




Principal Components Analysis: An Alternative Way for Removing Natural Growth Trends

DANIELA OLIVEIRA DA SILVA,¹  VIRGINIA KLAUSNER,¹ ALAN PRESTES,¹ HUMBERTO GIMENES MACEDO,¹
TUOMAS AAKALA,² and IURI ROJAHN DA SILVA¹

Abstract—In this article, we establish a new approach for removing natural growth trends from tree-ring samples, also called detrending. We demonstrate this approach using *Ocotea porosa* (Nees & Mart) Barroso trees. Nondestructive samples were collected in General Carneiro city, located in the Brazilian southern region (Paraná state). To remove natural tree growth trends, principal components analysis (PCA) was applied on the tree-ring series as a new detrending method. From this, we obtained the tree-ring indices by reconstructing the tree-ring series without the first principal component (PC), which we expect to represent the natural growth trend. The performance of this PCA method was then compared to other detrending methods commonly used in dendrochronology, such as the cubic spline method, negative exponential or linear regression curve, and the regional curve standardization method. A comparison of these methods showed that the PCA detrending method can be used as an alternative to traditional methods since (1) it preserves the low-frequency variance in the 566-year chronology and (2) represents an automatic way to remove the natural growth trends of all individual measurement series at the same time. Moreover, when implemented using the alternating least squares (ALS) method, the PCA can deal with tree-ring series of different lengths.

Keywords: Dendrochronology, tree ring, time series, natural records, principal components analysis.

1. Introduction

Studying tree rings allows us to infer environmental conditions and geophysical phenomena from recent years to millennia in the past (Fritts, 1976). In seasonal climates, tree rings can be dated at an annual resolution, and thus, the information extracted from tree rings has a high temporal resolution relative to

many other such biological archives. One important and common aim in tree-ring research is the reconstruction of climate variability at annual-to-decadal and longer time scales (Briffa et al., 1996), which provides us with a longer-term context of the modern climatic variability and changes. However, information stored in tree rings has been used for a number of different purposes, such as reconstructions of sunspot activity (Stuiver & Quay, 1980), cyclones (Miller et al., 2006), or volcanic eruptions (Piermattei et al., 2014).

A pertinent issue in research based on tree-ring widths is the extraction of the desired signal from the noisy ring-width data. Tree growth and the width of the tree rings are influenced by several physiological and biochemical factors that need to be taken into account when interpreting data from tree rings (Speer, 1971). Depending on the study objectives, the desired signal can consist, for instance, of year-to-year variability in the tree-ring widths and wood anatomy (Piermattei et al., 2014), sudden and sustained longer-term changes in the tree-ring widths (Maes et al., 2017), or longer-term, trend-like declines (Amoroso et al., 2012). In the context of climate reconstructions with a focus on annual-to-decadal and longer time scales, typically an important first step in the analysis is the removal of the natural growth trend and other non-climatic variability to maximize the climate signal (Briffa et al., 1996). A particular problem here is how to preserve the low-frequency signal (climate signal) when removing the natural tree growth trend (Helama et al., 2017).

Statistical methods can be applied on a tree-ring time series to remove the biological growth trend. This application is commonly referred to as detrending or standardization (Fritts, 1976). The choice of the detrending method depends on the

¹ Research and Development Institute-IP&D, Vale do Paraíba University-UNIVAP, São José dos Campos, SP, Brazil. E-mail: fys.dani@gmail.com

² School of Forest Sciences, University of Eastern Finland-UEF, Joensuu, Finland.

purpose of the study because it will influence the interpretation of the tree-ring data.

Detrending methods are separated into stochastic and deterministic methods, and each one has its specific advantages and disadvantages (Cook et al., 1990). Deterministic methods fit mathematical models to estimate the growth of the trees, while stochastic methods are more adaptable to the data characteristics (Cook et al., 1990). Common examples of detrending methods aimed at removing non-climatic signals include regional curve standardization (RCS) (e.g., Helama et al., 2004, 2017) or the negative exponential or linear regression curve (standard [STD]) (e.g., Cook et al., 1990; Lorenzi & Prestes, 2016; Prestes, 2009).

One common statistical method for extracting signals from noisy data is principal components analysis (PCA), which is one of the most popular methods for reducing the dimensionality of a data set, simplifying its analysis and interpretation (Regazzi 2000). In tree-ring research, PCA has been commonly used to deal with multicollinearity of the independent variables. As an example, Fekedulegn et al. (2002) applied principal components regression (PCR) to model the sensitivity of tree growth to temperature and precipitation. Enright (1984) applied PCA on tree-ring samples from different regions of Canada to assess the response of the trees to precipitation and temperature. Flower and Smith (2011) used PCA to develop a white spruce ring-width chronology, and also to reconstruct the average temperatures from June to July in the northern Canadian Rockies. The major disadvantage of PCA was that it requires time series of equal length, leading to loss of information when they are resized.

The choice of a detrending method may vary according to the tree-ring series characteristic that need to be preserved (Helama et al., 2004). In studies aimed at reconstructing past climates, choices in the detrending method risk missing the trees' responses to climatic variations, and consequently providing a misleading view of past climate (Shi et al., 2020). In this article, we intend to establish a new approach to remove natural growth trends using PCA. To solve the issues arising from the loss of data when resizing (i.e., trimming tree-ring series to equal length), we use a nonlinear optimization method, the alternating

least-squares method (ALS), which performs the best adjustment for smaller series. We then compare the method to three other commonly used tree-ring standardization methods based on cubic spline functions, regional curve standardization (RCS), and negative exponential functions. Following, we will decompose the tree-ring series into its principal components (PCs) to estimate their natural growth trends. After obtaining the chronology via PCA standardization, the PCA chronology will be compared to the chronology obtained via traditional methods.

2. Tree-Ring Data and Study Area

2.1. Study Area

Samples of imbuia trees obtained in General Carneiro city (26°24'0125" S, 51°24'0391" O), Paraná state, in southern Brazil were used in this article. The city has a territorial extension of 1070.30 km² and is located at an altitude of 983 meters (General Carneiro, 2020). It is also part of the mesoregion of the southeast of Paraná and the microregion of União da Vitória. Figure 1 shows the location of General Carneiro city in the Brazilian map of the south region, and its adjacent areas where the samples used in this study were collected.

2.2. Tree Species Description

Ocotea porosa (Nees & Mart) Barroso belongs to the Lauraceae family, and it is known as imbuia (Fig. 2). This species can reach heights of 10–20 m, and an average of diameter at breast height (DBH) of between 50 and 150 cm. In adulthood, the height values can reach up to 30 m, with a DBH of 320 cm or more. This tree is a species of a mixed ombrophilous forest (Carvalho 2003), which is characterized by a floristic mixture, comprising Australasian (*Drymis*, *Araucaria*) and Afro-Asian (*Podocarpus*) genera, with a physiognomy strongly marked by the predominance of *Araucaria angustifolia* (pine) in its upper stratum. Its area of occurrence coincides with a humid climate without a dry season, and with average annual temperatures ~18°C. Its

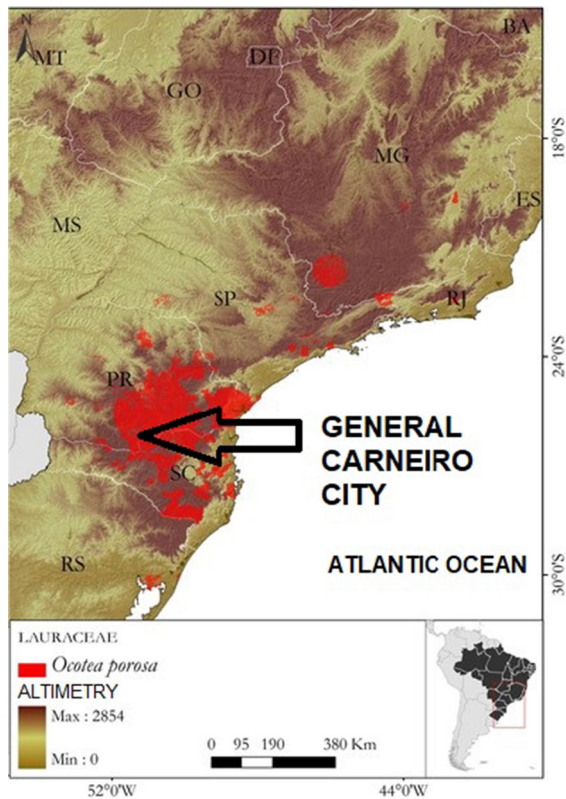


Figure 1

The sample collection location of Carneiro city, Paraná, southern Brazil, and the species occurrence map. Source: Modified from CNCFlora (2012)

environments predominate in the southern Brazilian Plateau on lands above 500–600 m of altitude, with disjunctions at higher points in the mountains of Mar and Mantiqueira (Berrêdo 2015).

Imbuia is native to the Brazilian states of Paraná (PR), Rio de Janeiro (RJ), Rio Grande do Sul (RS), Santa Catarina (SC), and São Paulo (SP), between latitudes 22°30' S (RJ) and 29°50' S (RS), and it is associated with *Araucaria angustifolia* (known as pinheiro-do-paraná) and is rare where there are no pine trees (Klein 1963). In the pine sub-forests, it constitutes the most abundant tree, being commonly found at a rate of 6 to 20 adult imbuias per hectare.

There are areas with high concentrations of imbuia, and this is due to several conditions such as soils of low natural fertility, high levels of aluminum, and medium and high levels of chemical fertility (Reitz et al., 1978). This species can be observed



Figure 2

Ocotea porosa (Nees & Mart) Barroso (imbuia) in the region of General Carneiro. Source: Laboratório de Registros Naturais (LRN)-UNIVAP

from the bottom of valleys to the top of the slopes (Marchesan et al., 2006; Carvalho, 1994).

