Intertype mark correlation function: A new tool for the analysis of species interactions

Alicia Ledo [a,*], Sonia Condés [a], Fernando Montes [b]

[a] *Escuela Técnica Superior de Ingenieros de Montes, Universidad Politécnica de Madrid, Ciudad Universitaria, sn, 28040 Madrid, Spain*
[b] *CIFOR-INIA, Ctra. de La Coruña km 7.5, 28040 Madrid, Spain*

ARTICLE INFO

ABSTRACT

The spatial pattern of the different species in complex ecosystems reflects the underlying ecological processes. In this paper a second order moment function is proposed and tested to analyse the spatial distribution of a mark, which could be a tree characteristic such as diameter or height, between two different types of points, which could be two different tree species. The proposed function was a conditional density function based on the intertype $K_{rs}(d)$ function, incorporating as test function the correlation of the marks between pairs composed of points of different types. The results obtained in simulated and real plots prove that the function is capable of revealing the scale at which spatial correlation of the mark between two types of points exists. The proposed function allows the spatial association between individuals of different species at different life stages to be identified. This analysis may reveal information on species ecology and interspecific interactions in forest ecosystems.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Analysing the spatial pattern of the different elements of the ecosystem is a key aspect of population ecology theory (Legendre and Fortin, 1989; Chave, 1999; McIntire and Fajardo, 2009). The way in which different species or populations coexist in complex ecosystems provides the key to understanding these communities and therefore any findings in this regard must be taken into account when developing guidelines for their management and conservation (Goreaud and Pélissier, 2003). Understanding the spatial variability of processes is a key aspect of ecology (Lortie et al., 2005) since a number of ecological processes such as establishment, competition and mortality tend to follow specific patterns in the stand. Hence, by analysing the resulting overall pattern it is possible to infer the underlying processes (Borcard and Legendre, 1994). In recent years, new spatial analysis methods based on the second order characteristics of the distribution of individuals have been developed to analyse the spatial structure of ecosystems. The $K(d)$ function (Ripley, 1977) has been employed to describe the second moment properties of the spatial pattern as a function of the inter-point distance in a wide range of ecological applications (Dixon, 2002; Wiegand et al., 2007; Picard et al., 2009). The different interactions among populations induce identifiable spatial structures (Goreaud and Pélissier, 2003). Many biological ques-

tions are concerned with the relationships between two or more point patterns, the analysis of which requires the use of multivariate methods (Dixon, 2002). Several functions have been developed to analyse bivariate spatial point processes (Lotwich and Silverman, 1982; Wiegand and Moloney, 2004). These functions can be used to analyse the spatial relationship between two species (Duncan, 1991; He and Duncan, 2000) or between individuals in different life stages (Pélissier, 1998; Mast and Veblen, 1999). In many cases, the analysis of the spatial arrangement of a variable (or mark), such as the diameter or height of the trees, provides a better insight into the biological processes (Picard et al., 2009). The spatial correlation of a mark associated with a point process can be analysed using mark correlation functions (Stoyan and Penttinen, 2000; Schlather, 2001; Law et al., 2009).

Many ecological questions such as the influence of the ecological processes and the environmental conditions on the species distribution can be better addressed by identifying the position and size of individuals of different species (Plotkin et al., 2002). The relationships between species and environmental factors, as well as species strategies and interactions, vary over the life history of species (Comita et al., 2007). For example, if two species have different dispersion strategies, spatial repulsion between the young individuals of both species could appear. However, if these species require similar environmental conditions to develop, the adult individuals of both species may show spatial aggregation linked to those environmental conditions. In such cases, if the spatial association among species is analysed, including in the analysis all the individuals of each species regardless of their life stage, inter-specific ecosystem-

species associations at a determined life stage may be hidden. For this reason, it may be necessary to analyse the spatial relationship between the individuals of both species according to the different life stages in order to reveal the existing inter-specific interactions in the ecosystem, restricting the analysis to a few dominant species (Comas and Mateu, 2007).

Within the framework of research currently being carried out by the authors on the characterization of the spatial structure and the processes involved in Tropical Cloud Forest, a new tool has been developed to analyse the relationship between two species, which takes into account the size class of the different individuals. The idea was to design a conditional density function, based on the intertype $L_{rs}(d)$ function (the symmetrical standardization of the $K_{12}(d)$ function; Dale, 1999), to assess the second order structure of the correlation of the marks between two species, i.e., an intertype mark correlation function.

The proposed function was tested using simulated bivariate marked distributions and an example using real data from a Tropical Montane Cloud Forest.

2. Materials and methods

2.1. The framework: second order moment functions

2.1.1. Univariate techniques

Ripley's $K(d)$ function (Ripley, 1977) gives the expected number of points within a given distance d of a randomly chosen point divided by the pattern intensity (λ) in the study area:

$$\lambda \hat{K}(d) = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\delta_{ij}(d)}{n}, \quad \delta_{ij}(d) = \begin{cases} 1 & \text{if } d_{ij} \leq d \\ 0 & \text{if } d_{ij} > d \end{cases}$$

where n is the number of individuals and d the distance at which the $K(d)$ function is evaluated.

When computing second order moment measures, the boundary effect must be taken into account (Ripley, 1981; Illian et al., 2008). In this study, we use Ripley's boundary effect correction (Ripley, 1977) using the explicit formulas given in Goreaud and Pélissier (1999). δ_{ij} is substituted by ω_{ij}, which is calculated as the inverse of the fraction contained within the plot of a circumference centred on point i and passing through point j (Ripley, 1981):

$$\lambda \hat{K}(d) = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\omega_{ij}(d)}{n}$$

The $K(d)$ function calculated for the empirical data is compared with the values for a null model, usually a Poisson (homogeneous or inhomogeneous) process. A Montecarlo simulation is usually used, consisting of 99 iterations of the null model, taking into consideration a band between the 95% quantile bounds (Ripley, 1981).

