Simulating spatial distributions of forest trees by using data from fixed area plots

G. STAMATELLOS* AND G. PANOURGIAS

School of Forestry and Natural Environment, University Forest Administration and Management Fund, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

* Corresponding author. E-mail: stamatel@for.auth.gr

Summary

Cost is a critical factor for managing ecosystems. Common forest inventories are usually carried out at regular time intervals (e.g. every 10 years) and are the basis for management planning. This study investigated the potential of utilizing the data of common forest inventories for simulating the spatial distribution of forest tree locations. Fixed area plot sample data were taken from the University Forest of Pertouli, which is an uneven-aged *Abies borisii regis* Matf. forest. Using as criterion the index of dispersion, the tree spatial distribution was characterized as aggregated. The Neyman Type A distribution, a typical distribution of aggregated populations, was a good fit to the data, while tree locations simulation was based on the Poisson cluster process. The simulations resulting from the application of fixed area plot sampling do not incorporate information about the real distances between the trees, but they can describe adequately the spatial patterns of their locations in two-dimensional space and seem to be useful tools for managers of forest ecosystems. For similar populations the detection of their aggregation seems not to be affected by a considerable decrease of the number (up to 36) and the size (up to 125 m²) of sample fixed-area plots. This method is cost effective and its use, potentially in combination with other methods, could be further investigated for its advantages.

Introduction

In a forest ecosystem, all tree species play a central role highlighted by their variation in size and spatial distribution (Lähde *et al.*, 1999). The relative locations of trees within a forest ecosystem can be illustrated graphically by a point configuration formed by the dimensionless trees on a horizontal plane, which is called the spatial pattern of trees (Tomppo, 1986). The estimated spatial trees' distributions provide indications of

the underlying biological processes and of the ecosystem management techniques. The spatial distribution of trees is an important characteristic of forest stands reflecting their history (Moeur, 1993), but also contributing to the formulation of a more effective future management approach. More specifically, the spatial distribution affects (1) the sampling design of forest inventory, (2) timber production, and (3) the need for silvicultural treatments in a stand (Tomppo, 1986).

From the ecological point of view the spatial pattern of plants has important effects on the interactions between plants, between plants and other organisms such as herbivores, and between other organisms such as herbivores and their predators (Dale, 1999). Spatial pattern is a critical determinant of habitat structure (Dale, 1999; Pommerening, 2002; Aldrich *et al.*, 2003).

An ideal statistical summary of a forest would be a complete description of the spatial distribution of its trees and the distribution of tree sizes (Schreuder *et al.*, 1993). The diameter distributions provide, in a simple and efficient manner, the size distribution of trees, while a survey of the spatial distribution of trees provides: (1) definition of the stand density per unit area, and (2) identification of the kind of spatial model (random, clustered or regular) (Pielou, 1960). At a next stage, a more detailed description of the pattern can be obtained by identifying different scales of patterns or by formulating an explicit model of the underlying process describing individual tree locations (Diggle, 1983).

Two methods of sampling are primarily used in surveys of the spatial distribution of tree locations: fixed-area plot (hereafter referred as 'plot') and distance sampling. The first is generally easy to apply in the forest and it may be reliable, but the detection of non-randomness may vary depending on the size of the sample plot. With distance sampling, the basic sampling unit is a point, and information is recorded in the form of distances to neighbouring events (Diggle, 1983). This sampling method provides more detailed information regarding the distribution of trees, but is expensive, difficult to implement, and is biased. Distances between points and their nearest tree neighbours can be very different and can exceed different ecological scales.

Cost is a crucial factor for managing forest ecosystems, e.g. in evaluating forest structure (Gadow and Pogoda, 2000) and in assessing and monitoring biodiversity (Albert and Gadow, 1998). On the other hand, inventory data must be collected usually by the plot sampling method to compile management plans (common forest inventory) and could also be used in forest management operations, thus reducing the cost. Therefore, it could be of interest to investigate the potential of utilizing these data for studying the spatial distribution of trees.

