

1987

## Random set theory and problems of modeling

Noel A. Cressie  
*University of Wollongong*, [ncressie@uow.edu.au](mailto:ncressie@uow.edu.au)

G M. Laslett

Follow this and additional works at: <https://ro.uow.edu.au/infopapers>



Part of the [Physical Sciences and Mathematics Commons](#)

---

### Recommended Citation

Cressie, Noel A. and Laslett, G M.: Random set theory and problems of modeling 1987, 557-574.  
<https://ro.uow.edu.au/infopapers/2427>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: [research-pubs@uow.edu.au](mailto:research-pubs@uow.edu.au)

---

## Random set theory and problems of modeling

### Abstract

The three- or four-dimensional world in which we live is full of objects to be measured and summarized. Very often a parsimonious finite collection of measurements is enough for scientific investigation into an object's genesis and evolution. There is a growing need, however, to describe and model objects through their form as well as their size. The purpose of this article is to show the potentials and limitations of a probabilistic and statistical approach. Collections of objects (the data) are assimilated to a random set (the model), whose parameters provide description and/or explanation.

### Keywords

set, theory, random, modeling, problems

### Disciplines

Physical Sciences and Mathematics

### Publication Details

Cressie, N. A. & Laslett, G. M. (1987). Random set theory and problems of modeling. *SIAM Review*, 29 (4), 557-574.

## RANDOM SET THEORY AND PROBLEMS OF MODELING\*

NOEL CRESSIE† AND G. M. LASLETT‡

**Abstract.** The three- or four-dimensional world in which we live is full of objects to be measured and summarized. Very often a parsimonious finite collection of measurements is enough for scientific investigation into an object's genesis and evolution. There is a growing need, however, to describe and model objects through their form as well as their size. The purpose of this article is to show the potentials and limitations of a probabilistic and statistical approach. Collections of objects (the data) are assimilated to a random set (the model), whose parameters provide description and/or explanation.

**Key words.** Boolean model, hitting function, random set limit theory, random fields

**AMS(MOS) subject classifications.** Primary 62M99; secondary 60D05

**1. Introduction.** In any scientific investigation, the relationship between data and theoretical models is very important. Which comes first is not always clear, since the collection, storage and retrieval of large data files already relies on a more or less vague underlying theory. In this article, the data will be recordings of objects, often images on a photographic plate or a TV screen. Such visual images may be analyzed per se, or they may be converted into numerical data by defining pixels, and recording for each pixel a gray level, or (color) frequency and intensity. A further conversion to a two-phase image might be made by recording "black" or "1" if the gray level of a pixel is above a certain threshold, or of a certain color. We will see how random-set models can be used to explain or describe them.

Theoretical models are not expected to represent the data exactly, but at the very least they act as a sorting device that directs the data analyst to efficient ways of extracting information. When a model has a component of randomness in it, there is an extra, although exploitable, source of inexactness. By definition, two realizations of the same random phenomenon will *not* be exactly the same. However, parameters estimated from two such realizations should be stable; the larger the realizations, the closer the two parameter estimates should become. Figures 1(a) and 1(b) show a small part of two artificially generated realizations of the two-dimensional Boolean model (with random parallelograms as the primary sets) discussed in §4. The parameters used for each generation were identical, and summarized by  $\lambda = 0.02$ ,  $E(P(S)) = 20.00$ ,  $E(|S|) = 15.92$  (see §4.4 for details). Generalized-least-squares-estimation techniques, described in §4, yield for Fig. 1(a):

$$\hat{\lambda} = 0.020, \quad \hat{E}(P(S)) = 16.27, \quad \hat{E}(|S|) = 15.47,$$

and for Fig. 1(b):

$$\hat{\lambda} = 0.017, \quad \hat{E}(P(S)) = 18.26, \quad \hat{E}(|S|) = 18.18.$$

The statistical and probabilistic techniques which ensure the existence of stochastic models, and efficient estimation of their parameters, are very well developed

---

\* Received by the editors July 28, 1983; accepted for publication (in revised form) November 23, 1986.

† Department of Statistics, Iowa State University, Ames, Iowa 50011. The work of this author was supported in part by National Science Foundation grant DMS-8503693, and was completed while he was a visitor at the Division of Mathematics and Statistics of the Commonwealth Scientific and Industrial Research Organization, Clayton, Australia.

‡ Division of Mathematics and Statistics, Commonwealth Scientific and Industrial Research Organization, Clayton VIC 3168, Australia.

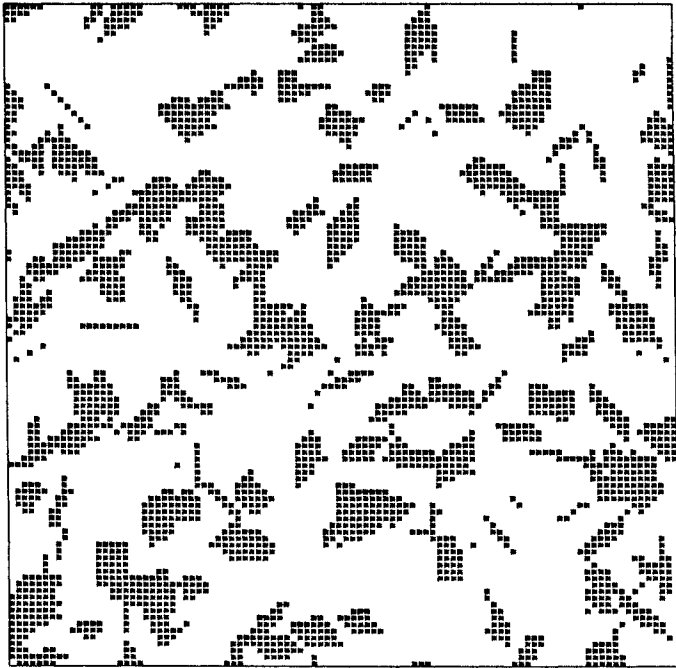


FIG. 1(a). A realization of a Boolean model whose primary sets are random parallelograms.

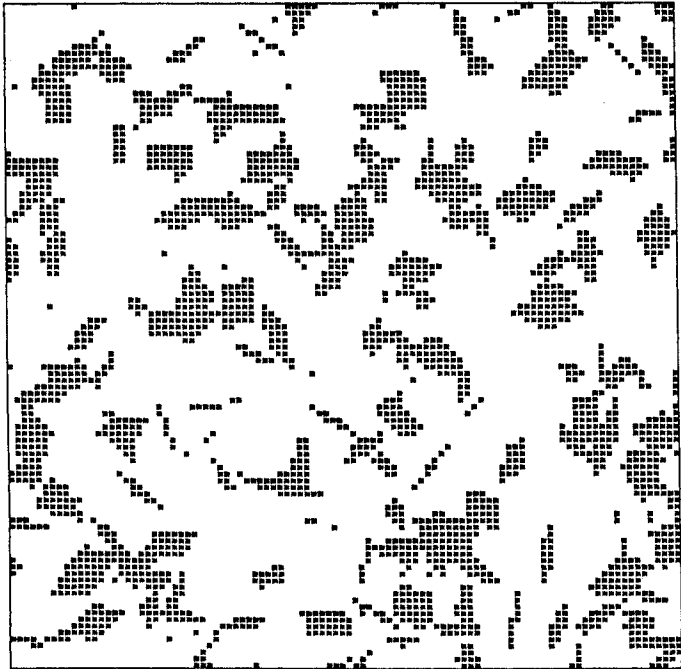


FIG. 1(b). Another realization of the same Boolean model that was used to generate Fig. 1(a).

for data that are modeled as independent and identically distributed *random variables*. Here the interrelationships between any two subcollections are extremely simple, and one only needs to determine the law of any individual to determine the law of the whole. In fact only the probability of the events “ $\{X \ni x\}$ , for all  $x \in (-\infty, \infty)$ ,” is needed. These techniques have been extended for the following:

- (a) A sample whose variables are vectors, or are elements of a Banach space, etc.
- (b) A collection of (often real-valued but also vector-valued) random variables whose dependence structure is Markov, or (strong or weak) stationary, etc.

