ELSEVIER

Ecological Modelling 108 (1998) 163–174

ECOLOGICAL MODELLING

The spatial pattern of a forest ecosystem

J. Mateu [a,*], J.L. Usó [a], F. Montes [b]

[a] *Department of Mathematics, Campus Penyeta Roja, University Jaume I of Castellón, Spain*
[b] *Department of Statistics and Operation Research, University of Valencia, Valencia, Spain*

Accepted 17 March 1997

Abstract

Statistical analysis of stands of trees as a whole need suitable methods of spatial statistics. Obviously, trees within a stand affect development and survival of their neighbours. They interact and therefore have to be considered as a system of dependent random variates from an unknown stochastic process. One such statistical model which considers the spatial dependence among trees in a forest and their characteristics is a marked point process. The 'points', called events in spatial statistics, are the tree positions and the 'marks' are tree characteristics such as crown lengths or tree species. A minimal prerequisite for any serious attempt to model an observed pattern is to test the hypothesis of complete spatial randomness (CSR). Concerning the fitting of parametric models to spatial point patterns, a class of models which seems potentially useful for describing the present type of data is the class of marked Gibbs (pairwise interaction) point processes. Essentially, these processes characterise the interaction between events by some parametrically specified function of distance. In this paper several statistical methods to test CSR are described and marked Gibbs processes are used to fit a model in two different forest ecosystems.

Keywords: Complete spatial randomness; F-function; G-function; Independence; K-function; Marked Gibbs processes; Pseudo-likelihood; Random labelling

1. Introduction

A spatial point pattern is a set of data $\{(x_i, y_i)\ i = 1,\ldots,n\}$ consisting of n locations in an essentially planar region. Examples include the locations of cell nuclei in a microscopic tissue section, trees in a forest, or cases of disease in a geographical region. A fundamental assumption in the analysis of such data is that they can usefully be regarded as a partial realisation of a stochastic point process (Cox and Isham, 1980). Many systems of individuals can also be described by attaching to the locations measurable quantities m_x

* Corresponding author. Tel.: +34 964 345775; fax: +34 964 345847.

0304-3800/98/$19.00
PII S0304-3800(98)00027-1

like, say, diameter of a tree. This last case is an example of marked points patterns.

The concept of complete spatial randomness (CSR) is fundamental to the quantitative description of a spatial pattern. We refer to the positions of the trees in question as events, to distinguish them from arbitrary points in the region of observation A. A formal definition of CSR is that the events in A constitute a partial realisation of a homogeneous, planar Poisson process (Diggle, 1983). This process incorporates a single parameter, λ, the intensity, or mean number of events per unit area. The actual number of events in A, n say, is an observation from a Poisson distribution with mean $\lambda|A|$, where $|A|$ denotes the area of the region A. If we consider n as fixed, we arrive at the following definition of CSR: (1) each of the n events is equally likely to occur at any point within A; (2) the n events are located independently of each other. Our interest in CSR is that it represents an idealized standard which, if strictly unattainable in practice, may nevertheless be tenable as a convenient first approximation. Most analysis begin with a test of CSR, and there are several good reasons for this. Firstly, a pattern for which CSR is not rejected scarcely merits any further formal statistical analysis. Secondly, tests are used as a means of exploring a set of data, rather than because rejection of CSR is of intrinsic interest. It is emphasized that ecologists often know CSR to be untenable but nevertheless use tests of CSR as an aid to the formulation of ecologically interesting hypotheses concerning pattern and its genesis. Thirdly, CSR acts as a dividing hypothesis to distinguish between patterns which are broadly classifiable as 'regular' or 'aggregated'. It is important to emphasize the value of graphical methods for these will almost always be informative and will sometimes make formal testing unnecessary.

In recent years, there has been considerable interest in studying the spatial pattern presented by different types of trees coexisting in the same region. A bivariate spatial point pattern is one in which the events are of two distinguishable types. One possible benchmark hypothesis for the assessment of interactions between the two types is that the two component patterns are determined independently of one another. If, on the other hand, a bivariate pattern arises through some form of labelling mechanism the question of whether or not the component patterns are statistically independent is not relevant. A more natural benchmark hypothesis is that the two component patterns are formed by random labelling, by which we mean that events are labelled independently, each event being labelled type 1 with probability P, and type 2 with probability $1-P$. Independence and random labelling are in general statistically distinct hypotheses: they coincide if and only if the superposition of the two types of event forms a completely random pattern, in which case both component patterns are also completely random.

In the analysis of point patterns the statistical modelling of locations of interacting individuals (trees in a forest) is studied in terms of Gibbs point processes. Gibbs point processes first appeared in the theory of statistical physics, where Gibbs distributions are applied to describe the equilibrium states of closed physical systems of interacting objects. Gibbs models are constructed from local interactions between the individuals. In this paper pairwise interaction processes will be considered, where only interactions among pair of objects are taken into account and a parametrized pair potential function is used to describe interactions between individuals. If the objects are marked, marked pairwise interaction processes will be applied where the pair potential function depends on marks, too. The parameters can be estimated by using some optimization method such as maximum likelihood estimation. The major problem in the application of this method is that the likelihood function contains an unknown scaling factor which depends on parameters of the pair potential function and which is intractable. Hence, the maximum likelihood method cannot be applied straightforwardly. In this paper, the pseudo-likelihood estimation method will be used.

