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

Christopher H. Guiterman∗,a,b, Ann M. Lyncha,c, Jodi N. Axelsonc

*aLaboratory of Tree-Ring Research, University of Arizona, 1215 E Lowell St. Box 210045,*

*Tucson, AZ, 85721*

*bThree Pines Forest Research, LLC, PO Box 225, Etna, NH, 03750*

*cU.S. Forest Service, Rocky Mountain Research Station, 1215 E Lowell St. Box 210045,*

*Tucson, AZ, 85721*

*dDept of Environmental Science, Policy & Management, University of California, Berkeley, Berkeley, CA 94720*

1

2

3

4

5

6

7

8

9

10

**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 pro- gram OUTBREAK, and builds on existing resources in the R computing en- vironment. ‘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 individ- ual 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.

11

12

13

14

15

16

17

18

19

20

21

22

23

24

*Key words:* Dendroecology, spruce budworm, Choristoneura, pandora moth,

Coloradia pandora Blake, larch-bud-moth

25

26

∗Corresponding Author

*Email addresses:* [chguiterman@email.arizona.edu](mailto:chguiterman@email.arizona.edu) (Christopher H. Guiterman),

[lyncha@email.arizona.edu](mailto:lyncha@email.arizona.edu) (Ann M. Lynch), [jodi.axelson@berkeley.edu](mailto:jodi.axelson@berkeley.edu) (Jodi N. Axelson)

*Preprint submitted to Tree-Ring Research*

*April 22, 2020*

**1. Introduction**

27

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 sup- pressed growth by subtracting a detrended and standardized climate series (a “control” chronology) from host individual-tree detrended and standardized ra- dial 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

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

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 in- herently 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 analy- ses. 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 com- puting 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 analyti- cal tests (Jevšenak and Levanič, 2018; Zang and Biondi, 2015). Tools for assess- ing 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 ro- bust 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).

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

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 and on the package website (h[ttps://chguiterman.github.io/dfoliatR/),](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.](https://github.com/chguiterman/dfoliatR_paper)

88

89

90

91

92

93

94

95

96

97

**2. Overview of the software**

98

The dfoliatR library requires two sets of tree-ring data to infer defoliation

and outbreak events:

99

100

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

101

102

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)](#_bookmark2) and then com- posite across multiple trees to identify outbreak events (Figure [2).](#_bookmark3)

103

104

105

106

107

108

109

110

*2.1. Identifying Defoliation of Trees*

The defoliate\_trees() function is the point of entry to the dfoliatR li- brary. It performs two processes, removing climate-related growth signals from the host-tree series and identifying tree-level defoliation events. The climatic or

111

4

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 gen- erates 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

GSI = H *−* (NH *−* NH)

*σ*H

*σ*NH

(1)

*i i*

*i*

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 chronolo- gies are brought to common variance by scaling the non-host chronology by its

mean (NH *≈* 1*.*0) and multiplying by the ratio of host and non-host standard

112

113

114

115

116

deviations ( *σ*H ), which approximates the variance of the host tree series.

117

*σ*NH

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 for a single year within the departure se- quence. 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 al- lowed one positive excursion on each side of the year of maximum departure. Duration is assessed across the entire sequence that includes up to two posi- tive excursions. The default duration is eight years, as is commonly used in WSBW studies (Swetnam and Lynch, 1989). Different species of defoliation in- sects 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. Like OUTBREAK, users are provided an option to omit the duration threshold in cases of an ongo- ing defoliation or outbreak event. This applies to negative growth sequences at

the recent end of the tree series, and should be used if the user has direct knowl-

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

5

edge of defoliation at the site during the sampling campaign. The advantage of allowing potentially short, series-end events 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.

