Evaluating height structure in Scots pine forests using marked point processes

Fernando Montes, Ignacio Barbeito, Agustín Rubio, and Isabel Cañellas

Abstract: In this study, the second-moment analysis of marked spatial point processes is applied to the characterization of the tree height distribution in two Scots pine (*Pinus sylvestris* L.) forests in the Central Mountain Range of Spain. The cumulative function $L_m(d)$ weighted by the normalized mark variance is proposed to analyse the second-order properties of marked point patterns. The empirical $L_m(d)$ was compared with two null models to assess the relationship between the spatial distribution of the trees and the tree height correlations: the first null model was used to characterize the spatial clustering of the trees and was derived from the complete spatial randomness model used with Ripley's $K(d)$ function. The second null model, which is derived from the random labelling model used with the intertype second-moment measure $K_{12}(d)$ (type 1 intensity conditioned to the type 2 intensity and vice versa), allows us to identify the mark correlations. The performance of the technique was assessed through simulated marked point patterns. The second-moment analysis showed that most of the analysed Scots pine stands have a uniform height distribution at small scale and greater heterogeneity at large scales, with the exception of an upper altitudinal stand, which exhibited heterogeneity at short distances. These results demonstrate the utility of second-moment analysis of marked point processes for characterizing height structure in forest stands and the interaction between the height and the spatial pattern of the trees.

Résumé : Dans cette étude, l'analyse du moment de second ordre des processus ponctuels marqués est utilisée pour caractériser la distribution des hauteurs d'arbre dans deux forêts de pin sylvestre (*Pinus sylvestris* L.) situées dans la cordillère centrale en Espagne. La fonction cumulative $L_m(d)$ pondérée par la variance normalisée est proposée pour analyser les propriétés du moment de second ordre des structures spatiales. La fonction empirique $L_m(d)$ est comparée à deux modèles de référence pour évaluer la relation entre la distribution spatiale des arbres et les corrélations entre les hauteurs d'arbre. Le premier modèle de référence, dérivé du modèle spatial entièrement aléatoire basé sur la fonction de Ripley $K(d)$, est utilisé pour caractériser le regroupement spatial des arbres. Le deuxième modèle de référence, dérivé du modèle de marquage aléatoire utilisé avec la mesure inter type du moment de second ordre $K_{12}(d)$ (intensité de type 1 conditionnelle à l'intensité de type 2 et vice-versa), permet d'identifier les corrélations. La performance de la technique est évaluée par l'intermédiaire des structures spatiales simulées. L'analyse du moment de second ordre montre que la plupart des peuplements de pin sylvestre analysés présentent une distribution uniforme des hauteurs à petite échelle et une plus grande hétérogénéité à grande échelle, à l'exception d'un peuplement situé à plus haute altitude dont l'hétérogénéité se manifeste sur de courtes distances. Ces résultats démontrent l'utilité de l'analyse du moment de second ordre des processus ponctuels marqués pour caractériser la structure de la hauteur des peuplements forestiers et l'interaction entre la hauteur et la répartition spatiale des arbres.

[Traduit par la Rédaction]

Introduction

The vertical stand structure constitutes an important aspect of forest ecosystems, reflecting key ecological processes such as intra- and inter-specific competition, shade tolerance of species, site conditions, and disturbance regime. Most studies have generally focused on two main aspects of vertical forest structure: height diversity and height differentiation (del Río et al. 2003). Foliage height diversity (FHD; MacArthur and MacArthur 1961) has been widely employed to characterize height diversity. However, the FHD is dependent on the class width and plot size, and there is no general agreement on what the class width should be or the sensitivity of the index to plot size (Ferris-Kaan et al. 1998; Varga et al. 2005). The use of indices based on the variance to quantify the diversity (Staudhammer and LeMay 2001) may solve the class-width problem associated with continuous variables. On the other hand, methods based on height differentiation (von Gadow 1993; Zenner and Hibbs 2000; Aguirre et al. 2003) present a picture of height variability at microscale, i.e., one tree and its neighbours. Second-moment analysis methods or geostatistical methods are also based on the distances between pairs of points, but they can provide information on a wider range of scales, taking into account all the inter-

F. Montes[1,2] **and A. Rubio.** Departamento de Silvopascicultura. Universidad Politécnica de Madrid, Ciudad Universitaria s/n. 28040 Madrid, Spain.

I. Barbeito and I. Cañellas. Centro de Investigación Forestal, Instituto Nacional de Investigación Agraria (CIFOR-INIA), Ctra. La Coruña, km 7,5, 28040 Madrid, Spain.

[1]Corresponding author (e-mail: fmontes@inia.es).
[2]Present address: Departamento de Silvopascicultura, Escuela Técnica Superior de Ingenieros de Montes, Universidad Politécnica de Madrid, Ciudad Universitaria s/n. 28040 Madrid, Spain.

Table 1. Stand-level variables in the experimental plots (extent 0.5 ha) in Navafría forest (N2–N5) and Valsaín forest (V1–V8).

Study site and plot	Plot size (m×m)	Age (years)	Trees/ha	Mean DBH (cm)	BA (m²/ha)	H (m)
Navafría						
N2	110×45.5	33	4488	14.1	41.8	12.0
N3	70.7×70.7	71	680	32.9	60.6	22.0
N4	70.7×70.7	99	364	40.6	48.2	20.4
N5	70.7×70.7	115	304	42.5	44.2	22.1
Valsaín						
V1	100×50	Uneven	584	20.2	34.4	13.6
V2	85×58.8	41	1668	16.0	41.0	14.3
V3	100×50	61	1322	20.5	48.5	16.4
V4	100×50	82	686	30.5	53.3	23.5
V5	70.7×70.7	106	550	34.6	54.0	22.1
V6	100×50	122	334	38.3	41.4	24.0
V7	70.7×70.7	Uneven	608	18.8	25.2	11.5
V8	100×50	Uneven	578	19.9	22.7	10.0

Note: Number of trees per hectare and diameter at breast height (DBH) were recorded for trees ≥10 cm DBH. BA, basal area; and H, mean height of trees with a DBH ≥ 10 cm.

tree distances and not just the minimum distance. The use of geostatistical methods implies the assumption that tree heights are a realization from a random field, independent of tree position, which does not always hold (Wälder and Stoyan 1996). Despite this shortcoming, some studies have shown that the variogram may prove useful for tree height autocorrelation analysis (Kuuluvainen et al. 1996; Kint 2003). However, these methods do not account for the spatial pattern of the trees or the influence of the latter on the vertical structure. Marked point processes have been applied to the analysis of the relationship between the spatial distribution of objects and their associated properties in fields as diverse as material research (Hirsch et al. 2000; Ballani et al. 2006), astrophysics (Beisbart and Kerscher 2000), or forestry (Penttinen et al. 1992; Gavrikov and Stoyan 1995). The analysis of marked point processes through their second-order characteristics may prove to be a powerful statistical tool for application in forestry (Stoyan and Penttinen 2000). However, marked point processes are not, as yet, commonly used in vertical forest structure analysis. Moreover, the methods developed for marked point processes have generally focused on the analysis of second-order characteristics for marks, separating the spatial pattern of the points (Schlather 2001). Our study demonstrates the potential of second-order analysis of marked point processes for analyzing forest vertical structure and the interaction between the height and the spatial pattern of the trees.

