



---

Recent Applications of Point Process Methods in Forestry Statistics

Author(s): Dietrich Stoyan and Antti Penttinen

Source: *Statistical Science*, Feb., 2000, Vol. 15, No. 1 (Feb., 2000), pp. 61-78

Published by: Institute of Mathematical Statistics

Stable URL: <https://www.jstor.org/stable/2676677>

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



JSTOR

*Institute of Mathematical Statistics* is collaborating with JSTOR to digitize, preserve and extend access to *Statistical Science*

# Recent Applications of Point Process Methods in Forestry Statistics

Dietrich Stoyan and Antti Penttinen

*Abstract.* Forestry statistics is an important field of applied statistics with a long tradition. Many forestry problems can be solved by means of point processes or marked point processes. There, the “points” are tree locations and the “marks” are tree characteristics such as diameter at breast height or degree of damage by environmental factors. Point process characteristics are valuable tools for exploratory data analysis in forestry, for describing the variability of forest stands and for understanding and quantifying ecological relationships. Models of point processes are also an important basis of modern single-tree modeling, that gives simulation tools for the investigation of forest structures and for the prediction of results of forestry operations such as plantation and thinning.

*Key words and phrases:* Point process, mark, modeling, ecology, intensity, variability indices, second order characteristic, correlation, single-tree model, Cox process, Gibbs process.

## 1. INTRODUCTION

A lot of time and money has been, and will be, spent on forest statistics. The original purpose was estimation of wood volume and other similar economically important characteristics for a (large) forest area. Later the reliability of inventories or, in statistical terms, the precision of the estimators, was studied. Modern technology, including remote sensing, digital aerial photographs and spectrometer imaging, is used nowadays to develop inventories, posing new statistical problems. Furthermore, today wood production is only one part of forestry; both government, industry and forest owners are interested in multiuse of forests. Therefore the objective of modern forestry includes a multitask problem field, for which forest ecology, landscape ecology and related statistical methods become increasingly important.

---

*Professor Dietrich Stoyan is Professor, Institut für Stochastik, TU Bergakademie Freiberg, Bernhard-von-Cotta-Str. 2, D-09596, Freiberg, Germany (e-mail: stoyan@orion.hrz.tu-freiberg.de). Dr. Antti Penttinen is with the University of Jyväskylä, Department of Statistics, P.O. Box 35, 40351 Jyväskylä, Finland (e-mail: penttine@jyu.fi).*

Forestry uses numerous methods for statistical analysis; see, for example, Avery and Burkhart (1983) and Schreuder, Gregoire and Wood (1993). A considerable part of them belong to spatial statistics that includes point process statistics as a particularly interesting area. There exist many successful applications of point process statistics in forestry that may serve as a pattern for other sciences.

As the name “point process” indicates, the origin of the theory was the study of random point sequences on the time axis. Such processes still play an important role, for example, in models of queueing or telecommunication. But today spatial point processes seem to be the dominating field, and sometimes methods of spatial point process statistics are even converted back to the temporal case; see Cannon and Cressie (1997). Important fields of application of spatial point processes are astronomy, geology, archaeology and materials science. But still none of these fields uses point process methods so intensively, and has stimulated the theory so much, as has forestry.

Typically point process analysis and modeling in forestry comes down to such details as locations of single trees and tree characteristics such as diameter at breast height (dbh), tree height, stem increment during a given time span, species code or de-

gree of damage by environmental factors. It is quite natural to apply here planar (marked) point process methods: the tree locations are “points” in the plane, while the characteristics are “marks.”

Data sets in this context take the form  $\{[x_i, y_i; m_i]\}$ , giving the locations  $\mathbf{x}_i = (x_i, y_i)$  and the marks  $m_i$  of all relevant trees in a window  $W$  of observation, which is typically a forest stand in a rectangle, polygon or circle. However, it is still laborious to obtain such data sets by classical measurement methods. New technological developments are changing the situation drastically. For example, distance measurements can be done using laser techniques and the satellite-based global positioning system (GPS). Already in the 1980s, point process data sets were determined manually from aerial photographs for small forest stands. At the moment, the resolution in digital aerial imaging in use reaches the level of 0.5 m to 1.5 m. Thus individual trees can be observed through image analysis with reasonable precision, with the exception of trees that are very close together or of small trees growing within the canopy of bigger trees. A compromise are *indirect observations*, such as aerial images combined with ground-based measurement. Even in this context, point process models may be helpful; see, for example, Dralle and Rudemo (1996). Marked point processes will thus become important components of what is called “multi-source” forest inventory (e.g., Schreuder, Gregoire and Wood, 1993) in modeling small-scale variation and restoration of indirect observations.

Several other spatial statistical methods are used in forestry. One example is an investigation in which a grid of observation points is placed in a forest, and for each grid point some close trees are measured. Geostatistical methods can be applied for analyzing such data.

Many of the methods described in this paper can be applied in statistical analyses of other ecosystems, in particular plant communities. Linking of forest structure and species occurrence in a variety of spatial and temporal scales is an objective in landscape ecology (see, e.g., Turner and Gardner, 1991). There (marked) point processes play a central role in small-scale modeling and, through smoothing or averaging transformations, also in large scale modeling. One of the advantages of point process models and their transformations is *consistency* on different scales.

Two of the first papers using ideas from forestry-oriented point process statistics are the book by König (1835) and the article by Svedberg (1922). While König tried to estimate tree density by a distance method, Svedberg used a Poisson process for

explaining spatial patterns of plant communities. The work of Matérn (1960) marks the beginning of modern point process statistics in forestry. For forestry purposes, he developed two hard-core process models and a cluster process model and gave expressions for the corresponding second order characteristics. A now classical survey paper on point process methods in forestry statistics is Warren (1972). Since then many new ideas have appeared and today there is a vast literature, where the books by Diggle (1983) and Cressie (1993) have been very influential. The following modern ideas may be seen as particularly important:

1. The use of *functions* of the character of density functions that describe variability and interaction in a given forest on various scales, replacing earlier numerical indices or cumulative functions.
2. The construction of *single-tree models* of whole forests, which enable computer simulation of forest development over long periods with consideration of the ecological situation of every individual tree.
3. The construction of advanced *stochastic models* for marked point patterns.

This paper reviews, for a broad readership, ideas and problems that arise in the application of point process statistics in forestry. After some remarks on ecological problems, the classical problem of measurement of tree density is discussed. Then the description of forest stand variability by indices and correlation functions and some problems of ecological modeling are described. Single-tree models are briefly explained, and the most attractive point process models for forestry today, Cox and Gibbs processes, are discussed.

## 2. SPATIAL PATTERNS AND ECOLOGICAL PROCESSES

A forest stand is the result of former land use, complex ecological processes and practical forestry. The corresponding spatial marked point pattern is an important source of information on the plant population. Although the underlying biological processes cannot be deduced directly from observed patterns, they themselves and their changes are highly indicative of certain processes; according to Tilman (1988), a central goal of ecology is to understand the causes of the patterns observed in the natural world. This task is very complicated when only a snapshot analysis based on one pattern is pos-

sible; repeated observations yield, of course, more information. It is possible that similar patterns result from different ecological processes. In particular, a completely random pattern, that is, a point distribution as in (or similar to) a Poisson process, may appear in the beginning of an ecological evolution (for example, in the case of natural regeneration of beeches or birches), but it may also be a result of multiphase ecological processes. Thus its appearance does not necessarily mean that there was (or is) no interaction or that the pattern is a result of the action of many small factors (Odum, 1971, page 215). These problems are described in the paper by Lepš (1990). The application of point process statistics may help to study the patterns, to pose statistical and ecological hypotheses and to investigate their agreement with empirical data.

Natural patterns of trees often show clumping or *clustering* in their starting phase, caused by environmental heterogeneity, seed dispersion and competition with other species; they are sometimes called *environmental*, *morphological* and *sociological* causes. Young populations show large variations both in space and time due to irregular seeding, high mortality, lack of competition and possible changes in the local environmental conditions. A typical study of such a population is Salonen, Penttinen and Särkkä (1992), who consider plant colonization of a bare peat surface over five years.

During the evolution of a forest there typically exists a tendency towards *regularity*. This can be mainly explained by competition among neighboring trees and by dependence of mortality on local population density. Also, environmental variables such as ground cover vegetation, light conditions, microclimate, soil characteristics, profile and ecological history of the forest area play an important role.

In a forest, both forms of variability, clustering and regularity, may coexist on various scales. For example, trees may be regularly distributed on a small scale (for older trees due to competition among neighbors and for young trees due to planting in rows), but be clustered on a larger scale (due to ecological heterogeneity).

Of course, the forester's work is decisive for forest development; even his personality has an influence; see Fuldner, Sattler, Zucchini and von Gadow (1996). The forester's influence often begins with planting, but he later becomes a competition manager, first by weeding out competing shrubs and thinning of young trees and then by removing certain trees to promote the growth of others. Figures 1 and 2 show two typical tree patterns. For the natural stand of young pines in Finland (Figure 1) the clustering is obvious; in contrast, in the old spruces

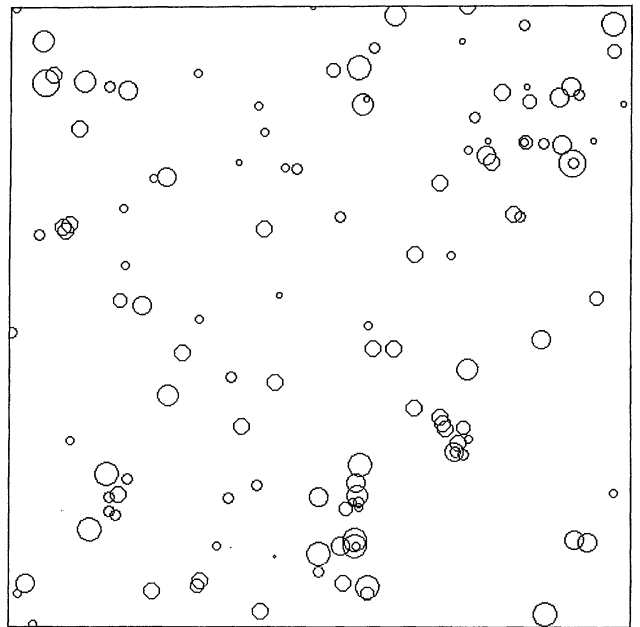


FIG. 1. 126 young pines in a 10 m × 10 m square plot in a Finnish forest. The trees are shown as circles, the diameters of which are proportional to the tree heights. This pattern is highly clustered; some trees are very close together.

