ELSEVIER Computational Statistics & Data Analysis 36 (2001) 85–105

COMPUTATIONAL STATISTICS & DATA ANALYSIS

www.elsevier.com/locate/csda

Gibbs point processes for studying the development of spatial-temporal stochastic processes

Eric Renshaw[a,*], Aila Särkkä[b]

[a] *Department of Statistics and Modelling Science, University of Strathclyde, Livingstone Tower, 26 Richmond Street, Glasgow G1 1XH, UK*

[b] *Department of Mathematical Statistics, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*

Received 1 July 1999; accepted 1 May 2000

Abstract

Although many studies of marked point processes analyse patterns in terms of purely spatial relationships, in real life spatial structure often develops dynamically through time. Here we use a specific space–time stochastic process to generate such patterns, with the aim of determining purely spatial summary measures from which we can infer underlying generating mechanisms of space–time stochastic processes. We use marked Gibbs processes in the estimation procedure, since these are commonly used models for point patterns with interactions, and can also be chosen to ensure that they possess similar interaction structure to the space–time processes under study. We restrict ourselves to Strauss-type pairwise interaction processes, as these are simple both to construct and interpret. Our analysis not only highlights the way in which Gibbs models are able to capture the interaction structure of the generating process, but it also demonstrates that very few statistical indicators are needed to determine the essence of the process. This contrasts markedly with the relatively large number of estimators usually needed to characterise a process in terms of spectral, autocorrelation or K-function representations. We show that the Strauss-type procedure is robust, i.e. it is not crucial to know the exact process-generating mechanism. Moreover, if we do possess additional information about the true mechanism, then the procedure becomes even more effective.

Keywords: Immigration–interaction–death; Inhibition; Marked Gibbs point processes; Maximum pseudo-likelihood; Pair potential functions; Space–time; Strauss-type processes

* Corresponding author. Tel.: +44-141-548-3591; fax:+44-141-552-2079.
E-mail address: eric@stams.strath.ac.uk (E. Renshaw).

0167-9473/01/$ - see front matter
PII: S 0167-9473(00)00028-1

1. Introduction

Although Whittle's (1954) pioneering paper put the study of spatial processes firmly on the statistical map, subsequent developments have mainly centered on purely *spatial* relationships. In real life, however, patterns often develop dynamically through *time*. So although a purely spatial analysis is ideal for achieving a snap-shot summary of spatial structure, such as classifying texture through parameter estimates for a Markov random field, to gain an understanding of the underlying evolutionary processes which generate pattern necessitates a spatial–temporal approach. Bartlett (1975) provided a mechanism for achieving this through the development of the spatial–temporal spectrum, though for some time afterwards there was considerable inertia against conducting space–time statistical analyses. In fairness this was most likely due to the difficulty of collecting appropriate data; modern electronic scanning and video devices have clearly made this far less of a problem.

One of the early biological experiments that provided the impetus for change was performed by Ford (1975), who studied the effect of between-plant competition on *Tagetes patula* L. (marigolds) planted on a regular lattice. A time-lapse film over a three-month period produced a (visual) matrix of continuous-valued variables in continuous time; for the purpose of statistical analysis measurements on plant height were recorded at intervals of 2, 4, 6 and 8 weeks. Ford and Diggle (1981) modelled these data by assuming that competition between individuals was for light. Interaction between neighbours was viewed as a spatial process in relation to differences in plant height, and simulations reproduced both bimodality in the frequency distribution of plant size and an even spatial distribution of large or surviving plants. Renshaw (1984) made theoretical progress by exploiting the underlying lattice structure using a frequency-domain approach. This holds a considerable advantage over space-domain techniques since, unlike sample autocovariances, periodogram values are effectively independent of each other. Further stimulii (Renshaw and Ford, 1983, 1984; Ford and Renshaw, 1984) are provided by the distinctive nature of various ecologically based spectra, and that use of sliding data windows can highlight the presence of nonstationarity.

Development of a theoretical spectrum is greatly helped by the assumption of a linear model. Renshaw (1984) assumes that in the absence of competition plants grow at constant rate λ, and that two plants a distance (r,s) apart compete at rate a_{rs} times their height difference. Then on denoting $\{X_{ij}(t)\}$ to be the set of plant heights at time $t \geq 0$ over $-\infty < i,j < \infty$, we have the linear stochastic equation

$$X_{ij}(t+\mathrm{d}t) = X_{ij}(t)(1+\lambda\,\mathrm{d}t) + \sum_{r,s=-\infty}^{\infty} a_{rs}\{X_{ij}(t) - X_{i+r,j+s}(t)\}\,\mathrm{d}t + \mathrm{d}Z_{ij}(t) + \mathrm{o}(\mathrm{d}t), \tag{1.1}$$

where $a_{rs} \geq 0$, $\sum a_{rs} < \infty$ and $\{\mathrm{d}Z_{ij}(t)\}$ is a white-noise sequence with mean 0 and variance $\sigma^2\,\mathrm{d}t$. Although the associated autocovariances can be expressed neatly only in terms of their generating function (a.g.f.), the paradigm between the a.g.f. and the spectrum means that the latter takes a simple form, namely that for

frequencies $\underline{\omega} = (\omega_1, \omega_2)$

$$f(\omega_1, \omega_2; t) = \tau^2 \exp(\psi t) + (\sigma^2/\psi)\{\exp(\psi t) - 1\}, \tag{1.2}$$

where

$$\psi \equiv \psi(\omega_1, \omega_2) = 2\left\{\lambda + \sum_{r,s=-\infty}^{\infty} a_{rs}[1 - \cos(r\omega_1 + s\omega_2)]\right\} \tag{1.3}$$

and τ^2 is the variance of the initial configuration $\{X_{ij}(0)\}$. This result is important for three reasons. First, it enables least-squares estimates to be obtained for the growth and spatial interaction parameters (Renshaw, 1984), with $a_{rs} > 0$ being viewed as competition and $a_{rs} < 0$ as "enhancement" at distance (r, s). Second, it enables us to predict the type of spatial structure that will result from given interaction parameters (Renshaw, 1994a). Third, Renshaw et al. (2000) show that taking the inverse Fourier cosine transform of $\psi(\omega_1, \omega_2)$ allows the construction of weights $\{a_{rs}\}$ which will produce any specified "target" spectrum.

Since here we are interested in phenomena that can be considered to be realisations of spatial marked point processes, the above system (1.1)–(1.3) is not appropriate since it is lattice based. Indeed, transposing this to the real plane (see Daley and Vere-Jones, 1988, Chap. 11) poses fundamental theoretical problems since it destroys the Fourier properties which give rise to such simple functional spectral structures. This difficulty is compounded if we impose a logistic-type reaction function to ensure that the process remains non-explosive. For the foreseeable future theoretical progress therefore appears unlikely, and so study of such spatial–temporal marked point processes will have to proceed via simulation; our aim is to determine optimal ways of characterising their development. Given recent work by several authors on autocovariance and spectral analysis of pure point and marked point processes, together with earlier work involving distance-based measures (e.g. Ripley's K-function), one possible route would be to see how sample autocovariance, spectral and distance functions change with respect to the reaction and diffusion components of the space–time model. However, each of these are many-valued descriptors, and so the question arises as to whether one can construct just a few summary statistics which are sensitive to change in the underlying process. We shall base these on marked Gibbs processes (see, for example, Stoyan et al., 1995), and study whether the interaction parameters of these purely spatial processes provide a useful and sensible characterisation of our space–time process.

The reason for considering Gibbs processes is that they have some similarities with the system (1.1)–(1.3). They are defined through the probability density function (p.d.f.)

$$f(\varphi) = \frac{1}{V}\exp\{-U(\varphi)\}, \tag{1.4}$$

where $\varphi = \{[x_1; m_1], [x_2; m_2], \ldots, [x_n; m_n]\}$ is a realisation of the point process Φ consisting of the locations x_i and marks m_i of the points $i=1, \ldots, n$, V is a normalising constant, and U denotes an energy (reaction) function (see Stoyan et al., 1995, Chap. 5). Now this representation has two key elements in common with our original space–time process. First, U can take exactly the same form as the right-hand side of

Eq. (1.1). Second, both the numerators in (1.4) and (1.2) take a similar exponential form; the main structural difference lies in the denominator, with the normalising constant V replacing the function ψ.

Our aims are therefore two-fold. First, we determine whether purely spatial Gibbs-type estimation procedures yield quantitative insight into underlying space–time stochastic generating mechanisms when their nature is known, and then unknown. Second, we assess whether parameter estimates relating to a given realisation change markedly over time.

