Available online at www.sciencedirect.com
ScienceDirect
Computational Statistics & Data Analysis 51 (2006) 1698–1718
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

The analysis of marked point patterns evolving through space and time

Aila Särkkä[a], Eric Renshaw[b,*]

[a]*Department of Mathematical Statistics, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden*
[b]*Department of Statistics and Modelling Science, University of Strathclyde, Livingstone Tower, 26 Richmond Street, Glasgow G1 1XH, UK*

Received 12 May 2005; received in revised form 6 July 2006; accepted 7 July 2006
Available online 28 July 2006

Abstract

A maximum pseudo-likelihood approach has previously been developed for fitting pairwise interaction models to patterns generated by growth–interaction processes that are sampled at *fixed time* points. This approach is now extended, not only by estimating the parameters of the process *through time*, but also by employing least squares estimation since likelihood based approaches are much more computationally demanding. First, simple stochastic models are used to demonstrate that least squares methods are as powerful as likelihood-based approaches, as well as being mathematically and computationally simpler. The algorithm generates simulations of the deterministic growth–interaction and stochastic immigration–death process, and through these the parameter estimates are determined. Logistic and linear growth are then combined with (symmetric) disc-interaction and (asymmetric) area-interaction processes, and between them these generate a variety of mark–point spatial structures. A robustness study shows that the procedure works well in that the presence, structure and strength of a growth–interaction process can be determined even when an incorrectly presumed model is employed. Thus, the technique is likely to prove to be very useful in general practical applications where the underlying process generating mechanism is almost certain to be unknown. Finally, the procedure is applied to the analysis of a new Swedish pine forest data set for which tree location and diameter at breast height were recorded in 1985, 1990 and 1996.

Keywords: Area interaction; Disc interaction; Growth–interaction process; Least squares estimation; Linear growth; Logistic growth; Robustness; Space–time processes; Swedish pine data

1. Introduction

In recent years there has been a surge of interest in spatio-temporal modelling, due in no small measure to the large amount of data generated by environmental scientists studying pollution and global climate monitoring through geographical information systems, remote sensing platforms, monitoring networks and computer simulation models. As a consequence, substantial progress has been made in the parallel development of methods of analysis involving geostatistical, hierarchical and multivariate time series approaches, together with implementation of space–time dynamic models and point process models.

Geostatistical models (e.g., Kyriakidis and Journel, 1999) are used to study underlying continuous spatial processes which are observed only at a finite set of locations, and typically yield space–time trends of some phenomenon of

* Corresponding author. Tel.: +44 141 548 3591; fax: +44 141 552 2079.
E-mail addresses: aila@math.chalmers.se (A. Särkkä), eric@stams.strath.ac.uk (E. Renshaw).

0167-9473/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2006.07.008

interest. Whilst multivariate time series methods (e.g., Bennett, 1999) involve the construction of specific space–time dynamic models. Brown et al. (2001), for example, consider a high-dimensional multivariate state-space time series model in which the cross-covariance structure is derived from the spatial context of the component series; as such its interpretation is essentially independent of the spatial locations at which the data are recorded. Algorithms are developed for estimating the parameters of various models by maximum likelihood, including a non-separable model due to Brown et al. (2000), and applied to a radar-rainfall data set. Both methodologies generate their own specific problems. Geostatistical approaches require complete specification of the joint space–time covariance structure, yet realistic covariance functions can be difficult to specify and implement. Whilst time-series methods specify dynamic models that are linked spatially, and it can be difficult to predict what happens at unmonitored sites because of the lack of a continuous spatial component in the model structure. However, geostatistical and time-series approaches can be combined in a statistical model that is both temporarily dynamic and spatially descriptive. Such *space–time dynamic models* have been considered by, for example, Goodall and Mardia (1994), Guttorp et al. (1994), Mardia et al. (1998), Meiring et al. (1998), and Wikle and Cressie (1999). Goodall and Mardia (1994) also provide an early review of spatio-temporal modelling, and outline an approach based on a general spatio-temporal state space process designed to model the evolution of spatial fields through time. Estimation is recursive, based on the Kalman filter, and is central to Mardia et al. (1998) who combine Kriging (spatial statistics) with the Kalman filter (multivariate time series) and thereby obtain a powerful modelling strategy called the Kriged Kalman filter. Wikle and Cressie (1999) also consider the spatio-temporal Kalman filter, using it to reduce dimension in the analysis of large spatio-temporal data sets. Wikle (2001) describes a parallel approach for modelling complicated dynamical spatio-temporal processes, based on a stochastic integro-difference equation, where the redistribution kernel is allowed to vary with space and/or time. Xu et al. (2005) utilise a similar approach within a hierarchical Bayesian framework. Whilst Stroud et al. (2001) model the mean function of the quantity of interest at each time point as a locally weighted mixture of regression surfaces which are then allowed to evolve through time. MacNab and Dean (2001) propose general additive mixed models for the analysis of geographic and temporal variability of mortality rates. Their aim is to identify temporal trends and produce series of smoothed maps from which spatial patterns of mortality risks can be monitored over time.

In environmental science the processes studied are often highly complex, involving multiple sources of data from a variety of different platforms, as well as various degrees of additional scientific knowledge. Such complexity is exacerbated by the difficulty of employing a full likelihood approach, though this can be ameliorated by using pseudo-likelihood procedures (e.g., Renshaw and Särkkä, 2001 (henceforth denoted as RS)). A related approximation technique involves extending Veccia's (1988) conditional density approach based on partial observations to space–time processes (Jones and Zhang, 1997); the nearest-neighbour conditioning vector is based on preliminary estimates of the space–time correlation function. Rather than specifying joint multivariate spatio-temporal covariance structures it may be much easier to factor the joint distributions into a series of conditional models and then link these together in a hierarchical (usually Bayesian) framework. Wikle (2003) reviews the use of such models, focussing on the spatio-temporal modelling of the intensity surface based on observations at fixed spatial locations. Hierarchical spatio-temporal models are also commonly employed in other disciplines, such as disease mapping. Waller et al. (1997), for example, extend the spatial models of Besag et al. (1991) to accommodate temporal effects, as well as space–time interactions, thereby providing a hierarchical framework for modelling regional disease rates over space and time. Whilst Knorr-Held and Richardson (2003) analyse space–time surveillance of meningoccocal disease for time-series of disease counts. Hierarchical space–time models have also been applied to mapping of the human brain by means of functional magnetic resonance images (Gössl et al., 2001). Lawson (2001) discusses modelling of spatio-temporal clustering in disease and space–time processes in relation to pollution sources and space–time scan statistics. Moreover, descriptive spatio-temporal statistical methods are also highly relevant in climatology (Wikle, 2002). Methods used include: evaluating empirical orthogonal functions; principal oscillation pattern analysis; spatio-temporal canonical correlation analysis; and space–time spectral analysis to infer relationships between spatial and temporal scales of variability.

Now most of the above studies are essentially motivated by the desire to make predictions of inferences about space–time structure based on sampling at fixed locations. In many ecological situations, however, the locations of the measurements play a fundamental role in the process generating mechanism, and so employing fixed location sampling strategies will result in a considerable loss of information. So if both the marks and their point locations develop in time, a far better approach would be to base estimation and inference strategies on spatio-temporal mark–point process

models. To date, spatio-temporal point process modelling has mainly been concentrated on detecting and describing space–time clustering and estimating the intensity of spatio-temporal point processes. Knox (1964) develops a test of significance for space–time clustering. Williams (1984) reviews later work on Knox's test and some related approaches. Diggle et al. (1995) extend existing second-order methods for purely spatial point process data to the spatio-temporal setting. Whilst Fenton et al. (2004) consider the effect of the under-reporting of cases on the space–time K-function. Intensity surfaces are often determined by a combination of a purely spatial population intensity and a risk surface that depends both on space and time, the aim being to identify areas with increased risk surface. Brix and Møller (2001), for example, study local spatial variation in weeds by using inhomogeneous log-Gaussian Cox process models, and they extend these to spatio-temporal models called log-Gaussian Cox birth processes. Whilst Brix and Diggle (2001) consider log-Gaussian Cox processes, where the underlying Gaussian process is a spatial Ornstein–Uhlenbeck process in order to monitor changes in the space–time incidence pattern of gastro-intestinal infections.

Our aim is to develop and apply models for marked point processes evolving *through time*, a key application being to further the study of tree growth in forests. In a purely spatial context, marked point processes have, of course, already been applied to forestry (e.g., Penttinen et al., 1992; Gavrikov and Stoyan, 1995; Stoyan and Penttinen, 2000). Whilst theoretical developments of spatial marked point processes can be found in Stoyan and Stoyan (1994) and Stoyan et al. (1995). Schlather et al. (2004) introduce two characteristics for stationary and isotropic marked point processes that can be used to study mark–point interactions and introduce a test in order to decide whether the marks are independent of the point locations.

The management of modern multipurpose forests, which often involve uneven-aged and mixed species stands, has an essential need of accurate and generally applicable single-tree growth functions. A large number of growth-yield and ecological models have been developed to describe and model forest dynamics assuming individual tree information (e.g., Botkin et al., 1972; Shugart and West, 1977; Pacala et al., 1993; Mickelson et al., 1998; Nyström, 2001) but only a few studies take distance-dependent spatial tree information into account. Though some simulation models have been developed (e.g., Pukkala 1989a,b; Pukkala and Kolström, 1991), they suffer the drawbacks that tree locations and sizes are not allowed to develop simultaneously, and that the underlying spatial growth models are based on so-called competition indices. The latter are computed 'separately' via the size and number of competing neighbours, but the actual locations of neighbours are not taken into account.

So given that geostatistical-type approaches to the analysis of marked point processes clearly have to be questioned (Wälder and Stoyan, 1996), all of the above concerns raise several serious issues. First, we have to recognise that many situations will involve a genuine dependence between point locations and their associated mark variables. For example, the relative positions of trees in a forest directly affect competition for light or nutrient; so any geostatistical analysis of tree sizes which does not allow for possible dependencies between points and marks must be questioned (Schlather et al., 2004). Second, in real life many spatial structures develop through time, and so point/mark development has to reflect this, rather than presuming some kind of purely spatial structure: whilst any spatial pattern can be expressed *mathematically* in terms of such a simultaneous or conditional scheme, the actual *process generating mechanism* will often be temporal. Third, only rarely will we know the exact process generating mechanism, so we need to generate a general and flexible stochastic model that can act as a realistic surrogate, yet have the potential for generating simulated events quickly. Fourth, as time progresses, new locations may be created and old ones destroyed, as happens, for example, in forest growth. The aim of this paper is therefore first to develop space–time models (Section 2) that satisfies each of these four attributes. Although one of these models was developed in RS, namely the symmetric interaction process with logistic growth function, we introduce a new non-symmetric interaction process which is more realistic in forestry situations. Since all marks may (potentially) interact with all others, we need to ensure that the models are as simple as possible in order to keep the level of spatial complexity, and hence the speed of computation, under reasonable control. For given the inherent level of mathematical intractability associated with marked point structures, we are forced into using a simulation based approach (Section 4). However, before we use this structure to develop a full spatio-temporal estimation procedure (Section 5), we first examine and compare likelihood and least squares estimation strategies (Section 3) for much simpler processes in order to demonstrate that the latter, whilst being much simpler to employ than the former, do not result in any appreciable loss of information. Note that the estimation procedure presented in this paper is very different from that in RS, where we used maximum pseudo-likelihood to estimate the interaction parameters of a purely spatial pairwise interaction model fitted separately at the observed time points. Here, the parameters of the full space–time model are estimated over all given time points simultaneously. Finally, we use our space–time estimation technique to analyse a real (as opposed to simulated) data set (Section 6).

