Dendrochronologia 64 (2020) 125772

Contents lists available at ScienceDirect

Dendrochronologia

journal homepage: www.elsevier.com/locate/dendro

Technical Note

"dendRoAnalyst": A tool for processing and analysing dendrometer data

Sugam Aryal [a,*], Martin Häusser [a], Jussi Grießinger [a], Zexin Fan [b], Achim Bräuning [a]

[a] Friedrich-Alexander-Universität Erlangen-Nürnberg, Institut Für Geographie, Wetterkreuz 15, 91058 Erlangen, Germany
[b] Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China

ARTICLE INFO

Keywords:
Tree-growth monitoring
Stem-variation analysis
R-package
Dendrometer
Stem-cycle approach
Zero-growth approach

ABSTRACT

Dendrometers are vital tools for studying the response of trees to intra-annual environmental changes in different temporal resolutions, ranging from hourly, daily to weekly time resolution. Dendrometers are increasingly used in forest management and tree physiological studies. Besides the data analysis, data processing is also challenging, time-consuming and potentially error-prone due to the immense number of measurements generated by self-registering electronic dendrometers. We present the package 'dendRoAnalyst' based on R statistical software to process and analyse dendrometer data using various approaches. This package offers algorithms for handling and pre-cleaning of dendrometer data before the application of subsequent data analytical steps. This includes identifying and erasing artefacts in dendrometer datasets not related to actual stem circumference change, identifying data gaps within records, and the possibility of change(s) in temporal resolution. Furthermore, the package can calculate different daily statistics of dendrometer data, including the daily amplitude of tree growth. The package dendRoAnalyst is therefore intended to facilitate researchers with a collection of functions for handling and analysing dendrometer data.

1. Introduction

Dendrometers are instruments vital for measuring the growth of plant stems in high spatio-temporal resolution. The first mentioned use of dendrometers was during the 1920s (Grosenbaugh, 1963). Since then, there has been significant technological development in the engineering of the devices. During the mid-20th century, various studies used dendrometers in forestry and agronomy (Clark et al., 2000). With the technological evolution of automatic loggers for the recording of high-resolution temporal data, there has also been development in the theory of data analysis approaches. Currently, three major concepts to analyse the data exist: The "daily approach (DA)" (Herzog et al., 1995; King et al., 2013; van der Maaten et al., 2013, 2016), the "stem-cycle approach (SC) "(Deslauriers et al., 2007, 2011; Downes et al., 1999; van der Maaten et al., 2016), and the "zero-growth approach (ZG)" (Zweifel et al., 2005; 2006; 2016).

Tree stems undergo diurnal swelling and shrinking due to water uptake by the roots and water loss by crown transpiration. In the DA, daily maximum, minimum, and corresponding times are extracted along with the amplitude of the daily records (van der Maaten et al., 2013). The SC, on the other hand, divides the diurnal cycle into three phases known as contraction, expansion, and radial increment (Deslauriers et al., 2007, 2011; Downes et al., 1999). During the daytime, when tree water loss by transpiration exceeds water uptake and transport through the stem, the stem diameter decreases. This phase is called "contraction" (Deslauriers et al., 2003; van der Maaten et al., 2016) or "shrinkage "(Downes et al., 1999). Conversely, the phase when transpiration ceased and water uptake dominates, the diameter increases until reaches to the previous maximum is called "expansion" (van der Maaten et al., 2016) or "recovery" (Downes et al., 1999). If the tree's diameter increases further, surpassing the previous day's maximum, it is supposed that this represents real tree growth by formation of new xylem cells, and hence this phase has been termed "radial increment" phase (Deslauriers et al., 2011; Downes et al., 1999). Deslauriers et al. (2003) referred to the term "expansion" as a combination of recovery and increment. Later, van der Maaten et al. (2016) re-defined it as a phase when the tree diameter increases but remains below the previous maximum. According to the ZG approach, stem growth is divided into two phases. The tree water deficit (TWD) is the part of the phase that includes reversible shrinking and expansion of the stem due to the loss and uptake of water, and irreversible stem expansion (GRO), considered as radial growth (Zweifel et al., 2005, 2016). The ZG approach is based on the assumption that tree growth stops when trees undergo shrinkage (Zweifel et al., 2016). Building upon this, an enveloping curve will be produced joining all

* Corresponding author.
E-mail address: sugam.aryal@fau.de (S. Aryal).

https://doi.org/10.1016/j.dendro.2020.125772
Received 20 August 2020; Accepted 27 September 2020
Available online 6 October 2020
1125-7865/

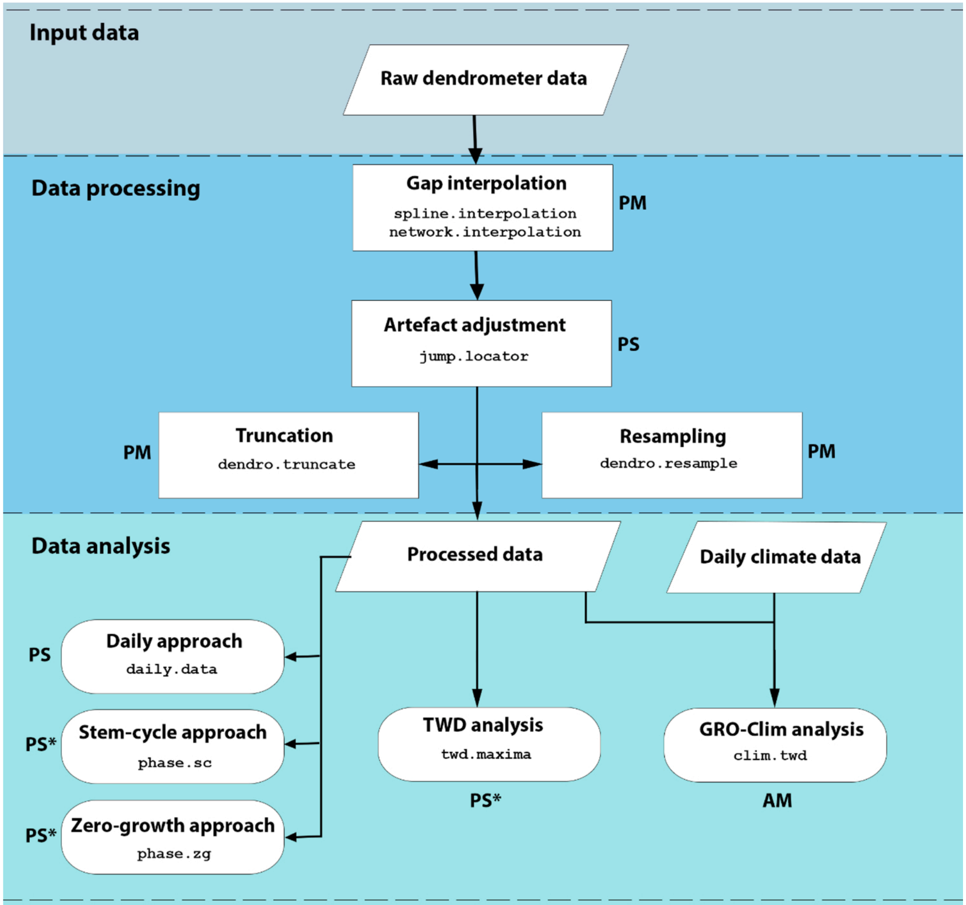

Fig. 1. Flow chart depicting the individual substeps of the package dendRoAnalyst. The functions are categorised in four groups. PM indicates the functions that can analyse perennial data containing multiple trees, PS indicates the functions suitable for perennial data containing a single tree, AM indicates the function that can handle annual data containing multiple trees and PS* indicates the functions that can be applied to perennial data for a single tree but unable to generate a graph for perennial data.

dendrometer data points, which exceed the previous maximum value (GRO-curve). Then, individual data points are subtracted from the GRO-curve to obtain the actual tree water deficit (TWD) values for individual data points, in a way that the TWD values for GRO phase are always zero (Zweifel et al., 2016).