2. The general space–time process and Gibbs process

Any attempt to construct simulated lattice realisations $\{X_{ij}(t)\}$ hits the problem that i and j take all integer values. Though they could be restricted to lie within the finite range $i,j=1,\ldots,N$, the ensuing edge-effects are likely to distort the space–time process quite severely. However, placing the process on a torus gives rise to a Turing ring process (Renshaw, 1994b; Smith and Renshaw, 1994; Turing, 1952), and Renshaw (1994a) shows (in one dimension) that the processes on the infinite line and finite ring have (effectively) identical spectra and thereby similar spatial characteristics; choice of N merely determines the maximum observable wavelength. The extension of this result to two (and higher) dimensions is immediate, whence by considering the plane as a lattice limit it is reasonable to presume that marked point processes on the plane and torus will also possess equivalent characteristics. However, when analysing real, as opposed to simulated, data sets, one might wish to employ edge corrections (see, for example, Ripley, 1988) rather than mapping the study area onto a torus.

An immediate paradigm to (1.1) is the immigration–growth–spatial interaction process on the unit torus T, in which the ith individual is of size (mark) $m_i(t)$ at time t, and is located at x_i. We assume that the process is simple, so that there is at most one individual at each location. New immigrants arrive randomly in time according to a Poisson process with rate α, have uniformly distributed locations on $U(0,1)^2$, and are assigned marks from $U(0,\varepsilon)$ for appropriately small $\varepsilon>0$. In the successive small time intervals $(t,t+\mathrm{d}t)$, each individual undergoes the deterministic incremental size change

$$m_i(t+\mathrm{d}t)=m_i(t)+\lambda m_i(t)(1-m_i(t)/K)\mathrm{d}t+\sum_{j\neq i}h(m_i(t),m_j(t);||x_i-x_j||)\,\mathrm{d}t, \tag{2.1}$$

where λ denotes the intrinsic rate of growth, K denotes the logistic non-spatial carrying capacity, $||x_i-x_j||$ denotes the distance between individuals (i.e. points) i and j, and $h(\cdot)$ is an appropriate spatial interaction function taken over all points $j\neq i$. If $m_i(t+\mathrm{d}t)\leq 0$ then the individual is deemed to have died and the point i is deleted. Here we have used logistic growth to prevent explosive birth, though this could clearly be replaced by any other suitable non-linear function. Note that if $h(\cdot)$ corresponds to competition then this process clearly remains bounded. However, if mutual enhancement is possible (corresponding to some $a_{rs}<0$ in (1.1)), then it

may become necessary to subject each individual to a simple death process at rate $\mu > 0$ in order to secure a proper equilibrium distribution.

We will study whether Gibbs point processes are able to capture the interaction structure of the generating space–time process (2.1). For inhibitive interaction (competition) it is usually sufficient to write the energy function U in (1.4) in terms of pairwise interactions through

$$U(\varphi) = \sum_i \beta(x_i, m_i) + \sum_{i<j} \phi(x_i, x_j, m_i, m_j), \tag{2.2}$$

where β: $T \times M \to (-\infty, \infty)$, M being the mark space, is the so-called chemical activity function and ϕ: $T \times T \times M \times M \to (-\infty, \infty)$ is an appropriate pair potential function. The chemical activity function can be interpreted as the ability of the system to receive a point at location x_i with mark m_i. The pair potential function, ϕ, describes the interaction between two points: in broad terms $\phi > 0$ indicates inhibition, $\phi < 0$ clustering, and $\phi = 0$ corresponds to the Poisson case with no interaction between points. In principle, pairwise interaction processes are appropriate models for both clustered and regular point patterns, but problems can occur in the clustered case since the models are not always well-defined. Furthermore, if we try to simulate highly attractive pairwise interaction processes, we will most likely generate one clump of points instead of a nicely clustered pattern: this was demonstrated for the Strauss process by Geyer and Thompson (1995). If the process is spatially stationary, as assumed here, then β is a function of the marks alone, whilst ϕ is a function of the marks and the inter-point distance $||x_i - x_j||$. The assumption of spatial homogeneity of the point pattern can be tested (see, for example, Diggle, 1983); if this homogeneity hypothesis is rejected then we can use inhomogeneous Gibbs point processes (Ogata and Tanemura, 1986; Baddeley and Turner, 2000; Nielsen and Jensen, 2000; Stoyan and Stoyan, 1998; Baddeley et al., 1999).

Diggle (1983), Ogata and Tanemura (1985), Takacs and Fiksel (1986), Särkkä (1993) and Goulard et al. (1996) have examples of pair potential models for binary point processes, i.e. processes with two types of points. In this case, and in the general multitype case, one can take a non-marked pair potential function with separate interaction functions for each type of interaction. In the binary case we would have three kinds of interactions: two within types and one between types. Further examples can be found, for example, in Stoyan et al. (1995), whilst others arise in statistical physics (e.g. Ogata and Tanemura, 1985).

Fiksel (1984), Ogata and Tanemura (1985), Baddeley and Møller (1989), Stoyan (1989) and Goulard et al. (1996) consider models for pair potential functions with continuous marks. Ogata and Tanemura suggest the use of potential functions that are similar in shape to the non-marked potentials, but with scales dependent on the marks. If the non-marked potential function is $\phi_\theta(||x_i - x_j||)$, for parameter θ, then the corresponding marked potential function is $\phi_\theta(||x_i - x_j||/z_{ij})$ where z_{ij} is a function of the marks m_i and m_j, e.g.

$$z_{ij} = \left(\frac{m_i + m_j}{2\bar{m}}\right)^c, \qquad z_{ij} = \left(\frac{m_i + m_j}{\bar{m}}\right)^{c/2} \quad \text{or} \quad z_{ij} = \frac{m_i^c + m_j^c}{2\bar{m}}, \tag{2.3}$$

where $\bar{m}$ is the mean mark and c is an "adjustment" parameter.

Baddeley and Møller (1989) and Goulard et al. (1996) suggest using a random disc process. Here a disc is centered at each point and has a radius that depends on the mark of the point. Then we can choose, for example, a Strauss-type pair potential function

$$\phi(||x_i - x_j||, m_i, m_j) = \begin{cases} \gamma & \text{if } ||x_i - x_j|| < m_i + m_j, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

So there is interaction between two points only if the discs connected to them overlap. This is the construct we choose to use here, though we include a mark scaling factor r so that $\phi(||x_i - x_j||, m_i, m_j) = \gamma$ if $||x_i - x_j|| < r(m_i + m_j)$.

The parameters of the pair potential functions can be estimated using, for example, maximum likelihood (ML) (Ogata and Tanemura, 1985; Penttinen, 1984; Geyer and Thompson, 1992; Geyer and Møller, 1994), or maximum pseudo-likelihood (MPL) (Jensen and Møller, 1991; Särkkä, 1993; Goulard et al., 1996; Baddeley and Turner, 2000). The MPL method is computationally easier than the ML method because no simulations (for the approximation of the normalising constant) are needed, and so we adopt this approach here. The MPL estimators satisfy unbiased estimation equations, are consistent, and are asymptotically Normal under suitable conditions. Moreover, if interaction between the points is weak then the MPL estimators are efficient. Although the pseudo-likelihood (PL) function includes an integral that has to be calculated numerically, in case of an exponential family model we can use standard software for fitting generalised linear models to maximise this function (Baddeley and Turner, 2000). Note, however, that the MPL method tends to overestimate interaction if this is strong between points.

To examine whether Gibbs parameter estimates can provide a useful summary of the development of the space–time process (2.1), our initial approach (Section 3) will be to employ a marked Gibbs process whose pair potential function is directly equivalent to the spatial interaction weights $h(\cdot)$ which are equivalent to those used in marked Gibbs processes (see Goulard et al., 1996). We will then investigate whether incorporating a version of (2.1) into the Gibbs estimation procedure also generates useful summaries. For then the pair potential function should be replaced by

$$\phi(||x_i - x_j||, m_i, m_j) = a(||x_i - x_j||)\{m_i - m_j\}, \tag{2.5}$$

which being non-symmetric in m_i and m_j is not of the usual Gibbs form, and so presents a much harder test. We shall demonstrate (Section 4) that in spite of this, the estimates of the interaction parameters of a symmetric Gibbs model still provide a useful summary of the underlying space–time process. This is a remarkable robustness property, which, for practical applications, means that it is not crucial to know the exact process generating mechanism.

3. Symmetric simulation experiment

We first generate point patterns according to (2.1) with a specific symmetric interaction term $h(\cdot)$ which enables us to match the interaction parameters of the estimating process with those of the generating process.

