Forest Ecology and Management 425 (2018) 59–67

Contents lists available at ScienceDirect

Forest Ecology and Management

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

Testing the efficacy of tree-ring methods for detecting past disturbances

Volodymyr Trotsiuk[a,b,c,*], Neil Pederson[d], Daniel L. Druckenbrod[e], David A. Orwig[d], Daniel A. Bishop[d,f,g], Audrey Barker-Plotkin[d], Shawn Fraver[h], Dario Martin-Benito[i,j]

[a] *Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Kamycka 129, Praha6–Suchdol, Prague 16521, Czech Republic*
[b] *Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland*
[c] *Institute of Agricultural Sciences, ETH Zurich, Universitätsstrasse 2, CH-8092 Zurich, Switzerland*
[d] *Harvard Forest, Harvard University, 324 North Main St., Petersham, MA 01366, United States*
[e] *Department of Geological, Environmental and Marine Sciences, Rider University, 2083 Lawrenceville Road, Lawrenceville, NJ 08648-3099, United States*
[f] *Tree Ring Laboratory of Lamont-Doherty Earth Observatory, POB 1000 Route 9w, Palisades, NY 10964, United States*
[g] *Division of Biology and Paleo Environment, Department of Earth and Environmental Sciences, Columbia University, New York, NY 10964, United States*
[h] *School of Forest Resources, University of Maine, 5755 Nutting Hall, Orono, ME 04469, United States*
[i] *Forest Research Centre (INIA-CIFOR), Ctra La Coruña km 7.5, 28040 Madrid, Spain*
[j] *Forest Ecology, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zurich, Universitätsstrasse 16, CH-8092 Zurich, Switzerland*

ARTICLE INFO

Keywords:
Absolute-increase
Boundary-line
Competition
Forest development
Forest dynamics
Radial-growth averaging
Time-series analysis
Release

ABSTRACT

The retrospective study of abrupt and sustained increases in the radial growth of trees (hereinafter 'releases') by tree-ring analysis is an approach widely used for reconstructing past forest disturbances. Despite the range of dendrochronological methods used for release-detection, a lack of in-depth comparison between them can lead researchers to question which method to use and, potentially, increases the uncertainties of disturbance histories derived with different methods.

Here, we investigate the efficacy and sensitivity of four widely used release detection methods using tree-ring width series and complete long-term inventories of forest stands with known disturbances. We used support vector machine (SVM) analysis trained on long-term forest census data to estimate the likelihood that *Acer rubrum* trees experiencing reductions in competition show releases in their tree-ring widths. We compare methods performance at the tree and stand level, followed by evaluation of method sensitivity to changes in their parameters and settings.

Disturbance detection methods agreed with 60–76% of the SVM-identified growth releases under high canopy disturbance and 80–94% in a forest with canopy disturbance of low severity and frequency. The median competition index change (CIC) of trees identified as being released differed more than two-fold between methods, from −0.33 (radial-growth averaging) to −0.68 (time-series). False positives (type I error) were more common in forests with low severity disturbance, whereas false negatives (type II error) occurred more often in forests with high severity disturbance. Sensitivity analysis indicated that reductions of the detection threshold and the length of the time window significantly increased detected stand-level disturbance severity across all methods.

Radial-growth averaging and absolute-increase methods had lower levels of type I and II error in detecting disturbance events with our datasets. Parameter settings play a key role in the accuracy of reconstructing disturbance history regardless of the method. Time-series and radial-growth averaging methods require the least amount of *a priori* information, but only the time-series method quantified the subsequent growth increment related to a reduction in competition. Finally, we recommend yearly binning of releases using a kernel density estimation function to identify local maxima indicating disturbance. Kernel density estimation improves reconstructions of forest history and, thus, will further our understanding of past forest dynamics.

* Corresponding author at: Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Kamycka 129, 165 21 Prague, Czech Republic.
E-mail addresses: trotsiuk@fld.czu.cz (V. Trotsiuk), neilpederson@fas.harvard.edu (N. Pederson), ddruckenbrod@rider.edu (D.L. Druckenbrod), orwig@fas.harvard.edu (D.A. Orwig), dbishop@ldeo.columbia.edu (D.A. Bishop), aabarker@fas.harvard.edu (A. Barker-Plotkin), shawn.fraver@maine.edu (S. Fraver), dmartin@inia.es (D. Martin-Benito).

https://doi.org/10.1016/j.foreco.2018.05.045
Received 27 March 2018; Received in revised form 17 May 2018; Accepted 19 May 2018
0378-1127/

1. Terminology

1.1. Tree level

Window length parameter – number of consecutive years used to calculate average growth for radial-growth averaging disturbance detection methods prior to and following a potential disturbance event. For the time-series approach, window length is the number of years used to calculate residuals.

Threshold parameter – minimum change in radial growth (absolute or relative depending on method) that must be exceeded for an increase in radial growth to be defined as a growth release (i.e., an abrupt increase in radial growth).

False positive – an event is detected by the dendrochronological method, but not by support vector machine analysis based on changes in the competition index before and after the event (type I error).

False negative – an event is detected by support vector machine analysis based on changes in competition index before and after the event, but not by the method (type II error).

1.2. Stand level

Disturbance severity – the proportion of trees responding to disturbance standardised by moving kernel density estimation function. Severity is related to the number of trees with a detected event and their temporal synchrony. Our definition is adopted and modified from Pickett and White (1985a).

Peak disturbance year – year with the greatest estimated disturbance severity for a specific event.

Accuracy – agreement between the severity of the disturbance identified by SVM analysis and that of the peak disturbance estimated from tree rings for the same event.

Precision – temporal agreement between the disturbance identified by SVM analysis and peak disturbance year estimated from tree rings.

2. Introduction

Reconstruction of past forest disturbances reveals the dynamics that have led to current forest composition, structure, and function. Tree-ring reconstructions of past disturbances surpass the length of time in contemporary forest inventories and, quite often, the era of local written records. Importantly, crossdating tree rings assigns a precise calendar year to each ring so that past centuries of forest dynamics can be investigated with annual resolution (Douglass, 1920). Increased precision in dating past disturbances allows ecologists a greater chance of correctly identifying agents of disturbance (Black et al., 2016). Relative to the lifespan of a tree, disturbance events are rapid processes that occur over the course of hours (e.g., windstorm) to months, seasons, or years (e.g., drought) (Pickett and White, 1985). Documenting disturbance with annual resolution, over centuries, and from the tree to continental scales is a powerful method that can shed much light on the mechanisms driving forest dynamics.

