



Modelling Spatial Patterns

B. D. Ripley

Journal of the Royal Statistical Society. Series B (Methodological), Volume 39, Issue 2 (1977), 172-212.

Your use of the JSTOR database indicates your acceptance of JSTOR's Terms and Conditions of Use. A copy of JSTOR's Terms and Conditions of Use is available at <http://www.jstor.org/about/terms.html>, by contacting JSTOR at jstor-info@umich.edu, or by calling JSTOR at (888)388-3574, (734)998-9101 or (FAX) (734)998-9113. No part of a JSTOR transmission may be copied, downloaded, stored, further transmitted, transferred, distributed, altered, or otherwise used, in any form or by any means, except: (1) one stored electronic and one paper copy of any article solely for your personal, non-commercial use, or (2) with prior written permission of JSTOR and the publisher of the article or other text.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Journal of the Royal Statistical Society. Series B (Methodological) is published by Royal Statistical Society. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/rss.html>.

Journal of the Royal Statistical Society. Series B (Methodological)
©1977 Royal Statistical Society

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2001 JSTOR

Modelling Spatial Patterns

By B. D. RIPLEY

Imperial College, London

[Read before the ROYAL STATISTICAL SOCIETY at a meeting organized by the RESEARCH SECTION on Wednesday, March 16th, 1977, Professor S. D. SILVEY in the Chair]

SUMMARY

Spatial point processes may be analysed at two levels. Quadrat and distance methods were designed for the sampling of a population in the field. In this paper we consider those situations in which a map of a spatial pattern has been produced at some cost and we wish to extract the maximum possible information. We review the stochastic models which have been proposed for spatial point patterns and discuss methods by which the fit of such a model can be tested. Certain models are shown to be the equilibrium distributions of spatial-temporal stochastic processes. The theory is illustrated by several case studies.

Keywords: SPATIAL PATTERN; POINT PROCESS; HARD-CORE MODEL; MARKOV POINT PROCESS; MATHEMATICAL MORPHOLOGY; CELL SHAPES

1. INTRODUCTION

PATTERNS of "small" objects in two or three dimensions or on the surface of the terrestrial or celestial spheres are commonplace; some examples are towns in a region, trees in a forest and galaxies in space. We will see that other spatial patterns such as a sheet of biological cells can be reduced to a pattern of points.

Such spatial patterns are often tested for "randomness", i.e. the null hypothesis of an underlying Poisson process is tested. We can divide the methods used into four classes:

- (i) those based on quadrat counts;
- (ii) "distance" or nearest-neighbour methods;
- (iii) second-order methods (including spectra);
- (iv) the "test-set" approach.

Methods of type (i) and (ii) seem to be the most appropriate in preliminary fieldwork. Here we suppose that a complete map of the spatial pattern is available. Our aim will be to find useful non-Poisson models and to establish methods of types (iii) and (iv) to test the fit of any proposed model. The main technique will be the estimation of the first two moment measures of the supposed underlying point process.

One of the drawbacks of existing methods is their inability to test for interactions at different scales *simultaneously* in formal significance tests (*cf.* Mead, 1974). Even in the analysis of contiguous quadrat counts the choice of simultaneous scales is restricted (Hill, 1973). We attempt to overcome this problem by validating tests with artificial data.

Usually spatial patterns will have evolved over a period of time; we would like our models to reflect this temporal aspect. In Section 4 we show that some of the models reviewed in Section 3 can arise as the equilibrium distributions of spatial-temporal stochastic processes. We use this fact in Section 5 to design procedures for the Monte Carlo simulation of some of our models.

It is impossible to write down the complete likelihood function for most of our models, so the usual methods for the estimation of unknown parameters cannot be used. This and other difficulties are discussed in several case studies.

2. SPATIAL POINT PROCESSES

Our models will represent the “small” objects by points or by spheres of fixed radius. (The latter may be specified by their centres; this is a special case of the basic transformation of stochastic geometry (Kendall, 1974a).) We assume that for a realization of the spatial pattern we are able to observe all the objects within a known set E which we will call the *sampling window*. (Alternatively E might be natural and unknown; it is then a separate interesting pattern recognition problem to estimate E from the data (*cf.* Grenander, 1973, 1977).) In practice it may be difficult or impossible to collect the data in this form; for instance, in a mature forest it may be much easier to measure nearest-neighbour distances.

Occasionally E may be a natural region but usually the pattern extends beyond E to an unknown extent. We assume E to be a homogeneous region (as far as causal factors are concerned) and that there is no preferred origin or direction. This assumption forces us to postulate that the pattern extends throughout the space (two- or three-dimensional space or a sphere, denoted by \mathcal{X}).

Formally a *model* is the distribution of a simple second-order point process strictly stationary under rigid motions (in the terminology of Ripley, 1976a). An informal description follows. Each realization is a countable set of points from \mathcal{X} (a finite set if \mathcal{X} is sphere). The assumption of simplicity excludes the possibility that two points might coincide in space. The important random variables are the $Z(A)$ counting the number of points in each Borel subset A of \mathcal{X} . If A is bounded then $Z(A)$ must be finite; we are assuming that $E(Z(A)^2)$ is also finite. The statistical properties are invariant under both translations and rotations.

The simplest parameter of a model is the *intensity* λ , the expected number of points per unit area; this summarizes the first moments of the family $\{Z(A)\}$. The second moment structure may be reduced to a function K defined on $(0, \infty)$ (or $(0, \pi]$ for a sphere). The following intuitive definitions of this function are explained in the Appendix.

(a) $\lambda^2 K(t)$ is the expected number of ordered pairs of distinct points less than distance t apart with the first point in a given set of unit area.

(b) $\lambda K(t)$ is the expected number of further points within t of an arbitrary point of the process.

(c) Under additional assumptions $g(t) = (\lambda^2 dK/dt)/c(t)$ is a joint density for the occurrence of two points distance t apart (where $c(t) = 2\pi t$, $4\pi t^2$ or $2\pi \sin t$ for R^2 , R^3 or the unit sphere respectively). Thus g is a *product density*. We can express the second moments of $\{Z(A)\}$ by

$$E\{Z(A)Z(B)\} = \lambda\nu(A \cap B) + \lambda^2 \int_{(0,\infty)} \nu_t(A \times B) dK(t)$$

where ν is Lebesgue measure on \mathcal{X} and

$$\nu_t(A \times B) = \int \sigma_t[\{y-x: x \in A, y \in B, d(x, y) = t\}] d\nu(x).$$

Here $d(x, y)$ is the distance between x and y (the great-circle distance for a sphere) and σ_t is the uniform probability on the surface of the sphere of radius t centred at the origin. (The existence of this decomposition and the uniqueness of K are established in Ripley, 1976a.)

We avoid technical difficulties by considering the distribution function K of a measure rather than its density; we shall see later that it is easier to use the cumulative function in statistical analyses (as we might have expected from real random variables). The choice of the power of λ in the definition of K is explained below; notice that K has the dimensions of an area or volume.

We may also reduce the higher moments of a model; we shall discuss third moments in an example.

We may specify a model by giving $P\{Z(A_1) = 0, \dots, Z(A_r) = 0\}$ for all finite collections $\{A_1, \dots, A_r\}$ of spheres of the form $b(x, t) = \{y: d(x, y) < t\}$ for any $x \in \mathcal{X}$ and $t > 0$ (Ripley,

1976b). The “first-order” part of this description would be to consider $p(t) = P\{Z(b(x, t)) > 0\}$ which is independent of the choice of x by stationarity. This provides another function on $(0, \infty)$ summarizing part of the model.

In Section 6 we will construct estimators of both K and p . In the next three sections we study classes of models.

3. MODELS

The basic model is the Poisson process defined (within our framework) by either or both of the following properties: $Z(A)$ has a Poisson distribution for all bounded Borel A and $\{Z(A_i)\}$ are independent whenever $\{A_i\}$ is a disjoint class of such sets. This model is specified by its intensity λ ; $K(t) = \pi t^2$ irrespective of λ (hence the scaling of K). Models which give “clustered” patterns will have K greater than this function whereas “self-inhibiting” models will have a smaller K .

We will describe our models on R^2 ; they may be defined on R^3 or a sphere with trivial modifications.

3.1. Cluster Processes

One of the earliest and most intensively studied classes of models is the class of cluster processes (see Neyman and Scott, 1972, for a historical survey.) Recently cluster processes have played an important role in the theory of infinitely divisible point processes (cf. Matthes, 1972). We have a parent process (usually Poisson) on a group G , either R^2 or the group of rigid motions, and a daughter process ϕ . We sample the parent process; with each point we associate an independent copy of ϕ translated by the chosen transformation. The cluster process is the superposition of the daughter processes. For a Poisson parent process the use of the whole group of rigid motions amounts to giving the daughter process an independent uniformly distributed rotation. For conditions for such a model to be well defined see Goldman (1967); Fisher (1972) gives an elementary survey.

The most important subclass consists of Neyman–Scott processes. Here the daughter process ϕ is a random number N of independent identically distributed points and the parent process is Poisson of intensity α .

If the density of the distribution of the distance between two arbitrary points of the daughter process exists and is denoted by f , we have:

$$\lambda = \alpha E(N),$$

$$g(t) = \lambda^2 + \alpha E\{N(N-1)\}f(t)/c(t).$$

The first term in the last expression comes from the consideration of two points from different clusters, the second from two points within a cluster.

3.2. Hard-core Models

Various models have been proposed for processes of the centres of non-overlapping circles or spheres of radius R . Matérn (1960) defined two such models. The first is obtained by sampling a Poisson process of intensity α and deleting any point which is within $2R$ of any other whether or not this has already been deleted. For this process

$$\lambda = \alpha \exp(-4\pi\alpha R^2)$$

and $g(t)$ exists and is given by

$$g(t) = \begin{cases} 0, & t \leq 2R, \\ \alpha^2 \exp\{-\alpha U(t)\}, & t > 2R, \end{cases}$$

where

$$U(t) = \text{meas} \{b(0, 2R) \cup b((t, 0), 2R)\},$$

$$b(x, r) = \{y: d(x, y) < r\}.$$

This model has a maximum intensity (as α is varied) of $(4\pi e R^2)^{-1}$ and so cannot model densely packed cores. This is 10 per cent of the theoretical bound $(\sqrt{12}R^2)^{-1}$ which is attained by a triangular lattice packing (see Rogers, 1964).

Matérn's second process was one of the first marked point processes; the points of a Poisson process of intensity α are independently marked with a uniformly distributed birth time on $(0, 1)$, so that we have a Poisson process on $R^2 \times (0, 1)$. A point is retained if no point within $2R$ has an earlier birth time. We can define two variants of this process by considering points which have already been deleted or not. Matérn took the former view; although this is less natural it does enable us to find

$$\lambda = (1 - e^{-\alpha c})/c, \quad c = U(0) = 4\pi R^2,$$

$$g(t) = \begin{cases} 0, & t \leq 2R, \\ \frac{2U(t)(1 - e^{-\alpha c}) - 2c(1 - e^{-\alpha(t)})}{cU(t)(U(t) - c)}, & t > 2R. \end{cases}$$

These formulae are summarized in Paloheimo (1971). This second model may be appropriate in botanical problems but not in competitive zoological situations such as nesting colonies of seabirds. We shall meet this dichotomy again.

These models may also be defined on a bounded set. If we wish to simulate their restrictions to a bounded set A we may do so by simulating the Poisson process on $\{y: \exists x \in A, d(x, y) \leq 2R\}$; notice that this is impossible for the second variant of the second model.

If we do consider a point process on a bounded subset of the plane we will have to impose boundary conditions. The natural choice is to assume that there are no points outside this set but this will be appropriate only if we know the true extent of the process to be modelled. In general, a more satisfactory solution is to take a sufficiently large square B and impose periodic boundary conditions. Formally we regard B as a torus. If we envisage the plane packed by a lattice of translations of B we will obtain a process which is strictly stationary under the group of translations parallel to the sides of B (if the process on the torus is strictly stationary, as all our examples are); however, this process is never stationary under rotations. We will only consider this process restricted to a subset of B ; it is then plausible that this directionality (which is the price we have to pay on a bounded set) will be negligible. It may be possible to obtain a limiting distribution for the restriction of the process to a bounded set A as the square B expands (cf. Ruelle, 1970). We ignore this directionality even though it strictly invalidates our reduction of the second moments.

Our third hard-core model may only be defined on a bounded set A . We sample a Poisson process on A and retain only those *samples* which contain no pair of points less than $2R$ apart.

Our three hard-core models differ even on spaces such as a torus or sphere where edge-effects are absent. This can be grasped most easily by considering their analogues on the discrete space of twelve sites arranged on a clock-face. The equivalent of hard-cores is to forbid the occupation of adjacent sites. It is possible to calculate the probabilities of various triples under all these processes (by enumerating all outcomes). Rejecting those samples which contain forbidden combinations generates all triples equiprobably. The analogues of both Matérn's models give different probabilities to $\{2, 4, 6\}$ and $\{2, 6, 10\}$. Of course Matérn's two models differ; their second moment properties expressed by g are different.

Bartlett (1974) considered an extension of Matérn's second model with variable core size. Suppose each point is marked independently with a birth time and a core size R chosen from

a prescribed distribution. Various definitions are possible; perhaps the most plausible is that a point is deleted if there is another with an earlier birth time (deleted or not) such that the discs of marked size centred on the points overlap. Bartlett ignored deleted points, deleting a point if an earlier point fell within the disc of marked size centred on the point.

The ssi process of Diggle, *et al.* (1976) is the variant of Matérn's second model conditioned on the total number of points.

3.3. Fixed-range Interactions

Throughout the rest of this section we will define point processes on a bounded set A by the Radon–Nikodym derivatives f of their distributions with respect to that of a Poisson process. Thus f is a function from the class of finite subsets of A to $(0, \infty)$. For each realization x we can think of $f(x)$ as a measure of how much more frequently the finite set of points x occurs in this process than in the Poisson process, in the same way that the density of a continuously distributed random variable heuristically measures the frequency of occurrence of a particular value. Almost always f will be defined up to a constant factor which cannot be given explicitly. We demand that f be *hereditary*, that is if $f(x) > 0$ then $f(y) > 0$ for all $y \subset x$.

Pairs of points whose distance apart are less than $r = 2R$ are called *neighbours*. In a hard-core model such pairs are excluded. Consider a Borel set B and its complement $A \setminus B$. Let

$$B^* = \{\xi; \xi \notin B, \exists \eta \in B \text{ with } d(\xi, \eta) < r\}.$$

If for every B the distribution of the process restricted to B conditional on the realization in $A \setminus B$ is the same as that conditional on that part of the realization in B^* the process is said to be *Markov* of range r . (In statistical mechanics such a process is said to have finite range (Preston, 1976b)). This concept extends that of a Markov random field on a finite lattice or graph (e.g. Besag, 1974; Preston, 1974). We have the following extension of the Markov field theorem:

$$f(x) = \prod_{y \subset x} \phi(y),$$

where $\phi(y) = 1$ unless every pair of points in y are neighbours (Ripley and Kelly, 1977, cf. Hammersley, 1975). Notice that the requirement that f be hereditary is a natural relaxation of what Hammersley (1972) called the “positivity condition”.

Suppose the conditional density of a point ξ given the realization in $A \setminus \{\xi\}$ depends only on the *number* of neighbours of ξ . Then

$$f(x) = ab^{\#(x)} c^{s(x)}$$

where $\#(x)$ is the number of points in x and $s(x)$ is the number of pairs of neighbours. Here b may be any positive number (which may be suppressed by increasing the intensity of the underlying Poisson process by a factor of b) but $0 \leq c \leq 1$ unless the total number of points is limited. The case $c = 0$ is our third hard-core model. For a fixed sample size and $c > 0$ this formula was given by Strauss (1975); the quoted result is taken from Kelly and Ripley (1976).

3.4. Pairwise Interactions

In the last class of models the interaction between pairs of points depends only on whether or not they are neighbours. In this subsection we let this interaction depend on the distance between the pair. We define

$$f(x) = ab^{\#(x)} \prod_{\substack{\xi, \eta \in x \\ \xi \neq \eta}} h\{d(\xi, \eta)\}$$

for constants a and b and an interaction function $h : (0, \infty) \rightarrow [0, \infty)$. If we consider realizations with two points we see that $h\{d(\cdot, \cdot)\}$ is a joint density for the pair of points. The product

form allows us to combine all these pairwise interactions; it was suggested by the formulae of statistical mechanics (e.g. Ruelle, 1969) where $-\log h(s)$ is proportional to the interaction energy between two particles distance s apart. We have to check whether for a given b and h this function is the Radon–Nikodym derivative of a probability for some choice of a . In general this need not be so. We shall use these models only for bounded functions h which vanish on $(0, r)$. Then $f(x) = 0$ if $\#(x)$ exceeds the finite number N of balls of radius $(r/2)$ which may be packed into A , and $f(x) \leq ab^N (\sup h)^{N^{1/2}}$, so f is integrable.

Note that the final models of the last subsection are obtained by setting

$$h(d) = \begin{cases} c, & d < r, \\ 1, & d \geq r. \end{cases}$$

3.5. Mixtures of Models

Classes of models have been proposed which are mixtures of simpler models, the most important being mixed Poisson and doubly stochastic Poisson (or Cox) processes. Such models cannot be fitted from a single sample unless parametric assumptions are made about the random mean measure; see Grandell (1972).

4. EVOLVING POINT PROCESSES

In this section we show how point processes on a bounded set A may arise as the equilibrium distributions of spatial-temporal stochastic processes. These are of two types; we may think of them as modelling botanical and zoological phenomena.

4.1. Spatial Birth-and-death Processes

This class of Markov jump processes was introduced by Preston (1976a). Suppose we are interested in a point process whose distribution is specified by a Radon–Nikodym derivative f with respect to that of the Poisson process of unit intensity on a bounded set A . The state space of our stochastic process will be $\mathcal{L} = \{x : f(x) > 0\}$, i.e. the class of finite sets which have positive “density”. Thus viewed at any fixed time the stochastic process is a point process. The process evolves in jumps, a finite number of which occur in each bounded time interval. These jumps are of two types, a “birth” in which a single point is added, or a “death” in which one of the points is deleted. The probability that a point within F , a measurable subset of A , is added to the finite set x between times t and $t+s$ is $s \int_F b(x, \xi) d\nu(\xi) + o(s)$, and the probability that the point ξ is deleted from $x \cup \xi$ in this time interval is $D(x, \xi) s + o(s)$. The functions b and D characterize a spatial birth-and death process; of course we have to impose conditions on b and D to ensure that the process exists, in particular that we do not have an infinite sequence of jumps in a finite time. Sufficient conditions and a formal construction are given by Preston.

We would like the point process which results from viewing our stochastic process at each fixed time to be that specified by f , i.e. we would like this to be the equilibrium distribution. Suppose the following condition holds:

$$b(x, \xi) f(x) = D(x, \xi) f(x \cup \xi) > 0 \quad \text{if } x \cup \xi \in \mathcal{L}. \quad (1)$$

This is a “detailed balance” condition which ensures that births from x to $x \cup \xi$ match deaths from $x \cup \xi$ to x . The stochastic process is then time-reversible and the point process specified by f is its unique equilibrium distribution. (This may be checked directly. The uniqueness is not given by Preston but follows from the indecomposability of \mathcal{L} ensured by (1) and the assumption that f is positive on the hereditary set \mathcal{L}).

