dendrometeR: analyzing the pulse of trees in R

Authors:

Ernst van der Maaten[a,*], Marieke van der Maaten-Theunissen[a], Marko Smiljanić[a], Sergio Rossi[b,c], Sonia Simard[d], Martin Wilmking[a], Annie Deslauriers[b], Patrick Fonti[e], Georg von Arx[e], Olivier Bouriaud[f]

[a] Institute of Botany and Landscape Ecology, University of Greifswald, Soldmannstr. 15, 17487 Greifswald, Germany.

[b] Département des Sciences Fondamentales, Université du Québec à Chicoutimi, 555 Boulevard de l'Université, Chicoutimi, QC, G7H2B1, Canada.

[c] Key Laboratory of Vegetation Restoration and Management of Degraded Ecosystems, Provincial Key Laboratory of Applied Botany, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, China

[d] Climate Dynamics and Landscape Evolution, GFZ German Research Centre for Geosciences, Telegrafenberg, 14473 Potsdam, Germany.

[e] Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland.

[f] National Forest Inventory, Forest Research and Management Institute, Bucharest, 128 Bd Eroilor, Voluntari, Romania.

* Corresponding author:
Tel: +49 (0)3834 86 4193, Fax: +49 (0)3834 86 4096
Email: ernst.vandermaaten@uni-greifswald.de and ernst.vandermaaten@gmail.com

Keywords: stem-size variation, data analysis, tree-growth monitoring, tree water status, dendrometer, radial growth

Paper type: Technical Note

Manuscript length: abstract – 153 words; main text – 2595 words; 3 figures

Abstract

Dendrometers are measurement devices proven to be useful to analyze tree water relations and growth responses in relation to environmental variability. To analyze dendrometer data, two analytical methods prevail: (1) *daily approaches* that calculate or extract single values per day, and (2) *stem-cycle approaches* that separate high-resolution dendrometer records into distinct phases of contraction, expansion and stem-radius increment. Especially the stem-cycle approach requires complex algorithms to disentangle cyclic phases. Here, we present a new R package, named dendrometeR, that facilitates the analysis of dendrometer data using both analytical methods. By making the package freely available, we make a first step towards comparable and reproducible methods to analyze dendrometer data. The package contains customizable functions to prepare, verify, process and plot dendrometer series, as well as functions that facilitate the analysis of dendrometer data (i.e. daily statistics or extracted phases) in relation to environmental data. The functionality of dendrometeR is illustrated in this note.

1. Introduction

Dendrometers are measurement devices used in plant sciences that can monitor size variation of plant organs like stems, roots, branches and fruits with high temporal and spatial resolution. In forest ecological and tree physiological research, these tools are increasingly used to study seasonal growth dynamics of trees (e.g., Duchesne et al., 2012; van der Maaten, 2013), to gain insights in environmental parameters driving tree growth (Biondi and Hartsough, 2010; Deslauriers et al., 2003; Köcher et al., 2012), and to monitor the water balance of trees (Giovanelli et al., 2007; Turcotte et al., 2011; Zweifel et al., 2005). Dendrometers continuously record stem-size variations without invasive sampling of the cambium (Drew and Downes, 2009), making them particularly suitable for long-term monitoring. Recorded signals comprise irreversible stem growth and reversible cycles of stem water depletion and replenishment (Herzog et al., 1995; Kozlowski and Winget, 1964; Tardif et al., 2001). Several approaches have been proposed to analyze the different components of these data (e.g., Deslauriers et al., 2003; Downes et al., 1999; Drew and Downes, 2009; Herzog et al., 1995; King et al., 2013). Among them, two major approaches can be identified: (1) *daily*, and (2) *stem-cycle approaches*. The daily approach characterizes the properties of the circadian cycle by calculating or extracting summary metrics per day (i.e. daily mean, minimum or maximum) (Bouriaud et al., 2005; King et al., 2013; van der Maaten et al., 2013), whereas the stem-cycle approach separates stem-size changes into the distinct phases of contraction, expansion and stem-radius increment (Deslauriers et al., 2003; Downes et al., 1999; Herzog et al., 1995). Although time series from daily and stem-cycle approaches are highly correlated (Deslauriers et al., 2007), only stem-cycle approaches can consider cycles that last longer than one day.