It is our aim in this paper to present random-set models, to analyze objects in  $d$ -dimensional Euclidean space  $\mathbf{R}^d$ ,  $d \geq 1$ . Section 2 compares random sets with random functions and specifies where their approaches diverge. Section 3 discusses the role of the hitting function  $\Pr \{X \cap B \neq \emptyset\}$ ,  $B$  a test set (or “trap,” according to Kendall (1974)), in characterizing the random closed set  $X$ . Its utility in building random-set models is severely limited; we shall indicate by example why the hitting function has really failed to be a useful analogue to the cumulative distribution function for random variables. Basically, the construction of any but the simplest set models becomes extremely difficult because the hitting function is usually intractable. Even random-set limit theory has avoided its use. Section 4 discusses the problem of estimating model parameters from data, in the case of the Boolean model; even here, little is known about optimal estimators and their statistical properties. The parallel-ogram data of Fig. 1(a) are analyzed. Section 5 brings together the various themes of the paper.

**2. Random sets and random functions.** In one sense, a random set is just a special case of a random function (or random field) that takes only the values 0 or 1. In fact, if any random function (Adler (1981)) is “sliced” at, say, a level  $u$  and looked at from above, then the boundaries of the slice trace out the boundary of a random set. Any analysis of the original random function should be equally possible on these “level sets” indexed by  $u$ , and conversely. Both random sets and random functions have the concept of covariation; however, beyond this there seems little in common in their analysis (see below and Adler (1981, p. 71)). The main reason is that (random-function) operations such as convolution and Fourier filtering are linear, whereas the (random-set) morphological transformations (i.e., transformations that affect shape, (Serra (1982, Part 1))) are highly nonlinear.

In random-function theory, the variogram of a random function  $\{Z(x): x \in \mathbf{R}^d\}$  is defined as  $\text{var}(Z(x) - Z(x + h))$ , which is usually considered to be a function only of the vector  $h$  (known as intrinsic stationarity (see Matheron (1963) and Cressie (1986)). Now suppose that  $X$  is a stationary random closed set (defined in §3), and let  $Z_X(x)$  denote its indicator function. Then it is not difficult to prove that  $Z_X$  is a random function satisfying intrinsic stationarity, with variogram  $2(p - \Pr \{x \in X, x + h \in X\})$ , where  $p = \Pr \{x \in X\}$ . If we write  $X_{-h}$  for the (random) set  $X$  translated to the point  $-h$ , then the variogram is  $2(p - \Pr \{x \in X \cap X_{-h}\})$ . Clearly  $K(h) \equiv \int \Pr \{x \in X \cap X_{-h}\} dx$  is the probabilistic analogue of the *geometrical covariance* of a (deterministic) set  $A \subset \mathbf{R}^d$ , defined as  $\text{meas}(A \cap A_{-h})$ . But  $K(h)$  contains information about the surface measure of  $X$  as  $h \rightarrow 0$ . Provided  $X$  is almost surely (a.s.) regular (see Serra (1982, p. 274)) then  $\{K(h) - K(0)\}/|h|$  exists; call it  $K'_\alpha(0)$ , where  $\alpha \equiv h/|h|$ ; in  $\mathbf{R}^2$ ,

$$-\frac{1}{2} \int_0^{2\pi} K'_\alpha(0) d\alpha = E\{\text{perimeter of } X\}.$$

Now knowledge gained from the behavior at the origin of the covariance of a random function usually relates to the behavior of its spectrum at very high frequencies. This is exactly what is happening in the above relation, where the random function is the indicator function of a random set; the left-hand side (lhs) can be interpreted simply as covariance behavior near the origin, and the right-hand side (rhs) pertains to the boundary of  $X$  where there is “high frequency,” i.e., where the 0-1 random function undergoes its most drastic change. But it is at this stage of an image analysis that a random-function approach fails to capture the full geometric complexities of the image. This is clear when we write  $X \cap X_{-h}$  in terms of the erosion operation  $X \theta \check{B} \equiv \bigcap_{a \in B} X_{-a}$ . The choice of “structuring element”  $B = \{0, h\}$ , which yields  $X \cap X_{-h}$ , is just one of many that could be made in order to structurally sort the random set  $X$ . The scope of geometric possibilities expands enormously through varying the structuring element  $B$ , and (linear) spectral analysis in random-function theory is just one of these.

**3. The hitting function.**

**3.1. Preliminaries.** A summary of the main definitions and results of random-set theory (Matheron (1975)) will be needed. Let  $E$  be a locally compact, Hausdorff and separable space, and define  $\mathcal{F}$  to be the set of all closed subsets of  $E$  (including the empty set  $\emptyset$ ). Let  $\mathcal{K}$  denote the set of all compact sets, and  $\mathcal{K}' \equiv \mathcal{K} \setminus \emptyset$  the set of all nonempty compact sets. For any set of sets  $\mathcal{R}$ ,  $C(\mathcal{R})$  denotes that subset whose sets are convex. For any  $A \subset E$ , define

$$\mathcal{F}_A = \{F \in \mathcal{F} : F \cap A \neq \emptyset\}, \quad \mathcal{F}^A = \{F \in \mathcal{F} : F \cap A = \emptyset\}.$$

For  $K$  compact and  $G_1, \dots, G_n$  open, generate sets of the form  $\mathcal{F}^K \cap \mathcal{F}_{G_1} \cap \dots \cap \mathcal{F}_{G_n}$ . It can be shown that this class of subsets of  $\mathcal{F}$  is a *base* for a topology on  $\mathcal{F}$  (called the hit-or-miss topology), and that the topological space is compact, Hausdorff and separable. In fact it can be shown (Matheron (1975, p. 28)) that all that is needed is hit-or-miss information either on the set of all compact sets, or on the set of all open sets; we will return to this point later. Equipped with a topology on  $\mathcal{F}$ , one can now be rigorous about convergence of a sequence of closed sets. Furthermore, by taking countable unions and intersections of the open sets of the topological space  $\mathcal{F}$ , a  $\sigma$ -algebra  $\Sigma$  on  $\mathcal{F}$  is generated.

A *random closed set* or RACS (which is often just called a random set) is defined as a measurable mapping  $X$  from a probability space  $(\Omega, \mathcal{A}, Q)$  into the measure space  $(\mathcal{F}, \Sigma)$ . Let  $\text{Pr}$  be the law of  $X$ , i.e., the probability induced on  $\Sigma$  by

$$\text{Pr}(v) \equiv Q(X^{-1}(v)), \quad v \in \Sigma.$$

Special cases are random variables, random vectors and point processes, while for its more general form, Matheron (1975) has defined a RACS whose realizations are elements of a locally compact, Hausdorff, and separable topological space.

**3.2. Hit-or-miss.** The hit-or-miss topology is basic to this theory of random sets. It was chosen because it reflects the way image data in  $\mathbf{R}^d$  are analyzed; i.e., its roots are in practical applications. Often there is little to be gleaned from an image or pattern in  $\mathbf{R}^d$  just by looking at it (although of course it is the first thing to be done). Clearly some sort of systematic probing is needed, which leads to the use of structuring elements  $B$  (chosen independently of the image) to check whether “ $B$  hits  $X$ ” ( $B \cap X \neq \emptyset$ ) or “ $B$  misses  $X$ ” ( $B \cap X = \emptyset$ ). Furthermore, suppose  $\mathcal{P}(\mathbf{R}^d)$ , the set of all subsets of  $\mathbf{R}^d$ , is equipped with a  $\sigma$ -algebra generated by

$\mathcal{P}_G = \{P \in \mathcal{P}(\mathbf{R}^d) : P \cap G \neq \emptyset\}$ ,  $G$  open. Then the equivalence

$$P \cap G \neq \emptyset \Leftrightarrow \bar{P} \cap G \neq \emptyset$$

shows that in order to study *any* random set with the  $\sigma$ -algebra generated by  $\mathcal{P}_G$ , it is equivalent to study its closure using the  $\sigma$ -algebra  $\Sigma$ . Hence we see also how the “hit-or-miss” approach virtually *demands* the study of random *closed* sets. This restriction of the type of sets under study is a strength of the approach, since it reflects the reality of the objects being modeled. For example, no experiment can hope to distinguish between  $X$  being a disk of the plane, or being only the set of irrational points in that disk.

**3.3. Choquet’s theorem and the hitting function.** It can be shown that all the interesting set transformations (dilation, erosion, opening, closing, convexification, etc.) of a RACS  $X$  are themselves RACS. Matheron (1975, p. 28) has shown that, provided the set transformation is upper or lower semicontinuous into  $\mathcal{F}$ , then the transform of the RACS  $X$  is also a RACS. Therefore, to analyze set data, all one needs is a “bagful” of random-set models, and the rest is in principle straightforward. But it is here where the random-set approach fails to fulfill its potential.