However, imbuia is seldom used in dendrochronological studies, but its dendrochronological potential is recognized (Stepka 2013). Considering its dendrochronological potential, imbuia is possibly the longest-lived tree species in the “araucaria forest,” with a lifetime that can exceed 500 years (Carvalho 1994). Anatomically, in the cross section of the wood, the presence of distinct growth layers is observed. It is characterized by the flattening of the fibers in the latewood, with cell walls that gradually thicken in the radial direction (Fig. 3). At the boundaries between the tree-ring growth, there is a sudden transition from cells with thick walls to those with thin walls which characterize the initial wood of the next ring (Tomazello Filho et al., 2004; Cosmo et al., 2009).

2.3. Sample Preparation

The samples of *Ocotea porosa* (Nees & Mart) Barroso were collected in the municipality of General Carneiro in January 2013. We obtained 64



Figure 3

Macroscopic cross section of *Ocotea porosa* (Nees & Mart) Barroso wood showing the tree rings. Source: Laboratório de Registros Naturais (LRN)-UNIVAP

samples from 21 trees at 1.3-m height using an increment borer. Healthy individuals were selected for sampling to minimize any influence of tree damage, disease, or insect pests on the tree-ring series. As a first step to obtaining the dendrochronological series, the samples were initially polished on the transversal surface using different sandpapers (from 50 to 600 grains), so that it is possible to better visualize the tree rings. Subsequently, they were examined under a stereomicroscope (6–40 \times magnification) and an optic fiber lighting system for the demarcation of the annual tree rings. To measure the tree rings, a Velmex measuring table with an accuracy of 0.001 mm (shown in Fig. 4) was used.

Ring widths were measured from the tree bark to the pith. The last ring formed until the date of collection corresponds to the year 2011. This is because the beginning of the formation of an imbuia ring occurs in spring/summer (considering the Southern Hemisphere); in this way, the ring corresponding to the year 2012 had not been yet completed its growth in January 2013. The oldest tree had 566 rings.

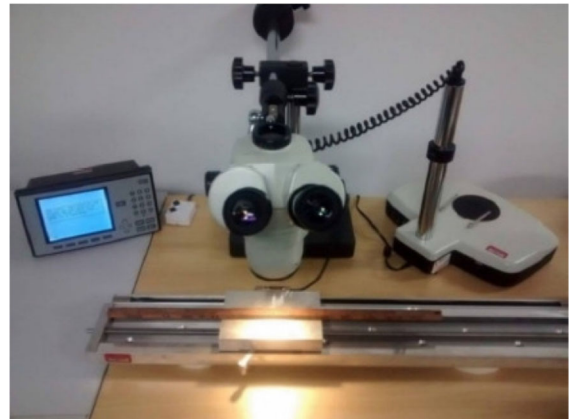


Figure 4

Velmex measurement table associated with a stereoscopic microscope. Source: Laboratório de Registros Naturais (LRN)-UNIVAP

An analysis to verify the tree-ring width time series similarities to reduce the error of the mean chronology was performed. To this end, a subset of trees was defined to represent each period of time that trees coexist. Following, we used a measure of squared Euclidean distance to compare the tree-ring series pairwise to find dissimilarities between series, and dendrograms were built from the resulting dissimilarity matrix. The grouping method was the Ward pair-group method (inner squared distance). Another reason to apply the grouping method is that this method is able to eliminate samples with counting errors or false rings and partially missing rings through a similarity group analysis. From this analysis, only 41 samples were chosen to develop the mean chronology. To improve accuracy, the cross-validation technique was also performed using a 50-year window to provide the dating control of the examined characteristics (Cook & Peters, 1997).

3. Methods to Develop the Mean Chronology

3.1. Removing the Growth Trend

The tree-ring chronologies based on the PCA detrending method were obtained using the following three steps. First, the natural growth trend was expected to be associated with the first PC. This hypothesis was based on the fact that the first PC

represents the direction of largest variance of the original tree-ring series (detailed in Appendix A). Second, the natural growth trend was removed from the tree-ring series by standardization methods. In our case, the tree-ring indices were computed using the first PC by three methods, i.e., by division, by subtraction, and by reconstructing the series without its first PC, in order to compare the results. Third, the chronology was then computed as a mean of the individual detrended series using the Tukey biweight robust mean method.

To perform PCA of the tree-ring series with different lengths, we estimated the values of the shortest series using the ALS method. After this procedure, all tree-ring series were padded using ALS in order to have the same length (566 years). These estimated values were removed after the principal components (PCs) were calculated, and therefore, they were not used to develop the mean chronology. Although PCA can be conducted with missing values, the resulting eigenvectors lack some of the usual statistical properties. The standard procedure is to delete individuals or variables containing missing observations and perform the PCA. However, this loss of information reduces the ability of PCA to detect patterns and can also introduce biases (Dray & Josse, 2015). Brief descriptions of PCA and ALS methods are given in Appendix A.

We then compared the PCA detrending to chronologies obtained by other commonly used detrending methods. In this, we included the 67% spline curve (SP67), regional curve standardization (RCS) (Melvin & Briffa, 2014a), and negative exponential curve detrending–STD (Helama et al., 2004). For the negative exponential detrending, we used a modified version in which a linear model was fitted if the negative exponential presented a zero slope. Finally, the four versions of chronologies were computed for each detrended method mentioned above using Tukey's biweight robust mean. If our hypothesis holds, the first PC should contain similar variation to the other detrending methods.

3.2. Standardization

The tree-ring indices were defined by reconstructing the tree-ring series without the first PC or by

dividing each measured tree-ring width, R_t , by its expected values; in other words, by its estimated growth trend G_t ,

$$I_t = R_t/G_t,$$

or by subtracting the expected values from each measured tree-ring width,

$$I_t = R_t - G_t,$$

where t stands for a year. For the other detrending methods, subtraction and division were applied in the same way.

3.3. Estimation of the Mean Chronology

We used Tukey's biweight robust mean to reduce the effect of outliers in the estimation of the mean chronology. This method reduces the weight of outliers by considering weighted averages assigned to each data element (i) of the tree-ring indices in a certain year, as follows

$$\bar{I} = \frac{\sum_{i=1}^n w_i I_i}{\sum_{i=1}^n w_i},$$

such that

$$w_i = \left(1 - \left[\frac{I_i - \bar{I}}{cS_i}\right]^2\right)^2$$

are the weights, with $S_i = \text{median}\{|I_i - \bar{I}|\}$. The constant c is often taken as 6 or 9 (Cook & Peters, 1997). Here, we used $c = 9$. This procedure is a process of iteratively reweighted least squares. The constraint takes the following form

$$\left[\frac{I_i - \bar{I}}{cS_i}\right]^2 < 1;$$

otherwise, it will be given a weight of zero and will not enter into the calculation of the biweight estimate at all. Next, we can test for convergence when, for example, an estimate of \bar{I} changes by no more than 10^{-3} from one iteration to another. To start the interaction for computing the final value of \bar{I} , the arithmetic mean or median can be used as an initial estimation (Cook & Peters, 1997).

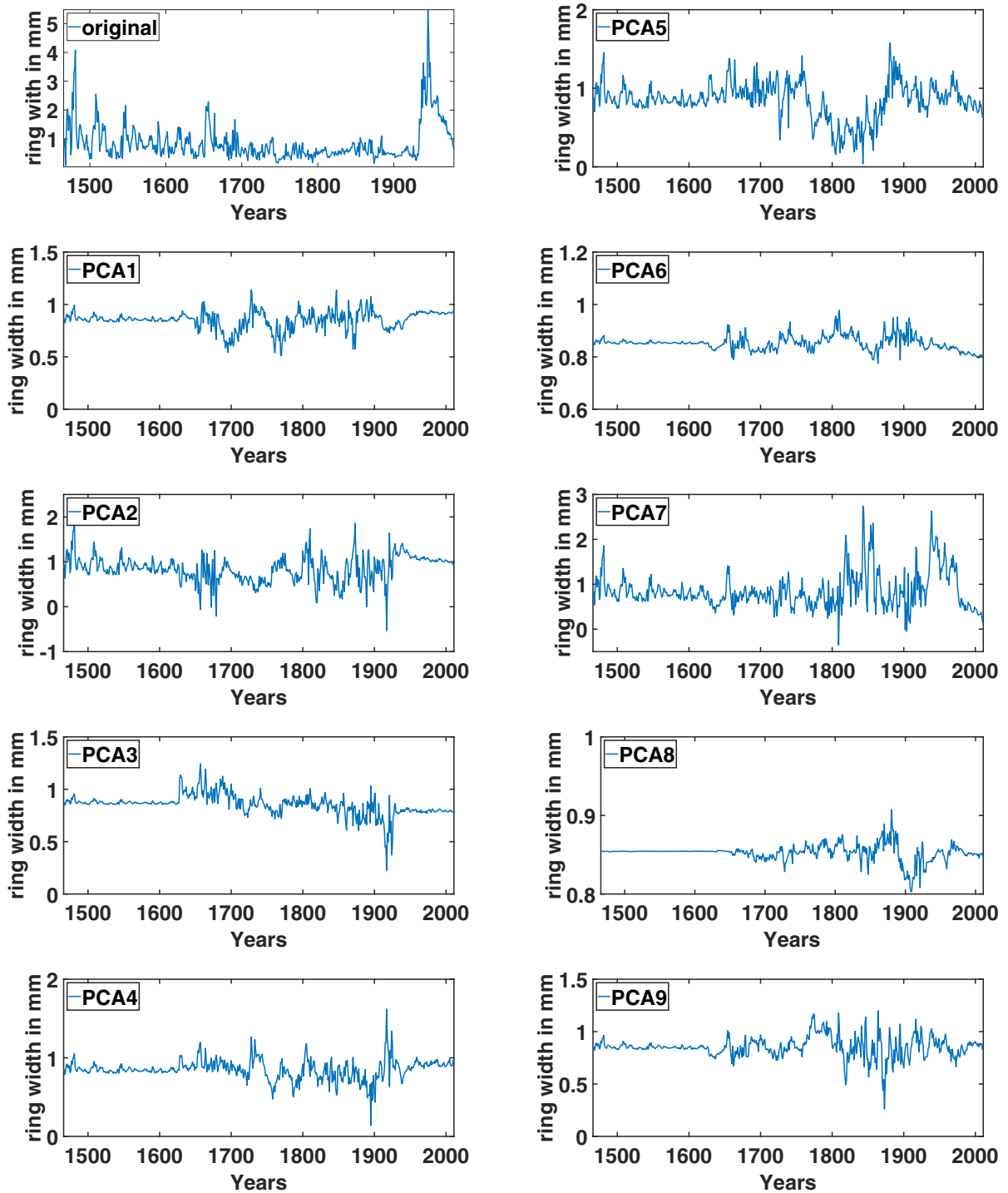


Figure 5

Different components decomposed by PCA for an example tree-ring width data set. Each decomposition is the projection in the principal component direction of the data that explains a maximal amount of variance—in other words, the axis that captures the most information of the data

