Modeling Geo-Biosphere Processes vol. 1, p. 145–161 Cremlingen 1992

Spatial Point Processes and their Applications to Biology and Ecology

D. Pfeifer, H.-P. Bäumer & M. Albrecht

Summary

Random spatial point patterns occur in several areas in biology and ecology, e.g. in connection with the spreading of insect larvae (Richter & Söndgerath 1990, Example 3.1), the distribution of trees in woodlands (Cressie 1991, Chapter 8.2: Spatial Data Analysis of Longleaf Pines; Mecke, Schneider, Stoyan & Weil 1990, p. 136 ff.; Ripley 1981, Chapter 8.6: Examples), the distribution of bird's nests (Stoyan, Kendall & Mecke 1989, Example 5.1; Ripley 1981, p. 180 ff.), or experiments with revitalization of erosed areas in the wadden sea (Pfeifer, Bäumer & Albrecht 1992), to mention some typical examples.

Although the mathematical theory of random point patterns and pattern processes derived from such models is quite well developed and also a sufficiently application-oriented literature is available, the general importance of the field for mathematical modeling in the applied sciences has seemingly not yet been fully recognized.

Therefore, some of the most important features of stochastic point process theory and some further generalizations thereof like Boolean models and mosaics will be presented in the sequel, with particular emphasis on possible applications in biology and ecology.

1 What is a point process?

Intuitively a random point process is a spatial point pattern in which in general the location of points in 2- or 3-dimensional space as well as the total number of points are random. A first simple mathematical approach to describe such an object, say ξ, could be a representation as a random vector with a random number of components, i.e.

$$\xi = (X_1, X_2, \ldots, X_N) \tag{1}$$

where $X_1, X_2, \ldots$ are random vectors with values in $\mathbf{R}^2$ or in $\mathbf{R}^3$, resp., describing the *position* of points, and N is an integer-valued random variable, giving the

ISSN 0938-9563

0938-9563/92/5011851/US$ 2.00 + 0.25

total *number* of points in the pattern. In this setup, however, it is quite difficult to describe the distribution of points in subsets $A \subset \mathbf{R}^2$ or $A \subset \mathbf{R}^3$, resp., since we would have to choose those components $X_{i_1}, X_{i_2}, \ldots$ of ξ which have values just in the set A. But since the corresponding subscripts $i_1, i_2, \ldots$ now are themselves random variables, depending also on A, a formally well-structured description becomes almost impossible.

For this reason it has turned out that a measure-theoretic approach is much better suited, i.e. a representation of point patterns by *random measures*

$$\xi = \sum_{k=1}^{N} \varepsilon_{X_k}, \tag{2}$$

where ε_x is the *Dirac-measure* concentrated in $x \in \mathbf{R}^2$ or $x \in \mathbf{R}^3$, resp., i.e.

$$\varepsilon_x(A) = \begin{cases} 1, & x \in A \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Here the random variable

$$\xi(A) = \sum_{k=1}^{N} \varepsilon_{X_k}(A) \tag{4}$$

describes the random number of points in the set $A \subset \mathbf{R}^2$ or $A \subset \mathbf{R}^3$, resp., while again $X_1, X_2, \ldots$ are the (random) locations of individual points.

A proof of existence and well-definedness of such objects usually requires highly advanced mathematics such as *topological measure theory* (cf. Daley & Vere-Jones 1988 or Stoyan, Kenall & Mecke 1989). In particular, the set of locally finite measures on the collection of Borel sets $\mathcal{B}^2$ or $\mathcal{B}^3$ in 2- or 3-space, resp., has to be endowed with a suitable σ-field to make the mapping ξ *measureable*.

More elementary treatments can be found e.g. in Cox & Isham (1980), Mathar & Pfeifer (1990), Chapter 3.4 or Topsøe (1990).

2 Spatial Poisson processes

One of the most important basic models in point process theory is perhaps the spatial *Poisson process*. It is obtained from (1) if the random vectors $X_1, X_2, \ldots$ are independent and identically distributed with distribution Q, say, and the number N of points in the pattern is independent thereof with a Poisson-distribution $\mathrm{P}(\mu)$, $\mu > 0$. The random variables $\xi(A)$ then are again Poisson-distributed with parameter (=mean) $E[\xi(A)] = \mu Q(A)$ since for all $n \in \mathbf{Z}^+$,

$$
\begin{aligned}
P(\xi(A)=n) &= \sum_{m=n}^{\infty} P(N=m)\cdot P\left(\sum_{k=1}^{m}\varepsilon_{X_k}(A)=n\right) \qquad (5)\\
&= \sum_{m=n}^{\infty} e^{-\mu}\frac{\mu^m}{m!}\binom{m}{n}Q(A)^n((1-Q)(A))^{m-n}\\
&= e^{-\mu}\frac{(\mu Q(A))^n}{n!}\sum_{m=n}^{\infty}\frac{[\mu(1-Q(A))]^{m-n}}{(m-n)!}\\
&= e^{-\mu}\frac{(\mu Q(A))^n}{n!}e^{\mu(1-Q(A))} = e^{-\mu Q(A)}\frac{(\mu Q(A))^n}{n!}
\end{aligned}
$$

(A good impression of realisations of such point processes can perhaps be obtained by considering raindrops scattered on a walkway.)

The mapping $E\xi : A \to E[\xi(A)]$ is also called the *intensity measure* of the Poisson process; it describes the mean number of points in a (Borel) set A, generated by a Poisson point pattern. Hence if $E\xi(A)$ is large for a set A, then most probably "many" points will be generated inside A, wheras if $E\xi(A)$ is small, then only "few" points will be generated inside A. (In our example, A might be a particular square on the walkway.)