To disentangle the different cyclic phases from dendrometer data, Deslauriers et al. (2011) presented an algorithm for the proprietary software SAS. For the free and open-source statistical software environment R (R Development Core Team, 2016) no such routine is available, yet. A steadily increasing offer and use of dendro-related R-packages like 'dplR' (Bunn, 2008), 'treeclim' (Zang and Biondi, 2015) and 'pointRes' (van der Maaten-Theunissen et al., 2015) are clearly highlighting the appreciation of the research community to make use of the extremely versatile R environment. Hence, a new R package was developed, named dendrometeR, that facilitates the analysis of sub-daily dendrometer data. Rather than a simple translation of the original SAS code (Deslauriers et al., 2011), dendrometeR presents an innovative and more comprehensive suite of customizable functions including functions for both daily and stem-cycle approaches. In this note, we describe and illustrate the functionality of the package.

2. Package functionality

The package dendrometeR contains functions (1) to prepare and verify dendrometer and environmental data formats for further processing in the package, (2) to perform gap-filling of dendrometer data, and to sequentially process dendrometer and environmental data for (3) daily statistics and (4) stem-cycle analysis (Fig. 1). Appropriate plotting functions allow to easily visualize gap-filled time series and the stem-cycle assignments.

FIGURE 1

2.1 Data formatting and verification

The package dendrometeR requests the input data to be formatted as a data frame with a timestamp as row names (in date-time format: `%Y-%m-%d %H:%M:%S` without daylight savings, e.g., time zone GMT), and dendrometer series (or environmental data) in columns; missing values should be indicated with `NA`. To facilitate a possibly needed transformation of raw dendrometer data, the package includes a vignette called 'Import dendrometer data'. It is highly recommended to consult this vignette, as it illustrates the transformation process for diverse raw data formats. The functions `is.dendro` and `dendro.resolution` can be used to verify the correct formatting and the time resolution of the input data. The function `is.dendro` returns `TRUE` when the data is in the required format, and `FALSE` if not. In the latter case, specific error messages on the nature of the problem (e.g., problems with timestamp, non-numeric data etc.) are returned as well. The temporal resolution of the data, which needs to be constant within a time series, can be obtained using `dendro.resolution`.

2.2. Gap filling

As there may be missing values in the dendrometer data, a function named `fill_gaps` is provided. This function employs an ARIMA model (cf. Deslauriers et al. 2011) to fill gaps of short duration (i.e. several hours). The ARIMA model cannot sensibly handle long gaps, i.e. lasting over more than a day. Optimal models are selected using the `auto.arima` function from the 'forecast' package (Hyndman and Khandakar, 2008). Optionally, seasonal components of ARIMA models can be included. In that case, AR-, I- and MA-components are checked across the seasonal oscillations within the data (for dendrometer data most likely to be daily). Although the inclusion of a seasonal component might increase the robustness and precision of the ARIMA model, it will also demand more computation resources, thereby slowing down the execution of `fill_gaps`. The output of the model can be smoothed using a user-defined smoothing parameter. As the ARIMA parameters, and thus the gap-filling, might be distinct for individual growing seasons, we deliberately designed `fill_gaps` for single growing seasons. Consequently, long dendrometer series should be splitted in individual growing seasons prior to gap-filling. To allow the usage of the function for datasets from the Southern Hemisphere, the input data may contain two consecutive calendar years at maximum. `fill_gaps` can work on multiple series simultaneously and returns a data frame with gap-filled dendrometer series. The output can be conveniently displayed for specified time windows using the `fill_plot` function.

2.3. Daily approach

For daily analyses, the function `daily_stats` can be applied on both dendrometer and environmental datasets. The function returns, depending upon the entry for argument `sensor` (i.e. a `numeric` or `"ALL"`), multiple statistics (mean, minimum, maximum, amplitude, and timing of minimum / maximum) for a specified sensor, or a single statistic (daily mean, minimum, maximum, or sum) for all sensors in a data frame. The option to calculate daily sums is included in `daily_stats` as it is relevant for environmental parameters like precipitation. An optional smoothing argument is included (`smooth.param`) to handle noisy datasets; it requires gap-free (or -filled) series.

