts. This was also observed with the other sampling approaches (Table 2). The disturbance growth chronology in Figure 3a identified periods of growth release pulses attributable to disturbance which are evident in the Figure 3b mean chronology. After disturbance correction, the spread of the individual chronologies was reduced as the growth release trends were removed and the post-CID chronologies exhibited greater similarity to the mean chronology which did not contain the growth release trends. The mean pre- and post-CID chronologies are displayed together with the SIBIU
instrumental temperature record (back to 1851) and the Central European (CEU) instrumental
temperature composite extending back to the 1770s (Figure 3d). The main differences between the
corrected and uncorrected chronologies become apparent with lower post-correction RW index values
in the first half of the 20th century and higher values from approximately 1770 until 1850. These results
also highlighted the improved agreement of the post-CID chronology with both the shorter SIBIU ($r_{pre-CID}$
= 0.27; $r_{post-CID}$ = 0.36) and longer CEU ($r_{pre-CID}$ = 0.14; $r_{post-CID}$ = 0.26) instrumental temperature series.

[insert Figure 3]

The change in correlation between individual pre-CID and post-CID chronologies and the SIBIU
temperature series for the common 1909-2009 period (Figure 4a) showed overall improvement of the
mean chronology as well as all individual chronologies with the exception of PLOT-16. Similar results
were obtained when evaluating the full length of each chronology (Figure 4c). A comparison of the pre-
and post-CID root-mean-square error (RMSE) results for the common 1909-2009 period (Figure 4b) and
the full length of overlap (Figure 4d) between individual PLOT chronologies and SIBIU indicated a RMSE
decrease in nearly all post-CID chronologies. This RMSE pattern largely mirrored the correlation change
results and indicated chronology improvement in the sense that lower RMSE results were observed in
the post-correction chronologies. The results from Figure 4c were also represented spatially in Figure 1.
The greatest degree of post-CID chronology correlation increase with instrumental temperatures was
observed in chronologies from the southeastern slope, which predominantly contained growth release
trends in the first half of the 20th century. Chronologies showing intermediate improvement were
located farther north and included chronologies from the northwestern slope which predominantly
contained disturbance related trends in the second half of the 19th century. The least improvement was
observed in chronologies from the northwestern (PLOT-1) and northernmost (PLOT-18) investigated
locations as well as on the southern ridge (PLOT-6) and in the valley (PLOT-16), with the latter two
chronologies exhibiting no late 19th / early 20th century disturbance trends. Supplementary Figure S3
highlights in greater detail this broad spatial and temporal split in the pattern of disturbance of the
northwest / southeast groups and the very large percentage of trees in each group affected by these
two major disturbance events. Individual chronologies developed according to the other sampling
strategies showed an overall pattern of post-CID improvement similar to the location based (PLOT) assessment (Table 3). The pre-CID and post-CID results from Table 3 along with their respective chronologies are displayed graphically in supplementary Figure S4.

The principal component (PC) time-series scores of the dominant modes of variance of the pre-CID dataset are presented in Figure 5a and include three PCs (loadings of the chronologies on each eigenvector are presented in Table 4). PC3 showed the strongest correlation with SIBIU Jun-Jul temperatures ($r = 0.45$) and PC1 correlated more weakly ($r = 0.33$), while PC2 was weakly negatively correlated with temperatures ($r = -0.24$). When compared to the disturbance chronology in Figure 3a, a strong correlation was observed with PC2 ($r = 0.91$). After CID correction, only two dominant PCs were identified. Although the first PC was uncorrelated with temperatures, a stronger relationship was observed between temperature and PC2 ($r = 0.56$) than was identified with any of the pre-CID PC scores. Conversely, PC1 significantly correlated with the disturbance chronology ($r = 0.50$), whereas no correlation was found with PC2.

<table>
<thead>
<tr>
<th>CHRON. (SUBSET)</th>
<th>PC1 (PRE-CID)</th>
<th>PC2 (PRE-CID)</th>
<th>PC3 (PRE-CID)</th>
<th>CHRON. (SUBSET)</th>
<th>PC1 (POST-CID)</th>
<th>PC2 (POST-CID)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLOT-10</td>
<td>0.911</td>
<td>0.163</td>
<td>0.275</td>
<td>PLOT-1</td>
<td>0.932</td>
<td>0.157</td>
</tr>
<tr>
<td>PLOT-5</td>
<td>0.907</td>
<td>0.216</td>
<td>0.168</td>
<td>PLOT-11</td>
<td>0.878</td>
<td>0.358</td>
</tr>
<tr>
<td>PLOT-19</td>
<td>0.885</td>
<td>0.302</td>
<td>0.265</td>
<td>PLOT-14</td>
<td>0.875</td>
<td>0.244</td>
</tr>
<tr>
<td>PLOT-2</td>
<td>0.826</td>
<td>0.350</td>
<td>0.301</td>
<td>PLOT-19</td>
<td>0.827</td>
<td>0.472</td>
</tr>
<tr>
<td>PLOT-1</td>
<td>0.813</td>
<td>0.528</td>
<td>-0.071</td>
<td>PLOT-6</td>
<td>0.791</td>
<td>0.508</td>
</tr>
<tr>
<td>PLOT-9</td>
<td>0.741</td>
<td>0.436</td>
<td>0.403</td>
<td>PLOT-10</td>
<td>0.791</td>
<td>0.478</td>
</tr>
<tr>
<td>PLOT-11</td>
<td>0.703</td>
<td>0.625</td>
<td>0.160</td>
<td>PLOT-2</td>
<td>0.777</td>
<td>0.530</td>
</tr>
<tr>
<td>PLOT-7</td>
<td>0.603</td>
<td>0.460</td>
<td>0.576</td>
<td>PLOT-8</td>
<td>0.772</td>
<td>0.374</td>
</tr>
<tr>
<td>PLOT-8</td>
<td>0.414</td>
<td>0.828</td>
<td>0.312</td>
<td>PLOT-5</td>
<td>0.721</td>
<td>0.402</td>
</tr>
<tr>
<td>PLOT-14</td>
<td>0.535</td>
<td>0.795</td>
<td>0.156</td>
<td>PLOT-7</td>
<td>0.718</td>
<td>0.601</td>
</tr>
<tr>
<td>PLOT-15</td>
<td>0.356</td>
<td>0.761</td>
<td>0.461</td>
<td>PLOT-9</td>
<td>0.714</td>
<td>0.585</td>
</tr>
<tr>
<td>PLOT-3</td>
<td>0.377</td>
<td>0.758</td>
<td>0.448</td>
<td>PLOT-16</td>
<td>0.112</td>
<td>0.876</td>
</tr>
<tr>
<td>PLOT-4</td>
<td>0.377</td>
<td>0.692</td>
<td>0.536</td>
<td>PLOT-4</td>
<td>0.337</td>
<td>0.858</td>
</tr>
<tr>
<td>PLOT-6</td>
<td>0.551</td>
<td>0.652</td>
<td>0.481</td>
<td>PLOT-13</td>
<td>0.367</td>
<td>0.858</td>
</tr>
<tr>
<td>PLOT-16</td>
<td>0.121</td>
<td>0.010</td>
<td>0.934</td>
<td>PLOT-17</td>
<td>0.438</td>
<td>0.815</td>
</tr>
<tr>
<td>PLOT-12</td>
<td>0.173</td>
<td>0.518</td>
<td>0.803</td>
<td>PLOT-12</td>
<td>0.539</td>
<td>0.788</td>
</tr>
<tr>
<td>PLOT-18</td>
<td>0.413</td>
<td>0.337</td>
<td>0.776</td>
<td>PLOT-15</td>
<td>0.523</td>
<td>0.780</td>
</tr>
<tr>
<td>PLOT-13</td>
<td>0.138</td>
<td>0.579</td>
<td>0.723</td>
<td>PLOT-3</td>
<td>0.529</td>
<td>0.772</td>
</tr>
<tr>
<td>PLOT-17</td>
<td>0.270</td>
<td>0.582</td>
<td>0.710</td>
<td>PLOT-18</td>
<td>0.587</td>
<td>0.635</td>
</tr>
</tbody>
</table>

**Table 4:** Principal Component (PC) Analysis loadings of location-based (PLOT) chronologies before (pre-CID) and after (post-CID) disturbance correction on the dominant eigenvectors (results in bold indicate the strongest loading of each chronology).
The restricted three site (PLOT3, 7 and 10) correlation response analysis assessing the relationship between pre-CID, post-CID and BI data, with SIBIU temperatures (Figure 6a) clearly shows disturbance trends in the RW data with various degrees of post-CID improvement. In contrast to the relatively narrow RW seasonal response, the BI chronology responded more strongly to a broader seasonal window displaying highest correlations with mean July-September temperatures ($r = 0.65$). The response of BI was stronger than post-CID RW with respect to the optimal season of each parameter. Although improvement of the post-CID chronologies (Figures 6b, c and d) was apparent especially before 1880 when the deviation of the pre-CID chronology from the instrumental record was reduced, the degree of improvement was limited, particularly as periods of weaker agreement remained as indicated by running correlations between the pre-/post-CID chronologies and SIBIU temperatures. In contrast, the BI chronologies more closely matched the instrumental trends with running correlations displaying a consistently strong relationship back into the 19th century.

