Modeling and simulation of tree spatial patterns in an oak-hickory forest with a modular, hierarchical spatial point process framework

Andrew J. Lister[a],*, Laura P. Leites[b]

[a] US Forest Service, 11 Campus Blvd, Ste. 200, Newtown Square, PA 19073, United States
[b] Penn State University, 312 Forest Resources Bldg., University Park, PA 16802, United States

ABSTRACT

Modeling and simulating tree spatial distribution in complex forests is important to ecologists and applied scientists who seek to both understand pattern-creating biological processes and create realistic model forests that can be used for hypothesis testing and sampling experiments. Several patterns of tree spatial distribution can co-occur in a forest. Clustering can occur due to localized patterns of growth and mortality of larger trees and corresponding regeneration of smaller trees, while trees of medium size can exhibit more uniform patterns. Inter-tree interaction may be characterized by asymmetry of competitive strength, with larger individuals having a disproportionate influence on smaller individuals.

Many point process modeling approaches exist, but few have incorporated hierarchical principles that describe inter-tree competition. Those that do sometimes assume symmetric interaction among trees, which can be unrealistic. None of the existing models allow for the use of different model types at each level of the hierarchy, something that could provide a more realistic representation of the patterns displayed by trees of different size. In this study, we model and simulate a forest using a novel, modular, hierarchical approach that allows for the use of different model types at each hierarchy level, and incorporates asymmetrical interactions as well as the effects of environmental covariates. The forest is a mid-successional 8-ha stem-mapped oak-hickory watershed in Pennsylvania, USA. Results suggest that asymmetrical interactions based on tree size do exist, and these are mediated by the effects of topography. The hierarchical models reproduce the spatial patterns found in the original data better than non-hierarchical versions of the same models. The flexibility afforded by the modularity of our modeling framework will allow simulation of forests with varying levels of complexity as well as the testing of ecological hypotheses about drivers of spatial pattern creation.

1. Introduction

Ecologists and forest managers model spatial patterns of forest trees for two main reasons: to understand the processes that lead to forest community development, and to understand the effects of stand spatial structure on growth of individual trees. Spatial patterns can be governed by a mixture of chance events and plant-plant interactions mediated by environmental gradients (Bormann & Likens, 1994). As competition theory suggests, patterns of soil resource availability and propagules influence initial stand composition, and as plants' zones of influence begin to overlap, competitive or facilitative hierarchies develop, with interaction strength related to life history strategy, plant size, and interplant distances (Bella, 1971; Brooker et al., 2008; Schwinning & Weiner, 1998; Wu et al., 1985). In mixed deciduous northern hardwood forests, these processes are thought to be mediated by the effects of topography and its effects on light and soil conditions

(Frey et al., 2007). Spatial pattern dynamics reflect these processes, with younger stands exhibiting a more clustered pattern, followed by stand homogenization as competition intensifies, followed by a different type of clustered pattern as mortality and gap dynamics become important (Larson et al., 2015; Raventos et al., 2010). Analysis of the relationship between tree size, location, and environmental gradients can thus advance the science of plant community ecology by supporting existing hypotheses or suggesting new ones in cases where observed patterns and relationships do not align with current theory.

Point process models are useful tools for studies of spatial patterns. They are particularly appealing because they can be used to describe the spatial or temporal structure of a phenomenon and to create simulated point patterns based on the spatial distribution of observed points. In point process modeling, the set of locations of events (e.g., tree locations) is seen as one realization of a stochastic event-generating process. The intensities or density functions of these processes can be

modeled using various families of statistical models that incorporate randomness, environmental covariates, and inter-point interaction functions.

The selection of an adequate point process model depends on the system under study. For example, a homogeneous Poisson process (HPP) model is typically used as a type of null model of spatial pattern; descriptors for observed patterns are compared against those generated from an HPP model in order to diagnose a non-random spatial distribution of points (Baddeley et al., 2015; Wiegand & Moloney, 2014). In systems where environmental trends like topographic gradients affect patterns, inhomogeneous Poisson process (IPP) models will better describe the observed patterns (e.g. Getzin et al., 2008). For applications like forest ecology where clustered patterns are common (Grabarnik & Särkkä, 2009; Stoyan & Penttinen, 2000), the family of cluster models might be more appropriate. Finally, a type of Markovian model commonly referred to as a Gibbs model relies on a pair potential function that specifies a set of symmetrical inter-point interaction strengths, and is often used for patterns where individuals are dispersed.

In order to incorporate in point process models the potential for asymmetric, hierarchical relationships that commonly exist in forests, Högmander and Särkkä (1999) proposed a modeling system in which the intensity of the points in the highest level of the hierarchy is modeled first using a Gibbs point process model, then the lower levels of the hierarchy are modeled as subsequent Gibbs processes, conditional on the intensity of the higher levels. Illian et al. (2009), Genet et al. (2014), and Grabarnik and Särkkä (2009) used a similar Gibbsian approach, defining hierarchies by reproductive strategy (the former), and tree size (the two latter). Although Gibbs models have many advantages, they are not suitable for the variety of clustered patterns that occur in forests (Baddeley et al., 2015; Illian et al., 2008; Stoyan & Penttinen, 2000). In addition, multi-type Gibbs models are parameterized with information about the relative strengths of interaction between individuals of different types, something that might not be known to the researcher a priori (Baddeley et al., 2015; Illian et al., 2009; Prokešová et al., 2006; Wiegand & Moloney, 2014). These difficulties suggest that Gibbs models are not always suitable for modeling forest stands where a combination of different pattern types, hierarchical relationships, and environmental trends coexist. We argue that a more flexible approach is needed.

To address the needs for modeling and simulating complex forests where hierarchy and different spatial patterns may co-occur, we propose a modular, hierarchical point process modeling framework (MHPPF). In this approach, it is assumed that trees of different size classes may display different spatial patterns and thus that different families of point processes and environmental variables should be used to describe them. This is achieved by the modularity of our approach: at each level of the tree size-based hierarchy a different model type and covariate can be used. More specifically, a multivariate inhomogeneous Poisson or cluster process is constructed by the superposition of independent point processes that are conditioned on the cumulative point intensities of the higher levels of the hierarchy and/or environmental covariates. Although the form of the hierarchy can be constructed following different theories or hypotheses, the approach we present here assumes that trees of different size classes interact asymmetrically in the direction of decreasing size. Asymmetric competition is incorporated by adding the intensity of the higher levels in the hierarchy as covariates in the modeling of lower hierarchy levels. That is, the locations of trees lower in the hierarchy are dependent on the locations of trees at higher levels, but not vice-versa. A generalized joint probability density associated with such a framework can be expressed as

$$f(x) = \alpha \prod_{i=1}^{j} \begin{cases} f_i(S_i|\theta_i) & \text{if } i = j \\ f_i(S_i|S_{i+1}...S_j, \theta_i) & \text{if } i \neq j \end{cases} \quad (1)$$

where x is the combined pattern of all levels, α is a normalizing constant, i is the hierarchy level (1 is the lowest and the smallest tree size class), j is the number of levels, S_i is the intensity of the level i pattern, and θ_i is the vector of parameters associated with the environmental covariates associated with the intensity of the level i pattern. Simulation is achieved by performing the modeling and simulation sequentially based on the hierarchy of point types, incorporating the simulated intensity surfaces from higher levels as inputs in the form of covariates to the simulation process at lower levels.