The $L(d)$ function (Besag in discussion of Ripley, 1977), which is normally used, linearizes the $K(d)$ and stabilizes the variance, allowing a better visual interpretation:

$$\hat{L}(d) = \sqrt{\frac{\hat{K}(d)}{\pi}} - d$$

$K(d)$ and $L(d)$ are cumulative functions, implying that for each distance d the value of the function accounts for all the pairs of points separated by distances less than d. If the only pairs of points which are taken into account are those with distances contained within a distance lag centred at d then the function is denominated an o-ring type, which is preferable for some analyses (Condit et al., 2000; Wiegand et al., 2007; Lawes et al., 2008; Law et al., 2009; Picard et al., 2009). The non-cumulative function has the advantage that it can isolate the analysis for specific distances (Wiegand and

Moloney, 2004), although the cumulative functions display better statistical properties in goodness-of-fit tests (Stoyan and Penttinen, 2000).

2.1.2. Bivariate techniques

In order to identify the spatial attraction or repulsion produced by the individuals of one type over the individuals of another type, Lotwich and Silverman (1982) developed the intertype $L_{12}(d)$ function:

$$\hat{L}_{12}(d) = \sqrt{\frac{\hat{K}_{12}}{\pi}} - d; \quad \hat{L}_{21}(d) = \sqrt{\frac{\hat{K}_{21}}{\pi}} - d$$

where $\hat{\lambda}_2\hat{K}_{12}(d) = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}(\omega_{ij}(d))/n_1$, $\hat{\lambda}_1\hat{K}_{21}(d) = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}(\omega_{ij}(d))/n_2$, $\hat{\lambda}_1 = n_1/A$, $\hat{\lambda}_2 = n_2/A$ and A is the extent of the study area.

The $L_{12}(d)$ function is asymmetric, so the pattern of attraction or repulsion may be different with $L_{12}(d)$ or with $L_{21}(d)$. The symmetric $L_{rs}(d)$ function (Dale, 1999) takes into account all the pairs of individuals of different types:

$$\hat{L}_{rs}(d) = \sqrt{\frac{1}{(\lambda_r + \lambda_s)\pi}\left(\sum_{i=1}^{n_r}\sum_{j=1}^{n_s}\frac{\omega_{ij}(d)}{n_r} + \sum_{i=1}^{n_r}\sum_{j=1}^{n_s}\frac{\omega_{ij}(d)}{n_s}\right)} - d$$

where n_r and n_s are the number of trees of each class, $\lambda_r = n_r/A$ and $\lambda_s = n_s/A$ respectively.

This function can be used to analyse the attraction or repulsion between two types of points at different scales. The most common null models are the toroidal shift and the random labelling models (Goreaud and Pélissier, 2003). The toroidal shift model keeps the position of the points of one class unchanged and shifts all the points of the other class by the same random vector. The random labelling model maintains the joint spatial pattern of both classes of points and randomly assigns the label indicating the class.

2.1.3. Analysis of marked point processes

In the analysis of marked point processes, the aim is to assess the spatial correlation of the marks as a function of the distance. Several conditional mean functions $\kappa_f(d)$ (in which $f(m_i, m_j)$ is a suitable test function and where m_i and m_j are the values of the mark at the points i and j respectively) have been described for analysing the spatial correlation of the mark (Stoyan and Penttinen, 2000).

In the $k_{mm}(d)$ function (Stoyan, 1984) $f(m_i, m_j) = m_i m_j$, and the $\kappa_f(d)$ is normalized by the squared mean mark ($\bar{m}^2$):

$$\hat{k}_{mm}(d) = \frac{\sum_{i=1}^{n}\sum_{\substack{j=1 \\ j \neq i}}^{n}((m_i m_j)/\bar{m}^2)I_{ij}(d+\delta, d-\delta)}{\sum_{i=1}^{n}\sum_{\substack{j=1 \\ j \neq i}}^{n}I(d+\delta, d-\delta)}$$

The $I_{ij}(d+\delta, d-\delta)$ function takes a value of 0 if the points are not within the distance interval $(d+\delta, d-\delta)$ or 1 if they are. The width of distance interval 2δ determines the number of pairs of points used to evaluate the test function in each lag and the number of distance lags in which the function is tested. If the marks are not space correlated, $\hat{k}_{mm}(d) = 1$. If positive correlation exists, the $\hat{k}_{mm}(d) > 1$ and if negative correlation exists, $\hat{k}_{mm}(d) < 1$.

The mark variogram, where $f(m_i, m_j) = (1/2)(m_i - m_j)^2$ (Cressie, 1993), is also commonly employed:

$$\hat{\gamma}(d) = \frac{1}{2}\left(\frac{\sum_{i=1}^{n}\sum_{\substack{j=1 \\ j \neq i}}^{n}(m_i - m_j)^2 I(d+\delta, d-\delta)}{\sum_{i=1}^{n}\sum_{\substack{j=1 \\ j \neq i}}^{n}I(d+\delta, d-\delta)}\right)$$

Cumulative functions can also be derived:

$$K_{mm}(d) = \int \kappa_{mm}(d)g(d)2\pi d d(d)$$

where $g(d)$ is the expected number of points at distance d from each randomly chosen point. One estimator of this function is the $L_m(d)$ cumulative function derived from Stoyan's $k_{mm}(d)$ (Dale et al., 2002):

$$\hat{K}_m(d) = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \omega_{ij}(d) \cdot \frac{m_i m_j}{\lambda^2}; \qquad \hat{L}_m(d) = \sqrt{\frac{\hat{K}_m(d)}{\bar{m}^2\pi}} - d$$

2.2. Intertype mark correlation function for bivariate marked point processes

The proposed function is a conditional mean function based on the intertype $\hat{L}_{rs}(d)$ function, incorporating as test function the correlation of the marks between pairs composed of points of different types (similar to the $k_{rx}(d)$ function employed by Montes and Cañellas, 2007) which would be a special case of the new function when only one of the two point patterns is marked. The new function can be defined as the expected correlation of the marks between pairs of points of different types within a given distance:

$$\hat{K}_{mm}^{rs}(d) = \frac{\sum_{i=1}^{n_r}\sum_{j=1}^{n_s}\omega_{ij}(((m_i - \bar{m}_r)/s_r)((m_j - \bar{m}_s)/s_s)) + \sum_{i=1}^{n_r}\sum_{j=1}^{n_s}\omega_{ji}(((m_i - \bar{m}_r)/s_r)((m_j - \bar{m}_s)/s_s))}{\sum_{i=1}^{n_r}\sum_{j=1}^{n_s}\omega_{ij} + \sum_{i=1}^{n_r}\sum_{j=1}^{n_s}\omega_{ji}}$$