DISCUSSION

Considering the relatively small area of the Calimani study area, it would be reasonable to assume that chronologies developed from the plots would be similar in the absence of disturbance and should therefore also express a very similar climate signal. However, despite the adequate replication of the different chronologies, a range of chronology trends were observed (Figure 2a) expressing substantial differences in correlation with temperature ranging from zero to ~0.5. The possibility of developing a climatically sensitive chronology by randomly choosing and sampling all trees in a specific plot would therefore depend on chance. An alternative approach, which randomly samples trees from the whole stand (Figure 2b), produced a more consistent and uniform outcome, although generally resulting in correlations of only ~0.3 with temperature.

A sampling strategy commonly applied for dendroclimatology favours the preferential selection of larger / wider (i.e. higher DBH) and presumed older trees in order to extend a chronology as far back in time as possible. The strong age / DBH relationship (Figure S2), would support this type of reasoning.
Forming chronologies by grouping series according to DBH, revealed that samples from the largest trees expressed the weakest temperature signal (Figure 2c). Although less clear-cut than the DBH results, there was also a tendency for chronologies composed of samples from old trees to produce a climatically weaker signal. Yet the 1st-differenced results indicate that there is no fundamental limitation in the ability of older trees to record climatic information and that, at least at high frequencies, the sensitivity of larger trees compared to smaller ones is actually greater. This observation demonstrates that the overall response of trees does not simply weaken with age but is instead related to the presence of disturbance related trends that bias the lower frequency expressed in the data after detrending.

It would therefore be reasonable to conclude that the weaker performance of the older (and generally larger) trees at decadal and longer timescales is at least, in part, related to the greater likelihood that older trees will be affected by some disturbance event during their life than younger trees. Although the Figure 2d results represent a common period of analysis in the 20th century, disturbance trends in earlier parts of the older chronologies would still affect the chronology trends in this recent period (i.e. by biasing the fit of the detrending functions). Therefore, without addressing trend biases, sampling the largest (and perhaps oldest) trees will likely produce chronologies with poor climatic sensitivity at decadal and longer timescales. This presents a problem as the oldest trees are also the most valuable for studying longer-term climate. Furthermore, none of these strategies can guarantee a good chronology response in terms of climate signal. We must therefore ask the question, why is this the case, and does a reasonable approach exist for optimizing (maximizing) the climatic potential of the population sample? If disturbance is an important factor influencing climate response, then disturbance correction may be an appropriate strategy to improve calibration.

Previous studies have identified wind and windstorm damage as the dominant determinant of large scale severe disturbance in the Romanian Carpathians and to a lesser extent insect outbreaks and snow damage (e.g. Griffiths et al., 2014; Popa, 2008; Svboda et al., 2014), which would account for the observed synchronous and temporally clustered nature of disturbance (Figure 3a) and the imprinting of disturbance trends in individual RW series on the mean chronologies (Figure 3b). There is also evidence that large scale severe windstorm events can impact the majority of trees over relatively large areas, as
for example during the 2004 event in the Slovakian Tatra Mountains (Western Carpathians) which
affected 12,000 ha of montane forest stands (e.g. Holeksa et al., 2016; Zielonka et al., 2010). The
reduced range and more uniform trends expressed in individual PLOT chronologies, which more closely
matched the mean (all 760 series) chronology after CID correction in Figure 3c compared to the pre-
periods of growth release), suggests that CID correction produced individual PLOT sub-chronologies that
more accurately approximate the larger-scale (regional population) chronology.

Compared to the lower mean correlation of the unfiltered pre-CID chronologies (Figure 3b, \( r = 0.28 \)), correlations of the unfiltered post-CID chronologies (Figure 3c, \( r = 0.41 \)) as well as the 1st
differenced pre-CID and post-CID chronology versions \( (r = 0.45 \) and 0.49 respectively) were all
considerably higher. This suggests that the high frequency climate signal in pre-CID chronologies was
unaffected by the presence of disturbance trends and that the weaker correlations of the unfiltered pre-
CID chronologies were related to the lower frequency trends, which was supported by the substantial
degree of unfiltered post-CID correlation improvement (Figure 3c). Furthermore, the long term trend of
the post-CID mean (all 760 series) chronology differed when compared to its pre-CID counterpart.
Specifically, the most apparent changes included a reduction of index values affected by growth release
in the first half of the 20th century and higher values before 1850 after correcting for non-climatic
growth suppression trends. Taken together, the above evidence suggests that a disturbance-free
chronology may not necessarily be achieved simply by collecting and averaging a very large number of
series.

The improvement in post-CID chronology running correlations (Figure 3e) against both SIBIU
and the longer CEU temperature series as well as the improved visual lower frequency trend agreement
with these instrumental records (Figure 3f) suggests that the CID-corrected chronology better
represented observed temperature trends. It should be pointed out that although CEU indicated
warmer temperature conditions before the mid-19th century than suggested even by the post-CID
chronology, early instrumental series (including those in central and eastern Europe) may contain a
positive warm bias as a result of measurement practices and the lack of screen use before the mid / late
19th century (Böhm et al., 2010; Moberg et al., 2003). Hence, it is unclear whether the post-CID chronology indices were still too low or the instrumental record contained an early period warm bias.

The correlation and RMSE change results (Figure 4) indicate that, with one exception, all chronologies showed some degree of improvement after CID correction. Specifically, nearly all chronologies exhibited improved agreement (i.e. greater similarity) with the reference SIBIU instrumental temperature series expressed by a correlation increase and reduced RMSE. However, CID correction may not necessarily produce a substantial degree of improvement in all cases. In some instances this may be a result of applying CID to chronologies that already expressed a strong climate signal and did not exhibit any considerable degree of disturbance related trends (e.g. PLOT18 in Figure 4). In other cases, where only very limited improvement was observed in weakly correlating chronologies with temperature (e.g. PLOT1 in Figure 4), other unidentified factors (not necessarily related to disturbance) are likely responsible. In general, however, CID correction resulted in climate signal improvement for RW data, which is true, not only for location-based sampling, but also the other sampling strategies (Table 2 and 3). Spatially, it appears that a high severity disturbance event around the 1840s and possibly others in the subsequent decades mainly affected the northern and western slopes, whereas another event around the 1910s mostly affected the eastern slope. We hypothesize that this distinct spatial pattern and segregation of areas affected by disturbance in these two cases may point to windstorms as the most likely disturbance agent and that the spatial configuration of this pattern may be indicative of the spatially distinctive impact of wind disturbance in these two instances.

The PC analysis (Figure 5) demonstrates that even extracting the dominant modes of variability as PC scores, will not separate the climatic and non-climatic signals (i.e. this approach does not guarantee best achievable results when the influence of disturbance is present). Though these are the results of a local-scale analysis, it is conceivable that temporally common disturbance trends can be present in chronologies even over a larger region (e.g. due to wind storms or large-scale insect outbreaks). The inability to isolate the climate signal was expressed by the significant correlation of both PC1 and PC3 with temperature, but to a lesser degree also through their correlation with the disturbance chronology, which was mainly represented by PC2. After CID correction, a clearer separation of the climatic and non-climatic signals was achieved with PC analysis as indicated by the
reduction from three dominant PCs to two and the increased correlation between PC2 and temperature. Importantly, however, though weaker (compared to pre-CID), the influence of the disturbance signal was reduced but not entirely removed by the CID procedure. This may be due to the relatively conservative threshold (3.29 sigma) applied in the identification of release events in order to minimize the likelihood of falsely identifying growth releases that are not disturbance related.

The parameter comparison for three sub-chronologies (PLOT 3, 7 and 10) in Figure 6 indicates that BI is not only the strongest temperature proxy but could potentially serve as a disturbance-free parameter, though further investigation in other locations and with additional species would be required to assess whether the decreased susceptibility of this parameter to disturbance is observed more generally. Kaczka and Czajka (2014) noted a similar (stronger than RW) summer temperature response of Norway spruce BI from Babia Góra in southern Poland. The importance of BI (and by extension maximum latewood density) to dendroclimatological research as a parameter that appears generally unaffected (or less affected) by disturbance and with a stronger climate signal is clearly emphasised by the evidence presented here. This may have implications for deriving chronologies free of disturbance with a stronger climatic signal as one possible way to by-pass the undesirable impact of disturbance on tree-ring data in dendroclimatic investigations. Furthermore, comparing RW and BI chronologies may represent an additional approach to the identification of disturbance trends in RW data.

A recent study by Rydval et al. (2016) demonstrated that disturbance related to anthropogenic activities (i.e. extensive logging) can induce growth trend biases in RW chronologies. The evidence presented herein demonstrates that natural disturbance can also potentially cause systematic chronology biases within closed-canopy forests. This can occur even if care is taken to select seemingly undisturbed sites as any evidence of disturbance occurring in the past (i.e. multiple decades or centuries ago) may have been erased from the landscape and may therefore no longer be visible at the time of sampling. By examining a very large number of samples, highly representative of the full stand population in this study, it is clear that the strength of the climate signal expressed in a chronology from a particular location can vary extensively and no sampling strategy can reliably ensure that the chronology produced from any set of collected RW samples will contain a well expressed climatic signal.
chronologies which express a sufficiently strong common population signal (i.e. assessed using the widely applied EPS metric) can result in chronologies poorly correlated with climate even when the relationship between climate and chronologies from other sets of samples from the same area is considerably stronger. This can arise when non-climatic trends occur synchronously in those samples that make up a chronology.

