ORIGINAL PAPER

Spatial pattern analysis of forest trees based on the vectorial mark

Honglu Xin[1,2] · Toby Jackson[3] · Yujie Cao[4] ·
Huanyuan Zhang[2] · Yi Lin[1] · Alexander Shenkin[2,5]

Abstract Analysis of spatial patterns to describe the spatial correlation between a tree location and marks (i.e., structural variables), can reveal stand history, population dynamics, competition and symbiosis. However, most studies of spatial patterns have concentrated on tree location and tree sizes rather than on crown asymmetry especially with direct analysis among marks characterizing facilitation and competition among of trees, and thus cannot reveal the cause of the distributions of tree locations and quantitative marks. To explore the spatial correlation among quantitative and vectorial marks and their implication on population dynamics, we extracted vertical and horizontal marks (tree height and crown projection area) characterizing tree size, and a vectorial mark (crown displacement vector characterizing the crown asymmetry) using an airborne laser scanning point cloud obtained from two forest stands in Oxfordshire, UK. Quantitatively and vectorially marked spatial patterns were developed, with corresponding null models established for a significance test. We analyzed eight types of univariate and bivariate spatial patterns, after first proposing four types. The accuracy of the pattern analysis based on an algorithm-segmented point cloud was compared with that of a truly segmented point cloud. The algorithm-segmented point cloud managed to detect 70–86% of patterns correctly. The eight types of spatial patterns analyzed the spatial distribution of trees, the spatial correlation between tree size and facilitated or competitive interactions of sycamore and other species. These four types of univariate patterns jointly showed that, at smaller scales, the trees tend to be clustered, and taller, with larger crowns due to the detected facilitations among trees in the study area. The four types of bivariate patterns found that at smaller scales there are taller trees and more facilitation among sycamore and other species, while crown size is mostly homogeneous across scales. These results indicate that interspecific facilitation and competition mainly affect tree height in the study area. This work further confirms the connection of tree size with individual facilitation and competition, revealing the potential spatial structure that previously was hard to detect.

Keywords Spatial pattern · Spatial correlation · Quantitative mark · Vectorial mark · Summary statistics

✉ Alexander Shenkin
alexander.shenkin@ouce.ox.ac.uk

[1] Institute of Remote Sensing and Geographic Information Systems, School of Earth and Space Science, Peking University, Beijing 100871, People's Republic of China

[2] Environmental Change Institute, School of Geography and the Environment, University of Oxford, South Parks Road, Oxford OX1 3QY, UK

[3] Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EA, UK

[4] College of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, People's Republic of China

[5] SICCS, Northern Arizona University, Flagstaff, AZ 86001, USA

Introduction

The spatial pattern of forest trees is the distribution and spatial arrangement of trees in a population and can be described by the spatial correlation of the marks

characterizing the structure of individual trees in a forest stand (Wang et al. 2010; Podlaski 2019). The spatial pattern can also characterize the spatial correlation of potential intraspecific and interspecific interactions among individuals (Pillay and Ward 2012; Erfanifard and Sheikholeslami 2017) and reveal information on stand history, population dynamics, competition and their survival in different environments (Zhang et al. 2010; Alvarez et al. 2011; Lin et al. 2011; Genet et al. 2014).

Spatial patterns can be divided into unmarked and quantitatively marked patterns depending on the involvement of mark and into univariate, bivariate and multivariate patterns according to the number of species involved. A mark is information that characterizes individual trees and can be qualitative or quantitative (Wiegand and Moloney 2013). A quantitative mark is a scalar that describes a size variable of an individual tree, such as tree height, diameter at breast height, and crown width; whereas a qualitative mark is a descriptor such as tree species or tree habit. Summary statistics have frequently been used to quantify statistical properties of spatial patterns and to provide a brief and concise description of patterns using numbers, functions and diagrams at multiple scales (Wiegand and Moloney 2004; Illian et al. 2008). Commonly used summary statistics are the pair correlation function, mark connection function and its variations. Specific null models are also used to randomize certain aspects of the data, while holding others constant according to the specific type of patterns (Wiegand and Moloney 2013), aiming to create patterns that are expected in the absence of a particular ecological mechanism of interest (Graves 1996).

However, the spatial correlations of tree location and facilitated or competitive interaction and of size and facilitated or competitive interaction are still not clear due to the lack of corresponding marks and summary statistics. Previous research with traditional census data has usually focused only on tree height and diameter at breast height, which lack insight into the three-dimensional structure of individual trees. Laser scanning technology, which can provide a three-dimensional structure of a forest tree, has provided a new solution (Henning and Radtke 2006; Packalen et al. 2013; Taubert et al. 2015; Lin et al. 2016; Fischer et al. 2020). Airborne laser scanning (ALS) technology can be used to collect a point cloud covering a large area, and individual trees can be delineated from the ALS point cloud through segmentation algorithms, from which tree marks can be further extracted (Zhao et al. 2014).

Aiming at exploring whether there are spatial correlations (1) among individual locations based on a given coordinates (i.e., if trees are spatially clustered, random, or segregated), (2) among quantitative marks representing tree size (i.e., if trees are significantly large or small at certain scales), (3) among vectorial marks representing individual facilitated or competitive interactions (i.e., if trees are spatially facilitated or competitive at certain scales), (4) among quantitative and vectorial marks (i.e., if facilitated trees are large or competitive trees are small at certain scales). Four types of patterns were investigated: (1) unmarked pattern, (2) quantitatively marked pattern, (3) vectorially marked pattern and (4) quantitatively and vectorially marked pattern. For the vectorial mark-related patterns, we developed new methods to analyze. All patterns were tested with the 1-ha plot both univariately and bivariately, so eight types of patterns are presented in this study. However, because the species information of the 9-ha plot was not specified, we did only a univariate pattern analysis for the 9-ha plot and thus present only four types of patterns. The analysis based on the 9-ha plot showcases the implementation of ALS for forest trees patterns analysis over a much wider area compared with traditional measurements.

Materials and methods

Data collection and mark extraction

Study area and data collection and preparation

A 1-ha plot and a 9-ha plot were established in a deciduous forest located in Wytham Woods (1°20′ W, 51°47′ N), Oxfordshire, UK. These two plots are part of the ForestGEO global network run by the University of Oxford (Butt et al. 2009; Calders et al. 2018). The mean annual temperature is 10 °C, the mean annual rainfall is 726 mm, and the mean annual radiation is 118 W/m^2. The plots are dominated by *Acer pseudoplatanus* (sycamore), *Fraxinus excelsior* (ash) and *Corylus avellana* (hazel), followed by *Quercus robur* (English oak) and *Crataegus monogyna* (common hawthorn).

We collected a terrestrial laser scanning (TLS) point cloud of the 1-ha plot with a RIEGL VZ-400 terrestrial laser scanner in leaf-off (December 2015, January 2016) condition at multiple locations, which were connected by reflective targets (Calders et al. 2016). The vegetation point density was about 11,000 points per m^2 according to the visual interpretation. We also collected an ALS point cloud with a Leica ALS-50 II scanner during leaf-on (June 2014) (Butt et al. 2009; Calders et al. 2018). The point cloud density was 6.0 points per m^2. The southwest coordinate of the 1-ha plot is located at latitude 51.7750579, longitude − 1.33904729, accurately measured with the differential GPS. We collected census data containing species information for the 1-ha plot, which was used for a bivariate pattern analysis. We also collected an ALS point cloud of the 9-ha plot on the side of the 1-ha plot.

Through careful visual inspection, 523 individual trees were manually extracted from the 1-ha TLS point cloud, then manually assigned species information using the census data. The 1-ha plot contains 364 sycamore trees and 159 other species. The average tree height is 15.42 m, median height is 16.08 m, average crown width is 6.18 m, and median crown width is 5.61 m.

Because an ALS point is most likely from the same tree as its spatially nearest TLS point, we tagged all ALS points with corresponding trees based on TLS points that had been georeferenced. These tagged ALS points are called "true ALS" hereafter. We also segmented the ALS point cloud of the 1-ha and 9-ha plot with automatic segmentation algorithms, including canopy height model-based approach, pit-free canopy height model-based approach, and point cloud–layer stacking seed point-based approach and a metabolic transporting distance-based approach (Millikan et al. 2019; Zaforemska et al. 2019; Hastings et al. 2020; Xin et al. 2021). The approach with highest accuracy was then selected for segmenting ALS. The segmented ALS points are hereafter called segmented ALS. The true ALS was used to assess the accuracy of the segmented ALS-based marks in the pattern analysis of the 1-ha plot.

Mark extraction

Tree height and crown projection area (CPA) were used as quantitative marks and were extracted from the segmented ALS point cloud. The height of a tree is the elevation from ground to the highest ALS point of the tree (Kwak et al. 2007). The CPA, an essential factor describing horizontal tree size, is the area of the convex hull of ALS points in a horizontal plane (Xu et al. 2013). However, the CPA is often overlooked in patterns analysis due to the inaccessibility of census data and is replaced by crown widths in two directions.

The essential mark we extracted from the tree crown is the crown displacement vector (CDV), a vectorial mark. CDV is the vector from tree top to crown centroid, which is the centroid of the convex hull of the ALS points. Note that the CDVs we used in this study are two-dimensional vectors in a horizontal plane, although they have vertical component. An essential assumption is that the x-, y-coordinates of the tree top are the same as that of the tree root and tree location; thus, CDV can represent the crown displacement. The ecological significance of CDV pairs of two trees lies in their characterization of the facilitation or competition between the two trees (Mead 1966; Brisson 2001).