3.1. Simulation algorithm for the space–time process

Inserting the disc potential function (2.4) into the non-linear growth-interaction process (2.1) leads to the following algorithm, in which (small) population updates to each of the n marks are made at the discrete set of times $h, 2h, \ldots$, for $h \ll 1$ (typically $h = 0.01$), and $\{Z\}$ denotes a sequence of independent pseudo-random $U(0,1)$ variables. A simple death-process with rate μ is incorporated to ensure that the process remains bounded, though for pure competition we can safely take $\mu = 0$. Since spatial interaction is assumed to occur over the unit torus, there are no edge-effects to consider.

cycle over time $t = h, 2h, \ldots$
cycle over marks $i = 1, \ldots, n$

1. test for deaths: if $Z \leq \mu h$ then $m_i \to 0$;
2. apply deterministic growth and spatial interaction:

$$\tilde{m}_i = m_i + \lambda m_i(1 - m_i/K_i)h - c \sum_{j=1, j\neq i}^{n} I\{||x_i - x_j|| < r(m_i + m_j)\}h; \quad (3.1)$$

3. reset -ve marks: $m_i = \max(0, \tilde{m}_i)$;
4. test for new marks: if $Z \leq \alpha h$ then $n \to n + 1$ with location x_{n+1}, where both coordinates $\sim Z$, carrying capacity $K_{n+1} \sim k_1 + k_2 Z$ and mark $m_{n+1} \sim Z$;
5. delete zero sites: if $m_i = 0$ then delete $(x_i; m_i)$ from list, relabelling all remaining marks with $j > i$ to $j - 1$, and set $n \to n - 1$.

3.2. Behaviour of the space–time process

We simulate the process (3.1) starting from an empty torus, with $\alpha = 5$, $\mu = 0.02$, $\lambda = 1$, $k_1 = 15$, $k_2 = 5$, $c = 0.2$ and $r = 0.005$. These parameter values were chosen since they give rise to a "reasonable" number of points ($\simeq 120$) in the unit square; our choice of c enables some (but not too much) interaction between the points. Fig. 1 shows the development of a simulation at times $t = 4$, 20, 100 and 200. As the bubble plot radius equals rm_i, two points interact if and only if their discs overlap. Here $\alpha = 5$ new immigrants are expected to arrive per unit time, and in the absence of spatial interaction (i.e. $c = 0$) each has expected lifetime $1/\mu = 50$ with a $U(15, 20)$ carrying capacity, the equilibrium population size being Poisson with parameter $\alpha/\mu = 250$. For $c > 0$ this equilibrium size will be lower. Transient development is clearly visible through the first three frames, with empty spaces being exploited by new immigrants; the third and fourth frames appear (visually) to be statistically similar, though as more than two expected lifetimes have elapsed between them, few points are common to both. Clearly, summary statistical indicators are needed if formal tracking of such processes is to be carried out throughout their transient and persistent phases. The histogram of marks at time $t = 200$ (Fig. 2c) shows that the "canopy" is strikingly divided into an "upper" regime of n_{up} established marks (say size ≥ 10), and a "lower" regime of n_{low} small marks (say size

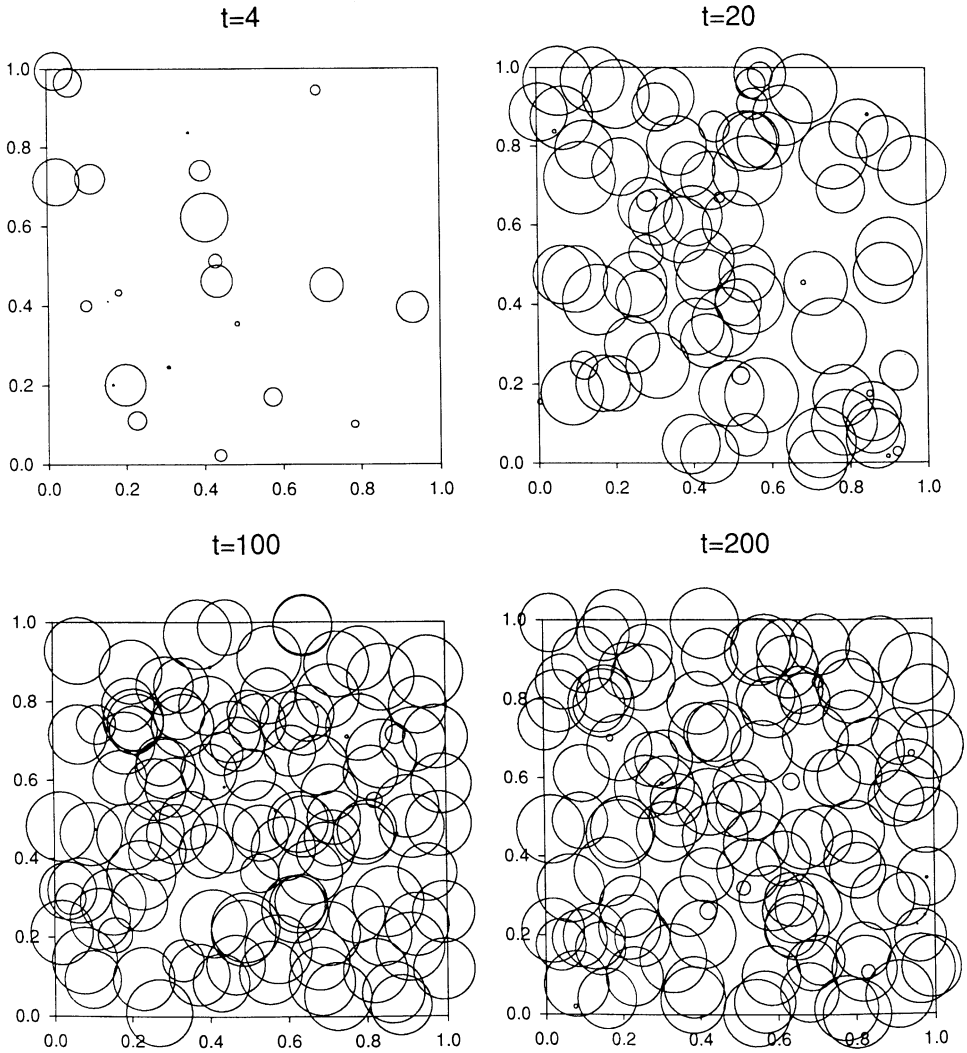

Fig. 1. Simulation of growth-interaction process (3.1) shown at times $t = 4, 20, 100$ and 200, with $\alpha = 5$, $\mu = 0.02$, $\lambda = 1$, $k_1 = 15$, $k_2 = 5$, $c = 0.2$ and $r = 0.005$; bubble plot radius is proportional to rm_i.

< 10) waiting the opportunity to develop as soon as the death of a nearby established mark creates sufficient space. Note that this bimodality mirrors that found in the geometrically based spatial interaction models simulated by Ford and Diggle (1981); see Section 1. Fig. 2a shows a steady rise in the number of points ($n = n_{\text{up}} + n_{\text{low}}$) until $t \simeq 100$ (i.e. two expected Poisson lifetimes), after which the increasing influence of death and spatial inhibition causes n to stabilise at around 130 points. Note that the lower canopy becomes established much faster than the upper canopy. The sum of the upper marks follows a similar path to n_{up}; that of the lower marks shows virtually no transient behaviour though it does exhibit considerable variability (Fig. 2b).

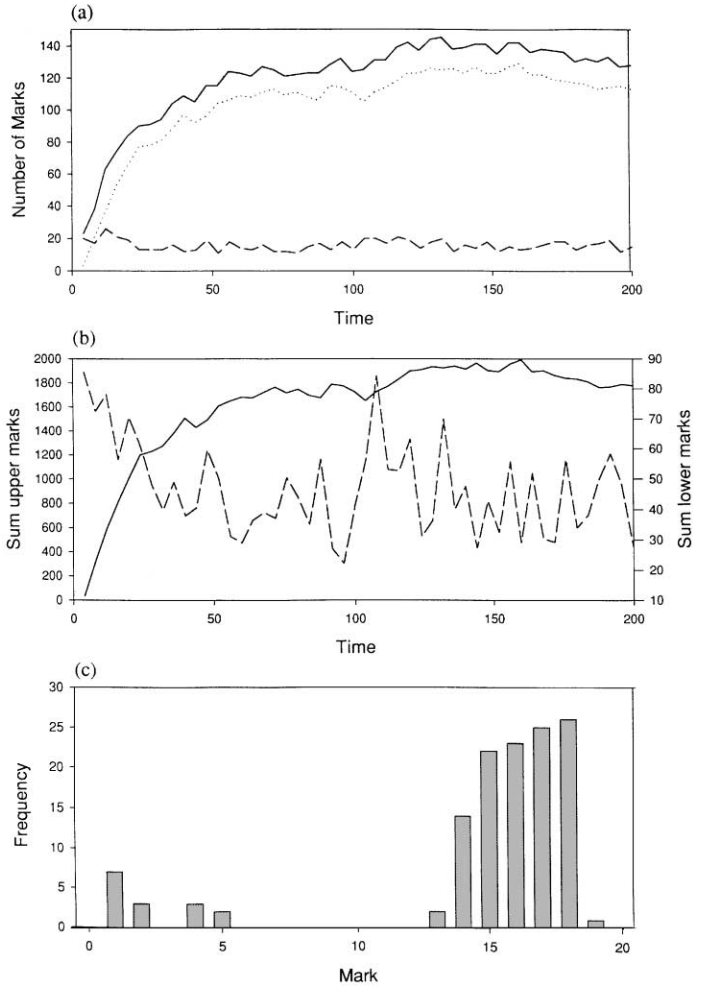

Fig. 2. Time-series plots corresponding to Fig. 1: (a) Total number of points n (–), upper canopy points n_{up} (···), and lower canopy points n_{low} (– – –); (b) sum of upper (–) and lower (– – –) canopy marks. The mark histogram (c) corresponds to $t = 200$.