In the current study, we lay out the methods for implementing this approach, demonstrate its application by modeling a mapped stand of trees in central Pennsylvania, USA, and compare results of this new approach to a method that does not incorporate hierarchical principles. We hypothesized that trees of different size classes would exhibit different types of spatial patterns, and that these patterns would be related to not only the presence of larger trees, but also to environmental covariates. Specifically, we hypothesized that in this mid-successional forest, smaller trees would be clustered, and larger trees would show a more hyper-dispersed pattern and that the MHPPF would allow us to simulate realistic tree spatial patterns in this complex forest.

2. Methods

2.1. Study site and data

Mapped tree data for this study were obtained from the Susquehanna Shale Hills Critical Zone Observatory (SSHCZO) (Kaye et al., 2015), an approximately 8-ha watershed found in the Ridge and Valley physiographic region of central Pennsylvania, USA (Fig. 1). The watershed is oriented east-west, with predominantly north- and south-facing aspects and an elevation range of 240–300 m above sea level. Tree information for 2050 trees was collected in 2012 and included geographic location obtained using a survey-grade GPS and data logger, species, and diameter at breast height (DBH) to the nearest 0.254 cm (0.1 in) for trees greater than 20.32 cm (8 in) (Kaye et al., 2015; Naithani et al., 2013; Wubbles, 2010).

To define each hierarchy level, trees were grouped into diameter classes, with the assumption that DBH is related to competitive strength and thus different sized trees may display different spatial patterns. To create tree size groupings, the Jenks optimization procedure found in the classInt package for version 3.3.3 of the R statistical software (Bivand, 2015; Jenks & Caspall, 1971; R Core Team, 2015) was used. We chose the Jenks method because it provides a site-specific way to classify DBH into natural groups that are internally homogeneous relative to other potential groupings, making it a generic approach for conducting this type of analysis in new forest ecosystems with different DBH distributions. The four DBH class boundaries identified by this method were as follows: 20.3–29.1 (n = 744), 29.2–37.6 (n = 682), 37.7–48.5 (n = 442), and ≥48.6 cm (n = 182). The convention we use here is that large diameter classes represent higher levels in the competitive hierarchy, i.e. level 4 is the highest level of the hierarchy and level 1 is the lowest.

To evaluate the potential effect of topographic gradients on the spatial patterns of the different size-class trees, we evaluated elevation (USGS, 2010), percent slope, and transformed aspect (Beers et al., 1966), all of which were based on a 3-m pixel digital elevation model (DEM) and generated using the Spatial Analyst extension of ArcGIS Desktop software (ESRI, 2014). In addition, two topographic indices were generated using the terrain function of the raster package of R (Hijmans, 2015): topographic position index (TPI) and topographic ruggedness index (TRI) (Wilson et al., 2007). TPI represents an index of relative slope position, and TRI is an index of local topographic complexity. All DEM-based variables were smoothed using the focal mean function applied in a 5 × 5 pixel window in R's raster package (Hijmans, 2015).

To incorporate potential tree-size-mediated competition effects, fixed-bandwidth kernel estimates of the intensity surfaces of levels 1–4
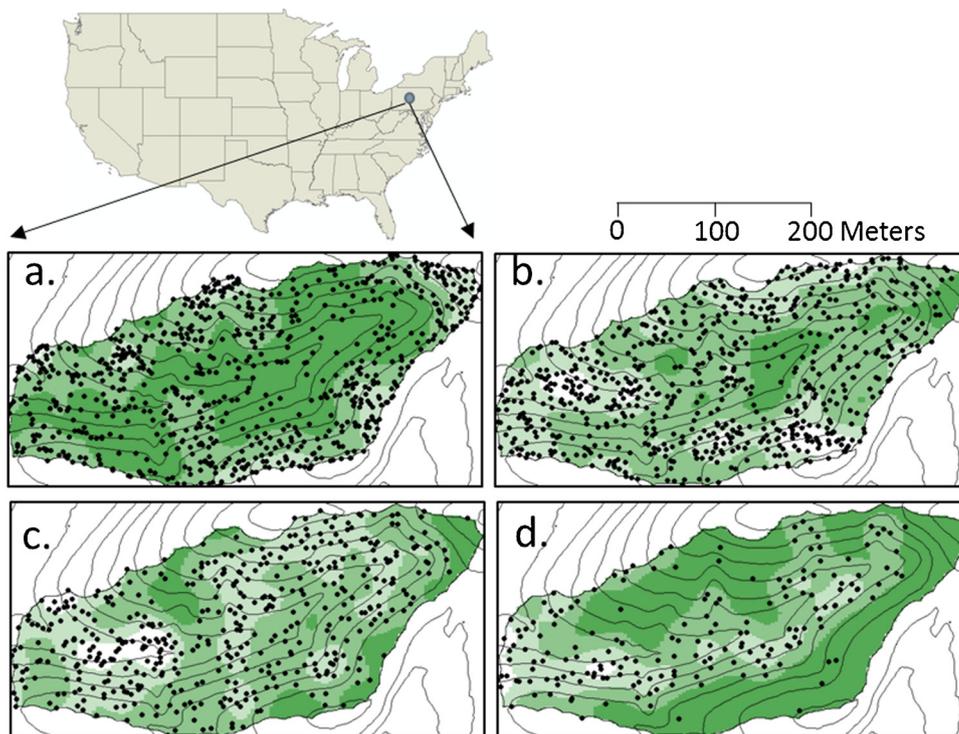
Fig. 1. Location of the study area with tree locations and elevation contours in Pennsylvania, USA. Locations of trees for each level of the hierarchy (DBH class) are superimposed on kernel density estimates of intensity of the point patterns. a = level 1 (20.3–29.1 cm), b = level 2 (29.2–37.6 cm), c = level 3 (37.7–48.5 cm), and d = level 4 (> 48.6 cm).

individually, 1 + 2, 1 + 2 + 3, 2 + 3, 2 + 3 + 4, and 3 + 4 were created. The kernel estimation method is a neighborhood-based spatial interpolation that weights neighbors using an isotropic Gaussian smoothing kernel with standard deviation of 15 m. The size of the smoothing kernel was chosen using a cross validation approach described by Berman and Diggle (1989). These intensity surfaces were then evaluated as covariates in the respective hierarchy-level model.

