Journal of Forestry Research (2025) 36:3
https://doi.org/10.1007/s11676-024-01799-3

ORIGINAL PAPER

Adjusting the significance of daily climate responses in tree-ring proxies

Max C. A. Torbenson[1,2] · James H. Stagge[3] · Feng Chen[4,5,6] · Emanuele Ziaco[1] · Jan Esper[1,2]

Received: 2 July 2024 / Accepted: 1 August 2024

Abstract Correlating tree-ring parameters with daily resolved climate data is becoming increasingly common for understanding the complex relationships between tree growth and the surrounding environment. However, with an increased number of calculated correlations, there is an inherent risk of spurious significance. In this study, we present an analysis using synthetic weather and tree-ring data mimicking the statistical properties of ten real-world sites across Europe to quantify the extent to which numerous comparisons may inflate maximum correlations. Comparisons of different tree-ring proxies, considering varying overlapping period lengths and seasons, revealed 95th percentile correlation differences reaching 0.25 by chance. Using synthetic tree-ring chronologies with an assigned non-signal ($r = 0.00$), spurious correlations can reach statistical significance in over 60% of tests. Correlation inflation is greater when: (1) the climate-proxy relationship is weaker; (2) comparison periods are shorter; and (3) the length of seasonal windows is longer. Autocorrelation in the proxy records does not appear to have a major effect. These findings indicate that caution should be exercised when computing high numbers of correlations with limited observations. We provide tables listing correlation inflations for precipitation- and temperature-sensitive tree-ring chronologies that can inform interpretations of significance.

Keywords Dendroclimatology · Signal–noise · Daily resolution · Significance · Uncertainty

Introduction

Tree rings represent a formidable proxy of pre-instrumental climate variability and centennial-to-millennial length records of tree growth form the basis for our understanding of Common Era climate (IPCC 2021). The foundation of dendrochronology is an external forcing on the limits of tree productivity, often observed to be climatic in nature (Fritts 1976; Schweingruber 1988). The relationship between such limiting factors and tree-ring chronologies, the average and

Project funding: This study was supported partly by the ERC Advanced Grant Monostar (AdG 882727) and the co funded EU project AdAgriF (CZ.02.01.01/00/22_008/0004635).

The online version is available at https://link.springer.com/.

Corresponding editor: Tao Xu.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11676-024-01799-3.

✉ Max C. A. Torbenson
mtorbens@uni-mainz.de

[1] Department of Geography, Johannes Gutenberg University, Becherweg Weg 21, 55128 Mainz, Germany

[2] Global Change Research Institute of the Czech Academy of Science, Bělidla 986/4a, 60300 Brno, Czech Republic

[3] Department of Civil, Environmental and Geodetic Engineering, Ohio State University, 470 Hitchcock Hall, 2070 Neil Ave, Columbus, OH 43210, USA

[4] Institute of International Rivers and Eco-Security, Yunnan University, Kunming 650091, People's Republic of China

[5] Institute of Desert Meteorology, Chinese Meteorological Administration, Urumqi 830002, People's Republic of China

[6] Southwest United Graduate School, Kunming 650500, People's Republic of China

standardized growth of multiple trees from a single location, has facilitated the reconstruction of summer temperatures (Esper et al. 2002; Anchukaitis et al. 2017; Büntgen et al. 2024), drought and precipitation (Cook et al. 2015; Tejedor et al. 2017; Stahle et al. 2020), and streamflow (Woodhouse et al. 2012; Chen et al. 2023; Torbenson et al. 2023a). These reconstructions have often been based on a simple relationship between tree growth and monthly- or seasonally resolved climate data (Cook and Kairiukstis 1990).

As the length and availability of daily resolved climate data, such as precipitation and temperature, has increased, the use of such variables to define "optimum" climate signals in tree-ring parameters have likewise become a more common practice (Caprio et al. 2003; Ziaco and Liang 2019; Howard et al. 2021). Comparing ring width and density measurements, tree-ring stable isotopes, and variability in quantitative wood anatomy to sub-monthly climate data has been further facilitated by the development of user-friendly computer packages (Jevšenak and Levanič 2018; Jevšenak 2020). While these advances undoubtedly have improved our understanding of the multi-faceted relationship between atmospheric conditions and biological activity (Jevšenak 2019), the statistical and practical significances of correlation also become increasingly difficult to interpret. Tree ring proxies are tested against daily climatic records using shifting temporal windows with lengths ranging from a few days up to several months. This approach generates thousands of climatic time series that can be compared to single tree-ring proxies. Several other factors, such as multiple tree-ring parameters and temporal autocorrelation, may further influence the analysis of significance. However, the increased number of calculated correlations not only risks an overestimation of tree ring-climate relationships, but also influences the interpretation of changing signal strengths when this type of analysis is performed for different periods.

The multiple comparisons problem represents a long-standing issue in statistics, with its origins in nineteenth century mathematics (Boole 1847). Fundamental to the issue is an exaggerated risk of type I error, in which a relationship is assumed to be significant based on an arbitrary *p* value but is false (Sokal and Rohlf 1981; Genovese et al. 2006). Advances have been made since the 1800s to develop correction techniques, including those outlined by Bonferroni (1937) and Šidák (1967). However, the inclusion of such corrections in dendroclimatology has been rare but see Meko and Woodhouse (2005) for a commendable example. The Bonferroni approach to correcting statistical significance has often been considered conservative (Conneely and Boehnke 2007), with an increased risk of type II errors, i.e., rejecting hypotheses that are true (Feise 2002). This may particularly occur when data are non-orthogonal, and other methods have been suggested for the ecological sciences (Peres-Neto 1999; Moran 2003; Nakagawa 2004), including the False Discovery Rate (FDR) introduced by Benjamini and Hochberg (1995). As the number of calculated correlations exceeds the thousands, traditional correction techniques may become unviable.

The analysis presented here is tailored to assess the potential impacts that multiple comparisons can have on dendroclimatology. Using synthetic time series of both the predictor tree-ring data and predicted and climate target data, we quantify the relative component of correlation values that can be expected to occur by chance from thousands of comparisons. The results are not meant to question previous interpretations of climate-tree growth relationships based on daily records but rather provide a methodological support for assessing significance of daily climate data correlations. We derive generalized values from our simulations applicable to dendroclimatic studies using daily data. We additionally quantify the difference in correlations stemming from natural climate variability and discuss how they can be accounted for when interpreting comparisons of tree-ring records and daily resolved climate data.

Materials and methods

In this study, we produced synthetic data that mimic proxy and target variables found in a traditional dendroclimatic setting. The analyses rely on simulated time series, but real-world examples are used to infer practical meaning of the results. A general month-to-multi-month precipitation or temperature signal was assumed (based on real-world observations), and daily synthetic data and tree-ring series were derived using the characteristics of these observations.

Tree-ring data, gridded climate observations, and baseline relationships

Ten tree-ring records from the European continent, recently developed or publicly available through the International Tree-Ring Data Bank (ITRDB; Zhao et al. 2019), were selected for analysis (Table 1; Fig. 1). Five of the records display statistically significant Pearson's correlations ($r = 0.45$ to 0.60) with growing season precipitation totals and five with maximum temperatures. The tree-ring data include several different species and variables but share a common period of 1930–2011. Total ring-width (TRW) represents the most common tree-ring variable in our dataset, but adjusted latewood width (LW_a; Meko and Baisan 2001), maximum density (MXD), and a stable isotope series were also included. Raw measurements for each parameter were standardized and detrended using an age-dependent spline (Cook and Peters 1981) to produce site chronologies. For the LW_a chronology, adjustments were performed on detrended individual series rather than the averaged chronology (Meko

Table 1 General characteristics and site locations of tree-ring data used in the analyses, with the corresponding study in which the data was first presented

General characteristics	Site name	Nat	Species	Variable	Lat	Lon	Season	Study
Precipitation	Koblenz	DE	QURO	TRW	50.36	7.61	JJ	This study
	Pefkas	EL	PINI	TRW	37.96	22.40	AMJ	Touchan et al. (2014)
	Eketånga	SE	PISY	LWa	56.65	12.77	July	This study
	Weinviertel	AT	QURO	TRW	48.50	16.40	JJA	Karanitsch-Ackerl et al. (2019)
	Puzz Atelli	FR	PIHA	TRW	42.14	9.12	JJA	Touchan et al. (2017)
Temperature	Torneträsk	SE	PISY	TRW	68.22	19.83	July	Homfeld et al. (2024)
	Beinn Bhreac	UK	PISY	TRW	57.03	–3.03	JA	Rydval et al. (2017)
	Val di Sole	IT	PICE	MXD	46.42	10.69	JAS	Cerrato et al. (2019)
	Czechia	CZ	QUSP	δ^{18}O	N/A	N/A	July	Büntgen et al. (2021)
	Gerber	ES	PIUN	MXD	42.39	0.58	AM	Büntgen et al. (2024)

Nat. = Country code, Lat./lon. = latitude/longitude in ° north/east for the locations of sites

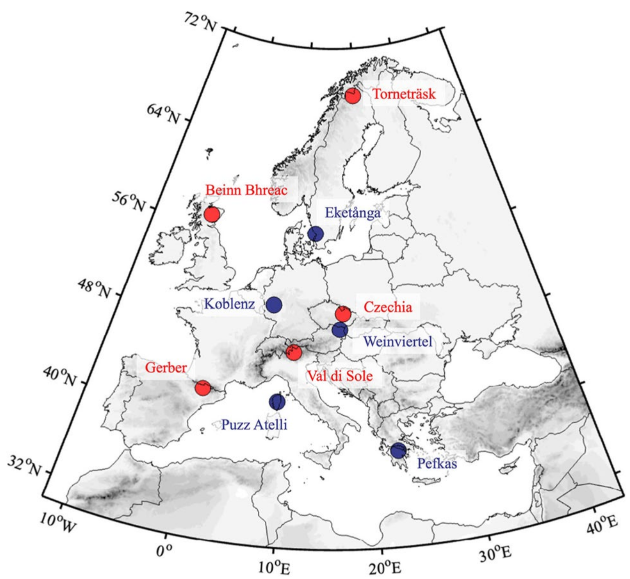

Fig. 1 Map indicating locations of five precipitation-sensitive chronologies (blue) and five temperature-sensitive chronologies (red) in Europe

and Baisan 2001; Torbenson et al. 2016). The stable oxygen isotope record (δ^{18}O) did not contain a significant age-related trend and was therefore not detrended (Urban et al. 2021). The final chronologies for analyses represent the average growth/variability of all series in their respective site collection.