How can the models be specified? What are the important events that make two random sets different? For a partial answer, we return to the hit-or-miss topology. If we can specify  $\Pr(X \in \mathcal{F}^K \cap \mathcal{F}_{G_1} \cap \dots \cap \mathcal{F}_{G_n})$  for all compact  $K$ , and all open  $G_1, \dots, G_n$ , for all integers  $n$ , in a consistent way, then  $X$  is well defined. Fortunately a great reduction of test sets is possible.

For any  $K \in \mathcal{K}$ , define the *hitting function*  $T$  as

$$T(K) = \Pr(X \in \mathcal{F}_K) = \Pr\{X \cap K \neq \emptyset\}.$$

Then  $T$  has the following properties (Matheron (1975, p. 29)):

- (i)  $T(\emptyset) = 0$  and  $0 \leq T \leq 1$ .
- (ii)  $T$  is increasing.
- (iii)  $T$  satisfies the following recurrence relations. For any  $n \geq 0$ , let  $S_n(B_0; B_1, \dots, B_n)$  denote the probability that  $X$  misses  $B_0$  but hits  $B_1, \dots, B_n$ . Then

$$S_0(B_0) = 1 - T(B_0) \geq 0$$

$$S_1(B_0; B_1) = T(B_0 \cup B_1) - T(B_0) \geq 0$$

⋮

$$S_n(B_0; B_1, \dots, B_n) = S_{n-1}(B_0; B_1, \dots, B_{n-1}) - S_{n-1}(B_0 \cup B_n; B_1, \dots, B_{n-1}) \geq 0.$$

That is,  $T$  is a Choquet capacity of infinite order. A powerful result, proved independently by Matheron (1971) and by Kendall (1974), is Choquet’s theorem in the context of random-set theory; it says that the converse of the above is true. In other words, if a given  $T$  on  $\mathcal{K}$  is a Choquet capacity of infinite order, there exists a *necessarily unique*  $P_T$  on  $\Sigma$  such that

$$P_T(\mathcal{F}_K) = T(K) \quad \text{for all } K \in \mathcal{K}.$$

An immediate example of its use is when the RACS  $X$  is an orderly point process (i.e., no more than one event at any location) in  $\mathbf{R}^d$ , a.s. locally finite. Let  $N(A)$  denote the number of points of the process in  $A \subset \mathbf{R}^d$ . Then Choquet’s theorem says the

point process is completely specified from

$$\begin{aligned} T(K) &= \Pr(X \cap K \neq \emptyset) \\ &= 1 - \Pr(X \cap K = \emptyset) \\ &= 1 - \Pr(N(K) = 0) \quad \text{for all } K \in \mathcal{K}. \end{aligned}$$

This observation that the point process is uniquely determined from  $\{\Pr(N(K) = 0): K \in \mathcal{K}\}$  was made by Ripley (1976) in his corollary on p. 989. It is tempting to bracket Choquet’s theorem with the result which says that a random variable  $X$  is well defined once the probabilities of events  $\{X \leq x\}$ , for all  $x \in (-\infty, \infty)$ , are specified in a consistent way. But the results, while being similar, are not identical. In fact, Choquet’s theorem for a random variable, where the RACS  $X$  is a one-point set in  $\mathbf{R}^1$ , involves test sets  $\{[a, b]: -\infty < a \leq b < \infty\}$ ; more work is needed to modify the necessary test sets down to  $\{(-\infty, x]: -\infty < x < \infty\}$ . It is in this domain, namely finding ways to reduce the hitting-function test sets down from the full complement  $\mathcal{K}$ , that results are scarce. We believe that this has greatly held back the development of random-set *models*.

If something extra is known about the random set  $X$ , say all its Minkowski functionals (e.g., volume, surface area, diameter, etc.) a.s. exist and are finite (see Serra (1982, Chap. V)), then in principle this extra knowledge should reduce the number of test sets needed (Molchanov (1984)). For example, Trader and Eddy (1981) considered a.s. compact convex sets, and were able to work with events  $\{X \subset C\}$ , for all  $C \in C(\mathcal{K})$ . But  $\Pr(X \subset C) = \Pr(X \cap C^c = \emptyset) = 1 - \Pr(X \cap C^c \neq \emptyset) = 1 - T(C^c)$ . Not only are the number of test sets reduced from that of Choquet’s theorem, but also  $\{C^c: C \in C(K)\}$  is not even contained in  $\mathcal{K}$ . Trader (1981) has demonstrated the quite general result that just as  $\{T(K): K \in \mathcal{K}\}$  determines the probability measure of a RACS  $X$ , so also does  $\Pr\{X \subset K\}$ , for all  $K \in \mathcal{K}$ ; i.e., so also does  $\{T(K^c): K \in \mathcal{K}\}$ . This is perhaps not so surprising since  $K^c$  is an open set, which in turn can be approximated by a sequence of compact sets, and the compact sets themselves are measure determining. Ripley (1981, §9.1) also discusses the problem of choice of test sets. The strongest result so far available is due to Salinetti and Wets (1986), who prove that  $\{T(U): U \in \text{set of all finite unions of closed balls in } E\}$  determines the probability measure of a RACS.

Those who wish to build models depending on sets more regular than those of  $\mathcal{K}$  struggle with the test sets of Choquet’s theorem and, even when a reduction is possible, it is not always easy to calculate the hitting function. Suppose that the random set  $X$  in  $\mathbf{R}^2$  is the random ray obtained by taking a random point on the unit circle, according to a distribution function given by  $F(\theta) = \Pr\{\text{point} \in \text{arc } [0, \theta]\}$ , and joining this point to the origin. The test sets are simply the arcs  $\{A(\theta): \theta \in [0, 2\pi]\}$  pictured in Fig. 2.

Now generate an independent random *vector*  $W$  in  $\mathbf{R}^2$ , according to a distribution function  $G(\cdot)$ . Then the random set  $X \oplus W = \{x + W: x \in X\}$  is well defined and has hitting function  $T_{X \oplus W}(K) = \int_{\mathbf{R}^2} T_X(K-w) dG(w)$ ,  $K \in \mathcal{K}$ ; recall  $K-w = \{k-w: k \in K\}$ . In fact the test sets can again be reduced to arcs, but no longer necessarily centered at the origin. Hence we need to be able to calculate  $\Pr\{X \cap A(\theta)_b \neq \emptyset\}$ , for any  $b \in \mathbf{R}^2$  and any  $\theta \in [0, 2\pi]$ . This is not a trivial task once the harmony of both  $X$  and the test set, being centered at the origin, is broken.

**3.4. Tumor-growth models.** An important naturally occurring phenomenon to try to characterize geometrically is that of tumor growth. Models have usually been



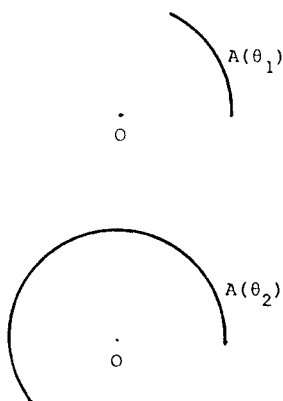


FIG. 2. Examples of test sets for random rays on the unit disc in  $\mathbf{R}^2$ .

formulated by mathematicians at the local (i.e., cell) level, leading to a very sophisticated theory of interacting particles. The simple growth model of Eden (1961) is generalized by Williams and Bjerknæs (1972), who show geometric shapes from their simulations, and make conjectures about the geometry of the boundary of the cancerous cells. For the most part, the subsequent literature has not dealt with the geometric problems, but rather with characterizing one-dimensional summaries (e.g., volume, leading front, radius) of the tumor through (sometimes stochastic) differential equations (Laird (1964); Burton (1966); Saidel, Liotta and Kleinerman (1976); Hanson and Tier (1982); LeCam (1982); Bartoszynski, Brown and Thompson (1982)). Thus in spite of comments made by clinicians (Rubin (1982)) that the tumor's appearance (i.e., fibrous versus solid) is extremely important in characterizing its growth, very little research has been devoted to working directly with the tumor as a geometric object. Some asymptotic results have appeared that prove, for the interacting-particle models mentioned above, that the tumor is asymptotically circular (Richardson (1973); Schurger (1979); Bramson and Griffeath (1981); Durrett and Liggett (1981)). However, there is little guidance as to what this means *in vivo*.

