

# Quantifying Spatial and Temporal Relationships Among Tree-Ring Records

Megan Heyman<sup>1</sup>, Scott St George<sup>2</sup> and Snigdhansu Chatterjee<sup>3</sup>

<sup>1</sup>*Department of Mathematics,*

*Rose-Hulman Institute of Technology, USA*

<sup>2</sup>*Department of Geography, Environment, and Society*

*University of Minnesota, USA*

<sup>3</sup>*School of Statistics*

*University of Minnesota, USA*

Received: 04 May 2020; Revised: 12 July 2020; Accepted: 13 July 2020

---

## Abstract

Tree growth rings contain yearly information about climate, extreme weather events, and other growing conditions. In this analysis, we model the relationship strength between tree-ring records with respect to location and time. We employ the discrete wavelet transformation on the ring width records in order to de-correlate the observations within each series while simultaneously retrieving time-scale information. Our model then describes correlations among the resulting wavelet coefficients at different temporal scales by distance. Statistical inference through a new version of the wild bootstrap indicates that the relationship strength decreases linearly as record pair distance increases, but the slopes differ across temporal scales.

*Key words:* Wavelet; Bootstrap; U-Statistics; Tree ring records.

**AMS Subject Classifications:** 62G05, 62G09

---

## 1. Introduction

Because instrumental measurements of temperature, precipitation, and other aspects of Earth's climate typically span only the past century or so (Harris *et al.*, 2020), we rely on surrogate information recovered from natural climate archives to extend our perspective on recent changes. Fallen snow accumulates on the surface of continental ice sheets or alpine and arctic glaciers, and builds year-by-year a frozen archive of atmospheric chemistry that can extend back several hundreds of thousands of years (Steiger *et al.*, 2017; Brook and Buizert, 2018). Sediment that sinks down to the bottom of lakes often traps windblown pollen and the remains of waterborne organisms, which in turn reflect environmental conditions across the broader region (Holmes *et al.*, 2016; Sandeep *et al.*, 2017). And the water that flows underground to form cave deposits leaves behind physical and chemical clues within the

resulting mineral structures (Arienzo *et al.*, 2017; Affolter *et al.*, 2019). Over the past several decades, networks of proxy climate records have served as the foundation for quantitative estimates of past temperature, drought severity, and other key climate metrics at the local, regional, continental, or global scale (Trouet *et al.*, 2018; Neukom *et al.*, 2019; Cook *et al.*, 2020).

Within the so-called ‘Common Era’ (the past two thousand years), the growth rings of trees are, by far, the most dominant source of past climate information (Emile-Geay *et al.*, 2017). Every year, trees across the world’s temperate and boreal forests form a new layer of wood about their stem. One of the most obvious signs in nature documenting the passage of time, those rings also encode information about the tree’s immediate environment within their physical and chemical structure (St George and Ault, 2014). For trees growing in cold alpine or arctic forests, the width and wood density of their annual ring can be excellent surrogates for growing season temperatures (Esper *et al.*, 2018; St George and Esper, 2019). And for trees in warmer and drier environs, because narrow rings tend to follow dry weather, tree-ring records can be used to estimate past changes in precipitation or drought extending back hundreds of years or more (Granato-Souza *et al.*, 2019; Opała-Owczarek and Niedźwiedź, 2019).

Because tree rings and other proxies can extend our perspective on climate change farther back in time than instrumental weather observations, they offer new opportunities to evaluate the time evolution of the dominant modes of climate variability. Towards that purpose, the wavelet transform is commonly applied to study multiscale, nonstationary processes occurring across space and time within tree-ring records (Fan and Bräuning, 2017; McKenzie *et al.*, 2018; Kasatkina *et al.*, 2019).

Climate is a multi-scale phenomenon. Some quasi-periodic and oscillatory patterns like the El Niño Southern Oscillation (ENSO) are sub-decadal in periodicity, while others like the Atlantic multi-decadal oscillation (AMO) or the Pacific decadal oscillation (PDO) have longer time scales. In this paper, our primary goal is to verify that at different temporal scales, the correlation between tree-ring records decay smoothly over space in a climatically homogeneous region. To that end, in this paper we explore the spatio-temporal patterns of paleoclimate records, as exhibited by Ponderosa pine (*Pinus ponderosa* Douglas ex C. Lawson) tree-ring records from four states of the USA. We restrict to only this subspecies of trees in order to eliminate differences due to species variation, and restrict regionally so that we may consider a homogeneous, contiguous region where the climatic patterns are similar. Additional details are provided later in this paper.

Our principal approach is to consider each three ring record as an individual functional observation over time, that exhibits quasi-periodic and oscillatory patterns according to the climatic conditions the tree has experienced in the past. A discrete wavelet transformation of these records allows us to deconstruct such functional time series into various temporal scales. We then consider correlations between the wavelet coefficients from two different trees, at different scales, and study the pattern of such correlations as a function of distance between the two trees.

Notice that the data that we analyze here has complex dependency patterns, hence

it is non-trivial to conduct inference on a statistical model for how correlations, across various wavelet scales, decays over spatial distance. To address this issue, we propose a novel resampling scheme, that generalizes the well-known *wild or external* bootstrap scheme.

The rest of the paper is as follows: In Section 2 we provide a detailed description of the dataset, and then Section 3 provides an overview of the discrete wavelet transform. Next, we present a new approach to describe spatial relationships among time-series records in Section 4. Namely, we apply the discrete wavelet transformation on tree-ring records and calculate three versions of correlation between pairs of wavelet coefficients within each time-scale. Afterwards, it is necessary to re-format the data for analysis; this is described in Section 5. With the re-formatted data, we describe a novel approach to modeling relationship strength across temporal scales in Section 6. The proposed linear model utilizes time-scale, distance, and elevation information among the records to predict strength of relationship across all record pairs. We also describe differences among obtaining estimates utilizing least squares, median regression, and Huber’s M estimation techniques. Finally, statistical inference is performed via an extension of the wild bootstrap in Section 7. The wild bootstrap is modified to utilize external random variables which are generated per tree-ring record, instead of for each case within the dataset. The random variables are incorporated together within the method to help account for dependence among cases utilizing the same record(s) while simultaneously producing the sampling distributions for the coefficient estimates. An illustrative theoretical insight into the properties of the proposed extension of the wild bootstrap is presented in Section 9. We collect some concluding remarks and ideas about future research in Section 10.

## 2. Tree-Ring Record Description

Tree-ring datasets are typically presented in the structure of a ‘chronology’ — a composite series made by averaging together measurements of tree-ring width, wood density, isotopic composition, or other anatomical or biogeochemical variables from several dozen or more trees at the same location (Cook and Peters, 1997). Compared to the initial measurements made on samples taken from individual trees, tree-ring chronologies offer several advantages (Coulthard *et al.*, 2020). They are easier than sample- or tree-level data to incorporate into regional syntheses, they have been pre-processed to remove the confounding influence of tree size or age (Bunn *et al.*, 2004), and they are often adjusted to minimize or eliminate biologically-driven persistence (Kannenberg *et al.*, 2019).

The International Tree-Ring Databank (ITRDB), an open-access database maintained by the National Oceanic and Atmospheric Administration in Boulder, Colorado, is the largest archive of freely-available tree-ring data worldwide (Zhao *et al.*, 2019). The ITRDB, which was established in 1974 as a permanent repository for digital tree-ring measurements, includes more than four thousand tree-ring datasets (each composed of measurements made on one dozen to more than one hundred trees) from all continents except Antarctica. Here we use a subset of the St George and Ault (2014) dataset, which used the ITRDB’s holdings as the foundation for their network of more than 2,200 age-corrected, quality-controlled tree-ring width chronologies from the Northern Hemisphere. Each of these time series, which have annual resolution but varying start and end dates, reports yearly tree growth across

an entire forest stand as a unitless index of tree-ring width (RWI). Our analysis focuses on tree-ring width chronologies derived from Ponderosa pine (*Pinus ponderosa* Douglas ex C. Lawson) forests in the Four Corners region of the southwestern United States (Figure 1). Ponderosa pine (coded as PIPO) is the fourth most common source species for tree-ring width measurements housed by the ITRDB (St George and Ault, 2014) and a range that extends from northern Mexico to southern Canada. Across the United States, Ponderosa pine grow best under high precipitation and low evapotranspirative demand, but throughout its range there are important regional differences in the species' sensitivity to climate (McCullough *et al.*, 2017). For that reason, we restricted our analysis to only include those tree-ring chronologies from Arizona, New Mexico, Colorado, and Utah (yielding a total of 97 series).