We analyzed eight types of patterns based on the quantitative and vectorial marks, four of which were new developed here, then developed corresponding summary statistics and null models. The patterns characterize the forest from several perspectives, including the spatial distribution of trees (i.e.,

if trees are spatially clustered, random or segregated), spatial correlation of tree size (i.e., if trees are significantly large or small at certain scales), spatial correlation of facilitation/competition (i.e., if trees are spatially facilitated or competitive at certain scales), and spatial correlation between tree size and facilitation/competition (i.e., if facilitated trees are large or competitive trees are small at certain scales), all considered both univariately and bivariately. In this study, we employed 19 Monte Carlo simulations with a confidence level of 95% to determine whether a pattern is significantly different from that produced by null models (Wiegand and Moloney 2013). The summary statistics can be tested for significance against the corresponding null model, and whether there is a structure in the pattern that does not exist in the null model (Wiegand and Moloney 2013).

In general, summary statistics can be divided into cumulative and noncumulative according to the mode of recruitment of individual trees, as shown in Fig. 1. Bandwidth b, a buffer zone threshold, equals half the value of ring width, used to determine whether a neighbour tree is distance r away from a focal tree. According to Illian et al. (2008), bandwidth can be set as.

$$d = 0.1/\sqrt{\lambda}, \tag{1}$$

where λ is the tree density of the plot. There are 523 trees in the 1-ha plot of this study, then $d = 0.1/\sqrt{523/10000} \approx 0.44m$. However, the values of summary statistics at different scales fluctuate strongly with the bandwidth of 0.5 m. Thus, we set bandwidth as 1.0 m for all the patterns for the 1-ha and 9-ha plots to avoid potential strong fluctuation. It is clear that the cumulative statistics gloss over some details of the spatial patterns shown in their noncumulative counterpart, while noncumulative
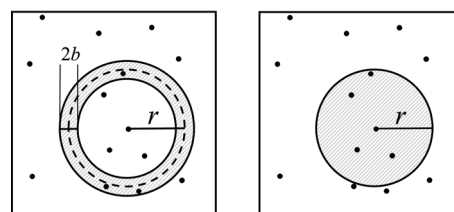

Fig. 1 Schematic diagram of **a** noncumulative and **b** cumulative summary statistics. Each *dot* stands for an individual tree, r stands for scale, and b stands for bandwidth. **a** Noncumulative summary statistics recruit trees at the distance of scale r with bandwidth b from the focal tree, that is, to recruit the trees located in a ring with a ring width of $2b$ centred at the focal tree, three trees are recruited in this example. **b** Cumulative summary statistics recruit all trees within the distance of scale r from focal tree, that is, to recruit the trees located in a disk with a radius of r centred at the focal tree; four trees are recruited in this example. The noncumulative and cumulative summary statistics iterate all the individual trees in the plot as the focal tree

statistics always show some stochastic noise. The question as to whether or not to use noncumulative or cumulative statistics depends on whether the goal is to reveal specific scale effects among the trees separated by distance r or to clarify the overall effects of all trees within distance r (Wiegand and Moloney 2013). In this study, aimed at revealing more details at different scales, all the results are based on the noncumulative summary statistics with a bandwidth b of 1 m.

Accuracy assessment of segmented ALS-based marks and pattern analysis

To analyze the eight types of spatial patterns in a larger area using ALS data, we thus used the algorithm-segmented ALS point cloud, instead of the matched true ALS. We assessed the accuracy of marks extracted from segmented ALS by comparing the mean error of segmented ALS-based quantitative marks with the true ALS-based marks, and the mean errors of magnitude of segmented ALS-based CDVs along the x- and y- coordinates were calculated. Meanwhile, to test the applicability of the segmented ALS for spatial pattern analysis, we calculated the summary statistics based on the segmented ALS and the true ALS and used cosine similarity (CS) and correctly detected rate (CDR) to assess accuracy of the indices.

Cosine similarity, a metric used to measure how similar the two multi-dimensional vectors are, calculates the cosine of the angle of the two multi-dimensional vectors (Nguyen and Bai 2011). CS is defined as

$$CS(v_s, v_t) = \frac{\sum_{i=1}^{m} v_s(i) \cdot v_t(i)}{\sqrt{\sum_{i=1}^{m} v_s^2(i)} \cdot \sqrt{\sum_{i=1}^{m} v_t^2(i)}}, \tag{2}$$

where m counts the number of scales r, and v_s and v_t, the vectors composed of summary statistics values at a series of scales, were calculated from segmented ALS and true ALS respectively. v_s and v_t equal to

$$v_s = [v_s(1), v_s(2), \ldots, v_s(m)] \tag{3}$$

$$v_t = [v_t(1), v_t(2), \ldots, v_t(m)] \tag{4}$$

The value of CS is always within the range of -1 and 1. The larger value of CS signifies that the segmented ALS-based summary statistics values are more similar to that based on true ALS, indicating the better results of segmented ALS-based analysis. Instead, the smaller value signifies worse analysis results. The CS, a direct index to evaluate the analytical results, compares the values based on segmented and true ALS.

We also used CDR to indirectly evaluate the applicability of segmented ALS on pattern analysis. There are three possibilities for the relative relationship between the summary

statistics value and its simulation envelopes produced by the null model, value above the simulation envelopes (positive correlation or clustered), below simulation envelopes (competition or segregated) and within simulation envelopes (no spatial correlation or random). We defined it as correctly detected if the segmented and the true ALS-based the summary statistics values have same relative relationship in comparison with simulation envelopes, otherwise, incorrectly detected. Thus, CDR is the ratio of correctly detected number of scales to the number of all scales as follows:

$$CDR = \frac{N_C}{N_C + N_I}, \tag{5}$$

where, N_C is the number of scales where spatial correlations are correctly detected, N_I is the number of scales where spatial correlations are incorrectly detected. CS evaluates the accuracy of the segmented ALS-based results in a direct way and CDR evaluates the accuracy in an indirect way.

Analysis 1: unmarked pattern

The ecological question of univariate unmarked pattern is to characterize the spatial distribution of trees, to test whether they are clustered, random or segregated distribution at a series of spatial scales. The ecological question of bivariate unmarked pattern is to characterize the spatial distribution of other species (one or more species) around the focal species, to test whether other species are aggregately, randomly or separately distributed surround focal species. Commonly used summary statistics are univariate and bivariate Ripley (1976)'s L function (cumulative) and pair correlation function (noncumulative) (Illian et al. 2008). The null model for univariate unmarked pattern is to randomize the two-dimensional tree locations in the horizontal plane. For bivariate unmarked pattern is to randomize the two-dimensional tree locations within each component pattern using a toroidal shift approach in the horizontal plane (Wiegand and Moloney 2013).

Analysis 2: quantitatively marked pattern

The ecological question of univariate quantitatively marked pattern is to characterize spatial correlation of tree size. The commonly used summary statistic is the univariate mark correlation function, its estimator is calculated as follows (Wiegand and Moloney 2013):

$$\hat{k}_{t,uni}(r) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n,\neq} t(m_i, m_j) \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}}{c_t \sum_{i=1}^{n} \sum_{j=1}^{n,\neq} k(||p_i - p_j|| - r) \cdot w_{i,j}}, \tag{6}$$

where $k_{t,uni}(r)$ is the value of univariate mark correlation function at a scale of r, n is the number of trees, $t(m_i, m_j)$

is a test function and has multiple expressions (Illian et al. 2008), c_t is the nonspatial mean of the test function, yielding the asymptotic value of test function, $||p_i - p_j||$ is the distance separating focal tree p_i and neighbor tree p_j, $k(||p_i - p_j|| - r)$ is a kernel function determining whether or not two trees are located distance r apart and equals 1 if the value $||p_i - p_j|| - r$ is smaller than bandwidth b and equals 0 if greater than bandwidth b, $w_{i,j}$ is the plot edge correction factor (Goreaud and Pélissier 1999). In this study, we used the periodic edge correction method to avoid any strange or spurious configurations of the trees along the plot borders; hence, $w_{i,j} = 1$. The estimators visit all pairs of trees separated by distance r and calculate the mean value of the test function over these pairs.

The ecological question of the bivariate quantitatively marked pattern is to characterize the spatial correlation of tree size of two species. The bivariate mark correlation function is calculated as follows (Illian et al. 2008):

$$\hat{k}_{t,biv}(r) = \frac{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2, \neq} t(m_i, m_j) \cdot C_{lm}(p_i, p_j) \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}}{c_t \cdot \sum_{i=1}^{n_1} \sum_{j=1}^{n_2, \neq} C_{lm}(p_i, p_j) \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}},$$
(7)

where $k_{t,biv}(r)$ is the summary statistics value at scale of r, n_1 is the number of focal species, n_2 is the number of other species, and $C_{lm}(p_i, p_j)$ is a judging function that yields a value of 1 if the focal tree p_j is of species l and neighbour tree p_i is of type m or other species, and 0 otherwise.