2. The basic spatio-temporal stochastic model

RS show that Gibbs point processes yield useful summary measures from which we can infer underlying generating mechanisms of space–time stochastic processes. They base their analysis around an immigration–growth–spatial interaction process on the unit torus T, in which the ith individual to enter the system is of size (mark) $m_i(t)$, and is located at x_i. The process is assumed to be 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 sufficiently small $\varepsilon > 0$. In the successive small time intervals $(t, t+\mathrm{d}t)$, each individual either dies 'naturally' according to a simple death process with *probability* $\mu\,\mathrm{d}t$, or else it undergoes the *deterministic* incremental size change

$$m_i(t+\mathrm{d}t) = m_i(t) + f\,(m_i(t))\,\mathrm{d}t + \sum_{j\neq i} h\left(m_i(t), m_j(t); \left\|x_i - x_j\right\|\right)\mathrm{d}t. \tag{1}$$

Here $\left\|x_i - x_j\right\|$ denotes the distance between different individuals (i.e., points) i and $j \neq i$, $f(\cdot)$ denotes the individual growth function in the absence of spatial interaction, and $h(\cdot)$ is an appropriate spatial interaction function taken over all points $j \neq i$. If $m_i(t+\mathrm{d}t) \leqslant 0$ then the individual is deemed to have died 'interactively' and the point i is deleted, as also happens for natural death. Note that although here we only study models with competition (repulsion) between the marks, the same interaction process can also be used to describe pattern formations undergoing attraction.

This combination of a deterministic growth/interaction term, coupled with a simple stochastic immigration–death process, results in a powerful, yet highly flexible and computationally fast, generator of space–time structure, and is especially suitable for analysing dynamic mark–point processes where the underlying point structure is constantly changing (here through natural and interaction-induced deaths, and immigration). As such, this structure is therefore ideal for our current analysis. For less complicated spatial structures, more statistically based procedures are, of course, possible. Marion et al. (2003), for example, describe how importance sampling can be applied to estimate likelihoods for spatio-temporal stochastic models of epidemics in plant populations, where observations comprise the set of diseased individuals at two or more distinct times. They overcome the problem of likelihood computation, imposed by the inherent lack of independence of the status of individuals in the population, by first partitioning the population into a number of sectors and then attempting to take account of this dependence within each sector, whilst neglecting that between sectors. Model (1), of course, involves no such assumption.

In the absence of spatial interaction, a linear (i.e., simple birth–death) growth function $f\,(m_i(t))$ leads to exponentially 'exploding' marks. RS therefore adopt the quadratic (i.e., logistic) growth function

$$f\,(m_i(t)) = \lambda m_i(t)\,(1 - m_i(t)/K)\,, \tag{2}$$

for intrinsic growth rate λ and (non-spatial) carrying capacity K, to ensure that mark size remains bounded. Whilst this simple growth structure suits our current needs admirably, since we shall be concentrating on spatial interaction effects, it could clearly be replaced by any other suitable nonlinear form. For example, taking the pure growth process

$$\mathrm{d}m_i(t)/\mathrm{d}t = f\,(m_i(t)) = \lambda(1 - m_i(t)/K)^2, \tag{3}$$

gives rise to $m_i(t) = K\lambda t/(K+\lambda t)$, and although this generates the same initial linear growth at rate λ, the subsequent asymptotic approach to K no longer involves passing through a point of inflexion. Other simple extensions can produce short-term fluctuations around different carrying capacities K_i, with the process intermittently 'flip-flopping' between them.

Note that including the simple death term μ not only ensures that the number of marks remains bounded above by a Poisson(α/μ) distribution, since we have interactive as well as natural death, but it also secures a proper equilibrium distribution over computationally feasible time periods. For if $\mu = 0$, a mark that becomes 'dominant' early on in the process is likely to dominate its immediate surroundings virtually indefinitely as it will effectively prohibit any neighbouring mark from developing to its full potential. The chance of it being simultaneously surrounded by a sufficiently huge number of immigrants to kill it, within any realistic finite time length, is remote.

There are clearly a large variety of 'obvious' choices that we can make for the spatial interaction function $h(\cdot)$. For example, paralleling the symmetric function (3.1) in RS gives (for $b > 0$) the purely symmetric function (A)

$$h\left(m_i(t), m_j(t); \left\|x_i - x_j\right\|\right) = -bI\left(\left\|x_i - x_j\right\| < r\left(m_i(t) + m_j(t)\right)\right), \tag{4}$$

where the indicator function $I(x) = 1$ if x is true and $I(x) = 0$, otherwise. Here, we assume that the area, in horizontal projection, over which a mark competes for resources can be represented by a disk of radius $rm_i(t)$. As soon as two disks overlap, i.e., $\|x_i - x_j\| < r\left(m_i(t) + m_j(t)\right)$, then competitive interaction takes place with force b. Since this function is symmetric, the larger and smaller of an interactive-pair are affected equally, and although this represents a reasonable approximation in the case of say a planted forest, where we have interaction between trees of similar size, it is clearly inappropriate if $m_i(t)$ and $m_j(t)$ are radically different. For then the smaller mark can exert an appreciable influence on the growth of the larger mark irrespective of how tiny it may be.

The asymmetric function (4.2) in RS, namely (B)

$$h\left(m_i(t), m_j(t); \|x_i - x_j\|\right) = b\left(m_i(t) - m_j(t)\right) I\left(\|x_i - x_j\| < c\right), \tag{5}$$

attempts to circumvent this problem by enabling a point j which lies within a disc of radius c centred on point i to enhance the growth of i if $m_j(t) < m_i(t)$ and impede it otherwise. Though this is again not really appropriate when we consider interaction between trees of radically different size, if we envisage λ as an "average growth rate under spatial interaction" then the concept of a small mark enhancing a larger neighbouring mark does make sense. To avoid this artificial construct of 'enhancement', we could use a one-sided function such as (C)

$$h\left(m_i(t), m_j(t); \|x_i - x_j\|\right) = -b \max\left[0, \left(m_j(t) - m_i(t)\right) I\left(\|x_i - x_j\| < c\right)\right], \tag{6}$$

though this unfortunately reflects the extreme case in which the smaller of the two marks m_i and m_j has no effect whatsoever on the larger. So of these three choices, model A is the most mathematically convenient, since it is directly applicable to equilibrium Gibbs processes, whilst model C is at least partly relevant to real-life scenarios. The asymmetric model B represents a half-way house between these two extremes and so has little to offer in the way of additional features. Moreover, neither B nor C are appropriate when we consider interaction between trees which are of approximately the same size, since the interaction terms would be negligible, and hence unrealistic.

To construct a model that is appropriate irrespective of the relative sizes of two interacting marks, we assume (as for model A) that the area over which a mark competes for resources can be represented by a disk. The amount of competition experienced by a given mark is then assumed to be a function of the extent to which its influence zone overlaps those of neighbouring trees (Staebler, 1951). Different functions of tree size for calculating the influence zone radius, and different procedures for weighting the overlapping areas with the relative sizes of competitors, have led to a proliferation of indices (for further discussion of this issue see Gadow and Gangying, 2001). One of the prototypes, which suits our current requirements, was to allow the interaction function to be proportional to the relative size of the influence zone (Gerrard, 1969). Let $D(x_i, s)$ denote the disk with centre x_i and radius s. Then on placing (D)

$$h\left(m_i(t), m_j(t); \|x_i - x_j\|\right) = -b \text{ area}\left\{D(x_i, rm_i(t)) \cap D\left(x_j, rm_j(t)\right)\right\} / \pi r^2 m_i^2(t), \tag{7}$$

we see that the smaller of two interacting marks will be affected substantially more than the larger, whilst marks of equal size are affected equally. This function clearly has particular application to forest processes, and doubtless to many other real-life growth–interaction scenarios.

Moreover, model D offers two further advantages over B and C. First, in the latter, marks and points affect the interaction function separately, whereas D involves marks and points jointly. Second, B and C (as well as A) are based on 'hard-core' interaction structures, in the sense that the interaction effect rises in a single step as soon as two disks of influence intersect. Though conversion to soft-core processes could easily be made by replacing $I(\cdot)$ by a smoothly declining interaction function. Model D, however, is naturally 'soft-core', and so requires no such modification. Note that although Møller and Waagepetersen (2003) also use an area-interaction construct, they do so in a purely symmetric setting, and so their process provides a closer parallel to model A than D.

3. Comparative estimation strategies

The RS estimation strategy is to simulate the process (1) over the unit time intervals $(n, n+1)$, using the spatial configuration at time n as the 'initial' space condition. This generates sets of parameter estimates at the 'snap-shot' times $n = 2, 3, \ldots$, which can then be combined to form overall estimated values together with associated estimated standard errors. Because this procedure involves separate time steps it is amenable both to maximum likelihood

and pseudo-likelihood estimation. Moreover, use of a semi-parametric interaction function, e.g.,

$$h\left(m_i, m_j; \left\|x_i - x_j\right\|\right) = \theta_v \quad \text{if } \gamma_{v-1}\left(m_i + m_j\right) < \left\|x_i - x_j\right\| \leqslant \gamma_v\left(m_i + m_j\right)$$

for appropriate step-wise parameters $\left(\theta_v, \gamma_v; v = 1, 2, \ldots\right)$, results in increased flexibility. Here, however, we wish not only to estimate simultaneously over all available sampling times, but also to determine exact point estimates for the parameter values. The ensuing increase in complexity in likelihood-based approaches suggests that a switch to least squares estimation would be computationally far more practical, especially since the process (1) for *existing* individuals is purely deterministic; *stochasticity* enters only through new immigrants and natural deaths, both of which can be handled separately.

To justify this stance, we need to show not only that little, if anything, is likely to be lost by moving to a least squares procedure, but also such that an approach readily lends itself to a full space–time analysis based on given sample times. Since randomness occurs solely through the simple immigration–death process, let us first use this to discuss and compare maximum likelihood and least squares approaches before we address full growth–interaction processes.