For each selected tree-ring chronology, data from the closest point in the gridded E-OBS network (v.23.1e / 0.25° spatial resolution; Cornes et al. 2018) were extracted. Daily values of precipitation and temperature for 1950–2020 were used as a baseline of climate variability at each location, with the exception of the Pefkas site (Greece) for which continuous data only exist for 1950–2004. The daily climate data were averaged into monthly and seasonal values for the initial comparison with the tree-ring data. Climate signals in the tree-ring chronologies were assessed through Pearson's correlation for the longest possible overlap with the monthly and seasonally averaged E-OBS data for each respective chronology. The month/season chosen does not necessarily represent the highest possible tree ring-climate relationship exhibited at any given site but provides a range of values from real-world settings. All relationships were positive (Table 2), with the exception for the Czech δ^{18}O chronology that displayed negative correlations with temperature—as is expected for the relationship between δ^{18}O and temperature (Torbenson et al. 2023b). Spearman's correlations were also calculated for each record.

Simulation of synthetic data based on real-world observations

Synthetic precipitation and maximum temperature time series of daily resolution were produced using the WeaGETS weather generator (Chen et al. 2010, 2012). Daily precipitation, minimum temperature, and maximum temperature data from E-OBS were entered into the generator to simulate the frequency of days with precipitation and the magnitude of each day, as well as distributions of maximum temperature. Daily precipitation totals for 71-year periods were generated and summed to monthly and seasonal totals. The model was run as a 3rd order Markov chain using a Gamma distribution for precipitation, no smoothing of resulting precipitation, and any daily value > 0.1 mm considered as a day with precipitation. WeaGETS does not consider leap years, and February 29th non-zero precipitation totals (when they occurred) were split evenly onto February 28th and March 1st. Because most European tree-ring records sensitive to hydroclimatic variability display the maximum correlation

Table 2 Chronology statistics

Correlations methods	Periods	Precipitation sensitive sites (r)					Temperature sensitive sites (r)				
		Koblenz	Pefkas	Eketånga	Weinviertel	Puzz Atelli	Torneträsk	Beinn Bhreac	Val di Sole	Czechia	Gerber
AR1		0.53	0.51	–0.05	0.35	0.15	0.58	0.59	0.28	0.58	0.23
Pearson's	Corr	0.45	0.46	0.51	0.55	0.60	0.48	0.53	0.53	0.58*	0.60
	31-year	0.23	0.23	0.17	0.16	0.14	0.18	0.17	0.16	0.14	0.13
	51-year	0.15	0.15	0.11	0.10	0.08	0.13	0.11	0.10	0.09	0.08
	71-year	0.11	0.11	0.08	0.07	0.06	0.09	0.08	0.07	0.07	0.06
Spearman's	Corr	0.44	0.47	0.56	0.55	0.51	0.42	0.56	0.52	0.59*	0.64
	31-year	0.25	0.24	0.17	0.19	0.21	0.23	0.19	0.19	0.15	0.15
	51-year	0.17	0.16	0.11	0.11	0.13	0.16	0.13	0.14	0.10	0.11
	71-year	0.13	0.12	0.08	0.08	0.09	0.13	0.10	0.10	0.08	0.09

Notes: AR1 = autocorrelation at year – 1 (for the period 1930 to the most recent year). Corr. = correlation between real-world tree-ring chronology and climate data from the closest E-OBS grid point (for the period 1950 to the most recent year). The 95th percentile value of the difference between A_x and M_x for three period lengths. *Correlations for the Czech isotope chronology are negative, and the series was therefore inverted

with summer precipitation (St George and Ault 2014), leap years are assumed to play a negligible role in the analysis.

Each tree-ring chronology was decomposed into reconstructed components (RCs) through singular spectrum analysis (SSA; Ghil et al. 2002), which allowed for the decomposing of a time series into spectral components of varying frequencies (Vautard et al. 1992). A window (M) of 20 years was used which is larger than that used by St George and Ault (2011) for which $M = 15$ as the intention was not to extract the strongest periodicities but to capture as complete spectral characteristics of the tree-ring series as possible. Nevertheless, M falls well within 1/3rd of the record length of consideration (i.e., 82 years; 1930–2011) and the resulting spectral components are statistically robust. Sinus waves of periodicities corresponding to the 20 resulting RCs were resampled according to their respective weight (normalized eigenvalue). These RCs were then added to the normalized WeaGETS generated weather data averaged for the target season (here after referred to as the assigned signal time-series; AS_n), e.g., JJA, and linked to each corresponding PTR_n. All daily values for AS_n were retained to facilitate further comparisons with windows of differing number of days. The WeaGETS generated data were weighted (depending on the length of comparison, see below), to produce pseudo tree-ring series (PTRs) that retained the spectral properties of the original chronology. The resulting series were scaled to have the same mean and variance as the original tree-ring data. The PTRs ($PTR_{1...n}$) were correlated with their respective assigned signal time-series ($AS_{1,\ ...,\ n}$) to assure that Pearson's or Spearman's correlations (here after referred to as the assigned correlation; A_x) fell within ten points (+/– 0.05) of the real-world example correlation, e.g., for a chronology exhibiting $r = 0.53$ with local precipitation, only PTRs correlating from 0.48 to 0.58 with the WeaGETS output were considered. Any PTR displaying a correlation outside this range was discarded. To assure that the spectral properties of the final PTRs were not significantly biased by the assigned signal, the power spectra of any given PTR was compared to that of the original tree-ring chronology. The discrete Fourier transform was computed for both series and used to produce a periodogram. A 95% confidence interval around the spectrum was calculated using the chi-squared distribution (Bloomfield 1976). Any PTR that displayed non-overlapping confidence intervals below 20-year periodicities was likewise discarded. The full procedure was repeated with new WeaGETS output until 10,000 pairs ($AS_{1,\ ...,\ 10{,}000}$ and $PTR_{1,\ ...,\ 10{,}000}$, with all daily values for each AS series retained) had been obtained for each of the ten chronologies (Fig. 2).

For the five temperature-sensitive chronologies, an additional modification was made to WeaGETS-generated maximum temperatures to include potential trends and low-frequency behavior in the data. Variability of any

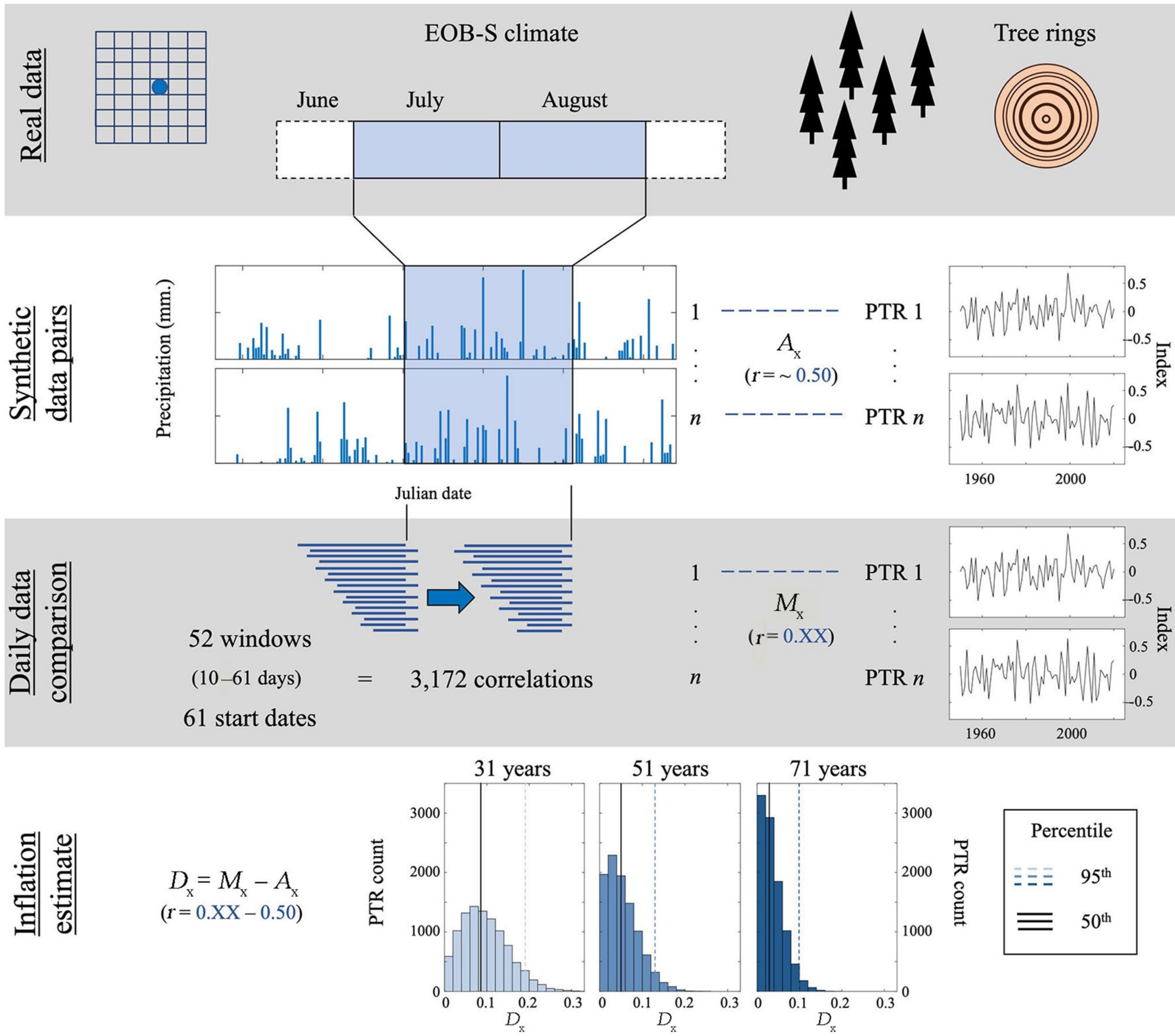

Fig. 2 Overview of the analysis presented in this paper exemplified by a July–August precipitation signal where the A_x is $r = 0.50$