The presupposition that collecting a large number of samples and avoiding disturbance-affected sampling locations can alleviate disturbance related biases in chronologies may be misleading because large-scale disturbances can affect whole stands and presumably even many stands in a region (e.g. due to wind disturbance or large-scale insect outbreaks). Nehrbass-Ahles (2014) performed an evaluation of sampling strategies, although it mainly assessed chronologies based on various sampling techniques in relation to the ‘full population’ and was also conducted in a managed stand that did not in fact display much climatic sensitivity. Such an approach, however, implicitly assumes that the population itself is unbiased in relation to its representation of the climate signal. Here we have demonstrated that the assumption of an unbiased population may not be justified. Evaluating chronologies in relation to the population (or rather a very large sample of the population) may therefore not represent a sound strategy in some cases as the possible influence of disturbance should also be taken into account. This finding provides some support for adopting strategies such as the careful selection (or screening) of samples at the local site level, or chronologies on the multi-site network scale, by assessing their climatic sensitivity in order to avoid including samples or chronologies significantly affected by disturbance in dendroclimatic analyses. Such screening practices have already been commonly applied in the development of reconstructions from large scale networks (e.g. Cook et al., 2013; Ljungqvist et al., 2016). Nevertheless, the use of methods such as CID may be preferable as this can reduce the risk of potential subjectivity and perhaps even expand the range of useable chronologies which may otherwise be deemed unsuitable for dendroclimatic analysis.

Although this study demonstrates this issue only at a single location, there is potential for systematic disturbance to affect RW chronologies in virtually any closed canopy forest ecosystem and such a possibility cannot be dismissed a priori. The issues highlighted and discussed here may for
example directly affect calibration strength of reconstructions as well as the possibility of making
inaccurate inferences about past climatic conditions from RW-based reconstructions that may include
disturbance related biases. It is important to be able to perform some assessment of possible
disturbance effects on RW chronologies because assessing the fidelity of reconstructed climate
estimates before the instrumental period is difficult. We therefore recommend that all future
dendrochronological studies investigating medium to low frequency climatic trends should perform
some form of disturbance assessment and that the CID method (Druckenbrod et al., 2013; Rydval et al.,
2016) represents a reasonable approach.

CONCLUSION

In this study, we have demonstrated that natural disturbance events can act as agents which
significantly and systematically affect tree growth, subsequently biasing mid- to long-term RW
chronology trends. These disturbance trends cannot be removed using conventional detrending
approaches without also removing lower frequency climatic information. In closed canopy forests, the
oldest (and dendroclimatologically most valuable) trees are more likely to contain an embedded
disturbance response. It is not possible to ensure that this response can be factored out or minimized
simply by adopting a subjective sampling strategy or relying on a very large sample size (with respect to
both trees and sites). Furthermore, sampling trees across a landscape may produce a record with a
complex range of disturbance histories rather than reducing the disturbance signals. This important
finding highlights the need to develop site selection and sampling approaches for closed-canopy forests
that are very different from those developed by Fritts (1976) for open-canopy forests. More specifically,
it is imperative to develop better methods to disentangle disturbance and climate signals.

Disturbance detection techniques could be used, at a minimum, to identify and assess the
effects of disturbance on RW chronologies and (if replication permits) exclude subsets substantially
affected by such trends which would therefore represent a poorer expression of longer term climatic
variability. This also provides justification for the application of approaches such as data screening in
order to exclude subsets of larger datasets, which are weakly correlated with climate, from climatic
analyses. An alternative approach would include the utilization of some sort of disturbance correction
procedure (e.g. CID) to improve the expression of the climate signal in disturbance affected RW series.

Finally, other tree-ring parameters, such as BI (or maximum latewood density), which may be less prone
to the effects of disturbance and often express a stronger climate signal than RW (Wilson et al. 2016),
could also be developed.

The findings of this study are broadly applicable and of relevance to RW chronologies from
closed canopy stands. Additional larger scale investigations including various species from other
locations would be beneficial in assessing the relevance of our findings. Certainly, consideration should
be given to the possibility of disturbance related trends affecting medium to low frequency growth
trends in RW chronologies. We therefore recommend that some form of evaluation of this potential
effect should be performed as part of any dendrochronological research utilizing RW data to investigate
climatic trends as it may be possible to reduce this limitation and improve the expression of the climate
signal in such data.

ACKNOWLEDGEMENTS

The study was supported by the institutional project MSMT (CZ.02.1.01/0.0/0.0/16_019/0000803) and
the Czech Ministry of Education (Project INTER-COST no. LCT17055). We thank the Călimani National
Park authorities, especially E. Cenuşă and local foresters, for administrative support and assistance in
the field.

REFERENCES

temperatures from tree rings: Part II, spatially resolved reconstructions. *Quaternary Science
Reviews* 163: 1-22.

Björklund JA, Gunnarson BE, Seftigen K et al. (2014a) Blue intensity and density from northern Fennoscandian tree rings, exploring the potential to improve summer temperature reconstructions with earlywood information. *Climate of the Past* 10(2): 877-885.

Björklund J, Gunnarson BE, Seftigen K et al. (2014b) Using adjusted Blue Intensity data to attain high-quality summer temperature information: A case study from Central Scandinavia. *Holocene* 25(3): 547-556.


Cook ER, Krusic PJ, Anchukaitis KJ et al. (2013) Tree-ring reconstructed summer temperature anomalies for temperate East Asia since 800 CE. Climate Dynamics 41(11-12): 2957-2972.


Figure 1: Site location and approximate distribution of sampling plots in Calimani National Park, Romania. Red shading represents post-disturbance correction correlation increase* of plot-based (PLOT) chronologies (see Table 1 for details) against June-July mean instrumental temperatures for the full length of chronologies – same representation as Figure 4C. (*note that chronology PLOT-16 shows a slight correlation decrease after disturbance correction).
Figure 2: Chronology plots and correlations with Jun-Jul mean temperatures from Sibiu for four 'sampling' methods including grouping according to (A) sample location (PLOT), (B) random sample selection (RAN), (C) diameter at breast height (DBH) (D) and recruitment age (AGE) – see Table 1 for additional details. Each chronology was truncated in the year where expressed population signal dropped below 0.85. (pre-CID indicates that chronologies were developed with series before correcting for disturbance trends using the Curve Intervention Detection method).
Figure 3: (A) Summary of Calimani disturbance history from Curve Intervention Detection (CID) analysis; (B) chronologies before disturbance correction (pre-CID) and (C) after disturbance correction (post-CID) and (D) pre-CID/post-CID chronologies with Jun-Jul temperatures from Sibiu (SIBIU) and the longer central/east Europe (CEU) Jun-Jul regional composite temperature series; (E) 51-year running correlations between instrumental and ring-width chronologies in (D); (F) as in (D) except smoothed with a 20 year low-pass Gaussian filter.
Figure 4: Comparing the (A, C) change in correlation and (B, D) root-mean-square error change of Calimani plot-based (PLOT) chronologies before disturbance correction (pre-CID) vs. after disturbance correction (post-CID) in relation to instrumental temperature data from Sibiu for the (A, B) 1909-2009 period and (C, D) full chronology length (max. back to 1851). (The green colour in A and C indicates the size of the correlation increase after disturbance correction whereas red colour (only PLOT16) indicates a correlation decrease.)
Figure 5: Amplitudes of the dominant principal components (PCs) from chronologies developed (A) before disturbance correction (pre-CID) and (B) after disturbance correction (post-CID), and their correlation with instrumental temperatures from Sibiu and the disturbance chronology in Figure 3A. (Scatterplots of significant relationships (p < 0.01) between the PCs and the disturbance chronology / Sibiu temperatures are represented in supplementary figure S5).
Figure 6: Comparison of the PLOT-3, PLOT-7 and PLOT-10 blue intensity (BI) and ring width (RW) chronologies developed before (pre-CID) and after (post-CID) disturbance correction with instrumental temperatures from Sibiu (SIBIU) over the 1851-2009 period showing (A) the combined correlation response of the Calimani chronologies against SIBIU temperatures; and the time-series of the RW and BI chronologies together with Jun-Jul and Apr-Sep SIBIU mean temperature respectively for (B) PLOT-3, (C) PLOT-7 and (D) PLOT-10. (Highlighted periods indicate where expressed population signal is < 0.85.)
**Supplementary figures**

**Figure S1:** Correlation response of chronologies composed of all 760 series from Calimani (PLOT-all) before (pre-CID) and after (post-CID) disturbance correction against mean monthly and seasonal temperatures from Sibiu for the 1851-2011 period.