If $\lambda=0$ in (3.1), then $m_i(t) \leq 1$ for all $i=1,2,\ldots$ and $t \geq 0$, whilst if $r=0$ (or $c=0$) then all points develop as independent logistic birth–death processes, with $m_i(t) > 1$ for all but the newest arrivals. This suggests the existence of a critical boundary $b(r, \lambda|c)$, such that for (r, λ) above the boundary (i.e. super-critical) both upper and lower canopies can co-exist, whilst conversely (i.e. sub-critical) the force of spatial interaction is too great to allow a subset of marks the opportunity of developing into an upper canopy. Consider a small lower canopy point j lying within the interaction disc of an upper canopy point i with mark $m_i \simeq K$, a fixed carrying capacity. Then the interaction radius $r(m_i + m_j) \simeq rK$, and the associated disc area is $\simeq \pi r^2 K^2$. So as α/μ is effectively an upper bound on the number of points contained in the unit torus, the expected number of points in the disc is at most $P = (\alpha/\mu)(\pi r^2 K^2)$. For

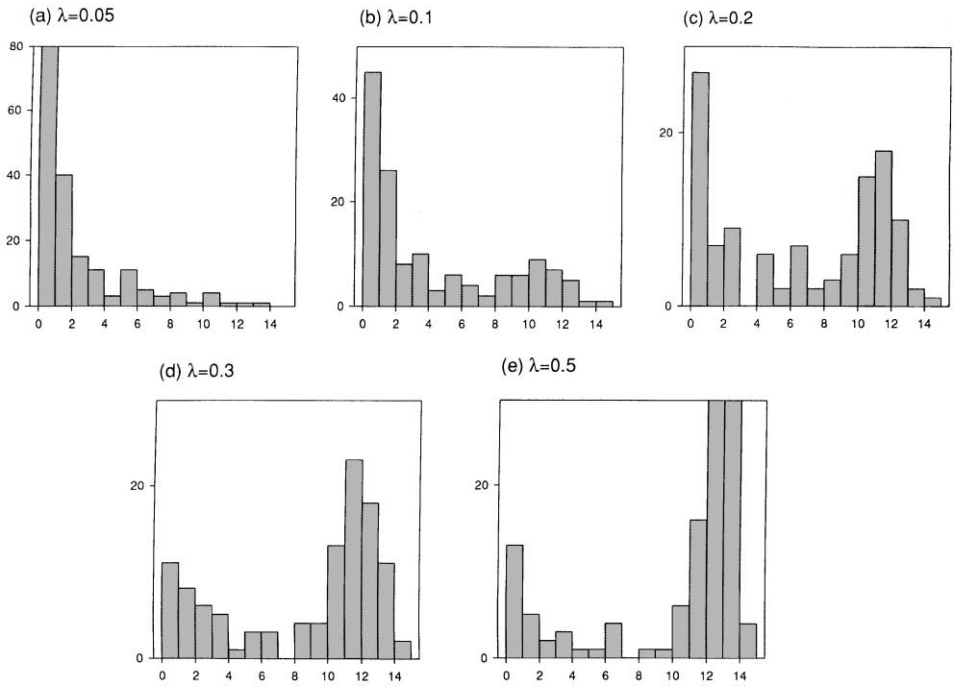

Fig. 3. Histograms of mark size at time $t = 500$ with $\alpha = 5$, $\mu = 0.02$, $K = 15$, $c = 0.2$ and $r = 0.005$ for λ equal to (a) 0.05, (b) 0.1, (c) 0.2, (d) 0.3 and (e) 0.5.

example, with $\alpha = 5$, $\mu = 0.02$, $r = 0.005$ and $K = 15$, we have $P = 4.42$ (say 4, given spatial effects). Now we see from (3.1) that in equilibrium

$$\lambda m_i(1 - m_i/K) - cP = 0, \tag{3.2}$$

which has roots

$$\tilde{m} = (K/2)[1 \pm \sqrt{(1 - 4cP/\lambda K)}]. \tag{3.3}$$

So a crude condition for a globally stable upper and a lower canopy to exist is that $4cP/[\lambda \min(K)] < 1$, i.e. on substituting for P,

$$\lambda > 4c\alpha\pi r^2 K/\mu, \tag{3.4}$$

which for our example yields $\lambda > 0.24$. Fig. 3 shows the histograms of mark size at time $t = 500$ (which is a sufficiently large burn-in period to ensure the process has converged to equilibrium) for $\lambda = 0$, 0.05, 0.1, 0.2, 0.3 and 0.5. These values contain 0.24, and we see that whilst (a) is J-shaped, i.e. no separate upper canopy exists, there is a slight semblance of a double canopy in (b), and this clearly gains in strength as λ increases ((c) to (e)). Note that the effect of spatial competition reduces the number of points from $\alpha/\mu = 250$ to 180, 139, 115, 112 and 122, respectively, and replacing α/μ in (3.4) by ~ 120 reduces the critical λ-value from 0.24 to 0.11, which shows remarkably good behavioural agreement with Fig. 3.

Here we have taken $\mu = 0.02 > 0$ to ensure a reasonable turn-round of points. However, when $\mu = 0$ points can die only through spatial inhibition, and the resulting

large increase in point numbers means that the equilibrium process takes far longer to develop. If we wish to be certain that equilibrium has been reached, development of the perfect simulation algorithm of Propp and Wilson (1996) for this spatial scenario could well prove useful (see Kendall and Møller, 2000).

At this point it is worth commenting on the role of the interaction distance $r(m_i + m_j)$; this will cover the maximum inter-point distance $1/\sqrt{2}$ when $r \geq 1/[\sqrt{2}(m_i + m_j)]$. So points in the upper canopy ($m_i \geq 10$) will affect all other points if $r \geq 1/10\sqrt{2} \simeq 0.07$. For our case with $r = 0.005$, two large marks with $m_i + m_j \sim 30$ will have interaction distance 0.15, with combined disc area 0.035 covering 1/28th of the torus; whilst one large and one small mark yield discs covering around 1/50th of the torus.

3.3. Estimation of the Gibbs interaction parameters

To develop summary Gibbs-type estimators we shall restrict ourselves to spatially stationary processes with pairwise interactions. The corresponding density function can be written as

$$f(\varphi;\mu,\theta) = \frac{1}{V(\mu,\theta)} \exp\left\{ -\sum_i \beta(m_i;\mu) - \sum_{i<j} \phi(||x_i - x_j||, m_i, m_j;\theta) \right\}. \quad (3.5)$$

The functions β and ϕ are as described after (2.2); the parameter vectors μ and θ are connected to the chemical activity and pair potential function, respectively.

