*dfoliatR*: An R package for detection and analysis of insect defoliation signals in tree rings

2020-03-12

We present a new R package to provide dendroecologists with tools to infer, quantify, analyze, and visualize growth suppression events in tree rings caused by insect defoliation. The ‘dfoliatR’ library is based on the FORTRAN V program OUTBREAK, and builds on existing resources in the R computing environment. ‘dfoliatR’ expands on OUTBREAK to provide greater control of suppression thresholds, additional output tables, and high-quality graphics. To use ‘dfoliatR’ requires standardized ring-width measurements from insect host trees and an indexed tree-ring chronology from local non-host trees. It performs an indexing procedure to remove the climatic signal represented in the non-host chronology from the host-tree series. It then infers defoliation events in individual trees. Site-level analyses identify outbreak events that synchronously affect a user-defined number or proportion of the host trees. Functions are available for summary statistics and graphics of tree- and site-level series.

# Introduction

Openning paragraph:

* General overview of importance of defoliation reconstructions to forest ecology and management
* Contributions of tree-ring research to these ends
* Lead-in for need for computing power and repeatable/open science. (In the burnr paper, we framed it around a need for coarse-scale syntheses of regions, etc. that requires more powerful and efficient programming)

The dfoliatR library adds to a growing suite R packages designed for dendrochronology. Stemming from the dplR library (Bunn 2008) that enables R users to read and write an array of tree-ring data formats, standardize ring widths, build and evaluate chronologies, perform quality control (to name a few), one can now also measure ring widths (Lara et al. 2015, Shi and Xiang 2019), perform and check crossdating (Bunn 2010), and perform many analytical tests (Zang and Biondi 2015, Jevšenak and Levanič 2018). Tools for assessing stand dynamics and disturbance analyses are under rapid development, with new packages for assessing growth and release events (TRADER: Altman et al. 2014), metrics of growth resilience (pointRes: Maaten-Theunissen et al. 2015), and fire history (burnr: Malevich et al. 2018). The key objective of dfoliatR is to provide tools to identify and analyse insect defoliation and outbreak events by building on the methods employed the FORTRAN program OUTBREAK (Swetnam and Lynch 1989). What sets dfoliatR apart from packages such as TRADER is that it explicitly performs an indexing procedure on host-tree series to remove climatic and other non-defoliation related signals represented by non-host tree species. Insect defoliation signals are identified in the disturbance index by the duration and magnitude of negative departures.

dfoliatR draws upon data formats in dplR that are commonly employed by other tree-ring libraries. It uses and outputs data formats that facilitate the use packages embodied by tidyverse (Wickham et al. 2019) that include efficient data manipulation (dplyr: Wickham et al. 2020) and graphics (ggplot2: Wickham 2016).

In this paper, we describe the statistical methods employed by dfoliatR, its availability, and run through analyses for a single site in New Mexico. Users need not have much experience in R to replicate the analyses and graphics below. All R code presented below is executable in an R session once the required libraries are installed and loaded. Support documentation in addition to this paper is provided within the package via standard help menus (accessed by typing ? before a function name) and on the package website (<https://chguiterman.github.io/dfoliatR/>), which includes up-to-date vignettes that describe in detail the functionality of the software. Code to create a preprint of this manuscript including the R scripts is available from <https://github.com/chguiterman/dfoliatR_paper>.

# Overview of the software

The dfoliatR library requires two sets of tree-ring data to identify defoliation and outbreak events:

* Standardized ring-width series for individual trees of the host species
* Standardized tree-ring chronology from a local non-host species

Users can develop these data sets in software of their choosing, such as dplR or ARSTAN. It is important that the host-tree data include only one tree-ring series per tree. Both dplR and ARSTAN have options for averaging multiple sample series into a tree-level series. The tree-ring series and chronology can be read into R via several available dplR functions.