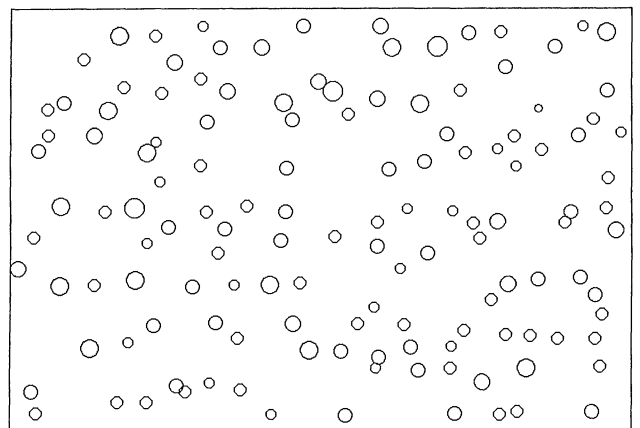


FIG. 2. 134 spruces of age 60 years in a 56 m × 38 m stand of Tharandter Wald (Germany). The trees are shown as circles, where the diameters are proportional to the bhd's.

stand (Figure 2) there is some form of regularity, mainly as a result of the forester's work.

### 3. FORESTRY SUMMARIES

The simplest summary characteristic of a forest is *intensity*  $\lambda$ , defined as the stand density or average number of trees per unit area. Multiplied by the area of the stand  $B$ , it yields the expected total number of trees,

$$(1) \quad E(\Phi(B)) = \lambda \nu(B).$$



Here  $\nu$  denotes the planar Lebesgue measure yielding the area of  $B$  and  $\Phi(B)$  denotes the number of points in the set  $B$ . The character  $\Phi$  also denotes the point process, that is, the set of all points or tree positions. Formula (1) holds if  $\Phi$  is stationary, that is, its distribution is translation invariant.

The estimation of the intensity  $\lambda$  is not a trivial problem in forestry. The application of the standard estimator,

$$(2) \quad \hat{\lambda} = \frac{\Phi(W)}{\nu(W)},$$

is practically difficult: counting the number of trees  $\Phi(W)$  in a large sampling window  $W$  is laborious. Therefore, various other stand density estimators have been developed that are based on sampling and often on fast measurement methods.

Particularly simple in practice are *distance* methods; see Ripley (1981) and Krebs (1989). The idea here is to measure distances, “tree–nearest neighboring tree” or “test point–nearest tree” distances, and to transform them into intensity estimators. The method goes back to König (1835) who suggested the estimator

$$\hat{\lambda}_K = \frac{1}{\bar{\delta}^2},$$

where  $\bar{\delta}^2$  is the empirical second moment of “tree–nearest neighboring tree” distances. His formula is precise in the case of a square lattice, while he remains aware of the randomness of forests. The estimator is sensitive to the underlying process. For example, in the case of a Poisson process it is natural to put

$$\hat{\lambda}_P = \frac{1}{4\bar{\delta}^2},$$

using the formula for the mean nearest neighbor distance  $\bar{\delta}$  of the homogeneous Poisson process. Many alternative estimators for  $\lambda$  have been proposed (cf. Ripley 1981, page 134), but none is particularly robust, thus restricting their use for forestry purposes.

Tree numbers alone do not tell enough about a forest; it is important to consider measures of tree quantities such as height, basal area or volume. This concept can be realized by means of marked point processes.

Let  $\Psi = \{[\mathbf{x}; m_i]\}$  be a marked point process with scalar marks  $m_i$  for simplicity. Their distribution  $M$  is called the *mark distribution*. In the stationary case,  $M$  can be interpreted as the distribution of the mark of the “typical” point. Statistically it is obtained as the empirical distribution of the marks of points in a sampling window  $W$ .

Many useful forest summaries are of the general form

$$\sum_{[\mathbf{x}; m] \in \Psi} f(\mathbf{x}, m)$$

for any nonnegative measurable function  $f$ . A particular case is  $f(\mathbf{x}, m) = \mathbf{1}_B(\mathbf{x})m$ , where  $\mathbf{1}_B$  denotes the indicator function of the set  $B$ . Then the sum is equal to the sum  $S(B)$  of the marks of all points in  $B$ , for example, the total wood volume in a stand.

For a stationary marked point process  $\Psi$ , the following formula, often called the Campbell theorem (see Stoyan, Kendall and Mecke, 1995, page 106) holds:

$$(3) \quad E\left(\sum_{[\mathbf{x}; m] \in \Psi} f(\mathbf{x}, m)\right) = \lambda \int \int f(\mathbf{x}, m) M(dm) d\mathbf{x},$$

where the integration is over  $\mathbb{R}^2$  and the set of all possible marks. In particular, (3) yields for the  $f(\mathbf{x}, m)$  above and a stationary marked point process,

$$E(S(B)) = \lambda \nu(B) \mu,$$

where  $\mu$  is the mean mark.

A forestry-originated method is Bitterlich sampling (Bitterlich, 1948, 1952) or *angle count sampling*. It is used in practical forestry as a rapid tool for measuring the *proportional mean basal area*  $a$ , which is defined as the mean of the sum of all cross-section areas at breast height of the trees in a given stand  $B$  divided by the area  $\nu(B)$  of the stand,

$$(4) \quad a = \frac{\pi}{4} E\left(\sum_{[\mathbf{x}; \beta] \in \Psi} \beta^2 \mathbf{1}_B(\mathbf{x})\right) / \nu(B);$$

here,  $\beta$  denotes the random dbh (=diameter at breast height) mark. The Campbell theorem (3) yields

$$(5) \quad a = \frac{\pi}{4} \lambda E(\beta^2),$$

independently of  $B$ . The mean basal area is closely related to the mean volume of trees, a central objective in forest inventories.

In the angle count method, sampling points  $\xi$  are chosen either randomly or following some sampling design. For each  $\xi$ , the trees that can be seen in an angle larger than a fixed size  $2\alpha$  are counted; see Figure 3. Their number is multiplied by  $\sin^2 \alpha$ , yielding the estimator  $\hat{a}(\xi)$  of  $a$  given as

$$\hat{a}(\xi) = \sum_{[\mathbf{x}; \beta] \in \Psi} b(\xi, \mathbf{x}, \beta)$$

with

$$b(\xi, \mathbf{x}, \beta) = \sin^2 \alpha \mathbf{1}_{\{\|\mathbf{x} - \xi\| < \beta/2 \sin \alpha\}}(\mathbf{x}, \beta).$$

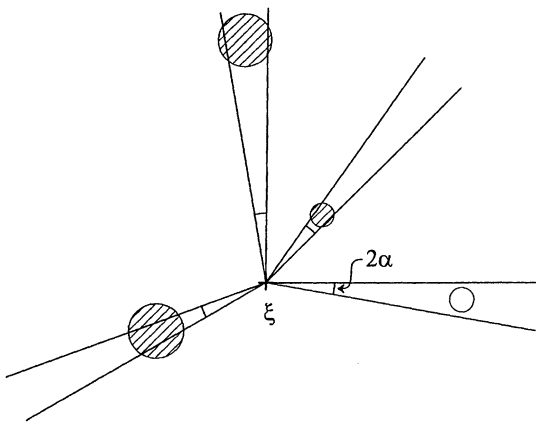


FIG. 3. Schematic explanation of Bitterlich sampling. The observer's position is in  $\xi$ . All hatched trees are counted, that is, all trees seen under an angle larger than  $2\alpha$  ( $\alpha$  fixed).

Again (3) gives

$$E(\hat{a}(\xi)) = \alpha,$$

that is, Bitterlich's estimator is unbiased.

It is of practical importance that when estimating the mean basal area, the measurement can be carried out using a simple piece of equipment called Relaskop and the investigator does not need to measure the dbh's. By the way, here the probability that a randomly selected sample point picks a given tree is related to its basal area, and Bitterlich's estimator can be shown to be a Horvitz–Thompson-type estimator; see, for example, Overton and Stehman (1995). Bitterlich sampling can also be used in the estimation of other tree quantities such as mean height.

Incidentally, Bitterlich sampling can be considered as sampling from a random field; see Penttinen (1988). Define the random variable

$$(6) \quad Y(\xi) = \sin^2 \alpha \sum_{[x, \beta] \in \Psi} \mathbf{1}_{\{\|\mathbf{x} - \xi\| < \beta/2 \sin \alpha\}}(\mathbf{x}, \beta),$$

where  $\xi$  is any sampling point of  $\mathbb{R}^2$ . Then  $\{Y(\xi), \xi \in \mathbb{R}^2\}$  is a random field with mean  $\alpha$ . This approach opens a way for application of geostatistical methods to forestry.

## 4. DESCRIPTION OF VARIABILITY

### 4.1 Classical Variability Indices

In the 1950s and 1960s, some numerical indices were created with the aim of describing aspects of variability in a forest by a single number; see Ripley (1981). Usually these indices were based on pairs: "tree-neighboring tree." The first index of this type is the *aggregation index* of Clark and Evans (1954),

which is defined as

$$(7) \quad R = \frac{\bar{\delta}}{E(D)}.$$

Here  $\bar{\delta}$  is the empirical mean of the distances from the trees to their nearest neighbors, while  $E(D)$  is the mean nearest neighbor distance in a Poisson forest of the same intensity  $\lambda$  as the forest under analysis. Values of  $R$  greater than 1 indicate that the pattern has a tendency to regularity, while  $R < 1$  points to clustering. The index  $R$  has been used to quantify and prove statistically the intuitively clear statement that in a forest one often finds increasing regularity (and thus increasing  $R$ ) with age.

A further index is the *coefficient of segregation*  $S$  of Pielou (1977). It describes the degree of mixing of trees of two species  $A$  and  $B$  in a forest, and, like  $R$ , it is based on nearest neighbor distances. Let  $N$  be the number of all trees, let  $m$  and  $n$  be the numbers of base trees of species  $A$  and  $B$ , respectively, and let  $r$  and  $s$  be the numbers of times trees of species  $A$  and  $B$  have served as nearest neighbors ( $NN$ ) of other trees, respectively. These numbers are set out in an obvious form in a  $2 \times 2$  table as follows:

Species of NN			
	A	B	
Species of A	$a$	$b$	$m$
base plant B	$c$	$d$	$n$
	$r$	$s$	$N$

The coefficient of segregation is then defined by

$$(8) \quad S = 1 - \frac{N(b+c)}{ms+nr}.$$

If the nearest neighbors are always of the same species as the base tree, then  $S = 1$ ; if all neighbors are of different species then  $S = -1$ . In the case of complete randomness of species distribution, one can expect values of  $S$  around 0.

A disadvantage of both indices is that they are "short-sighted," because they only consider nearest neighbors; what is beyond the nearest neighbors is ignored. However, these indices are still of value when applied in situ in a forest, when only nearest neighbor distances are measured, not tree locations.

### 4.2 Correlation Functions

Modern point process statistics uses *functions* instead of indices such as those described in the preceding section. These functions depend on the

intertree distance  $r$  and thus are in relation to various ecological scales in the patterns. Probably the most successful functions in describing tree density variability are the *second order characteristics*.