Most dendrometer manufacturers have dedicated individual software solutions to extract the instrumental raw data from a data logger. However, these software packages are usually not capable of performing complex statistical analyses or very time-consuming. There had been previous efforts to develop computer programs to make data extraction and analysis efficient and easy. Grosenbaugh (1963) developed the basic Fortran computer program DD2 to extract measurements from optical dendrometers. At present, besides the data analysis, data processing is also challenging due to the enormous numbers of measurements generated by self-registering electronic dendrometers. To make data analysis more convenient and interactive, Deslauriers et al. (2011) published a collection of algorithms written in the SAS platform based on the stem-cycle approach. Later, a package called 'dendrometeR' (van der Maaten et al., 2016) based on R Software (R Development Core Team, 2020) was created, which uses both DS and SC approaches for dendrometer data analysis. In the 'dendrometeR' package, there are functions for calculating data resolution, identifying missing values in the data set, and filling them using the ARIMA model (van der Maaten et al., 2016). However, users cannot apply the ZG approach in this package. Meanwhile, due to the intensive use of automatic dendrometers with the capacity of recording high-resolution data, there is a demand for a tool that can not only perform analysis applying all prevalent approaches but is also able to pre-process and clean the raw dendrometer data. Here, we present a package called 'dendRoAnalyst' based on the R software (R Development Core Team, 2020) that provides an opportunity to use all existing analysis approaches and additionally includes new tools for data (pre)processing and analysis.

2. Package functionality

2.1. Dataset preparation

The presented package requires a dataset with the first column consisting of time in extended date-time format (e.g. yyyy-mm-dd HH:MM:SS) without daylight savings. Individual dendrometer data has to be sorted from the second column onward. The dataset may contain data for more than one year, but every single device needs its column. The package is flexible to column names but is strict with their order. If the format of the time is not proper, it will generate an automated error message. In total, this package offers ten different functions that are useful for dendrometer data processing and analysis (Fig. 1).

2.2. Data processing

High-resolution dendrometers often exhibit data gaps caused by mechanical or environmental effects. Since they are electronic devices that need a battery or continuous power supply for operation, they can easily be disturbed, e.g. due to humid weather conditions. Furthermore, animals can also harm cables or the power source. This causes interruptions in the recordings, and leads to the generation of gaps in the data. This package offers two different functions to deal with such gaps. The first function spline.interpolation not only detects data gaps, but also interpolates missing values using a cubic spline. This function is capable of handling data of several dendrometers in one dataset. It locates the starting time of a gap, based on a user-defined resolution of the

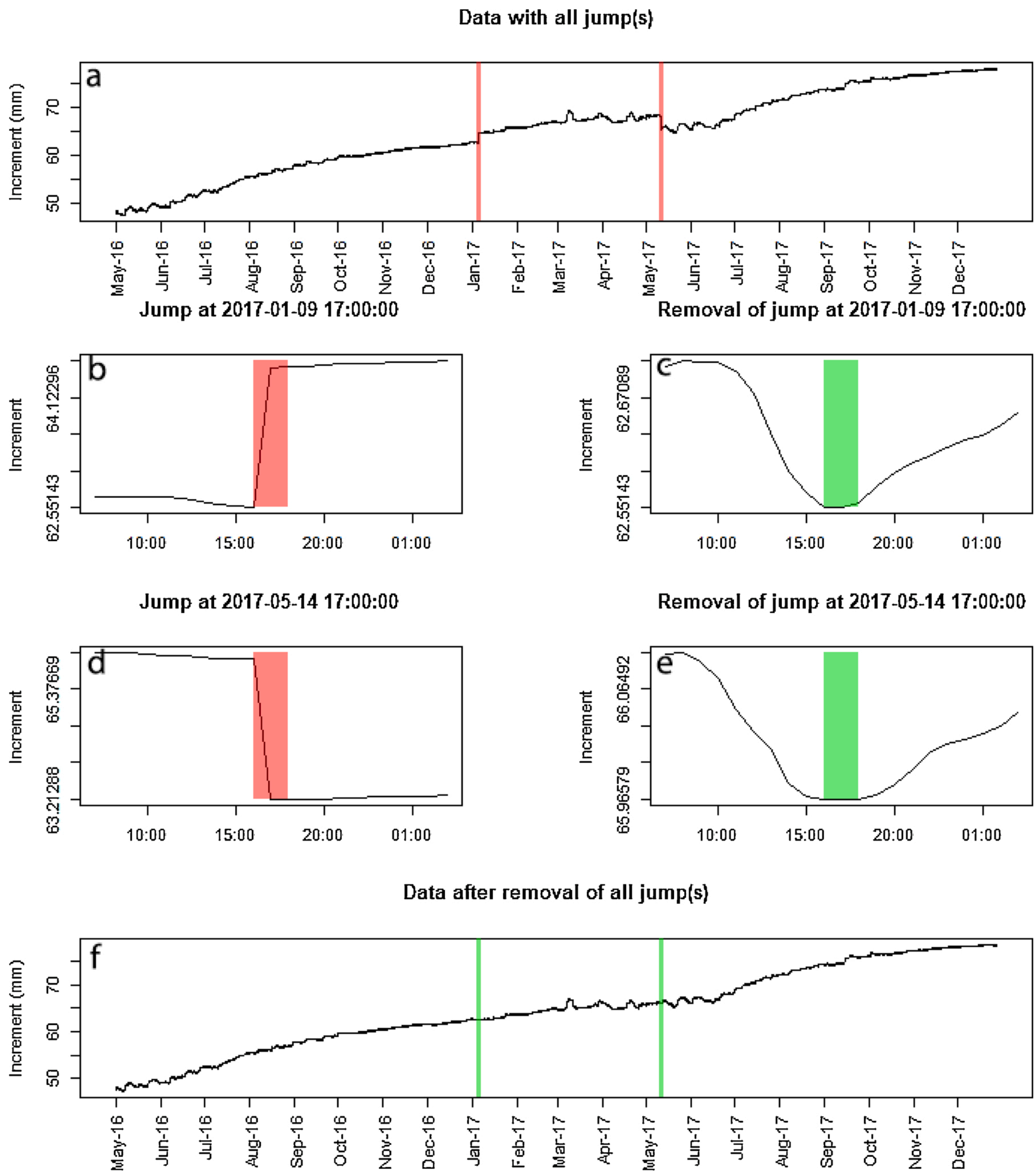

Fig. 2. Plots generated by the jump.locator function. (a) Original data with an indication of the located jumps (red bars). (b/d) The zoomed-in plot at the location of the individual jumps with the indication of the jump (red bar). (c/e) The zoomed-in plot after the removal of the jump with the indication of its location (green bar). (f) Data set after the removal of the jumps, the position of jump removal is indicated by green bars. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

data using the argument "resolution", and indicates the consecutive number of gaps for each position. If the parameter fill is false, it leaves the missing readings as 'NA', otherwise it fills all missing readings. spline.interpolation uses the whole dataset to predict missing readings and replaces the 'NA' values with the cubic spline interpolation method. The uncertainty of spline.interpolation increases with the number of consecutive gaps, so we recommend not using this method in a dataset with consecutive gaps lasting for more than 24 h. Generally, dendrometer data from the same site, and species often show similar stem circumference variations (Spannl et al., 2016). Based on this principle, a function network.interpolation is designed to fill the data gaps within one dendrometer dataset with the help of contemporaneous dendrometer data from the same site. This method of data interpolation is different from existing interpolation techniques included in previous software packages. The function included two options for network interpolation methods: 'linear' and 'proportional'. Users can select the desired methods using a parameter niMethod. When applying the 'linear' method, the missing value at a data point is computed using a linear model between previous data points and corresponding data points of the missing value in the reference dataset (Equation 1). If using the 'proportional' method, the average proportional change between previous data points and corresponding data points of a missing data point in the reference dataset provides the basis for predicting the missing value (Equation 1). To make predictions more robust, we included the bootstrap method in both interpolation methods. The 'proportional' method for example calculates the proportional change, and runs 500 iterations using the provided data. Subsequently, the mean proportional change is calculated and displayed for each missing location including the 95 % confidence level. By contrast, the 'linear' method calculates the r-square, intercept, and slope of the linear model by running 500 iterations and displays the predicted value with its corresponding r-squared value.