A commonly referenced set of guidelines for investigators creating tree-ring records may be found in Cook (1987). Generally, at a given location, multiple trees are chosen to be measured. Within each tree, between 1 and 5 cores (cylinders of heartwood extracted from the trunk of the tree) are obtained in such a way that the tree is not vitally harmed. The entire set of growth rings, combining both early and late wood, is measured within each core. Then, a B-spline is used to remove a growth trend from each individual core. Finally, all of the de-trended core series are combined into a single ring-width index (RWI) for the location. Thus, the records of tree-ring widths are *unitless* and composed of many trees and cores. It is also common practice to further clean records before analysis, by only maintaining observations with strong coherence between cores within the record location at any given time point.

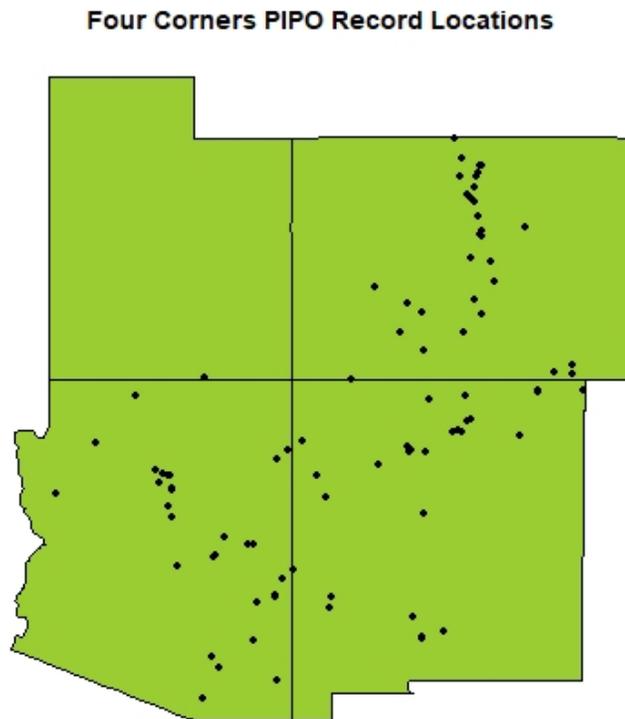
### 3. Discrete Wavelet Transform

In paleoclimatology and climate science in general, nearly all applications of wavelets are based on the continuous transform (Lau and Weng, 1995; Torrence and Compo, 1998). That method does not place any restriction on series length, but in principle should be applied only to data where the observations are made continuously through time. However, tree-ring records are discrete time series, with observations that are equally spaced and made at specific times (either once or twice within the growing season). For that reason, in this paper, we apply the discrete wavelet transform to a network of tree-ring width chronologies to test the potential of this other method to problems in climate science and paleoclimatology.

Our model aims to quantify spatial and temporal relationships *among* RWI records. However, we cannot ignore the temporal dependencies likely present *within* each record. We choose to employ wavelets to address the *within record* temporal dependency. Wavelets are an ideal choice since they overcome the limitations with respect to the time and frequency domains that are present in Fourier and windowed Fourier analyses. Let  $\mathbb{Z}$  be the set of integers  $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ . Suppose  $\psi \in \mathcal{L}_2(\mathbb{R})$  is a given function whose properties will be discussed later. Wavelets constitute the family of functions defined by

$$\psi_{j,k}(u) = 2^{j/2} \psi(2^j u - k), \quad j, k \in \mathbb{Z}. \quad (1)$$

For appropriate choices of  $\psi(\cdot)$ , these translations and dyadic dilations of  $\psi(\cdot)$  constitute an orthonormal basis of  $\mathcal{L}_2(\mathbb{R})$ . The dilation and shrinkage within the function allow a user to



**Figure 1: Ponderosa pine record locations. Locations of the 97 Ponderosa pine records in the Four Corners states analyzed by our model.**

discover signals which change over the time domain. In many situations, a second function, denoted as  $\phi(\cdot)$  and also called the *scaling function* since it may be derived from  $\psi(\cdot)$ , is used in conjunction with subsets of the family of wavelet functions  $\{\psi_{j,k}(\cdot)\}$  for representing functions in a relatively simpler and parsimonious way. The very basic overview of wavelets provided here is necessarily brief; more comprehensive discussions and many additional points of interest may be found in Daubechies (1992); Hubbard (1998); Ogden (1997); Vidakovic (1999); Wasserman (2006); Percival and Walden (2006); Nason (2008) and several other places.

In our case, we are working with discrete time points and are specifically choosing to employ the discrete wavelet transform (DWT). The DWT is a transformation of a vector of data,  $\mathbf{x}$ , which utilizes an orthonormal basis. Assume that the data vector,  $\mathbf{x} = (x_1, \dots, x_T)$ , is of length  $T = 2^J$  where  $J \in \mathbb{Z}^+ = \{1, 2, \dots\}$ . Consider the collection of functions

$$\left\{ \phi(\cdot), \psi_{j,k}(\cdot); j = 0, 1, \dots, J - 1, k = 1, 2, \dots, 2^j \right\}.$$

We evaluate these  $T = 2^J$  functions at each of  $t = 1, 2, \dots, T$ , and construct the matrix  $\mathbf{W}$ . The first row of  $\mathbf{W}$  is  $(\phi(1), \phi(2), \dots, \phi(T))$ . Then, the remainder of the matrix is generated from the bottom-up. That is, the  $T^{\text{th}}$  (last) row is given by  $(\psi_{0,1}(1), \psi_{0,1}(2), \dots, \psi_{0,1}(T))$ .

Rows  $T - 2$  and  $T - 1$  are  $(\psi_{1,1}(1), \psi_{1,1}(2), \dots, \psi_{1,1}(T))$  and  $(\psi_{2,1}(1), \psi_{2,1}(2), \dots, \psi_{2,1}(T))$ , respectively. Generally, sets of  $2^j$  rows are generated together for evaluations of the wavelet functions at time-scales,  $j = 0, 1, \dots, J - 1$ . A concrete example demonstrating this construction is provided in Section 3.1. With some amount of algebra and exploiting known properties of wavelets, for appropriate choice of the function  $\psi(\cdot)$  it can be seen that the matrix  $\mathbf{W}$  is an orthonormal basis system of  $\mathbb{R}^T$ , that further has all the multi-resolution properties for wavelets as defined in (1).

Based on this, the DWT of the data vector,  $\mathbf{x}$  is

$$\boldsymbol{\gamma} = \mathbf{W}\mathbf{x} \quad (2)$$

where the vector,  $\boldsymbol{\gamma}$ , contains the wavelet coefficients. One important item of note is that the wavelet coefficients convey coarse and fine level information about the series. Specifically,  $\boldsymbol{\gamma}$  may be written element-wise as

$$\boldsymbol{\gamma} = \left( \gamma_0 \quad \gamma_{J-1,1} \quad \dots \quad \gamma_{J-1,2^{J-1}} \quad \dots \quad \gamma_{1,1} \quad \gamma_{1,2} \quad \gamma_{0,1} \right)^T. \quad (3)$$

Here,  $\gamma_0$  is the coarse level coefficient corresponding to the scaling function  $\phi(\cdot)$ , and  $\gamma_{j,k}$  is the fine level coefficient corresponding to the wavelet function  $\psi_{j,k}(\cdot)$ , for  $\{(j, k) : j = 0, 1, \dots, J - 1, k = 1, 2, \dots, 2^j\}$ . A basic example of the DWT is provided in Section 3.1 for readers who are less familiar with the technique.

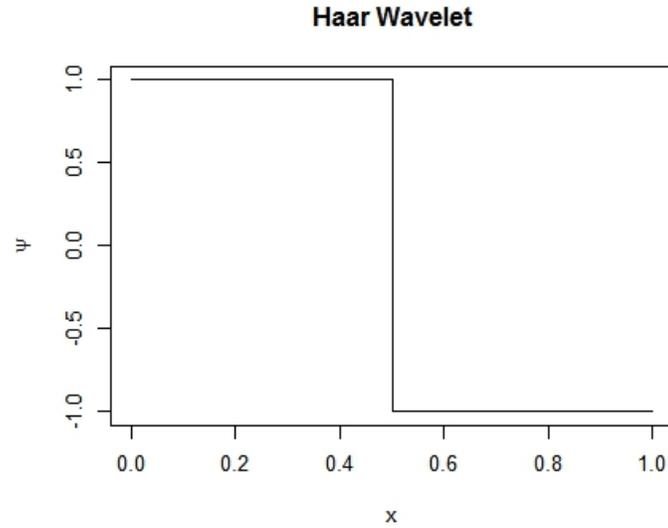
The DWT has an additional advantage for use in large datasets. The discrete wavelet transformation requires less stringent technical assumptions compared to Fourier transformation since the amplitude corresponding to each frequency does not have to be stable over time. Computationally, the DWT is faster than the fast Fourier transform (see Percival and Walden (2006) for details). As we motivate in Section 6, our model describes correlations among the RWI records by incorporating spatial distance and temporal information. In order to create the correlations between records, we need to begin by de-correlating the individual series. The DWT tends to produce independent (or slightly correlated, Vannucci and Corradi (1999)) coefficients from observed data series which are strongly correlated (see Chapter 10 in Percival and Walden (2006)).