2.2. Baseline analysis to characterize patterns

In order to choose the best point process family at each level of the hierarchy (DBH class), we conducted assessments using the spatstat package for R statistical software (Baddeley et al., 2015; R Core Team, 2015). We first determined if the observed patterns for each DBH class exhibited evidence for complete spatial randomness (CSR), hyper-dispersion or clustering by comparing the observed values of three second order pattern descriptors, the pair correlation statistic g (Wiegand & Moloney, 2014), Besag's L (Besag, 1977), and the nearest neighbor statistic G (Wiegand & Moloney, 2014), to simulation envelopes created from 99 realizations of a CSR pattern with the same number of points. If the observed functions fell within the simulation envelopes, this suggested CSR. If the functions fell above them, this suggested clustering. If they fell below, this suggested a hyper-dispersed or inhibited pattern (Baddeley et al., 2015). We used two-sided maximum absolute deviation tests (Baddeley et al., 2014) to assess the significance of any departure of the true value of the function from the simulation envelope within a 15-m analysis window.

Next, we assessed the relationships between the point intensity of each hierarchy level and the topographic covariates described above. We used two approaches for this assessment: relative distribution estimates (Baddeley, 2010; Baddeley & Turner, 2005) and Kolmogorov-Smirnov (KS) tests (Berman, 1986). The former describes the relationship between the point pattern and the covariates, and the latter tests if there is a non-random association between the point pattern and the covariate.

2.3. Model development and selection

Based on results of the baseline analyses, we implemented the MHPPF. We first modeled the highest level of the hierarchy (level 4) by evaluating the families of point process models suggested by the baseline analysis and considering only the topographic variables as covariates. Models for the next lower levels (3, 2, and 1, respectively) were built by choosing the best combination of point process model family, topographic covariates, and estimated intensity surfaces from points of the higher levels of the hierarchy. Model parameters were estimated using the spatstat implementation of Berman-Turner Maximum Likelihood (for Poisson processes) and the minimum contrast method (for cluster processes) (Baddeley et al., 2015).

To select the best model within each hierarchy level, we used the mimetic approach of Goreaud et al. (2004). In this approach, metrics for sets of simulated patterns created from each of the candidate models were compared to those from the original data using simulation envelope tests as guides rather than hard and fast rules (Baddeley et al., 2015; Wiegand & Moloney, 2014). Best models were chosen using three criteria. First, we preferred models where observed values were mostly contained within the simulation envelopes. Second, if two candidate models were similar in this respect, we chose the model where higher-level intensity surfaces were included. Finally, with these criteria met, we followed the parsimony principle and preferred models that were simpler. For example, homogeneous models are simpler than inhomogeneous models, Poisson models are simpler than cluster models, and using elevation is simpler than using elevation derivatives as covariates.

Simulations were performed with the Metropolis-Hastings simulation algorithm (Baddeley & Turner, 2006; Geyer & Moller, 1994). For simulating patterns below level 4, we used the chosen models from the next higher levels to create simulated patterns, then generated the associated point intensity surfaces from these simulated patterns, and finally used these simulated patterns as the covariates. This process is summarized in Fig. 2.

To assess whether the simulated patterns reflected the relationships between the observed point patterns and topography, we created relative distribution functions using only elevation as a topographic
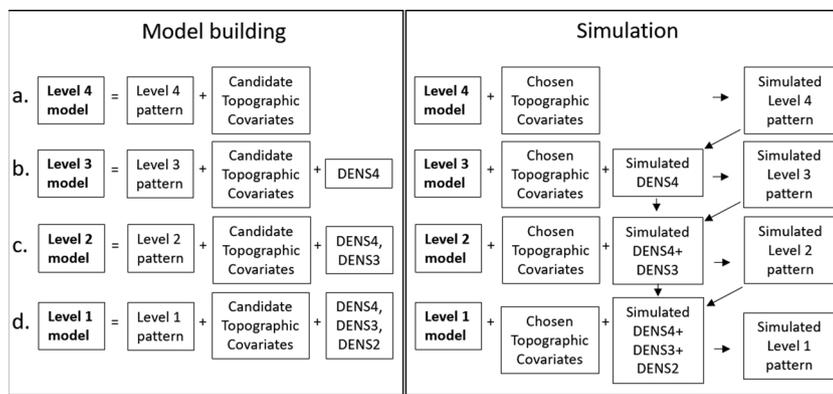
Fig. 2. Schematic diagram of the hierarchical modeling process. Processes for each level of the hierarchy, a–d (proceeding left to right), illustrate the sequential steps in generating simulated realizations of a hierarchical point process. Level 4 is the highest level, and represents the large diameter trees. DENS4, etc. represent the intensity surfaces generated from the point patterns. To generate the final, multi-level point pattern, simulated patterns of all of the levels are superimposed.
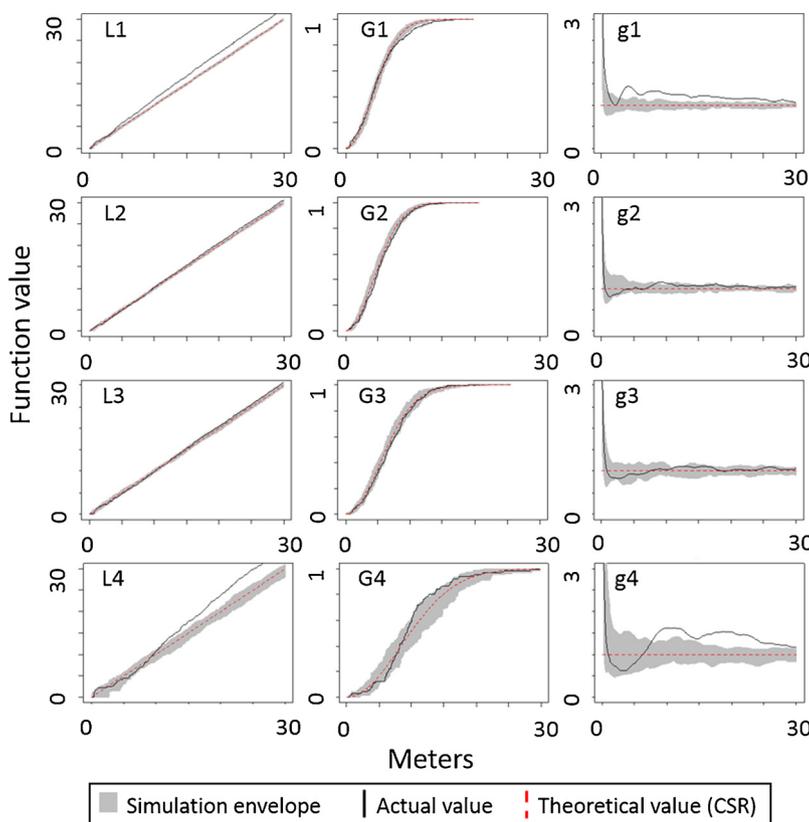