$$\mathrm{M_i} = \mathrm{M_{(i-1)}} \times m + c \tag{1a}$$

$$\mathrm{M_i} = \mathrm{M_{(i-1)}} \times \mathrm{M_{(i-1)}} \times p \tag{1b}$$

Where $\mathrm{M_i}$ is the missing value at the ith row in dataset M and $\mathrm{M_{(i\text{-}1)}}$ is a value at (i-1)th row in dataset M; m and c are the slope and intercept of the linear model between all the data at (i-1)th and ith row in the reference dataset; p is the average proportional change between all the data at (i-1)th and ith row in the reference dataset.

Occasionally, dendrometer installations have to be maintained, e.g. by exchanging exhausted logger batteries, or by adjusting dendrometer bands or needles after longer phases of stem growth. This may result in artefacts displayed as positive or negative jumps in the data. To automatically adjust them, the package includes the function jump.locator. It locates conscious jumps in the data and interactively adjusts them individually by producing a series of figures (Fig. 2). Although it is highly recommended not to use data with any gaps, the jump.locator function is capable to handle data with 'NA' values. It screens the whole dataset and locates their position, timing, and number based on the value of argument v that represents a user-defined threshold, such that fluctuations above it or below its negative are considered as a jump. The function was designed to first display a plot with the original data, indicating the location with red bars (Fig. 2). Then, it zooms to each position (Fig. 2) and asks for the user's input to confirm or to remove the jump. Subsequently, it displays a zoomed curve after each correction is applied (Fig. 2). Finally, the whole data after correction is displayed in a graph, indicating the corrected data locations by green bars (Fig. 2).

Dendrometers record a vast number of data every year. Processing and analysing continuously recorded data for many years is often complex and highly time-consuming. Truncated dendrometer data reduce the complexity and duration of calculations. Therefore, dendrometer data are generally analysed for yearly or seasonally truncated time intervals. Thus, the implemented function dendro.truncate provides an opportunity to truncate dendrometer data. The function requires a correctly formatted data frame df, the year for which data should be truncated CalYear, and an array containing two elements for the beginning and end of the period of truncation DOY. The function is capable to truncate data within and between years. If the user provides an array (with two numbers) instead of a single value for CalYear and a single value for DOY, it truncates data from the DOY of the first CalYear to the same DOY of the second CalYear. Conversely, if the user provides one value for CalYear and an array of two elements for DOY, it truncates the dataset from the first DOY to the second DOY within the same CalYear. Finally, if the user provides an array with two values for both DOY and CalYear, it truncates data from the first DOY of the first CalYear to the second DOY of second CalYear.

The temporal resolution of a data set may differ according to the objectives of the investigation. With the help of the function dendro.resample, users can change the resolution of their dataset based on a user-defined temporal frequency. For example, if resampled on a daily base, the function chooses only one value for each day. This function needs a data frame that contains the time in the first column and dendrometer recordings from the second column onwards. It can resample more than one dendrometer data collected for several years simultaneously. Based on the objective, the user can define either maximum, minimum, or mean values indicating either 'max', 'min', or 'mean' in the parameter value. Besides, indicating either 'H', 'D', 'W', or 'M' for the parameter by, users can resample data in hourly, daily, weekly or monthly frequency.

Table 1
Description of statistics generated by the daily.data function.

Name of columns	Description
DOY	Day of the year.
min	Minimum value for the corresponding day.
Time_min	Time when the minimum value was recorded for the corresponding day.
max	Maximum value for the corresponding day.
Time_max	Time when the maximum value was recorded for the corresponding day.
mean	Daily average value of the dendrometer reading.
median	Daily median value of the dendrometer reading.
amplitude	The difference between daily maximum and daily minimum.

2.3. Daily approach

With the function daily.data, our package offers a tool to calculate daily statistics of dendrometer data regardless of its temporal resolution. The provided daily statistics include daily maximum and minimum with their corresponding times, and daily amplitude (the difference between daily maximum and minimum). daily.data is capable of analysing datasets containing data for more than one year. The user can define the dendrometer data to be analysed by using the parameter TreeNum, which is an integer denoting the column number of dendrometer data excluding the first column (i.e., if dendrometer data is in the second column of data frame df, one should enter TreeNum = 1). This function can handle the data containing readings for more than 1 year. Furthermore, it can also compute similar statistics for climate data, if formatted accordingly. It furthermore generates a table that contains various daily statistics (Table 1).

2.4. Stem-cycle approach

The function phase.sc provides an algorithm to analyse data using the stem-cycle approach. For this, a data frame formatted following the methods mentioned above is required (see section 2.1) and TreeNum needs to be entered to determine which dendrometer data shall be analysed. The function includes another parameter called outputplot. If entering 'TRUE', it produces a plot with different phases based on the interval defined by argument days. The parameter days accepts an array with two elements. The elements denote the beginning and end of the plotting period in terms of the day of the year.

The smoothing of data is essential to ignore sudden fluctuations, which can be considered as noise. A higher smoothing value reduces the cycle amplitude and can affect the results of the analysis, whereas a lower smoothing value might not be able to remove the noise level (Deslauriers et al., 2011). The phase.sc function therefore further includes a parameter called smoothing, which defines the length of the spline in hours for the smoothing window of the dendrometer data. The function calculates the length of the spline using the resolution of the provided dataset and the input value for smoothing. For instance, if the resolution of the provided dataset is 60 min and the desired smoothing interval is five (i.e. 5 h =300 min), then it divides the smoothing value (in minutes) by the resolution of data (in minutes) to determine the spline window. The value of smoothing should, therefore, be chosen with care and be as low as possible. After smoothing the data, the function defines three phases: *Shrinkage*, when the dendrometer reading is less than the previous reading; *Expansion*, when the current reading is more than the previous reading, and *Increment*, when the current reading is higher than previous maxima. Mathematically, the method is expressed as the following equation (Eq. 2).

$$\begin{aligned} phase_{i+1} &= 1\ (Shrinkage)\quad (when\ x_{i+1} < x_i\ \&\ x_i < max_i) \\ phase_{i+1} &= 2\ (Expansion)\quad (when\ x_{i+1} > x_i\ \&\ x_i < max_i) \\ phase_{i+1} &= 3\ (Increment)\quad (when\ x_{i+1} < x_i\ \&\ x_i > max_i) \end{aligned} \tag{2}$$

Table 2
Description of statistics in 'SC_cycle', generated by the phase.sc function.