low frequency (> 10-year periodicities) in the first ten RCs (i.e., the components that explained the highest amount of variance; RC1-10 for $M = 20$) of the instrumental temperature data for 1950–2020 were considered. The original generated time-series were averaged to the temporal window of interest (e.g., JJA) and normalized, with any low frequency components identified added in the same way as for the PTRs. The difference between (a) the modified series and (b) the original generated series, seasonally averaged, was added to the original generated data to produce daily temperature series that incorporated the low frequency components of the target data. Comparisons between A_x and the maximum correlation (M_x) for temperature-sensitive chronologies were calculated with the low frequency adjusted temperature series.

Analysis of correlation inflation

Synthetic precipitation totals/temperature means for windows of 10-X days were calculated, with X representing the number of days covered by the assigned signal (e.g., $X = 61$ for a comparison with an assigned JJ signal and $X = 92$ for a JJA signal). These synthetic time-series are independent from the real-world data but behave in the same way in terms of day-to-day changes in temperature and precipitation probability. Correlations were calculated with the candidate climate variables ending on the first day of the assigned signal (e.g., June 1st / 152nd Julian Date; A_x), moving one day at a time and repeated until reaching the final day of the assigned window (e.g., August 31st / 243rd Julian Date) (Fig. 2). For a pseudo tree-ring record

(PTR_n) with an assigned JJA signal, 7,636 correlations were produced (83 windows and 92 days). Correlations were only calculated for windows that overlapped or were fully encompassed by the assigned window. Both Pearson's correlations and Spearman's ranked correlations were calculated to assess the effect each method had on the absolute difference in correlation. The analysis was performed for 31-, 51-, and 71-year periods.

The M_{xn} obtained from the analysis was compared to the A_{xn} for the same PTR_n (which was always within $r = +/- 0.05$ of the real-world data). Because the matrix of correlations calculated also included the seasonal window of AS_n, the M_{xn} value will always be equal to or greater than A_{xn}. The differences (D_x) between the M_x of each of the 10,000 PTRs and their corresponding A_xS were fitted with a Kernel density distribution. The 95th percentile of the distribution was then calculated for each site/chronology to provide an estimate of the magnitude of correlation inflation occurring 1 in 20 times, as well as the mean correlation inflation for the 10,000 comparisons. The 95th percentile represents an estimate of when the maximum correlation (M_x) based on daily-resolved data is stronger than the monthly/season correlation (A_x) beyond chance (at $p = 0.05$).

Simulation of synthetic data to test signal strength influence

In addition to the above simulations and comparisons, a theoretical ensemble of tree-ring chronologies with no direct real-world counterpart was also produced. For each climate variable, i.e., precipitation and maximum temperature, reconstructed components of all five locations were used and chosen at random and weighted by relative normalized eigenvalue. This set of chronologies was used as a basis to produce PTRs with various autocorrelation (at 1-year lag; AR1) magnitudes, A_x of different strengths ($r = 0.45$, 0.50, 0.55, and 0.60), and for three potential signal window lengths (one, two, and three months). An E-OBS grid point from central Germany (50.375°W; 9.875°N) was used for the climate comparison, for JJA precipitation and maximum temperature. The range of A_x is within the observed relationship between tree-ring records and local climate variables for the region (St George and Ault 2014), as well as the full dataset analyzed here and should be considered realistic. This suite of chronologies produced from PTRs with similar time series characteristics, allows for the isolation of a single parameter to be assessed as an influence on correlation inflation. Additionally, the same PTRs were produced for a scenario where $A_x = 0.00$ (+/− 0.05) to assess how strong spurious correlations can become from a large number of comparisons of initially unrelated series.

Results

For each of the ten chronologies, 10,000 PTRs mimicking the spectral properties of the real-world data were produced. Each PTR was correlated similarly ($r = +/- 0.05$) with the monthly totals/averages of precipitation/maximum temperature generated from the WeaGETS simulator to that of the real tree-ring chronology against real observed climate data. The daily data making up the totals/averages differed for each PTR but were generated based on the instrumental data. As such, these results should be interpreted as an estimate of shared variability stemming purely from a large number of comparisons.

Distribution of correlation inflation in precipitation-sensitive settings

The maximum Pearson's correlations (M_x) from comparison between PTRs and the WeaGETS generated daily precipitation data exceeded their respective assigned (A_x) correlations by up to $r = 0.23$ (for the site from Sweden, adjusted latewood width, July, 31 years; SE-LW_a-J-31) when considering the 95th percentile (Table 2; Fig. 3b). The lowest D_x value was recorded for FR-TRW-JJA-71 ($r = 0.06$). For all the precipitation-related tree-ring records, the differences between A_x and M_x decreased with length of comparison regardless of number of observations, as shown by darker lines representing longer record length (Fig. 3). The same pattern holds true when considering the mean (i.e., the 50th percentile), with average D_x values ranging from 0.02 to 0.12 (Supplementary Table 1). D_x was larger for sites with lower A_x, largely due to greater space to the $r = 1.00$ upper bound. This suggests that sites with lower "true" correlation may be more susceptible to inflated correlations due to repeated comparisons.

The differences between A_x and M_x for Spearman's correlations (D_x) were larger than those obtained using Pearson's correlations (Fig. 3). In some environments, parametric tests may not be suitable because the target (precipitation) data is not always normally distributed and/or they contain outliers. For example, of the 10,000 generated July precipitation time-series produced from WeaGETS for the Eketånga LW_a chronology 11%–9%, depending on number of years produced, fail a Lilliefors test for normality (Conover 1980). For the 10-day windows, less than 50% of the 310,000 generated precipitation sum series passed the same test when calculated for the 71-year period. In such cases, Spearman's correlation tests may be more appropriate.

Distributions of correlation inflation in temperature-sensitive settings

The same general patterns recorded for precipitation-sensitive chronologies are also apparent for the five

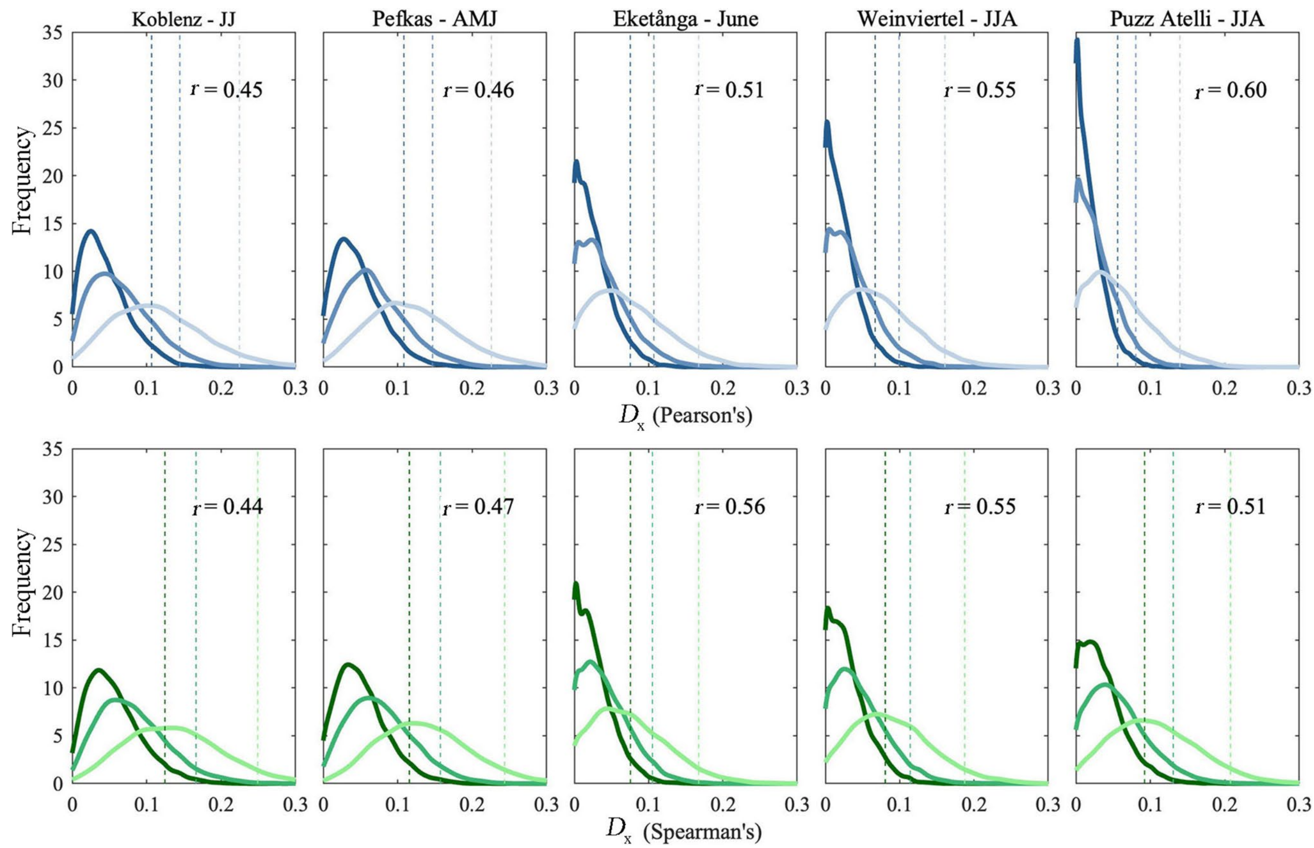

Fig. 3 Kernel density distributions of the difference between assigned and maximum Pearson's correlations (top) and Spearman's correlations (bottom) for five precipitation-sensitive chronologies analyzed ($n=10{,}000$). Dark blue/green = 71-year period, blue/green = 51-year period, light blue/green = 31-year period. The dashed lines represent the upper 5th percentile of values, of respective color. *R*-values indicate the Pearson's and Spearman's correlation between the real-world tree-ring and precipitation data