The $\hat{K}_{mm}^{rs}(d)$ function calculated for the empirical data should be compared with the values obtained through simulations of a suitable null model. The random marking null model was used to test independence in the mark distribution, keeping fixed the spatial pattern of both species and randomizing the mark. If the two types of points show similar normalized values of the mark at a certain scale, the empirical $\hat{K}_{mm}^{rs}(d)$ function will be above the quantile

bounds, and if the opposite occurs, it will be below the quantile bounds.

Hence, this function allows us to analyse the correlation of the mark between two different types of points at different scales and, therefore, it could reveal the scale at which spatial attraction or repulsion exists between trees of different species belonging to different life stages.

As the non-cumulative functions could be useful for exploratory data analysis given the ease with which they are interpreted (Stoyan and Penttinen, 2000), we also developed and tested the o-ring version of the proposed function, accounting for those pairs of points located at distances that fall within the lag $(d + \delta, d - \delta)$. The analyses were carried out for different δ and the results were compared with those obtained using the cumulative function.

2.3. Plot simulation

The performance of the proposed function was tested using simulated plots with simple inter-specific attraction or repulsion processes, generating the mark randomly or positively or negatively correlated. In our simulated plots, the two types of points correspond to two different tree species in a given forest stand and the mark represents the diameter at breast height (*DBH*). The aim is to determine whether or not $\hat{K}_{mm}^{rs}(d)$ is capable of detecting if spatial association between species according to their size exists and the scales at which it occurs. The different cases simulated were representative of different competition and facilitation processes that can take place in forest ecosystems, to assess whether the proposed functions reflect the patterns resulting from these processes. 99 simulations were carried out for each process. The plot size (1 ha square) and the intensity (200 points/ha) was the same for all the
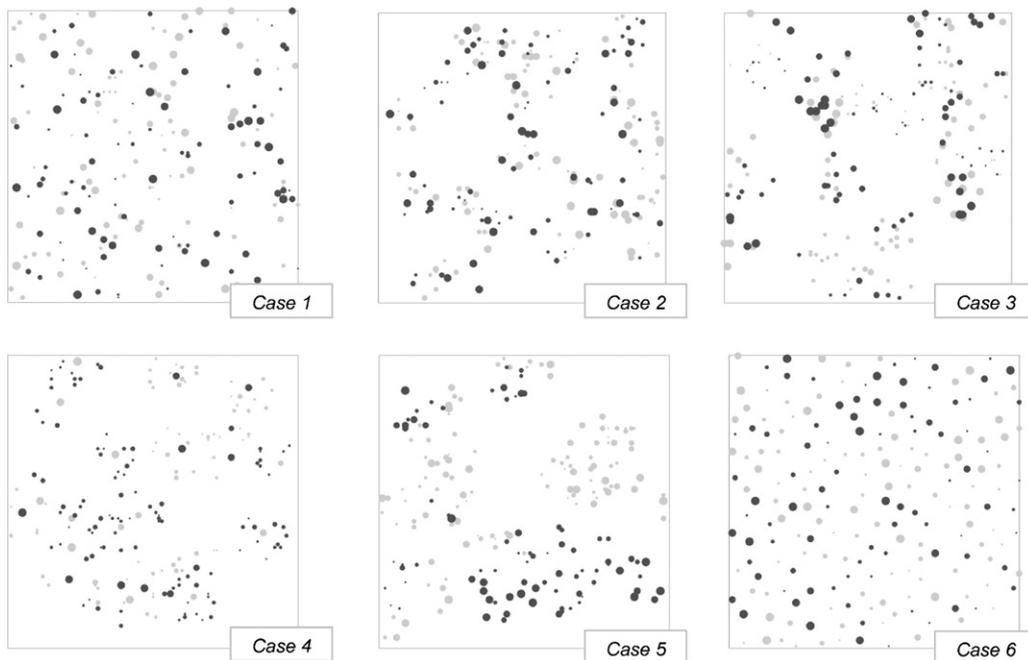

Fig. 1. A single realization of each of the six simulated bivariate marked point processes: (1) homogeneous Bernoulli process with independent marks, (2) Neyman–Scott process with independent marks, (3) Neyman–Scott process with positively correlated marks, (4) Neyman–Scott process with negatively correlated marks, (5) intra-specific Neyman–Scott process with intra-specific positive correlation of the mark and (6) Strauss Hard core process with independent marks. Each circle corresponds to a tree graduated by its diameter (mark). Each symbol corresponds to a different species (class).

Table 1
Simulated processes.

	Spatial pattern	Mark distribution
Case 1	Random	Independent
Case 2	Cluster—mixed	Independent
Case 3	Cluster—mixed	Positively correlated
Case 4	Cluster—mixed	Negatively correlated
Case 5	Cluster—monospecific	Intra-specific positively correlated
Case 6	Regular	Independent

simulated plots, comprising a similar number of points of each type, in order to obtain comparable envelope tests. These values would seem to provide a realistic compromise between a viable plot size in multi-species stands and having a sufficient number of pairs of individuals of different species in order to perform the analyses. The different cases of simulated plots are shown in Table 1. A single realization of each of the six simulated bivariate marked point processes is shown in Fig. 1.

Case 1: *Homogeneous Bernoulli process with independent marks*: The pattern was simulated through a homogeneous Bernoulli process, randomly generating the location of each event. Each event was randomly labelled for the species and its mark was generated following a uniform distribution within interval (0; 100), which was the most frequent *DBH* range in our real plots. This could be the case of a distribution with complete independence between species.