The hypothesis of CSR, independence and random labelling will be analyzed in Section 2. The use of a Gibbs model and the estimation method will be presented in Section 3. Finally, in Section 4 two forest ecosystems will be studied. The first ecosystem consists of a forest of two types of

mediterranean trees, *Quercus ilex L.* and *Pinus halepensis L.*, coexisting in the same area. The second one is concerned with a birch forest area in which locations and crown lengths are recorded.

2. Spatial statistical methods: identification of pattern

For testing an observed pattern against a benchmark hypothesis, in this case to test for departure from complete spatial randomness, several useful empirical functions are the empirical distribution functions of nearest neighbour distances and point to nearest event distances (Diggle, 1983). It will be used an informal combination of them to indicate the nature of any departure from CSR.

For n events in a region A, let d_i denote the distance from the ith event to the closest of the other $n-1$ events. These distances are called nearest neighbour distances, and typically include duplicate measurements between reciprocal nearest neighbour pairs. The empirical distribution function (EDF), $\hat{G}_1(s)$ say, represents the observed proportion of nearest neighbour distances d_i that are at most s,

$$\hat{G}_1(s) = n^{-1}\phi(d_i \leq s)$$

where ϕ means 'the number of'. For a completely random pattern, the underlying theoretical distribution function is given by

$$G(s) = 1 - \exp(-\pi\lambda s^2)$$

where λ is the intensity. To assess the significance or otherwise of departures from CSR, the conventional approach would be to find the sampling distribution of $\hat{G}_1(s)$ under CSR, but this is complicated. We therefore proceed as follows. Calculate EDFs, from $\hat{G}_i(s)$, $i = 2,\ldots,ns$, from each of $ns-1$ independent simulations of n events independently and uniformly distributed on A, and define upper and lower simulation envelopes,

$$U(s) = \max_{i=2,\ldots,ns}\{\hat{G}_i(s)\};$$

$$L(s) = \min_{i=2,\ldots,ns}\{\hat{G}_i(s)\}$$

Under CSR, and for each s, $P\{\hat{G}_1(s) > U(s)\} = P\{\hat{G}_1(s) < L(s)\} = ns^{-1}$. The graphical procedure then consists of plotting $\hat{G}_1(s)$, $U(s)$ and $L(s)$ against $G(s)$.

Similarly, for a grid of k sampling origins, let e_i denote the distance from the ith origin to the closest of the n points. Then, the EDF for point to nearest event distances is given by

$$\hat{F}_1(s) = k^{-1}\phi(e_i \leq s)$$

Under CSR, the theoretical distribution function $F(.)$ matches exactly $G(.)$ given above. By analogy with the procedure adopted for the nearest neighbour distances, upper and lower envelopes must be calculated based on a number of simulations under CSR.

An exact Monte-Carlo test of complete spatial randomness can be constructed as follows. Calculate

$$D = \max_{i=1,\ldots,ns}\left|\hat{F}(s_i) - \hat{G}(s_i)\right|$$

The value, D_1 say, of D for the observed data is compared with values D_2, …, D_{100} from $ns-1$ simulations of CSR. If D_1 ranks pth largest amongst D_2, …, D_{100}, the attained significance level of the test is $p\%$ (Barnard, 1963).

The second-order properties of a spatial point process describe variation in the relative frequency of pairs of events as a function of their positions. Under the assumption of constant local intensity, this function depends only on the relative positions of the two events. Under the further assumption that the underlying process involves no directional effects, it reduces to a function of distance only. Perhaps the most useful such function is the K-function (Ripley, 1977) defined by,

$$\lambda K(s)$$

= mean number of events within distance s of an arbitary event (excluding the arbitrary event).

For a bivariate process, in which the local intensities of type 1 and type 2 events are λ_1 and λ_2 respectively, a complete description of the second-order properties of the process requires us to define the K-functions $K_{ij}(s)$ by,

$\lambda_j K_{ij}(s)$ = mean number of type j events wihtin distance s of an arbitrary type i event.

Note that $K_{11}(s)$ and $K_{22}(s)$ correspond to $K(s)$ as previously defined, also that $K_{12}(s)=K_{21}(s)$.

The forms of these K-functions under our various benchmark hypothesis are very simple. Under CSR, $K(s)=\pi s^2$. Under independence, $K_{12}(s)=\pi s^2$, irrespective of the forms of $K_{11}(s)$ and $K_{22}(s)$. Finally, under random labelling, $K_{11}(s)=K_{22}(s)=K_{12}(s)$.

3. Marked Gibbs models

Gibbs point processes are applied to model point patterns with interacting objects. The theory first arose in the theory of statistical physics, where Gibbs distributions are use to describe the equilibrium states of closed physical systems. The theory of Gibbs processes is based on the idea of sampling realizations from a weight process, usually from a Poisson process, following a special type of probability density. Therefore the distribution of a Gibbs process is defined as the Radon–Nikodym derivative of its distribution with respect to a Poisson process with intensity one. Poisson process is chosen as a weight process because it is natural to compare point patterns with interacting events to a completely random one. The definitions in next subsection are given to marked Gibbs processes either on $\mathscr{R}^d\times K$ or on $W\times K$, where W is a bounded subset of $\mathscr{R}^d$.