In order to avoid technical difficulties, it is assumed here that the analyzed point process  $\Phi$  is stationary and isotropic; that is, its distribution is invariant with respect to translations and rotations. In this case all process characteristics of second order can be expressed by means of the intensity  $\lambda$  and Ripley's  $K$  function (Ripley, 1976, 1977). The quantity  $\lambda K(r)$  is interpreted as the mean number of points of  $\Phi$  in a disc of radius  $r$  centred at a "typical" point of  $\Phi$ , which is not counted.

The statistical estimation of  $K(r)$  follows that idea, by giving in turn all the points of a sample of  $\Phi$  the role of the "typical" point. If, as usually happens in forestry, only the points in a bounded window  $W$  of observation can be studied, edge correction is necessary to obtain precise estimates. The quantity  $\lambda^2 K(r)$  can be estimated by an unbiased estimator of Horvitz–Thompson type (Ohser and Stoyan 1981; Baddeley 1998). It is then divided by an estimate of  $\lambda^2$ . By no means is it optimal to use simply the squared  $\lambda$  of formula (2); instead, more elaborate intensity estimators should be applied that are adapted to  $W$  and  $r$ ; see Stoyan and Stoyan (1998b).

As it happens, one of the first applications of the  $K$  function was for a forestry example: Ripley (1977) analysed positions of redwood seedlings and found that they are more clustered than the points of a Poisson process. The book by Diggle (1983) is the reference for statistical analyses of point processes using the  $K$  function; many of the examples considered there have a forestry or ecology background.

In the case of the homogeneous Poisson process, the  $K$  function satisfies  $K(r) = \pi r^2$ . As in this particular case, the function  $K(r)$  for other point processes increases as  $r^2$ . Thus a popular choice is the  $L$  function,

$$(9) \quad L(r) = \sqrt{K(r)/\pi} \quad \text{for } r \geq 0,$$

also because the square root transformation stabilizes variances. Figures 4a and 4b show the  $L$  functions for the young pines and the spruces, respectively. The curves show the typical behavior for a cluster process and for a soft core process: values greater and smaller than  $r$ , respectively. Figures 5a and 5b will show the distributional differences more clearly.

The  $L$  function is a valuable tool for goodness-of-fit tests in point process statistics. In such tests the empirical (estimated)  $L$  function is compared to the  $L$  function of the model to be tested. If the

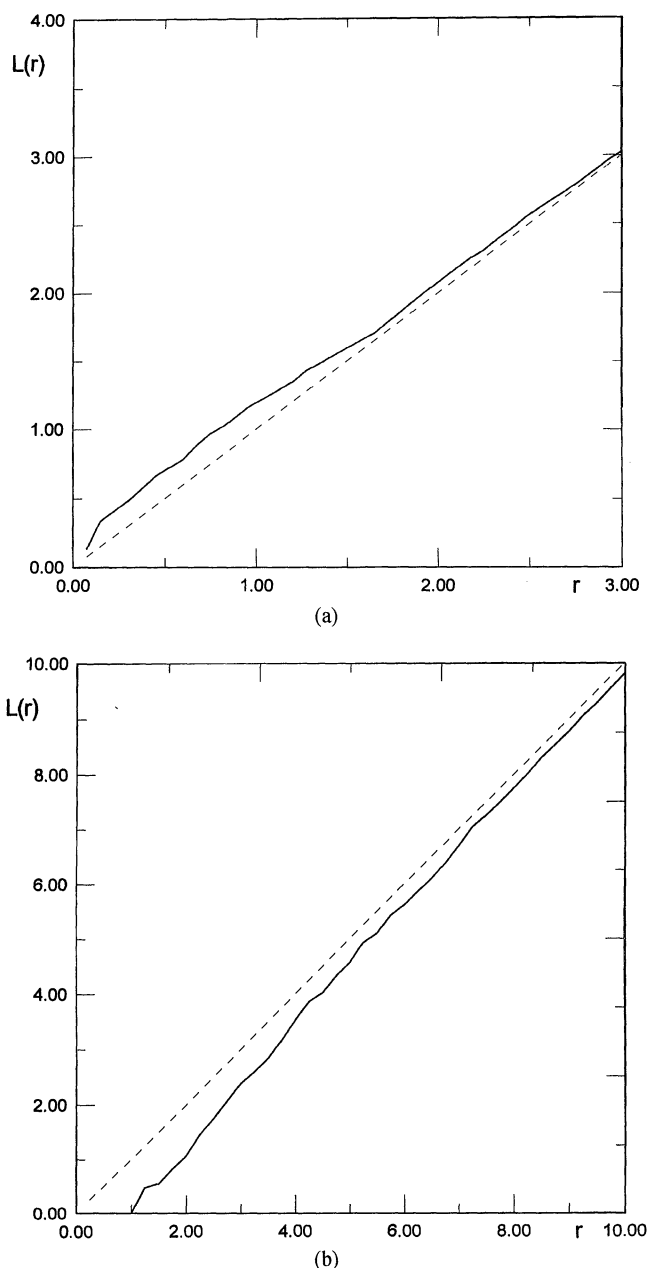


FIG. 4. The empirical  $L$  function for the young pines of Figure 1a and for the spruces of Figure 2b. The values of  $\hat{L}(r)$  for the pines are greater than  $r$  for  $r \leq 3$  m, which indicates clustering of the pattern. For the spruces of Figure 2,  $\hat{L}(r)$  vanishes for  $r < 1$  m because the minimum intertree distance is 1 m. The values of  $\hat{L}(r)$  smaller than  $r$  indicate some form of "repulsion" between the trees, which is in this case mainly a result of forester's work. Figure 5 will better represent the information obtained by second order characteristics.

null hypothesis is a Poisson process, then the  $L$ -test introduced by Ripley can be used; see Ripley (1988). For testing of other null hypotheses, typically simulation tests are used; see Diggle (1983); Ripley (1988) and Stoyan and Stoyan (1994).

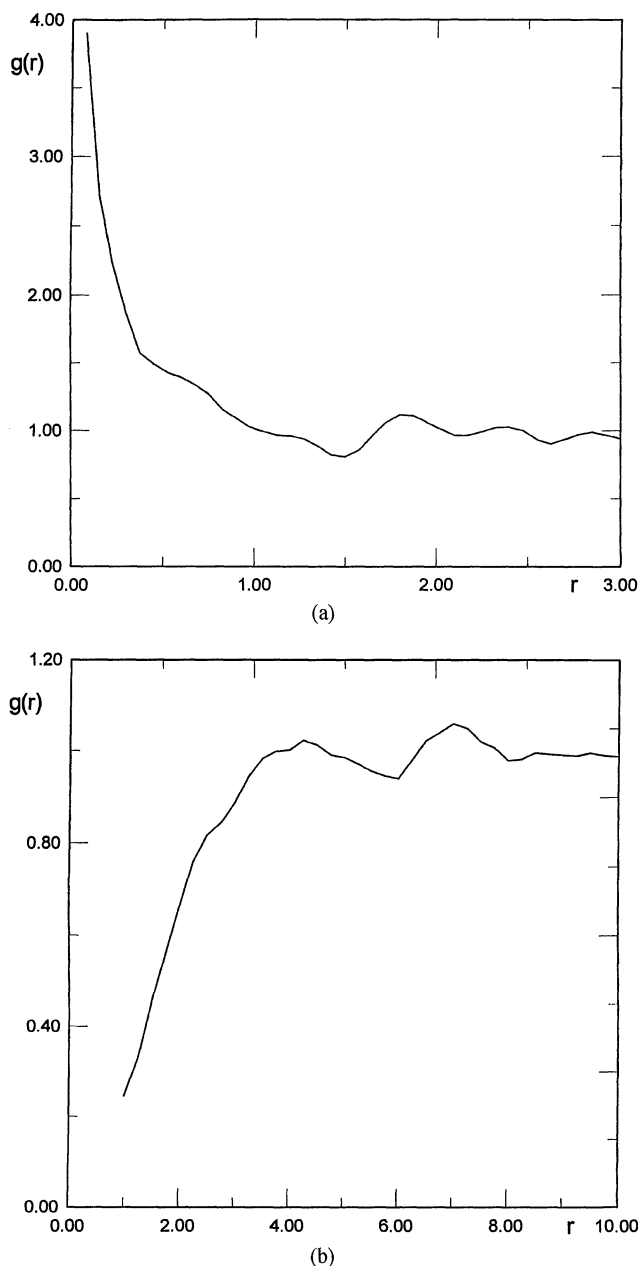


FIG. 5. The empirical pair correlation function for the young pines of Figure 1a and for the spruces of Figure 2b. For the pines, the function has the typical form of the pair correlation function of a cluster process. For the spruces of Figure 2, the function has the form of the pair correlation function of a point process with a hard-core distance and soft-core repulsion for larger distances.

There is another function that presents the information given by the  $K$  or  $L$  function in a clearer, easier way. The situation is similar to that in classical univariate statistics where the cumulative distribution function is a good descriptor of a distribution that is particularly useful in goodness-of-fit tests, but the probability density function (or its empirical counterparts such as kernel estimates, his-

tograms or stem-and-leaf plots) has proved to be more attractive for exploratory data analysis, because most people can more easily interpret a probability density function than a cumulative distribution function. A second order function of nature similar to that of a density function is the *pair correlation function*  $g(r)$ . It is related to the  $K$  function by

$$(10) \quad g(r) = \frac{d}{dr} K(r) / (2\pi r) \quad \text{for } r \geq 0.$$

An independent heuristic definition is as follows: consider two infinitesimally small discs of areas  $dx$  and  $dy$  at a distance  $r$ . Let  $p(r)$  denote the probability that each disc contains a point of the process. Then

$$p(r) = \lambda^2 g(r) dx dy.$$

For a completely random point process (i.e., a homogeneous Poisson process),  $g(r) \equiv 1$ . Values of the pair correlation function  $g(r)$  larger than one indicate that interpoint distances around  $r$  are relatively more frequent than they would be under a completely random point process; if this is the case for small values of  $r$ , typically there is clustering. Conversely, values of  $g(r)$  smaller than one indicate that the corresponding intertree distances are rare, which may indicate inhibition. It is even possible that  $g(r) = 0$ ; then the interpoint distance  $r$  has zero probability in the point process. The pair correlation function can take on any value between zero and infinity; for large  $r$  it tends to 1. It can be interpreted as a modern counterpart to the aggregation index of Clark and Evans. Section 14.4.2 in Stoyan and Stoyan (1994) discusses in detail the interpretation of pair correlation functions.

Figure 5 shows the estimated pair correlation functions for the two forest stands already discussed. For the Finnish young pines it has the typical form of the pair correlation function of a cluster process. If one accepted the hypothesis of circular clusters, then the pair correlation function would suggest a mean cluster of about one meter in diameter. However, Møller, Syversveen and Waagepetersen (1997) have shown that the simple Matérn cluster process with circular clusters of diameter one meter, that was used in Stoyan and Stoyan (1994) for these data, is probably not a good model; instead, their more flexible Cox process model, which will be described in Section 6.2, seems to give a better fit.