Case 2: *Neyman–Scott process with independent marks*: The spatial distribution of the points was simulated through a Neyman–Scott process (Diggle, 2002). The probability of a new individual being classified as parent event was 0.1; so the probability of being classified as offspring was 0.9. Parent events were established according to a Bernoulli distribution. The location of each offspring event was randomly generated, eliminating those individuals located outside a determined influence area around the parents. The radius of the influence area was 10 m. Random labelling was used for the species and the mark was generated within the interval (0; 100) following a uniform distribution. Thus, the two species appear in mixed clusters, and their diameters can be considered randomly distributed in each cluster. An example of this spatial distribution might be where two tree species require a similar environmental condition in order to recruit and develop.

Case 3: *Neyman–Scott process with positively correlated marks*: The spatial distribution of both types of point patterns was simulated as in Case 2 according to a Neyman–Scott process, but whereas for each parent event i, the mark (DBH_i) was randomly generated, for each individual offspring established within the area of influence around the parent event i, the mark was generated following a uniform distribution within the interval ($DBH_i - \Delta$; $DBH_i + \Delta$); Δ being equal to $0.1 \cdot DBH_i$. In this case, the two species were mixed in clusters and the diameters were similar for all the individuals within each cluster. An example of this spatial distribution could be where two species recruit under similar conditions and both species have similar requirements for development into adult trees. Another possible case would be where the development of a given species depends on the existence of another in that zone.

Case 4: *Neyman–Scott process with negatively correlated marks*: The spatial distribution of both point patterns was simulated as in Case 2. The marks for the parents followed a uniform distribution within the interval ($DBH_p - \Delta$; $DBH_p + \Delta$), where DBH_p was set equal to 70 and Δ equal to $0.3 \cdot DBH_p$. The mark of each individual offspring followed a uniform distribution within the interval ($DBH_o - \Delta$; $DBH_o + \Delta$), where DBH_o was set equal to 20 and Δ equal to $0.2 \cdot DBH_o$. This pattern may reflect facilitation between the adult trees of one species and the offspring of the other species.

Case 5: *Intra-specific Neyman–Scott process with intra-specific positive correlation of the mark*: The spatial distribution of parent and offspring events was simulated as in Case 2. The parents were randomly labelled and their diameter randomly generated following a uniform distribution within the interval (0; 100). Within each cluster, each new offspring was labelled following a binomial distribution with a 95% probability of being the same species as the parent. The mark of each individual offspring was generated as in Case 3, following a uniform distribution within the interval ($DBH_i - \Delta$; $DBH_i + \Delta$), DBH_i being the mark of the parent and Δ equal to $0.1 \cdot DBH_i$. The result is a cluster pattern where individuals of one species with similar diameters predominate within each cluster. This spatial distribution may appear when two species display an antagonistic response to a given habitat condition (e.g. light requirements).

Case 6: *Strauss Hard core process with independent marks*: The spatial pattern was created according to a hard-core process, which is a special type of Strauss process (Kelly and Ripley, 1976). The coordinates of the first event were randomly generated and each new event was randomly located. If the new event is within an exclusion buffer area around the pre-established events it is eliminated. Each event was randomly labelled and its mark was generated following a uniform distribution within the interval (0; 100). In this case, the two species display a regular distribution and the diameters are independent. This may reflect strong competition for resources.

These simulations are hypothetical, and hence artificial; but although the simulated process ought to be close to the reality; the main aim is to illustrate the performance of the new tools proposed.

2.4. Application to a Tropical Cloud Forest data set

The new function was used to analyse the species interaction according to different life stages in a 1 ha plot situated in an Andean Montane Cloud Forest, in North Peru, described in Ledo et al. (2009). We analysed the spatial association between a representative canopy species (*Oreopanax raimondii* Harms.) and three typical species of the different phyto-sociological layers: an under-canopy species (*Critoniopsis sevillana* (Cuatrec.) H. Rob.), a canopy species (*Delostoma integrifolium* D. Don) and an undercanopy o canopy (facultative) species (*Myrcianthes fimbriata* (Kunth) McVaugh.) was analysed. In this case study, the type corresponds to the different species and the mark is the diameter of the tree.

2.5. Deviation measure and significance test

To assess the significance of the deviation from the null model for each simulation of the different simulated processes as well as for the real plots, the *p*-value of the deviation test (Loosmore and Ford, 2006) was calculated. In addition, for the 99 simulations of each simulated process, the envelope test was performed using the maximum and minimum values of the $\hat{K}_{mm}^{rs}(d)$ function and the maximum and minimum values of the 95% quantile bounds obtained through the 99 iterations of the random marking null model for each simulation.

All the computations were done using a program developed using VisualBasic (accessible by contacting the authors).

3. Results

3.1. Simulated processes

The mean $\hat{K}_{mm}^{rs}(d)$ function and the mean quantile bounds of the null models for the 99 simulations of each process are shown in Fig. 2. The parameters used appear to be adequate for testing the

