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

Christopher H. Guiterman[*,a,b], Ann M. Lynch[a,c], Jodi N. Axelson[c]

[a]*Laboratory of Tree-Ring Research, University of Arizona, 1215 E Lowell St. Box 210045, Tucson, AZ, 85721*
[b]*Three Pines Forest Research, LLC, PO Box 225, Etna, NH, 03750*
[c]*U.S. Forest Service, Rocky Mountain Research Station, 1215 E Lowell St. Box 210045, Tucson, AZ, 85721*
[d]*Dept of Environmental Science, Policy & Management, University of California, Berkeley, Berkeley, CA 94720*

Abstract

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.

Key words: Dendroecology, spruce budworm, Choristoneura, pandora moth, Coloradia pandora Blake, larch-bud-moth

[*]Corresponding Author
Email addresses: `chguiterman@email.arizona.edu` (Christopher H. Guiterman), `lyncha@email.arizona.edu` (Ann M. Lynch), `jodi.axelson@berkeley.edu` (Jodi N. Axelson)

1. Introduction

Variation in the width and morphology of annual radial growth rings in trees permits dating and quantification of past forest insect defoliator outbreaks. Defoliation can be distinguished from climate- and other disturbance-related influences by comparing ring-width or other annually-resolved features in the wood of host species to non-host species or to climate records. The effect of defoliation on radial growth of trees has been recognized since the mid-1800s, but it was not until the 1980s that dendrochronology techniques were applied to identify and quantify defoliation events (Lynch, 2012; Speer, 2010; Swetnam et al., 1985). The first studies (Swetnam and Lynch, 1993, 1989; Swetnam et al., 1985) focused on developing historical reconstructions of western spruce budworm (WSBW; *Choristoneura freemani*). The methodology has since been successfully applied to a wide range of defoliator species, most of which are conifer herbivores, and has evolved in sophistication and application to a wide range of ecosystem situations (Lynch, 2012).

The main dendrochronological tool for inferring, dating, and quantifying defoliator outbreaks from tree-ring records has been the software routine OUT-BREAK (Holmes and Swetnam, 1986; Swetnam and Lynch, 1989; Swetnam et al., 1985). OUTBREAK computes indices (described later in detail) of suppressed growth by subtracting a detrended and standardized climate series (a "control" chronology) from host individual-tree detrended and standardized radial growth series after the host and non-host series have been brought to a common variance. The non-host chronology often consisted of a site chronology developed from non-host tree species, but a gridded climate data point series, like the North American Drought Atlas (Cook and Krusic, 2004) also suffices. If the host and non-host species respond similarly to climate (which can be tested), the derived series retains variability that the host and non-host series do not have in common, generally the insect signal, along with some unexplained variability (noise). The user defines a rule base specifying the magnitude and duration that a period of indexed growth suppression must meet or surpass for

2

a period of suppressed growth to be inferred as a defoliation event at the tree level.

Though powerful, OUTBREAK is outdated and increasingly difficult to use in modern computing environments. It was written in FORTRAN V with inherently severe restrictions, as RAM and disk space were limited at that time (256 kb and 10 MB, respectively) and FORTRAN conventions imposed very strict formatting, file naming, and output conventions. The program provides no graphical interface or capabilities, forcing users to import generated text files into spreadsheet or other software for assessing results and performing analyses. Furthermore, OUTBREAK can only handle one test at a time, creating barriers to batch operation and a large burden for researchers with a datasets including multiple sites. We developed the software library `dfoliaR` as an `R`- and `dplR`-based routine to overcome these issues.

`dfoliatR` adds to a growing suite of dendrochronology packages the `R` computing environment (R Core Team, 2019). Stemming from the `dplR` library (Bunn, 2008) that enables `R` users to read and write an array of tree-ring data formats, standardize ring width series, build and evaluate chronologies, and perform quality control (to name a few), one can now also measure ring widths from scanned images of prepared samples (Lara et al., 2015; Shi and Xiang, 2019), perform and check crossdating (Bunn, 2010), and perform many analytical tests (Jevšenak and Levanič, 2018; Zang and Biondi, 2015). 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 analyze insect defoliation and outbreak events by building on the methods employed by OUTBREAK. It capitalizes on the robust software already available in `R` by using `dplR` data formats for incoming tree-ring series and providing output data formats embodied by the `tidyverse` (Wickham et al., 2019) that include efficient data manipulation (Wickham et al., 2020) and graphics (Wickham, 2016).