Name of columns	Description
Phase	Cyclic phases. 1, 2, and 3 for *Shrinkage*, *Expansion*, and *Increment* respectively.
start	Time when the corresponding phase starts.
end	Time when the corresponding phase ends.
Duration_h	Duration of the corresponding phase in hours.
Duration_m	Duration of the corresponding phase in minutes.
Magnitude	Radial/circumferential change during the corresponding phase in millimetres.
rate	Rate of radial/circumferential change during the corresponding phase in micrometres per hour.
DOY	Day of the year for the corresponding phase.

Table 3
Description of output statistics in 'ZG_cycle', generated by the phase.zg function.

Name of columns	Description
DOY	Day of the year for the corresponding phase.
Phase	TWD for tree water deficit and GRO for irreversible expansion.
start	Time when the corresponding phase starts.
end	Time when the corresponding phase ends.
Duration_h	Duration of the corresponding phase in hours.
Magnitude	Radial/circumferential change during the corresponding 'GRO' phase in millimetres.
rate	Rate of radial/circumferential change during the corresponding 'GRO' phase measured in micrometres per hour.
Max.twd	Maximum TWD recorded for the corresponding TWD phase.
Max.twd.time	Time of occurrence of the maximum TWD value for each corresponding TWD phase.
Avg.twd	Average of TWD values for each TWD period.
STD.twd	Standard deviation of TWD values for each TWD period.

Table 4
Exemplary data set (the circumference in mm) nepa17 consisting of two trees with data gap between 2017-08-26 18:00:00 and 2017-08-26 23:00:00.

Time	T2	T3
2017−08-26 14:00:00	73.7598	58.07983
2017−08-26 15:00:00	73.7356	58.06337
2017−08-26 16:00:00	73.72882	58.05272
2017−08-26 17:00:00	73.72978	58.05127
2017−08-26 18:00:00	73.7327	58.05272
2017−08-26 23:00:00	73.76077	58.07402
2017−08-27 00:00:00	73.76561	58.07935
2017−08-27 01:00:00	73.77142	58.08177

Where x_i is a cumulative reading of dendrometer data at data point i; x_{i+1} is the cumulative reading of dendrometer data at data point i+1; and max_i is the maximum dendrometer reading between 1st to ith dendrometer data.

phase.sc can also analyse data for two consecutive years. The only precondition to use such a dataset is that the data frame must not contain the same days of two calendar years. For example, the dataset can contain data from 2017−06-01 to 2018−05-31, but should not contain data for further than this in 2018. The function is able to analyse a perennial dataset without the mentioned precondition. However, it only returns the analysed dataset for the whole period without generating any plot. It is necessary to define a period of plotting using days argument providing an initial and final day of the year. In the case of perennial data, there is a repetition of days of the year in every calendar year, which makes the plotting period arbitrary to identify. This function generates a list of two datasets. The first data frame 'SC_cycle' contains cyclic phases along with the beginning, end, duration, magnitude, and rate of each phase (see Table 2). The second data frame 'SC_phase' contains assigned phases for each data point.

2.5. Zero-growth approach

The package offers a function called phase.zg for analysing data using the zero-growth method. The format of the dataset should be the same as for phase.sc. First, phase.zg divides the data into two categories: Tree water deficiency (TWD), the reversible shrinkage and expansion of the tree stem when the current reading is less than previous maxima, and Increment (GRO), the irreversible expansion of the stem when the current reading is higher than previous maxima. Mathematically, Eq. 3 represents the method. Secondly, it applies a linear interpolation between each consecutive GRO period and forms a growth curve (GRO-curve). Then it determines the TWD value of each data point by subtracting the actual dendrometer reading from the corresponding value of the modelled data of the GRO-curve so that TWD value for the GRO phase is zero.

$$\begin{aligned} phase_{i+1} &= TWD \quad (when\ x_{i+1} < max_i) \\ phase_{i+1} &= GRO \quad (when\ x_{i+1} > max_i) \end{aligned} \tag{3}$$

Where x_i is the cumulative reading of dendrometer data at data point I; x_{i+1} is the cumulative reading of dendrometer data at data point i+1; max_i is the maximum dendrometer reading between 1st to ith dendrometer data.

Like phase.sc, the zero-growth function phase.zg is capable to analyse data with two consecutive calendar years and generates a list of two datasets. The first data frame 'ZG_cycle', contains cyclic phases along with the beginning, ending, duration, magnitude, and rate of each GRO phase along with maximum TWD value and its time of occurrence (see Table 3). The second data frame 'ZG_phase' contains the TWD value for each data point.

2.6. Additional functions for growth-climate analysis

We further included two additional functions that can investigate the growth of trees during extended climate events. For instance, trees experience shrinkage of circumference during extended dry periods. However, the severity of the impact may not be equal for all trees growing at the same site. Trees can reveal different responses to the same climate event, based on their functional type and age (Raffelsbauer et al., 2019). The two functions twd.maxima and clim.twd help to zoom to a particular climate event and assess the circumferential/radial change of the trees. The function clim.twd requires two data sets: a dendrometer data frame df and daily climate data Clim as an input. The user can define events by setting criteria with the parameters climThreshold and daysThreshold. The parameter climThreshold defines the threshold for climate below which is considered as an adverse climate.

Table 5
Data set after the application of gap finding using the function spline. interpolation.

Time	with `fill=FALSE`		with `fill=TRUE`	
	T2	T3	T2	T3
2017-08-26 15:00:00	73.7356	58.06337	73.7356	58.06337
2017-08-26 16:00:00	73.72882	58.05272	73.72882	58.05272
2017-08-26 17:00:00	73.72978	58.05127	73.72978	58.05127
2017-08-26 18:00:00	73.7327	58.05272	73.7327	58.05272
2017-08-26 19:00:00	NA	NA	73.7373	58.05525
2017-08-26 20:00:00	NA	NA	73.74305	58.05879
2017-08-26 21:00:00	NA	NA	73.74932	58.06318
2017-08-26 22:00:00	NA	NA	73.75544	58.06831
2017-08-26 23:00:00	73.76077	58.07402	73.76077	58.07402
2017-08-27 00:00:00	73.76561	58.07935	73.76561	58.07935
2017-08-27 01:00:00	73.77142	58.08177	73.77142	58.08177
2017-08-27 02:00:00	73.77433	58.08371	73.77433	58.08371

Table 6
Comparison of interpolated data by network.interpolation using the 'proportional' and 'linear' method. The shaded cells in columns 3 and 4 represent the interpolated value for each missing data using 'proportional' and 'linear' methods respectively.

Time	Original	niMethod='proportional'	niMethod='linear'
	T2	T2	T2
2017-01-02 10:00:00	62.53088	62.53088	62.53088
2017-01-02 11:00:00	62.51878	62.51878	62.51878
2017-01-02 12:00:00	62.49408	62.49408	62.49408
2017-01-02 13:00:00	62.47133	62.47133	62.47133
2017-01-02 14:00:00	62.46697	62.46697	62.46697
2017-01-02 15:00:00	NA	**62.46954**	**62.45247**
2017-01-02 16:00:00	NA	**62.4695**	**62.45006**
2017-01-02 17:00:00	NA	**62.46964**	**62.45102**
2017-01-02 18:00:00	NA	**62.46846**	**62.45151**
2017-01-02 19:00:00	NA	**62.47075**	**62.45005**
2017-01-02 20:00:00	NA	**62.4728**	**62.44908**
2017-01-02 21:00:00	NA	**62.47422**	**62.44812**
2017-01-02 22:00:00	NA	**62.47257**	**62.45391**
2017-01-02 23:00:00	NA	**62.48702**	**62.43506**
2017-01-03 00:00:00	NA	**62.49986**	**62.41387**
2017-01-03 01:00:00	NA	**62.51145**	**62.39274**
2017-01-03 02:00:00	62.55606	62.55606	62.55606
2017-01-03 03:00:00	62.56622	62.56622	62.56622
2017-01-03 04:00:00	62.57833	62.57833	62.57833
2017-01-03 05:00:00	62.60205	62.60205	62.60205
2017-01-03 06:00:00	62.61319	62.61319	62.61319

Table 7
Output of daily.data containing the daily statistics of dendrometer data.