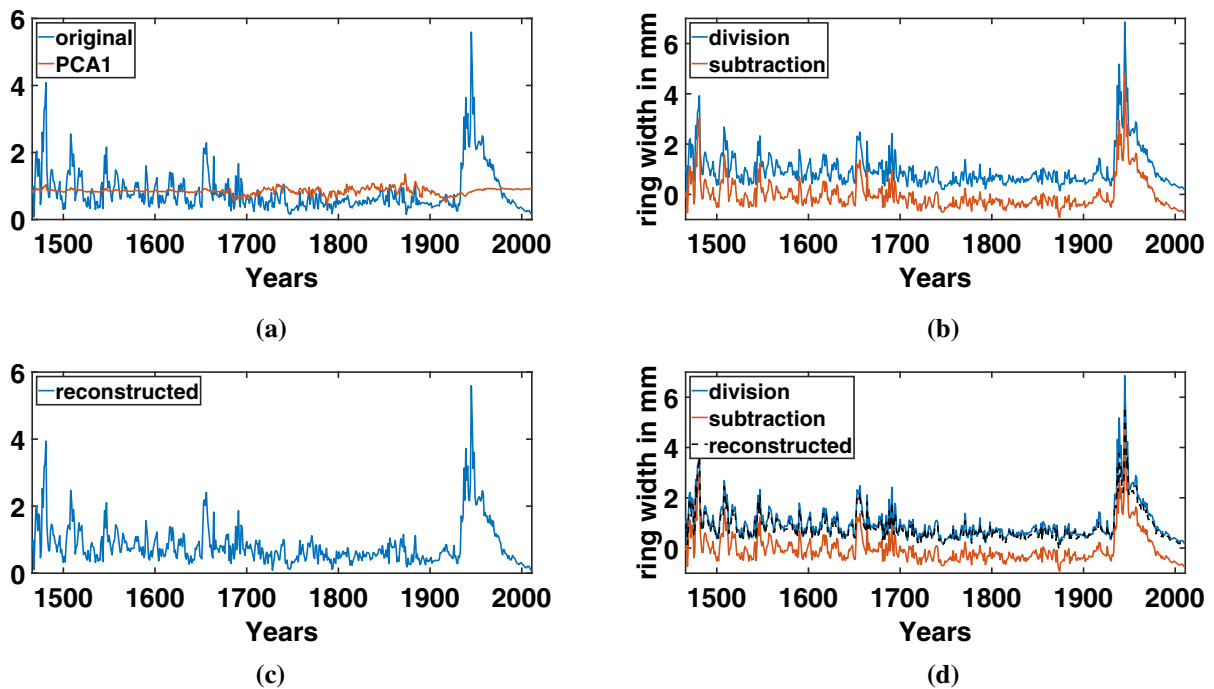


Figure 6

Tree-ring standardization using the first PC as the growth trend G_t (a). The tree-ring indices are calculated using the raw data divided (or subtracted) by G_t (b). Also, the reconstruction without the first PC is shown (c). The resulting index series, calculated as ratios, residuals, and reconstruction are shown together in the same plot (d)

4. Results and Discussion

4.1. Tree Growth Identified Based on PCA

Here, we illustrate the step-by-step use of PCA for tree-ring series detrending by considering the first PC as the natural growth trend of the tree. Figure 5 shows an example of tree-ring raw data and its nine decompositions. In this case, we use only the first PC, only the second PC, and so on. The reconstruction uses the projections on the eigenvectors of the covariance matrix as described in Appendix A.

Figure 6a shows the use of the first PC to estimate the growth trend, G_t . In Fig. 6b, the tree-ring indices are estimated by ratios and by residuals. Each PC explains a fraction of the total variance. The first PC represents $\sim 45\%$ of the total variance. Figure 6c shows the reconstruction of the ring width series using all the PCs, only disregarding the first PC. As depicted in Fig. 6d, the resulting indices by the three methods are displayed together for comparison. This procedure was applied to the 41 tree-ring series, but

the other tree graphs are not shown here. For instance, Fig. 7 presents three examples of tree-ring width series and their associated first PC used to identify growth trends using the same steps mentioned above. The visual comparison between the tree-ring indices as ratios and reconstruction demonstrates several similarities.

4.2. Tree Growth Identified by Other Methods

Four tree-ring standardization methods (PCA, SP67, RCS, and STD) are compared in this paper. Of these, the PCA, RCS and STD methods conserve low-frequency signals, while the SP67 method removes low-frequency signals. In Fig. 8, we provide four examples to illustrate the results using 67% n splines as growth model trends. According to Speer (1971) and Cook (1985), the cubic spline method is a stochastic method that uses a low-pass filter. Changing spline rigidity controls the frequencies of variance that are removed or preserved, and so the

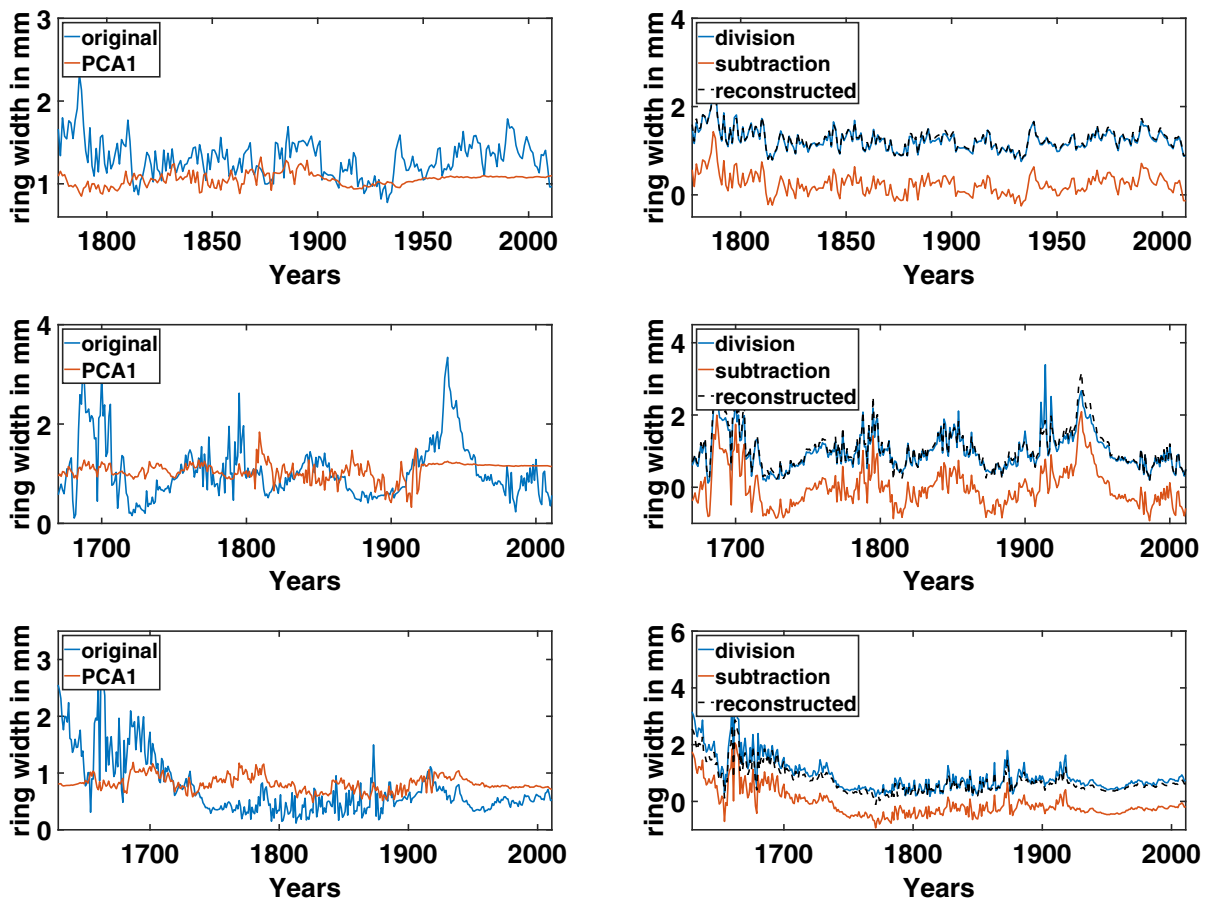


Figure 7

Growth trends and detrended series for three raw tree-ring width data sets using PCA. The left panels show the growth trend identified by PCA, and the right panels are the detrended series for each tree-ring data set using ratios, residuals, and reconstruction

signal at desired frequencies can be amplified. As an example, a 32-year-old spline preserves around half the range of variations that have a wavelength higher than 32 years (Helama et al., 2004), while a 50-year spline is more rigid (Brienen & Zuidema, 2005; Briffa et al., 1990; Lindholm et al., 1999). Here, we chose the 67% spline with a 50% frequency cutoff (SP67). As discussed by Helama et al. (2004), a 67% spline is a stiffer function, and no more than half of the amplitude with wavelengths of two-thirds of the length of each tree-ring series can be expected to be preserved in its resulting indices.