Cressie (1984) has proposed to build generations of the tumor via the iterative equation (for which the hitting function can be calculated):

$$(3.1) \quad Y_{t+1} = \bigcup_{a \in P_0(Y_t)} X^a \oplus a,$$

where  $P_0(A)$  is a Poisson process on the set  $A$  (i.e.,  $N(A)$  is a Poisson random variable, and conditional on  $N(A)$  the points of the process are distributed uniformly over the set  $A$ ), and where  $\{X^a: a \in P_0(A)\}$  is a collection of independent and identically distributed compact RACS. By appropriate tuning of the parameters of (3.1), the tumor can be made to grow or regress. Further details of this modeling of tumor growth using random sets (e.g., hitting function, simulation, fitting to image data, etc.) will be found in Cressie (1988, Chap. 9).

**3.5. Random set limit theorems.** Even such a simple transformation as random translation involves extremely complicated hitting function calculations (see §3.3). The situation becomes intractable when we try to calculate the hitting function of the sum of two independent copies of  $X$ , namely  $T_{X_1 \oplus X_2}$ ; by definition  $X_1 \oplus X_2$ , the Minkowski sum of  $X_1$  and  $X_2$ , is simply the union of random translates of  $X_1$ , the random translation vector ranging over the random set  $X_2$ .

An important set parameter to estimate is the "expected value"  $E(X)$  of the random set  $X$ . This can be well defined via selections (Aumann (1965); Artstein and Vitale (1975)). The natural estimator is

$$\bar{X}_n \equiv (X_1 \oplus X_2 \oplus \dots \oplus X_n)/n,$$

the Minkowski average of  $n$  independent copies  $X_1, X_2, \dots, X_n$ , of  $X$ . Limit theorems for this estimator are necessary for formal inference. Artstein and Vitale (1975), Cressie (1978), Hess (1979), Artstein and Hart (1981) and Puri and Ralescu (1983) prove strong laws of large numbers for  $\bar{X}_n$  under various conditions. They show that  $\bar{X}_n$  converges to the expected *convex hull* of  $X$ , with probability 1. Eddy (1982) used the union and intersection operations instead of the Minkowski sum to find analogues of univariate extreme-value limit theorems for RACS.

Cressie (1979b), Ljašenko (1979), Trader and Eddy (1981), Vitale (1981), Weil (1982) and Artstein (1984) show, under various conditions, that the rate of convergence in the strong law of large numbers is  $n^{-1/2}$ ; Cressie's central limit theorem is geometric in that limiting normalized sets are given, whereas the other authors' theorems are in terms of normalized Hausdorff distances, which lose the geometric subtleties of the limiting process. These latter methods are employed by Giné, Hahn and Zinn (1983) to carry over any probability result in a Banach space (law of large numbers, central limit theorem, law of the iterated logarithm, etc.) to Minkowski sums of compact convex random sets. The convexity condition has recently been dropped by Puri and Ralescu (1985).

All of the proofs of the above limit theorems are obtained by direct inspection of the random set  $(X_1 \oplus X_2 \oplus \dots \oplus X_n)/n$ ; the absence of the hitting function is notable. Norberg's (1984) result, which for convergence in distribution requires only convergence of hitting functions on a suitable countable class of bounded Borel sets, might possibly be used in cases where direct inspection of the random set is not possible.

**4. The Boolean model.** Data analysis when the data are sets is not a situation with which most statisticians feel comfortable. There are certain exploratory ways of looking at the data, but if one wants to consider them as being "representative" of a phenomenon, with information on interpretable "average" quantities (parameters) associated with the phenomenon, then one must turn to a model. This is true for studying any type of random variation, but it is particularly difficult in the case of random sets because of the dearth of tractable models available. Moreover, since the sets (in  $\mathbf{R}^d$ ) usually have to be probed in some one- or two-dimensional way, there is an extra source of "inexactness" in the inference process. Suppose a particular set model is used to represent a random phenomenon; inference from the probes to the model parameters (a part of stereology) is a hard problem in itself, quite apart from the problem of assimilating a model to the set data.

In this section we shall present what is arguably the most important set model, namely the Boolean model, and show how its properties can be used in the analysis of the data of Fig. 1(a). Generalizations to other models will be discussed, but it is clear that a fruitful path to broader classes of models has yet to be developed.

It appears that Solomon (1953) was the first person to consider this model in the literature (see also Matern (1960)). Marcus (1966), (1967) uses the Boolean model to examine the meteoroidal impact hypothesis for the origin of lunar craters, Dupuč (1980) considers the etching of tracks formed by the fission of randomly located uranium atoms in a fission material, Serra (1980) models ore-sintering, and Diggle (1981) uses it to model the incidence of heather (he calls it a random binary mosaic).

The Boolean model is obtained by fixing independent realizations of a random closed and bounded set (called the primary set)  $S$  in  $\mathbf{R}^d$ , at each point of a realization

$\{t_i\}$  of a homogeneous Poisson process in  $\mathbf{R}^d$ , and then taking the union. This results in

$$(4.1) \quad X = \bigcup_i (S_i \oplus t_i)$$

where  $S_i \oplus t_i$  is the  $i$ th realization of  $S$  fixed at  $t_i$ . Thus there are two sources of randomness in the model:

- (i) The Poisson process, characterized by the intensity  $\lambda$ ;
- (ii) The probability law of the (random) primary set  $S$ .

**4.1. The hitting function.** The hitting function of the Boolean model  $X$  can be constructed in several stages.

Consider the *homogeneous* Poisson process in  $\mathbf{R}^d$ , with constant intensity  $\lambda$ , which is characterized as follows. Let  $N(A)$  denote the number of points of the process in any subset  $A \subset \mathbf{R}^d$ . Then if  $A_1 \cap A_2 = \emptyset$ ,  $N(A_1)$  and  $N(A_2)$  are independent Poisson random variables with respective means  $\lambda|A_1|$  and  $\lambda|A_2|$ . Also,

$$\Pr\{N(dA) = 0\} \sim 1 - \lambda dx,$$

$$\Pr\{N(dA) = 1\} \sim \lambda dx,$$

where  $dx$  is the volume element of the infinitesimal region  $dA$ ; i.e.,  $dx = |dA|$ .

Now

$$\begin{aligned} Q_X(K) &= \Pr(X \cap K = \emptyset) \\ &= \Pr((S_i \oplus t_i) \cap K = \emptyset, \text{ for all } i), \end{aligned}$$

where  $S_i$  is a realization of a random closed and bounded set  $S$  and  $\{t_i\}$  are the points of the Poisson process. In the infinitesimal region of volume  $dx$  centered at  $x \in \mathbf{R}^d$ , two mutually exclusive events may happen:

- (i) There is no element of  $\{t_i\}$  in  $dx$ ; this occurs with probability  $1 - \lambda dx$ .
- (ii) There is one element of  $\{t_i\}$  in  $dx$  centered at  $x$ , but  $S \oplus x$  does not hit  $K$ ; this occurs with probability  $\lambda dx Q_S(K_{-x})$ .

To the order of magnitude ignored, this yields

$$1 - \lambda(1 - Q_S(K_{-x})) dx \sim \exp[-\lambda(1 - Q_S(K_{-x})) dx].$$

Hence,  $Q_X(K)$  is obtained by taking the product over all disjoint regions:

$$(4.2) \quad Q_X(K) = \exp\left[-\lambda \int (1 - Q_S(K_{-x})) dx\right].$$

This formula is fundamental in linking the hitting function of  $X$  to the hitting function of  $S$ .

Now define  $I_{S \oplus \check{K}}(\cdot)$  to be the indicator function of  $S \oplus \check{K} = \{s - k : s \in S, k \in K\}$ , where  $\check{K} \equiv \{-k : k \in K\}$ . But

$$|S \oplus \check{K}| = \int_{\mathbf{R}^d} I_{S \oplus \check{K}}(x) dx,$$

and so

$$\begin{aligned} E\{|S \oplus \check{K}|\} &= \int E(I_{S \oplus \check{K}}(x)) dx = \int T_S(K_x) dx = \int T_S(K_{-x}) dx \\ &= \int (1 - Q_S(K_{-x})) dx, \end{aligned}$$

where the integrals with respect to  $x$  are taken over  $\mathbf{R}^d$ . Finally then,

$$T_X(K) = 1 - Q_X(K),$$

where

$$(4.3) \quad Q_X(K) = \exp[-\lambda E\{|S \oplus \check{K}|\}].$$

**4.2. Properties of the Boolean model.** We will summarize the main characteristics of the Boolean model, from which it should be made clear why it is central among random-set models. More details can be found in Serra (1980).