Almost a century after the publication of a pioneering paper on the potential identification of past forest disturbance from tree rings (Marshall, 1927), a number of tree-ring-based disturbance-detection methods have been developed to differentiate disturbance-induced changes in tree growth from those caused by life-history traits, biometry, stresses, or climate variability. Briefly, an abrupt, large, and sustained increase in tree-ring width (radial growth) is inferred to be a release from tree-to-tree competition and is taken as evidence of past canopy disturbance (Lorimer, 1980). Disturbance detection methods were first formalized in the mid- to late-1980s so the frequency and severity of disturbance could be objectively quantified through time and synthesized into time series of canopy disturbance (Lorimer, 1985; Lorimer and Frelich, 1989). A series of methods were developed soon afterward that either built directly upon these original methods (Nowacki and Abrams, 1997; Fraver and White, 2005) or used new approaches (Black and Abrams, 2003; Druckenbrod, 2005; Druckenbrod et al., 2013; Lee et al., 2017). The growing interest in studying old-growth forests, ecological restoration, and forest conservation biology increased the use of these methods.

Several methods of disturbance detection have been compiled into the R package TRADER (Altman et al., 2014). The creation of TRADER allows for the opportunity to simultaneously compare several methods and modify the parameters and thresholds for each method. Yet, faced with the diversity of approaches and parameters, researchers are likely to ponder, "*How should one choose which method, parameters, and thresholds to use given particular research goals and specific forest conditions*?"

Developers of the various release-detection methods have independently discussed the strengths and weaknesses of their specific approach (Lorimer and Frelich, 1989; Black and Abrams, 2003, 2004; Fraver and White 2005; Druckenbrod et al., 2013). A few studies have examined the sensitivity to varying parameters and thresholds of the growth-averaging method (Rubino and McCarthy, 2004; Bouriaud and Popa, 2007; Stan and Daniels, 2010). To date, however, no work has provided a detailed comparison of the performance of the most widely used detection methods with forest inventory datasets of controlled or observed records of disturbance. A rigorous examination of these methods is critical to correctly identify and correctly date past disturbances (Rubino and McCarthy, 2004; Bouriaud and Popa, 2007; Copenheaver et al., 2009; Altman et al., 2014; McEwan et al., 2014; Pederson et al., 2014; Šamonil et al., 2015).

Our primary objective was to analyse the performance of four widely used disturbance-detection methods in a forest subjected to an experimentally-induced disturbance and a forest with minimal canopy disturbance. These four methods are: radial-growth averaging (Lorimer and Frelich, 1989; Nowacki and Abrams, 1997), boundary line (Black and Abrams, 2003, 2004), absolute increase (Fraver and White, 2005), and time series (Druckenbrod, 2005; Druckenbrod et al., 2013). Performance was assessed by a method's ability to detect a disturbance of known timing and magnitude. Our secondary objectives were to: (i) investigate the efficacy of these methods in reconstructing the timing and severity of disturbance at the stand level and (ii) gain insight into the sensitivity of each method to adjustments in their temporal parameters and growth thresholds. Our study will provide guidance for future tree-ring studies with respect to method selection and interpretation of results.

3. Materials and methods

3.1. Study sites

To examine how each method performed in forests with differing canopy disturbance, we used repeated forest census data and tree rings from two nearby forest stands. First, for a forest with severe disturbance we examined trees from an experiment designed to mimic the damage in upland forests caused by a hurricane (Cooper-Ellis et al., 1999). To examine how methods performed in forests with little to no canopy disturbance, we used each method on trees in a 3-ha study plot with repeated forest measurements since 1969 and no significant canopy disturbance (Eisen and Plotkin, 2015).

3.1.1. High severity disturbance forest

The hurricane manipulation experiment ("Hurricane pulldown") was located at the Harvard Forest, Petersham, Massachusetts, USA (72.20 °N, 42.49 °W, 300–315 m a.s.l.) in a forest dominated by red maple (*Acer rubrum*) and northern red oak (*Quercus rubra*) (Cooper-Ellis et al., 1999; Plotkin et al. 2013). The forest originated following a clear-cut in 1915 (Harvard Forest Archives, *unpub. data*). All trees ≥5 cm diameter at breast height (DBH) were tagged, spatially mapped and recorded as live or dead during inventory surveys (1990 before, and after the experiment, 1996, 2000, 2005 and 2010). In early October 1990, 276 trees were toppled in a northwesterly direction to effectively simulate the disturbance caused by the 1938 hurricane in New England. The Hurricane pulldown occurred over 0.8 ha of forest and was separated from a 0.6 ha control forest by a 30-m forest buffer. Similar to the impact of the 1938 hurricane, surveys immediately following the toppling of trees indicated that 80% of the canopy trees and two-thirds of all trees ≥ 5 cm DBH were damaged (Rowlands, 1941; Foster, 1988). In 2009, a total of 57 *Acer rubrum* trees from within the hurricane experiment and the adjacent control forest were cored at approximately breast height (1.3 m) to determine how implementation of the Hurricane pulldown experiment affected the growth of surviving trees. Increment cores from the site revealed tree ages ranging from 42 to 95 years (median age = 79 years).

3.1.2. Low severity disturbance forests

Vegetation and tree-ring sampling were conducted in long-term monitoring plots at the "Lyford plot" in the Harvard Forest. Lyford plot is a second-growth mixed northern hardwood forest dominated by northern red oak and red maple with relatively little disturbance (Eisen and Plotkin, 2015) and is used here as a contrast to the high severity Hurricane pulldown. All trees ≥ 5 cm diameter at breast height (DBH) were tagged, spatially mapped and recorded as live or dead during inventory surveys (1969, 1975, 1991, 2001 and 2011) (Foster and Barker Plotkin, 1999). In 2011, trees in three Lyford plot subplots were cored following the nested design of: (i) trees > 10 cm DBH out to 13 m from plot center and (ii) trees > 20 cm DBH from 13 to 20 m from center. Tree ages from cores ranged from 26 to 152 years (ages at breast height; median age = 93 years).