Thus we can consider our point process evolving in time by the addition and deletion of points which remain fixed throughout their lifetimes. The ratio

$$b(x, \xi) / D(x, \xi) = f(x \cup \xi) / f(x)$$

of the birth rate to the death rate is a measure of the “viability” of a point at ξ given the rest of the points x . For a Markov point process this ratio is $\prod_{y \in x \cup \{\xi\}} h(y \cup \xi)$ which depends only on the interaction of ξ with the set of its neighbours in x . For a pairwise interaction process we have $\prod_{y \in x} h\{d(\eta, \xi)\}$ which is a measure of the “crowding near ξ ”.

We have not yet shown that (1) imposes sufficient conditions on b and D for the existence of our stochastic process; indeed it does not. We can, however, always take $b(x, \xi) = 1$ for $x \cup \xi \in \mathcal{L}, 0$ otherwise. Then proposition 5.1 of Preston (1976a) shows that the process is well defined. Preston also considers conditions under which $P\{X(t) \in F\}$ converges to the probability given to the measurable subset F of \mathcal{L} by the equilibrium point process. If this point process has a bounded number of points, for instance any hard-core model, then the process converges in this sense to its unique equilibrium distribution given by f .

For the Strauss process, Kelly and Ripley (1976) considered the spatial birth-and-death processes given by $b(x, \xi) = 1$, $D(x, \xi) = \exp\{-t(x, \xi)\}$ and $b(x, \xi) = \exp\{t(\xi, x)\}$, $D(x, \xi) = 1$, where $t(\xi, x)$ denotes the number of neighbours of ξ in x . Both processes are well defined and converge to the Strauss process.

4.2. Diffusions

For pairwise interaction processes conditioned on the total number of points N we can find a model with continuously moving particles. The N particles perform (mathematical) diffusions with constant infinitesimal variance but with drift the vector sum of $\nabla_j \log h(d(x_j, x_i))$ over all the other points. We have to impose boundary conditions on A ; these will be either reflecting or periodic. This process is defined at least if h is positive and infinitely differentiable, when with periodic boundaries there is a unique equilibrium density f which is the solution of

$$\nabla_i \log f = \sum_{j \neq i} \nabla_j \log h\{d(x_j, x_i)\}, \quad i = 1, \dots, N$$

(J. T. Kent, personal communication); this is satisfied by our product formula for f . This process converges weakly to the pairwise interaction process conditioned on a total of N points. These assertions are very plausible for reflecting boundaries and more general interaction functions h .

We can combine this process with a spatial birth-and-death process to obtain a process with equilibrium distribution the complete pairwise interaction process. However, when using this type of model we will usually be interested in time-scales short compared with the lifetimes of the points.

5. SIMULATION

Our methods of testing the goodness of the fit of a model will depend on our ability to simulate a proposed model.

It is clear how one would simulate a Poisson point process on a bounded set A ; one chooses the number of points from a Poisson distribution with mean $\lambda\nu(A)$ and distributes these uniformly and independently. To obtain a uniformly distributed point on A we can enclose A in a rectangle and generate a uniformly distributed point on this rectangle by choosing its co-ordinates uniformly. If this point lies within A we accept it; if it does not we repeat the procedure until we generate a point within A .

Those point processes defined by a derivative f with respect to a Poisson process may in principle be simulated by rejection sampling if f is bounded (which might be achieved by a suitable choice of the Poisson intensity α or by restricting attention to samples of a given size). In this procedure f is calculated for a sample of the Poisson process and compared with a uniformly distributed random variable X on $(0, \sup f)$; the sample is accepted if X is smaller otherwise the process is repeated. If f has a non-continuous distribution we must refine this procedure. The only case we shall meet is that in which f takes the value 0 with positive probability; of course we reject all such samples. An example is our third hard-core

model; in this case the probability of an acceptable sample is about $\exp(-\alpha^2 \pi r^2)$, about 10^{-24} for the example of Section 7. (The number of neighbour pairs per unit area has mean about $\alpha^2 \pi r^2$ and is approximately Poisson, cf. Brillinger, 1976). In this case the rejection method is impracticable; this appears to be a common conclusion.

Our simulations were obtained from the spatial birth-and-death processes of Section 4. We could choose a stochastic spatial-temporal process with the required equilibrium distribution and simulate this over a long time interval to obtain a sample of a point process with distribution "close" to that required. We will, however, be interested in samples containing a fixed number of points N . Either we choose $D(x, \xi) = 1$ and

$$b(x, \xi) = f(x \cup \xi)/f(x) \quad \text{if } \#(x) = N-1,$$

both 0 otherwise, or if f is positive we may choose $b(x, \xi) = 1$ and $D(x, \xi) = f(x)/f(x \cup \xi)$ if $\#(x) = N$, 0 otherwise. Suppose we observe these processes only when jumps occur. We obtain a discrete-time Markov process in which births and deaths alternate. In the continuous-time processes all states with N points are left at unit rate, so an equilibrium distribution for the discrete-time process has a density proportional to f over states with N points (with respect to Lebesgue measure on $(R^2)^N$). It is easily checked that these Markov chains are indecomposable and so converge to a unique equilibrium distribution (e.g. Breiman, 1968, chapter 7).

Although it was convenient to use spatial birth-and-death processes to show that these Markov chains converge to the required point process distribution we can describe them independently. The first is defined by selecting one of the points at random and deleting it, then adding a point with a density proportional to $p(\xi) = f(x \cup \xi)/f(x)$. This random point might be generated by rejection sampling. This process (every second step of the Markov chain) converges to the point process defined by f conditional on N points. The second process is defined similarly; we add a uniformly distributed point then delete one point chosen at random according to a distribution with $P(\xi \text{ is deleted from } x)$ proportional to $f(x \setminus \xi)/f(x)$. Because this process does not involve rejection sampling it might seem preferable to use it; empirically it tends to add a point and delete it immediately, so converging to equilibrium very slowly. In our examples the first process was used.

To use any of these stochastic processes we have to know that they are in equilibrium. We can generate a starting pattern of points from a Poisson process or from a lattice of points if as in a hard-core model only certain finite sets of points are allowed. We can use the methods of the next section to decide whether the stochastic process has for practical purposes reached equilibrium. We can then sample the process at sufficiently long intervals; for our examples this was every $4N$ steps.

6. FITTING MODELS

In the introduction we mentioned several methods which have been used to test the fit of a Poisson model. In this section we consider how we might test the fit of a quite general model. Statistical inference from just one sample will be possible because our assumption of stationarity provides in-built replication.

Nearest-neighbour methods can essentially only test for the presence of pattern on a single scale (in the ecologists' terminology). The use of second or third nearest neighbours or Besag and Gleaves' (1974) T-square methods can extend the range of scales slightly, but not for highly clustered data (see Section 8). It seems to be virtually impossible to obtain analytical information on the distribution of distance method test statistics under most non-Poisson models. (For exceptions see Besag and Gleaves, 1974; Diggle, 1975; Diggle et al., 1976; Holgate, 1965a, b, c.) Bartlett's (1974, 1976) example of herring gull nest data is a case in which the scale of distance methods is appropriate, but ecologists are often interested in the interaction of clusters of objects (cf. Greig-Smith's contribution to Bartlett, 1971). Thus distance methods did not seem appropriate for the intensive analysis of a map.

Finding the distribution of the numbers of points in each of a set of quadrats is straightforward; it amounts to partially specifying the model. However, even a test of the Poisson null hypothesis for a collection of scales has been described only recently (Mead, 1974). The use of counts on a grid of contiguous quadrats disregards the fine structure of the pattern. A different type of model and analysis is appropriate to such a lattice process (Besag, 1974; Bartlett, 1976).

Bartlett (1964, 1976) considered the two-dimensional spectral analysis of a spatial point pattern. Spectral analysis is closely tied to the use of the space itself as the group governing stationarity and so is most appropriate when we assume stationarity under translations only (cf. Ripley, 1976a). The spectral estimate must be smoothed; the choice of the degree of smoothing seems necessarily to introduce an arbitrary factor into the analysis.

The two methods proposed here are based on estimators of the two one-dimensional summaries of a model described in Section 2.

An unbiased estimator of $\lambda^2 K(t)$ was described in Ripley (1976a). We give each ordered pair (x, y) of objects a weight $k(x, y)$ which is the reciprocal of the proportion of the perimeter of the circle centred on x and passing through y which is within the sampling window. Thus $k(x, y)$ is inversely proportional to the chance that a point the distance of y from x had of being recorded. Let

$$\lambda^2 \hat{K}(t) = \sum k(x, y)/\nu(E),$$

the sum being over ordered pairs of distinct data points less than t apart. (Remember $\nu(E)$ is the area of the sampling window.) This is an unbiased estimator of $\lambda^2 K(t)$ for

$$t \leq t_0 = \inf\{s : \exists x \in E, \partial b(x, s) \cap E = \emptyset\}.$$

Thus if the intensity λ is known, $\hat{K}(t)$ is an unbiased estimator of $K(t)$ for small enough t . If λ is unknown we estimate it by the obvious quantity $\hat{\lambda} = Z(E)/\nu(E)$ and define

$$\hat{K}(t) = \sum k(x, y) \nu(E) N^{-2}, \quad \text{if } Z(E) = N.$$

In practice the possible bias of \hat{K} seems to be small.

For a Poisson process we can calculate this bias exactly. Suppose we have a sample of N points. If Y is uniformly distributed on E then

$$\begin{aligned} E\{1_{(0 < d(x, Y) < t)} k(x, Y)\} &= \int_0^t \left\{ \int_{E \cap \partial b(x, s)} k(x, y) dm \right\} ds / \nu(E) \\ &= \pi [\min[t, \inf\{s : \partial b(x, s) \cap E = \emptyset\}]^2 / \nu(E)], \end{aligned}$$

where m is Lebesgue measure on $\partial b(x, s)$, the circumference of $b(x, s)$. Thus for $t \leq t_0$ the conditional expectation of each term in the sum is $\pi t^2/N^2$, so

$$E\{\hat{K}(t) | Z(E) = N\} = \pi t^2(N-1)/N.$$

This correction factor is possibly related to the use of $\{N/\nu(E)\}^2$ as an estimate of λ^2 whereas $N(N-1)\nu(E)^{-2}$ is unbiased.

The advantage of an unbiased estimator of K is that one can develop an intuition about the process from K without taking into account the shape of the sampling window. Bartlett (1964) suggested the use of the distribution of the interpoint distances uncorrected for edge effects. Usually we cannot even calculate this distribution or the mean of its empirical version.

The *test-set* method is based on the function

$$\hat{p}(t) = \nu[\{x : (x + tB) \cap (\text{data}) \neq \emptyset\}] / \nu(E)$$

for a *test set* or *structuring element*, B , usually a circle, hexagon or square centred at the origin. For a process that is stationary under translations, in particular for a model, this is a nearly

unbiased estimator of the probability that tB is not empty, the bias being introduced by edge-effects. Because $P(Z(B) > 0)$ is not additive in B it is not possible to introduce an edge-correction analogous to \hat{K} . We can obtain an unbiased estimator by considering only points within

$$E_{tB} = \{x: x + tB \subset E\}.$$

Then

$$\begin{aligned} E\{\hat{p}(t)\} &= E(\nu[\{x: x \in E_{tB}, (x + tB) \cap (\text{data}) \neq \emptyset\}]) / \nu(E_{tB}) \\ &= \int P\{Z(x + tB) > 0\} d\nu / \nu(E_{tB}) = P\{Z(tB) > 0\} \end{aligned}$$

by stationarity. Thus if we take B to be the unit circle we obtain an unbiased estimator of $p(t)$ (defined in Section 2).

The computation of \hat{p} is best done by an image analyser such as the *texture analyser* developed at Fontainebleau by Serra (Klein and Serra, 1972; Serra, 1972; Watson, 1973). These machines compute the proportion of those x 's on a fine grid for which $(x + tB)$ contains a datum point. This method can also be applied to general (non-point) spatial patterns; some of the relevant theory is given by Matheron (1967, 1975).

Both \hat{K} and \hat{p} provide useful summaries of the data even when no model is envisaged. Very little is known about the sampling fluctuations of either statistic even for a Poisson model. Some asymptotic results for \hat{K} have been given by Silverman (1976).

We are forced to use simulation to explore the sampling fluctuations, and we will use these simulations to generate the critical regions of our "goodness of fit" tests. Suppose we plot \hat{K} for the data, K for the model (if known) and \hat{K} for m simulations of this model. Sometimes it will be obvious that K for the data is radically different from the other plots and no formal test is then needed. If we had a particular interpoint distance in mind before observing the data we could compare the ordinates of \hat{K} at this distance for the data and the simulations. On our null hypothesis that the data were generated by the model we have $(m+1)$ independent samples of a random variable, so the probabilities that the data gave the upper or lower extreme value are each $1/(m+1)$ (cf. Hope, 1968, and Barnard's contribution to the discussion of Bartlett, 1963).

It is more reasonable to assume that we are interested in a certain range of interpoint distances. We can define the acceptance region of a test by requiring \hat{K} for the data to be within the envelope of \hat{K} for the simulations throughout this range. The size of this test must exceed $2/(m+1)$; it may be estimated by the proportion of a further batch of simulations which are rejected. We may consider one-sided tests with acceptance region the area above the lower envelope or below the upper envelope. These will be appropriate if we have alternatives which are "self-inhibiting" or "clustered" respectively. In general our two-sided tests are against all other models.

This discussion ignores the fact that we will have to fit some parameters of a family of models. Even in the Poisson case we have to choose the intensity. We side-step the problem of unknown intensity by conditioning on the total number of data points and generating simulations with the same number of points as the data. For Poisson models we can justify this conditioning by an appeal to a conditionality principle. In general all we can say is that conditional inference seems reasonable and that our revised procedures yield valid conditional tests.

There may be other parameters which we will have to fit from the data in a formal or informal way; of course this will affect our significance levels. As a general principle we should only fit families of models with few parameters compared to the number of data points, so the effect should be small. A counsel of perfection is to divide the data into two parts, using one to establish a suitable model which is then tested against the other.

We can use \hat{p} in exactly the same way. With either statistic we are essentially using the combination of a large number of tests at different ordinates which are highly inter-dependent. The suggested procedure is not ideal but needs little human judgement.

Both K and p are essentially cumulative functions; the reader may be wondering why we have chosen to estimate these rather than their densities. One reason is that we would have to smooth these estimates which it seems impossible to do in an objective way. A further reason is that the author's limited experiments suggest that tests based on such smoothed estimates are considerably less powerful than their cumulative analogues. This point is discussed further in the examples.

7. CELL SHAPES AND COMPARTMENTS

In this section we model a patch of biological cells in order to provide a solution to the following problem posed by Dr Francis Crick; the biological background is given by Crick and Lawrence (1975). At an early stage in the development of an adult insect there is a roughly circular patch of about 20 two-dimensional contiguous cells. These divide simultaneously, then the patch of daughter cells is split into two roughly equal compartments by an approximately straight line following the cell boundaries. The problem is to find the proportion of sister cell pairs which are in different compartments.

We approach this problem by first analysing the patch of cells at a single point in time. A sample patch of cells (with somewhat idealized boundaries) is shown in Fig. 1. Inspection

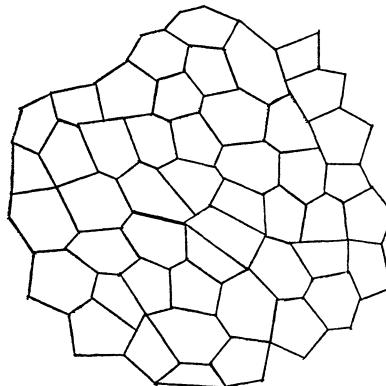
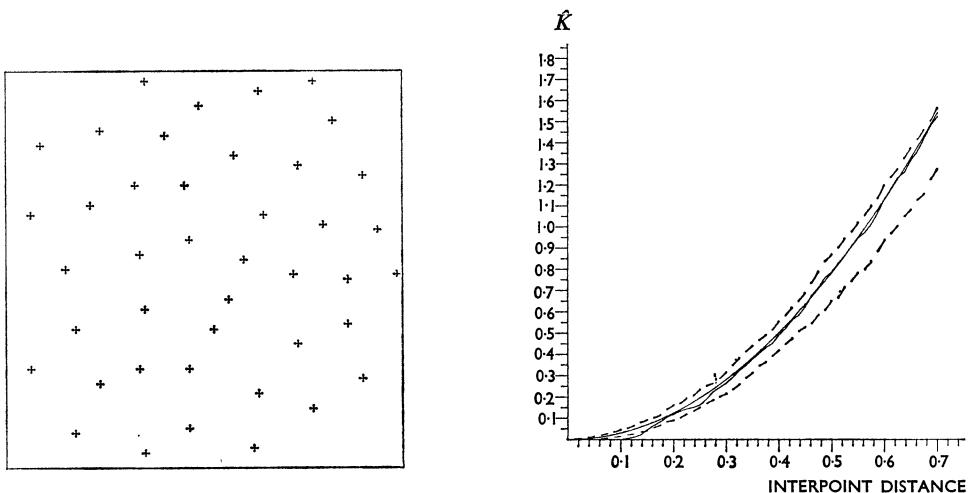


FIG. 1. A typical patch of cells.

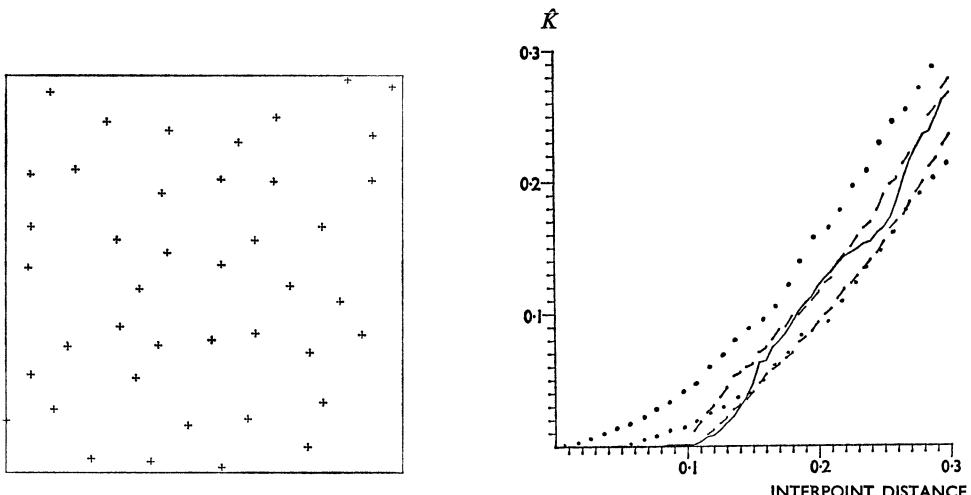
of this figure shows that it is reasonable to approximate the cells by the Dirichlet cells of their centres of gravity, except at the edges of the patch where the boundary is undefined. (The Dirichlet cell or Voronoi polygon of a cell centre is the set of points nearer to this centre than to any other.)