(i) *Porosity,  $q$ .* Porosity is the probability that a point of the space is in the complement of  $X$  (i.e., in the pores).

Let  $K = \{x\}$ . Then  $q \equiv Q_X(\{x\}) = \Pr\{x \in X^c\}$ ; i.e.,  $q = e^{-\lambda E\{|S|\}}$ .

(ii) *Covariogram of the pores,  $\chi(h)$ .* The covariogram is the probability that two points,  $h$ -apart, are both in the complement of  $X$ ; it measures dependence between pores.

Let  $K = \{x, x + h\}$ , and  $\chi(h) \equiv Q_X(\{x, x + h\})$ . Then

$$\begin{aligned} \chi(h) &= \Pr\{x, x + h \in X^c\} \\ &= \exp[-\lambda E\{|S \cup S_{-h}|\}] \\ &= q^2 e^{\lambda K(h)}, \end{aligned}$$

where

$$K(h) = E\{|S \cap S_{-h}|\} = \int \Pr\{x \in S \cap S_{-h}\} dx,$$

the probabilistic analogue of the geometrical covariance.

(iii) *Stationarity.* Since  $E\{|\cdot|\}$  ignores location,  $X$  is clearly stationary.

(iv) *Stability under dilation.* From (4.3), the RACS  $X \oplus L = \{x + l: x \in X, l \in L\}$ , for  $L$  any deterministic compact set, is also a Boolean model.

(v) *Cross-sections are Boolean models.* This is true because  $K$  in (4.3) may belong to a subspace of  $\mathbf{R}^d$ .

(vi) *Infinite divisibility with respect to union.* A RACS  $Y$  is said to be *infinitely divisible with respect to union* if for any integer  $m > 0$ ,  $Y$  is equivalent to  $\cup_{i=1}^m Y_i$  of  $m$  independent equivalent RACS  $\{Y_i; i=1, \dots, m\}$ . This shows the Boolean model to be a candidate model for limits of unions of RACS.

(vii) *The Boolean model with convex  $S$  is semi-Markov.* A RACS  $Y$  is said to be *semi-Markov* if for any  $K, L$  and  $M \in \mathcal{K}$ , and  $K$  and  $M$  separated by  $L$  (i.e., the segment joining any point of  $K$  to any point of  $M$  hits  $L$ ), the RACS  $Y \cap K$  and  $Y \cap M$  are conditionally independent given  $Y \cap L = \emptyset$ . Thus the Boolean model's probability law is really determined by "local" conditions. Widely separated parts of the set are only weakly dependent.

In terms of the functional  $Q$ , Matheron (1975, p. 132) has shown that if  $Y$  is infinitely divisible with respect to union, and if for any  $K, M, K \cup M \in C(\mathcal{K})$ ,

$$Q_Y(K \cup M)Q_Y(K \cap M) = Q_Y(K)Q_Y(M),$$

then  $Y$  is semi-Markov. Matheron (1975, p. 148) has furthermore provided the following important characterization theorem (presented here in  $\mathbf{R}^3$ ): Any RACS in  $\mathbf{R}^3$  which is stationary, infinitely divisible with respect to union and semi-Markov is equivalent to

$$X_1 \cup X_2 \cup X_3,$$

where  $X_1, X_2, X_3$  are stationary, independent and

$X_1 =$  Boolean model with primary (random) sets  $S$  convex.

$X_2 =$  union of cylinders with bases that are two-dimensional Boolean models with primary sets  $S$  (in  $\mathbf{R}^2$ ) convex.

$X_3 =$  union of cylinders with bases that are one-dimensional Boolean models with primary sets  $S$  (in  $\mathbf{R}^1$ ) convex.

Notice that if the  $S$  in  $X_2, X_3$  are a.s. points, then the associated cylinders become Poisson lines and Poisson planes respectively, yielding Poisson-flat processes studied extensively by Miles (1969).

(viii) *The union of two independent Boolean models with identically distributed primary sets  $S$  is a Boolean model.*

In fact, if the Poisson points of a Boolean model are *thinned* so that they now occur with an intensity  $\rho\lambda$  ( $\rho < 1$ ), leaving behind a Boolean model whose union is taken with another, independent Boolean model of intensity  $(1 - \rho)\lambda$  and the same primary set distribution, *then* the resulting random set is a Boolean model identically distributed to the original one.

**4.3. Generalizations of the Boolean model.** The choice of mathematical models available to the data analyst is often governed by their tractability rather than their applicability. When the data are sets, this leaning is even more pronounced. Serra (1982, Chap. XIII) has provided users with a menu of models and of examples for which they are appropriate. By far the most important groupings are those based on the Boolean model, which we will present and extend in this section.

The most general extension of the Boolean model considered thus far is the so-called *grain-germ-model* (Hanisch (1980)). It removes the Poisson assumption and the independence of the  $S_i$ 's, and allows for nonoverlapping of the grains. Let

$$(4.4) \quad X = \bigcup_{(t_i, S_i) \in \Phi} (S_i \oplus t_i),$$

where  $\Phi$  is a random marked point process with mark space  $\mathcal{H}$ . Let  $U$  be the set of all measurable functions from  $\mathbf{R}^d \times \mathcal{H}$  to  $[0, 1]$ , and  $V = \{1 - u : u \in U\}$ . Let  $\Phi$  be a random marked point process in  $\mathbf{R}^d$  with mark space  $\mathcal{H}$  and corresponding distribution  $P$  on  $M_{\mathcal{H}}, \mathcal{M}_{\mathcal{H}}$ , where  $M_{\mathcal{H}}$  is the set of all Radon counting measures  $\phi$  on  $\mathbf{R}^d \times \mathcal{H}$  with  $\phi(B \times \mathcal{H}) < \infty$  for any bounded Borel set  $B$ , and  $\mathcal{M}_{\mathcal{H}}$  is the corresponding  $\sigma$ -field. The functional  $G_P: V \rightarrow [0, 1]$ , which is given by

$$(4.5) \quad G_P(v) = \int_{M_{\mathcal{H}}} \prod_{(x, S) \in \phi} v(x, S) P(d\phi), \quad v \in V,$$

is called the generating functional of the point process  $\Phi$ . For  $X$  given by (4.4), Hanisch is able to relate  $Q_X$  to  $G_P$ :

For  $K \in \mathcal{H}$  let  $v_K$  be the mapping given by

$$v_K(x, S) = 1 - I_{S \oplus K}(x), \quad x \in \mathbf{R}^d, \quad S \in \mathcal{H}.$$

Then

$$(4.6) \quad Q_X(K) = G_P(v_K), \quad K \in \mathcal{H}.$$

Special cases of (4.4) yield models already studied in the literature:

(i) The marked point process  $\Phi$  a.s. yields *independent* markings (see Stoyan (1979) and Mase (1982)).

(ii) The point process is a cluster process where initial points are generated according to a homogeneous Poisson process, and final points are generated independently and identically around each initial point (see Neyman and Scott (1958)).

(iii) The point process is regionally independent, in particular a (not necessarily homogeneous) Poisson process with intensity function  $\{\lambda(x); x \in \mathbf{R}^d\}$ ; if we define the *weighted* measure of  $S \oplus \check{K}$  as

$$|S \oplus \check{K}|_\lambda \equiv \int_{\mathbf{R}^d} I_{S \oplus \check{K}}(x) \lambda(x) dx,$$

then

$$(4.7) \quad Q_X(K) = \exp [-E\{|S \oplus \check{K}|_\lambda\}].$$

Hypothesis testing for constant intensity function in the Boolean model is not, as yet, well developed.

(iv) The point process is regionally independent and stationary. Then this *must* be a Poisson process with  $\lambda(x) = \lambda$ , which yields the Boolean model (4.1). Serra (1982, p. 484ff) has a detailed discussion of the model and various of its associates. One interesting extension is to consider a series of independent Boolean models occurring from the “infinite past”; at each instant of time the Boolean model fills part of the space. At the next instant, some of the pores (and sets) will be covered by a set from the new Boolean model, and some of the sets (and pores) will remain exposed. This process is continued until “the present,” so that finally a tessellation of the space results. This tessellation is called the *dead-leaves model*.