3

In this paper, we describe the statistical methods employed by `dfoliatR`, compare results to those produced by OUTBREAK, and present an example analysis. Users need not have much experience in `R` to replicate the analyses and graphics as presented. The `R` code 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 various software routines. Code to generate a preprint of this manuscript including the `R` scripts, and tubular and graphical output is available from https://github.com/chguiterman/dfoliatR_paper.

2. Overview of the software

The `dfoliatR` library requires two sets of tree-ring data to infer 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` (Bunn, 2008) or ARSTAN (Cook and Holmes, 1996). 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.

At the heart of `dfoliatR` lies two functions: `defoliate_trees()` and `outbreak()`. These identify defoliation events on individual trees (Figure 1) and then composite across multiple trees to identify outbreak events (Figure 2).

2.1. Identifying Defoliation of Trees

The `defoliate_trees()` function is the point of entry to the `dfoliatR` library. It performs two processes, removing climate-related growth signals from

4

the host-tree series and identifying tree-level defoliation events. The climatic or non-defoliation signals in each host-tree series are characterized by a non-host chronology or climate reconstruction. `dfoiatR` removes that non-defoliation signal by subtracting the non-host series from each host-tree series, which generates a residual index. In 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 for each host tree as

$$\text{GSI}_i = \text{H}_i - \left(\text{NH}_i - \overline{\text{NH}}\right) \frac{\sigma_\text{H}}{\sigma_\text{NH}} \tag{1}$$

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 host and non-host chronologies are brought to common variance by scaling the non-host chronology by its mean ($\overline{\text{NH}} \approx 1.0$) and multiplying by the ratio of host and non-host standard deviations ($\frac{\sigma_\text{H}}{\sigma_\text{NH}}$), which approximates the variance of the host tree series.

Negative departures in the normalized GSI (NGSI) 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 (Swetnam and Lynch, 1989). Duration is assessed by examining sequences of negative NGSI 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 (Swetnam and Lynch, 1989). Different species of defoliation insects vary in the length of defoliation and the degree to which they can suppress tree growth. Researchers can, and 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 successive events and also by allowing potentially short-

5

duration events that occur at the end of the series. In cases where two defoliation events are separated by a single year, bridging will link them into a single event (**Show figure?**). We urge careful use of this option because there is no setting to limit the number or length 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 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.

2.2. Inferring Outbreak Events

Defoliation of one or a few trees does not constitute an outbreak. 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 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 et al., 1999). The second two parameters adjust the scale of defoliation considered to be 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). We urge users to carefully consider the choice of absolute numbers in situations where the number of trees represented in the series varies with time, or the choice of percentages when sample size is small.

3. Evaluation

3.1. Approach

We tested `dfoliatR` against OUTBREAK by comparing NGSI to corrected indices for individual trees and years, defoliation status for individual trees and years, and percentage of trees recording outbreaks at the site level, using standardized ring-width data from eight sites in British Columbia, Colorado, and New Mexico, and author-provided non-host site chronologies.

We detrended host data for both `dfoliatR` and OUTBREAK using AR-STAN v6.1 (Cook and Holmes, 1996) with default double detrending (128 year wavelength and a 50% smoothing spline). In both `dfoliatR` and OUTBREAK we used event thresholds of -1.28 normalized index, 8 years duration, and allowed for events at the end of series in seven of eight sites. These seven sites were sampled during ongoing outbreak events (Swetnam and Lynch, 1993). We found it necessary to be consistent in how we detrended and in what software (e.g., ARSTAN vs `dplR`) we employed because subtle differences in standardized ring-width indices created between the programs exacerbated differences between `dfoliatR` and OUTBREAK. In the end, we chose to only use the standardization output files from ARSTAN, which are easily read into `R` (and then `dfoliatR`) using the `dplR` package.

The `R` code to replicate our comparisons is available from https://github.com/chguiterman/dfoliatR_paper.