The intensity measure $E\xi$ is the characteristic quantity of a Poisson point pattern: its knowledge determines completely the distribution of ξ as a random measure and hence the distribution of all possible point patterns generated by ξ. This is mainly due to that fact that Poisson processes have *independent increments*, i.e. if the (Borel) sets A, B, $C,\ldots$ are pairwise disjoint, then the random variables $\xi(A), \xi(B), \xi(C),\ldots$ are independent. This can be seen as follows: let $I_k = \varepsilon_{X_k}(A)$, $k = 1,2,\ldots$, then $I_k = 1$ if the point X_k lies inside A. Correspondingly, $1 - I_k = \varepsilon_{X_k}(A^c)$, where A^c denotes the complement of A. Obviously, we have

$$
\xi(A) = \sum_{k=1}^{N} I_k, \qquad \xi(A^c) = \sum_{k=1}^{N}(1-I_k). \qquad (6)
$$

But for all $n, m \in \mathbf{Z}^+$,

$$
\begin{aligned}
P(\xi(A)=n, \xi(A^c)=m) &= P(\xi(A)=n, N-\xi(A)=m) \qquad (7)\\
&= P(\xi(A)=n \mid N=n+m)\cdot P(N=n+m)\\
&= P\left(\sum_{k=1}^{n+m} I_k = n\right)P(N=n+m)\\
&= \binom{n+m}{n}Q(A)^n(1-Q(A))^m e^{-\mu}\frac{\mu^{n+m}}{(n+m)!}\\
&= e^{-\mu(Q(A))}\frac{(\mu Q(A))^n}{n!}\cdot e^{-\mu(1-Q(A))}\frac{[\mu(1-Q(A))]^m}{m!}\\
&= e^{-\mu(Q(A))}\frac{(\mu Q(A))^n}{n!}\cdot e^{-\mu(Q(A^c))}\frac{[\mu(Q(A^c))]^m}{m!}
\end{aligned}
$$

such that indeed $\xi(A)$ and $\xi(A^c)$ are independent. Since the sets B, $C, \ldots$ are subsets of A^c, one sees that $\xi(A)$ and $\xi(B)$, $\xi(C), \ldots$ are independent, and so on.

If, in particular, $\mathcal{X} \subset \mathbf{R}^2$ or $\mathcal{X} \subset \mathbf{R}^3$, resp., is a closed and bounded set — typically a rectangle, a disc, a cube or a ball — and P is the *uniform distribution* over $\mathcal{X}$, then ξ is also called a *homogeneous* Poisson process over $\mathcal{X}$. I.e., $Q(A) = m(A \cap \mathcal{X})/m(\mathcal{X})$ for any Borel set $A \subset \mathbf{R}^2$ or $A \subset \mathbf{R}^3$, resp., where m denotes the *Lebesgue-measure*, which corresponds to area in $\mathbf{R}^2$ or volume in $\mathbf{R}^3$, resp. Hence the intensity measure $E\xi(A) = \mu \cdot m(A \cap \mathcal{X})/m(\mathcal{X})$ is invariant against shifts and rotations of the sets A within $\mathcal{X}$. (Point processes with the property of being rotation invariant about the origin are also called *isotropic.*) For instance, the distributional patterns created by *Arenicola marina* (beneath their excrements on the surface) or *Littorina littorea* in the wadden sea resembles very much the realisations of homogenous Poisson point processes over bounded regions of $\mathbf{R}^2$.

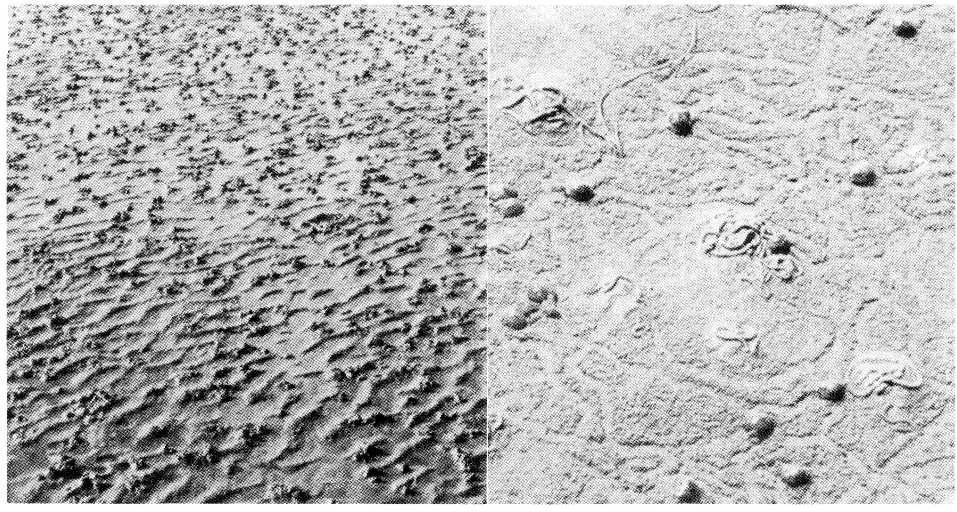

Photo 1: *Distribution of Arenicola marina and Littorina littorea.*

It is also possible to extend the construction of homogeneous Poisson processes to the whole space $\mathbf{R}^2$ or $\mathbf{R}^3$. Namely, if $\mathcal{X}_1 \subset \mathcal{X}_2 \subset \mathcal{X}_3 \subset \ldots$ are increasing closed and bounded sets with union $\mathbf{R}^2$ or$\mathbf{R}^3$, resp., then by a suitable limiting procedure, a point process (random measure) ξ can be constructed such that the restriction ξ_n of ξ to subsets of $\mathcal{X}_n$ is a homogeneous Poisson process over $\mathcal{X}_n$, for all n. For such a homogeneous Poisson process, the intensity measure $E\xi$ always is a positive multiple λ of Lebesgue-measure, i.e. $E\xi(A) = \lambda \cdot m(A)$ for all Borel sets $A \subset \mathbf{R}^2$ or $A \subset \mathbf{R}^3$, resp. Such a process will henceforth be called *homogeneous Poisson process with parameter* λ. The relation of the parameters μ_n of the restricted Poisson processes ξ_n over $\mathcal{X}_n$ and λ is here given by

$$\mu_n = E\xi_n(\mathcal{X}_n) = E\xi(\mathcal{X}_n) = \lambda \cdot m(\mathcal{X}_n) \quad \text{or} \quad \lambda = \frac{\mu_n}{m(\mathcal{X}_n)} \quad \text{for all} \quad n \in \mathbf{N}. \tag{8}$$

Note that in such models, the *total* number of points in $\mathbf{R}^2$ or $\mathbf{R}^3$ is always infinite, with probability one.