3.1. Interaction

Suppose that $\perp$ is a symmetric relation on W. Two points $\xi=[x,m_x]$ and $\xi=[y,m_y]$ on $W\times K$ are said to be neighbours if $x\perp y$ implies that $y\perp x$, and in addition x can not be neighbour of itself. In many applications individuals are neighbours if they are close enough each other. For a set $A\subset W\times K$ we define the neighbourhood V(A) by setting

$$V(A)=\{[x;m_x]|x\perp y \text{ for some } [y;m_y]\in A\}.$$

Let X^K denote the family of all locally finite counting measures on. A subset H of X^K is hereditary if $\psi_1\in H$ and $\psi_2\subset\psi_1$ implies that $\psi_2\in H$ (Ripley and Kelly, 1977). This condition has been presented independently in the construction of Markov random fields known as vacuum condition (Averintshev, 1973).

A function $f:X^K\to[0,\infty)$ is a (mark) Markov function if $H=\{f>0\}$ is hereditary and whenever $\psi\in H$ and $\xi\in\psi$ and the expression

$$\frac{f(\psi\cup\xi)}{f(\psi)+f(\psi\cup\xi)}$$

depends only on $\xi=[x,m_x]$ and $V(\xi)\cup\psi$.

A function $h:X^K\to[0,\infty)$ is a (mark) interaction function if $h(\psi)\neq 1$ implies that ξ and η are neighbours for every $\xi,\eta\in\psi$. A Markov function can also be characterized in terms of interaction.

3.1.1. Factorizing theorem (Ripley and Kelly, 1977)

A function $f:X^K\to[0,\infty)$ is a (mark) Markov function if and only if there is an interaction function h such that

$$f(\psi_1)=\prod_{\psi_2\subset\psi_1}h(\psi_2)$$

for all $\psi_1\in X^K$.

The factorizing theorem is known as Hammersley–Clifford theorem for Markov random fields (Grimmet, 1973).

3.2. Marked Gibbs processes

As stated above, when introducing the marked Gibbs point process model it is necessary to make a difference between a finite point process defined on a set $W\times K$, where W is a bounded subset of $\mathscr{R}^d$ and K is a given mark space, and an infinite (stationary) point process on $\mathscr{R}^d\times K$. The finite case is simpler: given a potential function, we only have to check that the probability measure is integrable. However, finite processes cannot be stationary due to boundary effects. On the other hand, in the infinite case, in addition to checking the integrability of the probability measure, it is necessary to prove the existence of the process by means of considering limits of some thermodynamical functions (see e.g. Ruelle, 1969; Preston, 1976).

As a compromise between these two approaches we consider point processes in the space $S=T\times K$, where T is a torus (obtained by identi-

fying opposite sides of a hyperrectangle $A \subset \mathscr{R}^d$). In applications, a rectangular sampling window can be mapped onto a torus and the mark space K is usually either a discrete set or a part of the real line.

For ease of notation, let's consider that X^K also denote the family of all locally finite counting measures ψ on S and F_S the smallest σ-algebra on X^K generated by the maps $\psi \to \psi(W \times L)$, where $\psi(W \times L)$ is the number of points of in a bounded Borel set W with marks in L. A marked point process Ψ is a random variable in X^K. The distribution of Ψ is denoted by P. For convenience, we write $\Psi = \{[x_i, m_i]\}$, where the x_i's are the locations of points of the process and the m_i's are the marks.

A marked point process on a torus is stationary if its distribution is invariant under shifts on the torus: the processes $\Psi = \{[x_i, m_i]\}$ and $\Psi_x = \{[x_i \oplus x, m_i]\}$ have the same distribution for all x in T, $\oplus$ denoting a shift on the torus. The distance between the points x and y defined on the torus is denoted by $d(x, y)$.

Further on we consider stationary marked point processes on S. For such processes the intensity measure has the form

$$\Lambda p = \lambda_p v \times M_p$$

where λ_p is the intensity of the process Ψ, v is the Lebesgue measure and M_p is the mark distribution of the process.

Let us fix a finite non-atomic measure τ on T (typically τ is the Lebesgue measure, yielding the unmarked Poisson process of unit intensity) and a measure M on K corresponding to the distribution of the marks. Given $\rho = \tau \times M$ we define the marked Poisson measure Q with intensity ρ on (X^K, F_S) as

$$Q(F) = e^{-\rho(S)} \sum_{n=0}^{\infty} \frac{1}{n(n-1)(n-2)\cdots 1} \int \cdots \int$$
$$1_F(\partial[x_1, m_1] + \cdots + \partial[x_n, m_n]) \rho(d[x_1, m_1])$$
$$\times \cdots \rho(d[x_n, m_n])$$

for $F \in F_S$

A probability measure P on (X^K, F_S) is called a marked Gibbs measure with total energy U if it is absolutely continuous with respect to the marked Poisson measure Q and the Radon–Nikodym derivative $f = \mathrm{d}P/\mathrm{d}Q$ is a Markov function. The measurable function $U: F_S \to (-\infty, \infty]$ must satisfy the following conditions (see Nguyen and Zessin, 1979): $U(\varnothing) = 0$ and $U(\psi_1) = \infty$ whenever $U(\psi_2)$ and $\psi_2 \subset \psi_1$ for $\psi_1, \psi_2 \in F_S$.