This study was carried out in two Scots pine (*Pinus sylvestris* L.) forests in the Central Mountain Range of Spain, each subject to a different forest management regime. The Scots pine constitutes one of the most widely studied species because of its broad distribution area (approximately 28×10^6 ha in Europe and northwestern Asia; Mason and Alía 2000), and the importance of its forests for timber (Kelly and Connolly 2000). However, under a climatic change scenario, the behaviour of the species in the analysed stands, located at the southern limit of its distribution area, is of great interest.

The objectives of this study are (*i*) to develop a method based on the second-moment measure of marked point proc-

esses to describe the interaction between the spatial pattern of the trees and the height structure of the stand and (*ii*) to obtain a precise characterization of the height structure of Scots pine stands at different developmental stages and under different ecological conditions in the Central Mountain Range of Spain.

Materials and methods

Experimental sites and data collection

The data in this study are derived from two managed Scots pine forests, Valsaín pinewood and Navafría pinewood, both situated on the northern slopes of the Sierra de Guadarrama, in the Central Mountain Range of Spain. A chronosequence trial consisting of various 0.5 ha plots covering the different developmental stages of the stand was established in each forest: five plots at Navafría and six in Valsaín. Plot N1 at Navafría was excluded from this analysis because of the high density of smaller trees. In Valsaín, two additional plots were established: one at the lower limit for Scots pine stands in the Central Mountain Range of Spain (1300 m), where the pinewood is replaced by Pyrenean oak (*Quercus pyrenaica* Willd.), and the other close to the tree line at 1700 m, where the pinewood is gradually replaced by alpine shrub, mainly composed of *Juniperus communis* var. *montana* Ait. and *Cytisus oromediterraneous* Rivas-Martínez (Table 1). The main differences between the forests are, on one hand, the more intensive thinning regime carried out in Navafría, and on the other, the method of regeneration, which takes the form of a shorter uniform shelterwood period in Navafría that leads to a more homogenous recruitment, whereas ithe shelterwood period in Valsaín is longer and results in the coexistence of the newly established stand and the previous stand, as in plot V1 (Montes et al. 2005).

In each rectangular 0.5 ha plot, the tree height (*H*) and diameter at breast height (DBH) were measured, and the position of all trees with DBH > 10 cm was mapped. The characterization of the stand in the different plots is shown in Table 1.

Vertical structure analysis

Second-moment analysis of marked point processes

The second-order characteristics of a point process can be described through the intensity λ and Ripley's $K(d)$ function (Ripley 1977):

$$[1] \qquad \lambda K(d) = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{\omega_{ij}(d)}{n}$$

where d is the distance, n the total number of trees in the study region, and $\omega_{ij}(d)$ gives 0 if stems i and j are more than distance d from each other, or if not, it gives the inverse of the fraction of a circumference centred on i (with the radius being the distance from i to j) that falls within the plot (Ripley 1977; Goreaud and Pélissier 1999).

The $L(d)$ function (eq. 2) is normally used instead of the $K(d)$ function because of its stable variance and expected 0 value under a Poisson distribution:

$$[2] \qquad \widehat{L}(d) = \sqrt{\frac{\widehat{K}(d)}{\pi}} - d$$

When describing the second-order characteristics of marked point processes, the interest is centered on the correlations of the marks conditioned to the corresponding points separated by a distance d. Where m_i is the value of the mark at point i and m_j is the value of the mark at point j, given a non-negative weighting function $f(m_i, m_j)$, the mean conditional density of $f(m_i, m_j)$ is defined as (Stoyan and Stoyan 1994):

$$[3] \qquad \kappa(d) = \frac{\rho_f^{(2)}(d)}{\rho^{(2)}(d)}$$

where $\rho^{(2)}(d)$ is the product density function for the unmarked process ($f(m_i, m_j) \equiv 1$) and $\rho_f^{(2)}(d)$ is the product density function of $f(m_i, m_j)$. The product density gives the mean $f(m_i, m_j)$ function for all pairs of points separated by a distance d. In practice, $f(m_i, m_j)$ is averaged over a bandwidth δ around the distance d, multiplied by the probability of two points being separated by a distance d (see Schlather (2001) and Hirsch et al. (2000) for details). Just as $\rho^{(2)}(d)$ coincides with the product density in the case of unmarked points, $\kappa(d)$ depends exclusively on the mark correlations (Beisbart and Kerscher 2000).

Several mark-weighted conditional correlation functions have been derived from $\kappa(d)$ using different weighting functions. If $f(m_i, m_j) = m_i \times m_j$, then Stoyan's $k_{mm}(d)$ function (Stoyan 1984) can be estimated as the $\kappa(d)$ function normalized by the squared mark mean

$$[4] \qquad \widehat{k}_{mm}(d) = \frac{\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \left[\frac{m_i m_j}{\overline{m}^2} I_{ij}(d + \delta, d - \delta) \right]}{\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} [I_{ij}(d + \delta, d - \delta)]}$$

where $I_{ij}(d + \delta, d - \delta)$ is an indicator function that takes the value 0 if point j is outside the ring defined by two circum-

ferences centred at point i with radius $d + \delta$ and $d - \delta$ or 1 if point j is within the ring (Wiegand and Moloney 2004).

If $f(m_i, m_j) = 0.5(m_i - m_j)^2$ then the $\kappa(d)$ function is called the mark variogram and is denoted by $\gamma(d)$ (Cressie 1993)

$$[5] \qquad \widehat{\gamma}(d) = 0.5 \left\{ \frac{\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq 1}}^{n} [(m_i - m_j)^2 I_{ij}(d + \delta, d - \delta)]}{\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} [I_{ij}(d + \delta, d - \delta)]} \right\}$$

The use of the $\kappa(d)$ function is generalized because $\rho_f^{(2)}(d)$ depends on the spatial correlation of the mark and the spatial clustering of the points, making it more difficult to manage (Beisbart and Kerscher 2000). However, in most cases, it is not only the marks that are of interest, but also the relationship between these marks and the spatial pattern of the points. This relationship may be established by separately analysing the mark correlation function and the spatial pattern of the unmarked point process.

The method proposed in this paper allows us to analyse the tree height correlation and the spatial pattern of the trees using a single cumulative function. Stoyan's $k_{mm}(d)$ and the mark variogram describe the correlation or variance of the mark for those pairs of points separated by a distance d. When we are interested in testing the significance of the deviations with respect to a certain null model, a cumulative function (which is evaluated for all pairs of points, one of these being located within a circle of radius d that is centred on the other) is preferable to the pair correlation function (Stoyan and Penttinen 2000). Two cumulative functions, called $\widehat{K}_m(d)$ and $\widehat{L}_m(d)$, analogue to Ripley's $K(d)$ function and the $\widehat{L}(d)$ function, respectively, are derived from Stoyan's $k_{mm}(d)$ in Dale et al. (2002)

$$[6] \qquad \widehat{K}_m(d) = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{m_i m_j}{\lambda^2 s(d_{ij})}$$

$$[7] \qquad \widehat{L}_m(d) = \sqrt{K_m(d)/\pi \overline{m}^2} - d$$

where λ is the intensity and $s(d_{ij})$ is an edge correction function.