(v) The points are not generated by a point process, but are fixed and finite in number in  $\mathbf{R}^d$ . The only source of randomness is from the  $S$ . Then in (4.5) the measure  $P$  becomes degenerate, and

$$G_P(v_K) = \prod_a E(v_K(a, S)).$$

Hence,

$$Q_X(K) = \prod_a Q_S(K_{-a}).$$

(vi) The point process is homogeneous Poisson and the  $S$  have the property that  $|S|$  is *not* random (although in general unknown)

$$\log q = -\lambda |S|,$$

$$\int \log (\chi(h)/q^2) dh = \lambda |S|^2,$$

where  $q$  and  $\chi(h)$  are given in §4.2, and can be estimated from the set data. Hence  $\lambda$  and  $|S|$  can also be estimated (although the statistical properties of these estimates are unknown).

(vii) The point process is homogeneous Poisson and the  $S$  can be written as  $B(1)R$ , where  $R$  is a random variable and  $B(1)$  is the closed unit ball. The parameter  $\lambda$  and the distribution of  $R$  can be estimated from the covariance  $\chi(h)$  (Serra (1980)), although little is known about the estimators’ statistical properties (see Dupac̆ (1980)).

**4.4. Statistical inference for the Boolean model.** In this section, we shall analyze the two-dimensional data of Fig. 1(a), an artificially generated Boolean model whose primary elements are uniformly oriented random parallelograms. More specifically, at each point of a simulated Poisson process with rate  $\lambda = .02$ , we fixed the bottom left-hand corner of a random parallelogram (the length of both sides was uniform on  $[2.5, 7.5]$ , and the angle between the adjacent sides was uniform on  $[0, \pi]$ ), which was

then uniformly oriented on  $[0, 2\pi]$ . Figures 1(a), (b) show two realizations of the same Boolean model, discretized onto a  $100 \times 100$  grid.

More generally, consider an a.s. convex primary set  $S$ , whose probability law is invariant under rotations of  $S$  about its "fixing" point. This yields a stationary isotropic Boolean model  $X$ . Under these circumstances, and for  $K$  convex, Steiner's formula (Mack (1954) and Serra (1982, p. 111)) in  $\mathbf{R}^2$  yields

$$(4.8) \quad E\{|S \oplus \check{K}|\} = E(|S|) + \frac{1}{2\pi} E(P(S))P(K) + |K|,$$

where  $P \equiv$  perimeter. When  $S = B(1)R$ ,  $E(P(S))/2\pi = E(R)$ , and  $E(|S|)/\pi = E(R^2)$ . Faced with one realization of what is believed to be a Boolean model, one wishes to make inferences about the convex set  $S$ . From (4.3) and (4.8)

$$(4.9) \quad Q_X(K) = \exp \left[ -\lambda \left\{ |K| + \frac{1}{2\pi} P(K)E(P(S)) + E(|S|) \right\} \right],$$

and hence estimators can be found for  $\lambda$ ,  $E(P(S))$  and  $E(|S|)$  by, for example, fixing  $K$  to be  $B(1)$ , the disk with unit radius, and considering a number of test sets  $B(t) \equiv tB(1)$  for various  $t$ . Theoretically,

$$\begin{aligned} -\log Q_X(B(t)) &= \lambda[\pi t^2 + E(P(S))t + E(|S|)] \\ &= \beta_0 + \beta_1 t + \beta_2 t^2. \end{aligned}$$

By scanning the image, one obtains an estimate  $\hat{Q}_X(B(t))$ , for various  $t$ , which, when regressed on  $1, t, t^2$ , yields  $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ . Image analyzers (Serra (1982)) are specifically built to do this type of task; however, it is not difficult to use conventional computers to analyze random-set data like Fig. 1(a). We computed both least squares and generalized least squares estimates of quadratic regression coefficients for Fig. 1(a). The set parameters in the above case of  $K = B(t)$  are easily obtained as

$$(4.10) \quad \begin{aligned} \hat{\lambda} &= \hat{\beta}_2/\pi, \\ \hat{E}(P(S)) &= \hat{\beta}_1/\hat{\lambda}, \\ \hat{E}(|S|) &= \hat{\beta}_0/\hat{\lambda}. \end{aligned}$$

Alternatively, choice of the three different types of test set:  $K_0$  the origin,  $K_1$  the straight line segment of length  $l$ ,  $K_2$  the closed square of side  $l$ , yields

$$(4.11) \quad \begin{aligned} Q_X(K_0) &= \exp[-\lambda E(|S|)], \\ Q_X(K_1) &= \exp \left[ -\lambda \left\{ E(|S|) + \frac{l}{\pi} E(P(S)) \right\} \right], \\ Q_X(K_2) &= \exp \left[ -\lambda \left\{ E(|S|) + \frac{2l}{\pi} E(P(S)) + l^2 \right\} \right], \end{aligned}$$

from which estimates could be obtained. Hall (1985) and Kellerer (1985) take essentially this approach; however, it leaves no degrees of freedom to assess model adequacy.

Because the original Boolean model has been discretized onto a square grid, we chose to analyze Fig. 1(a) with  $K = tC$ , where  $C$  is a square of side 2, and took  $t = 0, 1, 2, 3, 4$  (larger values of  $t$  led to  $\hat{Q}_X \equiv 0$ ). Then (4.9) gives

$$-\log Q_X(tC) = \lambda E(|S|) + \{4\lambda E(P(S))/\pi\}t + 4\lambda t^2.$$

Explicitly, the estimator for the left-hand side (lhs) is

$$\hat{Q}_X(tC) = \sum_{x_i \in E_t} \{1 - I_{X \oplus tC}(x_i)\} / N_t, \quad t = 0, \dots, 4,$$

where  $N_t$  is the number of pixels (at locations  $\{x_i: i = 1, \dots, N_t\}$ ) in the eroded mask  $E_t = E \ominus tC = \cap_{c \in C} E_{tc}$ , placed over the region of interest  $E$  (in Figs. 1(a) and 1(b),  $E$  is the bordered region). Clearly this is an unbiased estimator of  $Q_X(tC)$ . Write  $\underline{Y}' = (-\log \hat{Q}_X(tC): t = 0, \dots, 4)$ , and let  $X$  be the  $5 \times 3$  matrix whose  $t$ th row is  $(1, t, t^2); t = 0, \dots, 4$ . Then the usual ordinary-least-squares (ols) estimator is given by  $(X'X)^{-1} X' \underline{Y}$ , which estimates the coefficients of 1,  $t$ , and  $t^2$  without taking into account correlations between the elements of  $\underline{Y}$ . To exploit these correlations we took a generalized-least-squares (gls) approach to the estimation of the parameters  $\lambda$ ,  $E(P(S))$ , and  $E(|S|)$ . Table 1 gives the ols and gls estimates for the data of Fig. 1(a).

TABLE 1  
*Summary statistics, and final estimates for the simulated Boolean model data of Fig. 1(a); the true values of parameters are  $\lambda = .02$ ,  $E(P(S)) = 20.00$ ,  $E(|S|) = 15.92$ .*

$t$	$\hat{Q}_X(tC)$	Fitted quadratic (gls)	Fitted quadratic (ols)
0	0.72950	0.72985	0.72692
1	0.44138	0.44137	0.44430
2	0.22797	0.22678	0.22831
3	0.09948	0.09901	0.09863
4	0.03568	0.03673	0.03582
		Parameter estimates (gls)	Parameter estimates (ols)
	$\hat{\lambda}$	0.020	0.022
	$\hat{E}(P(S))$	16.27	14.69
	$\hat{E}( S )$	15.47	14.71

To give the gls estimator explicitly, some definitions are needed first. From the definition of  $\hat{Q}_X(tC)$  given above, it is easily seen that for the Boolean model,

$$\text{cov}(\hat{Q}_X(tC), \hat{Q}_X(uC)) = \sum_{x_i \in E_t} \sum_{x_j \in E_u} C_{t,u}(x_i - x_j) / N_t N_u,$$

where

$$C_{t,u}(h) = [\exp \{ \lambda E(|(S \oplus tC) \cap (S_{-h} \oplus uC)|) \} - 1] \times \exp \{ -\lambda E(|S \oplus tC|) - \lambda E(|S \oplus uC|) \}.$$

The first term in this expression for  $C_{t,u}(h)$  can be simplified for special cases of structuring element  $C$ . Regardless, the expression for the covariance depends on the unknown parameters; recall our objective here is to estimate these parameters. For this reason, it makes more sense to compute the empirical covariance function, and use it in gls estimation of the parameters.