3.1.3. Natural hurricane disturbance

Vegetation and tree-ring sampling were also conducted in the Harvard Forest tract ("Harvard tract") in Pisgah State Park, a mixed northern hardwood forest dominated by eastern hemlock (*Tsuga canadensis*) owned by the Harvard Forest. Pisgah State Park is approximately 36 km to the northwest of the Harvard Forest in southwest New Hampshire. The 1938 hurricane knocked down > 80% of the canopy trees in this tract of old-growth forest in Pisgah State Park (Rowlands, 1941; Foster, 1988), hereafter the 'Harvard tract'. Two tree-ring sampling plots were established in 2014 following the protocol used in the Lyford plot. Because the tract is old-growth, we added an additional nest to both plots where trees > 30 cm DBH and 20–30 m from plot center were surveyed and cored. Increment cores from the Harvard tract revealed 42–340 year old trees (median age = 93 years).

3.2. Tree core preparation

All cores were dried, sanded, cross-dated, and measured following standard dendrochronological methods (Stokes and Smiley, 1968). Each ring was measured to the nearest 0.001 mm and dating was verified with the program COFECHA (Holmes, 1983; Grissino-Mayer, 2001). A gypsy moth defoliation event (1981) was used as a marker ring. In the Hurricane pulldown experiment, dating control of all cores resulted in 15 crossdated trees from the control plot and 32 crossdated trees from the Hurricane pulldown plot. A total of 144 trees were analysed from the Lyford plot and 224 from the Harvard tract.

3.3. Disturbance detection methods

Disturbance detection methods used here can be divided into two broad groups: growth averaging (radial-growth averaging, boundary-line, absolute-increase) and time-series approach (time-series). Growth averaging methods involve comparing mean growth rates prior to and after any year *t* within an *a priori* number of years, hereafter 'window length', to determine if an abrupt and sustained increase in growth occurred after year *t*; further constraints are involved in the boundary-line and absolute-increase methods (Black and Abrams, 2003; Fraver and White, 2005). The time-series method identifies sequences of statistically extreme residual ring widths after accounting for the effects of any biological age trend and autocorrelation (Druckenbrod, 2005). All methods were originally designed and developed for various forest types or species in eastern North America, and are currently applied to a much wider range of forest conditions and species.

(i) Radial-growth averaging

Radial-growth averaging (GA) is one of the earliest developed and still commonly used methods and is based on running means of raw ring widths (Lorimer, 1980, 1985; Lorimer and Frelich, 1989). The original method averaged radial growth over the preceding 15-year period M_1 (including the target year *t*), and the average radial growth over the subsequent 15-year period, M_2 (excluding the target year *t*) to calculate the percentage growth change (PGC) for each annual ring as:

$$PGC = \frac{M_2 - M_1}{M_1} * 100 \tag{1}$$

The original percentage growth thresholds to detect growth releases in understory trees were ≥100% growth increase for a "major, sustained" release and 50–99% for a "moderate" release (Lorimer, 1985; Lorimer and Frelich, 1989). Later, the original window lengths were shortened to 10 years and the "moderate" release was reduced to a growth increase of 25% to derive disturbance history from old-aged canopy oak trees (Nowacki and Abrams, 1997). Different M_1 and M_2 window lengths, as well as growth thresholds, have been applied to meet species-specific or site-specific criteria (Rubino and McCarthy, 2004; Stan and Daniels, 2010). The default settings for the radial-growth averaging method in TRADER (Altman et al., 2014) are those proposed by Nowacki and Abrams (1997). Therefore, we use the TRADER default for the radial-growth averaging method.

(ii) Boundary-line

In the boundary-line (BL) method, the percentage growth change of each year for each tree (Eq. (1)) is scaled by its maximum potential observed growth, as defined by prior growth rates for that species growing at one or several locations (Black and Abrams, 2003). Growth pulses exceeding 20% of the prior growth boundary-line were classified as releases. The rationale of this method is that standardizing growth should account for the influence of site condition, species, size and tree age on the radial growth rates (Black and Abrams, 2003, 2004; Ziaco et al., 2012). Defining the boundary-line requires a large amount of data from a single species within similar site conditions, in some cases up to 50,000 radial increments, which may make the boundary line difficult to fit for certain stands and species.

(iii) Absolute-increase

While the two previous methods are based on the relative changes of ring-width averages, the absolute-increase (AI) method (Fraver and White, 2005) relies on the subtraction of the average pre-event growth rate (M_1) from the average post-event rate (M_2), using 10-year window lengths.

$$AI = M_2 - M_1 \tag{2}$$

The growth increase is determined as a release if the difference in growth rate exceeds a predetermined threshold for a given species. The method is meant to detect overhead canopy disturbances, making it similar to the 'major' releases referred to in radial-growth averaging and boundary-line methods. Fraver and White (2005) demonstrated the appropriate species-specific release threshold by testing it against empirical absolute increases found in different trees responding to dated canopy gaps. In cases where knowledge of species' growth potential is not available, they suggest selecting a threshold value equal to 1.25 times the standard deviation, or somewhat less than the 90th percentile, of all absolute increases. In the current study, we used the latter procedure (1.25 SD) to determine species-specific thresholds for the absolute increase method.

(iv) Time-series

Time-series (TS) analysis is central to the reconstruction of past climate using tree rings (Cook and Kairiukstis, 1990) but also for reconstructing past ecological changes (Druckenbrod, 2005). This release detection approach removes the long-term growth trend from a series, accounts for the autocorrelation present in width measurements of sequential tree-rings and uses intervention detection, enabling its release criteria to scale with a tree's growth rate similar to the boundary-line and absolute-increase methods. Release events are identified as sequences of unusually large, positive departures from autoregressive residuals over intervals of 9 to 30 years. Using Tukey's biweight mean as a robust estimate of location and scale, the detection criteria identifies any sequence with a scale > 3.29 (or 99.95% of the observations for a one-tailed analysis) as an outlier (Druckenbrod et al., 2013). A Hugershoff curve (Warren, 1980; Cook and Kairiukstis, 1990; Druckenbrod et al., 2018) is then fit at the start of the growth release. The flexibility of this curve captures both transient and sustained release events. Unlike growth averaging methods, this curve intervention detection approach allows time-series analysis to reconstruct not only the year of release but also the magnitude and duration of the subsequent growth change caused by the disturbance event (Rydval et al., 2015).

3.4. Analysis of disturbance methods

3.4.1. Tree level methods evaluation

To independently determine the likelihood of a growth release for each individual tree, we first used the census data to calculate the distance-weighted size competition index (CI) proposed by Hegyi (1974) as:

$$CI = \sum_{j=1}^{n} \left[\frac{\left(\frac{D_j}{D_i}\right)}{R_{ij}} \right] \tag{3}$$

where D_j is the DBH of competitor tree, D_i the DBH of focal tree, R_{ij} the distance between the neighbouring and focal trees, and n the number of trees included in the sample. The maximum radius for the competitor tree to be included was 10 m and the DBH ≥ 5 cm. CI was calculated in R library 'siplab' (García, 2014).