The above framework is based on  $T = 2^J$ , that is, the length of the observed data is a power of 2. There are several existing techniques to overcome this limitation, see for example Gong *et al.* (2018); Walden and Cristan (1998); Ogden (1997); Nason (2008), however, we choose not to use such advanced techniques here for clarity and tractability of the results.

### 3.1. Basic example of the DWT

Here we provide a basic example to demonstrate the time-scale information in wavelet coefficients that we have discussed and utilize in our analysis. Consider the most basic wavelet basis, proposed by Haar. The Haar wavelet basis appears as a step function (see Figure 2).

At the coarsest level, the step function spans across the entire vector of data. At finer levels, the step function is shrunk and translated across the data vector. Let's consider the



**Figure 2: Haar wavelet (on a continuous scale)**

the Haar basis when  $n=8$ , as generated using the `GenW()` within the `wavethresh` **R** package by Nason (2016).

$$\mathbf{W} = \begin{bmatrix} 2^{-3/2} & 2^{-3/2} & 2^{-3/2} & 2^{-3/2} & 2^{-3/2} & 2^{-3/2} & 2^{-3/2} & 2^{-3/2} \\ 2^{-1/2} & -2^{-1/2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2^{-1/2} & -2^{-1/2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2^{-1/2} & -2^{-1/2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2^{-1/2} & -2^{-1/2} \\ 2^{-1} & 2^{-1} & -2^{-1} & -2^{-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2^{-1} & 2^{-1} & -2^{-1} & -2^{-1} \\ 2^{-3/2} & 2^{-3/2} & 2^{-3/2} & 2^{-3/2} & -2^{-3/2} & -2^{-3/2} & -2^{-3/2} & -2^{-3/2} \end{bmatrix} \quad (4)$$

Since the vector of wavelet coefficients is found using  $\mathbf{W}\mathbf{x}$ , we can see that the scaling coefficient,  $\gamma_0$  is a weighted sum of the vector entries. The coarsest wavelet coefficient,  $\gamma_{0,1}$  is a difference in weighted sums between the first and second halves of the vector entries. The coarsest wavelet coefficients capture long-term signals in the data. The finest level coefficients,  $\gamma_{2,1}$ ,  $\gamma_{2,2}$ ,  $\gamma_{2,3}$ ,  $\gamma_{2,4}$ , are weighted pairwise differences between entries which are next to each other. These capture signal changes that are close in time.

It is clear with the specific choice of Haar basis that  $\mathbf{W}$  is an orthonormal matrix. Most generally, orthonormality is a wavelet basis property. Although it is easier to visualize and discuss, the Haar basis has an obvious discrete nature which creates non-smooth estimates that are difficult to interpret in practice. Other wavelet bases typically do not have closed-form solutions, but are more often used in statistical analysis. These include Daubechies (1992) and Chui and Wang (1992). With the guidance of several existing works that provide recommendations for selecting wavelet families, including Mandal *et al.* (1996),

Strickland and Hahn (1996), and Mojsilovic *et al.* (2000), the analysis we present here utilizes Daubechies Least Asymmetric family with 8 vanishing moments. Several software packages enable automatic implementation wavelet bases for analysis including `wavethresh` by Nason (2008), `wavelets` by Aldrich (2020), `waveslim` by Whitcher (2020) in **R** and the Wavelet Toolbox in Matlab (The MathWorks, 2020).

#### 4. Strength of Relationship Between RWI Records

Once we have attempted to account for the strong dependence *within* the RWI record by implementing the DWT, we consider modeling the relationships among records. Without assuming a specific type of relationship (for example linear) between two quantitative variables, there are several ways to compute correlation. We considered three correlation measures to describe the RWI data.

##### 4.1. Correlation measures

Generally, correlation measures the strength of relationship between two quantitative variables. Suppose that  $\mathbf{x} = (x_1 \ x_2 \ \dots \ x_n)^T$  and  $\mathbf{y} = (y_1 \ y_2 \ \dots \ y_n)^T$  are two vectors containing real-valued observations. Pearson's correlation (eqn. 5) provides a measure of the linear association between the vectors.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (5)$$

Although commonly utilized, Pearson's correlation is limited to quantifying the linear relationship. Spearman's rank correlation (6) provides a bit more flexibility, by describing the strength of monotonic relationship between two quantitative variables. Spearman's correlation is created by utilizing the *rank* transformation on the observed vectors. In this transformation, the observations within each vector are ordered from smallest to largest, and thus, changed to values within  $\mathbb{Z}^+$  where 1 corresponds to the smallest observation within each vector. This transformation doesn't depend on linearity and produces measures less affected by outliers.

$$r_s = \frac{\sum_{i=1}^n (\text{rank}(x_i) - \overline{\text{rank}(x)})(\text{rank}(y_i) - \overline{\text{rank}(y)})}{\sqrt{\sum_{i=1}^n (\text{rank}(x_i) - \overline{\text{rank}(x)})^2} \sqrt{\sum_{i=1}^n (\text{rank}(y_i) - \overline{\text{rank}(y)})^2}} \quad (6)$$

Kendall's tau correlation (7) also provides a measure of the strength of monotonic relationship between two quantitative variables. It specifically examines the discordance between pairs of observations, as opposed to the orderings among the individual observations as in Spearman's correlation. Discordance is quantified by obtaining the sign, positive or negative, of the difference between observations through the *sgn* transformation. Kendall's tau correlation has a bit of an advantage to Spearman's rank correlation with respect to

interpretability since describing discordant pairs is more straightforward than the sums of ranks.

$$\tau = \frac{\sum_{i=1}^n \sum_{j=1}^n \text{sgn}(x_i - x_j) \text{sgn}(y_i - y_j)}{n(n-1)} \quad (7)$$

## 4.2. Creating pairwise correlations of wavelet coefficients

Our model describes correlations between the  $\binom{97}{2}$  pairs of PIPO records within the Four Corners states. The correlations are between the series' wavelet coefficients, not the raw observations. This choice allows us to model the relationship strength of relationship with respect to spatial location, while also accounting for temporal separation. Below is an outline of the algorithm utilized to compute the correlations.

1. For the records in a given pair of locations, keep the years when both records have observations.
2. Retain the most recent  $T$  observations for each of the series obtained in (1). We require  $T = 2^J$  for  $J \in \mathbb{Z}^+$ .
3. Perform the DWT on each series obtained in (2) and obtain the resulting wavelet coefficient vectors.
4. Within the wavelet coefficients, calculate the time-scale correlation for any levels with 8 or more coefficients. For example, a correlation will be computed for the  $2^{J-1}$  finest scale coefficients. This would correspond to the same year time-scale relationship.

With the Four Corners data, we had enough observations to calculate up to 16-year correlation information. Each of Pearson, Spearman, and Kendall correlations were calculated and considered as candidate responses for the model. The Daubechies least asymmetric wavelet with 8 vanishing moments (see Daubechies (1992)) was selected for the basis in all of our DWT computations.

## 5. Re-formatting the Data for Analysis

For each of the five available time-scales (same-, 2-, 4-, 8-, and 16-year), the algorithm outlined in subsection 4.2 created  $\binom{97}{2}$  correlations. These correlations may be combined to compose vectors of the form

$$\mathbf{c}_j = \left( c_{1,2,j} \quad c_{1,3,j} \quad \dots \quad c_{96,97,j} \right)^T \quad (8)$$

where  $c_{i,k,j}$  denotes the correlation of records in the  $i^{\text{th}}$  and  $k^{\text{th}}$  locations. Time-scales, same-year (s), 2-year (2), 4-year (4), 8-year (8), 16-year (16), are denoted by the index  $j$  in eqn. 8. Many of the record pairings did not result in series long enough to compute the 8- or 16-

year correlations with our constraints. Those entries are missing. To construct the response in our model, we stack all of the time-scale correlation vectors into one vector

$$\mathbf{c} = (\mathbf{c}_s \quad \mathbf{c}_2 \quad \mathbf{c}_4 \quad \mathbf{c}_8 \quad \mathbf{c}_{16})^T. \quad (9)$$

We also construct the set of predictor variables in a similar manner. The structure of the wavelet coefficients lends to incorporating the time-scale information to the model. A set of indicator variables for each time scale were created with respect to the response vector

$$Y_{i,k,J_0} = \begin{cases} 1, & j = J_0 \text{ in } C_{i,k,j} \\ 0, & j \neq J_0 \text{ in } C_{i,k,j} \end{cases}. \quad (10)$$

The indicator variables for each time scale were stacked in a similar fashion as the correlations in eqn. 9 to create  $\mathbf{y}_s, \mathbf{y}_2, \mathbf{y}_4, \mathbf{y}_8, \mathbf{y}_{16}$  vectors, all of the same length as  $\mathbf{c}$ .