2.4 Stem-cycle approach

The stem-cycle processing includes three functions that need to be sequentially performed, i.e. `phase_def`, `cycle_stats` and `climate_seg`. The function `phase_def` identifies and assigns each timestamp to the three distinct phases of contraction, expansion and stem-radius increment for dendrometer series from a data frame with gap-free (or -filled) dendrometer data. Thereby, the function first searches for minimum and maximum points within a specified daily time window. Then, the original dendrometer series are offset back- and forward to make sure that the identified extrema are indeed extrema of the cyclic phases. A comparison between the original and offset series finally allows selecting all appropriate minimum and maximum values. The `phase_def` function can be customized in many different ways. For example, the minimum temporal distance and the minimum difference between consecutive minimum and maximum points (i.e. in *x* and *y* direction) can be specified using the arguments `minmaxDist` and `minmaxSD`, respectively. The argument `radialIncrease` allows to determine from which moment on data points should be assigned to the stem-radius increment phase: when data points are continuously above the previous maximum (`"max"`), when a single data point is above the previous maximum (`"min"`), or right in between `"min"` and `"max"` (`"mid"`). This highly flexible architecture of `phase_def` allows handling noisy and sub-hourly data as well, making it a more robust algorithm compared to the original SAS routine. The output of `phase_def`, a data frame with numbers indicating the different stem-cyclic phases, can be directly used as input for the `phase_plot` function. This plotting function creates graphs for single or multiple dendrometer series showing stem-cyclic phases (one color per phase). The time axis is automatically labeled depending upon the length of the dendrometer series. The output of `phase_def` is further used in `cycle_stats`, a function that defines stem cycles from the identified phases and that calculates statistics for all phases and cycles. These statistics include the timing and duration of each phase and cycle, as well as information on the magnitude and range of stem-size changes. The function works for single dendrometer series, which are defined by the argument `sensor`. We further included a smoothing option in `cycle_stats` (argument `smooth.param`; cf. Deslauriers et al. 2011) particularly for noisy datasets in which outliers may under- or overestimate the minimum and maximum stem size within phases and stem cycles. By default, no smoothing is performed.

The function `climate_seg` finally calculates means or sums, or extracts minimum or maximum values of environmental parameters for the stem-cyclic phases as defined using `cycle_stats`. Thereby, the function facilitates the analysis of dendrometer in relation to environmental data. For `climate_seg`, the temporal resolution of the environmental data should be equal to, or higher than that of the dendrometer data used to define the cyclic phases. Similarly, the period covered by data should be identical or longer.

The output of dendrometeR, being either daily statistics for dendrometer and environmental data (when using a daily approach) or segmented dendrometer and environmental data according to stem-cyclic phases, can be used in further analyses.

3. Illustrated example

The package dendrometeR includes dendrometer data from Canada and Germany, both raw and pre-processed, to exhaustively illustrate all functions on its integrated help pages. The Canadian series presents hourly dendrometer data for a coniferous tree (*Picea mariana* (Mill.) BSP) from Camp Daniel for the year 2008; the German series present half-hourly data for three broadleaved trees (*Fagus sylvatica* L.) from Hinnensee and Eldena for the years 2012 and 2015, respectively. For Eldena, also some temperature data is included in the package. Hence, we illustrate the functions of dendrometeR for Eldena in the following examples. Thereby, we will focus on the steps that need to be sequentially performed when using a stem-cycle approach, mainly because the usage and output of `daily_stats` is very straightforward (i.e. a single daily statistic for all sensors in a data frame, or multiple statistics for a specified sensor).

After processing the raw dendrometer data (named `dmEDraw`) using code from the vignette 'Import dendrometer data', and checking the import data using `is.dendro`, few missing records in the series should be filled using `fill_gaps`. In the following example, we introduce, after loading the data, some artificial gaps for demonstration purposes, fill these gaps and create a plot with gap-filled series:

```
> data(dmED)
> dmED[c(3189:3196, 3401:3419),1] <- NA
```

```
> dm.gpf <- fill_gaps(dmED, Hz = 0.01, season = TRUE)
> fill_plot(dmED, dm.gpf, sensor = 1, year = NULL, period = c(124, 134))
```

The argument `Hz` of `fill_gaps` is a smoothing parameter allowing to adjust the level of smoothing of the results of the ARIMA model; higher values mean rougher smoothing. With the argument `season` it can be indicated whether only non-seasonal (`season = FALSE`), or non- and seasonal models should be checked (`season = TRUE`). The output, in this case named `dm.gpf`, is directly used as input in `fill_plot`. This function creates a plot highlighting the filling of missing records in orange (Fig. 2). The argument `sensor` allows to specify a particular dendrometer (by column number), whereas `year` and `period` define the year and period (using day of year numbers for begin and end) to be plotted.