With respect to the spatial distribution, a regular distribution is the opposite of a clustered distribution; the values of the $K(d)$ function being lower for a regular distribution, whereas in the case of the height distribution, similar height values indicate regular structures. Therefore, to define an index in accordance with the concept of regularity in forest science, formulations based on the variance (such as the mark variogram), which increase with height heterogeneity, would be preferable to formulations based on the covariance of the mark (such as Stoyan's $k_{mm}(d)$ function or the $\widehat{L}(d)$ in Dale et al. (2002)), which show high covariance values for similar heights at a given scale. Thus, we propose the following L function estimator, weighted by the normalized mark variance:

$$[8] \quad L_m(d) = \sqrt{\frac{A}{2\pi s^2 n^2} \left[\sum_{\substack{i=1}}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \omega_{ij}(d)(m_i - m_j)^2 \right]} - d$$

$$[9] \quad s^2 = \frac{\sum_{i=1}^{n}(m_i - \overline{m})^2}{n}$$

where A is the plot area; n is the number of stems with DBH > 10 cm; m_i and m_j are the heights of trees i and j, respectively; and ω_{ij} is calculated as described for Ripley's $K(d)$ function. The $L_m(d)$ function has zero expectation under a Poisson point distribution and independent uncorrelated mark distribution. This function can be used to test a given null hypothesis by selecting an appropriate null model and comparing the value of the function for the observed data with its distribution under the model. If the trees are overdispersed or the heights are more homogeneous for the empirical data than for the null model for a certain distance, the empirical value will fall below the lower quantile of the model distribution. When the clustering or the height heterogeneity for the empirical data are significantly greater than for the null model for a certain distance, the empirical value will go above the upper quantile of the model distribution.

Separating the effect of the spatial clustering from the mark correlation

The value of $L_m(d)$ depends on both the spatial clustering of the trees and the height correlation. Thus, this function allows us to simultaneously assess the spatial distribution of the trees and the spatial structure of the tree height by comparing its value with two null models: one representing the null hypothesis of independence between the spatial location of trees and their height and the second representing the null hypothesis of random spatial distribution of trees.

To test the null hypothesis of independence between the spatial location of trees and their height, the heterogeneous random labelling null model was developed. This model involves random labelling while keeping tree location fixed and randomly assigning the heights of the trees (Goreaud and Pélissier 2003), which accounts for large-scale variation in the mark mean. To incorporate the large-scale variation of the mean height into the null model, a moving window (Wiegand and Moloney 2004) was used. The height of each tree was randomly selected from the empirical height distribution for the trees located within a circle of a specified radius (depending on the size of the plot, which in turn determines the allowed distance range for the $L_m(d)$ analysis).

Deviations from random position were tested through the heterogeneous mark mean Poisson model, which consists of spatial randomness accounting for large-scale variation in the mark mean. For this purpose, the coordinates for each tree were randomly assigned, and thereafter, the height of each tree was randomly selected from the empirical height distribution for the trees located within a circle of a specified radius (window radius r).

Simulated marked point processes

To assess the meaningfulness of the height structure characterization using the newly proposed techniques, six simulations of marked point processes with known spatial scales were generated for both the spatial pattern and autocorrelation of the marks.

The simulated stands consist of 100 trees distributed in a 10 m × 10 m plot. The spatial pattern was simulated using the following two processes:

1. Simple sequential inhibition process (Diggle et al. 1976; Cressie 1993): This model was used to simulate small-scale regularity. The X and Y coordinates of each new tree were randomly located within the plot, and the point was retained as long as it was not located within a circle of a determined radius (the so-called distance of repulsion) centred on any other previously established tree.

2. Heterogeneous Poisson process: This model was used to simulate facilitation between the previously established trees and the new establishments. It is based on the Matérn cluster processes (Matérn 1960; Stoyan and Penttinen 2000) and the homogeneous Neyman–Scott processes (Neyman and Scott 1952; Dixon 2002; Wiegand and Moloney 2004). Each new tree is labelled as a "parent" event if a random realization between zero and one is above a given threshold (facilitation probability), or it is labelled as "offspring" if the contrary occurs. In a second step, the spatial location of the tree was determined. If the tree had been labelled as a parent event, its X and Y coordinates were randomly located within the plot. If the tree had been labelled as offspring, the distance from its parent event (if it constituted the first offspring event of the cluster), or from the most recent offspring establishment in the case of subsequent offspring events, follows a uniform distribution from zero to a determined facilitation distance.

The heights for each tree were then randomly generated within a predefined height range, with probability increasing from the extreme values of the range towards a predefined mode, following a linear distribution. To allow for autocorrelaton in the tree height, the probability of maintaining or discarding a generated value was defined by a normal distribution with the mean being equal to the mean height for the trees previously labelled within a specified distance (radius of autocorrelation) and a standard deviation selected for each plot. The height range was set at 0–25 m, with a 10 m mode for all the simulated stands.

Two regular plots were generated using the simple sequential inhibition process with a 0.8 m distance of repulsion between trees: the first with height autocorrelation up to 5 m and a standard deviation around the local mean height of 6 m (plot S1) and the second with autocorrelation up to 2 m and identical standard deviation (plot S2).

In another two plots, a large-scale inhomogeneous pattern was simulated through the heterogeneous Poisson process, with a facilitation probability equal to 0.9 and a 5 m facilitation distance. In the same way as for the regular plots, the tree height was generated for the first plot with height autocorrelation up to 5 m and a standard deviation around the local mean height of 4 m (plot S3) and the second with autocorrelation up to 2 m and identical standard deviation (plot S4).

For the last two plots, a small-scale cluster spatial pattern was simulated, with a facilitation probability equal to 0.9 and a 2 m facilitation distance. The tree height was first generated with an autocorrelation range of 5 m and local standard deviation of 2 m in plot S5 and with autocorrelation up to 2 m and the same standard deviation in plot S6.

The $L_m(d)$ function was calculated for each simulation, as well as the confidence bounds for the heterogeneous random labelling and heterogeneous mark mean Poisson null models, setting the moving window radius of both null models at the height autocorrelation range for the simulation.

Results

Second-moment analysis of marked point processes for the simulated stands

Figure 1 shows the location of the trees and the height distribution in the simulated plots. Figure 2 shows the $L_m(d)$ function analysis, illustrating the interpretation of the empirical function behaviour with respect to the confidence bounds of the null models. As can be seen, the $L_m(d)$ function value depends both on the spatial pattern of the trees and the spatial structure of the mark. The empirical function registers a minimum at 0.8 m for plots S1 and S2 (simple sequential inhibition process) and has significantly higher values than those of the heterogeneous mark mean Poisson null model for distances under 2 m (approximately) for plots S5 and S6 (inhomogeneous with facilitation distance = 2 m). The inclusion of the large-scale trend of the mark in the heterogeneous mark mean Poisson null model leads to a wider confidence interval, especially for larger distances, keeping the empirical function within the quantile bounds for plots S3 and S4 (inhomogeneous at a scale of 5 m). However, for plots S1, S3 and S5, in which height shows a large-scale trend, the empirical function stays inside the heterogeneous random labelling null model bounds at short distances but falls outside the bounds at large scales in plots S1 and S5. This indicates the existence of large areas (with radii of approximately 5 m) where height tends to be homogeneous. On the other hand, for plots S2, S4, and S6, in which height was autocorrelated up to distances of 2 m, the $L_m(d)$ function reflects a significant regular height distribution with respect to the heterogeneous random labelling null model at short distances, i.e., the variability of the height for those distance lags is lower than the variability for the plot, forming homogeneous height patches.