Case 1: Pure immigration with fixed number of events, N. To set the scene consider the basic immigration process with rate α (i.e., the death rate $\mu = 0$), with arrivals occurring at times $t_1, t_2, \ldots, t_N$. Then the time $s_i = t_i - t_{i-1}$ between the $(i-1)$th and ith events follows the Exponential(α) distribution. So the likelihood

$$L(\alpha) = \alpha \mathrm{e}^{-\alpha t_1} \alpha \mathrm{e}^{-\alpha(t_2 - t_1)} \ldots \alpha \mathrm{e}^{-\alpha(t_N - t_{N-1})} = \alpha^N \mathrm{e}^{-\alpha t_N}, \tag{8}$$

which yields the maximum likelihood estimate (m.l.e.)

$$\hat{\alpha} = N/t_N. \tag{9}$$

To determine the associated variance we use $\ell(\alpha) \equiv \ln[L(\alpha)] = N \ln(\alpha) - \alpha t_N$ to form

$$\ell''(\alpha) = -N/\alpha^2 \simeq -N/\hat{\alpha}^2 = -t_N^2/N, \tag{10}$$

whence

$$\begin{aligned} E[\ell''(\alpha)] \simeq -(1/N) E\left(t_N^2\right) &= -(1/N)\left[Var\left(t_N\right) + E^2\left(t_N\right)\right] \\ &= -(N+1)/\hat{\alpha}^2, \end{aligned} \tag{11}$$

since $t_N = s_1 + \cdots + s_N$ where $E(s_i) = 1/\alpha$ and $Var(s_i) = 1/\alpha^2$. Thus,

$$Var\left(\hat{\alpha}\right) \simeq -1/E\left[\ell''\left(\hat{\alpha}\right)\right] \simeq \hat{\alpha}^2/(N+1) = (N/t_N)^2/(N+1). \tag{12}$$

Note that N and t_N are sufficient for determining both $\hat{\alpha}$ and $Var\left(\hat{\alpha}\right)$, since expressions (9) and (12) depend only on N and t_N. To generate numerical values we simply note that since t_N is the sum of N independent Exp(α) random variables it follows the Gamma(n, α) distribution, which can then be used directly to simulate m empirical values $t_N^{(j)}$ $(j = 1, \ldots, m)$. Taking the immigration rate $\alpha = 1$ and forming the estimates N/t_N, i.e., $\hat{\alpha}_j = N/t_N^{(j)}$ with $N = 100$ and $m = 1000$, yielded the simulated estimates

$$\overline{\alpha} = 1.013 \quad \text{and} \quad Var\left(\overline{\alpha}\right) = 0.0098237. \tag{13}$$

So not only does the mean lie close to the true value one, but also the variance is virtually identical to that predicted by (12), namely $\alpha^2/(N+1) = 0.0099$.

To construct the corresponding least squares estimates (l.s.e.), we first note that under pure immigration the deterministic population size $m(t) = m(0) + \alpha t$. So deterministically, the number of immigrants arriving during the stochastic inter-event $s_i = t_i - t_{i-1}$ is αs_i, since this corresponds to a population increase of one (i.e., $i - 1 \to i$). Thus, we can form the sum of squares

$$S \equiv \sum_{i=1}^{N} (\alpha s_i - 1)^2, \tag{14}$$

and hence the least squares estimate

$$\tilde{\alpha} = \sum_{i=1}^{N} s_i \Big/ \sum_{i=1}^{N} s_i^2 = t_N \Big/ \sum_{i=1}^{N} s_i^2. \tag{15}$$

Thus, unlike the maximum likelihood estimator (9), result (15) involves all the additional immigration times t_i. This exposes a fundamental difference between the two procedures, since under maximum likelihood the time to the Nth immigrant is a sufficient statistic. However, if this is all the information we have, then under least squares we have to use the single deterministic population size $m(t_N) = \alpha t_N = N$ to form

$$\check{\alpha} = N/t_N. \tag{16}$$

As this agrees exactly with the m.l.e. (9), in this sense no information is lost by using the deterministically based least squares estimators (15) and (16).

To determine the associated bias, we first note that t_N is the sum of N i.i.d. Exponential(α) random variables. So on writing $t_N = N/\alpha + \varepsilon$, it follows that $E(\varepsilon) = 0$ and $Var(\varepsilon) = N/\alpha^2$. Thus

$$\begin{aligned} E(\check{\alpha}) &= NE(1/t_N) = \alpha E\left[(1+\alpha\varepsilon/N)^{-1}\right] \\ &= \alpha E\left[1 - (\alpha\varepsilon/N) + (\alpha\varepsilon/N)^2 - \cdots\right] \\ &\simeq \alpha + \left(\alpha^3/N^2\right) Var(\varepsilon) = \alpha(1 + 1/N), \end{aligned} \tag{17}$$

i.e., the estimator exhibits slight upwards bias. Similarly, since $E\left(\check{\alpha}^2\right) = \alpha^2 E\left[(1+\alpha\varepsilon/N)^{-2}\right] \simeq \alpha^2[1+3/N]$, we have

$$Var(\check{\alpha}) = E\left(\check{\alpha}^2\right) - E^2(\check{\alpha}) \simeq \alpha^2/N, \tag{18}$$

which, to $O(1/N)$, agrees with the m.l.e. variance (12).

Case 2: Pure immigration over a fixed time $[0, T]$. Now consider the inverse situation of observing a variable number of events N in a fixed time period $[0, T]$. Then the likelihood function remains virtually unchanged. For as $t_N \leqslant T < t_{N+1}$, with probability Pr [no event in (t_N, T)] $= e^{-\alpha(T-t_N)}$,

$$L(\alpha) = \alpha e^{-\alpha t_1} \alpha e^{-\alpha(t_2-t_1)} \ldots \alpha e^{-\alpha(t_N - t_{N-1})} e^{-\alpha(T-t_N)} = \alpha^N e^{-\alpha T}. \tag{19}$$

Whence comparison with (8) and (9) immediately yields

$$\hat{\alpha} = N/T. \tag{20}$$

To determine the associated variance we note that

$$\begin{aligned} E\left[l''(\alpha)\right] = E\left[-T^2/N\right] &= -T^2 E\left[\frac{1}{(N-\alpha T) + \alpha T}\right] \\ &= -T^2 E\left[\frac{1}{\alpha T\{1 + (N-\alpha T)/\alpha T\}}\right] \\ &\simeq -\left(\frac{T}{\alpha}\right) E\left[1 - \left(\frac{N-\alpha T}{\alpha T}\right) + \left(\frac{N-\alpha T}{\alpha T}\right)^2 - \cdots\right] \\ &\simeq -\left(\frac{T}{\alpha}\right)\left[1 + \frac{1}{\alpha T}\right], \end{aligned} \tag{21}$$

since $N \sim \text{Poisson}(\alpha T)$ we have $E(N) = Var(N) = \alpha T$. Thus,

$$Var\left(\hat{\alpha}\right) \simeq -1/E\left[l''(\alpha)\right] = \alpha^2/(1+\alpha T) \simeq \alpha/T. \tag{22}$$

Since, for reasonably large N, $T \simeq N/\alpha$, this leads to $Var\left(\hat{\alpha}\right) \simeq \alpha/T \simeq \alpha^2/N$, in agreement with (18). So as far as estimation is concerned, it is immaterial whether we make observations over a fixed number of events N or a fixed time T. Moreover, under a deterministic approach neither N nor T varies, so the previous least squares estimates (15) and (16), bias (17) and variance (18) also apply here. That switching from N fixed and T varying, to N varying and T fixed, retains the estimate, is reassuring.

Case 3: Pure immigration with population size recorded at fixed times. In practice, we may only be able to record the total population sizes $n_0, n_1, \ldots, n_s$ at given sampling times $t_0, t_1, \ldots, t_s$. Since the number of immigrants arriving in a time interval of length τ is $\text{Poisson}(\alpha\tau)$, the associated likelihood is given by

$$L(\alpha) = \prod_{r=1}^{s} \frac{(\alpha\,(t_r - t_{r-1}))^{n_r - n_{r-1}}}{(n_r - n_{r-1})!} \mathrm{e}^{-\alpha(t_r - t_{r-1})}. \tag{23}$$

Whence for $t_0 = 0$, the log-likelihood takes the simple form

$$\ell(\alpha) = (n_s - n_0)\ln(\alpha) - \alpha t_s + c, \tag{24}$$

where c is independent of α. As the final population size is sufficient for α, this form is identical to the two previous loglikelihoods developed for Cases 1 and 2, and so

$$\hat{\alpha} = (n_s - n_0)/t_s. \tag{25}$$

To determine the corresponding least squares argument, we parallel (14) by writing

$$S \equiv \sum_{r=1}^{s} \left[\alpha\,(t_r - t_{r-1}) - (n_r - n_{r-1})\right]^2. \tag{26}$$

Whence forming $\mathrm{d}S/\mathrm{d}\alpha = 0$ yields the least squares estimate

$$\tilde{\alpha} = \left[\sum_{r=1}^{s} (t_r - t_{r-1})\,(n_r - n_{r-1})\right] \Bigg/ \sum_{r=1}^{s} (t_r - t_{r-1})^2. \tag{27}$$

So for regularly spaced sample times, say $t_1 = 1, \ldots, t_s = s$, we have

$$\hat{\alpha} = \sum_{r=1}^{s} (n_r - n_{r-1})/s = (n_s - n_0)/s, \tag{28}$$

in exact agreement with the likelihood estimate (25) for $n_0 = 0$ and $t_s = s$. Though for irregularly spaced sample times the least squares and m.l.e.'s will clearly differ.

These simple examples suggest that in population based scenarios, using least squares procedures may be just as informative as using maximum likelihood. Since these two techniques are, respectively, associated with deterministic and stochastic population approaches, and given that in general the former is far more mathematically (and indeed computationally) tractable, choosing least squares over maximum likelihood is likely to result in very little loss of information and a huge gain in algebraic transparency and computational simplicity. It is for this reason that our space–time model depends heavily on deterministic growth and interaction functions. For by restricting random effects to immigration and death, both of which are eminently tractable, we retain both the fundamental space–time characteristics of the underlying process and a stochastic presence. Moreover, it is computationally trivial to compute associated standard errors for the parameter estimates by running a given simulation model many times with different start seeds; bootstrap-type estimation procedures could also be used.

Case 4: Simple immigration–death with known event times. Let us now include the death components in our stochastic space–time simulation algorithm. The resulting (non-spatial) simple immigration–death process is a tour de force in the analysis of stochastic structure, since it is mathematically tractable (e.g., Renshaw, 1991), offers relatively simple

expressions for occupation probabilities and moments, yet easily lends itself to useful extensions. Jakeman et al. (1995) and Gillespie and Renshaw (2005, 2006), for example, use it to draw attention to the unusual properties of an exactly solvable population model which is generic to the recently developed area of quantum optics known as 'non-classical light'. Other theoretical developments include the inclusion of birth (Jakeman et al., 2003); whilst Gibson and Renshaw (1998, 2001) and Renshaw and Gibson (1998) use its underlying structure in the development of practically oriented Markov chain Monte Carlo procedures.

Suppose we record d deaths at times $w_1, \ldots, w_d$ together with n_i immigrants between w_{i-1} and w_i at times $t_{i,1}, \ldots, t_{i,n_i}$. Then to derive the likelihood density that, starting with r individuals at t_0, n immigrants arrive at $t_1, \ldots, t_n$ followed by a death at w, we multiply the probabilities of successive events and inter-event times to obtain

$$L(\alpha, \mu | r, n) = (r+n)\mu\alpha^n \mathrm{e}^{-\alpha(w-t_0)} \exp\{\mu[rt_0 + (t_1 + \cdots + t_n) - (r+n)w]\}. \tag{29}$$

Multiplying terms of this form corresponding to each of the d deaths then yields the full likelihood density, from which it is straightforward to deduce the m.l.e's

$$\begin{aligned} \breve{\alpha} &= \left(\sum_{i=1}^{d} n_i\right) \Big/ w_d, \\ \breve{\mu} &= d \Big/ \left[-\sum_{i=1}^{d} \left(t_{i,1} + \cdots + t_{i,n_i}\right) + \sum_{i=1}^{d} w_i + w_d \sum_{i=1}^{d} (n_i - 1) \right]. \end{aligned} \tag{30}$$

The likelihood is contained within a steeply rising narrow bell which is virtually elliptical in cross-section; this shape provides a useful touchstone to compare the parameter likelihood obtained using the MCMC procedure when observations are incomplete, such as when only death-times are recorded (Renshaw and Gibson, 1998; Gibson and Renshaw, 2001).