The pixels of Fig. 1 are regularly spaced, and so  $C_{t,u}(h)$  can be estimated for  $h$  in horizontal, vertical, and diagonal directions. Let

$$\hat{C}_{t,u}(h) \equiv \left[ \frac{\sum_{x_i, x_j \in M(h)} \{1 - I_{X \oplus tC}(x_i)\} \{1 - I_{X \oplus uC}(x_j)\}}{\sum_{x_i, x_j \in M(h)} 1} \right] - \hat{Q}_X(tC) \cdot \hat{Q}_X(uC),$$



where  $M(h) = \{(x_i, x_j): x_i \in E_l, x_j \in E_u, x_i - x_j = h\}$ . Write

$$\hat{\Sigma} = (\text{côv}(\hat{Q}_X(tC), \hat{Q}_X(uC)): t=0, \dots, 4; u=0, \dots, 4),$$

and

$$\underline{r} = (\hat{Q}_X(tC) - \exp\{-\beta_0 - \beta_1 t - \beta_2 t^2\}: t=0, \dots, 4).$$

Then our proposed gls estimator for  $\beta_0 = \lambda E(|S|)$ ,  $\beta_1 = \{4\lambda E(P(S))/\pi\}$ , and  $\beta_2 = 4\lambda$ , is obtained by minimizing  $\underline{r}' \hat{\Sigma}^{-1} \underline{r}$ . This was done using a quasi-Newton numerical optimization procedure.

Occasionally  $\hat{\Sigma}$  has proved not to be positive-definite. Difficulties of this sort are to be expected when empirical covariances are used; for example, a time series  $\{z_t: t = 1, \dots, n\}$  has the empirical covariance function  $\{\sum_{t=1}^{n-h} (z_{t+h} - \bar{z})(z_t - \bar{z}) / (n-h): h = 1, \dots, n-1\}$ , which is not assured of being positive-definite. The usual way around this problem is to fit a positive-definite model to the empirical covariance function. We have been experimenting with a displaced Weibull for the Boolean-model covariance, and will report on these results elsewhere. After we analyzed a number of images like Fig. 1, the general rule emerged that gls fitting of the equation (4.9) yielded superior estimates (to ols) of the Boolean-model parameters.

The estimation technique of (4.10) and (4.11) could be liberally described as a "method-of-moments." The matching of theoretical moments to sample values to estimate parameters is well known in the statistics of random variables and random vectors, although it is usually used only when other approaches such as maximum likelihood fail. This is because there is no general theory that will yield statistical properties of the estimators. The same is true here when the method is applied to random sets. Are the estimators from (4.10) and (4.11) biased? What are their variances and covariances? Under what conditions are they consistent and asymptotically Gaussian (normal)? Are the estimators in any way optimal? Dupač (1980), by assuming  $S = B(1)R$  where  $R$  is Gaussian, produces method-of-moments estimators where for the first time, approximate variances (but not biases) are presented. Ohser (1980) writes about the more general problem, and indicates how expected values of the Minkowski functionals of  $S$  could be estimated (via a method-of-moments). It should be possible, from straightforward but tedious application of his results, to develop estimators of precision.

Another possibility for formal inference appears in the paper by Baddeley (1980). Although he is not directly involved with the estimation of Boolean-model parameters, he presents a limit theorem for the empirical distribution function  $\hat{p}(t)$  of the scalar function  $p(t)$  associated, e.g., with the random set  $X$  dilated by  $tB(1)$ ,  $t \geq 0$ . Suppose  $p(t) = \Pr\{0 \in X \oplus tK\}$ ;  $K$  fixed compact. If  $X$  is a Boolean model (with primary set  $S$ ), then so is  $X \oplus tK$  (with primary set  $S \oplus tK$ ). Hence

$$\begin{aligned} p(t) &= 1 - \exp(-\lambda E|S \oplus tK|) \\ &= 1 - Q_X(tK). \end{aligned}$$

Thus, Baddeley's limit theorem should be able to be modified to prove weak convergence of the empirical process  $\{-\log \hat{Q}_X(tK): t \geq 0\}$  (used in obtaining  $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ , and hence  $\hat{\lambda}, \hat{E}(P(S)), \hat{E}(|S|)$ ; see (4.10)) to a Gaussian process. The design question of which  $t_1, \dots, t_n$  to choose for the regression of  $-\log \hat{Q}_X(tK)$  on  $t$ , is yet to be resolved.

In this section, we have concentrated on the two-dimensional problem. Estimation of the Boolean-model parameters in  $\mathbf{R}^3$  contains a further source of error, since the three-dimensional object is usually sliced up in some way, and two-dimensional sections are analyzed. Recent results that do not assume the primary

three-dimensional sets are necessarily spheres have been obtained by Tallis and Davis (1984).

**5. Conclusions.** Our aim has been to present the theory of random sets from a modeling point of view. The emphasis has been on demonstrating both its flexibility and its shortcomings. For example, summarizing an object in terms of a finite number of measurements (e.g., volume, surface area, etc.), and building a stochastic model based on these measurements is not very helpful if it is important to say something about the object's shape. In this situation a random (stochastic) set model would be more appropriate. But can we then find a meaningful definition of, say, the volume of the random set  $X$ ? Matheron (1975) has shown that a nonempty stationary random closed set  $X$  in  $\mathbf{R}^d$  is almost surely unbounded, causing some difficulty in defining one-dimensional summaries of a stationary  $X$ . Intuitively, we need a notion of volume per unit area, but it is not clear that a sensible definition can be found. Weil and Wieacker (1984) resolve this problem by giving a formula which involves the normalized limit as  $r \rightarrow \infty$  of  $X$  intersected with  $rC$ , where  $C$  is a compact convex set, and show that it does not depend on  $C$ . Their result is very much in the spirit of this article, since it hints at how these specific volumes may then be estimated.

We have explored the random set as a special case of a 0-1 random function, and demonstrated that the usual random function techniques handle the geometry of the random set rather crudely. New (nonlinear) operations are needed to probe and sort complicated sets; Matheron (1975) calls these morphological transformations.

Distribution theory for random-set models depends on the hitting function; important advances need to be made to reduce the number of test sets needed to determine the probability measure of  $X$ . The Boolean model and its various generalizations are a tractable suite of models from which hitting functions can be calculated, and parameters can be estimated. We have shown how to carry out statistical inference on the data of Fig. 1. Sound statistical principles are needed for first estimating set-model parameters, and then estimating the bias and precision of these estimators. For the Boolean model, we have used generalized least squares to solve the first problem; how to effectively solve the second is at present under investigation.

**Acknowledgment.** We would like to thank the referees for perceptive comments which led to substantial improvements in the paper.

#### REFERENCES

- R. ADLER (1981), *The Geometry of Random Fields*, John Wiley, New York.
- Z. ARTSTEIN (1984), *Convergence rates for the optimal values of allocation processes*, Math. Oper. Res., 9, pp. 348–355.
- Z. ARTSTEIN AND S. HART (1981), *Law of large numbers for random sets and allocation processes*, Math. Oper. Res., 8, pp. 482–492.
- Z. ARTSTEIN AND R. VITALE (1975), *A strong law of large numbers for random compact sets*, Ann. Probab., 3, pp. 879–882.
- R. AUMANN (1965), *Integrals of set-valued functions*, J. Math. Anal. Appl., 12, pp. 1–12.
- A. BADDELEY (1980), *A limit theorem for statistics of spatial data*, Adv. in Appl. Probab., 12, pp. 447–461.
- R. BARTOSZYNSKI, B. BROWN AND J. THOMPSON (1982), *Metastatic and systemic factors in neoplastic progression*, in Probability Models and Cancer, L. LeCam and J. Neyman, eds., North Holland, Amsterdam, pp. 253–264.
- M. BRAMSON AND D. GRIFFEATH (1981), *On the Williams–Bjerknes tumour growth model*, I, Ann. Probab., 9, pp. 173–185.
- A. BURTON (1966), *Rate growth of solid tumors as a problem of diffusion*, Growth, 30, pp. 157–176.
- N. CRESSIE (1978), *A strong limit theorem for random sets*, Adv. in Appl. Probab., Suppl. 10, pp. 36–46.