The null model of univariate quantitatively marked pattern randomizes the mark (i.e., tree height or CPA) all over the trees, which we called random labelling, and maintains the tree locations unchanged. That is, the marks are randomly shuffled over all trees. The null model of bivariate quantitatively marked pattern randomizes the mark within each species and maintains the tree locations unchanged.

Analysis 3: vectorially marked pattern

Under the assumption of environmental homogeneity, tree growth is not affected by environmental factors (e.g. illumination, nutrition and moisture), then the CDV pair is able to describe the facilitation or competition of two trees, which are separated by distance r.

We defined facilitation and competition of focal and neighbour trees by their CDVs as follows. An auxiliary vector from the focal tree to the neighbour tree was established. If the angle between the CDV of the focal tree and auxiliary vector is smaller than 45° and the angle between the CDV of neighbour tree and auxiliary vector is greater than 135°, then there is facilitation between the focal tree and the neighbour tree. If the angle between the CDV of focal tree and auxiliary vector is greater than 135° and the angle between the CDV of neighbour tree and auxiliary vector is smaller than 45°, then there is a competition between focal

tree and neighbour tree. Otherwise, we assume that there is no facilitation or competition between focal and neighbour trees. Note that the vectors are two dimensional in horizontal plane, and the vertical component was not considered.

We proposed a vectorially marked pattern, where the vectorial mark (i.e., CDV) is attached to the univariate or bivariate pattern. Analogous to the use of the mark correlation function for a qualitatively marked pattern, mark connection functions are developed to characterize the spatial relationship among the trees regarding CDV. The ecological question of vectorially marked pattern is to determine whether the CDV pairs of the individual trees separated by distance r are spatially correlated. Analogous to the summary statistics for bivariate unmarked pattern, the mark connection function for vectorially marked pattern is estimated with the ratio of the product densities.

The second-order product density characterizes the frequency of possible configurations of two trees in the plot (Illian et al. 2008). For a homogeneous pattern in which the properties of the pattern are the same in all directions and at all locations, the second-order product density for unmarked pattern can be estimated as follows:

$$\hat{\gamma}^{(r)} = \frac{n}{A} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n, \neq} k\left(\left||p_i - p_j\right|| - r\right) \cdot w_{i,j}}{2\pi r}$$
$$= \frac{1}{2\pi r A} \sum_{i=1}^{n} \sum_{j=1}^{n, \neq} k\left(\left||p_i - p_j\right|| - r\right) \cdot w_{i,j},$$
(8)

where A is the area of the plot, n/A is the tree density within the plot, and $\left(\left||p_i - p_j\right|| - r\right) \cdot w_{i,j}/2\pi r$ is the tree density around tree i at the distance of r, that is the tree density within the ring with the width of two times of bandwidth while the ring is centred on tree i.

The second-order product density for vectorially marked pattern can be estimated as follows:

$$\hat{d}_V(r) = \frac{1}{2\pi r A} \sum_{i=1}^{n} \sum_{j=1}^{n, \neq} t(v_i, v_j) \cdot V(v_i, v_j) \cdot k(||p_i - p_j|| - r) \cdot w_{i,j},$$
(9)

where v_i and v_j are the crown displacement vectors of the focal tree and neighbour tree, respectively. $V(v_i, v_j)$, a judging function, returns a value of 1 for facilitation of the focal and neighbour tree, which is determined by the CDV pair, a value of -1 for competitive interaction, and a value of 0 otherwise. The quantity $t(v_i, v_j)$ can be expressed as

$$t(v_i, v_j) = \left|v_i \cdot v_j\right|,$$
(10)

where $t(v_i, v_j)$ is the absolute value of dot product v_i and v_j, and $t(v_i, v_j) \cdot V(v_i, v_j)$ in Eq. 4 quantifies the degree of facilitation or competition of tree pairs. To some extent, $t(v_i, v_j)$ is a quantitative mark for the vector and is similar to the test function $t(m_i, m_j)$ in Eq. 1.

The function $\hat{d}'(r)$ counts all tree pairs, but $\hat{d}_V(r)$ only counts the tree pairs which are facilitated or competitive due to the judging function $V(v_i, v_j)$. Thus, the estimator of second-order product density for the unmarked pattern is accordingly modified as

$$\hat{d}(r) = \frac{1}{2\pi r A} \sum_{i=1}^{n} \sum_{j=1}^{n,\neq} \left| V(v_i, v_j) \right| \cdot k(||p_i - p_j|| - r) \cdot w_{i,j},$$
(11)

where $\left| V(v_i, v_j) \right|$ the absolute value of the judging function, equals 1 or 0. Therefore, the noncumulative mark connection function for the univariate vectorially marked pattern can be estimated as

$$
\begin{aligned}
\hat{q}_{V,\text{uni}}(r) &= \frac{\hat{d}_V(r)}{\hat{d}(r)} \\
&= \frac{\sum_{i=1}^{n} \sum_{j=1}^{n,\neq} t(v_i, v_j) \cdot V(v_i, v_j) \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}}{\sum_{i=1}^{n} \sum_{j=1}^{n,\neq} \left| V(v_i, v_j) \right| \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}}
\end{aligned}
$$
(12)

where the subscript V refers to the function judging the relative directions of CDVs. Analogous to Eq. 7, the estimator for the cumulative mark connection function for the univariate vectorially marked pattern can be estimated as

$$\hat{q}_{V,\text{uni}}^{\text{cum}}(r) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n,\neq} t(\mathbf{v}_i, \mathbf{v}_j) \cdot V(\mathbf{v}_i, \mathbf{v}_j) \cdot I(||p_i - p_j||, r) \cdot w_{i,j}}{\sum_{i=1}^{n} \sum_{j=1}^{n,\neq} \left| V(\mathbf{v}_i, \mathbf{v}_j) \right| \cdot I(||p_i - p_j||, r) \cdot w_{i,j}}$$
(13)

where the superscript cum refers to cumulative, the kernel function $I(||p_i - p_j||, r)$ equals 1 if the neighbour tree p_j is located within distance r from focal tree p_i and 0 otherwise. Compared with the noncumulative function, the cumulative function glosses over some details, but is less likely to be affected by scale because the cumulative functions involve more trees at the same scale, which reduces the potential stochastic noise (Wiegand and Moloney 2004). Note that expressions without superscript cum are all noncumulative summary statistics [e.g. $q_{V,\text{uni}}(r), q_{V,\text{biv}}(r), k_{t,V,\text{uni}}(r)$ and $k_{t,V,\text{biv}}(r)$].

Accordingly, to characterize the spatial correlation of facilitation or competition of two species, we developed the noncumulative and cumulative bivariate mark connection function for vectorially marked pattern as follows:

where n_1 is the number of focal tree species, n_2 is the number of neighbour tree species. $\hat{q}_{V,\text{biv}}(r)$ recruits neighbour trees of species 2 at distance r away from the focal tree of species 1, while $\hat{q}_{V,\text{biv}}^{\text{cum}}(r)$ recruits trees within distance r.

For the univariate vectorially marked pattern, the null model is to randomize CDVs of all the trees while the tree locations remain unchanged. But for bivariate vectorially marked pattern, the null model is to randomize vectorial marks of trees within each component pattern, because the two component patterns are a priori different (Wiegand and Moloney 2013).

Analysis 4: quantitatively and vectorially marked pattern

The quantitatively and vectorially marked pattern is a univariate or bivariate unmarked pattern attached a quantitative mark (e.g., tree height or CPA) and a vectorial mark (i.e., CDV). The ecological question is to determine the spatial correlation between quantitative mark and vectorial mark. For example, we may expect that tree pairs with facilitation would be taller or with larger CPA than expected given over all sizes of facilitated or competitive tree pairs. Accordingly, this pattern requires a null model where quantitative mark or vectorial mark fixed, while the other mark is randomly shuffled over trees. Note that the null model of univariate pattern shuffles marks over all trees, but that of the bivariate pattern shuffles marks within each species and avoids an interspecific shuffle.

We also built summary statistics for a quantitatively and vectorially marked pattern, analogous to the former functions. The estimator of non-normalized univariate mark correlation functions yields

$$\hat{k}_{t,V,\text{uni}}(r) = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n,\neq} t_{\text{sum}}(m_i, m_j) \cdot V(v_i, v_j) \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}}{\sum_{i=1}^{n} \sum_{j=1}^{n,\neq} \left| V(v_i, v_j) \right| \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}},$$
(16)

where $t_{\text{sum}}(m_i, m_j) = m_i + m_j$ if $V(v_i, v_j) = 1$, $t_{\text{sum}}(m_i, m_j) = 2m_{\max} - (m_i + m_j)$ if $V(v_i, v_j) = -1$, and $m_{\max}$ is the maximum value of the quantitative mark, and equals highest tree height or largest CPA.

$$\hat{q}_{V,\text{biv}}(r) = \frac{\hat{d}_V(r)}{\hat{d}(r)} = \frac{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2,\neq} t(\mathbf{v_i}, \mathbf{v_j}) \cdot V(\mathbf{v_i}, \mathbf{v_j}) \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}}{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2,\neq} \left| V(\mathbf{v_i}, \mathbf{v_j}) \right| \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}}$$
(14)

$$\hat{q}_{V,\text{biv}}^{\text{cum}}(r) = \frac{\hat{d}_V(r)}{\hat{d}(r)} = \frac{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2,\neq} t(\mathbf{v_i}, \mathbf{v_j}) \cdot V(\mathbf{v_i}, \mathbf{v_j}) \cdot I(||p_i - p_j||, r) \cdot w_{i,j}}{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2,\neq} \left| V(\mathbf{v_i}, \mathbf{v_j}) \right| \cdot I(||p_i - p_j||, r) \cdot w_{i,j}},$$
(15)