Second-moment analysis for the experimental plots

Figure 3 shows the $L_m(d)$ function calculated for the experimental data from the Navafría plots and the 95% quantiles of the heterogeneous random labelling and the heterogeneous mark mean Poisson null models, with a radius for the window of $r = 25$ m for plot N2 and $r = 35$ m for plots N3, N4, and N5. The $L_m(d)$ empirical function for plots N2 and N3 falls between the 95% lower quantiles of both null models for distances up to 2 and 15 m, respectively. The lower quantiles of both null models are quite similar in plots N4 and N5; significant differences were not found with either null model in the case of plot N4, whereas the height distribution in N5 showed notable regularity up to scales of approximately 15 m. The significant deviations

from the heterogeneous mark mean Poisson model found for the first distance lags in plots N2 and N4 may reflect the effect of the regular spatial distribution of stems at small distances.

In Fig. 4, the empirical $L_m(d)$ function of the plots located in the Valsaín forest is compared with the 95% quantiles of the function for the heterogeneous random labelling and heterogeneous mark mean Poisson models, using $r = 25$ m for both models except for plots V5 and V8, for which $r = 35$ m was used. Plots V2, V3, and V5 show a similar pattern, with the empirical function below the heterogeneous mark mean Poisson null model lower quantile up to distances of around 15, 20, and 25 m, respectively. In plot V6, the empirical function only falls below the heterogeneous mark mean Poisson model lower quantile for the shortest distances, up to 8 m. The $L_m(d)$ function for plot V4 is not significantly different from the heterogeneous mark mean Poisson model for almost any distance nor is it significantly different from the heterogeneous random labelling model for distances up to 20 m.

The empirical function for plot V1 falls below the lower quantiles in both null models. In Fig. 5, the $L_m(d)$ function has been calculated separately for larger trees (DBH > 25 cm) and younger strata trees (10 cm < DBH < 25 cm) in plot V1. As can be appreciated in Fig. 5, each strata displays a very different pattern. The empirical function for the larger trees is not significantly different from that of both null models. In the case of the younger strata, the height distribution is significantly more regular than is the case with the heterogeneous random labelling model for distances <8 m, probably as a result of height homogeneity within groups of young trees, whereas it is significantly more irregular than that found using the heterogeneous mark mean Poisson model for distances between 6 and 16 m, perhaps because of a clustered distribution. The low values given by the $L_m(d)$ function for the two strata together may be due to the spatial segregation of larger and younger trees in even-sized groups, which leads to a strong mark autocorrelation (and hence low variance) when the strata are analysed together. However, when each strata is analysed separately, only those trees belonging to the same cohort are included in the analysis, and the effect of segregation among even-sized groups disappears. Thus, the $L_m(d)$ only reflects the spatial structure of height within the strata.

The empirical function for plot V7 falls below the lower quantile of the heterogeneous mark mean Poisson model for all the analysed distances; however, it only falls below the lower quantile of the heterogeneous random labelling model for distances <25 m (see Fig. 4). In Fig. 6, the L_m function is analysed separately for Scots pine and Pyrenean oak in plot V7. The empirical function calculated for Scots pine falls below the heterogeneous mark mean Poisson model quantile from 5 to 15 m and below the heterogeneous random labelling model quantile for all the analysed range. However, the empirical function calculated for Pyrenean oak is situated within the bounds of the heterogeneous random labelling model, far above the heterogeneous mark mean Poisson model upper quantile bound.

In plot V8, the behaviour of the $L_m(d)$ function with respect to the bounds of both null models is similar to that of the younger strata in plot V1, although the empirical func-

Fig. 1. Spatial distribution and height of the trees (circumferences of the circles are proportional to the height) in the simulated stands. Regular spatial pattern with a large-scale trend (plot S1) (*a*) or with smal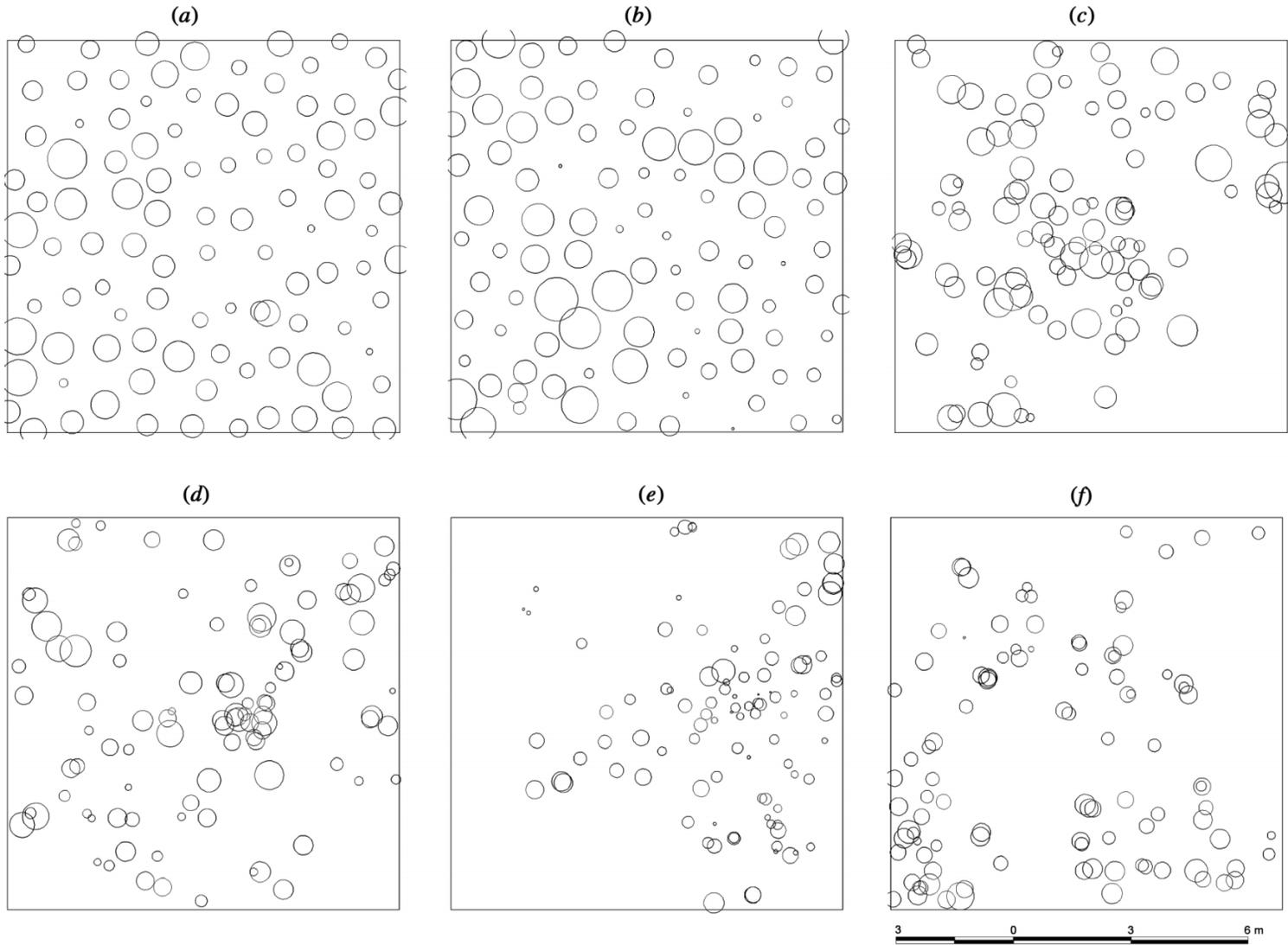l-scale autocorrelation in the mark distribution (plot S2) (*b*), spatial pattern inhomogeneous at large scale with large-scale trend (plot S3) (*c*), or with small-scale autocorrelation in the mark distribution (plot S4) (*d*) and small-scale cluster spatial pattern with large-scale trend (plot S5) (*e*), or with small-scale autocorrelation in the mark distribution (plot S6) (*f*).