In addition to the time-scale information, our model also incorporates spatial information in the form of the distance between locations in each pair. For each pair of records, Lambert's distance, Lambert (1942), in kilometers was calculated. The distance vector for all pairs may be denoted as

$$\tilde{\mathbf{d}} = (d_{1,2} \quad d_{1,3} \quad \dots \quad d_{96,97})^T. \quad (11)$$

Five  $\tilde{\mathbf{d}}$  vectors are stacked as in eqn. 9 to create  $\mathbf{d}$  of the same length as  $\mathbf{c}$ . Finally, along similar lines, the absolute difference in location elevations for each pair is computed and used to create the  $\mathbf{h}$  predictor.

## 6. Model Estimation

Taking the Kendall correlation as an example, the scatterplots in Figure 5 within the appendix indicate a negative relationship between the locations' distance and the observed correlation at each time-scale. Similar results also hold true for the other types of correlations. Therefore, we propose a linear model with an interaction between time-scale and distance and a linear term in elevation to describe correlation in eqn. 12.

$$\begin{aligned} \mathbf{c} = & \beta_0 \mathbf{y}_s + \beta_1 \mathbf{y}_2 + \beta_3 \mathbf{y}_4 + \beta_4 \mathbf{y}_8 + \beta_5 \mathbf{y}_{16} + \\ & \beta_6 \mathbf{y} + \beta_7 \mathbf{y}_2 \circ \mathbf{d} + \beta_8 \mathbf{y}_4 \circ \mathbf{d} + \beta_9 \mathbf{y}_8 \circ \mathbf{d} + \beta_{10} \mathbf{y}_{16} \circ \mathbf{d} + \\ & \beta_{11} \mathbf{h} + \boldsymbol{\epsilon}. \end{aligned} \quad (12)$$

With such a large dataset, it was difficult to pinpoint the best estimation method and type of correlation. Even with the pairwise correlations plotted on a gradient scale, (as in Figure 5), it is difficult to determine to see what is happening. Thus, we also considered different estimation techniques for our model – least squares, median regression, and the Huber M estimator.

### 6.1. Estimation techniques

The three estimation techniques we consider for this analysis are all related through an objective function,  $\rho$  which is maximized in order to obtain the estimates. The model in eqn. 12 is a linear model. If we combined all of the predictor vectors into a matrix,  $\mathbf{X}$ , and all of the parameters into a vector,  $\boldsymbol{\beta}$ , then the model could be written simply as  $\mathbf{c} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ .

Then, the estimates obtained through least squares, median regression, and Huber-M estimation are all solved by minimizing  $\rho(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ . Specifically, the least squares estimate is obtained through minimizing the squared  $\mathcal{L}_2$  norm as in eqn. 13 while the median regression estimate minimizes the  $\mathcal{L}_1$  norm (eqn. 14).

$$\arg \min_{\boldsymbol{\beta} \in \mathbb{R}^{k+1}} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 \quad (13)$$

$$\arg \min_{\boldsymbol{\beta} \in \mathbb{R}^{k+1}} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_1 \quad (14)$$

Huber-M estimate (eqn. 15) utilizes a combination of the  $\mathcal{L}_1$  and  $\mathcal{L}_2$  norms in the minimization problem, dependant on the size of the element within  $\mathbf{y} - \mathbf{X}\boldsymbol{\beta}$ . The  $\mathbf{I}_k$  within eqn. 15) is an indicator vector, the elements of which take a value of 1 when the magnitude of  $\mathbf{y} - \mathbf{X}\boldsymbol{\beta}$  is no larger than a specified  $k$ .

$$\arg \min_{\boldsymbol{\beta} \in \mathbb{R}^{k+1}} \left[ \left( 0.5 \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 \right) \circ \mathbf{I}_k + \left( k \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_1 + k^2/2(\mathbf{1} - \mathbf{I}_k) \right) \right] \quad (15)$$

### 6.2. Resulting estimates

Tables 1, 2, and 3 contain the estimates for the model proposed in eqn. 12 for each of the correlations and estimation methods. Within each correlation type, estimation method does not seem to make much difference in the obtained values. We do see differences amongst the point estimates across correlation type, but this was not surprising from our data exploration. Generally, we see that each correlation and method contain the same general patterns across the coefficients.

## 7. Statistical Inference

Continuing forward, the results examined utilize median regression to predict Kendall correlation. In this section, we specifically compare the slopes in record pair distance across time-scales. Figure 3 contains the estimated relationships while the absolute difference in elevation is held constant at 0 km.

Although the linear model seems to be a straight-forward way to describe the RWI correlations, the data structure is actually quite complicated. Each correlation,  $c_{i,k,j}$ , is based upon the  $i^{th}$  and  $k^{th}$  records. Thus, each record appears within  $5 \times 96$  observations within the response vector. There is not an intuitive argument to claim that the correlations

	Pearson			Spearman			Kendall		
	L.S.	Med.	Hub.	L.S.	Med.	Hub.	L.S.	Med.	Hub.
$\hat{\beta}_0$	6.597	6.806	6.663	6.431	6.630	6.480	4.614	<b>4.695</b>	4.623
$\hat{\beta}_0 + \hat{\beta}_1$	6.002	6.347	6.144	5.843	6.101	5.973	4.249	<b>4.297</b>	4.295
$\hat{\beta}_0 + \hat{\beta}_2$	5.953	6.493	6.295	5.805	6.360	6.113	4.257	<b>4.538</b>	4.424
$\hat{\beta}_0 + \hat{\beta}_3$	6.372	7.217	6.940	5.979	6.776	6.557	4.555	<b>4.926</b>	4.855
$\hat{\beta}_0 + \hat{\beta}_4$	6.511	7.120	6.930	5.744	6.352	6.074	4.497	<b>4.716</b>	4.663

**Table 1: Estimated intercepts. Each estimate is multiplied by  $1 \times 10^{-1}$  and bold estimates are associated with the model in which we perform statistical inference. L.S=least squares regression, Med.=Median regression, Hub.=Huber M estimator regression.**

	Pearson			Spearman			Kendall		
	L.S.	Med.	Hub.	L.S.	Med.	Hub.	L.S.	Med.	Hub.
$\hat{\beta}_6$	6.23	6.41	6.28	6.15	6.33	6.18	4.48	<b>4.56</b>	4.48
$\hat{\beta}_6 + \hat{\beta}_7$	4.70	4.97	4.79	4.63	4.74	4.73	3.46	<b>3.41</b>	3.48
$\hat{\beta}_6 + \hat{\beta}_8$	4.36	4.88	4.68	4.48	5.10	4.78	3.32	<b>3.71</b>	3.50
$\hat{\beta}_6 + \hat{\beta}_9$	5.77	6.53	6.14	5.30	5.88	5.69	4.10	<b>4.35</b>	4.31
$\hat{\beta}_6 + \hat{\beta}_{10}$	3.00	3.35	3.28	2.43	2.86	2.58	2.04	<b>2.22</b>	2.16

**Table 2: Estimated slopes in distance. Each estimate is multiplied by  $-1 \times 10^{-4}$  and bold estimates are associated with the model in which we perform statistical inference. L.S=least squares regression, Med.=Median regression, Hub.=Huber M estimator regression.**

are independent. For example, a certain record could happen to be less correlated with all of the other records. Perhaps at that location, an external event caused all of the tree species to have stunted growth.

There are a few existing methods to deal with dependence due to repeated measures. A common choice is to add an individual effect to the model. That is, add a set of indicator variables associated with record ID as fixed or random effects to the model fit. In our analysis, this translates to adding 96 indicator variables to the model. Although possible with the number of cases in our data, the addition of ID is difficult to interpret and our structure is more complicated, since each response is associated with two record IDs. Moreover, treating ID as a random quantity would create a mixed effects model. With few distributional assumptions placed on the model error structure, inference would still be an open area of research.