Since we do not presume knowledge of the reaction component $\beta(m_i;\mu)$, to determine the strength and range of the underlying spatial interaction process we adopt a semi-parametric estimation approach. First, the mark space is fixed as $[D_0, D]$, where D_0 and D denote the minimum and maximum mark, respectively. This interval is then divided into q sub-intervals (here we fix $q = 10$); these are generally chosen to be of equal length, though this is not necessary. For the resulting partition $D_0 < D_1 < \cdots < D_q = D$, the value of $\beta(m;\mu)$ on $(D_{i-1}, D_i]$ equals μ_i. If we wished to use the equivalent pair potential function to that used in the generator (3.1), then we would take $\phi(||x_i - x_j||, m_i, m_j) = \theta$ if $||x_i - x_j|| \leq r(m_i + m_j)$ and 0 otherwise. However, in practice the interaction distance, essentially r, will be unknown, and so we take a generalisation of the Strauss disc model (2.4) as the model for interaction, in which the strength of spatial interaction between points varies according to the extent to which discs centred on the points, and with radius proportional to mark size, overlap. The pair potential function therefore takes the form

$$\phi(||x_i - x_j||, m_i, m_j) = \begin{cases} \theta_1 & \text{if } ||x_i - x_j|| \leq \gamma_1(m_i + m_j), \\ \theta_2 & \text{if not but } ||x_i - x_j|| \leq \gamma_2(m_i + m_j), \\ \quad \ldots\ldots & \\ \theta_v & \text{if not but } ||x_i - x_j|| \leq \gamma_v(m_i + m_j), \\ 0 & \text{otherwise,} \end{cases} \quad (3.6)$$

where, for appropriately chosen v, $\theta_1,\ldots,\theta_v$ are the interaction parameters to be estimated, and $\gamma_1 < \cdots < \gamma_v$ are fixed (i.e. given) radii. To estimate the parameters μ and θ we use the MPL method. The underlying technique is to maximise the log PL function with respect to μ and θ, i.e.

$$\begin{aligned}\log PL(\mu,\theta;\phi) = &- \sum_{[x,m]\in\varphi} \sum_{i=1}^{q} \mu_i I(m \in (D_{i-1}, D_i]) \\ &- \sum_{i=1}^{v} \theta_i \sum_{[x,m],[y,l]\in\varphi; x\neq y} I(\gamma_{i-1}(m+l) < ||x-y|| \leq \gamma_i(m+l)) \\ &- \frac{1}{|D-D_0|} \int_T \int_{D_0}^{D} \exp\left\{ - \sum_{i=1}^{q} \mu_i I(m \in (D_{i-1}, D_i]) \right. \\ &\left. - \sum_{i=1}^{v} \theta_i \sum_{[y,l]\in\varphi} I(\gamma_{i-1}(m+l) < ||y-\xi|| \leq \gamma_i(m+l)) \right\} \mathrm{d}m\,\mathrm{d}\xi \end{aligned} \tag{3.7}$$

(here $\gamma_0 = 0$). Since we are primarily interested in the interaction parameters θ_i, we first differentiate (3.7) with respect to μ_i and write the μ_i in terms of the θ_i. On inserting these into (3.7), $\log PL$ is a function of the θ_i only, which are then computationally amenable to estimation. Note that although standard software is available to enable fast computation of MPL estimates for such exponential models (see Section 2), specifically designed programs (as used here) are easy to construct and quick to use.

Suitable choice of the range and number of θ_i-segments will depend on the underlying process. The interaction distance (range of segments) can be estimated by using application specific information and, for example, the pair correlation function (see Goulard et al., 1996, Example 2, p. 373). Another possibility, which is chosen here, is to use the generalisation of the Strauss-disc process (3.6) since then it is not so crucial to know the interaction distance exactly: γ_v can be chosen to be so large that $\max\{\gamma_v(m_i + m_j)\}$ is greater than the (global) interaction distance. If it is much greater than that, then the estimates of the largest interaction parameters, θ_v, $\theta_{v-1},\ldots$, would be close to zero, which would indicate that there is virtually no interaction between points separated by such long distances. One way of choosing v would be to take a high initial value and then combine those θ-parameters which lead to closely aligned estimates.

In our case, Fig. 3 corresponds to $r = 0.005$, and so we might well choose to cover this by selecting $v = 4$ and 10, i.e. $\gamma_1 = 0.0025,\ldots,\gamma_4 = 0.01$ and $\gamma_1 = 0.001,\ldots,\gamma_{10} = 0.01$. MPL-estimates for the process of Fig. 3, evaluated at time $t = 200$ for single simulation runs, are shown in Table 1, and even the 4-parameter summary (a) clearly contains considerable information. For all 5 λ-values there is clear evidence for inhibition at distances $||x_i - x_j|| < 0.005(m_i + m_j)$, with γ_1 and γ_2 both increasing as λ decreases, i.e. as the upper canopy ceases to dominate. Note that $\gamma_1 > \gamma_2$, with the γ_1-estimates showing a strong increase in value as $\lambda \to 0$. The 10-parameter summary (Table 1b) clearly gives more quantitative detail, but given that the initial

Table 1
MPL estimates of the interaction parameters θ_i of the pairwise interaction process with pair potential structure (3.6), for the process of Fig. 3, evaluated at time $t = 200$: (a) $v = 4$ and (b) $v = 10$

(a) $v = 4$			γ_i	
λ	0.0025	0.0050	0.0075	0.010
0.05	> 10	1.210	−0.294	0.035
0.1	> 10	0.748	−0.220	0.101
0.2	2.503	0.211	−0.055	−0.084
0.3	1.177	0.154	−0.010	−0.040
0.5	0.711	0.117	−0.046	−0.058

(b) $v = 10$					γ_i					
λ	0.001	0.002	0.003	0.004	0.005	0.006	0.007	0.008	0.009	0.010
0.05	> 10	> 10	> 10	2.067	0.597	−0.665	0.052	−0.106	0.056	0.020
0.1	> 10	> 10	> 10	1.438	0.222	−0.565	0.040	0.141	0.199	−0.041
0.2	> 10	2.562	1.656	0.314	−0.084	−0.113	−0.087	−0.151	−0.052	−0.057
0.3	1.547	1.689	0.800	0.874	−0.349	−0.278	−0.065	−0.110	−0.110	0.110
0.5	1.080	0.834	0.258	0.290	−0.045	−0.186	0.015	0.080	−0.058	−0.083

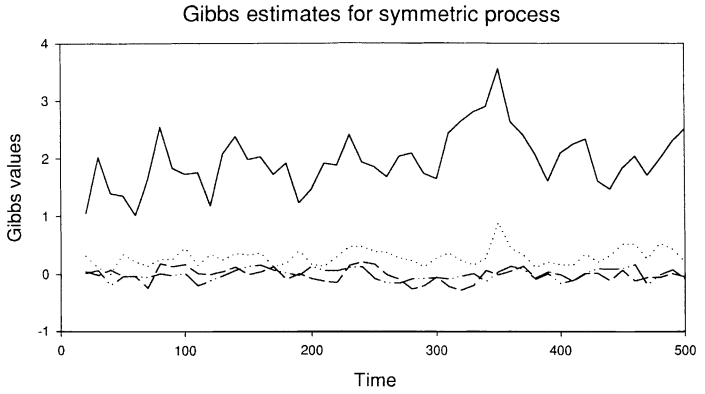

Fig. 4. Estimates θ_1 (–), θ_2(· · ·), θ_3 (– – –) and θ_4 (– · · · –) for the process corresponding to Fig. 3c, shown at times $t = 20, 30, \ldots, 500$.

objective is to present minimal summaries, taking $v = 4$ is reasonably "optimal" for this case. Note how the $v = 10$ case for γ_5 fades as λ increases beyond the threshold value 0.1: this is unlikely to be due to the degredation in performance of the MPL method as interaction increases, since in Goulard et al. (1996) the largest θ-value was ~ 2, and even that relatively high value did not lead to any overestimation.

Fig. 4 shows the temporal development of the Gibbs estimates over $t = 20, 30, \ldots,$ 500 for $\lambda = 0.2$ and $v = 4$, and demonstrates apparent oscillatory behaviour. We shall not pursue a time-series analysis of the γ_i here, but simply note that $\gamma_1 > 0$ dominates $\gamma_2 > 0$ (as above), with 95% confidence intervals (after allowing for a burn-in period

of length 200) of 2.124 ± 0.084 and 0.316 ± 0.030, respectively. In contrast, γ_3 is mildly negative (-0.045 ± 0.022) and γ_4 is effectively zero (-0.0082 ± 0.0186). Thus there is very strong evidence for inhibition, with $r = 0.005$ being clearly identified.

4. Asymmetric simulation experiment

If we choose $h(\cdot)$ to be non-symmetric, then the parallel pair potential function will not take the usual Gibbs form. So suppose we use instead a symmetric potential function that is of Gibbs form but does not correspond exactly to the generating space-time process. In Section 3 we saw that if we know the (symmetric) generating mechanism, then we are able to capture it by using the corresponding Gibbs process. Thus here we study whether we can capture an (asymmetric) generating mechanism by employing a different form for the pairwise interaction process. This example is of fundamental importance, since it demonstrates that we can infer spatial competition structure even when we do not know its underlying form.