Fig. 2. L_m function for the simulated plots data (thick line) and 95% quantiles of the heterogeneous random labelling null model (thin dotted line) and the heterogeneous mark mean Poisson null model with window radius $r = 4.5$ (a, c, and e) and $r = 2$ (b, d, and f) (thin solid line). Figures 2a–2f show the same plots as Figs. 1a–1f, respectively.
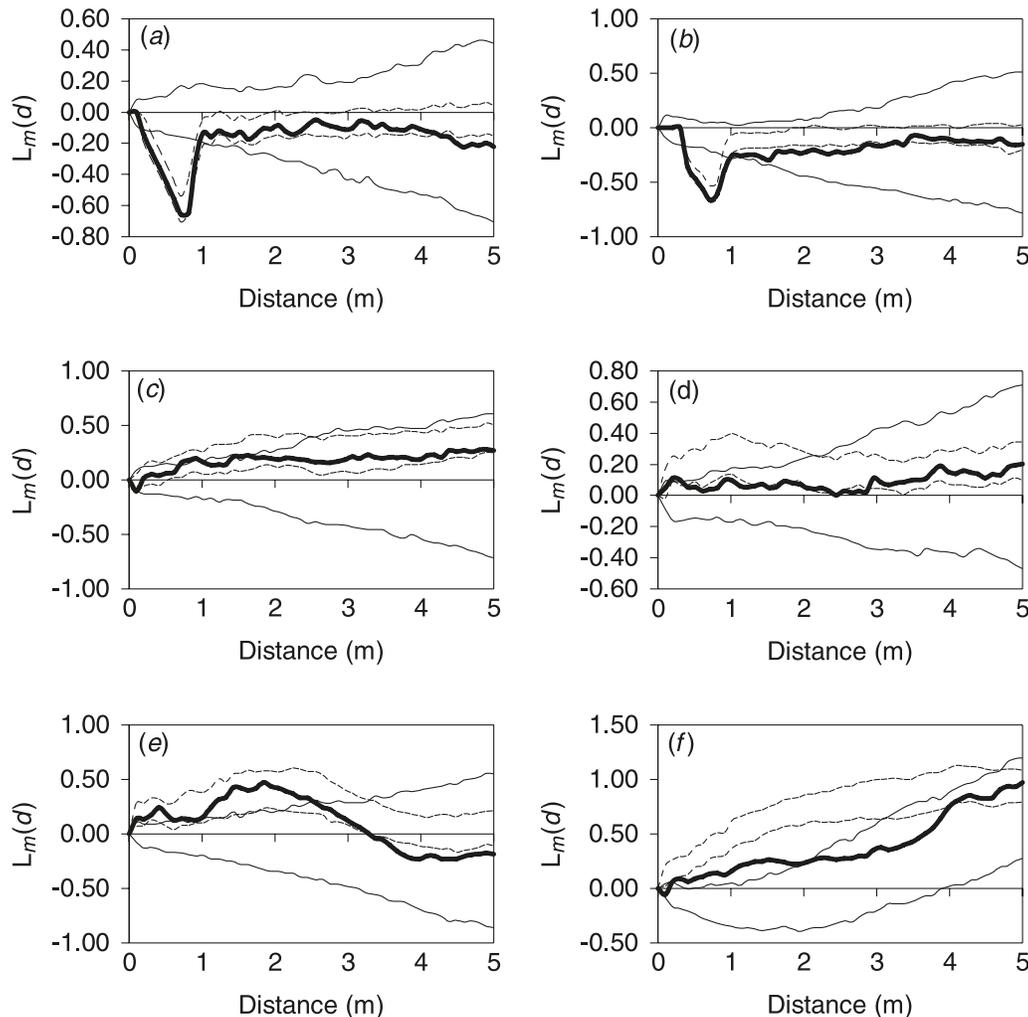

tion falls below the lower quantile of the heterogeneous random labelling model for distances <15 m and the range above the heterogeneous mark mean Poisson model upper quantile extends from 0 to 6 m (see Fig. 4).

Discussion

Marked-point processes constitute powerful tools for forest structure characterization and modelling. Despite this, their use is still relatively infrequent, in part because many of the spatial statistic techniques used in marked-point processes have only recently been developed (e.g., Renshaw and Särkkä 2001; Guan 2006) and, therefore, are not as yet available in accessible statistical packages. Furthermore, a number of practical aspects and problems are currently being addressed (Schlather 2001; Guan et al. 2007) and biologically meaningful models developed (Stoyan and Penttinen 2000; Comas and Mateu 2007). The most commonly used tools for describing the second-order characteristics of marked-point processes, such as the $k_{mm}(d)$ mark correlation function (Stoyan 1984) or the mark variogram (Cressie 1993), are based on the conditional correlation function.

The latter only reflect the correlation of the marks, whereas the effect of the spatial clustering is removed from the second-order measure (Beisbart and Kerscher 2000; Schlather 2001). Nevertheless, the relationship between the spatial pattern and the distribution of the qualitative variables of the trees, such as height, is of great interest for forest structural characterization. An $\widehat{L}_m(d)$ estimator based on the mark covariance was suggested in Dale et al. (2002). The present paper provides a new formulation for the cumulative $L_m(d)$ function weighted by the normalized mark variance, which allows us to analyse the interaction between the variable distribution and the point pattern. The proposed formulation of the $L_m(d)$ function would appear to be better for describing the concept of regularity in forest sciences than formulations based on the covariance of the mark, such as Stoyan's $k_{mm}(d)$ function or the $\widehat{L}_m(d)$ function in Dale et al. (2002). Moreover, for a Poisson process, the expected value of the $L_m(d)$ function will be zero if the variable distribution is independent of the point distribution. This technique may be used to analyse any variable related to the trees, such as growth, fruit production, or pest incidence, providing information about the underlying spatial processes. However, the

Fig. 3. L_m function for the experimental data (thick line) and 95% quantiles of the heterogeneous random labelling null model (thin dotted line) and the random position with moving window for the mark mean null model (thin solid line), both null models with radius for the window $r = 25$ m in plot N2 (*a*), and $r = 35$ m in plots N3 (*b*), N4 (*c*), and N5 (*d*) of Navafría.
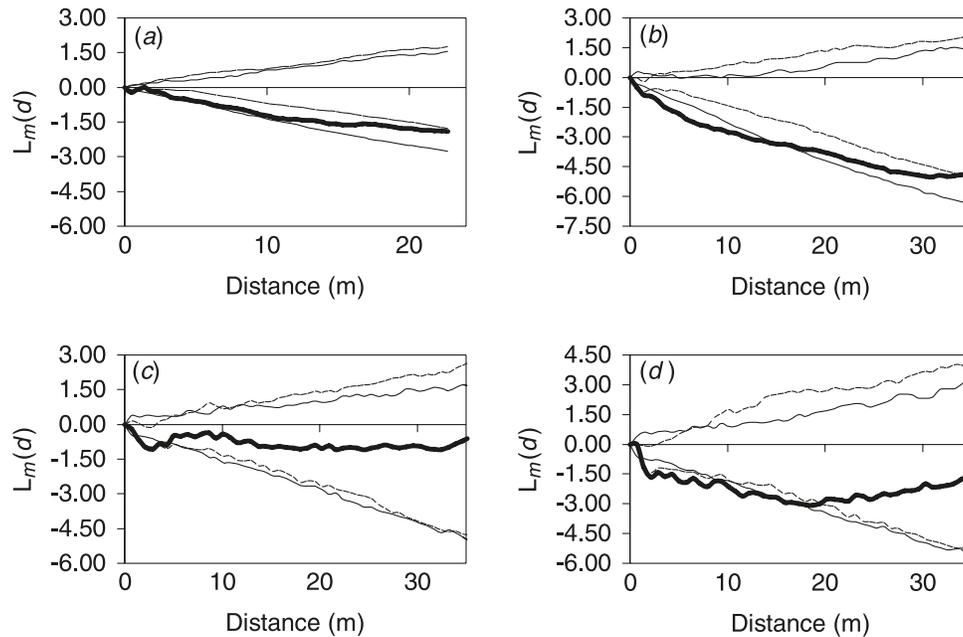

