

Article

Supplementary Material: Developmental dynamics of *Gilbertiodendron dewevrei* (Fabaceae) drive forest structure and biomass in the Eastern Congo Basin

Henry B. Glick^{1,†}0000-0002-2956-530X, Peter M. Umunay^{1,2,*}, Jean-Remy Makana³, Sean C. Thomas⁴, Jonathan D. Reuning-Scherer¹, and Timothy G. Gregoire¹

¹ School of the Environment, Yale University, New Haven, 06511, USA; henry.glick@yale.edu (H.G.); peter.umunay@yale.edu (P.U.); jonathan.reuning-scherer@yale.edu (J.R.-S.); timothy.gregoire@yale.edu (T.G.)

² United Nations Environment Programme, PO Box 30552-00100, Nairobi, Kenya

³ Faculty of Science, University of Kisangani, Democratic Republic of Congo; jeanremymakana@gmail.com

⁴ Institute of Forestry and Conservation, University of Toronto, Toronto, Ontario, M5S 3B3, Canada; sc.thomas@utoronto.ca

* Correspondence: peter.umunay@yale.edu

† These authors contributed equally to this work.

Abstract: Patterns of structural change associated with monodominant tropical forest complexes have remained enigmatic for decades. Here we extend previous efforts [1,2] in presenting the first longitudinal, local-scale analysis of forest dynamics in central Africa. Using four 10-ha censused plots measured across three time periods (959,312 stems ≥ 1 cm DBH), we analyzed changes in a number of biometrical attributes for four distinct forest types capturing the developmental gradient from mixed to *Gilbertiodendron dewevrei*-dominated forest. We modeled above-ground biomass (AGB), basal area (BA), and stem density across all species, and diameter at breast height (DBH), recruitment, and mortality for *Gilbertiodendron dewevrei*. We hypothesized that trends in these attributes would be consistent with a slow spread of *Gilbertiodendron dewevrei* into adjacent mixed forest. We identified statistically significant increases in AGB and BA across sites, and positive, though non-significant, increases in AGB and BA for most forest types. DBH and relative recruitment increased significantly for *Gilbertiodendron dewevrei* stems, while relative mortality did not. When looking from mixed to transitional to monodominant forest types, we found a statistically significant pattern of developmental aggradation and net expansion of monodominant forest. We do not attribute this to atmospheric forcing, but to a combination of (a) landscape-scale recovery or response to widespread disturbance (primarily historical fires), (b) *Gilbertiodendron dewevrei*'s ectomycorrhizal association, and (c) *Gilbertiodendron dewevrei*'s exceptional stress tolerance traits.

Keywords: *Gilbertiodendron dewevrei*, longitudinal, modeling, forest dynamics, Congo, Ituri

Citation: Glick, H.B.; Umunay, P.M.; Makana, J.-R.; Thomas, S.C.; Reuning-Scherer, J.D.; Gregoire, T.G. Developmental dynamics of *Gilbertiodendron dewevrei* (Fabaceae) drive forest structure and biomass in the Eastern Congo Basin. *Forests* **2021**, *12*, 738. <https://doi.org/10.3390/f12060738>

Received: 16 May 2021

Accepted: 29 May 2021

Published: 4 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Submitted to *Forests* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Conceptual framework

From the outset, this research sought to explore the temporal and spatial dynamics of *Gilbertiodendron dewevrei* by moving extensive tabular data into a spatial framework. At its heart, this was made possible by field-measured locations for every stem. Much of the statistical analysis that appears to be aspatial is intrinsically spatial by virtue of the fact that model inputs were derived from aggregated spatial data structures, discussed below. With one minor exception noted below, we conducted all analysis in a scripted framework using R version 3.5.3 [3] via RStudio version 1.1.463 [4]. We relied heavily on a number of R packages, which are referenced individually. The analytical workflow can be thematically grouped into: data pre-processing, creation of spatial data structures, and regression modeling, with quality control checking throughout. In the interest of reproducibility, we detail our methods in the following sections.

1. Data pre-processing

This research relied on data from the Ituri Forest Dynamic Plots curated by the Smithsonian Institution's Center for Tropical Forest Science (CTFS). The Ituri plots are part of the CTFS' Forest Global Earth Observation Network (ForestGEO), currently comprised of 67 long-term forest monitoring plots. In partnership with the CTFS, the Ituri plot data is made available through a collaborative effort between the Centre de Formation et de Recherche en Conservation Forestière (CEFRECOF, in the D.R.C.), the Wildlife Conservation Society (WCS, regional office in the DRC), and the Okapi Faunal Reserve (OFR, in the D.R.C). Researchers wishing to explore these core datasets can submit requests at <https://forestgeo.si.edu/sites/africa/ituri>. There are four Ituri plots (also referred to below as field sites) measuring 10 ha in size (200 m x 500 m), situated in two groups: (a) edoro1 and edoro2, and (b) lenda1 and lenda2. The field sites are named for the Edozo and Lenda Rivers that flow nearby, and the sites were established in 1994 based on specific forest conditions. The edoro2 site was selected as a reference plot. At the time of establishment, it was of mixed forest structure and species composition and did not contain any *Gilbertiodendron* species. The edoro1 site was selected because it had a comparable structure and composition to edoro2, with the exception of select, non-dominant patches of *Gilbertiodendron dewevrei*. The lenda1 and lenda2 plots were selected based on a mixed forest structure and species composition, but with *Gilbertiodendron dewevrei* dominance on portions of each site. At the lenda1 and lenda2 sites, $\geq 50\%$ of trees larger than 30cm diameter at breast height (DBH) were *Gilbertiodendron dewevrei* at the time the sites were established. Additional details on the field sites can be found in Makana *et al.* [5].

Our original tabular data inputs were .rdata files from the CTFS data repository. Each of 12 files documented one of three complete censuses for one of the four field sites. The duration of the first census was greater than subsequent censuses, due to the initial establishment of the field sites, the tuning of monitoring protocols, and field crew training. The first census was conducted between the 1994 (4.3% of all records), 1995 (17.9%), and 1996 (9%) field seasons; the second census was conducted between the 2001 (32.5%) and 2002 (0.6%) field seasons; and the third census was conducted in the 2007 (35.7%) field season. Here a census is defined as the measurement of a suite of attributes (including relative spatial location) on *all* stems of *all* trees ≥ 1 cm DBH (diameter at 1.3 m height). The field sites were surveyed by theodolite, and were divided into 20 m x 20 m quadrats and 5 m x 5 m sub-quadrats to ease the record keeping process (See also [5]).

We combined the 12 tabular inputs to produce a single master dataset before computing basal area from measured DBH, formatted time stamps, field site identifiers, and census identifiers. We relied on the status variable included with the original data in downstream analysis. The status variable contains one of three codes to indicate that a given stem is alive (A), dead (D), or prior (P). A code of P indicates that the referenced stem does not yet exist, and reflects the underlying CTFS plot data structure: each dataset for each field site contains all trees that have ever been documented at that site. For example, a dataset from the first census will contain a record for each tree documented in census one, but also a record for each new tree documented in the second or third censuses, even though those trees did not exist at the time of the first census. All records with P status ($n = 69,367$) were omitted from analysis since they effectively represented duplicates and had no diameter measurements. Any missing values in the master dataset's status variable were populated through interpretation of the codes variable originally distributed with the data. The codes variable provides a more nuanced indication of tree condition. Due to the temporal span of data collection and the use of different field crews, inconsistencies in the codes variable were resolved with the help of CTFS staff.

Species information was coded into the original data (sp variable), and primarily contained six letter codes in which the first four letters conveyed the genus and the last two letters conveyed the species (e.g., *Gilbertiodendron dewevrei* was coded as GILBDE). Trees were identified to the species level for $>99\%$ of records. Data cleaning was required to clean and unify typographic errors, and to re-code unique but unidentified species of known genera with seven letter codes indexing the genus and the unique but unidentified species (e.g., ACANSP0). There were 438 unique tree species identified across all sites and censuses. Our processed tabular dataset was comprised of 959,312 individual stem measurements from 342,893 unique trees across the four field sites and three censuses. It contained records of alive ($n = 872,333$, 90.9%), dead ($n = 66,995$, 7%), and missing stems (i.e., the stem tag could not be found, $n = 19,984$, 2.1%). Trees whose stem tags could not be located were treated as dead. Stem diameters ranged from 1 cm to 176.5 cm for live trees. Diameter measurements were missing from all dead and missing stems, and for 806 alive stems. These records were omitted from some analyses, detailed below.

2. Creation of spatial data structures

Where our analysis depended on locational information, we used two techniques to populate the (x,y) coordinate information missing from 13,102 non-prior records. In the first, we used individual stem identifiers stored in the tag variable to identify stems that contained at least one set of valid coordinates across the three censuses. We then associated these coordinates with the respective records missing coordinates ($n = 1$ tag). In the second technique, we made use of

the quadrat variable, which contains coded values specifying to which 20 m x 20 m quadrat a particular stem belonged. The quadrat codes are four digit codes in which the first digit reflects the field site, the second digit increments quadrats along the short axis of the field site (relative to the origin), and the third and fourth digits increment quadrats along the long axis of the plot (relative to the origin). For each record that lacked coordinates but possessed a quadrat code ($n = 13,101$), we parsed the quadrat codes and used the parsed locational identifiers to locate each stem within its quadrat in a spatially random manner. Randomization was performed across both x and y dimensions, and ensured that no two stems were located in the same location. We felt this technique was suitable for our needs, given that much of our analysis relied on aggregated spatial units (detailed below). While it was recorded in the field, sub-quadrat locational information was not available in the original CTFS files. We corrected the longitude of one record that possessed a typographic error in its location.

While the corners of all sites were captured in the field using a civilian grade Global Positioning System (GPS) transmitter, minor locational error led us to define spatial plot boundaries using the origins (in World Geodetic System of 1984, WGS84) and the known field plot sizes and orientations. The longitude and latitude of each stem in our dataset was spatially referenced in meters from the respective plot origin (SW corners for lenda1 and lenda2, and NW corners for edoro1 and edoro2). To obtain valid real-world (projected) coordinates of each stem, the known origin locations were projected to Zone 35 North of the Universal Transverse Mercator (UTM 35N) coordinate reference system (in meters), and the appropriate stem-specific offsets were applied to each record, relative to their respective plot origins.

In a two-phase process we spatially aggregated stem records into quadrats measuring 25 m by 25 m in size ($n = 160$ per site per census, $n = 1,920$ across all sites and censuses). The quadrat size was selected to provide a comparable scale to what was recorded in the field, but also to provide units that would help remove measurement artifacts associated with field-based quadrat delineations. We performed these aggregations using aspatial data that contained coded quadrat identifiers. In the first phase we used the `parallel` [6] and `doSNOW` [7] packages to perform embarrassingly parallel aggregation by species. For all stems of each species within each quadrat, within each site, within each census, we:

- (a) tallied the number of alive stems, the number of dead stems, and the total number of stems;
- (b) assigned a quadrat identifier, field site identifier, and census identifier;
- (c) identified the mean date of field work, ascribed this to each stem lacking a date attribute, and then computed the mean of all stem measurement dates; and
- (d) computed total basal area and above-ground biomass of living trees (status = A).