3.2. Findings

Since both programs employ Equation 1, `dfoliatR` and OUTBREAK compute identical growth suppression indices at 0.000 precision (Figure 3). In all there are 43,280 ring-width indices from 222 trees included in our test. Of these, the programs identified 11,530 years of defoliation. The programs agreed on 92% of the years, leaving 927 "difference" years in which only one program identified defoliation on an individual tree. The differences included 102 events on 85 trees. We carefully inspected each of these events in the full context of each tree's ring series, and categorized the differences as follows

7

- *Series-end events* (40% of the total) in which we allowed OUTBREAK to include "truncated outbreaks" (for seven sites) at the end of each series. In `dfoliatR`, this option is controlled by the "series_end_events" parameter to `defoliate_trees()`. In OUTBREAK, the option appears while changing the duration parameter (option 3). When selected, OUTBREAK will include any sequences of negative indices at the either the beginning or the end of each tree series. No thresholds are applied in identifying these as defoliation events. In `dfoliatR`, the negative sequences are evaluated as with any other possible defoliation, but the duration threshold is omitted. Each of the 13 events included in these differences did not meet the "max_reduction" parameter (-1.28 NGSI) in `dfoliatR` and were excluded. In two cases, OUTBREAK included events at the beginning of the series where `dfoliatR` does not allow truncated events. In four cases, OUTBREAK omitted the last year of the series because it was positive. In two cases, `dfoliatR` omitted possible events because it had already included a positive NGSI excursion after the "max_reduction" year, and since it will only allow one excursion on either side of the max year, the events were omitted due to short duration.

- *Sequential events* (36%) in which OUTBREAK omitted events that occur one year prior or one year following an identified event. OUTBREAK selects the event with the greatest negative index year. On two trees, OUTBREAK omitted two of three sequential events. `dfoliatR` allows all such events. While inspecting these differences, we added an option to `defoiate_trees()` that would "bridge" between sequential events, making them into single, long events. We felt that this was ecologically justified, especially for studies of WSBW, because the insect will occasionally maintain minor defoliation f individuals for decades even while not in an outbreak phase.

- *Undetermined differences* (22%) occurred in cases where OUTBREAK omitted events without clear cause.

- *Rounding* (2%) differences in the indices either omitted or cut short events on two trees. In both cases the indices were very close to zero.

Our inspections revealed what we believe are short-comings in how OUT-BREAK identifies defoliation events on individual trees. In every one of the 102 cases we inspected, we felt that `dfoliatR` provided a more biologically and statistically appropriate assessment of defoliation.

and use them to infer identical tree-level outbreak events, including onset and termination dates and date and magnitude of the maximum growth suppression when `dfoliatR` bridging matches the OUTBREAK protocols. Once the issues with ARSTAN (described earlier) were resolved, dfoliatR and OUTBREAK produced identical indices at 0.000 precision.

When inferring outbreak events at the tree level, our initial results did not always match OUTBREAK, leading to development of the bridging option. Identifying a minimum within a period while allowing for positive excursions is a tricky computation. OUTBREAK deliberately disallows back-to-back events – after a minimum value within a period of negative values, a second positive index terminates an event. OUTBREAK does not allow a period of highly negative values to be appended to an earlier outbreak if the intervening period is 1 year, even if those negative values fall below the minimum in the earlier period. Due in large part to reconstructions using OUTBREAK (see papers cited by Lynch, 2012), we now know considerably more about forest defoliator outbreak regimes than we did in the 1980s when OUTBREAK was written. We think that two prolonged events separated by a single year should in some situations be considered a single outbreak at both the tree and site levels. This is particularly relevant to western spruce budworm and spruce budworm (*C. fumiferana*), for which regimes have been reconstructed for several geographic areas and for which researchers and forest health experts have gained considerable knowledge (**Sanders et al. 1985, Brookes et al. 1987 and many later publications**). We now know that the greatest growth suppression often occurs late in the outbreak due to cumulative effects (**rrr**). We developed the bridging option to

permit these periods to be linked to preceding events separated by a single year.

Bridging operates at the tree level in `dfoliatR`, and these effects generally have minor effects on site-level reconstructions (**Fig.Bridging**). Occasionally site-level events are inferred differently (**Table.Bridging**), and this can affect return interval and duration statistics.

3.3. Best Practice Recommendations

We encourage authors to report software and library version numbers and download dates where applicable. This is standard practice for statistical packages, but has often been ignored for R- and dplR-based analyses. R-based software evolves over time, and results cannot always be replicated if the software versions are unknown. Similarly, significant changes (especially fixes) imply that re-analysis of data might differ from earlier published results and interpretations. Wherever possible, refereed publications supporting the software should be cited. Our experience with ARSTAN demonstrates the importance of adopting this practice. In the United States, analytic methods not documented thusly and supported by the literature might be questioned if the results are used in a National Environmental Protection Act (NEPA) process.