For the spruce data, a very different pair correlation function is obtained. This point pattern has a hard-core distance of about one meter (this is the



minimum intertree distance in the stand). Therefore, the pair correlation function vanishes for  $r$  between zero and one meter. For values greater than  $r = 1$  m, the pair correlation function is similar to that of so-called soft core processes. Probably the waves of the estimate for values of  $r$  larger than eight meters are only statistical fluctuations and do not give relevant information. Thus the range of correlation in that stand is about eight to ten meters, a value which has been often observed for older stands in the forestry literature; for the young pines the range may be about two meters. In general, for mature Boreal forests there were observed three partially overlapping scales: “hard-core” distances of zero to two meters, tree interaction and correlation with distances between one and ten meters and independence or influence of environmental factors, with distances over seven meters.

One should note that the pair correlation function does not directly define the range of mutual interaction between trees. A measure which does do this, the pair potential function, is discussed in Section 6.3. For example, a point pattern of trees with a short-range inhibition may show long-term correlation behavior.

The pair correlation function also appears in an important formula in point process statistics. The variance of the random number  $\Phi(B)$  of points  $\Phi$  in any set  $B$  is given by

$$(11) \quad \text{var}(\Phi(B)) = \lambda^2 \int_0^\infty \bar{\gamma}_B(r) 2\pi r g(r) dr + \lambda \nu(B) - (\lambda \nu(B))^2.$$

Here  $\bar{\gamma}_B(r)$  denotes the isotropized set covariance function of  $B$ ; see, for example, Stoyan and Stoyan (1994) for a definition. Thus the variance of the intensity estimator  $\hat{\lambda}$  defined by (2) can be calculated if  $g(r)$  is known.

The second order analysis can be extended to consider also inhomogeneous point patterns through modifications of the  $K$ - and  $g$ -functions; see Baddeley, Møller and Waagepetersen (1998). In Stoyan, Kendall and Mecke (1995) and in Stoyan and Stoyan (1994) the unisotropic stationary case is considered, that is between motion invariant and inhomogeneous.

### 4.3 Nearest Neighbor Distances

For many statisticians it was (and still is) natural to consider the nearest *neighbor distance* distribution function  $D$  instead of the pair correlation function or the  $L$  function; this distribution function is easier to understand and to measure in situ. It describes the distribution of the distance from a “typical” point of a point process to its nearest neighbor;

see the discussion in Stoyan, Kendall and Mecke (1995), Section 4.4. The estimation of  $D(r)$  follows that idea, by giving in turn all the points of a sample of  $\Phi$  the role of the “typical” point and checking whether its nearest neighbor is closer than  $r$ . Here edge correction is recommendable, and again Horvitz–Thompson estimators are helpful; see Chiu and Stoyan (1998). Baddeley and Gill (1997) used Kaplan–Meyer estimators. Figure 6 shows the statistical estimates of  $D$  for the two tree patterns of Figures 1 and 2. They also show the corresponding probability density functions  $d(r)$ , which were estimated by means of a kernel estimator. The distributional differences of the two tree patterns are clearly expressed by  $D$  and  $d$ .

In exploratory statistics it may be useful to apply in addition a function that could be called the *nearest neighbor pair hazard rate*,

$$(12) \quad h(r) = \frac{d(r)}{1 - D(r)} / (2\pi r \lambda) \quad \text{for } r \geq 0.$$

The ratio  $d(r)/1 - D(r)$  is the hazard rate corresponding to  $D$ , the statistical estimation of which is studied in Baddeley and Gill (1997). Because

$$\frac{d(r)}{1 - D(r)} dr$$

gives the probability of the appearance of the nearest neighbor of the typical point  $x$  in the annulus with radii  $r$  and  $r + dr$  centered at  $x$ , it seems natural to divide the hazard rate by the length of the corresponding perimeter  $2\pi r$ ; the division by  $\lambda$  normalizes with respect to the intensity. For a homogeneous Poisson point process,  $h(r) \equiv 1$ . Figure 6 shows estimated nearest neighbor pair rate functions for the tree patterns of Figures 1 and 2. These curves show in a particularly clear way the differences in the short range behavior of the two point patterns.

An alternative to the nearest neighbor distribution function is the *empty space statistic* established by Cox (1971) for forestry purposes. This statistic uses information on distances from sampling points to nearest trees. In the stationary case the sampling point can be taken as the origin  $o$ . Then define

$$1 - H_s(r) = P(\Phi(b(o, r)) = 0) \quad \text{for } r \geq 0.$$

The probability gives the proportion of those sampling points for which the distance to the nearest tree is larger than  $r$ . The distribution function  $H_s$  gives the distribution of that random variable; see Diggle (1983) and Ripley (1977, 1981). The character  $H_s(r)$  means “spherical contact distribution function,” as in Stoyan, Kendall and Mecke (1995)

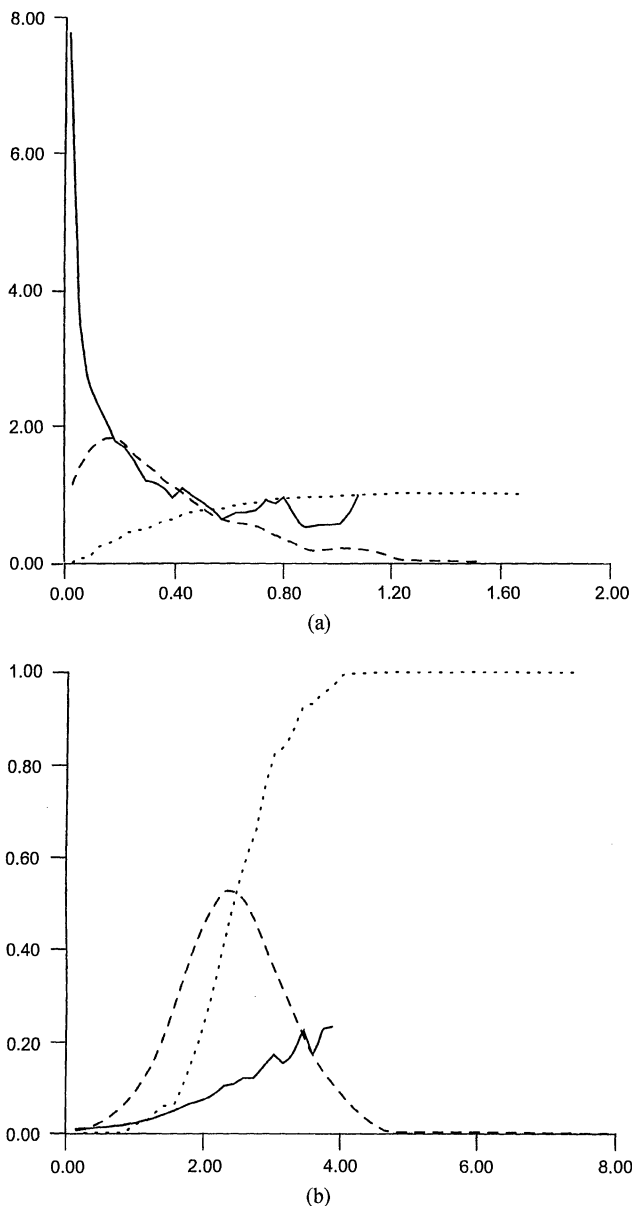


FIG. 6. The empirical nearest neighbor distribution function  $\hat{D}(r)$  (---) and the corresponding estimates of the density function  $\hat{d}(r)$  (---) and the nearest neighbor pair hazard rate  $\hat{h}(r)$  (—). For the pines of Figure 1a, all three functions have forms which are typical for cluster processes. For the spruces of Figure 2b these functions have a form which supports the idea that a soft-core process with an 1 m hard-core is a good model.

and originates in mathematical morphology as the “law of first contact;” see Serra (1982), page 488.

An extensive forestry application of the spherical contact distribution function can be found in Pohtila (1980), where the author analyzes tree patterns of different regeneration history: naturally generated, cultivated with different methods of establishment (burning-over together with planting,

ploughing with planting) and mixed (cultivated with plants of natural origin).

A compromise between the  $D$  and  $H_s$  statistics is the quotient

$$J(r) = \frac{1 - D(r)}{1 - H_s(r)}$$

introduced by Van Lieshout and Baddeley (1996), which turned out to be a valuable statistic for determining the range of interaction in point patterns.

Many statisticians today believe that second order characteristics such as the  $K$  function, or the pair correlation function  $g$ , give more interesting information on the tree distribution variability than  $D$  and  $H_s$ . The reason is that  $D$  and  $H_s$  are short-sighted, as they use only the relatively short nearest neighbor distances, whereas  $K$ ,  $L$  and  $g$  operate in a variety of scales. However, if the parameters of a point process were estimated by means of second order characteristics, then it is probably better to test the goodness-of-fit by means of a test in which the functions  $D$  or  $H_s$  are used and not by a test based again on a second order characteristic.

#### 4.4 Mark Correlation Functions

Not only the number of trees in a stand is of interest; their diameters (dbh), basal areas and other marks, are also important.

In many analyses, the marks are considered to be random but *independent*. The mathematical model is then an “independently marked point process.” This model is sometimes realistic. For example, statistical analyses suggest the assumption that the heights of the pines in Figure 1 and the dbh’s of the spruces in Figure 2 are independent; the independence of the spruces in Figure 2 is a result of forester’s work, which eliminated suppressed trees.

Correlations of the marks in a stationary marked point process can be described by mark correlation functions. Like the pair correlation function, these functions depend on the interpoint distance  $r$ . They are conditional mean values, and the condition is that there is an underlying point process point both in  $o$  and  $r$ , where  $r$  is any fixed point at a distance  $r$  from origin  $o$ . Let  $m(o)$  and  $m(r)$  be the corresponding marks. Their relationship is quantified by  $f(m(o), m(r))$ , where  $f$  is a suitable “test” function. Important examples are

$$f_1(m_1, m_2) = m_1 m_2$$

and

$$f_2(m_1, m_2) = \frac{1}{2}(m_1 - m_2)^2.$$

The mean value in the case of  $f_1$ ,  $\kappa_{f_1}(r)$ , is often normalized by division by the squared mean mark

$\mu^2$ , which then yields the *mark correlation function*  $k_{mm}(r)$ ; see Stoyan and Stoyan (1994). Other ways of describing mark correlation exist; see Schlather (1999). In the case of independent marks,  $k_{mm}(r) = 1$  for all  $r$ . Otherwise, the mark correlation function describes a certain type of correlation of the marks. If the marks are dbh's, then frequently the following behavior is observed. First,  $k_{mm}(r)$  takes on values smaller than one for small  $r$ ; it then increases and fluctuates around one for large  $r$ . This behavior may indicate that the price that trees have to pay for being close together is that their diameters tend to be smaller than the mean  $\mu$  in the whole forest. Penttinen, Stoyan and Henttonen (1992) and Gavrikov and Stoyan (1995) give forestry examples.