The above process yielded a tabular dataset that could be aggregated by quadrat and across species in the second phase of aggregation. For each unique quadrat within each unique census, we:

- (a) computed the total number of live and dead trees;
- (b) computed the grand total number of stems;
- (c) computed the total basal area and above-ground biomass;
- (d) computed, separately, the total above-ground biomass of the species that held the largest and second-largest proportions of AGB on the quadrat;
- (e) computed, separately, the proportions that the two elements of (d) were of the AGB from (c);
- (f) computed the total AGB of *Gilbertiodendron dewevrei* stems;
- (g) computed the proportion that (f) was of the AGB from (c);
- (h) computed the mean date of field work; and
- (i) classified each quadrat into one of four land cover classes: *Gilbertiodendron*-dominated forest, mixed forest, transitional forest falling between *Gilbertiodendron*-dominated and mixed forest structures, and riparian forest (Fig. 2 main text). We performed classification using the total basal area of *Gilbertiodendron* stems ≥ 5 cm DBH, relative to the total basal area of all stems ≥ 5 cm DBH. If *Gilbertiodendron* basal area represented $\geq 50\%$ of the total basal area, we classified the quadrat as *Gilbertiodendron*-dominated forest. If the *Gilbertiodendron* basal area represented $< 5\%$ of the total basal area, we classified the quadrat as mixed forest. We classified all other quadrats as transitional forest. Then, in a subsequent step, we reclassified a quadrat to riparian forest if it shared direct adjacency (i.e., direct intersection, or one or more shared boundary vertices) with polygonal hydrographic features that were manually digitized using imagery, old field maps, and local knowledge of the landscape.

As a sub-task of the above process, for each site, for each census, for each quadrat, we also (a) isolated all *Gilbertiodendron dewevrei* records, (b) extracted the DBH values and associated census identifier, and (c) computed the change in DBH between the maximum and minimum DBH values of each record. This information was held separately given its non-aggregated structure, to be used for annual increment growth modeling.

We converted our aspatial aggregated dataset into a spatial dataset by binding quadrat-specific records to their spatial counterparts. We created the spatially explicit counterparts by first defining tessellation objects (class `tess`)

using the spatstat package [8,9], and then (using custom functions), converting the tessellation objects to objects of class SpatialPolygonsDataFrame associated with the sp package [10,11]. Our SpatialPolygonsDataFrames were geographically referenced to UTM 35N (Fig. 1 main text). Downstream tests of spatial dependency relied on centroid coordinates from each quadrat, which were joined to the tabular records associated with each polygonal element. To ensure we did not obtain artifacts produced from the regularity of the original centroid coordinates, we randomly perturbed the coordinates up to a maximum of 12.49 m (one-half the quadrat width/length) across both x and y dimensions, adding unbiased variability (with a mean of 0) into the locations of the centroids.

3. Regression modeling

3.1. Linear mixed effects modeling overview

We used our aggregated quadrat data in linear mixed-effects modeling to determine whether there were meaningful and/or significant changes in a number of biometrics across time. Following Gregoire *et al.* [12] and Fitzmaurice *et al.* [13], mixed-effects modeling was selected as a means of addressing the longitudinal and (possible) spatial autocorrelation that was inherent in our spatially explicit, unbalanced, repeated measures data. Longitudinal data present us with three sources of variability: between-subject, within-subject, and measurement error. The general objective is to properly account for the first two sources, given the third, in order to obtain valid model inference. Thus, longitudinal data of this nature require that we both model the mean response across time, and model the covariance among values of the same subject, as the correlation implicit in the latter violates customary assumptions of independence that are central to linear modeling [13] (p. 24-25). To the analyst, longitudinal autocorrelation may feel like a nuisance, while it is better viewed as a beneficial, model strengthening attribute. “Accounting for the covariance among repeated measures usually increases efficiency or the precision with which the regression parameters can be estimated” [13] (p. 163).

The basic linear mixed-effects model (1) is an extension of a traditional multiple linear regression modeling framework, and is denoted by

$$\hat{Y} = X\hat{\beta} + Z\hat{\gamma} + \hat{\epsilon} \quad (1)$$

Mixed effects modeling is largely concerned with the estimation of subject-specific parameters (the random parameters), where in our usage, a subject is a quadrat. Note, however, that these parameters are not necessarily of direct interest (addressed below). Allow that $i = 1, \dots, n$ indexes unique subjects, and $j = 1, \dots, T_i$ indexes unique measurement periods for the i th subject. T_i is thus the total number of observations for the i th subject, leading to $T = \sum T_i$, or the total number of observations across all subjects and measurement periods. In equation (1), the vector of response values Y , is a function of a fixed effects term ($X\hat{\beta}$), a random effects term ($Z\hat{\gamma}$), and an error term ($\hat{\epsilon}$). The fixed effects term is comprised of X , a $T \times p + 1$ design matrix, and $\hat{\beta}$, a $T \times 1$ vector of coefficients, as would be the case in a generalized least squares linear model. Here, p is the number of fixed effect predictors. The random effects term is comprised of Z , a $T \times nq$ block diagonal matrix, and γ , and $nq \times 1$ vector of random coefficients, where q is the number of random effect covariates. Each of the Z_i blocks contains a $T_i \times q$ design matrix. The error term is a $T \times 1$ vector of unexplained residual variation. X and Z need not include the same covariates, but the latter is always equal to, or a subset of, the former.

The values of the response vector $Y|\gamma$ (2) reflect a normally distributed random variable with a mean derived from the summed mixed and random effects terms, and, in the simplest form, a spherical error variance-covariance matrix

$$Y|\gamma \sim N(X\beta + Z\gamma, \sigma^2 I) \quad (2)$$

Conditional on the X and Z matrices, γ and ϵ can be characterized by similar multivariate normal structures

$$\gamma \sim N(0_q, \sigma^2 B) \quad (3)$$

$$\epsilon \sim N(0_T, \sigma^2 W) \quad (4)$$

where $0_{..}$ is a vector of subscripted length. B is a $q \times q$ correlation matrix of the γ random effect coefficients. The variance of the q random effects coefficients in γ are estimated from the unexplained variability remaining after fitting the mean model (i.e., the fixed effects term), and thus are computed across all records. This means that B captures an element of *between*-subject correlation and is not subject-specific (i.e., no subscripted index), though it could be a scalar in the case where there is a single random effect parameter to estimate, i.e., where Z is a unit vector, reflecting random variation in the intercepts associated with each subject. W is a $T \times T$ block diagonal correlation matrix of all W_i , which capture *within*-subject error correlations. In both (3) and (4), the global $\hat{\sigma}^2$ can be thought of as scaling factor to the correlation values.

The error variance-covariance matrix for the modeled response values, conditioned on the fixed and random effects terms, is captured by the block diagonal matrix

$$\text{Cov}(Y) = \Sigma = \sigma^2 V = \sigma^2 (\mathring{Z} \mathring{B} Z' + W) \quad (5)$$

in which σ^2 is a scaling parameter, and \mathring{B} is the Kronecker product [14] of $I_n \otimes B$, a $T \times T$ block diagonal matrix in which all of the B matrices appear on the diagonal and, in the case of uncorrelated subjects (addressed below), the off diagonals are all zeros. $\mathring{Z} \mathring{B} Z'$ and W can both be viewed as correlation matrices that serve to scale the residual variance.

Notationally, (5) follows Gregoire *et al.* [12], who use W more liberally than Fitzmaurice *et al.*'s [13] R . Gregoire *et al.* indicate that it is possible to both induce a structure in W through the use of random effects terms, and specify a correlation structure for W . This is in some contrast to Fitzmaurice *et al.* [13] who feel that mixing these two covariance modeling strategies can be problematic, and that it's easier to provide that $W = \sigma^2 I$, which puts constant variance on the diagonal and zeros in all off-diagonal elements. The implied divergence in these points of view warrants a moment of our attention.

Fitzmaurice *et al.* [13] speak to three broad approaches for modeling the error covariance matrix of repeated measures data: (1) unstructured covariance, or allowing the modeling process to estimate all covariance parameters in an unstructured manner; (2) use of covariance pattern models, or specifying a controlled, structured covariance matrix through knowledge of the correlation structure or mechanistic processes of interest; and (3) random effects covariance structures, in which the inclusion of random effects terms induce a covariance structure in W . Each of these approaches has its merits and limitations, but Fitzmaurice *et al.* [13] (p. 195) do not promote combining random effects covariance structures with covariance pattern models because: (a) when the W covariance matrix has a non-diagonal pattern there can be issues with model misidentification that prohibit estimating B and W simultaneously. Patterning in W implies that there is some model mis-specification that should, conceptually, have been captured by either $X\beta$ or $Z\gamma$ (and the estimated B associated with γ). Second, (b), the residual errors no longer have the simple interpretation of measurement or sampling error, in part because they would contain some element of non-independence.

Here we elected to ignore Fitzmaurice *et al.*'s [13] suggestion for several reasons. First, since we regarded the implicit meaning of the error values as unimportant, and since computerized model fitting permits refitting of misidentified models with few keystrokes, the only risk to fusing covariance modeling strategies was the possibility that misidentified models could not be fit. We felt this was of low cost. Second, the structure of W , dictated by the random effects estimation process, is related to the within-subject correlation. Meanwhile, the off-diagonal blocks of W capture the between-subject correlation and, in our data structure, should not be influenced by temporally-based random effects. It is the off-diagonal blocks that allow us to relate subjects spatially, if necessary, and the lack of overlap between the temporal and spatial blocks in W should not confound the interpretation of the model residual error. Third, our choice was based on the potential need to account not only for longitudinal autocorrelation, but also for spatial autocorrelation and heteroscedasticity. Within R there are limited model fitting functions that can accomodate this framework. In theory, this trio could be addressed using spatio-temporal or crossed-random effects models (e.g., [15]). However, we were unable to locate tools that could fit such models while also accounting for non-constant error variance. The one exception is Pinheiro and Bates' nlme package [16,17], which provides utilities for simultaneously fitting random effects and covariance pattern models, while also permitting model adjustments to account for heteroscedasticity of the error variance. The limitation with nlme is that the grouping variable for the random effects and spatial autocorrelation terms must be identical, which, in the case of our longitudinal dataset, was not ideal. We did not know *a priori* whether our modeled Ituri data would present spatial autocorrelation or heteroscedasticity, but we elected to use the most versatile toolset available.

In addition, it is worth noting that while "failure to take account of the covariance among the repeated measures will results in incorrect estimates of the sampling variability and can lead to quite misleading scientific inferences" [13] (p. 163), this is specific to model metrics derived from the variance-covariance matrix of the errors (standard errors, confidence intervals, etc.). In fact, the "least squares estimators of the fixed-effect covariance parameters are unbiased even in the presence of an incorrectly specified covariance structure" [12] (p. 137), and in practice the analyst will often find that modifications to the covariance structure do not appreciably change the estimates of the fixed effect coefficients.

Mixed effects models of the structure presented above provide a way for the analyst to estimate what are effectively sub-models for each level of a grouping variable that may or may not be nested. The random effects that are estimated (i.e., the empirical best linear unbiased predictions), serve as offsets to the coefficients estimated for the fixed effects term, whether a model has a single random effect parameter (i.e., random intercepts by subject) or as many random effect parameters as there are fixed effect parameters (i.e., $Z = X$). These offsets reduce the over residual error by providing mean models for each level of some grouping element, leaving ϵ to capture what Gregoire *et al.* [12] refer to as

“unattributable deviations”, or “disturbances” (p. 138). A distinction can be made between modeling frameworks that view the random effects coefficients as of intrinsic interest, versus those that view these as nuisance parameters. Here our primary interest was in the mean models and so we regarded the random effects as nuisance parameters.

3.2. Linear mixed effects model fitting

Given the longitudinal study design, our primary interest was in modeling the change in biometrics over time, by forest type. Here we present a detailed account of our workflow for modeling changes in AGB, followed by a more abbreviated description of the models for BA and the number of alive trees, which followed identical workflows.