We recommend that researchers compare dfoliatR results with and without bridging, and see which option agrees best with known insect biology and ecology. Bridging may be a logical choice for insects with prolonged non-outbreak periods, such as western spruce budworm, but not for situations where impacted stands barely recover from one outbreak before another begins, such as occurs with pine processionary caterpillars (*Thaumetopoea pityocampa* (Lepidoptera: Thaumetopoeidae)) (**Carus 2004, 2009**).

4. 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")
```

²⁸¹ In each R session, dfoliatR can be loaded via

```
library(dfoliatR)
```

²⁸² Development versions of dfoliatR are available on GitHub and installed
²⁸³ using the devtools library,

```
devtools::install_github("chguiterman/dfoliatR")
```

²⁸⁴ Issues, bug reports, and ideas for improving dfoliatR can be posted to
²⁸⁵ https://github.com/chguiterman/dfoliatR/issues. As an Open Source library,
²⁸⁶ we welcome and encourage community involvement in future development. The
²⁸⁷ best ways to contribute to dfoliatR are through standard GitHub procedures
²⁸⁸ or by contacting the first author.

5. Example Usage

²⁹⁰ In dfoliatR we provide two sets of tree-ring data to aid users in exploring the
²⁹¹ functions, graphics, and outputs. Each set consists of individual host-tree series,
²⁹² standardized using 50-year splines with a 50% frequency response, and a local
²⁹³ non-host chronology. The non-host ring-width data were standardized using
²⁹⁴ 150-year splines with a 50% frequency response. The non-host chronologies
²⁹⁵ were averaged following standard procedures (Speer, 2010). Host trees from
²⁹⁶ Demijohn Peak (DMJ; 2902 m asl) in the San Juan Mountains of southern
²⁹⁷ Colorado include Douglas-fir (*Pseudotsuga menziesii*) compared against a local
²⁹⁸ non-host ponderosa pine (*Pinus ponderosa*) chronology (Ryerson et al., 2003).
²⁹⁹ The East Fork site (EF; 2580 m asl) in the Jemez Mountains of northcentral
³⁰⁰ New Mexico include Douglas-fir and white fire (*Abies concolor*) host trees and
³⁰¹ a ponderosa pine non-host chronology (Swetnam and Lynch, 1993).

5.1. Tree-Level Defoliation Events

303 To assess defoliation of trees, we begin with a set of standardized host-tree
304 series and a local non-host chronology or climate reconstruction. First, the
305 `dfoliatR` library needs to be loaded (after installation as above) and the in-
306 cluded datasets can be loaded. The function `defoliate_trees()` performs the
307 GSI indexing procedure on each host-tree series and then identifies defoliation
308 events.

```r
## Load the package
library(dfoliatR)

## Load data
data("dmj_h") # host tree series
data("dmj_nh") # non-host chronology
# View data structure for any data object via
# "View(dmj_h)" or "head(dmj_h)"

## Calculate indices, identify defoliation events
dmj_defol <- defoliate_trees(host_tree = dmj_h,
                             nonhost_chron = dmj_nh,
                             duration_years = 8,
                             max_reduction = -1.28,
                             bridge_events = TRUE,
                             series_end_event = TRUE,
                             list_output = FALSE)
```

309 The result of running `defoliate_trees()` is long-format (stacked) data
310 frame with five variables: "year", "series", "gsi", "ngsi", and "defol_status."
311 The "gsi" variable is the "corrected" tree-level growth suppression index. It was
312 normalized or scaled into z-scores to produce "ngsi", which is used to identify de-
313 foliation events. The "defol_status" column provides a set of character strings,

12

"nd" for non-defoliation year, "defol" for a defoliation year, "max_defol" for the year of maximum suppression (that acts as the basis for individual events), "bridge_defol" to identify years that link subsequent events (only one is present at DMJ), and "series_end_defol" to identify defoliation at the present-end of the series. Selecting `list_output = TRUE` provides a list-object of data frames, each with the `rwl` object that combines the host tree and non-host chronology and the columns created in `defoliate_trees()`. This option is not used by subsequent functions in `dfoliatR`, but researchers can examine it to check the results of the GSI calculation (Equation 1).

We can assess the results of running `defoliate_trees()` through graphical and table outputs. The function `get_defol_events()` will provide a list of every defoliation event for every tree, with the corresponding mean "ngsi" value. A summary table of the results for each tree is produced by `defol_stats()` (Table 1).