At the heart of dfoliatR lies two functions: defoliate\_trees() and outbreak(). These identify defoliation event son individual trees and then composite across multiple trees for a given site to identify outbreak events. It is up to the researcher to distinguish these separate spatial scales of analysis, dfoliatR then provides a set of functions to visualize and statistically summarize tree-level defoliation and site-level outbreak events, which will be discussed further.

## Identifying Defoliation of Trees

The defoliate\_trees() function will often be the point of entry to the dfoliatR library. It performs two processes: removes non-defoliation growth signals from the host-tree series and then identifies defoliation events. The climatic or non-defoliation signals in each host-tree series are represented by the non-host chronology or a climate reconstruction. dfoiatR removes that non-defoliation signal by subtracting the non-host series from each host-tree series, which generates a disturbance index. In the program OUTBREAK, this residual index was termed the “corrected” index. We call it the “growth suppression index” (GSI). The GSI is calculated the same as in OUTBREAK (following Swetnam et al. 1985, Swetnam and Lynch 1989) for each host tree as

where H and NH are the host-tree series and the non-host chronology, in year i, respectively. Only the common period between the host-tree series and the non-host chronology are used in Equation 1. The non-host chronology is scaled by its mean () and multiplied by the ratio of host and non-host standard deviations (), which approximates the variance of the host tree series.

Negative departures in the normalized GSI that surpass user-defined thresholds in duration and magnitude are defined as *defoliation events*. As in OUTBREAK, magnitude is assessed on a single year within the departure sequence. The default setting is -1.28 (in units of standard deviation), which was previously determined to be representative of WSBW effects (**cite**). Duration is assessed by examining sequences of negative GSI before and after the year of maximum departure. Each defoliation event is allowed one positive excursion on each side of the maximum departure year. Duration is assessed across the entire sequence that includes up to two positive excursions. The default duration is eight years, as is commonly used in WSBW studies (**cite**). Different species of defoliation insects vary in the length of defoliation and the degree to which they can suppress tree growth. Researchers should adjust the duration and magnitude parameters accordingly and critically evaluate the results.

Diverging from OUTBREAK, dfoliatR allows users to extend defoliation events by bridging between successive events and also by allowing potentially short-duration events that occur at the end of the series. In cases where two defoliation events are separated by a single year, bridging will connect them through time. We urge careful use of this option because there is no setting to limit the number of potentially bridged events. The series end option can be used in cases when the host trees were actively being defoliated at the time of sampling. This option eliminates the duration parameter for an event at the recent end of the series, but all other thresholds still apply. The advantage of this parameter is that it can aid in identifying the start-year for the current defoliation event or outbreak, which is both useful in management and allows the current event to be included in return-interval estimates.

## Identifying Outbreak Events

To determine when defoliation becomes an *outbreak event*, dfoliatR composites the individual tree defoliation series into a site-level chronology with the outbreak() function. Users have options to define the number and/or the proportion of trees required for an event to be considered an outbreak. Three parameters control the whether a defoliation event constitutes an outbreak: the minimum number of trees available, the minimum number of trees recording defoliation, and the percent of trees recording defoliation. The first allows the researcher to make a judgement call as to the confidence ascribed to reduced sample depth toward the ends of their chronologies, thus compensating for the “fading record problem” (Swetnam and Fritts?). The second two parameters adjust the scale of defoliation considered an outbreak. Absolute numbers of trees and percentages can be applied separately or in conjunction, following filtering conventions in tree ring fire history studies (Malevich et al. 2018).

# Availability and installation

The dfoliatR library (Guiterman et al. 2020) is provided free and open source from the Comprehensive R Archive Network (CRAN; <https://cran.r-project.org/>). To install dfoliatR from CRAN use

install.packages("dfoliatR")

In each R session, dfoliatR can be loaded via

library(dfoliatR)

Development versions of dfoliatR are available on GitHub and installed using the devtools library,

devtools::install\_github("chguiterman/dfoliatR")

Issues, bug reports, and ideas for improving dfoliatR can be posted to <https://github.com/chguiterman/dfoliatR/issues>. As an Open Source library, we welcome and encourage community involvement in future development. The best ways to contribute to dfoliatR are through standard GitHub procedures or by contacting the first author.