To fit the linear mixed effects model structures presented above, we used the `lme` function within the `nlme` package [16, 17]. We fit fixed and random effects terms simultaneously because the unbalanced nature of the data led to an interaction in which the estimation of the random effects coefficients influenced the estimates of fixed effects coefficients, and vice versa. For our fixed effects term we used an analysis of co-variance (ANCOVA) model [18] (p. 466) of the form

$$\text{lme}(\text{fixed} = \text{agb} \sim \text{year} + \text{class} + \text{year:class} + \text{alive} + \text{qmd}) \quad (6)$$

where *year*, *alive* (number of alive stems), *qmd* (quadratic mean diameter), and *ba* (basal area) are continuous and forest *class* is a factor variable with four levels. In R, for a single observation, model (6) is internally expanded to the more familiar form

$$Y_{ij} = \beta_0 + \beta_1 \text{year}_{ij} + \beta_2 \text{class}_{2ij} + \beta_3 \text{class}_{3ij} + \beta_4 \text{class}_{4ij} + \beta_5 \text{year}_{ij} \text{class}_{2ij} + \beta_6 \text{year}_{ij} \text{class}_{3ij} + \beta_7 \text{year}_{ij} \text{class}_{4ij} + \beta_8 \text{alive}_{ij} + \beta_9 \text{qmd}_{ij} + \epsilon_{ij} \quad (7)$$

where $i = 1, \dots, 640$ indexes unique subjects (quadrats), and $j = 1, \dots, 3$ indexes measurement periods. Equation (7) is simplified, of course, to $Y_i = X_i \beta + \epsilon_i$. Equation (7) provides forest class-specific intercepts and slopes. We could have modeled each forest type separately, but (7) provides equivalent results [19] but with the increased precision obtained through a pooled sample size. Because forest class was defined based on relative *Gilbertiodendron* BA, we included some form of basal area within the mean model as a structural variable that could help control between-site differences in the same forest classes. The other predictors were the result of standard model selection procedures (i.e., evaluation of Akaike Information Criterion [AIC][20], Bayesian Information Criterion [BIC][21], likelihood ratios, parameter significance, etc.).

The random effects design matrix (*Z*) contained a vector of ones and the *year* variable, set to estimate random intercept and slope coefficients by unique quadrat. This is captured in R with

$$\begin{aligned} \text{lme}(\text{fixed} = \text{agb} \sim \text{year} + \text{class} + \text{year:class} + \text{alive} + \text{qmd}, \\ \text{random} = 1 + \text{year} | \text{quadrat}) \end{aligned} \quad (8)$$

or, following (7),

$$\begin{aligned} \text{agb}_{ij} = \beta_0 + \beta_1 \text{year}_{ij} + \beta_2 \text{class}_{2ij} + \beta_3 \text{class}_{3ij} + \beta_4 \text{class}_{4ij} + \beta_5 \text{year}_{ij} \text{class}_{2ij} + \\ \beta_6 \text{year}_{ij} \text{class}_{3ij} + \beta_7 \text{year}_{ij} \text{class}_{4ij} + \beta_8 \text{alive}_{ij} + \beta_9 \text{qmd}_{ij} + \\ b_{1i} + b_{2i} \text{year}_{ij} + \epsilon_{ij} \end{aligned} \quad (9)$$

where the *i* subscript on the random effects terms (third line) reflects values specific to the *i*th subject.

We thoroughly tested the conditional residuals from (9) for spatial autocorrelation, as we expected, *a priori*, that our quadrat data would contain spatial dependencies. Using the `moran.test` function of the `spdep` package [11,22], we evaluated the estimates and variance of the estimates of Moran’s *I* under the assumption of randomization. Of the twelve census × site combinations, only edoro1 at census 3 ($I = 0.09 \pm 0.006$ [1 s.e.], $p = 0.002$), lenda 2 at census 3 ($I = 0.03 \pm 0.003$; $p = 0.06$), and lenda 2 at census 3 ($I = 0.06 \pm 0.005$; $p = 0.03$) contained statistically significant positive spatial autocorrelation. Using the `variogram` function from the `gstat` package [23,24], we visually evaluated variograms to explore the spatial autocorrelative patterns suggested by these test results. We detected no discernible patterns of spatial association, despite the numerical results.

To ensure that we were not mistaken, we used the `fit.variogram` function from the `gstat` package to fit a number of correlation functions to the spatially lagged variogram data built from each these three subsets. The models included, among others, exponential, Gaussian, linear, spherical, and Matern correlation structures. None of the models converged with reliable estimates. We also tried adding a patterned covariance model to our mixed effects model (9) through the `lme` function, which uses its own model fitting algorithms. Here too we had limited success, with the exception of a

rational quadratic correlation structure that suitably converged and provided a defined variance-covariance structure. However, the added complexity of this model was not compensated for by increased explanatory power, and it proved a less likely fit to our data than did (9), when comparing the models using AIC [20] and BIC [21]. An element dictating the success of adding a patterned covariance model to a linear mixed effects model, is that when using the `nlme::lme` function, the analyst is forced to use the same stratification (grouping) variable for both the random effects structure and the `corStruct` [16] (p. 234) object supplied to the function's correlation argument. To account for longitudinal dependency, our stratification variable for our random effects was a `quadrat` identifier. However, it would make little sense to examine spatial correlation structures across time. If we omit a grouping variable all together, then spatial association is evaluated across space and time simultaneously. While we individually tested each census' worth of data for spatial dependencies, it was fortunate that we did not detect marked evidence of spatial dependence, as our actual modeling framework would not have been able to accommodate such dependencies in an optimal manner.

Plots of the conditional residuals of (9) were evaluated for each census \times site combination and as a group, with respect to observation order, each of the covariates, and the fitted values, and with normal quantile plots. We considered both standardized "Pearson" residuals, in which the raw residuals associated with each subject were divided by their respective standard errors, and Cholesky residuals. Fitzmaurice *et al.* [13] (p. 237-239) argue that evaluating conventional residual plots associated with mixed effects models (i.e., using Pearson residuals) is often a poor choice given the potential issues of heteroscedasticity and correlation between residuals and covariates that may induce systematic trends in the residual plots. In regression frameworks we consider these issues to be the rule, as opposed to the exception, and see no reason why standardized or normalized residuals can't be used to evaluate these two behaviors. However, as a quality control measure, we explored the Cholesky residuals. Following Fitzmaurice *et al.* [13], Cholesky residuals were computed as

$$r_i^* = L_i^{-1}r_i \quad (10)$$

where L_i^{-1} is, for the i th subject, the inverse of the conjugate transpose of a Cholesky decomposition of the unexplained model error variance covariance matrix, such that $\hat{\Sigma}_i = L_i L_i'$, and where r_i are the raw conditional residuals for the i th subject.

There were no clear patterns of heteroscedasticity in the Cholesky residual plots, and color coding by subject confirmed our above finding that there was no apparent spatial autocorrelation. Again, for quality control purposes, we experimented by adding to (9) a model for undetected heteroscedasticity within Σ . Using the `varFunc` classes from the `nlme` package [16,17], we applied weighting functions to (9) as one would in a weighted least squares regression. However, the weighting functions provided by the `varFunc` classes [16] (p. 208) are more nuanced, and variance can be modeled as a function of one or more covariates, as well as by subject. We tested models using:

- (a) fixed variance (by *year*, *ba*, *qmd* and the proportion of *agb* that was *Gilbertiodendron dewevrei*);
- (b) different variances estimated by grouping stratum (site, forest classification, site \times classification, *quadrat*, census ID);
- (c) variance as a function of the power of a covariate (*year*, *ba*, *qmd*, *NumOfTrees*, *ba* or *qmd*) by *site*, *ba* or *qmd* by forest class, *ba* or *qmd* by *quadrat*, *ba* or *qmd* by site-census pairings);
- (d) variance as a function of the exponential or a covariate (*ba*, *qmd*, *NumOfTrees*, and each of these with the grouping variables *site*, *censusID*, *site* \times *censusID*, or *forestclass*).
- (e) variance as a function of a constant plus a power of a covariate (*year*, *ba*, *qmd*, *NumOfTrees*, and each of these with the grouping variables *site*, *censusID*, *site* \times *censusID*, or *forestclass*); and
- (f) variance as a combination of (b) and (c), (b) and (d), or (b) and (e).

During model evaluation and comparison, candidate models were fit using maximum likelihood (ML). Those that properly converged were compared using an ANOVA table and associated AIC and BIC measures. We define proper convergence as successful optimization *and* as the stable estimation of a variance-covariance matrix. Often when fitting complex mixed effects models with `nlme::lme`, models appear to converge correctly, while close scrutiny of the correlation coefficients from B reveal perfect correlation values of ± 1 , or close scrutiny of the variances from Σ reveals values of ≈ 0 [25]. These problems are associated with models that include estimated singular variance covariance matrices, and generally appear when trying to fit models that are too complex (particularly with respect to the random effects structure) to be supported by the data [26]. In our workflow, the top ranking model (type (d) above) with proper convergence was refit with restricted maximum likelihood (REML) to reduce possible bias associated with estimates derived from Σ . The top ranking variance sub-model for our mode of AGB model was captured by

$$Var(\epsilon_{ij}|X_{ij}) = \sigma^2 \exp(2\delta_{c_{ij}} X_{ij}) \quad (11)$$

where X_{ij} is a covariate value associated with the i th subject and the j th measurement occasion, and δ_c is an unrestricted parameter that is estimated from $\hat{\Sigma}$, indexed by $1, \dots, c$ unique field site identifiers ($n=4$). Here, $X = ba$. Thus, we extend (8) to

$$\begin{aligned} \text{lme}(\text{fixed} &= \text{agb} \sim \text{year} + \text{class} + \text{year:class} + \text{alive} + \text{qmd}, \\ \text{random} &= 1 + \text{year} | \text{quadrat}, \text{weights} = \text{varExp}(\text{form} = \text{ba} | \text{site})) \end{aligned} \quad (12)$$

and extend (9) to

$$\begin{aligned} \text{agb}_{ij} &= \beta_0 + \beta_1 \text{year}_{ij} + \beta_2 \text{class}_{2ij} + \beta_3 \text{class}_{3ij} + \beta_4 \text{class}_{4ij} + \beta_5 \text{year}_{ij} \text{class}_{2ij} + \\ &\quad \beta_6 \text{year}_{ij} \text{class}_{3ij} + \beta_7 \text{year}_{ij} \text{class}_{4ij} + \beta_8 \text{alive}_{ij} + \beta_9 \text{qmd}_{ij} + \\ &\quad b_{1i} + b_{2i} \text{year}_{ij} + \epsilon_{ij} \\ \text{Var}(\epsilon_{ij} | \text{ba}_{ij}) &= \sigma^2 \exp(2\delta_{c_{ij}} \text{ba}_{ij}) \end{aligned} \quad (13)$$