The distribution of the point process is given by a probability density function $f: W^n \times K^n \to [0, \infty)$ such that

$$f(\psi) = \frac{1}{Z} \exp(-U(\psi))$$

where Z is a normalizing constant. We restrict ourselves to pairwise interaction processes for which the total energy U can be written in terms of a mark pair potential function $\phi: T \times T \times K \times K \to (-\infty, \infty)$ and a mark chemical activity function $\alpha: K \to (-\infty, \infty)$ i.e.

$$U(\psi) = \sum_{i=1}^{n} \alpha(m_i) + \sum_{i<j} \phi(x_i, x_j, m_i, m_j)$$

If the process is homogeneous ϕ depends on the distance $d(x_i, x_j)$. This function characterizes interactions between the marked points and $\alpha(m)$ describes the ability of the system to receive a point with mark m. In addition to the total energy U we can consider the local energy at $[x, m]$ with respect to ψ.

$$E(x, m, \psi) = \alpha(m) + \sum_{[y,l] \in \psi} \phi(x, y, m, l)$$

which can be interpreted as the energy required to add the point $[x, m]$ to the configuration ψ.

3.3. Maximum pseudo-likelihood method

From now on we consider parametric Gibbs models for which the local energy can be written as

$$E_\theta(x, m, \psi) = \alpha(m) + \sum_{[y,l] \in \psi} \phi(x, y, m, l; \theta)$$

where θ is a parameter vector to be estimated. Since the distribution of the points of a data set can potentially be influenced by some points outside of the sampling window, we should take the effect of the unobserved points into account. Different strategies have been proposed to reduce the

edge effect but their careful treatment is beyond the scope of this paper. We'll assume that the point pattern is a realization of a Gibbs process on a torus T. Hence, we avoid the problem of boundary correction. Clearly, applicability of the estimation method considered here is not restricted to the cases for which the sampling window can be wrapped onto a torus.

The idea of the maximum pseudo-likelihood (MPL) method and its application in estimating parameters of Gibbs point processes is due to Besag (1978). This estimation procedure is a computationally easier alternative to the maximum likelihood method since the likelihood function contains a scaling factor which cannot be calculated explicitely (Mateu and Montes, 1994).

In order to provide a rigourous base to the use of the MPL method for analyzing point patterns, Besag et al. (1982) showed that a sequence of lattice processes approximating the point process converges to the point process. Using this limit procedure Ripley (1988) found a general form for the pseudo-likelihood (PL) function of unmarked processes. Jensen and Moller (1991) proposed a definition of the PL function as a limit of a product of conditional densities. Starting from the true likelihood it is proved that a sequence of degraded likelihoods gives in the limit the pseudo-likelihood.

3.3.1. Definition

The PL function is defined by

$$\log PL(\theta;\psi) = -\sum_{[x,m]\in\psi} E_\theta(x,m,\psi-[x,m]) - \int_T\int_K \exp(-E_\theta(\xi,t,\psi))M(\mathrm{d}t)\mathrm{d}\xi$$

Justification of the functional form of the PL function can be found in Grabarnik and Sarkka (1992). According to the above definition, the log PL function for a multitype Gibbs process (discretely marked) is

$$-\log PL(\theta;\psi) = \sum_{i=1}^{k} n_i\alpha(i) + \sum_{i=1}^{k}\sum_{j=1}^{k}\sum_{x\in\psi_i}\sum_{\substack{y\in\psi_j\\ y\neq x}}\phi(d(x,y),i,j) + \sum_{i=1}^{k} e^{-\alpha(i)}\int_T \exp\left(-\sum_{j=1}^{k}\sum_{x\in\psi_j}\phi(\mathrm{d}(x,\xi),i,j)\right)\mathrm{d}\xi$$

where n_i is the total number of events of species i. And the log PL function for a continuous marked Gibbs process with a bounded mark space K is

$$-\log PL(\theta;\psi) = \sum_{[x,m]\in\psi}\alpha(m) + \sum_{[x,m]\in\psi}\sum_{\substack{[y,l]\in\psi\\ y\neq x}}\phi(d(x,y),m,l) + \int_T\int_K \exp(-\alpha(t) - \sum_{[x,m]\in\psi}\phi(d(x,\xi),m,t))\,\mathrm{d}t\,\mathrm{d}\xi$$

Proofs can be found in Grabarnik and Sarkka (1992).

4. Spatial analysis of forest ecosystems

Let us consider two examples of forest ecosystems. Both data sets have been observed within a rectangular sampling window and therefore some corrections are needed to reduce the edge effects (Diggle, 1983). Here the toroidal edge correction is applied, i.e. the rectangle is mapped onto a torus by identifying opposite edges.

4.1. Ecosystem 1

A data set consisting of locations of two types of mediterranean trees, *Quercus ilex L.* (110 individuals, type 1-events) and *Pinus halepensis L.* (81 individuals, type 2-events), measured in a rectangular planar area of size 180×120 m^2. See Fig. 1.

Fig. 2 shows the EDF plot of nearest neighbour distances ($\hat{G}$ function) for the pattern consisting of the locations of *Quercus ilex L.*, together with the upper and lower envelopes from 99 simulations of CSR (the benchmark hypothesis). There is a clear excess of nearest neighbour distances between 4 and 5 and between 6 and 7 m which provides evidence for rejection of CSR in favour of a relative regularity of events in the above mentioned intervals. Fig. 2 also provides the plot of the EDF of point to nearest event distances ($\hat{F}$

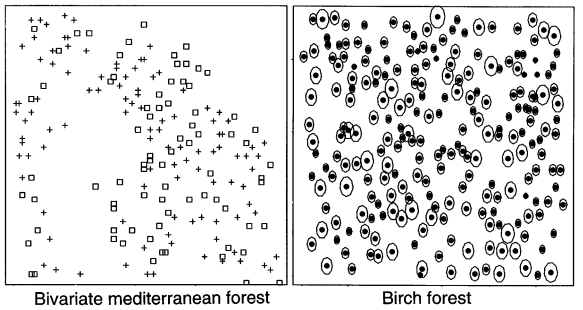

Fig. 1. Mediterranean forest: *Quercus ilex L.* (110 individuals, crosses) and *Pinus halepensis L.* (81 individuals, squares) in a rectangular plot of size 180×120 m^2. Birch forest (231 individuals) in a rectangular plot of size 280×190 m^2 . The sizes of the circles are proportional to crown lengths.

function) using a grid of 1100 points inside the surrounding polygon. This EDF for the data lies below the lower simulation envelope for part of its range (from 3.75 to 10 m) leading to emphatic rejection of CSR. This different behaviour of $\hat{G}$ and $\hat{F}$ functions is typical in this kind of patterns (Diggle, 1983).