# Example Usage

In dfoliatR we provide two sets of tree-ring data to aid users in exploring the functions, graphics, and outputs. Each set consists of Douglas-fir (*Pseudotsuga menziesii*) host-tree series, standardized with \_\_\_\_\_\_\_\_\_\_, and a local ponderosa pine (*Pinus ponderosa*) non-host chronology. The non-host ring-width data were standardized by \_\_\_\_\_\_\_\_ and the chronologies averaged following standard procedures (Speer 2010). Data from Demijohn Peak (DMJ; 2902 m asl), in the San Juan Mountains of southern Colorado, come from Ryerson et al. (2003). Data from the East Fork site (EF; 2580 m asl) in the Jemez Mountains of northcentral New Mexico were presented by Swetnam and Lynch (1993).

## Tree-Level Defoliation Events

## Site-Level Events

# Evaluation

* Describe testing of dfoliatR vs OUTBREAK?
* What to present?

# Extensions

* Describe how dfoliatR can be combined with other R libraries
  + Mapping, what else?

# References

Altman, J., P. Fibich, J. Dolezal, and T. Aakala. 2014. TRADER: A package for Tree Ring Analysis of Disturbance Events in R. Dendrochronologia 32:107–112.

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

Bunn, A. G. 2010. Statistical and visual crossdating in R using the dplR library. Dendrochronologia 28:251–258.

Guiterman, C., A. Lynch, and J. Axelson. 2020. DfoliatR: Detection and analysis of insect defoliation signals in tree rings.

Jevšenak, J., and T. Levanič. 2018. dendroTools: R package for studying linear and nonlinear responses between tree-rings and daily environmental data. Dendrochronologia 48:32–39.

Lara, W., F. Bravo, and C. A. Sierra. 2015. MeasuRing: An R package to measure tree-ring widths from scanned images. Dendrochronologia 34:43–50.

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

Malevich, S. B., C. H. Guiterman, and E. Q. Margolis. 2018. burnr: Fire history analysis and graphics in R. Dendrochronologia 49:9–15.

Ryerson, D. E., T. W. Swetnam, and A. M. Lynch. 2003. A tree-ring reconstruction of western spruce budworm outbreaks in the San Juan Mountains , Colorado , U . S . A . Canadian Journal of Forest Research 33:1010–1028.

Shi, J., and W. Xiang. 2019. MtreeRing: A shiny application for automatic measurements of tree-ring widths on digital images.

Speer, J. H. 2010. Fundamentals of tree-ring research. Page 333. The University of Arizona Press.

Swetnam, T. W., and A. M. Lynch. 1989. A tree-ring reconstruction of western spruce budworm history in the southern Rocky Mountains. Forest Science 35:962–986.

Swetnam, T. W., and A. M. Lynch. 1993. Multicentury, Regional-Scale Patterns of Western Spruce Budworm Outbreaks. Ecological Monographs 63:399–424.

Swetnam, T. W., M. A. Thompson, and E. K. Sutherland. 1985. Using dendrochronology to measure radial growth of defoliated trees. Page 39p.

Wickham, H. 2016. Ggplot2: Elegant graphics for data analysis. Springer.

Wickham, H., M. Averick, J. Bryan, W. Chang, L. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. Pedersen, E. Miller, S. Bache, K. Müller, J. Ooms, D. Robinson, D. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani. 2019. Welcome to the tidyverse. Journal of Open Source Software 4:1686.

Wickham, H., R. François, L. Henry, and K. Müller. 2020. Dplyr: A grammar of data manipulation.

Zang, C., and F. Biondi. 2015. Treeclim: An R package for the numerical calibration of proxy-climate relationships. Ecography 38:431–436.