Fig. 3. Plots of the boundaries of the simulation envelopes (grey, polygonal region) for the *L*, *G* and *g* functions from 99 simulated HPPs, the mean of the simulated values (red, dashed line) and the functions computed from the observed data (bold, black line). Figure labels indicate the function type (*L*, *G* or *g*) and hierarchy level (1–4); e.g., L4 is the *L* function for hierarchy level 4 (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

variable because we considered it a surrogate for the other DEM-based variables. For comparisons using second order statistics, we used homogeneous versions of the *L*, *G* and *g* functions.

To assess whether the characteristics of the combined, multi-type pattern reproduced the observed inter-level spatial relationships, we conducted bivariate tests using the homogeneous *L* and *g* functions (Baddeley et al., 2015). Every pair combination of hierarchical levels was evaluated (i.e. 4–3, 4–2, … 2–1). To evaluate whether the incorporation of higher-level intensity surfaces used to reflect asymmetric competitive effects improved upon non-hierarchical models, we repeated the bivariate assessment of the final model using another set of models. These new models used the same point process model type and topographic covariates previously selected, but did not include the intensity surfaces from the other levels. We then compared the non-hierarchical versions of the models against those from the MHPPF.

For all tests, 48 realizations of each candidate model were generated. This number was chosen due to the large number of candidate models being evaluated, time requirements when performing each simulation of large point patterns, and the fact that the parallel

processing algorithm we used to expedite our analysis used a server with 16 cores – thus 3 realizations were generated on each core and results were pooled to arrive at a total of 48 unique simulated patterns for each candidate model and for each test.

3. Results

3.1. Baseline analysis

The locations of trees from each hierarchy level, as well as these superimposed on kernel density estimates, can be seen in Fig. 1. Smaller trees (hierarchy levels 1 and 2) appear in most parts of the watershed, and seem to exhibit a non-random, clustered pattern. Level 3 shows a more hyper-dispersed pattern with fewer obvious clusters, and level 4 exhibits a more geographically-restricted pattern, with most trees in the central portion of the study area, corresponding with lower elevations.

Two-sided maximum absolute deviation tests using the homogeneous second order function (*g*, *G* and *L*) simulation envelopes indicated that the summary function for the observed pattern of levels 1
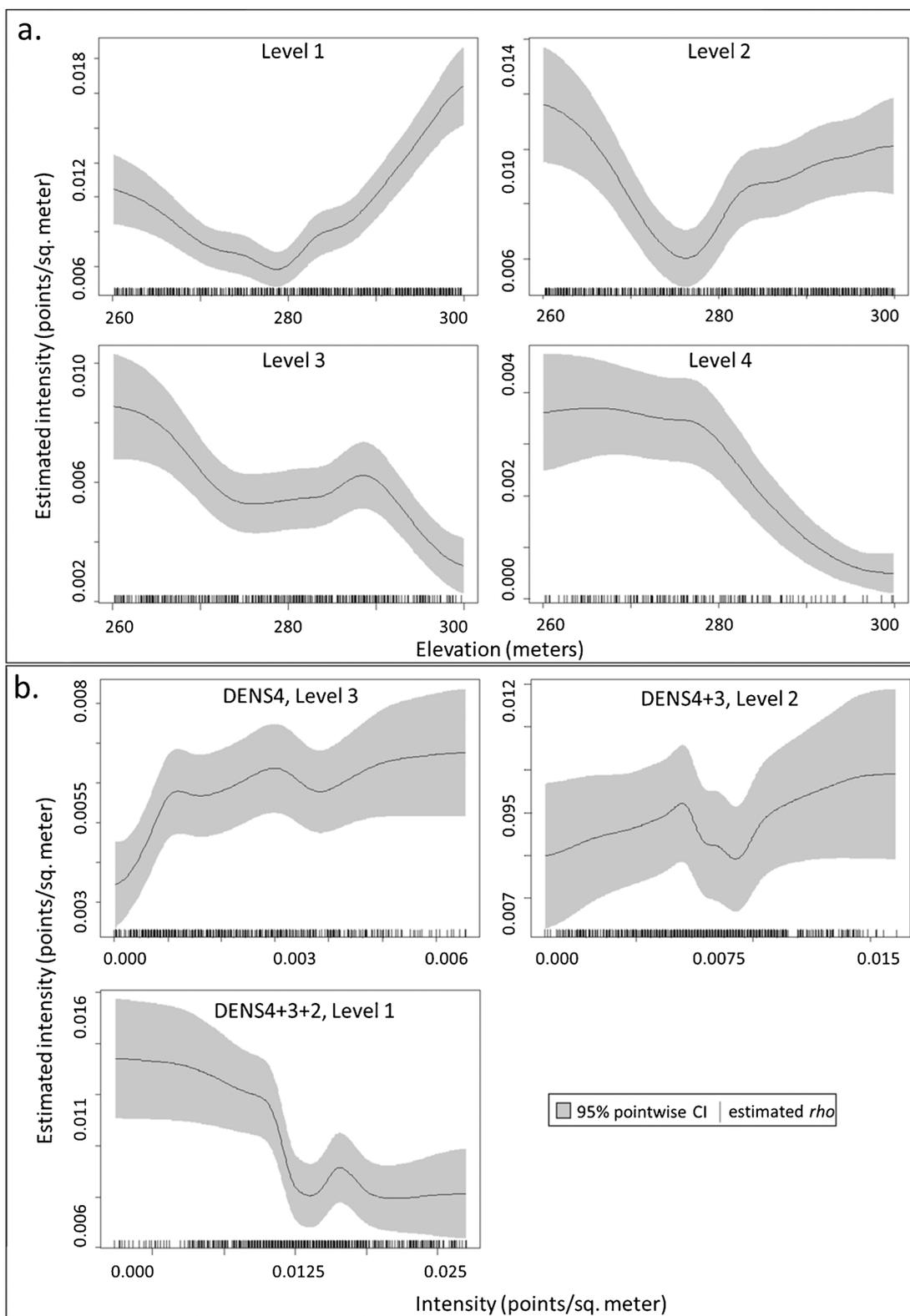
Fig. 4. Plots of estimated intensity (rho) vs. candidate covariates. a. Elevation above sea level (meters). Plots representing different levels of the hierarchy are labeled Levels 1–4. b. Point intensity (points/square meter) derived from the cumulative point intensity (density) surfaces of the higher levels of the diameter class hierarchy (DENS4 = intensity of level 4, DENS4 + 3 = cumulative intensity of level 4 and 3, DENS4 + 3 + 2 = cumulative intensity of level 4, 3, and 2). Grey polygonal areas represent the region bounded by the high and low values of the 95% pointwise confidence interval of the estimate of rho. Hash marks along X axis represent the frequency of observations along the range of X values.