**Figure S2:** Relationship between estimated tree age and diameter at breast height (DBH).
Figure S3: Summary of Calimani disturbance history from Curve Intervention Detection (CID) analysis (A) using series from plots from the northwest part of the stand (Figure 1 – PLOT 1, 2, 5, 6, 7, 10, 11, 18, 19) predominantly affected by disturbance in the mid-19th century and (B) from the southeast part for the stand (Figure 1 – PLOT 3, 4, 8, 9, 12, 13, 14, 15, 16, 17) predominantly affected by disturbance in the early 20th century. The before disturbance correction (pre-CID) and after disturbance correction (post-CID) chronologies of these two spatial groups are presented in (C) and (D) respectively.
Figure S4: Before disturbance correction (pre-CID) and after disturbance correction (post-CID) chronology plots and correlations with Jun-Jul mean temperatures from Sibiu for four ‘sampling’ methods including grouping according to (A) sample location (PLOT), (B) random sample selection (RAN), (C) diameter at breast height (DBH) (D) and recruitment age (AGE) – see Table 1 for additional details. Each chronology was truncated in the year where expressed population signal dropped below 0.85.
Figure S4 – continued

B

random chronologies (pre-CID)

random chronologies (post-CID)

1912-2009 (pre-CID)

1912-2009 (post-CID)
Figure S4 – continued

C  DBH-based chronologies (pre-CiD)

DBH-based chronologies (post-CiD)

smallest  1917-2009 (pre-CiD)  largest

smallest  1917-2009 (post-CiD)  largest

chronology
**Figure S4** – continued

![Graphs showing age-based chronologies (pre-CID) and (post-CID)](image)

- **Graph A:** Ring-width index vs. year for age-based chronologies (pre-CID).
- **Graph B:** Ring-width index vs. year for age-based chronologies (post-CID).

**Histograms showing correlation**

- **Graph C:** Correlation between different age groups and chronologies (pre-CID) for the years 1933-2009.
- **Graph D:** Correlation between different age groups and chronologies (post-CID) for the years 1933-2009.

The histograms display the correlation coefficients between the youngest and oldest age groups for each chronology.
Figure S5: Scatterplots of significant relationships (p < 0.01) between the disturbance chronology in Figure 3A / Jun-July average Sibiu temperatures and amplitudes of the dominant principal components (PCs) from chronologies developed before disturbance correction (pre-CID) and after disturbance correction (post-CID).
Curve Intervention Detection (CID) Matlab code

function \([ymn, varyh, df, w, ybar, se]=bisqmean_CID(y)\)

% Biweight mean for a vector of numbers.
% Last revised 2011-7-09
% Revised by Daniel Druckenbrod 2012-1-11
% Source: Mosteller and Tukey (1977, p. 205, p 351-352)
% Cook and Kairiukstis (1990, p. 125-126)
%****************  INPUT *************************
% y (? x 1)r vector of data -- say, indices for ? cores in a year
%********************  OUTPUT ************************
% ymn (1 x 1)r biweight mean
% varyh (1 x 1)r asymptotic standard dev of biweight mean - p. 208,
% df (1x1)r degrees of freedom
% w (? x 1)r final weights on values in y
% ybar (1 x 1)r arithmetic mean corresponding to ymn
% se (1 x 1)r standard error of ybar
%****************  NOTES ********************
% ybar and se just included in debugging to double check
% on closeness of ybar to ymn, se to sqrt(varyh)

sens = 0.001; % hard coded threshold of sensitivity for stopping
iterat
nits = 100; % max number of allowed iterations
[n,ny]=size(y);
if ny > 1;
   error('y should be a vector')
end
if any(isnan(y));
   error('y not permitted to have NaNs');
end;
if n<6; % if fewer than 6 sample size, use median
   ymn = median(y);
   w=[];
ybar=mean(y);
   se = sqrt(var(y)/n); % standard error of mean
df=[];
   varyh=NaN;
   return;
end;
ww = 1/n; % weight for even average
ybar = mean(y); % arith mean
%ybar=median(y);
se = sqrt(var(y)/n); % standard error of mean

nz=0;
ymn = ybar; % initial biweight mean as arith mean

for i = 1: nits;  % iterate max of nits times
    ymnnold = ymn;  % store old value of mean
    e = y-ymn; % deviations from mean
    S = median(abs(e));  % median abs deviation
    u = e / (6*S);  % scaled deviations
    w = (1 - u.^2).^2;  % compute weights
    L1 = abs(u)>=1; % flag huge errors
    L1s = sum(L1);
    if L1s>0
        nz=0;
        nz= nz(ones(L1s,1),:);
        w(L1)=nz;  % set weights on those obs to zero
        end
    w = w / sum(w); % adjust weights to sum to 1.0
    ymn = sum(w .* y); % compute biweight mean

    % Variance of estimate of biweight mean
    ui= e / (9*S);
    L2 = ui>1;
    ui(L2)=[];
    z =y(~L2);
    nz = length(z);
    nom1 = (z - ymn) .^2;
    nom2 = (1-ui .^2) .^4;
    nom = sum(nom1 .* nom2);
    den1 = sum((1-ui .^2) .* (1-5*ui .^2));
    varyh_hoaglin=(n^0.5)*(nom^0.5)/den1; % Dan: p. 417 3rd equation
    den2 = -1 + sum ((1-ui .^2) .* (1-5*ui .^2));
    % varyh = nom / (den1*den2); % variance of biweight mean
    % last eqn, p. 208
    varyh = n^.5*nom^.5 / ((den1*den2) .^ .5); % Dan: p.417 Kafadar approach
    df = 0.7 * (nz -1); % degrees of freedom

    % if little change in mean, exit loop
    if abs (ymn - ymnnold) < sens
        return
    end
end
% hugershoff.m
% This function fits a tree ring time series to the growth trend equation
% developed by Warren (1980) TRR.
% Function written Nov 12, 2013.
% Function last revised Nov 12, 2013.

function qq = hugershoff(beta,x)

% Assign parameters from beta vector.
 a=beta(1);
b=beta(2);
c=beta(3);
k=beta(4);
qq=a*((x).^b).*exp(-c*x)+k;
%qq=a*((x+1).^b).*exp(-c*(x+1))+k;
%q=log(a)+b*log(x)%-c*x;
%qq=exp("...")
function qq = nonlinear_exp(beta,x)

% Assign parameters from beta vector.
a = beta(1);
b = beta(2);
d = beta(3);

qq = a*exp(-b*x)+d;
function [col_header,rings]=ringwidth_import_999(header,varargin)
if nargin==1
     [filename,path]=uigetfile('*.txt','Select ".txt" file');
elseif nargin==2
     filename=varargin;
else
     disp('Too many parameters entered (DLD).')
end
% Read in header lines
headers=textread(filename, '%q',10)';
% disp([headers]) % Display 1st 10 words as screen output.
[label yr y0 y1 y2 y3 y4 y5 y6 y7 y8 y9]=...
textread(filename,...
'8s 4d 5d 5d 5d 5d 5d 5d 5d 5d 5d 5d',...
'headerlines',header);
% Place decadal format widths into one matrix
widths=[y0 y1 y2 y3 y4 y5 y6 y7 y8 y9];

% Extract unique labels of each core.
importedrows=length(label);
a=1;b=1;
while(a<=importedrows)
     core(b)=label(a);
     corestr(:,b)=strcmp(label(a),label);
     a=max(find(corestr(:,b)==1))+1;
     b=b+1;
end
% Find range of years over all cores and set as col 1 in rings.
% As it is difficult to know how many years are in the last row
% of measurements for a core, assume that the last decade has 10
% measurements.
rings=(min(yr):(max(yr)+10))';

% Transfer widths into vectors by core
for i=1:length(core)
     core_rows=find(corestr(:,i));
     core_yr=yr(core_rows);
     core_widths=widths(core_rows,:);
     % Find # of measurements in a row and assign to vector series.
     k=1;series=0;
     for j=1:length(core_rows)
         % Look for end of series flag
         flag=find(core_widths(j,:)==999);
         if flag>0
             msmts=flag-1;
     end
     series=series+msmts;
     rings(k)=series+10*(k-1);
     k=k+1;
end

elseif (ceil(core_yr(j)/10)*10)-core_yr(j)==0
    msmts=10;
else
    msmts=(ceil(core_yr(j)/10)*10)-core_yr(j);
end
series(k:(k+msmts-1))=core_widths(j,(1:msmts));
k=msmts+k;
end

% Determine start and end of series
sos=find(rings(:,1)==min(core_yr));
length(series)+sos-1;
eos=length(series)+sos-1;
% Assign series to output matrix and convert to 1/1000th of a mm
rings(sos:eos,i+1)=(series./100)';

% Remove 999 from end of series
for j=1:length(rings(1,:))
    for k=1:length(rings(:,1))
        if rings(k,j)==-99.99
            rings(k,j)=0;
        end
    end
end

% Construct column headers
col_header(2:(length(core)+1))=core;
col_header(1)={'Year'};
% This function imports decadal format tree ring data for manipulation
% as a matrix in Matlab. The number of header lines must be specified
% as an input by the user. The end of each series must be flagged
% with -9999. Measurements are stored as one thousandth of a
% millimeter. The filename can either be specified as an input or
% found using a gui. The LAST LINE of the input text file must also
% be blank!