Lewandowski and Gadow (1997) present a method based on the spatial distribution of neighbour-attributes for reproducing the structure of an heterogeneous mixed forest. Spatial stand structures were simulated by Pretzsch (1997) who used the aggregation index of Clark and Evans (1954) and the segregation index of Pielou (1977) to describe the structure of mixed beech-larch stands. A stand structure generator was designed for the modelling and reproduction of spatial stand structures. The pattern produced was the result of a combination of an inhomogeneous Poisson process and a hard-core process. Letcher et al. (1998) developed a spatially explicit simulation model for Picoides borearis, an endangered bird endemic to the south-eastern United States. Model results indicate the importance of considering the spatial distribution of territories in management plans. A non-simple sequential inhibition process of reconstructing the spatial pattern of trees is introduced by Hanus et al. (1998). Zenner and Hibbs (2000) modelled the forest structure (a three-dimensional phenomenon with horizontal and vertical components) by a method based on the concept of spatial tessellation of tree positions, where point patterns are converted into two-dimensional nearest neighbour triangles. Point processes and marked point processes are powerful tools for solving many forestry problems (Penttinen et al., 1992; Gavrikov and Stoyan, 1995; Stoyan and Penttinen, 2000). In the former the points are tree positions. In the latter the marks are qualitative or quantitative tree characteristics. Pommerening (2002) presents equation indices such as the Shannon index, the aggregation index of Clark and Evans, the contagion index, the coefficient of segregation of Pielou, the mingling index, the diameter differentiation index, the pair correlation and the mark correlation function for the quantification of stand structure. These equations have the advantage that they are easier to survey by means of common forest inventories and can be easily combined with the sampling method 'structural group of four' or traditional circular sample plots. A new method to reproduce spatial distributions of trees is introduced by Hui et al. (2003). This method is based solely on information about the uniform angle index and stand density which are easy to assess.

The purpose of the present study is the simulation of forest-tree locations by using sample data from fixed-area plots. A secondary aim is to investigate the effect of the number and size of sample plots on the detection of tree spatial pattern. In the next section, the mathematical frame of spatial distributions of trees is given and the simulation process is described.

Methods and data

We consider a forest area $A \subset R^2$ and a finite population N of forest trees whose centres are at the points $y_i \in A \subset R^2$. The precise locations of trees may be regarded as a realization of a spatial process of points in two-dimensional Euclidean space. By applying plot sampling the area is divided into many small plots (quadrates). Next, a sample selected systematically or occasionally at random of n plots of size α is taken and the number of the trees $N(\alpha) = X$ in each plot (count data), is recorded.

The density D, as the number of trees per unit area is:

$$D = \frac{1}{\alpha n} \sum_{i=1}^{n} X_{i} \quad \text{with variance}$$
$$\text{var}(D) = \frac{1}{\alpha^{2} n} \left[\sum_{i=1}^{n} \left(X_{i} - \overline{X} \right)^{2} \right] / (n-1)$$

Presence *f*, which is defined as the proportion of plots in which there is at least one tree, is:

$$f = \frac{1}{n} \sum_{i=1}^{n} z_i$$
 with variance

$$var(f) = f(1-f)/n$$

where z_i is 0 or 1 depending on the absence or presence of at least one tree, respectively. A population is said to be distributed at random if the relative location of an individual object is independent of the location of any other object within an infinitely large area (Reich and Davis, 1998). Under the null hypothesis of random spatial pattern trees, X has a Poisson distribution and var(X) = E(X). For alternative hypotheses involving clumps of trees we would expect var(X) > E(X), while for more regularly spaced trees var(X) < E(X) (Perry and Mead, 1979). The

index of dispersion (ID), which is defined as the ratio of the sample variance to the sample mean of X (Reich and Davis, 1998), is commonly used to characterize the spatial pattern of a population. Under the null hypothesis, ID has a chisquare distribution with (n-1) degrees of freedom (Hoel, 1943). This information can be used to construct the confidence limits for ID. If the null hypothesis H_0 of complete spatial randomness is rejected, then the next step is to fit an alternative model to the data, and to revise it if necessary.

The study was conducted in the management unit 'Vathy' of 192 25 ha at the Pertouli University Forest. The area is almost entirely covered with *Abies borisii regis* Matf. The number of stems is 296 ha⁻¹, the basal area 20.5 m² ha⁻¹, the volume 177.5 m³ ha⁻¹ and the diameter distribution is exponential (inverse J-distribution). The structure of stands is uneven-aged with two main strata (University Forest Service, 1978). The management of the forest is considered to follow a selective individual tree silviculture control method (Matis, 2000).