It is clear that if there is no spatial correlation among marks, the estimator of the non-normalized univariate mark correlation functions $\widehat{k}'_{t,V,uni}(r)$ equals the asymptotic value of $t_{sum}(m_i, m_j)$. Thus, to ensure the comparability at difference scales r, the estimator (Eq. 11) should be normalized by c, the asymptotic value of $t_{sum}(m_i, m_j)$. The expectation of all facilitated tree pairs yields

$$E(m_i + m_j) = 2\mu \tag{17}$$

and the expectation of all competitive tree pairs yields

$$E(2m_{max} - (m_i + m_j)) = 2m_{max} - 2\mu. \tag{18}$$

Thus,

$$c = \{n_{pos} \cdot E(m_i + m_j) + n_{neg} \cdot E[2m_{max} - (m_i + m_j)]\}/$$
$$n_{pos+neg} = (2\mu(n_{pos} - n_{neg}) + 2m_{max}n_{neg})/n_{pos+neg}, \tag{19}$$

where n_{pos} and n_{neg} are the numbers of facilitated or competitive tree pairs respectively, and $n_{pos+neg} = n_{pos} + n_{neg}$.

Furthermore, the estimator of normalized univariate and bivariate mark correlation function yields

$$\widehat{k}_{t,V,uni}(r) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n,\neq} t_{sum}(m_i, m_j) \cdot V(\mathbf{v_i}, \mathbf{v_j}) \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}}{c \cdot \sum_{i=1}^{n}\sum_{j=1}^{n,\neq} \left|V(\mathbf{v_i}, \mathbf{v_j})\right| \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}} \tag{20}$$

$$\widehat{k}_{t,V,biv}(r) = \frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2,\neq} t_{sum}(m_i, m_j) \cdot V(\mathbf{v_i}, \mathbf{v_j}) \cdot C_{lm}(p_i, p_j) \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}}{c \cdot \sum_{i=1}^{n_1}\sum_{j=1}^{n_2,\neq} \left|V(\mathbf{v_i}, \mathbf{v_j})\right| \cdot C_{lm}(p_i, p_j) \cdot k(||p_i - p_j|| - r) \cdot w_{i,j}}. \tag{21}$$

The test function $t_{sum}(m_i, m_j)$ for univariate pattern are calculated for pairs of trees that fulfill two conditions: the focal tree and neighbour tree are either facilitated or competitive, and the trees are separated by distance r.

Accordingly, the cumulative univariate and bivariate functions are estimated as follows:

$$\widehat{k}_{t,V}^{cum}(r) = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n,\neq} t_{sum}(m_i, m_j) \cdot V(\mathbf{v_i}, \mathbf{v_j}) \cdot I(||p_i - p_j||, r) \cdot w_{i,j}}{c \cdot \sum_{i=1}^{n}\sum_{j=1}^{n,\neq} \left|V(\mathbf{v_i}, \mathbf{v_j})\right| \cdot I(||p_i - p_j||, r) \cdot w_{i,j}} \tag{22}$$

$$\widehat{k}_{t,V,biv}^{cum}(r) = \frac{\sum_{i=1}^{n_1}\sum_{j=1}^{n_2,\neq} t_{sum}(m_i, m_j) \cdot V(\mathbf{v_i}, \mathbf{v_j}) \cdot C_{lm}(p_i, p_j) \cdot I(||p_i - p_j||, r) \cdot w_{i,j}}{c \cdot \sum_{i=1}^{n_1}\sum_{j=1}^{n_2,\neq} \left|V(\mathbf{v_i}, \mathbf{v_j})\right| \cdot C_{lm}(p_i, p_j) \cdot I(||p_i - p_j||, r) \cdot w_{i,j}}, \tag{23}$$

where the neighbour trees are within distance r of focal tree. Analogous to the cumulative form of functions for the vectorially marked pattern, cumulative functions for quantitatively and vectorially marked pattern reduce the potential

Table 1 Accuracy of quantitative and vectorial marks based on a segmented ALS point cloud, compared with a true ALS point cloud. Mean error is the average absolute value of difference between the segmented and the true ALS. CDV_x and CDV_y are the x and y components of crown displacement vector, respectively

Mark		Mean error (m)
Quantitative marks	Tree height	2.21
	Crown width	2.21
Vectorial mark	CDV_x	1.04
	CDV_y	1.14

stochastic noise, but gloss over some details, especially at the larger scales due to the higher numbers of tree pairs. Note that both the noncumulative estimator (Eqs. 19 and 20) and cumulative estimator (Eqs. 21 and 22) need c, the asymptotic value of $t_{sum}(m_i, m_j)$, because both of them contain the test function $t_{sum}(m_i, m_j)$, though noncumulative estimators only recruit trees at distance r, while cumulative estimators recruit trees within distance r.

The null models for the functions of this type of pattern are also random labelling. It requires a null model where the quantitative marks are fixed, but the vectorial marks are randomly shuffled over all trees (for univariate pattern) or within each component pattern (for bivariate pattern). Alternatively, another type of null model is to maintain the vectorial mark fixed, but to shuffle the quantitative marks over all trees (for univariate pattern) or within each component pattern (for bivariate pattern).

Results

Comparison of matched true ALS- and algorithm-segmented ALS-based results

Accuracy assessment of the four ALS point cloud segmentation results showed that the metabolic transport distance-based approach (Xin et al. 2021) yields a coefficient of determination (R^2) for tree height and crown width of detection of 0.9108 and 0.6193, respectively, and the overall R^2 of tree
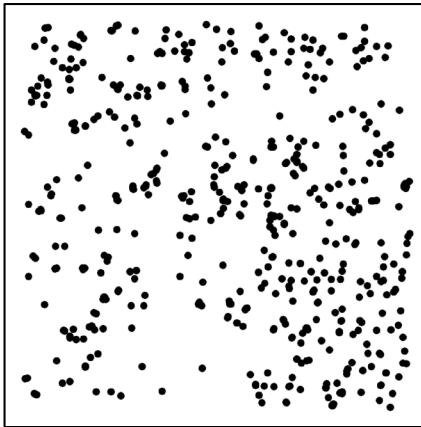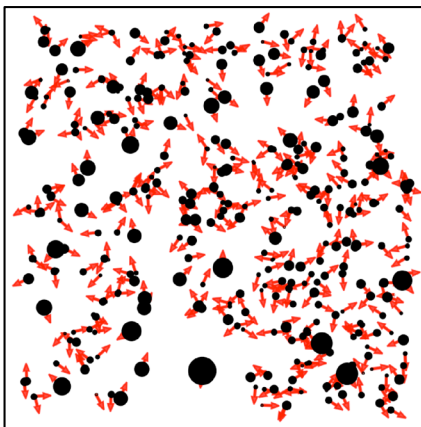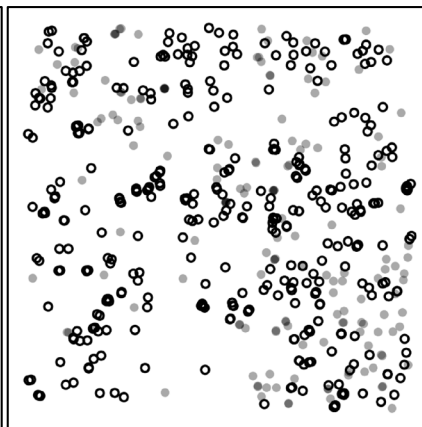
(a) Univariate unmarked pattern
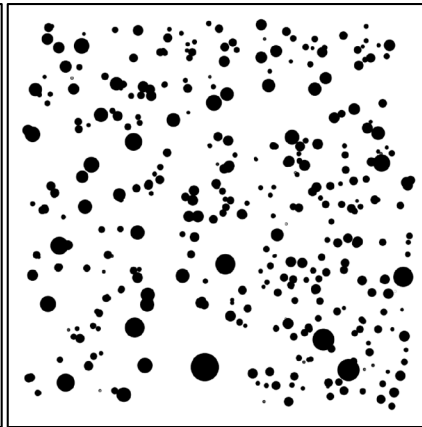
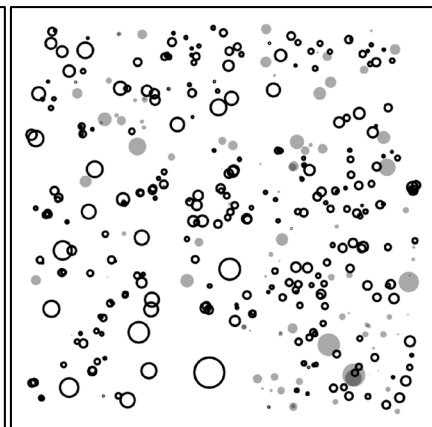**(b)** Univariate quantitatively marked pattern (tree height or CPA marked)
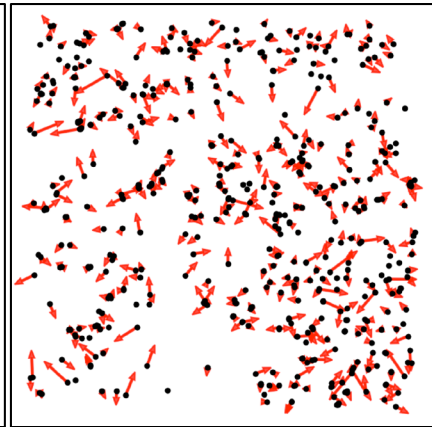
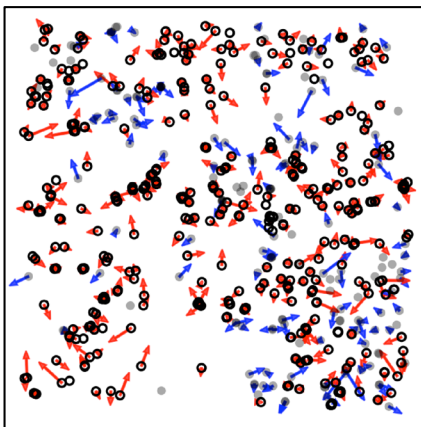**(c)** Univariate vectorially marked pattern (CDV marked)