% Function last revised May 24, 2012.

function [col_header,rings,flag]=ringwidth_import_9999(header,varargin)

if nargin==1
    [filename,path]=uigetfile('*.txt','Select *.txt file');
elseif nargin==2
    filename=varargin;
else
    disp('Too many parameters entered (DLD).')
end

% Read in header lines
headers=textread(filename, '%q',10)';
% disp([headers]) % Display 1st 10 words as screen output.
[label yr y0 y1 y2 y3 y4 y5 y6 y7 y8 y9]=textread(filename,...
    '%8s %4d %5d %5d %5d %5d %5d %5d %5d %5d %5d %5d',...
    'headerlines',header);
% Place decadal format widths into one matrix
widths=[y0 y1 y2 y3 y4 y5 y6 y7 y8 y9];

% Extract unique labels of each core.
importedrows=length(label);
a=1;b=1;
while(a<=importedrows)
core(b)=label(a);
corestr(:,b)=strcmp(label(a),label);
a=max(find(corestr(:,b)==1))+1;
b=b+1;
end

% Find range of years over all cores and set as col 1 in rings.
% As it is difficult to know how many years are in the last row
% of measurements for a core, assume that the last decade has 10
% measurements.
rings=(min(yr):(max(yr)+10))';

% Transfer widths into vectors by core
for i=1:length(core)
core_rows=find(corestr(:,i));
core_yr=yr(core_rows);
core_widths=widths(core_rows,:);
% Find # of measurements in a row and assign to vector series.
k=1;series=0;
for j=1:length(core_rows)
    % Look for end of series flag
    flag=find(core_widths(j,:)==-9999);
    if flag>0
        msmts=flag-1;
    end
end

function [col_header,rings,flag]=ringwidth_import_9999(header,varargin)
elseif (ceil(core_yr(j)/10)*10)-core_yr(j)==0
    msmts=10;
else
    msmts=(ceil(core_yr(j)/10)*10)-core_yr(j);
end

series(k:(k+msmts-1))=core_widths(j,(1:msmts));
k=msmts+k;
end

% Determine start and end of series
sos=find(rings(:,1)==min(core_yr));
length(series)+sos-1;
eos=length(series)+sos-1;
% Assign series to output matrix and convert to 1/1000th of a mm
rings(sos:eos,i+1)=(series./1000)';
end

% Construct column headers
col_header(2:(length(core)+1))=core;
col_header(1)={'Year'};
% This function extracts a single tree-ring time series from
% ringwidth_import.m and places it in a vector for time series
% analysis.
% The 'filename' used to load tree-ring data for processing should
% refer to
% the file containing data imported using the function
% ringwidth_import.m.
% Following the approach used in ARSTAN (Ed Cook, Columbia
% University),
% the function power transforms and removes the mean to create
% transformed
% residuals. The function then detrends with an iterative neg.
% exponential
% fit, or if that does not fit or fails to find a solution, then a
% linear
% regression with either a positive or negative slope is fit to the
% data.
% Using the maximum entropy model solution otherwise known as the Burg
% method, the autoregressive model that is the best fit for the series
% is
% determined. Using the best fit model, the function searches for
% autoregressive outliers iteratively. These outliers may either be
% pulse
% events (1 yr) or CSTs (> minimum no. of yrs). After the first pass,
% the outliers are removed and the series is reconstituted. The best
% order is then redetermined and the function searches for additional
% outliers. The # of iterations is set by the user (8 should be
% enough).
% This version uses a power transformation to minimize
% the heteroscedastic nature of my time series. 'fig' is a flag that
% specifies whether you want a figure (=1) or not (=0). Missing years
% are
% set to the average of neighboring rings. The central limit theorem
% is used to search the residuals for trend outliers. This version
% also
% uses David Meko's (University of Arizona) biweight mean code and
% currently runs with a window of 9 to 30 yrs. Estimated values for
% missing rings are removed in the output series. This version uses
% a
% modified Hugershoff curve with a potentially nonzero asymptote to
% detrend + and - disturbance events. It also returns the transformed
% standardized series.
% Function written Sep 10, 2002.
% Function last revised Jun 6, 2014.

function [YEARS,transformed,detrended,St,Str,Dtr,Atr,age,outs]=...
    v105pn(core,fig,iter)
    global PARAM; PARAM=0; % vector of parameters for best order AR model.
    global ORDER; ORDER=0; % best order of AR model determined by AIC.
    global YEARS; YEARS=0; % calendar years of tree growth from datafile
    % Load tree-ring data (returns vars *col_header* and *rings*)
    load filename.mat %Insert filename here
    % Find pointer to start and end of series
    sos=find(rings(:,(core+1))>0, 1);
    eos=find(rings(:,(core+1))>0, 1, 'last');
% Assign years and raw widths to respective vectors.
YEARS=rings(sos: eos, 1);
raw=rings(sos: eos, core+1);
disp(['Core: ' char(col_header(core+1)))
nyrs=length(YEARS);
disp(['Total no. of measured years: ' int2str(nyrs)])
disp(['First year is ' num2str(YEARS(1))])
disp(['Last year is ' num2str(YEARS(nyrs))])

% Estimate missing ring widths using mean of neighboring rings
mss=NaN(length(raw), 1);
if find(raw==0)
m1=find(raw==0);
disp(['Missing rings at years ' num2str([YEARS(m1)])])
for nm=1:length(m1)
prior=mean(raw(find(raw(1:m1(nm)), 1, 'last')));
subs=mean(raw(find(raw(m1(nm):length(raw)), 1, 'first')+m1(nm)-1));
mss(m1(nm))=mean([prior subs]);
end
raw=nansum([raw mss], 2);
end

% Power transformation.
diff=0;
for x=1:(length(YEARS)-1) % Calculate 1st differences
diff(x, 1)=raw(x+1);
diff(x, 2)=abs(raw(x+1)-raw(x));
end
s=1;
for q=1:(length(YEARS)-1)
if (diff(q, 1)==0) && (diff(q, 2)==0)
nz_diff(s,:)=diff(q, 1:2); % non-zero ring widths
s=s+1;
end
log_diff=[log(nz_diff(:, 1)) log(nz_diff(:, 2))];
X=[ones(length(log_diff(:, 1)), 1) log_diff(:, 1)];
bb = regress(log_diff(:, 2), X);
optimal_line = bb(2)*log_diff(:, 1)+bb(1);
optimal_pwr = 1-bb(2);
disp(['Optimal Power = ' num2str(optimal_pwr)])
if optimal_pwr <= 0.05
transformed=log10(raw);
tzero=log10(0.001);
disp('Series was log10 transformed')
elseif optimal_pwr>1
optimal_pwr=1;
transformed=(raw.^(optimal_pwr));
tzero=0.001.^(optimal_pwr);
disp('Series was power transformed with power = 1')
else
transformed=(raw.^(optimal_pwr));
disp(['Series was power transformed with power = ' ... num2str(optimal_pwr)])
tzero=0.001.^optimal_pwr;
end
transm=mean(transformed);

% Nonlinear detrending option.
% Function nlinfit employs nonlinear least squares data fitting by the
% Gauss-Newton Method.
crashed=zeros(nyrs,1);
wlnth=zeros(nyrs,1);
trendtype=0; % Neg exp = 1, neg linear reg = 2, or pos linear reg = 3
minyr=30; % minimum # of yrs to fit to nlinfit
if minyr>nyrs
    disp('Insufficient # of years to fit minimum nonlinear age
trend.')
end
b=zeros(nyrs,3);
mse=NaN(nyrs,1);
warning off
for i=minyr:nyrs
    try
        lastwarn('')
        beta = [.5 .1 1];
        xyrs = 1:i; % set years from 1 to length of series
        [b(i,1:3),~,~,~,mse(i)]=nlinfit(...
            xyrs(1:i),transformed(1:i),'nonlinear_exp',beta);
        crashed(i)=1;
        msgstr = lastwarn;
        wlnth(i)=length(msgstr);
        catch % Stops code from crashing because of problems fitting exp
curve
            crashed(i)=2;
    end
end
warning on
i_c=0;