$L_m(d)$ function requires mapped data of large areas, which can be difficult to obtain, especially in dense young stands. This shortcoming may be dealt with using subsampling estimators (Politis and Sherman 2001).

The simulated plot analysis shows that, by comparing the empirical function with the two proposed null models, it is possible to determine which part of the spatial variability in height is due to the pattern and which is due to the height distribution, as well as to relate the scale characteristics of both processes. This information may be completed with a test for independence between marks and points, as proposed by Guan (2006). The moving window seems to be an adequate method for incorporating large-scale height distribution variability into the null model (Wiegand and Moloney 2004; Montes and Cañellas 2007), which in turn allows deviations from the null model to be tested at scales within the scope of the experimental design by removing the noise from scales beyond the range of analysis. In most of the analysed stands, the heterogeneous random-labelling model seems to show much more irregularity across the analysed range of distances than the model underlying the empirical distribution, whereas the heterogeneous mark mean Poisson model seems to be appropriate for large-scale distributions. In the case of the height distribution for Pyrenean oak, the heterogeneous random-labelling model may describe the underlying process (Fig. 6*b*).

The second-moment analysis showed that the height pattern for the analysed Scots pine stands tends towards regularity at short distances, where the trees form patches of relative homogeneous height. All the analysed stands showed significantly greater regularity in height distribution than for the heterogeneous mark mean Poisson model at small scales except for plot V8, located near the upper limit of the altitudinal distribution for Scots pine in the Central Mountain Range of Spain. This plot showed an irregular dis-

tribution because of the low density and frequent disturbances (snow and windthrow) resulting from the mountain climate conditions. In the plot at Valsaín in which the regeneration period is coming to an end (V1), the $L_m(d)$ function indicates that tree height is very evenly distributed at scales under 25 m. The patchiness in the height distribution during the regeneration period is similar to the pattern reported in other studies of coniferous species. For example, in western hemlock (*Tsuga heterophylla* (Raf.) Sarg.) old-growth forests, the intertype $K_{12}(d)$ function (Moeur 1993) and point by point associations between seedling density and competition indices (Moeur 1997) show a pattern of seedling cohort clusters associated with canopy gaps. Similarly, using the $K_{rx}(d)$ function to assess the spatial relationship between the pattern of the mature strata trees and the sapling density, Montes and Cañellas (2007) reported a cluster structure for Scots pine saplings that showed a negative spatial relationship with the mature strata trees, and similar results were obtained for Norway spruce (*Picea abies* (L.) Karst.) saplings in *P. abies* and silver fir (*Abies alba* Mill.) mixed stands (Grassi et al. 2004). This height homogeneity is favoured by the competition and, so, was absent in plot V8, where density declined, and it was less noticeable in thinned mature stages (plot N4) and where fellings had been carried out (plots N5 and V6). In the $L_m(d)$ analysis for plot V1, it can be seen that, when the young cohorts develop in small gaps, the asymmetric competition from the larger tree strata leads to a higher structural heterogeneity in the young cohort. The height heterogeneity added to the cluster pattern at short scales in plot V8, similar to that found in other studies at the altitudinal tree line (Camarero et al. 2000), may be due to the smothering effect provided by the windward neighbours. This effect allays the extreme, high-altitude conditions within the cluster and results in "wave regeneration" (Satō and Iwasa 1993), which seems to favour regeneration

Fig. 4. L_m function for the experimental data (thick line) and 95% quantiles of the heterogeneous random labelling null model (thin dotted line) and the heterogeneous mark mean Poisson null model (thin solid line), both null models with a radius for the window of $r = 25$ m for plots V1 (*a*), V2 (*b*), V3 (*c*), V4 (*d*), V6 (*f*), and V8 (*h*) and $r = 35$ m for plots V5 (*e*) and V7 (*g*) of Valsaín.
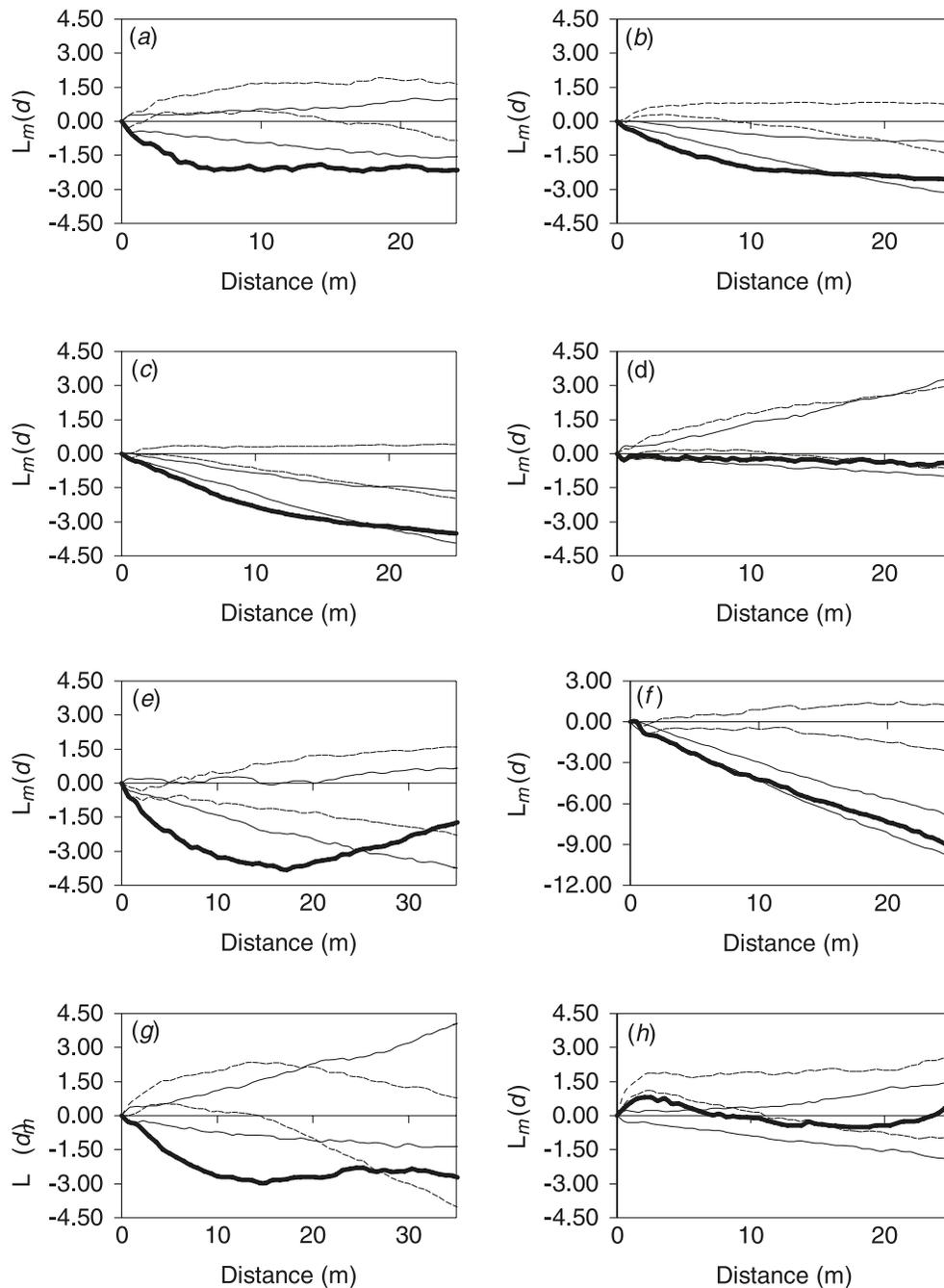
performance in the vicinity of the established tree groups. Figure 6 shows the different vertical development of Scots pine and Pyrenean oak in the ecotone between the pinewood and the broad-leaved forest. Other studies have reported a "trade-off" between competition and dispersal where competing species coexist (Skellam 1951; Pacala and Levin 1997; Amarasekare and Nisbet 2001; Murrell and Law 2003). The $L_m(d)$ function reveals differences in the vertical structure of the two species: whereas the Scots pine shows a high level of height homogeneity over the analysed range with a minimum of around 10 m for both null models, the