```
defol_stats(dmj_defol)
View(dmj_defol)
```

The `plot_defol()` function produces a "ggplot" graphics object with line segments showing the measured sequence of each series and a filled segment for each identified defoliation event (Figure 1). The defoliation segments are colored by their relative severity based on their average NGSI value. The default cut-off values between "Severe" and "Moderate" is the overall mean across all events and between "Moderate" and "Minor" is the first quartile. Users can define these breaks to suit their needs.

```
plot_defol(dmj_defol)
## This creates a ggplot object, so additions can
## be made to adjust plotting themes and
## aesthetics, like color.
## For more on the features of ggplot:
## https://ggplot2.tidyverse.org/
```

13

```
library(ggplot2)
plot_defol(dmj_defol) +
  scale_color_manual(values = c("red", "orange", "purple"))
```

These basic output functions aid in assessing the sensitivity of input param-
eters to `defoliate_trees()` including the duration and magnitude thresholds
for identifying defoliation events. Using `plot_defol()` also provides a direct
assessment of the between-tree variability in deflation. For further analyses, we
composite the host-tree defoliation series to the site level using `outbreak()`.

5.2. Site-Level Events

Once we have identified defoliation events on individual trees, we can infer
outbreak events at the site level. The function `outbreak()` acts to composite
tree-level defoliation series into a single chronology, with input parameters that
control thresholds in the number and proportions of trees recording a defoliation
event.

```
## Use the defol object "dmj_defol" created above
dmj_obr <- outbreak(dmj_defol,
                    filter_min_series = 3,
                    filter_min_defol = 1,
                    filter_perc = 25)
```

Input parameters to `outbreak()` include "filter_min_series" to control the
chronology cut-off points with regard to sample depth, "filter_min_defol" and
"filter_perc" to control the minimum number and percent of trees, respec-
tively, recording a defoliation event in a given year. `Outbreak()` produces a
new data frame with eight variables: "year", "num_defol", "percent_defol",
"num_max_defol", "mean_gsi", "mean_ngsi", and "outbreak_status." All of
these variables are populated regardless of an inferred outbreak event, providing
a continuous disturbance chronology. The "num_max_defol" variable counts
the number of trees recording their maximum defoliation in a given year. The

14

"mean_gsi" and "mean_ngsi" variables provide averages of these indices across all available trees. Finally, the "outbreak_status" column shows whether an outbreak event is inferred ("outbreak") or not ("not_obr").

The default plotting function to visualize results form `outbreak()` creates a three-panel graph showing the mean site-level chronology, the percent of trees recording a defoliation, and the sample depth over time (Figure 2).

```
plot_outbreak(dmj_obr, disp_index = "mean_gsi")
```

Inferred outbreak events are shown in the top panel of Figure 2 as the filled-in spaces. Users can change the time series in this panel with the "disp_index" parameter, choosing between the mean NGSI (the default) or GSI.

A summary table of the inferred outbreak events is generated by the `outbreak_stats()` function (Table 2).

```
dmj_obr_stats <- outbreak_stats(dmj_obr)
```

The output table from `outbreak_stats()` shows the start and end years of each event, along with their corresponding duration, the number and percent of trees in defoliation at the start of the event, the the number of trees recording the outbreak event (in part or full), the minimums of the mean GSI and NGSI indices, and two points of "peak defoliation." The first, "peak_outbreak_year," is the year in which the greatest number of trees is recording the outbreak, and the second, "peak_defol_year," is the year in with the lowest average NGSI. These output variables provide a array of options for assessing basic metrics of the insect defoliation regime, including interval analyses:

```
## calculate mean return interval for outbreak start years
intervals <- diff(dmj_obr_stats$start)
mean(intervals) ## 51.2
```

15

6. Conclusions

- Encourage use of `dfoliatR` over OUTBREAK – more acruate, greater user control, more easily conduct sensitivity tests to evaluate paramters, data, and ecological patterns
- Facilitates high-levelr statistical analyses by easily transfering to other libraries (MASS, etc.)
- A community resource that can be easily maintained and updated into the future.

References

Altman, J., Fibich, P., Dolezal, J., Aakala, T., 2014. TRADER: A package for Tree Ring Analysis of Disturbance Events in R. Dendrochronologia 32, 107–112. doi:10.1016/j.dendro.2014.01.004

Bunn, A.G., 2008. A dendrochronology program library in R (dplR). Dendrochronologia 26, 115–124. doi:10.1016/j.dendro.2008.01.002

Bunn, A.G., 2010. Statistical and visual crossdating in R using the dplR library. Dendrochronologia 28, 251–258. doi:10.1016/j.dendro.2009.12.001

Cook, E.R., Holmes, R.L., 1996. Guide for computer program arstan, The international tree-ring data bank program library version.

16

Cook, E.R., Krusic, P.J., 2004. The North American Drought Atlas.

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

Holmes, R.L., Swetnam, T.W., 1986. Dendroecology program library: Program outbreak user's manual. Laboratory of Tree-Ring Research, University of Arizona, Tucson.

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. doi:10.1016/j.dendro.2018.01.005

Lara, W., Bravo, F., Sierra, C.A., 2015. MeasuRing: An R package to measure tree-ring widths from scanned images. Dendrochronologia 34, 43–50. doi:10.1016/j.dendro.2015.04.002

Lynch, A.M., 2012. What Tree-Ring Reconstruction Tells Us about Conifer Defoliator Outbreaks, in: Barbosa, P., Letourneau, D.K., Agrawal, A.A. (Eds.), Insect Outbreaks Revisited. Blackwell Publishing Ltd, pp. 126–154. doi:10.1002/9781118295205.ch7