% Dissallow curve to be concave up and make sure nlinfit
% converges by making b(2) sufficiently large.
% constant b(3) must be >=0 in original mm
i_c=find(crashed==1 & b(:,1)>=0 & b(:,2)>0.001 & b(:,3)>=tzero &
    wlnth==0) & b(:,2)<0.5);
[mmse,imse]=min(mse(i_c));
if fig==1 % fig=1 if you want a figure as output
    figure('Position', [10 150 600 600])
    subplot(3,1,1)
    plot(YEARS,raw,'k','LineWidth',2)
    ylabel('\bf Ring width (mm)')
    fig1atext = {'Optimal power = ', num2str(optimal_pwr,4)};
    text(range(YEARS)/3+YEARS(1), max(raw)/1.2,fig1atext)
end
if i_c(imse)>0
    disp(['Lowest error from fit = ' num2str(mmse)])
    disp(['Best age trend fit from years ' num2str(YEARS(1)) ' to '
    ...
        num2str(YEARS(i_c(imse))))])
    disp(['Best fit extends for ' num2str(i_c(imse)) ' years'])
    best=b(i_c(imse),:);
end
trendtype=1;
y_exp=nonlinear_exp(best,xyrs);
detrended=transformed-y_exp;
disp('Initial Age Detrending')
disp(['Y = ', num2str(best(1),4), ' *exp(-', num2str(best(2),4),...
'*x)+', num2str(best(3),4)]);
if fig==1 % fig=1 if you want a figure as output
    subplot(3,1,2)
    [h312a, h312h1, h312h2] = plotyy(YEARS,[transformed
    y_exp'],YEARS(i_c),mse(i_c));
    set(h312h1(1),'LineWidth',2)
    set(h312h1(1),'Color',[0 0 0])
    set(h312h1(2),'Color',[.2 .2 1])
end
set(h312h2(1),'LineStyle','none', 'Marker','.','MarkerFaceColor',[1 .2
1.2])
    fig1btext = {{'Y = ', num2str(best(1),4), ' *exp(-',
num2str(best(2),4),'. .
'*x)+', num2str(best(3),4)]};
text(range(YEARS)/3+YEARS(1), max(transformed)/1.2,fig1btext)
    line([YEARS(i_c(imse)) YEARS(i_c(imse))],...
[y_exp(i_c(imse))+2]'Color','k','Marker','v','MarkerEdgeColor',[1 .2 .2]
'.MarkerFaceColor',[1 .2 .2])
set(get(h312a(1),'Ylabel'),'String',\bf Transformed width')
set(get(h312a(2),'Ylabel'),'String',\bf Error Term Variance')
subplot(3,1,3)
plot(YEARS,detrended,'k','LineWidth',2)
ylabel('\bf Transformed width')
xlabel('\bf Year')
end
else
    trendtype=2;
    xyrs=(1:nyrs)';
    % Linear detrending option used if neg. exponential curve dissallwd.
    [b,~,~,stats] = regress(transformed,...
    [ones(length(YEARS),1) xyrs]);
    if b(2)>=0; trendtype=3; end % Find positive age trends
    y_lin=b(2)*xyrs +b(1);
    detrended=transformed-y_lin;
disp('Initial Age Detrending')
disp(['Y = ', num2str(b(2)), ' * X + ', num2str(b(1))]);
if fig==1 % fig=1 if you want a figure as output
    subplot(3,1,2)
    h312b=plot(YEARS,transformed,'k',YEARS,y_lin,'k--');
    set(h312b(1),'LineWidth',2)
    fig1btext = {{'Y = ', num2str(b(2)), ' * X + ',
num2str(b(1))})
    text(range(YEARS)/3+YEARS(1), max(transformed)/1.2,fig1btext)
    ylabel('\bf Transformed width')
    subplot(3,1,3)
    plot(YEARS,detrended,'k','LineWidth',2)
ylabel('\bf Transformed width')
xlabel('\bf Year')
end
end
% Output age detrending info
age={char(col_header(core+1)); trendtype; YEARS(i_c(imse))};
Initial arrays.
next_iter=1; % Switch to determine whether next iteration is needed
St=detrended; % St will be the iterated series (standardized)
Atr=NaN(length(raw),1); % Age trend re-expressed in raw units
rline=NaN(length(raw),1); % Just the slope of the intervention
tline=NaN(length(raw),iter); % Slope and constant of the intervention
outs=zeros(iter,5);

for g=1:iter % Iterate AR model 'iter' times to remove all outliers
  if next_iter==1
    bckcasted=0;ar_estimate=0;residuals=0;area_t=0;
    iter_i=St; % Initial values of series for ith iteration.
    disp(' ');
    disp(['Statistics for AR model iteration ' int2str(q) ':']);

    % Calculate best AR model order and return in the following
    % order:
    % residuals (white noise) and ar model estimates
    [ar_white, ar_model]=ar_order(St);

    % Use new coefficients to prewhiten ORIGINAL series without
    % downweighted originals.

    % Backcast for pth order years of AR model.
    bckcasted=backcast(detrended);
    bckcasted=backcast(St);

    for g=ORDER:length(bckcasted) % g = observation year
      ar=0; % ar model estimate for order i, year g
      for k=1:ORDER % kth parameter of order ORDER
        if (g-ORDER)>0 % ensure obs yr > model order
          ar=PARAM(k)*(bckcasted(g-k))+ar;
        end
      end
      if g-ORDER>0 % calculate model estimate and residuals
        if detrended(g-ORDER)==0 % Set missing rings to ar
         disp(['Missing ring at year: ' int2str(YEARS(g-ORDER))])
        end
        ar_estimate(g-ORDER)=ar;
        residuals(g-ORDER)=(bckcasted(g)-ar);
      end
    end
    ar_estimate=ar_estimate';
    residuals=residuals';
  if fig==1 % fig=1 if you want a figure as output
    figure('Position', [600 150 600 600])
    subplot(4,1,1)
    h411=plot(YEARS,St,'k',YEARS,ar_estimate,'k-.');
    set(h411(1),'LineWidth',2)
    title(['\bf ' char(col_header(core+1)) ' iteration ' int2str(q)])
    ylabel('\bf Trans. width')
    xlabel('\bf Year')
end
axis([min(YEARS) max(YEARS) min(St)*.9 max(St)*1.1])
end

% Find release outliers
[dowres, mres, otype]=outlier_clt(residuals, fig);
f=find(dowres==0);
if otype==1 && ~isempty(f) % Pulse Outlier Detected
    St(f)=ar_estimate(f);
elseif otype>1 && length(f)>1 % Trend Outlier Detected
    w=ones(length(f),1) (1:length(f))';
    slope=regress(St(f),w);
disp(['Constant and slope = ' num2str([slope(1)
    slope(2)])])
end

% Fit Hugershoff curve to remainder of series
lngthw=min(f):length(St);
lngthwf=(max(f)+1):length(St);
lngthn=1:length(lngthw);
lngthn=lngthn(:);
opts = statset('nlinfit');
opts.FunValCheck = 'off';
opts.MaxIter = 400;
bw=nlinfit(lngthn,St(lngthw),'hugershoff', [.1 .5 .1
    .1],opts);
disp(['Hugershoff Parameters: ' num2str(bw)])
ar_est=ar_estimate(f(1));
rline(lngthw)=-bw(1)*(lngthn.^bw(2)).*exp(-bw(3)*lngthn)-
    bw(4);

% If nlinfit returns NaN, then try again with different
initial parameters.
if find(isnan(rline(lngthw)))>0; rline(lngthw)=0;
disp('Default initial parameters for Hugershoff curve failed')
disp('Fitting alternate, robust initial parameters [.1
    .5 .1 .1]')
    opts.RobustWgtFun = 'bisquare'
    bw=nlinfit(lngthn,St(lngthw),'hugershoff', [.1 .5 .1
    .1],opts);
disp(['Hugershoff Parameters: ' num2str(bw)])
ar_est=ar_estimate(f(1));
rline(lngthw)=-bw(1)*(lngthn.^bw(2)).*exp(-
    bw(3)*lngthn)-bw(4);
end

% If nlinfit returns NaN, then end outlier iterations and quit.
if find(isnan(rline(lngthw)))>0; rline(lngthw)=0;
disp('Unable to fit Hugershoff curve')
ar_est=0;
next_iter=0;
outs(q,1:5)=[0 0 0 0 0];
end

if f(1)>1
    St(lngthw)=rline(lngthw)+St(lngthw)+ar_est;
tline(lngthw,q)=-rline(lngthw);
elseif f(1)==1 % If trend occurs in 1st yr of series
    St(lngthw)=rline(lngthw)+St(lngthw);
tline(lngthw,q)=-rline(lngthw);
end
outs(q,1:5)=[YEARS(min(f)) YEARS(max(f)) slope(1) slope(2) otype];

if isempty(f) % Determine whether any outliers...
    next_iter=0; % were detected on this iteration
end

if q==iter && ~isempty(f)
    disp('Need to run additional iterations to resolve series!')
end

if fig==1 % fig=1 if you want a figure as output
    subplot(4,1,4)
    hold on
    if ~isempty(f) && min(f)>1 % Draw detrended regression
        line
        line([YEARS(min(f)) YEARS(max(f))], [ar_est ar_est],...
             'Color', [.6 .6 .6], 'LineStyle', '--', 'LineWidth', 2)
    else % Draw same line, but set to first year of series
        line([YEARS(1) YEARS(max(f))], [0 0],...
             'Color', [.6 .6 .6], 'LineStyle', '--', 'LineWidth', 2)
    end
    h414=plot(YEARS,iter_i,'k',YEARS,St,'k',YEARS,tline(:,q),'k');
    set(h414(1),'LineWidth',2)
    set(h414(3),'LineWidth',2,'Color', [.6 .6 .6])
    ylabel('f Trans. width')
    xlabel('f Year')
    ymin=min([min(iter_i) min(St)])*.9;
    ymax=max([max(iter_i) max(St)])*1.1;
    axis([min(YEARS) max(YEARS) ymin ymax])
    box on
    hold off
end

subplot(4,1,2)

h412=plot(YEARS,residuals,'k-.',YEARS,zeros(1,length(YEARS)),'k',...
           YEARS,mres);
set(h412(3),'Color', [.6 .6 .6])
set(h412(3), 'LineWidth', 2)
axis([min(YEARS) max(YEARS) min(residuals)*1.1 max(residuals)*1.1])
ylabel('f Residuals')
xlabel('f Year')

end

if fig==1 % fig=1 if you want a figure as output
    figure('Position', [1200 150 600 400])
    subplot(2,1,1)
    transDt=detrended-St; % transformed outlier series
    hpentult=plot(YEARS,transDt,'k',YEARS,St,'k--');
    set(hpentult(1), 'LineWidth', 2)
    set(hpentult(2), 'LineWidth', 2)
end