(d) Univariate quantitatively and vectorially marked pattern (tree height and CDV marked, or CPA and CDV marked)

(e) Bivariate unmarked pattern

(f) Bivariate quantitatively marked pattern (tree height or CPA marked)

(g) Bivariate vectorially marked pattern (CDV marked)

(h) Bivariate quantitatively and vectorially marked pattern (tree height and CDV marked, or CPA and CDV marked)
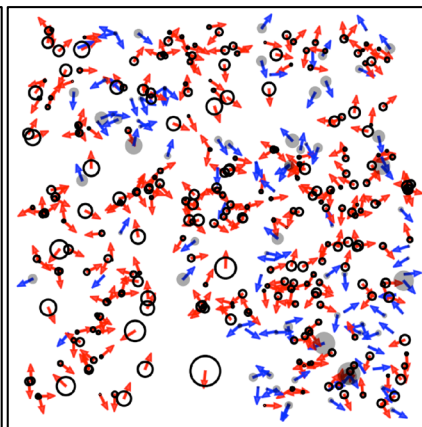
◄**Fig. 2** The eight types of spatial patterns of the 1-ha plot. Each *dot* or *circle* represents an individual tree. **a** Spatial distribution of 523 individual trees. **b** Disk area of the *black dot* is proportional to CPA. **c** The length of the red arrow is proportional to the length of CDV. **d** Similar to b and c, but the length of red arrow is constant. **e** Spatial distribution of sycamore (*open circles*, 364 individuals) and other species (*grey dots*, 159 individuals). **f** Similar to **b** and **e**. **g** Similar to **c** and **e**. *Red* and *blue arrows* are the CDVs of sycamore and other species, respectively. **h** Similar to **d** and **g**

height and crown width are 0.8045 and 0.4743, respectively, the mean error of tree height is 2.20 m, evaluated against the true ALS-based tree height and crown width. The other three approaches delineate understories with lower R^2 of tree height and crown width, and the mean errors of tree height ranged from 6.24 to 6.58 m (detailed by Xin et al. 2021). Thus, we employed the metabolic transport distance-based approach for the ALS point cloud segmentation of the 1-ha and 9-ha plots, which showed that there are 4380 trees in the 9-ha plot.

The validity of using segmented ALS in spatial pattern analysis was evaluated by the mean error of marks, CS and CDR. The mean errors of quantitative and vectorial mark are shown in Table 1. The error of the two quantitative marks was 2.21 m. The mean errors of the two components of the vectorial mark were 1.04 m and 1.14 m, implying that the segmented ALS-based structure parameters could be used for the spatial pattern analysis.

CS and CDR were used to assess the results for the eight types of segmented ALS-based pattern analysis in comparison to the true ALS-based results. From most patterns, values for assessment index CS were greater than 0.94, the average value of CS of all eight types of patterns was also 0.94, indicating a high similarity of the two lines of summary statistics. The CDR values of eight types of patterns vary from 70 to 86% (mean: 79%), which means about a number of 39.5 scales of all 50 scales were correctly detected.

In conclusion, according to the summary statistic lines and indices of CS and CDR, the algorithm segmented ALS has an excellent performance in the eight types of spatial patterns. We found that tree height-marked analysis usually get better results compared with CPA-marked analysis due to the coefficient of determination R^2 of tree height based on ALS segmentation algorithm is better than that of crown width. The performance test of segmented ALS, which used the true ALS as the criterion, is the precondition for segmented ALS-based analysis and also for further larger plot and larger scale analysis without census data.

The eight types of spatial patterns analysis

The spatial distributions of the locations and marks of trees in the 1-ha plot were shown in Fig. 2, we analyzed the univariate patterns (Fig. 2a–d) in Sect. 3.2.1 and bivariate patterns in Sect. 3.2.2. The eight types of spatial patterns we analyzed were listed in Table 2.

Univariate patterns

We analyzed four types of univariate patterns, as shown in the unmarked (Fig. 3a), quantitatively marked (Fig. 3b, c), vectorially marked (Fig. 3d), and quantitatively and vectorially marked (Fig. 3e, f) patterns. The trees showed strong segregation at scales of 1–3 m due to the close tree spacing, and rapidly turned to an aggregation at the scale of 5 m (Fig. 3a). Trees were significantly spatially aggregated at scales of 5–11 m, and peaked at scales of 4 and 11 m.

The tree height-marked quantitative pattern (Fig. 3b) showed a significant departure at scales of 1–3, 6–9 and 11 m under the test function.

$$t(m_i, m_j) = (m_i - \mu)(m_j - \mu), \tag{24}$$

where m_i and m_j are tree heights of the focal and the neighbor tree, respectively, and μ is the mean height. A significant departure from the simulation envelopes indicates a spatial correlation among the quantitative marks of the focal tree and neighbour tree. Values of $k_{t,uni}(r)$ lying above the simulation envelopes for null model indicate facilitation among the quantitative mark of focal tree and neighbour tree; that is, the focal tree and neighbour tree both are small trees or both are large trees. At scales of 1–3 m, Fig. 3b showed the positive correlation between the height of focal tree and neighbour tree, and at scales of 6–9 and 11 m, it showed a competitive interaction of the height. By contrast, the CPA-marked quantitative pattern (Fig. 3c) was less correlated with tree distribution, and it did not show significance at most of scales, except the slightly negatively departure from simulation envelopes at the scale of 3 and 6 m. The results showed that CPA is less likely to be affected by tree populations compared with tree height.

However, we are not yet able to determine whether they are small or large trees. Also, values below simulation envelopes indicate a negative correlation, and we are still not able to determine the size of focal and neighbour trees until joining the vectorially marked pattern.

Vectorially marked pattern (Fig. 3d) showed spatial correlation of facilitation among trees at the scales of 5–6 and 9–10 m; that is, CDVs of the focal tree and neighbour trees are more likely to point to each other, indicating the facilitation among trees at these scales. Thus, combined with Fig. 3b, we can estimate that both the focal and neighbour trees are potentially tall trees at a scale of 6 and 9 m, and combined with Fig. 3c, the focal tree is more likely to be with a large crown, while neighbours have a small crown.

Table 2 The eight types of spatial pattern analysis of forest trees based on qualitative, quantitative and vectorial marks, where the type iii, iv, vii and viii were newly developed. The new patterns analyzed the facilitation and competition and its impact on tree size. Each type has a specific ecological question, corresponding null model and summary statistics and analyzes the different types of spatial correlations among trees in the plot

Species	Mark	Spatial pattern type	Specific ecological question
Single species, or ignore species (Univariate)	Nonmark (unmark)	i. Univariate unmarked pattern (Wiegand and Moloney 2013)	Are trees spatially clustered, random or segregated?
	Tree height or CPA (quantitative mark)	ii. Univariate quantitatively marked pattern (Wiegand and Moloney 2013)	Are trees significantly large or small at certain scales?
	CDV (vectorial mark)	iii. Univariate vectorially marked pattern	Are trees spatially facilitated or competitive at certain scales?
	Tree height and CDV, or CPA and CDV (quantitative and vectorial mark)	iv. Univariate quantitatively and vectorially marked pattern	Are facilitated trees large or competitive trees are small at certain scales?
Two species, or one and other species (Bivariate)	Nonmark (unmark)	v. Bivariate unmarked pattern (Wiegand and Moloney 2013)	Are individuals of sycamore and other species spatially clustered, random or segregated?
	Tree height or CPA (quantitative mark)	vi. Bivariate quantitatively marked pattern (Wiegand and Moloney 2013)	Are individuals of sycamore or other species significantly large or small at certain scales?
	CDV (vectorial mark)	vii. Bivariate vectorially marked pattern	Are individuals of sycamore and other species spatially facilitated or competitive at certain scales?
	Tree height and CDV, or CPA and CDV (quantitative and vectorial mark)	viii. Bivariate quantitatively and vectorially marked pattern	Are facilitated individuals of sycamore and other species large or are competitive individuals small at certain scales?