Competition index change for each tree (CIC) was estimated from inventory data as the difference of CI from two subsequent inventories as:

$$\Delta CIC_t = \frac{(CI_t - CI_{t-1})}{N} \tag{4}$$

where CI_t is the competition index at measurement t, CI_{t-1} the competition index at measurement preceding to t, N the number of years between two subsequent inventories (t and $t-1$). CIC was calculated for each tree and measurement year within the Hurricane pulldown (1990, 1996) and the Lyford plot (1991, 2001), excluding the trees in the buffer zone (7 m). Thus, negative values of CIC indicate that competition around the focus tree has decreased. Later, we applied support vector machine (SVM) analysis to CIC and diameter change of the trees at Hurricane pulldown and control plots at the individual tree level to empirically determine the likelihood that individual trees had responded to the reduction in competition after the experimental hurricane disturbance. We identified the optimal separating hyperplane (line) between the two classes by maximizing the margin (distance between nearest points and hyperplane) between the classes using the linear kernel fit and C-classification (binary classification). To do so, SVM maps the input vectors into a n-dimensional feature space (where n is number of features) to construct the linear decision surface (Cortes and Vapnik, 1995). In our case SVM operates with a linear kernel to separate the repeated measure data into two classes, simply trees that likely experienced a growth releases versus those that did not. SVM analysis was performed with R library 'e1071' (Meyer et al., 2015).

Trees identified as released or not released through SVM analysis were used as the standard for comparing the efficacy of the four disturbance detection methods. Efficacy at the tree level was evaluated by the (i) correct classification of trees identified as having been released through SVM analysis and (ii) timing of the detected event compared to the year of the known event. We considered two types of false detections: (i) a false positive (type I error): where SVM analysis did not show the tree having a significant change in competition, but a release detection method identified a release, and (ii) a false negative (type II error): where SVM analysis classified a tree as released, but a disturbance was not detected by the release detection method.

3.4.2. Stand-level methods evaluation

We recorded the timing and severity of each disturbance event identified by each method at the stand level. To overcome limitations of temporal resolution incurred by the decadal binning of annual disturbances, we fit a kernel density estimation (KDE) function to reconstructed disturbance histories to better characterize the timing and severity of disturbance events at the stand level. The moving KDE function was fit to 15-year windows, and the value of the fitted function for a particular year was extracted from the distribution at the midpoint. Values were standardized by the maximum value calculated by fitting the KDE to the normal distribution with 1 standard error and a 15-year window (0.28184). Using outcomes from the KDE analysis, we derived a disturbance severity, which incorporates both the proportion of trees showing response and synchrony in the timing of response. Thus, peaks in detected disturbances represent the standardized proportion of trees responding to a disturbance. These calculations were performed using the 'density' function in the 'pracma' R package (Borchers, 2017) and 'findpeaks' function in the 'quantmod' R package (Ryan, 2008).

We then conducted a sensitivity analysis to quantify how different window lengths and thresholds for each method influenced the stand-level reconstruction of disturbance history (Table 1). Thresholds varied depending on the method used. For the growth averaging methods, moving averages were calculated for window lengths between 5 and 15 years with a 1-year time step. For the time-series method, window lengths from 1 to 22 years were used to calculate residuals (Table 1). We used the R package TRADER (Altman et al., 2014) for growth averaging methods, and an executable program for the time-series approach (Druckenbrod et al., 2013) in Matlab (Mathworks, 2014). All subsequent analyses were done using R statistical software ver. 3.0.3 (Team, 2017).

Table 1
Ranges of window lengths and thresholds used for the sensitivity analysis of disturbance detection methods.

Method	Window length ($M_1 = M_2$, in years)	Threshold
Radial-growth averaging	From 5 to 15 by 1 step	From 25% to 175% by 25% increments
Boundary line		From 20% to 80% by 10% increments
Absolute increase		From 70% to 130% of default absolute increase threshold
Time series	Time step of 1, 2, 3, 4, 5, 7, 12, 17, 22	From 70% to 130% of default time series threshold

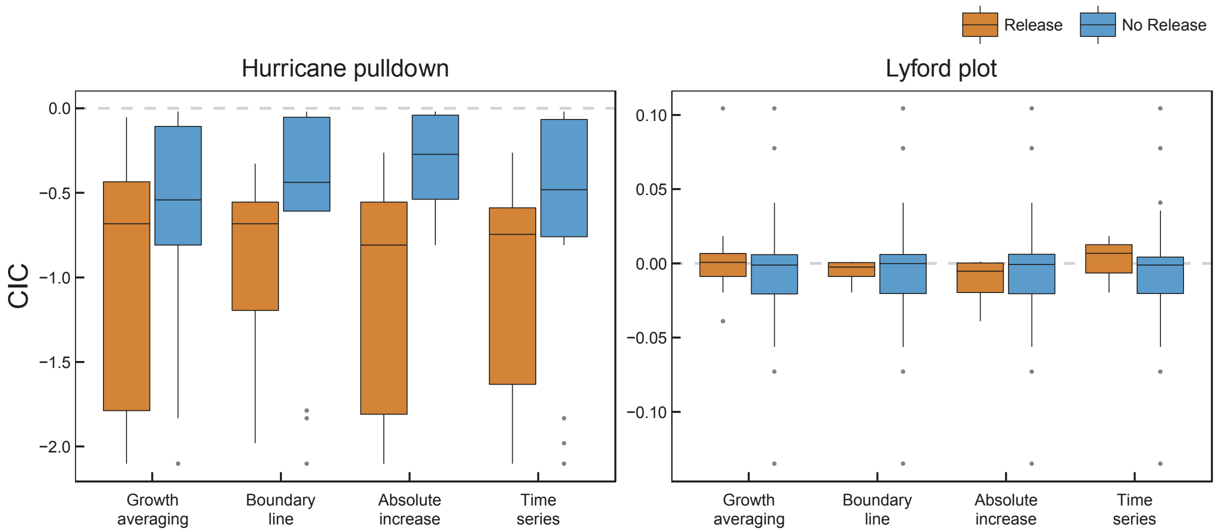

Fig. 1. Competition index change (CIC) of trees identified as released (orange) and not released (blue) based on the four disturbance detection methods at the Hurricane pulldown experiment (left panel) and the Lyford plot (minor 1938 hurricane damage, right panel). The solid black line represent median, the box represents 25 and 75 quantiles. CIC was calculated based on the inventory data (see methods). Note the different scales on the two y-axis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Results

4.1. Event detection at the tree level

Based upon changes in estimated tree competition in the Hurricane pulldown experiment, SVM analysis classified 87% of surviving trees as having been released from competition. A decrease of 0.325 of the CI, translating to approximately 33% loss of basal area around trees within five years, was determined to result in a significant chance of a growth release for trees in the pulldown plot versus those in the control plot (accuracy 0.94, sensitivity = 0.95, $p < 0.001$).