temperature-related records (Fig. 4), with the largest D_x values recorded for the 31-year periods. The overall differences are slightly smaller, however. The highest 95th percentile inflation for the five temperature-sensitive records was found for SE-TRW-J-31 ($r=0.18$; Table 2). Mean D_x ranged from 0.02 to 0.08, with the lowest values recorded for CZ-δ^{18}O-J-71 (Supplementary Table 1). Spearman's correlation differences (D_x) were overall lower for temperature-sensitive chronologies compared to those of the precipitation-sensitive records, although A_x values were also slightly higher. Similar to the precipitation comparison, Spearman's D_x was higher than Pearson's D_x across all periods and chronologies (Table 2).

Influence of signal strength and autocorrelation

The generic PTRs produced without real-world equivalences allowed for the isolation of various parameters such as signal strength and autocorrelation, and the assessment of their relative impact on increased correlations (Tables 3, 4). For all windows of comparisons (i.e., 31- to 71-year), there was a general decrease in D_x with increasing A_x (Fig. 5). Similarly, D_x decreased with shorter windows of A_x (e.g., one month compared to three months). These results agree with those of the real-world examples (Table 2). The results display greater differences between the 31-year and 51-year periods versus the 51-year and 71-year periods, suggesting a negative exponential relationship between spurious correlation inflation and comparison sample size. The 95th percentile for the highest values (precipitation; $A_x=0.45$; three months window) of the 31-year period (Pearson's $D_x=0.24$ and Spearman's $D_x=0.26$) are nine points higher than the same test for the 51-year period (Pearson's $D_x=0.15$ and Spearman's $D_x=0.17$) (Table 3). For the 71-year period, the drop was four points. Similar differences are also recorded for mean D_x (Supplementary Tables 2, 3).

The results for $A_x=0.00$ show the highest differences with 95th percentile M_x of all calculated comparisons, both for Pearson's and Spearman's correlation (Tables 3, 4). In several cases, the correlations between the PTRs and daily data exceeded $r=0.50$. As with the real-world examples, the magnitude of spurious correlation was the highest for the

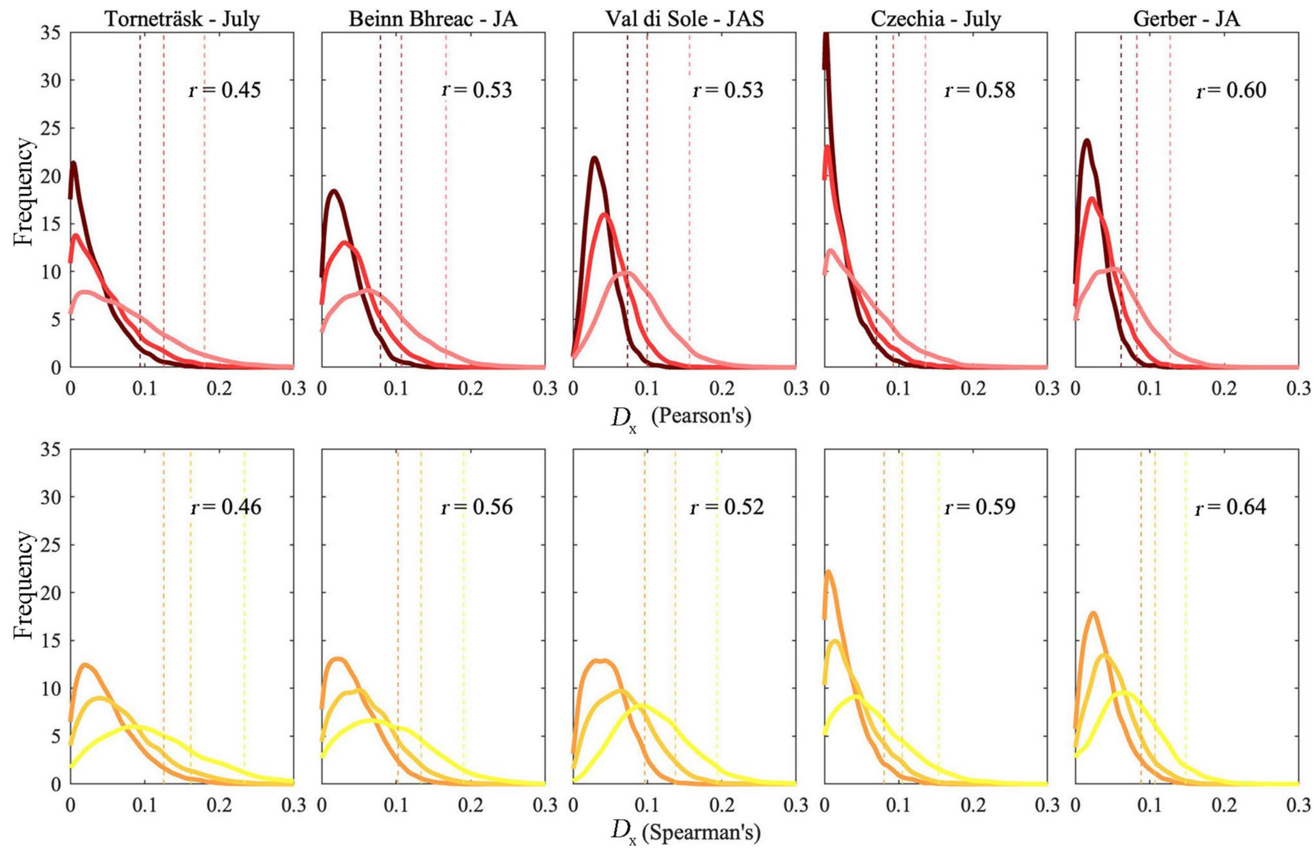

Fig. 4 Kernel density-fitted distribution of differences between assigned and maximum Pearson's correlations (top) and Spearman's correlations (bottom) for five temperature-sensitive chronologies analyzed (n = 10,000). Dark red/orange = 71-year period, red/orange = 51-year period, light red/yellow = 31-year period. The dashed lines represent the upper 5th percentile of values, of respective color. *R*-values indicate the Pearson's and Spearman's correlation between the real-world tree-ring and precipitation data

31-year period. For precipitation, the mean D_x values (Supplementary Tables 2, 3) indicate that spurious correlations up to 0.39 can be expected for a 31-year period on average. Of the 10,000 comparisons for three months windows, over 60% displayed correlations greater than the threshold for a standard p value < 0.05 ($r = 0.356$; 0.276; and 0.234 for 31-, 51-, and 71-year comparison, respectively). The data underlying the distributions for all results are provided through a data repository (see below).

Discussion

Our analyses are an assessment of the influence multiple comparisons can have on maximum or "optimum" correlation in common settings of dendroclimatic studies. We compared assigned monthly or seasonal signals to the correlations of daily data that can occur through PTRs, with the distribution and behavior of daily values at a given location being derived from the characteristics of real climate data from a respective location. Because the signals studied are assigned, and therefore deemed "true", any difference between M_x and A_x must be considered spurious. By extension, these values represent an estimate of type I error—a problem affecting many aspects of dendrochronology (e.g., Helama 2023).

It should be noted that it is just as unlikely that arbitrary months or seasons are the true signal in real-life data as it is for any other temporal window (Jevšenak 2019). Furthermore, correlations calculated for data of monthly and/or seasonal windows are not immune to the problem of multiple comparisons either. Although our study focuses on correlation inflation that arises specifically from comparisons with hundreds to thousands of non-independent time-series, these issues should always be considered. We acknowledge that our analysis is not exhaustive, as there are a large number of parameters that play a role in the correlations between tree-ring chronologies and climate variables. These include the assumption of non-bias as it relates to sign (Wise and Dannenberg 2019), and the potential role of non-stationary processes and of anthropogenic forcings on tree growth (Bogachev et al. 2024). Our results are limited

Table 3 Generalized correction information for the inflation of correlation stemming from multiple comparisons for precipitation-sensitive tree-ring chronologies