To circumvent problems introduced by biological edge-effects I took a centrally located sample square E containing the 42 cell centres shown in Fig. 2. The units were chosen to make E the unit square. The initial analysis is based on this sample. Fig. 3 shows \hat{K} for the data and K for the Poisson process. These are clearly different up to 0·1 but similar throughout the range (0·2, 0·7). The dashed lines are the envelope of \hat{K} for 99 simulations of the Poisson process conditioned to produce 42 points. The area above the lower envelope (for $0 < t < 0·7$) provides the acceptance region of a significance test of size about 7 per cent (estimated by further simulation); the Poisson null hypothesis is rejected; the part of the band in the range (0, 0·3) is shown enlarged on Fig. 5; the form of \hat{K} for the data shows that there are no pairs of cell centres closer than 0·1 apart, which suggests we should fit a hard-core model with inhibition distance $r = 0·1$.



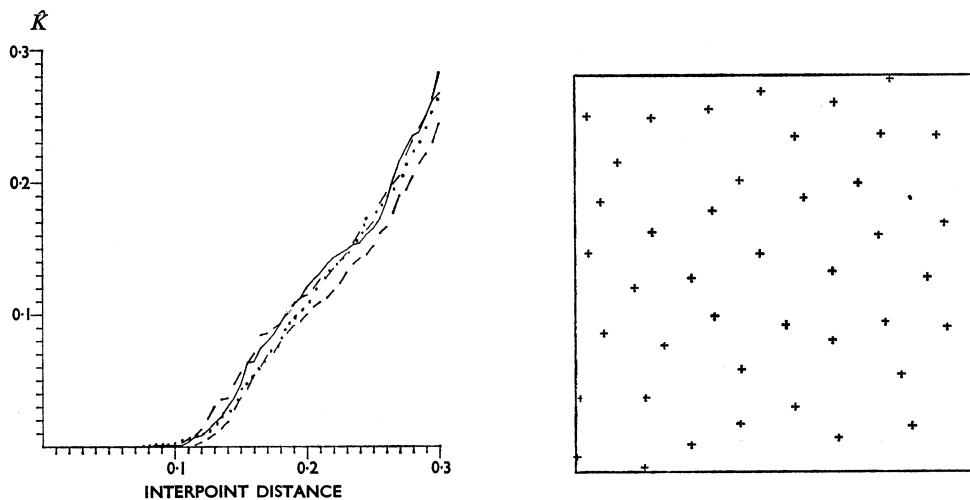
Left. FIG. 2. The sample square containing 42 points derived from the patch of biological cells.

Right. FIG. 3. \hat{K} plotted for the cell data and the envelope of the plots of \hat{K} for 99 simulations of the Poisson process conditioned to produce 42 points. The smooth curve is the theoretical value of K for the Poisson process.



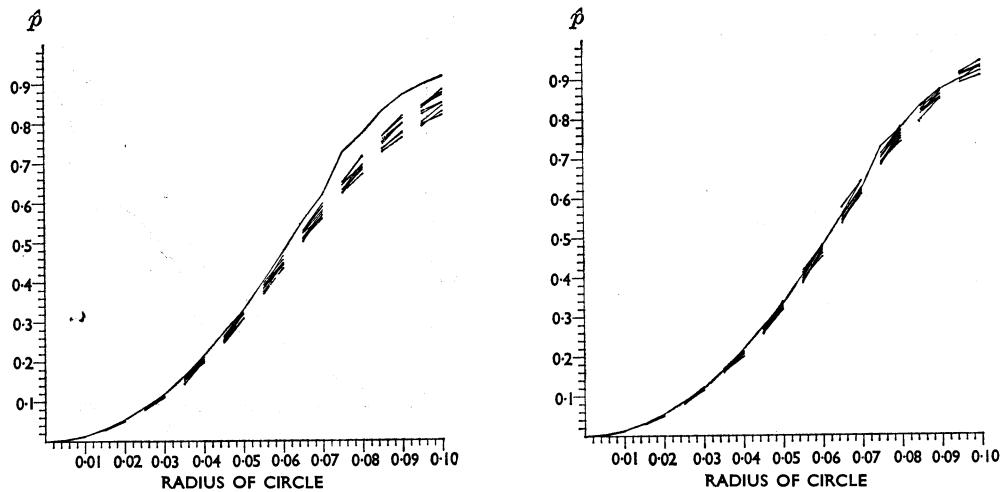
Left. FIG. 4. A typical simulation of the third hard-core model with periodic boundary conditions, containing 42 points.

Right. FIG. 5. \hat{K} plotted for the cell data and the envelopes of the plots of \hat{K} for 49 simulations of the hard-core model (the dashed curves) and of 99 simulations of the Poisson process (the dotted curves); all samples contained 42 points.



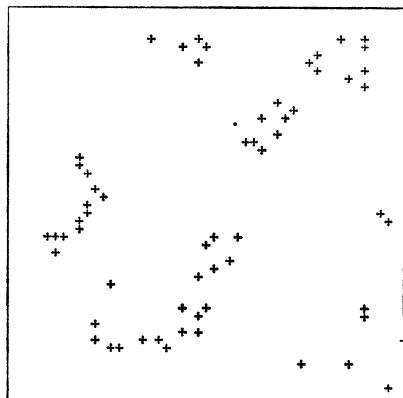
Left. FIG. 6. \hat{K} plotted for the cell data and the envelope of 20 simulations of the pairwise interaction process with periodic boundary conditions. The dotted curve is \hat{K} for another sample of 89 cell centres.

Right. FIG. 7. A typical simulation of the pairwise interaction process with repulsion in the range 0.10–0.14.

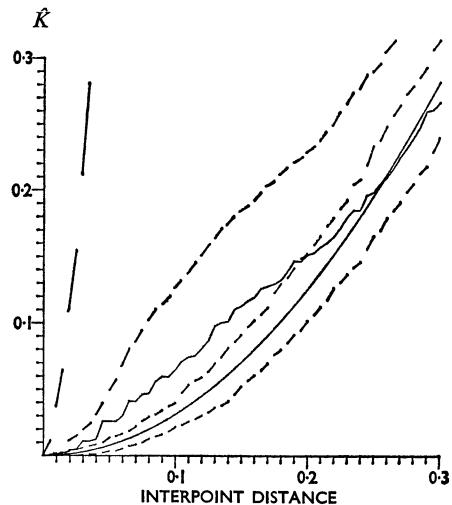


Left. FIG. 8. Plots of $\hat{\beta}$ for the cell data and 10 simulations (the dashed lines) of the hard-core model.

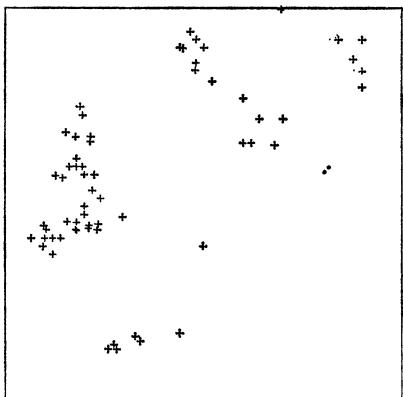
Right. FIG. 9. Plots of $\hat{\beta}$ for the cell data and 10 simulations (the dashed lines) of the pairwise interaction model.



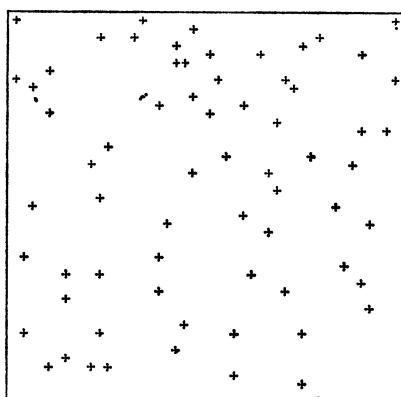
Left. FIG. 10. The positions of 62 Redwood seedlings.



Right. FIG. 11. The solid curves are the plots of \hat{K} for the Redwood data and of K for the Poisson process (the parabolic curve). The lower and upper pairs of dashed curves are the envelopes of the plots of \hat{K} for 99 simulations of the Poisson process and 20 samples of Strauss' model.



Left. FIG. 12. A typical simulation of Strauss' model, containing 62 points.



Right. FIG. 13. The centres of 69 towns in a 40-mile square in Spain.

For this value of r , Matérn's first process has a maximum intensity of about 11·7; his second model is biologically inappropriate. We fit the third hard-core model with periodic boundary conditions on E . A typical simulation of this process with 42 points is shown in Fig. 4, and Fig. 5 shows \hat{K} for the data compared with 49 simulations of this process. It is clear that the formal significance test rejects the hard-core model (at an estimated 20 per cent), on the basis of Fig. 5 one might be prepared to accept it. However, the simulations appear less regular than the data (compare Figs 2 and 4), having "holes" which would give rise to some unacceptably large Dirichlet cells.

Thus we will continue to search for a satisfactory model; we try a model with more general pairwise interactions. Unfortunately there seems to be no general way to estimate the interaction function h from the data. The curves of Fig. 5 show that the data have fewer pairs of points with distances in the range (0·1, 0·14) than the simulations of the hard-core model. This suggested we should introduce a repulsive force in this range; it is convenient to make this force constant and take $h(s) = \exp\{k(s-0·14)\}$ in this range, $h(s) = 0$ for $s \leq 0·1$, and $h(s) = 1$ for $s \geq 0·14$. There is a point process defined by h for each positive value of k . The "constant" in the product formula for f is an unknown function of k , hence so is the likelihood function for k and no formal statistical method may be used to estimate k . The value $k = 80$ was chosen to obtain a good fit of \hat{K} for the simulations to \hat{K} for the data. Fig. 6 shows the plots of \hat{K} for the data and the envelope of 20 simulations. Fig. 7 shows a typical simulation; notice that the "holes" have disappeared. This model corresponds to cells with a hard core and soft edge.

It might be objected that it is unreasonable to fit such a complex model with so many parameters to just one small sample. This objection is countered by the plot of \hat{K} on Fig. 6 for another sample of a patch of cells containing 89 centres. The agreement between the two samples and the simulations of our model is sufficient for us to accept this as our model.

The existence of "holes" shows up well in the analysis based on \hat{p} . Figs 8 and 9 show \hat{p} for the cell data and ten simulations (the dashed lines) of the hard core and final models respectively. The test set B here was a circle. The superior fit of the pairwise interaction model is clearly shown. The small number of simulations is due to the expense of this procedure in computer time; these figures took about 5 minutes compared with 30 seconds for Fig. 5.

7.1. The Next Generation

The exact nature of the transition of a patch of cells from one generation to the next is unknown. We seek a plausible model which will preserve the statistical properties of the pattern of the cell centres. We assume that each cell splits into two daughter cells and then we rescale all distances by $\sqrt{2}$ to allow for the expansion of the patch following the doubling of the number of cells.

In the first phase we replace each cell centre by two centres at the ends of a line segment centred on the parent cell centre. This segment has a uniformly distributed orientation independent for each parent, and a length equal to half the average cell diameter. This models the division of a cell along a "diameter". This set of $2N$ cell centres is then allowed to move under forces which are constant direct repulsions between each pair of neighbouring centres (less than 0·1 apart) until there are no neighbours.

The second phase is an approximation to the diffusion discussed in Section 4. We replace the diffusion by a random walk performed by the cells which are assumed to have hard cores which behave as perfectly elastic spheres should they collide. The motion proceeds in discrete steps. First each centre is moved a distance d_1 away from any other centre which was between 0·10 and 0·14 distant, and then a distance d_2 in a random uniformly distributed direction independently for each centre. If this perturbation causes the hard cores to collide then they bounce apart elastically. (This process is closely related to the Manx motion of Kendall, 1974b).

The essential parameter of this second phase is the relationship between d_1 and d_2 which must be connected with the parameter k of our model. Consider just two points distance d apart away from the boundary of E . Let $D = 2(d_1 + d_2)$; we will consider the distribution of d in the range $(0.10 + D, 0.14 - D)$. Let $p(d)$ denote the equilibrium density for d which must satisfy

$$p(d) = \int p(d - 2d_1 - 2d_2 \cos \theta) d\theta / 2\pi;$$

this is satisfied by $h(d) = c \exp \{k(d - 0.14)\}$ if

$$\exp(2kd_1) = I_0(2kd_2) \sim 1 + (kd_2)^2,$$

where I_0 is the modified Bessel function of the first kind of order zero. Since the equilibrium distribution conditional on two points must be proportional to h for our model, we assume the approximate relation $d_2^2 = 2d_1/k$. This process was simulated with $d_2 = 1/200$, $d_1 = 1/1,000$; the resulting pattern of cell centres after 50 steps produced functions \hat{K} which accorded well with those for the original data and our model. This process is an indecomposable Markov chain which thus converges to its equilibrium distribution (in the sense described for spatial birth-and-death processes).

7.2. Compartments

This modelling enables us to mark on a simulated patch of cells which pairs of cells are sisters. There seems to be no way in which this could have been done directly on Fig. 1. We divide the patch of daughter cells by a random straight line passing near the centre of the patch; cells are allocated to a compartment according to the side of this line on which their centres fall. Suppose we have a roughly circular patch of $2N$ daughter cells. The length of the dividing line between compartments will be proportional to $N^{1/2}$, which suggests that the average proportion of sister cell pairs which are in different compartments will be $cN^{-1/2}$ where c is approximately constant. From simulations on $N = 40$, c was estimated as about 0.4.

Another model for the formation of compartments which yielded the estimates quoted by Crick and Lawrence is discussed in Ripley (1976c).

8. OTHER EXAMPLES

8.1. A Redwood Plot

The data for this example were taken from Strauss (1975). The 62 points are shown in Fig. 10; they were taken from the clustered group of seedlings of his Fig. 1. This part of the plot was not square so for computational convenience I took a square mainly within the region of supposed clustering.

The plot of \hat{K} compared with the envelope of 99 simulations of the Poisson process producing 62 points is shown in Fig. 11. Again we choose our units to make E the unit square. Strauss suggested that the plants might interact up to $r = 0.08$ (about 6 feet), so the scale of the plot was chosen to cover $(0, 0.3)$. The ordinates at 0.08 provide an exact 1 per cent one-sided significance test which clearly rejects the Poisson model. The whole upper envelope gives a test of size about 3 per cent.

Strauss fitted the density $f(x) = ac^{s(x)}$ by an approximation method and estimated $\log c$ as 1.03. Samples from this density were simulated by the method described in Section 5; a typical sample is shown in Fig. 12. It is evident that this model does not fit the data, as the plots of \hat{K} on Fig. 11 also show. Simulations of this model for other values of c and r showed that it cannot be fitted to the data. It seems to produce tightly bound clusters rather than the loose clusters in the data.

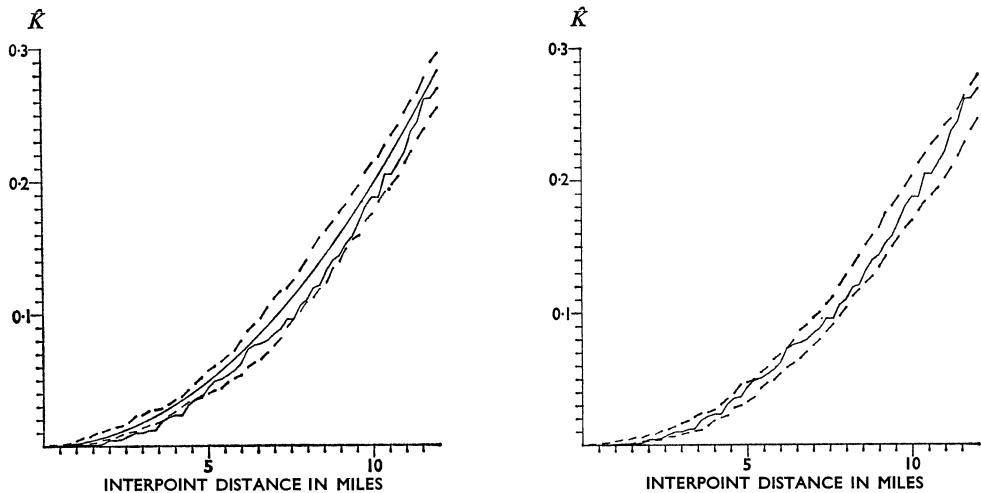
Notice that the plot of \hat{K} for the data crosses K for the Poisson process at about 0.25. In fact \hat{K} and K are very close throughout $(0.25, 0.7)$. For a cluster process the function $K(t) - \pi t^2$

is an increasing function if the parent process is Poisson (because $\text{cov}\{Z(E), Z(F)\} \geq 0$ for any E and F). Thus no such cluster process can fit the data. It appears that there is both clustering and inhibition between clusters. In addition, \hat{K} for the data is zero up to a distance corresponding to about 1 foot, suggesting some local inhibition (which is certainly not statistically significant).

I have been unable to find any model which fits these data. The number of *clusters* involved is too small to merit detailed analysis.

8.2. Central Places

Geographers are interested in the distribution of towns in homogeneous regions; Cliff and Ord (1975) give a survey and further references. The Poisson random pattern is usually contrasted with some form of uniform spacing. Fig. 13 shows the centres of 69 towns in a geographically homogeneous square on the Spanish plateau. Fig. 14 shows the plots of \hat{K}



Left. FIG. 14. \hat{K} for the Spanish towns compared with K for the Poisson process and the envelope of \hat{K} for 99 simulations of that process.

Right. FIG. 15. \hat{K} for the Spanish towns and the envelope of 20 simulations of Strauss' model with 69 points, $c = 0.5$ and an inhibition distance of 3.5 miles.

for the data and 99 simulations of the Poisson process. Here we are interested in a one-sided test so we take the acceptance region to be the set above the lower envelope; this has an estimated size of 3 per cent. There is a marginal deviation from the Poisson null hypothesis at an interpoint distance of about 3.5 miles.

Glass and Tobler plotted a histogram estimate of $g(r)/2\pi r$ with a bin size of 1 mile and an edge correction similar to ours. They failed to consider the sampling fluctuations of their estimate but claimed to reject the Poisson hypothesis and to fit a hard-core model with $r = 2R$ equal to 3.46 miles. This is absurd since Fig. 14 shows that there are pairs of towns closer than this distance (as does their Fig. 2!). Simulations showed that the unsmoothed histogram is too rough to be useful and that the non-Poisson behaviour is swamped by the sampling fluctuations (the details are given in Ripley, 1976c).

The shape of \hat{K} for the data suggested that we should fit an interaction model with interaction distance $r = 3.5$ miles. Strauss' model was fitted by comparing \hat{K} for the data with simulations. The plots for $c = 0.5$ are shown in Fig. 15; this model fits well, the choice of c being not at all critical.

The analysis of the towns in Wisconsin used by Brush (1953) gave very similar results with an interaction range of about 6 miles.