Another method to derive detrended series used here is RCS. The RCS method first realigns all time series of ring width by biological age to calculate an average ageing curve of the same tree species and

region (e.g., Helama et al., 2004, 2017). Thus, it is expected that only the ageing factor will be preserved in the average curve. Before the standardization is done, this curve is smoothed. The use of only one average curve to remove the natural growth trend from all samples has already been applied in studies by Huntington (1914), Briffa and Melvin (2011), Autin et al. (2015), and Biondi and Qeadan (2008).

It is possible to implement RCS in different ways. In our implementation, all series of tree-ring measurements were first aligned by their cambial age, which means that the first year of each series is set to the biological age of 1. The arithmetic average of measurements was then used to produce a curve of the mean tree-ring series. However, the obtained mean tree-ring series was noisy and needed to be

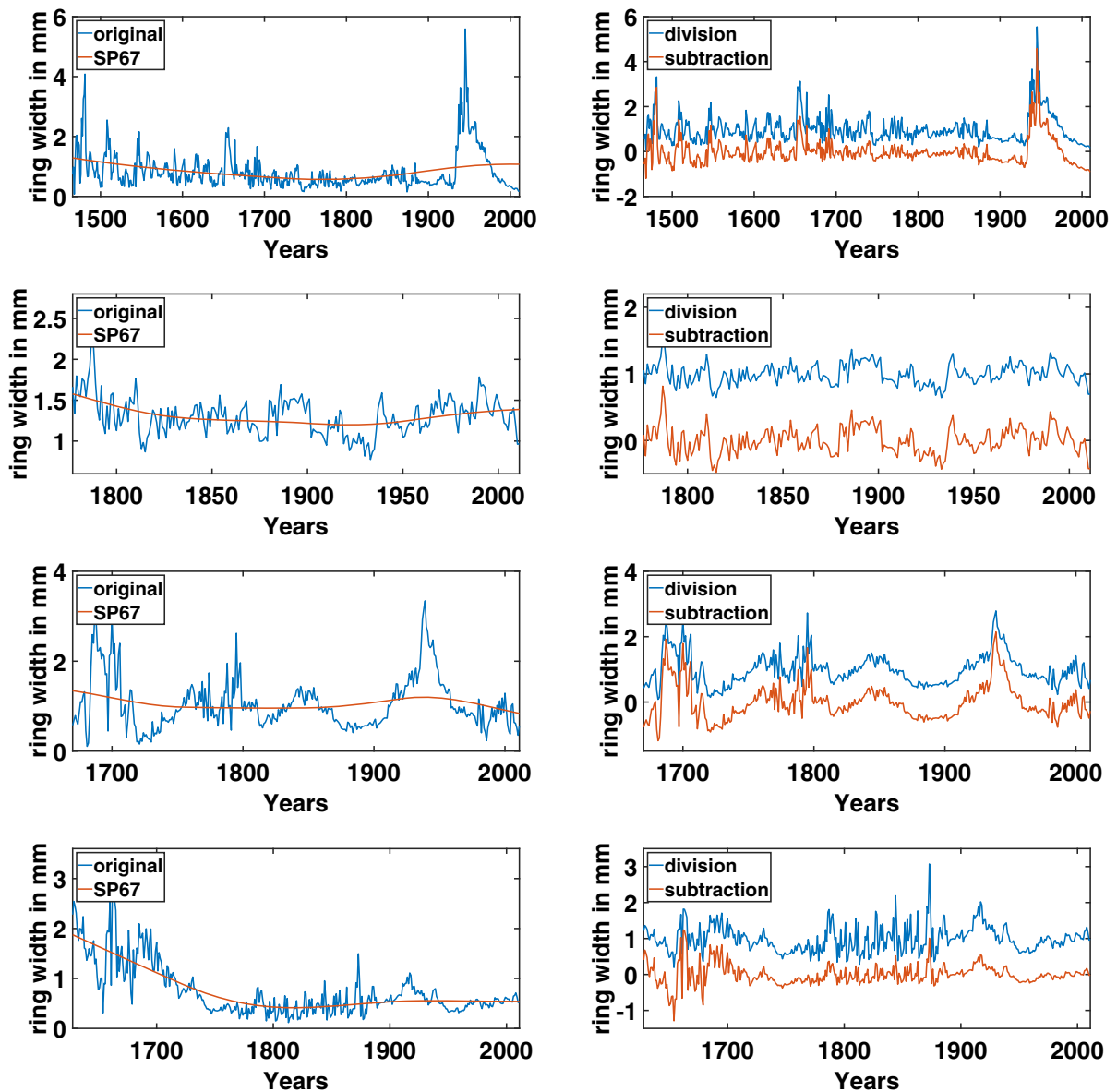


Figure 8
Four examples of the 67%*n* splines as growth model trend results

smoothed. We used the 67%*n* spline function to create a smooth RCS curve (Melvin & Briffa, 2014b), as shown in Fig. 9. Each tree-ring series was then divided by and subtracted by the RCS curve value for its particular biological age to give an index, as presented in Fig. 10. Finally, all tree-ring series after the RCS removal were realigned to their original calendar years to produce the RCS chronology

In the next method, it was necessary to calculate fitting functions (such as linear or negative exponential functions) until the best fit was found for each sample. The linear model is a special case of a negative exponential with a zero slope. In this article, the negative exponential is applied together with the linear regression. In analogy with the PCA detrending method, the following steps were also involved to

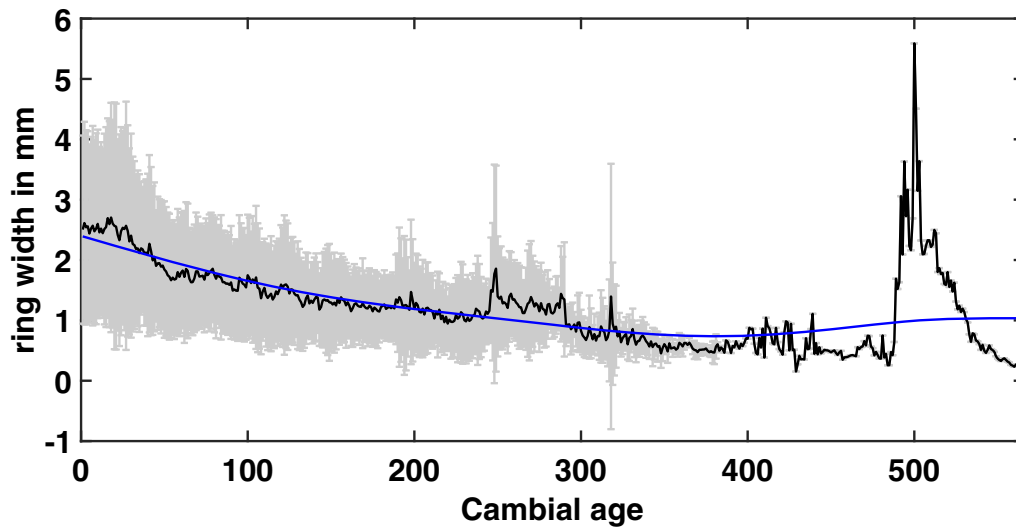


Figure 9

The stochastic derivation of regional curve standardization (RCS) for the tree-ring widths. RCS was based on 67%*n* spline fits to the mean values of all series. The bars show the standard deviation of each data according to the mean growth data

develop the chronology. First, the natural growth trend was associated with a negative exponential that was the best fit for the ring-width series. The R^2 coefficient of determination was calculated for each fit by an iterative method with a stoppage criterion of zero slope. The negative exponential with the largest amount of variance explained R^2 was then used for standardizing the ring width series (Fig. 11).

4.3. Chronologies

After the tree-ring indices were computed by division and subtraction, we observed an apparent heteroscedasticity in the imbuia series. To correct for this, it is a common practice to stabilize the variance using the natural growth trend removal by division (Helama et al., 2004).

Figure 12 presents the four versions of chronology developed based on PCA (both as ratios and reconstructed without the first PC), SP67, RCS, and STD. The chronologies show similar variations, despite the differences at the ends. The same kind of result was observed by Zhang and Chen (2017) using ensemble empirical mode decomposition (EEMD) to remove the tree growth trend. The chronology versions obtained by SP67 and STD were very similar ($r = 0.95, p < 0.01$), and the two PCA

versions were also similar ($r = 0.83, p < 0.01$, Table 1). However, these similarities are not as evident between the chronology versions obtained by SP67, RCS, and STD and the two PCA versions ($r = 0.66\text{--}0.86, p < 0.01$). This may be explained by the differences at the ends. As discussed by Zhang and Chen (2017), the SP67, RCS, and STD methods fail to fit the trend for the tree-ring series at the end, as shown in Figs. 8, 10, and 11.