The median competition index change (CIC) of trees identified as being released by the four disturbance detection methods ranged from −0.68 (boundary-line and radial-growth averaging) to −0.81 (absolute-increase) for Hurricane pulldown and was around −0.01 for the Lyford plot (Fig. 1). Trees identified as released by absolute-increase methods had a significantly lower CIC ($p < 0.001$) than those identified as not released, where negative values of CIC indicate decrease in competition. Radial-growth averaging method identified the highest number of growth releases (n = 13 Hurricane pulldown, n = 8 Lyford plot) while time-series identified the least (n = 8 Hurricane pulldown, and n = 3 Lyford plot).

The accuracy of the four disturbance detection methods (measured as trees with an identified release from those with a significant change in CIC identified as having been released through SVM analysis) was 60–76% for Hurricane pulldown, and 85–94% for the Lyford plot, the forest with the lowest disturbance rate (Fig. 2). False negatives were more common than false positives for the Hurricane pulldown, but false negatives were not detected in the Lyford plot (meaning that all trees estimated to have been released by the SVM analysis were also detected by the tree-ring method). The radial-growth averaging method had the highest prevalence of false positives in both stands (up to 15%). The time series method had the highest accuracy for the Lyford plot (94%), but showed the highest value of false negatives (36%) for the Hurricane pulldown.

4.2. Event detection at the stand level

The performance of each method regarding the timing and severity of detected disturbance events at the stand level was evaluated using data from the Hurricane pulldown experiment (Fig. 3) and the Harvard tract. While all methods identified peaks in disturbances within a year of the known event (1990 for the Hurricane pulldown and 1938 for the Harvard tract), the temporal offset of all series analyzed ranged from −6 years (six years prior to the event) to +7 years (7 years after the event). The absolute-increase method had the best temporal accuracy overall showing the lowest standard deviation (Hurricane pulldown = 1.88; Harvard tract = 1.31), positive or neutral skewness (1.85; −0.13 respectively), and low kurtosis for the Harvard tract (3.4; 0.04 respectively) in identifying the correct year of disturbance, suggesting correct identification of most releases. Here, a lower standard deviation shows greater temporal precision, positive skewness more detections in the years following the event and low kurtosis light tails or lack of outliers. In comparison, other methods had higher standard deviation (e.g. time-series method 2.7, boundary-line 2.5) or negative skewness (time-series method −0.96), reflecting a wider temporal spread of detected releases and releases detected several years before the actual disturbance.

The absolute-increase and the radial-growth averaging methods identified the highest severity of disturbance at the Hurricane pulldown experiment (i.e., peaks of the kernel density function, 47% and 49%, respectively). In contrast, the estimated severity of disturbance for the time-series and boundary-line methods showed roughly half those values (16% and 22% respectively). Differences among methods in estimating the severity for the 1938 event at the Harvard tract were lower, but still notable.

Our sensitivity analysis of all methods at the stand level showed that the boundary-line method was the most sensitive to changes in window length parameters and threshold levels within pre-established ranges (Table 1), showing the widest range of temporal detection (−1 to 5 years) (Fig. 3). Variation in detected severity of the event was greater than variation in temporal accuracy, and deviated from the default settings for up to 35% (radial-growth averaging), when within 14% from the default. In all methods, increasing the window length and minimum threshold for releases resulted in decreases of estimated disturbance severity ($p < 0.001$).

5. Discussion

Our comparison of four commonly used methods of disturbance detection from forests with well-documented past disturbance revealed a wide range of efficacy among the methods. Despite only minor

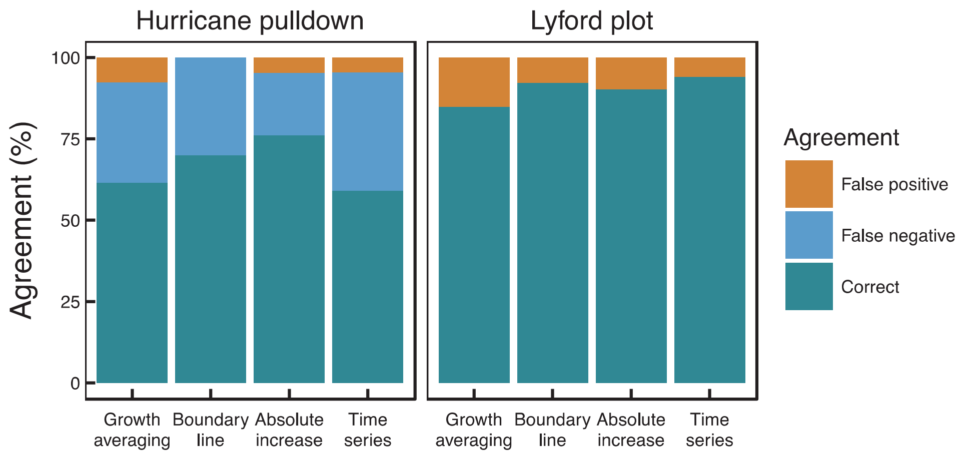

Fig. 2. Agreement between the classification of tree growth from the Hurricane pulldown experiment (left panel) and the Lyford plot (minor 1938 hurricane damage, right panel) using support vector machine analysis and the four disturbance detection methods tested for this study.

differences in the temporal accuracy in the detection of a disturbance event between methods, pronounced differences were observed in estimating the stand level severity of disturbance (up to 1.9 times higher). Performances improved (i.e. greater precision rates) and differences among methods were minor in a stand with a relatively low rate of canopy disturbance (Table 2). Our results showed that these uncertainties were greatest with the boundary-line and time-series methods.