When using  $f_2$ , the corresponding mean  $\kappa_{f_2}(r)$  is denoted by  $\gamma(r)$ . The function is called *mark variogram* (Cressie, 1993) because of a formal parallelity to geostatistical variograms; see also Section 7. It describes the differences of the marks of trees at intertree distance  $r$ . For small  $r$ , it gives information about tree interaction, while for larger  $r$  it describes the influence of environmental factors such as soil, humidity or altitude. A mark variogram can show a behavior different from that of a geostatistical variogram, and geostatistical ideas may be insufficient for its interpretation; see the discussion Walder and Stoyan (1996) and Stoyan and Walder (1999).

In the case of a multivariate point process, that is, a point process with discrete marks characterizing, for example, species, or with discretized marks, the *mark connection functions*  $p_{ij}(r)$  are helpful. They are closely related to  $K_{ij}$  functions as considered in Diggle (1983). The value  $p_{ij}(r)$  is the conditional probability that one of the points has mark  $i$  and the other has mark  $j$ , where the condition is the same as for the mark correlation function.

If the discrete marks are independent, then the functions are constant. Otherwise, for large  $r$ , they tend to fixed values  $p_{ij}$  with

$$p_{ij} = 2p_i p_j \quad \text{and} \quad p_{ii} = p_i^2,$$

where  $p_i$  is the probability that the typical point has mark  $i$ . These functions can be seen as a modern counterpart of Pielou's segregation coefficient. For example, they are used in Gavrikov and Stoyan (1995), where the marks characterize different degrees of damage by environmental factors.

Another second order approach for the study of clustering of a property of trees (called cases) in a heterogeneous point pattern of trees is the *spatial case-control* approach by Diggle and Chetwynd (1993).

## 5. GROWTH MODELS FOR FOREST STANDS

Foresters need models for predicting the yield of forests and consequences of their work. Tools of increasing quality for satisfying this need have existed, for hundreds of years. The development began with stand yield tables in China (seventeenth century) and Central Europe (eighteenth century). These tables helped predict growth in dependence of tree species and environmental conditions. They were developed and improved over the years and have finally led to extensive computer programs.

The best modern models are the so-called *single-tree models* or *individual tree-based models*. Newnham (1964) is one of the pioneering works. The state of modern research in this field can be found in Biging and Dobbartin (1992, 1995), Liu and Ashton (1995) and Pacala et al. (1996). In such models, all trees in a stand appear explicitly. The growth behavior of each tree is described in relation to its own parameters (age, size, species and others), to those of its neighbors (through competition indices), to the whole stand and to environmental factors. Only by such models does it seem possible to predict the results of the forester's work, such as planting and thinning, with sufficient precision. Figure 7 was obtained by means of the single-tree growth simulator SILVA 2.1, a program developed by Pretzsch (1993, 1997).

It is possible to start such models in arbitrary configurations. Lepš and Kindlmann (1987) began with patterns of a Poisson process, of a hard-core process and of a cluster process. For both non-Poisson patterns, a tendency towards a more random, Poisson-like pattern could be observed.

Many single-tree models use statistical relationships which mainly result from regression analysis. For example, Rathbun and Cressie (1994) studied a single-tree model in which growth of a given tree is a function of its diameter and of the competitive influences of its neighbors, measured by their distances and diameters. Typically, such models summarize statistical relationships; they do not yet go back to the true processes of growth. However, since the 1970s, a new approach has been established in forestry: system analysis and *ecological-physiological modeling* of growth. The aim is to explain the growth processes causally.

Since many elements of the growth processes in forests are still not explored on the level of these ecological-physiological models, today single-tree models which represent a compromise between the classical single-tree models and modern ideas are used. Such models are called "hybrid models"; see Kimmins (1993).

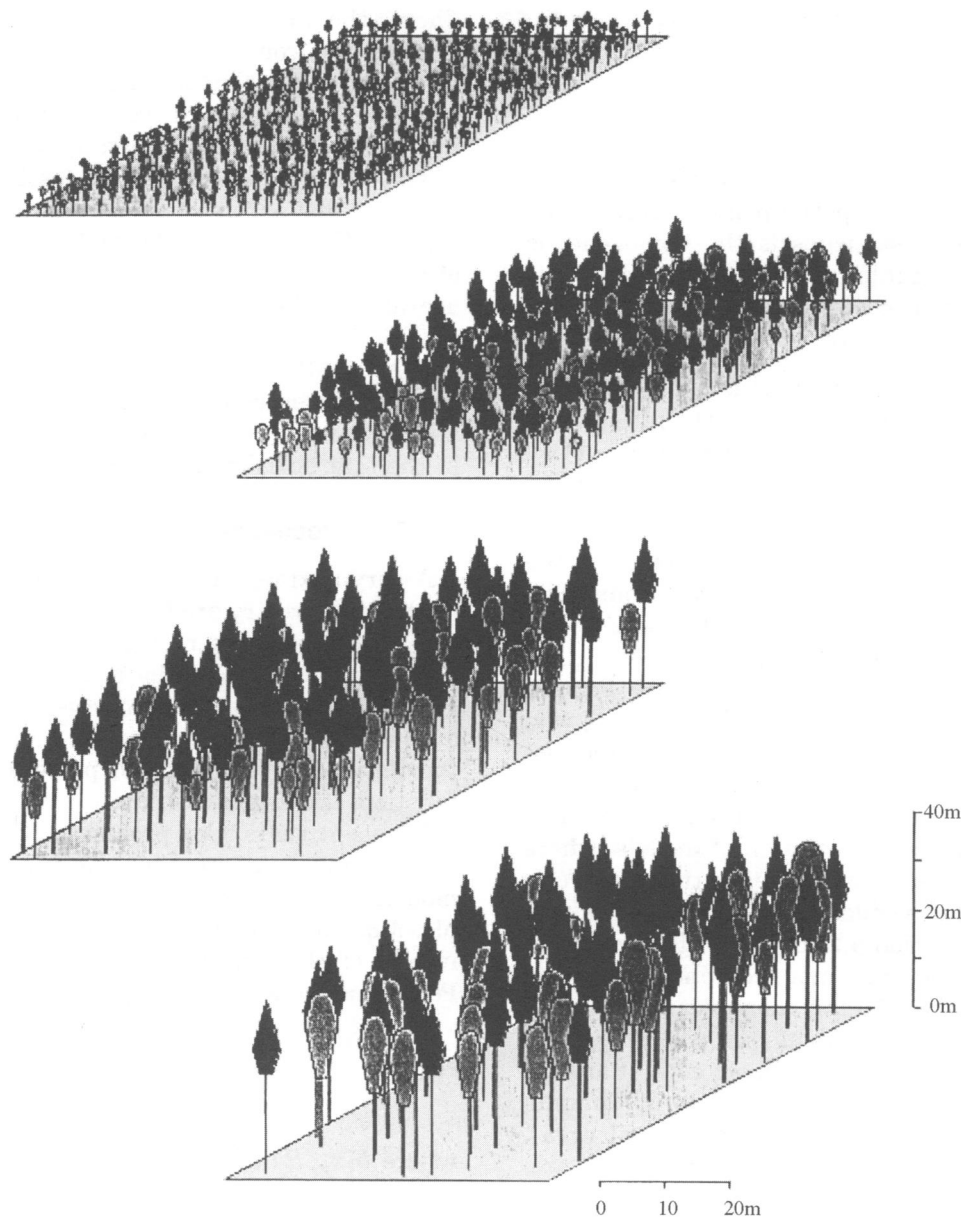


FIG. 7. Some steps of the evolution of a mixed forest with spruces and beeches, obtained by means of the single-tree model SILVA 2.1. Start configuration with spruces and beeches, aged 20 and 25 years, respectively and situations 20, 80 and 145 years later. (Published with the kind permission of H. Pretzsch.)

In the context of point processes, single-tree models can be seen as models of spatio-temporal point processes. They pose the following statistical problems:

1. Model validation, that is, checking that a given single-tree model produces realistic artificial forests.
2. Construction of geometrical models for tree crowns, which will be used as marks.
3. Modeling mortality of trees, the main random component in many forests.

4. Developing models for realistic start configurations based on a statistical analysis of any forest to be used in the simulation of its development.

## 6. POINT PROCESS MODELS

### 6.1 Poisson Process

Point process *models* play an important role in forestry statistics. They are needed both in confirmatory and exploratory data analysis and as starting configurations for single-tree models. In modern forestry statistics, two classes of point processes



are of particular interest: Cox and Gibbs processes. Probably, Gibbs processes will be the more interesting in the future because of their greater flexibility and realism.

Quite often a null model is used, since it is particularly simple or serves as a benchmark. The most important null model in spatial point process statistics and in forestry statistics is the homogeneous *Poisson process* with intensity  $\lambda$ , which corresponds to the hypothesis of *complete spatial randomness* and has the following properties:

1. The number of points in any region  $B$  follows the Poisson distribution with mean  $\lambda\nu(B)$ .
2. Given  $n$  trees in  $B$ , their positions behave as an independent sample from the uniform distribution in  $B$ .

Condition (2) says that no interaction exists in the point pattern. Incidentally, the homogeneous Poisson process is a particular case of both the Cox and Gibbs processes discussed below.

One of the earliest applications of the Poisson model in biological sciences is Svedberg (1922), who compared counts of plants with expected frequencies calculated from the Poisson distribution. Today the Poisson process is conventionally used as a starting point in the construction of models, where one of the first steps is just testing of the hypothesis of complete spatial randomness, usually by means of the  $L$ -test of Section 4.2. If that null model is in agreement with the data, more complicated models are not worth trying. An impression of the generality of complete spatial randomness in forestry data is given by Tomppo (1986) who detected that around 30% of the permanent inventory plots in Finland, presenting thinning forests in mineral soil areas, could be considered as realizations of Poisson processes.

The *inhomogeneous Poisson process* with intensity function  $\lambda(\mathbf{x})$  is defined as follows:

1. The number of points in any region  $B$  follows the Poisson distribution with mean  $\int_B \lambda(\mathbf{x}) d\mathbf{x}$ .
2. Given  $n$  trees in  $B$ , their positions can be considered as an independent sample from the distribution with density  $\lambda(\mathbf{x}) / \int_B \lambda(\mathbf{x}) d\mathbf{x}$ .

This process is fundamental in modeling of intensity in heterogeneous environment: often some concomitant variables exist and are regressed with respect to the intensity; this is analogous to modeling of the hazard through concomitant variables in survival analysis.

A sophisticated nonparametric Bayesian approach to the estimation of spatial Poisson intensity is suggested by Heikkinen and Arjas (1998).