- (1979a), *Random set limit theorems*, in Proc. Eighth Conference on Stochastic Processes and their Applications, Canberra, Australia, 1978; Adv. in Appl. Probab., 11, pp. 281–282.
- (1979b), *A central limit theorem for random sets*, Z. Wahrsch. Verw. Gebiete, 49, pp. 37–47.
- (1984), *Modelling sets*, in Springer Lecture Notes in Mathematics 1091 (Proc. International Conference in Multifunctions and Integrands, Catania, 1983), G. Salinetti, ed., Springer Verlag, New York, pp. 138–149.
- (1986), *Kriging nonstationary data*, J. Amer. Statist. Assoc., 81, pp. 625–634.
- (1988), *Statistics for Spatial Data*. John Wiley, New York, to appear.
- P. DIGGLE (1981), *Binary mosaics and the spatial pattern of heather*, Biometrics, 37, pp. 531–539.
- V. DUPAČ (1980), *Parameter estimation in the Poisson field of discs*, Biometrika, 67, pp. 187–190.
- R. DURRETT AND T. LIGGETT (1981), *The shape of the limit set in Richardson's growth model*. Ann. Probab., 9, pp. 186–193.
- W. EDDY (1982), *Laws of large numbers for intersection and union of random closed sets*, Technical Report No. 227, Department of Statistics, Carnegie-Mellon University, Pittsburgh, PA.
- M. EDEN (1961), *A two dimensional growth process*, in Proc. Fourth Berkeley Symposium on Mathematical Statistics and Probability, IV, J. Neyman, ed., University of California Press, Berkeley, CA, pp. 223–239.
- E. GINÉ, M. HAHN AND J. ZINN (1983), *Limit theorems for random sets: an application of probability in Banach space results*, in Springer Lecture Notes in Mathematics 990 (Proc. Fourth International Conference on Probability in Banach Spaces, Oberwolfach, 1982), A. Beck and K. Jacobs, eds., Springer Verlag, New York, pp. 112–135.
- P. HALL (1985), *Counting methods for inference in binary mosaics*, Biometrics, 41, pp. 1049–1052.
- K.-H. HANISCH (1980), *On classes of random sets and point process models*. Elektron. Informationsverarb. Kybernet., 16, pp. 498–502.
- F. HANSON AND C. TIER (1982), *A stochastic model of tumor growth*, Math. Biosci., 61, pp. 73–100.
- C. HESS (1979), *Théorème ergodique et loi forte des grands nombres pour des ensembles aléatoires*, Comptes Rendus Academie des Sciences Paris A, 288, pp. 519–522.
- A. M. KELLERER (1985), *Counting figures in planar random configurations*, J. Appl. Probab., 22, pp. 68–81.
- D. KENDALL (1974), *Foundations of a theory of random sets*, in Stochastic Geometry, E. Harding and D. Kendall, eds., John Wiley, New York, pp. 322–376.
- A. LAIRD (1964), *The dynamics of tumour growth*, British J. Cancer, 28, pp. 490–502.
- L. LECAM (1982), *On some mathematical models of tumor growth and metastasis*, in Probability Models and Cancer, L. LeCam and J. Neyman, eds., North Holland, Amsterdam, pp. 265–286.
- N. LJAŠENKO (1979), *Limit theorems for sums of independent random sets*, Theory Probab. Appl., 24, pp. 438–440.
- C. MACK (1954), *The expected number of clumps when convex laminae are placed at random and with random orientation on a plane area*, Proc. Cambridge Philos. Soc., 50, pp. 581–585.
- A. MARCUS (1966), *A stochastic model of the formation and survival of lunar craters*, Icarus, 5, pp. 165–200.
- (1967), *A multivariate immigration with multiple death process and applications to lunar craters*, Biometrika, 54, pp. 251–261.
- S. MASE (1982), *Properties of fourth-order strong mixing rates and its application to random set theory*, J. Multivariate Anal., 12, pp. 549–561.
- B. MATERN (1960), *Spatial variation*, Meddelanden fran Statens Skogsforskningsinstitut, 49:5.
- G. MATHERON (1963), *Principles of geostatistics*, Economic Geology, 58, pp. 1246–1266.
- (1971), *Random sets theory and its application to stereology*, J. Microscopy, 95, pp. 15–23.
- (1975), *Random Sets and Integral Geometry*, John Wiley, New York.
- R. MILES (1969), *Poisson flats in Euclidean spaces*, Adv. in Appl. Probab., 1, pp. 211–237.
- I. MOLCHANOV (1984), *A generalization of the Choquet theorem for random sets with a given class of realizations*, Theory Probab. Math. Statist. 28, pp. 99–106.
- J. NEYMAN AND E. SCOTT (1958), *Statistical approach to problems of cosmology*, J. Royal Statist. Soc. Ser. B, 20, pp. 1–43.
- T. NORBERG (1984), *Convergence and existence of random set distributions*, Ann. Probab., 12, pp. 726–732.
- J. OHSER (1980), *On statistical analysis of the Boolean model*, Elektron. Informationsverarb. Kybernet., 16, pp. 651–653.
- M. PURI AND D. RALESCU (1983), *Strong law of large numbers for Banach space valued random sets*, Ann. Probab., 11, pp. 222–224.
- (1985), *Limit theorems for random compact sets in Banach space*, Math. Proc. Cambridge Philos. Soc., 97, pp. 151–158.

- D. RICHARDSON (1973), *Random growth in a tessellation*, Math. Proc. Cambridge Philos. Soc., 74, pp. 515–528.
- B. RIPLEY (1976), *Locally finite random sets: Foundations for point process theory*, Ann. Probab., 4, pp. 983–994.
- (1981), *Spatial Statistics*, John Wiley, New York.
- H. RUBIN (1982), *Some remarks on cancer as a state of disorganization at the cellular and supracellular levels*, in Probability Models and Cancer, L. LeCam and J. Neyman, eds., North Holland, Amsterdam, pp. 211–219.
- G. SAIDEL, L. LIOTTA AND J. KLEINERMAN (1976), *System dynamics of a metastatic process from an implanted tumor*, J. Theoret. Biol., 56, pp. 417–434.
- G. SALINETTI AND R. WETS (1986), *On the convergence in distribution of measurable multifunctions (random sets), normal integrands, stochastic processes and stochastic infima*, Math. Oper. Res.
- K. SCHURGER (1979), *On the asymptotic geometrical behavior of a class of contact interaction processes with a monotone infection rate*, Z. Wahrsch. Verw. Gebiete, 48, pp. 35–48.
- J. SERRA (1980), *The Boolean model and random sets*, Computer Graphics and Image Processing, 12, pp. 99–126.
- (1982), *Image Analysis and Mathematical Morphology*, Academic Press, London.
- H. SOLOMON (1953), *Distribution of the measure of a random two-dimensional set*, Ann. Math. Statist., 24, pp. 650–656.
- D. STOYAN (1979), *Applied stochastic geometry: A survey*, Biometrical J., 21, pp. 693–715.
- G. TALLIS AND A. DAVIS (1984), *The nucleus problem*, Biometrical J., 26, pp. 95–100.
- D. TRADER (1981), *Infinitely divisible random sets*, Ph.D. dissertation, Department of Statistics, Carnegie-Mellon University, Pittsburgh, PA.
- D. TRADER AND W. EDDY (1981), *A central limit theorem for Minkowski sums of random sets*, Technical Report No. 228, Department of Statistics, Carnegie-Mellon University, Pittsburgh, PA.
- R. VITALE (1981), *A central limit theorem for random convex sets*, Technical Report, Department of Mathematics, Claremont Graduate School, Claremont, CA.
- W. WEIL (1982), *An application of the central limit theorem for Banach-space-valued random variables to the theory of random sets*, Z. Wahrsch. Verw. Gebiete, 60, pp. 203–208.
- W. WEIL AND J. WIEACKER (1984), *Densities for stationary random sets and point processes*, Adv. in Appl. Probab., 16, pp. 324–346.
- T. WILLIAMS AND R. BJERKNES (1972), *Stochastic model for abnormal clone spread through epithelial basal layer*, Nature, 236, pp. 19–21.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.