Case 5: Simple immigration–death with sampled event times. If we only have information on the population sizes $n_0, n_1, \ldots, n_s$ at times $t_0, t_1, \ldots, t_s$, then the likelihood is far less tractable. For although the probability generating function (p.g.f.) of the population size n_r at time t_r, given size n_{r-1} at the previous sample time t_{r-1}, takes the neat form

$$G(z; t) = \left[1 + (z-1)\mathrm{e}^{-\mu(t_r - t_{r-1})}\right]^{n_{r-1}} \exp\left\{(\alpha/\mu)(z-1)\left(1 - \mathrm{e}^{-\mu(t_r - t_{r-1})}\right)\right\}, \tag{31}$$

being a straight convolution of a pure death process and a Poisson process, the associated probabilities are far less pleasant. Moreover, even the least squares approach does not yield simple closed form expressions for the estimates, since it involves minimising a function which has to take into account not only the n_{r-1} individuals known to be present at time t_{r-1}, but also all the unobserved immigration and death events which occur during the subsequent time interval (t_{r-1}, t_r). The prospect of integrating out a likelihood over all unobserved growth and interaction events is totally impractical, and further justifies the decision to adopt a least squares approach. To effect this, we first assign a 'status' to each individual, depending on whether they die by 'natural causes' or 'spatial interaction', and ignore all 'unobservable events'. Although this causes a downwards bias in both α and μ, which could be corrected by using a simulated search procedure to determine the 'optimal' (α, μ) that gives rise to the l.s.e.'s $(\tilde{\alpha}, \tilde{\mu})$, here our main objective is to concentrate on growth and spatial interaction effects.

4. Least squares estimation based on the simulated process

The general simulation procedure involves using the *deterministic* incremental equation (1) to update each of the n 'alive' individuals at the discrete set of times $\mathrm{d}t, 2\,\mathrm{d}t, \ldots$ (typically $\mathrm{d}t = 0.01$). Non-positive marks are recorded as being 'interactively dead' if $m_i(t) > 0$ and $m_i(t + \mathrm{d}t) \leqslant 0$ (whence $m_i(t + \mathrm{d}t) \to 0$). To identify the status of each individual, we associate an indicator value $I_i = 0$ (natural death), $I_i = 1$ (alive) or $I_i = 2$ (interactive death) for each point $i = 1, \ldots, n$. Whilst on letting the process develop over the unit torus or square, with $\{Z\}$ denoting a sequence of independent $U(0, 1)$ random variables, stochasticity enters in two ways: (i) immigrants enter the system randomly at rate α, are assigned uniformly distributed locations on $(0, 1)^2$, and are given initial mark $\sim U(0, 1)$ and carrying capacity K; (ii) alive

individuals die from 'natural causes' at rate μ. Finally, we test to see whether a new immigrant arrives during $(t, t+\mathrm{d}t)$. Hence on starting with $n=0$ at $t=0$, and taking $\{Z\}$ to be a sequence of i.i.d. $U(0,1)$ pseudo-random variables, we have the

Simulation Algorithm.

cycle over time $t = \mathrm{d}t, 2\,\mathrm{d}t, \ldots$
  cycle over individuals $i = 1, \ldots, n$
    1. test for natural death: if $Z \leqslant \mu\,\mathrm{d}t$ and $I_i = 1$ then set $m_i = 0$ and $I_i = 0$;
    2. determine m_i from (1) over $I_i = 1$;
    3. test for interactive death: if $m_i \leqslant 0$ and $I_i = 1$ then set $m_i = 0$ and $I_i = 2$
  test for new immigrants: if $Z \leqslant \alpha\,\mathrm{d}t$ then add new mark $m_{n+1} = U(0,1)$
  with carrying capacity K and location $x_{n+1} = U(0,1)^2$, and set $I_{n+1} = 1$
  and $n \to n+1$
if t is (say) integer then output $\{(i, x_i(t), m_i(t), I_i) : i = 1, \ldots, n\}$.

This algorithm is easily altered to accommodate any variant of the immigration, growth and interaction regimes.

Suppose we observe the process $\{i, x_i(t), m_i(t) : i = 1, \ldots, N_t\}$ at times $t = 1, 2, \ldots, T$, where N_t denotes the number of observed individuals in the time period $(0, t)$. Then for given λ, K, b and r, and each $t = 1, \ldots, T-1$, we first run the deterministic integral routine (1) over the set $\Omega_t = \{i = 1, \ldots, N_t \text{ for which } I_i = 1\}$ in order to determine the values $\tilde{m}_i(t+1)$ predicted from $m_i(t)$; each $\tilde{m}_i(t+1) < 0$ is replaced by $\tilde{m}_i(t+1) = 0$. We then determine the least squares estimates $\hat{\lambda}$, $\hat{K}$, $\hat{b}$ and $\hat{r}$ by minimising the overall sum of squares

$$S \equiv \sum_{t=1}^{T-1} \sum_{i \in \Omega_t} \left[\tilde{m}_i(t+1) - m_i(t+1)\right]^2. \tag{32}$$

An estimate for α is clearly provided by

$$\hat{\alpha} = N_T / T; \tag{33}$$

though this is biased downwards since individuals that arrive and die within the same unit time interval cannot be observed. Whilst we may estimate μ via the likelihood

$$L(\mu) \equiv \prod_{i=1}^{n_T} \left(\mu \mathrm{e}^{-\mu t_i}\right) \times \prod_{j=1}^{m_T} \left(\mathrm{e}^{-\mu s_j}\right) \tag{34}$$

where $t_1, \ldots, t_{n_T}$ and $s_1, \ldots, s_{m_T}$, respectively, denote the lifetimes of the n_T individuals who have died from natural causes by time T (i.e., $I_i(T) = 0$) and the m_T individuals who are still alive at time T (i.e., $I_j(T) = 1$). For the log-likelihood $l(\mu) = \ln[L(\mu)]$ yields $\mathrm{d}l(\mu)/\mathrm{d}\mu = 0$ at

$$\hat{\mu} = n_T \Bigg/ \left[\sum_{i=1}^{n_T} t_i + \sum_{j=1}^{m_T} s_j\right]. \tag{35}$$

The corresponding interval estimates can be constructed by using these estimated values to simulate M realisations of the process, and then evaluating M sets of new parameter estimates based on the same output times $t = 1, 2, \ldots, T$.

5. Simulated examples for models A and D

The two spatial interaction functions in models A and D, i.e., (4) and (7), essentially differ in two respects. First, whereas A is symmetric, in that a large mark $m_j(t)$ affects a neighbouring small mark $m_i(t)$ to exactly the same extent as $m_i(t)$ affects $m_j(t)$, the non-symmetric function D is far more 'realistic' since the effect of the larger mark on the smaller is much greater than the effect of the smaller on the larger. Second, since the proportion of area component in (6) is at most one (when the smaller disk lies entirely within the larger), under D the interaction between

any two interacting marks $m_i(t)$ and $m_j(t)$ will always be bounded above by that under A. In both cases, however, 'small' marks can exist in harmony with 'large' marks provided the intersection of all the disks does not cover the entire torus. For then small marks can develop in the vacant spaces, and undergo a pure growth process until such time as their zones of influence intersect with others. In addition, if the growth component is strong enough to support one or more interacting neighbours, then a small mark of size m can also exist within say θ covering zones of influence.

For example, on comparing the downwards interaction force A with the logistic growth rate (2), we see that for m to survive we require

$$\mathrm{d}m/\mathrm{d}t = \lambda m(1 - m/K) - b\theta > 0. \tag{36}$$

Hence as $\mathrm{d}m/\mathrm{d}t = 0$ at

$$m = \left(m', m''\right) = (K/2)\left[1 \pm \sqrt{1 - 4b\theta/\lambda K}\right], \tag{37}$$

if $m < m'$ then m decays to zero, if $m' < m < m''$ then m grows to m'', whilst if $m > m''$ then m decreases to m''. This presupposes that $\sqrt{1 - 4b\theta/\lambda K}$ is real; whilst for m to be able to survive within even a single zone of influence we need $b < \lambda K/4\theta$. However, when $4b\theta/\lambda K \simeq 1$, $m' \simeq m'' \simeq K/2$, in which case new small marks will have size $m < m'$ and so will decay to zero. As $m' = 0$ at $b = 0$ and $m' = 1$ at $b \simeq \lambda/\theta$, it follows that for $m \sim U(0, 1)$ and $K \gg 1$, we require $b < \lambda/\theta$ for any of these marks to survive within θ covering disks. So with $\lambda = 1$, taking $b = 0.2$ allows most of the new immigrants to survive in the shadow of a single disk, and some to survive under as many as four. Since the interaction effect D is weaker, under D a new immigrant could survive even more easily than under A.

Although simulations of model A with logistic growth appear to produce an upper and lower canopy structure, as often observed in practice (and previously successfully simulated by Ford and Diggle, 1981), close inspection reveals that all the lower canopy marks are 'new'. This, however, is to be expected, since under logistic growth any $m > m'$ will increase to $m'' = K - m'$, and so all surviving marks will be of roughly the same size. Thus, choice of growth function is clearly crucial to the production of an established 'upper' and a 'lower' canopy structure. A simple way of achieving this is to replace logistic growth by the linear decay function

$$f\left(m_i(t)\right) = \lambda\left(1 - m_i(t)/K\right). \tag{38}$$

For whereas under logistic growth (2) we see that $f\left(m_i(t)\right) \simeq \lambda m_i(t) \simeq 0$ for $m_i(t) \simeq 0$, under the linear model (38) we have $f\left(m_i(t)\right) \simeq \lambda$, and so a very small mark will survive under the shadow of θ covering disks provided $\lambda > b\theta$.

Let us therefore now demonstrate the estimation procedure on two scenarios: (1) disk-interaction A under logistic growth; and, (2) area-interaction D under linear growth. To determine appropriate values for α and μ recall that in the absence of spatial interaction (i.e., $b = 0$) the number of points (N) contained in the unit torus is Poisson(α/μ); so $E(N) = \alpha/\mu = 250$ for $\alpha = 5$ and $\mu = 0.02$. Thus, as some points will suffer death by interaction (i.e., have status 2) when $b > 0$, the number of 'alive' points (status 1) present at any given time is likely to be substantially less than 250. Indeed, as a large mark $m_i(t) \simeq K$ has interaction diameter $2rm_i(t) \simeq 0.2$ for $K = 20$ and $r = 0.005$, regular rectangular packing (for example) allows at most 25 such marks before interaction occurs between them; in practice we observed mark numbers in the range [61, 76]. Hence taking $r = 0.005$ is likely to give rise to substantial spatial interaction, which is why this value was chosen for the simulation study in RS.