To avoid difficulties with edge effects for bounded sets $\mathcal{X}$ it is in general advantageous to deal with such more general homogeneous Poisson processes from the very start. This is actually no restriction since the former case can be reobtained in the way outlined above. Likewise, it is possible to construct Poisson point processes ξ with essentially *arbitrary* (locally finite) intensity measures $E\xi$ over $\mathbf{R}^2$ or $\mathbf{R}^3$, resp.

Due to the importance of homogensous Poisson processes in modeling "completely random" phenomena it is in general necessary to have statistical tools available for testing the hypothesis of homogeneity, and/or for the estimation of the intensity parameter λ if such a model is assumed to be valid. A classical testing procedure uses the so-called *index of dispersion* (see e.g. Richter & Söndgerath 1990, Chapter 3.1.5.). The idea here is to compare the empirical mean $\bar{\xi}_n$ and variance σ_n^2 obtained from the independent Poisson-distributed numbers of points $\xi(A_i)$, $1 \leq i \leq n$, within equally large (in area or volume) and pairwise disjoint observation windows $A_1, \ldots, A_n$, i.e.

$$\bar{\xi}_n = \frac{1}{n}\sum_{i=1}^{n} \xi(A_i), \quad \sigma_n^2 = \frac{1}{n-1}\sum_{i=1}^{n} (\xi(A_i) - \bar{\xi}_n)^2 \tag{9}$$

Since under the hypothesis of homogeneity, $\sigma_n^2/\bar{\xi}_n$ behaves like $E(\sigma_n^2)/E(\bar{\xi}_n) = 1$ almost surely for large values of n by the law of large numbers, it seems reasonable to use the index of dispersion $D_n = (n-1)\sigma_n^2/\bar{\xi}_n$ as a test statistic, which under the hypothesis of homogeneity is asymptotically χ^2-distributed (with $n-1$ degrees of freedom), for large values of n. This is due to the fact that by the independent increments property of Poisson processes, D_n behaves asymptotically like a sum of squares of independent, identically normally distributed random variables with zero mean and variance 1.

Unfortunately, by its asymptotic nature, the index of dispersion-test is only applicable for a sufficiently large number of observation windows together with a sufficiently large average number of points within each window. For instance, if the observation windows cover a set $\mathcal{X}$ but are so small that each of them contains only very few points (typically at most one), then

$$\begin{aligned} D_n &= \frac{\sum_{i=1}^{n} (\xi(A_i) - \bar{\xi}_n)^2}{\bar{\xi}_n} = \frac{\sum_{i=1}^{n} \xi(A_i)^2 - n\bar{\xi}_n^2}{\bar{\xi}_n} \approx \frac{n\bar{\xi}_n - n\bar{\xi}_n^2}{\bar{\xi}_n} \\ &= n - \sum_{k=1}^{n} \xi(A_i) = n - \xi\left(\bigcup_{k=1}^{n} A_i\right) = n - \xi(\mathcal{X}) \end{aligned} \tag{10}$$

which means that in this case, $n - D_n$ is approximately Poisson-distributed with mean $E\xi(\mathcal{X}) = \lambda \cdot m(\mathcal{X})$! Using the quantiles from a χ^2-distribution with $n-1$ degrees of freedom here would thus result in less frequent rejections of the null hypothesis for large n and hence produce erroneous results. This shows that the index

of dispersion-test in general is not very powerful since deviations from homogeneity in "small" areas of the point pattern cannot be properly detected.

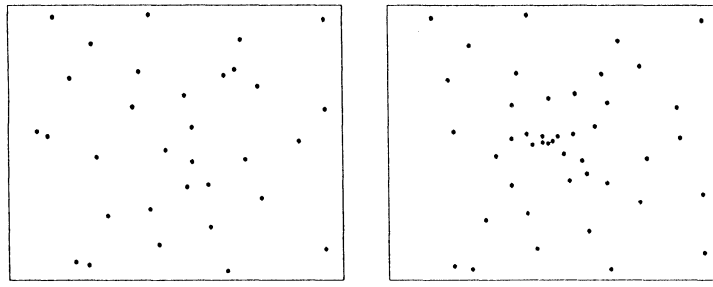

Fig. 1
Left picture: Realization of a spatially homogeneous Poisson process; Right picture: Realization of an inhomogeneous unimodal Poisson process.

Indeed it has turned out that test procedures based on distances between neighboured points in the pattern are much more powerful here (see e.g. Ripley 1981, 1988, or Stoyan, Kendall & Mecke 1989). Let $S_r(x)$ denote the disc with radius r in $\mathbf{R}^2$ or the ball with radius r in $\mathbf{R}^3$, resp., with center $x \in \mathbf{R}^2$ or $x \in \mathbf{R}^3$, resp. If x is a "typical" point of a homogeneous Poisson point process with parameter λ and Δ_x denotes the distance of the point x to its nearest neighbour in the point pattern, then the cumulative distribution function of Δ_x in terms of $r > 0$ is given by

$$\begin{aligned} P(\Delta_x \le r) &= P(\xi(S_r(x)) \neq 0) = P(\xi(S_r(0)) \neq 0) \\ &= 1 - \exp(-E\xi(S_r(0))) = \begin{cases} 1 - \exp(-\lambda\pi r^2), & x \in \mathbf{R}^2 \\ 1 - \exp(-\frac{4\lambda}{3}\pi r^3), & x \in \mathbf{R}^3 \end{cases} \end{aligned} \tag{11}$$