Cumulative functions

Envelope test
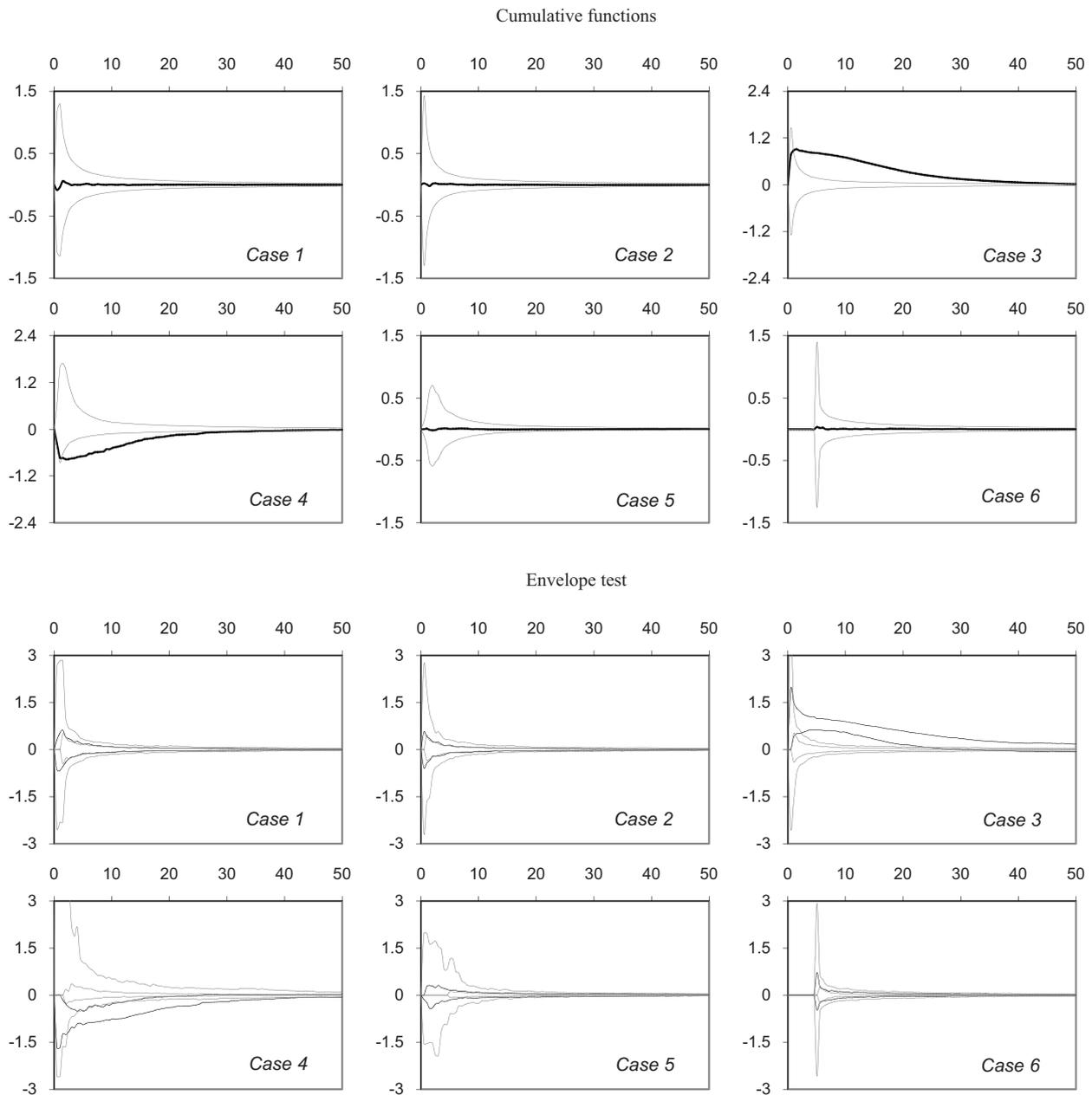
Fig. 2. (a) Mean empirical $\hat{K}_{mm}^{rs}(d)$ function for the 99 simulations of the six bivariate marked processes shown in Fig. 1 (black line) and quantile bounds of the random marking null model (grey lines). (b) Maximum and minimum values of the empirical $\hat{K}_{mm}^{rs}(d)$ function for the 99 simulations of the six bivariate marked processes (black lines) and the maximum and minimum values of the 95% quantile bounds obtained through the 99 iterations of the random marking null model for each simulation (grey lines).

function, avoiding extreme distributions so as to enable us to discern whether the function effectively reflects the simulated pattern.

The empirical $\hat{K}_{mm}^{rs}(d)$ function was compared with the random marking quantile bounds to test the mark correlation; successfully detecting the positive mark correlation in case 3 and negative correlation in case 4, whilst revealing no correlation in the cases where the marks were randomly assigned (Fig. 2). The significance test also showed values of $p < 0.05$ for cases 3 and 4 and values greater than 0.30 in other cases (Table 2). The envelope test indicates that the proposed function correctly describes the mark correlation in the simulations (Fig. 2).

The results obtained in the o-ring analysis are quite similar to the results obtained for the cumulative functions (Fig. 3). A δ of 2.5 m was chosen, instead of the 0.5 m lags employed for the cumulative functions, in order to increase the number of pairs of points within each distance lag, thereby smoothing out the function fluc-

tuations. The p-values and the envelope test indicate that the o-ring functions show a slightly greater tendency than the cumulative functions to detect deviations from the null model in some simulations of the cases where the marks were randomly assigned (Table 2 and Fig. 3).

3.2. Tropical Cloud Forest example

The analysis carried out using the Montane Cloud Forest data showed some significant pairwise interactions that reflect the feasibility of different ecological associations (Fig. 4).

The $\hat{K}_{mm}^{rs}(d)$ function reflected repulsion of the mark at small distances between the canopy species *O. raimondii* and the under-canopy *C. sevillana* ($p = 0.05$), with a trend towards spatial attraction at this scale (Fig. 4). Hence, the smaller trees of one species appear near the larger trees of the other species. A possible reason could be

Table 2
Mean values, percentiles at 95%, maximum and minimum *p*-values obtained, and percentage of correct significant deviation assessment (% acc) for the 99 simulations of the six bivariate marked processes shown in Fig. 1. $p \leq 0.05$ were considered significant deviations from the random marking null model, and $p > 0.25$ were considered non-significant deviations.