The power spectra of the tree-ring chronologies computed with different detrending methods show a slight distortion in the low-frequency part of the spectrum (Fig. 13). The first-order autocorrelations (AR1) are expected to indicate the low-frequency variability in chronologies (Cook and Peters 1997). As described by Helama et al. (2004), our results also showed that AR1 are lowest with the most flexible standardization curves (SP67) and are highest with stiffer functions, PCA, RCS, and STD (Table 2). Our results confirmed that RCS is a robust method for preserving long-period variations in the tree-ring chronologies ($AR1 = 0.81$), as discussed by several authors (Cook & Peters, 1997; Helama et al., 2002, 2004; Melvin & Briffa, 2014b). Moreover, the two PCA versions presented higher AR1 values than SP67 and STD. In this sense, it seems that most of the low-frequency variability in each ring series is

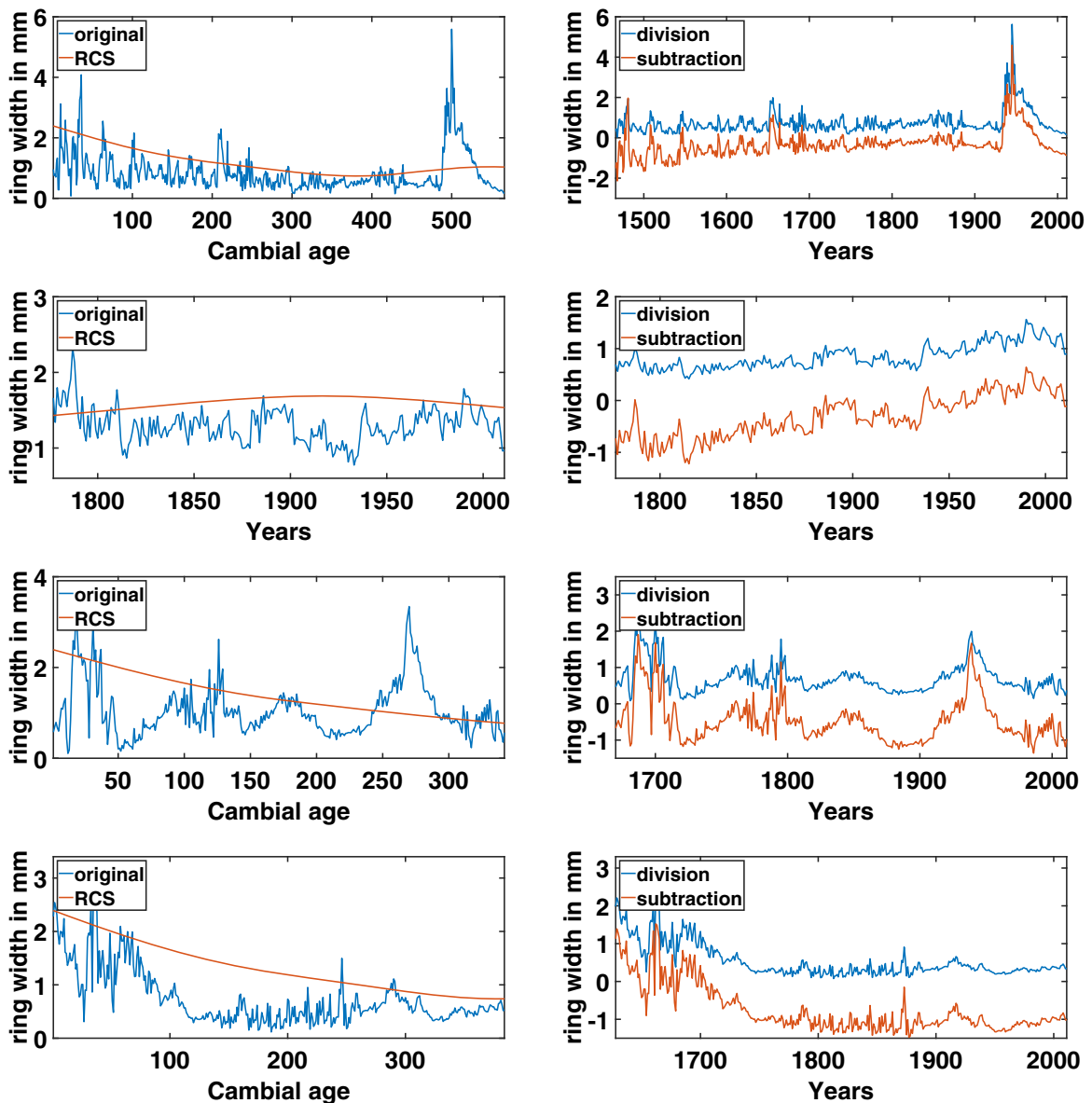


Figure 10

Example of four data sets, each one standardized using the RCS curve shown in Fig. 9 in two ways: by division and by subtraction

preserved by the chronology based on the PCA reconstruction without the first PC.

The removal of natural growth trends has been a subject of much research, especially for dendroclimatic studies, where the aim is to reconstruct past climatic variability. It is important to preserve low-frequency signals as they contain climatic signals present in the tree-ring series. Climate signals, which

have periods longer than the tree-ring series length, are impossible to be removed by all detrending methods. For example, our longest tree-ring series has 566 years; therefore, any climate signals which have periods longer than these 566 years are not removed by just fitting a curve on it (Zhang & Chen, 2017). The most common methods for preserving low- to medium-frequency signals are RCS (Briffa &

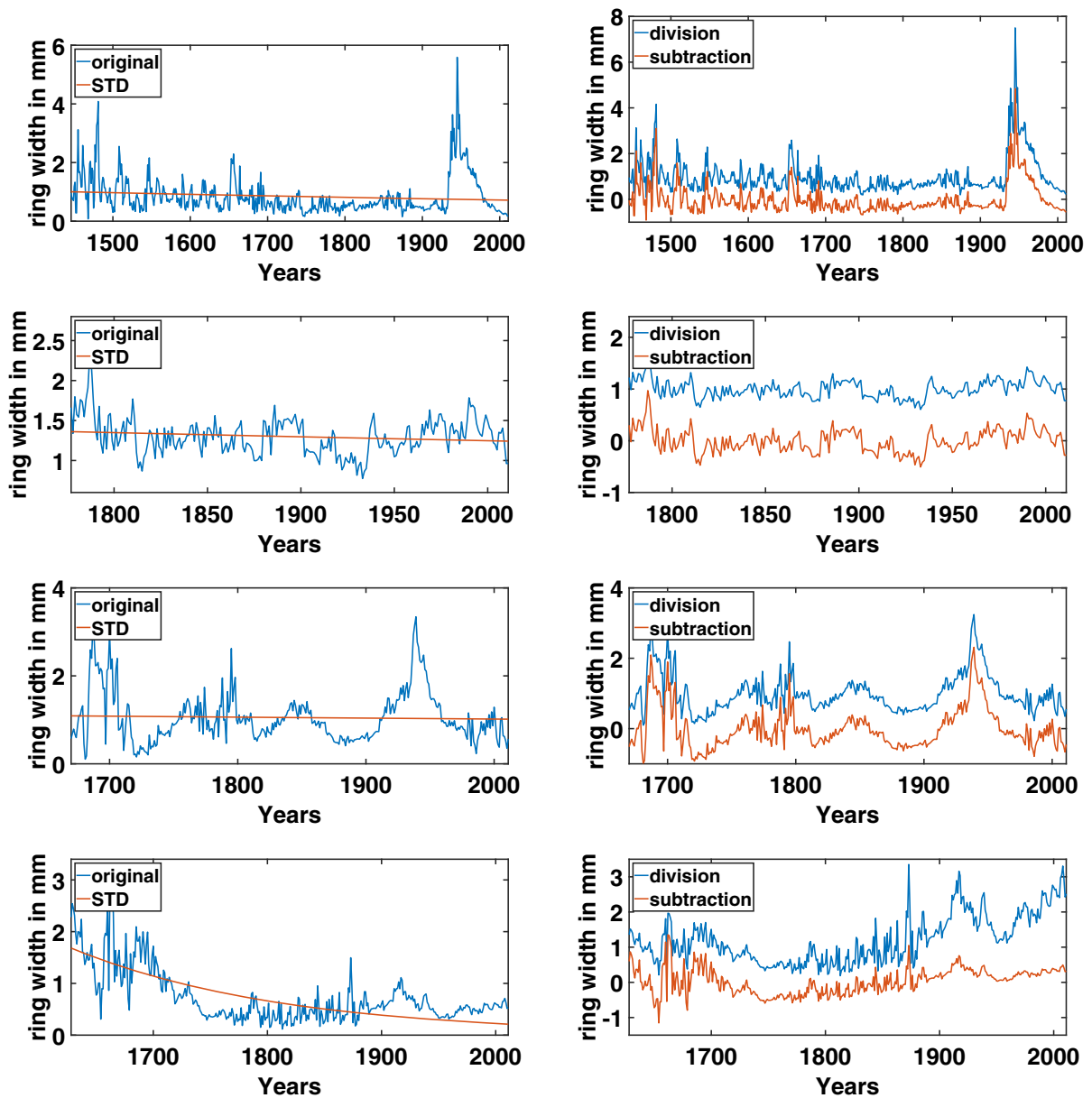


Figure 11
Four examples of the linear or negative exponential curve (STD) as growth model trend results

Melvin, 2011), EEMD (Zhang & Chen, 2017), and STD (Cook & Peters, 1997). In RCS, each series of tree indices is set relative to the RCS fitting, which enables the RCS to preserve medium-frequency variance (Melvin & Briffa, 2014b). If the main goal is the preservation and interpretation of long-time-scale variance in tree-ring chronologies, other

methods, such as flexible spline functions, end up removing part of those long-timescale variances, and consequently, are obviously not recommended for applications in which the low frequencies might contain relevant information. In summary, the use of detrending methods needs caution, and their limitations should not be overlooked, because it may lead

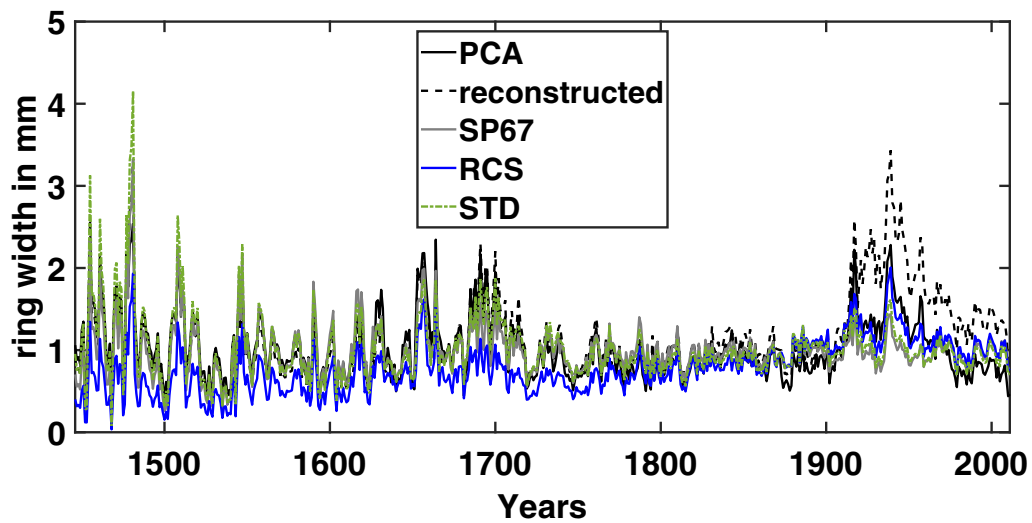


Figure 12

The chronologies developed based on PCA (both PCA as ratios and reconstructed without the first PC), 67%*n* spline curve (SP67), regional curve standardization (RCS), and linear or exponential curve (STD)

Table 1

Coefficients of correlation (*r*) and *p* values (*p*) between the chronologies using growth trend models based on PCA (PCA as a growth trend and series reconstruction without the first PC) and the 67%*n* spline curve (SP67), regional curve standardization (RCS), and linear or exponential curve (STD)

	Growth trend model				
	PCA	Reconstructed	SP67	RCS	STD
PCA					
<i>r</i>	1	0.83	0.82	0.65	0.85
<i>p</i> < 0.01	y	y	y	y	y
Reconstructed					
<i>r</i>	0.83	1	0.66	0.85	0.69
<i>p</i> < 0.01	y	y	y	y	y

“y” stands for yes

to the misinterpretation of the climate variability in the tree-ring chronology.