Tree height- and vector-marked quantitative and vectorial patterns (Fig. 3e) showed a significantly positive departure from envelopes at the scale of 9–12 m, which indicates t facilitation among focal and neighbour trees and that they are tall trees. These results further confirmed the pattern analysis results of Fig. 3b, d. CPA- and vector-marked quantitative and vectorial pattern (Fig. 3f) also showed significantly positive departure at scales of 9–12 m, indicating the facilitation and large CPA of focal and neighbour trees. Figure 3c showed significant departure at the scales of 3 and 6 m, while Fig. 3f did not show any significant departure at scales smaller than 9 m, indicating that the small CPA is not correlated with the tree interactions.

On the whole, the univariate pattern showed that at small scales (1–3 m), trees tend to be spatially segregated and have small crowns and short heights due to the competition among trees for nutrients. At larger scales (around 5–10 m), trees tend to facilitate with each other, resulting in a spatially clustered distribution and taller trees, but the CPA did not show much spatial correlation with tree interactions. Thus, the vectorially marked pattern and quantitatively and vectorially marked pattern indicate a specific spatial relationship of size and interaction of trees.

Bivariate patterns

The bivariate pattern analysis tests for an interspecific spatial correlation, and the summary statistics different from those of the univariate pattern. Bivariate pattern analysis goes through the specific species separately, while the univariate pattern goes through all trees without distinguishing them by species. Results of bivariate pattern analysis are shown in Fig. 3g–l, where the "two" species are dominant species (sycamore) and the other species. The unmarked pattern Fig. 3g did not show a significant departure from envelopes at small scales (1–15 m), indicating that sycamore and the other trees are randomly distributed with each other at these scales, and there was no significant spatial correlation of tree locations.

The tree height-marked quantitative bivariate pattern (Fig. 3h) showed significant departure at a scale of 3–10 m, indicating a positive correlation between the two species for tree height; that is, both focal sycamore and neighbor trees (other species) are tall trees or both are short trees. CPA-marked pattern (Fig. 3i) showed there is slightly significant departure at a scale of 11 m. However, the sizes of sycamore and other species are still not clear; more types of patterns are needed.

The undetermined spatial relationship of tree size of sycamore and other species could be clarified by a bivariate vectorial pattern (Fig. 3j). As shown in Fig. 3j, there is significant departure at a scale of 9 m, indicating facilitation between sycamore and the other species. When these results are combined with results of the quantitatively marked pattern, we can judge that both the sycamore and other species are with large trees. Moreover, we can further clarify using
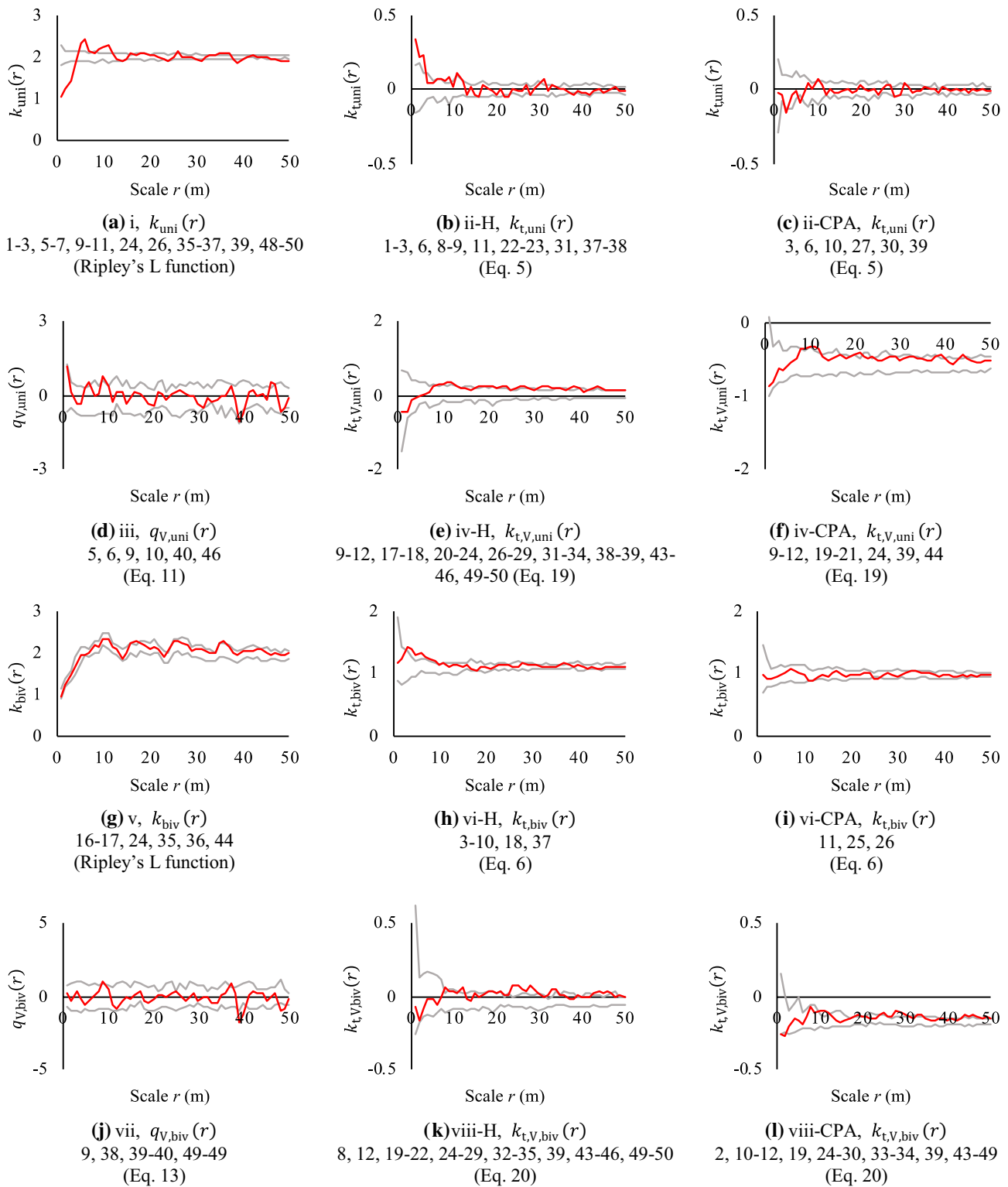
(a) i, $k_{uni}(r)$
1-3, 5-7, 9-11, 24, 26, 35-37, 39, 48-50
(Ripley's L function)

(b) ii-H, $k_{t,uni}(r)$
1-3, 6, 8-9, 11, 22-23, 31, 37-38
(Eq. 5)

(c) ii-CPA, $k_{t,uni}(r)$
3, 6, 10, 27, 30, 39
(Eq. 5)

(d) iii, $q_{V,uni}(r)$
5, 6, 9, 10, 40, 46
(Eq. 11)

(e) iv-H, $k_{t,V,uni}(r)$
9-12, 17-18, 20-24, 26-29, 31-34, 38-39, 43-46, 49-50 (Eq. 19)

(f) iv-CPA, $k_{t,V,uni}(r)$
9-12, 19-21, 24, 39, 44
(Eq. 19)

(g) v, $k_{biv}(r)$
16-17, 24, 35, 36, 44
(Ripley's L function)

(h) vi-H, $k_{t,biv}(r)$
3-10, 18, 37
(Eq. 6)

(i) vi-CPA, $k_{t,biv}(r)$
11, 25, 26
(Eq. 6)

(j) vii, $q_{V,biv}(r)$
9, 38, 39-40, 49-49
(Eq. 13)

(k) viii-H, $k_{t,V,biv}(r)$
8, 12, 19-22, 24-29, 32-35, 39, 43-46, 49-50
(Eq. 20)

(l) viii-CPA, $k_{t,V,biv}(r)$
2, 10-12, 19, 24-30, 33-34, 39, 43-49
(Eq. 20)

Fig. 3 Summary statistics (*red lines*) for the eight types of patterns in the 1-ha plot based on algorithm-segmented ALS, and the simulation envelopes (grey lines) produced by null models. The *x*-axes are the spatial scales with a step of 1 m, indicating the spatial extent; the *y*-axes show the values of dimensionless summary statistics. In each diagram, values above their envelopes indicate spatial aggregation (a and g), large tree size (**b**, **c**, **h** and **i**), facilitation (**d** and **j**), or s above their envelopes indicate a spatial segregation (a and g), small tree size (**b**, **c**, **h** and **i**), competition (**d** and **j**), or competition and small tree size (**e**, **f**, **k** and **l**). The scales with significant departures are listed below each diagram. The quantitative marks H and CPA are tree height and crown projection area, respectively. The test function used for pattern ii and vi is $t(m_i, m_j) = (m_i - \mu)(m_j - \mu)$, where m_i and m_j are the quantitative mark values of the focal and neighbour tree, respectively, and μ is the mean mark. Different patterns employ different null models

the quantitatively and vectorially marked pattern that both Fig. 3k, l showed a significant departure at a scale of 12 m, indicating a positive correlation of tree size with interspecific interactions among sycamore and other species. Thus, both bivariate vectorially marked pattern and bivariate quantitatively and vectorially marked pattern clarified the connection between tree size and interspecific interaction.