5.1. Simulation run 1—model A with logistic growth function

Fig. 1 shows a simulation run at times $t = 5$, 10, 50 and 100 with immigration rate $\alpha = 5$, birth rate $\lambda = 1$, death rate $\mu = 0.02$, carrying capacity $K = 20$, force of interaction $b = 1.5$ and radial factor $r = 0.005$. Since $b > \lambda/\theta$ even for $\theta = 1$, a new immigrant cannot survive unless it lies completely outwith all the existing zones of influence. So new marks can develop only within empty spaces. The bubble plot radii are $rm_i(t)$, so interacting points are those whose disks intersect. We see that even by $t = 5$ there is a mixture of large, and newly arrived, small marks; by $t = 10$ empty space is already being filled; and by $t = 50$ most points are competing with at least one neighbour. Since the expected lifetime of an isolated mark is $1/\mu = 50$, roughly $\frac{2}{3}$ of the points change between $t = 50$ and 100. Close inspection shows the existence of several struggling new immigrants lying within established disks, but as their initial size is $U(0, 1)$, their growth

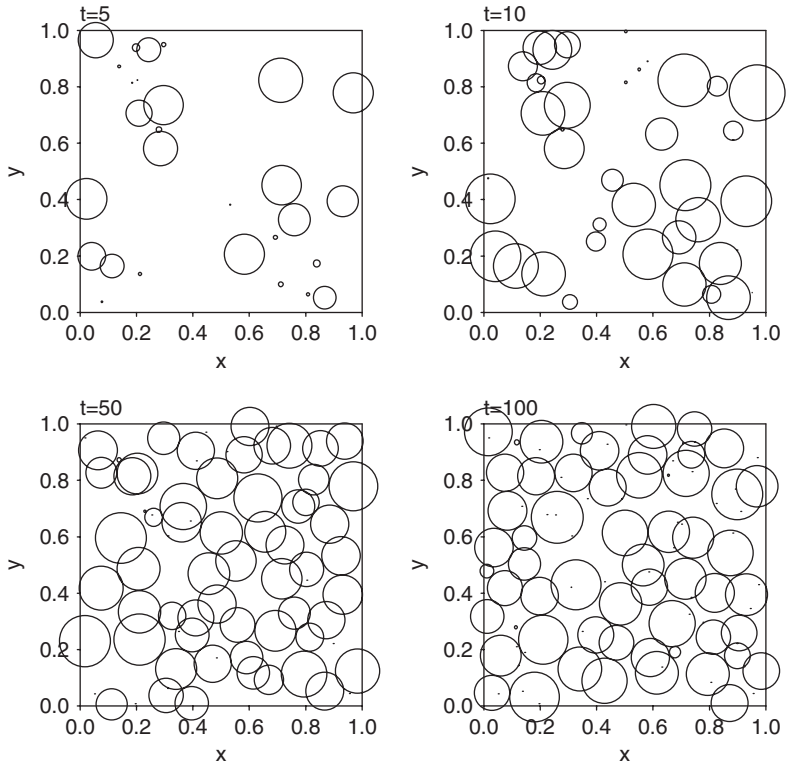

Fig. 1. Simulation of A with logistic growth shown at times $t = 5, 10, 50, 100$ for $\alpha = 5$, $\lambda = 1$, $\mu = 0.02$, $K = 20$, $b = 1.5$, $r = 0.005$ and bubble plot radius $rm_i(t)$.

rate $\lambda m_i(t)(1 - m_i(t)/K) < 0.95$ is dominated by the inhibitory pressure $b\theta \geqslant 1.5$, so they swiftly die. However, a new immigrant arriving in empty space can develop uninhibited until its zone of influence touches neighbouring zones, at which point it can 'mark time' until one of its neighbours dies and releases space for it to grow into. Note that b is large enough to ensure that the disks are unlikely to overlap so severely that they preclude the existence of empty space, even though at $t = 100$ almost all the established marks form a single clique. Moreover, in the limit $\mu = 0$ (i.e., there is no death by natural causes), as $b \to \infty$ zones of influence will touch but not overlap. Thus, there will always be empty space available for immigrants to establish themselves in, whereupon they exert pressure on neighbouring marks, and thereby all marks, causing a general reduction in mark size. So is there a critical value $b = b_0$ which divides the limiting behaviour into two types, namely the number of marks is bounded above for $b > b_0$ but becomes infinite for $b < b_0$? For such phase transition phenomena have already been shown to exist in other area-interaction type processes, an example being the two-type Poisson competition process (Widom–Rowlinson mixture model); see Häggström et al. (1999). This maximum packing (see Stoyan, 2003) phenomena is currently under investigation (Renshaw and Comas, 2006).

Estimating the parameters over the initial highly transient growth period for the 10 time points $t = 1, 2, \ldots, 10$ yielded

$$\hat{\alpha} = 4.30, \quad \hat{\mu} = 0.0407, \quad \hat{\lambda} = 1.001, \quad \hat{K} = 19.74, \quad \hat{r} = 0.00589, \quad \hat{b} = 0.440.$$

These values were obtained by repeatedly taking each parameter in turn, determining the sum of squares over (for example) $(\alpha - 2\delta, \alpha - \delta, \ldots, \alpha + 2\delta)$ for say $\delta = 0.1$, and then replacing α by the new minimising value until convergence was obtained. This operation was then repeated with successively smaller values of δ. Though these estimates for λ and K are accurate, those for the remaining parameters are not. As the time period $0 \leqslant t \leqslant 10$ is relatively short we would expect μ to be poorly estimated, simply because the likely number of deaths is so small; likewise for the

interaction parameter b, as there are few points to interact. Over $t = 10, 20, \ldots, 100$, which, since $1/\mu = 50$, provides a roughly equal split between the transient and equilibrium development, we obtained the estimates

$$\hat{\alpha} = 1.87, \quad \hat{\mu} = 0.013, \quad \hat{\lambda} = 0.834, \quad \hat{K} = 17.35, \quad \hat{r} = 0.0045, \quad \hat{b} = 0.90. \tag{39}$$

These are all biased downwards, most likely because of the large sampling intervals. For many individuals arrive and die unobserved, and although these have a considerable effect on spatial development they cannot be incorporated into the estimation algorithm.

Even though these estimates are biased, the success of this inverse estimation approach is also governed by the relative size of the estimated standard errors (e.s.e.'s). These are easily generated by simulating the process s times and obtaining fresh estimates in each case. For example, 10 simulations over $t = 10, 20, \ldots, 100$ based on $\alpha = 5$, $\mu = 0.02$, $\lambda = 1$, $K = 20$, $b = 1.5$ and $r = 0.005$ yielded

	α	μ	λ	K	r	b
Correct	5.0	0.02	1.0	20	0.005	1.5
Est. mean	1.884	0.0122	0.929	17.78	0.00476	0.955
Bias	−62%	−39%	−7%	−11%	−5%	−36%
e.s.e.	0.029	0.00018	0.079	0.47	0.00020	0.0987
e.s.e./est. mean	1.5%	1.5%	8.5%	2.6%	4.2%	10.2%

So in terms of precision the least squares estimation procedure clearly works well, with the estimates for α and μ being precise (as expected), and those for K and r showing little relative variability. Even the variability in $\hat{\lambda}$ and the interaction parameter $\hat{b}$ is low, bearing in mind the close inter-dependency between growth and interaction. It is the presence of strong downwards bias that causes the most concern. Note that the accurate estimation of the range of interaction r is promising. For not only have previous studies shown this to be difficult, but they often estimate other parameters based on an initial estimate of r.

5.2. Simulation run 2—model A with linear growth function

Given that the hard nature of the influence zones under $b = 1.5$ causes such strong interaction between the estimates, suppose we now let b take the less aggressive value 0.2 and replace logistic growth by the linear form (38). Then the resulting model A′ enables new immigrants to survive even when covered by several disks of influence. The associated pattern structure (Fig. 2) is clearly quite different from that under model A, with the number of marks being considerably greater (205 as opposed to 110), and mark size having a much wider spread but a smaller range; this is especially true for 'old' marks. Note that under A′ there is no longer a clear divide between the upper and lower canopy structure. Though the apparent existence of a lower canopy under A (suggested by bimodality in Fig. 2(c)) is really illusory, since this consists purely of young marks which then either grow into the upper canopy or die (Fig. 2(e)). In contrast, small marks can exist quite easily under A′, without succumbing to competitive pressure. So choosing the 'correct' growth function for any given biological/ecological growth–interaction process is clearly crucial if we are to mimic successfully the space–time construction of real-life processes. Moreover, it is also important to note that if pattern structure is too complex to be described by standard space or frequency domain approaches, then an alternative *test* procedure would be to fit the 'most appropriate' space–time model and see whether it produces significant parameter estimates.

5.3. Simulation run 3—model D with linear growth function

In order to obtain simulations using the area-interaction model D, we first require the area W of two intersecting disks, centred at x and y, with radii u and $v \leqslant u$, with centre-to-centre distance $d = \|x - y\|$:

(a) if the disks do not overlap, i.e., $d \geqslant u + v$, then $W = 0$;
(b) if one disk lies completely inside the other, i.e., $d < u - v$ then $W = \pi v^2$;

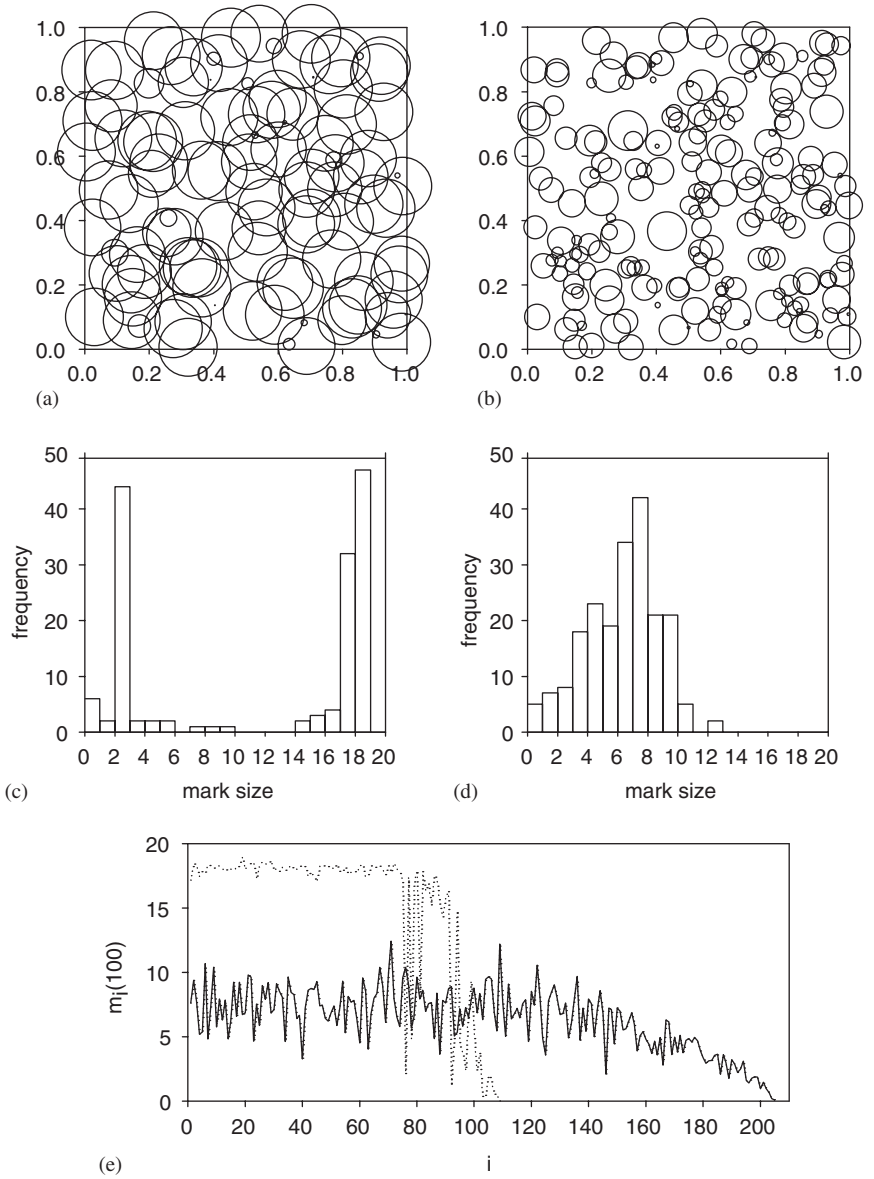

Fig. 2. Simulation of Model A (a) with logistic growth for $b = 1.5$ and (b) with linear growth for $b = 0.2$ at time $t = 100$ for $\alpha = 5$, $\lambda = 1$, $\mu = 0.02$, $K = 20$, $r = 0.005$ and bubble plot radius $rm_i(t)$; respective mark histograms (c) and (d); and, (e) (alive) mark size for logistic (dotted line) and linear (solid line) ordered by age.