Uncertainties in temporal accuracy of release events at tree level resulted in large uncertainties in reconstructed disturbance severity at the stand level (Fig. 3). Temporal precision was higher with the absolute-increase and radial-growth averaging methods and lower with boundary-line and time-series methods. All methods showed a substantial limitation in reconstructing the estimated severity of canopy disturbance. Although the radial-growth averaging and absolute-increase methods estimated a disturbance severity of just 50% of the trees in the Hurricane pulldown, that severity was more than twice the value estimated with the other two methods. While each method may detect the same number of trees showing a release around the time of a known disturbance, variations in temporal precision influenced the estimation of disturbance severity by nearly 1.9 times. It is clear that temporal precision of the various methods needs to be considered carefully when estimating disturbance severity and, equally important, when identifying agents of past disturbance events (Fritts, 1976; Black et al., 2016). Additional sources of information, such as growth suppression, tree injuries, death dates, may also be used to precisely estimate the date of disturbance and potentially determine the agent of disturbance that affected the forest.

Parameter setting (window lengths and thresholds) is among the most critical and still largely unsettled issues in disturbance analysis from tree rings as it influences the precision and accuracy in releases and ultimately lead to over- or under-estimations of severity of disturbance events (Rubino and McCarthy, 2004; Bouriaud and Popa, 2007; Copenheaver et al., 2014). Thus, it is important to consider the existing trade-off between the higher probability of obtaining more false-positives with short window lengths and low thresholds or more false-negative with longer window lengths and more strict thresholds. The sensitivity to changes in parameters also depends on the method and our sensitivity test showed that the absolute-increase and radial-growth averaging methods were least sensitive to changes in parameter selection.

Spotlight: Examining a legacy of thresholds used for radial-growth averaging analysis.

The original thresholds for the radial-growth averaging method was set at averages over two 15-year windows and a minimum growth increase of 50% between each window (Lorimer, 1985; Lorimer and Frelich, 1989). Within a decade, lower thresholds of 10 years and a 25% growth increase were designated as a theoretical sensitivity of canopy trees to changes in local competition (Nowacki and Abrams, 1997). Both methods have been long used and are popular within the community and the radial-growth averaging default in TRADER (Altman et al., 2014) is the 10-year 25% growth increase of Nowacki and Abrams (1997). Our sensitivity analysis gives us the opportunity to compare how release detection differs between these thresholds.

We found an improvement of accuracy to 73% (Lyford plot) and 96% (Hurricane pulldown) using the 15-year, 50% threshold compared to 61% and 85% respectively of the 10-year, 25% threshold (Table 2). One goal of the 15-year, 50% threshold was to take a more conservative approach that would avoid detecting disturbance that could be related to changes in other drivers, such as climatic variability (Lorimer, 1985). Our results support this approach. While some events are missed with the 15-year, 50% threshold, it is less likely to identify events that did not occur.

The boundary-line and absolute-increase methods are constrained by the quantity of data or expert knowledge required prior to using them for disturbance analysis. These two methods require a significant amount of *a priori* information while radial-growth averaging and time-series methods require the least amount. Considering these constraints, the application of the boundary-line and absolute-increase methods may be primarily limited to large datasets of a single species or to locations where growth information of many species is already available (Black and Abrams, 2003; Fraver and White, 2005; Ziaco et al., 2012). We view the *a priori* need for large data sets regarding expected species growth patterns or expert knowledge to be a limitation for studies in forests with high tree-species diversity or where high tree replication is not feasible (e.g. projects with short timeframes for completion, protected areas or rare species). Of these methods, radial-growth averaging appears to be the method requiring the least amount of *a priori* information while still providing accurate results. We also note that the time-series method is the only one that produces additional information regarding the magnitude and duration of release events (Rydval et al., 2015).

Finally, the aggregation of disturbances at the stand level from any of these four methods affects estimates of disturbance severity and timing and, thus our understanding of stand dynamics. Our results