Table 1

Kolmogorov-Smirnov (KS) test p-values to assess the relationship between patterns of points from different levels of the hierarchy and candidate covariates. *: < 0.05, **: < 0.01, ***: < 0.001, ns: > = 0.05. Dash mark indicates combinations for which the hierarchy level being evaluated is a component of the candidate covariate, and thus tests were not performed.

Level	DEM	SLOPE	ASPECT	TPI	TRI	DENS4	DENS4 + 3	DENS4 + 3 + 2	DENS3	DENS3 + 2	DENS2	DENS1
1	***	***	***	***	***	***	***	***	***	ns	**	–
2	ns	*	ns	**	*	ns	ns	–	ns	–	–	***
3	**	ns	ns	**	ns	**	–	–	–	–	**	***
4	***	**	***	***	**	–	–	–	***	ns	ns	***

Table 2

List of the combinations of model family and covariate inclusion scenarios evaluated at each level of the hierarchy.

LEVEL	MODEL TYPE	COVARIATE COMBINATION USED
level 4	HPP	
	IPP	DEM
	Inhomogeneous Matern	DEM
	Inhomogeneous Thomas	DEM
level 3	HPP	
	IPP	DEM, DENS4, DEM + DENS4
level 2	HPP	
	IPP	DEM, TRI, DENS4, DENS3, DENS(3 + 4): all combinations
	Inhomogeneous Matern	DEM, TRI, DENS4, DENS3, DENS(3 + 4): all combinations
level 1	HPP	
	IPP	DEM, SLOPE, DENS2, DENS3, DENS4, DENS234, DENS23, DENS24: all combinations
	Inhomogeneous Matern	DEM, SLOPE, DENS2, DENS3, DENS4, DENS234, DENS23, DENS24: all combinations

and 4 does not follow a CSR pattern for at least one of the summary functions (p < 0.05). The values for the function for the observed patterns fall above the average of the CSR-derived summary functions for the majority of the radii assessed, suggesting an aggregated pattern at multiple scales of observation (Fig. 3). For levels 2 and 3, however, results provide evidence of a pattern more closely resembling an HPP (Fig. 3).

An example of plots of the relative distribution functions developed to assess the relationships between the intensities of the patterns for each DBH class and the candidate covariates is shown in Fig. 4. The observed point patterns for all hierarchy levels were related to at least one of the topographic variables, with elevation playing a strong role (Fig. 4a). Fig. 4b shows relative distribution functions that reveal the relationship between intensity of points at higher levels of the hierarchy and the lower level patterns. The intensity of points at level 4 displays a positive correlation with that of level 3. However, the intensity of the combination of levels 4 and 3 do not display a strong relationship with that of level 2. Finally, the cumulative intensity of levels 4, 3 and 2 has a strong negative relationship with that of level 1.

Results of the KS tests can be seen in Table 1. Significant results indicate that there is evidence of a non-random relationship between patterns of points at a given level of the hierarchy with those of the candidate covariates. For the candidate topographic variables, Levels 1 and 4 show non-random relationships, whereas for levels 2 and 3, there are fewer and weaker significant results.

3.2. Model development and selection

Based on results of our baseline analysis, we evaluated for each hierarchy level the set of models presented in Table 2.

An example of the different second order statistical and relative distribution functions for two competing models for level 4 can be seen in Fig. 5. Fig. 5a–d are an example of the mimetic tests of the *L*, *g*, *G* and relative distribution functions, respectively, for the chosen model for hierarchy level 4, and e–h are those for an HPP model, which does not track the observed data as well. In many cases, however, the different models that included covariates performed similarly, or there was a lack of agreement among the summary functions as to which model was better. The models ultimately chosen for each hierarchy level are presented in Table 3.

Results of all of the bivariate *L* function tests that included level 1 can be seen in Fig. 6. Summary functions from the MHPPF (a–c) were closer to those of the observed data than were those from the non-hierarchical, no-competition models (d–f). This effect is particularly prominent with bivariate *L* functions for the level 1–4 pairs (c vs. f).
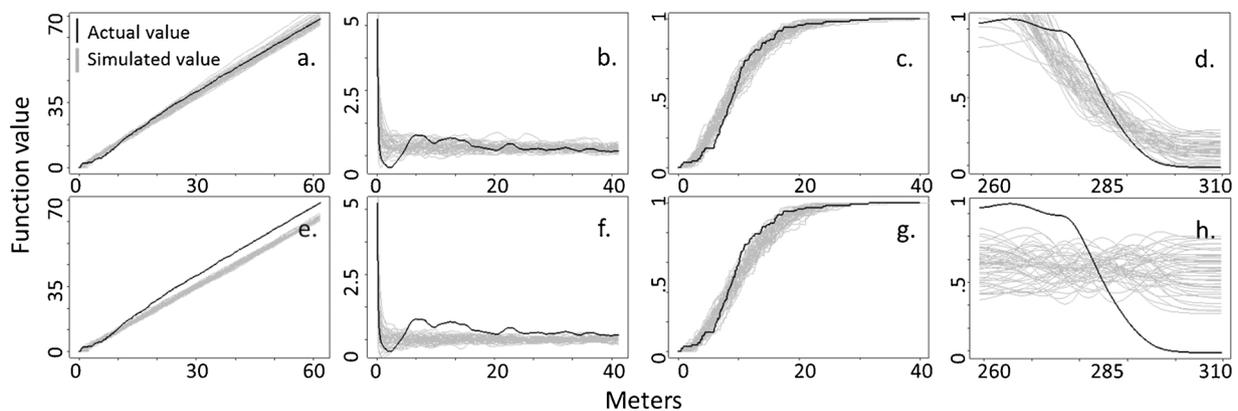

Fig. 5. Example of the second order statistical and relative distribution functions for the chosen level 4 model, and for an HPP model. a–d (the top row of figures) represent, respectively, the *L*, *g*, *G* and relative distribution functions for the level 4 model, and e–h (bottom row of figures) represent the corresponding functions for an HPP model. Black lines are the values for the observed data, and sets of grey lines are values for the functions derived from 48 simulations using the level 4 (top row) or HPP (bottom row) models.