Our analysis actually continues with the model from eqn. 12 and does not over-complicate it by adding an explicit record ID variable. Although the fits obtained should be unbiased for a correctly specified model, utilizing the asymptotic normal distribution to estimate parameter error is incorrect with dependent data. To estimate coefficient error, we will

	Pearson			Spearman			Kendall		
	L.S.	Med.	Hub.	L.S.	Med.	Hub.	L.S.	Med.	Hub.
$\hat{\beta}_{11}$	8.461	8.781	8.825	8.566	8.867	8.871	6.598	<b>6.358</b>	6.610

**Table 3: Estimated slope in Elevation Difference. Each estimate is multiplied by  $-1 \times 10^{-2}$  and the bold estimate is associated with the model in which we perform statistical inference. L.S.=least squares regression, Med.=Median regression, Hub.=Huber M estimator regression.**

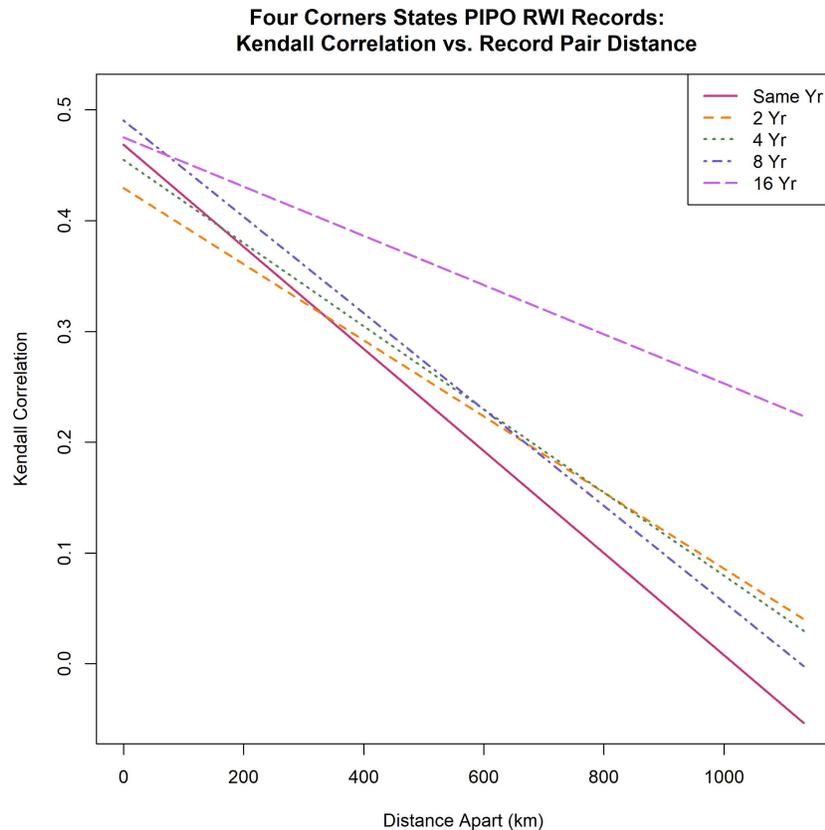
instead use a version of the wild bootstrap, originally proposed by Wu (1986), which utilizes random variables indexed by the record ID. We describe how a wild bootstrap response,  $\mathbf{c}_b$ , is obtained in our framework in the subsequent algorithm.

1. Generate  $n$ =number of records, independent random variables, each with mean 0 and variance 1. In our analysis, this translates to generating 97 random variables,  $U_1, U_2, \dots, U_{97}$ . Suppose that the index of these random variables corresponds to a location ID.
2. Create two vectors, each of the same length as the response,  $\mathbf{c}$ , containing the random variables generated in (1). The elements of the first vector correspond to the first location ( $i$  index in the  $\mathbf{c}$ ). The elements of the second vector correspond to the second location ( $k$  index in the  $\mathbf{c}$ ). Call these vectors  $\mathbf{u1}$  and  $\mathbf{u2}$ .
3. Create the wild bootstrap response,  $\mathbf{c}_b$ , as  $\mathbf{c}_b = \hat{\mathbf{c}} + \text{diag}(\mathbf{u1})\text{diag}(\mathbf{u2})\mathbf{r}$ .

In order to decide whether there was evidence that the linear relationship with record pair distance differed by time-scale, we implemented the described wild bootstrap to obtain bootstrap distributions of all model coefficients estimated in eqn. 12. These bootstrap sampling distributions are displayed in figure 6 within the appendix. Each distribution is centered at the original coefficient estimate (as expected), and all seem to be fairly light-tailed. Specifically, we explored whether there was evidence of a difference in the slopes in distance with respect to the same-year time-scale. With  $B = 10,000$  bootstrap samples and Bonferroni adjusted significance level of  $\alpha = 0.05$ , confidence intervals were created for each of the parameters:  $\beta_7, \beta_8, \beta_9$ , and  $\beta_{10}$ . These intervals are displayed in figure 4. The black bands in each figure are associated with the bootstrap confidence interval. We can see that all of the bootstrap confidence intervals are shorter than the typical normal intervals, which assume independence across observations. These intervals also indicate evidence that the slopes in distance, across different time-scales, differ from the same-year relationship.

## 8. Results

The results in Figure 4 imply that the linear relationship between Kendall correlation and record pair distance differs over time. Examination of the scatterplots within Figure 5, indicate a lack of data in the 8- and 16-year time scales, compared to the other scales. Our model inference finds significant relationships at these scales, even though there was little



**Figure 3: Fitted lines with Kendall correlation as the response and median regression**

evidence of a relationship from the exploratory scatterplots. This finding could be due to a lack of power at the higher time-scales with less data. Interpreting the results at these time-scales should be done with caution.

This analysis may have significant implications for researchers examining tree-ring chronologies. First and foremost, this study implies that these *Pinus ponderosa* records contain intricate information, not only with respect to climate, as typically studied, but also with respect to each other. The type information that these records carry also varies over time. The specific analysis presented here was for a subset of the *Pinus ponderosa* in the U.S., and the maximum record length in the set was 551 years, while the shortest record only contained 63 years of RWIs. If older tree-ring records were available, we would be able to expand the analysis further, and draw more concrete conclusions about the longer time-scales. Specifically, if the researcher is interested in studying a  $2^k$  time-scale in the RWI records, then a minimum of  $2^{3+k+1}$  record length in the majority of records would be required. But, as seen in our analysis, this minimum requirement creates an unwieldy correlation estimate at the longest allowable time-scale, so the researcher would likely prefer to have records of length  $2^{5+k+1}$  to seriously study the  $k^{\text{th}}$  time-scale.

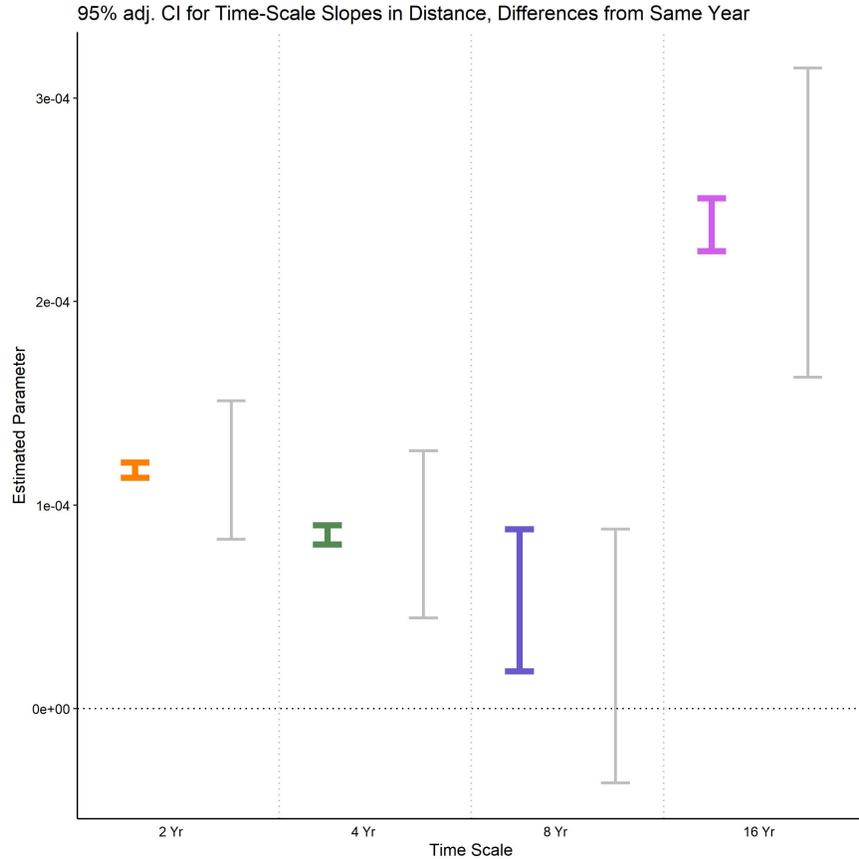


Figure 4: Bonferroni adjusted 95% confidence intervals for the time-scale slope parameters in distance. These are *differences* from the same-year slope. The traditional normal interval appears as light gray for comparison with our bootstrap intervals, in color.