(c) if $u - v \leqslant d < u + v$, which covers all the remaining cases (for which there can be 0, 1 or 2 centre points lying within the same circle), then

$$W = u^2 \cos^{-1}\left[\left(d^2 + u^2 - v^2\right) \Big/ 2du\right] + v^2 \cos^{-1}\left[\left(d^2 + v^2 - u^2\right) \Big/ 2dv\right] - (1/2)\sqrt{\left\{2\left(d^2u^2 + d^2v^2 + u^2v^2\right) - \left(d^4 + u^4 + v^4\right)\right\}}. \tag{40}$$

Now there is clearly a profound difference between models A and D. For under A the net growth rate drops immediately by b as soon as a new interacting point j affects i, and simulation runs confirm that a substantial number of points do die because of spatial competition. In contrast, under D the effect of a new interacting point builds up gradually, as the

proportion of overlapping area rises, and this enables interacting points to co-exist much more easily. Indeed, unlike model A 'old' marks sited in the lower canopy regime can remain permanently established, until they eventually die through natural causes. To enable direct comparison with model A, we retain the immigration rate $\alpha = 5$, birth rate $\lambda = 1$, death rate $\mu = 0.02$, force of interaction $b = 1.5$ and radial mark scale factor $r = 0.005$. Using 10 simulations over $t = 10, 20, \ldots, 100$ yielded the estimates

	α	μ	λ	K	r	b
Correct	5.0	0.02	1.0	20	0.005	1.5
Est. mean	2.538	0.0138	0.987	20.57	0.00562	0.637
Bias	−49.2%	−31.0%	−1.3%	2.9%	12.4%	−57.5%
e.s.e.	0.0187	0.000231	0.0282	0.208	0.000121	0.075
e.s.e./est. mean	0.74%	1.68%	2.85%	1.01%	2.15%	11.72%

Again there is strong downwards bias in the estimates for α, μ and b, though r is now over-estimated.

The presence of a strong bias effect is clearly generic to this time-sliced sampling procedure, and one way of correcting for it would be to proceed by first simulating the process over an initial coarse grid of parameter values, determining the corresponding least squares estimates, and then iterating this operation over successively finer grids centred on the minimising estimates until satisfactory convergence is achieved. As a simple illustration, suppose we just vary r or b over a coarse parameter grid $\{r_i, b_i\}$, keeping all other parameters fixed at the above estimated values, and obtain the estimates $\{\rho_i, \beta_i\}$. Write

$$\hat{\rho}_i = er_i + fb_i \quad \text{and} \quad \hat{\beta}_i = gr_i + hb_i,$$

and denote

$$S = \sum_i \left[(\rho_i - \hat{\rho}_i)^2 + \left(\beta_i - \hat{\beta}_i\right)^2 \right].$$

Then forming $\partial S/\partial e$, etc., yields simple equations for e, f, g, h which can then be used to generate initial estimates for r and b. For example, over the 3×3 parameter grid $r_i = (0.004, 0.005, 0.006)$ and $b_i = (1.0, 1.5, 2.0)$, we obtained $e = 0.93$, $f = 0$, $g = 23.99$ and $h = 0.86$, which give rise to the estimates $\hat{r} = 0.00458$ and $\hat{b} = 1.4141$. Even for this minimal case, these values lie much closer to the target values $r = 0.005$ and $b = 1.5$ than the 'point' values 0.0043 and 1.72.

Although this procedure is easily extended to the multivariate case, it is more efficient to replace the grid search by a Monte Carlo search procedure akin to that employed by Renshaw (2000). For $\{Z\}$ a set of i.i.d. uniformly distributed pseudo-random numbers on $(-0.5, 0.5)$:

0. initialise the parameters, e.g., put $r = 0.01$, $\mu = 0.1$, $\alpha = \lambda = b = 1$ and $K = 10$;
1. increment a randomly chosen parameter, say r, to $\tilde{r} = r + \delta_r Z$, for appropriate δ_r;
2. determine S and update (say) r to $\tilde{r}$ if S is reduced;
3. return to 1. until S reaches say 10^{-6}.

If required, the 'windows' $\delta_r, \delta_b, \ldots$ can be gradually reduced to enable a gain in efficiency. However, this approach appears to generate convergence problems which are currently under investigation.

5.4. Simulation run 4—robustness study

These simulation studies presume knowledge of the underlying spatial–temporal generating process. However, not only will the precise mechanism generally be unknown in practice, but it would almost certainly be more complex than our growth–interaction representation. As such, we should not only regard parameter estimates as being more qualitative than quantitative, but we also have to check whether our estimation procedure is robust with respect to an 'incorrect' choice of model structure.

Let us therefore use model A (i.e., logistic growth with hard-core symmetric interaction) to simulate the process, and the quite different model D′ (i.e., linear growth with soft-core asymmetric area-interaction) to estimate the resulting parameter values. Ten separate runs sampled at times $t = 10, 20, \ldots, 100$ yielded

	α	μ	λ	K	r	b
True	5.0	0.02	1.0	20	0.005	1.5
Mean	1.884	0.0125	3.546	18.95	0.00638	0.801
Bias	−62.3%	−37.5%	254%	−5.3%	27.6%	−46.6%
min.	1.79	0.0103	2.02	13.9	0.0014	0.50
max.	2.04	0.0139	5.10	20.7	0.0075	1.39
e.s.e.	0.029	0.00036	0.265	0.60	0.00057	0.094
e.s.e./mean	1.5%	2.9%	7.5%	3.2%	8.9%	11.7%

Comparison with the run 1 estimates shows that using the wrong estimation process does not affect $\hat{\alpha}$ and $\hat{\mu}$ (as expected), but $\hat{\lambda}$ is now severely over-estimated. A likely explanation is that linear growth produces more points than logistic growth, and area-interaction allows small marks to exist more easily than hard-disc interaction. So marks have to grow faster under model D′ (i.e., have a higher λ-value) if they are to produce a spatial structure similar to that grown under model A with all other parameters remaining unchanged. This also explains why the interaction radius r is over-estimated, and hence, since $\hat{r}$ and $\hat{b}$ are negatively correlated, why b is underestimated. The estimate for K shows a slight improvement. So, in qualitative terms, the procedure works well in determining the presence of a growth–interaction process, and in general practical applications looks likely to be robust with respect to an incorrect choice of the presumed underlying model.

Note that we have a classic identification issue here, since even a 'wrong' model may 'expose' interaction between the marks. In practice, however, we will often have prior ecological, physical, etc. information on the process under study, and this should always be used when formulating model choice. For example, in forestry a non-symmetric model will generally be more appropriate than a symmetric one (Gerrard, 1969). So even if model A (say) yields just as good a 'fit' as models D or D′, then if we know that interaction is non-symmetric then model A should be discounted.

6. Analysis of real data based on models A and D′

Having developed analyses based on simulated data sets, we shall now illustrate the application of this space–time modelling approach through the investigation of a Swedish pine forest data set (see Fig. 3). The sample area is located in a developing woodland near Stockholm, and data were recorded on tree location and diameter at breast height (cm) in a circular plot of radius 10 m on three different occasions in 1985, 1990 and 1996. These data contain all trees greater than 1.3 m in height or 10 cm in diameter. No change in soil structure is believed to have occurred during this 11-year period. Visually, there are clearly clumps of interacting trees interspersed with empty space. Note how the diameter distribution changes through time. At the start of the study (1985), there were more tall (taller than 1.3 m) and thin (diameter less than 10 cm) trees than later on, which explains the left-skewness of the diameter distribution. However, these non-observed trees were observed in the later years as their diameters then exceeded 10 cm, which explains both the subsequent right-skewness and the apparent increase in the number of trees.

In 1985 the average age of the trees was 22 years. So as the exact ages of the trees are unknown, we take 1941 as our base point for $t = 0$ and presume that each tree present in 1985 arrived at time $U(0, 44)$. Similarly, trees recorded for the first time in 1990 and 1996 are accorded arrival times of $U(44, 49)$ and $U(49, 55)$, respectively. The arrival time information is needed only to estimate the death rate μ *via* the maximum likelihood estimator (35). The immigration parameter α denotes the arrival rate over the whole 10 m disc, and is estimated by the mean number of arrivals per year between 1985 and 1996. To estimate the remaining parameters, the 1985 tree locations and diameters are taken as the starting configuration, after which trees develop according to our algorithm. We fit model A (symmetric interaction and logistic growth) and model D′ (area-interaction and linear growth) to these data. The resulting least squares