	Cumulative functions						O-ring functions					
	Mean	Pr 95	Pr 5	max	min	% acc	Mean	Pr 95	Pr 5	max	min	% acc
Case 1	0.675	0.97	0.35	1	0.34	100	0.594	0.92	0.27	0.99	0.18	99
Case 2	0.680	0.96	0.38	0.99	0.27	100	0.644	0.95	0.34	0.96	0.13	99
Case 3	0.010	0.01	0.01	0.01	0.01	100	0.010	0.01	0.01	0.01	0.01	100
Case 4	0.034	0.06	0.01	0.4	0.01	99	0.030	0.08	0.01	0.11	0.01	99
Case 5	0.553	0.93	0.3	0.99	0.26	100	0.611	0.92	0.26	0.99	0.13	97
Case 6	0.650	0.98	0.35	1	0.29	100	0.600	0.99	0.31	1	0.27	100
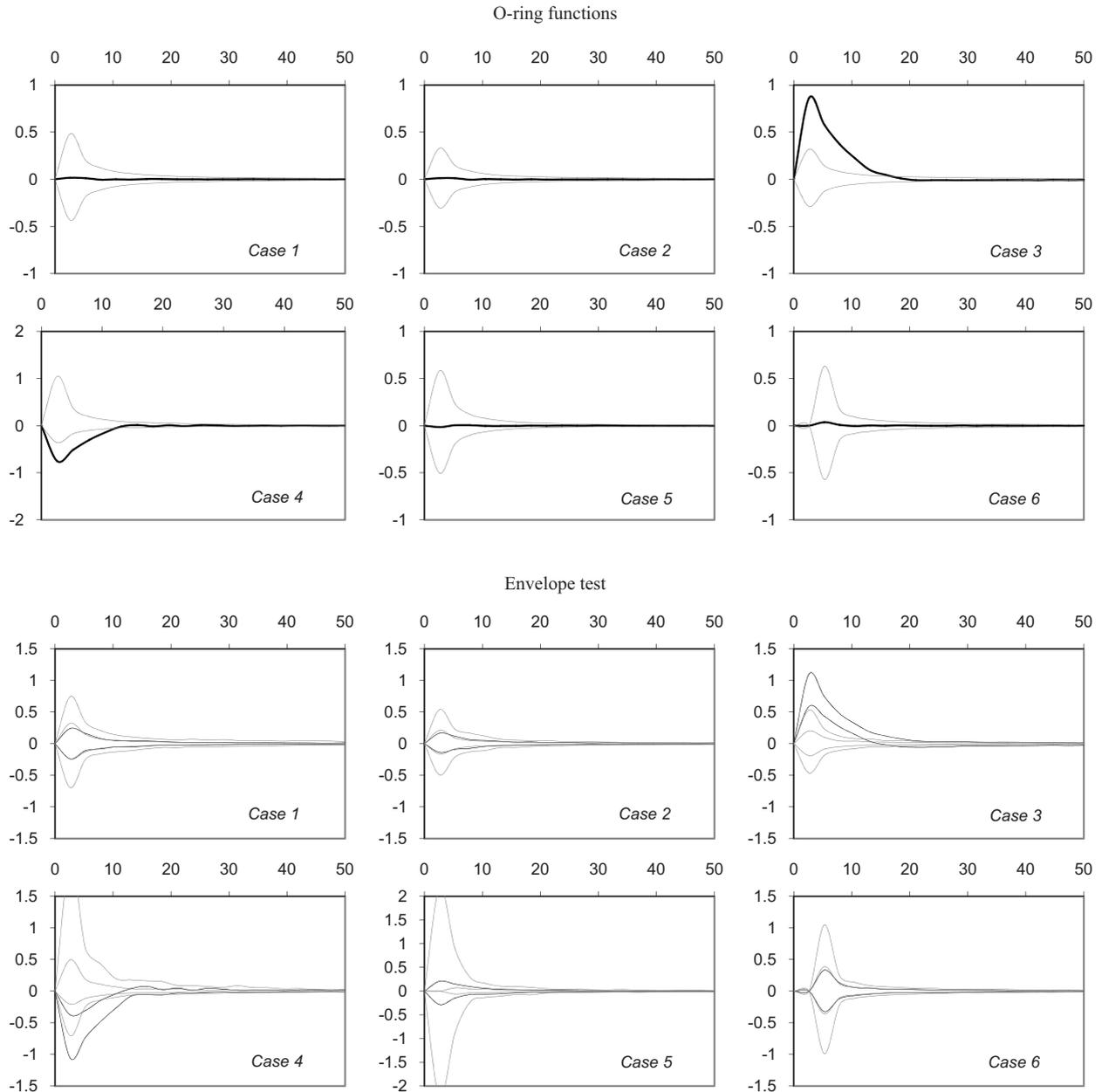
Fig. 3. (a) Mean empirical o-ring $\hat{K}_{mm}^{rs}(d)$ function for the 99 simulations of the six bivariate marked processes shown in Fig. 1 (black line) and quantile bounds of the random marking null model (grey lines). (b) Maximum and minimum values of the empirical o-ring $\hat{K}_{mm}^{rs}(d)$ function for the 99 simulations of the six bivariate marked processes (black lines) and the maximum and minimum values of the 95% quantile bounds obtained through the 99 iterations of the random marking null model for each simulation (grey lines).
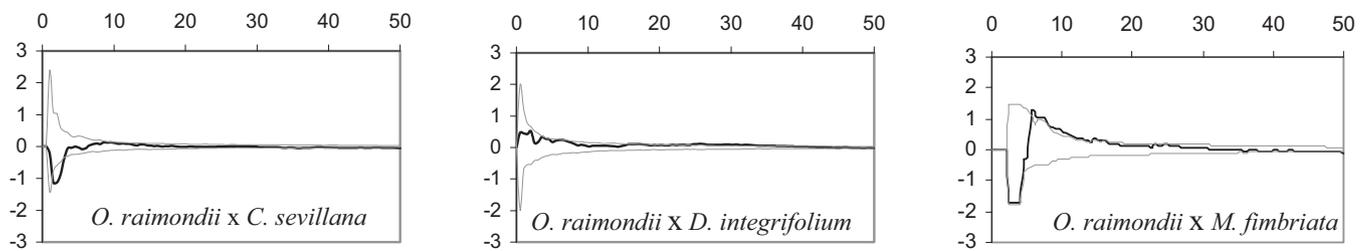
Fig. 4. Empirical $\hat{K}_{mm}^{rs}(d)$ function (black line) and quantile bounds of the random marking null model (grey lines) for the following species in the real study case: *Oreopanax raimondii* vs *Critoniopsis sevillana*; *Oreopanax raimondii* vs *Delostoma integrifolium* and *Oreopanax raimondii* vs *Myrcianthes fimbriata*.

that the *Critoniopsis* species are half shade-tolerant and the development of younger individuals of this species is favoured by the conditions under large *Oreopanax* trees, which do not have particularly dense foliage.