since the area of the disc $S_r(0)$ is πr^2 in $\mathbf{R}^2$ and the volume of the ball $S_r(0)$ is $4\pi/3r^3$ in $\mathbf{R}^3$. To get dimension-free results, it seems reasonable to represent the mean number $E\xi(S_r(x))$ of points in a typical disc or ball as

$$E\xi(S_r(x)) = \lambda v_d r^d, \qquad x \in \mathbf{R}^d, d = 2, 3 \tag{12}$$

where $v_d = m(S_1(0))$ is the volume (area) of the d-dimensional unit ball (disc), or equivalently,

$$L(r) =^d \sqrt{\frac{E\xi(S_r(x))}{\lambda v_d}} = r \tag{13}$$

(called *Ripley's L-function*). Inserting suitable statitical estimates for the quantities $E\xi(S_r(x))$ and λ here one obtains an empirical L-function, $\hat{L}(r)$, for which it is possible to develop simple confidence strips (see Ripley 1981, 1988, or Mecke,

Schneider, Stoyan & Weil 1990, Chapter 4.4 for a detailed discussion of this and related problems). Thus the hypothesis of a homogeneous Poisson process will be rejected if the empirical L-function leaves the confidence strip at least once. Ripley (1981), Capter 8.6, gives examples for point patterns of trees in New Zealand, Swedish pines, and bird's nests (eagles and peregrins).

When applying distance methods to point patterns observed only in windows $A \subset \mathbf{R}^2$ or $A \subset \mathbf{R}^3$, resp., edge effects have to be taken into consideration for those points in the pattern that are close to the boundary of A. It will therefore be necessary to use *edge corrections* in a statistical analysis based on nearest neighbour methods. Chapter 3 in Ripley (1988) is devoted solely to the discussion of such problems (see also Stoyan, Kendall & Mecke 1989, Chapter 4.6 and Cressie 1991).

The question of good estimates for the parameters of a point process is not only interesting for testing purposes alone. For instance, if an estimation of biomass is required for a species like Arenicola marina within a large area $\mathcal{X} \subseteq \mathbf{R}^2$, based on observations of a small window $B \subseteq \mathcal{X}$, say, then the expected total number of individuals, $E\xi(\mathcal{X}) = \lambda \cdot m(\mathcal{X})$, would be the typical value of interest here. But since $E\xi(B) = \lambda \cdot m(B)$, we have $\lambda = E\xi(B)/m(B)$, such that for a homogeneous Poisson process with parameter λ,

$$\hat{\lambda} = \frac{\xi(B)}{m(B)} \quad \text{with} \quad E(\hat{\lambda}) = \lambda, Var(\hat{\lambda}) = \frac{Var(\xi(B))}{m^2(B)} = \frac{\lambda}{m(B)} \tag{14}$$

(i.e. the number of points observed in B relative to its area or volume) is a reasonable estimate for λ. Indeed, by the independent increments property, it can be shown that $\hat{\lambda}$ is not only unbiased, but also sufficient — in the sense of mathematical statistics — and consistent (for large values of $m(B)$). Or, if we have again a division of a set $\mathcal{X}$ into pairwise disjoint observation windows $A_1, \ldots, A_n$ of equal size in area or volume as in the index of dispersion-test, and

$$S_n = \frac{1}{n} \sum_{k=1}^{n} J_k \quad \text{with} \quad J_k = \begin{cases} 0, & \text{if } \xi(A_k) \geq 1 \\ 1, & \text{if } \xi(A_k) = 0 \end{cases} \tag{15}$$

denotes the relative total number of *empty windows* , then

$$\hat{\hat{\lambda}} = -\frac{\log(S_n)}{c_n} \quad \text{with} \quad c_n = m(A_k) = \frac{m(\mathcal{X})}{n} \tag{16}$$

is another consistent estimate of λ since by the law of large numbers, for any window A_k,

$$S_n \approx P(\xi(A_k) = 0) = e^{-\lambda c_n} \quad \text{or} \quad \hat{\hat{\lambda}} = -\frac{\log(S_n)}{c_n} \approx \lambda \tag{17}$$

as requested; the (conditional) mean square error of $\hat{\hat{\lambda}}$ here is of the same order as the variance of $\hat{\lambda}$ when n is large, given that at least one empty window exists.

3 General point processes

In nature point patterns will seldom occur in a purely homogeneous form but frequently more clumpsy or more regular. In the first case we speak of point processes with *attraction*, in the second case of point processes with *repulsion*. Examples for such types of point patterns are: the spreading of insect larvae on leafs or the distribution of trees in young forests for attractive point patterns, and the occurence of bird's nests or the distribution of trees in older forests for repulsive point patterns. For instance, in young forests, the occurence of new trees is governed by a local spreading of seeds by existing trees, whereas elder trees prohibit the existence of close neighbours e.g. due to the extension of their crowns. Although point process theory is rich enough to allow modeling of practically all kinds of point patterns it is often justifiable to use two-stage models, due to the physical background of their emergence. Such a class of point patterns is the class of *cluster processes* which arise from some underlying point process — which is frequently homogeneous Poisson — by scattering daughter points around the parent points in a certain way. Usually the parent points themselves do not occur in the observed point pattern. *Neyman-Scott processes* are constructed right in this way; the daughter points may be distributed uniformly within discs or balls of random diameter (cf. Mecke, Schneider, Stoyan & Weil 1990, Chapter 4.5.d), or with a random number, independently and with identical distribution, before being shifted to a parent point as cluster center (cf. Stoyan, Kendall & Mecke 1989, Chapter 5.3, Ripley 1981, Chapter 8.4 or Cressie 1991, Chapter 8.5.3.). The case when the daughter points are shifted Poisson processes themselves has found particular interest in biology, for instance for modeling the spreading of larvae on leafs or in fields (see e.g. Stoyan, Kendall & Mecke 1989, p. 145, or Richter & Söndgerath 1990, Chapters 3.1.3 and 3.1.4). Here the parent points can be considered as the positions of egg masses (ovipositions) while the daughter points correspond to the positions of the larvae. Special subclasses of Neyman-Scott processes are the so-called *Thomas processes* in which the daughter points follow two- or three-dimensional normal distributions (see Richter & Söndgerath 1990, Chapter 3.1.2, Ripley 1981, Chapter 6.2, or Stoyan, Kendall & Mecke 1989, Chapter 5.3), and *Gauß-Poisson processes*, where each cluster consists of either zero, one, or two points only (cf. Stoyan, Kendall & Mecke 1989, p. 144 f.). In general, a statistical analysis of such models is less easy to perform than in the case of Poisson processes — see e.g. the discussions of these questions in Richter & Söndgerath (1990) or Cressie (1991, p. 666 ff.), where a model fitting to longleaf-pine data is performed. An example of a cluster process for the position of pines in a young forest is given in Mecke, Schneider, Stoyan & Weil (1990, p. 136).