Table 3
Selected models for each hierarchy level with model type and intensity function equations.

LEVEL	MODEL TYPE	MODEL FORM
level 4	IPP	$\lambda = e^{7.454-0.048*DEM}$
level 3	IPP	$\lambda = e^{-0.026-0.019*DEM}$
level 2	IPP	$\lambda = e^{-4.593-0.624*TRI-61.387*DENS4 + 66.176*DENS3}$
level 1	Inhomogeneous Matern	$\lambda = e^{-3.557 + 30.989*DENS2-86.543*DENS3-130.965*DENS4-0.044*SLOPE_*}$ $\sum_{i=1}^{\infty} 1_{(yi,\,10.2)}(x)^{a}$

[a] where yi is generated from the parent process, the intensity of which was estimated as 0.011.
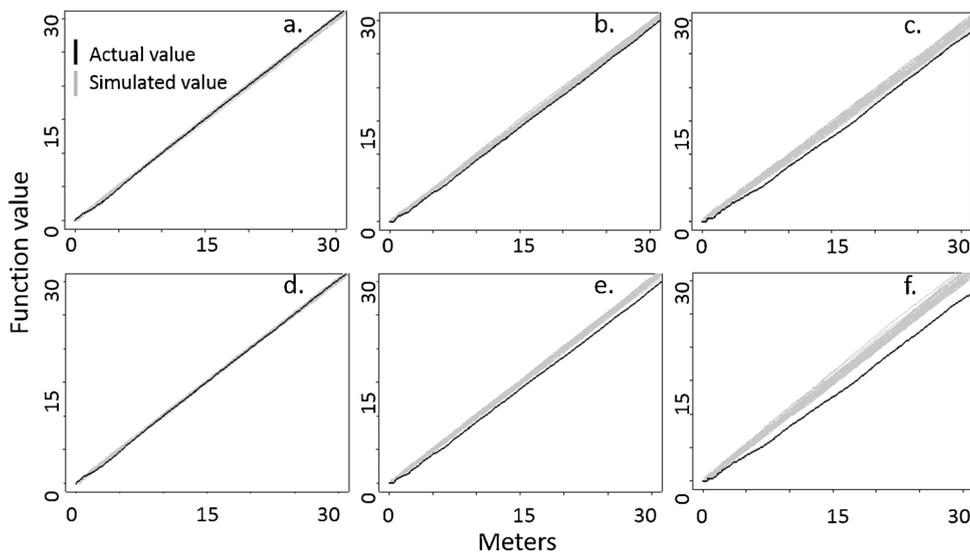

Fig. 6. Graphs of bivariate L functions for all bivariate combinations that included points of level 1 of the hierarchy. Level pairs for hierarchical models: a. Level 1-Level 2; b. Level 1-Level 3; c. Level 1-Level 4. Level pairs for non-hierarchical models: d. Level 1-Level 2; e. Level 1-Level 3; f. Level 1-Level 4. Black lines indicate the function associated with the observed data, grey lines represent functions associated with simulations using the hierarchical (a–c) or non-hierarchical (e-f) models.
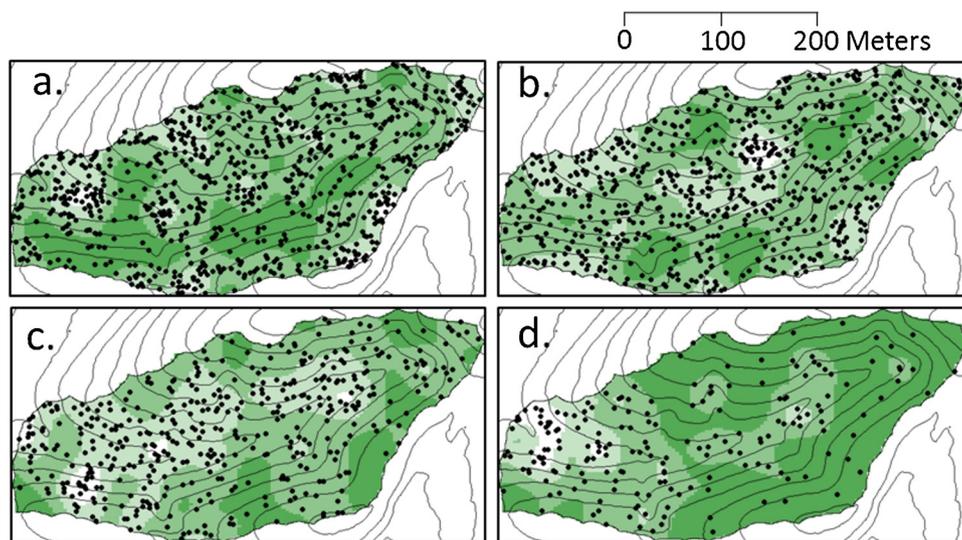

Fig. 7. Simulated locations of trees for each level of the hierarchy (DBH class) are superimposed on kernel density estimates of intensity of the point pattern created by the hierarchical modeling method. a = level 1 (20.3–29.1 cm), b = level 2 (29.2–37.6 cm), c = level 3 (37.7–48.5 cm), and d = level 4 (> 48.6 cm).

4. Discussion

In the study area, we assumed a traditional model of secondary forest development, and built our MHPPF accordingly. Implicit in Eq. (1) and Fig. 2 is the well-supported assumption that there is an underlying competitive hierarchy, mediated by environmental patterns, that leads to the spatial structure of many forest communities (Bormann & Likens, 1994). Important processes in forest succession commonly include competition, niche differentiation along environmental gradients, and some combination of continuous or episodic growth and mortality due to life history strategies and chance events. A useful feature of MHPPF is that it allows for the modeling of these and other types of interactions, like facilitation. In an arid environment, for example, we might hypothesize a mixture of competition for scarce moisture among dominant individuals, and a facilitative relationship between these and smaller individuals or other species that exploit shading or understory detritus accumulation proffered by the larger plants. Similarly, impacts of random events can be included using either the HPP or the IPP at any level in the process. Finally, variation in interaction strength among individuals based on any categorical

attribute, such as a combination of age and size, can be incorporated. Allowing for different model types at each hierarchy level, effects of random events, and incorporation of either positive or negative effects of other hierarchy levels as covariates thus allows researchers to determine whether observed patterns align with various types of hypotheses and assumptions about plant community development.