If we maintain that $i = 1, \dots, 640$ subjects, and $j = 1, \dots, 3$ measurement periods, but now add a stratification index of $c = 1, \dots, 4$ unique field sites, we arrive at the variance-covariance matrix Σ of unexplained residual variation. We illustrate one generic and two specific blocks from this block diagonal matrix:

$$\begin{aligned} \Sigma_i &= \sigma^2 \begin{bmatrix} Z_{i,1,1} BZ'_{i,1,1} + \exp(2\delta_{i,1,1} X_{2i,1,1}) & Z_{i,1,1} BZ'_{i,2,1} & Z_{i,1,1} BZ'_{i,3,1} \\ Z_{i,2,1} BZ'_{i,1,1} & Z_{i,2,1} BZ'_{i,2,1} + \exp(2\delta_{i,2,1} X_{2i,2,1}) & Z_{i,2,1} BZ'_{i,3,1} \\ Z_{i,3,1} BZ'_{i,1,1} & Z_{i,3,1} BZ'_{i,2,1} & Z_{i,3,1} BZ'_{i,3,1} + \exp(2\delta_{i,3,1} X_{2i,3,1}) \end{bmatrix} \\ \Sigma_{161} &= \sigma^2 \begin{bmatrix} Z_{161,1,2} BZ'_{161,1,2} + \exp(2\delta_{161,1,2} X_{2161,1,2}) & Z_{161,1,2} BZ'_{161,2,2} & Z_{161,1,2} BZ'_{161,3,2} \\ Z_{161,2,2} BZ'_{161,1,2} & Z_{161,2,2} BZ'_{161,2,2} + \exp(2\delta_{161,2,2} X_{2161,2,2}) & Z_{161,2,2} BZ'_{161,3,2} \\ Z_{161,3,2} BZ'_{161,1,2} & Z_{161,3,2} BZ'_{161,2,2} & Z_{161,3,2} BZ'_{161,3,2} + \exp(2\delta_{161,3,2} X_{2161,3,2}) \end{bmatrix} \\ \Sigma_{640} &= \sigma^2 \begin{bmatrix} Z_{640,1,4} BZ'_{640,1,4} \times \exp(2\delta_{640,1,4} X_{2640,1,4}) & Z_{640,1,4} BZ'_{640,2,4} & Z_{640,1,4} BZ'_{640,3,4} \\ Z_{640,2,4} BZ'_{640,1,4} & Z_{640,2,4} BZ'_{640,2,4} \times \exp(2\delta_{640,2,4} X_{2640,2,4}) & Z_{640,2,4} BZ'_{640,3,4} \\ Z_{640,3,4} BZ'_{640,1,4} & Z_{640,3,4} BZ'_{640,2,4} & Z_{640,3,4} BZ'_{640,3,4} \times \exp(2\delta_{640,3,4} X_{2640,3,4}) \end{bmatrix} \end{aligned} \quad (14)$$

While AGB was our primary focus, we also used linear mixed effects modeling to capture changes in BA, the number of alive stems, and DBH, across time. For each of these, our modeling framework followed the process noted above, with comparable mathematical structures. We modeled shifting BA with

$$\begin{aligned} \text{ba}_{ij} &= \beta_0 + \beta_1 \text{year}_{ij} + \beta_2 \text{class}_{2ij} + \beta_3 \text{class}_{3ij} + \beta_4 \text{class}_{4ij} + \beta_5 \text{year}_{ij} \text{class}_{2ij} + \\ &\quad \beta_6 \text{year}_{ij} \text{class}_{3ij} + \beta_7 \text{year}_{ij} \text{class}_{4ij} + \beta_8 \text{qmd}_{ij} + \beta_9 \text{alive}_{ij} + \beta_{10} \text{qmd}_{ij} + \beta_{11} \text{agb}_{ij} + \\ &\quad \beta_{12} \text{max1}_{ij} + \beta_{13} \text{max2}_{ij} + \\ &\quad b_{1i} + b_{2i} \text{year}_{ij} + \epsilon_{ij} \\ \text{Var}(\epsilon_{ij} | (\text{max1}_{ij})) &= \sigma^2 \exp(2\delta_{c_{ij}} \text{max1}_{ij}) \end{aligned} \quad (15)$$

where max1 represents the proportion of quadrat-level AGB captured by the species that holds the most AGB on each quadrat, max2 follows max1 but for the species that is second heaviest, and where $c = 1, \dots, 12$ indexes 12 unique $\text{site} \times \text{class}$ pairings. Here again, our variance covariance matrix of unexplained residual error is a $1,920 \times 1,920$ block diagonal matrix.

We modeled the number of live stems (of all species) on each quadrat with

$$\begin{aligned} \text{alive}_{ij} &= \beta_0 + \beta_1 \text{year}_{ij} + \beta_2 \text{class}_{2ij} + \beta_3 \text{class}_{3ij} + \beta_4 \text{class}_{4ij} + \beta_5 \text{year}_{ij} \text{class}_{2ij} + \\ &\quad \beta_6 \text{year}_{ij} \text{class}_{3ij} + \beta_7 \text{year}_{ij} \text{class}_{4ij} + \beta_8 \text{qmd}_{ij} + \beta_9 \text{ba}_{ij} + \\ &\quad b_{1i} + b_{2i} \text{year}_{ij} + \epsilon_{ij} \end{aligned} \quad (16)$$

$$\text{Var}(\epsilon_{ij} | X_{ij}) = \sigma^2 \delta_{1,c_{ij}}^2 (\delta_2 + |\text{ba}_{ij}|^{\delta_3})^2 \quad (17)$$

The reader will note that in (17) we are no longer employing (11) to model the variance, but are now employing the interaction of two variance functions (in R parlance, `nlme::varIdent` and `nlme::varConstPower`). This follows Pinheiro and Bates [16], and is a framework that can be extended to support any combination of variance models. Each of the δ_x parameters in (17) is estimated from $\hat{\Sigma}$. The first, δ_1 , is restricted to be positive and in our case, is indexed by $c = 1, \dots, 3$

unique censuses, with one δ being estimated for each census. The second, δ_2 , is a constant restricted to be positive while the third, δ_3 , is unrestricted. Pinheiro and Bates [16] (p. 212) note that when $\delta_3 > 0$, which is most always, variance will be approximately proportional to δ_2 at covariate values of 0, and increases with respect to the absolute value of the covariate. The combination of these two variance functions leads to the estimation of, potentially, six variance parameters — two (δ_2 and δ_3) for each of three census periods.

We modeled changes in the DBH of all *Gilbertiodendron* trees on each quadrat using a slightly different framework from that used for AGB, BA, or the number of alive stems. Here, instead of using aggregated values, we treated each tree as a subject. This greatly limited the explanatory variables we had available to us, but we felt this was the most sensible approach for capturing minor increment changes in tree diameter. Our best fitting model did not support a variance sub-model, and could only be fit with ML; REML was not supported. To understand the bias introduced into Σ -derived metrics from a model fit with ML, we computed the mean ratio of the fitted values (\hat{y}) from ML and REML models, for BA models. We found the mean ratio to be 0.000006, leading us to conclude that there was negligible bias in our DBH model's estimated coefficient standard errors. Our final DBH model was

$$\begin{aligned} dbh_{ij} = & \beta_0 + \beta_1 year_{ij} + \beta_2 class_{2ij} + \beta_3 class_{3ij} + \beta_4 class_{4ij} + \beta_5 year_{ij} class_{2ij} + \\ & \beta_6 year_{ij} class_{3ij} + \beta_7 year_{ij} class_{4ij} \\ & b_{0t} + b_{1t} year_{ijt} + \epsilon_{ij} \\ Var(\epsilon_{ij}) = & \sigma^2 \end{aligned} \quad (18)$$

where the t subscript on the random effects coefficients identifies individual identification tags, and where the remaining terms are as defined above.

Mean values and associated 95% confidence intervals reported in the main text were derived through bootstrapping. Using the `bootstrap.lme` function of the `lmeresampler` package [27], model observations were case resampled across all hierarchical levels ($n = 10,000$). In all cases we had only two levels: population and quadrat. The appropriate mixed effects model with heteroscedasticity sub-model was then fit to the resampled observations in a parallel computing framework. From the results we derived mean coefficient estimates and their 95% confidence intervals (percentile confidence intervals).

3.3. Beta regression modeling overview

Two of our response variables, the proportion of *Gilbertiodendron* stems that were classified as new recruits and the proportion that died, were continuous variables bounded by the unit interval [0-1]. Such data structures can lead to non-normal model error distributions and require a distinct modeling framework compared to that presented above. We elected to employ beta regression [28,29], which is effectively a generalized linear model (GLM), in that the response is related to the linear predictor (η) by way of a link function. However, unlike a logistic regression for binomially distributed data, here we are modeling data from a continuous distribution. Following Ferrari and Cribari-Neto [29] and Pereira and Cribari-Neto [30], we allowed that each of the y_1, \dots, y_n observations were beta-distributed random variables with density

$$f(y; \mu_i, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu_i \phi) \Gamma((1 - \mu_i) \phi)} y^{\mu_i \phi - 1} (1 - y)^{(1 - \mu_i) \phi - 1}, 1 < y < 1 \quad (19)$$

where γ is the gamma function, μ is the mean of y_i , and ϕ is a precision parameter. This is not the traditional parameterization of this distribution, but as Espinheira *et al.* [31] (p. 408) point out, this presentation is useful in that the $var(y_i) = \mu_i(1 - \mu_i)/(1 + \phi)$. Thus, for a fixed value of the mean, the variance of y_i shifts inversely to ϕ — as ϕ increases, the variance of y_i decreases [32]. ϕ is thus the inverse of dispersion. Because $var(y_i)$ is a function of the mean, heteroscedasticity can be easily accommodated using beta regression. When $\mu = 1/2$ and $\phi = 2$, the beta distribution equates to a standard uniform distribution. Following Ferrari and Cribari-Neto [29] and Cribari-Neto and Zeileis [33], a model for the mean response and variance of y_i can be captured by

$$g_1(\mu) = \mathbf{X}\boldsymbol{\beta} = \eta_1 \quad (20)$$

$$E(Y|X) = \mu = g_1^{-1}(\eta_1) \quad (21)$$

$$var(y_i) = \frac{\mu_i(1 - \mu_i)}{1 + \phi} \quad (22)$$

where $X\beta$ are the design matrix and vector of coefficients, respectively, and g_1 is a strictly monotonic and twice differentiable link function [32]. η_1 is the *linear predictor*, indexed to identify it from the precision model's link function shown below (25). Ferrari and Cribari-Neto [29] comment that a variety of link functions are possible in beta regression, but note that the logit link is particularly useful. We implemented a logit link for our μ sub-model (23) and a log link for our ϕ sub-model (24)

$$\text{logit } g_1(\mu) = \log(\mu/(1 - \mu)) \quad (23)$$

$$g_2(\cdot) = \log(\cdot) \quad (24)$$

Simas *et al.* [32] extend Ferrari and Cribari-Neto's [29] work by providing a non-linear beta regression framework, and, of particular interest here, a regression modeling structure for the precision parameter (ϕ). Similar to the mean model, Simas *et al.* [32] (p. 356) show how ϕ can be modeled using

$$g_2(\phi) = Z\theta = \eta_2 \quad (25)$$

$$E(\phi|\mu) = \phi = g_2^{-1}(\eta_2) \quad (26)$$

where $Z\theta$ are the design matrix and vector of coefficients for the ϕ model, and η_2 is the linear predictor specific to the precision parameter. Here, Z need not be identical to X , and could be a subset of X or have unique covariates. Cribari-Neto and Zeileis [34] refer to a beta regression model with dispersion covariates as a *variable dispersion beta regression model*, and from a conceptual point of view, shares similarities to Pinheiro and Bates' [16] variance functions. In the models we present below, we elected to model the variance using Simas *et al.*'s [32] approach.

In light of the inherent longitudinal autocorrelation, and minor spatial autocorrelation in our dataset, we endeavored to reduce the bias and possible inconsistencies in our variance covariance-based metrics (i.e., standard errors, confidence intervals, etc.) by applying a *post hoc* nonparametric sandwich estimator to our marginal beta regression model. Much of the literature on sandwich estimators appears in econometric journals, and their use has not yet become common in ecology. A wide variety of sandwich estimators exist [35], including those that are heteroscedasticity consistent, temporal autocorrelation consistent, and/or spatial autocorrelation consistent. A number of correction methods to account for cluster bias and/or heteroscedasticity-induced bias also exist, and are used in tandem with the sandwich estimators. In the present study we explored two sandwich estimators: Driscoll and Kraay's [36] heteroscedasticity-, temporal autocorrelation-, and cross-sectional autocorrelation-consistent estimator, and a simpler heteroscedasticity- and temporal autocorrelation-consistent estimator with bias correction [35,37,38]. In light of the fact that (a) our prior modeling efforts revealed low (though heretofore non-meaningful) levels of spatial autocorrelation in a selection of our field site \times census period combinations; and (b) ignoring spatial dependence in temporally autocorrelated datasets "typically leads to overly optimistic (anticonservative) standard error estimates" [39] (p. 282), we considered the Driscoll and Kraay's approach first, by way of the `vcovPL` function from the `sandwich` package [35,40,41].