Diverging from OUTBREAK, dfoliatR allows users to extend defoliation

events by bridging between sequential events. In cases where two defoliation events are separated by a single year, bridging will link them into a single event (**Show figure?**). The paramter was added to dfoliatR during the testing phase of development, when we realized that OUTBREAK deliberately omits sequential, or back-to-back events, even when these events surpass thresholds in duration and magnitude. Instead, OUTBREAK will select the one sequen- tial event with the greatest negative departure year. In every case we assessed (described below) we felt that the defoliation should be recorded on individual trees. Due in large part to reconstructions using OUTBREAK (see papers cited by Lynch, 2012), we now know considerably more about forest defoliator out- break 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 situ- ations be considered a single event. This is particularly relevant to WSBW and spruce budworm (*C. fumiferana*), for which regimes have been reconstructed for several geographic areas and for which researchers and forest health ex- perts 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. We urge caution using the op- tion, however, because it may not be appropriate for all studied insects, such as in situations where impacted stands barely recover from one outbreak before another begins, like pine processionary caterpillars (*Thaumetopoea pityocampa* (Lepidoptera: Thaumetopoeidae)) (**Carus 2004, 2009**).

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

6

*2.2. Inferring Outbreak Events*

Defoliation of one or a few trees does not constitute an outbreak. To de- termine 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 defolia- tion 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.

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

**3. Evaluation**

183

*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

184

185

186

187

188

189

190

191

192

7

we used event thresholds of -1.28 normalized index, 8 years duration, and al- lowed 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 generated between the programs transferred into differences between dfoliatR and OUTBREAK. In the end, we chose to only use the stan- dardization 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.](https://github.com/chguiterman/dfoliatR_paper)

[com/chguiterman/dfoliatR\_paper.](https://github.com/chguiterman/dfoliatR_paper)

193

194

195

196

197

198

199

200

201

202

203

* 1. *Findings*

Accross the 43,280 ring-width indices from 222 trees included in our test, we found that dfoliatR and OUTBREAK compute identical growth suppression indices at 0.000 precision (Figure [3).](#_bookmark4) We expected this outcome because both programs employ Equation 1. In all, the programs identified 11,530 years of defoliation on individual trees. 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

* + - *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” param- eter 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 eval- uated as with any other possible defoliation, but the duration threshold is

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

8

omitted. Each of the 13 events included in these differences did not meet the “max\_reduction” parameter (-1.28 NGSI) in dfoliatR and were ex- cluded. 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, mak- ing them into single, long events. We felt that this was ecologically justi- fied, especially for studies of WSBW, because the insect will occasionally maintain minor defoliation of individuals for decades even while not in an outbreak phase.

*Undetermined differences* (22%) occurred in cases where OUTBREAK omitted events without clear cause that dfoliatR correctly identified as defoliations.

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

223

224

225

226

227

228

229

230

231

•

232

233

234

235

236

237

238

239

240

241

•

242

243

244

•

245

246

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.

At the site-level, OUTBREAK and dfoliatR produce similiar time series of

247

248

249

250

251

9

percent trees defoliated (Figure [4),](#_bookmark5) which forms the basis for inferring outbreak occurance, intensity, and duration. Nearly all of the differences between the programs include either more trees or a longer duration of inferred outbreak by

dfoliatR. These arrise from the difference sin how dfoliatR identifies defoli-

ation events for individual trees, most often adding events omitted by OUT- BREAK. These differences translate to greater numbers of trees in defoliation during certain events, or events recorded by single trees during periods of low sample depth. In accordance with our conclusion that dfoliatR did a better job with identifying defoliation events on indivudla trees, we feel that dfoliatR did a better job in identifying OUTBREAK events and associated statistics at the site level.

252

253

254

255

256

257

258

259

260

261

262

**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-](https://cran.r-project.org/) [project.org/).](https://cran.r-project.org/) To install dfoliatR from CRAN use

263

264

265

266

**install.packages**("dfoliatR")

In each R session, dfoliatR can be loaded via

267

**library**(dfoliatR)

Development versions of dfoliatR are available on GitHub and installed

using the devtools library,

268

269

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

Issues, bug reports, and ideas for improving dfoliatR can be posted to [https://github.com/chguiterman/dfoliatR/issues.](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.