DOY	min	Time_Minimum	max	Time_Maximum	mean	median	Amplitude
1	62.266	00:00:00	62.469	07:30:00	62.430	62.454	0.2034
2	62.464	13:30:00	62.543	07:30:00	62.499	62.487	0.0794
3	62.528	00:00:00	62.992	14:00:00	62.725	62.625	0.4633
4	62.771	16:00:00	63.012	08:00:00	62.899	62.941	0.2411
5	62.794	14:30:00	62.943	08:30:00	62.852	62.831	0.1491
6	62.731	16:30:00	62.933	08:30:00	62.837	62.843	0.2014
7	62.791	14:00:00	62.900	08:00:00	62.839	62.829	0.1085
8	62.712	16:30:00	62.913	08:00:00	62.817	62.835	0.2014
9	62.730	16:30:00	62.897	08:00:00	62.815	62.806	0.1666
10	62.714	15:30:00	62.886	08:00:00	62.804	62.805	0.1724

The parameter daysThreshold defines the minimum duration (days) such that the period extending longer than this duration is considered as adverse periods. The clim.twd function first generates a figure and a table showing dry periods with newly created IDs. Users can choose one or more periods inserting respective ID numbers in the console. Finally, it simultaneously generates a figure and a data frame with all trees and their relative growth change. If we desire to plot two or more periods then it chooses different colours for the different IDs and different 'pch' (point signature) for different trees. It has to be stated, that clim.twd is only useful when daily climate data is available. However, daily climate data may not always available due to various reasons. In such cases, the second function, twd.maxima is a useful help/tool. Within twd.maxima, an algorithm defines the TWD periods and calculates the TWD value for all trough maxima within each TWD period incorporating the functions phase.sc and phase.zg. It generates a data frame with each TWD period with the exact time for each trough maximum and time difference of its occurrence from the beginning of each TWD period. Like in phase.sc, the user can modify the smoothing parameter to adjust the identification of trough maxima.

3. Illustrated example

For the demonstration of this package, we used a dataset of Chir pine (*Pinus roxburghii*) from Kathmandu, Nepal. The data was derived using an automatic logging band dendrometer (DRL26, EMS Brno) recording hourly circumference change from 2016 to 2017. For reference and comprehensibility, we included this dataset as a default called 'nepa' in the package. Following, we will demonstrate all the functions of the package based on this dataset in sequential order (Fig. 1).

Before starting data processing, the user should ensure that the format of the dataset fits within the prerequisites. The first step is to identify gaps in the dataset using either spline.interpolation, or network.interpolation. In our sample dataset (Table 4), a data gap is apparent where the dendrometer skipped four recordings between 2017−08-26 18:00:00 and 2017−08-26 23:00:00. While using the spline.interpolation function with the argument fill = FALSE, additional rows with missing times (shaded cells of first column of Table 5) are automatically generated and missing readings are filled with NA (shaded cells of columns 2 and 3 of Table 5). However, with fill = TRUE, the missing readings are filled applying the spline interpolation method (shaded cells of columns 4 and 5 of Table 5).

```
> data("nepa17")
```

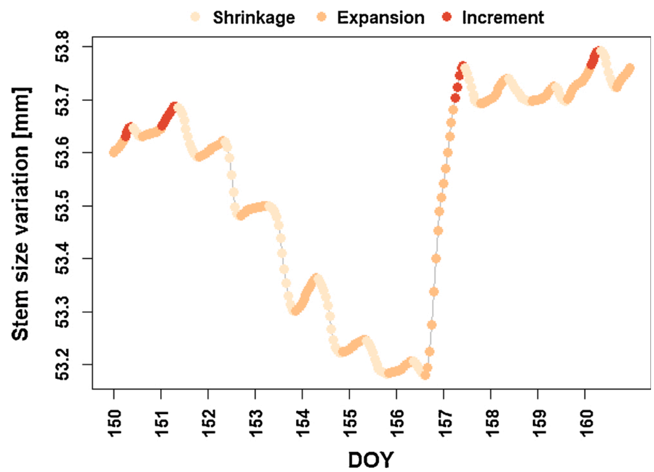

Fig. 3. Plot generated by the phase.sc function, which shows stem size variation for 10 days along with different phases of the dendrometer data (T2) in gf_nepa17.

> library(dendRoAnalyst)

> gf_nepa17 < -spline.interpolation(df = nepa17, resolution = 60, fill = FALSE)

To use the network.interpolation function, the respective dataset should be treated with spline.interpolation including fill = FALSE. For instance, we replaced in the sample dataset some reading in the data frame 'gf_nepa17′ (Table 6) for the demonstration and assigned a new name 'df1′. We also created a reference dataset df2, repeating the second and third columns of dataset 'gf_nepa17′ two times. We considered df1 as an input and df2 as a reference dataset.

data("gf_nepa17″)

> df1 < -gf_nepa17

> # Creating an example reference dataset.

> df2 < -cbind(gf_nepa17,gf_nepa17[,2:3],gf_nepa17[,2:3])

> # Creating gaps in dataset by replacing some of the reading with NA in dataset.

> df1[40:50,2]<-NA

> # Using proportional interpolation method.

> df1_NI<-network.interpolation(df = df1, referenceDF = df2, niMethod='proportional')

Table 8
Cyclic statistics using the stem-cycle approach. Data generated by the phase.sc function for T2 of gf_nepa17.

Phases	start	end	Duration_m	Duration_h	magnitude	rate	DOY
1	2017−01-01 01:00:00	2017−01-01 04:00:00	180	3.000	−0.005	−1.63144	1
2	2017−01-01 04:00:00	2017−01-01 05:00:00	60	1.000	0.001	1.11807	1
3	2017−01-01 05:00:00	2017−01-01 10:00:00	300	5.000	0.041	8.269678	1
1	2017−01-01 10:00:00	2017−01-01 13:00:00	180	3.000	−0.017	−5.6303	1
2	2017−01-01 13:00:00	2017−01-01 20:00:00	420	7.000	0.016	2.319211	1
3	2017−01-01 20:00:00	2017−01-02 10:00:00	840	14.000	0.131	9.390499	1
1	2017−01-02 10:00:00	2017−01-02 18:00:00	480	8.000	−0.032	−3.99042	2
2	2017−01-02 18:00:00	2017−01-03 05:00:00	660	11.000	0.029	2.643204	2
3	2017−01-03 05:00:00	2017−01-03 10:00:00	300	5.000	0.016	3.219348	3
1	2017−01-03 10:00:00	2017−01-03 13:00:00	180	3.000	−0.010	−3.48021	3