A random sample of 229 plots of 0.05 ha $(20 \times 25 \text{ m}^2)$ was taken and the trees with diameters between 13 and 87 cm at 1.3 m height in each plot were recorded. The data are not part of the common forest inventory but were taken for other research (Stamatellos, 1991). D, f and ID as they have been previously defined, were estimated and the null hypothesis of complete spatial randomness was tested. The following four distributions were fitted to the data: Neyman Type A, Thomas, Poisson and Negative Binomial. The Neyman Type A and Thomas distribution have been widely used by plant ecologists. Both distributions have two parameters which provide measures of the cluster density and average number of individuals per cluster (here the cluster is a set of plants). There are two assumptions associated with these distributions. The first is that the clusters are randomly distributed and second that the number of plants per cluster follows a Poisson distribution. The distinction between these two distributions is that Neyman's distribution models only associate the offspring (new generation) with an individual cluster, while the Thomas distribution models both the parent (mother tree) and offspring. The probability

function for the Neyman Type A distribution is given by Reich and Davis (1998):

$$P(X = 0) = \exp\{-m_1[1 - \exp(-m_2)]\}$$

and

$$P(X = k+1) = \frac{m_1 m_2 e^{-m_2}}{k+1} \sum_{t=0}^{k} \frac{m_2^t}{t!} P(X = k-t),$$

k = 0,1,...

where m_1 is the number of clusters per unit area, m_2 is the number of individuals per cluster and k, t have the same meaning as the X. The Thomas distribution is given by:

$$P(X=0) = e^{-m}$$

and

$$P(X = k) = \sum_{t=1}^{k} \frac{m^{t}}{t!} \frac{(t\lambda)^{k-t} e^{-t\lambda}}{(k-t)!}, k = 1, 2, ...$$

where m is the number of clusters per unit area, λ is the number of offspring and $(1 + \lambda)$ is the number of individuals (including the 'parent') per cluster. Parameter estimates are obtained using maximum likelihood (Johnson and Kotz, 1969).

The best-fitted model was identified and its parameters were estimated. The Thomas distribution is similar to Neyman's Type A distribution in terms of the probabilities generated. These distributions result, in the case of clustering, from a Poisson cluster process (Anderson, 1992). This process consists of the following components (Cressie, 1993; Reich and Davis, 1998): (1) parent process is Poisson with intensity c; (2) each parent produces a random number s of 'offspring'; and (3) positions of offspring relative to parent's location are independent and uniformly distributed with probability density function h(x, y). Based on these components the simulation is implemented in two steps: (1) let z be the total number of desired parents and offspring - Poissondistributed parents are generated with mean number of clusters equal to c; and (2) if z_1 is the number of parents $(z_1 < z)$, the remaining offspring $(z - z_1)$ can be assigned, with probability p (usually p = 1), to a randomly selected parent. Location of offspring relative to parent's position has distribution h(x, y), where x and y are the coordinates. Offspring are considered orphaned with a probability of 1 - p and are assigned to a random location in the region. Continue until all $z - z_1$ offspring are allocated. In the program agg (z, c, r, p) in S-plus code (Reich and Davis, 1998), the h(x, y) is equal to the uniform density function and allocates the offspring on the inside of a circle of radius r. In the simulation, we used z =282; c = 66; and r = 4, 5, 6 and 7 in four different model runs. In other words, 282 trees were distributed in an area of $100 \times 100 \text{ m}^2$ in 66 clusters of radius from 4 to 7 m. The z, c and r = 7 m resulted from the data of the plot sampling and the Neyman Type A distribution. Lastly, two similar clustered populations were considered and the effect of the number and the size of sampling plots on the ID were investigated. The simulation experiment was programmed in S-plus (Becker et al., 1988; Statistical Sciences, 1999).

Results and discussion

The species presence was estimated to be 0.9913 (~99 per cent of sample plots had at least one tree) and the density of the number of trees per unit area 14 0873 with standard errors of 0.0061 and 0.5699, respectively. The estimation of ID = 5.2804 > 1 was outside the estimated confidence interval (0.8249-1.1917). Consequently, the null hypothesis of the complete spatial randomness of trees was rejected and the spatial pattern was characterized as aggregated. In similar populations, the trees are found in clusters of various densities and sizes (Stamatellos et al., 2001). Some plants reproduce vegetatively so that offspring tend to cluster around parent plants. Plants that reproduce by seed may show a degree of aggregation, depending on the radius of seed dispersal (Thompson, 1958). In uneven-aged Siberian fir forest the conditions for germination and subsequent growth of seedling are very heterogeneous under the canopy. This explains why the smallest plants grow in clumps (Gavrikov and Stoyan, 1995). Site conditions, on the other hand, influence the adaptability and relative abundance of species on a given site. For example, the tree clumps' formation in a fir forest was especially associated with extreme climatic conditions (Colak, 2003). In this case, the aggregation could be explained by the manner of fir regeneration

and the applied system of silvicultural treatments (Zagas and Smiris, 1993).