The formal statistical test for CSR using a Monte-Carlo experiment to evaluate $\hat{F} - \hat{G}$ gives the following values:

$$D_1 = 0.209336,\ \min(D_i,\ i = 2,...,100)$$
$$= 0.042669,\ \max(D_i,\ i = 2,...,100) = 0.209172.$$

So the Monte-Carlo test suggests rejection of CSR with an attained significance level of 0.01 confirming the graphical assessment. We can conclude that the univariate pattern of *Quercus ilex L.* has been produced by a non-random and regular ecological mechanism with most of individuals being between 4–5 or 6–7 m.

Concerning type 2 data, we have also plotted the EDF $\hat{G}$ and $\hat{F}$ (see Fig. 2). There is a clear excess of nearest neighbour distances between 4 and 6 m. Finally, Fig. 2 also provides the plot of the EDF of point to nearest event distances ($\hat{F}$ function) using a grid of 1100 points inside the surrounding polygon. This EDF for the data lies below the lower simulation envelope for part of its range (from 4.5 to 10 m) leading both EDF's to emphatic rejection of CSR. We conclude that

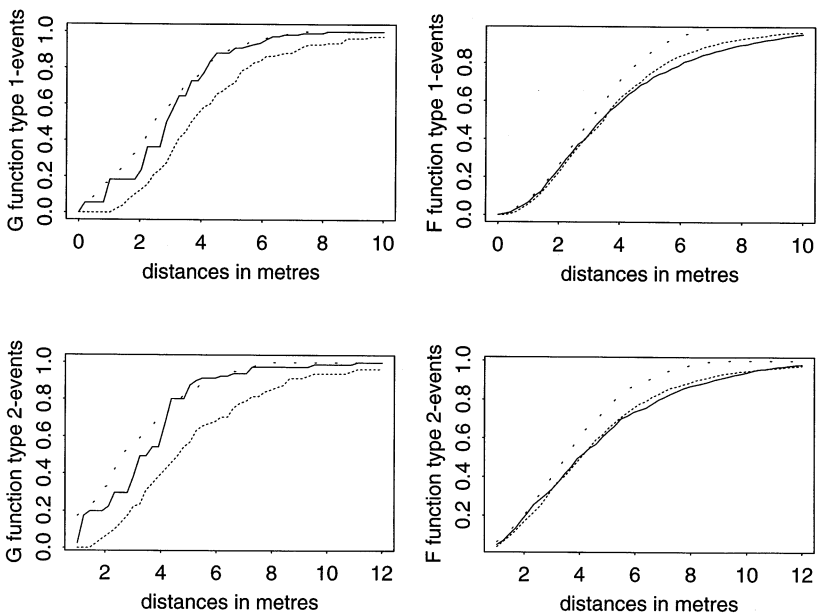

Fig. 2. EDF plot of nearest neighbour distances (G function) for type 1 and 2-events: $\hat{G}$ (solid line), Upper envelope under CSR (sparcely dotted line), Lower envelope under CSR (continuously dotted line). EDF of point to nearest event distances (F function) for type 1 and 2-events: $\hat{F}$ (solid line), Upper envelope under CSR (sparcely dotted line), Lower envelope under CSR (continuously dotted line).

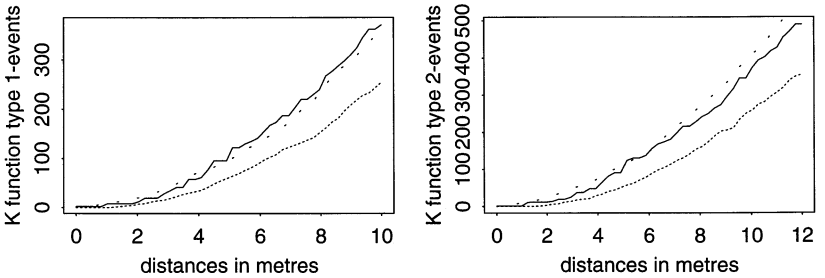

Fig. 3. K function for type 1 and 2-events (K_{11} and K_{22}, respectively): $\hat{K}_{ii}$ (solid line), Upper envelope under CSR (sparcely dotted line), Lower envelope under CSR (continuously dotted line).

Pinus halepensis L. pattern is produced by a regular mechanism with a 'normal' (under CSR) number of individuals being less than 4 m appart and with most of them being more than 4 m appart. The formal statistical test gives

$$D_1 = 0.243291,\ \min(D_i, i = 2,...,100)$$
$$= 0.071024,\ \max(D_i, i = 2,...,100) = 0.2418912.$$

Fig. 3 shows the K-function for the observed data together with the upper and lower envelopes from 99 simulations under the benchmark hypothesis of CSR for type 1 and type 2 individuals. It is obvious that type 1 and 2-events are not compatible with the CSR hypothesis.