Maaten-Theunissen, M. van der, Maaten, E. van der, Bouriaud, O., 2015. PointRes: An R package to analyze pointer years and components of resilience. Dendrochronologia 35, 34–38. doi:10.1016/j.dendro.2015.05.006

Malevich, S.B., Guiterman, C.H., Margolis, E.Q., 2018. burnr: Fire history analysis and graphics in R. Dendrochronologia 49, 9–15. doi:10.1016/j.dendro.2018.02.005

R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ryerson, D.E., Swetnam, T.W., Lynch, A.M., 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. doi:10.1139/X03-026

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

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

Swetnam, T., Allen, C., Betancourt, J., 1999. Applied historical ecology:

⁴³⁴ using the past to manage for the future. Ecological applications 9, 1189–1206.

⁴³⁵ Swetnam, T.W., Lynch, A.M., 1989. A tree-ring reconstruction of western

⁴³⁶ spruce budworm history in the southern Rocky Mountains. Forest Science 35,

⁴³⁷ 962–986. doi:10.1093/forestscience/35.4.962

⁴³⁸ Swetnam, T.W., Lynch, A.M., 1993. Multicentury, Regional-Scale Patterns

⁴³⁹ of Western Spruce Budworm Outbreaks. Ecological Monographs 63, 399–424.

⁴⁴⁰ Swetnam, T.W., Thompson, M.A., Sutherland, E.K., 1985. Using den-

⁴⁴¹ drochronology to measure radial growth of defoliated trees.

⁴⁴² Wickham, H., 2016. Ggplot2: Elegant graphics for data analysis. Springer.

⁴⁴³ Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François,

⁴⁴⁴ R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.,

⁴⁴⁵ Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V.,

⁴⁴⁶ Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to

⁴⁴⁷ the tidyverse. Journal of Open Source Software 4, 1686. doi:10.21105/joss.01686

⁴⁴⁸ Wickham, H., François, R., Henry, L., Müller, K., 2020. Dplyr: A grammar

⁴⁴⁹ of data manipulation.

⁴⁵⁰ Zang, C., Biondi, F., 2015. Treeclim: An R package for the numerical cali-

⁴⁵¹ bration of proxy-climate relationships. Ecography 38, 431–436. doi:10.1111/ecog.01335

18

Tables & Figures