8.3. Galaxies

The observed positions of galaxies in the celestial sphere and their estimated positions in space form vast data-sets of which parts may be regarded as homogeneous. Parts of various galaxy catalogues have been analysed by Peebles and his colleagues. They estimate the two-point correlation function w which is interpreted in the Appendix as $(g/\lambda^2 - 1)$. A power law $w(r) = cr^{-\alpha}$ was fitted to a histogram estimate of w (Peebles, 1974). There appears to be no general way to decide when a postulated λ and K summarize the first and second moments of a model. We note that for a Neyman–Scott cluster model $c(r)w(r)$ is a power law proportional to the density of the distribution between two cluster points, which cannot be of power law form (this is for galaxies in space). Thus no Neyman–Scott process corresponds to this choice of K .

Peebles and Groth (1975) considered three-point correlation functions for a point process on a sphere. This is a symmetric function z of three arguments such that

$$E\{Z(A)Z(B)Z(C)\}$$

$$= \lambda^3 \int_{A \times B \times C} \{1 + w(\theta_1 - \theta_2) + w(\theta_2 - \theta_3) + w(\theta_3 - \theta_1) + z(\theta_1 - \theta_2, \theta_2 - \theta_3, \theta_3 - \theta_1)\} d\theta_1 d\theta_2 d\theta_3$$

for all disjoint sets A , B and C . The function

$$\lambda^3 \{1 + w(\theta_1) + w(\theta_2) + w(\theta_3) + z(\theta_1, \theta_2, \theta_3)\}$$

is a density of the reduced third moment measure (cf. Krickeberg, 1973; Ripley, 1976a). This is a function of three arguments and is difficult to visualize. Nevertheless Peebles and Groth fitted a function $z(\theta_1, \theta_2, \theta_3)$ proportional to

$$w(\theta_1)w(\theta_2) + w(\theta_2)w(\theta_3) + w(\theta_3)w(\theta_1).$$

It is difficult to see what this would mean in terms of a possible model; the Gauss–Poisson processes (Newman, 1970; Milne and Westcott, 1972) have $z = 0$ but a similar form for the analogous four-point correlation function.

9. CONCLUSIONS

Our examples have indicated the usefulness of the statistics K and p in testing the fit of models. Analytical information on even their asymptotic distributions would be enlightening and might suggest more economical ways of generating significance tests. It might also be possible to obtain some results on the power of these tests; however, it would seem to be optimistic to hope that this might be possible for realistic alternatives. There is a shortage of suitable models, although the interaction models described here have broadened the field. These models would seem to be suitable for biological competition situations such as the nesting territories of birds (cf. the maps in Tubbs, 1974; and in Brown, 1976); indeed in this case the pattern is thought to result from pairwise competition.

The test set method of mathematical morphology has never before to my knowledge been applied to point processes. The theoretical unification between point processes and more general random sets expounded in Ripley (1976b) will I hope encourage applications; it would also be interesting to analyse patterns such as Fig. 1 directly as a random set.

ACKNOWLEDGEMENTS

I am indebted to all who have commented on this work, especially Julian Besag, Pierre Delfiner, Peter Diggle, Frank Kelly, John Kent, Bernard Silverman and Simon White. Francis Crick kindly provided me with the cell data; David Kendall introduced me to this problem and he and the referees made many useful comments on an earlier draft.

The computations were carried out by Fortran programs on the IBM 370/165 of the University of Cambridge Computing Service. Most of this work was carried out during the tenure of a Science Research Council studentship at the Statistical Laboratory, Cambridge.

APPENDIX

The interpretations of the reduced second moment measure $\lambda^2 K$ given in the text were obtained independently by Krickeberg (1976; lectures of 1974) and the author. We define the m th moment measure μ^m of a point process Z by

$$\mu^m\left(\prod_1^m A_i\right) = E\left(\prod_1^m Z(A_i)\right),$$

where A_i are bounded Borel sets. Let $P(\cdot \| x)$ be a Palm probability (Papangelou, 1974); this is a probability for each x such that

$$\int_E P(A \| x) d\mu^1(x)$$

exists and equals $\int_A Z(E) dP(z)$ for any Borel set E and measurable event A . We interpret $P(\cdot \| x)$ as the probability “given a point at x ”. It follows immediately that

$$\mu^m\left(\prod_1^m A_i\right) = \int_{A_1} \mu^{m-1}\left(\prod_2^m A_i \| x\right) d\mu^1(x).$$

Now suppose our point process is a *model*; we choose the unique invariant Palm probability considered by Papangelou. Then we have

$$\begin{aligned} K(t) &= \lambda^{-2} \mu^2[\{(x, y): x \in I, 0 < d(x, y) < t\}] \\ &= \lambda^{-1} \int_I \mu^1(b(x, t) \| x) d\nu(x) = \lambda^{-1} \mu^1(b(0, t) \| 0), \end{aligned}$$

where I is the unit cube. This is our interpretation (b).

Let K_0 be the distribution function of the measure representing $\nu \times \nu$. Suppose $\mu^2 \ll \nu \times \nu$ on $E_2 = \{(x, y): x \neq y\}$. (We know from Krickeberg (1973) that μ^2 is singular on the diagonal.) Then K is absolutely continuous with respect to K_0 , and if $g = \lambda^2 dK/dK_0$ then $f(x, y) = g\{d(x, y)\}$ is a version of $d\mu^2/d(\nu \times \nu)$ on E_2 . The factor $c(t) = dK_0/dt$, so

$$g(t) = (\lambda^2 dK/dt)/c(t).$$

Suppose A and B are cubes decreasing to the distinct points x and y respectively. Then

$$\mu^2(A \times B)/\nu(A) \nu(B) \rightarrow f(x, y), \quad \text{a.e.}$$

Thus if $d(x, y) = t$ then

$$g(t) = \lim_{A \downarrow x, B \downarrow y} \frac{E\{Z(A)Z(B)\}}{\nu(A) \nu(B)}$$

If we knew $E\{Z(A)Z(B)\} \sim P\{Z(A) > 0, Z(B) > 0\}$ we could regard $g(t)$ as the joint probability density for the occurrence of a pair of points distance t apart. Whereas we always have $E\{Z(A)\} \sim P\{Z(A) > 0\}$ (e.g. Leadbetter, 1972) this formula does *not* follow from our assumptions; a counter-example can be constructed by taking a Poisson cluster process on the group of rigid motions and a daughter process defined by the random set

$$(\{0, 1 + n^2 X: n = 1, 2, \dots\} \cap [0, 2]) \times \{0\}^{d-1},$$

where X is a uniformly distributed random variable on $(0, 1)$. This fact is often overlooked.

Under these assumptions we have

$$P\{Z(A) > 0, Z(B) > 0\} \sim \lambda^2 \{1 + w(t)\} \nu(A) \nu(B)$$

for “small” sets A and B about t apart, where $w = (g/\lambda^2 - 1)$ is Peebles’ “two-point correlation function”. Similar assumptions must be made to interpret the three-point analogue.

REFERENCES

- BARTLETT, M. S. (1963). The spectral analysis of point processes (with Discussion). *J. R. Statist. Soc. B*, **25**, 264–296.
- (1964). The spectral analysis of two-dimensional point processes. *Biometrika*, **51**, 299–311.
- (1971). Two-dimensional nearest-neighbour systems and their ecological application. In *Statistical Ecology*, Vol. I (G. P. Patil, ed.), pp. 179–194. Pennsylvania University Press.
- (1974). The statistical analysis of spatial pattern. *Advances Appl. Prob.*, **6**, 336–358.
- (1976). *The Statistical Analysis of Spatial Pattern*. London: Chapman and Hall.
- BESAG, J. (1974). Spatial interaction and the statistical analysis of lattice systems (with Discussion). *J. R. Statist. Soc. B*, **36**, 192–236.
- BESAG, J. and GLEAVES, J. T. (1974). On the detection of spatial pattern in plant communities. *Bull. Inst. Int. Statist.*, **45**, 153–158.
- BREIMAN, L. (1968). *Probability*. Reading, Mass.: Addison-Wesley.
- BRILLINGER, D. R. (1976). Estimation of the second-order intensities of a bivariate stationary point process. *J. R. Statist. Soc. B*, **38**, 60–66.
- BROWN, L. (1976). *British Birds of Prey*. London: Collins.
- BRUSH, J. E. (1953). The hierarchy of central places in Southwestern Wisconsin. *Geogr. Rev.*, **43**, 380–402.
- CLIFF, A. D. and ORD, J. K. (1975). Model building and the analysis of spatial pattern in human geography (with Discussion). *J. R. Statist. Soc. B*, **37**, 297–348.
- CRICK, F. H. C. and LAWRENCE, P. A. (1975). Compartments and polyclones in insect development. *Science*, **189** (4200), 340–347.
- DIGGLE, P. J. (1975). Robust estimation using distance methods. *Biometrika*, **62**, 39–48.
- DIGGLE, P. J., BESAG, J. E. and GLEAVES, J. T. (1976). Statistical analysis of spatial patterns by means of distance methods. *Biometrics*, **32**, 659–667.
- FISHER, L. (1972). A survey of the mathematical theory of multidimensional point processes. In *Stochastic Point Processes* (P. A. W. Lewis, ed.), pp. 468–513. New York: Wiley.
- GLASS, L. and TOBLER, W. R. (1971). Uniform distribution of objects in a homogeneous field: cities on a plain. *Nature*, **233** (5314), 67–68.
- GOLDMAN, J. R. (1971). Infinitely divisible point processes in R^n . *J. Math. Anal. Appl.*, **17**, 133–146.
- GRANDELL, J. (1972). Statistical inference for doubly stochastic Poisson processes. In *Stochastic Point Processes* (P. A. W. Lewis, ed.), pp. 90–121. New York: Wiley.
- GRENANDER, U. (1973). Statistical geometry; a tool for pattern analysis. *Bull. Amer. Math. Soc.*, **79**, 829–856.
- (1977). *Pattern Analysis; Lectures on Pattern Theory*, Vol. 2. Berlin: Springer-Verlag.
- HAMMERSLEY, J. M. (1972). Stochastic models for the distribution of particles in space. *Suppl. Adv. Appl. Prob.*, **47**–68.
- (1975). Ruminations on infinite Markov systems. In *Perspectives in Probability and Statistics* (J. Gani, ed.), pp. 195–200. New York: Academic Press.
- HILL, M. O. (1973). The intensity of spatial pattern in plant communities. *J. Ecol.*, **61**, 225–236.
- HOLGATE, P. (1965a). Some new tests of randomness. *J. Ecol.*, **53**, 261–266.
- (1965b). Tests of randomness based on distance methods. *Biometrika*, **52**, 345–353.
- (1965c). The distance from a random point to the nearest point of a closely packed lattice. *Biometrika*, **52**, 261–263.
- HOPE, A. C. A. (1968). A simplified Monte Carlo significance testing procedure. *J. R. Statist. Soc. B*, **30**, 582–598.
- KELLY, F. P. and RIPLEY, B. D. (1976). A note on Strauss' model for clustering. *Biometrika*, **63**, 357–360.
- KENDALL, D. G. (1974a). An introduction to stochastic geometry. In *Stochastic Geometry* (E. F. Harding and D. G. Kendall, eds.), pp. 1–9. London: Wiley.
- (1974b). Pole-seeking Brownian motion and bird navigation (with Discussion). *J. R. Statist. Soc. B*, **36**, 365–402.
- KLEIN, J. C. and SERRA, J. (1972). The texture analyser. *J. Micr.*, **95**, 349–356.
- KRICKEBERG, K. (1973). Moments of point processes. In *Stochastic Geometry* (E. F. Harding and D. G. Kendall, eds.), pp. 89–113. London: Wiley.
- (1976). The statistics of random processes. Lecture notes. Hanoi: Société Mathématique du Vietnam. (In Vietnamese.)
- LEADBETTER, M. R. (1972). On the basic results of point process theory. *Proc. 6th Berkeley Symp. Math. Statist. & Prob.*, Vol. III, pp. 449–462.
- MATÉRN, B. (1960). Spatial variation. *Meddelanden från Statens Skogsforskningsinstitut*, **49**, 5.
- MATHERON, G. (1967). *Éléments pour une Théorie des Milieux Poreux*. Paris: Masson.
- MATHERON, G. (1975). *Random Sets and Integral Geometry*. New York: Wiley.
- MATTHES, K. (1972). Infinitely divisible point processes. In *Stochastic Point Processes* (P. A. W. Lewis, ed.), pp. 384–404. New York: Wiley.

- MEAD, R. (1974). A test for spatial pattern at several scales using data from a grid of contiguous quadrats. *Biometrics*, **30**, 295–307.
- MILNE, R. K. and WESTCOTT, M. (1972). Further results on Gauss–Poisson processes. *Advances Appl. Prob.*, **4**, 151–176.
- NEWMAN, D. S. (1970). A new family of point processes which are characterized by their second-moment properties. *J. Appl. Prob.*, **7**, 338–358.
- NEYMAN, J. and SCOTT, E. L. (1972). Processes of clustering and applications. In *Stochastic Point Processes* (P. A. W. Lewis, ed.), pp. 646–681. New York: Wiley.
- PALOHEIMO, J. E. (1971). On a theory of search. *Biometrika*, **58**, 61–75.
- PAPANGELOU, F. (1974). On the Palm probabilities of processes of points and processes of lines. In *Stochastic Geometry* (E. F. Harding and D. G. Kendall, eds.), pp. 114–147. London: Wiley.
- PEBBLES, P. J. E. (1974). The nature of the distribution of galaxies. *Astron. & Astrophys.*, **32**, 197–202.
- PEBBLES, P. J. E. and GROTH, E. J. (1975). Statistical analysis of catalogs of extragalactic objects. V. Three point correlation function for the galaxy distribution in the Zwicky catalog. *Astrophys. J.*, **196**, 1–11.
- PRESTON, C. J. (1974). *Gibbs States on Countable Sets*. London: Cambridge University Press.
- (1976a). Spatial birth-and-death processes. *Bull. Inst. Int. Statist.*, **46** (in press).
- (1976b). *Random Fields*. Lecture Notes in Math. **534**. Berlin: Springer-Verlag.
- RIPLEY, B. D. (1976a). The second-order analysis of stationary point processes. *J. Appl. Prob.*, **13**, 255–266.
- (1976b). Locally finite random sets: foundations for point process theory. *Ann. Prob.*, **4**, 983–994.
- (1976c). *Stochastic Geometry and the Analysis of Spatial Pattern*. Ph.D. Thesis, Cambridge University.
- RIPLEY, B. D. and KELLY, F. P. (1977). Markov point processes. *J. London Math. Soc.* (in press).
- ROGERS, C. A. (1964). *Packing and Covering*. London: Cambridge University Press.
- RUELLE, D. (1969). *Statistical Mechanics*. New York: Benjamin.
- (1970). Superstable interaction in classical statistical mechanics. *Commun. Math. Phys.*, **18**, 127–159.
- SERRA, J. (1972). Stereology and structuring elements. *J. Micr.*, **95**, 93–103.
- SILVERMAN, B. W. (1976). Limit theorems for dissociated random variables. *Advances. Appl. Prob.*, **8**, 806–819.
- STRAUSS, D. J. (1975). A model for clustering. *Biometrika*, **62**, 467–475.
- TUBBS, C. R. (1974). *The Buzzard*. Newton Abbot, Devon: David and Charles.
- WATSON, G. (1973). Mathematical morphology. Tech. Report no. 21, Series 2, Dept. of Statistics, Princeton University.

DISCUSSION ON DR RIPLEY'S PAPER

Professor PHILIP HOLGATE (Birkbeck College, University of London): In his recent work and in the paper that he has presented this afternoon Dr Ripley has combined three of the main interests motivating this Society. He has contributed to the analytic probability theory of random point patterns. He has developed and studied models reflecting interesting physical and biological phenomena and he has considered the statistical problem of analysing data arising from them.

Fifteen years ago it was a common if not quite justified complaint among quantitative plant ecologists that statisticians ignored their problems. They were left alone to deal with a range of special techniques including nearest neighbour methods and the analysis of variance of blocks of contiguous quadrats that did not form part of everyday statistical currency and were consequently not supported by any accumulated theoretical knowledge and practical experience. As well as presenting his own substantial contribution to the subject, Dr Ripley's paper reminds us of the radical change that has taken place since then. An ecological botanist who is prepared to consider the pattern of his plants as a realization of a point process will have no difficulty in finding papers discussing the problem although he may have difficulty in understanding them. It is important that a genuine effort is made to convey findings such as those Dr Ripley has presented this afternoon to the practical scientists involved.

One problem raised by the paper is the adequacy of second-order analysis of point patterns. Such methods belong properly to the domain of processes that are linear and normal, which point patterns are not, as well as stationary. In another paper Dr Ripley (1966a) refers to psychological evidence that the human eye can only see the second-order aspects of a pattern, but that seems to me to be an added incentive to try to analyse the finer aspects by mathematical statistics.

Methods of doing this are indicated in the work cited in section 8.3. The term $z(\theta_1, \theta_2, \theta_3)$ is related to that part of the probability of finding points at angles $\theta_1, \theta_2, \theta_3$ that cannot be explained by the overall density or the pairwise interaction between infinitesimal areas at those angles, and should

therefore be as easy to explain to a practical scientist as a three factor interaction in a factorial experiment.

Consider the pattern arising from a bunch of larvae that hatch from a mass of eggs. If there are n of them they may crawl a distance ρ to the vertices of a regular n -gon. This would be reflected in atoms of probability for those particular point configurations, and if as would be the case in practice the effect were blurred by diffusion there would still be peaks in the Radon–Nikodym derivative with respect to the Poisson process. In general, variance of cluster size would probably make itself felt in the higher moments of the point pattern and its detailed distribution certainly would. The higher moments are also needed to assess statistical sampling fluctuations. Incidentally, the statistics proposed in my paper (Holgate, 1965a) are related to three point probabilities and they have proved informative on the few occasions that I have learnt of their use.

All statistical analysis should reflect as fully as possible all aspects of the probabilistic theory behind it. Further unity between the three aspects that I mentioned earlier could be obtained if it were possible to formulate models in a way that derived more directly from the general probability theory. One way of doing this would reflect the theory of n point figures developed by Miles (1970). Centres of figures could be supposed to form a Poisson process, and their size, shape and orientation to have specified distributions. Miles has shown how the last two attributes of figures can be characterized canonically. Formally this is just the Poisson cluster process, but it could be developed to reflect other aspects of pattern by using large figures. A structural condition on each cluster, for instance the requirement that no two of its points were closer than d , would not eliminate closer pairs from the process as a whole but would reduce their frequency.

Dr Ripley has pointed out that in higher dimensions there is no analogue to the characterization of a one-dimensional point distribution by the set of intervals between points. However, it is possible to associate with each point the infinite vector $(x_{i(1)}, x_{i(2)}, \dots)$ of distances to the first, second, ..., nearest neighbours. In a window containing n points this would be replaced by an $(n-1)$ -vector, and the set of vectors corresponding to the n points would comprise a matrix X . For some patterns at least, X provides a convenient and useful tool for statistical analysis. It would be of interest to ask *to what extent* the distribution of these vectors of distance measurements characterizes a random point pattern. One way of doing this would be to generate a matrix X whose rows had the appropriate distributions and then seek $n \times n$ matrices D and E such that D was a distance matrix of a two-dimensional distribution of points, and E not necessarily symmetric such that trace (EE^T) was a minimum subject to the condition that when the non-zero elements of the rows of $D+E$ were rearranged in ascending order, the result was X . Trace EE^T would measure the failure of the distance distributions to characterize the process and, if it were small, D would represent a realization of a point process which was almost so characterized. This approach would serve to establish a link between point processes and the theory of statistical map construction.