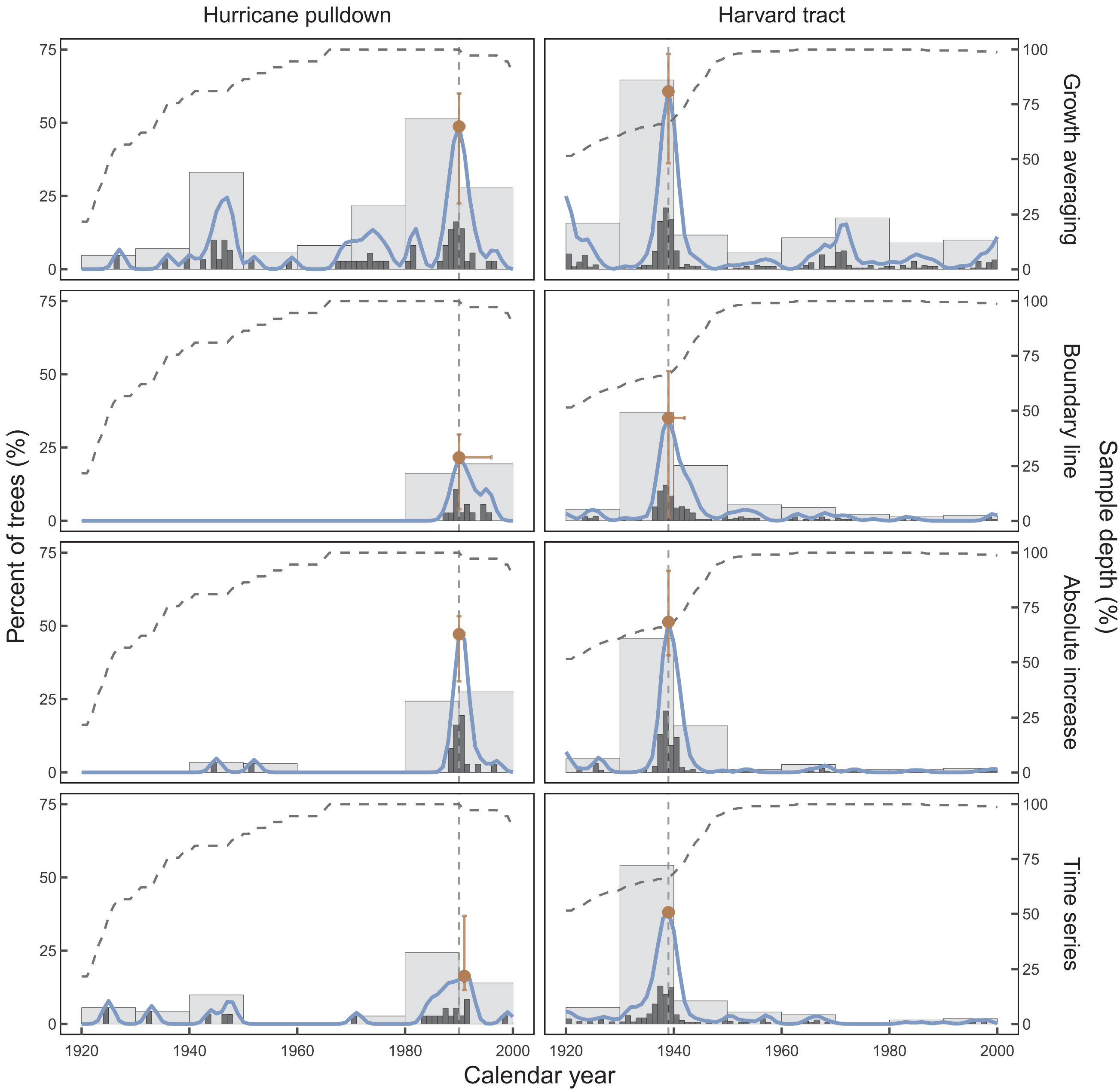

Fig. 3. Plot-level disturbance history in response to a simulated hurricane in 1990 (left panel) and natural hurricane at the Harvard tract in 1938 (right panel). The proportion of trees responding to disturbances is binned by year (black bars) and decade (grey bars). Peaks of disturbances (solid orange circles) were identified based on the standardized running kernel density estimation function (solid blue line). Accuracy, precision, and severity of release events (orange error bars) are identified here based on different window lengths. The dashed grey line shows sample depth (as a percent). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

indicate that annual binning of canopy disturbance, combined with kernel density estimation, would improve reconstructions of forest disturbance history and understanding of long-term forest dynamics. We find that decadal binning of disturbances may falsely suggest nearly continuous canopy disturbance. In contrast, annual binning revealed improved agreement with the documented events at our study forests and suggested more episodic disturbances.

Our study is among the first to compare the four methods of release detection; however, we are aware of possible limitations of our work, and suggest that future studies include: (i) a greater number of tree species with diverse functional traits (e.g. shade tolerance) and higher sample size, (ii) species-rich forests; (iii) a broader range of disturbance severity and agents; and (iv) use of different methods to estimate the likelihood that any given tree has been released from disturbance based on forest inventories. Nevertheless, our findings will assist on deciding which method to use and will advance the discussion and motivate researchers to conduct more in-depth comparison among the methods.

6. Conclusions

We found that the radial-growth averaging method and absolute-increase methods had lower levels of overall error in detecting canopy disturbance events in surviving understory trees with the original radial-growth averaging method producing fewer false positives than the default used in TRADER. Of the methods tested, radial-growth averaging requires the least amount of *a priori* information while returning

Table 2
Comparison of the efficacy of each method to various parameters of disturbance detection.

Parameters	Radial-growth Av.	Radial-growth Av. original	Boundary-line	Absolute-increase	Time-series
Correct					
Hurricane pulldown	61%	73%	70%	76%	59%
Lyford plot	85%	96%	95%	90%	94%
False Positive					
Hurricane pulldown	8%	4%	0%	5%	5%
Lyford plot	15%	4%	5%	10%	6%
False Negative					
Hurricane pulldown	31%	23%	30%	19%	36%
Lyford plot	0%	0%	0%	0%	0%
Temporal Precision of Disturbance (range, in years)	−1 to 1	−1 to 1	−1 to 5	0–1	−1 to 2
Calculated severity					
Hurricane pulldown	49%	41%	22%	47%	16%
Harvard tract	61%	49%	35%	52%	38%
Sensitivity to Parameter Thresholds	High	High	High	Low	High
Large Amount of Data Required	No	No	Yes	Yes	No
A priori Information Required	No	No	Yes	Yes	No
Additional Information Returned	No	No	No	No	Yes

reasonably accurate results. We note that the time-series method is the only one that produces additional information regarding the magnitude and duration of release events. The findings of this study can improve researchers' choice of which method to use for the dendrochronological reconstruction of the past disturbances.

Acknowledgements

We thank A. Ellison, J. McLachlan, R.-M. Muzika, M. Svoboda, J. Altman and others whose insight greatly improved the approach to our goals for this manuscript. Thanks to S. Irizarry for assistance with field data collection. Partial support for this research comes from the EF-1241930, which supports the PalEON Project (paleonproject.org). V. Trotsiuk was supported by Scientist Exchange Program, project (15-14840S) and Czech University of Life Sciences, Prague (CIGA No. , No.). D. Druckenbrod acknowledges support from a Rider University Research Sabbatical. D. Martin-Benito acknowledges support from Marie-Curie grant (EU-grant) and project AGL2015-73190-JIN from the Spanish . This is a publication of the Long Term Ecological Research (LTER) site.

Authors' contributions

VT, DD, DM-B and NP conceived ideas and designed the study. DO, DB, AB-P and NP performed the sampling. VT, DD, DM-B, DB and NP performed the dendrochronological analyses. VT, DD and DM-B performed data and statistical analysis. VT, DD, DM-B and NP wrote the manuscript and all authors commented on it. All authors contributed critically to the drafts and gave final approval for publication.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.foreco.2018.05.045.

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.

Black, B.A., Abrams, M.D., 2003. Use of boundary-line growth patterns as a basis for dendroecological release criteria. Ecol. Appl. 13, 1733–1749.

Black, B.A., Abrams, M.D., 2004. Development and application of boundary-line release criteria. Dendrochronologia 22, 31–42.

Black, B.A., Griffin, D., van der Sleen, P., Wanamaker, A.D., Speer, J.H., Frank, D.C., Stahle, D.W., Pederson, N., Copenheaver, C.A., Trouet, V., Griffin, S., Gillanders, B.M., 2016. The value of crossdating to retain high-frequency variability, climate signals, and extreme events in environmental proxies. Glob. Change Biol. 22, 2582–2595.

Borchers, H.W., 2017. pracma: Practical Numerical Math Functions; 2015. R package version, 2.

Bouriaud, O., Popa, I., 2007. Dendrochronological reconstruction of forest disturbance history, comparison and parametrization of methods for Carpathian Mountains. Analete ICAS 50, 135–151.

Cook, E.R., Kairiukstis, L.A., 1990. Methods of Dendrochronology: Applications in the Environmental Sciences. Springer, Netherlands.

Cooper-Ellis, S., Foster, D.R., Carlton, G., Lezberg, A., 1999. Forest response to catastropthic wind: results from an experimental hurricane. Ecology 80, 2683–2696.