Pyrenean oak, which mostly regenerates by sprouting, did not differ significantly from the heterogeneous random-labelling model, perhaps because of the half-shade tolerant behaviour of this species and the effect of the coppice structure on the intertree competition.

In conclusion, the second-moment analysis methods proposed provide a useful tool for the height structure characterization when mapped data are available. The two null models proposed allow us to identify the type of structure described by the empirical $L_m(d)$ function. However, further research into new null models that explain the joint structure
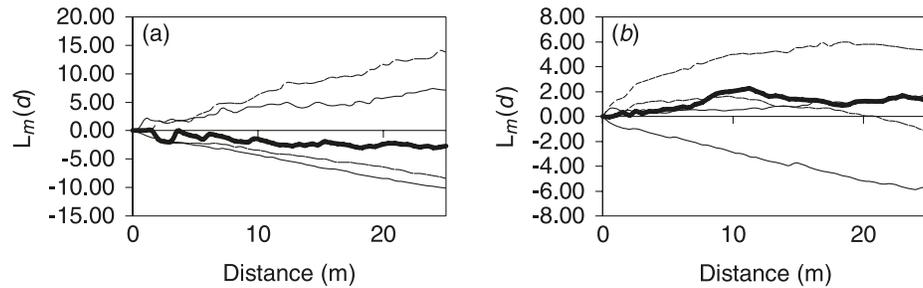
Fig. 5. L_m function for the experimental data (thick line) and 95% quantiles of the heterogeneous random labelling null model(thin dotted line) and the heterogeneous mark mean Poisson null model (thin solid line), both null models with a radius for the window of $r = 25$ m, for the cohort of trees with DBH > 25 cm (*a*) and the cohort of trees with 10 cm ≤ DBH < 25 cm (*b*) in plot V1.

Fig. 6. L_m function for the experimental data (thick line) and 95% quantiles of the heterogeneous random labelling null model (thin dotted line) and the heterogeneous mark mean Poisson null model(thin solid line), both null models with a radius for the window of $r = 35$ m, for Scots pine *Pinus sylvestris* (*a*) and for Pyrenean oak *Quercus pyrenaica* (*b*) in plot V7.
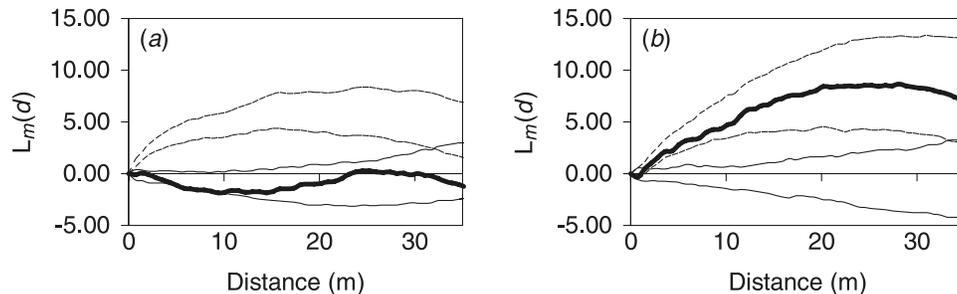
of two cohorts or two species may enhance our understanding of the underlying forest structure processes.

Acknowledgments

References

Aguirre, O., Hui, G., von Gadow, K., and Jiménez, J. 2003. An analysis of spatial forest structure using neighbourhood-based variables. For. Ecol. Manage. **183**: 137–145. doi:10.1016/S0378-1127(03)00102-6.

Amarasekare, P., and Nisbet, R.M. 2001. Spatial heterogeneity, source–sink dynamics, and the local coexistence of competing species. Am. Nat. **158**: 572–584. doi:10.1086/323586.

Ballani, F., Daley, D.J., and Stoyan, D. 2006. Modelling the microstructure of concrete with spherical grains. Comput. Mater. Sci. **35**: 399–407. doi:10.1016/j.commatsci.2005.03.005.

Beisbart, C., and Kerscher, M. 2000. Luminosity- and morphology-dependent clustering of galaxies. Astrophys. J. **545**: 6–25. doi:10.1086/317788.

Camarero, J.J., Gutiérrez, E., and Fortin, M.J. 2000. Spatial pattern of subalpine forest – alpine grassland ecotones in the Spanish Central Pyrenees. For. Ecol. Manage. **134**: 1–6. doi:10.1016/S0378-1127(99)00241-8.

Comas, C., and Mateu, J. 2007. Modelling forest dynamics: a perspective from point process methods. Biom. J. **49**: 176–196. doi:10.1002/bimj.200510268.

Cressie, N.A.C. 1993. Statistics for spatial data. John Wiley and Sons Inc., New York.

Dale, M.R.T., Dixon, P., Fortin, M.J., Legendre, P., Myers, D.E., and Rosenberg, M.S. 2002. Conceptual and mathematical relationships among methods for spatial analysis. Ecography, **25**: 558–577. doi:10.1034/j.1600-0587.2002.250506.x.

del Río, M., Montes, F., Cañellas, I., and Montero, G. 2003. Revisión: Índices de diversidad estructural en masas forestales. Invest. Agrar. Sist. Recursos For., **12**: 159–176.

Diggle, P.J., Besag, J.E., and Gleaves, J.T. 1976. Statistical analysis of spatial point patterns by means of distance methods. Biometrics, **32**: 659–667. doi:10.2307/2529754.

Dixon, P.M. 2002. Ripley's *K* function. *In* Encyclopedia of environmetrics. Vol. 3. *Edited by* A.H. El-Shaarawi and W.W. Piegorsch. John Wiley and Sons, Ltd., Chichester, UK. pp. 1796–1803.