Finally, I would suggest for future research the investigation of pattern arising from repulsion between points when the points, representing say territorial birds, have differing strengths. In this case the Dirichlet cells would be replaced by other cells obtained by moving their boundaries nearer to one of the competing points than to the other.

I found the paper extremely enjoyable and stimulating and it gives me great pleasure to propose a vote of thanks to Dr Ripley.

Mr J. E. BESAG (University of Durham): It is a pleasure to congratulate Brian Ripley on his excellent paper describing models and statistical techniques which will be of interest to all who investigate spatial point patterns. Of course, in practical applications, one must not expect too much: spatial statistical analysis usually provides helpful insight into situations rather than definitive conclusions.

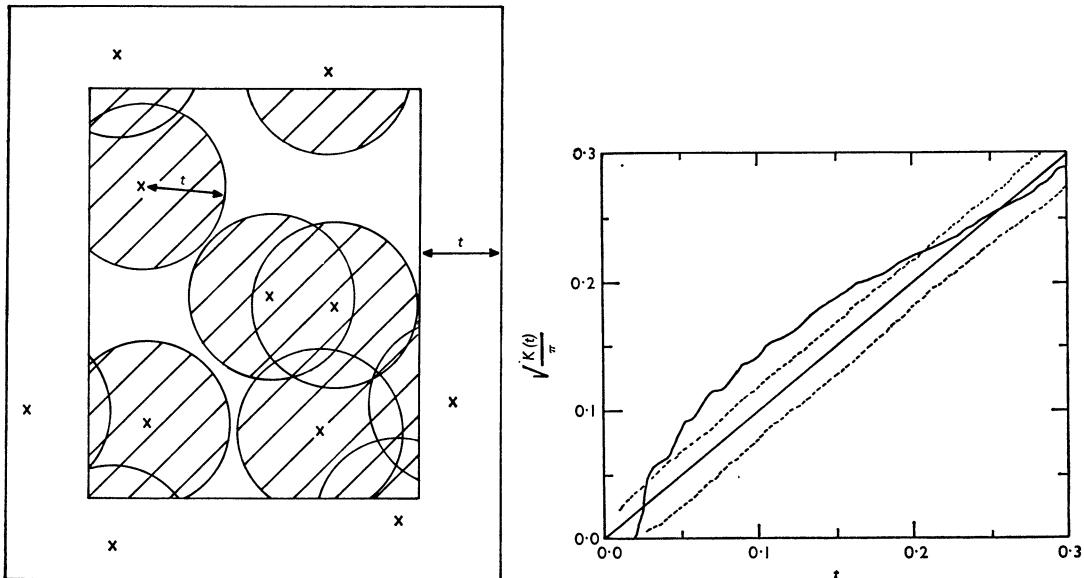
One regret is that Dr Ripley's notation, in a difficult subject, may have excluded the less mathematically inclined from his potential audience. A few diagrams in the earlier sections of the paper might have been helpful.

Though most of my detailed remarks concern K -plots, let me begin by saying how impressed I am by the results of Dr Ripley, together with Dr Kelly, in the area of Markov point processes. The very existence of such processes, and the ability to simulate them, is an important breakthrough in the modelling of point patterns.

Dr Ripley's equation for $g(t)$, in Section 3.1, enables one to calculate $K(t)$ for a reasonable variety of Poisson cluster processes. Such calculations are informative in themselves and, interestingly,

sometimes produce more transparent results in three dimensions than in two: for example, this occurs with the cluster process of Matérn (1971) and with variants thereof.

In Section 6, Dr Ripley describes unbiased estimation of $\lambda^2 K(t)$. Given sufficient data, one might prefer to achieve this by allowing a border within the sampling frame. This removes the problem in Dr Ripley's procedure that very large weights may be given to particular events near the boundary. For the unbiased estimation of $p(t)$, a diagram might again have been helpful: $\hat{p}(t)$ is merely the proportion of the inner area which is hatched in my Fig. D1.



Left. FIG. D1. Unbiased estimation of $p(t)$. $\hat{p}(t)$ is the proportion of the inner rectangle which is hatched.

Right. FIG. D2. Transformed K -plots for the redwood example. The solid curve corresponds to the data, the solid line to the theoretical Poisson and the dashed curves to the envelope of Dr Ripley's 99 Poisson simulations.

I found the analysis of Dr Crick's cell problem, in Section 7, a thoughtful and persuasive application of K - and p -plots. On the other hand, the particular subset of the redwood data, mentioned in Section 8, seems a rather odd choice (cf. Strauss, 1975, Fig. 1); also it is fairly clear that one needs rather larger data sets in order to model clustered patterns with any authority. For example, Steve Cook, at Durham University, has fitted the redwood plot within an envelope of 19 simulations of a "semi-deterministic cluster process" (Diggle, 1975), though we realize this process is strictly inappropriate. The analysis of the Spanish towns is perhaps more interesting, though I am somewhat unclear as to the purpose of statistical modelling in human geography. Dr Ripley, it should be said, is rather unfair to Glass and Tobler, who themselves dismissed their hard-core model and suggested that something softer is required. Markov point processes, as Dr Ripley demonstrates, can satisfy their requirement.

As a slight modification to K -plots, one might consider plotting \sqrt{K} against t , so that the theoretical Poisson model becomes a straight line. My Fig. D2 shows what happens to the redwood data and the associated Poisson envelope. The essentially constant width of the envelope away from the origin suggests a useful result concerning sampling fluctuations. Also, the transformed plot aids discrimination near the origin. However, at a more basic level, I am not entirely convinced that, in the *modelling* phase, K (or \sqrt{K}) is as informative as a smoothed estimate of Peebles' w -function (Section 8.3). Dr Ripley's plea for total objectivity, at the end of Section 6, suggests he might have us abandon histograms and even stem-and-leaf plots! I trust not.

At the meeting itself, I made a brief comparison between the \sqrt{K} -plot and the w -plot for a 3,000-event Monte Carlo realization of a very simple two-dimensional cluster process, the results for

which were kindly produced at very short notice by Tom Shanks, a member of the Durham University astrophysics group which is currently analysing galaxy catalogues using w -plots and other techniques. Incidentally, the number of events in the simulation was not untypical of that available in practice; unfortunately, galaxy data have complications of their own, notably the serious inhomogeneities which occur in the plates and the drastic loss of information in the two-dimensional representation of the original pattern.

In Section 8.3, Dr Ripley is unimpressed by Peebles' (1974) claim that w -plots for galaxy data may follow a power law over a substantial part of the range. Soneira and Peebles (1977) describe a model which they suggest is reasonably consistent with such behaviour. I would welcome Dr Ripley's comments.

Finally, I would like to say a little about Monte Carlo testing. Significance testing itself is not particularly popular these days, even among classical statisticians, but it surely has its place provided results are not interpreted too narrowly. Dr Ripley has described relatively sophisticated modelling applications of the Monte Carlo approach, applications in which one may gloss over possibly serious effects of parameter estimation. It seems to me that Monte Carlo tests also have their uses in much simpler spatial situations. Prompted by Dr Ripley's work, Peter Diggle and I have looked at some of these. Besides the obvious case of testing for departures from a planar Poisson process, we have applied the distribution-free Monte Carlo approach to the initial investigation of spatial interaction on a lattice, to Knox's (1964) space-time interaction test, to pattern transference, to species diversity and, most recently, to Mead's (1974) two-dimensional 4-within-16 randomization test for the detection of scales of pattern. The Monte Carlo approach has a number of advantages: it enables one to operate with the most interesting test statistic, it circumvents dubious distributional assumptions, it is manageable (especially with APL) and, not least, it is easy to explain. For example, epidemiologists are aware that Knox's Poisson approximation sometimes breaks down and I believe they are then prone to replace it by a normal approximation. In my own limited experience, in an ornithological example, the normal approximation may also be inappropriate.

Presumably, the main objection to the Monte Carlo approach is that different researchers will not obtain quite the same significance levels when using the same test statistic on the same set of data. This seems relatively unimportant at a stage when tests are regarded only as providing useful guidelines to further hypotheses and data collection, though, personally, I do prefer to compare the data with perhaps 99 or 199 simulations rather than 19. The last point can easily be implemented in the simple situations mentioned above.

I have greatly enjoyed Dr Ripley's paper and am very pleased to second the vote of thanks.

The vote of thanks was passed by acclamation.

Professor J. F. C. KINGMAN (Mathematical Institute, University of Oxford): It is a pleasure to join the congratulations to Dr Ripley for an able and comprehensive survey of an area of statistics which is only now assuming the importance which its many applications demand. Particularly welcome in this account is the variety of models, based on different postulates about underlying mechanisms. It is surely crucial that models should be chosen not only for their mathematical convenience, but because they reflect the scientist's insight into the nature of the phenomena observed.

From this point of view the discussion in Section 4 of spatio-temporal models is valuable, since any observed spatial pattern will have evolved over time. But a serious limitation, more serious than is often realized, is the restriction to models reversible in time. In the applications to statistical mechanics which have inspired some of the deepest work in this field, the restriction is natural since the laws of Newtonian mechanics are reversible. Not so those of biology or geography (so far as they exist at all), and the equilibrium distributions of non-reversible spatio-temporal processes, even in simple cases, can be strikingly different from those considered here (Kingman, 1969, Sections 5 and 6).

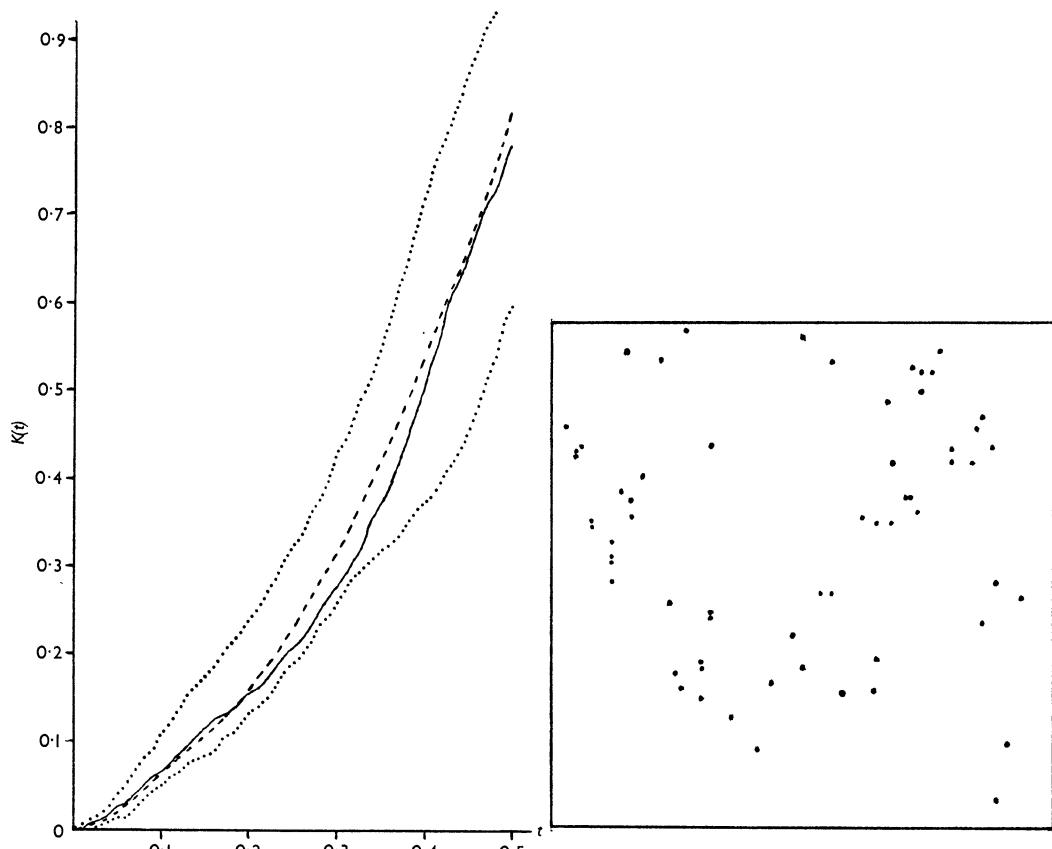
A similar argument concerns the cavalier dismissal of Cox (or doubly stochastic Poisson) processes. Certainly these are not appropriate to all applications, and in particular cannot represent many forms of clustering. But there are good theoretical reasons for supposing that some spatial patterns, such as those of reproducing and migrating populations, are likely to be Cox processes (Kingman, 1977). When this is so, it may be more sensible to make parametric assumptions about the rate process than parametric assumptions about the point process.

Finally, I would support Professor Holgate's warning against too heavy reliance on moment measures of low order. We all know that means and variances and covariances do not completely describe distributions and joint distributions, and in everyday practice we behave as though they did, and we know that the results will not usually be too misleading. This happy state does not extend to spatial point processes. For example, Kallenberg (1977) has constructed a point process in the plane, whose first and second moments coincide with those of a Poisson process, but all of whose realisations are regular lattices of points in the plane. The diagnostic value of estimators of quantities like $K(t)$ is necessarily rather limited.

Mr P. J. DIGGLE (University of Newcastle upon Tyne): My comments on tonight's extremely interesting paper relate to the function $K(t)$ and its estimator $\hat{K}(t)$. Firstly, I would agree that these provide useful summaries of model and data respectively. It is perhaps worth noting, therefore, that $K(t)$ can be derived explicitly for certain Poisson cluster processes, in particular those described respectively by Diggle (1975) and Matérn (1971) for which the daughter process consists of a Poisson-distributed number of points distributed uniformly and independently at random either on the circumference of, or more interestingly within, a circle of fixed radius a centred on the corresponding parent. In the latter case we obtain

$$K(t) = \begin{cases} \pi t^2 + (\pi \alpha a)^{-1} \{(2t^2/a - 2a) \cos^{-1}(t/2a) + \pi a - t(t^2/2a^2 + 1)\sqrt{1 - t^2/4a^2}\}, & 0 \leq t < 2a \\ \pi t^2 + \alpha^{-1}, & t \geq 2a \end{cases}$$

which is plotted, for $a = 0.07$ and $\alpha = 27$, in my Fig. D3.



Left. FIG. D3. A Matérn cluster process fitted to the redwood data. —, $\hat{K}(t)$ for data; ----, $K(t)$ for Matérn process;, envelope of $\hat{K}(t)$ from 19 simulations of Matérn process

Right. FIG. D4. A realization of the fitted Matérn cluster process.

With regard to the redwood plot, a comparison between the above $K(t)$ and $\hat{K}(t)$ for the data shows the discrepancy anticipated by Dr Ripley. However, Dr Ripley would seem to be guilty of the same failure to consider sampling fluctuations for which he proceeds to criticize Glass and Tober (1971). In fact, $\hat{K}(t)$ for the data lies entirely within the envelope generated by 19 simulations of a Matérn process with $a = 0.07$, 27 cluster centres and conditioned to produce 62 points in the unit square. Fig. D4 shows a realization of this process, and may be compared with Fig. 10 in the paper. Similar agreement is observed when a and the number of cluster centres are varied in the ranges 0.05–0.09 and 24–30 respectively, also for other Poisson cluster processes including an equally plausible model where by the dispersion of each point relative to its cluster centre has a symmetric radial Normal distribution, and the rather less plausible "semi-deterministic cluster process" mentioned earlier (Diggle, 1975). These results therefore suggest an inability to discriminate between the various models rather than a particularly satisfactory fit by the Matérn process, and my inclination is to endorse Dr Ripley's comment on the paucity of the data. Aggregated patterns are, of course, generally more awkward in this respect than are regular patterns.

Finally, I would remark that in the above investigation the *unweighted* empirical distribution function $\hat{F}(t)$ of inter-point distance tells much the same story, particularly for small t where tests based on $\hat{K}(t)$ are likely to be at their most sensitive. Also, $\hat{F}(t)$ is of course computationally much simpler. Whilst agreeing that insight into possible models is more readily available through $\hat{K}(t)$ rather than $\hat{F}(t)$, I wonder whether the former necessarily has any advantages in any formal assessment of the fit of model to data. Certainly, goodness-of-fit criteria arise more naturally for the latter.

Professor R. SIBSON (University of Bath): I should like to add my thanks and congratulations to Dr Ripley for a valuable and timely contribution to our understanding of the analysis of spatial point patterns. In his paper, he mentions the use of Dirichlet tessellations, and I intend to contribute a few footnotes on this topic.

The program developed by Dr P. J. Green and myself for computing plane Dirichlet tessellations is a FORTRAN implementation of a recursive algorithm which inserts the points one by one, growing a tile round each, and modifying neighbouring tiles as it does so. The storage requirement is linear and small—12 locations per point plus overheads. The time dependence is of the form $aN^1 + bN + \text{overheads}$ (N points) in the present version. We know how to reduce the dominant term to $N \log N$ but the relative values of the parameters a and b are such that we do not at present feel that this is worthwhile, since there is an associated storage cost. Actual run times on a CDC 7600 are about 0.78 sec for 1,000 points and about 11.6 sec for 10,000 points. A detailed account of the algorithm will appear shortly (Green and Sibson, 1977).

The speed of this program makes it unlikely that the tessellation step could constitute any kind of barrier in simulation work of the kind described by Dr Ripley. In addition to its direct use, the program can also play an indirect role by providing, for example, a way of calculating all the small distances in a configuration without calculating *all* the distances, which latter is an $O(N^2)$ computation. In many of the examples considered in Dr Ripley's paper, it is these small distances which are of primary interest.

Dr Ripley has reported his analysis of a cell pattern (Dr Crick's problem) as being carried out by replacing each cell by the tile (Dirichlet polygon) of its centroid; several speakers, particularly Dr Mollison, have questioned whether this Dirichlet tessellation actually fits the original cell pattern well enough. A first step in improving the fit would be to choose generating points more appropriate than the cell centroids, since in general the generating point of a tile is not its centroid. The generating points are, however, given by a simple geometrical construction in the case of regions which exactly form a Dirichlet tessellation. Consider a vertex of the tile whose generating point is to be found. In the non-degenerate case three tile edges will meet at that vertex, two of them part of the boundary of the chosen tile. Produce the third edge through the boundary into the tile, and then reflect it in the bisector of the angle between the other two edges. The generating point lies on the line so constructed. Since each tile has at least three vertices, this always overdetermines each generating point. In the case of an empirical cell pattern (with straight-line boundaries) the lines defined above need not all meet in a point, but it is easy to devise *ad hoc* methods for producing a good choice of point once the basic geometry is understood. It would be interesting to rework the analysis in Section 7 in the light of this.