Periods	Correlations methods	Various parameters	$A_x = 0.45$ (r)			$A_x = 0.50$ (r)			$A_x = 0.55$ (r)			$A_x = 0.60$ (r)			$A_x = 0.00$ (r)		
31-year	Pearson's	AR1	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m
		< 0.15	0.20	0.22	0.24	0.17	0.20	0.21	0.15	0.17	0.18	0.13	0.14	0.15	0.46	0.52	0.55
		0.15 to 0.30	0.20	0.22	0.24	0.18	0.19	0.21	0.15	0.17	0.18	0.13	0.15	0.15	0.46	0.51	0.54
		0.31 to 0.45	0.20	0.22	0.24	0.18	0.20	0.21	0.15	0.17	0.18	0.13	0.15	0.15	0.46	0.51	0.55
		> 0.45	0.20	0.22	0.24	0.17	0.19	0.21	0.15	0.17	0.18	0.13	0.15	0.15	0.46	0.51	0.55
	Spearman's	AR1	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m
		< 0.15	0.22	0.25	0.26	0.20	0.22	0.23	0.17	0.19	0.20	0.15	0.17	0.18	0.47	0.53	0.57
		0.15 to 0.30	0.22	0.25	0.26	0.20	0.22	0.23	0.17	0.20	0.20	0.15	0.17	0.17	0.48	0.54	0.57
		0.31 to 0.45	0.22	0.24	0.26	0.20	0.22	0.23	0.17	0.19	0.20	0.15	0.17	0.18	0.48	0.54	0.57
		> 0.45	0.22	0.25	0.26	0.20	0.22	0.23	0.17	0.19	0.20	0.15	0.17	0.18	0.48	0.53	0.56
51-year	Pearson's	AR1	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m
		< 0.15	0.13	0.14	0.15	0.11	0.12	0.13	0.09	0.11	0.11	0.08	0.09	0.10	0.36	0.40	0.43
		0.15 to 0.30	0.13	0.14	0.15	0.11	0.12	0.13	0.09	0.11	0.12	0.08	0.09	0.10	0.36	0.41	0.43
		0.31 to 0.45	0.13	0.14	0.16	0.11	0.13	0.13	0.09	0.11	0.11	0.08	0.09	0.10	0.36	0.40	0.44
		> 0.45	0.13	0.15	0.15	0.11	0.13	0.13	0.10	0.11	0.11	0.08	0.09	0.10	0.36	0.41	0.43
	Spearman's	AR1	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m
		< 0.15	0.14	0.16	0.17	0.13	0.14	0.15	0.11	0.12	0.13	0.09	0.10	0.11	0.37	0.42	0.45
		0.15 to 0.30	0.14	0.16	0.17	0.13	0.14	0.15	0.11	0.12	0.13	0.10	0.11	0.11	0.37	0.42	0.45
		0.31 to 0.45	0.15	0.16	0.17	0.12	0.14	0.15	0.11	0.12	0.13	0.09	0.10	0.11	0.37	0.42	0.44
		> 0.45	0.15	0.17	0.17	0.13	0.14	0.15	0.11	0.12	0.13	0.10	0.11	0.11	0.37	0.42	0.44
71-year	Pearson's	AR1	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m
		< 0.15	0.09	0.10	0.11	0.08	0.09	0.09	0.07	0.08	0.08	0.06	0.06	0.07	0.30	0.35	0.37
		0.15 to 0.30	0.09	0.11	0.11	0.08	0.09	0.10	0.07	0.08	0.08	0.06	0.06	0.07	0.31	0.34	0.37
		0.31 to 0.45	0.09	0.11	0.11	0.08	0.09	0.09	0.07	0.08	0.08	0.06	0.06	0.07	0.30	0.35	0.37
		> 0.45	0.10	0.11	0.11	0.08	0.09	0.10	0.07	0.08	0.08	0.06	0.07	0.07	0.30	0.34	0.37
	Spearman's	AR1	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m
		< 0.15	0.11	0.12	0.12	0.09	0.10	0.11	0.08	0.09	0.09	0.07	0.07	0.08	0.32	0.35	0.38
		0.15 to 0.30	0.11	0.12	0.12	0.09	0.10	0.11	0.08	0.09	0.09	0.07	0.08	0.08	0.31	0.36	0.38
		0.31 to 0.45	0.11	0.12	0.12	0.09	0.10	0.11	0.08	0.09	0.09	0.07	0.08	0.08	0.31	0.35	0.38
		> 0.45	0.11	0.12	0.13	0.09	0.10	0.11	0.08	0.09	0.09	0.07	0.08	0.08	0.31	0.36	0.38

Table 4 Generalized correction information for the inflation of correlation stemming from multiple comparisons for temperature-sensitive tree-ring chronologies

Periods	Correlations methods	Various parameters	$A_x = 0.45$ (r)			$A_x = 0.50$ (r)			$A_x = 0.55$ (r)			$A_x = 0.60$ (r)			$A_x = 0.00$ (r)		
31-year	Pearson's	AR1	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m
		0.15 to 0.30	0.18	0.20	0.21	0.15	0.18	0.19	0.13	0.15	0.16	0.12	0.13	0.14	0.42	0.48	0.51
		0.31 to 0.45	0.17	0.20	0.21	0.15	0.18	0.19	0.14	0.15	0.16	0.12	0.13	0.14	0.42	0.47	0.50
		> 0.45	0.18	0.20	0.22	0.16	0.18	0.19	0.14	0.15	0.16	0.12	0.13	0.14	0.41	0.48	0.51
	Spearman's	AR1	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m
		0.15 to 0.30	0.20	0.23	0.24	0.18	0.20	0.21	0.16	0.18	0.19	0.14	0.15	0.16	0.44	0.50	0.53
		0.31 to 0.45	0.20	0.23	0.24	0.18	0.20	0.21	0.16	0.18	0.19	0.14	0.16	0.16	0.43	0.49	0.53
		> 0.45	0.20	0.23	0.24	0.18	0.21	0.22	0.16	0.18	0.19	0.14	0.16	0.17	0.44	0.50	0.53
51-year	Pearson's	AR1	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m
		0.15 to 0.30	0.11	0.13	0.14	0.10	0.11	0.12	0.08	0.09	0.10	0.07	0.08	0.09	0.32	0.37	0.39
		0.31 to 0.45	0.11	0.13	0.14	0.10	0.11	0.12	0.09	0.09	0.10	0.07	0.08	0.09	0.33	0.37	0.40
		> 0.45	0.12	0.13	0.15	0.10	0.11	0.13	0.09	0.10	0.10	0.07	0.08	0.09	0.32	0.37	0.40
	Spearman's	AR1	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m
		0.15 to 0.30	0.13	0.15	0.16	0.11	0.13	0.14	0.10	0.11	0.12	0.09	0.09	0.10	0.34	0.38	0.41
		0.31 to 0.45	0.13	0.15	0.16	0.12	0.13	0.14	0.10	0.11	0.12	0.09	0.09	0.10	0.33	0.39	0.41
		> 0.45	0.13	0.15	0.17	0.11	0.13	0.14	0.10	0.11	0.13	0.08	0.09	0.11	0.33	0.39	0.42
71-year	Pearson's	AR1	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m
		0.15 to 0.30	0.08	0.09	0.10	0.07	0.08	0.08	0.06	0.07	0.07	0.05	0.05	0.06	0.27	0.31	0.34
		0.31 to 0.45	0.08	0.09	0.10	0.07	0.08	0.08	0.06	0.07	0.07	0.05	0.05	0.06	0.27	0.32	0.34
		> 0.45	0.08	0.09	0.10	0.07	0.08	0.09	0.06	0.07	0.08	0.05	0.05	0.06	0.27	0.32	0.34
	Spearman's	AR1	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m	1 m	2 m	3 m
		0.15 to 0.30	0.10	0.11	0.12	0.08	0.09	0.10	0.07	0.08	0.08	0.06	0.06	0.07	0.28	0.33	0.35
		0.31 to 0.45	0.09	0.11	0.12	0.08	0.09	0.10	0.07	0.08	0.09	0.06	0.07	0.07	0.28	0.33	0.35
		> 0.45	0.09	0.11	0.12	0.08	0.09	0.11	0.07	0.08	0.09	0.06	0.07	0.08	0.29	0.32	0.35

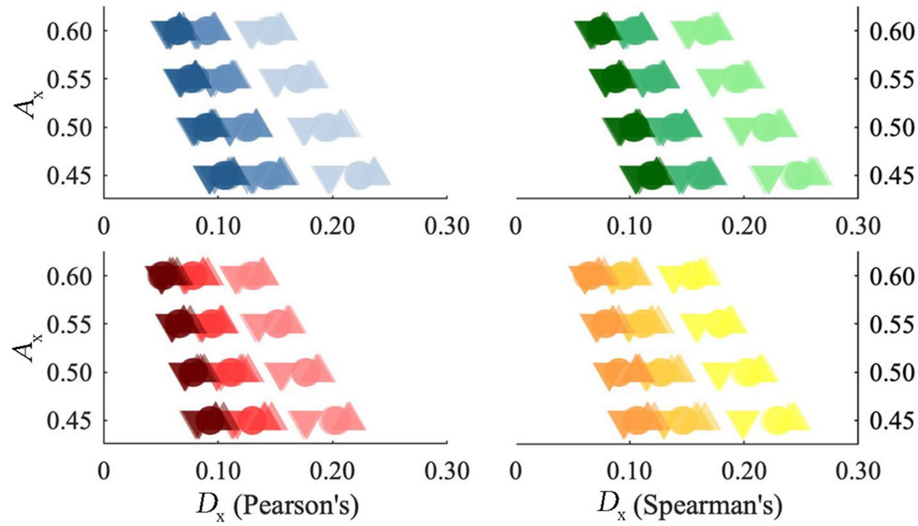

Fig. 5 Relationship between assigned signal and 95th percentile of D_x (maximum minus assigned correlation) for (a) Pearson's and (b) Spearman's correlations of precipitation-sensitive series, and (c) Pearson's and (d) Spearman's correlations of temperature-sensitive series. Up-pointing triangles show results for 31-year, circles 51-year, and down-pointing triangles 71-year periods. Four/three markers are plotted for each case, representing different AR1 values

to precipitation totals and maximum temperatures for one to three months. However, tree growth is likely related to longer windows of climate variability in some environments (Stahle et al. 2020; Pompa-García et al. 2021; Büntgen et al. 2024). In such settings, the number of calculated correlations is much greater than what we tested here, even if limiting the analysis to longer minimum windows (e.g., 21 days). Furthermore, we only assessed the inflated correlations within the assigned window of signal when it is possible for M_x to fall outside the "true" (assigned) window with a large enough number of comparisons.