In Fig. 4 it is observed that the trace for the data does not lie within the upper and lower envelopes from 99 simulations under the hypothesis of random labelling. If we explore Fig. 4 it is seen that the trace corresponding to the observed data lies within the upper and lower envelopes from 99 simulations under the hypothesis of independence of the two components for values of distance t less than about 7 m. For larger distances the data is not compatible with independence of the two components. For distances above 7 m both patterns become dependent but still generated by two different regular biological mechanisms.

As an appropriate model for this data set we consider a bivariate Gibbs point process for which the mark space consists of two elements, $K = \{1,2\}$, and the local energy $E_\theta(x, m, \psi)$ can be written in the form

$$E_\theta(x, 1, \psi)$$
$$= \alpha_1 + \sum_{[y,1]\in\psi} \phi_{11}(x, y; \theta_{11}) + \sum_{[y,2]\in\psi} \phi_{12}(x, y; \theta_{12})$$
$$E_\theta(x, 2, \psi)$$
$$= \alpha_2 + \sum_{[y,2]\in\psi} \phi_{22}(x, y; \theta_{22}) + \sum_{[y,1]\in\psi} \phi_{12}(x, y; \theta_{12})$$

where $(\alpha_1, \alpha_2, \theta_{11}, \theta_{22}, \theta_{12})$ are the parameters to be estimated by the MPL method, ϕ_{11} and ϕ_{22} describe interactions within the species 1 and 2, respectively, and ϕ_{12} interactions between the two species.

We have fitted a bivariate Strauss model with pair potential function

$$\phi_{ij}(x, y, \theta_{ij}) = \begin{cases} \infty & \text{if } d(x, y) < 1 \\ \theta_{ij} & \text{if } d(x, y) < 12 \\ 0 & \text{if } d(x, y) \geq 12 \end{cases}$$

where the hard-core distance is 1 m, the minimum inter-event distances, and interaction radii 12 m. Parameter values very close to zero indicate homogeneous or CSR pattern. Regular patterns are based on parameter values above 0.5 and cluster patterns are defined for parameter values below zero. The MPL estimates are $\hat{\alpha}_1 = 4.9$, $\hat{\alpha}_2 = 6.9$ and $\hat{\theta}_{11} = 0.671$, $\hat{\theta}_{22} = 0.523$, $\hat{\theta}_{12} = -0.231$. Type 2 trees pattern has more ability to receive a new marked event than type 1 pattern ($\hat{\alpha}_2 > \hat{\alpha}_1$). According to this model, the patterns formed by *Quercus ilex L.* and *Pinus halepensis L.* are generated by a regular inhibitory mechanism in determined intervals of distances and they are CSR patterns out of them. The dependence between both species is based on a certain attraction for distances above 7 m.

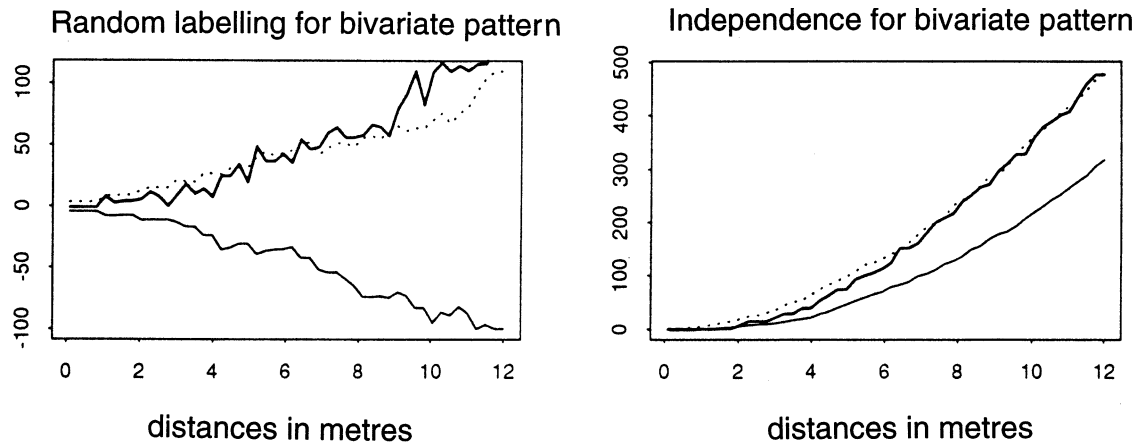

Fig. 4. Random labelling and independence: Empirical (wider solid line), Upper envelope under random labelling or independence (dotted line), Lower envelope under random labelling or independence (narrower solid line).

To see that the estimated model fits to the data we calculated the L_{11}, L_{12} and L_{22} functions (see e.g. Diggle, 1983) and the corresponding upper and lower envelopes from 99 simulations under estimated parameters. All the L-functions calculated from the data lie between the envelopes (Fig. 5) and therefore the estimated model fits well to the data.

4.2. Ecosystem 2

The second data set is concerned with a birch forest area in which locations (231 individuals) and crown lengths (lengths varying between 2 and 8 m) are recorded in a rectangular planar stand of size 280×190 m^2. See Fig. 1.