Professor Holgate, in proposing the vote of thanks, made two points which I shall take up. The first of these again concerns Dirichlet tessellations; he enquired about the possibility of modifying the basic construction so as to yield different cell sizes and/or curved cell boundaries. Weights may be attached to the points in order to achieve this, and there are at least three plausible ways in which they may be used to alter the boundary. The main restriction on how this may be done is the need to preserve the analogue of the property that the perpendicular bisectors of the sides of a triangle meet in a common point. Changing the distance ratio on the boundary from $1 : 1$ to $w_1 : w_2$ yields circular arcs; adding a constant to the distance, the constant differing from point to point, and then taking a $1 : 1$ ratio for these sums yields hyperbolic arcs; similarly, adding constants to the squared distances yields displaced straight line segments. This last case, which should be straightforward to implement, corresponds to moving the points out of the plane by varying amounts along a third axis perpendicular to the plane, constructing the three-dimensional tessellation of these points and then looking at its section by the original plane; the implementation of this last method would be a direct one and would not require the development of a three-dimensional tessellation algorithm.

The second of the points made by Professor Holgate concerns, in general terms, the handling of errors in distance matrices. This is essentially the subject matter of multidimensional scaling (whether "classical" or "non-metric") and there is an extensive literature dealing with the techniques themselves and with the distance geometry underlying them.

Finally, I will quote an example to illustrate the range of variation in the scale of the data we may need to handle. Dr Ripley has described an example with as few as 14 observations, and most of his problems have involved 20–100. Several speakers have mentioned larger problems: the scale to which these may extend was brought home to me forcibly by a solar physicist a couple of weeks ago when he mentioned that his photographs of the surface of the sun show a pattern of about 3 million cells within the range of the photographs!

Dr A. C. ATKINSON (Imperial College, London): It is a pleasure to congratulate Dr Ripley on a paper which is both interesting and useful. Like several other speakers, I was impressed by his constructive use of simulation in the analysis of data. In my comments I want to consider the reverse process, that of the application of Dr Ripley's techniques to the analysis of simulation results, in particular to the testing of pseudo-random numbers.

It is well known that successive k -tuples of numbers from linear congruential generators lie on the points of a lattice in k dimensions. For a bad generator the lattice cells are long and thin, giving rise to widely separated planes of points and consequently to directions in which sampling is sparse. This structure is not usually apparent in small samples, although it is clear if the whole cycle of the generator can be examined.

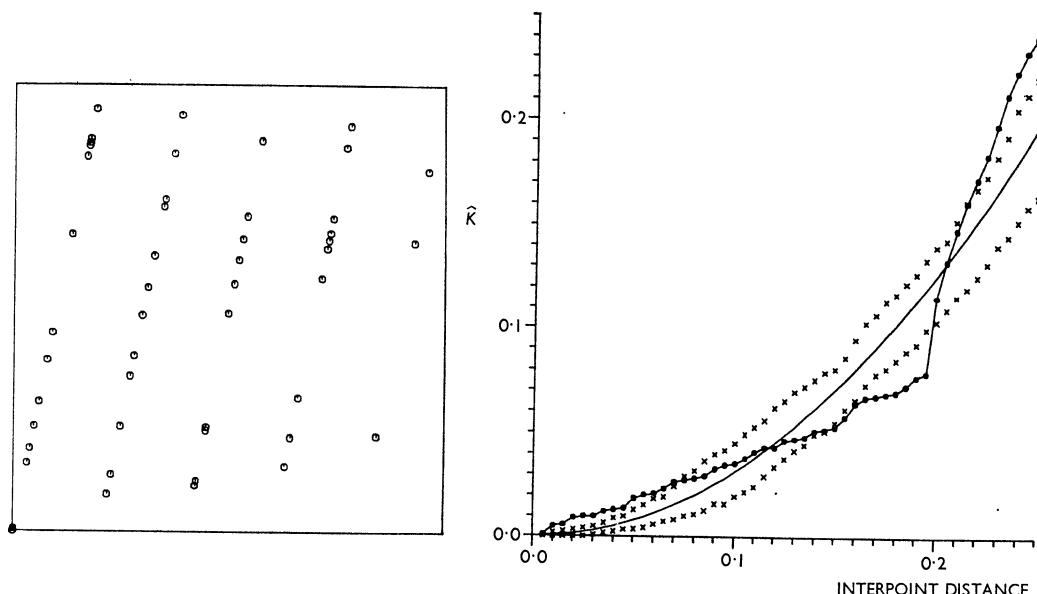
At my request Dr Ripley has used his techniques to investigate the behaviour of successive pairs of numbers produced by multiplicative generators modulo 2^{35} . I am grateful to him for the results and figures which follow. For the really dreadful generator with a multiplier of 5 a plot of 50 pairs is shown in Fig. D5. The non-Poisson nature of this distribution is picked up by \hat{K} , Fig. D6, the curve clearly showing the effect of clusters of distances between points on the same lines and on adjacent lines. For the better generator with a multiplier of 131 the plot of, in this case, 200 pairs is much better behaved. It looks like a Poisson process and \hat{K} is unable to tell the difference. In fact this is also a bad generator as was shown, for example, by Neave (1973) in an investigation of the generation of normal random variables.

To detect departures from expected behaviour Neave used samples of 1 million, so that it is hardly surprising that Dr Ripley's test fails on a sample of 200. The general point is that empirical tests, based on sampling, are rarely powerful enough to detect any but the very worst generators.

Fortunately, for linear congruential generators, there are powerful theoretical tests such as the lattice test of Marsaglia (1972) which, in effect, analyse the whole cycle without sampling. But, as far as I know, such tests are not available for other generators. It would therefore be interesting to know whether Dr Ripley's tests could be modified to detect an approximate lattice structure with some sparse directions.

A second question relates to the investigation of structures in higher dimensions. There seems to be no conceptual difficulty in extending Dr Ripley's tests in this way. What would be the effect on the computing time required?

Finally, I would like to congratulate Dr Ripley on what must surely be the first paper read to our Society to contain a reference to a work in Vietnamese.



Left. FIG. D5. Plot of 50 pairs of successive values from the generator $X_{i+1} = 5X_i \pmod{2^{35}}$.

Right. FIG. D6. \hat{K} for the data of Fig. D5.

Dr KEITH ORD (University of Warwick): I should like to join with earlier speakers in congratulating the author upon his presentation of a paper which combines both sound theory and constructive practice.

As is evident from the examples given, local inhibitory effects are often present and need to be incorporated into our models, despite the theoretical problems that arise. One difficulty with the hard core models of Section 3.2 is that while we can specify a minimum necessary distance between points, exceptions often occur, be they adjacent trees or twin cities, as in Glass and Tobler's example. A possible explanation for this is that it is the hinterland or surrounding area which is critical, rather than the distance between points. Such ideas are fundamental in geographical central place theory where towns of a given rank in the hierarchy partition the market area between them. Early expositions of this theory argued for a regular hexagonal grid although the more flexible grid of Dirichlet cells seems more appropriate.

Translating these ideas into hard core modelling we might allow a point to "live" provided that its Dirichlet cell was above a minimum size. The minimum size requirement produces a local inhibitory effect, but of a weaker kind as no distance restriction is imposed. Thus we may generate spatial birth or birth and death processes, depending on whether or not existing points can be "killed" by incursions into their territory. Theoretical properties of such a process are hard to come by although the maximum intensity is higher than that of the Matérn process for a common minimum cell size.

Others are better able to comment upon the computational feasibility of such schemes, but recent developments suggest that the computational problems are not insuperable. Extensions to non-convex cells are conceptually straight-forward but probably computationally prohibitive and lack scientific motivation.

Dr DENIS MOLLISON (Heriot-Watt University): I agree with other discussants in welcoming the use of the statistics K and p for point processes, and the Dirichlet tessellation to produce cells from such a process. It seems rather unfortunate though that Brian Ripley's principal example is not

basically a point process, but a cell process; this lays it open to several criticisms not applicable to his other interesting examples.

It might be expected that approximating Francis Crick's original cells by the Dirichlet cells of their centres of gravity would have a smoothing effect on their shapes and size distribution, both of which are relevant to the investigation described. With the help of Brian Ripley and the Green-Sibson tessellating program (not to mention the alternative technology of counting grid squares) I have examined the size distribution of the 21 cells for which Figs 1 and 2 provide a "before-and-after" comparison (see Fig. D7). For this sample, it turns out that tessellation reduces the proportional standard error s/\bar{X} from 0.20 to 0.12 (significant at the 5 per cent level in an F -test). It is

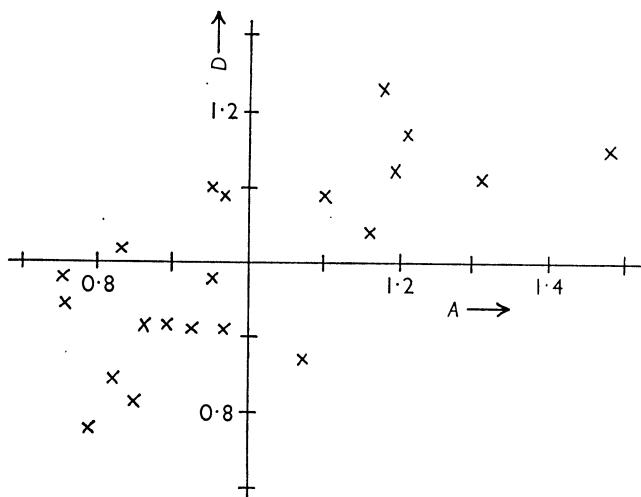


FIG. D7. Areas of Dirichlet cells (D) plotted against areas of original cells (A), for the 21 cells of Figs 1 and 2 whose Dirichlet cells do not suffer from edge effects (units such that $\bar{A} = \bar{D} = 1$).

particularly noticeable that the areas of the largest and smallest cells move sharply towards the mean. One wonders if this would have been true of the other five abnormally large cells in the full data set (see Fig. D8); this secondary mode may be of biological significance, perhaps representing undivided parent cells, in which case it would be a pity if it were lost in tessellating. It seems likely

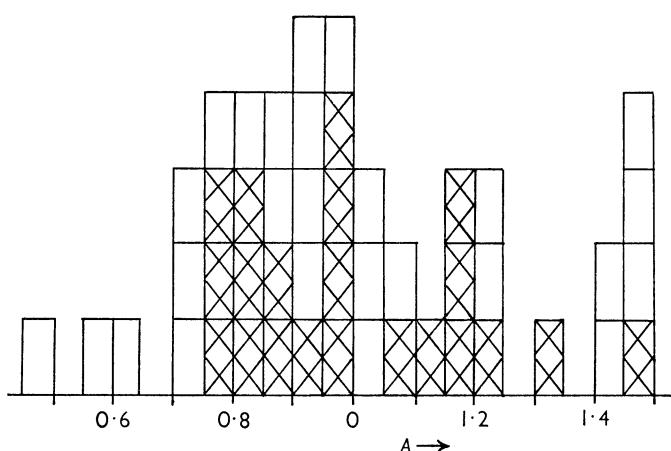


FIG. D8. Histogram of areas of the 48 original cells of Fig. 1 (the 21 cells used in Fig. D7 are shown by cross-hatching).

that extreme shapes would also be significantly smoothed by tessellating; both these effects clearly deserve fuller study than the back-of-envelope calculations described here.

I should also like to ask whether the estimate of c is sensitive to the particular method of assigning sisters, since that described in Section 7.1 seems a bit arbitrary.

Mr B. W. SILVERMAN (Statistical Laboratory, Cambridge): While the estimate \hat{K} is indeed a very important and useful diagnostic graph, the minimum interpoint distance d_{\min} is perhaps a more appealing statistic when testing the hypothesis that the data are Poisson against "hard core" alternatives. For the cell data and for both the unpublished examples given by Dr Ripley today it can be seen from the \hat{K} plots that d_{\min} is much greater than the values for the simulated data and hence that the Poisson hypothesis would be rejected by the use of this statistic; the \hat{K} plot can then be used as a diagnostic tool. The use of d_{\min} for statistical inference has several advantages. Since it is a univariate statistic rather than a function, the problem of testing to a given confidence level becomes very much easier and less controversial. If asymptotic results were obtained for the distribution of d_{\min} then a great deal of computing could be avoided since d_{\min} can be measured by eye direct from the data plots. A test based on d_{\min} does of course have less power against alternatives other than hard core models.

It is interesting to notice that Mr Besag's transformation leads to envelopes of simulations which are of approximately constant width. Suppose for simplicity that the observations are Poisson on a square of side $\sqrt{2}$ and that, to avoid edge effects, the square is made into a torus by identifying opposite edges. Let $\hat{L}(x) = \{\hat{K}(x)\}^{\frac{1}{2}}$ and $L(x) = \{K(x)\}^{\frac{1}{2}} = x$. It follows from the results of Silverman (1978) that the process $n\{\hat{L}(x) - L(x)\}$ converges weakly as n tends to infinity to a (non-Gaussian) process with mean zero and standard deviation proportional to $\sqrt{(1-x^2)}$. For values of x which are fairly small compared to the maximum possible interpoint distance this standard deviation is approximately constant. Therefore the behaviour of Mr Besag's simulations is not surprising. Results on the behaviour of weakly stationary processes should make it possible to construct approximate confidence bands for \hat{L} and hence to eliminate the need for simulation in the case of large samples from Poisson processes.

Mr JOHN KENT (University of Leeds): Dr Ripley has given us an interesting interpretation of pairwise interaction processes in terms of the equilibrium behaviour of diffusion processes. I have a few general comments to make about diffusions, in particular on the relationship between the *time reversibility* of a diffusion and its equilibrium behaviour.

For simplicity suppose the diffusion is defined on a bounded domain of R^p with generator

$$G = \sum \partial^2 / \partial x_i^2 + \sum \beta_i(x) \partial / \partial x_i,$$

and a reflecting boundary condition on the "smooth" boundary ∂D . The process has a probability transition density $p(t, x, y)$ satisfying backwards and forwards differential equations, $G_x p = \partial p / \partial t$ and $G_y^* p = \partial p / \partial t$, where G^* denotes the adjoint of G .

As $t \rightarrow \infty$, $p(t, x, y)$ converges to an equilibrium density $v(y)$ satisfying $G^* v = 0$. Furthermore, to find v it is sufficient to solve

$$G^* v(y) = \sum \partial / \partial y_i [\partial v / \partial y_i - \beta_i v] = 0 \quad (y \in D) \quad (1)$$

with a reflecting boundary condition on ∂D .

Now a diffusion in equilibrium viewed with time running backwards is still a diffusion. Thus, we may ask if the reversed process is the same as the original one. It turns out that a necessary and sufficient condition for the process to be reversible is

$$\partial / \partial y_i v(y) = \beta_i(y) v(y) \quad (i = 1, \dots, p) \quad (y \in D). \quad (2)$$

Thus to find the equilibrium density for a reversible process, it is only necessary to solve the first-order system of differential equations (2) for v , instead of the second-order equation (1). Further, the reflecting boundary condition is equivalent to the statement that (2) holds on ∂D .

In Dr Ripley's pairwise interaction process of Section 4.2, it seems necessary to suppose the interaction function $h(s)$ also satisfies $h'(0) = 0$, in order to ensure the drift is smooth everywhere. His equation in Section 4.2 is the same as (2), which he solves with periodic boundaries to get the equilibrium density given in Section 3.4.

I agree that one should be able to deal with reflecting boundaries and/or more general interaction functions, perhaps of the form

$$\begin{aligned} h(s) &= 0 \quad (s < \alpha), \\ h(s) &> 0 \quad \text{and "smooth"} \quad (s \geq \alpha), \end{aligned}$$

for some $\alpha > 0$, with the condition that the particles bounce apart (reflect) if they get too close. The solution of (2) is the same as before. However, there are some small mathematical complications in applying the theory now because the boundary ∂D contains sharp edges and corners.

Dr F. P. KELLY (University Engineering Department, Cambridge): I congratulate Dr Ripley on his development of the statistics \hat{p} and \hat{K} . These statistics (and the variant $\sqrt{(\hat{K}/\pi)}$ suggested by Mr Besag) are, I believe, very useful in summarizing certain aspects of the data in a manner which allows the eye to readily detect effects such as clustering or inhibition at a variety of scales. It may occasionally be helpful to test the significance of a function such as $\hat{K}(t)$ obtained from the data, and there is a fairly straightforward way to do this. Let $\hat{K}_1(t), \hat{K}_2(t), \dots, \hat{K}_n(t)$ be the functions obtained from n simulations of the null hypothesis model. Let j be the number of functions from the set $\mathcal{K} = \{\hat{K}(t), \hat{K}_1(t), \hat{K}_2(t), \dots, \hat{K}_n(t)\}$ which touch the envelope of the set \mathcal{K} . Then the test which rejects the null hypothesis if $\hat{K}(t)$ touches the envelope of the set \mathcal{K} has size exactly $j/(n+1)$. Of course, the size of this test is not known in advance, but then nor is the size of the test involving a further batch of simulations used by Dr Ripley. A more serious problem is that increasing the number of simulations n can be expected to decrease the size of the test, but it will also decrease the power of the test. A general method which can overcome these difficulties is to proceed as follows. Define the test statistic

$$f = F(\hat{K}; \hat{K}_1, \hat{K}_2, \dots, \hat{K}_n)$$

and let

$$f_i = F(\hat{K}_i; \hat{K}, \hat{K}_1, \hat{K}_2, \dots, \hat{K}_{i-1}, \hat{K}_{i+1}, \dots, \hat{K}_n), \quad i = 1, 2, \dots, n,$$

where F is a real valued function. List f, f_1, f_2, \dots, f_n in descending order (using a random arrangement to deal with tied values). The test which rejects the null hypothesis if f is amongst the first m in the list is a test of size $m/(n+1)$. When the test statistic $F(\hat{K}; \hat{K}_1, \hat{K}_2, \dots, \hat{K}_n)$ equals 1 or 0 depending on whether or not \hat{K} touches the envelope of the set \mathcal{K} , the method reduces to a randomized version of the test I first described; a generally more powerful test can be obtained by letting $F(\hat{K}; \hat{K}_1, \hat{K}_2, \dots, \hat{K}_n)$ equal the overall length of the intervals within which \hat{K} touches the envelope of the set \mathcal{K} . Another interesting test statistic would be a measure of the distance of \hat{K} from an average of $\hat{K}_1, \hat{K}_2, \dots, \hat{K}_n$. When the function F is symmetric in its last n arguments the random variables f, f_1, f_2, \dots, f_n are exchangeable under the null hypothesis; when F is a function which does not depend upon its last n arguments then f, f_1, f_2, \dots, f_n are independent under the null hypothesis and the method reduces to that of Barnard (1963) and Hope (1968).

In Section 8.1, Dr Ripley draws the conclusion that no cluster process can fit the redwood data. Is this conclusion based solely on the fact that in Fig. 11 the plot of $\hat{K}(t)$ crosses πt^2 at about $t = 0.25$? I ask this because, although $K(t) - \pi t^2$ is indeed increasing for a cluster process, it will be bounded above by a constant. Thus, given the proportional downward bias of $\hat{K}(t)$ suggested in Section 6 and by Fig. 3, we should expect $\hat{K}(t)$ to fall below πt^2 for large t even for a cluster process.