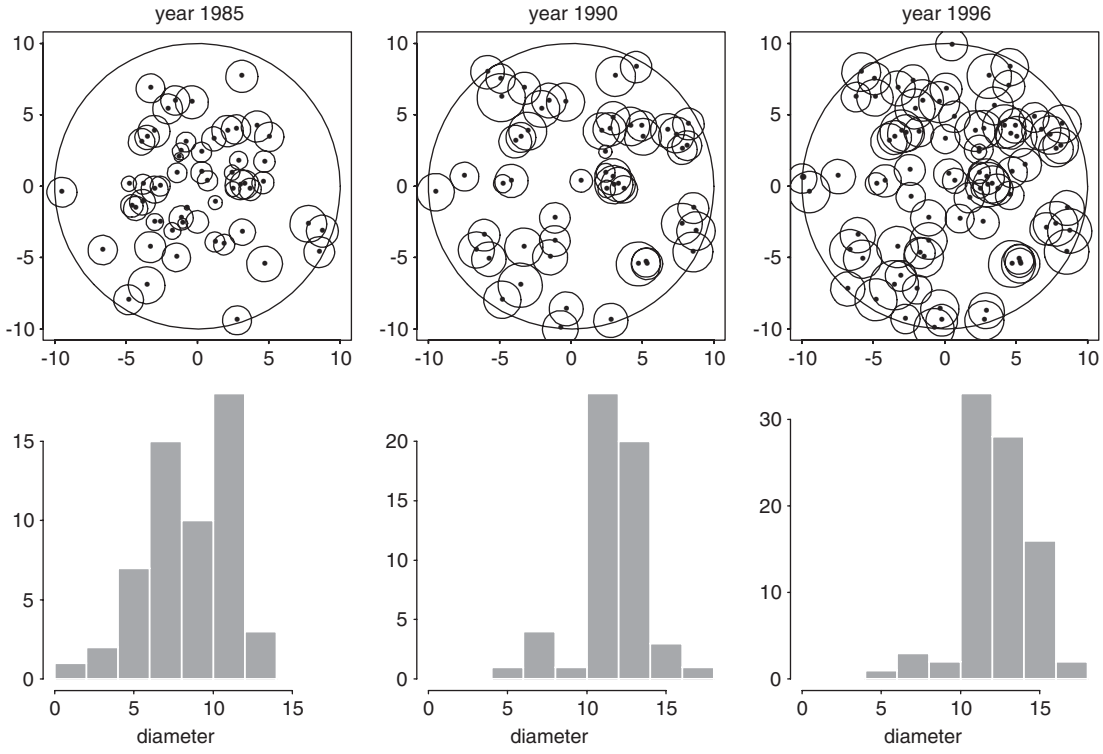

Fig. 3. Swedish space–time data and diameter histograms; circle radius is proportional to tree diameter.

parameter estimates are given by

Model	$\hat{\alpha}$	$\hat{\mu}$	$\hat{\lambda}$	$\hat{K}$	$\hat{r}$	$\hat{b}$
A	5.0	0.0119	3.08	14.2	0.097	1.12
D′	5.0	0.0119	0.478	96.4	0.178	0.044

Apart from $\hat{\alpha}$ and $\hat{\mu}$, which are estimated separately, the remaining parameter estimates clearly exhibit substantial differences reflecting the different model structures. Model D′ indicates a much larger range, and far weaker force, of interaction than model A. This is to be expected, since for a given mark–point pair hard-zone interaction exerts a greater effect than area-interaction.

To see whether the estimated models produce tree patterns of similar structure to the data, we used the estimated parameters to simulate patterns corresponding both to model A (Fig. 4) and model D′ (Fig. 5). Since not all small trees (i.e., diameter < 10 cm) are recorded—only those with height $\geqslant 1.3$ m—the mark histograms are truncated at 10 cm radius. Model D′ outperforms model A in two respects. First, it generates a similar number of marks to the data at the third time point; Fig. 4 contains too many. Second, the mark histograms for the data and D′ are both skewed right; A is skewed left. This observation suggests that the linear growth function is more appropriate than the logistic one (see Fig. 2). The fact that D′ produces more higher-diameter trees than the data can be overcome by reducing $\hat{K} = 96.4$ to a more biologically meaningful value. The number of (alive) trees with diameter $\geqslant 10$cm are: 97, 96 and 103 (model A); 63, 66 and 71 (model D′); and, 22, 49 and 81 (data) in 1985, 1990 and 1996, respectively. Residuals for the difference between the observed and expected (i.e., based on the model) diameters were also computed. Model D′ yields a narrower range for the residuals than model A.

Though our aim here is to *illustrate* that these models are at least qualitatively useful in practice, particularly in forestry, substantial future studies could be undertaken to explore further whether variations on them can produce processes which mimic better the (unknown) real-life generator for the data, and also whether more complete data

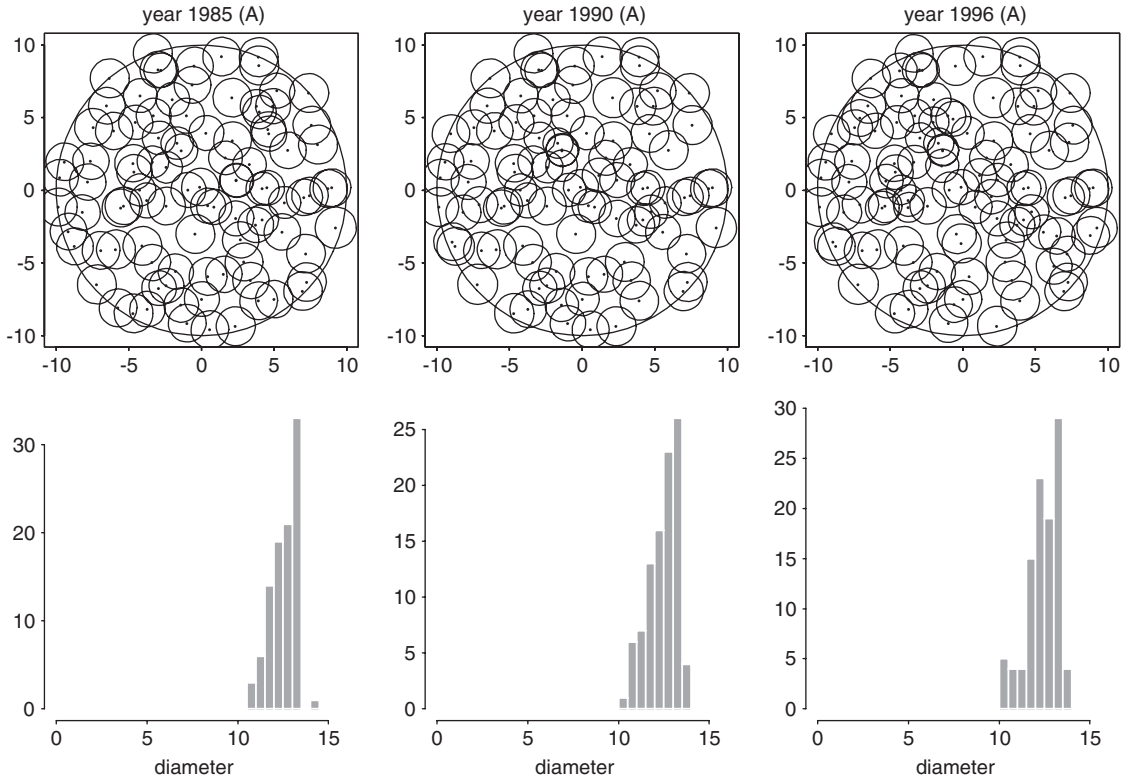

Fig. 4. Simulated patterns and mark histograms for the Swedish space–time data based on logistic growth and symmetric interaction (model A); circle radius is proportional to tree diameter.

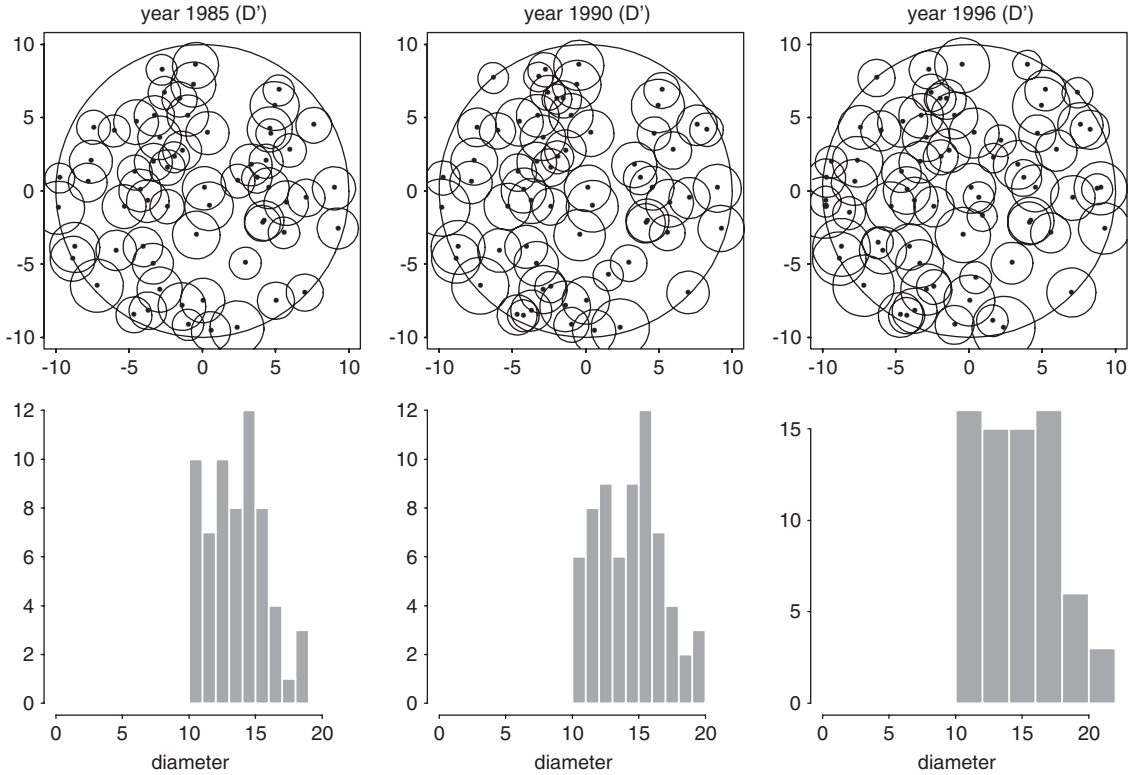

Fig. 5. As Fig. 4 but based on linear growth and asymmetric area-interaction (model D′).

collection would be of considerable help. For example, to establish the process we have used arbitrary arrival times uniformly distributed between 1941 and 1985; using a different start date might reduce the discrepancy between the data and model D′ for the number of alive trees in the first measured time period. Furthermore, edge effects could be taken into account by simulating over a larger region, whilst replication would enable the construction of standard errors for the estimates.

7. Summary and conclusions

In summary, whereas RS generate patterns from growth–interaction models and use them to estimate parameters based on a single pattern relating to a specific fixed time, in this current paper we utilise the full time-development of the process in order to construct parameter estimates. Since general likelihood methods are no longer as computationally feasible, analysis has proceeded through a least squares approach. Justification of this methodological switch is partially provided through the analysis of simple examples. The prospect of integrating out a likelihood over all unobserved growth and interaction events is mathematically impractical, and underpins our decision to adopt a least squares procedure. Investigating the use of MCMC to compute m.l.e.'s would be a subject of future research. Moreover, whilst our current approach provides good qualitative estimates, we have demonstrated the existence of bias problems which need to be solved in further studies. Bias also features in general applications since any choice of model is virtually bound to differ from the unknown real process generating mechanism; though extending the robustness study of RS would help to quantify the order of the induced error. Wherever possible, the growth–interaction processes should always be chosen so that they represent biologically reasonable simplifications of reality.

We state early on that since there are a large variety of possible growth and interaction functions to choose from, we have had to make a pragmatic choice of four simple processes. These neatly illustrate the points featured in our paper, and because A (symmetric hard-core interaction combined with logistic growth), and D′ (non-symmetric area-interaction and linear growth) exhibit the greatest 'behavioural distance' between these four models, we have concentrated on them. Furthermore, the main application we have in mind is to model the growth and spatial development of trees, and models A and D′ apply directly to this forestry context: specifically, model A applies to even-aged forests, where trees are roughly of the same size, and model D′ for naturally regenerating forests. Moreover, both the growth and interaction functions could easily be replaced by other functions, and one such choice is currently the subject of further investigation. Yet further work would involve the development of associated inference procedures: use of straight Monte Carlo techniques would be an obvious first approach.