Copenheaver, C.A., Black, B.A., Stine, M.B., McManamay, R.H., Bartens, J., 2009. Identifying dendroecological growth releases in American beech, jack pine, and white oak: within-tree sampling strategy. For. Ecol. Manage. 257, 2235–2240.

Copenheaver, C.A., Seiler, J.R., Peterson, J.A., Evans, A.M., McVay, J.L., White, J.H., 2014. Stadium woods: a dendroecological analysis of an old-growth forest fragment on a university campus. Dendrochronologia 32, 62–70.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Machine Learning 20, 273–297.

Douglass, A.E., 1920. Evidence of climatic effects in the annual rings of trees. Ecology 1, 24–32.

Druckenbrod, D.L., 2005. Dendroecological reconstructions of forest disturbance history using time-series analysis with intervention detection. Can. J. For. Res. 35, 868–876.

Druckenbrod, D.L., Neiman, F.D., Richardson, D.L., Wheeler, D., 2018. Land-use legacies in forests at Jefferson's Monticello plantation. J. Vegetation Sci. n/a–n/a.

Druckenbrod, D.L., Pederson, N., Rentch, J., Cook, E.R., 2013. A comparison of times series approaches for dendroecological reconstructions of past canopy disturbance events. For. Ecol. Manage. 302, 23–33.

Eisen, K., Plotkin, A.B., 2015. Forty years of forest measurements support steadily increasing aboveground biomass in a maturing, Quercus-dominant northeastern forest. J. Torrey Botanical Soc. 142, 97–112.

Foster, D.R., 1988. Disturbance history, community organization and vegetation dynamics of the old-growth pisgah forest, south-western new hampshire, U.S.A. J. Ecol. 76, 105–134.

Foster, D., Barker Plotkin, A., 1999. "Lyford Mapped Tree Plot at Harvard Forest since 1969." Harvard Forest Data Archive: HF032. Available online: http://harvardforest.fas.harvard.edu:8080.

Fraver, S., White, A.S., 2005. Identifying growth releases in dendrochronological studies of forest disturbance. Can. J. For. Res. 35, 1648–1656.

Fritts, H., 1976. Tree rings and climate. Academic, San Diego, Calif, pp. 567.

García, O., 2014. Siplab, a spatial individual-based plant modelling system. Comput. Ecol. Software 4, 215.

Grissino-Mayer, H.D., 2001. Evaluating crossdating accuracy: a manual and tutorial for the computer program COFECHA. Tree-ring research.

Hegyi, F., 1974. A simulation model for managing jack-pine stands. Growth Models Tree Stand Simul. 30, 74–90.

Holmes, R.L., 1983. Computer-assisted quality control in tree-ring dating and measurement. Tree-ring bulletin.

Lee, E.H., Wickham, C., Beedlow, P.A., Waschmann, R.S., Tingey, D.T., 2017. A likelihood-based time series modeling approach for application in dendrochronology to examine the growth-climate relations and forest disturbance history. Dendrochronologia 45, 132–144.

Lorimer, C.G., 1980. Age structure and disturbance history of a southern appalachian virgin forest. Ecology 61, 1169–1184.

Lorimer, C.G., 1985. Methodological considerations in the analysis of forest disturbance history. Can. J. For. Res. 15, 200–213.

Lorimer, C.G., Frelich, L.E., 1989. A methodology for estimating canopy disturbance frequency and intensity in dense temperate forests. Can. J. For. Res. 19, 651–663.

Marshall, R. (1927) The growth of hemlock before and after release from suppression. Harvard Forest.

Mathworks, I., 2014. MATLAB: R2014a. Mathworks Inc, Natick.
McEwan, R.W., Pederson, N., Cooper, A., Taylor, J., Watts, R., Hruska, A., 2014. Fire and gap dynamics over 300 years in an old-growth temperate forest. Appl. Veg. Sci. 17, 312–322.
Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., 2015. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien, 2015. R package version, 1.6-7.
Nowacki, G.J., Abrams, M.D., 1997. Radial-growth averaging criteria for reconstructing disturbance histories from presettlement-origin oaks. Ecol. Monogr. 67, 225–249.
Pederson, N., Dyer, J.M., McEwan, R.W., Hessl, A.E., Mock, C.J., Orwig, D.A., Rieder, H.E., Cook, B.I., 2014. The legacy of episodic climatic events in shaping temperate, broadleaf forests. Ecol. Monogr. 84, 599–620.
Pickett, S.T.A., White, P.S., 1985. The Ecology of Natural Disturbance and Patch Dynamics. Academic Press.
Plotkin, A.B., Foster, D., Carlson, J., Magill, A., 2013. Survivors, not invaders, control forest development following simulated hurricane. Ecology 94, 414–423.
Rowlands, W., 1941. Damage to even-aged stands in Petersham, Massachusetts by the 1938 hurricane as influenced by stand condition. MF thesis. Harvard University, Cambridge, Massachusetts.
Rubino, D.L., McCarthy, B.C., 2004. Comparative analysis of dendroecological methods used to assess disturbance events. Dendrochronologia 21, 97–115.
Ryan, J.A., 2008. quantmod: Quantitative financial modelling framework. R package version 0.3-5. URL < http://www.quantmod. com > URL < http://r-forge. r-project. org/projects/quantmod > .
Rydval, M., Druckenbrod, D., Anchukaitis, K.J., Wilson, R., 2015. Detection and removal of disturbance trends in tree-ring series for dendroclimatology. Can. J. For. Res. 46, 387–401.
Šamonil, P., Kotík, L., Vašíčková, I., 2015. Uncertainty in detecting the disturbance history of forest ecosystems using dendrochronology. Dendrochronologia 35, 51–61.
Stan, A.B., Daniels, L.D., 2010. Calibrating the radial-growth averaging method for detecting releases in old-growth forests of coastal British Columbia, Canada. Dendrochronologia 28, 135–147.
Stokes, M., Smiley, T., 1968. An introduction to tree-ring dating. University of Chicago, Chicago, Reprinted 1996. University of Arizona Press, Tucson. ()
Team, R.C., 2017. R: A Language and Environment for Statistical Computing.
Warren, W.G., 1980. On removing the growth trend from dendrochronological data. Tree-ring Bull. 1980, 35–44.
Ziaco, E., Biondi, F., Di Filippo, A., Piovesan, G., 2012. Biogeoclimatic influences on tree growth releases identified by the boundary line method in beech (Fagus sylvatica L.) populations of southern Europe. For. Ecol. Manage. 286, 28–37.