The equilibrium point process of Section 4.1 is robust against a certain form of departure from the assumptions of the underlying spatial birth-and-death process. This can be best illustrated when $D(x, \xi) \equiv 1$; then points have a lifetime which is exponentially distributed with unit mean. The equilibrium point process remains the same for any lifetime distribution with unit mean.

Mr Besag has suggested the statistic $\sqrt{(\hat{K}/\pi)}$, which is approximately linear for a Poisson process. Another statistic with this desirable property can be constructed directly by using $k(x, y)/d(x, y)$ rather than $k(x, y)$ in the development of Section 6. The statistic thus created has an interpretation in terms of the expected number of further points within t in a given direction from an arbitrary point of the process.

Dr F. HANSFORD-MILLER (Inner London Education Authority): I am becoming increasingly concerned at what seems to be the growing divergence between theoretical modelling and the real-life situation. May I pose the question—"Is modelling going down a blind alley?" It seems to me

that it is, with the real situation, with which it purports to be dealing, modified and almost subjugated altogether to the convenience of the model. And in this it is being supported and sustained by an ever growing use of computing time, with computers being fed by a kind of Computing Parkinson's Law.

Dr Ripley states, for example, in his report on work on redwood seedlings (Section 8.1), "This part of the plot was not square so for computational convenience I took a square mainly within the region of supposed clustering." In Section 8.2 he uses the distribution of "69 towns in a geographically homogeneous square on the Spanish plateau" and then uses his computer for "99 simulations of the Poisson process". But is even the Spanish meseta really so homogeneous over an area covered by 69 towns that any consideration of physical variations in the landscape and other relevant factors can be completely ignored? Further, in his spoken lecture, he refers to the distribution of eagles' nests in a plot of land which is portrayed with no scale of distance, no locational identification, no topographical configuration, and we are not even allowed to know its orientation. It seems to me these plotted points might just as well have been produced by some kind of random process to give an exercise in modelling. To make it a useful study of eagles' nests surely we need to know the geographical and other real life factors influencing the distribution. And similarly with the other examples.

Spatial patterns is not an easy area of study. I wrestled with it some 25 years ago well before the quantitative revolution in geography and when relevant research in the application of statistics in the field was practically non-existent. There was the paper to the Society on Crop Productivity in England (Kendall, 1939) and the warning that spatial units were modifiable units (Yule and Kendall, 1937) but not much else. There is often, also, the difficulty that in the real world situation data are sparse and not as one would like them, and especially is this the case with historical research such as that on the Geography of Religion in England in which I was engaged. Unlike many areas today, when material is plentiful, with mediaeval and sixteenth and seventeenth century data it is often a "take it or leave it" situation. I agree that to some extent spatial patterns are an end in themselves but I believe, also, that our main purpose in this area of research is to explain the distribution, and that statistics is a tool to this end.

Consider "The '2000' Ejected Ministers 1660–1662". These were the men who found they could not give "unfeigned assent and consent" to the Book of Common Prayer and rather than do so left their churches with their families. Actions of this kind have always interested me, and of course these men's actions led to the growth of Nonconformity and the Free Churches, who were once a considerable power in the land. First of all one can plot the distribution and I agree with Dr Ripley that visual inspection is important. Then one can relate the number of ejections in each county to the county area and, as the total number of benefices in each county vary and are obviously relevant to the number of ejections, relate to these similarly. In each case means and standard deviations can be calculated and, because I was particularly interested in the anomalous areas, the deviations from the mean in terms of the standard deviations. Histograms assist in identifying the counties with the greatest divergence from the mean and the geographical distribution is further clarified by showing on a county map of England the anomalies in each case, when it becomes clear that in terms of area Essex and Suffolk are the main above-average areas whilst as a proportion of benefices Lancashire, Cheshire, Staffordshire and Derbyshire form the region with the highest number of ejected ministers, with Northumberland, Durham and Cumberland forming another positive area, with London-dominated Middlesex being high in both categories. The reader is referred to Hansford-Miller (1968) for details.

This ends the easy part of the study. I now ask why there were these regional differences. What is their explanation? This is when the real hard work begins. I would suggest that even such a sophisticated statistical process as modelling is still only a tool to achieve other ends in research. It should not be seen to be, as I believe it is in danger of becoming, an end in itself. What we should be asking more often in all our research, and using modelling perhaps as one of our keys to unlock the mysteries, surely, is Why, Why, Why?

Mr A. J. GIRLING (University of Birmingham): In the case of pairwise interactions the product density $g(\cdot)$ of Section 2 of Dr Ripley's paper is λ^2 multiplied by the pair correlation function of liquid state physics. As such, numerous attempts have been made to find exact or approximate expressions for $g(\cdot)$ in various cases of interest (for a summary of recent work see Croxton, 1974).

In the case of a hard sphere liquid—the third hard-core model of Section 3.2—an excellent approximation exists (Percus and Yevick, 1957) which has the additional merit of giving an approximate expression in closed form for $g(\cdot)$ from which $K(\cdot)$ may be obtained by direct integration. Recent work in three dimensions (G. Yates and R. Jones, paper in preparation) has confirmed the effectiveness of this approximation in reproducing the detailed behaviour of the pair correlation function by comparing the results of extensive computer simulation with the approximate analytic expression. Thiele (1963) and Wertheim (1963) show that this is of the form

$$tg(t) = \lambda^2 \sum_{i=1}^3 P_i(t) \exp(\alpha_i t),$$

where the $P_i(t)$ are piecewise polynomials in t whose coefficients are functions of λ , and the α_i are functions of λ .

The discussion of the Percus-Yevick approximation in two dimensions has been less extensive, for obvious physical reasons. However, a similar analytical solution could easily be obtained and there is every reason to believe that it would perform at least as well as its three-dimensional analogue—if only because the corresponding approximation in one dimension is known to be exact.

It is to be hoped that the use of such approximate but explicit expressions will help to assess the suitability of a hard-core model and reduce the need for time-consuming simulations, which are prohibitively expensive when a large number of points is required.

Professor R. M. CORMACK (St Andrews University): Dr Ripley has made practical a valuable range of models. They seem particularly impressive in that they work well even when biological conditions on the data are either ignored by or are inconsistent with the models. In Strauss' paper the presence of redwood stumps is reported but not used. This suggests a cluster process, of which other discussants have found three to fit the data. Do these processes fit with the stumps as centres?

In the examples of nests of birds of prey discussed in the spoken version of the paper the underlying assumption of homogeneity is unlikely to hold. Spacing between nests will be imposed by the availability of suitable sites, on which the territorial behaviour underlying a hard-core model may be superimposed. Even if it is, there are difficulties in choosing the radius of the core to be the minimum observed distance between existing points—the model is likely to be overthrown in the next year by any increase in the population or merely by choice of a new nest site. It would be of interest to estimate the minimum size of core compatible with a given data set.

I hope that the author is concerned with the development and application of models to data, not to numbers (see Finney, 1974). His other examples show this concern.

Mr R. MEAD (Reading University): I found Dr Ripley's paper an interesting development in the theoretical study of spatial pattern. I am, however, concerned by a number of practical points in his paper and it is to these that I shall address my comments. I very much agree with him about the importance of considering methods of examining data in the form of a complete map of the spatial pattern for a population or a few large samples, since this is the form in which much data appears. However, I suspect the emphasis in analysing such spatial patterns should be primarily on describing the various patterns rather than testing the fits of various statistical models. Many users in ecology (see Greig-Smith's comments in Bartlett, 1971) are more concerned to use data on pattern to generate ecological hypotheses rather than to test statistical hypotheses. Should we not be concerned with testing the fit of a model only when the model has a biological (or other applied discipline) basis?

This brings me to the example in Section 7 on cells. Like Dr Mollison I am unhappy about the approximation of the cells by Dirichlet cells at their centres of gravity. This neatly turns a problem about cells into a problem about points. But will not any biological model be based on cells not points? And is the approximation good enough for the use to which it is subsequently put? My own feeling about this problem is that any model should take into account the size and shape of the observed cells, rather than of the derived Dirichlet cells.

Another practical point is whether equilibrium is a practically realistic assumption. I find it difficult to believe that social geographic models should be in equilibrium in such a rapidly changing society. And in ecology seasonal changes and environmental changes between years regularly produce irregular changes in parameters of most models fitted to data.

Finally, I want to suggest that we are in danger of allowing point process models to dominate our thinking about pattern to such an extent that we ignore the aspect of patchiness or mosaics, discussed in Pielou and unpublished work of Perry. (A nice, but probably inappropriate, analogy is with the wave and particle properties of light.) Point process models, like distance methods of investigating pattern, are concerned essentially with small scale pattern. To model larger scale pattern we need models incorporating ideas of mosaics, just as to detect such pattern we need large grids of contiguous quadrats. Incidentally, I believe Dr Ripley is wrong in asserting that counts from such grids ignore the fine structure of pattern; an extreme form of the contiguous grid of quadrats is a two-dimensional co-ordinate system, and such a system does certainly not lose the fine structure. In ecology there are aspects of pattern arising from the interaction of immediate neighbours which are clearly appropriately modelled by point processes. But there are other environmental and community aspects which operate on a larger scale and which require mosaic components of a model.

Professor A. M. WALKER (University of Sheffield): The following problem arose out of a query put to me a few years ago by a member of the staff of the Department of Human Biology and Anatomy at Sheffield. It concerned measurements of nearest-neighbour distances for a large collection of non-overlapping circles of equal size, which were obtained from microscopic examination of sections of nerve fibres. These nearest neighbour distances between circles were measured as "edge-to-edge" distances, and so could be expressed in terms of distances between the centres of the circles. Thus if a circle C had centre P , and the circle nearest to C had centre P' , then the nearest-neighbour distance for C would be $x = y - 2r$, where $y = PP'$ and r denotes the radius of the circles.

The problem, in its original form, was stated as that of finding the average nearest-neighbour distance for a large population of non-overlapping circles distributed "randomly" over a given area. The motivation underlying this was that the distribution under normal conditions was expected to be "random" and appreciable departure from "randomness" an indicator of some abnormality.

Clearly the first thing to do here was to try to specify a model corresponding to a "random" distribution. The difficulties here are well known (see, for example, Kendall and Moran, 1963), but Section 3.2 of Dr Ripley's paper provides a very helpful review of possibilities. If the radius r of the circles were small enough for their total coverage of a unit square to be small (with high probability), an acceptable approximation would be obtained by regarding them as points, generated by a Poisson process, then the obvious interpretation of randomness. This would give, by an elementary calculation, $E(X) \approx \frac{1}{2}\lambda^{-\frac{1}{2}}$, where X denotes the nearest-neighbour distance for an arbitrary point, and λ the intensity of the Poisson process. However, in the particular situation just described, it was obvious that this approximation would be quite inadequate, as the fraction of any area covered by the circles was typically of the order of 0.2.

Not being at all expert in this field, I adopted the rather simple-minded approach of defining a "random" distribution as follows. Suppose one has a unit area A , say a square, and a large number, n , of circles C_i , $i = 1, 2, \dots, n$, each of small radius r . Denote by P_i the centre of C_i . First, P_1 is given a uniform distribution over A ; then, conditional on P_1 , P_2 is given a uniform distribution over the remaining subset of A for which the condition that C_1 and C_2 do not intersect is satisfied; conditional on P_1, P_2 , P_3 is given a uniform distribution over the remaining subset defined by non-intersection of C_1, C_2, C_3 , and so on. With this "sequential" definition, I was able to obtain two approximations to the probability density of X , and correspondingly to $E(X)$, which I felt should be reasonably accurate. This opinion was confirmed by the results of some computer simulations. One of these approximations is quite simple, being given by

$$E(X) = r\{\pi/(\mu + 2\mu^2)\}^{\frac{1}{2}} e^{4(\mu+2\mu^2)} \Phi\{-2(2\mu + 4\mu^2)^{\frac{1}{2}}\}, \quad (1)$$

where $\mu = n\pi r^2$, the fraction of A covered by the circles, and $\Phi(\cdot)$ denotes the standard normal distribution function. Here μ must be fairly small, but a value of 0.2 seems not to be too large. Note that when $\mu \rightarrow 0$, $E(X) \sim \frac{1}{2}(\pi r^2/\mu)^{\frac{1}{2}}$, i.e. $\frac{1}{2}n^{-\frac{1}{2}}$, as it should.

In practice, of course, one would use $\bar{X} = (X_1 + \dots + X_n)/n$ to estimate $E(X)$. One would naturally like to have an approximation to $\text{var } \bar{X}$, but I did not manage to obtain this. One main difficulty here seems to be the correlation between the X_i .

The 'sequential' definition can be criticised because of an element of asymmetry, and one might argue that a proper definition of 'random' distribution would be given by the third 'hard-core' model described by Dr Ripley. However, I feel that for small μ , the difference between the two

definitions will be small, and that getting an approximation to $E(X)$ via the hard core model might be difficult, though I admit that I have not looked at this seriously. Also I think I am right in saying that $E(\bar{X})$ will be the same under both definitions because \bar{X} is invariant under all permutations of the suffices of the X_j .

I would be grateful if Dr Ripley, with his vastly greater knowledge of spatial processes, could provide any useful comments or advice on this particular problem.

The following contributions were received in writing, after the meeting.

Professor M. S. BARTLETT: I am sorry to miss the presentation of Dr Ripley's paper, owing to my departure for Australia the same morning. It thus reminds me of Dr Besag's paper on spatial *lattice* models presented to the Royal Statistical Society 3 years ago; this I also unfortunately missed by being in Australia.

Dr Ripley is concerned more with spatial *point* processes, and I was able to hear him describe some of his work at a recent seminar. He has in particular made use of two cumulative functions, the first being related to the covariance density (and hence to the spectral density), the other in effect to the nearest-neighbour distance from a random *point*. His exploration of significance levels for particular models by simulation is of considerable interest, in view of the intractability of the theory for some of the models, especially what I call "inhibitory" models (see, for example, Bartlett, 1976). It is perhaps worth-while underlining the danger of using nearest-neighbour distances from *individuals*,—for example, because of the repetition of small distances between points in a cluster, and the exclusion of any large distances between clusters, a difficulty avoided by Diggle, Besag and Gleaves (1976) by a directed distance, analogous to interval distance in a one-dimensional process.

Professor D. R. Cox (Imperial College). Dr Besag mentioned the possibility of linearizing some of the plots by considering \sqrt{K} versus t . One could take also K vs t^2 . The former has the supplementary advantage of exploiting the variance-stabilizing property of the square root transformation for Poisson and quasi-Poisson distributions. The second has the supplementary advantage that differences of K values are related, at least approximately, to Hoeffding U -statistics and that might provide an approach to their asymptotic distribution theory (Cox, 1965).

Prof. J. B. DOUGLAS (University of New South Wales and Imperial College, London): It is usually of interest to compare the behaviour of a new criterion with that of more familiar criteria on the same data sets—the results may be illuminating both for the new and the old criteria.

A traditional analysis of data such as the redwood seedlings, Fig. 10, is via quadrat counts: a grid of 144 square quadrats was superimposed, and the numbers of seedlings per quadrat counted visually, as well as corresponding counts for a quadrat area of four times this basic quadrat. The frequency distributions were:

Number of seedlings	0	1	2	3	4	5	6	7+
Unit quadrat	112	12	12	7	0	1	0	0
$4 \times$ unit quadrat	16	4	4	4	3	4	1	0

Visually, clustering is obvious in Fig. 10. What conclusions are suggested by estimating parameters in various discrete distributions? (There is no space to list alternative procedures; the above is a highly traditional procedure which a conservative modeller might wish to use.) The results are shown for maximum likelihood estimation in a number of models, together with "distribution free" estimates of Indices of Cluster Frequency (clusters per quadrat) and Cluster Size (individuals per cluster).

The data are not extensive: for the basic quadrat size far too high a proportion is in the zero class, for example, but all the "clustered" distributions do very much better than does the "random" Poisson. Estimates of the mean number of clusters per quadrat are about 0.3 (with a standard error of about 0.1) and of the number of individuals per cluster rather more than 1 (with a standard error of about 0.2): the negative-binomial distribution does not do very well. The evidence for clustering for this one type of analysis is, satisfactorily, strong, with the mean number of clusters per quadrat roughly increasing with the quadrat area.

Redwood seedling frequency fits

Assumptions	Unit quadrat		Goodness of fit		4 × Quadrat		Goodness of fit	
	Estimate	SE.	2 × RLL*	$\chi^2 \dagger$	Estimate	SE.	2 × RLL	χ^2
Poisson (λ)	0.4	0.1	52, 3	56, 2	1.7	0.2	29, 5	36, 4
Neyman (μ, ν)			7, 2	4, 3			6, 4	5, 5
Poisson (μ) clusters	0.3	0.1			1.2	0.3		
Poisson (ν) individuals	1.2	0.3			1.8	0.5		
Thomas (μ, λ)			6, 2	3, 3			6, 4	5, 4
Poisson (μ) clusters	0.24	0.04			0.8	0.2		
1 + Poisson (λ) individuals	0.7	0.2			1.1	0.3		
Negative-binomial (κ, ρ)			16, 2	12, 3			12, 4	8, 4
Gamma (κ) clusters	0.3	0.1			0.8	0.3		
Poisson (ρ) individuals	1.5	0.6			2.1	1		
Poisson ($\mu = \kappa \ln(1 + \rho)$) clusters	0.3	0.04			0.9	0.2		
log ($\alpha = \rho/(1 + \rho)$) indiv.	0.6	0.1			0.7	0.1		
and mean $\rho/\ln(1 + \rho)$ indiv.	1.7	0.2			1.9	0.4		
Poisson-binomial ($\mu, \nu, \tilde{\omega}$)			4, 2	1, 5			5, 4	5, 7
Poisson (μ) clusters	0.3	0.1			0.9	0.2		
Binomial ($\nu, \tilde{\omega}$) individuals	3, 0.5	0.1			5, 0.4	0.1		
Indices (see Douglas, 1975)								
of cluster frequency	0.5	0.1			1.5	0.6		
of cluster size	1.0	0.2			1.1	0.3		

* Twice relative log-likelihood, and "degrees of freedom" (Lindsey, 1974).

† Pearson χ^2 goodness of fit criterion, and degrees of freedom.

Mr J. P. RASSON (University of Namur, Belgium): I wish to consider the problem mentioned by Dr Ripley: given an unknown convex sampling window E in the plane within which N points independently and uniformly distributed have been observed, to reconstruct E by a statistical estimation.

We define

$$G = g(E) \text{ the centroid of } E, \quad S = s(E) = (E - g(E))/m(E) \text{ the shape of } E,$$

$$M = m(E) \text{ the area of } E, \quad H = H(\mathbf{x}) \text{ the convex hull of the sample } \mathbf{x}$$

and $g = g(H)$, $m = m(H)$, $s = s(H)$ similarly to G , M , S , for the convex set H .