Finally, in order to assess whether our growth–interaction process may be usefully applied to space–time forestry data, we fitted models A and D′ to a Swedish pine data set. Our preliminary analysis indicates that although model D′ appears to be more appropriate than model A for these data, a simple residual analysis suggests that further improvement to the model structure should be made if the technique is to be used in real-life forest scenarios. It would therefore be useful and timely to combine the expertise of mathematicians and foresters in order to generate more realistic growth and interaction functions that would enable direct application to the development of improved planting, thinning and extraction strategies (Comas, 2005).

Acknowledgements

We wish to thank Jun Yu of the Swedish University of Agricultural Sciences for providing the pine data. Aila Särkkä has been supported by the Swedish Research Council, the Swedish Foundation for Strategic Research, and the Wilhelm and Martina Lundgrens Foundation.

References

Bennett, R.J., 1999. Spatial Time Series. Pion, London.

Besag, J., York, J.C., Mollié, A., 1991. Bayesian image restoration, with two applications in spatial statistics (with discussion). An. Inst. Statist. Math. 43, 1–59.

Botkin, D.B., Janak, J.F., Wallis, J.R., 1972. Some ecological consequences of a computer model of forest growth. J. Ecol. 60, 849–873.

Brix, A., Diggle, P.J., 2001. Spatio-temporal prediction for log-Gaussian Cox processes. J. Roy. Statist. Soc. B 63, 823–841.

Brix, A., Møller, J., 2001. Space–time multi-type log Gaussian Cox processes with a view to modelling weed data. Scand. J. Statist. 28, 471–488.

Brown, P.E., Roberts, G.O., Kåresen, K.F., Tonellato, S., 2000. Blur-generated non-separable space–time models. J. Roy. Statist. Soc. B 62, 847–860.

Brown, P.E., Diggle, P.J., Lord, M.E., Young, P.C., 2001. Space–time calibration of radar-rainfall data. Appl. Statist. 50, 221–242.

Comas, C., 2005. Modelling forest dynamics through the development of spatial and temporal marked point processes. Ph.D. Thesis, University of Strathclyde.

Diggle, P.J., Chetwynd, A.G., Häggkvist, R., Morris, E.S., 1995. Second-order analysis of space–time clustering. Statist. Methods Med. Res. 4, 124–136.

Fenton, S.E., Clough, H.E., Diggle, P.J., French, N.P., 2004. An investigation of the effect of under-reporting of disease cases on the space–time *K*-function. In: Baddeley, A., Gregori, P., Mateu, J., Stoica, R., Stoyan, D. (Eds.). Proceedings of the Conference on Spatial Point Process Modelling and its Applications: Collecció treballs d'informática i tecnologia Núm. 20.

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

von Gadow, K., Gangying, K., 2001. Modelling Forest Development. Forestry Sciences. Kluwer, New York.

Gavrikov, V., Stoyan, D., 1995. The use of marked point processes in ecological and environmental forest studies. Environ. Ecol. Statist. 50, 331–344.

Gerrard, D.J., 1969. Competition quotient—a new measure of competition affecting individual forest trees. Michigan State University Agricultural Experimental Station Research Bulletin 20.

Gibson, G.J., Renshaw, E., 1998. Estimating parameters in stochastic compartmental models using Markov chain methods. IMA J. Math. Appl. Med. Biol. 15, 19–40.

Gibson, G.J., Renshaw, E., 2001. Inference for immigration–death processes with single and paired immigrants. Inverse Problems 17, 455–466.

Gillespie, C., Renshaw, E., 2005. The evolution of a batch-immigration death process subject to counts. Proc. Roy. Soc. London A 461, 1563–1581.

Gillespie, C., Renshaw, E., 2006. The evolution of a single–paired immigration death process. J. Appl. Probab., submitted for publication.

Goodall, C., Mardia, K.V., 1994. Challenges in multi-variate spatio-temporal modelling. Proceedings of the XVIIth International Biometric Conference, Hamilton, Canada.

Gössl, C., Auer, D.P., Fahrmeir, L., 2001. Bayesian spatiotemporal inference in functional magnetic resonance imaging. Biometrics 57, 554–562.

Guttorp, P., Meiring, W., Sampson, P.D., 1994. A space–time analysis of ground-level ozone data. Environmetrics 5, 241–254.

Häggström, O., van Lieshout, M.N.M., Møller, J., 1999. Characterization results and Markov chain Monte Carlo algorithms including exact simulation for some spatial point processes. Bernoulli 5, 641–658.

Jakeman, E., Phayre, S., Renshaw, E., 1995. The evolution and measurement of a population of pairs. J. Appl. Probab. 32, 1048–1062.

Jakeman, E., Hopcraft, K.I., Matthews, J.O., 2003. Distinguishing population processes by external monitoring. Proc. Roy. Soc. London A 459, 623–639.

Jones, R.H., Zhang, Y., 1997. Models for continuous stationary space–time processes. In: Gregoire, T.G., Brillinger, D.R., Diggle, P.J., Russek-Cohen, E., Warren, W.G., Wolfinger, R.D. (Eds.), Modelling Longitudinal and Spatially Correlated Data. Springer, New York, pp. 289–298.

Knorr-Held, L., Richardson, S., 2003. A hierarchical model for space–time surveillance data on meningococcal disease incidence. Appl. Statist. 52, 169–183.

Knox, E.G., 1964. Detection of space–time interactions. Appl. Statist. 13, 25–29.

Kyriakidis, P.C., Journel, A.G., 1999. Geostatistical space–time models: a review. Math. Geol. 31, 651–684.

Lawson, A.B., 2001. Statistical Methods in Spatial Epidemiology. Wiley, New York.

MacNab, Y.C., Dean, C.B., 2001. Autoregressive spatial smoothing and temporal spline smoothing for mapping rates. Biometrics 57, 949–956.

Mardia, K.V., Goodall, C.R., Redfern, E.J., Alonso, F.J., 1998. The Kriged Kalman filter (with discussion). Test 7, 217–285.

Marion, G., Gibson, G., Renshaw, E., 2003. Estimating likelihoods for spatio-temporal models using importance sampling. Statist. Comput. 13, 111–119.

Meiring, W., Guttorp, P., Sampson, P.D., 1998. Space–time estimation of grid-cell hourly ozone levels for assessment of a deterministic model. Environ. Ecol. Statist. 5, 197–222.

Mickelson Jr., J.G., Civco, D.L., Silander Jr., J.A., 1998. Delineating forest canopy species in the northeastern United States using multi-temporal TM imagery. Photo. Eng. Remote Sensing 64, 891–904.

Møller, J., Waagepetersen, R.P., 2003. Statistical Inference and Simulation for Spatial Point Processes. Chapman & Hall, New York.

Nyström, K., 2001. Growth models for young stands—development and evaluation of growth models for commercial forests in Sweden. Ph.D. Thesis, Silvestria 180, SLU.

Pacala, S.W., Canham, C.D., Silander, J.A., 1993. Forest models defined by field measurements: I. The design of a northeast simulator. Canad. J. Forest Res. 23, 1980–1988.

Penttinen, A., Stoyan, D., Henttonen, H.M., 1992. Marked point processes in forest statistics. Forest Sci. 38, 806–824.

Pukkala, T., 1989a. Prediction of tree diameter and height in a Scots pine stand as a function of the spatial pattern of trees. Silva Fennica 23, 83–99.

Pukkala, T., 1989b. Predicting diameter growth in even-aged Scots pine stands with a spatial and non-spatial model. Silva Fennica 23, 101–116.

Pukkala, T., Kolström, T., 1991. Effect of spatial pattern of trees on the growth of Norway spruce stand. A simulation model. Silva Fennica 25, 117–131.

Renshaw, E., 1991. Modelling Biological Populations in Space and Time. Cambridge University Press, Cambridge.

Renshaw, E., 2000. Applying the saddlepoint approximation to bivariate stochastic processes. Math. Biosci. 168, 57–75.

Renshaw, E., Comas, C., 2006. Space–time generation of high intensity patterns using growth–interaction processes, in preparation.

Renshaw, E., Gibson, G.J., 1998. Can Markov Chain Monte Carlo be usefully applied to stochastic processes with hidden birth times? Inverse Problems 14, 1581–1606.

Renshaw, E., Särkkä, A., 2001. Gibbs point processes for studying the development of spatial–temporal stochastic processes. Comput. Statist. Data Anal. 36, 85–105.

Schlather, M., Ribeiro, P., Diggle, P.J., 2004. Detecting dependence between marks and locations of marked point processes. J. Roy. Statist. Soc. B 66, 79–93.

Shugart, H.H., West, D.C., 1977. Development of an Appalachian deciduous forest succession models and its application to assessment of the impact of chestnut blight. J. Env. Man. 5, 161–179.

Staebler, G.R., 1951. Growth and spacing in an even-aged stand of Douglas fir. Master's Thesis, University of Michigan.

Stoyan, D., 2003. Hard problems with random systems of hard particles. Proceedings of the 54th Session of the International Statistical Institute, Berlin.

Stoyan, D., Penttinen, A., 2000. Recent applications of point process methods in forestry statistics. Statist. Sci. 15, 61–78.

Stoyan, D., Stoyan, H., 1994. Fractals, Random Shapes and Point Fields. Wiley, Chichester.

Stoyan, D., Kendall, W.S., Mecke, J., 1995. Stochastic Geometry and its Applications. second ed. Wiley, Chichester.

Stroud, J.R., Müller, P., Sansó, B., 2001. Dynamic models for spatiotemporal data. J. Roy. Statist. Soc. B 63, 673–689.

Veccia, A.V., 1988. Estimation and model identification for continuous spatial processes. J. Roy. Statist. Soc. B 50, 297–312.

Wälder, O., Stoyan, D., 1996. On variograms in point process statistics. Biometrical J. 38, 895–905.

Waller, L.A., Carlin, B.P., Xia, H., Gelfand, A.E., 1997. Hierarchical spatio-temporal mapping of disease rates. J. Amer. Statist. Assoc. 92, 607–617.

Wikle, C.K., 2001. A kernel-based spectral approach for spatio-temporal dynamic models. Proceedings of the First Spanish Workshop on Spatio–Temporal Modelling of Environmental Processes (METMA), Benicassim, Spain, pp. 28–31.

Wikle, C.K., 2002. Spatio-temporal Methods in Climatology. In: El-Shaarawi, A.H., Jureckova, J. (Eds.), Encyclopedia of Life Support Systems, to appear.

Wikle, C.K., 2003. Hierarchical models in environmental science. Int. Statist. Rev. 71, 181–199.

Wikle, C.K., Cressie, N., 1999. A dimension-reduced approach to space–time Kalman filtering. Biometrika 86, 815–829.

Williams, G.W., 1984. Time–space clustering of disease. In: Cornell, R.G. (Ed.), Statistical Methods for Cancer Studies. Marcel Dekker, New York, pp. 167–227.

Xu, B., Wikle, C.K., Fox, N.I., 2005. A kernel-based spatio-temporal dynamical model for nowcasting radar precipitations. J. Amer. Statist. Assoc. 100, 1133–1144.