FIGURE 2

The gap-filled dendrometer data can be used as input in `phase_def` to define phases of contraction, expansion and stem-radius increment. Example code reads:

```
> dm.phase <- phase_def(dm.gpf, resolution = dendro.resolution(dm.gpf),
shapeSensitivity = 0.6, minmaxDist = 0.2, minmaxSD = 2, radialIncrease =
"max")
```

The `phase_def` argument `resolution` specifies the resolution of the dendrometer data (in seconds), and defaults to the resolution of the dendrometer data (`dm.gpf`). `shapeSensitivity` specifies the time window (i.e. proportion of a day) within which extrema points are searched for in the dendrometer data. It further defines the offsetting of dendrometer series back and forth to assure that the identified extrema are indeed the extrema of cyclic phases: offsetting is fixed to `(1 - shapeSensitivity) / 2` day ratios. The arguments `minmaxDist` and `minmaxSD` allow to specify the minimum temporal distance (i.e. in *x* direction) and the minimum difference, expressed in standard deviations (in *y* direction), between consecutive minimum and maximum points. Here, these arguments are set to 0.2

day and 2 standard deviations. `radialIncrease` is set to `"max"`, meaning that the stem-radius increment phase is first defined when dendrometer records are continuously above the previous maximum. The output `phase_def` (a data frame with numbers indicating the different stem-cyclic phases) is named `dm.phase` here, and can be directly used as input in `phase_plot` as follows:

```
> phase_plot(dm.gpf, dm.phase, sensor = 1, period = c(145, 151), colPhases
= c("#fdcc8a", "#fc8d59", "#d7301f"), pch = 16, main = "Sensor Beech03
(2015)")
```

The `phase_plot` function creates a plot showing the three distinct phases of contraction, expansion and stem-radius increment for a period as defined in `period`, while using colors as specified in the `colPhases` argument (Fig. 3). `colPhases` defaults to the first three colors of the current palette. Additional graphical parameters (e.g., points, axis, text and color options) can be added to `phase_plot` as for the high-level plotting function `plot`.

FIGURE 3

The output of `phase_def` can further be used as input for `cycle_stats`. This function defines the actual stem cycles, and calculates statistics for them as well as for all individual stem-cyclic phases. The function can be called as follows:

```
> dm.stats <- cycle_stats(dm.gpf, dm.phase, sensor = 1)
```

The output of `cycle_stats` is a list containing a data frame named 'cycleStats' with information on the timing (begin and end) and duration (in hours and minutes) of all phases and cycles, as well as on the magnitude and range of stem-size changes. The function `climate_seg` finally calculates means or sums, or extracts minimum or maximum values of environmental parameters for the stem-cyclic phases as defined using `cycle_stats`. Example code to run the function reads:

```
> data(envED)
> clim.phase <- climate_seg(envED, dm.stats, value = "mean")
```

Next to the output of `cycle_stats` (`dm.stats`), the function requires environmental data as input. This environmental data should cover at least the same period as the dendrometer data, should have the same (or a higher) temporal resolution, and should be similarly formatted (verify using `is.dendro`). The `climate_seg` argument `value` allows to specify whether means (`"mean"`), sums (`"sum"`), minimum (`"min"`) or maximum values (`"max"`) should be calculated or extracted. As the example data includes air and soil temperature parameters, `"mean"` was selected here. The output of `climate_seg` is a data frame with the environmental data segmented for all phases and cycles.

4. Package availability

The dendrometeR package is available as an add-on package in R, and can be downloaded from the Comprehensive R Archive Network website (CRAN: http://cran.r-project.org/web/packages/dendrometeR). To install dendrometeR from the R console, type 'install.packages("dendrometeR")'. dendrometeR requires the packages 'forecast' (Hyndman, 2015), 'pspline' (Ripley, 2015) and 'zoo' (Zeileis and Grothendieck, 2005).

The package dendrometeR is designed with entry-level users of R in mind, and comes with extensive documentation including example code for all functions. In addition, a vignette describes how input data can be formatted and verified. The package documentation is accessible from the R console using the command '?dendrometeR', or directly from the integrated help pages for users of the RStudio software (RStudio, 2015).

5. Outlook

The package dendrometeR contains customizable functions to import, verify, process and plot high-resolution dendrometer data. Further, it facilitates analyses of dendrometer data in relation to environmental parameters. By making the package freely available in the open source R statistical software, we made a first step towards homogenized analyses of dendrometer data. In the future, new functions may be added to dendrometeR depending upon suggestions of the research community.