There, the intensity function is approximated by a piecewise constant function generated through the Voronoi tessellation of an “instrumental” Poisson process with positively correlated values in the cells. This spatial correlation plays the role of smoothing. The intensity is estimated from the posterior distribution using a Markov chain Monte Carlo (MCMC) sample of size  $N$ ,  $\hat{\lambda}^{(1)}(\mathbf{x}), \dots, \hat{\lambda}^{(N)}(\mathbf{x})$  of piecewise constant intensity estimators through averaging. The final estimator for the Poisson intensity is  $\hat{\lambda}(\mathbf{x}) = (1/N) \sum_{m=1}^N \hat{\lambda}^{(m)}(\mathbf{x})$ . Heikkinen and Arjas (1999) give a forestry application with a partially observed concomitant variable to data earlier analyzed by Rathbun (1996), who used kriging in the interpolation of the covariate.

## 6.2 Cox Processes

A very flexible and popular class of point process models for forestry are Cox processes. A Cox process can be seen as the result of a two-stage random mechanism; therefore the name “doubly stochastic Poisson process” is frequently used. In the first step a nonnegative function  $\lambda(\mathbf{x})$  is sampled, and then the inhomogeneous Poisson process with intensity function  $\lambda(\mathbf{x})$  is generated.

In recent years, Cox processes of increasing complexity and realism have been suggested. Early models were Poisson cluster processes, for example, *Matérn's cluster process*. Here  $\lambda(\mathbf{x})$  is built by means of the points  $\mathbf{x}_1, \mathbf{x}_2, \dots$  of a homogeneous Poisson process of intensity  $\lambda_p$ ,

$$(13) \quad \lambda(\mathbf{x}) = \mu \sum_{i=1}^{\infty} \mathbf{1}_{b(\mathbf{x}_i, R)}(\mathbf{x}),$$

where  $b(\mathbf{x}_i, R)$  is the disc with radius  $R$  centered at  $\mathbf{x}_i$ . The points of the Cox process appear here in independent (and possibly superposed) clusters centered at the  $\mathbf{x}_i$ 's. The model parameters are  $\lambda_p$ ,  $\mu$  and  $R$ . This model had appeared already in Matérn (1960) and was later used in several papers, perhaps not always very successfully; see the critique in Møller, Syversveen and Waagepetersen (1997, 1998).

The random intensity function  $\lambda(\mathbf{x})$  in Matérn's cluster process is a piecewise constant function, which cannot be more than a rough and discontinuous approximation to any unknown continuous intensity function; in forests one would expect smoother intensity functions. Such a function appears in the *Thomas process*, where

$$(14) \quad \lambda(\mathbf{x}) = \mu \sum_{i=1}^{\infty} \phi_{\sigma}(\mathbf{x} - \mathbf{x}_i),$$

and  $\phi_\sigma$  is the density function of the two-dimensional symmetric normal distribution with variance  $\sigma^2$ . This model and its generalizations are applied to forestry data in Diggle (1983), Cressie (1993) and Stoyan and Stoyan (1996).

A similar process appeared in the forestry literature as an ad hoc model; see Lepš and Kindlmann (1987):

$$(15) \quad \lambda(\mathbf{x}) = \lambda_0 \min \left\{ 1, \sum_{i=1}^n \exp(-a \|\mathbf{x} - \mathbf{x}_i\|^2) \right\}.$$

The  $n$  points  $\mathbf{x}_i$  here are random or deterministic points. The particular form results from the interpretation of the min-term as a probability; the Cox process is seen as the result of a location-dependent thinning of a stationary Poisson process of intensity  $\lambda_0$ , where the min-term is the retaining probability of a point at location  $\mathbf{x}$ .

In the (mathematical) statistical literature other models have appeared recently and have been applied to forestry problems. Ickstadt and Wolpert (1997) and Wolpert and Ickstadt (1998) studied the case

$$(16) \quad \lambda(\mathbf{x}) = \int \phi_\sigma(\mathbf{x} - s) \Gamma(ds),$$

where  $\Gamma$  is a random Gamma measure, a particular completely random measure. This model can be seen as a generalized nonstationary Thomas process. While a usual cluster process can be interpreted as a model where all clusters have the same weight, in the model of Wolpert and Ickstadt the clusters have random weights.

Møller, Syversveen and Waagepetersen (1997, 1998) suggested the “Log Gaussian Cox process,” where

$$(17) \quad \lambda(\mathbf{x}) = \exp(Z(\mathbf{x}))$$

with a Gaussian random field  $\{Z(\mathbf{x})\}$ . This random field may represent unobserved growth conditions while the exponential link function models the response of the trees.

For the cases shown in (13) to (17), which belong to stationary processes, intensity and pair correlation functions can be calculated analytically. Wolpert and Ickstadt (1998); Møller, Syversveen and Waagepetersen (1998) as well as Heikkinen and Arjas (1998, 1999) showed how these models can be used in the context of hierarchical modeling in Bayesian statistics for inhomogeneous Poisson processes. Using MCMC methods, they came to a posteriori estimates of  $\lambda(\mathbf{x})$  for given forests.

We revisit again the paper by Rathbun and Cressie (1994). These authors model the appearance of young trees (additionally recruited into an

already existing population) by a Cox process. Its random intensity is

$$\lambda(\mathbf{x}) = \exp(\beta_0 + \beta_1 d(\mathbf{x}) + \beta_2 W(\mathbf{x}) + \Psi(\mathbf{x})).$$

Here  $\beta_0$ ,  $\beta_1$  and  $\beta_2$  are model parameters;  $d(\mathbf{x})$  is the distance of the location  $\mathbf{x}$  to the nearest path in the forest;  $W(\mathbf{x}) = \sum_{(i)} (1/r_i(\mathbf{x}))$ , where  $r_i(\mathbf{x})$  is the distance from  $\mathbf{x}$  to the  $i$ th nearest tree (under the already existing trees) and where the sum is over all trees within 30 m of  $\mathbf{x}$ ;  $\Psi(\mathbf{x})$  is a Gaussian random field. The model parameters were estimated by means of the EM algorithm where the unknown realization of  $\Psi$  is taken as missing data.

It should be noted that  $\{\lambda(\mathbf{x})\}$ , the random Poisson intensity, is a continuous parameter random field. Such fields are the objective of geostatistical methods, which offer a further possibility for point processes statistics. Another use of geostatistics in connection with a Cox process is given when concomitant variables, observed in a set of sampled locations, are regressed with the intensity; see Rathbun (1996).

All these Cox models have a weak point: they do not guarantee a hard-core distance between trees and are therefore too variable at short distances. In mature forests there is typically some degree of regularity at short distances and there usually is a positive hard-core distance.

### 6.3 Gibbs Processes

A widespread family of point process models are *Gibbs point processes* with pairwise interactions. In a bounded window  $W$  they are defined through the density function formula

$$(18) \quad f(\mathbf{x}) = C \exp \left\{ - \sum_{i < j} \phi(\|\mathbf{x}_i - \mathbf{x}_j\|) \right\}$$

for a point pattern  $\mathbf{x} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  of  $n$  points, where  $n$  is fixed (see, e.g., Ripley, 1977; Stoyan, Kendall and Mecke, 1995);  $C$  is a normalizing constant. There are more general models with random  $n$  and also Gibbs point processes in the whole  $\mathbb{R}^2$ . In these generalizations, the density (18) should be considered with respect to the homogeneous Poisson process with unit intensity.

The function  $\phi$ , called the pair potential function, is such that  $\phi(\|\mathbf{x}_i - \mathbf{x}_j\|) = 0$  if and only if the two points  $\mathbf{x}_i$  and  $\mathbf{x}_j$  are not mutually interacting. Note that if  $\phi(r) = 0$  for all  $r > 0$ , then a pattern of  $n$  independent uniform points in  $W$  is obtained.

One of the assets of Gibbs point processes in applications is that interpretation is possible in terms of interaction. Indeed, inspecting the pair-potential function, positive values of  $\phi(r)$  indicate inhibition

on a scale defined by  $r$  and values of magnitude zero mean vanishing interaction; negative values of  $\phi(r)$  show attraction at this  $r$ .

A simple example of a Gibbs process is the Strauss process (Strauss, 1975; Kelly and Ripley, 1976), for which

$$\phi(r) = \begin{cases} \gamma, & \text{if } r < R, \\ 0, & \text{otherwise.} \end{cases}$$

Accordingly, points with mutual distance less than a fixed radius  $R$  are neighbors (and interacting). If  $S(\mathbf{x})$  is the number of neighboring pairs, then (18) can be rewritten in the form

$$(19) \quad f(\mathbf{x}) = C \exp\{-\gamma S(\mathbf{x})\}.$$

This model holds only for nonnegative  $\gamma$  corresponding to inhibitive interaction. If  $\gamma = \infty$ , a hard core model results.

What is especially nice in applications of Gibbs processes is that these processes are easy to simulate. MCMC simulation methods such as the Metropolis–Hastings algorithm suggested by Geyer and Møller (1994) or spatial birth-and-death process (Ripley, 1977; Stoyan, Kendall and Mecke, 1995) are commonly used. Conditional simulation, for example, keeping some tree positions fixed, is straightforward.

The most important applications of Gibbs point processes in forestry are in modeling of short-range competition between trees. Therefore, in forthcoming forest simulators, Gibbs processes will play an important role. Tomppo (1986) is an example of the extensive use of Gibbs type processes in forestry. Other applications can be found in Ogata and Tanemura (1984, 1985), Penttinen, Stoyan and Henttonen (1992), Stoyan, Kendall and Mecke (1995) and Särkkä and Tomppo (1998). Inhomogeneous Gibbs processes where the point density follows a deterministic trend are also useful; see Ogata and Tanemura (1986), Stoyan and Stoyan (1998a), Jensen and Nielsen (1998) and Baddeley, Møller and Waagepetersen (1998).

When modeling point patterns by means of Gibbs processes, a parametric family of pair potentials is usually introduced. Then (18) serves as the likelihood function, but its optimization is not straightforward due to the unknown scaling factor  $C$ . This problem can be bypassed by applying the Markov chain Monte Carlo maximum likelihood method; see, for example, Diggle et al. (1994) and Geyer (1998). In any event, the use of the maximum likelihood method is computationally intensive. Alternative estimation methods such as the pseudolikelihood method proposed by Besag (1978) (see Ripley, 1988; Goulard, Särkkä and Grabarnik, 1996), as

the Takacs–Fiksel method (see Stoyan and Stoyan, 1994), can also be used. Nonparametric approaches have also been suggested for pair potential estimation by Diggle et al. (1994) and, in the Bayesian frame, by Heikkinen and Penttinen (1998).