Driscoll and Kraay [36] present a simulation study evaluating the performance of ordinary least squares (OLS) and seemingly unrelated regressions (SUR) estimators of parameter standard errors, as compared to their proposed, consistent, panel estimator. Their results indicate that for finite samples, their sandwich estimator performed markedly better than either OLS or SUR in the presence of combined temporal and spatial autocorrelation, with some evidence that the depth of the temporal dimension of the observations is associated with improved confidence interval coverage (i.e., more intervals capture the true value). Berger *et al.* [35] comment that in their suite of simulations, "empirical coverages falling short of 0.95 are typically due to underestimated standard errors and would lead to inflated type I errors in partial Wald tests of the coefficients" (p. 19). Poor performance may occur when the temporal dimension is too limited. When " T is small or when there is only a single cross section, the problem of consistent nonparametric covariance matrix estimation appears to be much less tractable" [36] (p. 559). Hoechle's [39] modeling simulation results confirm that Driscoll and Kraay's [36] estimator clearly depends on large-sample asymptotics with respect to the temporal dimension, as the longer the dimension the "better calibrated" the standard errors. Hoechle's [39] (p. 299) results also suggest that with as few as five time steps and an autocorrelation coefficient of 0.125, Driscoll and Kray's standard errors are superior to OLS and Roger's [42] clustered standard errors. This superior performance holds "irrespective of whether a panel dataset is balanced" [39] (p. 290), which our data is not. Importantly, though, in the absence of autocorrelation, Driscoll and Kray's estimator does not perform as well as others. A naive comparison with Berger *et al.*'s [35] figure 5 (p. 29) suggest that with a data structure like ours ($n = 480$, a maximum group size of 3, and, in the case of recruitment, a temporal autocorrelation coefficient of $\rho = 0.18$ using a lag of 1), we could expect the confidence intervals for our parameters to have $< 80\%$ coverage.

These latter points are important with respect to the present study, as our temporal dimension was, at most, three time steps, and because we had so little spatial autocorrelation across all sites and time periods that we were not able to model it. This, and suggestions from betareg package author A. Zeileis (personal comm. July 2019) led us to explore the `betareg::vcovCL` function, which employs a simpler strategy to capture dependence within, but not between, clusters [35]. Sandwich variance covariance estimators that capture one-dimensional clustering have existed for decades [37,38]. Liang and Zeger's [38] approach avoids "the need for multivariate distributions by only assuming a functional form for the marginal distribution at each time. The covariance structure across time is treated as a nuisance" (p. 20). This is not conceptually dissimilar from how random effects are often treated in mixed effects modeling. To obtain consistent estimates of the variability, this approach requires that subjects are independent, but "no restrictions are placed on the form of the autocovariances for a given individual" [37] (p. 431). More recently, [43] extended sandwich estimators for clustered standard covariances to multiple, non-nested dimensions, as has been implemented in `vcovCL` function of the `sandwich` package [35,40,41].

For both our models of recruitment and mortality we employed `vcovCL` with a single dimension of clustering, being the `quadrat` identifiers, which link observations temporally, as in our mixed effects models. As noted above, bias correction can be separated into cluster bias correction and heteroscedasticity bias correction. The cluster bias correction corrects for the finite sample size, and is defined, following Berger *et al.* [35] (p. 6) as

$$G/(G - 1) \quad (27)$$

where G represents the number of different clusters ($n = 480$ in our study). Of the two heteroscedasticity bias correction options available for models of the R class `betareg`, we employed type "HC1", which Beger *et al.* [35] show to perform better than "HC0". The HC1 bias correction corrects for heteroscedasticity *within* each cluster using

$$n/(n - k) \quad (28)$$

where n is the number of observations and k is the number of model parameters to be estimated. In that we have a maximum of three observations per cluster, we can't expect any notable improvement from this bias correction method.

3.4. Beta regression modeling fitting

To fit the beta regression models discussed above, we used the `betareg` function from the `betareg` package [33,34,44]. We employed a logit link in all of our models, and followed Bayer and Cribari-Neto [45] in executing a two-stage modeling process in which we first modeled the mean response before considering a model for variable dispersion. Model selection was performed using two less common metrics: HQ_c and R^2_{LR} . Bayer and Cribari-Neto [45] present the results of a simulation study in which a large number of metrics used to evaluate the fit of variable dispersion beta regression models were considered. For samples sizes ≥ 50 they found that the HQ_c metric proposed by McQuarrie and Tsai [46] (p. 34-35) performed best among information criteria. HQ_c is a likelihood-based metric not too dissimilar from the Akaike information criterion (AIC), and incorporates a finite sample correction. Following Bayer and Cribari-Neto's [45] (p. 734) notation

$$HQ_c = -2\ell(\hat{\beta}, \hat{\gamma}) + \frac{2nk\log(\log(n))}{n - k - 1} \quad (29)$$

where $\hat{\beta}$ and $\hat{\gamma}$ are the maximum likelihood estimators of β and γ , ℓ is the log of the likelihood function, n is the number of observations, and k is the number of parameters to be estimated in the model. We created a custom function that returned the HQ_c value for any number of beta regression model objects supplied.

We also employed a pseudo coefficient of determination (R^2_{LR}) originally proposed by Nagelkerke [47] and Long [48], and recommended by Bayer and Cribari-Neto [45]. In the case of fixed dispersion models, Bayer and Cribari-Neto [45] found R^2_{LR} to perform similarly to other pseudo R^2 metrics. However, in a variable dispersion framework, R^2_{LR} was more sensitive to the ϕ sub-model, taking "significantly larger values" when the dispersion model was correctly modeled. We consider this to be desirable, in that the analyst has a lower chance of settling on a poorly fitting ϕ model. Nagelkerke [47] discusses how this metric can be interpreted as one would interpret the traditional coefficient of determination. Following Nagelkerke's [47] (p. 691) notation

$$1 - \exp\left[-\frac{2}{n}\{l(\hat{\beta}) - l(0)\}\right] \quad (30)$$

where $l(\hat{\beta})$ is the maximized log-likelihood of the model to be evaluated (presumably with covariates), $l(0)$ is the maximized log-likelihood of a beta regression model with only an intercept in each of the μ and ϕ sub-models, and n is the number of observations. As with HQ_c , we created a custom R function that would return the R^2_{LR} value for any number of beta regression model objects supplied.

In addition to HQ_c and R^2_{LR} , we used a typical array of residual plots to identify gross problems with model fit. Following the recommendations from Espinheira *et al.* [31], we used “standardized weighted residuals 2” type residuals when evaluating our models. These residuals proved superior in some applications because they can be used to investigate model misspecification while also identifying influential data points. This comes from the fact that these residuals implicitly take observation leverage into account. Following Espinheira *et al.* [31] (p. 409), these residuals are computed using

$$r_t^{ww} = \frac{r_t^*}{\sqrt{\phi^{-1}(1 - h_{tt})}} \quad (31)$$

$$r_t^* = \frac{y_t^* - \hat{\mu}_t^*}{\sqrt{\phi v_t}} \quad (32)$$

where y_t^* is the t th value of the logit response, $\hat{\mu}_t^*$ is the estimated value of y_t^* , ϕ is the estimated precision parameter, v_t is the variance of y_t^* , and h_{tt} is the t th diagonal element from the hat (projection) matrix. The “standardized weighted residuals 2”, r_t^{ww} , are implemented in R through the `residuals.betareg` method of the `betareg` package.

We concluded our beta regression model selection process by employing a misspecification test originally proposed by Ramsey [49] (p. 361). Ramsey’s “Regression Specification Error Tests”, or RESET, was advanced by Pereira and Cribari-Neto [30] as a way to test the suitability of both fixed and variable dispersion 0-1 inflated beta regression models, and suggested by Bayer and Cribari-Neto [45] as an element of the model selection process for non 0-1 inflated beta regressions. The test is based on the idea that “omitted variables and other forms of model misspecification can be proxied by some (unknown) analytic function of the linear predictor used in the definition of the” regression model [30] (p. 635). In practice, we can regress on the response vector, our original covariates *and* different powers of the fitted values from that original model. If the addition of the power covariates can approximate omitted explanatory terms, the augmented model will appear significantly better than the original model using likelihood ratio or Wald tests. If the model is suitably specified, the inclusion of the new explanatory terms will not improve the model fit. Pereira and Cribari-Neto’s [30] numerical simulations indicate that a RESET performed solely on the μ sub-model, is generally better performing than that applied to both the μ and ϕ sub-models. To extend (20),

$$g_1(\mu) = \eta_1 + A_1 \tau_1 \quad (33)$$

where A_1 is a matrix of testing variables used to augment the mean model, and τ_1 is a vector of estimated coefficients associated with those testing variables. Broadly, A could contain a number of perturbed terms, but Pereira and Cribari-Neto’s [30] (p. 655) “strong” recommendation is to augment the mean model with the squared values of the fitted linear predictor. We followed this guidance in our testing.

The proportion of *Gilbertiodendron* recruits within each quadrat was modeled using stems ≥ 20 mm in diameter. This threshold was established to avoid excessive skewing of the size class distribution, given that there were, at times, very large numbers of new recruits that subsequently died. The edoro2 field site was established as a control site, and contained no *Gilbertiodendron* in any of the censuses. Thus, we omitted these records in our analysis. From the edoro1, lenda2, and lenda2 field sites, we omitted any quadrat that lacked *Gilbertiodendron* in *every* census. Since we did not employ zero-one inflated beta regression, 0 and 1 values were manually altered to 0.001 and 0.999, respectively, to accommodate the `betareg` model fitting algorithms. As described above, the remaining records represented quadrat-level aggregations. We employed the following variable dispersion beta regression model

$$\text{logit}(\widehat{\text{pctRecruits}}) = \eta_1 = \hat{\beta}_0 + \hat{\beta}_1 \text{year}_{ij} + \hat{\beta}_2 \text{class}_{2ij} + \hat{\beta}_3 \text{class}_{3ij} + \hat{\beta}_4 \text{class}_{4ij} + \hat{\beta}_5 \text{year}_{ij} \text{class}_{2ij} + \quad (34)$$

$$\hat{\beta}_6 \text{year}_{ij} \text{class}_{3ij} + \hat{\beta}_7 \text{year}_{ij} \text{class}_{4ij} + \hat{\beta}_8 n\text{Trees}_{ij} + \hat{\beta}_9 ba_{ij} + \hat{\beta}_{10} \text{max}_{2ij} + \hat{\beta}_{11} \text{pctMort}_{ij}$$

$$\text{log}(\widehat{\phi|\mu}) = \eta_2 = \hat{\beta}_{\phi 0} + \hat{\beta}_{\phi 1} \text{year}_{ij} + \hat{\beta}_{\phi 2} \text{class}_{2ij} + \hat{\beta}_{\phi 3} \text{class}_{3ij} + \hat{\beta}_{\phi 4} \text{class}_{4ij} + \hat{\beta}_{\phi 5} \text{year}_{ij} \text{class}_{2ij} + \quad (35)$$

$$\hat{\beta}_{\phi 6} \text{year}_{ij} \text{class}_{3ij} + \hat{\beta}_{\phi 7} \text{year}_{ij} \text{class}_{4ij} + \hat{\beta}_{\phi 8} \text{site}_{2ij} + \hat{\beta}_{\phi 9} \text{site}_{3ij} + \hat{\beta}_{\phi 10} \text{mort}_{ij} + \epsilon_{ij}$$

$$\text{var}(\epsilon_{ij}) = \frac{\text{logit}^{-1}(\eta_1)[1 - \text{logit}^{-1}(\eta_1)]}{1 + \phi} \quad (36)$$

where $nTree$ represents the number of standing stems, $pctMort$ represents the proportion of *Gilbertiodendron* stems that were dead, padded by 0.001 as with the $pctRecruits$ variable (noted above); $site$ represents a categorical variable of three field sites; and the remaining variables are as defined above. We implemented this model in R using

$$\begin{aligned} &\text{betareg}(pctRecruits \sim \text{year} + \text{class} + \text{year:class} + \text{ntrees} + \text{ba} + \\ &\quad \text{max2} + \text{pctMort} \mid \text{year} + \text{class} + \text{year:class} + \\ &\quad \text{pctMort} + \text{site}, \text{data} = \text{datasetName}) \end{aligned} \quad (37)$$