Ferris-Kaan, R., Peace, A.J., and Humphrey, J.W. 1998. Assessing structural diversity in managed forests. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Gavrikov, V.L., and Stoyan, D. 1995. The use of marked point processes in ecological and environmental forest studies. Environ. Ecol. Stat. **2**: 331–344. doi:10.1007/BF00569362.

Goreaud, F., and Pélissier, R. 1999. On explicit formulas of edge correction for Ripley's *K*-function. J. Veg. Sci. **10**: 433–438. doi:10.2307/3237072.

Goreaud, F., and Pélissier, R. 2003. Avoiding misinterpretation of biotic interactions with the intertype K_{12}-function: population independence vs. random labelling hypothesis. J. Veg. Sci. **14**: 681–692. doi:10.1658/1100-9233(2003)014[0681:AMOBIW]2.0.CO;2.

Grassi, G., Minotta, G., Tonon, G., and Bagnaresi, U. 2004. Dynamics of Norway spruce and silver fir natural regeneration in a mixed stand under uneven-aged management. Can. J. For. Res. **34**: 141–149. doi:10.1139/x03-197.

Guan, Y. 2006. Tests for independence between marks and points of a marked point process. Biometrics, **62**: 126–134. doi:10.1111/j.1541-0420.2005.00395.x. PMID:16542238.

Guan, Y., Sherman, M., and Calvin, J.A. 2007. On asymptotic properties of the mark variogram estimator of a marked point process. J. Stat. Plann. Inference, **137**: 148–161. doi:10.1016/j.jspi.2005.10.004.

Hirsch, D.M., Ketcham, R.A., and Carlson, W.D. 2000. An evaluation of spatial correlation functions in textural analysis of metamorphic rocks. Geol. Mater. Res. **2**(3): 1–42.

Kelly, D.L., and Connolly, A. 2000. A review of the plant communities associated with Scots pine (*Pinus sylvestris* L.) in Europe, and an evaluation of putative indicator/specialist species. Invest. Agrar. Sist. Recurosos For. **9**(Suppl. 1): 15–39.

Kint, V. 2003. Structural development in ageing Scots pine (*Pinus sylvestris* L.) stands in western Europe. Ph.D. thesis, Faculteit Landbouwkundige en Toegepaste Biologische Wetenschappen, Melle-Gontrode, Belgium.

Kuuluvainen, T., Penttinen, A., Leionen, L., and Nygren, M. 1996. Statistical opportunities for comparing stand structural heterogeneity in managed and primeval forests: an example from boreal spruce forest in southern Finland. Silva Fenn. **30**: 315–328.

MacArthur, R.M., and MacArthur, J.W. 1961. On bird species diversity. Ecology, **42**: 594–598. doi:10.2307/1932254.

Mason, W.L., and Alía Miranda, R. 2000. Current and future status of Scots pine (*Pinus sylvestris* L.) forest in Europe. Invest. Agrar. Sist. Recurosos For. **9**(Suppl. 1): 317–335.

Matérn, B. 1960. Spatial variation. Meddelanden fran Statens Skogsforskningsinstitut 49(5). [2nd ed. 1986.] Springer-Verlag, New York. Lect. Notes Stat. 36.

Moeur, M. 1993. Characterizing spatial patterns of trees using stem-mapped data. For. Sci. **39**: 756–775.

Moeur, M. 1997. Spatial models of competition and gap dynamics in old-growth *Tsuga heterophylla* / *Thuja plicata* forest. For. Ecol. Manage. **94**: 175–186. doi:10.1016/S0378-1127(96)03976-X.

Montes, F., and Cañellas, I. 2007. The spatial relationship between post-crop remaining trees and the establishment of saplings in *Pinus sylvestris* stands in Spain. Appl. Veg. Sci. **10**(2): 151–160.

Montes, F., Sánchez, M., del Río, M., and Cañellas, I. 2005. Using historic management records to characterize the effects of management on the structural diversity of forests. For. Ecol. Manage. **207**: 279–293. doi:10.1016/j.foreco.2004.10.031.

Murrell, D.J., and Law, R. 2003. Heteromyopia and the spatial coexistence of similar competitors. Ecol. Lett. **6**: 48–59. doi:10.1046/j.1461-0248.2003.00397.x.

Neyman, J., and Scott, E.L. 1952. A theory of the spatial distribution of galaxies. Astrophys. J. **116**: 144–163. doi:10.1086/145599.

Pacala, S.W., and Levin, S.A. 1997. Biologically generated spatial pattern and the coexistence of competing species. *In* Spatial ecology: the role of space in population dynamics and interspecific interactions. *Edited by* D. Tilman and P. Kareiva. Princeton University Press, Princeton, N.J. pp. 204–232.

Penttinen, A.K., Stoyan, D., and Henttonen, H.M. 1992. Marked point processes in forest statistics. For. Sci. **38**: 806–824.

Politis, D.N., and Sherman, M. 2001. Moment estimation for statistics from marked point processes. J. R. Stat. Soc. Ser. B Stat. Methodol. **63**: 261–275. doi:10.1111/1467-9868.00284.

Renshaw, E., and Särkkä, A. 2001. Gibbs point processes for studying the development of spatial-temporal stochastic processes. Comput. Stat. Data Anal. **36**: 85–105. doi:10.1016/S0167-9473(00)00028-1.

Ripley, B.D. 1977. Modelling spatial patterns (with discussion). J. R. Stat. Soc. Ser. A, **39**: 172–212.

Satō, K., and Iwasa, Y. 1993. Modelling of wave regeneration in subalpine Abies forests, population dynamics with spatial structure. Ecology, **75**: 1538–1550. doi:10.2307/1940081.

Schlather, M. 2001. On the second-order characteristics of marked point processes. Bernoulli, **7**: 99–117. doi:10.2307/3318604.

Skellam, J.G. 1951. Random dispersal in theoretical populations. Biometrika, **38**: 196–218. PMID:14848123.

Staudhammer, C.L., and LeMay, V.M. 2001. Introduction and evaluation of possible indices of stand structural diversity. Can. J. For. Res. **31**: 1105–1115. doi:10.1139/x01-033.

Stoyan, D. 1984. On correlations of marked point processes. Math. Nachrichten, **116**: 197–207. doi:10.1002/mana.19841160115.

Stoyan, D., and Penttinen, A. 2000. Recent applications of point process methods in forestry statistics. Stat. Sci. **15**: 61–78. doi:10.1214/ss/1009212674.

Stoyan, D., and Stoyan, H. 1994. Fractals, random shapes and point fields. John Wiley and Sons, Ltd., Chichester, UK.

Varga, P., Chen, H.Y.H., and Klinka, K. 2005. Tree-size diversity between single- and mixed-species stands in three forest types in western Canada. Can. J. For. Res. **35**: 593–601. doi:10.1139/x04-193.

von Gadow, K. 1993. Zur Bestandesbeschreibung in der Forsteinrichtung. Forst Holz, **21**: 601–606.

Wälder, O., and Stoyan, D. 1996. On variograms in point process statistics. Biom. J. **38**: 895–905. doi:10.1002/bimj.4710380802.

Wiegand, T., and Moloney, K.A. 2004. Ring, circles, and null-models in point pattern analysis in ecology. Oikos, **104**: 209–229. doi:10.1111/j.0030-1299.2004.12497.x.

Zenner, E., and Hibbs, D. 2000. A new method for modelling the heterogeneity of a forest structure. For. Ecol. Manage. **129**: 75–87. doi:10.1016/S0378-1127(99)00140-1.