4.1. Model (3.6)

Given the success of using a (non-temporal) Gibbs model with pair potential function (3.6) for simulations developed through the temporal form (3.1), which employs (3.6) as its spatial interaction component, i.e. in (2.1) we take

$$h(m_i(t), m_j(t); ||x_i - x_j||) = -cI(||x_i - x_j|| < r(m_i(t) + m_j(t)), \tag{4.1}$$

a key question to ask is whether the same pair potential function will still produce meaningful results when it is applied to a different type of spatial interaction. A natural first choice is to adapt the process (2.1), with (2.5) replacing (4.1), since theoretical spectral properties are known for the (approximating) linearised lattice process. Let us therefore take $a(\cdot)$ in (2.5) to be the indicator function $a(||x_i - x_j||) = b$ if $||x_i - x_j|| \leq B$ (say) and 0 otherwise, whence (3.1) becomes

$$\tilde{m}_i = m_i + \lambda m_i(1 - m_i/K)h + b \sum_{j=1, j\neq i}^{n} (m_i - m_j) I\{||x_i - x_j|| < B\}h. \tag{4.2}$$

Note that we could replace $a(||x_i - x_j||)$ by any (non-marked) pair potential function which is either a step function or a continuous function of the distance $||x_i - x_j||$ (see, for example, Stoyan et al., 1995).

In spite of some obvious similarities between (3.1) and (4.2), fundamental differences do exist between them. First, whilst the interaction function in (3.1) is symmetric in $m_i(t)$ and $m_j(t)$, that in (4.2) is asymmetric. Second, in (4.1) (which is the interaction function used in (3.1)) the interaction radius varies according to the weighted mark sum $r(m_i(t) + m_j(t))$ and the interaction coefficient $-c < 0$ is constant; whilst in (2.5) the interaction radius B is constant, and the interaction coefficient varies as the weighted mark difference $b(m_i(t) - m_j(t))$ and so may take positive or negative values. Thus in (4.1) the presence of an interacting neighbour at x_j deflates $\mathrm{d}m_i(t)/\mathrm{d}t$ by c, no matter how small $m_j(t)$ is, and the interaction radius

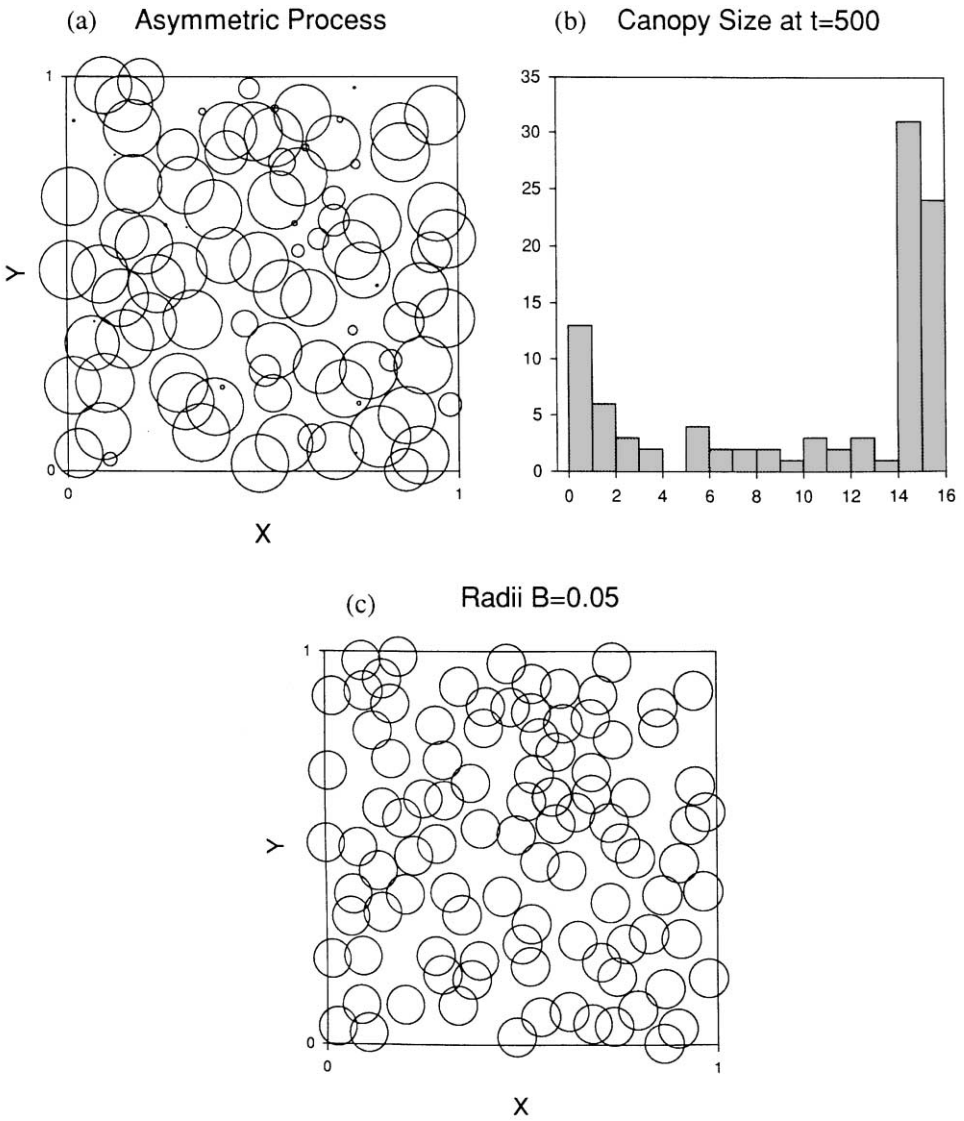

Fig. 5. Simulation of the asymmetric process (4.2) shown at time $t = 500$ with $\alpha = 5$, $\lambda = 0.2$, $\mu = 0.02$, $K = 15$, $b = 0.2$ and $B = 0.05$: (a) bubble plot radius proportional to rm_i (to enable direct comparison with Fig. 1); (b) mark histogram; (c) bubble plot radius proportional to B.

centred on x_i is at least $rm_i(t)$. Whilst in (2.5) the same neighbour changes $\mathrm{d}m_i(t)/\mathrm{d}t$ by the mark-dependent amount $b(m_i(t) - m_j(t))$, and the interaction radius B remains constant.

Fig. 5 shows the result of a simulation run at $t = 500$ for $\alpha = 5$, $\mu = 0.02$, $\lambda = 0.2$ and $K = 15$ (as above) with $b = 0.2$ and $B = 0.05$. On comparing the histogram plots of mark size (Fig. 5b) with that of the earlier process (Fig. 2c), we immediately see that although both exhibit upper and lower canopies their fine detail differs. For in

the latter process all interacting points experience inhibition, and the resulting upper canopy histogram is roughly Normal with all the $m_i \leq K$; whilst in the former case the most dominant marks experience enhancement and this is reflected in a large number of marks both just below and above K. In Fig. 5a the bubble plot radius equals bm_i, to enable marks to be compared with Fig. 1; in Fig. 5c the radius equals B, so interacting points are those whose discs overlap. If required, the impression that under model (3.1) points interact with more neighbours than under (4.2) could be checked by comparing the number of overlapping discs. There is a good reason why they should, since, unlike model (3.1), established marks ($m_i \simeq K$) will suppress nearby newly arrived ones ($m_i \simeq 1$).

On comparing

$$h_2(m_i(t), m_j(t); ||x_i - x_j||) = b(m_i(t) - m_j(t))I(||x_i - x_j|| < B) \tag{4.3}$$

with $h(\cdot)$ in (4.1), we may argue the estimation procedure as follows. First, dominant marks are unlikely to decline except through death (at rate μ). So spatial inhibition will exert its main effect through the inhibition of small marks $m_i \simeq 0$ by large marks $m_j \simeq K$. Thus we may approximate (4.3) by

$$h_2(m_i(t), m_j(t); ||x_i - x_j||) \simeq -bKI(||x_i - x_j|| < B), \tag{4.4}$$

whence comparison with (4.1) yields $bK \simeq c$ and $B \simeq Kr$. On taking $\gamma_i = 0, 0.001, 0.002, \ldots, 0.010$ and paralleling the analysis over $t = 10, 20, \ldots, 500$, with the same burn-in period of 200, we see that almost all the γ_1-values lie at or near to the enforced upper bound of 20. Although this contrasts markedly with $\hat{\theta}_1$ shown in Fig. 4, it simply implies that for this non-Gibbs process points seldom lie within $\gamma_1 = 0.001(m_1 + m_j) \simeq 0.03$ (2 large), 0.015 (1 large, 1 small) and 0.002 (2 small) of each other. The 99% confidence intervals (following burn-in) for the θ_i are

$$\theta_2\text{: } 1.67 \pm 0.29,\ \theta_3\text{: } 0.28 \pm 0.14,\ \theta_4\text{: } 0.15 \pm 0.12,\ \theta_5\text{: } 0.16 \pm 0.11,$$
$$\theta_6\text{: } 0.08 \pm 0.08,\ \theta_7\text{: } 0.08 \pm 0.06,\ \theta_8\text{: } 0.09 \pm 0.06,\ \theta_9\text{: } 0.05 \pm 0.05,$$
$$\theta_{10}\text{: } 0.00 \pm 0.05,$$

and they confirm an effective r-value of around $B/K = 0.003$, as required. Note that this is a comparative, and not a truly statistical, inference, given the dependence between successive estimates in our single realisation. It is not surprising that this inference is less clear than that for the Gibbs process, since as seen above the true interaction function is based upon $m_i - m_j$, whilst the assumed 'estimation' function is based on the radically different form $m_i + m_j$. Nevertheless, the technique still provides an extremely useful ball-park estimator.