5. Conclusions

In this article, the PCA method was introduced to remove the natural growth trends of *Ocotea porosa* (Nees & Mart) Barroso (imbuia) tree-ring series collected in General Carneiro city, southern Brazil.

For comparison, the tree-ring data was also detrended by the cubic spline method (SP67), by negative exponential or linear regression curve (STD), and by the regional curve standardization (RCS) method. The PCA detrending results are similar to the results obtained with these methods, especially with RCS and STD. Similar characteristics were observed in the frequency domain between the four versions of chronology detrended by PCA and the traditional methods. Only a slight distortion in the low-frequency part of the spectrum was observed. Our results showed that the PCA method can be used as an alternative to the traditional ones. The first PC identified by PCA can be considered as the natural growth trend for each tree-ring series. Therefore, one advantage of the method is that it can be used to remove more complicated trend patterns, while other methods usually assume a specific function. Still, at the same time, it can remove real patterns if the data includes a low number of samples. Another important result from our PCA using ALS is that we were able to deal with tree-ring series with different lengths using an optimization algorithm. Therefore, there is no need to re-dimension the tree-ring series for the same length, which would cause loss of data. In summary, the PCA fulfilled our purpose and showed some advantages: (1) it preserves low-frequency

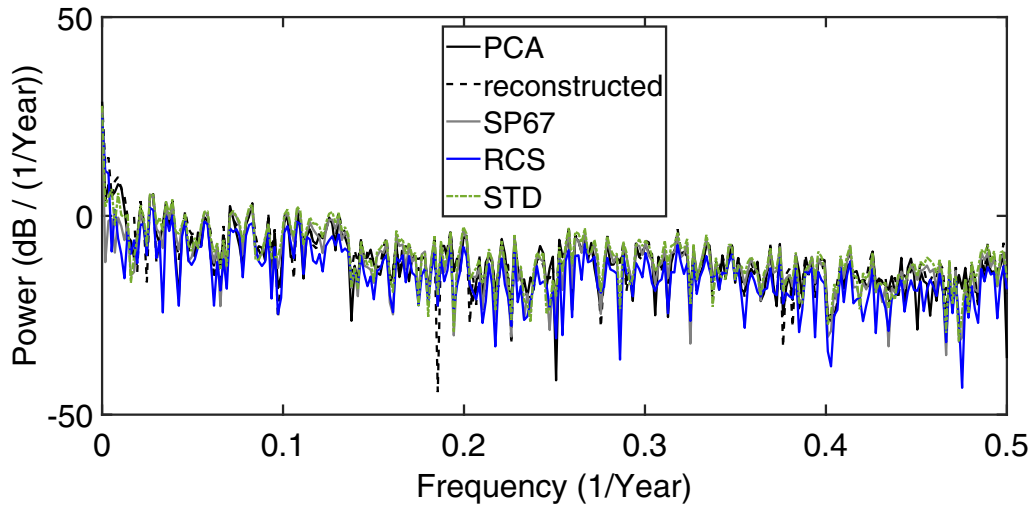


Figure 13
Spectral analysis of the four versions of chronology

Table 2

First-order autocorrelations (AR1) in chronologies constructed using growth trend models of PCA (both PCA as ratios and reconstructed without the first PC), 67%*n* spline curve (SP67), regional curve standardization (RCS), and linear or exponential curve (STD)

	Growth trend model				
	PCA	Reconstructed	SP67	RCS	STD
AR1	0.74	0.83	0.62	0.81	0.67

variance, (2) it applies a single adjustment to all collected samples (natural tree growth trend unique model), and (3) it minimizes the “end effect.”

Acknowledgements

We appreciate FAPESP support by the projects FAPESP–2009/02907-8 and CNPq (305249/2018-5) and CAPES and FVE (Fundação Vale Paraibana de Ensino) for the scholarships provided to Daniela Oliveira da Silva.

Author Contributions DOS contributed significantly to preparing the introduction, dataset and methodology, and conclusions. VK developed the idea of this manuscript. AP was

responsible for the acquisition of the dataset and its preconditioning. HGM contributed to the methodology description. TA collaborated in the revision of the text and description of the results. IRS contributed to the data analysis.

Funding

This research was supported by FAPESP–(2009/02907-8) and CNPq (305249/2018-5).

Availability of data

Data can be obtained from the authors upon request.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Appendix A: Principal Components Analysis

Let $X \in \mathbb{R}^{n \times m}$ be the data matrix whose m columns represent the variables x_1, x_2, \dots, x_m , zero-centered by hypothesis, and the n rows represent the observations of each of them, as described by

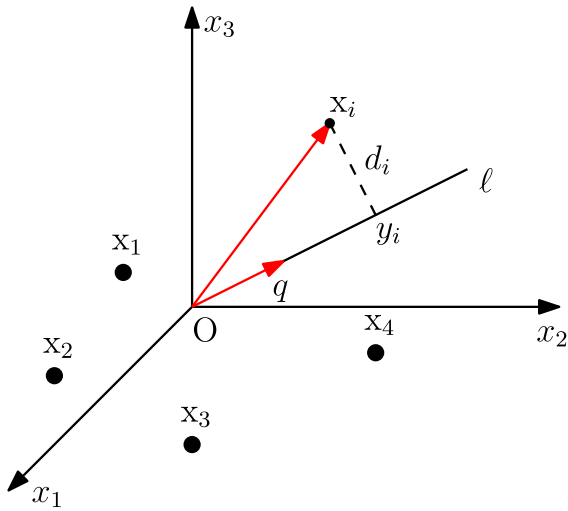


Figure 14

Representation of some observations, restricted to variables x_1, x_2 and x_3 , and the direction q that maximizes the variance of the projections y_i

$$X = \begin{matrix} & x_1 & x_2 & \dots & x_m \\ \begin{matrix} \text{1st} \\ \text{2nd} \\ \vdots \\ \text{n-th} \end{matrix} & \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} & = & \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \end{matrix}, \quad (1)$$

where $X_i \in \mathbb{R}^m$ is a row vector that corresponds to the i th observation, with $1 \leq i \leq n$. The idea in the standard derivation of PCA is to find a direction determined by a unit vector $q \in \mathbb{R}^m$ such that the variance of projections y_i associated to observations X_i is maximum, as illustrated in Fig. 14.

The projection of X_i onto the unit vector q is given by $y_i = X_i q$, so that the variance of all projections is defined by

$$V = \frac{1}{n-1} \sum_{i=1}^n y_i^2 = q^T \underbrace{\left[\frac{1}{n-1} \sum_{i=1}^n X_i^T X_i \right]}_{C_X} q = q^T C_X q, \quad (2)$$

where $C_X \in \mathbb{R}^{m \times m}$ is the covariance matrix associated to the original set of variables and can be rewritten in a more convenient way as follows:

$$C_X = \frac{1}{n-1} \sum_{i=1}^n X_i^T X_i = \frac{1}{n-1} X^T X. \quad (3)$$

The covariance matrix C_X has some properties that form the basis to the understanding of PCA, and, in order to visualize them, it is necessary to expand C_X in Eq. (3) considering that the observations are of the form $X_i = [x_{i1} \ x_{i2} \ \dots \ x_{im}]$, resulting in

$$C_X = \begin{bmatrix} \text{var}(x_1) & \text{cov}(x_1, x_2) & \dots & \text{cov}(x_1, x_m) \\ \text{cov}(x_2, x_1) & \text{var}(x_2) & \dots & \text{cov}(x_2, x_m) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(x_m, x_1) & \text{cov}(x_m, x_2) & \dots & \text{var}(x_m) \end{bmatrix}. \quad (4)$$

Equation (4) shows that C_X is a symmetric matrix due to the symmetry of covariance, and therefore, it has an orthonormal set of m eigenvectors, and it is orthogonally diagonalizable, so that there is an orthogonal matrix $Q \in \mathbb{R}^{m \times m}$ such that

$$D_X = Q^T C_X Q \iff C_X = Q D_X Q^T, \quad (5)$$

where $D_X = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$ is a diagonal matrix with the eigenvalues of C_X ordered in such a way that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$ and $Q = [q_1 \ q_2 \ \dots \ q_m]$, whose column $q_i \in \mathbb{R}^m$ is the respective eigenvector of λ_i . Besides that, C_X is a positive semidefinite matrix such that if $y \in \mathbb{R}^m$ is a nonzero vector, then

$$y^T C_X y = \frac{1}{n-1} \sum_{i=1}^n (X_i y)^2 \geq 0,$$

where equality can only be obtained when $X_i = 0$ since $y \neq 0$. As a consequence, the eigenvalues of C_X are nonnegative, which is consistent because they represent the variance of each principal component, as will be seen later. Finally, note that the diagonal elements of the matrix correspond to the variances of each variable x_1, x_2, \dots, x_m , whereas the off-diagonal elements correspond to all possibilities of covariance between these variables.