Regular point patterns can, among others, be modelled by so-called *simple sequential inhibition processes* (SSI-Processes). For instance, in *Matern's model I*, a homogeneous Poisson process ξ with parameter λ is thinned in such a way that all pairs of points which are closer in distance than a fixed value $\delta > 0$ are deleted. The probability $p(x)$ that a point x in the original pattern is retained is hence given by

$$p = \exp(-\lambda \cdot m(S_\delta(x))) = \exp(-\lambda \cdot v_d \delta^d) \tag{18}$$

(i.e. the probability that there is no point in a disc or ball of radius δ) where again $d = 2, 3$ is the dimension of the space and v_d denotes the Lebesgue measure of the unit circle or unit ball, resp. Thus the intensity measure $E\xi_\delta$ of ξ_δ is given by

$$E\xi_\delta(A) = \lambda \cdot p \cdot m(A) = \lambda \cdot \exp(-\lambda v_d \delta^d) m(A) \tag{19}$$

for all Borel sets A ; note, however, that ξ_δ itself is *not* a Poisson process. Likewise, the probability $p(x, y)$ that points located at x and y are both retained in ξ_δ is

$$p(x,y) = \begin{cases} \exp(-\lambda \cdot m(S_\delta(x) \cup S_\delta(y))) & \text{if } \|x - y\| \geq \delta \\ 0 & \text{otherwise} \end{cases} \tag{20}$$

(see Cressie 1991, Chapter 8.5.4) where $\|x - y\|$ is the Euklidean distance between x and y. A simple BASIC-program to simulate some SSI-process is given in Richter & Söndgerath (1990, Table 3.2).

SSI-processes describe quite well the spatial distribution of species whose individuals claim certain minimal areas for their existence. By numerical differentiation of Ripley's empirical L-function, considering the so-called *pair correlation function* $L'(L/r)^{d-1}$, it is possible to determine approximately the tpyical values of δ (see e.g. Mecke, Schneider, Stoyan & Weil 1990, pp. 128 and 137, or Stoyan, Kendall & Mecke 1989, p. 120 ff).

For a survey over the many models which exist in this field we refer to Cressie (1991), Chapter 8.5.4 or Mecke, Schneider, Stoyan & Weil (1990), Chapter 4.5.

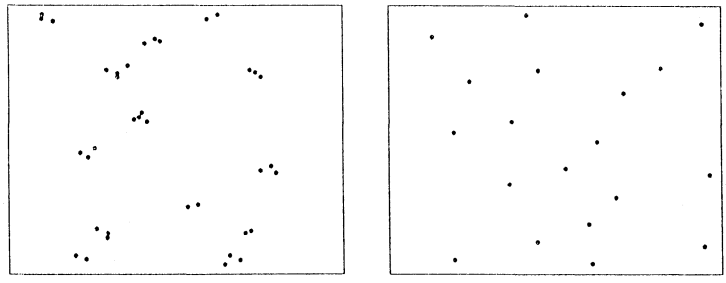

Fig. 2
Left picture: Realization of a cluster process;
Right picture: Realization of a SSI-process.

There is seemingly a basic difficulty for applied scientists how to choose an appropriate model when being faced with non-homogeneous phenomena, in particular in the fields of biology and ecology. Unfortunately, it is possible in most cases to specify different mathematical setups which produce very similar pictures in simulation studies — for instance, the left picture in fig. 2 could as well be generated by a multimodal non-homogeneous Poisson process. Mathematics alone cannot

provide the right answer here; instead, *some* knowledge about the mechanism how the point patterns under consideration emerge is necessary. Therefore, selecting statistical as well as deterministic models in biology and ecology has to be done in a very careful way; especially, if the models are used for forecasting purposes.

4 Boolean models

Besides mere (idealized) point patterns often random patchy spots can be observed in nature, for instance in the distribution of plants on the surface, or in mussel banks as of *Mytilus edulis*.

Photo 2: *Mussel bank of Mytilus edulis.*

The construction of cluster processes in the preceding section allows a statistical modeling of such phenomena in a very simple way in that instead of daughter *points* other suitable geometric objects (called *grains*) such as discs, rectangles, balls etc. are scattered around the parent points of the underlying point pattern. Objects created in this way are usually called *Boolean models* or *Poisson grain models*, if the underlying point process is homogeneous Poisson; the "random sets" Ξ which occur in this setup are, in particular, investigated in *mathematical morphology* (for a survey of this field, see e.g. Ripley 1981, Chapter 9, Ripley 1988, Chapter 6, Stoyan, Kendall & Mecke 1989, Chapters 1.4, 3 and 6, or Cressie 1991, Chapter

9). Statistical parameters which are of interest here are, among others, the area or *volume fraction* p which is given by

$$p = 1 - \exp(-\lambda E(m(\Xi_0)) \tag{21}$$

where again λ is the parameter of the Poisson process and Ξ_0 is the "typical" random set (grain) which serves as a sample for the daughter objects. In the case of discs with random radius R, we have, for instance, $E(m(\Xi_0)) = \pi \cdot E(R^2)$ as the average area covered by a grain. Similarly, the probability p_G that a typical grain of such a Boolean model is isolated (i.e., not covered by any other grain of the pattern) is given by

$$p_G = 1 - \exp(-\lambda E(m(\Xi_0 \oplus (-\Xi_0))) \tag{22}$$

where $A \oplus (-B) = \{x - y \mid x \in A, y \in B\}$ is the so-called *dilatation operation* for subsets A, B of $\mathbf{R}^2$ or $\mathbf{R}^3$, resp. (see Stoyan, Kendall & Mecke 1989 for a detailed discussion of these and related questions from morphology and stereology, or Mecke, Schneider, Stoyan & Weil 1990).