Differences between assigned and maximum correlations

Several different tree-ring proxies were used as real-world examples but there was no clear influence on the results. Ultimately, the original tree-ring data and resulting PTRs are only a manifestation of different time-series characteristics including signal, autocorrelation, and series length. Whereas the presented analysis could have been performed on strictly synthetic data, however, the real-world examples provide realistic boundaries in terms of tree-ring variables and climatology. They also provide practical examples of how testing the influence of multiple comparisons on correlation patterns for a single site chronology could be undertaken.

The difference between assigned and maximum correlations varied greatly across the analyzed data and periods. As expected, fewer years of comparison produced higher differences between A_x and M_x (Figs. 3, 4), with 1-in-20 comparisons inflating correlation by up to 25 points. Such inflation will undoubtedly have an impact on the interpretation of results. For example, a correlation difference between M_x and A_x of 0.18 for a setting with an assigned signal of $r=0.5$ produced an inflated explained variance of 21% in a simple linear regression model, and even a D_x of 0.10 would result in 10+% of additional explained variance. Even for the longest period of comparison ($n=71$), D_x was above 0.10 for several of the examples analyzed here. Uncertainty ranges for any reconstruction based on such relationships would therefore considerably underestimate the true amount of unexplained variance.

The D_x values presented here can be further generalized, as visualized in Fig. 6. Some of these generalizations may be expected, but Fig. 5 and Tables 3, 4 provide quantifications for these relationships. Inflation is less severe for shorter windows of assigned signal for which the number of comparisons is considerably lower. An assigned signal of a three-month window produces over 7,500 correlations in the presented setup, but less than 700 are calculated for a single month A_x. Similarly, there is a negative relationship between A_x and D_x. When the starting point of correlation is high, there is less room for improvement as there is a ceiling in correlation, i.e., $r=1.00$. This pattern is highlighted for the unrelated PTRs ($A_x=0.00$), for which a considerable number of the synthetic data shows correlations with daily windows that would be considered statistically significant if the multiple comparisons are not considered.

Recommendations for interpretation

The analysis presented here is not necessarily easy to reproduce, but we provide generalized values from this study in Tables 3, 4. We do not suggest that these values are universal or absolute but rather that they can provide a guideline for assessing the role multiple comparisons play in overestimation of significance as it pertains to dendroclimatological studies. These estimates do, however, represent the general range of values one might expect in a similar setting and the results from "real-world" data (Table 2) align well with the

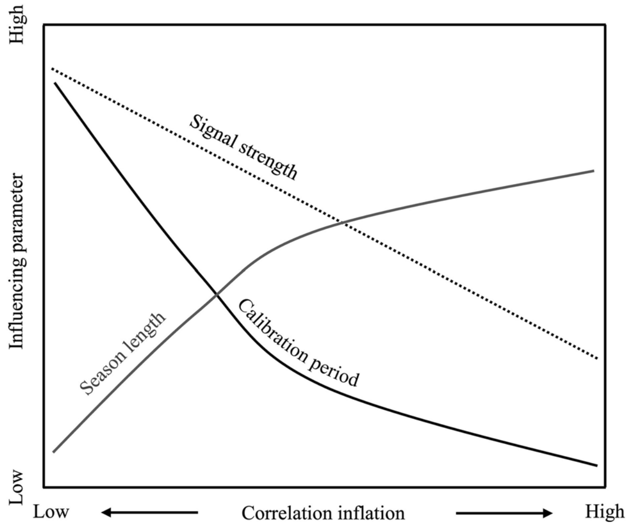

Fig. 6 Diagram of the general patterns of correlation inflation drivers suggested by the results

generalized inflations in Tables 3, 4. Care should be taken when interpreting the relationship between tree-ring chronologies and daily resolved climate data, especially when monthly or seasonal correlations are weak, and themselves potentially considered statistically significant due to multiple comparisons.

The high D_x values for PTRs with an assigned signal of $r = 0.00$ indicate the magnitude of the multiple comparisons problem in dendroclimatology. Inflated correlations of $r > 0.50$ were recorded for the 95th percentiles (Tables 3, 4). Such correlations would be highly statistically significant if using p values of single-comparison tests. Ultimately, the high number of time-series pairs caused inflation of correlation that rivals that which may be expected in European dendroclimatic settings. Therefore, the use of daily data could, in the worst case scenario, even mask any "true" relationship between tree-growth and climate. Conversely, the generalized estimates presented can provide statistical support in cases for which tree-ring chronologies display strong correlations with climate of specific sub-monthly or sub-seasonal windows. For example, a record with a correlation of $r = 0.68$ with precipitation totals for a 13-day window but only $r = 0.45$ with its strongest seasonal (1- or 2-months) window would occur less than one-in-twenty times by chance. We suggest that these estimates may be used to assess if a relationship between a tree-ring proxy and climate variability for a specific daily-resolved window is significantly different from the strongest monthly/seasonal correlation. If the difference (D_x) is less than corresponding values in Tables 3, 4, it should be treated with caution. Furthermore, adjustments to significance levels (such as FDR, Benjamini and Hochberg 1995; Jafari and Ansari-Pour 2019) can be performed using the outputs of these tests.

Beyond these values, we strongly urge future studies to consider whether or not the tests applied are appropriate for the data. We show that Pearson's correlation produces lower differences (D_x) between A_x and M_x compared to Spearman's correlation. Although Pearson's correlation does not necessarily assume normality, it can be sensitive to outliers (e.g., Zuur et al. 2010)—issues that are sometimes overlooked in the field of dendroclimatology. The issue of outliers and non-normality is greater for averages of shorter windows, and thus likely to affect daily data more than longer seasonal windows. Regardless, the larger differences for Spearman's correlation for PTRs across assigned precipitation signals indicate that the true problem is perhaps greater than what is suggested by the Pearson's values.

Lastly, we also want to highlight alternative approaches to utilizing information from daily resolved data without greatly increasing the number of correlations, with notable early examples, including Woodhouse and Meko (1997) and Yuan et al. (2003). More recently, similar studies that compare the frequency of specific events within a season have been published (Howard and Stahle 2020; Howard et al. 2023; Lee and Dannenberg 2023). We speculate that such methods are more robust than moving windows of correlations, as they limit the number of comparisons performed. Other approaches that explicitly consider the timing of physiological activity (e.g., Unterholzner et al. 2024) may also decrease the risk of overinterpretation. By confining the analysis to windows within or near the growing season or testing monthly-to-seasonal windows prior to the use of higher resolution data, the number comparisons can be lowered and the risk of spurious relationships subsequently limited.

Future directions for study

Our results do not show any significant differences for varying autocorrelation (or serial correlation; AR1) in the tree-ring record (Tables 3, 4). However, autocorrelation does have an effect on the distribution and regression of any time-series (Dawdy and Matalas 1964), and it is possible that autocorrelation in the "target", i.e., the assigned climate signal, could play a role in the relationship studied. The small differences between precipitation and temperature D_x displayed may hint at this, as autocorrelation differs between the two climate variables. Therefore, local climatology may play a role in the inflation of correlation with daily data. Outside of Europe, daily resolved climate data, or monthly for that matter, may be much more limited in its temporal extent. Our results suggest that the issue of multiple comparisons may be even greater in such places, as the exaggeration of correlations are of greater magnitude for shorter windows (e.g., 31 compared to 71 year-long correlations; Fig. 3). Ultimately, these comparisons do not require real-world data beyond the characteristics of the tree-ring series and the local climatology. However, in places like the United States, which is home to a wealth of tree-ring data (St George and Ault 2014; Stahle et al. 2020) as well as long records of daily weather observations (Menne et al. 2015), work similar to ours could provide complimentary results, as it would incorporate environments that either dendrochronologically or climatologically fall beyond the conditions of sites tested here.

The problem of multiple comparisons is also affecting ecological interpretations of tree productivity through tree rings, including phenological changes (Yang et al. 2017). A tree-ring chronology that displays its highest correlation with June 12th to June 29th precipitation totals for period 1 and with, e.g., June 25th to July 8th for period 2, is not necessarily showing evidence of a robust shift in signal. Approaches to test such differences, e.g., based on Fisher (1921), as implemented by Meko et al. (2011) or bootstrapping, as presented by Jevšenak (2019) would need to consider the large number of comparisons produced. Although we do not explicitly address such questions, our results hint at how this issue could influence interpretations of temporally changing correlations based on daily-resolved climate data. Future studies should attempt to confine this error and provide estimates for associated uncertainties.

Our results represent the first thorough analysis of how multiple correlations can affect statistical significance as it pertains to daily-resolved climate data in the field of dendroclimatology. However, they are only a small contribution to a larger scientific discussion on the adjustment of p values (Jafari and Ansari-Pour 2019). That discussion, in itself, is part of ongoing conversations that concern how researchers perceive and use statistical significance (Lin et al. 2013; Wasserstein and Lazar 2016). Beyond the statistical considerations presented here, any robust interpretation will undoubtedly benefit from expert knowledge and inferences of growing season length, biophysical factors, and local climatology.

Conclusion

We present an analysis of the multiple comparisons problem as it pertains to dendroclimatology and the use of subseasonal climate data. The differences between the assigned signal of the PTR and the maximum recorded correlation are considerable. As expected, the difference increases as the number of observations decreases. The issue appears greater for precipitation-sensitive records compared to those with a temperature signal, although only marginally. Our results indicate that the magnitude of difference stemming from inflated correlations decreases as the assigned signal strength increases. Most of these limitations can be overcome by testing the potential of inflated correlation on a site-by-site basis.

Ultimately, the responsibility of interpreting statistical and practical significances of correlations between tree-ring chronologies and climate data lay firmly in the hands of the user. The development of increasingly computationally demanding methods for comparing proxy data with high-resolution climate observations is embraced by the wider tree-ring community. However, if basic statistical considerations are not made, there is an inherent risk for overinterpretation. It is clear that a high number of correlations between initially unrelated time-series of similar characteristics to those of real-world tree-ring and climate data can lead to perceived statistical significance. We hope that the results presented here may help limit such issues.