## 9. The Theoretical Framework

We illustrate the main ideas of the underlying theoretical framework in this section, and omit much of the technical details. Fix a pair of locations. Since the wavelet decomposition de-correlates the data, we consider independent observations  $\{(Y_{i,1}, Y_{i,2}, \mathbf{x}_i) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^p; i = 1, \dots, n\}$  corresponding to this pair of locations. We assume that  $\mathbb{E}Y_{i,1} = \mu_1$ ,  $\mathbb{E}Y_{i,2} = \mu_2$ ,  $\mathbb{V}Y_{i,1} = \sigma_1^2$ ,  $\mathbb{V}Y_{i,2} = \sigma_2^2$  and  $Cor(Y_{i,1}, Y_{i,2}) = \rho(\mathbf{x}_i)$ . In other words, our data consists of bivariate response vectors  $(Y_{i,1}, Y_{i,2})$  related to the pair of locations, and associated covariate  $\mathbf{x}_i \in \mathbb{R}^p$  that is taken to be non-random in this paper, and we model the correlation between the two responses as a function of the covariate. The bivariate response is allowed to be heteroscedastic, and the two means and two variances may also be functions of the covariate, but that structure is not relevant for present purposes. We assume that  $\mathbb{E}Y_{i,j}^{8k} < \infty$  for  $j = 1, 2$  and a sufficiently high positive integer  $k$ . We also write  $\rho \equiv \rho(\mathbf{x}_i)$  below, since there is no cause of confusion here.

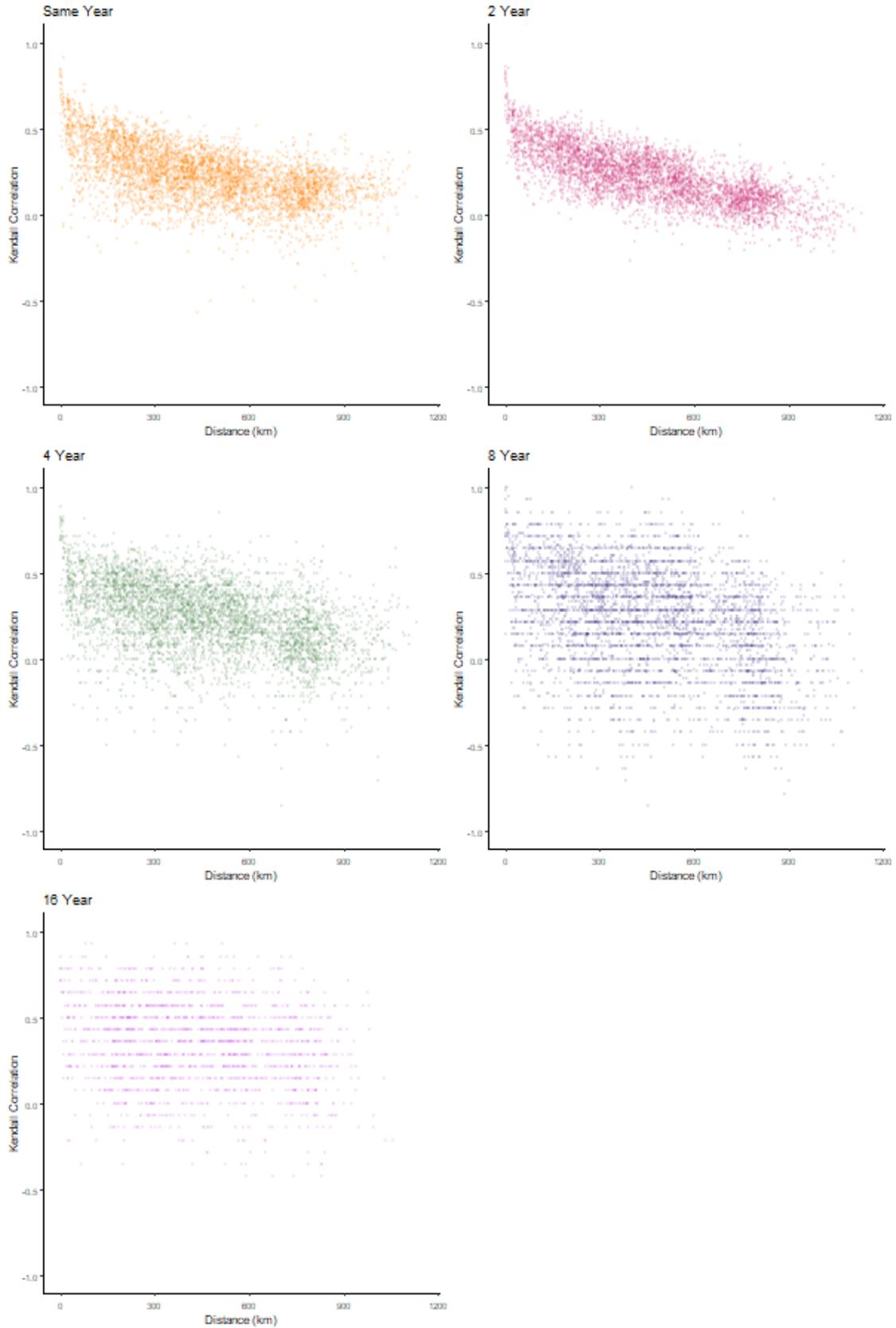


Figure 5: Pairwise Kendall correlations between wavelet coefficients at each time-scale

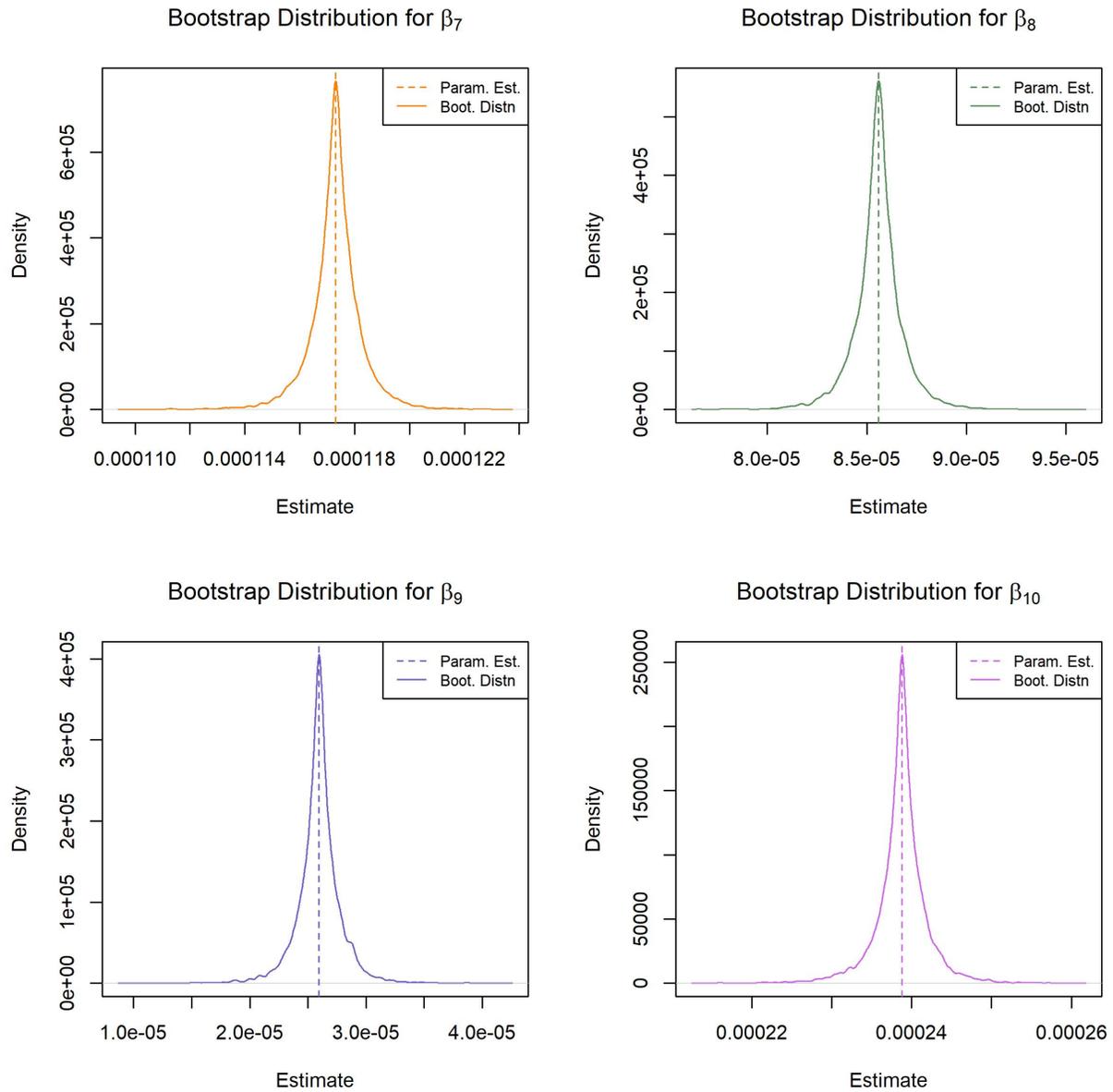


Figure 6: Bootstrap sampling distributions of time-scale slopes in distance. These are estimated *differences* from the same-year time-scale slope. The original estimate is the dashed vertical line.