Gibbs point processes have been successfully used as models for forests but, as noted earlier, this class of processes has difficulties in producing clustered patterns in sufficient variety as needed in forestry applications. Heuristically this means that the Gibbs models above do not take into account the limits of natural resources because the neighborhood property does not depend on the realization.

## 7. MODELS FOR MARKED POINT PROCESSES

In the preceding section, little was said about marks. The reader should remember that marks describe trees and may be qualitative or quantitative. Qualitative marks characterize, for example, tree species or their origin (natural or planted), while quantitative marks are characteristics such as dbh or degree of damage (if given in percentages).

An important particular case of qualitatively marked point processes are binary point processes, processes with two marks, 1 and 2. Two examples of binary point process models are *linked* and *balanced Cox processes*, where the random intensity functions of the subprocesses of 1-points and 2-points,  $\lambda_1(\mathbf{x})$  and  $\lambda_2(\mathbf{x})$ , satisfy

$$\lambda_1(\mathbf{x}) = a\lambda_2(\mathbf{x})$$

and

$$\lambda_1(\mathbf{x}) + \lambda_2(\mathbf{x}) = c,$$

respectively, with suitable positive constants  $a$  and  $c$ ; see Diggle and Milne (1983). The linked Cox process models positive association between the two process components, whereas the balanced Cox process models negative association. Møller, Syversveen and Waagepetersen (1998) suggest a bivariate log-Gaussian Cox process for which the component processes  $\log \lambda_j(\mathbf{x})$  are correlated Gaussian random fields,  $j = 1, 2$ .

The Gibbs process can be generalized to the binary point process case. If  $\{\mathbf{x}_i\}$  stands for the pattern of 1-points and  $\{\mathbf{y}_j\}$  for 2-points, then the density is of the form

$$f(\mathbf{x}, \mathbf{y}) = C \exp \left\{ - \sum_{i \neq k} \phi_{11}(\|\mathbf{x}_i - \mathbf{x}_k\|) - \sum_{j \neq l} \phi_{22}(\|\mathbf{y}_j - \mathbf{y}_l\|) - \sum_{i, j} \phi_{12}(\|\mathbf{x}_i - \mathbf{y}_j\|) \right\},$$

assumed to have  $n_j$  points of type  $j$  ( $j = 1$  and  $2$ ), and the  $\phi_{ij}(r)$ 's are interaction functions measuring interaction between pairs of points of type  $i$  and  $j$ , for  $i, j \in \{1, 2\}$ . Such processes are studied in Ogata and Tanemura (1985) for modeling forests. These authors show how to estimate parameters in parametrized pair potential functions via the maximum likelihood method.

The marks are frequently introduced as follows: an unmarked point process  $\Phi$  and an independent random field  $\{Z(\mathbf{x})\}$  are given. If the  $n$ th point of  $\Phi$  is located at  $\mathbf{x}_n$ , then its mark is

$$m(\mathbf{x}_n) = Z(\mathbf{x}_n),$$

that is, the marks come directly from the random field. This *random field mark model* may be a good approximation to marked point processes in forestry, in particular if  $\{Z(\mathbf{x})\}$  describes a tree variable which is in close correlation to environmental factors following a stationary and isotropic random field. For very short distances, the model is not correct because it ignores interaction of trees. The aim of statistical analysis for this model is, in addition to estimation of point process characteristics, to determine distributional characteristics of the random field. This can be done by geostatistical methods, since for this model the mark variogram introduced in Section 4.4 equals the variogram of  $\{Z(\mathbf{x})\}$ . If the random field mark model does not hold, the mark variogram may have a form which is different from that of a variogram of a random field; for example, it may not be negative definite; see Walder and Stoyan (1996) and Kuuluvainen, Penttinen, Leinonen and Nygren (1996).

The random field model can be refined by replacing the unmarked point process by a binary point process, where the marks describe the social positions of trees, dominant or suppressed. Then the marks may be determined by

$$m(\mathbf{x}_n^i) = f_i(Z(\mathbf{x}_n^i)) \quad \text{for } i = 1, 2$$

with suitable functions  $f_i$ . This models the extent of resource exploitation of the environment which trees are able to carry out depending on their social position; see Stoyan and Walder (1999). It is an important statistical problem to investigate the random field given the marks.

Marked Gibbs processes with quantitative marks are not yet frequently used. A good example of a forestry-oriented paper is Goulard, Sarkka and Grabarnik (1996). They fit a marked Gibbs process model to the spruces of Figure 2, where the marks are the dbh's; the case without marks was considered in Stoyan, Kendall and Mecke (1995).

Statistics and simulation of marked Gibbs processes are much more complicated than in the unmarked case since not only the pair potential (which, of course, depends on the marks) but also a "primary" mark distribution, which cannot be measured directly, has to be determined and used. The observed marks belong to another distribution. Here, the Gibbs process density is a density with respect to an independently marked Poisson process.

A trick used by Ogata and Tanemura (1985) can be applied in the estimation of parameters of pair potential of the form

$$\phi\left(\frac{\|\mathbf{x} - \mathbf{y}\|}{w(m(\mathbf{x}), m(\mathbf{y}))}\right)$$

with a suitable scaling function  $w$ . Degenhardt (1999) applied this method and obtained ecologically interpretable pair potential parameters for various German forests.

## 8. SUMMARY AND CONCLUSIONS

As the paper has shown, point processes and marked point processes are powerful tools in modern forestry statistics. The last decades have seen a rapid development of theoretical statistical methods and stochastic models for point processes, which is to some extent inspired solely by forestry applications. Of particular importance are the modern single-tree models, which have been developed mainly by forest scientists who now learn that they need the support of point process statisticians.

The recent development inspires one also to think that the borders between separate areas of spatial statistics, relevant for forest science, will fade away. Maybe we have not been farsighted when speaking, for example, about point process statistics and geostatistics as different disciplines for problem solving. Instead, these topics will appear as complementary, to supporting building blocks in modeling; see the discussion on Cox processes in the recent paper.

Some problems discussed in this paper will remain important research themes also in the future. One of them is modeling of inhomogeneous or spatially nonstationary forests and the development of adequate statistical characteristics and methods. A particular question is to determine the form by which the local interaction of trees (e.g., described by a pair potential depending on species and age) is influenced by factors (such as soil, humidity or environment) that may generate global inhomogeneity.



A further problem is automation of production of tree position maps and the full use of the possibilities of GPS and GIS (geographical information systems). Stable and fast statistical and image-analytical methods for the construction of such maps based on blurred and noisy images of forests have to be developed.

Finally, stochastic modeling has to be combined with biological ideas with the aim of obtaining realistic variability predictions based on incomplete measurements. Particularly important is good statistical support in modeling two essential processes in forest development that are probably highly influenced by random factors: mortality and regeneration.

In all these directions (and various others), forestry researchers need the help of professional spatial statisticians. Unfortunately, foresters do not always know or accept this and develop their own methods; at least in Europe, forestry research concerning problems of spatial statistics is rather conservative. On the other hand, the statistical community should note that much of published statistical work is not accessible by, or intelligible to, foresters.

## ACKNOWLEDGMENTS

The authors thank Jan Lepš very much for many valuable hints and remarks on an early draft of this paper. They also thank Dietrich Quednau and Hans Pretzsch for many comments on a lecture by the first author in Weihenstephan in May 1997 and later discussion. Helmut Stoyan's remarks on the first chapters have improved them. Furthermore, we thank Minna Pulkkinen for helpful comments; her forestry expertise has been very valuable. Two anonymous referees and the Editors greatly improved the paper by their constructive comments.

This work was supported by the Deutsche Forschungsgemeinschaft and by the Academy of Finland (projects 39372 and 37206).

## REFERENCES

- AVERY, T. E. and BURKHART, H. E. (1983). *Forest Measurements*, 3rd ed. McGraw-Hill, New York.
- BADDELEY, A. J. (1998). Spatial sampling and censoring. In *Stochastic Geometry: Likelihood and Computation* (O. E. Barndorff-Nielsen, W. S. Kendall and M. N. M. van Lieshout, eds.) 37–78. CRC Press/Chapman and Hall, London.
- BADDELEY, A. J. and GILL, R. D. (1997). Kaplan–Meier estimators of distance distributions for spatial point processes. *Ann. Statist.* **25** 263–292.
- BADDELEY, A. J., MØLLER, J. and WAAGEPETERSEN, R. (1998). Non- and semiparametric estimation of interaction in inho-

- mogeneous point patterns. R-98-2019, Dept. Mathematics, Inst. Electronic Systems, Univ. Aalborg.
- BESAG, J. (1978). Some methods of statistical analysis for spatial data. *Bull. Internat. Statist. Inst.* **47** 77–92.
- BIGING, G. S. and DOBBERTIN, M. (1992). A comparison of distance-dependent competition measures for height and basal area growth of individual conifer trees. *Forest Science* **38** 695–720.
- BIGING, G. S. and DOBBERTIN, M. (1995). Evaluation of competition indices in individual tree growth models. *Forest Science* **41** 360–377.
- BITTERLICH, W. (1948). Die Winkelzählprobe. *Allgemeine Forst- und Holzwirtschaftliche Zeitung* **59** 4–5.
- BITTERLICH, W. (1952). Die Winkelzählprobe. *Forstwissenschaftliches Centralblatt* **71** 215–225.
- CANNON, A. and CRESSIE, N. (1997). Temporal analogues to spatial  $K$  functions. *Biometrical J.* **37** 351–373.
- CHIU, S. N. and STOYAN, D. (1998). Estimates of distance distributions for spatial patterns. *Statist. Neerlandica* **52** 239–246.
- CLARK, P. J. and EVANS, F. C. (1954). Distance to the nearest neighbor as a measure of spatial relationships in populations. *Ecology* **35** 445–453.
- COX, F. (1971). Dichtebestimmung und Strukturanalyse von Pflanzenpopulationen mit Hilfe von Abstandsmessungen. *Mitt. Bundesforsch. Anst. Forst- und Holzwirtsch., Reinbek b. Hamburg* **87** 1–184.
- CRESSIE, N. (1993). *Statistics for Spatial Data*, rev. ed. Wiley, New York.
- DEGENHARDT, A. (1999). Description of tree distribution and their development through marked Gibbs processes. *Biometrical J.* **41** 457–470.
- DIGGLE, P. J. (1983). *Statistical Analysis of Spatial Point Patterns*. Academic Press, London.
- DIGGLE, P. J. and CHETWYND, A. G. (1993). Second-order analysis of spatial clustering for inhomogeneous populations. *Biometrics* **47** 1153–1163.
- DIGGLE, P. J., FIKSEL, T., GRABARNIK, P., OGATA, Y., STOYAN, D. and TANEMURA, M. (1994). On parameter estimation for pairwise interaction point processes. *Internat. Statist. Rev.* **92** 99–117.
- DIGGLE, P. J. and MILNE, R. K. (1983). Bivariate Cox processes: some models for bivariate spatial point-patterns. *J. Roy. Statist. Soc. Ser. B* **45** 11–21.
- DRALE, K. and RUDEMO, M. (1996). Stem number estimation by kernel smoothing of aerial photos. *Canad. J. For. Res.* **26** 1228–1236.
- FÜLDNER, K., SÄTTLER, S., ZUCCHINI, W. and VON GADOW, K. (1996). Modellierung personenabhängiger Auswahlwahrscheinlichkeiten bei der Durchforstung. *Allg. Forst- und Jagdzeitung* **167** 159–162.
- GAVRIKOV, V. L. and STOYAN, D. (1995). The use of marked point processes in ecological and environmental forest studies. *Environ. Ecol. Statist.* **2** 331–344.
- GEYER, C. (1998). Likelihood inference for spatial point processes. In *Stochastic Geometry: Likelihood and Computation* (O. E. Barndorff-Nielsen, W. S. Kendall and M. N. M. van Lieshout, eds.) 79–40. CRC Press/Chapman and Hall, London.
- GEYER, C. J. and MØLLER, J. (1994). Simulation and likelihood inference for spatial point processes. *Scand. J. Statist.* **21** 359–373.
- GOULARD, M., SÄRKKÄ, A. and GRABARNIK, P. (1996). Parameter estimation for marked Gibbs point processes through the maximum pseudo-likelihood method. *Scand. J. Statist.* **23** 365–379.