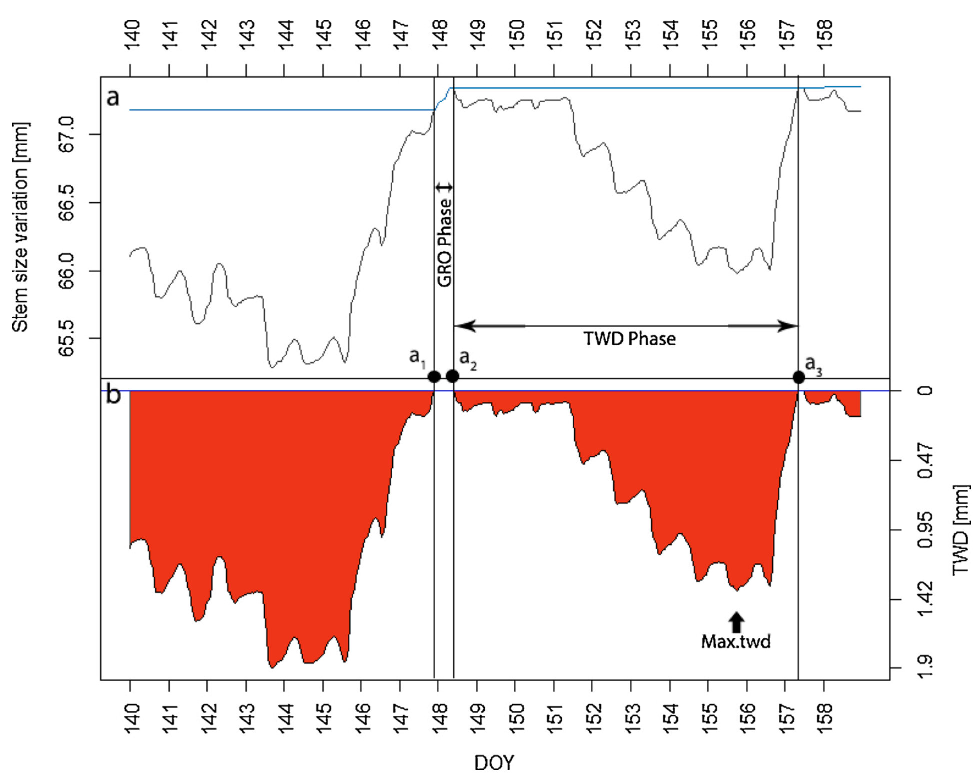

Fig. 4. Plot generated by the phase.zg function for T2 of gf_nepa17: (a) stem size variation (black line) and GRO curve (blue line) for 18 days along with TWD for each data point for that period. (b) GRO phase (from a_1 to a_2), TWD phase (from a_2 to a_3), and the maximum TWD value (Max.twd) for the TWD phase. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Table 9
Cyclic statistics for the zero-growth approach as generated by the phase.zg function for T2 of the default dataset gf_nepa17.

DOY	Phases	start	end	Duration_h	magnitude	rate	Max.twd	Max.twd.time	Avg.twd	STD.twd
1	GRO	2017−01-01 01:00:00	2017−01-01 09:00:00	8	0.205	25.600	NA	NA	NA	NA
2	TWD	2017−01-01 09:00:00	2017−01-02 00:00:00	15	NA	NA	0.022	2017−01-01 13:00:00	0.011	0.007
2	GRO	2017−01-02 00:00:00	2017−01-02 09:00:00	9	0.078	8.715	NA	NA	NA	NA
3	TWD	2017−01-02 09:00:00	2017−01-03 02:00:00	17	NA	NA	0.080	2017−01-02 15:00:00	0.047	0.028
3	GRO	2017−01-03 02:00:00	2017−01-03 09:00:00	7	0.084	11.966	NA	NA	NA	NA
3	TWD	2017−01-03 09:00:00	2017−01-03 14:00:00	5	NA	NA	0.109	2017−01-03 14:00:00	0.034	0.041
3	GRO	2017−01-03 14:00:00	2017−01-03 16:00:00	2	0.207	103.476	NA	NA	NA	NA
4	TWD	2017−01-03 16:00:00	2017−01-04 02:00:00	10	NA	NA	0.066	2017−01-03 18:00:00	0.033	0.022
4	GRO	2017−01-04 02:00:00	2017−01-04 09:00:00	7	0.067	9.545	NA	NA	NA	NA
19	TWD	2017−01-04 09:00:00	2017−01-19 07:00:00	358	NA	NA	0.811	2017−01-18 19:00:00	0.434	0.171

> # Using linear interpolation method.

> df1_NI<-network.interpolation(df = df1, referenceDF = df2, niMethod='linear')

The consecutive step is then applied to locate jumps in the dataset, which might not be related to environmental influences. For example, there are two jumps in the dataset 'nepa'. The first jump is positive (more than the provided value for v, i.e. 1 mm) whereas the second is negative (less than negative of provided value for v, i.e. 1 mm).

> jump_free_nepa<-jump.locator(df = nepa, TreeNum = 1, v = 1)

jump.locator produces a series of graphs (Fig. 2). The first graph shows all the jumps in the data indicated by red bars (Fig. 2). The small panels of Fig. 2 show the individual jumps before and after the adjustment (Fig. 2). Fig. 2f displays the sample data after the adjustment of all jumps with the indication of the location of jumps by green bars. (=jump-free series).

The dataset 'nepa' consists of data for two years: 2016 and 2017. Nevertheless, we require data only for 2017 from the day of year 1–365 for further analysis. The output of jump.locator can be used as the input file for this function.

> nepat2.2017=dendro.truncate(df = ju.gp.nepat2, CalYear = 2017, DOY = c(1365))

After performing all the mentioned steps above, the dataset is now suitable for the analysis of the different approaches (DA, SC, and ZG). The function daily.data can calculate the daily statistics of the dendrometer data. To use this function, the parameters TreeNum needs to be provided along with the input dataset that denotes the number of the column whose daily statistics are to be calculated.

> data ("gf_nepa17")

> daily17 < -daily.data(df = gf_nepa17, CalYear = 2017, TreeNum = 1)

This results in an output data frame illustrating various daily statistics of the dendrometer data (Table 7).

For a demonstration of the phase.sc function, we insert gap-filled dendrometer data of a test-site in Kathmandu for the year 2017 as follows:

> sc.phase<-phase.sc(df = gf_nepa17, TreeNum = 1, smoothing = 4, outputplot = TRUE, days = c(150,160))

This command generates a list containing two data frames and a plot showing different phases from the day of year 150–160 (Fig. 3). The first data frame 'SC_cycle' contains cyclic phases along with various statistics and the second data frame 'SC_phase' contains assigned phases for each data point. The data frame 'SC_cycle' contains the beginning, ending, duration, magnitude, and rate of each phase (Table 8).

Similarly, the data 'gf_nepa17′ can also be applied in the function phase.zg to analyse the data using ZG approach.