56
% Shows final iterated series in original units (mm presumably)
if trendtype==1 % negative exponential trend
    Stt=y_exp'*St; % Size trend & first detrending
    if optimal_pwr <= 0.05
        Str=10.^(Stt);% Size trend in original (raw) units
        Atr=10.^(y_exp');% Age trend in original (raw) units
    else
        Stt(Stt<=0)=0; % Set neg values to zero
        Str=(Stt).^(1/optimal_pwr);
        Atr=(y_exp').^(1/optimal_pwr);
    end
elseif trendtype==2 || trendtype==3 % linear regression trend
    Stt=y_lin+St; % Size trend & first detrending
    if optimal_pwr <= 0.05
        Str=10.^(Stt);% Size trend in original (raw) units
        Atr=10.^(y_lin);% Age trend in original (raw) units
    else
        Stt(Stt<=0)=0; % Set neg values to zero
        Str=(Stt).^(1/optimal_pwr);
        Atr=(y_lin).^(1/optimal_pwr);
    end
else
    disp('Error in trend type designation')
end

raw(mss>0)=NaN; % Remove estimated values of missing rings
Str(mss>0)=NaN; % Remove estimated values of missing rings
Dtr=raw-Str; % Remove estimated values of missing rings

if fig==1 % fig=1 if you want a figure as output
    subplot(2,1,2)
    h_end=plot(YEARS,Dtr,YEARS,Str,'k--',YEARS,raw,'b');
    set(h_end(1),'Color',[.6 .6 .6])
    set(h_end(1),'LineWidth',2)
    set(h_end(2),'LineWidth',2)
    set(h_end(3),'LineWidth',2)
    legend('Disturbance index','Standardized series','Original series',...'
    Location','NorthWest')
    legend('boxoff')
    ylabel('f Ring width (mm)')
    xlabel('f Year')
end

% ar_order.m
% This subfunction is based on series_ar.m and determines the autoregressive parameters for the best model order as calculated using AIC criteria. The function returns the residuals, and AR model estimate of the best order found with AIC criteria.
function [out_res, out_est]=ar_order(series)
    global PARAM
global ORDER
global YEARS

% Initializes variables for autoregressive modeling.
ar_param=0; residuals=0;

% Calculate Autoregressive parameters for orders 1 through 10.
for ar_order=1:10
    ar_param(ar_order,1:ar_order+1)=-arburg(series,ar_order);
end

%Remove first column of minus ones from ar_param.
ar_param(:,1)=[ ];

% Calculate residuals for particular AR order model.
for i=1:10 % i = ar model order
    for g=1:length(YEARS) % g = observation year
        ar=0; % ar model estimate for order i, year g
        for k=1:length(ar_param(1,:)) % kth parameter of order i
            if (g-k)>0 % ensure obs yr > model order
                ar=ar_param(i,k)*(series(g-k))+ar;
            end
        end
        if g-i>0 % calculate residuals
            residuals(g,i)=(series(g)-ar);
            ar_estimate(g,i)=ar;
        end
    end
end

% Calculate the total variance of the residuals by model order
resid_var(i)=var(residuals(:,i));

% Calculate variance of the residuals of a particular AR order model.
% Reference Box & Jenkins & Reinsel 1994 pp. 200-201.
% Using Akaike Information Criteria
% Equation now uses natural log and simply 'n' in the denominator.
% t+1 (or # of params+1) is a penalty factor for estimating the mean.
aic=0;
for t=1:length(resid_var)
aic(t)=log(resid_var(t))+(2*(t+1))/length(YEARS);
end

% Find the first minimum AIC order (ie first saw-tooth-shaped dip).
% If AIC values monotonically decrease, set best_order=9.
best_order=0;
for s=2:length(aic)
    if((aic(s)>aic(s-1)) && (best_order==0))
        best_order=s-1;
    end
end

if best_order==0;
    best_order=9;
end

ORDER=best_order;
PARAM=ar_param(best_order,1:best_order);
disp(['AR Model Parameters: ' num2str(PARAM)])
out_res=residuals(:,best_order);
out_est=ar_estimate(:,best_order);

% outlier_clt.m %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% This subfunction determines the auto regressive outliers in the
% residuals that are greater than a given number of std devs using the
% central limit theorem.
% 99% of the observations lie within 2.58(std_res)
% 97.5% of the observations lie within 2.24(std_res)

function [dres, rmr, type]=outlier_clt(in,fig2)
global YEARS

% initialize variables
type=0; % Type of outlier detected (1=pulse, 2=trend)
lngth=length(YEARS);
a=9; b=30;
% a=9; b=30; %b=lngth/3; %b=lngth-40; % min and max of trend window
if b>lngth/4
    b=floor(lngth/4);
    disp(['Maximum outlier detection length reduced to ' num2str(b) ' due to low ring #'])
end
lt=a; % Length of trend
window=0;
nse=3.29;
% nse=1.96; % 95 pct CI
% nse=2.58; % 99 pct CI8
% nse=3.29; % 99.9 pct CI
dres=zeros(lngth,1); % downweighted residuals
mr=zeros(lngth,1); % residuals mean in window
rmr=zeros(lngth,1);
rmu=0;
rshat=0;

% initialize masked to ones.
marker=zeros(length(YEARS),1);
masked=zeros(length(YEARS),1);

std_res = (var(in))^0.5; % Calculate std dev of residuals

% for u=1:length(YEARS) % Detect pulse outliers
%     rres=in(u)/(nse*std_res); % calculates relative residuals
%     if  rres>=1.0
%         dres(u)=1;
%         type=1;
%         disp(['Positive Pulse in ' int2str(YEARS(u))])
%         disp(['Outlier value = ' num2str(rres)])
%     elseif rres<=-1.0
%         dres(u)=1;
%         type=1;
%         disp(['Negative Pulse in ' int2str(YEARS(u))])
%         disp(['Outlier value = ' num2str(rres)])
%     else
%         disp('No pulse outliers detected')
%     end
%     psi=rres;
%     The code below simply produces psi = rres. Why did Ed
%     code it
%     % this way in his robar function?
%     % psi=rres*exp(-exp(3.0*(abs(rres)-3.0)))
%     else
%     disp('No pulse outliers detected')
if a <= b
    for v = a:b
        z = v - a + 1;
        window = in(u: (u + v - 1));
        mr(u, z) = mean(window);
    end
    % [muhat(z), sigmahat(z)] = normfit(mr(:, z)) % Arithmetic mean
    % Uses Tukey's bi-weight mean instead (Hoaglin 1983, Meko's code)
    [ynm(z), varyh(z), ~, ~, ~, se] = bisqmean_CID(mr(:, z));
    [mam(z), imax(z)] = max(mr(:, z)); % Find max. means & their locations
    [mim(z), imin(z)] = min(mr(:, z)); % Find min. means & their locations
end

poso = (mam - ynm) ./ varyh; % Determines # deviations from mean of means
nego = (ynm - mim) ./ varyh;
[relmam, rimax] = max(poso); % Find max. of positive dev.s & their locations
[relmim, rimin] = max(nego); % Find max of negative dev.s & their locations

if (poso(rimax) >= nse) && (relmam >= relmim) % Comment && for only + outliers
    type = 2;
    lt = rimax + a - 1; % length of trend
    dres(imax(rimax): (imax(rimax) + lt - 1)) = ynm(rimax);
    disp(['Release detected in ' int2str(YEARS(imax(rimax))) ... '
        for ' int2str(lt) ' years'])
    rmu = ynm(rimax);
    rshat = varyh(rimax);
    rmr = mr(:, rimax);
    disp(['rmu = ' num2str(rmu)])
    disp(['rshat = ' num2str(rshat)])
else
    % Comment elseif for only + outliers
end

else
    rmu = ynm(1);
    rshat = varyh(1);
    rmr = mr(:, 1);
    disp(['No trend outliers detected up to ' int2str(b) ' yrs'])
end
if fig2==1 % fig=1 if you want a figure as output
    subplot(4,1,3)
    hold on
    hist(rmr)
    h413 = findobj(gca,'Type','patch');
    set(h413,'FaceColor','k')
    box on
    title('f Histogram of Running AR Residual Means')
    line([rmu-nse*rshat rmu+nse*rshat],[10 10],'Color',[.6 .6 .6])
    plot(rmu,10,'o','Color',[.6 .6 .6])
    ylabel('f Frequency')
    xlabel(['f' int2str(lt) 'f Yr Residual Means'])
    hold off
end

% backcast.m
% This subfunction estimates the first elements of a series for which
% residuals could not be calculated owing to the use of ar modeling.
function bckcasted=backcast(seriesb)
  global PARAM
  global ORDER
  global YEARS

  % Invert time series for backcasting.
  flipped=flipud(seriesb);
  % Add in backcasted AR estimates as new values at end of inverted
  % series.
  for g=(length(YEARS)+1):(length(YEARS)+ORDER) % g = backcasted years
    ar=0; % ar model estimate for order i, year g
    for k=1:ORDER % kth parameter of order ORDER
        ar=PARAM(k)*(flipped(g-k))+ar;
    end
    flipped(g)=ar;
  end
  % Re-invert series and return as output.
  bckcasted=flipud(flipped);
  % disp('Backcasted Values:')
  % for h=ORDER:-1:1
  %     disp(['Year -' int2str(h) ': ' num2str(bckcasted(h))])
  % end
This function can be used to process multiple series (whereas v105pn.m is used to process single series). The function calculates each autoregressive outlier for each core in a dataset and lumps those results by tree. This version also returns the transformed standardized series for each core in St. The ‘filename’ used in this function and in function v105pn.m to load the data file must match and should contain data imported using function ringwidth_import.m.