The EDF plot based on the $\hat{G}$ function, Fig. 6, now shows the deficiency of small nearest neighbour distances which is typical of regular or inhibitory patterns. Using again a grid of 1100 points in the sampling window to calculate the $\hat{F}$ function, the position of this EDF is near or above the upper simulation envelope which typifies a regular pattern and again contrast with the behaviour of $\hat{G}$. An effective summary of the regularity in these data is a plot of the K-function which is shown in Fig. 6. There is a fairly well defined inhibition distance of 3.16 (minimum inter-event distance). Also, in the range $t \leq 14$ the K-function for the data lies below the lower envelope from 99 simulations of CSR.

A test for CSR using a Monte-Carlo experiment to evaluate $\hat{F} - \hat{G}$ leads to rejection of CSR with a comfortably attained significance level of 0.01.

Let a bounded set $K = [2,8]$ be the mark space for the crown lengths recorded in 231 positions of sampled birches in the stand. We model interactions among crowns by the mark pair potential function

$$\phi_{c_1,c_2}(x, y, m, l) =$$

$$\begin{cases} \infty & \text{if } d(x, y) < 3.16 \\ c_1 ml \exp(-c_2 d(x, y)) & \text{if } 3.16 \leq d(x, y) < 14 \\ 0 & \text{if } d(x, y) \geq 14 \end{cases}$$

which depends only on two parameters: c_1 and c_2. The chemical activity function considered was $\alpha(m) = \exp(-m\lambda)$ where λ is the intensity of the pattern and can be estimated in the obvious way by $\hat{\lambda} = n/|A|$, where n denotes the number of events in the region of observation.

The MPL estimates gave $\hat{c}_1 = 0.6$ and $\hat{c}_2 = 0.13$ indicating a regular or inhibitory pattern amongst birches. The goodness-of-fit of the model was studied by means of the L_{mm} statistic corresponding to the mark correlation measure (Penttinen et al., 1992). The L_{mm} statistic calculated from the data lies between the envelopes (Fig. 7) and therefore the estimated model fits well to the data.

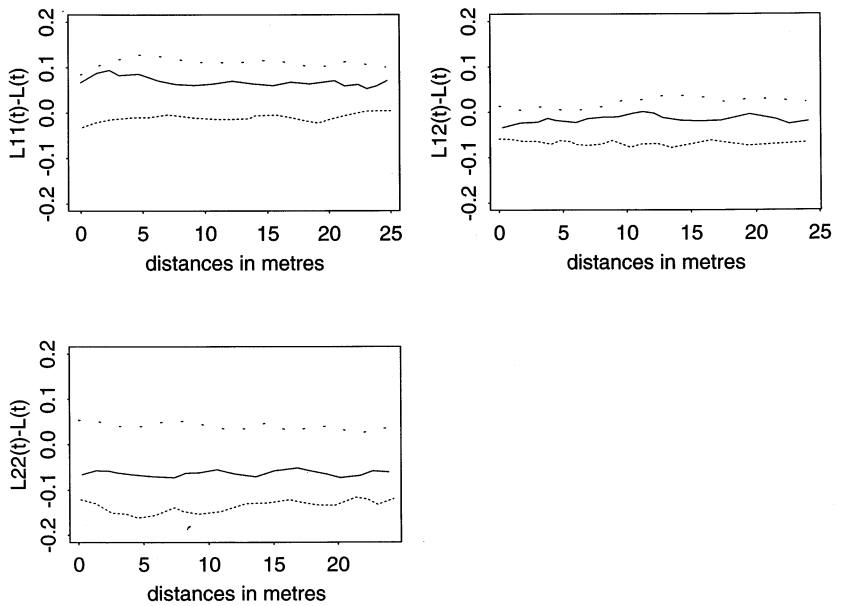

Fig. 5. Goodness-of-fit for bivariate Strauss model: $L_{11}(t) - L(t)$ (type 1-events), $L_{12}(t) - L(t)$ (combination of two types) and $L_{22}(t) - L(t)$ (type 2-events). Estimated (solid line), Upper envelope under CSR (sparcely dotted line), Lower envelope under CSR (continuously dotted line).

5. Discussion

In the literature the MPL method has been applied mainly to analyze unmarked point patterns, and only little attention has been paid to marked patterns. Obviously, the class of marked Gibbs point processes extends the applicability of point process models to such practical problems, where the interest is not only in the point pattern but also in variables associated to the locations. Models with marks are able to describe various kinds of interaction between the objects and may therefore be more realistic.

It is not necessary that the chemical activity is constant in the continuously marked case as it has been the case for a long time. Here we have proposed a chemical activity as a function of marks. However, we have not considered a parametrized chemical activity function $\alpha(m;\mu)$ depending on a set of parameters μ. This would be a nice idea to incorporate in pseudo-likelihood functions.

In this paper we have used a toroidal edge correction. This periodic boundary is commonly used for computer experiments in statistical mechanics in order to calculate physical quantities associated with a particular potential function. It is known to be highly reliable even for processes with very strong interactions. Nevertheless, it also has some disadvantage: periodic boundaries can introduce undesirable artefacts; for example, toroidal distances can be arbitrarily small even when the underlying process has a positive hard-core distance. Instead of this method one may use the one introduced by Ripley (1977) which corrects the bias using areas of parts of circles.

For non-homogeneous patterns we might use a varying local intensity, $\lambda(x)$ say, in the chemical activity function instead of a constant. An estimation given by Diggle (1985) has the following form

$$\hat{\lambda}(x) = \sum_{i=1}^{n} \left(1 - \frac{d_i^2}{2h_0^2}\right)^2$$

where the parameter h_0 specifies the width of the kernel whereas d_i is the distance from event i to x.