> zg.phase<-phase.zg(df = gf_nepa17, TreeNum = 1, outputplot = TRUE, days = c(140,158))

This function also generates a list containing two data frames and a

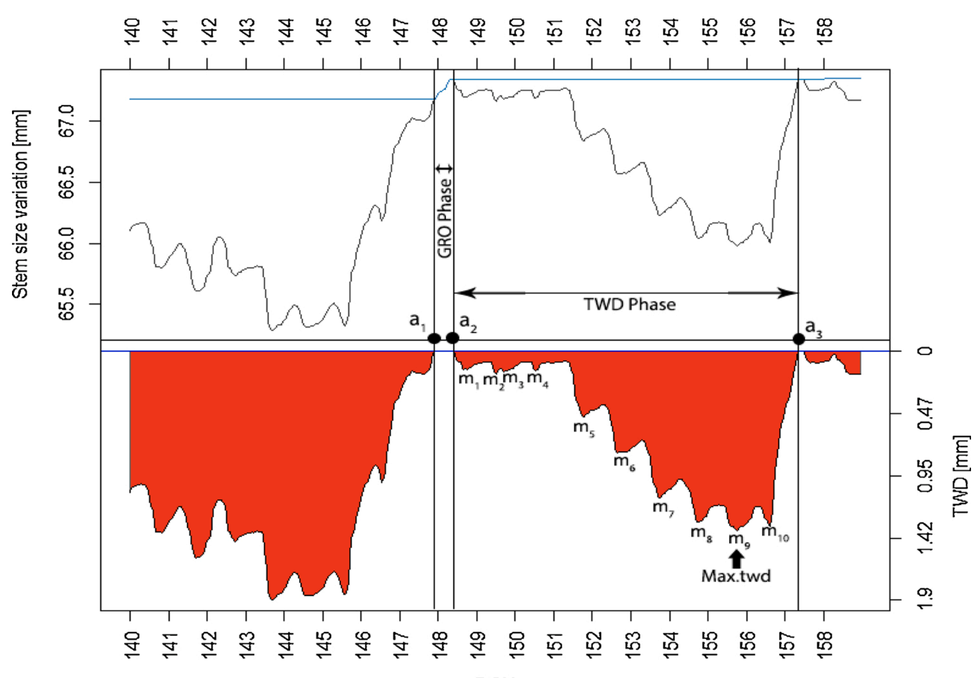

Fig. 5. Illustration of the output plot of the function twd.maxima for T2 of the default dataset gf_nepa17. It shows the TWD phase (from a$_2$ to a$_3$) and all the local maxima (m$_1$, m$_2$, …., m$_{10}$) within it.

Table 10
Example of the output table of the function twd.maxima for T2 of the default dataset gf_nepa17.

start.time	end.time	Time	TWD	duration_from_start	twd.number
2017−01-01 08:00:00	2017−01-01 23:00:00	2017−01-01 12:00:00	0.02239	180	1
2017−01-02 08:00:00	2017−01-03 01:00:00	2017−01-02 14:00:00	0.080369	300	2
2017−01-03 08:00:00	2017−01-03 13:00:00	2017−01-03 12:00:00	0.070597	180	3
2017−01-03 15:00:00	2017−01-04 01:00:00	2017−01-03 17:00:00	0.066112	60	4
2017−01-04 08:00:00	2017−01-19 06:00:00	2017−01-04 17:00:00	0.250679	480	5
2017−01-04 08:00:00	2017−01-19 06:00:00	2017−01-05 15:00:00	0.261033	1800	5
2017−01-04 08:00:00	2017−01-19 06:00:00	2017−01-06 17:00:00	0.36139	3360	5
2017−01-04 08:00:00	2017−01-19 06:00:00	2017−01-07 15:00:00	0.329621	4680	5
2017−01-04 08:00:00	2017−01-19 06:00:00	2017−01-08 17:00:00	0.44983	6240	5
2017−01-04 08:00:00	2017−01-19 06:00:00	2017−01-09 17:00:00	0.466453	7680	5

plot showing different phases from the day of year 140–158 (Fig. 4). The first data frame 'ZG_cycle' contains cyclic phases along with various statistics and the second data frame 'ZG_phase' contains assigned phases for each data point. The data frame 'ZG_cycle' contains the beginning, ending, duration, magnitude, and rate of each phase (Table 9).

The function twd.maxima uses both phase.sc and phase.zg functions. It first identifies all the TWD phases for the provided data. Then it identifies all the local maxima within each TWD phase and calculates the TWD values for each local maximum.

> twd_max<-twd.maxima(df = gf_nepa17, TreeNum = 1, days = c(148,158))

This command produces a data frame and a plot (Fig. 5) with the TWD phases and the local maxima contained within them. The output data frame (Table 10) contains the beginning (start.time) and end (end.time) of each TWD phase (a_2 and a_3 in Fig. 5), the time of occurrence (Time) and the corresponding TWD values (TWD) of each local maximum (m_1 to m_{10} in Fig. 5) within each TWD phase. It also contains the difference of time (duration_from_start) for each local maximum from the beginning of each TWD phase (a_2 in Fig. 5).

Furthermore, the function clim.twd requires two datasets as input. A dataset df containing dendrometer data with one or more trees and Clim containing a daily climate parameter. In this example, we take the gap filled dendrometer data ('gf_nepa17′) and daily rainfall data ('ktm_rain17′) of Kathmandu for 2017. The first column of Clim must be the day of the year as integers not as long time format as in the dendrometer data.

> data(gf_nepa17)

> data(ktm_rain17)

> relative_dry_growth<-clim.twd(df = gf_nepa17,Clim = ktm_rain17, dailyValue='max', climThreshold = 0, daysThreshold = 2,

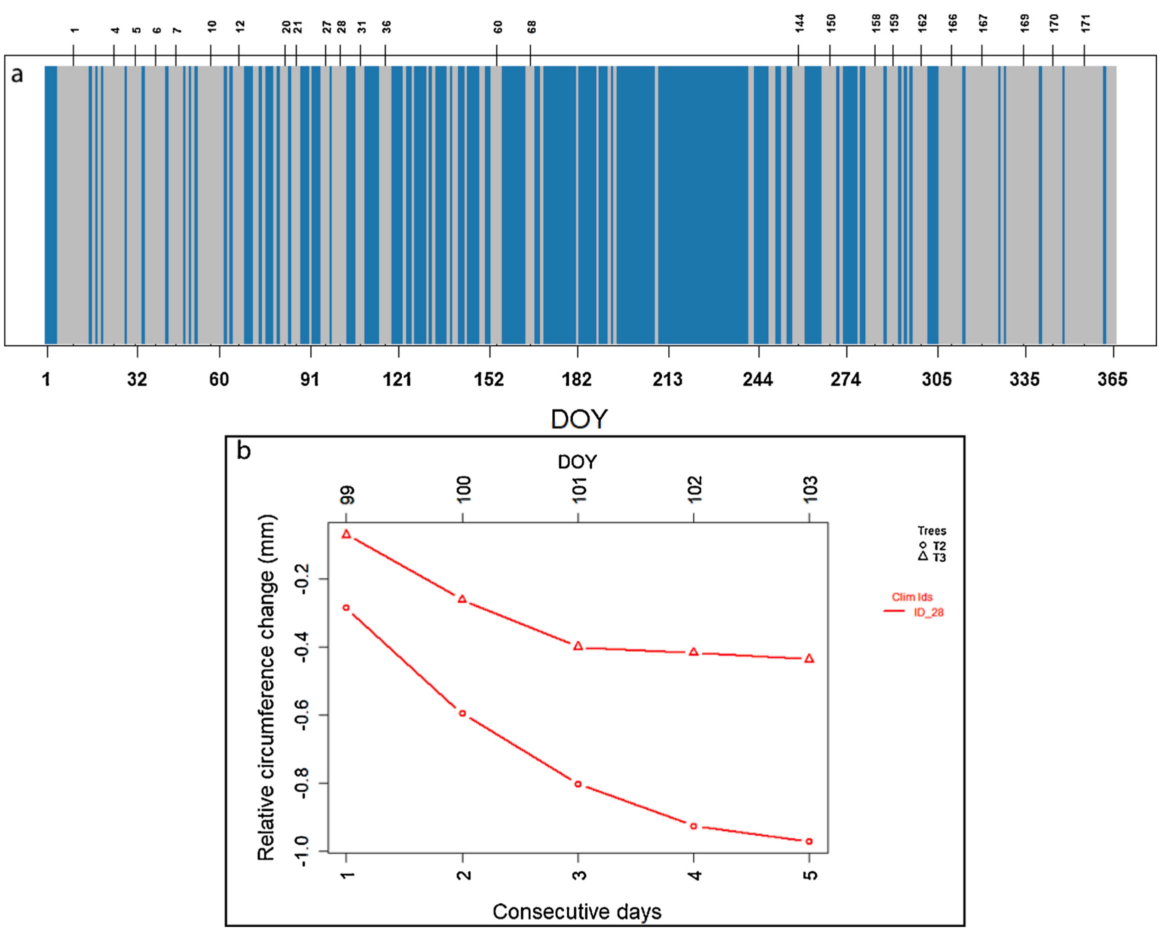

Fig. 6. Output plots of the function clim.twd. The upper plot (a) shows the no-rain periods in grey colour with assigned IDs along the upper x-axis and the lower plot (b) shows the relative circumferential change of the trees T2 and T3 for the no-rain period ID 28.