Univariate pattern analysis ignores the species information, and only focuses on one species or all species, while the bivariate pattern analysis focuses on two species or on one focal species and one other species. There is less significant departure of the bivariate pattern than that of the univariate pattern according to the results, indicating that the interspecific interaction is weaker than the intraspecific interaction in the 1-ha plot. Moreover, trees tend to be competitive and compete with each other at small scales, and are facilitated at large scales. Note that a significant departure occurred more randomly at scales greater than 20 m, showing weaker

spatial correlation at these scales and less ecological significance at larger distances among trees; thus, we only focused on scales less than 20 m in this study.

Extending the analysis to the 9-ha plot based on the ALS point cloud

The ALS has the ability to acquire a point cloud in a large area, so we also analyzed the spatial pattern of the 9-ha plot based on the ALS point cloud, as shown in Fig. 4.

The univariate pattern analysis of the 9-ha plot provided a wider perspective for exploring the spatial correlation of trees (Fig. 4). Similar to the analytical methods for ecological significance of the 1-ha plot, the unmarked pattern (Fig. 4a) showed that trees are aggregated at the scale of 6–8 m, and the tree height-marked quantitative pattern showed (Fig. 4b) that the heights of focal and neighbour trees are positively correlated, while the height- and
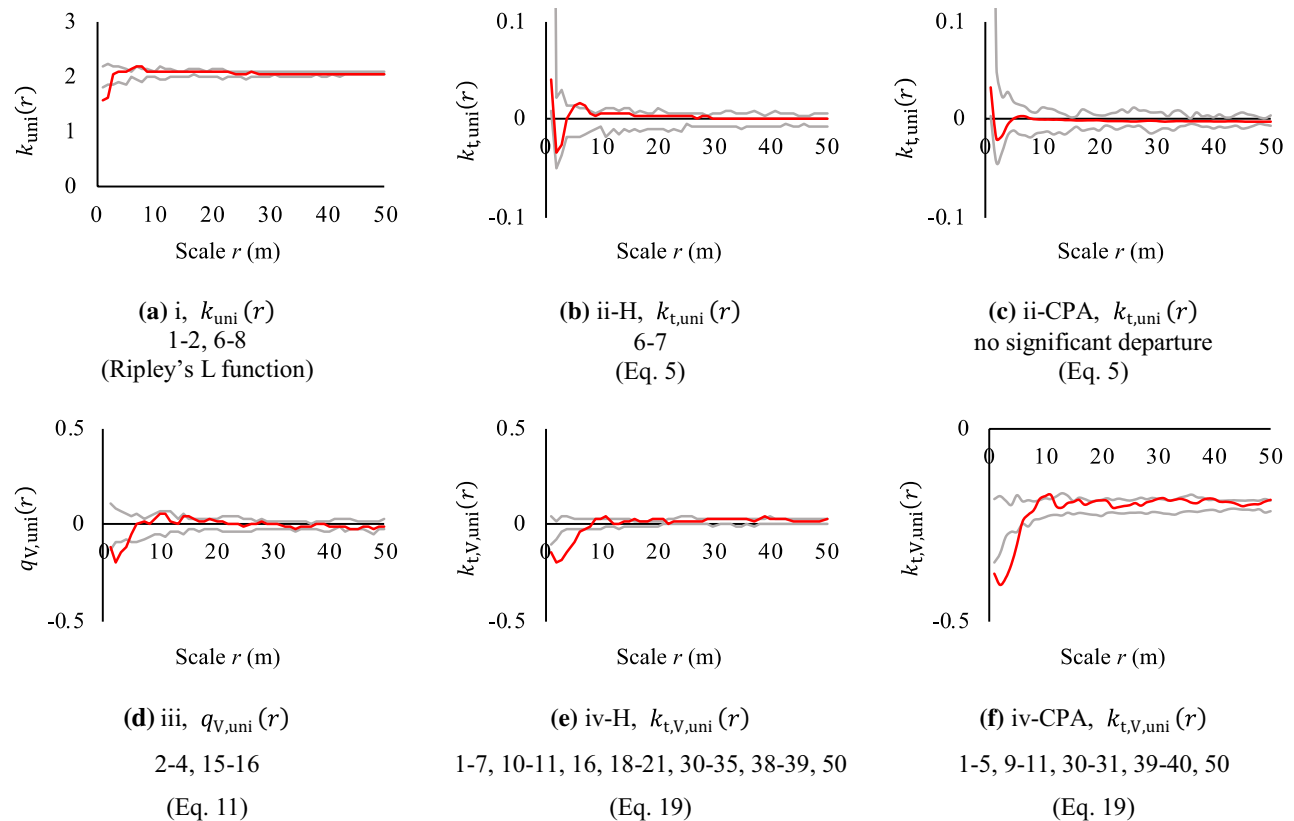

(a) i, $k_{uni}(r)$
1-2, 6-8
(Ripley's L function)

(b) ii-H, $k_{t,uni}(r)$
6-7
(Eq. 5)

(c) ii-CPA, $k_{t,uni}(r)$
no significant departure
(Eq. 5)

(d) iii, $q_{V,uni}(r)$
2-4, 15-16
(Eq. 11)

(e) iv-H, $k_{t,V,uni}(r)$
1-7, 10-11, 16, 18-21, 30-35, 38-39, 50
(Eq. 19)

(f) iv-CPA, $k_{t,V,uni}(r)$
1-5, 9-11, 30-31, 39-40, 50
(Eq. 19)

Fig. 4 Summary statistics (*red lines*) for the four types of univariate patterns in the 9-ha plot. The simulation envelopes (*grey lines*) were produced by null models. The *x*-axes are the spatial scales with a step of 1 m; *y*-axes show the values of dimensionless summary statistics. In each diagram, values above their envelopes indicate spatial aggregation **a**, large tree size (**b** and **c**), facilitation (**d**), or facilitation and large tree size (**e** and **f**). Values above their envelopes indicate spatial segregation **a**, small tree size (**b** and **c**), competition (**d**), or competition and small tree size (**e** and **f**). The scales with significant depar-

tures are listed below each diagram. The quantitative marks H and CPA are the tree height and crown projection area, respectively. The test function used for pattern ii and vi is $t(m_i, m_j) = (m_i - \mu)(m_j - \mu)$, where m_i and m_j are the quantitative mark values of the focal and neighbour tree, respectively, and μ is the mean mark. The scales with significant departures are listed below each diagram. Different patterns employ different null models. The bivariate pattern analysis was not conducted due to the lack of species information

vector- marked pattern (Fig. 4e) showed that the trees tend to be shorter and interacting competitively; thus, there is competitive inhibition among trees at the scale of 6–7 m. However, near 2 m, trees tend to be separately distributed (Fig. 4a) and competitive (Fig. 4d). Based on the quantitatively and vectorially marked pattern analysis (Fig. 4e, f), trees tend to be short and with small crowns at scales near 2 m due to competitive inhibition. Compared with the 1-ha plot (Fig. 3), the 9-ha plot (Fig. 4) tended to have less significant departures from the simulation envelopes. The reason lies in the number of trees; the 9-ha plot contains 4380 trees, while the 1-ha plot contains 523 trees, so there is less potential stochastic noise in the 9-ha plot, and the values of summary statistics of the 9-ha plot show relatively less departure than in the 1-ha plot. Note that the cumulative summary statistics also have less stochastic noise compared with the noncumulative summary statistics, though the noncumulative may gloss over some details.

Discussion

Previous spatial analytical methods have described the scattering of individual trees and the spatial distribution of tree size, whereas our newly proposed spatial analysis methods use new quantitative and vectorial marks to determine the driving factors that led to the spatial patterns. That is, traditional methods only analyze unmarked and quantitatively marked patterns such as a significant segregated distribution of certain species adults and a significant clustered distribution of subadults and juveniles at extreme small scales (Nguyen et al. 2014; Gu et al. 2019), a wide clustered distribution of some typical species at small scales (Nguyen et al. 2016), or a significantly smaller tree size of a minor species around the dominant species (Wiegand and Moloney 2013). In this study, we proposed new types of patterns based on the new marks CPA and CDV, which have not been used in other ecological studies, to further clarify the spatial correlations of tree facilitation or competition, of tree size and facilitation or competition. We found significant facilitation among the trees in the plot, which led to a larger tree size, and had a greater impact on tree height than crown size.

Trees could be spatially facilitated or competitive at any scale. The univariate and bivariate pattern of the 1-ha plot (Fig. 3) and the univariate pattern of the 9-ha plot (Fig. 4) showed that trees are more likely to spatially interact at smaller scales, and the significance of the effect (i.e., the departure from null models) is weaker at large scales. This result is consistent with other reports that tree spatial correlation diminishes with distance (Miao et al. 2014; Erfanifard et al. 2016; Moreno-Fernández et al. 2020). In fact, all the proposed univariate and bivariate patterns, including the quantitatively and vectorially marked related patterns, could

be further expanded to trivariate patterns, which incorporate spatial correlations for three species and also need corresponding summary statistics and null models (Velázquez et al. 2016).

Tree height, diameter at breast height and crown width have been commonly used marks in ecological studies, while crown width can be considered as a simplified form of CPA. In this study, we used CPA as a quantitative mark, measuring the vertical size of individuals. We also proposed the vectorial mark, CDV, a new type of mark that is a measure of the crown centre displacement from the stem of an individual tree. More three-dimensional marks could also be potentially extracted from ALS point cloud. They could be quantitative, describing the size of trees from different dimensions, or vectorial, representing the crown asymmetry more accurately, or other new types of marks. In addition, terrestrial laser scanning point clouds could also be used to provide more detailed three-dimensional structural marks, although the cost will be greater.