270

271

272

273

274

10

**5. Example Usage**

275

In dfoliatR we provide two sets of tree-ring data to aid users in exploring the functions, graphics, and outputs. Each set consists of individual host-tree series,

standardized using 50-year splines with a 50% frequency response, and a local non-host chronology. The non-host ring-width data were standardized using 150-year splines with a 50% frequency response. The non-host chronologies were averaged following standard procedures (Speer, 2010). Host trees from Demijohn Peak (DMJ; 2902 m asl) in the San Juan Mountains of southern

Colorado include Douglas-fir (*Pseudotsuga menziesii*) compared against a local

non-host ponderosa pine (*Pinus ponderosa*) chronology (Ryerson et al., 2003). The East Fork site (EF; 2580 m asl) in the Jemez Mountains of northcentral New Mexico include Douglas-fir and white fire (*Abies concolor* ) host trees and a ponderosa pine non-host chronology (Swetnam and Lynch, 1993).

276

277

278

279

280

281

282

283

284

285

286

287

*5.1. Tree-Level Defoliation Events*

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

288

289

290

291

292

293

294

11

*## 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)"*

The result of running

frame with five variables:

defoliate\_trees() is long-format (stacked) data

“year”, “series”, “gsi”, “ngsi”, and “defol\_status.”

295

296

The “gsi” variable is the “corrected” tree-level growth suppression index. It was normalized or scaled into z-scores to produce “ngsi”, which is used to identify de- foliation events. The “defol\_status” column provides a set of character strings, “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).](#_bookmark0)

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

12

*## 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)

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).](#_bookmark2) 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.

314

315

316

317

318

319

320

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().

321

322

323

324

325

*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

326

327

328

329

13

**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/* **library**(ggplot2) **plot\_defol**(dmj\_defol) **+**

**scale\_color\_manual**(values = **c**("red", "orange", "purple"))

**defol\_stats**(dmj\_defol)

**View**(dmj\_defol)

control thresholds in the number and proportions of trees recording a defoliation

event.

330

331

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 “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).](#_bookmark3)

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

**plot\_outbreak**(dmj\_obr, disp\_index = "mean\_gsi")

Inferred outbreak events are shown in the top panel of Figure [2](#_bookmark3) 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).](#_bookmark1)

347

348

349

350

351

14

*## Use the defol object "dmj\_defol" created above*

dmj\_obr <- **outbreak**(dmj\_defol,

filter\_min\_series = 3,

filter\_min\_defol = 1,

filter\_perc = 25)

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:

352

353

354

355

356

357

358

359

360

**6. Conclusions**

361

* 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.

362

363

364

365

366

367

368

**Acknowledgments**

369

We are grateful to late Richard Holmes of the Laboratory of Tree-Ring Re- search (LTRR) for his achievements in developing OUTBREAK and many other tree-ring software programs, and we thank Thomas Swetnam for his insights

and encouragement. The dfoliatR project was supported by the Agnes Haury

370

371

372

373

15

*## calculate mean return interval for outbreak start years*

intervals <- **diff**(dmj\_obr\_stats**$**start) **mean**(intervals) *## 51.2*

Visiting Scholars Fellowship at the LTRR and by the University of California Berkeley. Additional support was provided by the LTTR and the United States Forest Service Rocky Mountain Research Station. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement

by the U.S. Government.

374

375

376

377

378

**References**

379

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](https://doi.org/10.1016/j.dendro.2014.01.004)

Bunn, A.G., 2008. A dendrochronology program library in R (dplR). Den- drochronologia 26, 115–124. [doi:10.1016/j.dendro.2008.01.002](https://doi.org/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](https://doi.org/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.

Cook, E.R., Krusic, P.J., 2004. The North American Drought Atlas. Guiterman, C., Lynch, A., Axelson, J., 2020. DfoliatR: Detection and anal-

ysis of insect defoliation signals in tree rings.

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

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

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

16

Insect Outbreaks Revisited. Blackwell Publishing Ltd, pp. 126–154. [doi:10.1002/9781118295205.ch7](https://doi.org/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](https://doi.org/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](https://doi.org/10.1016/j.dendro.2018.02.005) R Core Team, 2019. R: A language and environment for statistical comput-

ing. R Foundation for Statistical Computing, Vienna, Austria.

Ryerson, D.E., Swetnam, T.W., Lynch, A.M., 2003. A tree-ring reconstruc- tion of western spruce budworm outbreaks in the San Juan Mountains, Col- orado, U.S.A. Canadian Journal of Forest Research 33, 1010–1028. doi:[10.1139/X03-](https://doi.org/10.1139/X03-026) [026](https://doi.org/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: using the past to manage for the future. Ecological applications 9, 1189–1206.