Then (g, m, s) is a minimal sufficient statistic for the problem, and the likelihood can be factorized as follows

$$f_{G,M,S}(g, m, s) = f_{G,M,S}(g | m, s) \cdot f_{M,S}(m, s).$$

We first estimate \hat{S} to be the shape which maximizes $f_{M,S}(m, s)$. It turns out (Ripley and Rasson, 1977) that $\hat{S} = s$. Replacing S by \hat{S} , we obtain a reduced problem for which we seek estimates of G and M which are equivariant with respect to the group of translations. If we fix M , the estimate of G minimizing the mean square error is g , which is independent of M and thus optimal uniformly in the unknown M . Since M is invariant by translation, any equivariant estimate of M depends only on (m, s) . We next find a constant c such that $E(cm) = M$. This constant is provided by the relation

$$E(m) = (1 - E(V_{N+1})/(N+1))M,$$

where V_{N+1} is the number of vertices of the convex hull of a sample of $N+1$ points. $E(V_{N+1})$ depends only on S and in this reduced problem (where $S = \hat{S}$) can be calculated directly.

Our complete estimate is

$$\hat{E} = cms + g,$$

a dilatation of the convex hull about its centroid.

The problem and its solution are in fact invariant with respect to the group of affine transformations of the plane. However, we prefer not to use the full force of this, because the methods used above yield a similar solution for the Ziezold dual of the problem, in which the points distributed in E become lines which do not cut E . This dual problem is not affinely invariant.

Dr SERRA and Mr CH. LANTUEJOU (Centre de Morphologie Mathématique, Ecole Nationale Supérieure des Mines de Paris): For a stationary point process, the measure of variance density C_0 (Gel'fand and Vilenkin, 1964) satisfies the relation

$$E\{(Z(A) - \lambda |A|)(Z(B) - \lambda |B|)\} = \langle C_0, 1_A * 1_B \rangle, \quad (1)$$

where $|A|$ denotes the Lebesgue measure of A , \check{B} the transposed set of B and 1_A the characteristic function of A . Hence

$$E\{Z(A)Z(B)\} = \langle C_0 + \lambda^2, 1_A * 1_B \rangle. \quad (2)$$

Let us assume that the process is isotropic. Choosing the circle $b(x, t)$ for A , and for B the boolean dx located at point x , one relates (2) to the function $K(t)$:

$$2\pi \int_0^t r C_0(dr) - \lambda = \lambda^2(K(t) - \pi t^2) = \lambda^2 K_1(t). \quad (3)$$

Equation (3) suggests that one subtracts πt^2 from $K(t)$. More generally, it seems that the increments only need to be stationary. The parabolic term πt^2 is related to the central moments and not especially to the Poisson case.

Notice that $K_1 \equiv 0$ in a Poisson process.

The behaviour of $K_1(t)$ as $t \rightarrow +\infty$ reflects the existence of clusters. For instance, in the Neyman–Scott process

$$\lambda^2 K_1(\infty) = \alpha E\{N(N-1)\}$$

$\lambda^2 K_1(\infty)$ is the range of the process (Matheron, 1973) or, as well, the value of the power spectrum at the origin. The range may be infinite. Consider, for instance, the intersections of Poisson lines in R^2 . For this point process, we have

$$\lambda^2 K_1(t) = 8\pi\theta^3 t.$$

The aligned points appear as infinite clusters (θ density of the Poisson lines).

We are rather sceptical as to the statistical inference for models. Two different models may yield the same $K(t)$ (e.g. doubly stochastic Poisson and Neyman–Scott process).

The author replied later, in writing, as follows:

It is extremely welcome to have as many as 20 discussants to my paper. It is obviously not possible for me to reply in detail to each of them; I hope the discussion will continue elsewhere as the discussants try out their ideas and report on them. I will only take up a few points which ran through the discussion.

Certain of the points made by contributors concerned my presentation at the meeting. I mentioned the use of the minimum interpoint distance and statistics from Dirichlet tessellations and presented some examples on the distributions of birds of prey.

Modelling

“When a biologist believes there is information in an observation it is up to us to get it out” (Fisher, in Finney, 1974). Several discussants have touched on the aims of modelling. In a spatial context these are usually either “is a proposed mechanism consistent with the data” or “devise a model to give an estimate of a particular number”. More generally a model will help to give insight into the data. Of course inappropriate models may “fit” the numbers; there is nothing especially spatial about this. Remember though that applied mathematics involves the art of approximation; *an*

answer from a crude model may be better than none from an elaborate scheme. An instance is the assumption of homogeneity; in the birds of prey example the nest sites are limited, but negligibly so on the scale of the figures shown.

There needs to be a close collaboration in model building with applied co-workers. This should decide when point process models are appropriate. Despite Mr Mead's comments I believe one of the advantages of my methods over distance methods to be their ability to investigate large-scale patterns. Large grids of contiguous quadrats merely discretize the problem but their conventional analysis tied to block sizes in powers of two obscures the fine structure. As mentioned in Section 9 it would be desirable to be able to analyse more general random sets. The technology and models are not yet available.

Models

Dr Ord considered the possibility of replacing hard cores by an absolute bound on the area of the associated Dirichlet cell. Both the Matérn-type birth process and the simulation of the spatial birth-and-death process require the generation of a point uniformly over the area in which no bound would be violated. This is computationally feasible with the Green-Sibson algorithm and should not be appreciably slower than the corresponding hard-core model.

I am grateful to Mr Girling for reminding us of the work done by liquid state theorists on the models of Sections 3.2–3.4. I had rejected their approximations to $g(t)$ when I had no means of checking their validity, so it is good to know of their accuracy. I fear simulation will still be needed to assess the sampling fluctuations. A method of simulating these processes by a random walk approximation to the diffusion of Section 5 (as in Section 7) is known as Metropolis' method.

Professor Kingman advocates Cox processes on theoretical grounds. I have yet to examine a data-set which is of pure cluster type. Although such processes must exist they seem to be rarer than is often supposed. If Cox processes are uncommon then Professor Kingman's argument can be reversed to throw doubt on the postulates of his as yet unpublished paper.

Adequacy of second-order methods

Professors Holgate and Kingman have questioned the adequacy of second-order methods. I agree but would suggest that they are more adequate than the use of nearest-neighbours or the second-order analysis of a grid of quadrats. Professor Kingman implies that Kallenberg's example is special to spatial point processes. Similar examples abound; in time series to every spectral density there corresponds a process with every realization a sine wave, yet polyspectra are virtually ignored. It is possible to estimate the third moment structure but the computation must be of order N^3 , being prohibitive even for moderate N . Also, as the astrophysics example shows, the results may be difficult to interpret.

My answer to similar doubts was the statistic \hat{p} . For the figures in the paper this was computed from points on a fine grid (100 by 100 for small radii). An alternative is Monte Carlo integration. Suppose a sample of m uniformly distributed points is generated. Calculate the distances r_i and d_i from the i th point to the nearest points of the data and boundary. Then

$$\hat{p}(t) = \sum_1^m 1(r_i < t, d_i < t) / \sum_1^m 1(d_i < t)$$

is an unbiased estimator of $p(t)$. With random points far fewer evaluations seem to be needed to produce a repeatable estimate, at least for regular patterns. Usually 200 points suffices.

Dr Serra and Mr Lantuejoul remark that K does not determine the model. However, there are models which can be viewed both as Cox processes and Poisson cluster processes, a device used by Bartlett (1963). Some summary of the model is needed for a goodness-of-fit test. The psychological experiments mentioned by Professor Holgate (of which I am now sceptical) supported the choice of K .

Mr Silverman has taken up my suggestion of d_{\min} . His desire for an asymptotic distribution theory has been partially fulfilled by Saunders and Funk (1977) whose results show that as λ increases $P(d_{\min} > x/\lambda) \rightarrow \exp(-cx^2)$ (where $c = \pi/2\nu(E)$ for a reasonable window) for a Poisson process of intensity λ .

Edge-effects and transformed plots

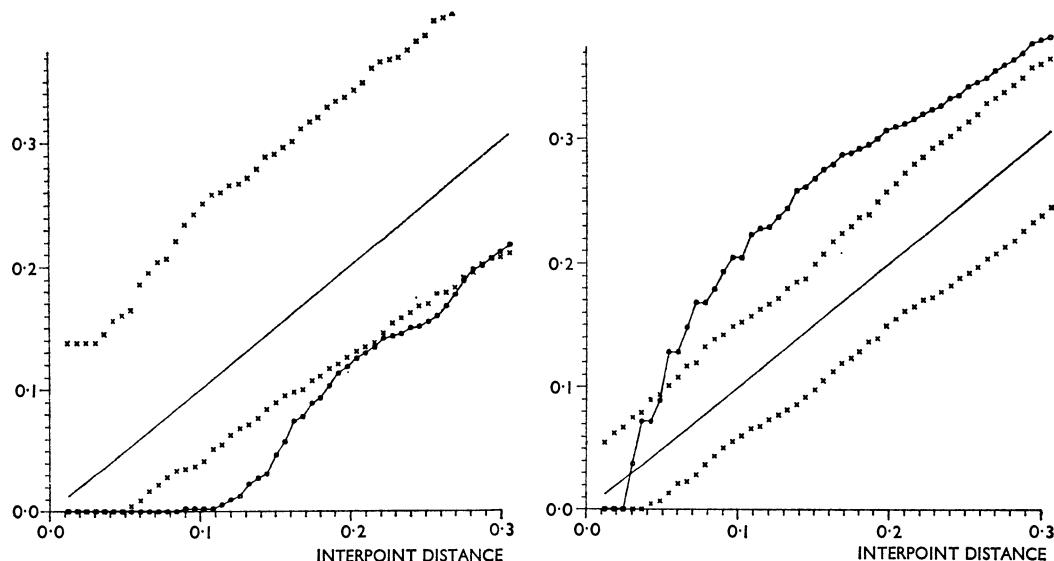
Messrs Besag and Diggle suggest that the edge-correction may at times be undesirable. Clearly for a large data-set a border will be adequate. My program in fact uses a border of $\frac{1}{2}$ per cent of the

side of the window to overcome numerical instability in the calculation of $k(x, y)$ with x very near the boundary. Note that for a square, k cannot exceed 4 in the region of unbiasedness. I agree that \hat{K} may be equally useful when fitting models. The choice may be affected by the computational penalty of \hat{K} ; my timings suggest a factor of 1.8 on the CDC 6400 here (which has slow trigonometric functions).

The use of a cumulative plot is also related to the size of the data-set. For small samples smoothing of a density plot is necessary which must obscure change-points such as the hard-core radius.

Mr Besag, Professor Cox and Dr Kelly have suggested transformations of the K -plots aimed at linearization for the Poisson process. The plot of $\hat{K}(t)$ vs t^2 compresses the small values of t , usually the range of interest. I have used Besag's plot of $\sqrt{(\hat{K}/\pi)}$ on a number of occasions. Its effectiveness as a variance-stabilizing device follows from the decomposition of the fourth moments as shown by Mr Silverman.

Frank Kelly's plot amounts to estimating $\int_0^t g(s) ds / \lambda^2$ by $\int_0^t (d\hat{K}/ds)/c(s) ds$ (where the derivative is interpreted as a generalized function). This increases the weight given to small distances and usually results in wide confidence bands. Figs 16 and 17 show Kelly plots for the cells and redwoods. They



Left. FIG. 16. Kelly plot for the cell data and 99 Poisson simulations.

Right. FIG. 17. Kelly plot for the redwoods and 99 Poisson simulations.

are comparable with Figs 5 and 11 of the paper and Besag's Fig. D2. The problems with Glass and Tobler's plot still appear in Kelly's cumulative version. I am unhappy about his interpretation; a "direction" would have to be interpreted as an infinitesimal linear strip rather than an angular sector.

Dr Serra, Mr Lantuejoul and Mr Diggle (privately) have suggested subtracting the Poisson mean curve. The disadvantage is in the interpretation and that it fails to emphasize that $K(t)$ is non-negative.

Cells

Dr Mollison refers to the areas of the cells in my Fig. 2. At the meeting I showed Table D1 of the means and standard deviations of those points in Figs 2, 4 and 7 which have complete Dirichlet cells. Although Fig. 4 was chosen to be a "typical" realization it has the smallest standard deviation of the 17 I have computed for this model. Mollison's F -test of course assumes independence; this is far from the case.

Several speakers questioned the accuracy of the transformation to a point process. In fact the data are not so precise as Fig. 1 would suggest; this was traced from a fuzzy photograph in which the cell boundaries were hard to distinguish (and only rediscovered late in the investigation).

Whether my model is biologically plausible is a matter of biological debate. It is a gross simplification but I feel it to be better than my original idea based on a hexagonal lattice and the biologists' ideas which ignored interactions and the need to fill the space. There was also an imminent deadline. I must challenge those who disagree with a point process model to produce a workable alternative.

TABLE D1

Statistics of the areas of the complete Dirichlet cells in percentages of the total window area

Figure no.	2	4	7
Mean	2.12	2.34	2.43
Standard deviation	0.29	0.38	0.24

Redwoods

The redwood data I feel to be the least satisfactory of my examples. I wanted to include a clustering example; this was the best of a bad bunch. I tried to fit several Poisson cluster processes including those mentioned by Mr Diggle. The reason that I failed was that inspection of Fig. 10 led me to try up to 10 clusters, whereas Steve Cook and Peter Diggle have used over 20.

I am very grateful to Professor Douglas for his alternative analysis. His indices give about 50 clusters which suggests that quadrat methods are detecting a different form of clustering. The matter would be resolved if we knew where were the cluster centres mentioned by Professor Cormack. As far as I know they were not recorded and so are not data!

Dr Kelly has mentioned considering $N/(N-1) \hat{K}$. This is still above πt^2 throughout the range (0.25, 0.7). My rejection of Poisson cluster processes was based on an (insufficiently) large set of simulation trials.

Clustering is a complex phenomenon and more work needs to be done. The availability of a statistic at all scales may well prove useful.

Applications

Dr Atkinson has introduced an entirely different field of application. The question of a satisfactory string of pseudo-random numbers raises both philosophical and practical questions. His contribution is salutary to those of us who use spatial simulations. Perhaps the safest approach is to use two generators with mutually prime moduli for the two co-ordinates of our points. The extension of my methods to higher dimensions involves primarily the calculation of the Euclidean distances. This and storage requirements lead me to predict a linear increase with the number of dimensions, although for small numbers of dimensions the fixed overheads are considerable.

Both Professor Kingman and Dr Atkinson mentioned processes of skewed lattices. In the latter case the directions are approximately known; this anisotropy may be the key to a more powerful test.

Professor Walker has introduced an application which is very similar to one on which I was consulted; the analysis of a cross-section of glass fibre rods on a texture analyser. His model is that termed an SSI process by Diggle *et al.* (1976); see Section 3.2. His approximation for the expected nearest-neighbour distance is new to me and very interesting. He is right in guessing that this model is barely distinguishable from my third hard-core model; even for my cell process it is impossible to distinguish realizations of \hat{K} and \hat{p} for the two models. His conjecture on $E(\bar{X})$ appears to be false; for three points on the clock face with eight points I calculated values of $\frac{8}{45}$ and $\frac{1}{6}$ respectively. Walker does not mention the aim of his investigation. For the glass fibres we were checking for anisotropy in known directions by P (rectangle of sides r, s meets data).

I wholeheartedly endorse Professor Holgate's point that we should attempt to introduce practical scientists to the body of knowledge now available. Unfortunately computer graphics packages are far from interchangeable so I cannot make a standard program available. I will, however, endeavour to help anyone who is interested in using these techniques.

REFERENCES IN THE DISCUSSION

- COX, D. R. (1965). On the estimation of the intensity function of a stationary point process. *J.R. Statist. Soc. B*, **27**, 332–337.
- CROXTON, C. (1974). *Liquid State Physics*. Cambridge University Press.
- DOUGLAS, J. B. (1975). Clustering and aggregation. *Sankhyā*, **B**, **37**, 398–417.
- FINNEY, D. J. (1974). Problems, data and inference (with Proceedings—Presidential Address). *J. R. Statist. Soc. A*, **137**, 1–18.
- GREEN, P. J. and SIBSON, R. (1977). Computing Dirichlet tessellations in the plane. *Computer J.* (in press).
- GEL'FAND, I. M. and VILENKN, N. YA. (1964). *Generalized Functions*, Vol. 4: *Applications of Harmonic Analysis* (A. Feinstein, translator). New York: Academic Press.
- HANSFORD-MILLER, F. H. (1968). A geographical and statistical analysis of religious nonconformity in England and Wales, 1550–1720. Ph. D. Thesis, University of London.
- KALLENBERG, O. (1977). A counterexample to R. Davidson's conjecture on line processes. *Math. Proc. Camb. Phil. Soc.* (in press).
- KENDALL, M. G. (1939). The geographical distribution of crop productivity in England. *J. R. Statist. Soc. A*, **102**, 21–48.
- KENDALL, M. G. and MORAN, P. A. P. (1963). *Geometrical Probability*. London: Griffin.
- KINGMAN, J. F. C. (1969). Markov population processes. *J. Appl. Prob.*, **6**, 1–18.
- (1977). Remarks on the spatial distribution of a reproducing population. *J. Appl. Prob.* (in press).
- KNOX, G. (1964). Epidemiology of childhood leukaemia in Northumberland and Durham. *Brit. J. Prev. & Soc. Med.*, **18**, 17–24.
- LINDSEY, J. K. (1974). Comparison of probability distributions. *J. R. Statist. Soc. B*, **36**, 38–47.
- MARSAGLIA, G. (1972). The structure of linear congruential sequences. In *Applications of Number Theory to Numerical Analysis* (S. K. Zaremba, ed.) pp. 249–285. New York: Academic Press.
- MATÉRN, B. (1971). Doubly stochastic Poisson processes in the plane. In *Statistical Ecology* (Patil, Pielou and Waters, eds), Vol. 1, pp. 195–213. Pennsylvania State University Press.
- MATHERON, G. (1973). *Théorie des Variables Regionalisées*. Paris: Masson.
- NEAVE, H. R. (1973). On using the Box-Muller transformation with multiplicative congruential pseudo-random number generators. *Appl. Statist.*, **22**, 92–97.
- PERCUS, J. and YEVICK, G. (1957). Analysis of classical statistical mechanics by means of collective coordinates. *Phys. Rev.*, **110**, 1–13.
- RIPLEY, B. D. and RASSON, J-P. (1977). Finding the edge of a Poisson forest. *J. Appl. Prob.* **14** (in press).
- SAUNDERS, R. and FUNK, R. M. (1977). Poisson limits for a clustering model of Strauss. *J. Appl. Prob.*, **14** (in press).
- SILVERMAN, B. W. (1978). Distances on circles, toruses and spheres. *J. Appl. Prob.*, **15** (in press).
- SONEIRA, R. M. and PEEBLES, P. J. E. (1977). Is there evidence for a spatially homogeneous population of field galaxies? *Astrophys. J.*, **211**, 1–15.
- THIELE, E. (1963). Equation of state for hard spheres. *J. Chem. Phys.*, **39**, 474–479.
- WERTHEIM, M. (1963). Exact solution of the Percus-Yevick integral equation for hard spheres. *Phys. Rev. Letters*, **10**, 321–323.
- YULE, G. U. and KENDALL, M. G. (1937). *Introduction to the Theory of Statistics*. London: Griffin.