4.2. Another symmetric Gibbs model

If, however, we have information that the interaction radius does not depend on marks (unlike (3.6)), and the true interaction function is indeed based upon $m_i - m_j$, there is no reason why we cannot replace (3.6) by

$$\phi(||x_i - x_j||) = \theta_1|m_i - m_j| \quad \text{if } ||x_i - x_j|| \leq \gamma_1, \text{ etc.} \tag{4.5}$$

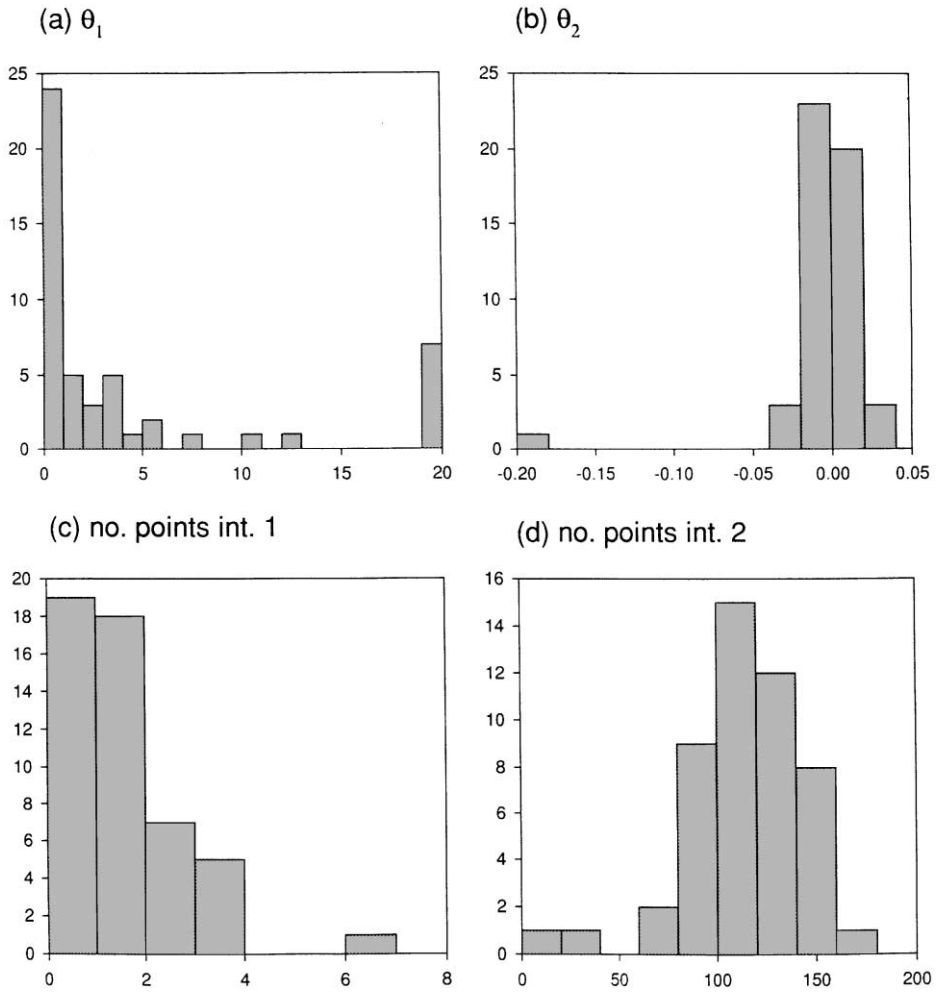

Fig. 6. Histograms of estimates θ_1 and θ_2 for Fig. 5, with the pair potential function (4.5): (a) $\hat{\theta}_1$ and (b) $\hat{\theta}_2$; number of points in (c) inner disk (i) and (d) outer annulus (ii).

Note that if $m_i \simeq 0$, $m_j \simeq 0$ or K then both $m_i + m_j$ and $|m_i - m_j| = 0$ or K; whilst if $m_i, m_j \simeq K$ then $m_i + m_j \simeq 2K$ but $|m_i - m_j| \simeq 0$. So interacting large marks affect the two estimation procedures very differently. Indeed, under the first regime the estimation process places greatest weight on two interacting large marks, whilst in the second regime maximum emphasis is placed on a large mark inhibiting a small mark. Moreover, whereas in the first regime mark size affects the range of interaction, now mark difference affects the strength of interaction. Fig. 6 shows the result of using this new pair potential function on these same simulated data over $t = 10, 20, \ldots, 500$ for just 2 disk radii, namely (i) $\phi = \theta_1 |m_i - m_j|$ if $||x_i - x_j|| \leq 0.05$, and (ii) $\phi = \theta_2 |m_i - m_j|$ if not but $||x_i - x_j|| \leq 0.1$. Figs. 6a and b show the distribution of the 500 θ_1- and θ_2-estimates, and there is clearly very strong interaction over distances up to the fixed interaction radius 0.05, and almost no

interaction thereafter. The values 20 in Fig. 6a occur when there are no point-pair distances in disk (i). Even though the value of θ_1 shows considerable variation, it is always clearly positive (minimum $\simeq 0.25$) and larger than θ_2, which is always very close to zero. There are two reasons for the variation in θ_1. First, the more pairs of points there are within distance 0.05, the closer θ_1 is to zero (weaker inhibition). Second, the estimates of θ_1 vary considerably for a given number of close pairs, due to variation in mark difference. The value $\hat{\theta}_2 = -0.18$ in Fig. 6b corresponds to the initial observations at $t = 10$; $\hat{\theta}_1$ and $\hat{\theta}_2$ converge quickly thereafter (the number of contributing pairs stabilises after $t = 40$). Figs. 6c and d show the respective number of contributing pairs, and given the small number of contributing pairs in the former case 0.05 is the smallest value one can reasonably take here.

Dividing the interval 0.05–0.1 into subintervals to enable the interaction to be studied more carefully produces the same result, namely that interaction drops after 0.05 (there are too few interacting pairs to split the first interval). Specifically, after allowing for a burn-in period of length 50, 99% confidence intervals corresponding to the intervals 0.05–0.06, …, 0.09–0.10 for $\theta_2, \ldots, \theta_6$ are given by

$$0.010 \pm 0.014,\ 0.004 \pm 0.015,\ 0.005 \pm 0.010,\ 0.001 \pm 0.012, -0.003 \pm 0.009. \tag{4.6}$$

Clearly no advantage is gained by increasing the number of θ parameters from 2 to 6.

5. Conclusions

Because of its simplicity, the Strauss-type disc process (2.4) is a natural choice for the interaction process in the Gibbs model which underlies our initial estimation procedure, if there are interactions depending on continuous marks. Equally, it is a natural choice for the interaction component in the space–time stochastic process (3.1). This model can be used, for example, to describe interaction between trees (discs or marks being influence zones of trees; see Goulard et al., 1996) or between territorial animals. We have shown it produces similar bimodality of canopy structure as that observed by Ford and Diggle (1981). Table 1 not only highlights the way in which the parallel Gibbs model catches the interaction structure of the generating process, namely $r = 0.005$, but also demonstrates that very few statistical indicators (here at most 3 or 4) are needed to capture the essence of the process. This is in total contrast to the relatively large number of autocorrelation or periodogram estimators that would be needed to characterise the process in terms of time-series type estimators, and is far more concise than the production of successive graphplots resulting from K-function analyses. Fig. 4 demonstrates the variability exhibited by our indicators over time, and highlights the inherent danger of placing too much credence in snap-shot analyses taken at a single-time point.

In our second model, (4.2), the range of interaction does not depend on the marks, but only on the interpoint distance. Furthermore, the strength of interaction relates to

mark difference; relevant scenarios include the growth of agricultural crops, colonising plants, and naturally regenerating forests. Even though this process is substantially different from the first one, (3.1), the Strauss-disc model is still able to describe the interaction structure quite well. According to the estimated model, there is interaction up to the distance $0.003(m_i + m_j)$, which corresponds to the fixed interaction radius 0.05 (for a large mark inhibiting a small one) used in the simulated realisation. This shows that the Strauss-type procedure is robust, i.e. it is not crucial to know the exact process generating mechanism.