Figure 1 shows the fit of distributions: Poisson, Thomas (p-value = 0.0811), Neyman Type A (p-value = 0.1405) and Negative Binomial distribution (p-value = 0.0341) for the count data. Poisson distribution was fitted for comparison. As it was expected, only the Thomas and Neyman Type A distributed adequately described the count data (a > 0.05). The $\chi^2_{\nu-1}$ test of goodnessof-fit (with 19 d.f.) has shown that the Neyman Type A distribution is a better fit to the data. This distribution is more effective for quadrate counts of plant populations (Anderson, 1992). For model constructions that simulate tree crowns as discrete geometrical objects, the aggregated distribution of tree locations was simulated using the Neyman Type A distribution (Leblanc and Chen, 1997). Parameter estimates of the Neyman distribution were $\hat{m}_1 = 3.29$ and $\hat{m}_2 = 4.28$. Four simulations were generated by a Poisson cluster process with different clustering radii of trees. These are illustrated in Figure 2. The clustering degree increases when the clustering radius decreases resulting in more extensive voids.

Although plot sampling provides inexpensive information, which can be used for characterizing the spatial pattern of trees, it has the disadvantage of being dependent on the size of the sample plot. This is a concern in relation to the size of clusters in the aggregated populations. Aggregation would be detected when the size of sample plot is approximately the same as the average size of clusters (clumps; Elliott, 1971). Therefore, the effect of sampling plot size and sample size upon the estimate of ID was indicatively studied. Two spatial tree variations from the same Neyman Type A distribution were simulated, which were only different as to the clustering degree (r = 7 m and r = 5 m). The trees were allocated to an area of 300 m × 300 m, which was considered as the population. The sampling was conducted by square-shaped plots of an area of 250 m² and 125 m² and was repeated 100 times. The results showed that in all cases ID > 1 (Table 1). The estimates of ID in all 100 replications were found to be outside the confidence intervals of ID = 1. Consequently, detection of the aggregated tree population in such cases is possible. The two sizes of the sampling plots have not affected the

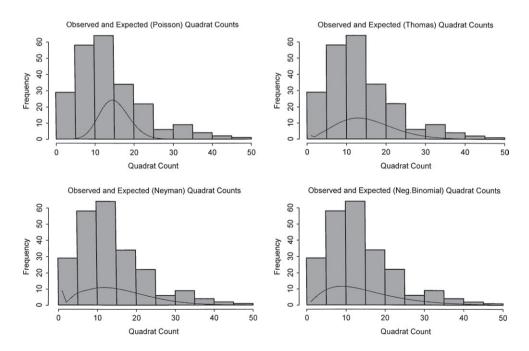


Figure 1. Observed and expected frequency distribution from sampling an aggregated spatial pattern.

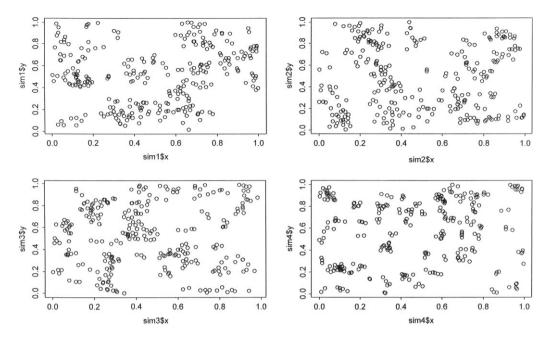


Figure 2. Simulating 1-ha forest (100 m \times 100 m) with four clustering radii of trees locations (sim1 (r = 7 m), sim2 (r = 6 m), sim3 (r = 5 m) and sim4 (r = 4 m)).

Table 1: The index of dispersion (ID) relevant to the number and size of sampling plots in two spatial layouts of trees

Aggregated population	No. of plots	Size of plots (m ²)	Density (trees per plot)	Dispersion index	<i>p</i> -value for ID
Spatial layout 1	36	250	6.911	4.192	0.0026
$(r = 7 \text{ m})^*$	72	125	3.146	2.623	0.0020
	36	125	3.115	2.311	0.0021
Spatial layout 2	36	250	6.484	6.386	0.0108
$(r = 5 \text{ m})^*$	72	125	3.309	4.621	0.0032
	36	125	3.067	4.190	0.0019

The results are based on 100 simulations.

characterization of the spatial pattern of trees, although the IDs have been decreased by almost 35 per cent. The decrease of the number of sampling plots to half (from 72 to 36) in the plots of 125 m² has not affected the detection of the aggregated tree population.