In order to maximize the variance in Eq. (2), it is necessary to solve an optimization problem involving the quadratic form with one constraint, as stated in

$$\begin{aligned} \max_q \quad & V = q^T C_X q \\ \text{subject to} \quad & q^T q = 1 \end{aligned} \quad (6)$$

The standard way to approach such a problem is by the Lagrange multiplier method that allows one to combine an objective function with a sequence of constraints, producing the Lagrange function, which in this case, is given by

$$\mathcal{L}(q, \lambda) = q^T C_X q - \lambda(q^T q - 1). \quad (7)$$

By differentiating (7) and equating the result to zero, considering that $\frac{\partial q^T q}{\partial q} = 2q^T$ and $\frac{\partial q^T C_X q}{\partial q} = 2q^T C_X$, one gets

$$2q^T C_X - \lambda 2q^T = 0 \iff C_X q = \lambda q. \quad (8)$$

Equation (8) is a major result since it indicates that the direction q that maximizes the variance V is given by the eigenvectors of the covariance matrix C_X , and as a consequence, the numerical value of V is given by the respective eigenvalue λ , as indicated by

$$V = q^T C_X q = q^T \lambda q = \lambda \underbrace{q^T q}_1 = \lambda. \quad (9)$$

Nevertheless, since the matrix C_X has a set of m orthonormal eigenvectors, which one maximizes V ? This question can be answered by observing Eq. (9), where one can conclude that the eigenvector with the largest eigenvalue is the one that maximizes it. Since the eigenvalues are ordered, one can make the following statements:

- (a) q_1 is the direction of largest variance λ_1 of the original data set;
- (b) q_2 is the direction of the second-largest variance λ_2 and so on.

An important point that deserves attention is the perpendicular distances d_i of observations X_i relative to line ℓ , as illustrated in Fig. 14. Mathematically, d_i is given by

$$d_i = \|X_i^T - y_i q\|_2 = \|X_i^T - (X_i q) q\|_2. \quad (10)$$

The goal is to verify that the variance D of these distances is minimized when the variance of the projections V is maximized. To do this, it is necessary to square Eq. (10), allowing us to express d_i^2 in matrix form as follows:

$$d_i^2 = \|X_i^T - (X_i q) q\|_2^2 = X_i X_i^T - (X_i q)^2. \quad (11)$$

From Eq. (11) one can obtain a closed expression for variance D , as in

$$D = \frac{1}{n-1} \sum_{i=1}^n d_i^2 = \frac{1}{n-1} \sum_{i=1}^n X_i X_i^T - V. \quad (12)$$

Note that Eq. (12) indicates that when V is maximum, D is minimum due to the minus sign, since the other term is constant. This explains why the eigenvectors seem to fit the cloud of observations when they overlap in the same plot.

With the largest variance directions determined by the eigenvalues of C_X , all that remains is to project the observations X_i in the eigenspace of C_X . Since the projection X_i in the direction of q_j is given by $y_{ij} = X_i q_j$, by doing this for all eigenvectors indexed by $1 \leq j \leq m$, there is a new representation of X_i given by the row vector $Y_i = [X_i q_1 \ X_i q_2 \ \dots \ X_i q_m]$. Thus, for all observations indexed by $1 \leq i \leq n$, one can get a new representation for the original data set, denoted by the matrix $Y \in \mathbb{R}^{n \times m}$, as given by

$$\begin{aligned} Y &= \begin{matrix} & y_1 & y_2 & \dots & y_m \\ \begin{bmatrix} X_1 q_1 & X_1 q_2 & \dots & X_1 q_m \\ X_2 q_1 & X_2 q_2 & \dots & X_2 q_m \\ \vdots & \vdots & \ddots & \vdots \\ X_n q_1 & X_n q_2 & \dots & X_n q_m \end{bmatrix} \end{matrix} \\ &= \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \begin{bmatrix} q_1 & q_2 & \dots & q_m \end{bmatrix} = XQ, \end{aligned} \quad (13)$$

where the new variables y_1, y_2, \dots, y_m are named *principal components*. It should be noted that the entry $y_{ij} = X_i q_j$ corresponds to the *score* of the i th observation in the j th principal component, whereas the eigenvector q_j is the so-called *coefficient* vector of this component. This nomenclature comes from the fact that the entry y_{ij} is written as a linear combination whose coefficients are the entries of q_j , as follows:

$$y_{ij} = q_{1j}x_{i1} + q_{2j}x_{i2} + \dots + q_{mj}x_{im} = \sum_{r=1}^m q_{rj}x_{ir}. \tag{14}$$

In general, the new representation of the original data set as given by the Eq. (13) is not made considering all the columns of matrix Q . The idea is to reduce the dimension of the original data set disregarding the columns that represent the principal components with the smallest variances, resulting in a new set of variables that represents the original ones with a certain degree of accuracy. For example, suppose that only the first $k \ll m$ principal components are used to represent a data set in an m -dimensional space. The accuracy of this representation is defined in terms of how much the total variance remains in the k -dimensional space and is given as a percentage as follows:

$$\frac{100}{\text{tr}(C_X)} \cdot \sum_{j=1}^k \lambda_j = 100 \cdot \frac{\lambda_1 + \lambda_2 + \dots + \lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_m}, \tag{15}$$

where it is quite common to get high values with only two principal components. Note that with Eq. (13), it becomes possible to obtain the original data matrix X by multiplying both sides on the right by Q^T , as indicated by

$$YQ^T = X \underbrace{QQ^T}_I \iff X = YQ^T. \tag{16}$$

One of the main results of PCA can be noticed when calculating the covariance matrix C_Y associated to the principal components, as given by

$$C_Y = \frac{1}{n-1} Y^T Y = Q^T C_X Q. \tag{17}$$

Substituting Eq. (5) into Eq. (17), it turns out that C_Y is a diagonal matrix,

$$C_Y = D_X = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_m \end{bmatrix}, \tag{18}$$

whereby matrix equality follows that $\text{var}(y_j) = \lambda_j$, that is, the eigenvalues of C_X are the variances of the principal components. Moreover, note that there is no

covariance between them, indicating that they are unrelated.

A.1 Alternating Least Squares

The alternating least squares (ALS) is a matrix factorization algorithm that plays an important role in the context of collaborative filtering (CF) in recommendation systems. Specifically, CF is a technique used to make predictions about the interests of a user based on their preferences over a set of available items (Zhou et al., 2008), which can be movies and/or songs in streaming platforms, for example. In this sense, to illustrate the ALS, suppose that a video streaming platform has n_u users and n_m movies available which can be rated. In that scenario, it is convenient to define a user-movie matrix $R \in \mathbb{R}^{n_u \times n_m}$ whose entry r_{ij} represents the rating score of movie j by user i . Unfortunately, the matrix R has a lot of missing entries since a user only rates a subset of the available movies. The purpose of a recommendation system is to estimate these missing entries based on the previous ratings of the users.

The idea is to perform a low-rank approximation of the user-movie matrix R by introducing a user feature matrix $U \in \mathbb{R}^{n_f \times n_u}$ and a movie feature matrix $M \in \mathbb{R}^{n_f \times n_m}$ in such a way that $r_{ij} \approx u_i^T m_j \forall i \in \{1, \dots, n_u\}, \forall j \in \{1, \dots, n_m\}$ or $R \approx U^T M$, where $u_i \in \mathbb{R}^{n_f}$ and $m_j \in \mathbb{R}^{n_f}$ are the user and movie feature vectors, respectively (Zhou et al., 2008). The n_f is the dimension of the feature space. In this sense, the matrices U and M are determined by minimizing the least squares error of the known ratings, i.e.,

$$\min_{U, M} f(U, M) = \min_{U, M} \sum_{(i,j) \in I} (r_{ij} - u_i^T m_j)^2 + \lambda \left(\sum_i n_{u_i} \|u_i\|^2 + \sum_j n_{m_j} \|m_j\|^2 \right), \tag{19}$$

where I denotes the index set of known ratings, n_{u_i} the number of ratings of user i , and n_{m_j} the number of ratings of movie j . The term involving λ in Eq. (19) is a regularization term to prevent overfitting in ALS (Zhou et al., 2008). It is noteworthy that the objective function in Eq. (19) is non-convex due to the term $u_i^T m_j$, but if one of the matrices is fixed, U for example, then it becomes a convex function of M and

vice versa. In this way, the approach used to solve this minimization problem is to fix U and optimize M , then fix M and optimize U until convergence is established by some stopping criterion. Actually, this fact explains the word *alternating* in the ALS name. Specifically, to do the optimization, it suffices to perform a matrix differentiation in Eq. (19) considering that one of the matrices is held constant and equates the result to zero. In this sense, suppose that M is given, then the columns of matrix U will be

$$u_i = (M_{I_i} M_{I_i}^T + \lambda n_{u_i} E)^{-1} M_{I_i} R^T(i, I_i), \quad i \in \{1, \dots, n_u\}, \quad (20)$$

where E is the $n_f \times n_f$ identity matrix, $M_{I_i} = [m_{ij}]_{j \in I_i}$, and $R(i, I_i)$ is the i th row vector of R considering only the columns with indexes in I_i (Zhou et al., 2008). Similarly, by carrying out the same procedure, but considering that U is given, it results

$$m_j = (U_{I_j} U_{I_j}^T + \lambda n_{m_j} E)^{-1} U_{I_j} R(I_j, j), \quad j \in \{1, \dots, n_m\}, \quad (21)$$

where $U_{I_j} = [u_i]_{i \in I_j}$, and $R(I_j, j)$ is a column vector of R considering only the rows with indexes in I_j , which represents the set of users who rated movie j .