Acknowledgements We appreciate the wealth of computer code made publicly available by the larger dendrochronological community, with special appreciation given to Drs. K.J. Anchukaitis, E.R. Cook, J. Jevšenak, and D.M. Meko. We also acknowledge all contributors of data to the ITRDB, and T. Larsson for assistance during sampling of the Eketånga collection. MCAT and JE are supported by the ERC Advanced Grant Monostar (AdG 882727) and the co-funded EU project AdAgriF (CZ.02.01.01/00/22_008/0004635).

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability Outputs from the simulation runs are provided in an OSF depository (https://osf.io/qsyzg). These include the full set of differences between A_x and M_x for the theoretical chronologies, which make up Tables 3, 4.

References

Anchukaitis KJ, Wilson R, Briffa KR, Büntgen U, Cook ER, D'Arrigo R, Davi N, Esper J, Frank D, Gunnarson BE, Hegerl G, Helama S, Klesse S, Krusic PJ, Linderholm HW, Myglan V, Osborn TJ, Zhang P, Rydval M, Schneider L, Schurer A, Wiles G, Zorita E (2017) Last millennium Northern Hemisphere summer temperatures from tree rings: part II, spatially resolved reconstructions. Quat Sci Rev 163:1–22. https://doi.org/10.1016/j.quascirev.2017.02.020

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc

Ser B Methodol 57(1):289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bloomfield P (1976) Fourier analysis of time series: an introduction. John Wiley and Sons, New York

Bogachev MI, Grigoriev AA, Pyko NS, Gulin AN, Grigorieva AV, Chindyaev AS, Kayumov AR, Tishin DV (2024) Detection and evaluation of anthropogenic impacts on natural forest ecosystems from long-term tree-ring observations. For Ecol Manag 558:121784. https://doi.org/10.1016/j.foreco.2024.121784

Bonferroni CE (1937) Volume in Onore di Racarddo dalla Volta. Universita di Firenze, Teoria statistica delle classi e calcolo delle probabilita, p 62

Boole G (1847) The mathematical analysis of logic. Henderson & Spalding, London, UK, pp 85

Büntgen U, Urban O, Krusic PJ, Rybníček M, Kolář T, Kyncl T, Ač A, Koňasová E, Čáslavský J, Esper J, Wagner S, Saurer M, Tegel W, Dobrovolný P, Cherubini P, Reinig F, Trnka M (2021) Recent European drought extremes beyond Common Era background variability. Nat Geosci 14:190–196. https://doi.org/10.1038/s41561-021-00698-0

Büntgen U, Reinig F, Verstege A, Piermattei A, Kunz M, Krusic P, Slavin P, Štěpánek P, Torbenson M, del Castillo EM, Arosio T, Kirdyanov A, Oppenheimer C, Trnka M, Palosse A, Bebchuk T, Camarero JJ, Esper J (2024) Recent summer warming over the western Mediterranean region is unprecedented since medieval times. Glob Planet Change 232:104336. https://doi.org/10.1016/j.gloplacha.2023.104336

Caprio JM, Fritts HC, Holmes RL, Meko DM, Hemming DL (2003) A Chi-square test for the association and timing of tree ring-daily weather relationships: a new technique for dendroclimatology. Tree Ring Res 59(2):99–113. https://doi.org/10.1080/02827580308621

Cerrato R, Salvatore MC, Gunnarson BE, Linderholm HW, Carturan L, Brunetti M, De Blasi F, Baroni C (2019) A *Pinus cembra* L. tree-ring record for late spring to late summer temperature in the Rhaetian Alps. Italy Dendrochronologia 53:22–31. https://doi.org/10.1016/j.dendro.2018.10.010

Chen J, Brissette FP, Leconte R (2010) A daily stochastic weather generator for preserving low-frequency of climate variability. J Hydrol 388(3–4):480–490. https://doi.org/10.1016/j.jhydrol.2010.05.032

Chen J, Brissette FP, Leconte R (2012) WeaGETS–a Matlab-based daily scale weather generator for generating precipitation and temperature. Procedia Environ Sci 13:2222–2235. https://doi.org/10.1016/j.proenv.2012.01.211

Chen F, Man WM, Wang SJ, Esper J, Meko D, Büntgen U, Yuan YJ, Hadad M, Hu M, Zhao XE, Roig FA, Fang OY, Chen YP, Zhang HL, Shang HM, Yu SL, Luo X, He DM, Chen FH (2023) Southeast Asian ecological dependency on Tibetan Plateau streamflow over the last millennium. Nat Geosci 16(12):1151–1158. https://doi.org/10.1038/s41561-023-01320-1

Conneely KN, Boehnke M (2007) So many correlated tests, so little time! Rapid adjustment of P values for multiple correlated tests. Am J Hum Genet 81(6):1158–1168. https://doi.org/10.1086/522036

Conover WJ (1980) Practical nonparametric statistics, 2nd edn. John Wiley, New York, NY, USA, pp 466–467

Cook ER, Kairiukstis L (1990) Methods of dendrochronology. Springer, New York, NY. https://doi.org/10.1007/978-94-015-7879-0

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

Cook ER, Seager R, Kushnir Y, Briffa KR, Büntgen U, Frank D, Krusic PJ, Tegel W, van der Schrier G, Andreu-Hayles L, Baillie M, Baittinger C, Bleicher N, Bonde N, Brown D, Carrer M, Cooper R, Čufar K, Dittmar C, Esper J, Griggs C, Gunnarson B, Günther B, Gutierrez E, Haneca K, Helama S, Herzig F, Heussner KU, Hofmann J, Janda P, Kontic R, Köse N, Kyncl T, Levanič T, Linderholm H, Manning S, Melvin TM, Miles D, Neuwirth B, Nicolussi K, Nola P, Panayotov M, Popa I, Rothe A, Seftigen K, Seim A, Svarva H, Svoboda M, Thun T, Timonen M, Touchan R, Trotsiuk V, Trouet V, Walder F, Ważny T, Wilson R, Zang C (2015) Old World megadroughts and pluvials during the Common Era. Sci Adv 1(10):e1500561. https://doi.org/10.1126/sciadv.1500561

Cornes RC, van der Schrier G, van den Besselaar EJM, Jones PD (2018) An ensemble version of the E-OBS temperature and precipitation data sets. J Geophys Res Atmos 123(17):9391–9409. https://doi.org/10.1029/2017jd028200

Dawdy DR, Matalas NC (1964) Analysis of variance, covariance and time series (Chapter 8-III). In: Te Chow V (ed) Handbook of Applied Hydrology. McGraw-Hill, New York, United States

Esper J, Cook ER, Schweingruber FH (2002) Low-frequency signals in long tree-ring chronologies for reconstructing past temperature variability. Science 295(5563):2250–2253. https://doi.org/10.1126/science.1066208

Feise RJ (2002) Do multiple outcome measures require p-value adjustment? BMC Med Res Methodol 2:8. https://doi.org/10.1186/1471-2288-2-8

Fisher RA (1921) On the 'probable error' coefficient of correlation deduced from a small sample. Metron 1:3–32

Fritts HC (1976) Tree rings and climate. Academic Press, New York, Academic Press, London, UK, p 567

Genovese CR, Roeder K, Wasserman L (2006) False discovery control with p-value weighting. Biometrika 93(3):509–524. https://doi.org/10.1093/biomet/93.3.509

Ghil M, Allen MR, Dettinger MD, Ide K, Kondrashov D, Mann ME, Robertson AW, Saunders A, Tian Y, Varadi F, Yiou P (2002) Advanced spectral methods for climatic time series. Rev Geophys. https://doi.org/10.1029/2000RG000092

Helama S (2023) Distinguishing type I and II errors in statistical tree-ring dating. Quat Geochronol 78:101470. https://doi.org/10.1016/j.quageo.2023.101470

Homfeld IK, Büntgen U, Reinig F, Torbenson MCA, Esper J (2024) Application of RCS and signal-free RCS to tree-ring width and maximum latewood density data. Dendrochronologia 85:126205. https://doi.org/10.1016/j.dendro.2024.126205

Howard IM, Stahle DW (2020) Tree-ring reconstruction of single-day precipitation totals over eastern Colorado. Mon Weather Rev 148(2):597–612. https://doi.org/10.1175/mwr-d-19-0114.1

Howard IM, Stahle DW, Torbenson MCA, Griffin D (2021) The summer precipitation response of latewood tree-ring chronologies in the southwestern United States. Int J Climatol 41(5):2913–2933. https://doi.org/10.1002/joc.6997

Howard IM, Stahle DW, Dettinger MD, Poulsen C, Ralph FM, Torbenson MCA, Gershunov A (2023) A 440-year reconstruction of heavy precipitation in California from blue oak tree rings. J Hydrometeorol 24(3):463–477. https://doi.org/10.1175/jhm-d-22-0062.1

IPCC (2021) Climatic change 2021: The physical science basis. Contributions of working group i to the sixth assessment report of the intergovernmental panel on climate change. Masson-Delmotte, V., et al. (Eds.). Cambridge University Press, Cambridge, UK. https://doi.org/10.1017/9781009157896

Jafari M, Ansari-Pour N (2019) Why, when and how to adjust your P values? Cell J 20(4):604–607. https://doi.org/10.22074/cellj.2019.5992

Jevšenak J (2019) Daily climate data reveal stronger climate-growth relationships for an extended European tree-ring network. Quat Sci Rev 221:105868. https://doi.org/10.1016/j.quascirev.2019.105868

Jevšenak J (2020) New features in the dendroTools R package: bootstrapped and partial correlation coefficients for monthly and daily climate data. Dendrochronologia 63:125753. https://doi.org/10.1016/j.dendro.2020.125753