Similarly, to capture trends in mortality we modeled the proportion of *Gilbertiodendron* stems that were initially dead or had died between censuses (i.e., the mortality of previously documented stems). In modeling stem mortality, we employed the following variable dispersion beta regression model

$$\text{logit}(\widehat{pctRecruits}) = \eta_1 = \hat{\beta}_0 + \hat{\beta}_1 \text{year}_{ij} + \hat{\beta}_2 \text{class}_{2ij} + \hat{\beta}_3 \text{class}_{3ij} + \hat{\beta}_4 \text{class}_{4ij} + \hat{\beta}_5 \text{year}_{ij} \text{class}_{2ij} + \hat{\beta}_6 \text{year}_{ij} \text{class}_{3ij} + \hat{\beta}_7 \text{year}_{ij} \text{class}_{4ij} + \hat{\beta}_8 \text{ba}_{ij} + \hat{\beta}_9 \text{pctRec}_{ij} + \hat{\beta}_{10} \text{site}_{2ij} + \hat{\beta}_{11} \text{site}_{3ij} \quad (38)$$

$$\text{log}(\widehat{\phi|\mu}) = \eta_2 = \hat{\beta}_{\phi_0} + \hat{\beta}_{\phi_1} \text{year}_{ij} + \hat{\beta}_{\phi_2} \text{class}_{2ij} + \hat{\beta}_{\phi_3} \text{class}_{3ij} + \hat{\beta}_{\phi_4} \text{class}_{4ij} + \hat{\beta}_{\phi_5} \text{year}_{ij} \text{class}_{2ij} + \hat{\beta}_{\phi_6} \text{year}_{ij} \text{class}_{3ij} + \hat{\beta}_{\phi_7} \text{year}_{ij} \text{class}_{4ij} + \hat{\beta}_{\phi_8} \text{pctRec}_{ij} + \epsilon_{ij} \quad (39)$$

$$\text{var}(\epsilon_{ij}) = \frac{\text{logit}^{-1}(\eta_1)[1 - \text{logit}^{-1}(\eta_1)]}{1 + \phi} \quad (40)$$

where all variables are as defined above.

We used the `boot` function from the `boot` package [50,51] with custom functions to bootstrap the beta regression models and their corrected variance-covariance matrices from `vcovCL`. We did this because of inherent instability in $\hat{\Sigma}$ given the small number of observations per group. Bootstrapping was performed through selection of random records (case resampling), which were then used to build models and variance-covariance matrices. Bootstrapping was performed in parallel using a `SNOW` cluster [52]. Bootstrapping was repeated four times – one for each forest class ($n = 40,000$ total). This ensured that the reference class (the withheld forest class) was the class of interest, yielding coefficient and standard error estimates that were more reliable than those derived by adding variances and covariances from the model error matrix. Previous testing indicated that this latter approach was unstable. We report BCa (bias-corrected) 95% confidence intervals [53] derived from the `boot.ci` function.

Model coefficients from mixed effects and beta regression modeling, are presented in Supplementary Table 1.

Table 1: Estimated coefficients and their standard errors (in parentheses) for models of above-ground biomass (AGB), basal area (BA), the number of living stems (A), diameter at breast height (DBH), the proportion of *Gilbertiodendron dewevrei* stems that were new recruits (pctRecruits), and the proportion of *Gilbertiodendron dewevrei* stems that were dead (pctMortality). Estimates are derived through bootstrapping as discussed above, with standard errors from percentile and BCa confidence intervals for linear mixed effects and beta regression models, respectively. The upper half of the table contains parameter estimates for the mean models while the lower half contains parameter estimates associated with residual variability and dispersion. (C) = categorical variable and (I) = interaction term. ProportionAGBForMax1st = the proportion of AGB held by the heaviest species in each quadrat; ProportionAGBForMax2nd = the proportion of AGB held by the second heaviest species in each quadrat; QMD = quadratic mean diameter; and the remaining entries are either self-explanatory or are described elsewhere.

Parameter	AGB	BA	Alive	DBH	pctRecruits	pctMortality
Intercept	1.039 (6.992)	-3.633 (0.214)	404.757 (39.232)	-1446.835 (69.89)	-2.906 (0.167)	-3.78 (0.122)
A	-0.015 (0.008)	0.004 (0)				
BA	19.783 (2.14)		-41.948 (19.243)		-0.181 (0.034)	-0.075 (0.027)
Mixed forest (C)	-1.569 (0.762)	-0.004 (0.041)	-9.939 (6.797)	-744.067 (206.603)	0.248 (0.13)	-0.33 (0.293)
Transitional forest (C)	-1.755 (0.85)	0.005 (0.039)	-6.804 (5.457)	44.787 (101.948)	0.028 (0.084)	-0.293 (0.115)
Riparian forest (C)	-2.715 (0.648)	-0.059 (0.039)	-14.384 (5.652)	-119.224 (141.763)	0.136 (0.099)	0.195 (0.107)
sqrt(BA)			714.156 (59.96)			
NumOfTrees					0.002 (0)	
pctMortality					3.651 (1.043)	
pctRecruits						1.471 (0.456)
ProportionAGBForMax1st		-0.161 (0.098)	-38.094 (15.294)			
ProportionAGBForMax2nd		-0.24 (0.153)	-35.406 (24.18)		0.657 (0.343)	
QMD	-0.952 (1.067)	0.57 (0.016)	-114.623 (3.356)			
edoro2 (C)	-0.244 (0.41)		12.809 (2.778)			
lenda1 (C)	1.987 (0.636)		-2.434 (5.444)			0.126 (0.076)
lenda2 (C)	1.421 (0.698)		3.294 (6.194)			0.357 (0.08)
Year	-0.034 (0.056)	0.029 (0.001)	-5.615 (0.218)	0.803 (0.035)	0.056 (0.008)	-0.015 (0.01)
Year:Mixed forest (I)	-0.021 (0.021)	-0.003 (0.001)	0.416 (0.217)	0.318 (0.104)	0.024 (0.03)	0.161 (0.052)
Year:Transitional forest (I)	0.034 (0.037)	-0.003 (0.002)	1.248 (0.431)	-0.061 (0.051)	-0.002 (0.016)	0.008 (0.021)
Year:Riparian forest (I)	-0.026 (0.029)	0.005 (0.002)	-1.263 (0.318)	0.046 (0.071)	-0.018 (0.019)	0.006 (0.02)
varExp (BA edoro1)	0.427 (0.033)	0.002 (0)				
varExp (BA edoro2)	0.688 (0.046)					
varExp (BA lenda1)	0.215 (0.036)					
varExp (BA lenda2)	0.492 (0.031)					
varExp (A)						
varPower (QMD Census1)			-1.415 (0.267)			
varPower (QMD Census2)			-0.982 (0.141)			
varPower (QMD Census3)			-1.008 (0.141)			
phi (Intercept)					1.737 (0.192)	3.094 (0.075)
phi Mixed forest (C)					-0.48 (0.196)	-0.173 (0.351)
phi Transitional forest (C)					0.188 (0.132)	0.138 (0.125)
phi Riparian forest (C)					-0.29 (0.159)	-0.213 (0.106)
phi pctMortality					-3.79 (1.762)	
phi pctRecruits						-0.882 (0.49)
phi lenda1 (C)					0.244 (0.201)	
phi lenda2 (C)					0.876 (0.197)	
phi Year					-0.02 (0.009)	0.014 (0.009)
phi Year:Mixed forest (I)						-0.222 (0.062)
phi Year:Transitional forest (I)						-0.003 (0.021)
phi Year:Riparian forest (I)						-0.017 (0.019)

4. Mechanistic synthesis

Collectively, our findings support the hypothesis that *Gilbertiodendron* is a controlling determinant of the structure of the Ituri Forest and, by extension, large portions of forested central Africa. However, attributing patterns of change to particular causal mechanism is problematic [54]. To some extent, Hart *et al.*'s [55, p. 557] sentiment that the “factors permitting the expansion of *G. dewevrei* remain mysterious”, still rings true. We did not directly explore the mechanisms by which this species succeeds, but are of the view that a complement of factors is at work [54]. Although early explorations

of tropical monodominance targeted isolated drivers [56], *Gilbertiodendron* dominance likely emerged, and is perpetuated through, at least three interconnected channels: (a) recovery or response from rare large-scale disturbance (fires), (b) ectomycorrhizal associations, and (c) stress tolerance traits.

Analyses of charcoal remains and pollen records indicate that fires have been present in the Ituri region periodically over the past 2,500 years, with a notable dry period prompting increased frequency and extent of fires in the range of 3,000–2,000 BP [55,57]. Fires were more common in current mixed species forest structures than in *Gilbertiodendron*-dominated forest in the Ituri region, as samples from the latter reflect the presence of fire only twice in the last 1,000 years, and not at all in the last 500 years. Hart *et al.* [55] found that moist evergreen forest was predominate in the Ituri forest over the past two millennia, though the species composition has shifted in critical ways. They found through charcoal analysis that seven of eleven of the most abundant modern tree species, including *Gilbertiodendron*, were not present in the Ituri Forest as recently as two centuries ago. This is in contrast to Torti *et al.* [56], who identified continued presence of *Gilbertiodendron*-dominated forest in the Sangha River region (west of Ituri, in Republic of the Congo) over the past 2,700 years. Though, like Ituri, these findings also indicate an absence of fire over the past two millennia. Tovar *et al.*'s [58] recent work also identified very little change in the vegetative composition across a number of sites in the past 2,700 years, though *Gilbertiodendron* pollen made up only a small proportion of all pollen at different points in time (possibly due to limitations in pollen dispersal).

Hart *et al.* [55] propose that drier periods (see also [58]) may have led to the establishment of a mosaic of different forest types, perhaps through opportunistic channels and refugia [57,59]. Oliver and Larson [60] provide support for this idea, acknowledging the ubiquity of fire as an allogenic disturbance agent. They speak to the manner in which fires, whether numerous or large, can lead to a spatial mosaic of remnant conditions, and they reference Davis [61] and Lopez-Portillo *et al.* [62] in demonstrating that even tropical forests can burn during unusually dry periods. However, the expansion of *Gilbertiodendron* is still remarkable given (a) what we would expect of the species' growth and reproductive characteristics, and (b) that no historical *Gilbertiodendron* was identified in any of Hart *et al.*'s [55] 279 charcoal-producing Ituri test pits, dated to 130 ± 70 to 4190 ± 160 years before present. These factors suggest that *Gilbertiodendron* possesses physiological and life-history traits that afford it a competitive advantage as a persistent monodominant, and possibly, that its colonization is more recent than its physiology and life-history might otherwise suggest [55].