The types and forms of the models that, according to our selection criteria, best described the tree spatial pattern at each hierarchy level align with our hypotheses of pattern formation in this system. For example, for larger trees (levels 4 and 3), the best-performing model was a simple IPP driven by elevation (Table 3). In both cases, higher elevation corresponded with decreased point intensity, indicating that larger trees occur more commonly in the valley than at the ridge tops. According to parameter estimates of the chosen models (Table 3), the elevation effect is larger for the trees in level 4 than for the trees in level 3, suggesting that effects of topography-related biological processes differ by tree size. This is also clearly seen in Figs. 1 and 7. The selected levels 4 and 3 models also align with evidence that soil quality and light differences along topographic gradients strongly influence patterns of tree growth and establishment in the SSHCZO (Frey et al., 2007; Naithani et al., 2013; Smith et al., 2017; Wubbles, 2010).

Our chosen level 2 model is consistent with our initial hypothesis of pattern formation from competition effects mediated by topography, and with the findings of Smith et al. (2017), who found higher mean per tree biomass increment in swales due to their higher level of soil organic matter. The level 2 model indicates that as local topographic complexity (TRI) increases, as would tend to occur around swales, density of the patterns of level 2 trees decreases. At the same time, KS tests (Table 1) and relative distribution functions indicated that there is a significant, positive relationship between TRI and the pattern of large trees. It is thus not completely clear from our analysis if competition from larger individuals or topography most affects the pattern of smaller trees. However, we performed a likelihood ratio test to determine if a more complex model that contained both the impacts of the intensity of larger trees and of TRI was superior to one with just TRI; results indicated that the complex model that included impacts from the point intensities of levels 3 and 4 was superior (p < 0.005), further supporting the hypothesis of multiple drivers of pattern.

Bormann and Likens (1994) and Canham et al. (2004) point out that in mature forests, canopy gaps and associated higher light and soil resource availability lead to clustered patterns of smaller trees competing to reach the canopy. Our chosen level 1 model agrees with this hypothesis; cluster locations (parent points) were created by an HPP, with locations of trees within clusters having a negative relationship with the presence of the largest tree size classes. There was a negative relationship between the points within clusters and slope, as well, suggesting that topography could play an impact, perhaps indirectly, since KS tests (Table 1) and relative distribution functions indicate a significant, positive relationship between slope percentage and larger trees.

Analytically distinguishing a true cluster process from an IPP driven by environmental gradients is difficult because second order statistical functions, like the *L*, *g* and *G* functions, from both types of patterns can have similar forms (Wiegand & Moloney, 2014). Similarly, it is difficult to distinguish the effects of hierarchical competitive interactions from those of environmental trends because there is often collinearity between different levels of the hierarchy and environmental covariates. However, we found that including the effects of the patterns of larger trees as covariates improved the performance of the overall model with respect to the maintenance of bivariate spatial relationships when compared to models with only topography (Fig. 6). This, as well as the aforementioned results from the models of the individual levels, further illustrates how the MHPPF can be used to incorporate effects of multiple processes in models of stand spatial pattern.

The advantage of the MHPPF over existing methods is that it allows for a certain plasticity and modularity that does not exist in other

techniques that rely on only one family of point process models. Relying solely on Gibbsian models to construct a hierarchical framework, as did Grabarnik and Särkkä (2009), Högmander and Särkkä (1999), and Illian et al. (2009), would not have allowed us to superimpose processes of very different types in a straightforward manner. The modularity of the MHPPF allows the practitioner to model any combination of CSR, clustering, or hyper dispersion for patterns of different types, with any combination of hierarchical relationships, depending upon the unique requirements of the ecosystem under study.

5. Conclusions

We analyzed multiple, co-occurring tree-size related spatial patterns in a mature, second growth forest by developing a novel, hierarchical modeling framework that allows for each of the spatial patterns to be described by a different point process model family based on the dominant biological processes at each level of the hierarchy. The spatial pattern of large trees was related to elevation, while the pattern of mid-sized trees was related to topographic complexity and competition effects. The spatial pattern of the smallest-sized trees was clustered and affected by competition from larger-size trees and slope. The MHPPF can be used to examine hypotheses about ecological processes affecting stand structure and create realistic, simulated tree patterns to use in other applications like sampling experiments or distance-dependent tree growth models.

References

Baddeley, A., 2010. Analysing Spatial Point Patterns in R–Workshop Notes, Version 4. CSIRO, Canberra, Australia.

Baddeley, A., Diggle, P.J., Hardegen, A., Lawrence, T., Milne, R.K., Nair, G., 2014. On tests of spatial pattern based on simulation envelopes. Ecol. Monogr. 84, 477–489.

Baddeley, A., Rubak, E., Turner, R., 2015. Spatial Point Patterns–Methodology and Applications with R. Chapman & Hall/CRC Interdisciplinary Statistics. Boca Raton, FL, 824 pp. .

Baddeley, A., Turner, R., 2005. Spatstat: an R package for analyzing spatial point patterns. J. Stat. Softw. 12, 1–42.

Baddeley, A., Turner, R., 2006. Monitoring spatial point patterns in R. In: In: Baddeley, A., Gregori, P., Mateu, J., Stoica, R., Stoyan, D. (Eds.), Case Studies in Spatial Point Process Modeling, vol. 185. Springer, New York, pp. 23–74.

Beers, T., Dress, P., Wensel, L., 1966. Aspect transformation in site productivity research. J. For. 64, 691–692.

Bella, I.E., 1971. A new competition model for individual trees. For. Sci. 17, 364–372.

Berman, M., 1986. Testing for spatial association between a point process and another stochastic process. J. R. Stat. Soc. Ser. C (Appl. Stat.) 35, 54–62.

Berman, M., Diggle, P., 1989. Estimating weighted integrals of the second-order intensity of a spatial point process. J. R. Stat. Soc. Ser. B (Methodol.) 51, 81–92.