Let  $\epsilon \ll 0.25 \min(\sigma_1^2, \sigma_2^2)$  be a small constant, and let  $\mathcal{A}$  be the event

$$\mathcal{A} = \left\{ \sum_{i=1}^n (Y_{i,1} - n^{-1} \sum_{i=1}^n Y_{i,1})^2 > \epsilon, \text{ and } \sum_{i=1}^n (Y_{i,2} - n^{-1} \sum_{i=1}^n Y_{i,2})^2 > \epsilon \right\}.$$

On  $\mathcal{A}$  we define our estimator of  $\rho$  to be the usual sample correlation

$$\hat{\rho} = \left[ \left\{ \sum_{i=1}^n (Y_{i,1} - n^{-1} \sum_{i=1}^n Y_{i,1})^2 \right\} \left\{ \sum_{i=1}^n (Y_{i,2} - n^{-1} \sum_{i=1}^n Y_{i,2})^2 \right\} \right]^{1/2} \sum_{i=1}^n (Y_{i,1} - n^{-1} \sum_{i=1}^n Y_{i,1}) (Y_{i,2} - n^{-1} \sum_{i=1}^n Y_{i,2}).$$

On its complement of  $\mathcal{A}$ , we take  $\hat{\rho} = 0$ . We have to separate the cases where the sample variances are sufficiently high and where they are not using  $\mathcal{A}$ , since we do not make any distribution assumptions like Gaussianity in this paper. It can be shown that  $1 - \mathbb{P}[\mathcal{A}] = O(n^{-4k})$ , consequently we only discuss  $\hat{\rho}$  under  $\mathcal{A}$ , the other case is negligible.

**Theorem 1:** Under the conditions stated above, we have the following results:

- (a) For every fixed pair of locations, the correlation estimator  $\hat{\rho}$  is consistent and the distribution of  $n^{1/2}(\hat{\rho} - \rho)$  converges weakly to a Gaussian distribution.
- (b) If  $\rho(\mathbf{x}) = \beta^T \mathbf{x}$  for some  $\beta \in \mathbb{R}^p$ , then the ordinary least squares estimator for  $\beta$  computed from all pairs of locations is consistent and asymptotically Gaussian.
- (c) If  $\rho(\mathbf{x}) = \beta^T \mathbf{x}$  for some  $\beta \in \mathbb{R}^p$ , and  $\hat{\beta}$  is the ordinary least squares estimator for  $\beta$ , then the distribution of  $n^{1/2}(\hat{\beta} - \beta)$  is consistently approximated by the proposed version of the wild bootstrap scheme, conditional on the data, for almost all sample paths.

**Proof:** [Proof of Theorem 1]

Here, we only provide a brief outline of the main arguments for part (a) of Theorem 1, in order to not overload this paper with algebraic details.

We define the following terms

$$\begin{aligned} T_1 &= n^{-1} \sigma_1^{-2} \sum_{i=1}^n (Y_{i,1} - n^{-1} \sum_{i=1}^n Y_{i,1})^2, \\ T_2 &= n^{-1} \sigma_2^{-2} \sum_{i=1}^n (Y_{i,2} - n^{-1} \sum_{i=1}^n Y_{i,2})^2, \\ T_3 &= n^{-1} \sigma_1^{-1} \sigma_2^{-1} \sum_{i=1}^n (Y_{i,1} - n^{-1} \sum_{i=1}^n Y_{i,1}) (Y_{i,2} - n^{-1} \sum_{i=1}^n Y_{i,2}). \end{aligned}$$

In terms of these notations, we have under  $\mathcal{A}$  that  $\hat{\rho} = (T_1 T_2)^{-1/2} T_3$ .

In the following, the notation  $R_a$  denotes a remainder for all choices of  $a$ , with the property that  $\mathbb{E}R^{2k} < \infty$ . We do not explicitly report the algebra relating to such remainder terms, those are routine. It can be worked out that on  $\mathcal{A}$ , we have the following

$$\begin{aligned} T_3 &= \rho + n^{-1/2}R_1 + n^{-1}R_2, \\ T_1 &= 1 + n^{-1/2}R_3 + n^{-1}R_4, \\ T_2 &= 1 + n^{-1/2}R_5 + n^{-1}R_6, \end{aligned}$$

consequently we can express  $(T_1T_2)^{-1/2}$  as

$$(T_1T_2)^{-1/2} = 1 + n^{-1/2}R_7 + n^{-1}R_8.$$

It is in the above expression, the condition dictating  $\mathcal{A}$  is required, for establishing that  $R_8$  is indeed a random variable satisfying  $\mathbb{E}R^{2k} < \infty$ .

Using the above terms, we now have that

$$\begin{aligned} \hat{\rho} &= (T_1T_2)^{-1/2}T_3 \\ &= \rho + n^{-1/2}R_9, \end{aligned}$$

where  $R_9$  is a smooth function of all of the following terms:

$$\begin{aligned} T_4 &= n^{1/2} \left( n^{-1} \sum_{i=1}^n Y_{i,1} - \mu_1 \right), \\ T_5 &= n^{1/2} \left( n^{-1} \sum_{i=1}^n Y_{i,2} - \mu_2 \right), \\ T_6 &= n^{1/2} \left( n^{-1} \sum_{i=1}^n (Y_{i,1} - \mu_1)^2 - \sigma_1^2 \right), \\ T_7 &= n^{1/2} \left( n^{-1} \sum_{i=1}^n (Y_{i,2} - \mu_2)^2 - \sigma_2^2 \right), \\ T_8 &= n^{1/2} \left( n^{-1} \sum_{i=1}^n (Y_{i,2} - \mu_2)(Y_{i,2} - \mu_2) - \rho\sigma_1\sigma_2 \right). \end{aligned}$$

The rest of the analysis, along with the justification for the novel resampling scheme used in this paper, now follows using fairly routine algebra, and the use of Lyapunov's Central Limit Theorem.  $\square$

## 10. Future Considerations

In the above analysis, we build a model with a response variable containing pairwise correlations between wavelet coefficients at different time-scales. Then, we perform a modified version of the wild bootstrap for statistical inference.

We utilized a small subset of the available data within the international tree-ring databank. The analysis could be further modified or generalized in many ways. Multiple species

within the same area could be included. It could be of interest to correlate records over a larger land area, or even across continents. Finally, incorporating information, such as temperature and precipitation, may help researchers to better understand past climate. However, all these come with additional challenges, for example, the distribution of tree-ring records over the world is patchy and inadequate in many regions, dependence of the data over different tree species and subspecies, between various climatic regions and other features need to be carefully modeled. Additionally, many data records are incomplete, or are unevenly spaced over time. The theoretical details behind the new resampling scheme deserves further study.

## Acknowledgments

Support for this collaboration was provided by the United States National Science Foundation (Awards 1622483 (St. G and SC), 1602512 (St. G), 1737918 (SC), 1939916 (SC) and 1939956 (SC)). We thank the Editors and the reviewers for their constructive comments that helped improve this paper.

## References

- Affolter, S., Häuselmann, A., Fleitmann, D., Edwards, R. L., Cheng, H., and Leuenberger, M. (2019). Central europe temperature constrained by speleothem fluid inclusion water isotopes over the past 14,000 years. *Science Advances*, **5** (6), eaav3809.
- Aldrich, E. (2020). *wavelets: Functions for Computing Wavelet Filters, Wavelet Transforms and Multiresolution Analyses*. R package version 0.3-0.2.
- Arienzo, M. M., Swart, P. K., Broad, K., Clement, A. C., Pourmand, A., and Kakuk, B. (2017). Multi-proxy evidence of millennial climate variability from multiple Bahamian speleothems. *Quaternary Science Reviews*, **161**, 18–29.
- Brook, E. J. and Buizert, C. (2018). Antarctic and global climate history viewed from ice cores. *Nature*, **558** (7709), 200–208.
- Bunn, A. G., Sharac, T. J., and Graumlich, L. J. (2004). Using a simulation model to compare methods of tree-ring detrending and to investigate the detectability of low-frequency signals. *Tree-Ring Research*, **60** (2), 77–90.
- Chui, C. K. and Wang, J. (1992). On compactly supported spline wavelets and a duality principle. *Transactions of the American Mathematical Society*, **330** (2), 903–915.
- Cook, E. R. (1987). The decomposition of tree-ring series for environmental studies. *Tree-Ring Bulletin*, **47**, 37–59.
- Cook, E. R. and Peters, K. (1997). Calculating unbiased tree-ring indices for the study of climatic and environmental change. *The Holocene*, **7** (3), 361–370.
- Cook, E. R., Solomina, O., Matskovsky, V., Cook, B. I., Agafonov, L., Berdnikova, A., Dolgova, E., Karpukhin, A., Knysh, N., Kulakova, M., *et al.* (2020). The european russia drought atlas (1400–2016 ce). *Climate Dynamics*, **54** (3), 2317–2335.
- Coulthard, B. L., St George, S., and Meko, D. M. (2020). The limits of freely-available tree-ring chronologies. *QSRv*, **234**, 106264.
- Daubechies, I. (1992). Ten Lectures on Wavelets. In *CBMS-NSF Regional Conference Series In Applied Mathematics*.