Connell and Lowman [63] identify two broad classes of tropical monodominance based on whether the dominant canopy species persists beyond a single generation in the absence of disturbance (type I; e.g., [64]), or whether it is supplanted by other species after a single generation (type II, e.g., [65]). Newbery *et al.* [66] add a third type corresponding to transient monodominants that depend on disturbance to gain or retain dominance (e.g., [67]). *Gilbertiodendron*-dominated forest of the Ituri region reflects type I monodominance [63,68], and a number of characteristics support multi-generational persistence. One life history trait that may afford *Gilbertiodendron* a competitive advantage over competitors is its ectomycorrhizae association (EM) with fungal symbionts [56,58,63,69]. An EM association generally affords its host improved protection against natural enemies and harmful physical factors, and can provide access to nutrients before they are available to vesicular-arbuscular mycorrhizal (VAM) associations [63]. These benefits may provide EM host species an advantage over VAM host species, allowing the former to both replace themselves and displace VAM host species over time [56,59,63]. Not all tropical monodominant species possess an EM association [56,69], however *Gilbertiodendron* [63,69] and *Julbernardia seretii* both do, while some other local species do not [56,68,69]. An EM association is neither a necessary nor a sufficient explanation of *Gilbertiodendron* monodominance when considered in isolation [56]. However, it may afford the species a competitive advantage over others [59] in the region's nutrient-poor soils [69].

While historical landscape disturbance might have presented an opportunity for *Gilbertiodendron* to gain a small foothold [55,57,59], and EM association may enhance success relative to other established competitors [56,58,63,69], *Gilbertiodendron*'s stress-tolerance traits [56] may be of greater importance in maintaining and expanding the species' territory. Seed mortality is high for both *Gilbertiodendron* and *Julbernardia seretii*, a common canopy dominant in the mixed species forest [68]. The production of intermittent mast crops does help to offset losses through satiation of mammal predators, but pre-dispersal seed predation by specialized beetles leads to low germination rates for both of these species [58,68]. However, Hart [68] found that successful germination led to increased survival rates for *Gilbertiodendron*, relative to *Julbernardia seretii*, across both the short- and long-term. This is at least partly attributable to tolerance of extremely low light levels. Photosynthetically active radiation in the *Gilbertiodendron*-dominated understory is less than half that of mixed forest, and given the homogeneity of the canopy, it is more consistently limited as well [56]. In both controlled nursery experiments [69] and field experiments [70], *Gilbertiodendron* has proven to remain stable in a range of light conditions, from full sunlight [69] to a photosynthetic photo flux density of $2.8 \text{ mol/m}^2/\text{day}$ (approximately 5.3% of full sunlight) required for seedling maintenance [70]. If we consider that the highest rates of seed survival for *Gilbertiodendron*

are in the transition zones at monodominant patch boundaries [68], we can begin to see a logical set of mechanisms by which this species achieves dominance and spreads across the landscape:

(1) *Gilbertiodendron* is initially present in the area, perhaps through historical refugia. (2) For productive trees, seed survival and germination are greatest in areas adjacent to monodominant forest. (3) Tolerance of low (and high) light conditions, paired with increased access to nutrients through an EM association, enables seedlings to continue to grow slowly [71] beneath early-emergent, faster growing, or more light-demanding species that also occupy the transition zones. These traits also permit success through opportunistic channels, such as within brighter canopy gaps. (4) In time, *Gilbertiodendron* overtakes competitors while effectively suppressing their future growth through densification of the canopy and reduction of understory light [69]. This combination of factors indirectly suggests that *Gilbertiodendron* does not necessarily possess enhanced resistance to pathogens, parasites, or animal predators [63], but instead, that it has locally enhanced access to certain resources (i.e., soil nutrients), and that it indirectly increases stress on competitors [56,69]. That such a suite of mechanisms may drive forest composition and structure is not unique among tropical forests [72,73].

Changes in the species composition and structural dynamics of tropical forests have often been attributed, whether directly or indirectly, to changes in climate and CO₂ fertilization [54,74–76]. We certainly can't exclude the effects of anthropogenically altered atmospheric chemistry, but on a local scale we see no reason to believe there would be a difference in the atmospheric conditions between *Gilbertiodendron*-dominated and mixed species forest types at Ituri. The effects of a changing climate are more likely to have an effect through the shifting cycles of temperature and precipitation [54], as has been common in central Africa since the Holocene [59]. These, in turn, may influence the duration and frequency of fire, drought, and inundation, which may have more marked effects on floristic composition than the changing atmosphere itself. However, according to Hart *et al.* [55], the patterns of floristic change over the past four millennia do not actually reflect any directional climatic change. Recent findings [77] suggest that elevation, perhaps through its relationship with climate, is responsible for differences between patterns and trends in Paleotropical and Neotropical forests, though again, differences in elevation across the four Ituri plots are insufficient to offer an explanatory mechanism. Chave *et al.* [72] remark that developmental changes in species composition are generally slower than those in stand structure, though on a centurial or geologic time scale, it is possible that the type of monodominance we see in *Gilbertiodendron* forest today could appear and disappear regularly [55,56].

Data Availability Statement: The research relied on data from the Ituri Forest Dynamic Plots curated by the Smithsonian Tropical Research Institute's Forest Global Earth Observatory (ForestGEO), currently comprised of 67 long-term forest monitoring plots. The data is available to researchers upon reasonable request through the ForestGEO data request platform, currently located at <https://forestgeo.si.edu/sites/africa/ituri>, and through correspondence with the plot principal investigators: Corneille E. N. Ewango (corneilleewango@gmail.com) and Jean-Remy Makana (jeanremymakana@gmail.com). Limitations of use and requirements associated with publication of analyses derived from the Ituri plot data can be found at the aforementioned URL.

References

1. Makana, J.R.; Ewango, C.N.; McMahon, S.M.; Thomas, S.C.; Hart, T.B.; Condit, R. Demography and biomass change in monodominant and mixed old-growth forest of the Congo. *Journal of Tropical Ecology* **2011**, *27*, 447–461. doi:<https://doi.org/10.1017/S0266467411000261>.
2. Lewis, S.L.; Lopez-Gonzalez, G.; Sonké, B.; Affum-Baffoe, K.; Baker, T.R.; Ojo, L.O.; Phillips, O.L.; Reitsma, J.M.; White, L.; Comiskey, J.A. Increasing carbon storage in intact African tropical forests. *Nature* **2009**, *457*, 1003. doi:<https://doi.org/10.1038/nature07771>.
3. R Core Team. *R: A language and environment for statistical computing*; R Foundation for Statistical Computing, 2019.
4. RStudio Team. *RStudio: Integrated development environment for R*; RStudio, Boston, MA, 2019.
5. Makana, J.R.; Hart, T.; Liengola, I.; Ewango, C.; Hart, J.; Condit, R. Ituri forest dynamics plots, Democratic Republic of Congo. In *Tropical Forest Diversity and Dynamism: Findings from a Large-Scale Plot Network*; University of Chicago Press: Chicago, USA, 2004; chapter 28, pp. 492–505.
6. R Core Team. Package parallel. In *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing, 2019.
7. Microsoft Corporation.; Weston, S. *doSNOW: Foreach Parallel Adaptor for the 'snow' Package*; Revolution Analytics, Comprehensive R Archive Network (CRAN), 2017.
8. Baddeley, A.; Rubak, E.; Turner, R. *Spatial Point Patterns: Methodology and Applications with R*; Chapman and Hall/CRC Press: London, UK, 2015.
9. Baddeley, A.; Turner, R. spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software* **2005**, *12*, 1–42. doi:<https://doi.org/http://dx.doi.org/10.18637/jss.v012.i06>.
10. Pebesma, E.J.; Bivand, R.S. Classes and methods for spatial data in R. *R News* **2005**, *5*.
11. Bivand, R.S.; Pebesma, E.J.; Gomez-Rubio, V. *Applied Spatial Data Analysis with R*, 2nd ed.; Springer: New York, NY, 2013.
12. Gregoire, T.G.; Schabenberger, O.; Barrett, J.P. Linear modelling of irregularly spaced, unbalanced, longitudinal data from permanent-plot measurements. *Canadian Journal of Forest Research* **1995**, *25*, 137–156. doi:<https://doi.org/10.1139/x95-017>.