Acknowledgements

This study is based upon work from COST Action FP1106 STReESS, supported by COST (European Cooperation in Science and Technology). EM and SS were further supported by the Helmholtz Association within the frame of the Virtual Institute of Integrated Climate and Landscape Evolution Analysis (ICLEA). MW was funded by DFG Wi2680/2-1. PF and GA have been supported by a grant from the Swiss State Secretariat for Education, Research and Innovation SERI (SBFI C14.0104). OB acknowledges support by a grant of the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, project number PN-II-ID-PCE-2011-3-0781.

References

Biondi, F., Hartsough, P., 2010. Using automated point dendrometers to analyze tropical treeline stem growth at Nevado de Colima, Mexico. Sensors 10, 5827-5844.

Bouriaud, O., Leban, J.M., Bert, D., Deleuze, C., 2005. Intra-annual variations in climate influence growth and wood density of Norway spruce. Tree Physiol. 25, 651-660.

Bunn, A.G., 2008. A dendrochronology program library in R (dplR). Dendrochronologia 26, 115-124.

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 Québec (Canada). Trees 17, 477-484.

Deslauriers, A., Rossi, S., Anfodillo, T., 2007. Dendrometer and intra-annual tree growth: what kind of information can be inferred? Dendrochronologia 25, 113-124.

Deslauriers, A., Rossi, S., Turcotte, A., Morin, H., Krause, C.l., 2011. A three-step procedure in SAS to analyze the time series from automatic dendrometers. Dendrochronologia 29, 151-161.

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.

Drew, D.M., Downes, G.M., 2009. The use of precision dendrometers in research on daily stem size and wood property variation: a review. *Dendrochronologia* 27, 159-172.

Duchesne, L., Houle, D., D'Orangeville, L., 2012. Influence of climate on seasonal patterns of stem increment of balsam fir in a boreal forest of Québec, Canada. Agric. For. Meteorol. 162-163, 108-114.

Giovannelli, A. et al., 2007. Evaluation of drought response of two poplar clones (*Populus* x *canadensis* Monch 'I-214' and *P-deltoides* Marsh. 'Dvina') through high resolution analysis of stem growth. J. Exp. Bot. 58, 2673-2683.

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.

Hyndman, R.J., Khandakar, Y., 2008. Automatic time series forecasting: the forecast package for R. J Stat Softw 26, 1-22.

Hyndman, R.J., 2015. forecast: Forecasting functions for time series and linear models. R package version 6.2, http://github.com/robjhyndman/forecast.

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.

Köcher, P., Horna, V., Leuschner, C., 2012. Environmental control of daily stem growth patterns in five temperate broad-leaved tree species. Tree Physiol., 32: 1021-1032.

Kozlowski, T.T., Winget, C.H., 1964. Diurnal and seasonal variation in radii of tree stems. Ecology 45, 149-155.

R Development Core Team, 2014. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

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/package=pspline

RStudio, 2015. RStudio: Integrated development for R. RStudio, Inc., Boston, MA.

Tardif, J., Flanningan, M., Bergeron, Y., 2001. An analysis of the daily radial activity of 7 boreal tree species, North-western Quebec. Environ. Monit. Assess. 67, 141-160.

Turcotte, A., Morin, H., Krause, C., Deslauriers, A., Thibeault-Martel, M., 2009. The timing of spring rehydration and its relation with the onset of wood formation in black spruce. Agric. For. Meteorol. 149, 1403-1409.

van der Maaten, E., 2013. Thinning prolongs growth duration of European beech (*Fagus sylvatica* L.) across a valley in southwestern Germany. Forest Ecol. Manag. 306, 135-141.

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.

van der Maaten-Theunissen, M., van der Maaten, E., Bouriaud, O. 2015. pointRes: An R package to analyze pointer years and components of resilience. Dendrochronologia 35, 34-38.

Zang, C., Biondi, F., 2015. treeclim: an R package for the numerical calibration of proxy-climate relationships. Ecography 38, 431-436.

Zeileis, A., Grothendieck, G. 2005. zoo: S3 Infrastructure for Regular and Irregular Time Series. J. Stat. Softw., 14, 1-27.

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.

Figure captions

Fig. 1 Schematic overview of the functions included in dendrometeR.

Fig. 2 Example of a plot created with the `fill_plot` function. Gap-filled records are indicated in orange. Data is presented for a European beech tree from Eldena (Germany) for selected days in 2015.

Fig. 3 Example of a plot created with the `phase_plot` function. The stem-cyclic phases of contraction, expansion and stem-radius increment are indicated by different colors. A sequential color scheme from ColorBrewer (http://colorbrewer.org) was used. Data is presented for a European beech tree from Eldena (Germany) for selected days in 2015.