In the case of the interaction between the species *O. raimondii* and *D. integrifolium*, both belonging to the canopy strata, positive correlations of the mark ($p = 0.07$) were found at short distances but also at larger distances (Fig. 4). Hence, individuals at a similar life stage for these two species may appear spatially associated at two different scales, which may relate to factors which exist at the younger and canopy stages respectively.

Negative correlation of the mark at small distances followed by positive correlation ($p = 0.04$) was found between the *O. raimondii* and the facultative *M. fimbriata* (Fig. 4). This may reflect the co-occurrence of the younger trees of both species (the positive correlation), and the existence of high differentiation between species at nearest neighbour scale (the negative correlation).

4. Discussion

This paper presents a new second order moment technique that allows us to characterize the spatial correlations of the marks between two point patterns. Analyses of the spatial pattern of the different species in the ecosystem can provide valuable information regarding the role of each species or the regeneration and development strategies of each (Legendre and Fortin, 1989; Mateu et al., 1998). Sometimes the spatial pattern of interaction between species varies according to the different life stages. Furthermore, the habitat associations differ among life stages, so associations were not often consistent across stages (Comita et al., 2007). The technique proposed may be useful to identify inter-specific species associations at different life stages, which would be complex and awkward using the existing bivariate unmarked functions or the mark correlation functions (Comas and Mateu, 2007).

The results obtained with the simulated plots reveal the capability of the proposed function to identify the scale of correlation of the mark (which may be the age or diameter of the trees) between species. Hence, it may be used to detect the presence of spatial association between the individuals of both species at similar or at different life stages and thus can provide us with useful information about the ecological processes taking place in the ecosystem. Moreover, the proposed function standardizes the mark with respect to the mean for each species, which allows us to compare two species with different diameter ranges. As the $\hat{K}_{mm}^{rs}(d)$ function has been designed as a conditional density function, it isolates the correlation of the mark from the spatial pattern, reflecting only the correlation of the mark (Beisbart and Kerscher, 2000). The random marking model seems to be the most appropriate for testing independence in the mark distribution, conditioned to the spatial pattern of both types of points (Montes et al., 2008). Nevertheless, different null models might be used, depending on the context of the study and the biotic interactions under analysis (Goreaud and Pélissier, 2003).

The cumulative and the o-ring functions gave similar results in the deviation test for the simulated processes, although the distance lag must be increased in the case of the o-ring functions in order to stabilize the empirical functions. The cumulative functions commonly show better statistical properties than the o-ring functions in goodness-of-fit tests (Stoyan and Penttinen, 2000). The p-values estimated by the deviation tests would appear to suggest that the cumulative functions are more appropriate to identify significant correlations in forests with a large number of species, like the forest analysed in the case study, because the number of trees may be relatively small for some species.

The proposed intertype mark correlation function has been developed within the framework of research on species ecology and forest structure in a Montane Cloud Forest in the North Peruvian Andes. The hypotheses for explaining the high diversity values in tropical areas are mainly based on the analysis of the pattern of seedlings or young trees with respect to the parent trees (Janzen, 1970; Connell, 1971), the relationship of young trees with certain environmental characteristics (Grubb, 1977), or the absence of relationship between younger and older trees (Hubbell, 2001). Both Plotkin et al. (2002) and Wiegand et al. (2007) proposed a double-cluster analysis in tropical forests to gain a better understanding of the ecological processes at work. The analysis of the inter-species relationships in our real plots using the proposed tools reveals different inter-specific interactions and species associations between species of the different phyto-sociological layers. These functions may also be useful to check the validity of stochastic models with parameters relating to spatial relationships that can play an important role in understanding and predicting the behaviour of complex systems (Fortin et al., 2003). However, it is necessary to be cautious when interpreting results, since similar patterns may arise as a result of different ecological processes (Stoyan and Penttinen, 2000). The ecological meaning must also be kept in sight when using statistical tools. Point processes and marked point processes are powerful tools in modern forestry statistics, but more research into spatial pattern statistics and null models is still necessary.

5. Conclusions

A second order moment function was developed to analyse the positive or negative spatial association of the mark at different scales between two classes of marked points. The proposed $\hat{K}_{mm}^{rs}(d)$ function assesses the correlation of the marks between individuals belonging to two classes at different scales. Random marking would be an appropriate null model for testing independence in mark distributions.

The proposed technique overcomes some of the limitations of the existing bivariate functions and the mark correlation functions for the analysis of mark correlations between two point patterns and could provide valuable information on species ecology and interspecific interactions in forest ecosystems.

Acknowledgements

References

Beisbart, C., Kerscher, M., 2000. Luminosity- and morphology-dependent clustering of galaxies. The Astrophysical Journal 545, 6–25.

Borcard, D., Legendre, P., 1994. Environmental control and spatial structure in ecological communities: an example using Oribatid mites (Acari, Oribatei). Environmental and Ecological Statistics 1, 37–53.

Chave, J., 1999. Study of structural, successional and spatial patterns in tropical rain forests using TROLL, a spatially explicit forest model. Ecological Modelling 124, 233–254.

Comas, C., Mateu, J., 2007. Modelling forest dynamics: a perspective from point process methods. Biometrical Journal 49, 176–196.

Comita, L.S., Condit, R., Hubbel, S.P., 2007. Developmental changes in habitat associations of tropical trees. Journal of Ecology 95, 482–492.

Condit, R., Ashton, P.S., Baker, P., Bunyavejchewin, S., Gunatilleke, S., Gunatilleke, N., Hubbell, S.P., Foster, R.B., Itoh, A., LaFrankie, J.V., Lee, H.S., Losos, E., Manokaran, N., Sukumar, R., Yamakura, T., 2000. Spatial patterns in the distribution of tropical tree species. Science 288, 1414–1418.