- HEIKKINEN, J. and ARJAS, E. (1998). Nonparametric Bayesian estimation of a spatial Poisson intensity. *Scand. J. Statist.* **25** 435–450.
- HEIKKINEN, J. and ARJAS, E. (1999). Modeling a Poisson forest in variable elevation: a nonparametric Bayesian approach. *Biometrics* **55** 738–745.
- HEIKKINEN, J. and PENTTINEN, A. (1999). Bayesian smoothing in the estimation of the pair potential function of Gibbs point processes. *Bernoulli* **5**. To appear.
- ICKSTADT, K. and WOLPERT, R. L. (1997). Multiresolution assessment of forest inhomogeneity. *Case Studies in Bayesian Statistics. Lecture Notes in Statist.* **3** 371–386. Springer, Berlin.
- JENSEN, E. B. V. and NIELSEN, L. S. (1998). Inhomogeneous Markov point processes by transformation. Research Report 2, Lab. Computational Stochastics, Dept. Mathematical Sciences, Univ. Aarhus.
- KELLY, F. P. and RIPLEY, B. D. (1976). A note on Strauss's model for clustering. *Biometrika* **63** 357–360.
- KIMMINS, J. P. (1993). Scientific foundations for the simulation of ecosystem function and management in FOTCYTE-11. Information Report NOR-X-328, Forestry Canada, Northwest Region Northern Forestry Centre.
- KÖNIG, G. (1835). *Die Forst-Mathematik*. Beckersche Buchhandlung, Gotha.
- KREBS, C. J. (1989). *Ecological Methodology*. HarperCollins, New York.
- KUULUVAINEN, T., PENTTINEN, A., LEINONEN, K. 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 Fennica* **30** 315–328.
- LEPŠ, J. (1990). Can underlying mechanisms be deduced from observed patterns? In *Spatial Processes in Plant Communities* (F. Krahulec, A. D. Q. Agnew and J. H. Willems, eds.) 1–11. Academia, Prague.
- LEPŠ, J. and KINDLMANN, P. (1987). Models of the development of spatial pattern of an even-aged plant population over time. *Ecological Modelling* **39** 45–57.
- LIU, J. and ASHTON, P. S. (1995). Individual-based simulation models for forest succession and management. *Forest Ecology and Management* **73** 157–175.
- MATÉRN, B. (1960). Spatial variation. *Meddelanden från Statens Skogsforskningsinstitut* **49** 1–144. *Lecture Notes in Statist.* **36**. Springer, Berlin. [2nd ed. (1986)]
- MØLLER, J., SYVERSVEEN, A. R. and WAAGEPETERSEN, R. P. (1997). Log Gaussian Cox Processes: a statistical model for analysing stand structural heterogeneity in forestry. In *Proceedings First European Conference for Information Technology in Agriculture* (H. Kure, I. Thysen and A. R. Kristensen, eds.) 339–342. Dept. Mathematics and Physics, Royal Veterinary and Agricultural Univ., Denmark.
- MØLLER, J., SYVERSVEEN, A. R. and WAAGEPETERSEN, R. P. (1998). Log Gaussian Cox processes. *Scand. J. Statist.* **25** 451–482.
- NEWMHAM, R. M. (1964). The development of a stand model for Douglas-fir. Ph.D. thesis, Univ. British Columbia, Vancouver.
- ODUM, E. P. (1971). *Fundamentals of Ecology*. Saunders, Philadelphia.
- OGATA, Y. and TANEMURA, M. (1984). Likelihood analysis of spatial point-patterns. *J. Roy. Statist. Soc. Ser. B* **46** 496–518.
- OGATA, Y. and TANEMURA, M. (1985). Estimation of interaction potentials of marked spatial point-patterns through the maximum-likelihood method. *Biometrics* **41** 315–338.
- OGATA, Y. and TANEMURA, M. (1986). Likelihood estimation of interaction potentials and external fields of inhomogeneous spatial point patterns. In *Proceedings Pacific Statistical Congress 1985* (I. S. Francis, B. F. J. Manly and F. C. Lam, eds.) 150–154.
- OHSEER, J. and STOYAN, D. (1981). On the second-order and orientation analysis of planar point processes. *Biometrical J.* **23** 523–533.
- OVERTON, W. S. and STEHMAN, S. V. (1995). The Horvitz–Thompson theorem as a unifying perspective for probability sampling: with examples from natural resource sampling. *Amer. Statist.* **49** 261–268.
- PACALA, S. W., CANHAM, C. D., SAPONARA, J., SILSANDER, J. A., KOBE, R. K. and RIBBENS, E. (1996). Forest models defined by field measurements: estimation, error analysis and dynamics. *Ecological Monogr.* **66** 1–43.
- PENTTINEN, A. (1988). A random field approach to Bitterlich sampling. *Ann. Acad. Sci. Fenn. Ser. A I Math.* **13** 259–268.
- PENTTINEN, A. K., STOYAN, D. and HENTTONEN, H. M. (1992). Marked point processes in forest statistics. *Forest Sci.* **38** 806–824.
- PIELOU, E. C. (1977). *Mathematical Ecology*. Wiley, New York.
- POHTILA, E. (1980). Spatial distribution development in young tree stands in Lapland. *Comm. Inst. Forestalis Fennicae* **98** 1–35.
- PRETZSCH, H. (1993). *Analyse und Reproduktion räumlicher Bestandesstrukturen. Versuche mit dem Strukturgenerator STRUGEN. Schriften aus der Forstlichen Fakultät der Universität Göttingen und der Nieders. Forstl. Versuchsanstalt* **114**. J. D. Sauerländers, Frankfurt.
- PRETZSCH, H. (1997). Analysis and modeling of spatial stand structures. Methodological considerations based on mixed beech-larch stand in Lower Saxony. *Forest Ecol. Manag.* **97** 237–253.
- RATHBUN, S. L. (1996). Estimation of Poisson intensity using partially observed concomitant variables. *Biometrics* **52** 226–242.
- RATHBUN, S. L. and CRESSIE, N. (1994). A space-time survival point process for a longleaf pine forest in Southern Georgia. *J. Amer. Statist. Assoc.* **89** 1164–1174.
- RIPLEY, B. D. (1976). The second-order analysis of stationary point processes. *J. Appl. Probab.* **13** 255–266.
- RIPLEY, B. D. (1977). Modelling spatial patterns (with discussion). *J. Roy. Statist. Soc. Ser. B* **39** 172–212.
- RIPLEY, B. D. (1981). *Spatial Statistics*. Wiley, New York.
- RIPLEY, B. D. (1988). *Statistical Inference for Spatial Processes*. Cambridge Univ. Press.
- SALONEN, V., PENTTINEN, A. and SÄRKKÄ, A. (1992). Plant colonization of a bare peat surface: population changes and spatial patterns. *J. Vegetation Sci.* **3** 113–118.
- SÄRKKÄ, A. and TOMPPA, E. (1998). Modelling interactions between trees by means of field observations. *Forest Ecol. Manag.* **108** 57–62.
- SCHLATHER, M. (1999). On the second order characteristics of marked point processes. *Bernoulli*. To appear.
- SCHREUDER, H. T., GREGOIRE, T. G. and WOOD, G. B. (1993). *Sampling Methods for Multiresource Forest Inventory*. Wiley, New York.
- SERRA, J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press, London.
- STOYAN, D., KENDALL, W. S. and MECKE, J. (1995). *Stochastic Geometry and Its Applications*, 2nd ed. Wiley, Chichester.
- STOYAN, D. and STOYAN, H. (1994). *Fractals, Random Shapes and Point Fields*. Wiley, Chichester.
- STOYAN, D. and STOYAN, H. (1996). Estimating pair correlation functions of planar cluster processes. *Biometrical J.* **38** 259–271.

- STOYAN, D. and STOYAN, H. (1998a). Non-homogeneous Gibbs process models for forestry—a case study. *Biometrical J.* **40** 521–531.
- STOYAN, D. and STOYAN, H. (1998b). Improving ratio estimators of point process statistics. Research report, TU Bergakademie, Freiberg.
- STOYAN, D. and WÄLDER, O. (1999). On variograms in point process statistics II. *Biometrical J.* To appear.
- STRAUß, D. J. (1975). A model for clustering. *Biometrika* **63** 467–475.
- SVEDBERG, T. (1922). Ett bidrag till de statistiska metodernas användning inom växtbiologien. *Svensk. Botanisk. Tidskrift* **16** 1–8.
- TILMAN, D. (1988). *Plant Strategies and the Dynamics and Structure of Plant Communities*. Princeton Univ. Press.
- TOMPPÖ, E. (1986). Models and methods for analysing spatial patterns of trees. *Comm. Inst. Forestalis Fennicae* **138**.
- TURNER, M. G. and GARDNER, R. H. (1991). *Quantitative Methods in Landscape Ecology*. Springer, Berlin.
- VAN LIESHOUT, M. N. M. and BADDELEY, A. J. (1996). A non-parametric measure of spatial interaction in point-patterns. *Statist. Neerland.* **50** 344–361.
- WÄLDER, O. and STOYAN, D. (1996). On variograms in point process statistics. *Biometrical. J.* **38** 395–905.
- WARREN, W. G. (1972). Point processes in forestry. In *Stochastic Point Processes, Statistical Analysis, Theory and Application* (P. A. W. Lewis, ed.) 85–116. Wiley, New York.
- WOLPERT, R. L. and ICKSTADT, K. (1998). Poisson/Gamma random field models for spatial statistics. *Biometrika* **85** 251–267.