However, if we do possess additional information, here (expression (4.5)) that the strength of interaction is based upon $m_i - m_j$, and that the interaction distance is constant, then the procedure becomes far more effective. The estimates ((4.6) and Fig. 6) show a clear cut-off at distance 0.05: as before, note the variability in the estimates $\hat{\theta}_1$.

Note that the model (4.5) does not have exactly the same interaction structure as the pattern generating process. In the estimation process the interactions are symmetric, whilst in the generating process they are asymmetric: in (4.5) two interacting marks will result in the larger being enhanced and the smaller being diminished, whilst the estimation procedure assumes that both will be diminished. As hinted earlier, matching the estimation and generating processes necessitates the use of the non-symmetric pair potential function (2.5), which is not admissible under the usual Gibbs process framework.

One possible solution might be to utilise the time-dependent nature of the data by adapting the approach of Guyon (1995) (see also Bayomog, 1996; Högmander and Särkkä, 1999) applied to Markov random fields, since this allows non-symmetric interactions through time (as opposed to space). In this we could replace (2.5) by

$$\phi(||x_i - x_j||, m_i(t), m_j(t-1)) = a(||x_i - x_j||)\{m_i(t) - m_j(t-1)\}, \tag{5.1}$$

which means that we are now much closer to the true generating mechanism. However, for most marks $m_j(t)$ and $m_j(t-1)$ will be broadly similar, whence on inserting (5.1) into (3.5) we see that the sum of the pair potentials $\phi(x_i, x_j, m_i(t), m_j(t-1))$ and $\phi(x_i, x_j, m_i(t-1), m_j(t))$ lies near to zero, which means that only weak interaction estimates can be obtained. Note that if we relate points at the two times $t-s$ and t, then individuals may die and arrive inbetween times. This does not present any analytic difficulty, but as s increases we diverge away from the original generating mechanism. Using (5.1) with the time delays $s = 1$, 2, 5, 10 and 20 and the two interaction parameters 0–0.05 and 0.05–0.1 produced effectively zero estimates, so an alternative strategy is clearly necessary. A promising line of attack is to acknowledge the reality that the right-hand side of (4.2) drives the first derivative of $m_i(t)$. So we could consider the pattern as being a two-mark process in which the pair potential function is based on $\phi(x_i, x_j, m_i(t), m_j(t), m_i'(t), m_j'(t))$ where $m_k'(t)$ denotes, for example, $m_k(t) - m_k(t-1)$. This work is currently under investigation and will be reported elsewhere.

Acknowledgements

The financial support of the Adlerbertska Foundation, the Swedish Natural Science Research Council and the Royal Swedish Academy of Sciences is highly appreciated. This work is part of the EU grant ERB-FMRX-CT96-0096, and was partially completed whilst Aila Särkkä was visiting the Department of Biostatistics at Emory University.

References

Baddeley, A., Møller, J., 1989. Nearest–neighbour Markov point processes and random sets. Int. Statist. Rev. 57, 89–121.

Baddeley, A.J., Møller, J., Waagepetersen, R., 1999. Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. Statist. Neederlandica, to appear.

Baddeley, A., Turner, T.R. 2000. Practical maximum psuedolikelihood for spatial point patterns. Austral. New Zealand J. Statist. to appear.

Bartlett, M.S., 1975. The Statistical Analysis of Spatial Pattern. Chapman & Hall, London.

Bayomog, S., 1996. Estimation of a Markov field dynamic. Prépublication du SAMOS No. 52, Université Paris 1, Pantheon Sorbonne.

Daley, D.J., Vere-Jones, D., 1988. An Introduction to the Theory of Point Processes. Springer, New York.

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

Fiksel, T., 1984. Estimation of parameterized pair potentials of marked and non-marked Gibbsian point processes. Electron Inform. Kybernet 20, 270–278.

Ford, E.D., 1975. Competition and stand structure in some even-aged monocultures. J. Ecol. 63, 311–333.

Ford, E.D., Diggle, P.J., 1981. Competition for light in a plant monoculture modeled as a spatial stochastic process. Ann. Bot. 48, 481–500.

Ford, E.D., Renshaw, E., 1984. The interpretation of process from pattern using two-dimensional spectral analysis: modeling single species patterns in vegetation. Vegetatio 56, 113–123.

Geyer, C.J., Møller, J., 1994. Simulation procedures and likelihood inference for spatial point processes. Scand. J. Statist. 21, 359–373.

Geyer, C.J., Thompson, E.A., 1992. Constrained Monte Carlo maximum likelihood for dependent data (with discussion). J. Roy. Statist. Soc. B 54, 657–699.

Geyer, C.J., Thompson, E.A., 1995. Annealing Markov chain Monte Carlo with applications to ancestral inference. J. Amer. Statist. Assoc. 90, 909–920.

Goulard, M., Grabarnik, P., Särkkä, A., 1996. Parameter estimation for marked point processes through the maximum pseudo-likelihood method. Scand. J. Statist. 23, 365–379.

Guyon, X., 1995. Random Fields on a Network. Springer, New York.

Högmander, H., Särkkä, A., 1999. Multitype spatial point patterns with hierarchical interactions. Biometrics 55, 1051–1058.

Jensen, J.L., Møller, J., 1991. Pseudolikelihood for exponential family models of spatial point processes. Ann. Appl. Probab. 3, 445–461.

Kendall, W.S., Møller, J., 2000. Perfect Metropolis-Hastings simulation of locally stable point processes. Adv. Appl. Probab., to appear.

Nielsen, L.S., Jensen, E.B.V., 2000. Inhomogeneous Markov point processes by transformation. Bernoulli, to appear.

Ogata, Y., Tanemura, M., 1985. Estimation of interaction potentials of marked spatial point patterns through the maximum likelihood method. Biometrics 41, 421–433.

Ogata, Y., Tanemura, M., 1986. Likelihood estimation of interaction potentials and external fields of inhomogeneous spatial point patterns. In: Francis, I.S., Manly, B.F.J., Lam, F.C. (Eds.), Proceedings of Pacific Statistical Congress 1985, pp. 150–154.

Penttinen, A., 1984. Modelling interaction in spatial point patterns: parameter estimation by the maximum likelihood method. Jyväskylä Studies in Computer Science, Economics and Statistics, Vol. 7, University of Jyväskylä.

Propp, J.G., Wilson, D.B., 1996. Exact sampling with coupled Markov chains and applications to statistical mechanics. Random Struct. Algorithms 9, 223–252.

Renshaw, E., 1984. Competition experiments for light in a plant monoculture: an analysis based on two-dimensional spectra. Biometrics 40, 717–728.

Renshaw, E., 1994a. The linear spatial-temporal interaction process and its relation to $1/\omega$-noise. J. Roy. Statist. Soc. B 56, 75–91.

Renshaw, E., 1994b. Non-linear waves on the Turing ring. Math. Sci. 19, 22–46.

Renshaw, E., Ford, E.D., 1983. The interpretation of process from pattern using two-dimensional spectral analysis: methods and problems of interpretation. Appl. Statist. 32, 51–63.

Renshaw, E., Ford, E.D., 1984. The description of spatial pattern using two-dimensional spectral analysis. Vegetatio 56, 75–85.

Renshaw, E., Phayre, S., Jakeman, E., 2000. The development of space-time interaction processes with given spectral structure. Inverse Problems, to appear.

Ripley, B.D., 1988. Statistical Inference for Spatial Processes. Cambridge University Press, Cambridge.

Särkkä, A., 1993. Pseudo-likelihood approach for pair potential estimation of Gibbs processes. Jyväskylä Studies in Computer Science, Economics and Statistics, Vol. 22, University of Jyväskylä.

Smith, M.A., Renshaw, E., 1994. Parallel-prefix remapping for efficient data-parallel implementation of unbalanced simulations. Adv. Parallel Comput. 9, 215–222.

Stoyan, D., 1989. Statistical inference for a Gibbs point process of mutually non-intersecting discs. Biometrical J. 31, 153–161.

Stoyan, D., Kendall, W.S., Mecke, J., 1995. Stochastic Geometry and its Applications, 2nd Edition. Wiley, Chichester.

Stoyan, D., Stoyan, H., 1998. Non-homogeneous Gibbs process models for forestry — a case study. Biometrical J. 40, 521–531.

Takacs, R., Fiksel, T., 1986. Interaction pair potential for a system of ant's colonies. Biometrical J. 28, 1007–1013.

Turing, A.M., 1952. The chemical basis of morphogenesis. Philos. Trans. Roy. Soc. London B 237, 37–72.

Whittle, P., 1954. On stationary processes in the plane. Biometrika 41, 434–449.