Connell, J.H., 1971. On the role of the natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. In: den Boer, P.J., Gradwell, G. (Eds.), Dynamics of Populations. Centre for Agricultural Publishing and Documentation, Wageningen, the Netherlands, pp. 298–312.

Cressie, N.A.C., 1993. Statistics for Spatial Data. Wiley, New York.

Dale, M.R.T., 1999. Spatial Pattern Analysis in Plant Ecology. Cambridge University Press.

Dale, M.R.T., Dixon, P., Fortin, M.J., Legendre, P., Myers, D.E., Rosenberg, M.S., 2002. Conceptual and mathematical relationships among methods for spatial analysis. Ecography 25, 558–577.

Diggle, P.J., 2002. Statistical Analysis of Spatial Point Patterns. Oxford University Press, Oxford, UK.

Dixon, P.M., 2002. Ripley's K function. In: El-Shaaraui, A.H., Piegorsch, W.W. (Eds.), The Encyclopedia of Environmetrics. Willey, New York, pp. 1796–1803.

Duncan, R., 1991. Competition and the coexistence of species in a mixed Podocard stand. Journal of Ecology 79, 1073–1084.

Fortin, M.J., Boots, B., Csillag, F., Remmel, T.K., 2003. On the role of spatial stochastic models in understanding landscape indices in ecology. Oikos 102 (1), 203–212.

Goreaud, F., Pélissier, R., 1999. On explicit formulas of edge effect correction for Ripley's K-function. Journal of Vegetation Science 10, 433–438.

Goreaud, F., Pélissier, R., 2003. Avoiding misinterpretation of biotic interactions with the intertype K12-function: population independence vs. random labelling hypotheses. Journal of Vegetation Science 14, 681–692.

Grubb, P.J., 1977. The maintenance of species-richness in plant communities: the importance of the regeneration niche. Biological Reviews 52, 107–145.

He, F., Duncan, R.P., 2000. Density-dependent effects on tree survival in an old-growth Douglas fir forest. Journal of Ecology 88, 676–688.

Hubbell, S.P., 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press.

Illian, J., Penttinen, A., Stoyan, H., Stoyan, D., 2008. Statistical Analysis and Modelling of Spatial Point Patterns. Wiley.

Janzen, D.H., 1970. Herbivores and the number of tree species in tropical forests. The American Naturalist 104, 501–529.

Kelly, F.P., Ripley, B.D., 1976. A note on Strauss' model for clustering. Biometrika 63, 357–360.

Law, R., Illian, J., Burslem, D.F.R.P., Gratzer, G., Gunatilleke, C.V.S., Gunatilleke, I.A.U.N., 2009. Ecological information from spatial patterns of plants: insights from point process theory. Journal of Ecology 97, 616–628.

Lawes, M.J., Griffiths, M.E., Midgley, J.J., Boudreau, S., Eeley, H.A.C., Chapman, C.A., 2008. Tree spacing and area of competitive influence do not scale with tree size in an African rain forest. Journal of Vegetation Science 19, 729–738.

Ledo, A., Montes, F., Condés, S., 2009. Species dynamics in a Montane Cloud Forest: identifying factors involved in changes in tree diversity and functional characteristics. Forest Ecology and Management 258, S75–S84.

Legendre, P., Fortin, M.J., 1989. Spatial pattern and ecological analysis. Vegetatio 80, 107–138.

Loosmore, N.B., Ford, E.D., 2006. Statistical inference using the G or K point pattern spatial statistics. Ecology 87 (8), 1925–1931.

Lortie, C.J., Ellis, E., Novoplansky, A., Turkington, R., 2005. Implications of spatial pattern and local density on community-level. Oikos 109, 495–502.

Lotwich, H.W., Silverman, B.W., 1982. Methods for analysing spatial processes of several types of points. Journal of the Royal Statistical Society. Series B 44, 406–413.

Mast, J.N., Veblen, T.T., 1999. Tree spatial patterns and stand development along the pine–grassland ecotone in the Colorado Front Range. Canadian Journal of Forest Research 29, 575–584.

Mateu, J., Usó, J.L., Montes, F., 1998. The spatial pattern of a forest ecosystem. Ecological Modelling 108, 163–174.

McIntire, E.J.B., Fajardo, A., 2009. Beyond description: the active and effective way to infer processes from spatial patterns. Ecology 90 (1), 46–56.

Montes, F., Barbeito, I., Rubio, A., Cañellas, I., 2008. Evaluating height structure in Scots pine forest using marked point processes. Canadian Journal of Forest Research 38, 1924–1934.

Montes, F., Cañellas, I., 2007. The spatial relationship between post-crop remaining trees and the establishment of saplings in Pinussylvestris stands in Spain. Applied Vegetation Science 10, 151–160.

Pélissier, R., 1998. Tree spatial patterns in three contrasting plots of a southern Indian tropical moist evergreen forest. Journal of Tropical Ecology 14, 1–16.

Picard, N., Bar-Hen, A., Mortier, F., Chadoeuf, J., 2009. Understanding the dynamics of an undisturbed tropical rain forest from the spatial pattern of trees. Journal of Ecology 97, 97–108.

Plotkin, J.B., Chave, J., Ashton, P.S., 2002. Cluster analysis of spatial patterns in Malaysian tree species. The American Naturalist 16, 629–644.

Ripley, B.D., 1977. Modelling spatial patterns (with discussion). Journal of Royal Statistical Society 2, 172–212.

Ripley, B.D., 1981. Spatial Statistics. Cambridge University Press.

Schlather, M., 2001. On the second-order characteristics of marked point processes. Bernoulli 7, 99–117.

Stoyan, D., 1984. On correlations of marked point processes. Mathematische Nachrichten 116, 197–207.

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

Wiegand, T., Gunatilleke, S., Gunatilleke, N., Okuda, T., 2007. Analyzing the spatial structure of a Sri Lankan tree species with multiple scales of clustering. Ecology 88, 3088–3102.

Wiegand, T., Moloney, K.A., 2004. Rings, circles and null-models for point pattern analysis in ecology. Oikos 10, 209–229.