- Emile-Geay, J., McKay, N. P., Kaufman, D. S., Von Gunten, L., Wang, J., Anchukaitis, K. J., Abram, N. J., Addison, J. A., Curran, M. A., Evans, M. N., *et al.* (2017). A global multiproxy database for temperature reconstructions of the Common Era. *Scientific Data*, **4**, 170088.
- Esper, J., St George, S., Anchukaitis, K., D'Arrigo, R., Ljungqvist, F. C., Luterbacher, J., Schneider, L., Stoffel, M., Wilson, R., and Büntgen, U. (2018). Large-scale, millennial-length temperature reconstructions from tree-rings. *Dendrochronologia*, **50**, 81–90.
- Fan, Z.-X. and Bräuning, A. (2017). Tree-ring evidence for the historical cyclic defoliator outbreaks on *Larix potaninii* in the central Hengduan mountains, SW China. *Ecological Indicators*, **74**, 160–171.
- Gong, K., Braverman, A., and Chatterjee, S. (2018). On a technique for evaluating the quality of earth system models. In *Proceedings of the 8th International Workshop on Climate Informatics: CI 2018*, pages 93 – 96.
- Granato-Souza, D., Stahle, D. W., Barbosa, A. C., Feng, S., Torbenson, M. C., de Assis Pereira, G., Schöngart, J., Barbosa, J. P., and Griffin, D. (2019). Tree rings and rainfall in the equatorial Amazon. *Climate Dynamics*, **52 (3-4)**, 1857–1869.
- Harris, I., Osborn, T. J., Jones, P., and Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, **7 (1)**, 1–18.
- Holmes, N., Langdon, P. G., Caseldine, C. J., Wastegård, S., Leng, M. J., Croudace, I. W., and Davies, S. M. (2016). Climatic variability during the last millennium in western Iceland from lake sediment records. *The Holocene*, **26 (5)**, 756–771.
- Hubbard, B. B. (1998). *The World According to Wavelets: The Story of a Mathematical Technique in the Making*. A K Peters, Ltd., Wellesley, MA, 2nd edition.
- Kannenbergh, S. A., Novick, K. A., Alexander, M. R., Maxwell, J. T., Moore, D. J., Phillips, R. P., and Anderegg, W. R. (2019). Linking drought legacy effects across scales: From leaves to tree rings to ecosystems. *Global Change Biology*, **25 (9)**, 2978–2992.
- Kasatkina, E. A., Shumilov, O. I., and Timonen, M. (2019). Solar activity imprints in tree ring-data from northwestern Russia. *Journal of Atmospheric and Solar-Terrestrial Physics*, **193**, 105075.
- Lambert, W. D. (1942). The distance between two widely separated points on the surface of the earth. *Journal of Washington Academy of Sciences*, **32 (5)**, 125–130.
- Lau, K.-M. and Weng, H. (1995). Climate signal detection using wavelet transform: How to make a time series sing. *Bulletin of the American Meteorological Society*, **76 (12)**, 2391–2402.
- Mandal, M. K., Panchanathan, S., and Aboulnasr, T. (1996). Choice of wavelets for image compression. In Chouinard, J.-Y., Fortier, P., and Gulliver, T. A., editors, *Information Theory and Applications II: 4th Canadian Workshop Lac Delage, Québec, Canada, May 28–30, 1995 Selected Papers*, pages 239–249, Berlin, Heidelberg. Springer Berlin Heidelberg.
- McCullough, I. M., Davis, F. W., and Williams, A. P. (2017). A range of possibilities: Assessing geographic variation in climate sensitivity of ponderosa pine using tree rings. *Forest Ecology and Management*, **402**, 223–233.

- McKenzie, S. M., Slater, G., Kim, S.-T., Pisaric, M. F., and Arain, M. A. (2018). Influence of seasonal temperature on tree-ring  $\delta^{13}\text{C}$  in different-aged temperate pine forests. *Forest Ecology and Management*, **419**, 197–205.
- Mojsilovic, A., Popovic, M. V., and Rackov, D. M. (2000). On the selection of an optimal wavelet basis for texture characterization. *IEEE Transactions on Image Processing*, **9** (12), 2043–2050.
- Nason, G. (2008). *Wavelet Methods in Statistics with R*. Springer, 1st edition.
- Nason, G. (2016). *wavethresh: Wavelets Statistics and Transforms*. R package version 4.6.8.
- Neukom, R., Steiger, N., Gómez-Navarro, J. J., Wang, J., and Werner, J. P. (2019). No evidence for globally coherent warm and cold periods over the preindustrial common era. *Nature*, **571** (7766), 550–554.
- Ogden, R. T. (1997). *Essential Wavelets for Statistical Applications and Data Analysis*. Birkhäuser, Boston, 1st edition.
- Opala-Owczarek, M. and Niedźwiedź, T. (2019). Last 1100 yr of precipitation variability in western central Asia as revealed by tree-ring data from the pamir-alay. *Quaternary Research*, **91** (1), 81–95.
- Percival, D. B. and Walden, A. T. (2006). *Wavelet Methods for Time Series Analysis*. Cambridge University Press, 1st edition.
- Sandeep, K., Shankar, R., Warriar, A. K., Yadava, M., Ramesh, R., Jani, R., Weijian, Z., and Xuefeng, L. (2017). A multi-proxy lake sediment record of Indian summer monsoon variability during the holocene in southern india. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **476**, 1–14.
- St George, S. and Ault, T. R. (2014). The imprint of climate within Northern Hemisphere trees. *Quaternary Science Reviews*, **89**, 1–4.
- St George, S. and Esper, J. (2019). Concord and discord among Northern Hemisphere paleotemperature reconstructions from tree rings. *Quaternary Science Reviews*, **203**, 278–281.
- Steiger, N. J., Steig, E. J., Dee, S. G., Roe, G. H., and Hakim, G. J. (2017). Climate reconstruction using data assimilation of water isotope ratios from ice cores. *Journal of Geophysical Research: Atmospheres*, **122** (3), 1545–1568.
- Strickland, R. N. and Hahn, H. I. (1996). Wavelet transforms for detecting microcalcifications in mammograms. *IEEE Transactions on Medical Imaging*, **15** (2), 218–229.
- The MathWorks, I. (2020). *Wavelet Toolbox*. Natick, Massachusetts, United States.
- Torrence, C. and Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, **79** (1), 61–78.
- Trouet, V., Babst, F., and Meko, M. (2018). Recent enhanced high-summer north Atlantic jet variability emerges from three-century context. *Nature Communications*, **9** (1), 1–9.
- Vannucci, M. and Corradi, F. (1999). Covariance structure of wavelet coefficients: theory and models in a Bayesian perspective. *Journal of the Royal Statistical Society B*, **61** (4), 971–986.
- Vidakovic, B. (1999). *Statistical Modeling by Wavelets*. Wiley, New York, NY, 1st edition.

- Walden, A. T. and Cristan, A. C. (1998). Matching pursuit by undecimated discrete wavelet transform for non stationary time series of arbitrary length. *Statistics and Computing*, **8**, 205–219.
- Wasserman, L. (2006). *All of Nonparametric Statistics*. Springer.
- Whitcher, B. (2020). *waveslim: Basic Wavelet Routines for One-, Two-, and Three-Dimensional Signal Processing*. R package version 1.8.2.
- Wu, C. F. J. (1986). Jackknife, Bootstrap and Other Resampling Methods in Regression Analysis (with discussion). *Annals of Statistics*, **14**, 1295.
- Zhao, S., Pederson, N., D’Orangeville, L., HilleRisLambers, J., Boose, E., Penone, C., Bauer, B., Jiang, Y., and Manzanedo, R. D. (2019). The international tree-ring data bank (itrd) revisited: Data availability and global ecological representativity. *Journal of Biogeography*, **46** (2), 355–368.