Jevšenak J, Levanič T (2018) dendroTools: R package for studying linear and nonlinear responses between tree-rings and daily environmental data. Dendrochronologia 48:32–39. https://doi.org/10.1016/j.dendro.2018.01.005

Karanitsch-Ackerl S, Mayer K, Gauster T, Laaha G, Holawe F, Wimmer R, Grabner M (2019) A 400-year reconstruction of spring–summer precipitation and summer low flow from regional tree-ring chronologies in North-Eastern Austria. J Hydrol 577:123986. https://doi.org/10.1016/j.jhydrol.2019.123986

Lee CC, Dannenberg MP (2023) Frequencies of multivariate air masses drive tree growth. JGR Biogeosciences 128(3):e2022JG007064. https://doi.org/10.1029/2022jg007064

Lin MF, Lucas HC Jr, Shmueli G (2013) Research commentary—too big to fail: large samples and the *p*-value problem. Inf Syst Res 24(4):906–917. https://doi.org/10.1287/isre.2013.0480

Meko DM, Baisan CH (2001) Pilot study of latewood-width of conifers as an indicator of variability of summer rainfall in the North American monsoon region. Int J Climatol 21(6):697–708. https://doi.org/10.1002/joc.646

Meko DM, Woodhouse CA (2005) Tree-ring footprint of joint hydrologic drought in Sacramento and Upper Colorado river basins, western USA. J Hydrol 308(1–4):196–213. https://doi.org/10.1016/j.jhydrol.2004.11.003

Meko DM, Touchan R, Anchukaitis KJ (2011) Seascorr: a MATLAB program for identifying the seasonal climate signal in an annual tree-ring time series. Comput Geosci 37(9):1234–1241. https://doi.org/10.1016/j.cageo.2011.01.013

Menne MJ, Williams CN Jr, Vose RS (2015) United States historical climatology network daily temperature, precipitation, and snow data. Carbon Dioxide Information Analysis Center. Oak Ridge National Laboratory, Oak Ridge, United States

Moran MD (2003) Arguments for rejecting the sequential Bonferroni in ecological studies. Oikos 100(2):403–405. https://doi.org/10.1034/j.1600-0706.2003.12010.x

Nakagawa S (2004) A farewell to Bonferroni: the problems of low statistical power and publication bias. Behav Ecol 15(6):1044–1045. https://doi.org/10.1093/beheco/arh107

Peres-Neto PR (1999) How many statistical tests are too many? The problem of conducting multiple ecological inferences revisited. Mar Ecol Prog Ser 176:303–306. https://doi.org/10.3354/meps176303

Pompa-García M, Camarero JJ, Colangelo M, González-Cásares M (2021) Inter and intra-annual links between climate, tree growth and NDVI: improving the resolution of drought proxies in conifer forests. Int J Biometeorol 65(12):2111–2121. https://doi.org/10.1007/s00484-021-02170-5

Rydval M, Gunnarson BE, Loader NJ, Cook ER, Druckenbrod DL, Wilson R (2017) Spatial reconstruction of Scottish summer temperatures from tree rings. Int J Climatol 37(3):1540–1556. https://doi.org/10.1002/joc.4796

Schweingruber F (1988) Tree rings: basics and applications of dendrochronology. Paul Haupt, Bern, Switzerland

Šidák Z (1967) Rectangular confidence regions for the means of multivariate normal distributions. J Am Stat Assoc 62(318):626–633. https://doi.org/10.1080/01621459.1967.10482935

Sokal RR, Rohlf FJ (1981) Biometry: the principles and practices of statistics in biological research, 2nd edn. W.H. Freeman and Company, New York, United States

St George S, Ault TR (2011) Is energetic decadal variability a stable feature of the central Pacific Coast's winter climate? J Geophys Res 116(D12):D12102. https://doi.org/10.1029/2010jd015325

St George S, Ault TR (2014) The imprint of climate within Northern Hemisphere trees. Quat Sci Rev 89:1–4. https://doi.org/10.1016/j.quascirev.2014.01.007

Stahle DW, Cook ER, Burnette DJ, Torbenson MCA, Howard IM, Griffin D, Diaz JV, Cook BI, Williams AP, Watson E, Sauchyn DJ, Pederson N, Woodhouse CA, Pederson GT, Meko D, Coulthard B, Crawford CJ (2020) Dynamics, variability, and change in seasonal precipitation reconstructions for North America. J Clim 33(8):3173–3195. https://doi.org/10.1175/jcli-d-19-0270.1

Tejedor E, Saz MA, Esper J, Cuadrat JM, de Luis M (2017) Summer drought reconstruction in northeastern Spain inferred from a tree ring latewood network since 1734. Geophys Res Lett 44(16):8492–8500. https://doi.org/10.1002/2017gl074748

Torbenson MCA, Stahle DW, Villanueva Díaz J, Cook ER, Griffin D (2016) The relationship between earlywood and latewood ring-growth across North America. Tree Ring Res 72(2):53–66. https://doi.org/10.3959/1536-1098-72.02.53

Torbenson MCA, Brázdil R, Stagge JH, Esper J, Büntgen U, Vizina A, Hanel M, Rakovec O, Fischer M, Urban O, Treml V, Reinig F, Martinez Del Castillo E, Rybníček M, Kolář T, Trnka M (2023a) Increasing volatility of reconstructed Morava River warm-season flow. Czech Republic J Hydrol Reg Stud 50:101534. https://doi.org/10.1016/j.ejrh.2023.101534

Torbenson MCA, Büntgen U, Esper J, Urban O, Balek J, Reinig F, Krusic PJ, Martinez del Castillo E, Brázdil R, Semerádová D, Štěpánek P, Pernicová N, Kolář T, Rybníček M, Koňasová E, Arbelaez J, Trnka M (2023b) Central European agroclimate over the past 2000 years. J Clim 36(13):4429–4441. https://doi.org/10.1175/jcli-d-22-0831.1

Touchan R, Anchukaitis KJ, Shishov VV, Sivrikaya F, Attieh J, Ketmen M, Stephan J, Mitsopoulos I, Christou A, Meko DM (2014) Spatial patterns of eastern Mediterranean climate influence on tree growth. Holocene 24(4):381–392. https://doi.org/10.1177/0959683613518594

Touchan R, Anchukaitis KJ, Meko DM, Kerchouche D, Slimani S, Ilmen R, Hasnaoui F, Guibal F, Julio Camarero J, Sánchez-Salguero R, Piermattei A, Sesbou A, Cook BI, Sabir M, Touchane H (2017) Climate controls on tree growth in the Western Mediterranean. Holocene 27(10):1429–1442. https://doi.org/10.1177/0959683617693901

Unterholzner L, Castagneri D, Cerrato R, Știrbu MI, Roibu CC, Carrer M (2024) Climate response of a glacial relict conifer across its distribution range is invariant in space but not in time. Sci Total Environ 906:167512. https://doi.org/10.1016/j.scitotenv.2023.167512

Urban O, Ač A, Kolář T, Rybníček M, Pernicová N, Koňasová E, Trnka M, Büntgen U (2021) The dendroclimatic value of oak stable isotopes. Dendrochronologia 65:125804. https://doi.org/10.1016/j.dendro.2020.125804

Vautard R, Yiou P, Ghil M (1992) Singular-spectrum analysis: a toolkit for short, noisy chaotic signals. Phys D Nonlinear Phenom 58(1–4):95–126. https://doi.org/10.1016/0167-2789(92)90103-T

Wasserstein RL, Lazar NA (2016) The ASA statement on *p*-values: context, process, and purpose. Am Stat 70(2):129–133. https://doi.org/10.1080/00031305.2016.1154108

Wise EK, Dannenberg MP (2019) Climate factors leading to asymmetric extreme capture in the tree-ring record. Geophys Res Lett 46(6):3408–3416. https://doi.org/10.1029/2019gl082295

Woodhouse CA, Meko D (1997) Number of winter precipitation days reconstructed from southwestern tree rings. J Climate 10(10):2663–2669. https://doi.org/10.1175/1520-0442(1997)010%3c2663:NOWPDR%3e2.0.CO

Woodhouse CA, Stahle DW, Villanueva Díaz J (2012) Rio Grande and Rio Conchos water supply variability over the past 500 years. Clim Res 51(2):147–158. https://doi.org/10.3354/cr01059

Yang B, He MH, Shishov V, Tychkov I, Vaganov E, Rossi S, Ljungqvist FC, Bräuning A, Grießinger J (2017) New perspective on spring vegetation phenology and global climate change based on Tibetan Plateau tree-ring data. Proc Natl Acad Sci USA 114(27):6966–6971. https://doi.org/10.1073/pnas.1616608114

Yuan YJ, Jin LY, Shao XM, He Q, Li ZZ, Li JF (2003) Variations of the spring precipitation day numbers reconstructed from tree rings in the Urumqi River drainage, Tianshan Mts. over the last 370 years. Chin Sci Bull 48(14):1507–1510. https://doi.org/10.1360/02wd0251

Zhao SD, Pederson N, D'Orangeville L, HilleRisLambers J, Boose E, Penone C, Bauer B, Jiang Y, Manzanedo R (2019) The international tree-ring data bank (ITRDB) revisited: data availability and global ecological representativity. J Biogeogr 46(2):355–368. https://doi.org/10.1111/jbi.13488

Ziaco E, Liang EY (2019) New perspectives on sub-seasonal xylem anatomical responses to climatic variability. Trees 33(4):973–975. https://doi.org/10.1007/s00468-018-1786-9

Zuur AF, Ieno EN, Elphick CS (2010) A protocol for data exploration to avoid common statistical problems. Meth Ecol Evol 1(1):3–14. https://doi.org/10.1111/j.2041-210x.2009.00001.x