Simple inhibition processes provide a useful class of models for modelling regular point patterns. The non-simple sequential inhibition process proposed by Hanus *et al.* (1998) has the

potential to simulate stands of very different spatial patterns and can be applied during a common forest inventory; however, its application presupposes the availability of equations to predict the maximum crown width of an open-growth tree.

The method for the reproduction of tree locations that was introduced by Hui et al. (2003) does not require measurement of distances between neighbouring trees. This is also an advantage of the plot sampling methods used in

^{*}r = the radius of circles inside which the cluster trees are located.

the present research. Because in common forest practice measured distances are expensive and rare (Hui *et al.*, 2003), these methods are both cost effective. A combination of the method that was introduced by Hui *et al.* (2003) with the plot sampling applied in this research could be further investigated for its advantages. In such a case the stand density would be estimated by using plot sampling.

Conclusions

Assuming that the locations of trees are generated by some underlying random mechanism, it is possible to simulate an average situation of forest tree spatial distribution. The simulations resulting after the application of plot sampling do not incorporate information about the real distances between the trees, but can describe adequately the spatial patterns of their locations in twodimensional space and seem to be useful tools for the managers of forest ecosystems. For similar spatial distributions the detection of the aggregated tree population seems not to be affected by a notable decrease in the number and the size of sampling plots. A combination of plot sampling with the method proposed by Hui et al. (2003) for the reproduction of tree distribution could be further investigated regarding its advantages for practical use.

Acknowledgements

The authors very much thank Dr Robin Reich and the anonymous referee who, with their constructive suggestions and comments on the initial manuscript, have greatly improved this paper. Special thanks are due to Dr Helen McKay for editorial assistance.

References

- Albert, M. and Gadow, K.V. 1998 Assessing biodiversity with new neighborhood-based parameters. In *Data Management and Modelling Using Remote Sensing and GIS for Tropical Forest Land Inventory*. Y. Iaumonier, B. King, C. Legg and K. Rennolls (eds). Proceedings of the FIMP-INTAG International Conference, October 26–29, 1998, Jakarta, Indonesia, pp. 433–445.
- Aldrich, P.R., Parcer, G.R., Ward, J.S. and Michrer, C.H. 2003 Spatial dispersion of trees in an old-growth

- temperate hardwood forest over 60 years of succession. For. Ecol. Manage. 180, 475–491.
- Anderson, J.C. 1992 Determination of dispersion patterns in spatial point process. Ph.D. dissertation, Colorado State University.
- Becker, R.A., Chambers, J.M. and Wilks, A.R. 1988 The New S Language. Wadsworth and Brooks/Coke, Pacific Grove, CA.
- Clark, P.J. and Evans, F.C. 1954. Distance to nearest neighbour as a measure of spatial relationships in populations. *Ecology* 35, 445–453.
- Çolak, A.v. 2003 Erste Untersuchungen über die Rottenstruktur von *Abies bornmülleriana* Mattf. in Nordwestanatolien [A preliminary study on tree clumps of *Abies bornmülleriana* Mattf. in North Anatolia]. *Allg. Forst-u. J-Ztg.* 174 (4), 117–122.
- Cressie, N.C. 1993 Statistics for Spatial Data. John Wiley & Sons, Inc.
- Dale, M.R.T. 1999 Spatial Pattern Analysis in Plant Ecology. Cambridge University Press, Cambridge.
- Diggle, P.J. 1983 Statistical Analysis of Spatial Point Patterns. Academic Press, London.
- Elliott, J.M. 1971 Some Methods for the Statistical Analysis of Samples of Benthic Invertebrates. Freshwater Biology Association, Ambleside.
- Gadow, K.v. and Pogoda, P. 2000 Assessing forest structure and diversity. *Man. For.* 1, 1–8.
- Gavrikov, V. and Stoyan, D. 1995 The use of marked point processes in ecological and environmental forest studies. *Environ. Ecol. Stat.* 2, 331–344.
- Hanus, M.L., Hann, D.W. and Marshall, D.D. 1998 Reconstructing the spatial pattern of trees from routine stand examination measurements. For. Sci., 44, 125–133.
- Hoel, P.O. 1943 On indices of dispersion. *Ann. Math. Statist.* 14, 155–162.
- Hui, G.Y., Albert, M. and Chen, B.W. 2003 Reproduction der Baumverteilung im Bestand unter Verwendung des Structurparameters Winkelma β [Reproduction of the spatial distribution of trees using the structural parameter uniform angle index (Winkelma β)]. *Allg. Forst-u. J-Ztg.* 174 (5–6), 109–116.
- Johnson, N. and Kotz, S. 1969 *Discrete Distributions*. John Wiley and Sons.
- Lähde, E., Laiho, O., Norokorpi, Y. and Saksa, T. 1999 Stand structure as the basis of diversity index. *For. Ecol. Manage.* 115, 213–220.
- Leblanc, S.G. and Chen, J.M. 1997 NDVI Directionality in boreal forests: a model interpretation of measurements. *Can. J. Rem. Sens.* 23, 368–379.