showPlot = TRUE)

clim.twd generates a plot (Fig. 6) and a table, showing all the periods that are included by climThreshold and daysThreshold with a unique ID assigned to each of them. Then users can plot (Fig. 6) the relative circumferential/radial change for all the dendrometers setting the ID of the desired periods in the console. Finally, it returns a table containing a beginning, end and, ID of each period along with the number of days for which the periods exist and the circumferential/radial change of all the trees for the corresponding days. In this case, the function showed a relative circumferential change of T2 and T3 for the period 28, which extends from the day of the year 99–103 (Fig. 6).

4. Availability

The package is freely accessible from the CRAN-Server via https://cran.r-project.org/package=dendRoAnalyst. It can be downloaded directly in the R console by using install.packages("dendRoAnalyst"). 'dendRoAnalyst' requires the packages pspline (Ripley, 2015) and zoo (Zeileis and Grothendieck, 2005).

5. Perspectives

The package initially contains various functions that are crucial not only for data analysis but also for data processing and management. It offers an opportunity for users to use all major approaches in the current state of the art of dendrometer data analysis. The functions of this package are simple to use with minimum input from the users. As this field of research is rapidly growing with the implementation of new methodologies, we look forward to suggestions from the scientific community to improve the package with the integration of more methods for further analyses.

Funding

This research was supported by the project "YunForest: Growth and physiological responses of forests to climate change along latitudinal and altitudinal gradients in Southwest China (Yunnan Province)" funded by Deutsche Forschungsgemeinschaft (DFG; project numbers: BR1895/31-1 and GR3799/4-1. MH was supported by the project "Climatic impact on functional wood anatomy, seasonal wood formation, and tree growth along altitudinal gradients on Corsica (BR 1895/27-1)". AB further benefitted from data generated during project BR 1895/29-1 "Tree physiological and structural properties as response and effect traits of biotic-atmospheric interactions in natural and anthropogenic eco-systems in South Ecuador". Data were collected during the project NEPALRING supported by the Federal Ministry of Education and Research (BMBF; FKZ: 01DP15009).

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgments

We express our sincere gratitude to the Department of Hydrology and Meteorology (DHM) of Nepal for providing the example meteorological data.

References

Clark, N.A., Wynne, R.H., Schmoldt, D.L., 2000. A review of past research on dendrometers. For. Sci. 46, 570–576.

Deslauriers, A., Morin, H., Urbinati, C., Carrer, M., 2003. Daily weather response of balsam fir (*Abies balsamea* (L.) Mill.) stem radius increment from dendrometer analysis in the boreal forests of Quec (Canada). Trees 17, 477–484. https://doi.org/10.1007/s00468-003-0260-4.

Deslauriers, A., Anfodillo, T., Rossi, S., Carraro, V., 2007. Using simple causal modeling to understand how water and temperature affect daily stem radial variation in trees. Tree Physiol. 27, 1125–1136. https://doi.org/10.1093/treephys/27.8.1125.

Deslauriers, A., Rossi, S., Turcotte, A., Morin, H., Krause, C., 2011. A three-step procedure in SAS to analyze the time series from automatic dendrometers. Dendrochronologia 29, 151–161. https://doi.org/10.1016/j.dendro.2011.01.008.

Downes, G., Beadle, C., Worledge, D., 1999. Daily stem growth patterns in irrigated *Eucalyptus globulus* and *E. nitens* in relation to climate. Trees 14, 102–111. https://doi.org/10.1007/PL00009752.

Grosenbaugh, L.R., 1963. Optical dendrometers for out-of-reach diameters: a conspectus and some new theory. For. Sci. 9, a0001–47. https://doi.org/10.1093/forestscience/9.s1.a0001.

Herzog, K.M., Häsler, R., Thum, R., 1995. Diurnal changes in the radius of a subalpine Norway spruce stem: their relation to the sap flow and their use to estimate transpiration. Trees 10, 94–101. https://doi.org/10.1007/BF00192189.

King, G., Fonti, P., Nievergelt, D., Büntgen, U., Frank, D., 2013. Climatic drivers of hourly to yearly tree radius variations along a 6°C natural warming gradient. Agric. For. Meteorol. 168, 36–46. https://doi.org/10.1016/j.agrformet.2012.08.002.

R Development Core Team, 2020. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/.

Raffelsbauer, V., Spannl, S., Peña, K., Pucha-Cofrep, D., Steppe, K., Bräuning, A., 2019. Tree circumference changes and species-specific growth recovery after extreme dry events in a montane rainforest in southern Ecuador. Front. Plant Sci. 10, 342. https://doi.org/10.3389/fpls.2019.00342.

Ripley, B., 2015. Pspline: Penalized Smoothing Splines. R Package Version 1. 0-17. R Port of S Original by Jim Ramsey. https://CRAN.R-project.org.

Spannl, S., Volland, F., Pucha, D., Peters, T., Cueva, E., Bräuning, A., 2016. Climate variability, tree increment patterns and ENSO-related carbon sequestration reduction of the tropical dry forest species Loxopterygium huasango of Southern Ecuador. Trees 30, 1245–1258. https://doi.org/10.1007/s00468-016-1362-0.

van der Maaten, E., Bouriaud, O., van der Maaten-Theunissen, M., Mayer, H., Spiecker, H., 2013. Meteorological forcing of day-to-day stem radius variations of beech is highly synchronic on opposing aspects of a valley. Agric. For. Meteorol. 181, 85–93. https://doi.org/10.1016/j.agrformet.2013.07.009.

van der Maaten, E., van der Maaten-Theunissen, M., Smiljanić, M., Rossi, S., Simard, S., Wilmking, M., Deslauriers, A., Fonti, P., Gv, Arx, Bouriaud, O., 2016. dendrometeR: Analyzing the pulse of trees in R. Dendrochronologia 40, 12–16. https://doi.org/10.1016/j.dendro.2016.06.001.

Zeileis, A., Grothendieck, G., 2005. zoo : S3 Infrastructure for Regular and Irregular Time Series. J. Stat. Soft. 14, 1–27. https://doi.org/10.18637/jss.v014.i06.

Zweifel, R., Zimmermann, L., Newbery, D.M., 2005. Modeling tree water deficit from microclimate: an approach to quantifying drought stress. Tree Physiol. 25, 147–156. https://doi.org/10.1093/treephys/25.2.147.

Zweifel, R., Zimmermann, L., Zeugin, F., Newbery, D.M., 2006. Intra-annual radial growth and water relations of trees: implications towards a growth mechanism. J. Exp. Bot. 57, 1445–1459. https://doi.org/10.1093/jxb/erj125.

Zweifel, R., Haeni, M., Buchmann, N., Eugster, W., 2016. Are trees able to grow in periods of stem shrinkage? New Phytol. 211, 839–849. https://doi.org/10.1111/nph.13995.