13. Fitzmaurice, G.; Laird, N.; Ware, J. *Applied Longitudinal Analysis*; Wiley Series in Probability and Statistics, John Wiley & Sons, Inc.: Hoboken, NJ, 2004.
14. Graybill, F.A. *Introduction to Matrices with Applications in Statistics*; Wadsworth Publishing: Belmont, CA, 1969.
15. Bates, D.; Mächler, M.; Bolker, B.M.; Walker, S.C. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **2015**, *67*, 1–48. doi:<https://doi.org/10.18673/jss.v067.i01>.
16. Pinheiro, J.C.; Bates, D. *Mixed-Effects Models in S and S-PLUS*; Springer-Verlag: New York, NY, 2009.
17. Pinheiro, J.; Bates, D.; DebRoy, S.; authors, E.; Heisterkamp, S.; Van Willigen, B.; core Team, R. Package 'nlme'. Report, Central R Area Network, 2019.
18. Neter, J.; Kutner, M.; Wasserman, W.; Nachtsheim, C. *Applied Linear Statistical Models*; Irwin Series in Statistics, McGraw-Hill/Irwin, 1996; p. 1408.
19. Gregoire, T.G. Nonparallel regressions with indicator variables. *International Journal of Statistics and Probability* **2015**, *4*, 46–54.
20. Akaike, H. Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*; Parzen, E.; Tanabe, K.; Kitagawa, G., Eds.; Springer, 1998; pp. 199–213. doi:https://doi.org/10.1007/978-1-4612-1694-0_15.
21. Schwarz, G. Estimating the dimension of a model. *The Annals of Statistics* **1978**, *6*, 461–464. doi:<https://doi.org/10.1214/aos/1176344136>.
22. Bivand, R.S.; Wong, D.W.S. Comparing implementations of global and local indicators. *TEST* **2018**, *27*, 716–748. doi:<https://doi.org/10.1007/018-0599-x>.
23. Pebesma, E.J. Multivariable geostatistics in S: The gstat package. *Computers & Geosciences* **2004**, *30*, 683–691. doi:<https://doi.org/10.1016/j.cageo.2004.03.001>.
24. Gräler, B.; Pebesma, E.J.; Heuvelink, G. Spatio-Temporal Interpolation using gstat. *The R Journal* **2016**, *8*, 204–218.
25. Bolker, B.M.; Others. *GLMM FAQ*; Ben Bolker via [github.io](https://github.com/bbolker/glmm-faq), 2019.
26. Bates, D.; Kliegl, R.; Vasishth, S.; Baayen, R.H. Parsimonious mixed models. *arXiv Preprint Server* **2018**, *arXiv:1506.04967v2*, 1–21.
27. Loy, A.; Steele, S. *Imeresampler: Bootstrap methods for nested linear mixed-effects models*; A. Loy via [github.io](https://github.com/aloymat/imeresampler) [previous version available from Comprehensive R Archive Network (CRAN)], 2019.
28. Kieschnick, R.; McCullough, B.D. Regression analysis of variates observed on (0, 1): Percentages, proportions and fractions. *Statistical Modelling* **2003**, *3*, 193–213. doi:<https://doi.org/10.1191/10.1191/0962280203sm530a>.
29. Ferrari, S.; Cribari-Neto, F. Beta regression for modelling rates and proportions. *Journal of Applied Statistics* **2004**, *31*, 799–815. doi:<https://doi.org/10.1080/0266476042000214501>.
30. Pereira, T.L.; Cribari-Neto, F. Detecting model misspecification in inflated beta regressions. *Communications in Statistics - Simulation and Computation* **2014**, *43*, 631–656. doi:<https://doi.org/10.1080/03610918.2012.712183>.
31. Espinheira, P.L.; Ferrari, S.L.P.; Cribari-Neto, F. On beta regression residuals. *Journal of Applied Statistics* **2008**, *35*, 407–419. doi:<https://doi.org/10.1080/02664760701834931>.
32. Simas, A.B.; Barreto-Souza, W.; Rocha, A.V. Improved estimators for a general class of beta regression models. *Computational Statistics & Data Analysis* **2010**, *54*, 348–366. doi:<https://doi.org/10.1016/j.csda.2009.08.017>.
33. Cribari-Neto, F.; Zeileis, A. Beta regression in R. Report, Vienna University of Economics and Business, 2009.
34. Cribari-Neto, F.; Zeileis, A. Beta regression in R [betareg vignette]. Report, Universidade Federal de Pernambuco, Universität at Innsbruck, 2019.
35. Berger, S.; Graham, N.; Zeileis, A. Various versatile variances: An object-oriented implementation of clustered covariances in R. Report, Working Papers in Economics and Statistics, 2017–12, 2017.
36. Driscoll, J.C.; Kraay, A.C. Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics* **1998**, *80*, 549–560. doi:<https://doi.org/10.1162/003465398557825>.
37. Arellano, M. Practitioner's Corner: Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics* **1987**, *49*, 431–434. doi:<https://doi.org/10.1111/j.1468-0084.1987.mp49004006.x>.
38. Liang, K.Y.; Zeger, S.L. Longitudinal data analysis using generalized linear models. *Biometrika* **1986**, *73*, 13–22. doi:<https://doi.org/10.1093/biomet/73.1.13>.
39. Hoechle, D. Robust standard errors for panel regressions with cross-sectional dependence. *The Stata Journal* **2007**, *7*, 281–312. doi:<https://doi.org/10.1177/1536867X0700700301>.
40. Zeileis, A. Econometric computing with HC and HAC covariance matrix estimators. *Journal of Statistical Software* **2004**, *11*, 1–17. doi:<https://doi.org/10.18637/jss.v011.i10>.
41. Zeileis, A. Object-oriented computation of sandwich estimators. *Journal of Statistical Software* **2006**, *16*, 1–16. doi:<https://doi.org/10.18637/jss.v016.i01>.
42. Rogers, W. Regression standard errors in clustered samples. *Stata technical bulletin* **1993**, *13*, 19–23.
43. Cameron, A.C.; Gelbach, J.B.; Miller, D.L. Robust inference with multiway clustering. *Journal of Business & Economic Statistics* **2011**, *29*, 238–249. doi:<https://doi.org/10.1198/jbes.2010.07136>.
44. Cribari-Neto, F.; Zeileis, A. Beta regression in R. *Journal of Statistical Software* **2010**, *34*, 1–24.
45. Bayer, F.M.; Cribari-Neto, F. Model selection criteria in beta regression with varying dispersion. *Communications in Statistics - Simulation and Computation* **2017**, *46*, 729–746. doi:<https://doi.org/10.1080/03610918.2014.977918>.
46. McQuarrie, A.D.R.; Tsai, C.L. *Regression and Times Series Model Selection*; World Scientific: New Jersey, 1998.
47. Nagelkerke, N.J. A note on a general definition of the coefficient of determination. *Biometrika* **1991**, *78*, 691–692. doi:<https://doi.org/10.1093/biomet/78.3.691>.
48. Long, J.S. *Regression models for categorical and limited dependent variables*, 2nd ed.; SAGE Publications: Thousand Oaks, CA, 1997.
49. Ramsey, J.B. Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society: Series B (Methodological)* **1969**, *31*, 350–371. doi:<https://doi.org/10.1111/j.2517-6161.1969.tb00796.x>.
50. Canty, A.; Ripley, B. *boot: Bootstrap R (S-Plus) Functions*; Comprehensive R Archive Network (CRAN), 2017.

51. Davidson, A.C.; Hinkley, D.V. *Bootstrap Methods and Their Applications*; Cambridge University Press: Cambridge, UK, 1997.
52. Tierney, L.; Rossini, A.J.; Li, N.; Sevcikova, H. *snow: Simple Network of Workstations*; Central R Area Network, 2018.
53. Efron, B. Better bootstrap confidence intervals. *Journal of the American statistical Association* **1987**, *82*, 171–185. doi:https://doi.org/10.1080/01621459.1987.10490131
54. Lewis, S.L.; Malhi, Y.; Phillips, O.L. Fingerprinting the impacts of global change on tropical forests. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **2004**, *359*, 437–462. doi:https://doi.org/10.1098/rstb.2003.1432.
55. Hart, T.B.; Hart, J.A.; Dechamps, R.; Fournier, M.; Ataholo, M. Changes in forest composition over the last 4000 years in the Ituri basin, Zaire. In *The Biodiversity of African Plants*; van der Maesen, L.J.G.; van der Burgt, X.M.; van Medenbach de Rooy, J.M., Eds.; Springer, Dordrecht: Wageningen, The Netherlands, 1996; pp. 545–560. doi:https://doi.org/10.1007/978-94-009-0285-5.
56. Torti, S.D.; Coley, P.D.; Kursar, T.A. Causes and consequences of monodominance in tropical lowland forests. *The American Naturalist* **2001**, *157*, 141–153. doi:https://doi.org/10.1086/318629.
57. Maley, J. A catastrophic destruction of African forests about 2,500 years ago still exerts a major influence on present vegetation formations. *IDS bulletin* **2002**, *33*, 13–30. doi:https://doi.org/10.1111/j.1759-5436.2002.tb00003.x.
58. Tovar, C.; Harris, D.J.; Breman, E.; Brncic, T.; Willis, K.J. Tropical monodominant forest resilience to climate change in Central Africa: A *Gilbertiodendron dewevrei* forest pollen record over the past 2,700 years. *Journal of Vegetation Science* **2019**, *30*, 575–586. doi:https://doi.org/10.1111/jvs.12746.
59. Hart, T.B. Monospecific dominance in tropical rain forests. *Trends in Ecology & Evolution* **1990**, *5*, 6–11. doi:https://doi.org/10.1016/0169-5347(90)90005-X.
60. Oliver, C.D.; Larson, B.C. *Forest Stand Dynamics*; John Wiley & Sons Inc: Hoboken, NJ, 1996.
61. Davis, J.B. Burning another empire. *Fire Management Notes* **1984**, *45*, 12–17.
62. Lopez-Portilla, J.; Keyes, M.R.; Gonzales, A.; Cabrera, E.; Sanchez, O. Los incendios de Quintana Roo: Catastrofe ecologia o evento periodico? *Ciencia Y Desarrollo* **1990**, *16*, 43–57.
63. Connell, J.H.; Lowman, M.D. Low-diversity tropical rain forests: Some possible mechanisms for their existence. *The American Naturalist* **1989**, *134*, 88–119. doi:https://doi.org/10.1086/284967.
64. Fonty, É.; Molino, J.F.; Prévost, M.F.; Sabatier, D. A new case of neotropical monodominant forest: *Spirotropis longifolia* (Leguminosae-Papilionoideae) in French Guiana. *Journal of Tropical Ecology* **2011**, *27*, 641–644. doi:https://doi.org/10.1017/S0266467411000356.
65. Anderson, J.A.R. The destruction of *Shorea albida* forest by an unidentified insect. *Empire Forestry Review* **1961**, pp. 19–29.
66. Newbery, D.; Van Der Burgt, X.; Worbes, M.; Chuyong, G. Transient dominance in a central African rain forest. *Ecological Monographs* **2013**, *83*, 339–382. doi:https://doi.org/10.1890/12-1699.1.
67. Demenois, J.; Ibanez, T.; Read, J.; Carriconde, F. Comparison of two monodominant species in New Caledonia: Floristic diversity and ecological strategies of *Arillastrum gummiferum* (Myrtaceae) and *Nothofagus aequilateralis* (Nothofagaceae) rainforests. *Australian Journal of Botany* **2017**, *65*, 11–21. doi:https://doi.org/10.1071/BT16125.
68. Hart, T.B. Seed, seedling and sub-canopy survival in monodominant and mixed forests of the Ituri Forest, Africa. *Journal of Tropical Ecology* **1995**, *11*, 443–459. doi:https://doi.org/10.1017/S0266467400008919.
69. Hall, J.S.; Harris, D.J.; Saltonstall, K.; Medjibe, V.d.P.; Ashton, M.S.; Turner, B.L. Resource acquisition strategies facilitate *Gilbertiodendron dewevrei* monodominance in African lowland forests. *Journal of Ecology* **2019**. doi:https://doi.org/10.1111/1365-2745.13278.
70. Makana, J.R.; Thomas, S.C. Effects of Light Gaps and Litter Removal on the Seedling Performance of Six African Timber Species. *Biotropica: The Journal of Biology and Conservation* **2005**, *37*, 227–237. doi:https://doi.org/10.1111/j.1744-7429.2005.00030.x.
71. Hart, T.B.; Hart, J.A.; Murphy, P.G. Monodominant and species-rich forests of the humid tropics: Causes for their co-occurrence. *The American Naturalist* **1989**, *133*, 613–633. doi:https://doi.org/10.1086/284941.
72. Chave, J.; Condit, R.; Muller-Landau, H.C.; Thomas, S.C.; Ashton, P.S.; Bunyavejchewin, S.; Dattaraja, H.S.; Davies, S.J.; Esufali, S.; Ewango, C.E. Assessing evidence for a pervasive alteration in tropical tree communities. *PLoS biology* **2008**, *6*, e45. doi:https://dx.doi.org/10.1371/journal.pbio.0060045.
73. Henkel, T.W. Monodominance in the ectomycorrhizal *Dicymbe corymbosa* (Caesalpiniaceae) from Guyana. *Journal of Tropical Ecology* **2003**, *19*, 417–437. doi:https://doi.org/10.1017/S0266467403003468.
74. Phillips, O.L.; Malhi, Y.; Higuchi, N.; Laurance, W.F.; Núñez, P.V.; Vásquez, R.M.; Laurance, S.G.; Ferreira, L.V.; Stern, M.; Brown, S. Changes in the carbon balance of tropical forests: Evidence from long-term plots. *Science* **1998**, *282*, 439–442. doi:https://doi.org/10.1126/science.282.5388.439.
75. Baker, T.R.; Phillips, O.L.; Malhi, Y.; Almeida, S.; Arroyo, L.; Di Fiore, A.; Erwin, T.; Higuchi, N.; Killeen, T.J.; Laurance, S.G. Increasing biomass in Amazonian forest plots. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **2004**, *359*, 353–365. doi:https://doi.org/10.1098/rstb.2003.1422.
76. Brienen, R.J.; Phillips, O.L.; Feldpausch, T.R.; Gloor, E.; Baker, T.R.; Lloyd, J.; Lopez-Gonzalez, G.; Monteagudo-Mendoza, A.; Malhi, Y.; Lewis, S.L.; others. Long-term decline of the Amazon carbon sink. *Nature* **2015**, *519*, 344. doi:https://doi.org/10.1038/nature14283.
77. Hubau, W.; L., L.S.; Phillips, O.L.; Affum-Baffoe, K.; Beeckman, H.; Cuní-Sánchez, A.; Ewango, C.E.N.; Fauset, S.; Sheil, D.; Sonke, B.; Sullivan, M.J.P.; Sunderland, T.C.H.; Thomas, S.C.; Abernethy, K.A.; Adu-Bredu, S.; Amani, C.A.; Baker, T.R.; Banin, L.F.; Baya, F.; Begne, S.K.; Bennett, A.C.; Benedet, F.; Bitariho, R.; Bocko, Y.E.; others. Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature* **2020**, *579*, 80–87. doi:https://doi.org/10.1038/s41586-020-2035-0.