Letcher, B.H., Priddy, J.A., Walters, J.R. and Crowder, L.B. 1998 An individual-based, spatially-explicit simulation model of the population dynamics of the endangered red-cockaded woodpecker, *Picoides* borealis, Biol. Conserv. 86, 1–14.

- Lewandowski, A. and Gadow, K.v. 1997 Ein Heuristischer Ansatz zur Reproduction von Waldbeständen [A method for reproducing uneven-aged forest stands]. *Allg. Forst-Jagdzg.* 168 (9), 170–174.
- Matis, C. 2000 Forest Biometry. II. Dendrometry. Pigasos, Thessaloniki [In Greek].
- Moeur, M. 1993 Characterizing spatial patterns of trees using stem-mapped data. *For. Sci.* 39, 756–775.
- Penttinen, A., Stoyan, D. and Henttonen, H.M. 1992 Marked point process in forest statistics. For. Sci. 38, 806–824.
- Perry, J.N. and Mead, R. 1979 On the power of the index of dispersion test to detect spatial pattern. *Biometrics* 35, 613–622.
- Pielou, E.C. 1960 A single mechanism to account for regular, random and aggregated populations. J. Ecol. 49, 574–584.
- Pielou, E.C. 1977 Mathematical Ecology. Wiley, New York.
- Pommerening, A. 2002 Approaches to quantifying forest structures. *Forestry* 75, 305–324.
- Pretzsch, H. 1997 Analysis and modeling of spatial stand structures: methodological considerations based on mixed beech–larch stands in Lower Saxony. *For. Ecol. Manage.* 97, 237–253.
- Reich, R.M. and Davis, R. 1998 Quantitative spatial analysis (Course notes). Colorado State University.
- Schreuder, H.T., Gregoire and Wood, G.B. 1993 Sampling Methods for Multiresource Forest Inventory. John Wiley & Sons.

- Stamatellos, G. 1991 Research of forest volume estimation possibilities with two-stage sampling designs. Ph.D. dissertation, Aristotle University of Thessaloniki.
- Stamatellos, G., Karamanolis, D. and Zagas, Th. 2001 The spatial pattern of forest tree diameter classes in the frame of structure and management of a forest ecosystem. *Proceedings: Ecological Protection of the Planet Earth*, Xanthi 5–8 June, Vol. 1. V. Tsichrigis and P. Tsalides (eds), pp. 409–415.
- Statistical Sciences 1999 S-PLUS 2000 *Programmer's Manual*. Data Analysis Products Division, MathSoft, Seattle, WA.
- Stoyan, D. and Penttinen, A. 2000 Recent applications of point process methods in forestry statistics. *Statist. Sci.* **15**, 61–78.
- Thompson, H.R. 1958 The statistical study of plant distribution patterns using a grid of quadrats. *Aust. J. Bot.* 6, 322–342.
- Tomppo, E. 1986 Models and methods for analyzing spatial patterns of trees. Commun. Instit. Forestal. Fenn. No. 138.
- University Forest Service 1978 Management Plan of Pertouli University Forest During 1979–1988. Pertouli, Greece.
- Zagas, Th. and Smiris, P. 1993 Silvicultural research in fir stands at the University forest of Pertouli. Scientific Annals of the Department of Forestry and Natural Environment, Vol. LST, Thessaloniki.
- Zenner, E.K. and Hibbs, D.E. 2000 A new method for modeling the heterogeneity of forest structure. *For. Ecol. Manage.* **129**, 75–87.

Received 13 January 2003