Another important quantity for Boolean models is the so-called *spherical contact distribution function* $H(r)$ which is given by

$$H(r) = 1 - \frac{P(\Xi \cap S_r(x) = \emptyset)}{1-p} = 1 - \frac{P(\Xi \cap S_r(0) = \emptyset)}{1-p}, \quad r \geq 0 \tag{23}$$

where again p is the area or volume fraction. It is essentially the distribution function of the distance from a point x chosen "randomly" outside Ξ, measured to the nearest point of Ξ. The spherical contact distribution function can be used for testing the model assumptions. For instance, in a Boolean model with discs in the plane, the logarithm of $1 - H(r)$ has the form

$$-\ln(1 - H(r)) = a \cdot r + b \cdot r^2, \quad r \geq 0 \tag{24}$$

with parameters $a \in \mathbf{R}, b \geq 0$ such that quadratic regression methods become applicable. Stoyan, Kendall & Mecke (1989) discuss such an example for lichen on a stone (p. 86 ff.).

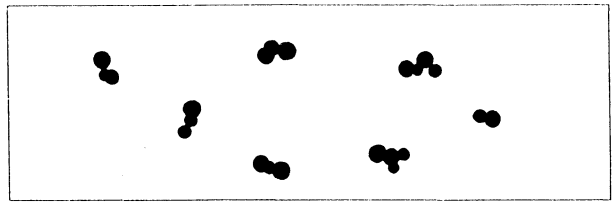

Fig. 3: *Realization of a Boolean model with discs of random radius.*

For practical applications such as the estimation of area covered by mussel banks in the wadden sea or an estimation of their total biomass, it is necessary to estimate

the area or volume fraction p from the data given through Ξ. At least three different techniques have been developed for this purpose over time. In the *point-count method*, a grid of distinct deterministic points $x_1 \ldots, x_n$ is selected from the image, and p is estimated by

$$p_C = \frac{1}{n}\sum_{k=1}^{n} \varepsilon_{x_k}(\Xi) = \frac{1}{n}\#\{k \mid x_k \in \Xi\} = \text{ fraction of grid points in } \Xi. \tag{25}$$

In the *lineal method*, which applies to planar models only an array of n parallel line segments of fixed length ℓ is laid over the image, and p is estimated by

$$p_L = \frac{L}{n\ell} \tag{26}$$

where L is the random total length of line segments intersecting Ξ. The latter method is of particular interest in marine ecology due to its technical simplicity. Both estimators p_C and p_L are unbiased for *stationary* (i.e. distributionally shift-invariant) and *isotropic* random sets Ξ as it is the case for Poisson grain models with discs of random radius as grains. In the *area method* the area fraction p is measured directly by image-analyzers. For a more detailed discussion, see Stoyan, Kendall & Mecke (1989), Chapter 6.3.

Similarly, estimations of the spherical contact distribution function are possible by point-count methods via

$$\hat{H}(r) = \frac{\#\{i \mid x_i \notin \Xi, \Xi \cap S_r(x_i) \neq \emptyset\}}{n(1-p_C)}, \quad r \geq 0 \tag{27}$$

i.e. the fraction of points outside the grid which are closer to Ξ than r.

5 Mosaics

Another possibility to create random planar or spatial geometric objects by point processes is given by *tesselations* or *mosaics*. Here $\mathbf{R}^2$ or $\mathbf{R}^3$, resp. is subdivided into polygons or polyhedra in an appropriate way. For instance, cell structures or cracks in dried surfaces can be modeled by means of mosaics; more references to applications in biology and ecology can be found in Stoyan, Kendall & Macke (1989, p.260 f.).

Planar *Poisson-Voronoi-mosaics* are obtained from homogeneous Poisson processes in the following way: each point x of the pattern is related to a polygon which consists of those points y in $\mathbf{R}^2$ that have a distance $\|y-x\|$ to x which is not larger than the distance $\|y-z\|$ for any other point z of the pattern (see Mecke, Schneider, Stoyan & Weil 1990, Chapter 3); spatial Poisson-Voronoi mosaics are constructed in a similar way.

A large amount of formulae for characteristic quantities of a "typical" cell in a mosaic is available meanwhile by stereological considerations (see Stoyan, Kendall & Mecke 1989, or Mecke, Schneider, Stoyan & Weil 1990); some of them are given