Besag, J., 1977. Contribution to the discussion on Dr. Ripley's paper. J. R. Stat. Soc. Ser. B 39, 193–195.

Bivand, R., 2015. classInt: Choose Univariate Class Intervals. https://CRAN.R-project.org/package=classInt ed. 0.1–232015. .

Bormann, F.H., Likens, G., 1994. Pattern and process in a forested ecosystem-disturbance. Development and the Steady State Based on the Hubbard Brook Ecosystem Study. Springer-Verlag Co., New York 253 pp.

Brooker, R.W., Maestre, F.T., Callaway, R.M., Lortie, C.L., Cavieres, L.A., Kunstler, G., Liancourt, P., Tielbörger, K., Travis, J.M.J., Anthelme, F., Armas, C., Coll, L., Corcket, E., Delzon, S., Forey, E., Kikvidze, Z., Olofsson, J., Pugnaire, F., Quiroz, C.L., Saccone, P., Schiffers, K., Seifan, M., Touzard, B., Michalet, R., 2008. Facilitation in plant communities: the past, the present, and the future. J. Ecol. 96, 18–34.

Canham, C., LePage, P., Coates, K., 2004. A neighborhood analysis of canopy tree competition: effects of shading versus crowding. Can. J. For. Res.-Revue Canadienne De Recherche Forestiere 34, 778–787.

ESRI, 2014. ArcGIS Desktop. 10.3.1 ed, Redlands, CA. .

Frey, B., Ashton, M., McKenna, J., Ellum, D., Finkral, A., 2007. Topographic and temporal patterns in tree seedling establishment, growth, and survival among masting species of southern New England mixed-deciduous forests. For. Ecol. Manage. 245, 54–63.

Genet, A., Grabarnik, P., Sekretenko, O., Pothier, D., 2014. Incorporating the mechanisms underlying inter-tree competition into a random point process model to improve spatial tree pattern analysis in forestry. Ecol. Model. 288, 143–154.

Getzin, S., Wiegand, T., Wiegand, K., He, F., 2008. Heterogeneity influences spatial patterns and demographics in forest stands. J. Ecol. 96, 807–820.

Geyer, C.J., Moller, J., 1994. Simulation procedures and likelihood inference for spatial point-processes. Scand. J. Stat. 21, 359–373.

Goreaud, F., Loussier, B., Ngo Bieng, M.A., Allain, R., 2004. Simulating realistic spatial structure for forest stands: a mimetic point process. In: Interdisciplinary Spatial Statistics Workshop 2004. ENGREF and University of Paris I Panthéon-Sorbonne, Paris, France, Paris, France.

Grabarnik, P., Särkkä, A., 2009. Modelling the spatial structure of forest stands by multivariate point processes with hierarchical interactions. Ecol. Model. 220, 1232–1240.

Hijmans, R.J., 2015. Raster: Geographic Data Analysis and Modeling. ed. 2.6-7. https://CRAN.R-project.org/package=raster.

Högmander, H., Särkkä, A., 1999. Multitype spatial point patterns with hierarchical interactions. Biometrics 55, 1051–1058.

Illian, J., Møller, J., Waagepetersen, R., 2009. Hierarchical spatial point process analysis for a plant community with high biodiversity. Environ. Ecol. Stat. 16, 389–405.

Illian, J., Penttinen, A., Stoyan, H., Stoyan, D., 2008. Statistical Analysis and Modelling of Spatial Point Patterns. John Wiley & Sons, Ltd., Hoboken, NJ 534 pp.

Jenks, G.F., Caspall, F.C., 1971. Error on choroplethic maps - definition, measurement, reduction. Ann. Assoc. Am. Geogr. 61, 217–244.

Kaye, M., Smith, L., Eissenstat, D., Wubbles, J., Adams, T., Osborne, J., 2015. Susquehanna Shale Hills Critical Zone Observatory Tree Survey (2012 updates). EarthChem Library (ECL) http://dx.doi.org/10.1594/IEDA/100516.

Larson, A., Lutz, J., Donato, D., Freund, J., Swanson, M., HilleRisLambers, J., Sprugel, D., Franklin, J., 2015. Spatial aspects of tree mortality strongly differ between young and old-growth forests. Ecology 96, 2855–2861.

Naithani, K.J., Baldwin, D.C., Gaines, K.P., Lin, H., Eissenstat, D.M., 2013. Spatial distribution of tree species governs the spatio-temporal interaction of leaf area index and soil moisture across a forested landscape. PLoS One 8, e58704.

Prokešová, M., Hahn, U., Jensen, E.V., 2006. Statistics for locally scaled Point processes. In: In: Baddeley, A., Gregori, P., Mateu, J., Stoica, R., Stoyan, D. (Eds.), Case Studies in Spatial Point Process Modeling, vol. 185. Springer, New York, pp. 99–123.

R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria ed 3.3.3.

Raventos, J., Wiegand, T., De Luis, M., 2010. Evidence for the spatial segregation hypothesis: a test with nine-year survivorship data in a Mediterranean shrubland. Ecology 91, 2110–2120.

Schwinning, S., Weiner, J., 1998. Mechanisms determining the degree of size asymmetry in competition among plants. Oecologia 113, 447–455.

Smith, L., Eissenstat, D., Kaye, M., 2017. Variability in aboveground carbon driven by slope aspect and curvature in an eastern deciduous forest, USA. Can. J. For. Res. 47, 149–158.

Stoyan, D., Penttinen, A., 2000. Recent applications of point process methods in forestry statistics. Stat. Sci. 15, 61–78.

USGS, 1/9 arc-second 2010 15 × 15 minute IMG, 2010, US Geological Survey, Reston, VA, https://thor-f5.er.usgs.gov/ngtoc/metadata/waf/elevation/1-9_arc-second/img/ned19_n40x75_w078x00_pa_south_2007_meta.xml.

Wiegand, T., Moloney, K.A., 2014. Handbook of Spatial Point-Pattern Analysis in Ecology CRC Press. Taylor & Francis Group, Boca Raton, FL 512 pp.

Wilson, M., O'Connell, B., Brown, C., Guinan, J., Grehan, A., 2007. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. Mar. Geodesy 30, 3–35.

Wu, H.I., Sharpe, P.J.H., Walker, J., Penridge, L.K., 1985. Ecological field theory: a spatial analysis of resource interference among plants. Ecol. Model. 29, 215–243.

Wubbles, J., 2010. Tree Species Distribution in Relation to Stem Hydraulic Traits and Soil Moisture in a Mixed Hardwood Forest in Central Pennsylvania, Horticulture. Penn State University, State College, PA p. 39.