We examined the performance of the pattern analysis models only for 1-ha and 9-ha plots, but the two plots are both mixed forest. More types of forests, such as a coniferous forest or tropical forest, need to be tested to evaluate the broad applicability of the proposed models (Carrer et al. 2018). Nevertheless, it is still a challenge to derive species information from the ALS point cloud, which is required by bivariate or further trivariate pattern analysis. Fortunately, we could obtain species information from census data or simultaneous remote sensing data (e.g., spectral or hyperspectral images).

Conclusions

Quantitative and vectorial marks were derived from ALS point clouds, and the spatial patterns of two forest stands in Oxfordshire, UK was analyzed. To reveal details on the three-dimensional structure and facilitation and competition of trees, we used newly proposed marks including CPA, which measures the horizontal size of trees, and CDV, which measures the crown displacement of trees. New types of patterns suitable for these new marks were developed, including the vectorially marked pattern and the quantitatively and vectorially marked pattern. Correspondingly, we developed new summary statistics and null models and assessed the utility of the newly proposed marks. The newly proposed patterns were useful in clarifying the spatial correlation of intraspecific and interspecific facilitated or competitive interactions and the spatial correlation of tree size and facilitated or competitive interactions at a series of scales. They also helped to connect the spatial relationship of quantitative and vectorial marks. ALS-based marks were introduced for

spatial pattern analysis, and their applicability shown. We perfected the spatial pattern system and further clarified the spatial correlations of individual trees in forests. This study thus provides a new perspective to reveal the relationship between facilitation and competition and tree size at a series of scales in forests, and gain insight into the ecological driving factors of tree distribution and in relation to tree size and location.

References

Alvarez L, Epstein H, Li J, Okin G (2011) Spatial patterns of grasses and shrubs in an arid grassland environment. Ecosphere 2(9):1–30

Brisson J (2001) Neighborhood competition and crown asymmetry in Acer saccharum. Can J for Res 31(12):2151–2159

Butt N, Campbell G, Malhi Y, Morecroft M, Fenn K, Thomas M (2009) Initial results from establishment of a long-term broadleaf monitoring plot at Wytham Woods. University of Oxford Report, Oxford

Calders K, Burt A, Origo N, Disney M, Nightingale J, Raumonen P, Lewis P (2016) Large-area virtual forests from terrestrial laser scanning data. In: 2016 IEEE international geoscience and remote sensing symposium (IGARSS), pp 1765–1767.

Calders K, Origo N, Burt A, Disney M, Nightingale J, Raumonen P, Akerblom M, Malhi Y, Lewis P (2018) Realistic forest stand reconstruction from terrestrial LiDAR for radiative transfer modelling. Remote Sens 10(6):933

Carrer M, Castagneri D, Popa I, Pividori M, Lingua E (2018) Tree spatial patterns and stand attributes in temperate forests: the importance of plot size, sampling design, and null model. For Ecol Manag 407:125–134

Erfanifard Y, Saborowski J, Wiegand K, Meyer K (2016) Efficiency of sample-based indices for spatial pattern recognition of wild pistachio (Pistacia atlantica) trees in semi-arid woodlands. J Forestry Res 27(3):583–594

Erfanifard Y, Sheikholeslami N (2017) Competitive interactions of Persian oak coppice trees (Quercus brantii var. persica) in a pure dry woodland revealed through point pattern analysis. Folia Geobot 52(1), 113–127.

Fischer F, Labrière N, Vincent G, Hérault B, Chave J (2020) A simulation method to infer tree allometry and forest structure from airborne laser scanning and forest inventories. Remote Sens Environ 251, 112056.

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

Goreaud F, Pélissier R (1999) On explicit formulas of edge effect correction for Ripley's K-function. J Veg Sci 10(3):433–438

Graves N (1996) Null models in ecology. Smithsonian Institution Press, Washington

Gu L, O'Hara K, Li WZ, Gong ZW (2019) Spatial patterns and inter-specific associations among trees at different stand development stages in the natural secondary forests on the Loess Plateau. China Ecol Evol 9(11):6410–6421

Hastings J, Ollinger S, Ouimette A, Sanders-DeMott R, Orwig D (2020) Tree species traits determine the success of LiDAR-based crown mapping in a mixed temperate forest. Remote Sensing 12(2):309

Henning J, Radtke P (2006) Ground-based laser imaging for assessing three-dimensional forest canopy structure. Photogramm Eng Remote Sens 72(12):1349–1358

Illian J, Penttinen A, Stoyan H, Stoyan D (2008) Statistical analysis and modelling of spatial point patterns (Vol. 70). John Wiley Sons.

Kwak D, Lee W, Lee J, Biging G, Gong P (2007) Detection of individual trees and estimation of tree height using LiDAR data. J for Res 12(6):425–434

Lin YC, Chang LW, Yang KC, Wang HH, Sun IF (2011) Point patterns of tree distribution determined by habitat heterogeneity and dispersal limitation. Oecologia 165(1):175–184

Lin Y, Wei T, Yang B, Knyazikhin Y, Zhang YH, Sato H, Fang X, Liang XL, Yan L, Sun SL (2016) TLS-bridged co-prediction of tree-level multifarious stem structure variables from worldview-2 panchromatic imagery: a case study of the boreal forest. Int Digital Earth 10(7):701–718

Mead R (1966) A relationship between individual plant spacing and yield. Ann Bot 30:301–309

Miao N, Liu SR, Yu H, Shi ZM, Moermond T, Liu Y (2014) Spatial analysis of remnant tree effects in a secondary Abies-Betula forest on the eastern edge of the Qinghai-Tibetan Plateau, China. For Ecol Manag 313:104–111

Millikan P, Silva C, Rodriguez L, Oliveira T (2019) Automated individual tree detection in Amazon tropical forest from airborne laser scanning data. Cerne 25(3):273–282

Moreno-Fernández D, Ledo A, Cañellas I, Montes F (2020) Strategies for modeling regeneration density in relation to distance from adult trees. Forests 11(1):120

Nguyen H, Bai L (2010) Cosine Similarity Metric Learning for Face Verification. Computer Vision. ACCV. Springer Berlin Heidelberg, 709–720.

Nguyen H, Uria-Diez J, Wiegand K (2016) Spatial distribution and association patterns in a tropical evergreen broad-leaved forest of north-central Vietnam. J Veg Sci 27(2):318–327

Nguyen H, Wiegand K, Getzin S (2014) Spatial patterns and demographics of Streblus macrophyllus trees in a tropical evergreen forest, Vietnam. J Trop For Sci, pp 309–319.

Packalen P, Vauhkonen J, Kallio E, Peuhkurinen J, Pitkänen J, Pippuri I, Strunk J, Maltamo M (2013) Predicting the spatial pattern of trees by airborne laser scanning. Int J Remote Sens 34(14):5154–5165

Pillay T, Ward D (2012) Spatial pattern analysis and competition between Acacia karroo trees in humid savannas. Plant Ecol 213(10):1609–1619

Podlaski R (2019) Models of the fine-scale spatial distributions of trees in managed and unmanaged forest patches with Abies alba Mill. and Fagus sylvatica L. For Ecol Manag 439:1–8

Ripley B (1976) The second-order analysis of stationary point processes. J Appl Probab 13(2):255–266

Taubert F, Jahn M, Dobner H, Wiegand T, Huth A (2015) The structure of tropical forests and sphere packings. Proc Natl Acad Sci 112(49):15125–15129

Velázquez E, Martínez I, Getzin S, Moloney K, Wiegand T (2016) An evaluation of the state of spatial point pattern analysis in ecology. Ecography 39(11):1042–1055

Wang XG, Wiegand T, Hao ZQ, Li BH, Ye J, Lin F (2010) Species associations in an old-growth temperate forest in north-eastern China. J Ecol 98(3):674–686

Wiegand T, Moloney K (2004) Rings, circles, and null-models for point pattern analysis in ecology. Oikos 104(2):209–229

Wiegand T, Moloney K (2013) Handbook of spatial point-pattern analysis in ecology. CRC Press

Xin HL, Malhi Y, Coomes D, Lin Y, Liu B L, Yang Q L, Shenkin A (2021) Individual tree detection and crown segmentation based on metabolic theory from airborne laser scanning data. J Appl Remote Sens 15(3), 034504.

Xu WH, Su ZF, Feng ZK, Xu H, Jiao YQ, Yan F (2013) Comparison of conventional measurement and LiDAR-based measurement for crown structures. Comput Electron Agr 98:242–251

Zaforemska A, Xiao W, Gaulton R (2019) Individual tree detection from UAV Lidar data in a mixed species woodland. ISPRS Geospatial Week.

Zhang ZH, Hu G, Zhu JD, Luo DH, Ni J (2010) Spatial patterns and interspecific associations of dominant tree species in two old-growth karst forests. SW China Ecol Res 25(6):1151–1160

Zhao D, Pang Y, Li ZY, Liu LJ (2014) Isolating individual trees in a closed coniferous forest using small footprint LiDAR data. Int J Remote Sens 35(20):7199–7218