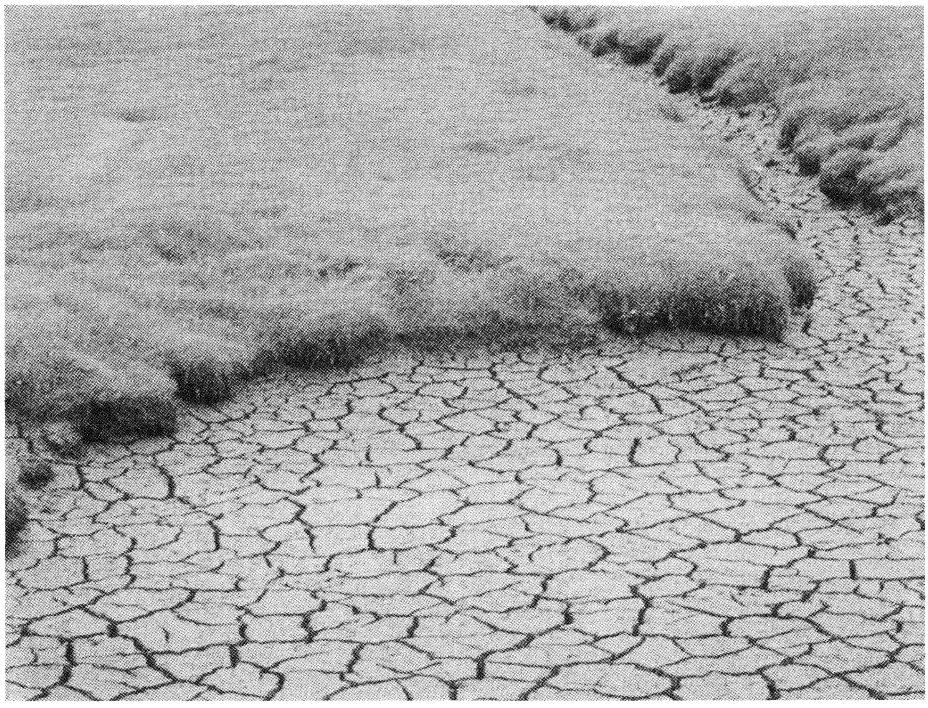

Photo 3: *Crack pattern of dried soil in the wadden sea.*

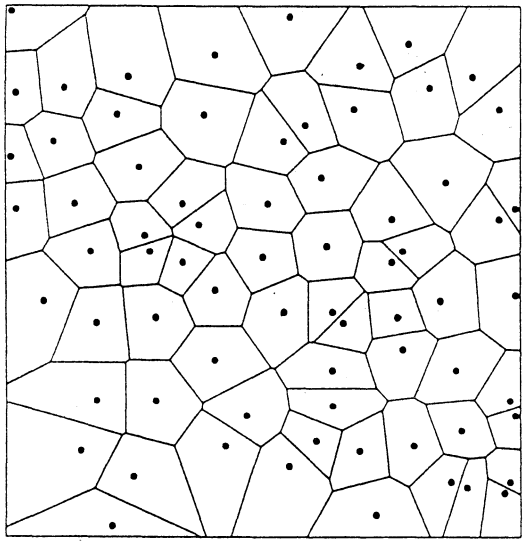

Fig. 4: *Realization of a Poisson-Voronoi-mosaic.*

below:

planar Poisson-Voronoi tesselations:

$$\begin{aligned} \text{mean area} &= \frac{1}{\lambda} \\ \text{mean circumference} &= \frac{4}{\sqrt{\lambda}} \\ \text{mean side length} &= \frac{2}{3\sqrt{\lambda}} = \frac{1}{6} \cdot \text{mean circumference} \end{aligned} \tag{28}$$

spatial Poisson-Voronoi tesselations:

$$\begin{aligned} \text{mean volume} &= \frac{1}{\lambda} \\ \text{mean surface} &= \sqrt[3]{\frac{256\pi}{3}}\Gamma\left(\frac{5}{3}\right)\lambda^{-2/3} = 5.821\lambda^{-2/3} \\ \text{mean average breadth} &= \frac{1}{15}\sqrt[3]{\frac{16}{9}\pi^5}\Gamma\left(\frac{1}{3}\right)\lambda^{-1/3} = 1.458\lambda^{-1/3} \end{aligned} \tag{29}$$

where Γ denotes the *Euler Gamma-function.*

The distribution of the number K of sides of a typical planar Poisson-Voronoi cell is given below. Note that the average value $E(K)$ is exactly 6.

n	3	4	5	6	7	8
$P(K = n)$	0.011	0.107	0.259	0.294	0.199	0.090

from Stoyan, Kendall & Mecke (1989), Table 10.4

6 Dynamic point patterns

So far only static point patterns have been considered in the preceding sections, without taking into consideration the possibility of changes in the image over time. The study of time-dependent point patterns, however, is of great importance especially in the fields of biology and ecology, in particular if forecasting of processes is desired. First approaches in theory are spatial *birth-death processes* (see Stoyan, Kendall & Mecke 1989, Chapter 5.5.5, Mecke, Schneider, Stoyan & Weil 1990, Chapter 4.5 c) or Cressie (1991, p. 678f.) which for instance have been used in modeling earthquake epicentres and sink-holes. Jetsche (1991) consideres high-dimensional Markovian birth-death processes for modeling biological populations with the additional possibility of migration, however not in the setup of point processes as such.

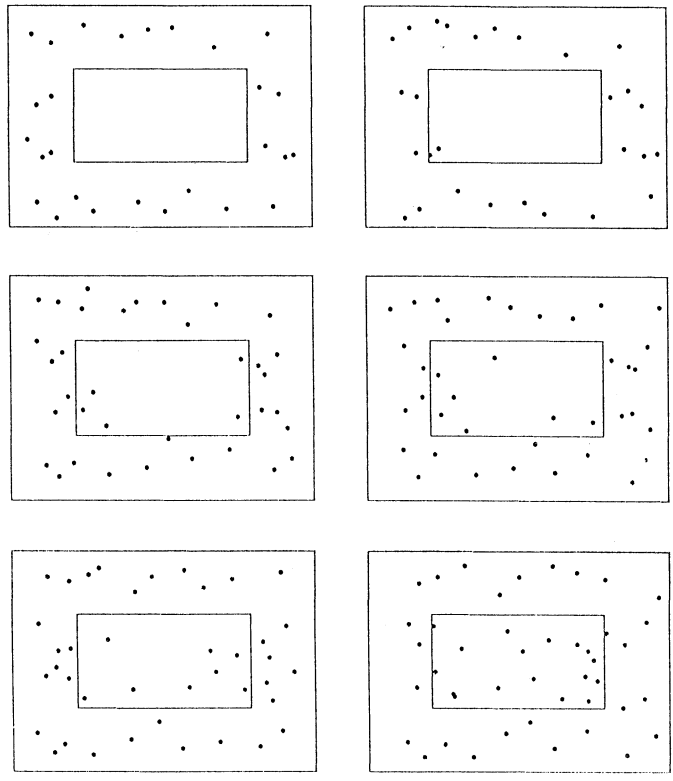

Fig. 5: *Simulation of equilibrium after erosion.*

A simple approach to a simultaneous study of birth, death and migration of points (individuals of a species) within the framework of Poisson processes has been given in Pfeifer, Bäumer & Albrecht (1992), motivated by revitalization experiments for erosed areas in the wadden sea. The approach is formally similar to (2), i.e. point processes $\xi(t)$ are considered, which now depend on the time t:

$$\xi_t = \sum_{k=1}^{N(t)} \mathbf{1}_{\{T_k > t\}} \varepsilon_{X_k(t)}, \quad t \geq 0, \tag{30}$$

where $\{N(t)\}_{t\geq 0}$ is a one-dimensional Poisson process which governes the occurence of new points, and $\{T_k\}_{k\in\mathbf{N}}$ is a family of random variables, statistically independent thereof, which correspond to the life-times of the individual particles. Here $\mathbf{1}_A$ denotes the *indicator random variable* of the event A, given formally by the relation $\mathbf{1}_A(x) = \varepsilon_x(A)$. Finally, the random variables $X_1(t), X_2(t), \ldots$ denote the position of points at time t, as before.

In this model it is possible to study the long-time behaviour of the system with very little effort since all the point patterns $\xi(t)$ are Poisson point patterns. For

instance, it is possible to simulate extinction, asymptotic equilibrium or explosion of populations, according to the choice of the model parameters. In particular, the total number $M_t = \xi_t(\mathbf{R}^2)$ or $M_t = \xi_t(\mathbf{R}^3)$, resp., of particles at time t forms a (in general non-homogeneous) birth-death process with birth-death rates given by

$$\beta_n(t) = \lambda(t)(1 - F(t)), \quad \delta_n(t) = \frac{nf(t)}{1 - F(t)}, \quad t \geq 0, n \in \mathbf{Z}^+, \tag{31}$$

where $\lambda(t) = \frac{d}{dt}E[N(t)]$ as usual denotes the intensity of the Poisson process $\{N(t)\}$ and $f(t) = \frac{d}{dt}F(t)$, $t \geq 0$, denotes the density of the life-time distribution (see Pfeifer, Bäumer & Albrecht 1992, section 3).

Acknowledgements

A part of the models presented in this paper has been used and generalized in the research project *Ökosystemforschung Niedersächsisches Wattenmeer (Pilotphase), Beiträge der Angewandten Statistik zur Bearbeitung von Maßstabsfragen und zur Versuchsplanung für die Analyse räumlicher Strukturen und dynamischer Vorgänge im Watt*, granted by the Bundesministerium für Forschung und Technologie, Förderkennzeichen FZ 03F0023F. The photos 1 and 2 are reprints from Helbing & Wonneberger (1990), with kind permission of the Nationalparkverwaltung "Niedersächsisches Wattenmeer". The photo 3 was taken by M. Herlyn, NLWA-Forschungsstelle Küste.

References

COX, D.R. & ISHAM, V. (1980): Point Processes. Capman and Hall, London.

CRESSIE, N. (1991): Statistics for Spatial Data. Wiley, N.Y.

DALEY, D.J. & VERE-JONES, D. (1988): An Introduction to the Theory of Point Processes. Springer Series in Statistics. Springer, N.Y.

HELBING, C. & WONNEBERGER, K. (1990): Nationalpark Niedersächsisches Wattenmeer. Editor: Nationalparkverwaltung "Niedersächsisches Wattenmeer", Wilhelmshaven.

JETSCHKE, G. (1991): Stochastic modeling of reaction-migration systems. In: Analyse dynamischer Systeme in Medizin, Biologie und Ökologie. 4. Ebernburger Gespräch, Bad Münster, April 1990. Informatik Fachberichte No. 275, 247–253. Springr, N.Y.

MATHAR, R. & PFEIFER, D. (1990): Stochastik for Informatiker. Leitfäden und Monographien der Informatik. Teubner, Stuttgart.

MECKE, J., SCHNEIDER, R.G., STOYAN, D. & WEIL, R.R. (1990): Stochastische Geometrie. DMV-Seminar, Vol. 16. Birkhäuser, Basel.

PFEIFER, D., BÄUMER, H.-P. & ALBRECHT, M. (1992): Moving point patterns — the Poisson case. Paper presented at: Jahrestagung der Gesellschaft für Klassifikation, Universität Dortmund, April 1992. To appear.

RICHTER, O. & SÖNDGERATH, D. (1990): Parameter Estimation in Ecology. The Link between Data and Models. VCH Verlagsgesellschaft, Weinheim.

RIPLEY, B.D. (1981): Spatial Statistics. Wiley, N.Y.

RIPLEY, B.D. (1988): Statistical Inference for Spatial Processes. Cambridge Univ. Press, Cambridge.

STOYAN, D., KENDALL, W.S. & MECKE, J. (1989): Stochastic Geometry and Its Applications. Wiley, N.Y.

TOPSØE, F. (1990): Spontane Phänomene. Stochastische Modelle und ihre Anwendungen. Vieweg, Braunschweig u. Wiesbaden.

Addresses of authors:
Prof. Dr. Dietmar Pfeifer
Dr. Mathias Albrecht
Carl von Ossietzky Universität Oldenburg
Fachbereich Mathematik
Postfach 2503
W-2900 Oldenburg, F.R.G.
Dr. Hans-Peter Bäumer
Carl von Osietzky Universität Oldenburg
Rechenzentrum
Postfach 2503
W-2900 Oldenburg, F.R.G.