By using Eqs. (20) and (21) in the ALS algorithm, one can find the matrix $\tilde{R} = U^T M$ that fits the known entries of matrix R and fills the missing ones.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Amoroso, M. M., Daniels, L. D., & Larson, B. C. (2012). Temporal patterns of radial growth in declining *Austrocedrus chilensis* forests in Northern Patagonia: The use of tree-rings as an indicator of forest decline. *Forest Ecology and Management*, 265, 62–70.
- Autin, J., Gennaretti, F., Arseneault, D., & Bégin, Y. (2015). Biases in RCS tree ring chronologies due to sampling heights of trees. *Dendrochronologia*, 36, 13–22.
- Berrêdo, V. D. (2015). Vulnerabilidade de biomassa às mudanças climáticas: O caso da mata atlântica no estado do paran . Ph.D. thesis, UFRJ/COOPE. Programa de Planejamento Energ tico, Rio de Janeiro.
- Biondi, F., & Qeadan, F. (2008). A theory-driven approach to tree-ring standardization: Defining the biological trend from expected basal area increment. *Tree-Ring Research*, 64(2), 81–96. <https://doi.org/10.3959/2008-6.1>.
- Brienen, R. J. W., & Zuidema, P. A. (2005). Relating tree growth to rainfall in Bolivian rain forests: A test for six species using tree ring analysis. *Ecophysiology*, 146(1), 1.
- Briffa, K., Bartholin, T., Eckstein, D., Jones, P. D., Karl n, W., Schweingruber, F. H., & Zetterberg, P. (1990). A 1,400-year tree-ring record of summer temperatures in Fennoscandia. *Nature*, 346, 434–439.
- Briffa, K. R., Jones, P. D., Schweingruber, F. H., Karl n, W., & Shiyatov, S. G. (1996). Tree-ring variables as proxy-climate indicators: Problems with low-frequency signals. In P. D. Jones, R. S. Bradley, & J. Jouzel (Eds.), *Climatic variations and forcing mechanisms of the last 2000 years* (pp. 9–41). Berlin: Springer.
- Briffa, K. R., & Melvin, T. M. (2011). *A closer look at regional curve standardization of tree-ring records: Justification of the need, a warning of some pitfalls, and suggested improvements in its application* (pp. 113–145). Dordrecht: Springer.
- Carvalho, P. E. R. (1994). *Esp cies florestais Brasileiras: Recomenda es silviculturais, potencialidades e uso da madeira*. Colombo: EMBRAPA.
- Carvalho, P. E. R. (2003). *Esp cies arb reas Brasileiras: Embrapa informa o tecnol gica* (Vol. 1). Colombo: Embrapa Florestas.
- CNCFLORA. (2012). Ocotea porosa in lista vermelha da flora brasileira vers o 2012.2 centro nacional de conserva o da flora. [http://cncflora.jbrj.gov.br/portal/pt-br/profile/Ocotea porosa](http://cncflora.jbrj.gov.br/portal/pt-br/profile/Ocotea%20porosa)
- Cook, E., Briffa, K., Shiyatov, S., & Mazepa, V. (1990). Tree-ring standardization and growth-trend estimation. Methods of dendrochronology: Applications in the environmental sciences.
- Cook, E. R. (1985). A time series analysis approach to tree ring standardization. Master's thesis, University of Arizona.
- Cook, E. R., & Peters, K. (1997). Calculating unbiased tree-ring indices for the study of climatic and environmental change. *The Holocene*, 7(3), 361–370.
- Cosmo, N. L., Gogosz, A. M., Nogueira, A. C., Bona, C., & Kuniyoshi, Y. S. (2009). Morfologia do fruto, da semente e morfo-anatomia da da pl ntula de *Vitex megapotamica* (Spreng.) Moldenke (Lamiaceae). *Acta Botanica Brasiliica*, 23(52), 389–397.
- Dray, S., & Josse, J., (2015). Principal component analysis with missing values: A comparative survey of methods. *Plant Ecology*, 216(5), 657–667. <https://hal.archives-ouvertes.fr/hal-01260054>
- Enright, N. J. (1984). Principal components analysis of tree-ring/climate relationships in white spruce (*Picea glauca*) from Schefferville, Canada. *Journal of Biogeography*, 11(4), 353–361.
- Fekedulegn, D., Colbert, J., Hicks, R., & Schuckers, M. (2002). *Coping with multicollinearity: An example on application of principal components regression in dendroecology*. Pap: USDA Forest Service Northeast. Res. Station. Res.
- Flower, A., & Smith, D. (2011). A dendroclimatic reconstruction of June–July mean temperature in the northern Canadian rocky mountains. *Dendrochronologia*, 29, 55–63.
- Fritts, H. C. (1976). *Tree rings and climate*. Tucson: The University of Arizona Press.
- General Carneiro, P. (2020). Dados gerais. [Online; Accessed 20-Outubro-2019]. <http://www.generalcarneiro.pr.gov.br/municipio/dados-gerais/>

- Helama, S., Lindholm, M., Timonen, M., & Eronen, M. (2004). Detection of climate signal in dendrochronological data analysis: A comparison of tree-ring standardization methods. *Theoretical and Applied Climatology*, *79*, 239–54.
- Helama, S., Lindholm, M., Timonen, M., Meriläinen, J., & Eronen, M. (2002). The supra-long Scots pine tree-ring record for Finnish Lapland: Part 2, interannual to centennial variability in summer temperatures for 7500 years. *The Holocene*, *12*(6), 681–687.
- Helama, S., Melvin, T. M., & Briffa, K. R. (2017). Regional curve standardization: State of the art. *The Holocene*, *27*(1), 172–177.
- Huntington, E. (1914). The climatic factor as illustrated in Arid America. No. 192. Carnegie Institution of Washington, Washington.
- Klein, R. M. (1963). Importância prática da fitossociologia para a silvicultura brasileira. *Anais do I Simpósio de reflorestamento da região da araucária* (pp. 1–61). Brasil: Curitiba.
- Lindholm, M., Eronen, M., Timonen, M., & Meriläinen, J. (1999). A ring-width chronology of Scots pine from northern Lapland covering the last two millennia. *Annales Botanici Fennici*, *36*, 119–126.
- Lorensi, C., & Prestes, A. (2016). Dendroclimatological reconstruction of spring-summer precipitation for Fazenda Rio Grande, PR, with samples of *Araucaria angustifolia* (Bertol.) Kuntze. *Revista Árvore*, *40*, 347–354.
- Maes, S. L., Vannoppen, A., Altman, J., Van den Bulcke, J., Decocq, G., De Mil, T., et al. (2017). Evaluating the robustness of three ring-width measurement methods for growth release reconstruction. *Dendrochronologia*, *46*, 67–76.
- Marchesan, R., Mattos, P. P., Bortoli, C., & Rosot, N. C. (2006). Caracterização física, química e anatômica da madeira de *Ocotea porosa* (Nees & C. Mart.) Barroso. Tech. Rep. 161, Colombo, PR, comunicado Técnico.
- Melvin, T. M., & Briffa, K. R. (2014a). Crust: Software for the implementation of regional chronology standardisation: Part 1. signal-free rcs. *Dendrochronologia*.
- Melvin, T. M., & Briffa, K. R. (2014b). CRUST: Software for the implementation of regional chronology standardisation: Part 1. Signal-free RCS. *Dendrochronologia*, *32*(1), 7–20.
- Miller, D. L., Mora, C. I., Grissino-Mayer, H. D., Mock, C. J., Uhle, M. E., & Sharp, Z. (2006). Tree-ring isotope records of tropical cyclone activity. *Proceedings of the National Academy of Sciences*, *103*(39), 14294–14297.
- Piermattei, A., Crivellaro, A., Carrer, M., & Urbinati, C. (2014). The “blue ring”: Anatomy and formation hypothesis of a new tree-ring anomaly in conifers. *Trees*, *29*(2), 613–620.
- Prestes, A. (2009). Relação sol-terra estudada através de anéis de crescimento de coníferas do holoceno recente e triássico. Ph.D. thesis, Instituto Nacional de Pesquisas Espaciais, São José dos Campos.
- Regazzi, A. J. (2000). Análise Multivariada, notas de aula INF 766. Departamento de Informática da Universidade Federal de Viçosa.
- Reitz, R., Klein, R. M., & Reis, A. (1978). *Projeto madeiras de Santa Catarina*. Itajaí: Herbário Barbosa Rodrigues.
- Shi, F., Yang, B., Linderholm, H. W., Seftigen, K., Yang, F., Yin, Q., et al. (2020). Ensemble standardization constraints on the influence of the tree growth trends in dendroclimatology. *Climate Dynamics*, *54*(7–8), 3387–3404.
- Speer, J. H. (1971). *Fundamentals of Tree-Ring Research*. Library of Congress Cataloging-in-Publication Data.
- Stepka, T. F. (2013). Modelagem do crescimento e dendro-cronologia em árvores nativas de *Araucaria angustifolia*, *Cedrela fissilis* e *Ocotea porosa* no sul do Brasil. Ph.D. thesis, Universidade Federal do Paraná, Paraná.
- Stuiver, M., & Quay, P. D. (1980). Changes in atmospheric carbon-14 attributed to a variable sun. *Science*, *207*(4426), 11–19.
- Tomazello Filho, M., Lisi, C. S., Hansen, N., & Cury, G. (2004). Anatomical features of increment zones in different tree species in the state of São Paulo, Brazil. *Scientia Forestalis*, *66*, 46–55.
- Zhang, X., & Chen, Z. (2017). A new method to remove the tree growth trend based on ensemble empirical mode decomposition. *Trees*, *31*, 405–413.
- Zhou, Y., Wilkinson, D., Schreiber, R., & Pan, R. (2008). Large-scale parallel collaborative filtering for the Netflix prize (pp. 337–348).

(Received September 21, 2020, revised May 26, 2021, accepted May 27, 2021, Published online June 14, 2021)