The tests using upper and lower envelopes are based on 99 simulations to produce a one-sided

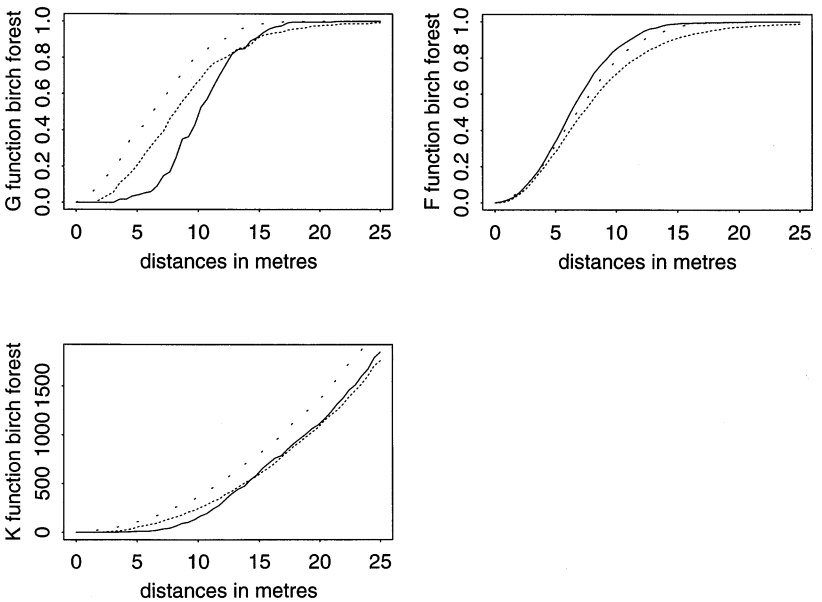

Fig. 6. EDF plots for birch forest: nearest neighbour distances (G function), point to nearest event distances (F function) and K function. $\hat{G}$, $\hat{F}$ and $\hat{K}$(solid line), Upper envelopes under CSR (sparcely dotted line), Lower envelopes under CSR (continuously dotted line).

test with significance level 1% (or 2% in a one-sided test). However, many published papers are based on 19 simulations to get significance levels of 5 or 10% in one or two-sided tests, respectively.

This paper has presented several statistical methods which are, we believe, of varying importance for the forest statistician. These methods may be included into the everyday toolbox of exploratory methods for analyzing mapped tree data.

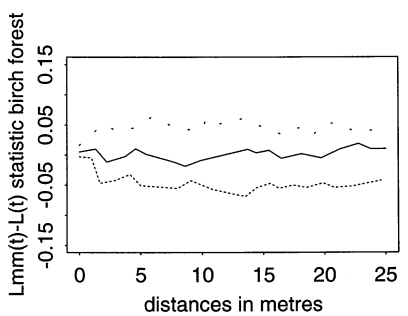

Fig. 7. Goodness-of-fit for birch forest: $L_{\mathrm{mm}}(t) - L(t)$. Estimated (solid line), Upper envelope under CSR (sparcely dotted line), Lower envelope under CSR (continuously dotted line).

References

Averintshev, M.B., 1973. Gibbsian representation of random fields whose conditional probabilities may vanish (in Russian). Problemy Peredachi Informatsii 11 (4), 86–96.

Barnard, G.A., 1963. Contribution to the Discussion of Prof. Bartlett's paper. J. R. Stat. Soc. B 25, 295.

Besag, J., 1978. Some methods of statistical analysis for spatial data. Bull. Int. Stat. Inst. 47 (2), 77–92.

Besag, J., Milne, R., Zachary, S., 1982. Point process limits of lattice processes. J. Appl. Prob. 19, 210–216.

Cox, D.R., Isham, V., 1980. Point processes. Chapman and Hall, London.

Diggle, P.J., 1983. Statistical Analysis of Spatial Point Patterns. London: Academic Press.

Diggle, P.J., 1985. A Kernel method for smoothing point process data. J. R. Stat. Soc. C 34, 138–147.

Grabarnik, P., Sarkka, A., 1992. On parameter estimation of marked Gibbs point processes. Preprints from the Department of Statistics, University of Jyvaskyla, p. 4.

Grimmet, G.R., 1973. A theorem about random fields. Bull. Lond. Math. Soc. 5, 81–84.

Jensen, J.L., Moller, J., 1991. Pseudolikelihood for exponential family models of spatial point processes. Ann. Appl. Prob. 1, 445–461.

Mateu, J., Montes, F., 1994. Modelling Pairwise Interaction point processes. 9th International Workshop on Statistical Modelling, Exeter, UK.

Nguyen, X.X., Zessin, H., 1979. Integral and differential characterization of the Gibbs process. Math. Nachr. 88, 105–115.

Penttinen, A., Stoyan, D., Henttonen, H., 1992. Marked point processes in forest statistics. For. Sci. 38 (4), 806–824.

Preston, C.J., 1976. Random Fields. Lecture Notes in Mathematics 534. Berlin-Heidelberg; New York, Springer.

Ripley, B.D., 1977. Modelling spatial patterns (with discussion). J. R. Stat. Soc. B 2, 172–212.

Ripley, B.D. 1988. Statistical inference for spatial processes. Cambridge University press, Cambridge.

Ripley, B.D., Kelly, F.P., 1977. Markov point processes. J. Lond. Math. Soc. 15, 188–192.

Ruelle, F., 1969. Statistical Mechanics. Benjamin, New York.