Swetnam, T.W., Lynch, A.M., 1989. A tree-ring reconstruction of western spruce budworm history in the southern Rocky Mountains. Forest Science 35, 962–986. [doi:10.1093/forestscience/35.4.962](https://doi.org/10.1093/forestscience/35.4.962)

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

Swetnam, T.W., Thompson, M.A., Sutherland, E.K., 1985. Using den- drochronology to measure radial growth of defoliated trees.

Wickham, H., 2016. Ggplot2: Elegant graphics for data analysis. Springer. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.,

Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to

the tidyverse. Journal of Open Source Software 4, 1686. [doi:10.21105/joss.01686](https://doi.org/10.21105/joss.01686)

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

17

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

435 of data manipulation.

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

437 bration of proxy-climate relationships. Ecography 38, 431–436. doi:10.[1111/ecog.01335](https://doi.org/10.1111/ecog.01335)

18

438 **Tables & Figures**

Table 1: Summary statistics for individual tree defoliation events.

series

first

last

years

num\_events

tot\_years

mean\_duration

DMJ01 DMJ02 DMJ03 DMJ04 DMJ06

DMJ07 DMJ11 DMJ14 DMJ15 DMJ16

DMJ17 DMJ22 DMJ23 DMJ24 DMJ25

DMJ26

DMJ27

1810

1750

1830

1720

1700

1710

1900

1675

1730

1746

1733

1720

1675

1895

1680

1700

1710

1996

1996

1996

1996

1996

1996

1997

1996

1996

1996

1996

1996

1997

1996

1996

1996

1996

187

247

167

277

297

287

98

322

267

251

264

277

323

102

317

297

287

4

6

4

8

6

7

2

9

4

7

6

5

8

3

5

6

3

50

75

42

90

81

96

19

112

64

86

85

56

96

28

70

79

34

12

12

10

11

14

14

10

12

16

12

14

11

12

9

14

13

11

19

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

1753

1825

1849

1881

1959

1987

1699

1769

1840

1865

1895

1970

1996

20

17

16

17

15

12

10

1

6

11

7

8

7

9

33.3

46.2

78.6

46.7

53.3

41.2

52.9

1690

1754

1831

1852

1886

1960

1989

3

7

12

13

14

15

15

1692

1755

1826

1853

1885

1965

1995

0.121

0.343

0.500

0.252

0.262

0.328

0.378

-2.407

-1.611

-1.304

-1.994

-1.945

-1.830

-1.640

20

DMJ27

DMJ26

DMJ25

DMJ24

DMJ23

DMJ22

DMJ17

DMJ16

DMJ15

DMJ14

DMJ11

DMJ07

DMJ06

DMJ04

DMJ03

DMJ02

DMJ01

1700

1800

1900

2000

Severe

Moderate

Minor

Figure 1: Default dfoliatR graphics for individual trees, produced by the ‘plot\_defol()’ func- tion. 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 defoli-

ation intensities, and use standard ggplot2 graphics parameters to adjust styles and themes.

21

2.0

1.5

1.0

0.5

100

75

50

25

0

15

10

5

1700

1800

1900

2000

Year

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.

22

Sample depth

% defoliated

GSI

DMJ

EFK

FCC

FNL

6

3

0

−3

MLC

OPK

RCC

TRL

6

3

0

−3

−3

0

3

6

−3

0

3

6

−3

0

3

6

−3

0

3

6

OUTBREAK NCOR

Figure 3: Comparison of disturbance indices between dfoliatR and OUTBREAK. Diagonal

lines show equal values (1:1).

23

dfoliatR NGSI

dfoliatR

OUTBREAK

DMJ

EFK

100

75

50

25

0

FCC

FNL

100

75

50

25

0

MLC

OPK

100

75

50

25

0

RCC

TRL

100

75

50

25

0

1600

1700

1800

1900

2000

1600

1700

1800

1900

2000

Year

Figure 4: Comparison of reconstructed western spruce budworm outbreaks computed by dfo-

liatR and OUTBREAK. Input parameters were identical between programs.

24

Percent trees defoliated