Table 1: Summary statistics for individual tree defoliation events.

series	first	last	years	num_events	tot_years	mean_duration
DMJ01	1810	1996	187	4	50	12
DMJ02	1750	1996	247	6	75	12
DMJ03	1830	1996	167	4	42	10
DMJ04	1720	1996	277	8	90	11
DMJ06	1700	1996	297	6	81	14
DMJ07	1710	1996	287	7	96	14
DMJ11	1900	1997	98	2	19	10
DMJ14	1675	1996	322	9	112	12
DMJ15	1730	1996	267	4	64	16
DMJ16	1746	1996	251	7	86	12
DMJ17	1733	1996	264	6	85	14
DMJ22	1720	1996	277	5	56	11
DMJ23	1675	1997	323	8	96	12
DMJ24	1895	1996	102	3	28	9
DMJ25	1680	1996	317	5	70	14
DMJ26	1700	1996	297	6	79	13
DMJ27	1710	1996	287	3	34	11

Table 2: Summary statistics for inferred outbreak events.

| start	end	duration	num_trees_start	perc_trees_start	peak_outbreak_year	num_trees_outbreak	peak_defol_year	min_gsi	mean_ngsi
1680	1699	20	1	33.3	1690	3	1692	0.121	-2.407
1753	1769	17	6	46.2	1754	7	1755	0.343	-1.611
1825	1840	16	11	78.6	1831	12	1826	0.500	-1.304
1849	1865	17	7	46.7	1852	13	1853	0.252	-1.994
1881	1895	15	8	53.3	1886	14	1885	0.262	-1.945
1959	1970	12	7	41.2	1960	15	1965	0.328	-1.830
1987	1996	10	9	52.9	1989	15	1995	0.378	-1.640
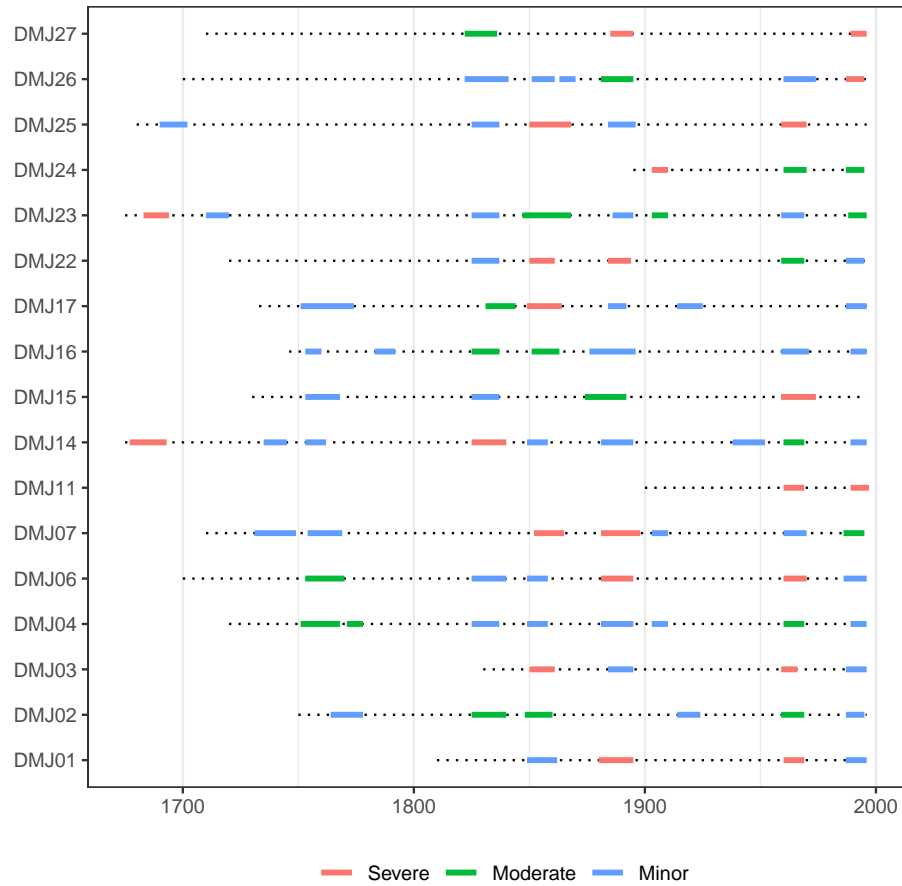
Figure 1: Default dfoliatR graphics for individual trees, produced by the 'plot_defol()' function. Dotted lines represent the series length for each tree and colored segments show periods of defoliation. Users can define the Cut-off values to determine severe-moderate-minor defoliation intensities, and use standard ggplot2 graphics parameters to adjust styles and themes.
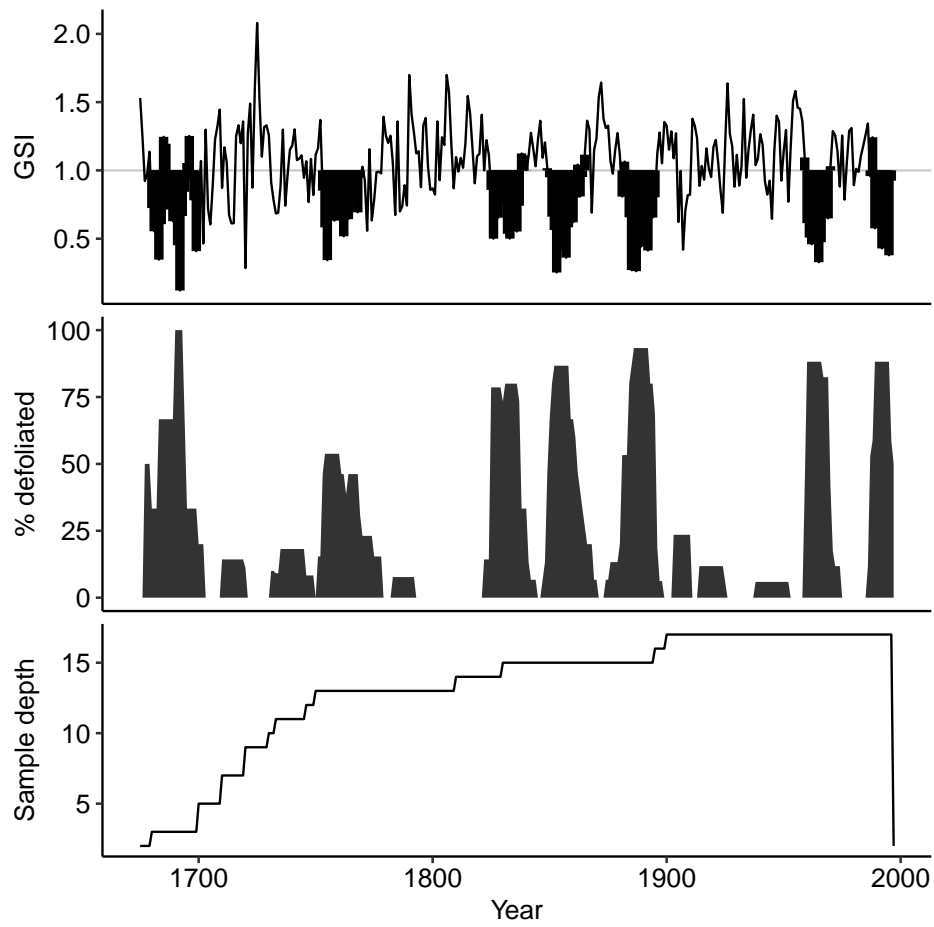
Figure 2: Default graphic to show outbreak events at the DMJ site. Top panel shows the mean GSI through time, with inferred outbreak events filled. The middle panel shows the percent of trees defoliated, and the bottom panel shows the number of trees through time.
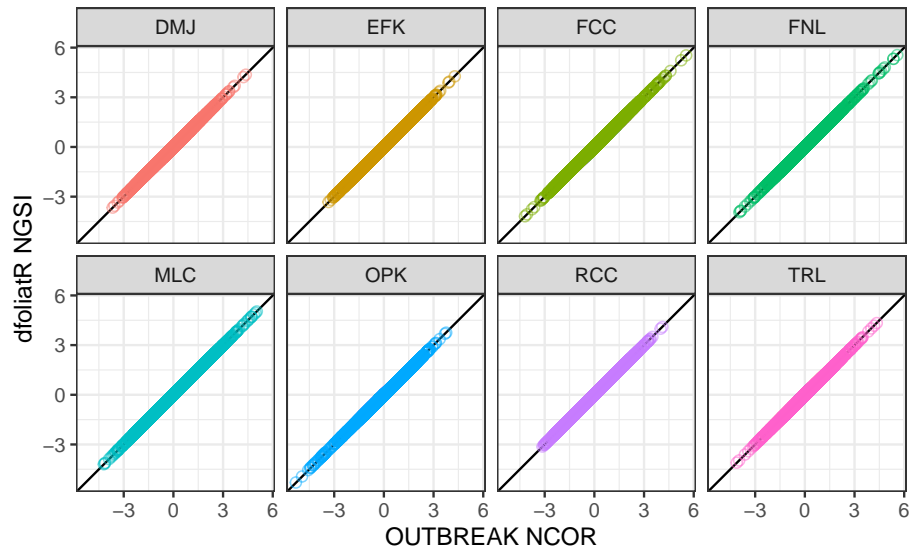
Figure 3: Comparison of disturbance indices between dfoliatR and OUTBREAK. Diagonal 1:1 lines show equal values.