function [yrs,tres,det,St,Straw,Dtraw,sigDtraw,Atraw,out,dbh_rel,age_rel] ... =v105pn_chron

% Load tree-ring data (returns vars *col_header* and *rings*)
load filename.mat

iter=8; % maximum number of iterations per series.
years=rings(:,1); % years for entire chronology
rings(:,1)=[]; % remove year column
col_header(1)=[]; % remove year label from array
ncores=size(rings,2); % # cores in group
nyrs=size(rings,1); % total # of years in chronology
expval=NaN(1,ncores); % value of last year that neg exp curve fits
yrs=NaN(nyrs,ncores); % years for each cores
dbh_rel=NaN(nyrs,ncores); % dbh at release for each core
age_rel=NaN(nyrs,ncores); % age at release for each core
tres=NaN(nyrs,ncores); % Transformed residuals for each core
det=NaN(nyrs,ncores); % Detrended series for each core
St=NaN(nyrs,ncores); % Undisturbed series for each core in transformed units
Straw=NaN(nyrs,ncores); % Undisturbed series for each core
Atraw=NaN(nyrs,ncores); % Age series for each core
Dtraw=NaN(nyrs,ncores); % Disturbance series for each core
agestats=NaN(2,ncores); % power and trend type for transformed core
agestats=cell(3,ncores); % power and trend type for transformed core
agestats(3,:)={'0000'};
out=NaN(iter,7,ncores); % Outlier statistics

for i=1:ncores
    disp(' ') disp(['Series #' num2str(i) '-------------------------------------'])
    % Find pointer to start and end of series
    s=find(rings(:,i)>0,1);
    e=find(rings(:,i)>0,1,'last');
    [yrs(s:e,i),tres(s:e,i),det(s:e,i),St(s:e,i),Straw(s:e,i),Dtraw(s:e,i),
     Atraw(s:e,i),agestats(:,i),out(1:iter,1:5,i)]=v105pn(i,0,iter);
    b=find(out(:,5,i)==2);% find all releases for a core
    if b>0
for c=1:length(b) % iterate through each release
    startyr=find(years==out(b(c),1,i));
    dbh_rel(startyr,i)=sum(rings(s:startyr,i))/1000;
    age_rel(startyr,i)=length(rings(s:startyr,i));
end
end
clear b
clear startyr
end

figure('Position', [10 5 700 800])
subplot(2,1,1)
nanrings=cumsum(rings,1); % cumulative dbh
coreage=rings;
coreage(find(coreage>0))=1;
coreage=cumsum(coreage,1); % count age of each core
coreage(coreage==0)=NaN;
av_ca=nanmean(coreage,2);
v_dbh=nanstd(nanrings,1,2)/1000;
hold on
fill([years; flipud(years)],[v_dbh+av_dbh; flipud(av_dbh-v_dbh)],[.7 .7 .7],'EdgeColor','none')
plot(years,av_dbh,'k',years,dbh_rel,'k--o')
ylabel('bf Av. Inside Diameter (m)')
xlabel('bf Year')
hold off
subplot(2,1,2)
hold on
fill([years; flipud(years)],[v_ca+av_ca; flipud(av_ca-v_ca)],[.7 .7 .7],'EdgeColor','none')
plot(years,av_ca,'k',years,age_rel,'k--o')
ylabel('bf Av. Age')
xlabel('bf Year')
rel=[0 0 0 0];
sup=[0 0 0 0];
d=1; % release counter
f=1; % suppression counter
for a=1:size(Dtraw,2) % # of cores
    b=find(out(:,5,a)==2); % find all releases for a core
    if b>0
        for c=1:length(b) % iterate through each release
            startyr=out(b(c),1,a);
            endyr=out(b(c),2,a);
           Dt_diff=out(b(c),6,a)-out(b(c),7,a);
gc=(Dt_diff)/out(b(c),7,a);
rel(d,1:4)=[a startyr endyr gc];
d=d+1;
end
end
clear b
g=find(out(:,5,a)==3);
if g>0
    for h=1:length(g)
startyr2=out(g(h),1,a);
endyr2=out(g(h),1,a);
Dt_diff2=out(g(h),6,a)-out(g(h),7,a);
gc2=(Dt_diff2)/out(g(h),7,a);
% Dt_inc=Dt_diff*(endyr2-startyr2+1);
sup(f,1:4)=[a startyr2 endyr2 gc2];
f=f+1;
end
end
clear g

clear g

rings(find(~rings))=NaN; % Convert rings matrix zeros to NaNs

% Find and average all cores with pos or neg outliers
subset=unique([rel(:,1); sup(:,1)]);
subset=subset(find(subset));% Find & remove nonzeros if no pos or neg outliers found
if subset % Only graph if interventions found.
sigDtraw=Dtraw(:,subset);
sigDtm=nanmean(Dtraw(:,subset),2);

    depth=size(Dtraw,2)-sum(isnan(Dtraw,2)); % Total sample depth
    subdepth=size(Dtraw(:,subset),2)-sum(isnan(Dtraw(:,subset)),2);

    figure('Position', [10 5 700 800])
    subplot(4,1,1)
    h_end=plot(years,nanmean(rings-Atraw,2),'k',years,nanmean(Straw-Atraw,2),'k--');
    set(h_end(1),'LineWidth',2)
    set(h_end(2),'LineWidth',2)
    legend('Mean Ct + Dt','Mean Ct','Location','NorthWest')
    legend('boxoff')
    ylabel('f Residuals (mm)')
    xlabel('f Year')

    subplot(4,1,2)
    [AX,H1,H2] = plotyy(years,sigDtm,years,depth);
    set(H1,'LineWidth',2)
    set(H1,'Color','k')
    set(H2,'Color','k')
    set(AX(1),'ycolor','k')
    set(AX(2),'ycolor','k')
    set(get(AX(1), 'Ylabel'), 'String', '{\bf Mean Dt)', '(mm)')
    set(get(AX(2), 'Ylabel'), 'String', '{\bf Sample Size}')
    set(AX(2), 'ylim', [0 ceil((ncores+1)/10)*10])
    set(AX(1), 'XTickLabel', [])
    bounds=xlim;
    box off

    subplot(4,1,3)
    mindecade=bounds(1);
    maxdecade=bounds(2);
    edges=[mindecade:10:maxdecade]; % Bin outliers by decade
    pCCT=histc(rel(:,2),edges);
    nCCT=histc(sup(:,2),edges);

    hold on
    bar(edges+5,pCCT,'k')
    bar(edges+5,-nCCT,'k')
    xlim([bounds(1) bounds(2)])

    subplot(4,1,4)
    h_end=plot(years,nanmean(rings-Atraw,2),'k',years,nanmean(Straw-Atraw,2),'k--');
    set(h_end(1),'LineWidth',2)
    set(h_end(2),'LineWidth',2)
    legend('Mean Ct + Dt','Mean Ct','Location','NorthWest')
    legend('boxoff')
    ylabel('f Residuals (mm)')
    xlabel('f Year')

    subplot(4,1,5)
    h_end=plot(years,nanmean(rings-Atraw,2),'k',years,nanmean(Straw-Atraw,2),'k--');
    set(h_end(1),'LineWidth',2)
    set(h_end(2),'LineWidth',2)
    legend('Mean Ct + Dt','Mean Ct','Location','NorthWest')
    legend('boxoff')
    ylabel('f Residuals (mm)')
    xlabel('f Year')

    subplot(4,1,6)
    h_end=plot(years,nanmean(rings-Atraw,2),'k',years,nanmean(Straw-Atraw,2),'k--');
    set(h_end(1),'LineWidth',2)
    set(h_end(2),'LineWidth',2)
    legend('Mean Ct + Dt','Mean Ct','Location','NorthWest')
    legend('boxoff')
    ylabel('f Residuals (mm)')
    xlabel('f Year')
hold off
ylabel('\bf +Dt Initiation Yrs')

subplot(4,1,4) % show cores that are open grown initially
agenum=cell2mat(agestats(2,:)); % Bin outliers by year
edges=mindecade:10:maxdecade;

for i=1:size(yrs,2);
fyr(i)=yrs(find(rings(:,i)>0,1,'first'),i); end
opngrp=firstyr(agenum==1);
disp('Trees that likely established in open conditions')
disp(col_header(agenum==1));
cldcan=firstyr(agenum==2|agenum==3);

% Establishment dates binned by decade
o=histc(opngrp,edges); o=o(:); % open oaks
u=histc(cldcan,edges); u=u(:); % understory oaks
Y=[u o];
bar(edges+5,Y, 'stacked')
axis([mindecade maxdecade 0 max(o+u+5)])
ylabel('\bf Tree Recruitment')
xlabel('\bf Year')

ColorOrder2=[0 0 0; 1 1 1];
colormap(ColorOrder2)

else
    sigDtraw=0;
sigDtm=0;
end